ММТ 2025

12 ДАХЬ УДААГИЙН МОНГОЛЫН МЭДЭЭЛЛИЙН ТЕХНОЛОГИЙН ХУРАЛ

ШУТИС

ММТК
Монголын Мэдээллийн Технологийн Консорциум

МЭДЭЭЛЛИЙН ТЕХНОЛОГИ, ЭЛЕКТРОНИКИЙН СУРГУУЛЬ
МОНГОЛ УЛСЫН ИХ СУРГУУЛЬ

5-р сарын 23, 2025 он, Улаанбаатар хот

ХИЙМЭЛ ОЮУН

Монголын Мэдээллийн Технологийн
12 дахь удаагийн хурлын эмхэтгэл

- Хиймэл оюун
- Кибер аюулгүй байдал
- Дижитал конвержинс
- Холбоо харилцаа, сүлжээ
- Тооцоолол, мэдээлэл технологийн хэрэглээ
- Эх хэл боловсруулалт
- Компьютерын хараа
- Дохио боловсруулалт, автоматжуулалт

mmt.edu.mn

ЗОХИОН БАЙГУУЛАГЧ

ИВЭЭН ТЭТГЭГЧ БАЙГУУЛЛАГА

ХАМТРАГЧ БАЙГУУЛЛАГУУД

Хамрах хүрээ

1. Хиймэл оюун
2. Кибер аюулгүй байдал
3. Дижитал конвержинс
4. Холбоо харилцаа, сүлжээ
5. Тооцоолол, мэдээллийн технологийн хэрэглээ
6. Эх хэл боловсруулалт
7. Компьютерын хараа
8. Дохио боловсруулалт, автоматжуулалт

МОНГОЛЫН МЭДЭЭЛЛИЙН ТЕХНОЛОГИ-2025

Монгол улс, Улаанбаатар хот
2025 оны 05 сарын 23-ны өдөр

©Copyright 2025, ММТ

Энэхүү эмхэтгэлийг хуулбарлах, олшруулах, хэвлэхийг
“Зохиогчийн эрхийн тухай хууль”-аар хориглоно.

www.mmt.edu.mn

Хавтасны дизайн: П.Гантуяа, О.Оргил
Хэвлэлийн эх бэлтгэсэн: П.Гантуяа
Редактор: Б.Наранчимэг

ISBN 978-9919-521-81-3

МЭНДЧИЛГЭЭ

Монголын Мэдээллийн Технологийн Консорциумаас анх 2013 онд санаачилан зохион байгуулсан “Монголын мэдээллийн технологи” эрдэм шинжилгээний хурал нь мэдээллийн технологийн чиглэлээр сургалт, судалгааны үйл ажиллагаа явуулж буй сургууль, байгууллагуудын эрдэмтэн, судлаач, салбарын ажилтнуудын судалгааны үр дүнгээ хэлэлцүүлэх, санал бодлоо солилцох, мэдлэг түгээх чухал ач холбогдолтой хурал болдог билээ.

Мэдээллийн технологи хамгийн хурдан өөрчлөгдөн хөгжиж байдаг бөгөөд ололт амжилт нь бусад салбараас илүү хурдан амьдралд нэвтэрч бүх хүний хэрэглээ болдог онцлогтой. Мэдээллийн технологийн нэг салбар болох хиймэл оюун сүүлийн жилүүдэд хурдацтай хөгжиж байна. Тухайлбал, Web of Science-д бүртгэлтэй сэтгүүлд тус чиглэлээр 2020 онд 26,578 өгүүлэл хэвлэгдсэн бол 2024 онд 36,342 өгүүлэл буюу 5 жилийн дотор ойролцоогоор 10,000 өгүүллээр нэмэгдсэн үзүүлэлт үүнийг илтгэж байна.

Түүнчлэн “МУИС-ийн стратеги төлөвлөгөө 2025-2030”-д сурах, ажиллах, таатай орчныг бүрдүүлэх тэргүүлэх чиглэлийн хүрээнд мэдээллийн технологийн ололт, амжилтыг түүчээлэн хэрэгжүүлж, хиймэл оюун ухаан, их өгөгдлийг сургалт, судалгаа, үйл ажиллагаанд ашиглахаар тусгасныг тэмдэглэн хэлье.

Дэлхийд төдийгүй Монгол Улсад хиймэл оюуныг онцолж байгаа энэ цаг үед “Монголын мэдээллийн технологи – 2025” эрдэм шинжилгээний хурлыг “Хиймэл оюун” сэдвийн хүрээнд зохион байгуулахаар болсон нь эрдэмтэн судлаач, профессор багш нарт салбар дундын судалгааг өргөжүүлэн хамтран ажиллах, судалгааны шинэ аргуудыг нэвтрүүлэх зэрэг шинэ боломжийг олгоно гэдэгт итгэлтэй байна.

“Монголын мэдээллийн технологи – 2025” эрдэм шинжилгээний хурлын үйл ажиллагаанд амжилт хүсье.

МУИС-ийн ректор, профессор Б.Очирхуяг

ӨМНӨТГӨЛ

Хавар хаврын энэ өдрүүдэд манай мэдээллийн технологийн судлаачид уламжлалт “Монголын Мэдээллийн Технологи” эрдэм шинжилгээний хуралдаа оролцон, хийж бүтээсэн эрдэм судлалын ажлын үр дүнгээ нийтийн хүртээл болгож, цаашдын хамтын ажиллагааны үндэс суурийг бэхжүүлдэг.

Энэ жилийн “ММТ2025” хурлыг МУИС-ийн Мэдээллийн Технологи, Электроникийн Сургууль (МТЭС) гардан зохион байгуулж байна. Хурлын голлох сэдэв нь “Хиймэл оюун”. Энэ сэдвийг сонгосон нь учиртай. Дэлхий нийтээр хиймэл оюунаар амьсгалж байгаа цаг мөчид энэ удаагийн хурал болох гэж байна. Хиймэл оюун хэдийнээ хүмүүний амьдралын салшгүй хэсэг, дотны туслагч, танин мэдэхүйн үүд хаалга, бүтээмжийг дээшлүүлэгч хэрэгсэл болжээ. Хиймэл оюуны хөгжил өөрөө техник технологиос гадна нийгэм эдийн засаг, хууль эрх зүй, ёс суртануун гээд олон чиглэлд шинэ шинэ асуудлыг дэвшүүлж, дэлхийн улсууд түүнд яаж хандах, ямар бодлого зохицуулалт хэрэгтэй вэ гэдэгт анхаарлаа хандуулж байна. Манай орны хувьд ч мөн ялгаагүй төрийн бодлогоо тодорхойлох, хөгжлийн стратеги төлөвлөгөөг гаргах ажил өрнөж байна. Хувийн хэвшил технологийн энэ дэвшлийг түүчээлж, шинэ шинэ хэрэглээг нэвтрүүлсээр байна. Боловсролын салбарынханд оюутан, сурагчдын бүтээлч ажлын цаана хүн, машины аль нь байна? Хоёул байгаа бол хувь оролцоо нь юу вэ? гэдгийг ялгахад бэрхтэй болж байна. Сүүлийн үед яль ч салбарын эрдмийн бүтээлд хиймэл оюуны элементийг тусгах явдал ихээхэн нэмэгдсэн. Хиймэл оюуныг тойрсон энэ бүх асуудал өөрөө шинжлэх ухааны судлагдахууны хувьд судлаачдын анхаарлыг ихээр татаж байна. Монголын мэдээлэл технологийн консорциумаас уламжлал болгон зохион байгуулдаг эрдэм шинжилгээний энэхүү хурлаар хиймэл оюун болон өмнөх жилүүдийн уламжлалт сэдвүүдийг мөн тунгаан хэлэлцэх болно.

Хурлын бэлтгэл ажилд профессор Б.Ганбат захиралтай МУИС-ийн МТЭС-ийн хамт олон, түүний дотор дэд профессор Ч.Алтангэрэлийн ахалсан МКУТ-ийн хамт олон зохион байгуулалтын, дэд профессор Б.Наранчимэг, дэд профессор Г.Гэрэлмаа нарын ахалсан хөтөлбөрийн хороод хурлын бэлтгэлийг хангахын тулд хичээн зүтгэснийг цохон тэмдэглэмээр байна. Хуралд өмнөх жилүүдээс бараг 2 дахин их буюу нийт 100 бүтээл ирсэн. Тэдгээрийг шүүн тунгааж, хурлын аман (45) болон үзүүлэн (34) илтгэлийг шалгаруулах амаргүй ажилд зохион байгуулагч сургуулиас гадна ШУТИС, МУБИС, АШУҮИС гээд манай томоохон их сургуулиуд, ШУА-ийн шилдэг судлаачид ихээхэн хүчин чармайлт гаргалаа. Тэдэнд талархал илэрхийлье.

Энэхүү эрдэм шинжилгээний хуралд хэлэлцэгдэх эрдэм шинжилгээний өгүүллүүдийг эмхэтгэн цаасан болон цахим хэлбэрээр та бүхний гар дээр өргөн барьж байна. ММТ2025-д бүтээлээ ирүүлсэн судлаачиддаа Мэдээллийн

Технологийн Консорциумын нэрийн өмнөөс болон хувиасаа чин сэтгэлийн талархал илэрхийлж, цаашдын эрдэм судлалын ажилд нь амжилт хүсье.

Хүндэтгэсэн,

Проф. А.Эрдэнэбаатар,
Монголын Мэдээллийн Технологийн Консорциумын
гүйцэтгэх захирал

ММТ 2025 эрдэм шинжилгээний хурал зохион байгуулах бүрэлдэхүүн

ХУРЛЫН ДАРГА

Б.Ганбат /МУИС, МТЭС-ийн захирал, профессор/

ХӨТӨЛБӨРИЙН ХОРОО

Дарга: Б.Наранчимэг /МУИС, МТЭС, МКУТ-ийн дэд профессор/
Орлогч дарга: Х.Загарзүсэм /ШУТИС, МХТС-ийн ЭНБдарга/
Нарийн бичгийн дарга: Б.Гэрэлмаа /МУИС, МТЭС, ЭХИТ-ийн дэд профессор/

ХӨТӨЛБӨРИЙН ХОРООНЫ ГИШҮҮД:

Б.Ганбат /МУИС, МТЭС, захирал/
А.Эрдэнэбаатар /ШУТИС, МХТС, Компьютерын ухааны салбар, зөвлөх профессор/
Н.Оюун-Эрдэнэ /МУИС, МТЭС, МКУТ-ийн профессор/
Д.Цэдэвсүрэн /МУБИС, СЦШХ-ийн дарга/
Г.Гантуяа /Кибер халдлага, зөрчилтэй тэмцэх нийтийн төв, Мэдээлэл, дүн шинжилгээний газрын дарга/
А.Алтангэрэл /ШУТИС, МХТС, Мэдээллийн технологийн салбарын эрхлэгч/
А.Одгэрэл/ШУТИС, МХТС, Электроникийн салбарын эрхлэгч/
З.Буянхишиг /ШУТИС, МХТС, Холбооны салбарын эрхлэгч/
А.Хүдэр /ШУТИС, МХТС, Компьютерын ухааны салбарын эрхлэгч/
Б.Мөнхбаяр /ШУТИС, МХТС, Мэдээллийн сүлжээ, аюулгүй байдлын салбарын эрхлэгч/
Ч.Мөнхнасан /ШУТИС, МХТС, ахисан түвшний салбарын эрхлэгч/
П.Жаргалбат /АШУҮИС, ФМЗ-ийн тэнхимийн эрхлэгч/
Ю.Отгонбаатар /Хүрээ МХТДС, Тэнхимийн эрхлэгч/
Б.Баянмөнх /Монголын эрүүл мэндийн мэдээллийн технологийн нийгэмлэг, дарга/
Д.Ууганбаатар /ШУА-ийн МТТХ-ийн ЭША, дэд профессор/
Д.Болормаа/СЭЗИС, АТС-ийн профессор/
Б.Сувдаа /МУИС, МТЭС, МКУТ-ийн дэд профессор/
У.Ганбаяр /МУИС, МТЭС, ЭХИТ-ийн ахлах багш/
Р.Мижиддорж /МУБИС, МЗТ-ийн дэд профессор/
Х.Золзаяа /МУИС, МТЭС, ЭХИТ-ийн ахлах багш/
Р.Амартүвшин /МУИС, МТЭС, ЭХИТ-ийн дэд профессор/
Д.Оргил /ХААИС, ИТС, МФМТ-ийн ахлах багш/
Ц.Ганбат /Отгонтэнгэр ИС, ХЭЗС, БПХТ-ийн дэд профессор/
М.Мэнд-Амар /МУИС, МТЭС, ХМТ-ийн дэд профессор/

Ц.Жавзмаа /АШУҮИС, Био-Анагаахын сургууль, Физик, мэдээлэл зүйн тэнхимийн дэд профессор/

ЗОХИОН БАЙГУУЛАХ ХОРОО

Дарга: Ч.Алтангэрэл /МУИС, МТЭС, МКУТэнхимийн эрхлэгч, дэд профессор/

ЗОХИОН БАЙГУУЛАХ ХОРООНЫ ГИШҮҮД:

С.Батбилэг /МУИС, МТЭС, ХМТ, дэд профессор/
М.Золжаргал /МУИС, МТЭС, МКУТ, дэд профессор/
П.Гантуяа /МУИС, МТЭС, МКУТ, дэд профессор/
Г.Анхбаяр /МУИС, МТЭС, ХМТ, дэд профессор/
Г.Махгал /МУИС, МТЭС, ХМТ, ахлах багш/
К.Есенбек /МУИС, МТЭС, МКУТ, багш/
Б.Наранбат /МУИС, МТЭС, МКУТ, багш/
А.Мөнхзул /МУИС, МТЭС, МКУТ, багш/
Э.Адъяабат /МУИС, МТЭС, ЭХИТ, багш/
Б.Аззаяа /МУИС, МТЭС, МКУТ, дадлагажигч багш/
О.Билгүүн / МУИС, МТЭС, МКУТ, дадлагажигч багш /
О.Шинэтулга /МУИС, МТЭС, ЭХИТ, лабораторийн сургагч/
Д.Тэмүүжин /МУИС, МТЭС, МКУТ, лабораторийн сургагч/
Г.Сарнай /МУИС, МТЭС, МКУТ, лабораторийн сургагч/
Э.Билгүүн /МУИС, МТЭС, МКУТ, лабораторийн сургагч/
Г.Батдэлгэр /МУИС, МТЭС, МКУТ, лаборарант/
М.Амарзаяа /МУИС, МТЭС, МКУТ, тэнхимийн туслах/
П.Пүрэвханд /МУИС, МТЭС, МКУТ, тэнхимийн туслах/
Н.Баярсайхан /МУИС, МТЭС, ЭХИТ, тэнхимийн туслах/

ММТ2025 ХИЙМЭЛ ОЮУН Эрдэм шинжилгээний хурлын хөтөлбөр

Цаг	Хөтөлбөр		
БҮРТГЭЛ			
08:30-09:00	Хуралд оролцогчдын бүртгэл		*МУИС-ийн Номын сан, V давхар 502 тоот*
НЭЭЛТИЙН ҮЙЛ АЖИЛЛАГАА			
09:00-09:10	Зочид төлөөлөгчдийг танилцуулах		*МУИС-ийн Номын сан, V давхар 502 тоот*
09:10-09:25	Нээлтийн үг 1. МУИС-ийн ректор, профессор Б.Очирхуяг 2. ЦХИХХЯамны сайд Ц.Баатархүү 3. Монгол Улсын Ерөнхийлөгчийн зөвлөх, дэд профессор Ч.Лодойравсал 4. ММТКонсорциумын гүйцэтгэх захирал, профессор А.Эрдэнэбаатар		
09:25-09:30	ММТ-2025 ЭШХ-ын хөтөлбөрийг танилцуулах, дурсгалын зураг татуулах		
ҮНДСЭН ХУРАЛДААН			
09:30-09:50	Уригдсан илтгэл-1, *П.Мөнхбат* Их өгөгдөл, хиймэл оюуны үндэсний стратеги		*МУИС-ийн Номын сан, V давхар 502 тоот*
09:50-10:10	Уригдсан илтгэл-2, *С.Сайнбаяр* Large Language Models		
10:10-10:30	Уригдсан илтгэл-3, *Ч.Алтангэрэл* Challenges in Language Resources		
10:30-11:00	ЦАЙНЫ ЗАВСАРЛАГА		
САЛБАР ХУРАЛДААН			
11:00-12:30	**Салбар хуралдаан 1:** ХИЙМЭЛ ОЮУН I *МУИС-ийн Номын сан, V давхар 502 тоот*	**Салбар хуралдаан 2:** КИБЕР АЮУЛГҮЙ БАЙДАЛ *МУИС-ийн Номын сан, II давхар 202 тоот*	**Салбар хуралдаан 3:** ДИЖИТАЛ КОНВЕРЖИНС *МУИС-ийн Номын сан, II давхар 203 тоот*
ҮЗҮҮЛЭН ИЛТГЭЛ			
12:30-13:00	ҮЗҮҮЛЭН ИЛТГЭЛ I		*МУИС-ийн Номын сан, V давхар 502 тоот*
ӨДРИЙН ХООЛНЫ ЗАВСАРЛАГАА			
13:00-14:00	ӨДРИЙН ХООЛ		*МУИС-ийн Номын сан, V давхар 502 тоот*
САЛБАР ХУРАЛДААН			
14:00-15:15	**Салбар хуралдаан 4:** ХИЙМЭЛ ОЮУН II *МУИС-ийн Номын сан, V давхар 502 тоот*	**Салбар хуралдаан 5:** ХОЛБОО ХАРИЛЦАА, СҮЛЖЭЭ *МУИС-ийн Номын сан, II давхар 202 тоот*	**Салбар хуралдаан 6:** ТООЦООЛОЛ, МЭДЭЭЛЛИЙН ТЕХНОЛОГИЙН ХЭРЭГЛЭЭ *МУИС-ийн Номын сан, II давхар 203 тоот*

Цаг	Хөтөлбөр		
15:20-16:00	ЦАЙНЫ ЗАВСАРЛАГА		
ҮЗҮҮЛЭН ИЛТГЭЛ			
15:20-16:00	ҮЗҮҮЛЭН ИЛТГЭЛ II		
САЛБАР ХУРАЛДААН			
16:00-17:15	**Салбар хуралдаан 7:** ЭХ ХЭЛ БОЛОВСРУУЛАЛТ *МУИС-ийн Номын сан, V давхар 502 тоот*	**Салбар хуралдаан 8:** КОМПЬЮТЕРЫН ХАРАА *МУИС-ийн Номын сан, II давхар 202 тоот*	**Салбар хуралдаан 9:** ДОХИО БОЛОВСРУУЛАЛТ, АВТОМАТЖУУЛАЛТ *МУИС-ийн Номын сан, II давхар 203 тоот*
17:15-18:00	ХЭЛЭЛЦҮҮЛЭГ, ХААЛТ		*МУИС-ийн Номын сан, V давхар 502 тоот*
18:00-20:00	ХҮНДЭТГЭЛИЙН АРГА ХЭМЖЭЭ		*Танан төв IV давхар, Танан event hall*

ҮНДСЭН ХУРАЛДААН: *МУИС-ийн Номын сан, V давхар 502 тоот* Хуралдаан даргалагч: Б.Ганбат /МУИС, Профессор/		
09:30-09:50	Их өгөгдөл, хиймэл оюуны үндэсний стратеги *П.Мөнхбат, ЦХИХХЯ, Инновацын бодлого зохицуулалтын газрын дарга*	
09:50-10:10	Large Language Models *С.Сайнбаяр, Meta AI*	
10:10-10:30	Challenges in Language Resources *Ч.Алтангэрэл, МУИС, МКУТ, Тэнхимийн эрхлэгч*	
Салбар хуралдаан 1: ХИЙМЭЛ ОЮУН I *МУИС-ийн Номын сан, V давхар 502 тоот* Хуралдаан даргалагч: Ч.Алтангэрэл /МУИС, Дэд профессор/		
11:00-11:15	CLOS network: A parameter-efficient alternative to linear layers in neural networks *Orgil Jargalsaikhan, Dorjkhand Battulga, Dolgorsuren Dulamjav, Bayrzul Sandagjav*	1
11:15-11:30	Криптовалютын үнийн хэлбэлзлийг таамаглах олон моделын харьцуулсан судалгаа *Б.Оршихбаяр, Эрдэнэбат Батцэцэг*	2
11:30-11:45	Хиймэл оюун ухаан ашиглан хорт URL илрүүлэх *П.Цэрэнпунцаг, Ж.Эрдэнэсувд, А.Буяннэмэх, Б.Мөнхбаяр*	3
11:45-12:00	Гүн сургалтанд ашиглагддаг өгөгдлийн форматуудын ашигт байдлын талаарх экспериментал судалгаа *Баттогтохын Доржнамжаа, Довдонгийн Энхзол*	4
12:00-12:15	Боловсролын салбар дахь хиймэл оюуны ухааны нөлөөллийн судалгаа *Гүндсамбуу Болд, Удвалцэцэг Батжав*	5

Цаг	Хөтөлбөр	
12:15-12:30	Өвчлөх эрсдлийг машин сургалтаар тооцох аргачлалын зохиомж *Хорлоогийн Дорж, Сандагийн Цогтсайхан, Зундуйн Цолмон, Ганболдын Амарсанаа*	6
Салбар хуралдаан 2: КИБЕР АЮУЛГҮЙ БАЙДАЛ *МУИС-ийн Номын сан, II давхар 202 тоот* Хуралдаан даргалагч: У.Ганбаяр /МУИС, Дэд профессор/		
11:00-11:15	Свитчийн порт мироринг функц халдлага илрүүлэх системийн пакетийн алдагдалд нөлөөлөх нь *Б.Энхбаяр, Г.Манлайбаатар, Д.Баттулга, Б.Өсөхбаяр, Н.Угтахбаяр*	7
11:15-11:30	Сүлжээний халдлага илрүүлэх, сэргийлэх нээлттэй эхийн системийг ашиглах судалгаа *Эрдэнийн Энх-Од, Ямхины Дашдорж, Хүчитийн Уянга*	8
11:30-11:45	Халдлага илрүүлэх системийн онцлогийг XAI ашиглан шинжилгээ хийх нь *Н.Угтахбаяр, П.Баясгалан, Г.Манлайбаатар, Д.Баттулга, Б.Өсөхбаяр*	9
11:45-12:00	Орчин үеийн системийн нийлүүлэлтийн гинжин хэлхээний аюулгүй байдлын судалгаа *Г.Ганчимэг, Т.Отгонбаяр*	10
12:00-12:15	Вэб аппликейшнууд дахь олон шатлалт баталгаажуулалт (MFA)-ыг тойрох халдлагын туршилт ба хамгаалалтын аргачлалуудыг боловсруулах нь *Лхагваагийн Одончимэг, Шүрэнцэцэгийн Нандиндулам*	11
12:15-12:30	Seeing the Unseen: Visualizing and Enriching Cyber attack *Nyamjargal Tulgabaatar, Ganbayar Uuganbayar, Adiyabat Enkhjargal and Telmuun Tumnee*	12
Салбар хуралдаан 3: ДИЖИТАЛ КОНВЕРЖИНС *МУИС-ийн Номын сан, II давхар 203 тоот* Хуралдаан даргалагч: Л.Ажнай /АШУҮИС, Профессор/		
11:00-11:15	Хүний биеийн 3D хөдөлгөөнийг дүрслэх мэдээллийг шахах арга *Өлзий-Ундралын Оргил, Ганболдын Амарсанаа*	13
11:15-11:30	Ерөнхий физикийн лабораторийн хичээлд сургалтын Moodle системийг ашигласан нь *Баярсайханы Алтанхүү, Шаравын Мөнхжаргал, Самбуугийн Мөнхцэцэг*	14
11:30-11:45	Чанарт төвлөрсөн Ажайл хөгжүүлэлтийн менежментийн арга болон түүний дэмжих орчин *Баттулгын Цэлмэг, Накашима Цуёоши*	15
11:45-12:00	Монгол улсын нийтийн өргөдлийн системд хийсэн үнэлгээ *Сайннямбуугийн Баясгалан, Самбуугийн Уянга, Сайнбилэгийн Алунгоо, Алтанцэцэгийн Мөнхзул*	16

Цаг	Хөтөлбөр	
12:00-12:15	Эрүүл Мэндийн Мэдээлэл зүйн хөтөлбөрийн шинэчлэлт *Пүрэвдолгор Лувсанцэрэн, Жавзмаа Цэнд, Ажнай Лувсан, Одгэрэл Болдбаатар, Баасандорж Чилхаасүрэн, Оюунцэцэг Сандаг, Ахыт Тилуебай, Нямдаваа Угандаваа, Энх-Үрэл Энхжаргал*	17
Салбар хуралдаан 4: ХИЙМЭЛ ОЮУН II *МУИС-ийн Номын сан, V давхар 502 тоот* Хуралдаан даргалагч: Д.Золзаяа /ШУТИС, Дэд профессор/		
14:00-14:15	Simplified graph neural network for node ranking *Amarsanaa Ganbaatar, Dorjnamjirmaa Badraa, Dalaijargal Purevsuren*	19
14:15-14:30	Физикээс Сэдсэн Граф Нейрал Сүлжээний хувилбарууд, тэдгээрийн харьцуулалт *Д.Тэмүүжин, П.Далайжаргал*	20
14:30-14:45	Төрийн цахим үйлчилгээг мэдлэгийн графт анхааралтай сүлжээгээр санал болгох туршилт *Мөнхжаргалын Билгүүндалай, Гантулгын Сарнай, Ганболдын Амарсанаа*	21
14:45-15:00	Өсөн Нэмэгдэх Стохастик Блок Загвар *Г.Гантулга, П.Далайжаргал*	22
15:00-15:15	Машин сургалт ба граф суурьт хамтын шүүлтүүрийг ашиглан бүтээгдэхүүн санал болгох системийг сайжруулах нь *Энхтөрийн Сэргэлэнбаатар, Тэгшээгийн Түвшинсайхан, Ишдоржийн Цэрэн-Онолт, Батжаргалын Долгорсүрэн*	23
Салбар хуралдаан 5: ХОЛБОО ХАРИЛЦАА, СҮЛЖЭЭ *МУИС-ийн Номын сан, II давхар 202 тоот* Хуралдаан даргалагч: З.Буянхишиг /ШУТИС, Дэд профессор/		
14:00-14:15	5G NSA сүлжээн дэх чадлын үр ашигтай нөөц хуваарилалт *Ванданцэрэнгийн Нарантуяа, Найманнарангийн Чулуунбанди, Ламжавын Эрдэнэбаяр*	24
14:15-14:30	Холбооны зориулалттай цацаргагч фидерээс энерги хуримтлуулах тородойл загвар *Батболд Дэлгэрмөрөн, Батбаатар Загарсүрэн, Гантас Энхзул, Дашбаатар Анударь, Баярсайхан Пүрэвцэрэн, Пүрэвдорж Ууганбаяр*	25
14:30-14:45	Био-анагаахын радио долгион дээр элэгний хорт хавдрыг илрүүлэх мэдрүүр антены судалгаа *Хүрэлсүх Есөн-Эрдэнэ, Батбаатар Загарсүрэн, Баярсайхан Пүрэвцэрэн, Пүрэвдорж Ууганбаяр*	26

Цаг	Хөтөлбөр	
14:45-15:00	Дрон унагаах төхөөрөмжид зориулсан өндөр үзүүлэлттэй антены хөгжүүлэлт *Ганбаатар Дөлмандах , Баярсайхан Пүрэвцэрэн , Пүрэвдорж Ууганбаяр*	27
15:00-15:15	5G сүлжээний хуваалт хоорондын хэрэглэгч баталгаажуулалт *Г.Манлайбаатар, Н.Угтахбаяр, Д.Баттулга, Б.Өсөхбаяр*	28
Салбар хуралдаан 6: ТООЦООЛОЛ, МЭДЭЭЛЛИЙН ТЕХНОЛОГИЙН ХЭРЭГЛЭЭ *МУИС-ийн Номын сан, II давхар 203 тоот* Хуралдаан даргалагч: Д.Цэдэвсүрэн /МУБИС, Профессор/		
14:00-14:15	Албан хэрэг хөтлөлтийн мэдээллийн систем нэвтрүүлснээр гарах үр ашиг *Баярмагнай Даваахүү, Самбуу Уянга, Алтанцэцэг Мөнхзул*	29
14:15-14:30	(K-means) боловсролын салбарын оновчтой шийдвэр гаргалтад хэрэглэх *Г.Батдэлгэр, Г.Баттөр, С.Батбилэг, Т.Цэрэннадмид, К.Есенбек*	30
14:30-14:45	Нарны теоретик потенциалд үүлнээс үүсэх нөлөөллийг IEEE-ийн загвараар тооцоолж үнэлэх нь *Отгонбаярын Уянга, Wooyoung Song*	31
14:45-15:00	Хиймэл оюунд ашиглах сүлжээний халдлагын сан үүсгэх *П.Баясгалан, Г.Манлайбаатар, Д.Баттулга, Н.Угтахбаяр*	32
15:00-15:15	Монгол улс дахь нээлттэй эхийн сангийн аюулгүй байдал, хэрэглээ *Индра Өлзийбадрах, Ганбаяр Ууганбаяр, Угтахбаяр Найдансүрэн*	33
Салбар хуралдаан 7: ЭХ ХЭЛ БОЛОВСРУУЛАЛТ *МУИС-ийн Номын сан, V давхар 502 тоот* Хуралдаан даргалагч: М.Золжаргал /МУИС, Дэд профессор/		
16:00-16:15	Нийгмийн сүлжээн дэх стандарт бус галиглагдсан бичвэрийг Кирилл бичвэрт хөрвүүлэх загварын судалгаа *Бямбадоржийн Золзаяа, Соном-Очирын Өлзийбаяр, Аюушийн Алтангэрэл, Энхбаярын Мөнхсүх, Kim Hyunchul*	34
16:15-16:30	Balancing Retrieval and Forgetting: A Comparative Analysis of RAG and Machine Unlearning For Hallucination *Taivanbat Bayarsaikhan, Altangerel Chagnaa*	36
16:30-16:45	Flickr30k-MN: Англи-Монгол зургийн тайлбарын өгөгдөл, орчуулгын үнэлгээ *Очирбатын Билгүүн, Чагнаагийн Алтангэрэл*	37
16:45-17:00	Ухаалаг хөтөч *Мөнхбаяр Түмэн-Аюуш, Чулуунжин Мичидгоо, Сүхбаатар Батмөнх, Соном-Очир Өлзийбаяр*	38
Салбар хуралдаан 8: КОМПЬЮТЕРЫН ХАРАА *МУИС-ийн Номын сан, II давхар 202 тоот*		

Цаг	Хөтөлбөр	
Хуралдаан даргалагч: Б.Сувдаа /МУИС, Дэд профессор/		
16:00-16:15	Гүн сургалтын арга ашиглан зурагнаас ой модыг илрүүлж тоолох нь *Т. Оюунбаатар, Т.Цэрэннадмид, Х.Оюундолгор*	39
16:15-16:30	Микроскопын дүрснээс Enterobius vermicularis-ийн өндөгийг гүн сургалт ашиглан илрүүлэх нь *Тэгшээгийн Түвшинсайхан, Доржсүрэнгийн Тэмүүлэн. Наранцогтын Гиймаа, Батжаргалын Долгорсүрэн*	40
16:30-16:45	Бодит цаг хугацаанд дамжуурга дээрх цементийг тоолох систем *Цолмонгийн Энхзаяа, Батсуурийн Сувдаа, Ганболдын Ууганбаяр*	41
16:45-17:00	Веб камер ашиглан нүдний хөдөлгөөнөөр виртуал аяллыг удирдах *Рэнцэндорж Жавхлан, Намсрайдорж Мөнхцэцэг, Намсрай Оюун-Эрдэнэ, Эрдэнэбаатар Цэцэгжаргал, Цоодол Лхамролом*	42
Салбар хуралдаан 9: ДОХИО БОЛОВСРУУЛАЛТ, АВТОМАТЖУУЛАЛТ *МУИС-ийн Номын сан, II давхар 203 тоот* Хуралдаан даргалагч: Ц.Тэнгис /ШУТИС, Дэд профессор/		
16:00-16:15	Монтгомери үржүүлэгчийн дизайныг FPGA дээр хэрэгжүүлэх нь *Болдбаярын Намуунаа, Батсүхийн Алтанцоож*	43
16:15-16:30	PQC mceliece алгоритмд зориулсан эрэмбэлэх алгоритмын хэрэгжүүлэлт *Баатарсүрэнгийн Энхзаяа, Батсүхийн Алтанцоож*	44
16:30-16:45	90нм технологи ашигласан LDO-ийн дизайн болон хэрэгжүүлэлт *Цогтгэрэлийн Энхтуяа, Батсүхийн Алтанцоож*	45
16:45-17:00	Хүний хоолойн өнгөнд суурилсан уушгины хатгалгааг оношлох судалгаа *Дорждэрэмийн Балжинням, Бямбадоржийн Золзаяа*	46
17:00-17:15	Object Pose Estimation for Random Bin Picking Using Synthetic Data *Munkhtulga Byambaa, Lodoiravsal Choimaa, Gou Koutaki*	47

ҮЗҮҮЛЭН ИЛТГЭЛ 1 *МУИС-ийн Номын сан, V давхар*		
1	Хуулийн хэлний магадлалт биграмм загвар *Г.Чинбаяр, Ц.Ганбат*	48
2	3D, 360 технологи дээр суурилсан сургалтийн цахим контент боловсруулалт *Б.Батшагай, М.Жавхлан, Ц.Дэвжрэх*	49
3	Ангиллын зарим аргаар өсвөр насныхны сэтгэл хөдлөл, зан үйлийг хүйсээс хамааруулан судлах *Батнаст Ганчөдөр, Ахыт Тилуебай, Жавзмаа Цэнд, Баярмаа Ваанчиндорж, Нямдаваа Уугандаваа, Пүрэвдолгор Лувсанцэрэн, Ажнай Лувсан, Баасандорж Чилхаасүрэн, Энх-Үрэл Энхжаргал, Булгансүрэн Дагвасамбуу*	50
4	Analysis of Traffic Congestion in UB city, Mongolia using CNN and Agent-Based Modelling *Erdenetuya Erdenebileg, Altanzaya Tumen*	52

5	Хиймэл оюун ухаанд суурилсан RedCAD программаар чихрийн шижингийн ретинопатийн эрт илрүүлэг хийсэн үр дүн *Пүрэвдорж, Баясгалан, Эрдэнэбаатар, Түмэнжаргал*	53
6	Пост квант криптографийг эмбэддэд системд хэрэгжүүлэх нь *Мөнхтөрийн Мөнхбагана, Нэмэхбаярын Давхарбаяр, Дэмчигсүрэнгийн Ганбаатар, Бат-Эрдэнийн Мөнхбаяр*	54
7	Монгол бичгийн хэвлэмэл текстийг машин сургалт болон гүн сургалтаар таних нь *Даваацэрэнгийн Цэцэнбилэг, Энхтөрийн Чойжилням, Батсуурийн Сувдаа*	56
8	Уул уурхайн салбарын кибер аюулгүй байдлын төлөвшлийг тодорхойлох судалгаа *Хүүхээгийн Мөнхтөр, Дэмчигийн Болормаа*	57
9	AI ашигласан чихрийн шижингийн эрсдэлийг урьдчилан таамаглах, сэргийлэх ухаалаг загвар *Э.Түвшинжаргал, Б.Золзаяа*	58
10	Морин хуурын нумны барилын онцлог, цахилгаан дохионы үзүүлэлтүүд *Цагаан Баатархүү, Оюунбилэг Бат-Энх, Оюунбилэг Баттөгс*	59
11	Монгол үндэсний хээ угалзын зарим дүрсийг математик загварчлалаар вектор график болгон зурах *П.Гантуяа, Г. Багаболд, А.Мөнхзул*	60
12	Хиймэл оюун ашигласан яриа таних загвар *Т.Отгонбаатар, Э.Адъяабат*	61
13	Нууц Үгийн Халдлагад Зориулсан Солонгон Хүснэгт (Rainbow table)-ийг Монгол Хүний Нэрийн Санд Тулгуурлан Үүсгэх нь *А.Ангараг, Б.Өсөхбаяр, Н.Угтахбаяр*	62
14	Фрэймийн тоог ихэсгэх байдлаар дүрс бичлэгийн чанарыг сайжруулах *Т. Болдбаяр, Эрдэнэбат Батцэцэг*	63
15	Биоангаахын хэрэглээнд зориулсан бага чадлын хоёр шатлалт CMOS үйлдлийн өсгүүрийн загварчлал ба шинжилгээ *Дамдингийн Баярцэцэг, Даланбаярын Болормаа*	64
16	Машин сургалтын аргаар экскаваторын шанаганы шүд илрүүлж, хүдрийн бүхэллэг тодорхойлох судалгаа *А.Түвшинжаргал, Д.Энхбаатар, Б.Гэрэлмаа, Б.Ганбат*	65
17	Био анагаахын SAR ADC-ийн Sample and hold болон компараторын параметрын хамаарлын судалгааг *Анхбаярын Наранмандах, Алтансүхийн Сугаррагчаа, Даланбаярын Болормаа*	66
18	Оновчтой сплайн муруй ашигласан дифференциал роботын зам төлөвлөлт *Д.Нанзадрагчаа, М.Баярпүрэв*	67
19	Маркетингийн идэвхжүүлэлтэд мэдлэгийн граф ашиглах нь *Нандин-Эрдэнэ Энхмягмар, Энхтуул Бөхсүрэн, Мөнхцэцэг Намсрайдорж*	68
ҮЗҮҮЛЭН ИЛТГЭЛ 2 *МУИС-ийн Номын сан, V давхар*		
20	Монгол бичвэрийг хураангуйлах аргын судалгаа *Бөхбатын Амартүвшин, Нямын Гантөмөр, Бямбадоржийн Золзаяа*	69
21	Багшийн дижитал мэдлэг, чадварын оюутны үнэлгээний кластер шинжилгээ	70

	Д.Мөнхбаатар, Д.Цэдэвсүрэн, Л.Пүрэвдолгор, Ч.Баасандорж, Ч.Галбадрах, Д.Булгансүрэн	
22	ESP32 Платформ дээр Matter Протокол Ашиглан Гэрэл, Унтраалгыг Холбох нь *С.Батбаяр, Н.Оюун-Эрдэнэ*	71
23	Монголын эрх зүйн бичиг баримтыг гүнзгий сургалт ашиглан шинжилж, шийдвэр гаргахад туслах системд нэвтрүүлэх талаарх судалгаа *Халтархүү Гармаабазар, Билиг Батжа*	72
24	Нарны эрчим хүчний тоолуурын тоолуурын системийн хөгжүүлэлт *Аюурзанын Одгэрэл*	73
25	Монгол модон барын бүтээлийг гар бичмэлийн бүтээлээс ялган таних *Пэрэнлэйлхүндэв Гантуяа, Батсуурь Сувдаа, Ганболд Ууганбаяр*	74
26	Семантик холбоо ашигласан яриа таних систем *Э.Адъяабат, М.Баярпүрэв, Д.Бямбажав*	75
27	Хиймэл оюун ухаанд суурилсан арьс оношилгооны аппликейшн *Болор-Эрдэний Пүрэвжаргал, Түвшинтөгсийн Энхжин, Алтансүхийн Сүхбат*	76
28	CCTV/IP Камерын Сүлжээний Аюулгүй Байдлын Шинжилгээ (Монгол Улсын кейс судалгаа) *О.Энхтуул, Б.Дэнсмаа*	77
29	Эмийн жор бичээч ухаалаг систем *Үүрцайхын Мөнхтуяа, Гантулгын Магван-Эрдэнэ, Батжаргалын Долгорсүрэн*	78
30	A User-Centric Tool for Design Flaw Detection in UML XMI Models Using Pre-Trained Machine Learning Evaluators *Bilguun Erdenebaatar, Lkhamrolom Tsoodol, Oyun-Erdene Namsrai, Batnyam Battulga, Naranchimeg Bold*	79
31	Two Wheel Drive Autonomous Robot Navigation using GPS *Temuulen Batbold, Munkhtulga Byambaa*	80
32	Бакалаврын судалгааны ажлын гүйцэтгэлд нөлөөлж буй хүчин зүйлсийг шинжилж, машин сургалтын арга ашиглан амжилтыг таамаглах нь *М.Амарзаяа, Б.Наранбат, М.Золжаргал*	81
33	Сүлжээний урсгалын шинжилгээнд суурилсан аюул илрүүлэх нь (radmin болон w32.rahack-ийн жишээн дээр) *Есөнжингийн Энхзаяа, Дондогмэгдийн Бямбадорж*	82
34	IoT орчинд нээлттэй портуудын аюулгүй байдлын шинжилгээ *Батчулуун.Жавхлантөгс, Дондогмэгд Бямбадорж, Бат-Эрдэнэ Мөнхбаяр*	83

ҮНДСЭН ХУРАЛДААН

УРИЛГЫН ИЛТГЭЛ

ҮНДСЭН ХУРАЛДААН: *МУИС-ийн Номын сан, V давхар 502 тоот* Хуралдаан даргалагч: Б.Ганбат /МУИС, Профессор/	
09:30-09:50	Их өгөгдөл, хиймэл оюуны үндэсний стратеги *П.Мөнхбат, ЦХИХХЯ, Инновацын бодлого зохицуулалтын газрын дарга*
09:50-10:10	Large Language Models *С.Сайнбаяр, Meta AI*
10:10-10:30	Challenges in Language Resources *Ч.Алтангэрэл, МУИС, МКУТ, Тэнхимийн эрхлэгч*

Их өгөгдөл, хиймэл оюуны үндэсний стратеги

П.Мөнхбат
ЦХИХХЯ, Инновацын бодлого зохицуулалтын газрын дарга

Хураангуй

Олон улсын их өгөгдөл, хиймэл оюуны чиг хандлагыг Монголын эдийн засагт хэрхэн ашиглах, стратегийн үе шат, хүрэх үр дүн, хиймэл оюуны дата төв, экосистемийг хөгжүүлэх талаар илтгэл хэлэлцүүлнэ.

Товч намтар

МУИС-ийн Эдийн засаг судлалын бакалавр зэрэгтэй. Дэлхийн банканд төслийн зөвлөх, “Оюу Толгой” ХХК-д бүс нутгийн стратегийн шинжээч, зөвлөхөөр ажиллаж байсан.

Large Language Models

С.Сайнбаяр
Meta AI

Хураангуй

Сүүлийн жилүүдэд хиймэл оюун үсрэнгүй хөгжиж байгаа ба үүний гол хөдөлгүүр нь эх хэлний том загвар юм. Энэ яриа нь хэлний том загварын эхлэл түүхээс эхлэн хамгийн сүүлийн үеийн ололт дэвшилт, мөн ирээдүйн хөгжийн чиг хандлагыг судалгааны талаас нь хураангуйлан тайлбарлах болно.

Товч намтар

Сүхбаатарын Сайнбаяр нь 2011, 2013 онуудад Токиогийн Их Сургуулиас бакалавр болон магистрын зэрэг хүртсэн ба 2018 онд Нью-Йоркийн Их Сургуульд хиймэл оюуны чиглэлээр докторын зэрэг хамгаалсан. Түүнээс хойш Мета компанид эрдэмтэн судлаачаар ажиллаж, хиймэл оюуны сэдвээр нийт 50 гаруй эрдэм шинжилгээний ажил хэвлүүлсэн нь 9000 давсан ишлэл цуглуулжээ. Түүнчлэн 2018 онд Монголд хиймэл оюуны DLUB зуны сургалтыг санаачлан эхлүүлж жил бүр зохион байгуулж байна.

Challenges in Language Resources

Ч.Алтангэрэл
МУИС, МКУТ, Тэнхимийн эрхлэгч

Хураангуй

Орчин үеийн ухаалаг технологиуд, хиймэл оюуны загвар, ялангуяа том хэлний загваруудын гүйцэтгэл нь хүний түвшинд хүрч байгаа амжилтын гол хүчин зүйл нь түүнийг сургах өгөгдөл буюу төрөл бүрийн хэлний нөөц юм. Хэлний нөөц гэж юу болох тухай, түүний хэмжээ, чанар зэрэг хүчин зүйлсээс гадна хэл, соёлын ялгаатай байдал ч гэсэн хэрхэн нөлөөлдөг хүчин зүйл болох талаар танилцуулна. Мөн Монгол хэлний хувьд хэлний нөөцийн хувьд олон улсын түвшинтэй харьцуулахад хэр түвшинд байгаа, тулгарч байгаа асуудал гэх мэтийг дүгнэн хэлэлцэх болно.

Товч намтар

МУИС-д 2001 онд бакалавр, 2003 онд магистр зэрэг, БНСУ-ын Улсаны Их Сургуульд 2008 онд Эх хэл боловсруулалт чиглэлээр докторын зэрэг хамгаалсан. 2001 оноос эхлэн туслах ажилтан, сургалтын инженер, багш, дэд профессорын албыг хашин өнөөг хүртэл тасралтгүй ажиллахдаа олон тооны магистр, докторантыг удирдаж хамгаалуулжээ. Эх хэл боловсруулалт, үг зүй, өгүүлбэр зүйн задлуур, машин орчуулга, утга зүйн сүлжээ, мэдлэгийн цахим сан, яриа боловсруулалт, бичвэрээс ярианд хувиргах, бичвэр дуудлага хувиргуур зэрэг чиглэлээр эрдэм шинжилгээний ажил хэвлүүлж, дотоод гадаадын судалгааны хэд хэдэн төсөл удирдан болон хамтран хэрэгжүүлжээ.

САЛБАР ХУРАЛДААН

CLOS Network: A Parameter-Efficient Alternative to Linear Layers in Neural Networks

1st Orgil Jargalsaikhan
Computer Science Department, MHTS, SHUTIS
Mongolian University of Science and Technology
Ulaanbaatar, Mongolia
orgilsict@must.edu.mn

2nd Dorjkhand Battulga
Computer Science Department, MHTS, SHUTIS
Mongolian University of Science and Technology
Ulaanbaatar, Mongolia
battulgdorjhand@gmail.com

3rd Dolgorsuren Dulamjav
Telecommunications Department, MHTS, SHUTIS
Mongolian University of Science and Technology
Ulaanbaatar, Mongolia
dolgorsuren@must.edu.mn

4th Bayrzul Sandagjav
Computer Science Department, MHTS,
SHUTIS Mongolian University of Science and Technology Ulaanbaatar, Mongolia
bayrzul@must.edu.mn

***Abstract*—In this paper, we propose utilizing the CLOS network architecture to replace traditional linear layers in deep learning models, including transformers. The CLOS network, commonly used in networking systems, is adapted to neural networks to reduce parameter sizes while maintaining model performance. Our experiments show that the CLOS network achieves comparable accuracy and loss to the conventional linear layer, with up to 40% fewer parameters. However, this efficiency comes at the cost of increased processing time, which is 1.5x to 3x slower. Despite this trade-off, the CLOS network offers a promising alternative for parameter reduction in various architectures, particularly for memory-constrained environments.**

***Index Terms*—Linear layer, CLOS network, Transformers, Parameter efficiency**

I. Introduction

Linear layers are fundamental building blocks of modern deep learning architectures, including models such as transformers, convolutional neural networks, and fully connected networks. Despite their simplicity and wide usage, linear layers can contribute to significant model size, which can be a limiting factor in memory-constrained environments or when deploying models at scale. Addressing this issue by reducing the number of parameters while maintaining model performance has become an active area of research.

The CLOS network, initially designed for efficient communication in high-performance networking systems, can potentially replace linear layers in neural networks. The CLOS architecture features a multi-stage interconnection of smaller switches, which can be adapted to neural network designs to reduce the number of trainable parameters. This approach promises to make neural networks memory-efficient while maintaining comparable accuracy and loss performance.

This work proposes replacing conventional linear layers with a CLOS network-based architecture across different neural network models, including transformers. Our experimental results demonstrate that the CLOS network can reduce the number of parameters by up to 40% in these models without compromising performance metrics such as accuracy and loss. However, this reduction comes at the cost of increased computational time, with our findings showing a 1.5x to 3x slower processing time compared to traditional linear layers.

This paper is organized as follows: Section II reviews related work on parameter-efficient neural network designs. Section III describes the implementation of the CLOS network within neural network architectures. Section IV presents the experimental setup and results. Finally, Section V discusses the findings and potential future research directions.

II. Related Work

Reducing Parameter Count in Deep Neural Networks. This area focuses on methods such as pruning and quantization, which aim to decrease the size of neural networks without compromising accuracy. Notable works include: "Deep Compression" (2015) by Han et al. [1], which effectively reduces parameters by removing less essential weights, and Jacob et al.'s (2018) quantization work [2], which uses quantized weights to enhance model deployment in hardware-constrained environments. These methods align with our goal of reducing parameters in deep models but typically concentrate on pruning or compression rather than architectural changes like CLOS networks.

Efficient Neural Network Architectures. Research in this category, such as the "Inception" architecture by Szegedy et al. [3] and SqueezeNet by Iandola et al. (2016) [4], proposes factorized convolutions and smaller, efficient networks. These works aim to maintain high performance while reducing parameter count, providing a conceptual framework similar to our application of the CLOS network for parameter efficiency in neural networks.

Alternatives to Fully Connected Layers. Approaches such as random projections and low-rank factorization (Denil et al., 2013) [5] aim to replace fully connected layers with more efficient alternatives. These techniques reduce parameter sizes

and computational costs, aligning with our goal of exploring CLOS networks as a potential substitute for fully connected layers.

Application of CLOS Networks to Machine Learning. While still an emerging area, exploring CLOS networks for fast inference presents a promising future for machine learning. This approach showcases how traditional network architectures can be adapted to reduce communication overhead, providing a clear path for applying CLOS networks to improve memory efficiency in neural networks.

Memory-Efficient Deep Learning Methods. Papers like "Training Deep Nets with Sublinear Memory Cost" (2016) by Chen et al. [6] and dynamic neural networks such as CondenseNet (Huang et al., 2018) [7] focus on reducing memory usage and dynamically learning which connections are necessary. These methods complement our goal by providing additional strategies for memory-efficient architectures.

Structured Matrices for Efficient Neural Networks. Sindhwani et al.'s "Structured Transforms for Small-Footprint Deep Learning" (2015) [8] introduces structured matrix approximations, such as circulant or Toeplitz matrices, to reduce a neural network's parameter count. Similarly, Cheng et al.'s "An Exploration of Parameter Redundancy in Deep Networks with Circulant Projections" (2015) [9] explores structured matrices to compress fully connected layers, aligning with our application of CLOS networks.

Efficient Transformer Models. Vaswani et al.'s "Attention is All You Need" (2017) [10] introduces the Transformer architecture, which relies heavily on dense layers, making it a prime candidate for CLOS networks. Dehghani et al.'s "Universal Transformers" (2019) [11] improves efficiency through shared weights, a concept we extend with CLOS's multi-stage structure.

III. Proposed Method

We propose a novel implementation of the CLOS network architecture for neural networks, which we call the CLOS Layer. This approach aims to reduce computational complexity while maintaining model expressiveness.

A. CLOS Layer Architecture

The CLOS Layer is designed as a three-stage network, inspired by the CLOS network topology used in telecommunications. It consists of input, middle, and output switches, represented by weight matrices W_1, W_2, and W_3, respectively. Figure 1 illustrates the CLOS Layer architecture, where each stage processes data through interconnected switches, reducing the total number of parameters compared to a standard linear layer's $W \in \mathbb{R}^{N \times M}$.

B. Algorithm

The forward pass of the CLOS Layer is defined as follows:

1) **Input:** Tensor $x \in \mathbb{R}^{B \times C \times N}$, where B is batch size, C is number of channels (2 or 3), and N is input features
2) Reshape x into (b_{in}, b_1) blocks
3) Compute first stage: $x_1 = xW_1 + b_1$

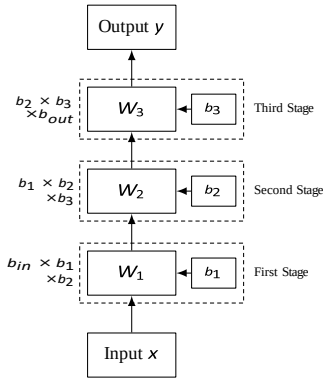


Fig. 1. CLOS Network Layer Architecture: A three-stage structure with input (W_1), middle (W_2), and output (W_3) switches, reducing parameters through structured interconnections.

4) Compute second stage: $x_2 = x_1 W_2 + b_2$
5) Compute third stage: $y = x_2 W_3 + b_3$
6) Reshape y to (B, C, M), where M is output features
7) **Output:** Tensor $y \in \mathbb{R}^{B \times C \times M}$

Where:

- $W_1 \in \mathbb{R}^{b_{in} \times b_1 \times b_2}$
- $W_2 \in \mathbb{R}^{b_1 \times b_2 \times b_3}$
- $W_3 \in \mathbb{R}^{b_2 \times b_3 \times b_{out}}$
- b_1, b_2, b_3 are bias vectors (if used)
- $b_{in}, b_1, b_2, b_3, b_{out}$ are switch sizes determined by factorizing input and output feature dimensions

The algorithm supports 2D and 3D inputs, adapting the tensor operations accordingly.

C. Complexity Analysis and Non-blocking

The CLOS Layer, inspired by the multi-stage CLOS network architecture, achieves both computational and parameter efficiency compared to standard linear layers. This section analyzes its time complexity and parameter efficiency, elucidating how the non-blocking property contributes to reducing parameter counts in neural networks.

The time complexity of the CLOS Layer is $O(N \cdot \max(b_{in}, b_1, b_2, b_3, b_{out}))$, where N is the input feature dimension, and $b_{in}, b_1, b_2, b_3, b_{out}$ are switch sizes determined by factorizing input and output dimensions. This is lower than the $O(N^2)$ complexity of a standard fully connected layer when switch sizes are chosen such that $\max(b_{in}, b_1, b_2, b_3, b_{out}) \ll N$. The reduced complexity arises from the CLOS Layer's decomposition of a large matrix multiplication into three smaller, sequential operations.

Parameter Efficiency. A standard linear layer with a weight matrix $W \in \mathbb{R}^{N \times M}$ requires $N \cdot M$ parameters. In contrast, the CLOS Layer employs three weight matrices: $W_1 \in \mathbb{R}^{b_{in} \times b_1 \times b_2}$, $W_2 \in \mathbb{R}^{b_1 \times b_2 \times b_3}$, and $W_3 \in \mathbb{R}^{b_2 \times b_3 \times b_{out}}$. The total parameter count for the CLOS Layer is given by:

$$\text{Parameters}_{\text{CLOS}} = (b_{in} \cdot b_1 \cdot b_2) + (b_1 \cdot b_2 \cdot b_3) + (b_2 \cdot b_3 \cdot b_{out})$$

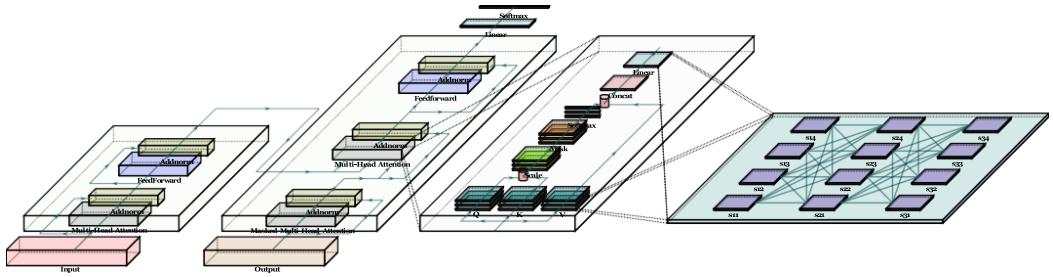


Fig. 2. CLOS Transformer Model Architecture: Integration of CLOS Layers into a Transformer framework, replacing traditional linear layers to achieve parameter efficiency.

By selecting appropriate switch sizes b_{in}, b_1, b_2, b_3, b_{out}, the CLOS Layer achieves a parameter count of $O(N \cdot k)$, where $k \approx \max(b_{in}, b_1, b_2, b_3, b_{out})$ and typically $k \ll M$. This is because the switch sizes are derived by factorizing N and M into smaller dimensions, ensuring that each stage operates on a subset of the data with reduced connectivity requirements.

For example, consider a layer with $N = 784$ and $M = 512$, as used in the first layer of our MNIST experiments (Section IV.A). A standard linear layer requires:

$$N \cdot M = 784 \cdot 512 = 401,408 \text{ parameters}$$

For the CLOS Layer, we select switch sizes $b_{in} = 28$, $b_1 = 28$, $b_2 = 16$, $b_3 = 16$, and $b_{out} = 32$, based on factorizing N and M to balance computational efficiency and parameter reduction (e.g., $b_{in} \cdot b_1 \approx \sqrt{N}$, $b_{out} \cdot b_3 \approx \sqrt{M}$). The CLOS Layer's parameter count is:

$$(b_{in} \cdot b_1 \cdot b_2) + (b_1 \cdot b_2 \cdot b_3) + (b_2 \cdot b_3 \cdot b_{out})$$

$$(28 \cdot 28 \cdot 16) + (28 \cdot 16 \cdot 16) + (16 \cdot 16 \cdot 32)$$

$$12,544 + 7,168 + 8,192 = 27,904 \text{ parameters}$$

This represents a parameter reduction of:

$$\frac{401,408 - 27,904}{401,408} \times 100 \approx 93.04\%$$

Thus, the CLOS Layer reduces the parameter count by approximately 93% compared to the linear layer. The switch sizes were chosen to approximate the square roots of the input and output dimensions (e.g., $\sqrt{784} \approx 28$, $\sqrt{512} \approx 22.6$), ensuring that each stage's matrix dimensions are significantly smaller than $N \times M$, while the multi-stage structure preserves the layer's expressiveness.

Non-blocking Property. CLOS networks, introduced by Charles Clos in 1953, are multistage switching networks designed for high-performance computing and telecommunications. Their three-stage architecture consists of:

Input stage
Middle stage
Output stage

A CLOS network is strictly non-blocking if the number of middle-stage switches (m) satisfies:

$$m \geq 2n - 1$$

where n is the number of input/output ports per switch in the input and output stages. This ensures that any unused input port can connect to any unused output port without conflicts, providing multiple paths for data routing. In the context of neural networks, the non-blocking property allows the CLOS Layer to efficiently route data through smaller weight matrices, reducing parameter redundancy. Unlike linear layers, which require dense connections for all input-output pairs, the CLOS Layer leverages the multi-stage structure to perform equivalent transformations with fewer parameters. For instance, the middle stage's switch count (m) corresponds to the size of $b_2 \cdot b_3$, which constrains the parameter count while maintaining connectivity.

Comparison with Structured Matrices. To contextualize the CLOS Layer's efficiency, we compare it with structured matrix approaches, such as circulant matrices (Cheng et al., 2015) [9]. A circulant matrix reduces parameters to $O(N)$ by enforcing a cyclic structure, but this can limit expressiveness for complex models. In contrast, the CLOS Layer's parameter count is $O(N \cdot k)$, where k depends on switch sizes. While circulant matrices may require fewer parameters for small N, the CLOS Layer's multi-stage design offers greater flexibility for large-scale models like transformers, as switch sizes can be tuned to balance efficiency and performance. Table I illustrates this comparison for a layer with $N = 784$ and $M = 512$.

TABLE I
PARAMETER COUNT COMPARISON

Method	Parameters	Example ($N = 784$, $M = 512$)
Linear	$N \cdot M$	401,408
Circulant	N	784
CLOS	$\sum (b_i \cdot b_j \cdot b_k)$	27,904

The CLOS Layer's parameter efficiency is further enhanced by the factor of 4 constraint (Section III.D), which ensures that switch sizes remain balanced, approximating $\sqrt{N \cdot M}$.

This constraint, derived empirically, aligns with common neural network expansion/reduction ratios, optimizing parameter reduction without compromising model stability.

In summary, the CLOS Layer's parameter efficiency arises from its multi-stage, non-blocking architecture, which decomposes large matrix operations into smaller, interconnected switches. The theoretical analysis demonstrates that the CLOS Layer significantly reduces parameters compared to linear layers, with tunable switch sizes offering flexibility over other structured matrix approaches.

D. Implementation Details

The CLOS Layer has been implemented as a custom PyTorch module, leveraging the efficiency of advanced tensor operations, particularly the `einsum` function for optimized matrix multiplications. This implementation autonomously determines the optimal switch dimensions based on the input and output feature spaces while allowing manual configuration when necessary.

Furthermore, we have developed a methodology to systematically replace all linear layers within a given model with their CLOS Layer counterparts, subject to a specific constraint. This substitution is executed only when the ratio of input size to output size (or vice versa) does not exceed a factor of 4 due to most model expansion and reduction parameters. This constraint can be formally expressed as:

$$\max\left(\frac{\text{input_size}}{\text{output_size}}, \frac{\text{output_size}}{\text{input_size}}\right) < 4 \tag{1}$$

The factor of 4 constraint was determined empirically through extensive experimentation. It ensures that the switch sizes remain balanced, optimizing parameter efficiency by maintaining factorization close to $N \approx M$. This value aligns with common expansion and reduction ratios in neural networks, ensuring stable performance across diverse architectures while maximizing parameter reduction.

This approach ensures that the CLOS Layer is applied judiciously, maintaining model efficiency while exploiting the benefits of the CLOS network architecture in neural network design.

In Figure 3, we illustrate the relationship between input/output features and parameter size, highlighting the significant reduction achieved by the CLOS Layer, as quantified in Section III.C.

IV. Experimental Setup and Results

We trained traditional linear layers and transformer-related networks, evaluating the CLOS Layer's performance on both simple and complex datasets. First, we targeted the MNIST dataset to assess accuracy and processing time in small custom models. Subsequently, we applied the CLOS Layer to deep transformer models such as Vision Transformer (ViT) on CIFAR10, analyzing its scalability. The following subsections discuss the training setups and results.

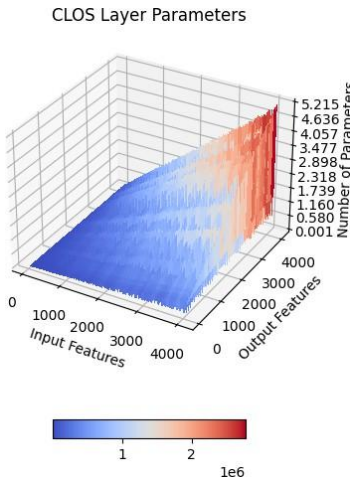


Fig. 3. Parameter Size of CLOS Network: Comparison of parameter counts across input and output dimensions, demonstrating up to 93% reduction compared to linear layers (see Section III.C).

A. Training Setup and Results on MNIST Dataset

We built two custom models with identical architectures, where both network weights were initialized from scratch and trained on CPU and GPU platforms. The MNIST dataset images were reshaped to $Batch \times 1 \times 784$ to feed directly into both traditional linear and CLOS layers. Both models had hidden units of $[784, 512, 256, 128]$ for the first two layers and an output layer of size 10. We used the SGD optimizer with a learning rate of 0.001, ReLU activation, cross-entropy loss, and trained for 10 epochs.

The CLOS Layer reduced the parameter count from 542,144 to 336,128, a 38% reduction. Table II shows that the CLOS Layer achieved an accuracy of 98.4% and a loss of 0.057, compared to the linear layer's 98.5% accuracy and 0.055 loss. However, processing times were 1.5x slower on GPU (2.25 vs. 1.50 minutes) and 3x slower on CPU (10.29 vs. 3.43 minutes).

TABLE II
MNIST Results: Parameters, Accuracy, Loss, and Processing Time

Model	Params	Acc. (%)	Loss	GPU (min)	CPU (min)
Linear	542,144	98.5	0.055	1.50	3.43
CLOS	336,128	98.4	0.057	2.25	10.29

Figures 4 and 5 illustrate the accuracy and loss curves for the CLOS and linear networks, respectively, confirming their similar convergence behavior despite the CLOS Layer's reduced parameter count.

B. Training Setup and Results on ViT Model

We implemented two parallel Vision Transformer (ViT) models with identical architectures, training both from scratch on CPU and GPU platforms. The CIFAR10 dataset was reshaped to $Batch \times 3 \times 224 \times 224$ to facilitate direct input into

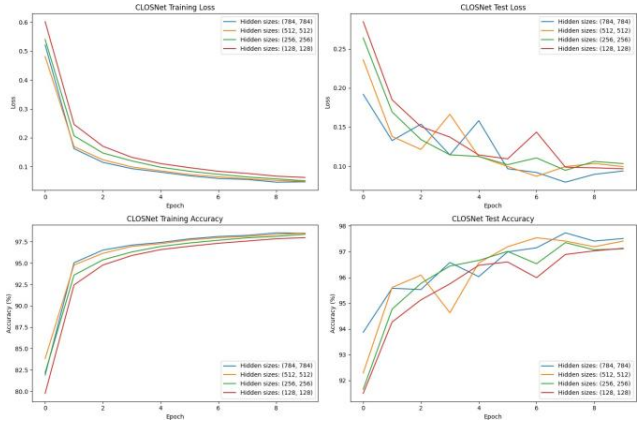


Fig. 4. Simulation Results for the CLOS Network on MNIST: Accuracy and loss curves over 10 epochs, showing comparable performance to the linear layer with 38% fewer parameters.

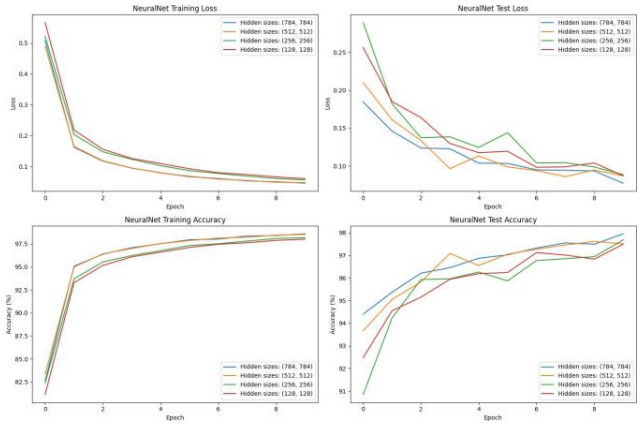


Fig. 5. Simulation Results for the Linear Neural Network on MNIST: Accuracy and loss curves over 10 epochs, serving as a baseline for comparison with the CLOS Layer.

both the traditional linear and CLOS Transformer attention layers.

The training process utilized the following hyperparameters:

- Batch size: 64
- Training epochs: 120
- Learning rate: 3×10^{-4}
- Learning rate decay (gamma): 0.7
- Random seed: 42
- Optimizer: Adam optimizer
- Loss function: Cross-entropy loss
- Learning rate scheduler: Step scheduler with step size 1

Performance evaluation was conducted on both computational platforms:

- **GPU Processing Time**: 01:37 minutes per epoch (Linear) vs. 02:54 minutes per epoch (CLOS)
- **CPU Processing Time**: 03:43 minutes per epoch (Linear) vs. 09:18 minutes per epoch (CLOS)

The CLOS Layer reduced the parameter count from 86.2M to 51.7M, a 40% reduction. Table III shows that the CLOS Layer achieved an accuracy of 91.2% and a loss of 0.312, compared to the linear layer's 91.5% accuracy and 0.308 loss, demonstrating comparable performance with significantly fewer parameters.

TABLE III
CIFAR10 RESULTS: PARAMETERS, ACCURACY, LOSS, AND PROCESSING TIME

Model	Params	Acc. (%)	Loss	GPU (min/epoch)	CPU (min/epoch)
Linear	86.2M	91.5	0.308	01:37	03:43
CLOS	51.7M	91.2	0.312	02:54	09:18

V. CONCLUSION

Our research demonstrates that CLOS networks can effectively serve as alternatives to traditional linear layers in neural network architectures. Through comprehensive experiments on both simple networks (MNIST dataset) and complex architectures like Vision Transformers (CIFAR10 dataset), we validated that the CLOS architecture maintains model performance while significantly reducing parameter counts. Key findings include:

- **Parameter Efficiency:** The CLOS Layer reduced parameters by 38% on MNIST (from 542,144 to 336,128) and by 40% on CIFAR10 (from 86.2M to 51.7M), making it suitable for memory-constrained environments.
- **Performance Metrics:** The CLOS Layer achieved comparable accuracy and loss to traditional linear layers (e.g., 98.4% accuracy and 0.057 loss on MNIST vs. 98.5% and 0.055 for linear layers; 91.2% accuracy and 0.312 loss on CIFAR10 vs. 91.5% and 0.308 for linear layers).
- **Performance Trade-off:** The improved parameter efficiency comes at the cost of increased computational overhead, with processing times 1.5x to 3x slower than conventional approaches.
- **Architecture Versatility:** The CLOS network structure can be successfully integrated into various neural network architectures, including transformers, demonstrating its flexibility.
- **Equivalent Convergence:** Both CLOS and traditional architectures demonstrate similar learning characteristics, confirming that parameter reduction does not compromise the network's ability to learn effectively.

Future work should focus on optimizing the computational efficiency of CLOS networks to address the processing time limitations while maintaining the parameter reduction benefits. Additionally, exploring hardware-specific implementations that could leverage the unique structure of CLOS networks may yield further improvements in both memory and computational efficiency. The CLOS network approach represents a promising direction for creating more parameter-efficient neural network architectures, particularly for deployment scenarios where model size is a critical constraint.

REFERENCES

[1] S. Han, H. Mao, and W. J. Dally, "Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding," *arXiv preprint arXiv:1510.00149*, 2015.

[2] B. Jacob *et al.*, "Quantization and Training of Neural Networks for Efficient Integer-Arithmetic-Only Inference," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 2704–2713.

[3] C. Szegedy *et al.*, "Going Deeper with Convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2015, pp. 1–9.

[4] F. N. Iandola *et al.*, "SqueezeNet: AlexNet-level Accuracy with 50x Fewer Parameters and ¡0.5MB Model Size," *arXiv preprint arXiv:1602.07360*, 2016.

[5] M. Denil *et al.*, "Predicting Parameters in Deep Learning," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2013, pp. 2148–2156.

[6] T. Chen *et al.*, "Training Deep Nets with Sublinear Memory Cost," *arXiv preprint arXiv:1604.06174*, 2016.

[7] G. Huang *et al.*, "CondenseNet: An Efficient DenseNet Using Learned Group Convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 2752–2761.

[8] V. Sindhwani *et al.*, "Structured Transforms for Small-Footprint Deep Learning," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2015, pp. 3088–3096.

[9] Y. Cheng *et al.*, "An Exploration of Parameter Redundancy in Deep Networks with Circulant Projections," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015.

[10] A. Vaswani *et al.*, "Attention is All You Need," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2017, pp. 5998–6008.

[11] M. Dehghani *et al.*, "Universal Transformers," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2019.

Криптовалютын үнийн хэлбэлзлийг таамаглах олон моделын харьцуулсан судалгаа

Б.Оршихбаяр
Мэдээллийн технологийн салбар
ШУТИС МХТС
Монгол улс, Улаанбаатар
orshikhbayar@gmail.com

Эрдэнэбат Батцэцэг
Компьютерын ухааны салбар
ШУТИС МХТС
Монгол улс, Улаанбаатар
battsetseg_e@must.edu.mn

Хураангуй—Энэхүү судалгаагаар Bitcoin (BTC/USDT)-ийн үнийн хэлбэлзлийг 2020 оноос хойших Yahoo Finance-ийн түүхэн дата дээр суурилан олон төрлийн цагаан хайрцаг болон хар хайрцаг загваруудыг харьцуулан үнэлсэн болно. LSTM, GRU, ARIMA, Prophet, SARIMAX, Random Forest, XGBoost зэрэг нийт 8 төрлийн загварыг сонгон авч, MAE, RMSE, MAPE, R^2 зэрэг гүйцэтгэлийн үзүүлэлтүүдээр харьцуулалт хийлээ. Судалгааны үр дүнд SARIMAX загвар хамгийн бага MAPE (2.92%) үзүүлсэн бол GRU загвар хамгийн өндөр R^2 (0.94)-тэй гарсан.

Түлхүүр үг—криптовалют, үнийн таамаглал, LSTM, GRU, SARIMAX

I. Удиртгал

Криптовалютын зах зээл нь сүүлийн жилүүдэд эрчимтэй хөгжиж, санхүүгийн салбарын нэгэн томоохон бүрэлдэхүүн хэсэг болж хувирсан. Түүний үнийн хэлбэлзэл, хөрөнгийн урсгалд үзүүлж буй нөлөө нь зөвхөн хөрөнгө оруулагч төдийгүй судлаачдын анхаарлыг татаж байна. Үнийн хэлбэлзлийг урьдчилан таамаглах нь зах зээлийн хандлагыг тодорхойлоход тустай төдийгүй эрсдэлийг бууруулах гол арга хэрэгслийн нэг юм. Санхүүгийн зах зээл, тэр дундаа криптовальютын зах зээл нь өндөр хэлбэлзэлтэй, олон хүчин зүйлсээс хамааралтай тул уламжлалт статистик аргачлалаар таамаглахад хүндрэлтэй байдаг. Харин гүн сургалтын загварууд, ялангуяа, GRU зэрэг давталтат нейрон сүлжээний аргууд нь цаг хугацааны цуваа өгөгдлийн хамаарлыг үр дүнтэй тусгаж, илүү нарийвчлалтай таамаглал гаргах боломжийг бүрдүүлдэг тул практикт өргөнөөр ашиглагдаж байна.

II. Ижил төстэй судалгаа

1. AlimohaddiAlimohammadi болон Nosratbakhsh (2023)-ийн хийсэн судалгаагаар LSTM, GRU, ARIMA, Prophet, Orbit, SARIMAX, Random Forest, XGBoost зэрэг загваруудыг харьцуулж, Bitcoin (BTC)-ийн үнийн таамаглал дээр Prophet болон Orbit загварууд тогтвортой өндөр үр дүн үзүүлснийг онцолсон байна. Энэхүү судалгаагаар Prophet загвар нь Accuracy (71%), F1-score (73%) болон MAPE (0.40%) зэрэг үзүүлэлтүүдээр бусад загвараас илүү сайн гүйцэтгэлтэй байсан бол SARIMAX нь Accuracy (70%) өндөр боловч MAPE (5.48%) үзүүлэлтээр муу байжээ.

Уг судалгаанд мөн Random Forest, GRU, LSTM, XGBoost загварууд нь Accuracy болон F1-score үзүүлэлтээр харьцангуй сул байсан боловч үнийн таамаглалд ашиглагддаг өөр бусад хэмжүүрүүдээр (MAPE, SMAPE, MASE, MSLE) Prophet загвар хамгийн өндөр нарийвчлалтайг харуулсан байна.

2. Bouteska, A., Abedin, M. Z., Hajek, P., & Yuan, K. (2024). Cryptocurrency price forecasting – a comparative analysis of ensemble learning and deep learning methods. International Review of Financial Analysis, 92, 103055.Энэхүү судалгаанд криптовалютын (Bitcoin, Ethereum, Ripple, Litecoin) ханшийг таамаглахад ансамбль сургалтын болон гүн сургалтын хэд хэдэн загварыг харьцуулжээ
Үр дүнгээр GRU (хаалгатай давталтат нэгж), энгийн RNN болон LightGBM зэрэг загварууд нь бусад алгоритмууд болон санамсаргүй худалдан авах/борих зэрэг суурь стратегиудаас илүү өндөр нарийвчлалтай байгааг тогтоосон байна
Энэ нь криптовалютын ханшийн хэлбэлзлийг таамаглахад гүн сургалтын аргууд, ялангуяа GRU загварын үр дүнтэй байдал өндөр болохыг харуулжээ.

3. Katina, J., Katin, I., & Komarova, V. (2024). Cryptocurrency price forecasting: a comparative analysis of autoregressive and recurrent neural network models. Entrepreneurship and Sustainability Issues, 11(4), 425–436. Katina Bitcoin, Ethereum, Dogecoin, Polygon, Toncoin зэрэг хэд хэдэн криптовалютын үнийн өгөгдөлд уламжлалт ARIMA, улирлын ARIMA (SARIMA) загваруудыг гүн сургалтын LSTM, GRU нейрон сүлжээний загваруудтай харьцуулсан байна.
Загвар бүрийн параметрийг тухайн өгөгдөлд тохируулж, гүйцэтгэлийг MSE, MAE, RMSE, MAPE зэрэг алдааны хэмжигдэхүүнээр үнэлэхэд LSTM, GRU зэрэг дүрсэлсэн дараалын нейрон сүлжээний загварууд нь криптовалютын үнийн хэв маягийг илүү сайн тусган, уламжлалт ARIMA төрлийн аргуудыг бодвол нарийвчлал өндөртэй таамаглал гаргасан байна. Үүний зэрэгцээ цагийн цувааны хэлбэлзэл, үл нийцэх шинж чанарт гүн сургалтын аргууд илүү тохирч буйг дүгнэжээ.

4. Rodrigues, F., & Machado, M. (2023). High-frequency cryptocurrency price forecasting using machine learning models: A comparative study. Information, 16(4), 300.
Энэ ажилд минутын түвшний өндөр давтамжийн өгөгдлөөр олон төрлийн криптовалютын (Bitcoin,

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Ethereum гэх мэт 10 нэр төрөл) ханшийг 60 минутын зайтай урьдчилан таамаглах загваруудыг харьцуулсан байна. Уламжлалт цагийн цувааны ARIMA (SARIMA оролцуулаад) загвар, шугаман регресс, санамсаргүй ой зэрэг машин сургалтын аргууд болон гүн сургалтын LSTM, GRU нейрон сүлжээний гүйцэтгэлийг харьцуулахад GRU сүлжээ нь хамгийн бага алдаатай, өндөр нарийвчлалтай таамаглал гаргасан (жишээлбэл, Bitcoin-ийн хувьд MAPE ~0.09% байсан) нь тогтоогджээ. Ингэснээр богино хугацааны (цаг, минутын) ханшийн хэлбэлзлийг загварчлахад GRU нь уламжлалт загварууд болон бусад машин сургалтын аргуудыг давуу байдлаар гүйцэтгэсэн байна.

5. Seabe, P. L., Moutsinga, C. R. B., & Pindza, E. (2023). Forecasting cryptocurrency prices using LSTM, GRU, and bi-directional LSTM: A deep learning approach. Fractal and Fractional, 7(2), 203.
Энэ судалгаанд зах зээлд хамгийн өргөн хэрэглээтэй гурван криптовалют болох Bitcoin, Ethereum, Litecoin-ий ханшийг урьдчилан тооцоолоход зориулагдсан гурван төрлийн давталтат нейрон сүлжээний (RNN) загваруудыг туршжээ: энгийн LSTM, GRU, болон хоёр чиглэлт LSTM (Bi-LSTM). Туршилтын үр дүнгээр хоёр чиглэлт LSTM загвар нь гурван валютын аль бүгд дээр хамгийн өндөр нарийвчлалтай байсан бөгөөд LSTM, GRU хоёуланг давж гарсан байна (жишээлбэл, Bi-LSTM загварын MAPE: Bitcoin-д 0.036, Litecoin-д 0.041, Ethereum-д 0.124 байв). Тус судалгаа нь цуваа өгөгдлийн хугацааны дараалын хоёул чиглэл дэх хамаарлыг ашигласан нейрон сүлжээ (Bi-LSTM) нь криптовалютын үнийн таамаглалд оновчтой болохыг харуулж байна.

6. Tripathy, N., Nayak, S. K., & Prusty, S. (2024). A comparative analysis of Silverkite and inter-dependent deep learning models for Bitcoin price prediction. Frontiers in Blockchain, 7, 1346410.
Tripathy нар нь Bitcoin-ийн үнийн таамаглалд зориулан дөрвөн төрлийн загварыг (LSTM, хоёр чиглэлт LSTM, LSTM-GRU холимог загвар, мөн Facebook-ийн Prophet) LinkedIn-ийн Greykite сангийн Silverkite алгоритмтай харьцуулан туршсан байна. 2012–2021 оны түүхэн өгөгдөл дээр хийсэн сорилтын үр дүнд хоёр чиглэлт LSTM (Bi-LSTM) загвар нь хамгийн бага алдаатай, өндөр нарийвчлалтай (RMSE ~0.815) байж, дараа нь LSTM-GRU нийлмэл загвар (RMSE ~0.917) орсон бөгөөд эдгээр нь хоёулаа Prophet болон Silverkite загваруудаас илт давуу байгааг харуулжээ. Энэхүү ажлаар гүн сургалтын дэвшилтэт аргууд, ялангуяа Bi-LSTM нь Bitcoin-ийн үнийн хэлбэлзлийг таамаглахад маш үр ашигтай болохыг нотолсон байна.

7. Angelo, M. D., Fadhiilrahman, I., & Purnama, Y. (2023). Comparative analysis of ARIMA and Prophet algorithms in Bitcoin price forecasting. Procedia Computer Science, 227, 490–499.
Энэхүү судалгаа нь 2019–2021 оны Биткойны ханшийн өдөр тутмын өгөгдөлд суурилан уламжлалт ARIMA цагийн цувааны загварыг Facebook-ийн хөгжүүлсэн Prophet загвартай харьцуулсан байна. Зохиогчид өгөгдлийг өдөр, долоо хоног, сарын чанартай хэсгүүдэд хуваж таамаглал хийж, хоёр загварын үр дүнг үнэлжээ. Үр дүнгээс үзэхэд Prophet загвар нь нийт гүйцэтгэлээрээ ARIMA-гаас илүү байсан ба таамаглалын тайлбарцал (R^2) нь Prophet-ийн хувьд ~0.94, ARIMA-ийн хувьд ~0.68 байснаар илэрхий давуу (өндөр) байв. Мөн Prophet нь бүх хугацааны интервал дээр илүү бага алдааны хэмжигдэхүүнүүдийг үзүүлсэн тул криптовалютын үнийн таамаглалд уламжлалт ARIMA-аас оновчтой болох нь тогтоогджээ.
8. Kaur, R., Uppal, M., Juneja, S., Singh, J., Gupta, S., & Alroobaea, R. (2024). Development of a cryptocurrency price prediction model: leveraging GRU and LSTM for Bitcoin, Litecoin and Ethereum. PeerJ Computer Science, 10, e2675.
Kaur болон хамтран судлаачид нь зах зээлийн үнэлгээгээр тэргүүлэгч гурван криптовалютын (Bitcoin, Ethereum, Litecoin) ханшийг таамаглахын тулд давталтат нейрон сүлжээний хоёр архитектурыг (LSTM ба GRU) ашиглан загвар боловсруулжээ. Загваруудыг түүхэн ханшийн өгөгдөл дээр сургаж, үнэн зөвийг нь MSE, MAE, RMSE, MAPE зэрэг үзүүлэлтээр хэмжихэд GRU дээр суурилсан загвар нь бүх валютын хувьд LSTM загвараас бага алдаатай, илүү өндөр нарийвчлалтай байсан (жишээлбэл, BTC ханшийн хувьд GRU загварын MAPE ~3.54%, LSTM загварынх ~5%-тай байв). Энэхүү ажлын дүгнэлтээр GRU төрлийн нейрон сүлжээ нь криптовалютын үнийн таамаглалд LSTM-ээс илүү үр дүнтэй болох нь батлагдсан байна.

9. Yamak, P. T., Yujian, L., & Gadosey, P. K. (2019). A comparison between ARIMA, LSTM, and GRU for time series forecasting. In Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence (ACAI) (pp. 1–8). ACM.
Энэ харьцуулсан судалгаанд Bitcoin-ийн үнэ цэнийг богино хугацаанд (өдөр тутмын түвшинд) таамаглахын тулд уламжлалт ARIMA загварыг гүн сургалтын LSTM, GRU сүлжээтэй жишин үнэлжээ. Сонирхолтой нь, энэхүү 2019 оны туршилтаар ARIMA загвар нь нейрон сүлжээний аргуудыг гүйцэтгэлээр бага зэрэг давж, хамгийн бага алдаатай таамаглал гаргасан (MAPE ~2.76%, RMSE ~302.5) болохыг тогтоосон байна. Харин LSTM ба GRU загваруудын дотроос GRU нь LSTM-ээс илүү сайн үр дүн үзүүлсэн (GRU-ийн MAPE ~3.97%, LSTM-ийн MAPE ~5% орчим) бөгөөд энэ нь зарим нөхцөлд уламжлалт цагийн цувааны аргууд ч криптовалютын ханшийг таамаглахад гүн сургалтын загвартай өрсөлдөхүйц байж болохыг харуулжээ.

III. Өгөгдөл

Судалгаанд ашигласан өгөгдөл нь BTC/USDT валютын хослолын өдрийн хаалтын үнэ бөгөөд Yahoo Finance платформоос татан авсан. Өгөгдлийн хүрээ нь 2022 оны 1-р сараас 2025 оны 1-р сар хүртэл байна.

Өгөгдлийг дараах байдлаар боловсруулсан:
Pandas ашиглан түүхий өгөгдлийг цэвэрлэж, огноо болон хаалтын үнэ (Close) багануудыг ялган авсан. Null утгуудыг шалгаж, шаардлагатай бол устгасан. Үнийн өгөгдлийг MinMaxScaler (0–1) ашиглан масштабжуулсан. 60 хоногийн үнэ дээр үндэслэн

дараагийн 1 өдрийн үнэ таамаглах supervised өгөгдөл (X, y) үүсгэсэн.
Өгөгдлийг 80:20 харьцаагаар сургалтын болон туршилтын багцад хуваасан:
Сургалтын багц: 1460 өдрийн дата
Туршилтын багц: 366 өдрийн дата

TABLE I. ӨГӨГДЛИЙН ЖИШЭЭ ХҮСНЭГТ

Date	BTC-USD-Open	BTC-USD_High	BTC-USD_Low	BTC-USD_Close	BTC- USD_Volume
2020-01-01	7194.89	7254.33	7174.94	7200.17	18565664997
2020-01-02	7202.55	7212.15	6935.27	6985.47	20802083465
2020-01-03	6984.43	7413.71	6914.97	7344.88	28111481032
2020-01-04	7345.37	7427.39	7309.51	7410.65	18444272175
2020-01-05	7410.45	7544.50	7400.53	7411.32	19725074095

ӨГӨГДЛИЙН АНХНЫ ХЭМЖЭЭ БА ЦЭВЭРЛЭГЭЭ

Судалгаанд ашигласан өгөгдөл нь 2022 оны 1-р сараас 2025 оны 4-р сарын хоорондох нийт 1,200 гаруй хоногийн түүхэн үнийн болон техникийн үзүүлэлтүүдийг агуулж байсан. Энэхүү өгөгдөлд дараах цэвэрлэгээ, урьдчилсан боловсруулалтыг хийсэн:

NaN (дүнгүй) болон Inf (хязгааргүй) утгуудыг шалгаж, нийт 48 мөрийн өгөгдлийг хассан. Эдгээр нь ихэвчлэн Bollinger Bands, Log Return, EMA зэрэг үзүүлэлтүүдийг тооцоолох эхний өдрүүдэд үүссэн хоосон мөрнүүд байсан.

Тасралтгүй цуваа шаарддаг GRU нейрон сүлжээний хувьд, 60 хоногийн үргэлжилсэн өгөгдөл шаарддаг тул тасралттай өгөгдөл бүхий цэгүүдийг автоматаар шүүж хассан.

Outlier буюу хэт өндөр/бага утгууд (топ 0.5%, доод 0.5%) шалгаж, ноцтой зөрүү үүсгэсэн 5 мөрийг үл тооцсон.

Эцсийн дүнд анхны 1,200 мөрөөс нийт 53 мөрийг хасч, 1147 мөр цэвэр, сургалтад бэлэн өгөгдөл үлдсэн. Энэ өгөгдөлд скейлинг (MinMaxScaler) болон 60 хоногийн дараалсан SEQ_LENGTH-тэй цуваа үүсгэж сургалт, шалгалтын датасет болгон хуваасан.

IV. СУДАЛГААНЫ АРГА ЗҮЙ

Судалгаанд Yahoo Finance-ээс татан авсан Bitcoin (BTC/USDT) криптовалютын 2020 оноос хойших өдрийн хаалтын үнийн дата ашигласан. Энэхүү өгөгдлийг ашиглан нийт 8 төрлийн таамаглалын загвар болох GRU, LSTM, ARIMA, Prophet, Orbit, Random Forest, SARIMAX болон XGBoost загваруудыг харьцуулан шалгасан.

Үнийн урьдчилсан таамаглалын чанарыг үнэлэхдээ Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), мөн R^2 (детерминацийн коэффициент) зэрэг үзүүлэлтүүдийг ашигласан. Эдгээр загварын сонголт нь машин сургалтын түгээмэл хэрэглэгддэг уламжлалт статистик загваруудаас эхлээд хамгийн сүүлийн үеийн гүн сургалтын (GRU, LSTM) аргуудыг хамарсан болно.

Загваруудыг яагаад сонгосон талаар:

Энэхүү судалгаанд сонгогдсон GRU, LSTM, ARIMA, Prophet, Orbit, Random Forest, SARIMAX, XGBoost загварууд нь цуваа өгөгдлийн таамаглал хийхэд өргөн хэрэглэгддэг, алгоритмын шинж чанар, гүйцэтгэлээр ялгаатай тул харьцуулсан судалгаа хийхэд оновчтой гэж үзсэн. Гүн сургалтын загварууд (GRU, LSTM) нь цаг хугацааны хамаарлыг барьж авах чадвартай бол, статистик загварууд (ARIMA, SARIMAX, Prophet) нь тайлбарлахад хялбар, улирлын болон чиг хандлагын бүрдэл хэсгийг сайн загварчилдаг. XGBoost, Random Forest зэрэг ensemble загварууд нь олонлог дээр суурилсан хүчтэй гүйцэтгэлтэй, хар хайрцаг төрлийн шийдэл юм.

Загваруудын онцлог ба тохиромжтой өгөгдөл:

GRU: Хэлбэлзэл ихтэй өгөгдөлд тохиромжтой, бага нөөц ашигладаг гүн сургалтын RNN бүтэц.
LSTM: Өнгөрсөн мэдээллийг урт хугацаагаар хадгалах чадвартай боловч GRU-гээс илүү төвөгтэй.
ARIMA/SARIMAX: Улирлын шинж, чиг хандлага, гадаад хувьсагчид дээр сайн. Богино хугацааны таамаглалд үр дүнтэй.
Prophet: Facebook-ийн боловсруулсан, тренд, улирал, баяр зэрэг оролтыг автоматаар загварчилдаг.
Orbit: Байесийн хандлагатай цаг хугацааны загвар.
XGBoost/Random Forest: Feature суурьтай, хар хайрцаг төрлийн алгоритм. Цаг хугацааны цуваа руу тохируулахад нэмэлт feature engineering шаарддаг.

Яагаад GRU, Prophet загварыг онцолсон бэ?

Судалгааны үр дүнгээс харахад GRU загвар нь хамгийн өндөр R^2 (0.94), бага MAE, RMSE үзүүлэлттэй гарсан бөгөөд үнийн огцом хэлбэлзлийг дагаж илүү нарийвчлалтай хариу гаргаж байсан. Харин Prophet нь энгийн, хурдан таамаг гаргах боломжтой, автомат hyperparameter тохиргоотой тул бодит хэрэглээнд нэвтрүүлэхэд тохиромжтой гэдгээрээ онцлог байлаа.

Улирлын шинж, баяр ёслол, хүчин зүйлсийг загвартаа оруулсан эсэх:
SARIMAX болон Prophet загваруудад улирлын шинж чанар, чиг хандлага, баярын өдрийн нөлөөллийг оруулах боломжтой байдаг. Судалгаанд Prophet загварт тухайн криптовалюттай холбоотой онцгой өдрүүдийн (holiday effects) нөлөөллийг тусгасан бол SARIMAX загварт улирлын бүрэлдэхүүн орсон. Бусад гүн сургалтын загваруудад энэ мэдээллийг шууд өгөгдөлд feature байдлаар нэмж оруулах шаардлагатай байдаг.

TABLE II. GRU ЗАГВАРЫН АРХИТЕКТУРЫН ХҮСНЭГТ

Давхаргын нэр	Тайлбар
Input Layer	Shape: [Time steps, Features]
GRU Layer 1	Units: 64, Activation: ReLU, Dropout: 0.2
GRU Layer 2	Units: 32, Activation: ReLU, Dropout: 0.1
Dense Output Layer	Units: 1, Activation: Linear

Prophet загварын архитектур:Prophet нь Facebook-ийн хөгжүүлсэн цаг хугацааны цуврал өгөгдөлд суурилсан таамаглалын статистик загвар бөгөөд дараах үндсэн бүрэлдэхүүн хэсгүүдтэй:
Тренд бүрэлдэхүүн (Trend): Хугацааны урт хугацааны чиг хандлагыг логистик өсөлтийн функцээр илэрхийлдэг.

Улирлын шинж (Seasonality):Улирлын давтамжтай хэлбэлзлийг Fourier цуваагаар загварчилдаг. Өдөр, долоо хоног, сар бүрийн давтамжийг тус тусдаа тусгаж болно.

Хувьсах шинж (Holiday effects):Тухайн криптовалюттай холбоотой онцгой өдрүүдийн нөлөөг тусгаж өгдөг. Prophet нь автомат hyperparameter тохируулгатай, хэрэглэгчийн оролцоогүйгээр өгөгдөлд зохицох параметрүүдийг өөрөө сонгодог. Энгийн, тайлбарлах боломжтой, хурдан таамаглал хийх боломжтой гэдгээрээ онцлогтой.

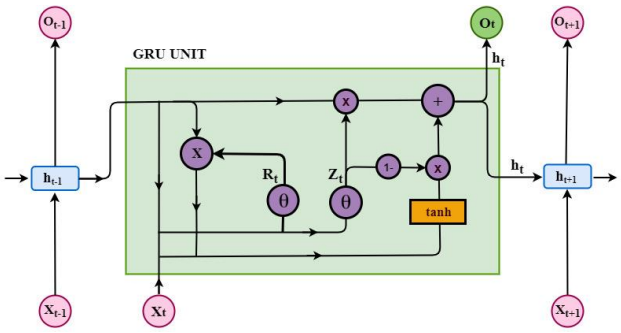


Fig. 1. Prophet загварын архитектур

V. ҮР ДҮН БА ХЭЛЭЛЦҮҮЛЭГ

Судалгааны ажилдаа сонгож авсан загваруудыг сургаж үнэлгээний 4 аргаар харьцуулж үнэлэхэд доорх үр дүн гарч байна.

Энэхүү судалгаанд машин сургалтын загваруудын таамаглалын үр дүнгийн чанарыг дараах дөрвөн гол үнэлгээний үзүүлэлтээр үнэлсэн. Эдгээр нь бүгд тоон утгатай бодит үнэ (actual) болон загвараар гарсан таамагласан утга (prediction) хоёрын зөрүү дээр суурилдаг:

1. Mean Absolute Error (MAE) – Дундаж алabsolute алдаа

Энэхүү үзүүлэлт нь бодит болон таамагласан утгуудын зөрүүгийн дундажийг абсолют (эерэг) хэлбэрээр авдаг.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Тайлбар: MAE нь бүх зөрүүг ижил жинтэй авч үздэг бөгөөд таамаглалын нийт дундаж алдааг энгийн байдлаар харуулдаг.

2. Root Mean Squared Error (RMSE) – Алдааны квадрат дундаж язгуур

RMSE нь зөрүүг квадрат болгон томруулж, дараа нь язгуур авснаар том зөрүүнүүдэд илүү их жин өгдөг үзүүлэлт юм.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Тайлбар: RMSE нь загвар том алдаа гаргах үед хурдан өсдөг тул хэт хэлбэлзэлтэй өгөгдөлд мэдрэмтгий байдаг.

3. Mean Absolute Percentage Error (MAPE) – Хувиар илэрхийлсэн дундаж алдаа

Таамаглалын алдааг бодит утгатай харьцуулан хувиар илэрхийлдэг үзүүлэлт.

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

Тайлбар: MAPE нь бүх алдааг бодит утгатай харьцуулдаг тул харьцангуй хэлбэрээр хэмждэг. Гэхдээ 0-д ойр утгад эмзэг.

4. R^2 Score (Детерминацийн коэффициент)

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Тайлбар: R^2 нь 1.0-д ойр байх тусам сайн гэж үзнэ. Энэ нь тухайн загвар нийт хэлбэлзлийн хэдэн хувийг зөв тайлбарлаж байгааг илэрхийлнэ.

TABLE III. ХАРЬЦУУЛСАН ҮНЭЛГЭЭНИЙ ҮР ДҮН

Model	MAE	RMSE	MAPE	R^2 Score
Gru	3330.52	4294.89	4.25%	0.9411
Lstm	6059.72	8879.25	13.35%	0.8561
Arima	4002.89	4401.24	3.95%	-0.3222
Prophet	2654.34	3304.48	7.44%	0.9792
Orbit	16670.16	23383	18.28%	-0.7632
Random_forest	13221.54	19422.83	14.29%	-0.2135
Sarimax	2876.05	3516.64	2.92%	0.1559
Xgboost	13739.15	20273.14	14.79%	-0.3221

Үр дүнгийн график:

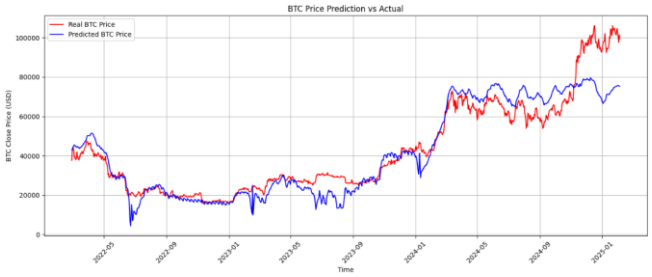


Fig. 2. LSTM загварын бодит болон таамагласан BTC үнийн харьцуулалт

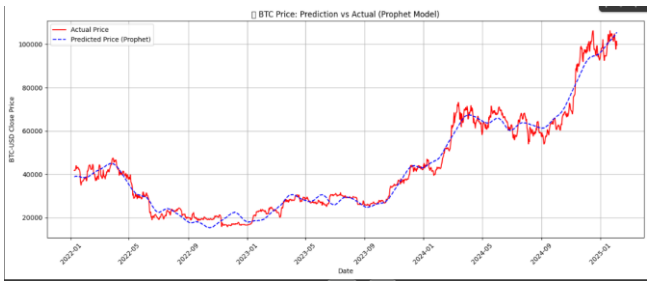


Fig. 3. Prophet загварын бодит болон таамагласан BTC үнийн харьцуулалт

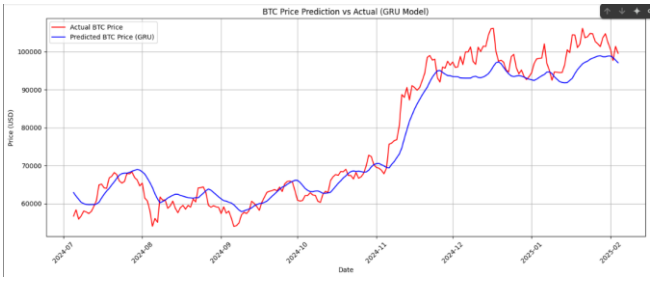


Fig. 4. GRU загварын бодит болон таамагласан BTC үнийн харьцуулалт

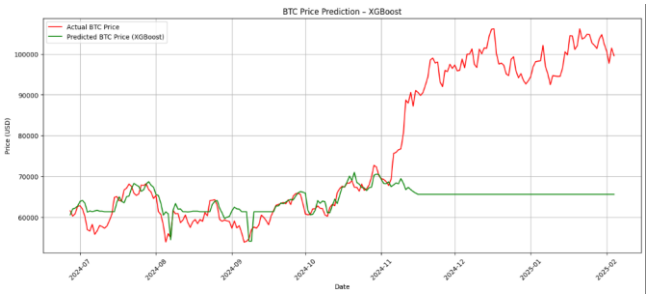


Fig. 5. XGBoost загварын бодит болон таамагласан BTC үнийн харьцуулалт

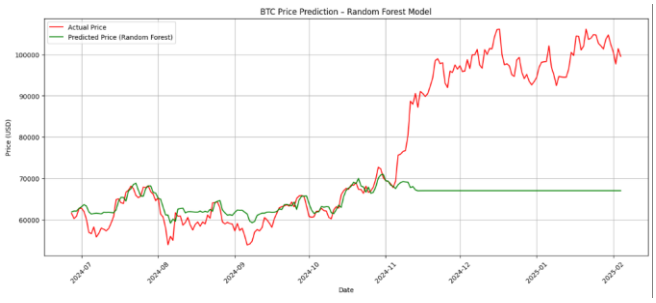


Fig. 6. Random Forest загварын бодит болон таамагласан BTC үнийн харьцуулалт

Үр дүнгээс харахад, Prophet загвар хамгийн өндөр R^2 (0.9792)-тай байсан бол, хамгийн бага MAPE үзүүлэлттэй (2.92%) загвараар SARIMAX шалгарав. Гүн сургалтын загваруудаас GRU-ийн үзүүлэлтүүд илүү сайн (R^2: 0.9411, MAPE: 4.25%) байсан нь LSTM (R^2: 0.8561, MAPE: 13.35%) загвараас мэдэгдэхүйц давуу байна. Уламжлалт статистик загваруудаас SARIMAX, Prophet илүү сайн үр дүн үзүүлсэн бол Orbit, Random Forest болон XGBoost загваруудын үзүүлэлт харьцангуй сул байлаа.

Үүнээс үзвэл, криптовалютын үнийн таамаглал хийхэд статистик загварууд тодорхой нарийвчлалтай байдаг ч GRU зэрэг гүн сургалтын аргууд нь цаг хугацааны динамик өөрчлөлтийг илүү сайн барьж авах чадвартай байгаа нь харагдаж байна.

Зураг 1-5-т LSTM, Prophet, GRU, XGBoost болон Random Forest загваруудын бодит болон таамагласан үнийн графикийг харуулав. Графикаас харахад Prophet болон GRU загварууд нь таамаглалын муруй нь бодит үнийн муруйтай илүү ойролцоо, тогтвортой байгааг харуулж байна.

VI. Дүгнэлт

Судалгаагаар олон төрлийн машин сургалтын болон статистик загваруудыг ашиглан криптовалютын ханшийн хэлбэлзлийг урьдчилан таамаглах боломжийг харьцуулан үнэлсэн. Үр дүнгээс харахад GRU загвар нь гүн сургалтын аргууд дундаас хамгийн өндөр R^2 болон бага алдаатай байсан бөгөөд огцом хэлбэлзэл бүхий өгөгдөлд илүү зохицож байв. Prophet загвар нь тайлбарлах боломжтой, хурдан тооцоологдох шинжээрээ хэрэглээнд тохиромжтой байлаа. SARIMAX загвар богино хугацааны нарийвчлалтай таамаглал хийхэд оновчтой байсныг баталсан. Цаашид эдгээр загваруудыг хослуулан ашиглаж, гадаад хүчин зүйлс, техникийн индикаторуудыг нэмж оруулснаар үнийн таамаглалын чанарыг илүү сайжруулах боломжтой юм.

TABLE IV. GRU МОДЕЛИЙН ГАРГАСАН ТОП 5 ТААМАГЛАЛ

Огноо	Жинхэнэ үнэ	Таамагласан үнэ
2024-03-31	94419.757812	91662.007812
2022-06-02	96886.875000	91868.820312
2022-05-25	98107.429688	92378.281250
2022-12-26	98236.226562	92974.335938
2022-12-01	16,967.13	16,969.77

2024 оны таамагласан үнийг бодит үнэтэй харьцуулах:
GRU загварын гаргасан 2024 оны таамаглалыг бодит BTC ханштай харьцуулахад 2024-03-31-ний жишээн дээр:
Жинхэнэ үнэ: 94419.75 USD
Таамагласан үнэ (GRU): 91662.00 USD
Харьцааны алдаа: ~2.93% орчим

Ном зүй

[1] Yahoo Finance, Bitcoin Historical Data, https://finance.yahoo.com

[2] https://github.com/Orshikhbayar/CryptoPredictions

[3] Activestate.com

[4] Researchgate.net

[5] facebook.github.io

[6] medium.com

[7] Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System.

[8] Brownlee, J. (2017). *Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python*. Machine Learning Mastery.

[9] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). *Dropout: A simple way to prevent neural networks from overfitting*. Journal of Machine Learning Research, 15(1), 1929–1958. https://jmlr.org/papers/v15/srivastava14a.html

[10] ideas.repec.org - Bouteska, A., Abedin, M. Z., Hajek, P., & Yuan, K. (2024). Cryptocurrency price forecasting – a comparative analysis of ensemble learning and deep learning methods. International Review of Financial Analysis, 92, 103055.

[11] ideas.repec.org - Katina, J., Katin, I., & Komarova, V. (2024). Cryptocurrency price forecasting: a comparative analysis of autoregressive and recurrent neural network models. *Entrepreneurship and Sustainability Issues*, *11*(4), 425–436.

[12] https://www.mdpi.com/2078-2489/16/4/300#:~:text=recurrent%20neural%20networks%20,MSE Rodrigues, F., & Machado, M. (2023). High-frequency cryptocurrency price forecasting using machine learning models: A comparative study. *Information*, *16*(4), 300.

[13] https://www.mdpi.com/2078-2489/16/4/300#:~:text=models%20like%20ARIMA,trading%20activity%2C%20thereby%20ensuring%20a Seabe, P. L., Moutsinga, C. R. B., & Pindza, E. (2023). Forecasting cryptocurrency prices using LSTM, GRU, and bi-directional LSTM: A deep learning approach. Fractal and Fractional, 7(2), 203.

[14] https://www.frontiersin.org/journals/blockchain/articles/10.3389/fbloc.2024.1346410/full- Tripathy, N., Nayak, S. K., & Prusty, S. (2024). A comparative analysis of Silverkite and inter-dependent deep learning models for Bitcoin price prediction. *Frontiers in Blockchain*, *7*, 1346410.

[15] https://www.scribd.com/document/820878591/1-s2-0-S1877050923017179-main#:~:text=Abstract%20Abstract%20The%20purpose%20of,and%20Prophet%20data%20between%20algorithms- Angelo, M. D., Fadhiilrahman, I., & Purnama, Y. (2023). Comparative analysis of ARIMA and Prophet algorithms in Bitcoin price forecasting. *Procedia Computer Science*, *227*, 490–499.

[16] https://www.researchgate.net/publication/389908963_Development_of_a_cryptocurrency_price_prediction_model_leveraging_GRU_and_LSTM_for_Bitcoin_Litecoin_and_Ethereum#:~:text=have%20explored%20cryptocurrency%20price%20prediction,The- Kaur, R., Uppal, M., Juneja, S., Singh, J., Gupta, S., & Alroobaea, R. (2024). Development of a cryptocurrency price prediction model: leveraging GRU and LSTM for Bitcoin, Litecoin and Ethereum. *PeerJ Computer Science*, *10*, e2675.

[17] uowojournals.org - Yamak, P. T., Yujian, L., & Gadosey, P. K. (2019). A comparison between ARIMA, LSTM, and GRU for time series forecasting. *In Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence (ACAI)* (pp. 1–8). ACM.

Хиймэл оюун ухаан ашиглан хорт URL илрүүлэх

П. Цэрэнпунцаг
Монгол Улс, Улаанбаатар хот, ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль,
Мэдээллийн Сүлжээ Аюулгүй Байдлын Салбар
purewdorjtserenpuntsag34@gmail.com

Ж. Эрдэнэсувд
Монгол Улс, Улаанбаатар хот, ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль,
Мэдээллийн Сүлжээ Аюулгүй Байдлын Салбар
erkajargalsaihn@gmail.com

А. Буяннэмэх
Монгол Улс, Улаанбаатар хот, ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль,
Мэдээллийн Сүлжээ Аюулгүй Байдлын Салбар
buynaane@gmail.com

Б. Мөнхбаяр (Ph.D)
Монгол Улс, Улаанбаатар хот, ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль,
Мэдээллийн Сүлжээ Аюулгүй Байдлын Салбар
munkhbayar.b@must .edu.mn

***Хураангуй* — Орчин үед цахим орчинд халдлагын тоо өсөн нэмэгдэж, кибер аюулгүй байдлын асуудал чухлаар тавигдаж байна. Хорт URL нь цахим халдлагын гол хэрэгсэл болж, хэрэглэгчдийн хувийн мэдээлэл алдагдах, системийн хэвийн үйл ажиллагаа доголдох эрсдэл үүсгэдэг. Энэхүү судалгааны ажил нь хиймэл оюун ухааныг ашиглан хорт URL илрүүлэх үр дүнтэй веб хуудас хийхэд зорьсон болно. Бид kaggle-аас болон өөрсдийн үүсгэсэн өгөгдлийн багц дээр туршилт хийсэн. Туршилтын үр дүнд AI модел 97% үзүүлэлттэйгээр ажилласан. Энэхүү судалгаа нь кибер аюулгүй байдлыг сайжруулахад хувь нэмэр оруулж, интернет хэрэглэгчдийг хамгаалахад чухал ач холбогдолтой юм. Цаашид энэхүү загварыг сайжруулж, бодит цагийн хорт URL илрүүлэх системд нэвтрүүлэх боломжтой.**

Түлхүүр үг— Фишинг, Кибер аюулгүй байдал, URL шинжлэх, Хиймэл оюун ухаан

I. Удиртгал

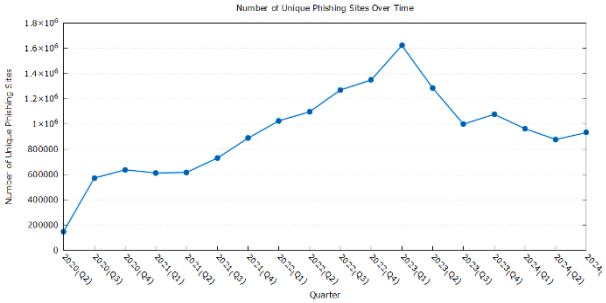
Сүүлийн жилүүдэд кибер аюулгүй байдлын асуудал дэлхий даяар хурцдаж, кибер халдлагын тоо эрс нэмэгдэж байна. Хортой вэб холбоосууд (URL) нь кибер гэмт хэрэгтнүүдийн гол зэвсгүүдийн нэг бөгөөд хэрэглэгчдийг хуурч, хортой програм татуулах, хувийн мэдээллийг хулгайлах, санхүүгийн хохирол учруулахад ашиглагддаг[1]. Тиймээс хортой URL-уудыг эрт илрүүлж, хэрэглэгчдийг хамгаалах үр дүнтэй аргуудыг хөгжүүлэх шаардлага улам бүр нэмэгдэж байна.

Хиймэл оюун ухаан (AI) болон машин сургалтын (ML) аргууд нь хорт URL-уудыг илрүүлэх өндөр нарийвчлалтай, үр дүнтэй шийдэл болж байна. Судалгаанаас үзэхэд уламжлалт дүрэмд суурилсан аргачлалууд нь шинэ төрлийн URL-уудыг илрүүлэхэд хангалтгүй байгаа бол AI суурилсан аргачлалууд нь их хэмжээний өгөгдлөөс үл мэдэгдэх хэв маягийг олж илрүүлэх чадвартай байдаг[2]. Тухайлбал, гүн сургалтын сүлжээнүүд (Deep Learning), логистик регресс (Logistic Regression), санамсаргүй ой (Random Forest) зэрэг алгоритмууд нь хорт URL илрүүлэхэд үр дүнтэй болох нь олон судалгаагаар батлагдсан [3].

Энэхүү судалгааны ажлын хүрээнд хиймэл оюун ухаан ашиглан хорт URL илрүүлэх веб системийг хөгжүүлэх бөгөөд уг систем нь бодит цагийн хяналт, өндөр нарийвчлалтай байх болно. . Системийн гүйцэтгэлийг нарийвчлал (Accuracy), оновчтой байдал (Precision), мэдрэг байдал (Recall) болон F1 үнэлгээ (F1-score) зэрэг гол шалгуур үзүүлэлтүүдээр хэмжиж, туршилтын үр дүнг үнэлэх болно. Энэхүү ажлын үр дүн нь кибер аюулгүй байдлын салбарт хиймэл оюун ухааныг ашиглах боломжийг өргөжүүлэхээс гадна хэрэглэгчдийн онлайн орчны аюулгүй байдлыг нэмэгдүүлэхэд чухал хувь нэмэр оруулах юм.

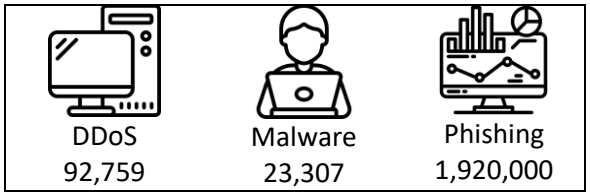
II. Сэдэв сонгосон үндэслэл

2024 оны 3-р улирлын байдлаар дэлхий даяар нийт 932,000 гаруй өвөрмөц фишинг вэбсайт илэрсэн бөгөөд өдөр бүр ойролцоогоор 3.4 тэрбум фишинг холбоос (URL) илгээгдэж байдаг [4][5][6] ба үүнийг зураг 1-т үзүүлэв.



Зураг 1: 2020 Q2-2024 Q3 оны байдлаар дэлхий даяар илэрсэн өвөрмөц фишинг сайтуудын тоо [4].

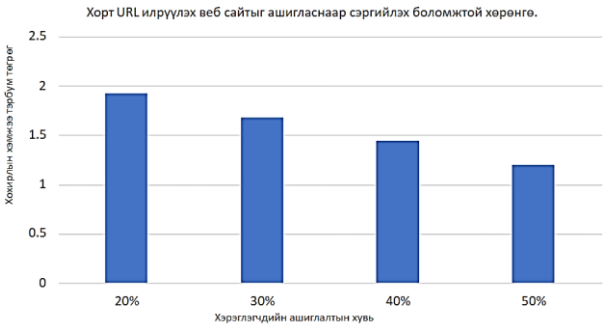
Монгол Улсын хувьд өдөрт дунджаар 40,000–50,000 кибер халдлага бүртгэгдэж байгаа бөгөөд үүний 1.9 сая орчим нь фишинг төрлийн халдлага эзэлж байна. Энэ нь Зүүн Өмнөд Азийн бүс нутгийн дунджаас тав дахин өндөр үзүүлэлттэй байгааг [7] зураг 2-т харуулав. Түүнчлэн, түгээмэл халдлагын төрлүүдийн дунд фишинг халдлага хамгийн өндөр тохиолдолтой байгаа нь хортой вэбсайт илрүүлэх, урьдчилан сэргийлэх технологийн хөгжлийн ач холбогдлыг улам нэмэгдүүлж байна.



Зураг 2: Монгол улсын халдлагын тоо [7].

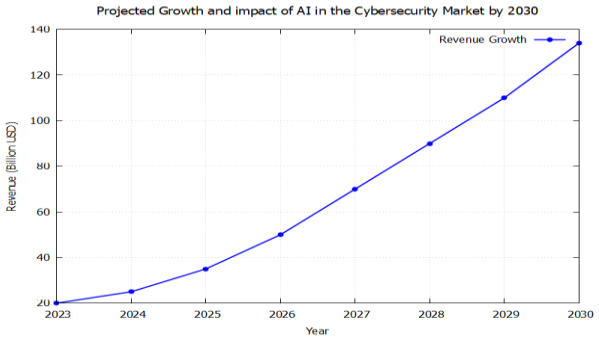
XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Фишинг халдлагын улмаас хохирсон иргэдийн тоо Монгол Улсад 2018 онд 659, 2019 онд 737, 2023 онд 8818, 2024 онд 7560 болж эрс нэмэгдсэн бөгөөд нийт хохирлын хэмжээ 2.4 тэрбум төгрөгт хүрээд байна [9]. Хэрэв иргэд хорт URL илрүүлэх веб системийг ашиглавал 1.2 тэрбум төгрөг хүртэлх хохирлоос урьдчилан сэргийлэх боломжтой болохыг зураг 3-т тооцоолсон байна.



Зураг 3: Сэргийлж чадах хохирлын хэмжээ.

Нөгөө талаас, кибер аюулгүй байдлын хиймэл оюун ухааны зах зээл 2023 онд 24.3 тэрбум ам.долларaар үнэлэгдсэн бөгөөд 2030 он гэхэд 134 тэрбум ам.долларт хүрэх төлөвтэй байгаа[8] зураг 4-өөс харж болох ба AI технологийн кибер аюулгүй байдалд гүйцэтгэх үүрэг, ач холбогдол тасралтгүй өсөн нэмэгдэж буйг харуулж байна.



Зураг 4: Кибер аюулгүй байдал дахь хиймэл оюун ухааны зах зээл [8].

Эдгээр хүчин зүйлсийг харгалзан үзэхэд, хортой вэб холбоосыг илрүүлэхэд хиймэл оюун ухааны боломжийг ашиглах нь цаг үеийн зайлшгүй шаардлагад нийцэж байна. Иймд энэхүү судалгаагаар AI суурилсан фишинг URL илрүүлэлтийн системийг хөгжүүлснээр гадаад болон Монгол Улсад фишинг халдлагаас урьдчилан сэргийлэх, иргэдийн хохирлыг бууруулах, улмаар кибер аюулгүй байдлын түвшинг дээшлүүлэхэд чухал хувь нэмэр оруулах зорилготой юм.

III. Өмнө судлагдсан ажил

A. The role of artificial intelligence in detecting and preventing cyber and phishing attacks[11]

Хиймэл оюун ухаан (ХОУ)-ын үүрэг кибер болон фишинг халдлагыг илрүүлэх, эсэргүүцэхэд чухал хэрэгсэл болж байна. Энэ судалгаа нь ХОУ-д суурилсан шийдлүүдийг ашиглан кибер аюулгүй байдлыг сайжруулах боломжийг судалж, машин сургалт (ML), байгалийн хэлний боловсруулалт (NLP), том хэлний загвар (LLM) аргуудыг ашиглан бодит цаг хугацаанд хортой үйлдэл болон URL-уудыг илрүүлэх боломжийг харуулсан. ХОУ нь шинэ аюулуудыг цаг хугацааны явцад суралцаж, илүү үр дүнтэй болж, хуурамч эерэг илрүүлэлтийг бууруулж, аюулгүй байдлыг сайжруулдаг. Мөн ХОУ-ийг хүний хяналттай хослуулан автомат хариу үйлдлүүд болон мэргэжлийн шинжилгээг ашиглан эрсдэлийг үр дүнтэй бууруулахыг онцолсон. Энэхүү судалгаа нь ХОУ-ийг кибер халдлагатай тэмцэх, мэдээллийг хамгаалах шинэ стратеги болгон ашиглах боломжийг харуулж байгаа бөгөөд энэ нь бидний төслийн зорилгыг гаргаж ирж байна.

B. AI-based Malware and Ransomware Detection Models[12]

Энэхүү судалгаа нь аюултай програм хангамж болон ransomware халдлагыг илрүүлэхэд зориулсан машин сургалт болон гүн сургалтын аргуудыг ашигласан. Манай төсөл нь фишинг URL-уудыг илрүүлэхэд зориулсан хиймэл оюун ухаан ашигладаг бөгөөд энэ судалгаанд ашигласан аргачлалууд нь фишинг төрлийн халдлагыг илрүүлэхэд хэрэглэгддэг. Тус судалгааны аргачлал нь хэрэглэгчдийн аюулгүй байдлыг хангахын тулд маш олон төрлийн аюултай URL-ууд болон програм хангамжийн аюулыг илрүүлэхэд тохиромжтой. Эдгээр аргачлалуудыг манай төслийн хэрэгжүүлэлттэй зарим хувилбар дээр ашиглаж болох бөгөөд энэ нь бидэнд үр дүнтэй хэрэгжүүлэлт хийх боломжийг олгож байна. Хэдийгээр бидний судалгаа болон энэ судалгаа нь өөр өөр чиглэлүүдэд чиглэсэн боловч судалгааны аргачлалууд нь хоорондоо төстэй бөгөөд бидний илрүүлэх аргачлалыг сайжруулахад олон санаа, зөвлөмжүүдийг санал болгож байгаа мөн эдгээр санаанууд нь бидний төслийг илүү үр дүнтэй, нарийвчлалтай болгоход туслах боломжтой.

C. Detecting phishing sites using ChatGPT[13]

Энэхүү судалгаа нь фишинг сайтуудыг илрүүлэхэд зориулсан ChatPhishDetector нэртэй системийг танилцуулсан бөгөөд вэб сайтуудаас мэдээлэл цуглуулан, ChatGPT зэрэг том хэлний загваруудыг (LLMs) ашиглан фишинг илрүүлэх шинэ аргачлалыг санал болгож байсан бөгөөд 98.7% - тай зөв ажилласан байна. Судалгааны хүрээнд том хэлний загваруудын олон талт боломжийг кибер аюулгүй байдалд үр дүнтэй ашиглаж болохыг харуулсан нь бидний төслийн чиг хандлагатай нийцэж байна. Манай төсөл мөн LLM технологийг ашигладаг боловч вэб сайт бүрэн задлан шинжлэхийн оронд зөвхөн хортой болон фишинг URL-уудыг илрүүлэхэд төвлөрдөгөөрөө ялгарах онцлогтой. Иймд тус судалгааны ажил нь бидний хэрэгжүүлж буй системийн хөгжүүлэлтэд арга зүйн болон техникийн чухал санаа, туршлагыг өгч байгаа юм.

IV. Онолын хэсэг

A. Phishing болон хорт URL

Фишинг халдлага гэдэг нь интернет хэрэглэгчдийг хуурч, тэдний хувийн мэдээлэл (нууц үг, санхүүгийн мэдээлэл, нэвтрэх эрх гэх мэт)-ийг

хулгайлах зорилготой цахим халдлага юм. Энэ төрлийн халдлагыг ихэвчлэн хуурамч имэйл, вебсайт, текст мессеж, онлайн сүлжээ ашиглан гүйцэтгэгддэг.

Фишинг халдлагын гол хэрэгсэл нь хортой URL (Malicious URL) бөгөөд энэ нь хэрэглэгчийг зохисгүй вебсайт руу чиглүүлж, тэдний мэдээллийг хулгайлах, эсвэл төхөөрөмжид нь хортой программ суулгах зорилготой байдаг.

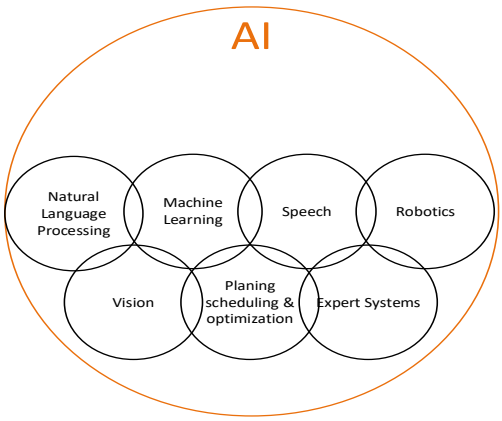
Үндсэндээ URL (Uniform Resource Locator) нь интернет дэх тодорхой нөөцийг олох өвөрмөц танигч юм. Энэ нь сервер дээр байрлах вэб хуудас, зураг, видео эсвэл аудио баримт бичиг болон бусад онлайн контентод хандах боломжийг бидэнд олгодог вэб хаягийн үүрэг гүйцэтгэдэг. Тодорхой байршлыг олоход тусалдаг физик хаягтай төстэй юм.[9]

B. Хиймэл оюун ухаан (AI)

Ярианы болон бичгийн хэлийг харж, ойлгох, хариу үйлдэл үзүүлэх, мэдээлэлд дүн шинжилгээ хийх, зөвлөмж гаргах гэх мэт хүний оюун ухаантай холбоотой танин мэдэхүйн үйл ажиллагааг дуурайх чадвартай машин, компьютер бүтээх технологийг ашиглахыг хэлдэг өргөн хүрээний салбар юм [10]. Хиймэл оюун ухаан бол систем бус системд хэрэгжсэн технологийн цогц юм үүнийг зураг 5-т харуулав.

1. Фишинг халдлагыг илрүүлэх (Phishing Detection) - д машин сургалтын алгоритмууд (SVM, Random Forest, Neural Networks), гүн сургалтын аргууд (CNN, RNN, Autoencoders), NLP (байгалийн хэл боловсруулалт) ашиглан имэйл, вэб контент шинжлэх зэрэг аргууд орно.

2. Хорт URL илрүүлэх (Malicious URL Detection)-д URL-ийн бүтэц, контент, хостинг, серверийн мэдээлэл дээр үндэслэн хортой эсэхийг ангилах болон AI загваруудыг ашиглан бодит болон хуурамч URL-ийг ялгах юм.



Зураг 5: Хиймэл оюун ухаан (AI) [7].

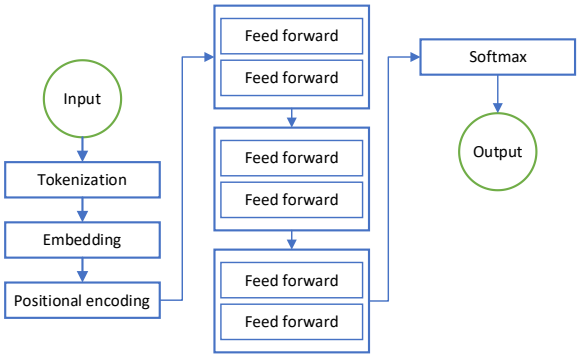
C. Том хэлний загвар (LLM)

LLM нь хүний хэл, текстийг боловсруулах, ойлгоход олон параметр бүхий мэдрэлийн сүлжээний техникийг ашигладаг хиймэл оюун ухааны алгоритмын нэг төрөл юм. Текст үүсгэх, машин орчуулга, хураангуй бичих, текстээс зураг үүсгэх, машин кодчилол, чат-ботууд эсвэл ярианы хиймэл оюун ухаан зэргийг LLM гүйцэтгэдэг[14].

LLM загварууд нь URL-ийн хэлбэр, домэйн, query string, obfuscation зэрэгт суурилсан сэжигтэй элементүүдийг таних боломжтой. Уламжлалт ML аргууд гар аргаар feature extraction хийх шаардлагатай байдаг бол LLM нь контекст ойлгож өөрөө гүйцэтгэж, сурж чаддаг[17].

Урьд нь бүртгэгдээгүй шинэ URL-ийг уламжлалт арга таних боломж бага. Харин LLM нь сэжигтэй байдал, утга, бүтэц дээр үндэслэн таамаглал дэвшүүлж, шинэ төрлийн халдлагыг илрүүлэх чадвартай болсон[18].

LLM-ийг өөрийн dataset дээр fine-tune хийж, эсвэл prompt инженерчлэл ашиглан шууд хариу гаргуулж болох ба энэ нь хөнгөн бөгөөд бодит хэрэглээнд нэвтрүүлэхэд тохиромжтой[19].

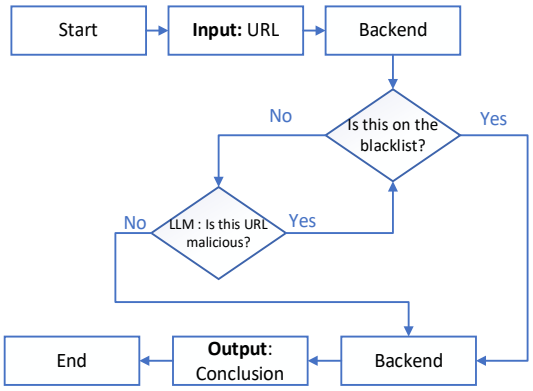


Зураг 6: Large Language Model (LLM) [14].

V. Санал болгох механизм

Том хэлний загвар (LLM) болох GPT модел нь байгалийн хэлний ойлголт, үүсгэлийн хувьсгалыг авчирсан. Эдгээр нь гүнзгий хэлний ойлголттой, хүнийхтэй төстэй текст үүсгэх чадвартай, нөхцөл байдлыг ойлгодог бөгөөд хүчирхэг асуудал шийдвэрлэх ур чадвартай тул олон салбарт (Жишээ нь: хайлтын систем, хэрэглэгчдийн үйлчилгээ, орчуулга) үнэлж баршгүй ач холбогдолтой болсон. Үүний зэрэгцээ, LLM нь аюулгүй байдлын салбарт анхаарал татаж, эмзэг байдлыг илрүүлэх, аюулгүй байдалтай холбоотой даалгавруудад ашиглах боломжтойг харуулсан [15].

Бидний судалгааны ажлын үр дүнд холбоосуудыг хортой эсвэл хоргүй гэж ангилах боломжтой веб хэрэгсэл хийхийг зорьсон бөгөөд гаралтын үр дүн нь “хортой” эсвэл “хоргүй” мөн яагаад тэгж шийдсэн гэж байх ба хортой байвал өөрсдийн үүсгэсэн хар жагсаалтын санд бүртгэж авна. Уг санал болгож буй механизмын ажиллах зарчмыг зураг 7-т үзүүлэв.



Зураг 7: Санал болгож буй механизм.

VI. Хэрэгжүүлэлт

A. *Өгөгдлийн багц /Dataset/*

Шаардлагатай өгөгдлийн багцыг бүрдүүлэхийн тулд бид kaggle [20] хуудас дээрх өгөгдлийн багцыг, мөн 500 ширхэг Монгол холбоос /URL/-ыг domains-monitor [21] болон github/Phish-Database- ээс [22] илэрсэн бүх .mn top level domain-тай холбоосуудыг цуглуулж баяжуулсан. Бид өгөгдлийг төрөлжсөн байдлаар сонгож цуглуулсан ба үүнд URL богиносгох үйлчилгээ ашигласан болон obfuscated гэх мэт төрлийн холбоосуудыг багтаасан.

B. *Шинж чанараар задлах /Feature Extraction/*

Бидний хийж буй судалгааны ажилд холбоос /URL/-г хортой эсвэл хоргүй гэдгийг ямар шинж чанарууд дээр нь үндэслэн таних вэ гэдэг чухал ойлголт.[16] Бид судалгааны ажилд дараах 3 бүлгийн шинж чанаруудыг тодорхойлж, LLM моделдоо эдгээр шинж чанарыг ашиглан холбоос /URL/-ын аюулгүй байдлыг үнэлэхийг зааварласан. Үүнд:

Үг зүйд суурилсан шинж чанарууд /Lexical based features/: Энэ бүлэгт холбоос /URL/-ын урт, домайн нэрийн урт, замын урт /path length/, тусгай тэмдэгтүүд агуулж байгаа эсэх гэх мэт холбоос /URL/-ын үг зүйд суурилсан буюу харагдах байдал талын шинж чанарууд орно.

Домайнд суурилсан шинж чанарууд /Domain based feature extraction/: Энэ бүлэгт тухайн холбоос /URL/ ын домайнд суурилсан шинж чанарууд буюу домайн нэрийн бичлэг /DNS record/, домайн нэрийн насжилт, домайн нэрийн статус гэх мэт шинж чанарууд багтана.

Контендэд суурилсан шинж чанарууд /Content based feature extraction/: Энэ бүлэгт тухайн холбоос /URL/ ын контент буюу агуулгад суурилсан шинж чанарууд багтана. Жишээлбэл өөр хуудас руу дахин чиглүүлж байгаа эсэх, javascript код ашиглан хулганы баруун товчыг идэвхгүй болгосон эсэх гэх мэт. Энэхүү бүлгийн шинж чанарууд тухайн холбоос /URL/ дээрх вэб кодыг бүрэн ачаалсаны дараа шинж чанаруудыг гаргаж авах тул бусад шинж чанаруудтай харьцуулахад бага зэргийн хүндрэлтэй.

Шинж Чанарын Бүлэг	**Шинж Чанар ба Тайлбар**
Үг зүйд суурилсан	Урт URL болон зам: хэт урт бол сэжигтэй
	Тусгай тэмдэгтүүд: @, %, -, гэх мэт
	IP хаяг ашигласан URL
	HTTPS протокол ашиглаагүй URL
	"login", "verify" гэх мэт сэжигтэй үгс
Домайнд суурилсан	Шинэ эсвэл сэжигтэй домайн насжилт
	DNS бүртгэл байхгүй эсвэл WHOIS мэдээлэл нууцлагдсан
Контентэд суурилсан	Дахин чиглүүлэлт хийдэг URL
	JavaScript-ээр хэрэглэгчийн үйлдэл хязгаарласан
	Нууцлагдсан iframe, хуурмаг формууд ашигласан

C. *Аргачлал*

Бидний хийсэн судалгааны ажлын илрүүлэлтийн хэсэгт Google компанийн Gemini 2.0 Flash загварыг ашиглан өмнө үүсгэсэн өгөгдлийн багцыг туршсан. Холбоос бүр дээр шинж чанарыг тодорхойлох зааварчилгаатай prompt ашиглан холбогдох URL-уудыг үнэлж, хортой эсвэл хоргүй гэж ангилахын зэрэгцээ, яагаад ингэж ангилсан шалтгааныг нэг өгүүлбэрээр тайлбарлуулах зарчмаар LLM загварыг удирдсан. Уг prompt нь URL-уудын үг зүйд суурилсан, домайнд суурилсан болон контентод суурилсан шинж чанарууд дээр тулгуурлан URL-ийн аюулгүй байдлыг үнэлэх зааварчилгааг багтаасан. Ашигласан prompt-ын Prompt Template 1-д харуулав.

Prompt Template 1
Analyze the following URL: {INPUT_URL} and determine whether it is 'benign' or 'malicious'. Base your judgment on established web security practices, taking into account lexical features (such as URL length, the presence of special characters, IP address usage, or suspicious words like "login", "verify", etc.), domain-based features (including the age of the domain, presence or absence of DNS records, and registration details), and content-based behaviors (such as redirections, embedded JavaScript that disables user actions, or hidden iFrames). Provide only a one-sentence explanation that highlights the most relevant reason for your classification.

VII. Үр дүн

Үр дүнг тооцоолохдоо доорх хүчин зүйлсийг тусгаж үзнэ.

- TP хортой холбоос /URL/-г зөв тодорхойлсон тоо
- TN хоргүй холбоос /URL/-г зөв тодорхойлсон тоо
- FP хоргүй холбоос /URL/-г буруу тодорхойлсон тоо
- FN хортой холбоос /URL/-г буруу тодорхойлсон тоо

Accuracy - Бүх зөв ангилсан тохиолдлын нийт өгөгдөлд харьцах хувь.

Precision - Загвар эерэг гэж ангилсан тохиолдлуудаас үнэн зөв байх магадлал.

Recall - Жинхэнэ эерэг бүх тохиолдлоос хэдийг нь зөв олж илрүүлсэн гэдгийг тодорхойлно.

F1 score – Precision ба Recall-ийн хоорондын тэнцвэрийг харуулдаг.

Гадаадын 1500 холбоос /URL/-г шалгаж, LLM моделийг ашиглан ангилалт хийсэн бөгөөд үр дүнг дараах хүснэгтүүдэд харуулав.

1-р хүснэгт. Гадаад 1500 холбоос /URL/-н ангилал

Confusion Matrix			
	Benign (Predicted)	Phishing (Predict)	Total (Actual)
Benign (Original)	1271 (TN)	23 (FP)	1294
Phishing (Original)	20 (FN)	186 (TP)	206
Total (Predicted)	1291	209	1500

2-р хүснэгт. Гадаад 1500 холбоос /URL/-г тодорхойлсон үр дүн

Score	
Accuracy	97.13%
Precision	89%
Recall	90.29%
F1 score	89.64%

Өгөгдлийн багцыг Монгол холбоос /URL/-уудаар өргөтгөж, нэмэлт 500 холбоосыг ангилсан үр дүнг дараах хүснэгтүүдэд үзүүлэв.

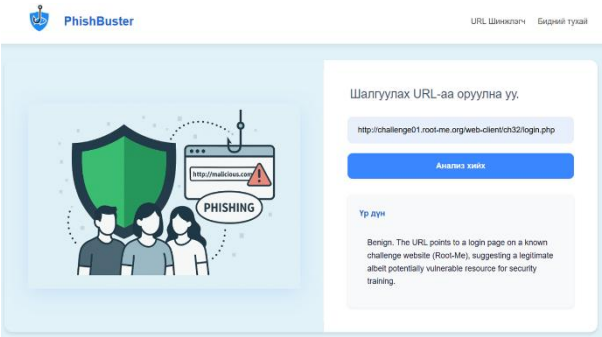
3-р хүснэгт. Монгол 500 холбоос /URL/-н ангилал

Confusion Matrix			
	Benign (Original)	Phishing (Original)	Total (Actual)
Benign (Original	413 (TN)	65 (FP)	478
Phishing (Original	3 (FN)	19 (TP)	22
Total (Actual)	416	84	500

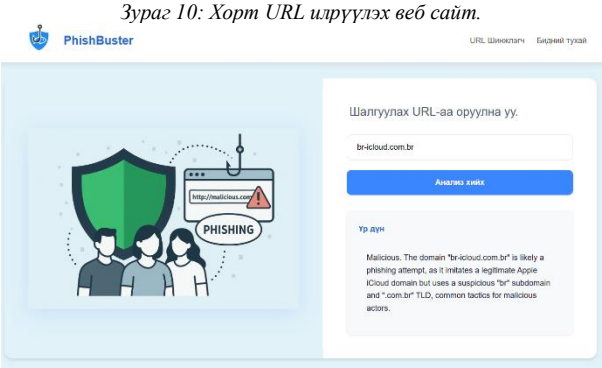
4-р хүснэгт. Монгол 500 холбоос /URL/-г тодорхойлсон үр дүн

Score	
Accuracy	86.40%
Precision	23%
Recall	86.36%
F1 score	36.10%

Бид хэрэглэгчдэд энгийн ойлгомжтой, орчин үеийн загвартай веб хуудсанд суурилсан хортой холбоос (URL)-ыг илрүүлэх системийг боловсруулсан. Энэхүү систем нь хэрэглэгчийн оруулсан холбоосыг шинжлэн, AI загварын тусламжтайгаар ангилж, аюултай эсэхийг тайлбар бүхий хариултаар буцаан харуулдаг. Системийн хэрэглээ, гүйцэтгэх зарчим болон интерфэйсийн ерөнхий бүтэц, үйл ажиллагааг зураг 10 ба зураг 11-д тусгав.



Зураг 10: Хорт URL илрүүлэх веб сайт.



Зураг 11: Хорт URL илрүүлэх веб сайт.(Илрүүлэхдээ тайлбар өгнө.)

ДҮГНЭЛТ

Энэхүү судалгааны үр дүнд, хиймэл оюун ухааны загвар нь хортой URL-уудыг өндөр үр дүнтэй илрүүлж чадсан бөгөөд гадаад холбоос /URL/-г 97.13%, монгол холбоос /URL/-г 86.405 нарийвчлалтай(accuracy) ажилласан нь загвар ихэнх URL-уудыг зөв ангилж байгааг харуулж байгаа юм. Precision (89%) болон Recall (90.29%) үзүүлэлтүүд нь загвар хортой URL-уудыг оновчтой илрүүлж, хэрэглэгчдийг цахим халдлагаас хамгаалах боломжтойг баталж байна.

Манай туршилтын үр дүнгээс харахад Precision болон F1 үзүүлэлт харьцангуй бага гарсан нь хуурамч эерэг (False Positive, FP) илрүүлэлт өндөр байсантай холбоотой байна. Үүний гол шалтгаануудын нэг нь богиносгосон холбоосуудыг автоматаар аюултай гэж ангилсан явдал бөгөөд мөн Монгол домэйн (.mn) бүхий, кирилл үсгийг латин галиг руу хөрвүүлсэн холбоосууд (жишээ нь: biirbeh.mn) нь англи үгийн сангийн шалгуураар зөв тодорхойлогдохгүй байсан тул FP тоо нэмэгдэхэд нөлөөлсөн. Үүний улмаас Precision болон F1 score нь accuracy онооноос харьцангүй бага гарсан.

Бусдаар загварын алдаа харьцангуй бага бөгөөд Confusion matrix-ээс харахад хортой болон энгийн URL-уудыг зөв ангилах чадвар өндөр байна. Цаашид бодит цагийн илрүүлэлтийг сайжруулах, худал эерэг (FP) болон худал сөрөг (FN) тохиолдлуудыг бууруулах чиглэлээр хөгжүүлэх боломжтой.

НОМ ЗҮЙ

[1] Sahoo, D., Liu, C., & Hoi, S. C. (2017). Malicious URL Detection using Machine Learning: A Survey. *arXiv preprint arXiv:1701.07179*.

[2] Ma, J., Saul, L. K., Savage, S., & Voelker, G. M. (2009). Beyond blacklists: Learning to detect malicious web sites from suspicious URLs. *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1245-1254.

[3] Zhang, X., Hu, C., & Wang, W. (2020). A Deep Learning-Based Approach for Detecting Malicious URLs. *IEEE Access, 8*, 45129-45140.

[4] Statista, “Number of phishing domain names worldwide.” [Online]. Available: https://www.statista.com/statistics/266155/number-of-phishing-domain-names-worldwide/. [Accessed: Mar. 23, 2025].

[5] AAG IT, “The latest phishing statistics.” [Online]. Available: https://aag-it.com/the-latest-phishing-statistics/. [Accessed: Mar. 23, 2025].

[6] Comparitech, “Phishing statistics & facts.” [Online]. Available: https://www.comparitech.com/blog/vpn-privacy/phishing-statistics-facts/. [Accessed: Mar. 23, 2025].

[7] Монгол Улсын Цахим Гэрчилгээний Үндэсний Төв, “PUBCERT — Монгол Улсын цахим гэрчилгээний үйлчилгээний төв.” [Online]. Available: https://www.pubcert.mn/mn/. [Accessed: Mar. 23, 2025].

[8] All About AI. (2024). AI Statistics in Cybersecurity. *All About AI Resources.*

[9] AKIpress, “Cyber crimes account for 26% of all crimes registered in Mongolia,” AKIpress News, Nov. 2023. [Online]. Available: https://m.akipress.com/news/. [Accessed: Mar. 23, 2025].

[10] D. J. Lemay, R. B. Basnet, and T. Doleck, “Examining the relationship between threat and coping assessment of phishing detection among college students,” Journal of Internet Services and Information Security, 2025.

[11] I. Naseer, "The role of artificial intelligence in detecting and preventing cyber and phishing attacks," European Journal of Advances in Engineering and Technology, vol. 11, no. 9, pp. 82–86, 2024. [Online]. Available: https://www.researchgate.net/publication/385077487_The_role_of_artificial_intelligence_in_detecting_and_preventing_cyber_and_phishing_attacks [Accessed: Mar. 24, 2025]

[12] B. Marais, T. Quertier, and S. Morucci, "AI-based Malware and Ransomware Detection Models," :2207.02108, 2022. [Online]. Available: https://arxiv.org/abs/2207.02108 [Accessed: Mar. 24, 2025].

[13] T. Koide, H. Nakano, and D. Chiba, "ChatPhishDetector: Detecting Phishing Sites Using Large Language Models," IEEE Access, vol. 12, pp. 154381–154400, 2024. [Online]. Available: https://arxiv.org/pdf/2306.05816 [Accessed: Mar. 24, 2025]

[14] A. Raj, “Large Language Models: Key Concepts & Examples,” Vitalflux, Mar. 21, 2024. [Online]. Available: https://vitalflux.com/large-language-models-concepts-examples/ [Accessed: Mar. 24, 2025]

[15] Yifan Yao, Jinhao Duan, Kaidi Xu, Yuanfang Cai, Zhibo Sun, Yue Zhang, A survey on large language model (LLM) security and privacy: The Good, The Bad, and The Ugly, High-Confidence Computing, Volume 4, Issue 2, 2024, 100211, ISSN 2667-2952, https://doi.org/10.1016/j.hcc.2024.100211. [Accessed: Mar. 24, 2025]

[16] D. Sahoo, C. Liu, S.C.H. Hoi, “Malicious URL Detection using Machine Learning: A Survey”. CoRR, abs/1701.07179, 2017 Japan, p. 301, 1982].

[17] M. Rao and M. Ali, “Detecting Malicious URLs using Transformer-based Language Models,” in Proc. IEEE Int. Conf. Big Data, Osaka, Japan, Dec. 2022, pp. 1142–1150. doi: 10.1109/BigData55660.2022.10020203.

[18] O. K. Sahingoz, E. Buber, O. Demir, and B. Diri, “Machine learning based phishing detection from URLs,” *Expert Systems with Applications*, vol. 117, pp. 345–357, Mar. 2019, doi: 10.1016/j.eswa.2018.09.029.

[19] Y. Chen, H. Jin, J. Zhang, and W. Wang, “Prompting GPT for Phishing Email Detection: A Zero-shot Learning Approach,” *arXiv preprint*, arXiv:2302.06892, 2023. [Online]. Available: https://arxiv.org/abs/2302.06892

[20] Kaggle, "Datasets for machine learning and data science," *Kaggle*. [Online]. Available: https://kaggle.com

[21] Domains Monitor, "Monitor domain registrations and detect phishing," *Domains Monitor*. [Online]. Available: https://domains-monitor.com/

[22] Phishing Database, "Open-source phishing links and threat intelligence," *GitHub*. [Online]. Available: https://github.com/Phishing-Database

Гүн сургалтанд ашиглагддаг өгөгдлийн форматуудын ашигт байдлын талаарх экспериментал судалгаа

Баттогтохын Доржнамжмаа
Мэдээлэл, компьютерын ухааны тэнхим
МТЭС, МУИС
21b1num0267@stud.num.edu.mn

Довдонгийн Энхзол*
Мэдээлэл, компьютерын ухааны тэнхим
МТЭС, МУИС
enkhzol@num.edu.mn

Хураангуй—Хиймэл оюун ухаанд суурилсан программ хангамжууд бидний амьдралын хэвшилд сүүлийн жилүүдэд ихээр нэвтэрч, ажлын бүтээмжийг өсгөхөд өргөн хэрэглэгдэж байна. Хиймэл оюуны дэд салбар болох Гүн сургалтын загварт тулгуурласан программын үр дүнгийн оновчтой байдлын нарийвчлал, боловсруулалтын хурд, санах ойн үр ашигтай хэрэглээнд тухайн загварыг сургасан болон, таамаглахад ашиглагдаж байгаа өгөгдлийн формат нь ихээхэн нөлөөлдөг. Энэ судалгаанд бид нийтлэг өгөгдлийн төрөл болох TFRecord, HDF5, JPEG, LMDB, TIFF, TAR, Pickle, ZIP, PNG, NumPy гэсэн арван төрлийн форматыг ашиглан 1001 ижил зургийн багц дээр туршилт хийж, объект илрүүлэх хугацаа болон илрүүлсэн объектын тоог харьцуулан шинжилснээ танилцуулав. Туршилтын үр дүнд илрүүлэлтийн хугацаа болон нийт илрүүлсэн объектын тоо формат бүрд ялгаатай гарсан. HDF5, TIFF, Pickle, NumPy форматууд ашиглахад гүн сургалтын Yolo загвар нь илүү олон объект илрүүлсэн бол TFRecord, JPEG, ZIP форматууд илрүүлэлтийн хурдны хувьд давуу талтай байв. Харин TAR, PNG форматууд илрүүлэлтийн хувьд удаан байсан нь туршилтаар ажиглагдсан. Улмаар өгөгдлийн форматыг зөв сонгох нь гүн сургалтын системийн гүйцэтгэлд хэрхэн нөлөөлдгийг харуулж, өгөгдлийн менежментийн оновчтой стратеги боловсруулах нь чухал ач холбогдолтой болохыг тодотгов.

Түлхүүр үг—өгөгдлийн формат, өгөгдлийн хадгалалт, обьект илрүүлэлт, гүн сургалт, YOLO

I. Удиртгал

Гүн сургалтын хэрэглээ нэмэгдэж байгаа нь том хэмжээний өгөгдлийг үр ашигтай хадгалах, боловсруулах асуудлыг бий болгож байна. Жишээ нь, ImageNet, COCO зэрэг dataset-үүд нь хэдэн зуун GB хэмжээтэй байдаг бөгөөд эдгээрийг хурдан унших, хадгалах формат сонгох нь системийн гүйцэтгэлд ихээхэн нөлөөлдөг [1,2]. Өмнөх судалгаанууд нь өгөгдлийн форматуудын санах ойн хэрэглээ [3], pipeline performance [4], distributed training дахь өгөгдлийн нөлөө [5] гэх мэт асуудлуудыг судалсан боловч, объект илрүүлэлтийн гүйцэтгэл болон шахалтын харьцааны уялдаа сайн судлагдаагүй байна. Энэ судалгаагаар бид 1001 зургийн өгөгдлийн багц дээр дээрх 10 өгөгдлийн форматыг туршиж, форматуудын хадгалах үр ашиг болон гүн сургалтад үзүүлэх нөлөөллийг хэмжих зорилготой.

II. Ижил төстэй судалгаа

Microscaling Data Formats for Deep Learning судалгааны ажил нь гүн сургалтын загваруудын тооцоолол болон хадгалах сангийн зардлыг бууруулах зорилготой MX (Microscaling) форматыг санал болгож, туршилтаар үнэлсэн байна. Судалгааны үр дүнд MX формат нь FP32-той харьцуулахад 4 дахин бага хадгалах сан ашиглаж, тооцооллыг 2-3 дахин хурдасгах боломжтой болохыг харуулсан бөгөөд загварын нарийвчлалд багахан хэмжээний алдагдалтай боловч хэрэглээнд нэвтрэхүйц гэж нотолсон. Wang нар [4] өгөгдлийн pipeline-ыг оновчлох талаар судалсан боловч, зөвхөн structured data (CSV, JSON) дээр төвлөрсөн. Zhang нар [5] HDF5 форматын санах ойн хэрэглээг судалсан ч, бодит deep learning workload дээр тест хийгээгүй. Бидний судалгаа эдгээрийг бодитоор үнэлж, Deep Learning дахь өгөгдлийн форматуудын гүйцэтгэлийг туршихад анхаарч байна.

III. Туршилтын хэсэг

A. Өгөгдлийн багц

Туршилтад 1001 зургийг car-object-detection dataset ээс сонгон авсан. Өгөгдлийн багц нь автомашиныг илрүүлэх, танихад зориулсан өгөгдөл агуулдаг. Энэ нь YOLO (You Only Look Once) зэрэг объект илрүүлэх алгоритмуудыг турших, сургахад тохиромжтой. Зургийг 640x640 pixel хэмжээтэй, RGB форматаар хадгалсан. Зургийн дундаж хэмжээ 95.7KB.

B. Туршилтын орчин

Туршилтыг дараах техник болон програм хангамж дээр хийсэн:

1) Техник хангамж:

- CPU: Intel(R) Core(TM) i5-1035G1 CPU @ 1.00GHz 1.19 GHz
- RAM: 8.00 GB (7.71 GB usable)
- SSD: Windows-SSD

Зураг 1. Өгөгдлийн багц

Хүснэгт I
Өгөгдлийн форматуудын харьцуулалт

Формат	Илрүүлсэн обьект (тоо)	Илрүүлсэн хугацаа (сек)	Хэмжээ (GB)
tfrecord	917	120.12	1.14
hdf5	1203	101.42	0.736
jpg	917	106.12	0.0976
lmdb	917	125.74	10.0
tiff	1230	114.07	1.14
tar	917	161.33	0.099
pickle	1230	121.71	1.14
zip	918	116.79	0.0976
png	1203	138.92	1.25
numpy	1203	113.37	0.735

2) Програм хангамж::

- OS: Windows
- Python 3.8.10
- PyTorch 1.9.0
- CUDA 11.3
- OpenCV 4.5.3

C. Гүн сургалтын загвар

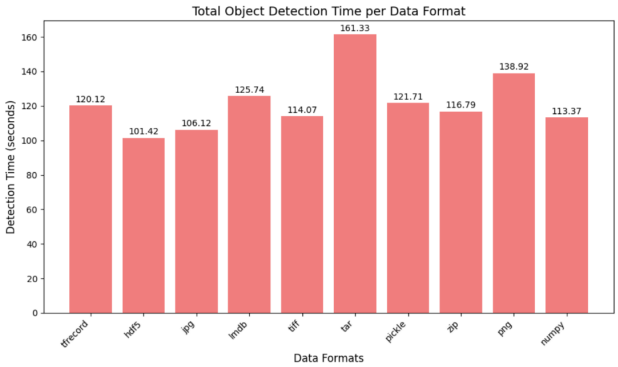
YOLO (You Only Look Once) загварын шинэ хувилбар болох YOLOv8 нь илүү нарийвчлалтай, хурдан, дотроо Nano (YOLOv8n), Small (YOLOv8s), Medium (YOLOv8m), Large (YOLOv8l), Extra Large (YOLOv8x) гэсэн таван үндсэн хувилбартай бөгөөд YOLOv8n(Nano) загвар нь хамгийн хөнгөн хувилбар юм. Энэхүү загварыг сонгох болсон гол шалтгаан нь хурдан ажиллагаа,бага хэмжээ,бага нөөц зарцуулалт, бодит цагийн ажиллах чадвар, хангалттай өндөр нарийвчлалтай байдал зэрэг хүчин зүйлүүд юм. Загварыг pretrained weights-тэй ашигласан бөгөөд нэмэлт сургалт хийгээгүй.

IV. Шинжилгээ

Өгөгдөл хадгалах форматын сонголт нь машин сургалтын загваруудын гүйцэтгэлд чухал нөлөө үзүүлдэг. Тухайлбал, том хэмжээтэй файлууд хадгалах хугацааг уртасгаж, илрүүлэлтийн хурдыг удаашруулдаг бол шахалт сайтай форматууд санах ойн багтаамжийг бууруулдаг. Энэхүү туршилтаар 1001 зургийг 10 төрлийн өөр форматаар хадгалж, тус бүрт обьект илрүүлсэн. Зургийг боловсруулах, обьект илрүүлэх хугацаа нь өгөгдлийн формат бүрийн хувьд ялгаатай байсан. Мөн ялгаатай форматуудаар хадгалагдсан ижил хэмжээний ижил өгөгдлөөс илрүүлж байгаа обьектын тоо формат бүрийн хувьд ялгаатай байна.

A. Обьект илрүүлэх хугацаа

Нийт 1001ширхэг зургаас бүх обьектыг илрүүлж дуусах хугацаа формат бүрт ялгаатай байна. Өгөгдлийн шахалт ба декомпрессийн хувьд tar, zip, pickle зэрэг форматууд нь өгөгдлийг шахсан байдалтай хадгалдаг. Илрүүлэх үйл явцын эхэнд эдгээр файлуудыг задлах хэрэгтэй болдог тул нийт хугацаа нэмэгдэх боломжтой.PNG нь lossless шахалт ашигладаг тул өгөгдлийг унших үед CPU болон санах ойд ачаалал өгч, удаашруулж болно. JPG нь lossy шахалттай тул өгөгдлийг илүү хурдан уншиж болох боловч чанарын алдагдал нь илрүүлэх үр дүнд нөлөөлөх магадлалтай. Файлын хандалтын хувьдHDF5, TFRecord, LMDB зэрэгформатууд нь өгөгдлийгсанах ойд оновчтой хадгалах,хурдан унших боломжийг олгодог. Энэ нь нийт илрүүлэх хугацааг бууруулах боломжтой.Харин PNG, JPG,TIFF зэрэг форматууд нь дан зураг бүрийг тусад нь хадгалдаг бөгөөд олон жижиг файлтай харьцах нь I/O үйл ажиллагааны саатлыг үүсгэж болно

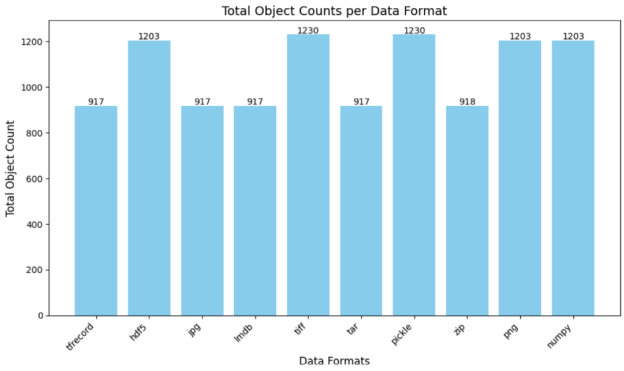


Зураг 2. Обьект илрүүлсэн хугацаа

B. Илрүүлсэн обьект

Ижил 1001 зураг агуулсан боловч өөрөөр форматаар хадгалсан үед илрүүлсэн обьектын тоо ялгаатай гарсан. JPG формат нь lossy compression(өгөгдлийг шахах үед мэдээлэл алдах) ашигладаг тул обьектийн тодорхой дүрслэл, өнгө, бүтэц өөрчлөгдөж болно. Энэ нь жижиг обьектуудыг илрүүлэх магадлалыг багасгаж, нийт илрүүлсэн обьектын тоонд нөлөөлж болно.PNG, TIFF, Numpy, HDF5, TFRecord зэрэгформатууд lossless com

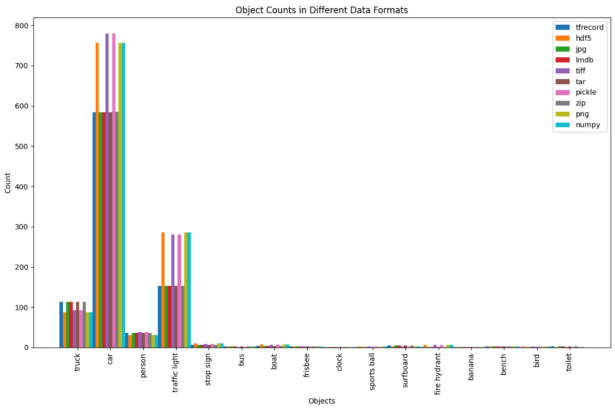
pression ашигладаг тул мэдээлэл алдагдахгүй, үүнээс шалтгаалж илүү олон обьект илрэх магадлалтай. TIFF, HDF5, Numpy зэрэг форматууд нь өндөр битийн нарийвчлалтай (16-bit, 32-bit гэх мэт) өгөгдлийг хадгалах чадвартай. Энэ нь гүн сүүдэр, өнгөний ялгарал сайтай байх боломжийг олгож, обьект илрүүлэлт илүү нарийвчлалтай байх магадлалыг нэмэгдүүлнэ. Харин JPG, TFRecord, ZIP гэх мэт форматууд 8-bit өнгөний нарийвчлалтай хадгалахад хязгаарлагдах тул жижиг, бүдэг обьектууд алдагдах магадлалтай. HDF5, TFRecord, Numpy форматууд нь мета өгөгдлийг (label, segmentation mask гэх мэт) хадгалах боломжтой бөгөөд энэ нь зарим алгоритмуудад нэмэлт давуу тал болж, илүү олон обьект илрүүлэхэд тусалж болно. ZIP, TAR, PICKLE зэрэг форматууд нь өгөгдлийг багцалж хадгалдаг тул зарим тохиолдолд мета өгөгдөл алдагдах эсвэл унших үед алдаа үүсч болно. Зарим форматууд (ZIP, TAR) нь нэг удаад бүх файлыг задлах шаардлагатай байдаг тул санах ойд оруулахад алдаа гарч эсвэл зарим зураг уншигдахгүй байх боломжтой. Энэ нь нийт илрүүлсэн обьектын тоог багасгах магадлалтай. TFRecord, HDF5, Numpy зэрэг форматууд нь алгоритмд шууд цуглуулгаар уншигдах хэлбэрээр хадгалагддаг тул өгөгдлийн бүрэн бүтэн байдал илүү найдвартай байна.



Зураг 3. Нийт илрүүлсэн обьектын тоо формат бүрийн хувьд

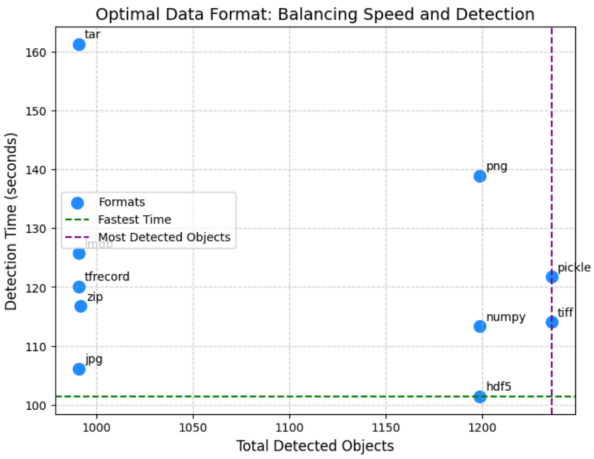
C. Илрүүлэлтийн тоо болон хугацааны хамаарал

- JPG (1000 обьект, 105 секунд) → Энэ нь маш хурдан боловч илрүүлэлтийн чанар дундаж байна.
- ZIP, TFRecord (1000-1020 обьект, 115-125 секунд) →JPG-тай ойролцоо хурдтай боловч удаан байна.
- TIFF, Numpy, Pickle (1200-1220 обьект, 120 135 секунд) → Эдгээр форматууд нь илүү нарийвчлалтай, гэхдээ илрүүлэх хугацаа бага зэрэг урт байна.
- PNG (1180 обьект, 140 секунд) → Нилээн олон обьект илрүүлсэн боловч хамгийн удаан хугацаа шаардаж байна.



Зураг 4. Обьектын ангилал бүрийн хувьд илрүүлэлт

- TAR (980 обьект, 160 секунд) → Энэ формат хамгийн бага обьект илрүүлсэн бөгөөд хамгийн удаан байна.



Зураг 5. Обьект илрүүлэлтийн тоо болон хугацаа

Нэг секундэд хэдэн обьект илрүүлж байгаа хэмжүүрийг оновчтой байдлын оноо гэж үзье.

$$\text{efficiency_scores} = \frac{\text{Total Detected Objects}}{\text{Detection Time (seconds)}} \quad (1)$$

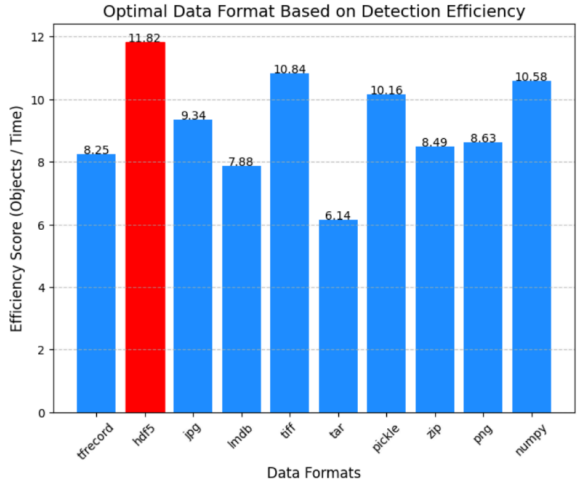
Их утгатай байх тусам өгөгдлийн формат илүү үр ашигтай гэж үзвэл HDF5 форматыг оновчтой гэж туршилт харуулж байна.

D. Хязгаарлалт

- Зөвхөн зургийн өгөгдөл дээр туршсан
- Нэг төрлийн object detection загвар ашигласан
- Тодорхой техник хангамж дээр хийгдсэн

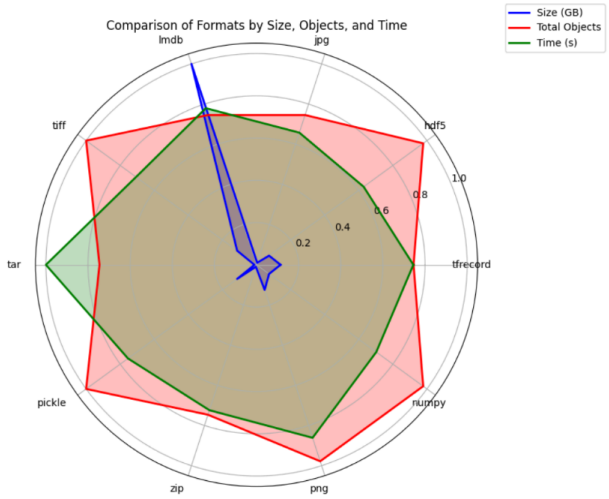
E. Цаашдын судалгаа

- Өөр төрлийн их өгөгдөл дээр туршилт хийх
- Distributed training системд харьцуулалт хийх
- Санах ойн хэрэглээг нарийвчлан судлах
- Өөр загваруудад туршилт хийх



Зураг 6. Өгөгдлийн форматын оновчтой байдал

Зураг 7. Өгөгдлийн хэмжээ, илрүүлэлсэн тоо, илрүүлсэн хугацаа



Зураг 8. Торон диаграм

V. Дүгнэлт

Энэхүү судалгаагаар гүн сургалтын хэрэглээнд ашиглагддаг өгөгдлийн хадгалалтын форматууд объект илрүүлэлтийн гүйцэтгэлд хэрхэн нөлөөлж байгааг системтэйгээр судаллаа. TFRecord, HDF5, JPEG, LMDB, TIFF, TAR, Pickle, ZIP, PNG, NumPy гэсэн 10 төрлийн форматыг ашиглан 1001 ижил зургаас илрүүлсэн объектын тоо, илрүүлэлтийн хурдыг харьцуулан шинжилсэн. Судалгааны үр дүнгээс харахад өгөгдлийн форматын сонголт нь илрүүлэлтийн нарийвчлал болон хурданд мэдэгдэхүйц нөлөөлж байв. HDF5, TIFF, Pickle, NumPy форматууд хамгийн олон объект илрүүлсэн бол TFRecord, JPEG, ZIP форматууд илрүүлэлтийн хурднаас давуу талтай байлаа. Харин TAR, PNG форматууд илрүүлэлтийн хувьд харьцангуй удаан байсан нь ажиглагдсан бөгөөд энэ нь өгөгдөл унших, шахалт задлах процессын сааталтай холбоотой байж болох юм. Форматуудын илрүүлэлтийн үр ашгийг тооцохын тулд Оновчтой байдлын оноо (Detection Efficiency Score) гэсэн хэмжүүрийг ашиглаж, илрүүлсэн нийт объектын тоог илрүүлэлтийн хугацаанд харьцуулан тооцоолсон. Үүний үр дүнд HDF5 формат нь илүү оновчтой үзүүлэлттэй гарч, илрүүлэлтийн нарийвчлал, хурдыг тэнцвэртэй хадгалсан хамгийн үр ашигтай сонголт болох нь тогтоогдов. Энэхүү судалгааны үр дүн нь өгөгдлийн форматыг зөв сонгох нь гүн сургалтын системийн гүйцэтгэлд хэрхэн нөлөөлдөгийг харуулж, өгөгдлийн менежментийн оновчтой стратеги боловсруулахад чухал ач холбогдолтой болохыг баталж байна. Цаашдын судалгаанд өгөгдлийн форматын гүйцэтгэлд нөлөөлөх техник хангамжийн онцлог, дата багцын нарийн төвөгтэй байдал, гүн сургалтын загваруудын бүтэц зэрэг хүчин зүйлсийг гүнзгий судлах шаардлагатай

Ном зүй

[1] UAVs for the Environmental Sciences
[2] Parallel Data Analysis Directly on Scientific File Formats
[3] Machine Learning and Deep Learning frameworks and libraries for large-scale data mining:a survey
[4] A survey on deep learning for big data
[5] Benchmarking the Linear Algebra Awareness of TensorFlow and PyTorch
[6] Deep Lake: a Lakehouse for Deep Learning
[7] Microscaling Data Formats for Deep Learning
[8] Comparative Study of Deep Learning Framework in HPC Environments
[9] A Review of Yolo Algorithm Developments
[10] Microscaling Data Formats for Deep Learning

Боловсролын Салбар Дахь Хиймэл Оюуны Ухааны Нөлөөллийн Судалгаа

1st Гүндсамбуу Болд
Компьютерын Ухааны салбар
Шинжлэх Ухаан Технологийн Их Сургууль
Улаанбаатар хот, Монгол улс
gundsambuu@must.edu.mn

2nd Удвалцэцэг Батжав
Мэдээллийн Технологийн салбар
Нэст Эдүкейшн АЙТ сургууль
Улаанбаатар хот, Монгол улс
udvaltsetseg@nhs.edu.mn

***Хураангуй*— Өнөө үед хиймэл оюун ухаан (AI) нь дэлхий даяар олон салбарт шинэчлэл авчирч буй томоохон технологиудын нэг бөгөөд боловсролын салбар ч энэ хувьслаас хоцрохгүй байна. 2023 оны байдлаар AI-ийн дэлхийн зах зээл 189 тэрбум ам.долларт хүрсэн бол 2033 он гэхэд 4.8 их наяд ам.долларт хүрч, огцом тэлэх төлөвтэй байна [1]. Энэ хурдацтай өсөлтөд боловсролын салбар чухал үүрэгтэй оролцох болно. AI нь сургалтын үйл явцыг илүү үр дүнтэй, хүртээмжтэй болгох, багш, суралцагчдын ачааллыг хөнгөлөх, хичээлийн агуулгыг хувийн хэрэгцээнд нийцүүлэх зэрэг олон боломжийг нээж байна. Гэвч давуу талуудын зэрэгцээ ёс зүй, өгөгдөл хамгаалалт зэрэг шинэ сорилтууд ч тулгарч байгаа билээ. Энэ өгүүлэлд бид хиймэл оюуны боловсрол дахь үүрэг, боломж, сорилт, Монголын нөхцөлд нийцсэн хэрэглээг судалж, цаашдын хөгжлийн чиг хандлагын талаар дүн шинжилгээ хийнэ.**

Түлхүүр үг—хиймэл оюун ухаан, боловсролын технологи, ухаалаг сургалтын систем

I. Удиртгал

Технологийн хурдацтай хөгжил бидний өдөр тутмын амьдрал, ажиллах арга барил, харилцааны хэлбэрийг эрс өөрчилж, бараг бүх салбарт хиймэл оюун ухаан нэвтэрч байна. Гэрээсээ хүссэн бараа бүтээгдэхүүнээ захиалах, ухаалаг утсаараа дэлхийн мэдээг агшин зуурт мэдэж, тэр ч байтугай өөрийгөө жолооддог автомашин ашиглах боломжтой болсон энэ цаг үед боловсролын салбар ч энэхүү дэвшлээс хоцрохгүй байна. ЮНЕСКО-гийн 2023 оны судалгаагаар дэлхий даяар боловсролын байгууллагын 72% нь хиймэл оюуныг сургалтын процесст туршиж эсвэл хэрэглэж эхэлсэн байна [2]. Хиймэл оюун ухаан боловсролын салбарт олон талын шинэчлэл, боломжуудыг авчирч байгаа бөгөөд түүний гол чиглэлүүдийг авч үзье.

A. Хувь хүнд чиглэсэн сургалт

Хиймэл оюун ухаан нь суралцагчийн мэдлэгийн түвшин, суралцах хурд, сонирхол, суралцах хэв маягийг тодорхойлсны үндсэн дээр тухайн суралцагчид тохирсон хичээлийн агуулгыг санал болгох боломжийг бүрдүүлдэг. Тухайлбал, Khan Academy платформ хиймэл оюун ашиглан суралцагчийн сул талыг тодорхойлж, түүнд тохирсон дасгал ажлуудыг автоматаар санал болгох функцтэй. Мөн хэлний сургалтад чиглэсэн Duolingo аппликейшн хувь хүний сонирхол, түвшинд нийцүүлэн хичээлээ боловсруулдгаараа олонд танигдаж, дэлхий даяар 500 сая хэрэглэгчтэй болсон байна. 2023 оны судалгаагаар ийм төрлийн систем ашигласан суралцагчдын сурлагын гүйцэтгэл дунджаар 35 хувиар нэмэгдсэн [3] нь хувь хүнд чиглэсэн сургалтын үр нөлөөг харуулж байна.

B. Багшийн ажлыг хөнгөвчлөх автоматжуулалт

Багш нар хиймэл оюун ухаан ашиглан хичээлийн төлөвлөгөө боловсруулах, бичвэр, тестийг автоматаар засах, суралцагчийн ахиц дэвшлийг шинжлэх боломжтой болж байна. Grammarly, Turnitin зэрэг платформууд хиймэл оюун ашиглан суралцагчийн бичсэн эссээг шалгаж, үг хэллэг, өгүүлбэрийн бүтэц, логик дараалал, найруулгын алдааг дүгнэх чадвартай. Мөн AI-тай тестийн системүүд ч суралцагчийн гүйцэтгэлийг автоматаар үнэлэн оноо гаргах боломжийг олгож байна. Үүнээс гадна GradeScope зэрэг систем шалгалтыг автоматаар засаж, багш нарын ажлын ачааллыг 70 хүртэлх хувиар бууруулсан тухай судалгаа гарсан байна [4]. Энэ нь багш нарт суралцагч бүрт илүү анхаарал хандуулж, зөвлөгөө өгөх боломжийг нэмэгдүүлж байна.

C. Виртуал туслах, чатботын дэмжлэг

Хиймэл оюун ухаанд суурилсан виртуал туслахууд нь суралцагчдад 24/7 тусламж үзүүлж, асуултад шуурхай хариулах, нэмэлт мэдээлэл өгөх боломжийг олгож байна. Тухайлбал, Испанийн Мурсиагийн Их Сургууль хиймэл оюунт чатбот нэвтрүүлснээр 38,708 асуултад 91 хувийн нарийвчлалтай хариулж, оюутны суралцах идэвхийг нэмэгдүүлжээ [4].

Khan Academy-ийн Khanmigo AI, ChatGPT, Egune, Socratic, Google Bolo зэрэг системүүд сурагчдад виртуал туслахын үүргийг гүйцэтгэж байна. Эдгээр хиймэл оюуны туслахууд нь суралцагчийн тавьсан асуултад хариулах, тодорхой сэдвийг ойлгоход нь дэмжлэг үзүүлэх, нэмэлт дасгал ажил санал болгох зэргээр олон талын туслалцаа үзүүлдэг. Энэ нь суралцагч бүр өөрийн хэмнэлээр, өөрт тохирсон хэлбэрээр хичээллэх боломжийг бүрдүүлж байна.

D. Хүртээмжтэй байдал

Хиймэл оюун ухаан нь алслагдсан бүс нутагт амьдарч буй суралцагчид болон гадаад хэлний мэдлэггүй, хараа, сонсголын бэрхшээлтэй хүмүүст боловсролыг хүртээмжтэй хүргэхэд чухал үүрэг гүйцэтгэж байна. Google Translate, Speech-to-Text AI зэрэг системүүд нь орчуулга хийх, аудио мэдээллийг текст болгох боломжийг олгож, хэлний бэрхшээлийг даван туулахад дэмжлэг үзүүлдэг. OpenAI Whisper, Чимэгэ бичээч зэрэг технологи нь аудио бичлэг, видеог текст болгон хөрвүүлж, сонсголын бэрхшээлтэй хүмүүст мэдээлэл авах нөхцөлийг бүрдүүлж байна. Харин Microsoft-ийн Seeing AI апп нь харааны бэрхшээлтэй суралцагчдад бичвэр, дүрсийг дуудан тайлбарлаж, тэдний суралцах үйл явцыг дэмжиж байна. Энэ мэт технологийн дэвшил нь боловсролыг хүн бүрт тэгш хүртээмжтэй хүргэхэд үнэтэй хувь нэмэр оруулж байна.

II. Хиймэл оюун ухааны боловсролын салбарт үзүүлэх давуу тал

Хиймэл оюун ухаан нэвтэрч буй салбар бүртээ ахиц дэвшил, давуу тал, өөрчлөлтийг авчирч байна. Боловсролын салбар ч үүнээс үл хоцорч, сургалтын чанар, хүртээмж, үр дүнд чухал хувь нэмэр оруулж байна. Юуны

өмнө хиймэл оюун ухаан нь багш нарын ажлын ачааллыг бууруулж, заах арга зүйг шинэ түвшинд хүргэхэд дэмжлэг үзүүлж байна. Тухайлбал, автомат үнэлгээ, шалгалт, бодлогын даалгавар засах системүүд нь багшийн цагийг хэмнэхээс гадна үнэлгээний алдааг бууруулж, нийт суралцагчдад ижил стандартаар оноо өгөх боломжийг бүрдүүлдэг. Мөн хиймэл оюунд суурилсан аналитик системүүд нь суралцагчийн гүйцэтгэлд дүн шинжилгээ хийж, анхаарал хандуулах шаардлагатай агуулгыг тодорхойлсноор багш нар тухайн суралцагчид чиглэсэн арга хэмжээг авах нөхцөл бүрдэнэ.

Үүнээс гадна онлайн болон алсын зайн сургалт хиймэл оюуны тусламжтайгаар илүү хүртээмжтэй болж, боловсролын тэгш бус байдлыг бууруулахад хувь нэмэр оруулж байна. Ялангуяа Монголын хөдөө орон нутагт хүн ам сийрэг суурьшсан байдлаас шалтгаалж үүсдэг боловсролын тэгш бус байдлыг шийдвэрлэхэд энэхүү технологи чухал үүрэг гүйцэтгэнэ. ЮНЕСКО-гийн зүгээс интернэт сүлжээнд холбогдох боломжгүй бүс нутгийн суралцагчдад зориулан Kolibri гэх оффлайн горимд ажилладаг боловсролын платформыг нэвтрүүлэх ажлыг дэмжиж байгаа бөгөөд уг платформ өнөөдрийн байдлаар дэлхийн 200 гаруй улс оронд 10 гаруй сая хүүхдэд хүрчээ [2].

Мөн хиймэл оюуны технологи нь гадаад хэл, программчлал, STEM хичээлүүдийг илүү сонирхолтой, үр дүнтэй сурах боломжийг бүрдүүлж байна. Тухайлбал, Duolingo, Khan Academy зэрэг платформууд нь хиймэл оюуны тусламжтайгаар суралцагчийн сонирхлыг татахуйц, идэвхтэй оролцоог дэмжсэн сургалтын орчин бүрдүүлснээр сая сая хэрэглэгчдэд боловсролыг хүртээмжтэй хүргэж байна [5].

III. Хиймэл оюуны боловсролын салбарт учруулж буй сорилт бэрхшээл

Хиймэл оюун ухаан боловсролын салбарт олон боломжийг нээж байгаа ч үүний зэрэгцээ тодорхой сорилт, бэрхшээлүүдийг ч дагуулж байна. Технологийн хурдацтай хөгжил, хүртээмжийн ялгаа, ёс зүйн болон мэдээллийн нууцлалтай холбоотой асуудлууд нь хиймэл оюуныг боловсролд үр дүнтэй, шударга ашиглахад саад болж байна.

A. Ёс зүй, оюуны өмчийн асуудал

Хиймэл оюуныг даалгавар, бие даалт, тайлан бичихэд зохисгүй ашиглах тохиолдол өсөж буй нь боловсролын ёс зүйд сөргөөр нөлөөлж байна. 2023 онд Turnitin компани хиймэл оюунаар бичигдсэн агуулгыг илрүүлэх шинэчлэлээ танилцуулж, 200 сая гаруй бичвэрийг шалгасан бөгөөд тэдгээрийн 11 хувь нь 20-иос дээш хувьтайгаар AI-аар бичигдсэн болохыг тогтоосон байна [6].

B. Мэдээллийн нууцлал, хамгаалалт

Суралцагчдын хувийн мэдээлэл, сурах хэв маяг, гүйцэтгэл зэрэг мэдээлийг AI системүүд цуглуулж, боловсруулж байдаг. Гэвч эдгээр мэдээлэл хэрхэн хадгалагдаж, хамгаалагдаж байгаагүй тодорхойгүй байх нь нууцлалын эрсдлийг үүсгэж байна. Жишээ нь, 2022 онд АНУ-ын нэгэн боловсролын платформ кибер халдлагад өртөж, 820 мянган хэрэглэгчийн мэдээлэл алдагдсан тохиолдол гарсан байна [7]

C. Хүртээмжийн ялгаа ба дэд бүтцийн хомсдол

Хиймэл оюуныг үр дүнтэй ашиглахад өндөр хурдны интернэт болон технологийн төхөөрөмж зайлшгүй шаардлагатай. Харамсалтай нь, хөгжиж буй орнууд, хөдөө орон нутгийн сургуулиудын хувьд энэ нь томоохон бэрхшээл хэвээр байна. ЮНЕСКО-гийн судалгаагаар дэлхий даяар сурагчдын 46% нь AI ашиглахад шаардлагатай интернэт болон төхөөрөмжгүй орчинд амьдардаг гэжээ [8]. Мөн Дэлхийн банкны 2023 оны тайланд хөгжиж буй орнуудын 60% нь сургуулийн орчинд интернэтийн хүртээмж хангалтгүй хэмээн дурдсан байна [9].

D. Багш нарын ур чадварын бэлтгэл хангалтгүй байдал

Хиймэл оюуны технологийг үр дүнтэй ашиглахад багш нарын мэдлэг, ур чадвар маш чухал. Гэсэн ч олон багш хиймэл оюуныг сургалтад ашиглах боломжтой гэж үзсэн ч хэрхэн зөв ашиглах талаарх арга зүй, сургалт дутагдалтай байна. EdTech Hub-ийн 2023 оны судалгаанд оролцсон багш нарын олонх нь хиймэл оюуны хэрэгслүүдийг ашиглах хэрэгцээг хүлээн зөвшөөрсөн ч, үр дүнтэй хэрэглээний талаарх хангалттай мэдээлэлгүй байгаагаа илэрхийлжээ [10].

Эдгээр сорилтыг төр засгийн бодлого, сургалт, хамтын ажиллагааны түвшинд шийдвэрлэх нь AI-г боловсролд үр дүнтэй нэвтрүүлэх үндэс болох юм.

IV. Монгол улсын боловсролын салбар дахь хиймэл оюун ухааны нэвтрэлт

Манай улсад үндэстний онцлогт тохирсон хиймэл оюун бүхий системүүд аль хэдийн хэрэглээнд нэвтрээд эхэлсэн. Тухайлбал, Болор софт компаниас хөгжүүлсэн ярианаас текст рүү буулгах, текстийг дуу бичлэг болгож хувиргах программууд зах зээлд гарч, хэрэглэгчдэд нээлттэй болсон. Хиймэл оюун ухааны технологи нь салбар бүрт шинэ боломжуудыг үүсгэж, боловсролын салбарт нэвтрүүлэх нь тэргүүлэх ач холбогдолтой. Гэвч үүнийг хэрэгжүүлэхэд тодорхой сорилт, бэрхшээлүүдийг даван туулах шаардлагатай байна.

Монгол Улсын өргөн уудам газар нутаг, хүн амын тархай бутархай байршил нь боловсролын тэгш хүртээмжийг хангахад хүндрэл учруулж байна. Хиймэл оюун ухааны технологи ашигласнаар алслагдсан бүс нутгийн суралцагчдад цахим сургалтад хамрагдах боломжийг олгож, боловсролын чанар, хүртээмжийг сайжруулах боломжтой. Гэхдээ хиймэл оюуныг үр дүнтэй ашиглахын тулд өндөр хурдны интернэт, ухаалаг төхөөрөмжүүд шаардлагатай. Тухайлбал, манай зарим бүс нутагт дэд бүтцийн хүртээмж хангалтгүй байгаа нь асуудал болж байна. Монгол улсын Боловсролын Яамны 2024 онд зохион байгуулсан ‘Боловсролын салбарын цахим хуваагдал’ уулзалтад нийт сургуулийн 77.5% нь интернэтэд холбогдсон [11] гэж мэдээлсэн. Гэвч холбогдсон сүлжээний хурд, тогтвортой байдал нь бүс нутгаас хамааран харилцан адилгүй байна. Иймд Засгийн газар бүх боловсролын байгууллагыг өндөр хурдны интернэтээр хангах, сурагч бүрийг ухаалаг төхөөрөмжөөр хангах зорилт тавьж, энэ чиглэлд ажиллаж байна. Хэрэв энэхүү төлөвлөгөө амжилттай хэрэгжвэл, оршин суугаа газраас үл хамааран хиймэл оюунд суурилсан системийг ашиглах боломж бүрдэнэ.

Хиймэл оюунд суурилсан сургалтын хөтөлбөр, виртуал туслахуудыг хөгжүүлснээр багш нарын ажлыг

хөнгөвчилж, сургалтын чанарыг дээшлүүлэх боломж бүрдэнэ. Боловсролын яам нь багш нарт хичээл төлөвлөх, үнэлгээ хийх, нэмэлт контент боловсруулахад туслах AI виртуал туслах хөгжүүлэх төслийг хэрэгжүүлж, үүнийг хариуцан ажиллах үндэсний багийг бүрдүүлэн ажлыг эхлүүлжээ [12]. Мөн хиймэл оюун ухааныг боловсролын системд нэвтрүүлэх хөтөлбөрт хамрагдахаар олон улсын Code.org байгууллагын төлөөлөлтэй уулзалт хийсэн байна [13]. Эдгээр төслийн үр дүн амжилттай хэрэгжвэл, хиймэл оюун ухааны нөлөө манай улсын боловсролын салбарт үр дүнтэй байх нь дамжиггүй. Нөгөө талаас хиймэл оюун ухааныг сургалтад үр дүнтэй ашиглахын тулд багш нарын мэдлэг, ур чадварыг дээшлүүлэх шаардлага тулгарч байна. Багш нарт зориулсан дижитал ур чадварын сургалтуудыг зохион байгуулж эхэлсэн ч, энэ чиглэлээр илүү өргөн хүрээтэй, тогтвортой сургалтууд хэрэгтэй байна.

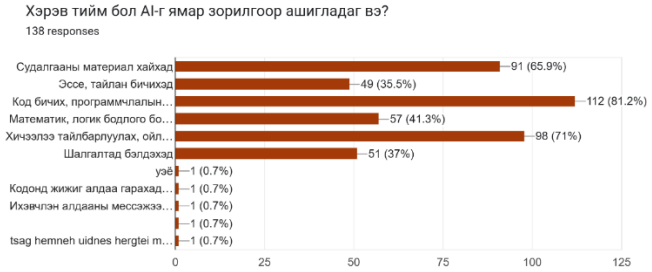
Мөн их, дээд сургуулиудад хиймэл оюун ухааны хөтөлбөрүүдийг шинээр боловсруулж, оюутнуудыг уг салбарт бэлтгэх ажил эхэлсэн нь хиймэл оюун ухааны салбарыг хөгжүүлэхэд чухал хувь нэмэр оруулах боломжийг бүрдүүлж байна. Тухайлбал, эдгээр хөтөлбөрүүд нь оюутнуудад хиймэл оюун ухааны үндсэн мэдлэг, ур чадварыг олгохоос гадна, түүнийг практик ашиглалтад хэрхэн зөв хэрэглэх талаар зааварчилгаа өгч, салбарын хөгжилд шинэ залуу мэргэжилтнүүдийг бэлтгэх ач холбогдолтой юм.

Хэдий улсын хэмжээнд хиймэл оюун ухааны хөгжлийг дэмжих үндэсний стратеги боловсруулах ажил эхэлсэн [14] ч, боловсролын салбарт энэ технологийг хэрхэн зөв зохицуулах, хэрэглээтэй холбогдох бодлого, аргачлал, зохицуулалт нь хараахан тодорхой биш байна. Боловсролын салбар дахь хиймэл оюун ухааны нэвтрэлтийг үр дүнтэй болгож, зөв зохицуулалт хийх нь тухайн салбарын хөгжлийг түргэтгэх гол түлхүүр болно.

V. Судалгааны үр дүн

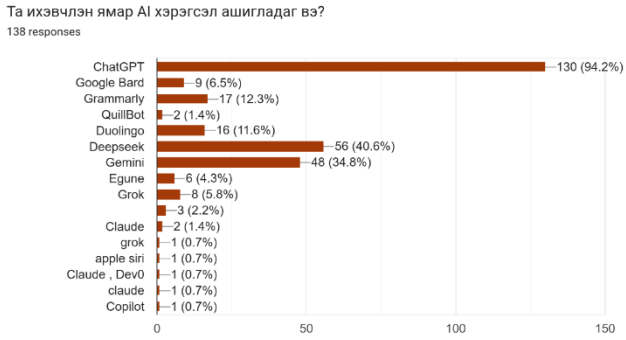
Боловсролын салбарт хүчтэй нэвтэрч, сургалтын чанар, хүртээмжид шинэ боломжуудыг авчирсан хиймэл оюун ухааны технологи нь манай улсын их, дээд сургуулийн оюутнуудын дунд өргөнөөр ашиглагдаж байна. Энэхүү судалгаа нь оюутнууд хиймэл оюун ухааныг хичээлдээ хэрхэн, ямар түвшинд ашиглаж буйг тодорхойлох, түүний үр нөлөө, давуу болон сул талуудыг судлахад чиглэсэн. Энэхүү судалгааг 2025 оны 4-р сарын эхний долоо хоногт цахимаар авсан бөгөөд нийт 144 орчим оюутан оролцов. Судалгааг ерөнхий мэдээлэл, хиймэл оюун ухаан ашиглах байдал, хиймэл оюуны үр нөлөө, хиймэл оюун ухаан ашиглахтай холбоотой бэрхшээл, ёс зүйн асуудал гэсэн дөрвөн хэсэгт хуваж, нийт арван таван асуулгатайгаар авав. Судалгаанд оролцогчдын ихэнх нь 21–23 насны бакалаврын түвшний оюутнууд бөгөөд инженер, технологи, компьютерын ухааны чиглэлээр 1-5-р курсийн суралцагчид байв.

Хиймэл оюун ухаан ашиглах байдал хэсэгт AI хичээлдээ ашиглаж байсан эсэх, ямар зорилгоор ашигладаг, ашиглалтын давтамж, ямар AI хэрэгсэл ашигладаг талаар асуув. Оролцогчдын 95.1 хувь нь хиймэл оюун ухааныг хичээлдээ ашиглаж байсан гэж хариулсан бөгөөд нийтлэг зорилго нь код бичих, программчлалын тусламж авах (81.2%), хичээлээ тайлбарлуулах ойлгох (71%), судалгааны материал хайх (65.9%), шалгалтад бэлдэх (37%) эсэх, тайлан бичих (35.5%) зэрэг байв. AI-г өдөр бүр (50%) болон долоо хоногт хэд хэдэн удаа (43.5%) хэрэглэдэг байна.



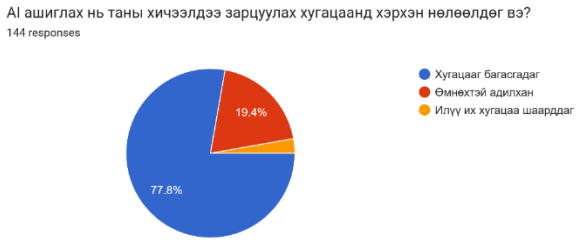
Зураг 1. AI-г ашиглах зорилго

Хамгийн өргөн хэрэглэгддэг хиймэл оюун ухааны хэрэгслүүд нь ChatGPT (94.2%), Deepseek (40.6%), Gemini (34.8%), Grammarly (12.3%), Duolingo (11.6%) бөгөөд Google bard, Grok, Copilot, Egune зэрэг бусад платформуудыг давхар ашигладаг байна.



Зураг 2. Ашигладаг AI хэрэгслүүд

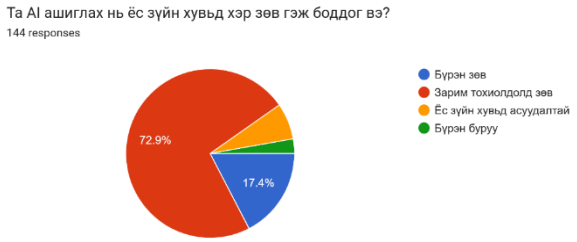
Хиймэл оюун ухааны үр нөлөө хэсэгт AI хичээлд хэр үр дүнтэй тусалдаг, AI ашиглах нь хичээлд зарцуулах хугацаанд хэрхэн нөлөөлдөг, AI ашигласнаар суралцах арга барилд ямар өөрчлөлт гарсан талаар судалсан бөгөөд AI-ийн хичээлд үзүүлэх нөлөөллийг оролцогчдын 52.8% нь эергээр буюу маш үр дүнтэй гэж үнэлсэн байна. Хиймэл оюун ашиглан хичээлээ тайлбарлуулснаар 77.1% нь хичээл илүү ойлгомжтой болсон гэж үзжээ. Үүний үр дүнд 77.8% нь хичээлд зарцуулах хугацаа багассан, 41% нь бүтээлч байдал сайжирсан гэх эерэг өөрчлөлт гарсан хэдий ч зарим оюутнууд бие дааж хийх чадвар буурсан гэж тэмдэглэсэн нь анхаарал татаж байна. Дээрх хариултуудыг үндэслэн, AI нь оюутнуудад багшийн тайлбарыг авах боломжгүй үед хурдан хариулт авч, цаг хэмнэхэд туслах хэрэгсэл болж байна.



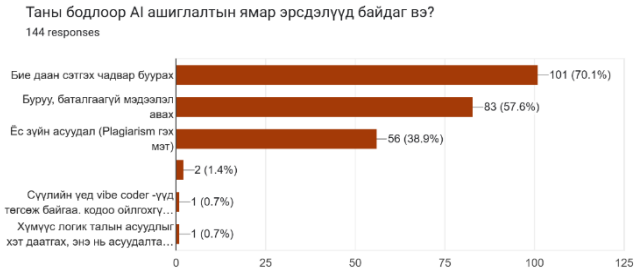
Зураг 3. AI хичээлд зарцуулах хугацаанд нөлөөлөх байдал

AI ашиглахтай холбоотой бэрхшээл, ёс зүйн талаарх хэсэгт AI ашиглах нь ёс зүйн хувьд хэр зөв тухай, AI-с хамааралтай болох эрсдэл, ямар ашиглалтын эрсдэл байгаа талаар мөн AI-г боловсролын салбарт хэрхэн

сайжруулж ашиглах талаар ямар саналтай байгаа тухай тандав. Асуултад нийт оролцогчдын 72.9% нь зарим тохиолдолд ашиглах нь зүйтэй гэж үзсэн бол 34.7% нь хиймэл оюунаас их хэмжээгээр хамааралтай болох эрсдэлтэй гэж тэмдэглэсэн байна. Оролцогчид гарч болох гол эрсдэлүүдэд бие даан сэтгэх чадвар буурах, баталгаагүй мэдээлэл авах, хуулбарлалт гэх мэтийн ёс зүйн асуудал, бичиж буй кодоо ойлгохгүй хиймэл оюун ухаанаар бичүүлэх зэрэг асуудлуудыг дурдсан байна. Хиймэл оюун ухааны хэрэглээг сургуулийн орчин хааж хорих бус амжилттай туршлагуудыг хэрэгжүүлж суралцагчдын чадамжийг илүү тэлэх талаас нь дэмжихийг хүссэн саналуудыг гаргасан байна



Зураг 4. AI ашиглах ёс зүйн асуудал



Зураг 5. AI ашиглалтын эрсдэлт хүчин зүйлс

VI. Дүгнэлт

Монгол улс боловсролын салбарт хиймэл оюун ухааны технологийг нэвтрүүлснээр боловсролын чанар, хүртээмжийг сайжруулах өргөн боломжтой. Гэхдээ энэ боломжийг хэрэгжүүлэхэд дэд бүтэц, хүний нөөцийн чадавх, бодлогын тодорхойгүй байдал зэрэг сорилтууд үүсэж байна.

AI технологи нь оюутнуудын суралцах үйл явцыг хялбар, хүртээмжтэй болгох боломжийг олгож байгаа ч, түүнтэй зэрэгцэн хэт хамаарал, ёс зүйн зөрчил, мэдээллийн үнэн зөв байдал зэрэг эрсдэлүүдийг дагуулж байна. Тиймээс оюутнуудад AI-г зохистой ашиглах талаар бодлого, сургалт, зааварчилгаа хэрэгтэй байна.

Судалгаанаас харахад, Монгол оюутнууд AI-ийг виртуал туслах хэрэгсэл байдлаар ихэвчлэн ашиглаж байна. AI-ийг шууд шийдэл гаргагч бус, туслах хэрэгсэл болгон ашиглаж, хувьчлагдсан сургалт, ёс зүйн хяналтыг нэвтрүүлбэл боловсролд илүү үр дүнтэй хувьсгал авчирч чадна.

Судалгаанаас харахад AI нь оюутнуудад цаг хэмнэж, хичээлийг хялбар, ойлгомжтой болгох давуу талтай ч бие даан сэтгэх чадвар буурах, баталгаагүй мэдээлэл, ёс зүйн асуудал зэрэг эрсдэл дагуулж байна. Оюутнуудын 94.2% ChatGPT-г ашигладаг нь энэ хэрэгслийн түгээмэл байдлыг харуулж байгаа ч түүний хэрэглээг тэнцвэржүүлэх шаардлага тодорхой байна.

Ашигласан материал

[1] “AI market projected to hit $4.8 trillion by 2033, emerging as dominant Frontier Technology,” UN Trade and Development (UNCTAD), https://unctad.org/news/ai-market-projected-hit-48-trillion-2033-emerging-dominant-frontier-technology.

[2] Global Education Monitoring Report 2023: Technology in education: A tool on whose terms? 2023. doi: 10.54676/uzqv8501.

[3] 2023 Global Education Outlook, https://www.holoniq.com/notes/2023-global-education-outlook (accessed Mar. 8, 2025).

[4] N. Verma, How effective is AI in education? 10 case studies and examples, Axon Park, https://axonpark.com/how-effective-is-ai-in-education-10-case-studies-and-examples/ (accessed Apr. 1, 2025).

[5] J. Zhang, Y. Zhang, and M. Zhang, The application of AI in online learning and its impact on student engagement, in 2021 IEEE Int. Conf. on Artificial Intelligence and Education (ICAIE), Beijing, China, 2021, pp. 35–40

[6] Turnitin, 2023 Turnitin Wrapped, Dec. 2023. [Online]. Available: https://www.turnitin.com/blog/2023-turnitin-wrapped.

[7] J. O’Neill and C. Bamberger, “Personal data of 820,000 NYC students compromised in Hack,” New York Post, https://nypost.com/2022/03/26/nyc-students-have-personal-data-hacked/.

[8] UNESCO, Artificial Intelligence in education: Challenges and opportunities, UNESCO Report, 2022. [Online]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000380307.

[9] World Bank, Digital Progress and Trends report 2023 World Bank Group, 2023, https://www.worldbank.org/en/publication/digital-progress-and-trends-report.

[10] EdTech Hub, Teachers’ readiness for using AI in education: Survey results, EdTech Hub Research, 2023. [Online]. Available: https://edtechhub.org/research/teachers-ai-education-2023

[11] ‘Боловсролын салбарын цахим хуваагдал’ уулзалтыг зохион байгуулж байна, Монгол улсын засгийн газар, Боловсролын яам https://moe.gov.mn/post/158018.

[12] ‘Хиймэл оюунд суурилсан багшийн виртуал туслахыг хөгжүүлэх, өгөгдөл, мэдээллийг боловсруулах үндэсний баг аждаа эхэллээ, Боловсролын ерөнхий газар, https://edu.gov.mn/public/article/2495

[13] ‘Хиймэл оюун ухааныг боловсролын системд нэвтрүүлэх хөтөлбөрт хамрагдахаар ярилцав, Монгол улсын засгийн газар, Боловсролын яам, https://moe.gov.mn/post/132776

[14] “НҮБ-ын Хөгжлийн хөтөлбөр Монгол Улсын их өгөгдөл, хиймэл оюуны үндэсний стратегийн анхны чуулга уулзалтыг хамтран зохион байгуулаа,” НҮБ-ын хөгжлийн хөтөлбөр, https://www.undp.org/mn/mongolia/press-releases/nub-yn-khogzhliyn-khotolbor-mongol-ulsyn-ikh-ogogdol-khiymel-oyuuny-undesniy-strategiyn-ankhny-chuulga-uulzaltyg-khamtran.

Өвчлөх эрсдэлийг машин сургалтаар тооцох аргачлалын зохиомж

Хорлоогийн Дорж
Мэдээллийн Технологи, Электроникийн Сургууль
Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол улс
18b1num1221@stud.num.edu.mn

Сандагийн Цогтсайхан
Био-Анагаахын сургууль
Анагаахын шинжлэх ухааны үндэсний их сургууль
Улаанбаатар хот, Монгол улс
tsogtsaikhan.s@mnums.edu.mn

Зундуйн Цолмон
Мэдээллийн Технологи, Электроникийн Сургууль
Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол улс
tsolmonz@num.edu.mn

Ганболдын Амарсанаа
Мэдээллийн Технологи, Электроникийн Сургууль
Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол улс
amarsanaag@num.edu.mn

***Хураангуй*—Эрүүл мэндийн их өгөгдлийн шинжилгээг үйлчлүүлэгчид үзүүлэх тусламж үйлчилгээг чанартай, үр дүнтэй болгоход тухайлбал, өвчин, эмгэг төлөвийг урьдчилан таамаглах, эрсдэлийн үнэлгээ хийхэд өргөн ашиглаж байна.**

Бид хүн амын эрүүл мэндийн их өгөгдлийн дүн шинжилгээнд ашиглах боломжтой өгөгдлийн инженерчлэл, машин сургалт ашигласан эрсдэлийн үнэлгээний загвар хөгжүүлэх зорилго тавин энэ судалгааг хийлээ.

Судалгаанд 2024 оны 10-р сараас 2025 оны 3-р сард АШУҮИС-ийн Монгол-Япон эмнэлгийн Эрүүл мэндийг дэмжих төвөөр үйлчлүүлэн эрүүл мэндийн урьдчилан сэргийлэх үзлэг шинжилгээнд хамрагдсан 961 хүний өгөгдөл болон үзлэг шинжилгээний дүнгээр оношлогдсон хоол боловсруулах өвчний оношнуудын өгөгдлийг ашиглав. Энэхүү аргачлал нь өгөгдөл цуглуулах, боловсруулах, цэвэрлэх, машин сургалтад бэлтгэх болон сургах гэсэн үе шатуудаар хэрэгжүүлнэ.

Өгөгдөл бага учир үр дүнг үнэлэх боломжгүй хэдий ч бид одоогийн байдлаар өвчин тус бүр дээр ач холбогдол бүхий утгуудыг түүж авч, эрсдэлийн бүлгүүдийг үүсгэн хувь хүн бүрд эрсдэлийн үнэлгээг гаргах, логистик регрессийн аргаар машин сургах боломжтой аргачлалыг зохиомжиллоо.

***Түлхүүр үгс*—хүн амын эрүүл мэнд, таамаглах, машин сургалт, өгөгдөл боловсруулалт, цэвэрлэх, ROC муруй**

I. Удиртгал

Эрүүл мэндийн их өгөгдлийн шинжилгээг үйлчлүүлэгчид үзүүлэх тусламж үйлчилгээг чанартай, үр дүнтэй болгоход болон эрүүл мэндийн удирдлагыг боловсронгуй болгоход, тухайлбал, өвчин, эмгэг төлөвийг урьдчилан магадлах (prediction), эрсдэлийн үнэлгээ (risk estimation), эмчилгээ, нийтийн эрүүл мэндийн үр дүнг хянах (improved monitoring), үйлчилгээг чирэгдэлгүй болгох (communication assistance), дүрс танилтыг стандартжуулах (standardized image recognition), лабораторийн үйл ажиллагааг автоматжуулах (automated lab-work), эмийн харилцан үйлчлэлийн хяналт (drug interaction detection), бодгальжсан буюу нарийвчилсан анагаах ухаан (personalized or presicion medicine) зэрэг чиглэлээр өргөн ашиглаж байна [11], [12].

Анагаахын шинжлэх ухааны үндэсний их сургууль (АШУҮИС)-ийн Био-Анагаахын сургууль, Анагаах ухааны сургууль, Монгол уламжлалт анагаах ухааны олон улсын сургуулийн судлаачдын хамтарсан судлаачдын баг нь бодгальжсан буюу давхраажуулсан анагаах ухааны (personalized or stratified medicine) аргуудыг ашиглан эрүүл мэндийн эрсдэлийг үнэлэх, эмчилгээ оношилгооны чанар үр дүнг сайжруулах чиглэлээр сүүлийн жилүүдэд хийсэн судалгааны үр дүнгээрээ статистик тооцооллын тодорхой алгоритмуудыг боловсруулсан [4], [6]. Энэхүү алгоритм нь ихэвчлэн тухайн эмч, мэргэжилтнүүд SPSS статистикийн программ хангамж дээр өгөгдлийг гараараа тооцон түүн дээр дүн шинжилгээ хийдэг байсан бөгөөд хүний алдаа гарах магадлал өндөр, нэг удаа тооцоолох боломж нь хязгаарлагдмал байсан учир үүнийг Python програмчлалын хэлийн сангуудын тусламжтайгаар бодит хугацаанд, машин сургалтын аргаар эрсдэлийг цаг алдалгүй үнэлэх, таамаглах боломжтой болсон.

Энэхүү эрсдэлийг үнэлснээрээ тухайн хүн ямар өвчнөөр өвчлөх магадлалтай аль эрсдэлийн бүлэгт байгааг мэдэх, өвчнөөс урьдчилан сэргийлэх, мөн бүх өгөгдлүүд дээрх тооцоолсон статистик үр дүн дээр эмнэлгийн байгууллага болон шийдвэр гаргагчид улс үндэстний хэмжээнд илүү тархацтай өвчин дээр арга хэмжээ авах цаашлаад Монгол улсын нийт хүн амын эрүүл мэндийг дэмжих чухал ач холбогдолтой юм.

II. Эрүүл мэндийн өгөгдлийн бүрдүүлэлт

Хүн амын эрүүл мэндийн эрсдлийг үнэлэхийн тулд АШУҮИС-ийн анагаах ухааны судалгааны ёсзүйн хорооны олгосон зөвшөөрлийн дагуу АШУҮИС-ийн Монгол-Япон эмнэлгийн Эрүүл мэндийг дэмжих төвөөр үйлчлүүлэн эрүүл мэндийн урьдчилан сэргийлэх үзлэг шинжилгээнд хамрагдсан нь 961 хүний өгөгдөл ашиглан хувь хүн болон нийт хүн амын эрсдэлийг үнэлэх

аргыг алгоритм руу хөрвүүлэх машин сургалт ашиглан тооцоолохын тулд цуглуулсан жишиг өгөгдөл юм.

Хүснэгт I
Асуумжийн олон улсын нэршил болон асуултын тоо

Асуумжийн нэр	Олон улсын нэршил	Нийт асуумжийн тоо
Нийгэм, хүн ам зүйн өгөгдөл	Social and Demographic	10
Хооллолт	Food Frequency Questionnaire	52
Хөдөлгөөний идэвх	Global Physical Activity Questionnaire	19
Нойрны байдал	Insomnia Severity Index	7
Зан төрх	Temperament	80
Стресст өртөмтгий байдал	International Stress Management Association	26
Архаг туйлдалд	Mitchel test	17
Никотины хамаарал	Fagerstrom test	9
Архи согтуурруулах ундааны хэрэглээ	Alcohol Use Disorders Identification Test	11

Энэхүү өгөгдөл нь хүн амын эрүүл мэндийг үнэлэх олон улсын 9 асуумж болон эмнэлзүйн шинжилгээний 1 үр дүн, тухайн хүний өвчний түүх болон эмчийн дүгнэлт бүхий өгөгдлийн багцуудаас бүрдэнэ. Үүнд 221 асуулт мөн OMICS-ийн шинжилгээний хариу бүхий 69 багана өгөгдөл, болон 206 өвчний онош бүхий 961 хүн дээр энэ жишиг өгөгдөл багтана. Тус өгөгдлийг бүрдүүлэхдээ төслийн веб сайт[1] дээр байрлах судалгаа цэс рүү орж өөрийн мэдээллийг бөглөн асуумжуудад хариулан цаашлаад шинжилгээ өгч, өвчний түүх болон онош зэргийг салбарын эмч нараар гарган авч бүрдүүлсэн.

Эдгээр өгөгдлүүд нь тийм, үгүй буюу хоёр утгатай, 1-8 хүртэлх эрэмбэлсэн (1. Улаан дарс огт хэрэглэдэггүй гэх мэт) утга, болон үргэлжилсэн тоон утга (өдөрт хэрэглэдэг давсны хэмжээ гэх мэт) бүхий 3 өөр төрлийн өгөгдөл бүхий багцуудаас бүрдэж байгаа бөгөөд тус бүрийг машин сургалт, статистикийн тооцоо хийхэд ашиглахын тулд 1 болон 0 хэлбэрт хувирган тооцоог хийнэ.

III. Өвчний эрсдэлийг тооцох арга

Хувь хүн болон нийт хүн амын эрсдэлийг тооцоолсноор урьдчилсан үзлэгээр илэрч амжаагүй шинж тэмдгүүдийг олж илрүүлэх маш чухал ач холбогдолтой бөгөөд үүнийг статистик болон компьютерын ухааны ойлголтуудыг хослуулан ашигласнаар хувь хүн дээрх асуумж болон шаардлагатай шинжилгээг өгснөөр өөрийн эрсдэлийг үнэлүүлэх боломжтой болно. Хувь хүн өөрийн эрсдэлийг үнэлснээрээ тухайн өвчинд шаардлагатай эмчилгээ болон тусламжийг цаг алдалгүй авах ач холбогдолтой юм.

A. Эрсдэлийг тооцох асуудал

Эрсдэлийг тооцооход дараах алхмуудаас бүрдэнэ. Pearson-ний хи квадрат арга болон Fisher-ийн нарийвчилсан тестийн ашиглан P-утгыг бодож гаргана [4], [6], [8]. Үүний дараагаар тус багана дээр харьцангуй эрсдэл (Relative risk) [6], харьцуулсан харьцаа (Odds ratio) [6], эерэг үр дүнгийн үзүүлэлт (Positive predictive value) [8]-үүдийг тооцон авна. Харин эцэст нь хувь хүн болон нийт эрүүл мэндийн бүлгийг эрсдэлийн давхарга болон ангилна.

P утга нь өөр хоорондоо ялгаатай байдал нь санамсаргүй байдлаар тохиолдож болох магадлалыг тодорхойлдог бөгөөд утга 0.05-аас бага бол тухайн хоёр багана статистикийн хувьд ач холбогдолтой хэмээн үзнэ [5], [4], [8]. Ач холбогдолтой багануудыг ялган авч өвчин тус бүр дээр логистик регресс ашиглан машин сургана [2].

Харин статистик тооцоолол хийхдээ ач холбогдол бүхий баганыг тухайн өвчний оноштой харьцуулсан 2X2 хэмжээтэй хүснэгт үүсгэнэ. Үүнд, өвчин илэрсэн ба шинж тэмдэг илэрсэн, өвчин илрээгүй ба онцлогтой, өвчин илэрсэн ба онцлоггүй, өвчин илрээгүй болон онцлоггүй гэсэн 4 хэсэгт ангилан дээрх тооцооллуудыг хийнэ.

Хүснэгт II
Өвчинтэй болон онцлогтой хүмүүсийн матриц

	Өвчинтэй	Өвчингүй
Онцлогтой	a	c
Онцлоггүй	b	d

Жишээлбэл, дээрх хүснэгтэд тамхи татдаг (онцлогтой), уушгины өвчтэй (өвчинтэй) гэж харьцуулан харах боломжтой бөгөөд тамхи татдаг ба уушгины өвчтэй (a), тамхи татдаг ба уушгины өвчингүй (c), тамхи татдаггүй ба уушгины өвчтэй (b), тамхи татдаггүй ба уушгины өвчингүй (d) бүлэг гэж ангилна.

Харьцангуй эрсдэлийг тооцоолохдоо өвчинтэй ба онцлогтой хүмүүс (a)-ийг нийт өвчинтэй хүмүүс (a+b)-т харьцуулсан харьцааг өвчингүй ба онцлогтой хүмүүс (c)-ийг нийт өвчингүй хүмүүс (c+d)-т харьцаанд хуваасаар харьцангуй эрсдэлийг (1) тооцоолно [6].

$$RR = \frac{a/(a+b)}{c/(c+d)} \tag{1}$$

$$OR = \frac{a*d}{b*c} \tag{2}$$

$$PPV = \frac{a}{a+c} \tag{3}$$

Харьцуулсан харьцааг (2)-р томьёогоор, эерэг үр дүнгийн үзүүлэлтийг (3)-р томьёогоор тус бүрд тооцоолж олно.

Харин үүний дараагаар ач холбогдолтой багануудад 1 гэж сонгосон бол тухайн багана бүрийн эрсдэлийг нэмснээрээ хувь хүний тус өвчин дэх эрсдэлийн үнэлгээг гарган авах боломжтой болно.

B. Өгөгдлийн боловсруулалт

Логистик регресс болон статистикийн тооцоолол хийхийн тулд бүх утгуудыг 1 болон 0-ээр илэрхийлсний дараагаар тооцоолох болон таамаглах боломжтой юм

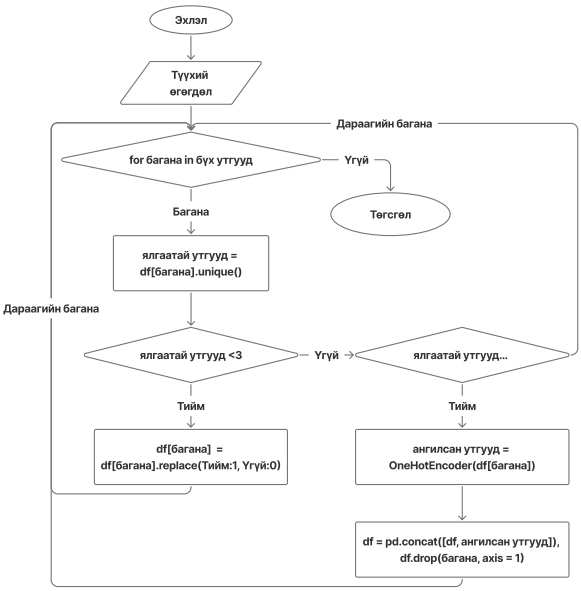
[1] https://www.bigdata.mnums.edu.mn/

[1]. Харин дээрх цуглуулсан өгөгдлүүд нь 3 үндсэн өөр хэлбэрээр илэрхийлэгдсэн бөгөөд эдгээрийг хувиргах нь энэхүү аргын нэг томоохон хэсэг юм.

Хүснэгт III
Өгөгдлийн төрлийг харуулсан жишээ

Тамхи татдаг эсэх ? (Nominal)	**Хооллолтын төрөл (Ordinal)**	**Нас (Continuous)**
Тийм	Хагас цагаан хоолтон	25
Үгүй	Ердийн хоолтон	30
Үгүй	Бүтэн цагаан хоолтон	21
Үгүй	Ердийн хоолтон	36
Тийм	Бүтэн цагаан хоолтон	28
Тийм	Ердийн хоолтон	43
Тийм	Ердийн хоолтон	57

Тийм болон үгүй гэсэн сонголттой асуумжуудыг хувиргах нь хамгийн хялбар. Тухайн асуумжид тийм гэж хариулсан бол 1, харин үгүй гэж хариулсан тохиолдолд 0 гэж хувиргана.



Зураг 1. Өгөгдлийг ангилах алгоритм

Харин эрэмбэлсэн хариултуудыг 1 болон 0 утгуудад хувиргахад one hot encoding аргыг ашиглана [9]. Гэхдээ энгийн one-hot-encoding-гийн ангилал биш анагаах ухаанд эрэмбэлсэн утгуудыг ангилдаг аргыг ашиглан эрэмбэлсэн найм хүртэлх утгуудыг энэхүү зохиомжид хувирган ашигласан. Жишээ нь, ”Алкохолын хэрэглээ” асуулт 8 хариултаас сонгох боломжтой бөгөөд үүнийг 1 ба бусад, 1-ээс 2 болон бусад гэх мэт өсөх байдлаар one-hot-encoding аргыг тохируулан хэрэгжүүлсэн.

Үргэлжилсэн тоон утга бүхий өгөгдлүүдийг хувиргахад илүү төвөгтэй арга ашиглана. Receiver Operating Characteristic (ROC) муруйг ашиглан тооцоолол хийх бөгөөд нэг талд тухайн өвчин эсвэл оношь, нөгөө талд онцлогийг ашиглан sklearn.metrics санг ашиглан тооцоолсон [10]. Хэрэв Area Under Curve (AUC) нь 95%-аас дээш бол тус багана дээр Youden index мэдрэг (sensitivity) болон өвөрмөц (specifity) чанаруудыг ашиглан (5)-аар томьёогоор тооцоолон олж хамгийн их утга бүхий мөрийн утгыг хуваалтын цэг (cut-off point) хэмээн үзэж тус цэгээс их буюу тэнцүү бол 1, бага бол 0 болгон хувиргана.

$$Youden = Sensitivity + Specifity - 100 \qquad (4)$$

Дээрх аргуудыг зураг 1-т харуулсан алгоритмыг ашиглан тус бүр ангилан түүнийгээ байнга тооцоолол хийхээс зайлсхийж номинал болон эрэмбэлсэн өгөгдлийг PostgreSQL олон улсын асуумж бүрээр хүснэгт үүсгэн user id-г foreign key-ээр авна. Харин жинхэнэ өгөгдлийн жишээн дээр PostgreSQL өгөгдлийн санд түүхий өгөгдлийг хадгалан машин сургах болон тооцоолохын өмнө дээрх алхмуудыг дахин тооцоолж зөвхөн үр дүнг хадгалах байдлаар ажиллаж болно.

IV. Хэрэгжүүлэлт, үр дүн

A. Архитектур ба алгоритм

Энэхүү зохиомж нь дараах архитектур болон алгоритмуудын зураг 2 болон 3-ийн дагуу ажиллаж байгаа, харин үүнийг сайжруулж шинэчлэх бүрэн боломжтой бөгөөд заавал энэхүү загварын дагуу ажиллуулах ёстой биш юм.

B. Ашигласан технологи

Pandas – Өгөгдөлтэй ажиллах preprocessing буюу боловсруулалт хийхэд dataframe ашиглан түүн дээр хоосон өгөгдлийг хасах, өөрчлөх, мөн өвчин болон оношилгооны dataframe-уудыг хооронд нь нэгтгэхэд ашигласан сан.

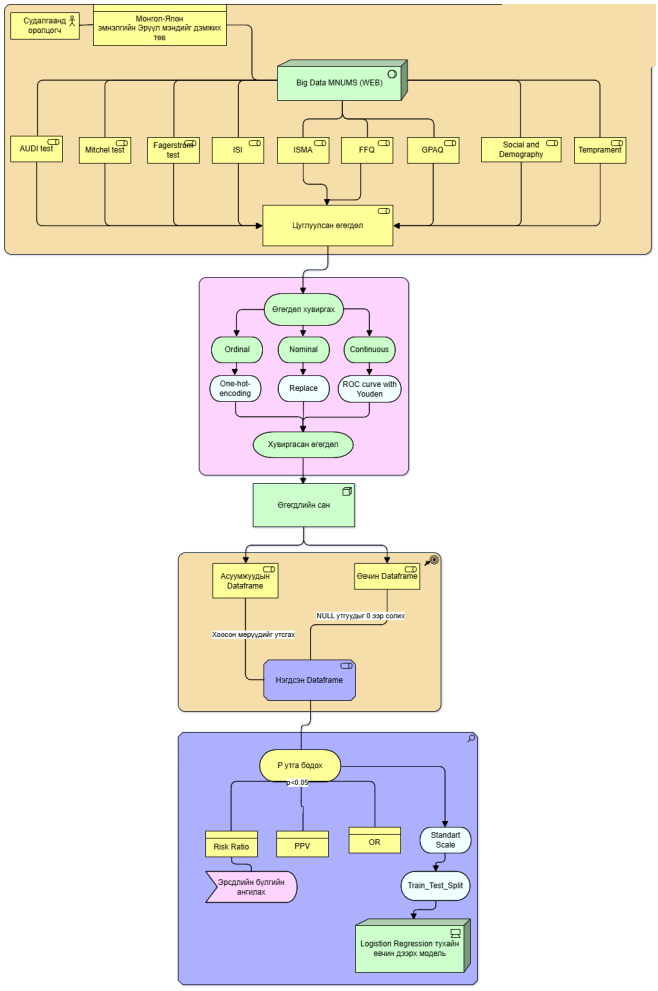
Numpy – Зарим математикийн тооцооллыг харуулахад ашигласан. Жишээ нь, аливаа тоог тэгт хуваах нь өөрөө асуудалтай сэдэв учраас үүнийг тус сангийн inf буюу хязгааргүй гэсэн илэрхийллээр илэрхийлэх гэх мэт.

Sqlalchemy – Веб сайт дээр байрлах Google Forms-оос татаж авсан excel өгөгдлийг PostgreSQL байрлах өгөгдлийн санд хадгалах, түүн дээрх өгөгдлийг ашиглахын тулд энэхүү санг ашигласан.

Scipy.stats – Тус сангаас p утгыг тооцоолохын тулд хи квадрат тест (Chi-square test) болон fisher test ашиглан бодуулсан.

Sklearn – Машин сургалтад ашиглах логистик регресс болон өгөгдлийг боловсруулах StandartScaler, өгөгдлийг хуваах train test split, логистик регрессийн тестийн үр дүнг харах classification report зэргийг ашигласан болно.

PostgreSQL - Ангилсан асуумж бүрийг дахин тооцоолохоос зайлсхийж хувиргасан олон улсын асуумж бүрийг хүснэгтэд хийж түүн өгөгдлийн сангаас өгөгдөл татан авч тооцоолол хийхэд ашигласан.



Зураг 2. Зохиомжийн архитектур

Зураг 3. Эрсдэлийг тооцоолох болон машин сургалтыг хэрэгжүүлэх алгоритм

C. Хийсвэр туршилт

Цуглуулсан өгөгдлийг ашиглан хийсвэр туршилт хийн статистикийн тооцоолол хийсэн үр дүнг хүснэгт 4-р харуулав.

Энэ мэтчилэн 206 өвчин дээр тооцоолол хийж тус бүрд нь таамаглах боломжтой модель сургасан бөгөөд үүнийг харьцуулахын тулд олон гишүүнт логистик регрессийн арга болон автомат машин сургалтын арга болох AutoML H20-аар [13] гарган авсан үр дүнгүүдийг тус бүр харьцуулав. Хоорондын үр дүнг харьцуулахын тулд хоол боловсруулах тогтолцооны өвчнүүдээс хамгийн их тохиолдол бүхий K29.7 буюу гастрит, тодорхойгүй (ходоод) өвчин дээр модель сургасан.

Энгийн логистик регрессийг ашиглан модель сургасан бөгөөд тестийн 289 өгөгдөл дээр гаргасан үр дүнг Хүснэгт 5-д харуулав.

Харин олон хүчин зүйлийн логистик регрессийн аргыг ашиглан бүх өвчнийг таамаглах боломжтой нэг модель үүсгэсэн бөгөөд үүнийг ашиглан зөвхөн K29.7 дээрх тестийн үр дүнг Хүснэгт 6-д харуулав [14].

Хүснэгт IV
K29.7 өвчин дэх ач холбогдолтой баганууд дээрх хийсэн тооцоолол

Ач холбогдол бүхий багана	**P утга**	**OR**	**RR**	**PPV**
Бүтээгдэхүүний давс натрийн хэмжээг хардаг уу?	0.04	2.24	0.56	0
Төмсний хэрэглээ (Хоногт 3-аас дээш ба бусад)	0.03	2.59	0.46	0
Сүү, цагаан идээний хэрэглээ (Хоногт 3-аас дээш ба бусад)	0.05	2.61	0.43	0
Кофены хэрэглээ (7 хоногт 5-6 удаа ба бусад)	0.02	2.36	0.65	0
Кофены хэрэглээ (Хоногт 1 удаа ба бусад)	0.02	2.44	0.51	0
Кофены хэрэглээ (Хоногт 2-3 удаа ба бусад)	0.01	3.03	0.41	0
Цагаан архины хэрэглээ (7 хоногт 5-6 удаа ба бусад)	0.02	11.16	0.09	0
Цагаан архины хэрэглээ (Өдөрт 1 удаа ба бусад)	0.00	56.06	0.02	0
Цагаан архины хэрэглээ (Өдөрт 2-3 удаа ба бусад)	0.00	inf	0.00	0
Тамхи татдаг уу (Үгүй ба бусад)	0.04	3.09	0.81	0
B_L3_Urine_AC_<30mg_g (Цусны шинжилгээний хариу)	0.02	2.39	0.55	0

Хүснэгт V
K29.7 өвчний логистик регрессийн үр дүн

	Precision	**Recall**	**f1-score**	**support**
0.0	0.98	0.76	0.86	278
1.0	0.09	0.64	0.16	11
accuracy			0.75	289
macro avg	0.54	0.70	0.51	289
weighted avg	0.95	0.75	0.83	289

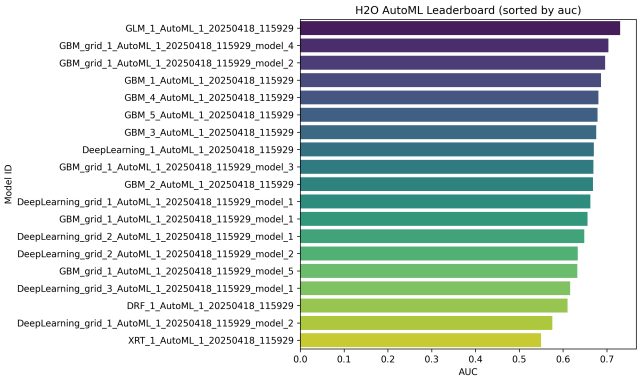
Хүснэгт VI
Олон хүчин зүйлийн логистик регрессийг ашиглан K29.7 өвчин дээрх тестийн үр дүн

	Precision	**Recall**	**f1-score**	**support**
0.0	0.95	0.94	0.94	274
1.0	0.11	0.13	0.12	15
accuracy			0.90	289
macro avg	0.53	0.54	0.53	289
weighted avg	0.91	0.90	0.90	289

H20 автомат машин сургалтын арга ашиглан 20 модель зэрэг сурган тус өгөгдөл дээр туршилт хийхэд үр дүн зураг 4-т харуулав [13]. Тус үр дүнгээс хамгийн сайн таамагласан модель нь Generalized Linear Model байсан бөгөөд ойролцоогоор 72%-тайгаар таамаглаж байсан.

V. Дүгнэлт

Аливаа өвчнийг олон улсын асуумж болон эмнэлзүйн үзүүлэлт, шинжилгээ гэх мэт эрүүл мэндийн өгөгдөл дээр машин сурган түүнийг таамаглах боломжтой бөгөөд хамгийн чухал хэсэг нь өгөгдөл цуглуулах юм [11]. Машин сургалтад өгөгдлийн хэмжээ хэр байхаас шалтгаалж



Зураг 4. AutoML H20 ашиглан сургасан үр дүн

тухайн моделийн чанар, үр дүн шууд хамааралтай байдаг. Хэрэв хангалттай хэмжээний өгөгдлийг цаашдаа цуглуулан энэхүү зохиомжоор туршилт хийн хэр үнэн зөв таамаглаж байгааг жинхэнэ туршилтаар мөн шалгах боломжтой болно.

Дараагийн чухал асуудал нь өгөгдлийг боловсруулах бөгөөд ялангуяа эрэмбэлсэн өгөгдлүүдийг хэрхэн ангилахаас шалтгаалж эрсдэлийн бүлэг болон үнэлгээ зэрэг нь өөрчлөгдөх ч боломжтой тул үүн дээр анхаарах нь нэн тэргүүний асуудал билээ.

Машин сургалт ашиглан сургасан модель тус бүрийн үр дүнг ажиглахад одоогийн байдлаар олон хүчин зүйлийн логистик регрессийн арга нь хамгийн өндөр таамаглах чадамжтай байгаа хэдий ч өгөгдлийн хэмжээ ихсэх үед өөрчлөгдөх боломжтой.

Зүүлт

References

[1] Kotsiantis, Sotiris, Kanellopoulos, Dimitris, Pintelas, P.. (2006). Data Preprocessing for Supervised Learning. International Journal of Computer Science. 1. 111-117.

[2] Schober, Patrick, Vetter, Thomas. (2021). Logistic Regression in Medical Research. Anesthesia, Analgesia. 132. 365-366. 10.1213/ANE.0000000000005247.

[3] Zhang, Wei, et al. “A Novel Hybrid Model for Data Classification.” Information Sciences, vol. 465, 2018, article 1903. ScienceDirect, https://doi.org/10.1016/j.ins.2018.10.1903.

[4] Munkh-Undrakh Batmunkh et al. “Sex-adjusted Approach to Baseline Variables Demonstrated Some Improved Predictive Capabilities for Disease Severity and Survival in Patients with Coronavirus Disease 19.” Inform Med Unlocked, vol. 31, 10 June 2022, article 100982, https://doi.org/10.1016/j.imu.2022.100982. PubMed Central, PMC9186409.

[5] Kumar, Anil, et al. ”Clinical Outcomes and Laboratory Findings in COVID-19 Patients: A Retrospective Analysis.” International Journal of Infectious Diseases, vol. 104, 2021, pp. 50–56. PubMed Central, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7939265/.

[6] Simon, Stephen D. “Understanding the Odds Ratio and the Relative Risk.” Journal of Andrology, vol. 22, no. 4, July–Aug. 2001, pp. 533–536. Wiley Online Library, https://doi.org/10.1002/j.1939-4640.2001.tb02212.x.

[7] Vinay, Sachin. “Standardization in Machine Learning.” ResearchGate, Mar. 2021,

[8] Bayarsaikhan, Solongo, et al. ”Survival, CD4 T Lymphocyte Count Recovery and Immune Reconstitution Pattern during the First-Line Combination Antiretroviral Therapy in Patients with HIV-1 Infection in Mongolia.” PLOS ONE, vol. 16, no. 3, 8 Mar. 2021, e0247929. Public Library of Science, https://doi.org/10.1371/journal.pone.0247929. https://www.researchgate.net/publication/349869617_STANDARDIZATION_IN_MACHINE_LEARNING.

[9] Samuels, Jamell Alvah. ”One-Hot Encoding and Two-Hot Encoding: An Introduction.” ResearchGate, Jan. 2024,https://www.researchgate.net/publication/377159812_One-Hot_Encoding_and_Two-Hot_Encoding_An_Introduction.

[10] Fawcett, Tom. ”An Introduction to ROC Analysis.” Pattern Recognition Letters, vol. 27, no. 8, June 2006, pp. 861–874. ScienceDirect, https://doi.org/10.1016/j.patrec.2005.10.010.

[11] Batko K, Ślęzak A. The use of Big Data Analytics in healthcare. Journal of big data. 2022;9(1):3.

[12] Cirillo D, Valencia A. Big data analytics for personalized medicine. Current Opinion in Biotechnology. 2019;58:161-167.

[13] Doe J, Smith J. Automated Machine Learning for Predictive Modeling. Proceedings of AutoML. 2020;paper 61.

[14] Erkan A. Using Multinomial Logistic Regression to Examine the Relationship Between Children’s Work Status and Demographic Characteristics. ResearchGate. 2016;publication 311908200.

Свитчийн Порт Мироринг Функц Халдлага Илрүүлэх Системийн Пакетийн Алдагдалд Нөлөөлөх нь

Б.Энхбаяр, Г.Манлайбаатар, Д.Баттулга, Б.Өсөхбаяр*, Н.Угтахбаяр
Мэдээллийн Технологи, Электроникийн Сургууль, Монгол Улсын Их Сургууль
Электроник, Холбооны Инженерчлэлийн Тэнхим
E-mail: enkhbayar.ba@num.edu.mn, *Corresponding author: usukhbayar@num.edu.mn

Хураангуй—Сүлжээгээр дамжин ирж буй халдлага, дайралтуудыг үр дүнтэй хянадаг, илрүүлдэг халдлага илрүүлэх системүүдийн хэрэгцээ шаардлага орчин үеийн кибер орчинд улам бүр өсөн нэмэгдэж байна. Байгууллагууд ихэвчлэн этернет свитчүүдийн порт мироринг (port mirroring) функцийг ашиглан сүлжээний урсгалыг салаалан халдлага илрүүлэх системүүд рүү дамжуулах шийдлийг ашигладаг бөгөөд энэ арга нь өргөн хэрэглэгддэг. Гэвч сүлжээний урсгалыг свитчийн порт мироринг ашиглан салаалах арга Халдлага илрүүлэх системүүд рүү ирэх сүлжээний урсгалд үзүүлэх нөлөөллийн талаар хангалттай судалгаанууд хийгдээгүй байна. Энэхүү судалгааны ажлын зорилго нь порт мироринг функцээр сүлжээний урсгалыг салаалан дамжуулснаар үүсэж буй пакет алдагдлыг сүлжээ салаалагч (Network TAP) төхөөрөмж ашигласан үетэй харьцуулан туршиж хэмжин үр дүнг тодорхойлсон ба цаашлаад порт мироринг функцийн пакетийн алдагдал Халдлага илрүүлэх системийн пакетийн алдагдалд хэрхэн нөлөөлж байгааг судалсан. Ингэснээр халдлага илрүүлэх систем нь пакетийн алдагдал үүссэн үед сүлжээгээр дамжиж буй халдлагын пакетуудийг гүйцэд хүлээж авч чадахгүй байгаа нь сүлжээнд ирж буй халдлага, дайралтуудыг бүрэн илрүүлэхэд нөлөөлж болохыг тодорхойлох боломжтой болох юм. Бидний судалгааны ажлын явцад хийсэн туршилтын үр дүнд порт мироринг функц ашиглан дамжуулах үед сүлжээ бага ачаалалтай байхад ч гэсэн пакет алдагдал үүсэж байсан бөгөөд энэ нь халдлага илрүүлэх системийн гүйцэтгэлд шууд нөлөөлдөг болохыг тодорхойлсон.

Түлхүүр үгс—Сүлжээний урсгал, пакет алдагдал, порт мироринг, гүйцэтгэл, сүлжээний халдлага илрүүлэх систем (ХИС), сүлжээ салаалагч

I. Удиртгал

Сүүлийн жилүүдэд интернэт болон сошиал медиа хэрэглээ дэлхий даяар эрчимтэй өссөөр байна. Үүний дотор байгууллагуудын дижитал шилжилт, үүлэн тооцоолол, IoT төхөөрөмжийн нэвтрэлт зэрэг нь сүлжээний орчныг улам бүр төвөгтэй болгож байна. Үүнийг дагаад сүлжээний аюулгүй байдлыг хамгаалах асуудал өнөө цагийн чухал асуудлуудын нэг болж байна. Мөн цахим гэмт хэрэгтнүүдийн халдлагын арга техник нарийсаж, улам бүр хөгжиж байгаа нь сүлжээний аюулгүй байдлыг хангах стратегийг сайжруулах зайлшгүй шаардлагыг бий болгож байна. Сүлжээгээр дамжин хийгдэж буй халдлагуудыг илрүүлэх, урьдчилан сэргийлэх болон хурдан хугацаанд хариу арга хэмжээ авахын тулд сүлжээний хамгаалалттай холбоотой системүүдийг сайжруулах, тэдгээрийн гүйцэтгэлийг нэмэгдүүлэх хэрэгтэй байна. Сүлжээний Халдлага Илрүүлэх Систем (ХИС)-үүд нь мэдээллийн аюулгүй байдлыг хамгаалах үндсэн хэрэгслүүдийн нэг юм. Халдлага илрүүлэх системийн үр ашиг нь зөвхөн ашиглаж буй алгоритм, техникээс гадна тоног төхөөрөмжийн хүчин чадал (CPU, RAM, Network card гэх мэт), сүлжээний урсгал боловсруулах аргачлал бүхий сан (Libpcap, AF_Packet), үйлдлийн систем (OS)-ийн оновчлол зэрэг олон хүчин зүйлээс хамаардаг болохыг судалж тодорхойлсон [1]–[4]. Сүлжээний ХИС-ийн гүйцэтгэлийг сайжруулах чиглэлээр олон төрлийн судалгаа хийгдсэн байдаг. Жишээлбэл халдлага илрүүлэх систем нь өндөр хурдтай, олон төрлийн сүлжээний урсгалыг хянах үед пакетуудийг богино хугацаанд хүлээн авч боловсруулж чадахгүй алддаг болохыг [1] судалгаагаар тогтоосон. Сүлжээгээр дамжиж буй пакетуудийг хянах, халдлагыг илрүүлэх, дүн шинжилгээ хийх зорилгоор бүх урсгалыг салаалуулах найдвартай арга хэрэгтэй [8]. Сүлжээний ХИС сүлжээний урсгалыг бүрэн хянахын тулд свитчийн порт мироринг функцийг суурилуулалт, тохируулахад хялбар, нэмэлт төхөөрөмж шаардлагагүй гэх мэт давуу талаас үүдэн түгээмэл ашигладаг. Мөн сүлжээг салаалуулах Network TAP (цаашид "сүлжээ салаалагч" гэх) ашиглах болох ч өртөг өндөр, суурилуулалт хийх шаардлагатай тул түгээмэл ашигладаггүй. Тиймээс энэхүү судалгааны ажлаар сүлжээний ХИС-ийн гүйцэтгэлд свитчийн порт мироринг функцээр сүлжээний урсгалыг салаалуулан дамжуулах үеийн пакетийн алдагдал халдлага илрүүлэх системийн гүйцэтгэлд хэрхэн нөлөөлж буйг судална.

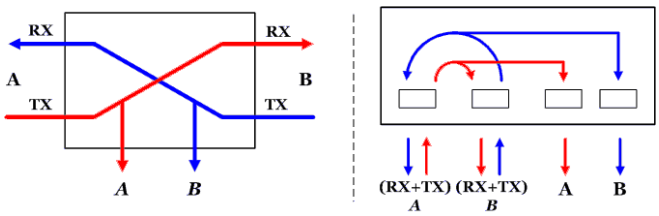
II. Сүлжээний урсгал салаалан дамжуулалт

Сүлжээний хяналт нь сүлжээний администраторуудад нарийн төвөгтэй компьютерын сүлжээний агшин зуурын төлөв байдал, урт хугацааны чиг хандлагыг мэдэх боломжийг олгодог механизмуудын багц юм [7]. Сүлжээний хяналт нь идэвхтэй эсвэл идэвхгүй байж болно. Сүлжээний идэвхгүй хяналт нь урсгалд нөлөөлөхгүйгээр шугамаас өгөгдлийг уншдаг. Идэвхтэй сүл-

жээний хяналт нь сүлжээгээр дамжиж буй өгөгдлийг өөрчлөх сонголтыг нэмдэг. Илүү их өгөгдөл цуглуулах тусам өгөгдлийг хадгалах, боловсруулахад илүү их технологийн шаардлага тавигддаг [9].

A. Сүлжээ салаалагч (TAP)

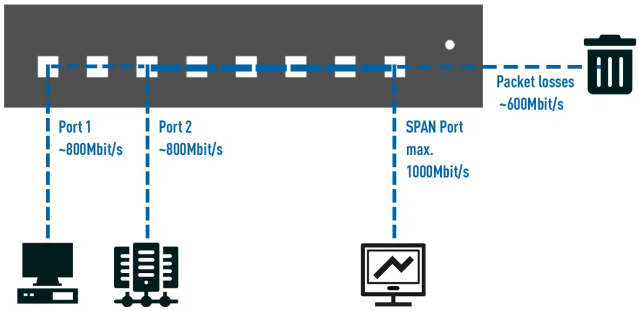
Сүлжээ салаалагч буюу сүлжээний Test Access Point (TAP) төхөөрөмж нь сүлжээгээр дамжиж буй пакетуудийг хянах, дүн шинжилгээ хийх зорилгоор сүлжээний холболт дээр физик түвшний интерфейсийг үүсгэдэг. Энэ нь pass-through буюу дамжуулагч төхөөрөмж байдлаар ажиллаж, холбогдсон сүлжээний бүх мэдээллийг тасалдуулахгүйгээр хуулбарлах чадвартай байдаг идэвхгүй төхөөрөмж [5]. Сүлжээ салаалагч төхөөрөмжийг сүлжээний хоёр зангилааны хооронд (жишээлбэл, свитч ба рүүтер хооронд) суурилуулдаг. Энэ нь сүлжээний кабелийг хоёр хэсэгт хуваах, холболтыг сүлжээ салаалагч төхөөрөмжөөр дамжуулдаг. Сүлжээ салаалагч нь сүлжээний урсгалыг хоёр чиглэлд бүртгэж, Intrusion Detection System (IDS) эсвэл статистикийн траффик үүсгэгч гэх мэт хяналтын төхөөрөмж рүү илгээдэг [5]. Зураг 1[1]-д сүлжээ салаалагч төхөөрөмжийн холболтын схемүүдийг дүрслэн харуулсан байна. Энэхүү төхөөрөмж нь мэдээж хүчдэл ашиглах ба үүний тусламжтай кабелаар дамжиж байгаа пакетуудийг хүчдэлийн уналтаас сэргийлдэг. Мөн Open Systems Interconnection (OSI) загварын 1-р түвшинд ажилладаг тул ямар нэг чиглүүлэлтийн шийдвэр гаргадаггүй.



Зураг 1: Этернет TAP холболтын схем

B. Свитчийн порт мироринг функц

Сүлжээний төхөөрөмжийн тодорхой портуудаас урсах пакетуудийг хуулбарлан өөр порт руу дамжуулж, түүнийг хянах, шинжлэх боломжийг олгодог функц юм. Үүнийг SPAN гэж нэрлэх нь ч бий. Порт мироринг нь суурилуулалт, удирдлага хийхэд хялбар, нэмэлт төхөөрөмж шаардлагагүй гэх мэтээс шалтгаалан түгээмэл ашиглагддаг. Гэсэн хэдий ч энэ нь уг төхөөрөмжийн нөөцөөс шалтгаалан пакет алдагдал үүсэж болох сул талтай. Свитч нь OSI-ийн 2-р түвшинд ажилладаг бөгөөд фремүүдийг чиглүүлэх болон порт мироринг ажиллуулахад хангалттай нөөцгүй байдаг [7]. Свитч нь үндсэн функцийг нэн тэргүүнд тавьдаг бөгөөд ачаалал ихтэй үед порт мироринг зөв ажиллахгүй байж болно [7]. Тодруулбал свитчийн урсгалд хандах хандалт нь урсгал хуулбарлан илгээж буй портын багтаамжаар хязгаарлагддаг. Хэрэв уг портоор дамжин өнгөрөх урсгал хэт их байвал порт нь пакетуудийг хаях бөгөөд ингэснээр зарим мэдээлэл алдагдах болно [5]. Жишээлбэл, 800 Mbps холбоос бүр дээр fullduplex урсгалыг харахын тулд урсгалын нэг салааг илгээж буй порт нь 1600 Mbps хүчин чадалтай байх шаардлагатай. Хэрэв уг салааг илгээж байгаа порт 1000Gbps бол 600Mbps урсгал алдагдах эрсдэлтэй. Үүнийг зураг 2-д дүрслэн харуулсан ба энэхүү зургийг Neox Networks[2] веб сайтаас авсан болно.



Зураг 2: Порт мироринг функцийг ашиглах үед үүсэж болох пакетийн алдагдал

III. Судалгааны үр дүн

Судалгааны хүрээнд хийсэн туршилтын эхний хэсэгт порт мироринг функцийг ашиглан свитч болон сүлжээ салаалагч төхөөрөмжүүдийн сүлжээний урсгал хуулбарлах гүйцэтгэл, пакет алдагдлын түвшинг тодорхойлно. Харин туршилтын хоёрдугаар хэсэгт бид эхний туршилтын үр дүнд үндэслэн сүлжээний ХИС-ийн гүйцэтгэлд хэрхэн нөлөөлж байгааг авч үзнэ.

Туршилтын гүйцэтгэхдээ линукс үйлдлийн системийн (Ubuntu Server 24.04) орчинд шаардлагатай программ хангамжуудыг суулгаж тусгаарласан орчин үүсгэсэн. Туршилт хийх үед дамжиж байгаа пакет бүрийн илүү нарийвчлалтай тооцох үүднээс свитчээс хоост төхөөрөмж рүү илгээгдэж байгаа пакет бүрийг зогсоож (CDP, DTP гэх мэт протоколуудыг идэвхигүй болгосон.) хэмжилт авсан. Туршилтыг свитч болон хоост төхөөрөмжүүдийн сүлжээний интерфейсийн хурдыг хязгаарлан тогтоох замаар туршсан. Өндөр ачаалалтай үед пакет алдагдал өндөр байх магадлалтай.

Туршилтад "Iperf3", "Pytbull-ng" программуудыг ашигласан бөгөөд эдгээр программуудыг ажиллуулах болон пакетуудын тоог хэмжиж авах үүрэг бүхий шель скриптийн тусламжтайгаар гүйцэтгэсэн. Эхний сүлжээний ачааллын нэмэгдүүлэхдээ "Iperf3" программыг "P" сонголтын тусламжтайгаар өөрчлөн туршсан. Туршилтад ашигласан төхөөрөмжүүдийн техникийн үзүүлэлтийг хүснэгт I, топологийг зураг 3, 5-т харуулав. 3

[1]https://dgonzalez.net/papers/roc/node4.html

[2]https://www.neox-networks.com/en/solutions/network-tap-versus-span-mirror-port/

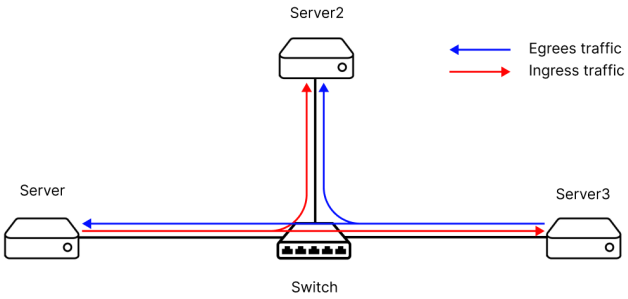
ширхэг мини компьютер, Cisco Catalyst 3650-X свитч, сүлжээ салаалагч төхөөрөмжүүдийг ашигласан.

A. Туршилт 1

Энэ хэсэгт порт мироринг функцийг ашиглан үүссэн хэмжилтүүд дээр судалгаа хийх зорилгоор хийсэн туршилтыг танилцуулна.

Туршилтын хэсэгт сүлжээ салаалагч болон свитчийн порт мироринг функцүүдийн пакет алдагдлыг харьцуулан туршина. Туршилтыг хоёр хэсэгтэй гүйцэтгэсэн. Үүнд:

- Свитч дээр порт мироринг тохируулж сүлжээгээр дамжиж буй пакетийг хяналтын төхөөрөмж рүү дамжуулж пакетийн алдагдлыг хэмжих туршилт
- Сүлжээ салаалагч төхөөрөмжийг ашиглан сүлжээгээр дамжиж буй пакетуудийг хяналтын төхөөрөмж рүү дамжуулж пакетийн алдагдлыг хэмжих туршилт

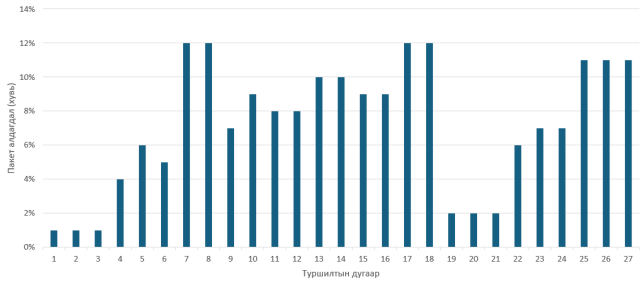


Зураг 3: Свитч төхөөрөмж дээр хийсэн туршилтын топологи

Зураг 3-т 1-р туршилтаар байгуулсан сүлжээний топологийг харуулсан ба сервер 1 төхөөрөмж нь сүлжээний урсгал үүсгэх үүрэгтэй программ хангамж, сервер 3 хүлээн авагч ба сервер 2 төхөөрөмж нь порт мироринг функцийг ашиглан сервер 1 болон сервер 3 төхөөрөмж хооронд дамжих бүх урсгалыг хяналтын төхөөрөмжид хүлээн авч байгааг шалгах юм. пакетуудийг дамжуулахдаа "Iperf3" программыг "P" сонголтын тусламжтайгаар зэрэгцээ дамжууллыг өөрчлөн туршсан.

Туршилтыг хийхдээ илгээгч төхөөрөмжөөс урсгалыг үүсгэн хүлээн авч төхөөрөмж рүү илгээнэ. Үүнийг свитч дээр салаалан монитор буюу хяналтын төхөөрөмж болон хүлээн авагч төхөөрөмж рүү зэрэг илгээнэ, хүлээн авагч хүлээж авсны дараа хариу пакет илгээхэд мөн адил монитор төхөөрөмж болон хүлээн авагч төхөөрөмж рүү илгээдэг байхаар топологийг байгуулж хэмжилтийг хийсэн.

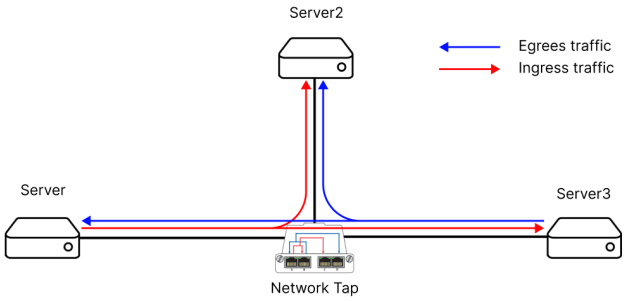
Хэмжилтийг хийхдээ төхөөрөмж мини компьютер бүрд илгээсэн болон хүлээн авсан пакетийн тоо авсан. пакетийн хэмжээг тооцоолохдоо "netstat" командыг "i" болон "n" сонголтуудтай ашигласан. Үүнийг ашиглахдаа өмнө хадгалагдсан дамжсан пакетийн тооноос хэмжилт хийсний дараах пакетийн тоо гарган авах замаар хэмжилтийг хийсэн. Хэмжилтийг нэг төхөөрөмж свитч



Зураг 4: Свитч дээр үүсэн пакет алдагдлын хувь

дээр хийхдээ хурдны хувьд 10Mbps, 100Mbps, 1Gbps тус бүрд нь параллелийн утгыг орчлон 3 удаагийн хэмжилтийг хийсэн. Жишээ нь 10Mbps дээр параллелийн утга нь 1 байх үед 3 удаагийн давтамжтай. Ийм маягаар нийт свитч дээр 27 хэмжилтийн утга авсан. Эдгээр хэмжилтийн үр дүнг зураг 4 диаграммаар дүрслэн харуулсан бөгөөд дунджаар 7% пакетийн алдагдал үүсэж байсан.

Яг ижил туршилтыг свитчийн оронд сүлжээ салаалагчийг зураг 5-т харуулж байгаачлан топологийг байгуулж хэмжилт хийсэн. Ижил аргачлал болон тэнцүү тооны туршилт гүйцэтгэх үед туршилтын үр дүн нь 0% буюу ямар ч пакетийн алдагдал үүсэхгүй байсан.



Зураг 5: Сүлжээ салаалагч ашиглан хийсэн туршилтын топологи

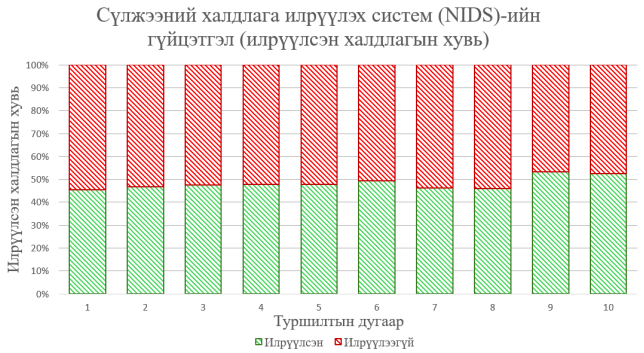
B. Туршилт 2

Эхний туршилтаас бид сүлжээ салаалагч физик төхөөрөмжийг свитчийн порт мироринг функцтэй харьцуулан пакет алдагдлыг хэмжсэн. Порт мироринг функц нь сүлжээний урсгалыг хуулбарлан дамжуулахдаа үүсгэж байгаа пакетийн алдагдал нь ХИС-ийн илрүүлэх гүйцэтгэлд хэрхэн нөлөөлж байгааг турших болно. Туршилтыг гүйцэтгэхдээ зураг 3-т харуулж байгаа топологид туршсан. хяналтын төхөөрөмжид Snort ХИС суулгаж бэлтгэсэн. Мөн энэхүү туршилтад халдлагын урсгал үүсгэх үүрэгтэй "Pytbull-ng" программыг ашигласан. Сервер 1 дээр "Iperf3" болон "Pytbull-ng" программуудыг зэрэг ажиллуулах бөгөөд нэг туршилт 1350 секундийн хугацаанд үргэлжилнэ. Энэ хугацаанд тодорхой пакетийн алдагдал үүсэхийг

Хүснэгт I: Туршилтад ашигласан төхөөрөмжүүдийн техникийн үзүүлэлт

№	Төхөөрөмжийн нэр	CPU	RAM	Hard	Network interface
1	Mini Computer 1	Intel(R) C3558R 2.40GHz	16 GB DDR4	256GB SSD	Gigabit Ethernet
2	Mini Computer 2	Intel(R) C3558R 2.40GHz	16 GB DDR4	256GB SSD	Gigabit Ethernet
3	Mini Computer 3	Intel(R) C3558R 2.40GHz	16 GB DDR4	256GB SSD	Gigabit Ethernet
4	Cisco Catalyst 3650-X	1.2 GHz, Dual-Core	16MB	N/A	12x Gigabit Ethernet
5	Сүлжээний салаалагч	N/A	N/A	N/A	3x Gigabit Ethernet

туршилт 1-ээр баталсан. Нийт 1547 ширхэг халдлагын пакетийг илрүүлэх дүрмүүдийг хяналтын төхөөрөмж дээрх Snort илрүүлэхээр тодорхойлж өгсөн. Энэхүү туршилтыг нийт 10 удаа гүйцэтгэхэд дунджаар 11% -ийн пакетийн алдагдал үүсэж байсан бол дунджаар илрүүлэх нийт халдлагын 48%-ийг илрүүлж байсан. Энэхүү туршилтын үр дүнг зураг 6-д харуулав.

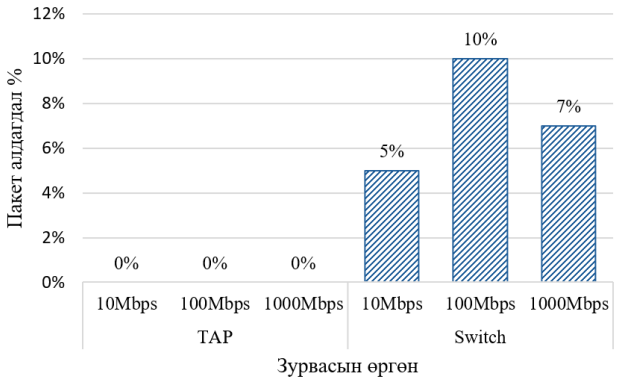


Зураг 6: Халдлага илрүүлэлтийн хувь

Туршилтыг свитчийн оронд сүлжээ салаалагч ашиглаж давтан гүйцэтгэсэн. Туршилтын үр дүнд 48%-ийн гүйцэтгэл нь 98% хүртэл өсөж байсан. Энэ нь ямар нэг пакет алдагдаагүй үед ХИС-ийн гүйцэтгэл пакет барих сан, системийн нөөц, программын хувилбар гэх мэт олон хүчин зүйлээс шалтгаалан тодорхой хувийн алдагдал үүсэх буюу дутуу илрүүлж болохыг [1]–[4] судалгаануудийн үр дүнгээс харж болно.

C. Үр дүн

Туршилтын үр дүнгээс үзэхэд сүлжээ салаалагч төхөөрөмж нь өгөгдлийг физик түвшинд хуулбарладаг тул пакетуудийг алдаагүй, бүрэн бүтэн дамжуулж байсан. Тодруулбал "Iperf3" программын тусламжтай ялгаатай 3 зурвасын өргөний хувьд пакет алдагдлыг 2 арга техник дээр туршиж хэмжилт хийхэд сүлжээ салаалагч төхөөрөмж ашиглах үед зурвасын өргөн бүрд ижил 0%-ийн пакет алдагдал үүсэж байсан бол свитчийн порт мироринг функц нь зурвасын өргөн тус бүрд 5%, 10%, 7% пакетийн алдагдал үүсэж байна. Зураг 7-т сүлжээ салаалагч төхөөрөмж болон свитчийн порт мироринг функц ашиглан сүлжээний ачааллын үед үүсэж болох пакет алдагдлыг хэмжихэд гарсан үр дүнг харуулж байна. Энэхүү туршилтад үндэслэн ийм хэмжээний пакетийн алдагдал үүсэж үед халдлагын хэдэн хувийг илрүүлэхийг туршилт 2-т туршсан бөгөөд үүний үр дүнд 11%-ийн пакетийн алдагдал нь 48% хүртэл сүлжээний ХИС-ийн халдлага илрүүлэх гүйцэтгэлийг бууруулсан.



Зураг 7: Хэмжилтийн үр дүнд үүссэн пакет алдагдлын диаграмм

IV. Дүгнэлт

Свитчийн порт мироринг функц нь сүлжээний өгөгдлийг хянах, шинжлэхэд тохиромжтой, зардал багатай шийдэл бөгөөд свитчийн үндсэн функц тул нэмэлт тоног төхөөрөмж шаардлагагүйгээр хэрэгжүүлж болно. Эдгээр давуу талууд нь энэхүү шийдлийг сүлжээний халдлага илрүүлэхэд түгээмэл ашиглах шалтгаан болсон. Сүлжээ салаалагч нь пакетуудийг физик түвшинд хуулбарлах тул пакет алдагдалгүй, өндөр нарийвчлалтай болохыг туршилт 1-ийн үр дүнгээс харж болно. Өөрөөр хэлбэл сүлжээний урсгал хуулбарлан хяналт хийх шинжлэх бол сүлжээ салаалагч төхөөрөмж нь өндөр гүйцэтгэлтэй буюу пакет алдагдал 0% байна. Харин сүлжээний ХИС-д свитчийн порт мироринг функцийг ашиглахдаа буруу тохируулах, хүчин чадал багатай свитч ашиглах зэрэг нь пакетийн алдагдалд хүргэх бөгөөд 1-р туршилтын үр дүнд 7%-ийн пакет алдагдал үүссэн. Энэ нь сүлжээний ХИС-ийн гүйцэтгэлийг 48% хүртэл бууруулж болохыг энэхүү судалгааны ажлаар харууллаа.

Товчилсон үгс

ХИС	Халдлага Илрүүлэх Систем
TAP	Test Access Point
IDS	Intrusion Detection System
NIDS	Network Intrusion Detection System
OS	Operating System
SPAN	Switched Port Analyzer
OSI	Open Systems Interconnection

Ном зүй

[1] Qinwen Hu, Se-Young Yu, Muhammad Rizwan Asghar, "Analysing performance issues of open-source intrusion detection systems in high-speed networks," Journal of Information Security and Applications, vol. 51, pp. 102426, 2020, Elsevier.

[2] Borys Lypa, Ivan Horyn, Natalia Zagorodna, Dmytro Tymoshchuk, Taras Lechachenko, "Comparison of feature extraction tools for network traffic data," arXiv preprint arXiv:2501.13004, 2025.

[3] Qinwen Hu, Muhammad Rizwan Asghar, Nevil Brownlee, "Evaluating network intrusion detection systems for high-speed networks," in 2017 27th International Telecommunication Networks and Applications Conference (ITNAC), pp. 1–6, 2017, IEEE.

[4] Adeeb Alhomoud, Rashid Munir, Jules Pagna Disso, Irfan Awan, Abdullah Al-Dhelaan, "Performance evaluation study of intrusion detection systems," Procedia Computer Science, vol. 5, pp. 173–180, 2011, Elsevier.

[5] Diego Gonzalez Gomez, "Receive-only UTP cables and Network Taps," 2006, Citeseer.

[6] Cisco Press, "Cisco Field Manual: Catalyst Switch Configuration," 2003, Citeseer.

[7] Jakub Svoboda, Ibrahim Ghafir, Vaclav Prenosil, "Network monitoring approaches: An overview," Int J Adv Comput Netw Secur, vol. 5, no. 2, pp. 88–93, 2015.

[8] Jian Zhang, Andrew Moore, "Traffic trace artifacts due to monitoring via port mirroring," in 2007 Workshop on End-to-End Monitoring Techniques and Services, pp. 1–8, 2007, IEEE.

[9] L. Cottrell, "Passive vs. Active Monitoring [online]," 2015, [cit. 2015-04-21], https://www.slac.stanford.edu/comp/net/wanmon/passive-vsactive.html.

[10] Garland Technology, "Packet Loss | TAP vs SPAN [online]," [cit. 2024-12-26], https://www.garlandtechnology.com/tap-vs-span-packet-loss.

[11] S. Chakrabarti, Mohuya Chakraborty, Indraneel Mukhopadhyay, "Study of snort-based IDS," in Proceedings of the International Conference and Workshop on Emerging Trends in Technology, pp. 43–47, 2010.

[12] Alka Gupta, Lalit Sen Sharma, "Performance analysis and comparison of snort on various platforms," International Journal of Computer Information Systems and Industrial Management Applications, vol. 12, pp. 10–10, 2020.

Сүлжээний Халдлага Илрүүлэх, Сэргийлэх Нээлттэй Эхийн Системийг Ашиглах Судалгаа

Эрдэнийн ЭНХ-ОД
Мэдээллийн сүлжээ, аюулгүй байдлын салбар
ШУТИС, МХТС
Монгол улс, Улаанбаатар
odko1988@gmail.com

Ямхины ДАШДОРЖ
Мэдээллийн сүлжээ, аюулгүй байдлын салбар
ШУТИС, МХТС
Монгол улс, Улаанбаатар
dashdorj@must.edu.mn

Хүчитийн УЯНГА
Мэдээллийн сүлжээ, аюулгүй байдлын салбар
ШУТИС, МХТС
Монгол улс, Улаанбаатар
uyangaakh@must.edu.mn

***Хураангуй*— Сүлжээний урсгалын хэмжээ болон төрөл зүйлийн хувьд өсөн нэмэгдэж буй байдал, мөн идэвхтэй халдагчдын тоо огцом өссөнтэй холбогдуулан төрийн болон хувийн хэвшлийн байгууллагууд сүлжээний халдлага илрүүлэх, сэргийлэх систем (IDS/IPS)-ийг ашиглах зайлшгүй шаардлагад хүрээд байна. Гэвч бидэнд хамгийн түрүүнд тулгардаг асуудал бол төсөв хөрөнгө байдаг нь арилжааны системүүдийг худалдан авч ашиглахад хүндрэлтэй болгож байна. Энэхүү асуудлыг шийдвэрлэх зорилгоор бид энэхүү өгүүлэлдээ нээлттэй эхийн системүүд болох Snort, Suricata, Bro-ийг онцлон авч харьцуулалт, шинжилгээ хийсэн. Энэхүү харьцуулалт, шинжилгээний үр дүнд аль системийг сонгох, системийн тохиргоог сайжруулах, мөн ирээдүйн судалгаанд чиглэл өгөхөд практик ач холбогдолтой гэж үзэж байна. Судалгааны үр дүн нь сүлжээний аюулгүй байдлын практик инженерүүд болон академик судлаачдад бодит хэрэглээний мэдээлэл, техникийн үндэслэлтэй шийдвэр гаргах үндэс суурийг бүрдүүлнэ гэдэгт итгэлтэй байна.**

Түлхүүр үг*— *Open source IDS/IPS, Denial-of-Service attack, Hping3 tool, Suricata

I. Удиртгал

Дэлхий дахинд кибер орон зайн хэрэглээ жил ирэх тусам ихсэж байна. Олон улсын статистик, судалгааны нэгдсэн сайт (Statista)-ын 2025 оны 2 дугаар сарын байдлаар гаргасан тайланд дэлхий даяар 5.56 тэрбум интернет хэрэглэгч байгаа нь дэлхийн хүн амын 67.9 хувийг эзэлж байна. Үүний 5,24 тэрбум нь буюу дэлхийн хүн амын 63,9 хувь нь сошиал медиа хэрэглэгчид бөгөөд үлдсэн 2.63 тэрбум хүн офлайн хэвээр байна [1].

SecureWorks-ийн 2024 оны 11 дүгээр сарын тайланд дурдсанаар дэлхий даяар кибер гэмт хэргийн хохирол 2025 он гэхэд 10.5 их наяд ам.долларт хүрнэ гэж таамаглаж байна. Энэ үр дүн нь санхүүгийн томоохон алдагдалд хүргэж болзошгүй бөгөөд үүнээс урьдчилан сэргийлэхийн тулд албан байгууллагууд IDS болон IPS зэрэг аюулгүй байдлын дэвшилтэт шийдлүүдийг ашиглан арга хэмжээ авах шаардлагатай байгааг харуулж байна [2].

Судалгаанаас үзэхэд IDS шийдлийн халдлага илрүүлэх түвшин нь 80%-иас 98% хүртэл байдаг байна. IPS шийдлийг нэвтрүүлсэн байгууллагын зөрчилд хариу үйлдэл үзүүлэх хугацаа 50%-иас 75% хүртэл буурдаг байна. IPS нь халдлагыг илрүүлж, хамгаалах явцыг автоматжуулснаар илрүүлэхээс эхлээд хариу арга хэмжээ авах хүртэлх хугацааг бууруулдаг юм.

Verizon-оос гаргасан Мэдээллийн зөрчлийн судалгааны тайланд дурдсанаар, зөрчлийн 70 орчим хувь нь гаднаас чиглэсэн халдлагууд байсан бөгөөд IDS болон IPS нь хортой программ, ransomware зэрэг гадны аюулыг илрүүлэхэд чухал үүрэг гүйцэтгэдэг талаар дурдсан [3].

Тиймээс энэхүү өгүүлэлдээ сүлжээний халдлага илрүүлэх, халдлагаас сэргийлэх нээлттэй эхийн системүүдийг аюулгүй байдлын шийдэл болгон ашиглах давуу талуудыг судалгаа, шинжилгээнд тулгуурлан гаргаж, ажиллах зарчмыг харуулан туршилтын ажлыг гүйцэтгэх болно.

II. Нээлттэй Эхийн IDS/IPS-Ийн Тухай

Чөлөөт болон нээлттэй эхийн программ хангамж гэдэг нь хэн ч эх кодыг үзэх, судлах, үнэлэх дүгнэх болон шинэчлэн өөрчлөх боломжтой байдаг программ хангамж юм. Эх код нээлттэй байдаг учир хэн ч сайн дурааараа программ хангамжийн дизайныг сайжруулах үйл ажиллагаанд оролцох боломжийг олгодог бөгөөд хаалттай эхийн программ хангамж нь хувилах эрх нь хязгаарлагдмал, эх код нь хэрэглэгчдэд харагддаггүй, патенттай өмчийн программ хангамжаас ялгаатай юм. Нээлттэй эхийн программ хангамжийг ашиглахын давуу талууд нь программ хангамж бүтээх зардлыг бууруулах, хамгаалалт болон тогтвортой байдлыг өндөрсгөх, хэрэглэгчдэд өөрсдийн техник хангамжаа илүү сайн удирдах боломж олгодог зэрэг юм.

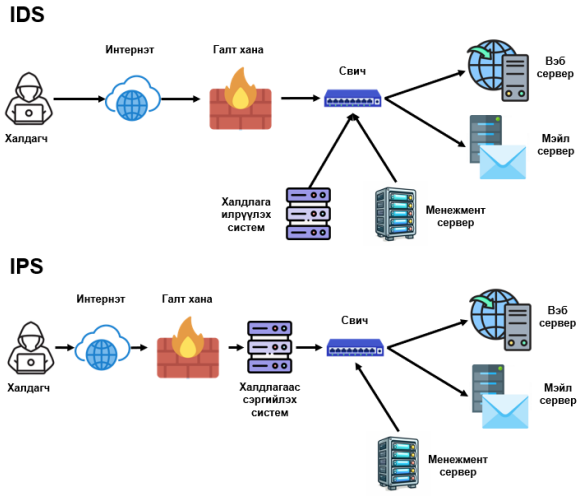


Fig. 1. IDS/IPS-ийн ажиллах зарчим

Мөн нээлттэй эхийн үйлдлийн систем болох Linux нь өнөөдөр олон хүнд ашиг тусаа өгч, серверүүд, суурин компьютер, ухаалаг утсууд зэрэг олон төхөөрөмжүүдийг ажиллуулсаар байна.

Байнгын хувьсан өөрчлөгдөж буй цахим аюулгүй байдлын орчинд хэрэглэгчдийн нууц мэдээлэл, эмзэг датаг хамгаалах нь байгууллага болон хувь хүний хувьд

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

хамгийн чухал зорилт болоод байгаа билээ. Энэхүү аюулгүй байдлын тогтолцооны бүрэлдэхүүн хэсэг нь халдлага илрүүлэх систем (IDS) ба халдлагаас урьдчилан сэргийлэх систем (IPS) юм. IPS/IDS системүүд нь хоёулаа сүлжээнд үүсэж болзошгүй аюулыг тодорхойлох, удирдах, хариу арга хэмжээ авах үүрэгтэй байдаг.

Intrusion Detection System (IDS) – IDS нь сүлжээнд заналхийлсэн аливаа халдлага, сэжигтэй тохиолдлыг илрүүлж анхааруулга, мэдэгдлийг өгдөг болов ч тэдгээрийг блок хийх эсвэл зогсоох арга хэмжээ авдаггүй.

A. Халдлага илрүүлэх систем (IDS)

IDS нь сүлжээний урсгалыг сэжигтэй үйлдэл, аюулгүй байдлын нийтлэг зөрчлийг зөрчиж буй эсэхийг хянаж, болзошгүй халдлага, сэжигтэй үйлдэл сүлжээнд илэрсэн үед мэдээлэх, анхааруулах зарчмаар ажилладаг.

IDS-д гурван үндсэн төрөл байдаг:

Сүлжээнд суурилсан IDS (NIDS): Network based IDS (Сүлжээний орчинд шаардлагатай хэсэгт суурилагдаж сүлжээг бүхэлд нь хянах зориулалттай.)

Хостод суурилсан IDS (HIDS): Host based IDS (Ямар нэгэн хост, сервер дээр тусгайлан суурилагдаж, тухайн суусан хост, серверийн сүлжээнд анализ хийх зарчмаар ажилладаг.)

Вэбд суурилсан IDS (WIDS): Web based IDS (Вэб программууд болон түүний сүлжээний урсгал руу чиглэсэн халдлагыг илрүүлэхэд онцгойлон анхаарч ажилладаг.)

IDS нь сүлжээнд заналхийлсэн аливаа халдлага, сэжигтэй тохиолдлыг илрүүлж анхааруулга, мэдэгдлийг өгдөг болов ч тэдгээрт хаалт хийх эсвэл зогсоох арга хэмжээ авдаггүй.

B. Халдлагаас сэргийлэх систем (IPS)

IDS системээс ялгаатай нь сэжигтэй үйлдлийг илрүүлээд зогсохгүй тэдгээр халдлагаас сэргийлэх, хамгаалах арга хэмжээг авах боломжтой.

IPS нь мөн адил сүлжээнд суурилсан эсвэл хост дээр суурилсан байж болно. IPS нь сүлжээний пакетуудыг шалгаж, сэжигтэй үйлдэл илэрсэн тохиолдолд тухайн пакетыг блок хийх, холболтыг дахин эхлүүлэх, холбогдож буй эх үүсвэрийг хаах зэрэг яаралтай арга хэмжээг автоматаар авдаг ба энэ нь цаашлаад сүлжээнд ажиллаж буй аливаа сервер, сүлжээний төхөөрөмж зэргийг сүлжээний түвшний түгээмэл халдлагад өртөх, цаашлаад Application системийн түвшний халдлагад өртөхөөс урьдчилан сэргийлдэг энгийн бөгөөд үр дүнтэй аюулгүй байдлын системүүд юм.

C. Халдлага илрүүлэлтийн хувьд давуу болон сул талууд

Өгөгдөлд суурилсан (Signature based) илрүүлэлтийн үед худал эерэг үр дүн бага гардаг ч зөвхөн мэдэгдэж байгаа халдлагын өгөгдөл нь илэрч, шинэ бөгөөд хараахан тогтоогдоогүй аюул занал сүлжээний цоорхойг бий болгодог. Гажуудалд суурилсан (Anomaly based) илрүүлэлтийн үед илүү олон худал эерэг үр дүн гарч ирдэг ч зөв тохируулагдсан тохиолдолд алдаагүй, урьд өмнө мэдэгдээгүй аюулыг хүртэл илрүүлж чаддаг.

SNORT

Snort-ийг 1998 онд Мартин Рош бүтээсэн бөгөөд сүлжээний халдлага илрүүлэх, сэргийлэх нээлттэй эхийн систем юм. Snort-ийг гурван өөр аргаар хэрэглэж болно; tcpdump, пакет бүртгэгч эсвэл сүлжээний халдлагыг илрүүлэх, урьдчилан сэргийлэх систем гэх мэт пакет sniffer. Пакет sniffer болгон ашиглах үед Snort нь сүлжээний пакетуудыг уншиж, консол дээр харуулах ба пакет бүртгэгчийн хувьд Snort нь пакетуудыг диск рүү бүртгэх болно. Халдлага илрүүлэх горимд энэ нь сүлжээний урсгалыг хянаж, хэрэглэгчийн тодорхойлсон дүрмийн дагуу урсгалыг шинжлэх болно [6].

BRO

Bro нь 1998 онд Верн Пакссон үүсгэн байгуулагдсан бөгөөд нээлттэй эхийн UNIX дээр суурилсан сүлжээний халдлагыг илрүүлэх систем юм. Bro идэвхгүй сүлжээний урсгалыг хянаж, хортой урсгалыг хайж байдаг. Энэ нь эхлээд сүлжээний траффикийг задлан шинжлэх замаар халдлагыг илрүүлж, дараа нь үйлдлийг хортой гэж үзсэн загвартай харьцуулахад чиглэсэн анализаторуудыг ажиллуулдаг. Шинжилгээнд тодорхой халдлага (гарын үсэг, үйл явдал) болон ер бусын үйл ажиллагаа (гажиг) илрүүлэх зэрэг орно.

Bro нь ихэвчлэн сүлжээний гол уулзвар дээр байрладаг бөгөөд энэ нь ирж буй болон гарч буй бүх урсгалыг хянах боломжтой. Bro нь сүлжээний урсгалыг цуглуулах, шүүх, дүн шинжилгээ хийх зэрэг функцуудыг хангадаг. Энэ нь түгээмэл протоколуудын нарийвчилсан дүн шинжилгээ хийх чадвартай бөгөөд энэхүү шинжилгээний үр дүн нь ажиглагдсан үйл ажиллагааг дүрсэлсэн хэд хэдэн үйл явдал юм [6].

SURICATA

Suricata бол 'Нээлттэй мэдээллийн аюулгүй байдлын сангаас' боловсруулсан нээлттэй эх сурвалжийн халдлагыг илрүүлэх, сэргийлэх систем юм. Бета хувилбар нь 2009 оны 12-р сард гарсан бол анхны тогтвортой хувилбар нь 2010 оны 7-р сард гарсан. Suricata нь халдлагыг илрүүлэх талбарт шинэ санаа, технологийг нэвтрүүлэх зорилгоор бүтээгдсэн.

Нээлттэй Мэдээллийн Аюулгүй байдлын Сан (OISF) нь Suricata-д халдлагыг илрүүлэх, сэргийлэх дүрмийн багцыг өгдөг бөгөөд аюулгүй байдлын оновчтой түвшинг хадгалах үйл явцыг Suricata хөдөлгүүрээр хялбаршуулдаг [6].

Сүлжээнд галт хана нь зөвхөн заасан портууд дээр ирж буй болон гарах сүлжээний траффикийг зөвшөөрөх замаар сүлжээний аюулгүй байдлыг хангах чухал үүрэг гүйцэтгэдэг. Галт хана нь тодорхой порт эсвэл хууль ёсны порт руу илгээсэн сүлжээний траффик эсвэл аливаа халдлага, зөрчлийг илрүүлэх зорилгоор бүтээгдээгүй. [8].

Жишээлбэл, TCP порт 80 дээр дотогшоо нэвтрэх боломжийг олгодог галт ханын дүрэм нь дотоод вэб серверт алсаас хандах боломжийг олгоно. Халдагчид вэб сервер рүү зөвшөөрөлгүй хандахдаа HTTP портыг ашиглаж болно. Энэ хувилбарт IDS нь вэб траффикийн гарын үсэг болон сайн мэддэг халдлагын гарын үсгийн мэдээллийн сангийн хоорондх харьцуулалт хийх замаар хууль ёсны траффик (зөвшөөрөгдсөн холболтууд) болон вэб сервер рүү халдахыг оролдсон хоорондын ялгааг

тодорхойлж чадна. Энд IDS нь сүлжээний администраторт ийм халдлагын талаар мэдэгдэж, зохих арга хэмжээ авахыг анхааруулж, чухал үүрэг гүйцэтгэдэг [9].

Нөгөөтэйгүүр, IPS нь довтолгооны оролдлогын үед зохих арга хэмжээг автоматаар авах замаар чухал үүрэг гүйцэтгэдэг. Жишээлбэл, вэб сервертэй холболтыг таслах / хаах. Тиймээс IDS нь халдлагын оролдлого эсвэл хортой үйлдэл гарсны дараа бүртгэл үүсгэж, сүлжээний администраторт сэрэмжлүүлэх замаар идэвхгүй үйлдэл хийдэг бол IPS нь эдгээр үйл ажиллагааг идэвхтэй хянаж, сүлжээг хамгаалахын тулд эдгээр үйл ажиллагааны эсрэг арга хэмжээ авдаг. Тиймээс IDS болон IPS нь компьютерийн сүлжээг хамгаалах сүлжээний түвшний хамгаалалтын үүрэг гүйцэтгэдэг [10].

TABLE I. НЭЭЛТТЭЙ ЭХИЙН IDS/IPS-ИЙН ЖАГСААЛТ

д/д	*Халдлага илрүүлэх, сэргийлэх систем*	*Үнэлгээ*
1	Snort	9.1
2	Suricata	9
3	Zeek (Bro)	8.9
4	Maltrail	8.5
5	Security Onion	7.6
6	Kismet	7
7	Psad	6.8
8	Sagan	6.3

[a.] Нийт 10 оноонооc үнэлсэн болно.

Дээрх жагсаалтыг үндсэн бүтэц, гүйцэтгэл, дүрмийн зохион байгуулалт, пакетыг унших, сүлжээний хамгаалалтын хяналт, пакетыг гүнзгий судлах чадвар, зэрэг 9 ерөнхий үзүүлэлтээс гаргасан.

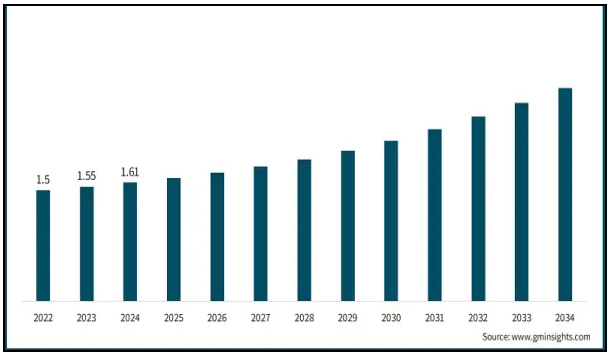


Fig. 2. Худалдааны IDS/IPS-ийн зах зээлийн өсөлтийн үзүүлэлт, 2022-2034, тэрбум ам,доллараар [4]

Дээрх зурагт бид дэлхийн зах зээлд зарагдаж байгаа хаалттай эхийн IDS/IPS-ийн судалгааг жил ирэх тутам тогтмол нэмэгдэх төлөвтэй байгааг харуулсан.

III. Туршилтын Ажил

Уг туршилтын ажлаар үйлдлийн системүүдийг үүсгэж, халдлага илрүүлэх, сэргийлэх системийг суулган тохиргоо хийж, сүлжээнд туршилтаар халдлага хийн үр дүнг харах болно.

TABLE II. SURICATA БОЛОН SNORT ХАРЬЦУУЛАЛТ

Онцлог шинж чанар	*Suricata*	*Snort*
Архитектур	Олон тооны урсгалд ажилладаг	Нэг урсгалтай (Snort 3 нь олон урсгалтай үед ажилладаг)
Гүйцэтгэл	Их ачаалалтай орчинд ерөнхийдөө илүү сайн.	Хязгаарлагдмал нөөцтэй орчинд илүү сайн боловч ачаалал ихтэй орчинд Suricata-аас удаан байж болно.
IDS/IPS	Дэвшилтэт IDS/IPS боломжуудтай.	Дүрэмд суурилсан илрүүлэлт, протоколын шинжилгээнд анхаарал хандуулдаг.
Дүрмийн менежмент	Дүрмийн багцыг удирдахад Suricata-Update-г ашигладаг.	Төхөөрөмж, үйлдлийн систем болон гуравдагч талын хэрэгслүүдтэй илүү нийцтэй байдаг.
Сүлжээний аюулгүй байдлын хяналт (NSM)	JSON (EVE) зэрэг форматаар баялаг NSM өгөгдлийг үүсгэдэг.	Мөн NSM өгөгдлийг үүсгэх боломжтой боловч Сурикатагийн мэдээллийн формат нь илүү уян хатан байдаг.
Packet Capture	Нөхцөлт PCAP хадгалах, зай хэмнэх боломжийг олгодог.	Ихэвчлэн бүх пакетуудыг анхдагч байдлаар хүлээн авдаг.
Deep Packet Inspection (DPI)	DPI болон автомат протокол илрүүлэх боломжийг олгодог.	DPI-д тусдаа модуль эсвэл хэрэгсэл шаардлагатай.
Гүйцэтгэлийн нэмэлт зардал	CPU болон санах ойн нөөцийг их хэмжээгээр зарцуулж болно. Ялангуяа өндөр зурвасын өргөнтэй сүлжээнд ажиллана	Нэг урсгалтай шинж чанараас болж ачаалал ихтэй орчинд ажиллах чадюаргүй.
Вэб интерфэйс	Суурилагдсан вэб интерфэйс байхгүй, голчлон командын мөрийн хэрэгсэл ашиглана.	Тохиргоо хийх, хянах зориулалттай вэб дээр суурилсан интерфэйстэй байдаг.

Дээрх хүснэгтээс үндэслэн халдлага илрүүлэх, халдлагаас сэргийлэх нээлттэй эхийн систем болох Suricata-г сонгох болсон шалтгаан нь суулгаж ашиглахад хялбар, тохиргоо хийхэд маш ойлгомжтой, бодит цаг

хугацаанд аюулыг илрүүлэх, нөөцийг үр ашигтай ашиглах, аюулын олон талт дүн шинжилгээ хийх зэрэг давуу талуудтай юм.

Уг систем (Suricata)-ийн халдлага илрүүлэх чадварыг судалгаанаас харахад секундэд 82,000 гаруй пакет боловсруулах чадвартай. Энэ нь нэг минутын дотор маш олон тооны болзошгүй халдлагыг илрүүлдэг бөгөөд пакетын хоцрогдолд орохгүй гэсэн үг юм.

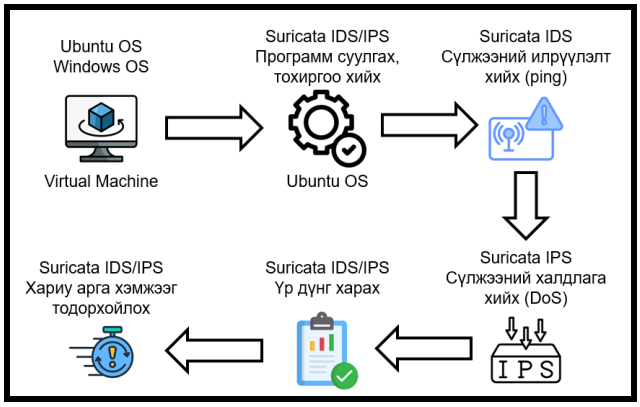


Fig. 3. Туршилтын ажлын ерөнхий топологи зураг

Туршилтын ажлыг гүйцэтгэхэд дараах төхөөрөмж, үйлдлийн систем, халдлагын хэрэгслийг ашигласан. Үүнд:

a) Зөөврийн компьютер (Dell G15, i5-13450, Memory 16gb, Disk 512gb SSD)

b) Виртуаль үйлдлийн систем (VMWare Workstation Pro-д Ubuntu, Windows 10 үүсгэх)

c) Нээлттэй эхийн систем (Suricata)

d) Халдлагын хэрэгсэл (hping3)

e) Халдлагын төрөл (Denial-of-Service /DoS/)

A. *Туршилтын орчныг бэлдэх, халдлага илрүүлэх, сэргийлэх системийг суулгах*

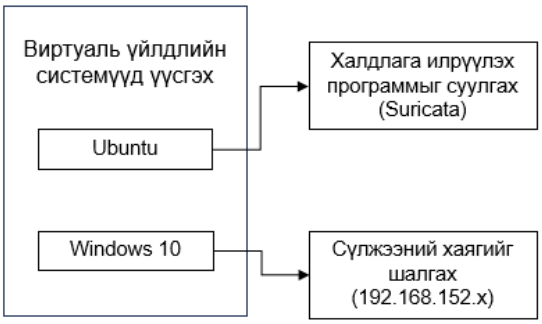


Fig. 4. Виртуаль орчинд үйлдлийн системүүд болон IDS/IPS-ийг суулгах

Командууд: [5]

```
sudo add-apt-repository ppa: oisf/suricata-stable
sudo apt update
sudo apt-get update && sudo apt-get install suricata
```

B. *IDS/IPS тохиргоо хийх*

Fig. 5. Халдлага илрүүлэх, сэргийлэх системд тохиргоо хийх

Командууд: [5]

```
cd /etc/suricata/
sudo nano suricata.yaml
```

Тохиргоо хийсэн байдал:

```
address-groups:
  HOME_NET: "[192.168.152.0/24]"
```

```
af-packet:
  - interface: ens33
```

```
nfq:
  mode: accept
  repeat-mark: 1
  repeat-mask: 1
```

```
  route-queue: 2
```

C. *Suricata-д тохиргоо хийх*

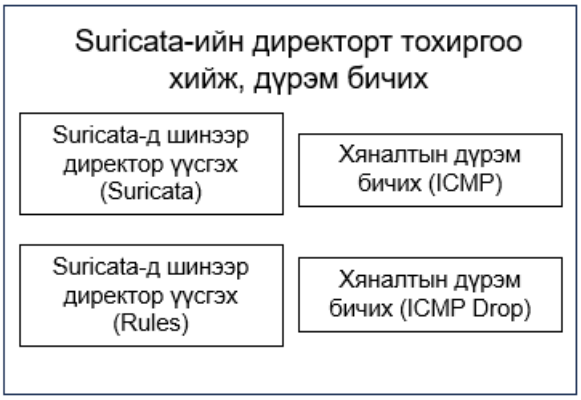


Fig. 6. Suricata болон rules директор шинээр үүсгэж, хяналтын дүрмүүдийг бичих

Командууд: [5]

```
cd /var/lib/
sudo mkdir suricata
cd suricata
sudo mkdir rules
sudo mv custom.rules /var/lib/suricata/rules
```

cd /var/lib/suricata/rules

sudo nano custom.rules

Тохиргоо хийсэн байдал:

```
alert icmp any any -> $HOME_NET any (msg:"ICMP uildel ilerlee!"; sid:123; rev:1;)
```

```
drop icmp any any -> $HOME_NET any (msg:"ICMP Drop hiilee"; sid:124; rev:1;)
```

Сүлжээний халдлага илрүүлэх системийн хувьд виртуаль орчинд үүсгэсэн Windows 10 үйлдлийн системээс сүлжээний урсгал (Ping) шалгахад дараах байдлаар харагдаж байна.

```
03/12/2025-02:30:40.503411  [**] [1:123:1] ICMP uildel ilerlee! [**] [Classifica
tion: (null)] [Priority: 3] {ICMP} 192.168.152.136:8 -> 192.168.152.135:0
03/12/2025-02:30:40.503442  [**] [1:123:1] ICMP uildel ilerlee! [**] [Classifica
tion: (null)] [Priority: 3] {ICMP} 192.168.152.135:0 -> 192.168.152.136:0
```

Сүлжээний халдлагаас сэргийлэх системийн хувьд дараах байдлаар харагдаж байна.

```
03/12/2025-02:30:40.503411  [**] [1:123:1] ICMP uildel ilerlee! [**] [Classifica
tion: (null)] [Priority: 3] {ICMP} 192.168.152.136:8 -> 192.168.152.135:0
03/12/2025-02:30:40.503442  [**] [1:123:1] ICMP uildel ilerlee! [**] [Classifica
tion: (null)] [Priority: 3] {ICMP} 192.168.152.135:0 -> 192.168.152.136:0
03/12/2025-02:53:37.011301  [Drop] [**] [1:124:1] ICMP Drop hiilee [**] [Classif
ication: (null)] [Priority: 3] {ICMP} 192.168.152.136:8 -> 192.168.152.135:0
```

D. *Сүлжээний халдлага хийж турших*

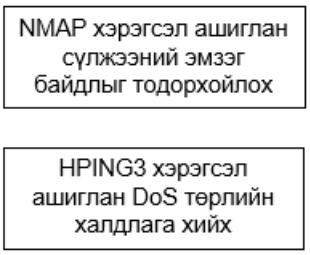


Fig. 7. Nmap хэрэгсэл ашиглан сүлжээний эмзэг портыг олж, халдлага үйлдэж турших

Бидний сонгон авсан халдлагын хэрэгсэл (Hping3) нь эх сурвалжийн IP хаягийг хууран мэхлэхийн зэрэгцээ зорилтот хэрэглэгч рүү их хэмжээний TCP траффик илгээж, энэ нь санамсаргүй эсвэл бүр тодорхой хэрэглэгчийн тодорхойлсон эх сурвалжаас гаралтай мэт харагдуулдаг.

Командууд: [5]

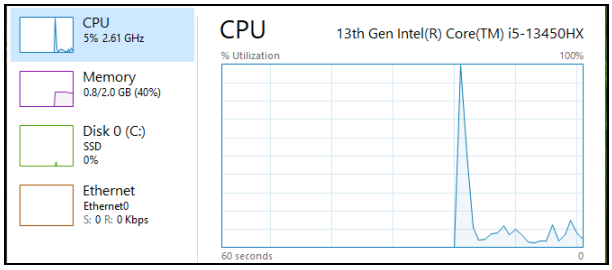
nmap -Pn 192.168.152.136

Гарсан үр дүн:

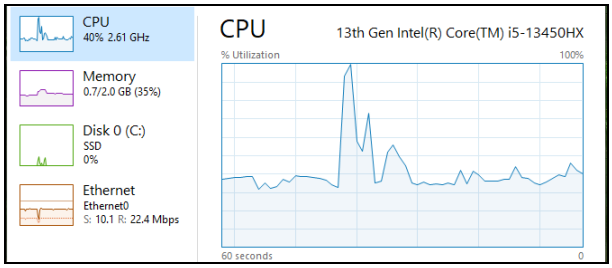
```
Nmap scan report for 192.168.152.136
Host is up (0.00059s latency).
Not shown: 997 closed tcp ports (conn-refused)
PORT     STATE SERVICE
135/tcp open  msrpc
139/tcp open  netbios-ssn
445/tcp open  microsoft-ds
```

sudo hping3 -S –flood -V -p 135 192.168.152.136

Өмнө:



Дараа:



E. *Сүлжээний халдлагыг Suricata ашиглан хянах*

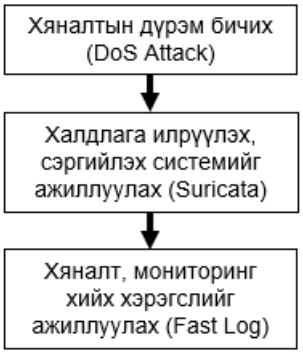


Fig. 8. DoS халдлагыг илрүүлэх дүрмийг бичиж, лог бүртгэлийн хэсэгт үр дүнг харах

Командууд: [5]

alert tcp any any -> $HOME_NET any (msg:"Possible Denial-of-Service attack!"; flags:S; flow:stateless; threshold:type both, track by_dst, count 200, seconds 1; sid:10001; rev:1;)

sudo suricata -c /etc/suricata/suricata.yaml -i ens33

sudo tail -f /var/log/suricata/fast.log

Гарсан үр дүн:

```
03/12/2025-22:47:44.088124  [**] [1:10001:1] Possible Denial-of-Service attack!
[**] [Classification: (null)] [Priority: 3] {TCP} 192.168.152.135:31052 -> 192.1
68.152.136:135
03/12/2025-22:47:45.086664  [**] [1:10001:1] Possible Denial-of-Service attack!
[**] [Classification: (null)] [Priority: 3] {TCP} 192.168.152.135:52762 -> 192.1
68.152.136:135
```

Халдлагаас сэргийлэхдээ сүлжээн дэх шаардлагагүй портууд (Жишээ нь: 135, 139, 445 портууд)-ыг хааснаар DoS болон DDoS халдлага хийх боломжийг хязгаарласан.

```
Nmap scan report for 192.168.152.136
Host is up (0.00059s latency).
Not shown: 997 closed tcp ports (conn-refused)
```

IV. Судалгааны Үр Дүн

Дээрх туршилтаар бид сүлжээний халдлага илрүүлэх, халдлагаас сэргийлэх нээлттэй эхийн системд өөрт тохирох, дурын дүрмүүдийг бичиж, хүссэн хяналтаа зохион байгуулах боломжтойг харуулж байна.

Туршилтын хэрэглэгч рүү зөвхөн нэг хостоос халдлага хийхэд процессорын ачаалал өмнө нь 5 хувь, сүлжээний ачаалал 0 Mbps байсан бол халдлагын үед процессорын ачаалал 40 хувь, сүлжээний ачаалал 22.4 Mbps болж өссөн нь процессор болон сүлжээний ачаалал дунджаар 14 дахин нэмэгдсэн байна. Үүнээс дүгнэхэд олон тооны хостоос зэрэг халдлага хийсэн тохиолдолд тухайн хэрэглэгч эсвэл серверийг үйлчилгээнээс бүр мөсөн гаргах боломжтой гэсэн үг юм. Уг халдлагыг Distributed Denial of Service гэж нэрлэдэг.

Туршилтаар хийсэн халдлагаас гадна Kali Linux үйлдлийн системийн хэрэгслүүдийг ашиглан дараах халдлагуудыг хийх боломжтой. Жишээ нь: Metasploit – Сүлжээний тоог гаргах, хорлонтой ачааллыг сервер рүү илгээх хэрэгсэл, THC Hydra - Алсын зайнаас хост/серверийн нууц үгийг задлахад ашигладаг онлайн нууц үг тайлах хэрэгсэл, Armitage - Сервер рүү халдахад зориулсан Metasploit-ийн графикал дэлгэц, WafW00f – Галт ханын төхөөрөмж илрүүлэх хэрэгсэл, Fierce – Домайн серверийн хэрэглэгчдийн тоог гаргах зэрэг байна.

Сүлжээний орчинд нийт 65535 порт байдгаас өргөн хэрэглэгддэг нь 14 порт байдаг. Үүнээс ашиглагдахгүй байгаа портуудыг хаах нь халдлагад өртөх эрсдлийг тодорхой хувиар бууруулна.

Уг нээлттэй эхийн системийг зөвхөн халдлага илрүүлэхэд ашиглахаас гадна хамгаалах байдлаар ашиглах боломжтой юм.

V. Дүгнэлт

Дэлхий дахины хэмжээнд кибер аюулгүй байдлыг хангах асуудал хурцаар тавигдаж байгаа өнөө үед байгууллагын болон хувийн сүлжээний аюулгүй байдлыг хамгаалах, халдлагаас урьдчилан сэргийлэх маш амар, хялбар шийдэл бол нээлттэй эхийн IDS/IPS шийдэл юм.

Хувь хүн болон албан байгууллагууд зардал гаргалгүй ашиглах программ хангамжийн эх кодуудыг өөрсдөдөө тохируулах, хөгжүүлэлт хийх, хамгаалалтын дүрмүүдийг зохиох боломжийг олгоно.

Уг халдлагын эсрэг системийг илүү хөгжүүлэн хэрэглэгчид ээлтэй (User-friendly) интерфэйстэй болгон ашиглах боломжтой.

VI. Ном зүй

[1] https://www.statista.com/statistics/617136/digital-population-worldwide [Хандсан: 2025.03.08]

[2] https://www.secureworks.com/centers/boardroom-cybersecurity-report-2024 [Хандсан: 2025.03.08]

[3] https://www.verizon.com/business/resources/reports/dbir/ [Хандсан: 2025.03.08]

[4] https://www.gminsights.com/industry-analysis/intrusion-detection-prevention-system-ids-ips-market [Хандсан: 2025.03.14]

[5] https://docs.suricata.io/en/suricata-7.0.2/rules/index.html

[6] Rødfoss, J. T. (2011). Comparison of open source network intrusion detection systems (Master's thesis).

[7] Asad, H., & Gashi, I. (2018). Diversity in Open Source Intrusion Detection Systems. Lecture Notes in Computer Science, 267–281. doi:10.1007/978-3-319-99130-6_18

[8] Sperotto, Anna, Gregor Schaffrath, Ramin Sadre, Cristian Morariu, Aiko Pras, and Burkhard Stiller. "An overview of IP flow-based intrusion detection." IEEE communications surveys & tutorials 12, no. 3 (2010): 343-356.

[9] Li, Wan, and Shengfeng Tian. "Preprocessor of intrusion alerts correlation based on ontology." In 2009 WRI International Conference on Communications and Mobile Computing, vol. 3, pp. 460-464. IEEE, 2009.

[10] Sivaraman, Vijay, Hassan Habibi Gharakheili, Arun Vishwanath, Roksana Boreli, and Olivier Mehani. "Network-level security and privacy control for smart-home IoT devices." In 2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), pp. 163-167. IEEE, 2015.

Халдлага илрүүлэх системийн онцлогийг XAI ашиглан шинжилгээ хийх нь

Н.Угтахбаяр
Монгол Улсын Их Сургууль
ugtakhbayar@num.edu.mn

П.Баясгалан
Монгол Улсын Их Сургууль
ЭХИТ

Г.Манлайбаатар
Хил хамгаалах ерөнхий газар

Д.Баттулга
Монгол Улсын Их Сургууль
ЭХИТ

Б.Өсөхбаяр*
Монгол Улсын Их Сургууль
usukhbayar@num.edu.mn

***Хураангуй*— Хиймэл оюун бидний өдөр тутмын амьдралын энгийн хэрэглээнд өргөн ашиглагдаж буйгаас гадна кибер аюулгүй байдлыг хангах шийдлүүдэд болон кибер халдлага, дайралт хийхэд ашиглагдаад багагүй хугацаа өнгөрсөн. Судлаачид энэхүү чиглэлд судалгааны ажил цөөнгүй хийж байгаагаас гадна жил ирэх тутам судалгааны ажлын тоо улам бүр нэмэгдсээр байна. Сүлжээний орчинд кибер халдлага илрүүлэхэд халдлага илрүүлэх системийн үүрэг чухал байдаг бөгөөд машин сургалтын алгоритм ашигласан халдлага илрүүлэх систем нь зөвхөн өндөр нарийвчлалтайгаар халдлага танихаас гадна тухайн халдлагыг таньж буй үр дүн нь тодорхой байх шаардлага бий болсон. Уг судалгааны ажлын хүрээнд Naïve Bayes ашигласан халдлага илрүүлэх системийн онцлог сонголтыг explainable AI ашиглан тайлбарлана. Үүний үр дүнд кибер аюулгүй байдлын мэргэжилтэн, судлаачдын судалгааны үр дүнг баталгаажуулах боломж нэмэгдэнэ. Бид Naïve bayes, SHAP (SHapley Additive exPlanations) ашиглан CSE-CIC-IDS2018 сангаас халдлагад гол нөлөө үзүүлж буй онцлогыг тодорхойлсон бөгөөд үүний үр дүнд халдлага таних хугацааг 2.6 дахин бууруулсан.**

Түлхүүр үг— Халдлага илрүүлэх систем, Explainable AI, SHAP аргачлал, Naive Bayes алгоритм, онцлог сонгох

I. Удиртгал

Хиймэл оюуны хэрэглээтэй холбоотой кибер халдлагын тоо, шинж чанар дэлхий дахинд хурдацтай хөгжиж, нэмэгдэж байна. Уг халдлагыг зогсоох, хохирол багатай даван туулах хэрэгсэлүүдийн нэг нь сүлжээний халдлага илрүүлэх систем бөгөөд сүлжээгээр дамжин өнгөрч буй халдлагын мэдээлэлд үндэслэн халдлагыг урьдчилан мэдэх, боломжит хариу арга хэмжээ авах нөхцөл бүрдүүлэх, халдлагын тухай мэдээллийг тодорхойлох зэрэг олон давуу талыг уг систем бий болгодог. Судлаачид бодит сүлжээний орчин үүсгэн олон төрлийн халдлагын туршилт хийж уг урсгалд дүн шинжилгээ хийхдээ хиймэл оюун, машин сургалтын арга аргачлалыг ашиглах нь түгээмэл байна [1]. Хэдийгээр энэ чиглэлээр олон судалгааны ажил хийгдэж, амжилттай үр дүн үзүүлж байгаа боловч шинэ тутам гарч буй халдлагыг илрүүлэх, zero-day халдлагыг таних зэрэг нь багагүй нөөц шаардсан хэвээр байна.

Халдлага илрүүлэх системийг илрүүлж буй аргачлалд нь тулгуурлан сигнатурт суурилсан болон гажигт суурилсан гэж үндсэн хоёр төрөлд ангилна [2]. Сигнатурт суурилсан аргач-лал нь zero-day халдлагыг судлан тухайн халдлагынонцлогыг агуулсан буюу тухайн халдлагыг таних боломжтой сигнатурыг үүсгэх замаар халдлагын мэдээлэл бүхий сигнатурын сан үүсгэж уг сантай урсгалын мэдээллийг тулгаснаар халдлагыг таньж илрүүлдэг [3]. Уг аргачлал нь гажигт суурилсан аргатай харьцуулахад илүү хурдан боловч юмсын интернэтийн орчин шиг шинэ, олон төрлийн халдлага бий болж буй нөхцөлд дангаараа төдийлөн үр дүнтэй биш ч бүрэн судлагдсан халдлагыг алдаагүй та-них магадлал нь гажигт суурилсан системээс илүү юм. Гажигт суурилсан системийн хувьд тухайн илрүүлэх урсгалын шинж чанар, өгөгдлийн мэдээлэл,давтагдах байдал зэрэг хэд хэдэн мэдээллийг ашиглан боловсруулалт хийсний үр дүнд халдлага мөн эсэхийг тодорхойлдог [4]. Машин сургалт, хиймэл оюун ашиглан гүйцэтгэж буй судалгааны ажлын хувьд судлаачдын үр дүн халдлагыг таньж буй танилтын хувиар тодорхойлогдож байгаа боловч тухайн моделийг шинэ халдлагад хэр оновчтой ажиллах зэргийг тодорхойлох боломж хомс буюу “хар хайрцаг”-тай төстэй байгаа нь судлаачийн ур чадвар болон сонгон авч буй сургалтын өгөгдөл зэргээс шууд хамааралтай байна [5]. Уг асуудлаас гадна халдлага таних танилтыг өндөр байлгах зорилгоор аль болох оновчтой онцлог сонгох шаардлагатай болж буйтай холбоотой онцлогийн тоо төдийлөн амжилттай цөөрөхгүй байгаа юм. Онцлогийн тоо цөөрөхгүй байгаа нь сүлжээний урсгал нэмэгдэх, өндөр хурдны сүлжээний орчинд ашиглахад хиймэл оюун, машин сургалт ашигласан халдлага илрүүлэх системийг бодит цагийн горимд халдлага илрүүлэхэд хүндрэлтэй болгодог. Машин сургалтын алгоритм ашиглан сүлжээний халдлага илрүүлэх олон шийдлийг судлаачид танилцуулж байгаа хэдий ч тооцооллын өндөр үр ашиг, магадлалын үндэслэлтэй шийдвэр гаргах чадвараараа Naive Bayes зэрэг Машин сургалтын (ML) загварууд илүү байгаа нь зарим судлаачдын судалгааны үр дүнгээр баталгаажиж байна [2][6]. Ихэнх судалгааны ажил KDD-99, NSL-KDD-99 эсвэл тухайн судлаачийн боловсруулсан санг ашиглаж байна. Гэвч CSE-CIC-IDS2018 зэрэг олон онцлог бүхий олон хэмжээст бүтэц бүхий өгөгдлийн хувьд онцлогийг зөв ялган авах, тасралтгүй сайжруулах нь зайлшгүй шаардлагатай ажил ба үүнийг XAI-ийн тусламжтайгаар гүйцэтгэх нь илүү үр дүнтэй болгоно [7].

Корреляц зэрэг онцлог сонгох уламжлалт аргууд нь статистикийн шинжлэх ухааны аргачлалуудыг ашиглаж байгаа бөгөөд ихэнх судлаачдын хувьд сонгогдсон онцлогууд нь яагаад чухал болохыг тайлбарлаж чадах хэдий ч баталгаажуулж чаддаггүй [8]. Иймд, сонгогдсон онцлогууд нь яагаад чухал болохыг тайлбарлах зайлшгүй шаардлага их хэмжээний урсгал, олон хэмжээст мэдээлэлтэй ажилладаг нөхцөлийн нэг болох бодит цагийн горимд халдлага таних шаардлагатай халдлага илрүүлэх системд байна. SHAP болон бусад XAI техникүүд нь загварын таамаглалд онцлог бүрийн оруулж буй хувь нэмрийг тооцоолсноор тухайн онцлог хэр чухал болохыг тайлбарлах боломжийг бүрдүүлдэг [9]. Судлаачид SHAP-ийг гүн сургалтын арга зэрэг неороны мэдээллийг мэдэх боломжгүй загваруудад хэрэглэж

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

байгаа боловч Naive Bayes шиг магадлалын ангилагчуудад онцлог сонголтыг оновчтой болгоход ашиглагдах боломжтой эсэхийг хангалттай судлаагүй байна.

Энэхүү судалгааны ажлын хүрээнд дараах хоёр үр дүнд хүрсэн.

1. SHAP-ийг Naive Bayes загварт нэгтгэн CSE-CIC-IDS2018 өгөгдөлд ялгаатай онцлогуудыг тодорхойлох аргачлал боловсруулсан.
2. SHAP-д суурилсан онцлог сонголт нь илрүүлэлтийн нарийвчлалыг сайжруулж, тооцооллын ачааллыг бууруулж байгааг туршилтаар нотолсон.

II. Судлагдсан байдал

Машин сургалтын алгоритмыг халдлага илрүүлэх системд ашигласан судалгааны ажил цөөнгүй байдаг бөгөөд бид уг ажлын хүрээнд онцлог сонголтыг сайжруулах чиглэлийн судалгааны ажлыг илүү сонгосон ба халдлага илрүүлэх системд машин сургалтын алгоритмын хэрэглээ, халдлага илрүүлэх системийн онцлог сонгох, XAI-ийг халдлага илрүүлэх системд ашиглах нь гэсэн үндсэн гурван бүлгийн хүрээнд судалгаа хийсэн.

Naïve bayes-ийн алгоритмыг халдлага илрүүлэх системд ашигласан судалгааны ажил цөөнгүй байхаас гадна тасралтгүй олон жил энэ чиглэлийн судалгааны ажил хийгдэж байгаагийн нэг нь [6] ажил бөгөөд уг ажлаар халдлага илрүүлэх системийн онцлогийн хамаарал, үр дүнтэй байдлыг тодорхойлсон бол [2] ажлаар судлаачид халдлага илрүүлэх системд олон төрлийн алгоритмуудыг ашиглан харьцуулж, чухал онцлогийг тодорхойлсон байна. Харин [11] ажлаар судлаачид CSE-CIC-IDS2018 санг танилцуулсан бөгөөд уг сан нь орчин үеийн халдлага илрүүлэх системийг үнэлэх гол зорилготой бөгөөд AWS-д суурилсан халдлагын орчин бэлдэж холбогдох мэдээлэл цуглуулсан сан юм. Харин онцлог сонгох чиглэлээр корреляц, рекурсив алгоритмуудыг ашигласан үндсэн хоёр чиглэлийн судалгаа түгээмэл хийгдсээр байна [12].

Кибер аюулгүй байдалд ашиглагдаж буй машин сургалтын загваруудыг ойлгомжтой болгох зорилгоор XAI хэрэглэгдэх нь сүүлийн хэдэн жил эрчимтэй өсч байна. [9] судалгаагаар SHAP (Shapley Additive Explanations) хэмээх тоглоомын онолын үндэслэлтэй, загварын таамаглалыг тайлбарлах аргачлалыг танилцуулсан бөгөөд энэ нь улмаар хортой програм илрүүлэх зорилгоор гүн сургалтын аргаар суралцсан загваруудыг тайлбарлахад өргөн хэрэглэгдэх болсон [13]. Мөн [10] судалгаанд LIME (Local Interpretable Model-agnostic Explanations) аргыг ашигласан бөгөөд уг судалгааны үр дүнд олон хэмжээст өгөгдөлд тогтворгүй байдалтай байдаг нь түүний нэг сул тал гэж дүгнэжээ.

[14] ажлаар SHAP-ыг ашиглан ransomware халдлагыг илрүүлэхэд онцлогуудыг тайлбарлуулсан бөгөөд сургалтын модель, онцлог хоорондын хамаарлыг хазайлтаар тодорхойлсон. Харин [15] ажилд LIME-ийг Random Forest загвартай хослуулан ботнет илрүүлэхэд ашиглаж, тухайн судалгаагаар ботнетийг илрүүлэхэд нөлөө бүхий онцлогуудыг тайлбарлуулан гаргасан.

Хэдийгээр дээрх судалгааны ажлууд нь XAI-ийг олон төрөлд ашигласан боловч XAI-ийг онцлог сонгохд ашиглаагүй байна. Өөрөөр хэлбэл энэ чиглэлийн судалгаанууд нь XAI-г сургасан машин, хиймэл оюуны үр дүнг тайлбарлуулах зорилгоор post-hoc хэблэрээр ашигласан байна. Бидний судалсан сангуудад SHAP-ийг Naive Bayes загвартай системтэй хослуулан, CSE-CIC-IDS2018 өгөгдлийн сантай ажиллаж халдлага илрүүлэх системд ашиглах онцлог сонгосон шийдэл байхгүй байна.

III. Аргачлал

A. Туршилтын сан

Бид уг судалгаанд CSE-CIC-IDS2018 санг ашгласан бөгөөд уг сангийн хувьд bruteforce, DoS, DDoS, botnet, web, infiltration, portscan зэрэг халдлагын мэдээлэл агуулсан бөгөөд AWS-ийн орчинд Виндоус, Линукс үйлдлийн систем ашиглан гүйцэтгэсэн 80 гаруй төрлийн онцлог бүхий сан юм[1] [16]. Бид уг сангийн тэгш хэмт бус (imbalanced) өгөгдлийн онцлогийг машин сургалтад нөлөөлөх нөлөөллийг бууруулах зорилгоор SMOTE аргачлал ашигласан [17].

B. Өгөгдөл боловсруулах

Бид туршилтын сангийн хувьд хоосон талбаруудыг утга нөхөх арга ашиглан гүйцээх ба ангиллын мэдээллийг Naïve Bayes алгоритмд нийцүүлэн тоон хэлбэрт хувиргах зорилгоор one-hot энкодинг ашигласан. Үүнээс гдадна бусад талбаруудад нормчлох аргачлал ашигласан. [2] судалгааны ажлын хүрээнд гүйцэтгэсэн аргачлалыг ашиглан Naïve bayes машин сургасан бөгөөд үүний дараа SHAP XAI ашиглан тус бүрийн шинжид хамаарах утгуудыг тооцоолсон. (1) томъёогоор SHAP-ийн ерөнхий ажиллагааг харуулав.

$$g(z') = \phi_0 + \sum_{j=1}^{M} \phi_j z'_j, \qquad (1)$$

Энэхүү туршилтыг 7 дахь үеийн Intel i7 2.5GHzCPU, 16GB DDR4 RAM, 256GB SSD хатуу диск бүхий энгийн персонал компьютер дээр гүйцэтгэсэн бөгөөд үйлдлийн систем нь Ubuntu, ашигласан програмчлалын хэл Python.

Үр дүн харьцуулах шийдлийг [18] ажлын Аргачлал бүлгийн D хэсэгт заасан алгоритмын шийдлийг шууд ашигласан бөгөөд тооцооллын хэмжүүрээр confusion матриц үүсгэсэн. Үүнийг үүсгэхдээ дараах утгуудыг ашигласан ба Table 1-т зааснаар боловсруулав. Үүнд:

TABLE I. Confusion матриц үүсгэх

	Халдлага гэж тодорхойлсон	***Хэвийн гэж тодорхойлсон***
Халдлага	TP	FN
Хэвийн	FP	TN

True positive (TP) – Сан дахь халдлагыг халдлага гэж зөв ангилсан тоо.

[1] https://www.unb.ca/cic/datasets/ids-2018.html

True negative (TN) - Сан дахь хэвийн урсгалыг хэвийн урсгал гэж зөв ангилсан тоо.

False positive (FP) – Сан дахь хэвийн урсгалыг халдлага гэж буруу ангилсан тоо.

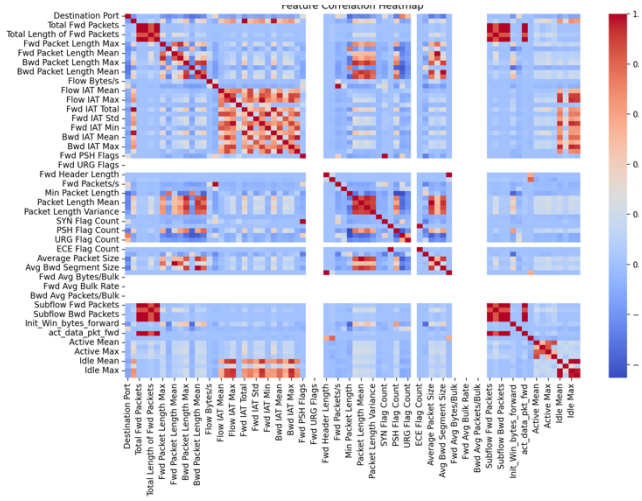
False negative (FN) - Сан дахь халдлагыг хэвийн урсгал гэж буруу ангилсан тоо.

Танилтын нарийвчлалд дараах томъёоллыг ашиглав.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

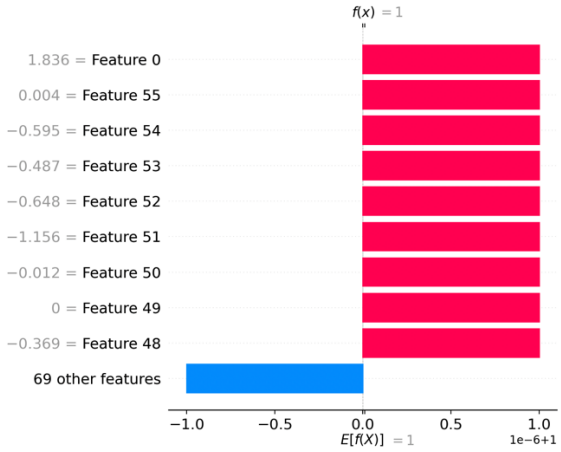
IV. Туршилтын үр дүн

Уг туршилтаар сангийн 80 гаруй онцлогийг бүгдийг ашиглан Naïve Bayes-ийн машин сургаж халдлагыг илрүүлсэн бөгөөд үүнээс халдлагад хамаарал бүхий онцлогийн матрицийг heatmap ашиглан Зураг 1-т харуулав. Уг туршилтын үр дүнд пакетийн нийт дамжуулсан хэмжээ, урт, үндсэн урсгалаас салбарласан урсгалын мэдээлэл, пакетийн хамгийн бага хэмжээ болон дундаж хэмжээ зэрэг нь халдлагыг илрүүлэхэд өндөр нөлөөтэй гэж харагдаж байна.



Зураг 1 Онцлогийн коррелацийн мэдээлэл

Уг туршилтын дараагийн үр дүнд SHAP-ийн үнэлгээг онцлог бүрээр бодоход 9 онцлог халдлагын мэдээллийг тодорхойлоход чухал байсан нь Зураг 2-т харагдаж буй SHAP онооны хэсгээс харагдаж байна.



Зураг 2 SHAP ашиглан онцлогт шинжилгээ хийсэн үр дүн

Дээрх 9 онцлогийг сургасан Naïve Bayes машин болон нийт онцлогийг сургасан Naïve Bayes машины үр дүнг харьцуулан Хүснэгт 2-т харуулав.

TABLE II. Үр дүнгийн харьцуулалт

Модель	*Онцлог сонгосон аргачлал*	*Accuracy*	*Detection time*
Naïve bayes	CSE-CIC-IDS2018 сангийн бүх онцлог	98.4%	3.7 ms
Naive bayes	Сонгогдсон 9 онцлог - Total Fwd Packets - Total Length of Fwd Packets - Packet Length Variance - SYN Flag count - Subflow Fwd Packets - Subflow Bwd Packets - Act_data_pkt_fwd - Idle mean - Idle max	93.6%	1.4 ms

Дээрх хүснэгтээс харахад SHAP ашиглан халдлага илрүүлэх системийн онцлог сонгох нь бусад уламжлалт онцлог сонгох аргачлалтай харьцуулахад илүү үр дүнтэйгээс гадна онцлогийн тоог оновчтой бууруулснаар буурсан онцлогтой харьцуулахад танилтын хувь харьцангуй өндөр байхаас гадна их хэмжээний сүлжээний урсгалыг цаг алдалгүй шинжилхэд шаардлагатай хугацаа буурч байгаа сайн талтай байна.

V. Дүгнэлт

Уг судалгаагаар тэгш хэмт бус өгөгдлөөс онцлог шинжийг оновчтой сонгоход SHAP болон Naïve Bayes ашигласан шийдлийг танилцууллаа. Уг судалгааны үр дүнд халдлагыг CSE-CIC-IDS2018 сангийн 9 онцлогийн тусламжтайгаар дундажаар 1.4ms-ийн хугацаанд 93.6% таних боломжтой болохыг харууллаа. Уг судалгааны үр дүн хангалттай гарсан хэдий ч судалгааны ажлыг сайжруулан бусад алгоритм болон сангуудын түвшинд хийж харьцуулах, бодит сүлжээний урсгалыг CIC flow ашиглан халдлага илрүүлэх машинд сургаж танилтын хурдыг нарийн тооцоолох зэрэг ажлуудыг хийж гүйцэтгэнэ.

Ном зүй

[1] Y. Xin et al., “Machine learning and deep learning methods for cybersecurity,” IEEE Access, vol. 6, 2018

[2] Ugtakhbayar, N., Usukhbayar, B., Baigaltugs, S. (2020). A Hybrid Model for Anomaly-Based Intrusion Detection System. In: Pan, JS., Li, J., Tsai, PW., Jain, L. (eds) Advances in Intelligent Information Hiding and Multimedia Signal Processing. Smart Innovation, Systems and Technologies, vol 157. Springer, Singapore.

[3] Pathan ASK. The state of the art in intrusion prevention and detection. Auerbach publications;2014.

[4] M Naga Surya Lakshmi DYR. A complete study on intrusion detection using data mining techniques. IJCEA. 2015;9.

[5] J. R. Goodall et al., “Situ: Identifying and explaining suspicious behavior in networks,” IEEE Trans. Vis. Comput. Graph., vol. 25, no. 1,pp. 204–214, Jan. 2019.

[6] Rish, I. (2001). An empirical study of the naive Bayes classifier. IJCAI.

[7] Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. IEEE S&P.

[8] Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. Computers & Electrical Engineering.

[9] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. NeurIPS.

[10] Ribeiro, M. T., et al. (2016). "Why should I trust you?" Explaining the predictions of any classifier. ACM SIGKDD. G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*

[11] Sharafaldin, I., et al. (2018). Toward generating a new intrusion detection dataset. ICISSP.

[12] Ambusaidi, M. A., et al. (2016). Building an intrusion detection system using a filter-based feature selection algorithm. IEEE TCC.

[13] Kacper Sokol, Peter Flach. (2020). Explainability fact sheets: a framework for systematic assessment of explainable approaches. FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency

[14] Apruzzese, G., et al. (2023). On the effectiveness of XAI in cybersecurity. Computers & Security.

[15] Alqahtani, H., et al. (2022). Explainable intrusion detection for IoT networks. IEEE IoT Journal.

[16] Iman Sharafaldin, Arash Habibi Lashkari, and Ali A. Ghorbani, “Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization”, 4th International Conference on Information Systems Security and Privacy (ICISSP), Portugal, January 2018

[17] Akdeas Oktanae Widodo., et al. (2024). Machine Learning-Based Intrusion Detection on Multi-Class Imbalanced Dataset Using SMOTE. Procedia Computer Science 234. 578–583

[18] Угтахбаяр, Өсөхбаяр, Нямжав, Байгалтөгс.Халдлага илрүүлэх системийн хосолмол загва-рын судалгаа. Монголын мэдээллийн технологи2017. 2017;1:86-90.

ОРЧИН ҮЕИЙН СИСТЕМИЙН НИЙЛҮҮЛЭЛТИЙН ГИНЖИН ХЭЛХЭЭНИЙ АЮУЛГҮЙ БАЙДЛЫН СУДАЛГАА

Г.Ганчимэг
Мэдээлэл, Холбооны Технологийн Сургууль,
ШУТИС, Улаанбаатар, Монгол улс
ganaa@must.edu.mn

Т.Отгонбаяр
Мэдээлэл, Холбооны Технологийн Сургууль,
ШУТИС, Улаанбаатар, Монгол улс
senarius.s@gmail.com

Хураангуй—Бараг бүх программ хангамжийн төсөлд гуравдагч талын сан, веб сервис эсвэл хэсэгчлэн авч ашигласан кодын хэсэг багтдаг. Программ хангамжийг хөгжүүлэх ийм арга нь функцийг хурдан хөгжүүлэх, одоо байгаа кодыг дахин ашиглах боломжийг олгодог боловч нийлүүлэлтийн гинжин хэлхээний довтолгооны хаалгыг нээж өгдөг. Иймээс дэлхий даяар программ хангамжийн нийлүүлэлтийн гинжин хэлхээний (supply chain) халдлагын талаарх санаа зовоосон асуудал улам бүр нэмэгдээд зогсохгүй, advanced persistent threat (APT) аргатай холбоотой халдлагын аргууд улам боловсронгуй болж байна. Ийм учраас программ хангамжийн аюулгүй байдлын стратегийг хөгжүүлэлтийн эхний үе шатнаас нь дахин үнэлэх шаардлагатай. Энэхүү судалгааны зорилго нь зөвхөн программ хангамж хөгжүүлэгчдэд орчин үеийн программ хангамжийн нийлүүлэлтийн гинжин хэлхээний аюулаас хамгаалах, эрсдлийн хүрээг багасгах зорилготой юм.

Түлхүүр үг— Software Supply Chain, Security, Threat, Threat Mitigation

I. Удиртгал

Хакерууд программ хангамжийн нийлүүлэлтийн гинжин хэлхээг ашиглан системийн админ болон системийн орчны хувьсагчдын мэдээллийг хулгайлан нууцаар бүхэл систем болон сүлжээ рүү нэвтрэх нь жил ирэх бүр нэмэгдэж байна. Халдлагын хүрээ нь зөвхөн нэг системийг сонгохоос эхлэн бүхэл бүтэн сүлжээний дэд бүтэц хүртэл байна [1-6]. Уламжлалт аргаар системийг халдлагаас хамгаалахад ихэвчлэн хост, сүлжээний аюулгүй байдал, firewall, anti-mailware зэрэг арга техникуудэд түлхүү анхаардаг [7]. Дээрх аргууд нь яг систем хөгжүүлэлтийн үеийн аюулгүй байдлын баталгаа бүрэн болж чаддаггүй [6-7]. Орчин үед систем болон сүлжээний аюулгүй байдлын стандарт ерөнхийдөө сайжирсан нь хакерууд программ хангамжийн нийлүүлэлтийн гинжин хэлхээг онилон системд халдах эрсдлийг ихэсгэж байгаа.

Түүнчлэн программ хангамжийн нийлүүлэлтийн гинжин хэлхээ нь өөрөө аюулгүй байдлын арга техникээр төдийлөн сайн хангагдаагүй байдаг нь нөлөөлж байна. Энэ сэдвийн хүрээнд хамгийн том жишээ бол “SolarWinds” гэж нэрлэгдсэн халдлага бөгөөд хакерууд тухайн системийн шинэ хувилбарт аюултай код нэмсэн бөгөөд тухайн шинэ хувилбарыг татан ашигласан халдлагад өртсөн системийн хүрээ нь дэлхий даяар бөгөөд дэлхий дамнасан корпорациуд төрийн байгууллагууд хүртэл багтаж, нийт халдварласан хэрэглэгчдийн тоо 18000 гэсэн тооцоо гарсан байдаг [5]. Орчин үеийн нөхцөлд нийлүүлэлтийн гинжин хэлхээгүй программ хангамж гэж байх бараг боломжгүй тул тухайн системийн аюулгүй байдал болон системд ашиглагдаж буй бүх бүрэлдэхүүн хэсгүүдийн аюулгүй байдал нь өөрөө орчин үеийн нөхцөлд системийн аюулгүй байдлын тулгамдсан асуудлуудын нэг болж байгаа тул системийн хөгжүүлэлтийн амьдралын мөчлөгийн үед заавал аюулгүй байдлын бодлого тодорхойлох болон аюулгүй байдлын сканнер болон солюшн ашиглах шаардлага зайлшгүй тулгарч байгаа юм.

II. Нийлүүлэлтийн гинжин хэлхээний тухай

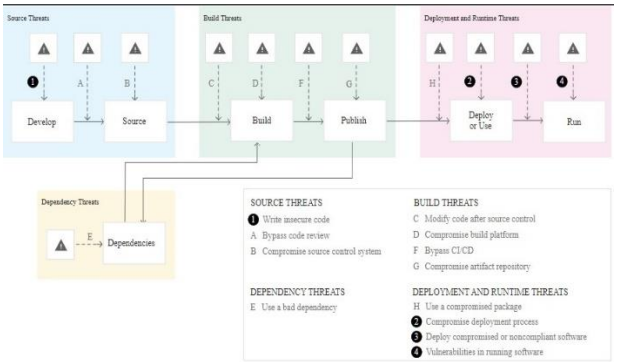
Программ хангамжийн нийлүүлэлтийн гинжин хэлхээ нь эх код, хоёртын файлууд болон бусад бүрэлдэхүүн хэсгүүдээс бүрдэнэ. Үүнд программыг бий болгох, савлах, ашиглахад зориулсан хөгжүүлэлтийн баг, багаж хэрэгсэл, үйл явц, үүнийг танай байгууллагын дотор болон гадна талд ажиллуулахад ашигладаг дэд бүтцийг багтаасан болно. Энгийнээр хэлбэл, программ хангамжийн нийлүүлэлтийн гинжин хэлхээ нь хөгжүүлэлтээс эхлээд үйлдвэрлэл, шинэчлэлт, сайжруулалт хүртэл тухайн кодыг хөндөж байгаа бүх зүйл юм [4]. Ерөнхийдөө программ хангамжийн нийлүүлэлтийн гинжин хэлхээ нь программ хангамжийг хөгжүүлэх бүх амьдралын мөчлөгийн туршид программ хангамжийг бүтээх, үйлдвэрлэх, түгээхэд шаардагдах бүх үйл ажиллагааг нэгтгэдэг гэж болно. *Үүнд [1-6]:*

- Программ хангамжийн хөгжүүлэлтийн амьдралын мөчлөгийн (SDLC) үйл явцын элементүүд, үүнд таны үүсгэсэн код, түүний хамаарал, программ хангамжийг хөгжүүлэх, багц суулгах, туршилтын орчин болон программ хангамжийг ажиллуулахад ашигладаг дотоод болон гадаад программ хангамжууд орно.
- Системд нэвтрэх процесс болон бодлого, тест, *"review"*, хяналт.
- Хөгжүүлэлтэнд ашиглаж байгаа бусад системүүд болон сангууд тэдгээрийг ажиллуулах, *"build"* хийх.

- Нээлттэй эхийн эсвэл гуравдагч этгээдийн программ хангамжийг энтерпрайс системийн нэг хэсэг болгон ашигласан.
- Нээлттэй эхийн платпормыг энтерпрайс системд ашигласан.
- Вэндоруудын мэргэжлийн үйлчилгээнүүд, зөвлөгөө өгөх эсвэл хөгжүүлэлтийн үйлчилгээнүүд бүтэн болон хэсэгчилсэн.
- Энтерпрайс-ууд партнер ашиглан дата хадгалах болон боловсруулах.
- Үүлэн технологийн үйлчилгээнүүд. *Үүнд:* IaaS, PaaS эсвэл SaaS
- Өдийг хүртэл дата болон системийн эрх нь хүчинтэй байгаа хуучин хамтран ажиллагчид, партнерууд.

III. Аюулын бүдүүвч

Аюулын бүдүүвчийн загварчлалын тухайд мэдлэг, сургалт, эрсдлийн үнэлгээ гэсэн уламжлалт гэсэн 3 үндсэн ашиглах арга байдаг [1]. Хэрэв бид зорилго, оролцогч, довтолгоо, TTP (тактик, арга техник, процедур), нийлүүлэлтийн гинжин хэлхээнд хамаарах аюулын үүрэг гүйцэтгэгч гэх мэт өргөн хэрэглэгддэг ойлголтуудыг авч үзвэл, өргөн хэрэглэгддэг STIX [34] аюулын бүдүүвч илүү тохиромжтой болно. Гэсэн хэдий ч энэ сэдвийн хүрээнд бид илүү нийлүүлэлтийн гинжин хэлхээг чиглэсэн *"Supply chain Levels for Software Artifacts"* (SLSA) загварт илүү анхаарлаа хандуулах ба, (SLSA) нь мэдээлэл технологийн салбарын том тоглогчдын зөвшилцлийн дагуу [2] зөвлөмж тогтоон, баталж нийлүүлэлтийн гинжин хэлхээний аюулгүй байдлын удирдамжуудын багц үүсгэдэг фремворк юм. SLSA нь нийлүүлэлтийн гинжин хэлхээг нийтлэг халдлагаас хамгаалахад тусалдаг [3].



Зураг 1. SLSA-д суурилсан аюулын бүдүүвч загвар. Энэ нь ердийн программ хангамжийн нийлүүлэлтийн гинжин хэлхээ болон хөгжүүлэлтийн амьдралын мөчлөгийг харуулсан бөгөөд бүх холбоос дээр тохиолдож болох халдлагын жишээг агуулдаг [3]

Source Threat: Энэ төрлийн аюулууд нь эх кодны бүрэн бүтэн байдалд нөлөөлдөг. Нууцлалын шаардлага хангахгүй код бичих нь программын аюулгүй байдлын цоорхой үүсэх, хүсэгүй үр дүн болон программын цоорхой үүсэх шалтгаан болдог. Мөн хөгжүүлэгчийн хөгжүүлэлтэнд ашиглаж байгаа нууцлалын шаардлага хангаагүй workstation энэ төрөлд хамаарагдана. Энэ тохиолдолд эрсдэл бууруулах аргууд:

- *Хөгжүүлэгчийн workstation дээр нууцлалын бодлогод баримталсан тохиргоо хийх*
- *Cloud workstation заавал нууцлалын бодлогодоо тулгуурлан тохируулах*
- *Local орчинд код скан хийх*
- *Local болон Cloud орчинд аюулгүй байдлын сканнер, api ашиглан realtime feedback авах*

A: Эх кодны контрол менежер лүү чанаргүй код оруулах. Үүнд зөвхөн хортой, асуудалтай код хамаарахгүй, давхар хөгжүүлэгчийн санамсаргүй оруулсан кодоор дамжин халдлагад өртөх боломж нээж өгсөн хэсгүүд хүртэл бас хамаарна. *Жишээ нь:* cross-site scripting. *Эрсдэл бууруулах аргууд:*

- *Хүний оролцоотой код review хийх*
- *Код скан болон линтинг эх кодны контрол менежер болон IDE -д интеграци хийх*

B: Эх кодны контрол менежер нь нууцлалын шаардлага хангаагүй байх. Build pipe-line эх кодны контрол менежер, болон системийн бусад хэсэг рүү хандах хандалтын эрхийг хязгаарлах, давхар MFA ашиглах. Кодны бүрэн бүтэн байдлыг үнэлэхдээ, давхар ашиглаж байгаа script болон тохиргоо эдгээрийг ашиглаж байгаа build, deploy хэсгүүдийг давхар шалгах хэрэгтэй. Энэ шалгалтын хэсгүүдийг Эх кодны контрол менежер болон код review хэсэгт нэгтгэх, ингэснээр source threat эрсдлүүдийг бууруулах болно.

Build threats: Энэ төрлийн эрсдлүүд программийг build болон package хийж байх үед хамаарах, эсвэл эцсийн хэрэглэгчдийг хуурах замаар асуудалтай хувилбарыг ашиглуулах хамаарна.

Deployment and runtime threats: Тухайн build хийхдээ ашиглаж байгаа package хувилбарыг бүрэн зааж өгөөгүй байх. Нэг хувилбарын 1-ээс олон build хийхэд ирээдүйд хийж байгаа build package хооронд зөрүү гарах. Package хувилбар авто сонголт хийхдээ гэмтэлтэй эсвэл зориудаар өөрчилөн системийн аюулгүй байдлыг алдагдуулахаар тохируулсан хувилбар сонгогдох магадлалтай, энэ байдлыг ашиглан хакер build буруу өөрийн өөрчилсөн package сонгуулах боломжтой.

Dependency threats: Package нь программд шууд ашиглаж байгаа болон тухайн package *(пакежуудын)* доош чиглэсэн рекурсив package багтана.

SBOM: Ерөнхийдөө тухайн систем яг ямар хэсгүүдээс бүрдсэн юм бэ гэдгийг харуулсан SPDX, CycloneDX болон SWID өргөтгөлтэй файл юм. Тухайн файлын гол зорилго нийлүүлэлтийн гинжин хэлхээний нээлттэй байдлыг болон аюулгүй байдлыг сайжруулах зорилготой. Ихэвчлэн SPDX болон CycloneDX формат нь түлхүү ашиглагддаг. SBOM файлын үйл ажиллагаанд ашиглагдах элементүүд үндсэн 3 төрөлд хуваагдсан байдаг *(Хүснэгт 1).*

Хүснэгт 1. SBOM файлын категори болон элементүүд

Хамгийн бага элементүүд	
Өгөгдлийн талбарууд	Компонент болгоны үндсэн мэдээллүүд болон түүнийг мөрдөх боломж: Нийлүүлэгч, компонентийн нэр, хувилбар, бусад өвөрмөц утга танигч, компонентүүдийн хамаарлын мэдээлэл, зохиогч,цагийн тэмдэг.
Автоматжуу-лалтын	Автоматжуулалт, үүнд автоматаар үүсгэх болон хөгжүүлэлтийн орчинд

дэмжлэг	масштаб хийх зорилгоор машинаар унших боломж. SBOM үүсгэх, хэрэглэхэд ашигладаг өгөгдлийн формат SPDX, CycloneDX, SWID шошго хэлбэрээр.
Бодит хэрэглээ ба процесс	SBOM хүсэлтүүдийн үйлдлүүдийг зарлах, үүсгэх ба хэрэглээ. Үүнд: Давтамж, гүний мэдээлэл, мэдэгдэж буй үл мэдэгдэх, тархалт ба хүргэлт, Хандалтын хяналт, Алдааг зохицуулах

IV. Нийтлэг халдлагын төрлүүд

Хорлонтой нээлттэй эхийн код: Нээлттэй эхийн кодын эрсдэл нь аюул заналхийлэгчид олон нийтэд нээлттэй кодын модульд хорлонтой код оруулах үед үүсдэг бөгөөд үүнийг үл тоомсорлон хөгжүүлэгчид тодорхой функцуудыг гүйцэтгэхийн тулд кодын үнэгүй блокуудыг хайж, улмаар өөрсдийн дотроо нэмэн ашигладаг [14]. *"Typo squatting"* нь модулиудыг суулгах явцад үсгийн алдааны давуу талыг ашигладаг бол *"combo squatting"* нь модулиудын нэр олон үгнээс бүтсэн үед тухайн үгний дараалын алдааг ашигладаг жишээ нь "nmap-python" гэж бичсэн "python-nmap" [23] *(зураг 4)*. Хорлонтой модулиуд нь дуурисан модулиудын үндсэн функцуулыг агуулах ба нэмээд ихэвчлэн системийн функц, "shell script", "boot" зэргийг ажиллуулан системд хохирол учруулах эсвэл эмзэг дата цуглуулах зориулалтай байдаг [10]. Нээлттэй эхийн энэ асуудал нь нээлттэй биш программ хангамжид ч нөлөөлж болзошгүй, учир нь тухайн программыг хөгжүүлэгчид бүтээгдэхүүндээ нээлттэй эх кодыг ашиглахгүй байх нь орчин үеийн нөхцөлд бараг боломжгүй юм [9].

Hijack the Updates: Халдагчид эцсийн хэрэглэгчдэд зориулсан программ хангамжийг тогтмол шинэчлэх явцад хорлонтой кодыг ашиглан систем рүү нэвтэрч, чухал мэдээлэлд хандах боломжтой [24]. Аюул заналхийлэгчид хохирогчийн сүлжээнд давуу эрх, байнгын хандалтыг олж авахын тулд программ хангамжийн вендор-уудыг хакердах замаар төлөвлөгөөгөө гүйцэтгэнэ. Үндсэн арга нь тухайн вендоруудын шинэчлэлтүүдийг ашиглан эцсийн хэрэглэгчдийн бүх үндсэн системийн хамгаалтууд болох "firewall", "router" гэх мэт тойрох замаар халдлагаа үйлдэнэ [25]. Ийм төрлийн халдлагын жишээ нь "NotPetya", тухайн программ нь Украины нутаг дэвсгэрээс цааш тархаж, олон улсын тээвэрлэлт, санхүүгийн үйлчилгээ, эрүүл мэндийн үйлчилгээ зэрэг дэлхийн томоохон чухал салбаруудад аюул үүсгэсэн [13].

Undermine the Codesigning *(Код зохиогчийн таних тэмдгээр дамжих)*: Код зохиогчийн таних тэмдэг болон кодын бүрэн бүтэн байдлыг баталгаажуулахын тулд кодын тэмдэглэгээг (*Code-signing)* ашигладаг. Халдагчид өөрөө өөртөө гарын үсэг зурах гэрчилгээ *(sefl-sign certificate)*, гарын үсэг зурах, шалгах системийг эвдэх болон цоорхойг ашиглах, эсвэл холбоотой аккоунтын буруу тохиргоог ашиглах замаар зорилгодоо хүрдэг. Дээрх аргыг хослуулан аюул заналхийлэгчид итгэмжлэгдсэн үйлдвэрлэгчийн дүрд хувирч, шинэ хувилбарт хортой код оруулах замаар эсвэл программ хангамжийн шинэчлэлтийг дуурайн эцсийн хэрэглэгчийн системд нэвтэрнэ [11-12]. Мөн аль хэдийн олон нийтэд ил болсон алдаагаар дамжин уг аргыг ашиглаж болно. *Жишээ нь*: Microsoft Windows CryptoAPI модуль нь Elliptic Curve Cryptography (ECC) функцын алдаагаар дамжин сертификетний баталгаатай эсэхийг хуурах арга [15-16].

Халдлагын төрлүүдээс ажиглахад нийлүүлэлтийн гинжин хэлхээний халдлага улам эрчимжин илүү өргөн хүрээнд хохирол учруулах боломжтой болж байгаа тул зайлшгүй халдлагаас сэргийлсэн хариу арга хэмжээ авах нь орчин үеийн системийн хөгжүүлэлтийн нөхцөл байдал болон суурилуулалтын орчин нөхцөлөөс давхар хамааралтай болж байгаа юм.

V. Аюулгүй байдлын фремворкууд

Нийлүүлэлтийн гинжин хэлхээний аюулгүй байдлыг баталгаажуулах үйл явцыг хялбарчлахын тулд олон тооны аюулгүй байдлын тогтолцоог санал болгосон байдаг. Эдгээр нь өөрсдийн гэсэн шилдэг туршлага, багаж хэрэгслийн сонголт, дизайныг санал болгодог. Өөрөөр хэлбэл программ хангамж дахь нийлүүлэлтийн гинжин хэлхээний artifact-уудын бүрэн бүтэн байдлыг хангахад чиглэсэн аюулгүй байдлын бүрэн цогц шийдлүүд юм. Аюулгүй байдлын фреймворкуудыг дараах байдлаар ангилдаг. *Үүнд:*

SCIM: Artifact баталгаажуулах үйл явц хэрхэн явагдахыг тодорхойлдог. Фремворк нь SCIMEvidence, SCIMPolicy, SCIM-Store хэсгүүдээс бүрдэх бөгөөд стандарт artifact data model болон exchange хийх format тодорхойлон тухайн artifact хэсгийг SCIMPolicy ашиглан үнэлсэний дараа SCIM-Store ашиглан model болон бодлогуудаа хадгалдаг. Энэ нь нийлүүлэлтийн гинжин хэлхээний холбоосуудын хооронд нотлох баримтыг жигд дамжуулах боломжийг олгодог [27]. SCIM хөгжүүлэлт нь зогссон бөгөөд залгамж халаа болон Supply Chain Integrity, Transparency, and Trust (SCITT) прожект хөгжүүлэгдэж байгаа бөгөөд одоогоор draft шатандаа явж байгаа юм.

Хүснэгт 2. Аюулгүй байдлын фремворкууд

Вендор	**Фреймворк**
Microsoft	Supply chain integrity model (SCIM)
Google	Supply chain levels for software arifacts (SLSA)
Linux Foundation	The cloud native computing foundation's (CNCF)

SLSA: 4 түвшин бүхий аюулгүй байдлын баталгаа болон хэмжүүртэй байдаг бөгөөд программ хөгжүүлэлтийн capability maturity model (CMM) моделтэй төстэй. 1 дүгээр түвшин нь хамгийн үндсэн аюулгүй байдлын нөхцөлүүдийг шаарддаг бол шат ахих тусам аюулгүй байдлын шаардлага болон нөхцөл чангардаг [9].

CNCF: Эх код, материал, build pipelines, artifacts болон deployment гэсэн үндсэн 5 алхамаас бүрдэнэ [28].

Frameworks		SCIM	SLSA 4	CNCF
Transparency	Artifacts	✓	✓	✓
	Operations	✓	✓	✓
	Actors		✓	✓
Validity	Artifacts	✓	✓	✓
	Operations	✓	✓	✓
	Actors		✓	✓
Separation	Artifacts		✓	✓
	Operations		✓	✓
	Actors			✓

Зураг 2. SCIM, SLSA, CNCF ба Аюулгүй байдлын дизайны шинж чанаруудаар хуваасан байдал

Аюулгүй байдлын дизайны шинж чанарууд: Халдлагын эрсдэлийг бууруулахын тулд нийлүүлэлтийн гинжин хэлхээний бүрэлдэхүүн хэсгүүдийг хамгаалсан байх ёстой. Дэд хэсгүүдийг хамгаалсны үр дүнд халдлагад өртсөн ч системийн бүрэн бүтэн байдлаа хадгалж чадаж байж нийлүүлэлтийн гинжин хэлхээг аюулгүй гэж үзэж болно. Илүү аюулгүй болгохын тулд шинж чанараар 3 хэсэг болгон ангилж болно [29]. *Үүнд:*

Transparency: Нийлүүлэлтийн гинжин хэлхээний хэсэг нэг бүрийн мэдээллийн хүртээмжтэй байдлын хүрээнд ил тод байлгах нь эргээд аюулгүй байдлын эрсдлээс сэргийлдэг. Энэ хүрээнд SBOM ашиглан бүх artifact ил тод болгосноор тухайн metadata өгөгдлүүдийг ашиглан Cyber Supply Chain Risk Mangement (C-SCRM) фремворкуудаар скан хийх болон эрсдэлийн менежмент анализ хийх бүрэн боломжтой болох юм.

Validity: Үйл ажиллагаа, artifacts бүрэн бүтэн байдал болон тэдгээрийн authentication аюулгүй байдал хангагдсан нийлүүлэлтийн гинжин хэлхээний хувьд гадны халдлагаас үүдсэн өөрчлөлт хийгдэх боломжийг хязгаарласан байх хэрэгтэй. Үүний хүрээнд зөвхөн эрх бүхий этгээд тухайн өөрчлөлтүүдийг хийх нөхцөлөөр хязгаарлана гэсэн үг юм.

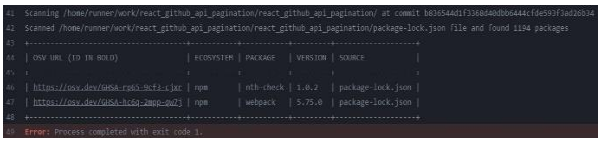
Separation: Аюулгүй нийлүүлэлтийн гинжин хэлхээ нь хэсэгчилсэн шинж чанартай байдаг бөгөөд тэдгээрийн хоорондын холболт нь салшгүй хэсэг юм. Гэхдээ тухайн холбоо нь үргэлж биш зөвхөн хэрэгтэй үед хийгдэх нь зүйтэй. Mirroring, version locking, containers гэх мэт ашиглан separation хийж болно. Npm хувьд scope, proxy ашиглаж болно. Intoto ашиглан хөгжүүлэлтэнд оролцож байгаа этгээдүүдээр separation хийх боломжтой.

VI. Халдлагын эрсдэл бууруулах стратеги

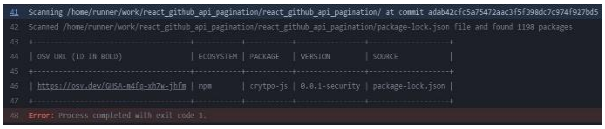
Автоматжуулалт/Automation & linter: Ерөнхийдөө аль хэдийн мэдэгдэж байгаа секюрити цоорхойнуудыг ашиглахаас зайлсхийхийн тулд программд ашиглаж байгаа бүх модиулуудаа сканнердах *(аюулгүй байдлын сканнер)*, ба давхар тасралтгүй интеграцийн (CI) хэрэгслийг ашиглах хэрэгтэй.

- Модулиудын түгжээний файлууд нь автоматаар үүсгэгддэг бөгөөд хүний нүдээр унших, засвар үйлчилгээ хийх зориулалттай биш тул *(зураг 3-7)*. “Lock file injection” [27-28]-ээс сэргийлэхийн тулд заавал “linter” ашиглан шалгах хэрэгтэй.
- “Arbitery command execution” [29] ашигласан модулиудыг аль болох ашиглахаас зайлсхийх хэрэгтэй эсвэл нэмэлт команд ашиглан [30] алгасах, бүр дараагийн алхам нь прожектийн “root” дотор ".npmrc" төслийн файл *(зураг 6)* үүсгэн "ignore-scripts" мөр нэмэх арга хэмжээ авч болно эсвэл глобал npm тохиргоонд энэ өөрчлөлтийг нэмсэнч болно.
- "Dependency confusion attack" [32] энэ төрлийн халдлагаас сэргийлэхийн тулд программд ашиглагдаж буй модулиудыг сохроор шинэчлэн суулгахаас зайлсхийх хэрэгтэй. Дотоод хувийн багцын нэрийг ашиглах тохиолдолд илүү өндөр хувилбартай нийтийн багцын нэртэй давхцаж байгаа нь багц менежерийг нийтийн багц руу шилжүүлэхэд хүргэж болзошгүй бөгөөд аюулгүй байдлын болзошгүй эрсдэлд хүргэнэ [31].
- Trojans source [32] халдлагаас зайлсхийж, хоёр чиглэлтэй (BIDI) юникод хувьсагчдыг шалгах эсвэл идэвхгүй болгох *(зураг 3 ба 5)*, GitHub нь (BIDI) ашигласан байна гэж анхааруулга харуулах боловч зөвхөн файл тус бүрийн хувьд харуулдаг.

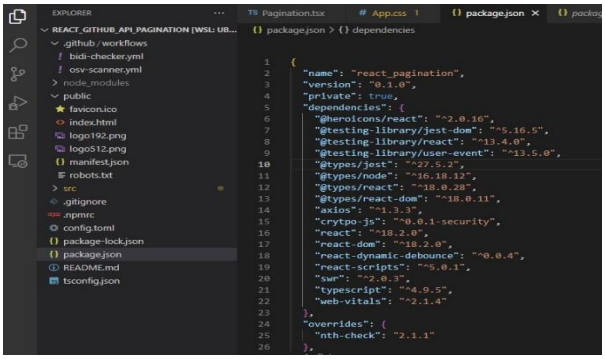
Зураг 3. BIDI checker github action ажилласны үр дүнд Trojans source илрүүлсэн байдал, ашигласан жишээ код [33]



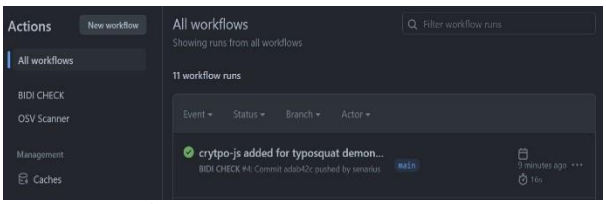
Зураг 4. osv-scanner хуучны модулиудын цоорхойг илрүүлсэн байдал, ашигласан жишээ код [33]



Зураг 5. osv-scanner typosquat ашигласан модуль илрүүлсэн байдал, ашигласан жишээ код [33]



Зураг 6. github дээр action автоматжуулалт нэмсэн байдал, nth-checker хуучин хувилбарыг хүчээр package.json файл дотор "override" дээшлүүлэн аюулгүйн байдлын эрсдлийг зассан байдал, “./.npmrc” тохиргоо нэмэн arbitrary command execution эрсдлээс сэргийлсэн, “./config.toml” osv-scanner дээрх зарим нуулцалын алдааг энэ тохиргоонд бичиж алгасах боломжтой, ашигласан жишээ код [33]



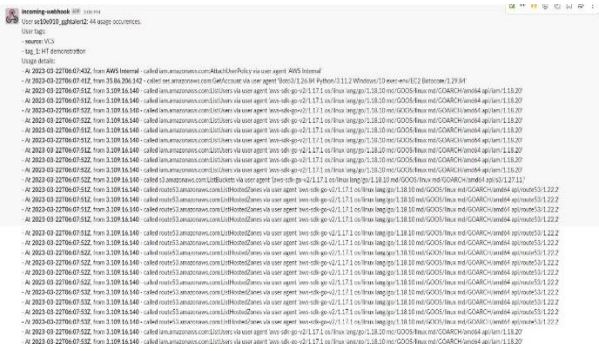
Зураг 7. After remediation all action works without error, demonstration source code included in [33]

Defense in Depth: Defense in Depth (DiD) нь аюулгүй байдлын ерөнхий байдлыг нэмэгдүүлэхийн тулд давхар хамгаалалтыг нэмэх зарчим юм. Өөрөөр хэлбэл, хэрэв халдлага нь хамгаалалтын нэг механизмыг доголдуулахад хүргэсэн бол бусад арга хэмжээ нь халдлагыг цаашид таслан зогсоох, бүр урьдчилан сэргийлэх арга хэмжээ авах болно [22].

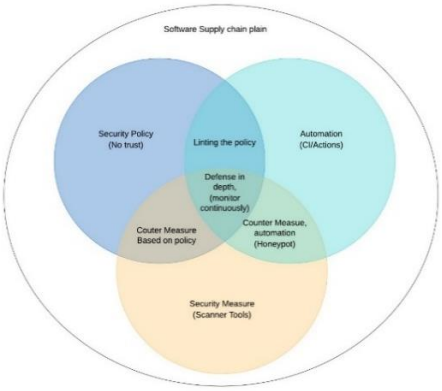
Zero Trust: Zero Trust (ZT) нь хэзээ ч аль нэг системийн "resource" итгэл үзүүлэхгүй, харин "resource"-т хандах хүсэлт тус бүрийн параметрүүдийг тасралтгүй хянах зарчим дээр тулгуурлан ажилладаг бөгөөд "resource" болгонд хандалт хийх хатуу протоколууд бий болгох зорилготой юм. ZT-ийн хүрээнд "resource" гэдэг нь *(программ, сүлжээний хандалт, дата хадгалалт, гэх.мэт)* хамаарна [8]. Энэ зарчмын хүрээнд дотоод сүлжээ болон дотоод хандалт, гадаад сүлжээ, гадаад сүлжээгээр дамжсан хандалтаас илүү аюулгүй гэж ялгаж авч үзэхгүй. Бид дотоод сүлжээ хэзээ нэгэн цагт заавал кибер халдалгад өртөнө гэж бодон үргэлж "authenticate" хийнэ гэж ойлгох хэрэгтэй. Энэ зарчмын хувьд multi factor authentication (MFA) [19-20], чиг үүргийн тусгаарлалтыг *(micro service, user access, service account)* эдгээр дээр заавал хэрэгжүүлэх болон хэрэглээний хяналтыг *(monitoring)* нэмэх шаардлагатай.

Honeypot: "Honeypot" нь кибер халдагчдыг системд нэвтэрсэн үед эрт илрүүлэх болон мэдэгдэх түүнчлэн эрт хариу арга хэмжээ авч халдлагын цар хүрээг хянахад зориулсан жижиг систем. "Honeypot" гол үүрэг нь ихэвчлэн үндсэн системээс тусдаа байршин халдлагад өртсөн тохиолдолд халдлагын бай болох зорилготой ба давхар мэдээлэл цуглуулах болон шууд хөгжүүлэгчдэд мэдэгдэх үүрэгтэй [21]. Хөгжүүлэгчдэд мэдэгдэхийн тулд урхи болгон микро вэб үйлчилгээг бэлтгэж, уг системийн орчны хувьсагчийг зорилтод эх код руу байршуулах замаар ажилладаг. Кибер халдалгад өртсөн тохиолдолд халдагчдын анхаарлыг өөртөө татан, үндсэн системд цохилт өгөхөөс, аюул учруулахаас өмнө тухайн довтолгооны талаар эрт сэрэмжлүүлж, халдагч болон тэдний аргын тухай мэдээллийг хамгаалалтын нэг хэсэг болгон цуглуулах болно. 8 дугаар зурагт үзүүлснээр AWS-н "access-key-id" болон "access-key-secret"-ийг үндсэн системийн эх код дотор өгөөш болгож, "Honeytoken"-ийн хувьд AWS дээр вэб сервисээ "lambda" хэлбэрээр байршуулан "slack" руу "webhook" хэлбэрээр мэдээллэх байдлаар тус тус тохируулсан болно [34-35].

Зураг 8. Халдлагад өртсөн мэдэгдэх лог Slack webhook ашиглан realtime-р мэдэж болно, ашигласан жишээ код “.env.aws” [33]



Зураг 9. Бодит байдлын туршилт AWS-ийн credentials зориудаар ил тавин шалгасан байдал. Зориудаар aws-ийн credentials-уудыг ил тавихад 4мин хугацаанд 100+ өөр халдлагын лог бүртгэгдсэн.



Зураг 10. Халдлагын эрсдэл бууруулах стратегийг Venn диаграмм дээр буулган уялдаа холбоог харуулсан байдал

Халдлагын эрсдлийг бууруулах стратеги баримталснаар хөгжүүлэлтийн бүхий л үе шатуудад тодорхой түвшний эрсдлийг бууруулах бөгөөд халдлагын төрлүүдийг танин мэдэж зөвхөн тухайн халдлагад зориулсан арга хэмжээ авах нь халдлагын эрсдэл бууруулах хамгийн эхний бөгөөд түгээмэл үндсэн хэлбэр юм. Мөн систем үргэлж ямар нэгэн байдлаар халдлагад өртөнө гэсэн бодолтой байх ба системээ сегментүүдэд хуваан, тухайн сегментүүдийн хандалтыг багасгах замаар халдлагын цар хүрээг бууруулах хэрэгтэй. Эцэст end-to-end шифрлэлтийг хийн, тухайн шифрлэлтийг бусдад ил тод байдлаар баталгаажуулах нь чухал юм.

Дүгнэлт

Аюулгүй байдлын сканнер, хамгаалалтын хэрэгслүүд, автоматжуулалт, аюулгүй байдлын бодлогыг хэрэгжүүлэн системийг байнга хянаж ашигласнаар нийлүүлэлтийн гинжин хэлхээний аюулгүй байдлын хүрээнд болзошгүй халдлагын нөлөөллийг үр дүнтэй бууруулж чадна. Аюулгүй байдлын бодлого нь 10 дугаар зурагт үзүүлснээр "Zero Trust" ба "Defense in Depth" (DiD) арга барилыг баримтлах хэрэгтэй. Мөн программ хангамжийн

шаардлага, төсөв, хөгжүүлэлтийн багийн туршлага зэрэг хүчин зүйлс дээр үндэслэн аюулгүй байдлын арга хэмжээ, аюулгүй байдлын сканнер хэрэгсэл зэргийг сонгох хэрэгтэй. Автоматжуулалт нь хөгжүүлэлтийн багийн тодорхойлсан аюулгүй байдлын арга хэмжээний аудитор, аюулгүй байдлын хэрэгжүүлэлтийн стандартын үүргийг гүйцэтгэнэ. Хэдийгээр эдгээр аргууд нь нийлүүлэлтийн гинжин хэлхээний халдлагын магадлалыг мэдэгдэхүйц бууруулж чаддаг ч аюулгүй байдлын бүх систем нь хамгийн сул холбоос шиг хүчтэй байдаг тул бүрэн урьдчилан сэргийлэх баталгаа байхгүй гэдгийг анхаарах нь чухал юм. Гэсэн хэдий ч бидний санал болгож буй арга нь халдлагад өртсөн тохиолдолд хөгжүүлэгчид халдлагыг багасгах үр дүнтэй сөрөг арга хэмжээ болж чадна.

НОМ ЗҮЙ

[1] P. Ladisa, H. Plate, M. Martinez, and O. Barais. Taxonomy of Attacks on OpenSource Software Supply Chains, 2022

[2] "Threat Model" SLSA. https://slsa.dev/spec/v0.1/threats (Accessed: 2023-03-07).

[3] "Requirements." SLSA. https://slsa.dev/spec/v0.1/ requirements (Accessed: 2023-03-07).

[4] "2.1 Digital Supply Chain Security" Securing the Nation's Critical Infrastructures: A Guide for the 2021-2025 Administration. (n.d.). United States: CRC Press.

[5] R. Alkhadra, J. Abuzaid, M. AlShammari, and N. Mohammad, "Solar winds hack: In-depth analysis and countermeasures," in Proc. 2021 12th Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT), pp. 1–7, doi: 10.1109/ICCCNT51525. 2021.9579611.

[6] U.S. Department of Homeland Security, "Security in the softwarelifecycle: Making software development processes—and software produced by them—more secure. DRAFT Version 1.2. ," 2006, (Accessed: 2023-01-28). [Online]. Available: http://www.cert.org/books/secureswe/SecuritySL.pdf.

[7] R. A. Khan, S. U. Khan, H. U. Khan, and M. Ilyas, "Systematic mapping study on security approaches in secure software engineering," IEEE Access, vol. 9, pp. 19139–19160, 2021, doi: 10.1109 (Accessed. 2023-02-26).

[8] Scott Rose, et al., "Zero Trust Architecture," NIST SP 800-207 (August 2020), https://doi.org/10.6028/NIST.SP.800-207.

[9] T. Herr, J. Lee, W. Loomis, and S. Scott, "Breaking trust: Shades of crisis across an insecure software supply chain," https://www.atlanticcouncil.org/in-depth-research-reports/report/breakingtrust-shades-of-crisis-across-an-insecure-software-supply-chain/, 2020, (Accessed: 2023-03-13).

[10] "Malicious packages found to be typo-squatting in Python Package Index." https://snyk.io/blog/malicious-packages-found-to-be-typo-squatting-in-pypi/. (Accessed: 2023-03-10).

[11] ENISA, "Valid Digital Certificates Code Signing Malware," European Union Agency for Cybersecurity, June 30, 2018, https://www.enisa.europa.eu/ publications/info-notes/valid-digital-certificates-code-signing-malware.

[12] CISA, "Alert (AA20-014A): Critical Vulnerabilities in Microsoft Windows Operating Systems," Cyber and Infrastructure Security Agency, January 14, 2020, https://us-cert.cisa.gov/ncas/alerts/aa20-014a.

[13] Fayi, Sharifah. (2018). What Petya/NotPetya Ransomware Is and What Its Remidiations Are. 10.1007/978-3-319-77028-4_15.

[14] Wei, Zhaohui. (2011). Research on the Application of Open Source Software in Digital Library. Procedia Engineering. 15. 1662-1667. 10.1016/j.proeng.2011.08.310.

[15] CERT/CC Vulnerability Note VU#849224, https://kb.cert.org/vuls/id/849224/

[16] CERT/CC Vulnerability Note VU#491944, https://kb.cert.org/vuls/id/491944/

[17] A distributed vulnerability database for Open Source, https://osv.dev/

[18] OpenSSF, https://securityscorecards.dev/

[19] Williamson, Joseph & Curran, Kevin. (2021). Best Practice in Multi-factor Authentication. Semiconductor Science and Information Devices. 3. 10.30564/ssid.v3i1.3152.

[20] "Mfa is a necessary chore! exploring user mental models of multi-factor authentication technologies," https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1649&context=hicss-53, 2020, (Accessed: 2023-02-10).

[21] Titarmare, Neha & Hargule, Nayankumar & Gupta, Anand. (2019). An Overview of Honeypot Systems. International Journal of Computer Sciences and Engineering. 7. 394-397. 10.26438/ijcse/v7i2.394397.

[22] Stytz, Martin. (2004). Considering Defense in Depth for Software Applications. Security & Privacy, IEEE. 2. 72 - 75. 10.1109/MSECP.2004.1264860.

[23] Vu, D.L., Pashchenko, I., Massacci, F., Plate, H., Sabetta, A.: Typosquatting and combosquatting attacks on the python ecosystem. In: 2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW). pp. 509–514. IEEE (2020)

[24] Yang, Jeong & Lee, Young & McDonald, Arlen. (2022). SolarWinds Software Supply Chain Security: Better Protection with Enforced Policies and Technologies. 10.1007/978-3-030-92317-4_4.

[25] Cybersecurity and Infrastructure Security Agency (CISA), Defending against Software Supply Chain Attacks https://www.cisa.gov/sites/default/files/publications/defending_against_software_supply_chain_attacks_508_1.pdf, 2021, (Accessed: 2023-03-14).

[26] George Karantzas, Constantinos Patsakis, 2021, an Empirical Assessment of Endpoint Detection and Response Systems against Advanced Persistent Threats Attack Vectors. https://doi.org/10.3390/jcp1030021.

[27] Bos, A. M. (2023). A Review of Attacks against Language-Based Package Managers. ArXiv. /abs/2302.08959

[28] Arbitrary command execution, vulnerability, https://nvd.nist.gov/vuln/detail/cve-2022-23812, (Accessed: 2023-03-21).

[29] Npm docs, ignore script option, https://docs.npmjs.com/cli/v7/commands/npm-find-dupes#ignore-scripts, (Accessed: 2023-03-18).

[30] 3 Ways to Mitigate Risk Using Private Package Feeds, Secure your hybrid supply chain, (2021) https://azure.microsoft.com/mediahandler/files/resourcefiles/3-ways-to-mitigate-risk-using-private-package-feeds/3%20Ways%20to%20Mitigate%20Risk%20When%20Using%20Private%20Package%20Feeds%20-%20v1.0.pdf, (Accessed: 2023-03-19).

[31] A Confusing Dependency, vulnerability, https://cve.mitre.org/cgi-bin/cvename.cgi?name=2021-29427, (Accessed: 2023-02-21).

[32] Boucher, N., & Anderson, R. (2021). Trojan Source: Invisible Vulnerabilities. ArXiv. /abs/2111.00169

[33] Demonstration, Otgonbayar Tuukhbaatar, https://github.com/ senarius/react_github_api_pagination

[34] STIX: Assets Affected in an Incident. Available online: http://stixproject.github.io/documentation/idioms/affected-assets/, (Accessed: 2023-03-21).

[35] Honeytoken notifier system, GGCanary, GitGuardian, Available online: https://github.com/GitGuardian/ggcanary, (Accessed: 2023-03-22).

Вэб аппликейшнууд дахь олон шатлалт баталгаажуулалт(MFA)-ыг тойрох халдлагын туршилт ба хамгаалалтын аргачлалуудыг боловсруулах нь

Лхагваагийн Одончимэг
Мэдээллийн Сүлжээ, Аюулгүй Байдлын салбар
ШУТИС, МХТС
Улаанбаатар, Монгол
odno@must.edu.mn

Шүрэнцэцэгийн Нандиндулам
Мэдээллийн Сүлжээ, Аюулгүй Байдлын салбар
ШУТИС, МХТС
Улаанбаатар, Монгол
nnandia370@gmail.com

***Хураангуй*— Олон шатлалт баталгаажуулалт (MFA) нь вэб аппликейшнуудын аюулгүй байдлыг нэмэгдүүлэх гол механизм боловч халдагчид үүнийг тойрох шинэ аргачлалуудыг тасралтгүй хөгжүүлж байна. Энэхүү судалгааны зорилго нь MFA-г тойрох халдлагын аргуудыг тодорхойлох, турших болон хамгаалалтын үр дүнтэй арга хэмжээнүүдийг боловсруулах явдал юм. Судалгааны хүрээнд фишингийн аргаар баталгаажуулалтын токен хулгайлах, MITM (Man-in-the-Middle) халдлага хийх, сесс хулгайлах, мөн нөөц баталгаажуулалтын сул талуудыг ашиглах зэрэг MFA-г тойрох түгээмэл аргуудыг туршсан. Туршилтын үр дүнгээс харахад SMS-д суурилсан OTP баталгаажуулалт халдлагад өртөмтгий байгааг тогтоолоо. Мөн халдлагын үр нөлөөг нэмэгдүүлэхэд нийгмийн инженерчлэл болон фишингийн хэрэгслүүдийг ашиглаж болох нь ажиглагдсан. Судалгааны үр дүнд үндэслэн MFA хамгаалалтыг сайжруулах хэд хэдэн аргачлалыг санал болгож байна. Энэхүү судалгаа нь олон шатлалт баталгаажуулалтын хамгаалалтын цоорхойг багасгах, MFA хэрэгжүүлэлтийг сайжруулахад хувь нэмэр оруулах зорилготой юм.**

Түлхүүр үг— Вэб аппликейшн, MFA, фишинг, нийгмийн инженерчлэл

I. УДИРТГАЛ

Сүүлийн жилүүдэд олон шатлалт баталгаажуулалт (MFA) нь вэб аппликейшнуудын аюулгүй байдлыг хангах гол хэрэгслүүдийн нэг болж хөгжсөн. Энэ нь зөвхөн нууц үг ашиглах хамгаалалтаас давж, нэмэлт баталгаажуулалтын давхаргыг нэвтрүүлснээр хэрэглэгчдийн мэдээллийг хамгаалах үр дүнтэй аргачлал болж байна. Гэсэн хэдий ч халдлагын арга техникүүд тасралтгүй хөгжиж байгаатай холбоотойгоор MFA-г тойрох боломжтой шинэ халдлагууд илэрсээр байна. Иймд эдгээр халдлагыг тодорхойлох, тестлэх, хамгаалах аргачлалыг судлах нь мэдээллийн аюулгүй байдлыг сайжруулахад чухал ач холбогдолтой юм.

MFA нь хэрэглэгчийг баталгаажуулахын тулд хэд хэдэн төрлийн танилт хийх аргыг (жишээ нь, нууц үг, нэг удаагийн код, биометр, физик төхөөрөмж гэх мэт) хослуулан ашигладаг. Гэвч кибер халдлагын шинэ арга техникүүдийн хөгжлийн улмаас MFA-г тойрох хэд хэдэн халдлагын төрлүүд бий болсон. Үүнд:

Фишинг халдлага: Хуурамч вэбсайтууд болон хортой программ ашиглан хэрэглэгчийн MFA кодыг хулгайлах.

Сесс булаах: Нэгэнт баталгаажсан хэрэглэгчийн сессийг булаах замаар халдагч системд хандах.

Man-in-the-Middle (MitM) халдлага: Шифрлэгдээгүй эсвэл хамгаалалт муутай холбоосыг ашиглан MFA баталгаажуулалтын мэдээллийг олж авах.

SIM Swapping: Хэрэглэгчийн SIM картын хяналтыг авах замаар нэг удаагийн кодыг хянах.

Эдгээр халдлагууд нь MFA хамгаалалтын арга хэмжээг сайжруулах шаардлагатайг харуулж байна. 2024 оны байдлаар, томоохон (10000-аас дээш ажилтантай) компаниудын 87%, дунд хэмжээний (26-100 ажилтантай) компаниудын 34%, жижиг (25 хүртэл ажилтантай) бизнесүүдийн 27% нь MFA ашигладаг болохыг тогтоосон[1][2]. Бусад компаниуд нь 2FA ашигладаг эсвэл өөр нэмэлт баталгаажуулалт ашигладаггүй гэж мэдээллэсэн. Одоогийн байдлаар хамгийн сайн хэмээн нэрлэгдэж буй MFA баталгаажуулалтын хувьд автоматжуулсан кибер халдлагын 99.9%, фишинг халдлагын 86%, тодорхой чиглэсэн халдлагын 76%-ийг зогсоодог [3] ба энэ нь хангалттай хэмжээний үзүүлэлт биш юм. Ихэнх байгууллагууд зөвхөн MFA баталгаажуулалтад итгэн бусад сайжруулалтыг хэрэгжүүлэхгүй байх хандлагатай ба үүнээс үүдэн халдагчдад боломж олгодог.

Судалгааны гол зорилго нь вэб аппликейшнууд дахь MFA-г тойрох халдлагыг тестлэх, тэдгээрийн үр нөлөөг тодорхойлох, хамгаалалтын аргачлалыг боловсруулахад оршино.

Судалгааны гол асуултууд:

1. Вэб аппликейшн дахь MFA-г тойрох ямар төрлийн халдлагууд хамгийн өргөн тархсан бэ?
2. Эдгээр халдлагууд хэрхэн ажилладаг бөгөөд ямар нөхцөл байдалд амжилттай хэрэгжих магадлалтай вэ?
3. MFA-г тойрох халдлагуудаас сэргийлэх үр дүнтэй хамгаалалтын арга техникүүд юу вэ?

Энэхүү судалгаа нь вэб аппликейшн хөгжүүлэгчид болон аюулгүй байдлын мэргэжилтнүүдэд MFA хамгаалалтын сул талыг тодорхойлох, сайжруулах арга хэмжээг боловсруулахад туслах болно.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

II. СУДЛАГДСАН БАЙДАЛ

A. MFA-ийн үр дүнтэй байдлын үнэлгээ

Өмнөх судалгаануудын ихэнх нь MFA нь уламжлалт нууц үг дээр суурилсан баталгаажуулалтаас илүү найдвартай гэдгийг баталсан байдаг. Ometov et al. (2018) [4] судалгаандаа MFA-ийн түгээмэл хэрэглэгддэг аргууд болох SMS OTP, TOTP (Time-based One-Time Password), биометр баталгаажуулалт болон токен ашиглалт нь фишинг халдлагад бага өртөмтгий боловч хэрэглэгчийн туршлага, хэрэгжүүлэх өртөг өндөртэй болохыг тогтоосон.

B. MFA-г тойрох халдлагууд

Fassl et al. (2021) [5] судалгаандаа хуурамч вэбсайт ашиглан хэрэглэгчдийг хуурч MFA кодыг авах фишинг халдлагын үр дүнг шинжилсэн. Тэдний судалгаа нь хүнээс хамааралтай хүчин зүйлс MFA-г хэрхэн сулруулж болохыг харуулсан.

Өөр нэг судалгаанд Kiivan et al. (2020) [6] вэб аппликейшнд MitM (Man-in-the-Middle) халдлагыг туршиж, шифрлэгдээгүй холбоосууд болон хяналтгүй хандалтын токенууд MFA-г тойроход чухал үүрэг гүйцэтгэж байгааг тогтоосон.

C. SIM Swapping халдлагын судалгаа

MFA-г тойрох түгээмэл аргуудын нэг болох SIM Swapping халдлагыг Gupta et al. (2019) [7] нар гүнзгий судалж, үүрэн холбооны операторуудын баталгаажуулалтын сул талууд энэ төрлийн халдлагыг амжилттай хэрэгжүүлэх үндсэн шалтгаан болдог гэдгийг онцолсон.

1-Р ХҮСНЭГТ.СУДАЛГААНЫ АЖЛУУДЫН ЯЛГААТАЙ БАЙДАЛ

Судалгаа	***Судлагдсан MFA төрөл***	***Тойрох халдлагын арга***	***Үр дүн***
Bonneau et al. (2018)	SMS OTP, TOTP, биометр, токен	Ерөнхий хамгаалалтын үнэлгээ	Биометр, токен хамгаалалт сайн ч хэрэгжүүлэх өртөг өндөр
Krombholz et al. (2021)	TOTP, SMS OTP	Фишинг халдлага	Хэрэглэгчийн мэдлэггүй байдал MFA сулруулах хүчин зүйл
Oli et al. (2020)	SMS OTP, TOTP	MitM, Session Hijacking	Шифрлэгдээгүй холбоос, хяналтгүй токен нь халдлагыг хөнгөвчилдөг
Gupta et al. (2019)	SMS OTP	SIM Swapping	Үүрэн холбооны операторын баталгаажуулалт хангалтгүй

Эдгээр судалгаануудаас харахад MFA нь уламжлалт баталгаажуулалтаас илүү найдвартай ч тодорхой төрлийн халдлагад өртөмтгий хэвээр байгааг үзүүлж байна. Иймд MFA-г тойрох халдлагын эсрэг илүү найдвартай хамгаалалтын аргачлалуудыг боловсруулах нь энэ судалгааны гол зорилтуудын нэг юм.

III. ТУРШИЛТЫН АРГА ЗҮЙ

A. Туршилтын орчин

Судалгааны хүрээнд MFA-г тойрох халдлагуудыг туршихын тулд тестийн орчин бэлтгэж, дараах хэрэгслүүдийг ашигласан.

Үйлдлийн систем: Kali Linux
Penetration testing хэрэгслүүд:

- Burp Suite – Вэб траффик хянах, халдлагын симуляци хийх
- Evilginx2 – Reverse Proxy ашиглан фишинг хийх
- Gophish – Фишинг кампанит ажил үүсгэх
- Сүлжээний урсгалыг хянах, халдлагын үеийн мэдээллийг задлан шинжлэх

Тестийн орчны сервер нь AWS cloud сервер дээр байрласан.

B. Туршсан MFA төрлүүд

Туршилтад 2-р хүснэгт дэх олон шатлалт баталгаажуулалтын төрлүүдийг хамруулсан бөгөөд тус бүр дээр MFA-г тойрох боломжтой эсэхийг судалсан

2-Р ХҮСНЭГТ.ТУРШИЛТАД АШИГЛАСАН БАТАЛГААЖУУЛАЛТЫН ТӨРЛҮҮД

MFA төрөл	***Тайлбар***	***Тойрох халдлагын боломж***
TOTP (Time-based One-Time Password)	Google Authenticator, Authy зэрэг апп ашиглан нэг удаагийн код үүсгэх	Фишинг халдлага, Reverse Proxy (Evilginx2)
SMS-based OTP	Нэг удаагийн кодыг SMS-ээр илгээх	SIM Swapping, Man-in-the-Middle (MitM) халдлага
Push Notification (Approval-based)	Хэрэглэгчид баталгаажуулалтын хүсэлт илгээх (Microsoft Authenticator, Duo Security)	Session Hijacking, Push Notification Spamming (Fatigue Attack)
FIDO2/WebAuthn (Security Key, Biometric)	Физик токен эсвэл биометр ашиглан баталгаажуулах	Фишингээс хамгаалагдсан боловч USB токенийг хулгайлах, төхөөрөмжийн найдвартай байдлаас хамаарна

C. Туршилтын аргачлал

1. Фишинг халдлага (Reverse Proxy ашиглах)

Evilginx2 ашиглан хууль ёсны вебсайтын хуулбар үүсгэх;

Хэрэглэгчийн нууц үг болон нэг удаагийн MFA кодыг хулгайлах оролдлого хийх;

2. Session hijacking (Сесс булаах халдлага)

Burp Suite ашиглан хэрэглэгчийн нэвтэрсэн сессийг олж авах;

MFA-г тойрон шууд серверт хандах боломжтой эсэхийг шалгах;

3. SIM Swapping халдлага

Үүрэн холбооны үйлчилгээ үзүүлэгчийн сул талуудыг ашиглан хэрэглэгчийн дугаарыг өөр SIM карт руу шилжүүлэх туршилт хийх;

4. Push Notification Fatigue (Хэрэглэгчийг залхаах халдлага)

Хэрэглэгчид олон тооны баталгаажуулалтын хүсэлт илгээж, эцэст нь санамсаргүйгээр зөвшөөрүүлэх боломжтой эсэхийг тестлэх;

5. Man-in-the-Middle (MitM) халдлага

Wireshark болон Burp Suite ашиглан вэб урсгалыг хянах;

HTTP траффикаар дамжуулан нууцлалгүйгээр илгээсэн MFA кодыг барих оролдлого хийх;

D. Туршилтын хязгаарлалт

Ёс зүйн асуудал: Бодит хэрэглэгчдийн мэдээлэлд халдаагүй, зөвхөн симуляцийн орчинд туршсан.

Бодит хамгаалалттай системүүд дээрх үр дүн: Зарим компаниудын хамгаалалтын нэмэлт механизмууд (device fingerprinting, IP-based protection) тестийн үр дүнд нөлөөлж болохыг тооцсон.

FIDO2/WebAuthn-д тулгарсан хүндрэл: Фишингээс хамгаалагдсан энэхүү технологийг шууд тойрох нь хүндрэлтэй байсан ч хэрэглэгчийн төхөөрөмж болон токен найдвартай байдалд суурилсан эрсдэл байсаар байгааг тэмдэглэсэн.

Энэхүү аргачлалаар дамжуулан MFA-г тойрох халдлагуудын амжилтын түвшинг тодорхойлж, хамгаалалтын үр дүнтэй арга хэмжээг боловсруулах боломж бүрдсэн.

IV. ТУРШИЛТ БА ҮР ДҮН

A. MFA тойрох халдлагын тестийн ерөнхий үр дүн

3-Р ХҮСНЭГТ.ХАЛДЛАГУУДЫН АМЖИЛТЫН ТҮВШИНГ ХЭМЖСЭН БАЙДАЛ

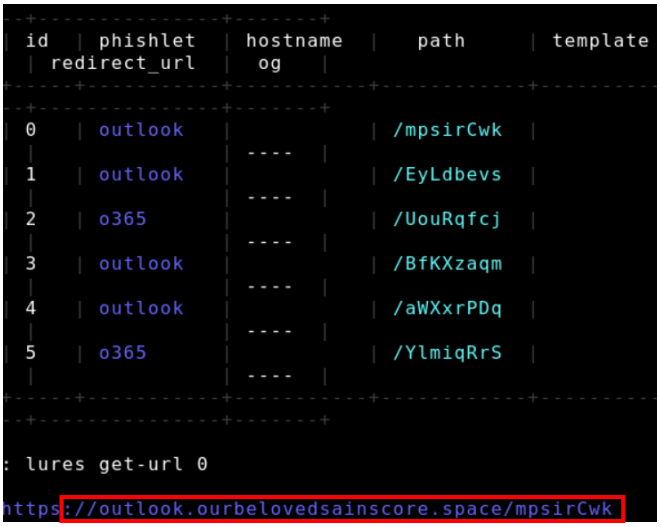
MFA төрөл	***Тойрох халдлагын төрөл***	***Амжилт-тын хувь (%)***	***Тайлбар***
SMS-based OTP	Фишинг (Reverse Proxy)	85%	Evilginx ашиглан амжилттай халдсан
SMS-based OTP	SIM Swapping	0%	Үүрэн операторын хамгаалалтуудыг шалган өөр хэрэглэгчийн SIM авах боломжийг судалсан
TOTP (Google auth)	Фишинг (MitM Proxy)	60%	Session Hijacking ашиглан сесс хулгайлсан тохиолдолд амжилт өндөр
Push Notif (Microsoft auth)	Push Spam (Fatigue Attack)	40%	Хэрэглэгч олон хүсэлтээс залхаж, зөвшөөрсөн
FIDO2 / WebAuth	Бие даасан халдлага боломжгүй	0%	Фишинг болон MitM халдлагад өртөөгүй

Гол дүгнэлт: SMS-based болон TOTP MFA нь фишинг, MitM халдлагад өндөр өртөмтгий байсан. Харин FIDO2/WebAuthn MFA болон SIM Swapping нь эдгээр халдлагад өртөөгүй.

B. Фишинг халдлагын үр дүн (Evilginx2 ашигласан тест)

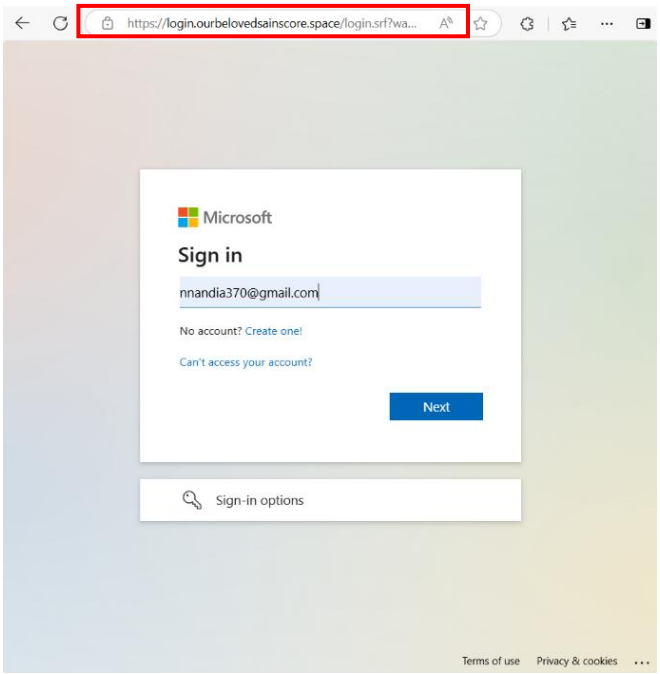
Тестийн алхмууд:

1. Evilginx2 ашиглан жинхэнэ вебсайтын хуулбар үүсгэсэн.

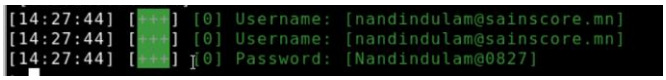


1-р зураг. Хуурамч холбоос үүсгэсэн байдал

Холбоосоор дамжин 2-р зурагт үзүүлж буй хуурамч вэб хуудас харагдах ба Gophishing хэрэгсэл ашиглан хүссэнээрээ өөрчлөх боломжтой.



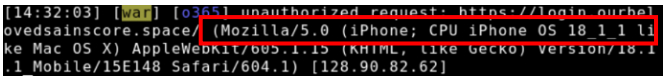
2-р зураг. Хуурамч вэб хуудас



3-р зураг. Хохирогчийн нэвтрэх нэр, нууц үг зэргийг авсан байдал

Хэрэглэгчийн нэвтрэх нэр, нууц үг зэргийг авах нь хангалтгүй ба MFA хэрэгжүүлсэн тохиолдолд туршиж байгаа учир баталгаажуулалтын кодыг авах нь чухал юм.

2. Хуурамч вэбсайтаар дамжуулж MFA кодыг хулгайлах оролдлого хийсэн.



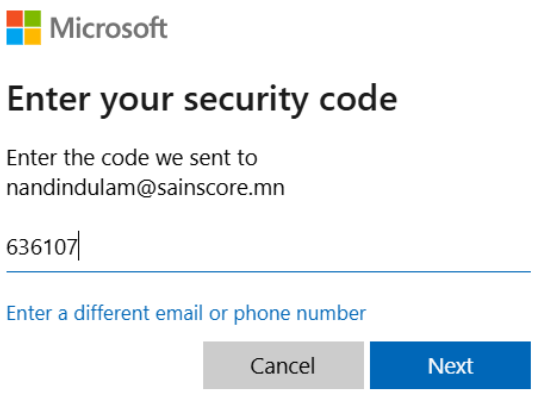
4-р зураг. MFA кодыг хулгайлах оролдлого

Туршилтаар MFA кодыг авч чадаагүй зөвхөн ямар төхөөрөмж ашиглан MFA хийж буйг олж авсан учир session хулгайлахаар шийдсэн. Өөрөөр хэлбэл, аль хэдийн MFA кодоо хийгээд нэвтэрсэн сесс-ийг авахыг оролдсон гэсэн үг.

3. Session hijacking ашиглан MFA-г тойрох боломжтой эсэхийг туршсан.

Хэрхэн сесс авсан болохыг 6-р зурагт үзүүлэв.

Тестийн үр дүн:

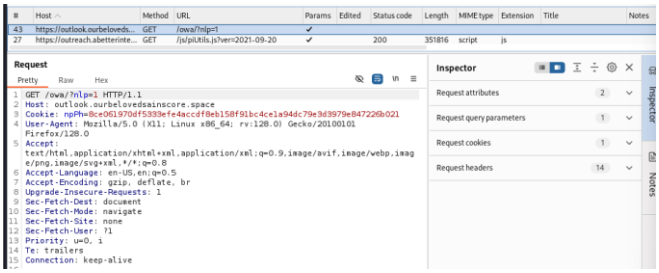


5-р зураг. Фишинг халдлагын үр дүнд сесс аван хандаж чадсан.

C. Session Hijacking тестийн үр дүн (Burp Suite ашигласан тест)

Тестийн алхмууд:

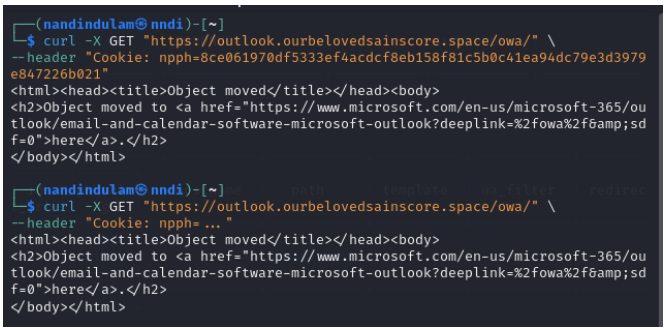
1. Burp Suite ашиглан хэрэглэгчийн сессийг хянах.



6-р зураг. Хэрэглэгчийн cookie-г авсан.

2. Хулгайлагдсан сессээр нэвтрэх оролдлого хийх.

Тестийн үр дүн:



7-р зураг. Cookie нэвтрэх оролдлого хийх.

7-р зурагт хэрэглэгчийн cookie нь идэвхтэй байгаа эсэхийг шалгаж байна. Мөн session нь олон платформ дээр ажиллах эсэхийг тестлэж үзсэн.

Дээрх сесс нь Outlook Web App (OWA) руу очих үед "Object moved" гэсэн хариу авсан байна. Энэ нь домэйн (outlook.ourbelovedsainscore.space) Microsoft-ийн албан ёсны Outlook сайт руу шилжсэн гэсэн үг юм. Өөрөөр хэлвэл сесс ашиглах хугацаа дууссан байна.

Зөвхөн туршилтын энэ тохиолдолд гарсан үр дүн хэдий ч цаашид burpsuite прокси сервер ашиглан хуурамч вэбсайт руу хандаж буй хэрэглэгчийн cookie, session-ийг хулгайлах бүрэн боломжтойг харуулж байна.

V. ХЭЛЭЛЦҮҮЛЭГ

MFA нь хэрэглэгчийн аккаунтын 99.9%-ийг халдлагаас хамгаалдаг хэмээн Microsoft компаны албан ёсны вэб хуудас дээр зарлагдсан мэдээллийг дээрх энгийн туршилтаар үгүйсгэж болох юм.

Тэгвэл MFA хангалтгүй байх шалтгаан болон тэдгээрийг сайжруулж болох шийдлүүдийг авч үзье.

A. Фишинг болон MitM халдлагад өртөх эрсдэл

Шийдэл: FIDO2/WebAuthn зэрэг фишингийн эсрэг хамгаалалттай MFA аргуудыг хэрэгжүүлэх.

B. OTP болон SMS баталгаажуулалтын сул тал

Шийдэл: Тусгай аппликейшнд суурилсан баталгаажуулалт ашиглах. Кодыг богино хугацаанд ашиглах боломжтой болгож, нэг удаа хэрэглэгддэг байдлыг хангах.

C. Биометрийн мэдээлэлтэй холбоотой аюулгүй байдлын асуудлууд

Шийдэл: Хэрэглэгчийн зан төлөвийн үнэлгээ хийн хэвийн бус үйлдлийг илрүүлэх AI суурилсан системийг нэвтрүүлэх.

D. MFA ашиглахад төвөгтэй байдал

Шийдэл: Хэрэглэгчийн байршил, төхөөрөмжийн итгэлцэл дээр үндэслэн уян хатан MFA бодлого хэрэгжүүлэх.

ДҮГНЭЛТ

Энэхүү судалгаагаар вэб аппликейшнууд дахь олон шатлалт баталгаажуулалтыг тойрох (MFA bypass) халдлагын аргуудыг судалж, бодит туршилтуудыг хийснээр хамгаалалтын сул талуудыг тодорхойллоо. Судалгааны үр дүнд фишингийн халдлага, MITM халдлага, сесс хулгайлах, fallback баталгаажуулалтын

эмзэг байдал зэрэг олон төрлийн MFA тойрох техникүүд халдлагад өртөмтгий болохыг тогтоосон.

Судалгааны үр дүнд үндэслэн фишингээс хамгаалах FIDO2/WebAuthn, нэмэлт биометр болон төхөөрөмжийн итгэлцэлд суурилсан баталгаажуулалт зэрэг хамгаалалтын арга шийдлүүдийг санал болгож байна. Эдгээр шийдлүүд нь MFA-ийн аюулгүй байдлыг сайжруулж, халдлагад өртөх магадлалыг бууруулах боломжтой.

Цаашдын судалгааны чиглэлд хиймэл оюун ухаанд суурилсан MFA хамгаалалт, динамик буюу хэрэглэгчийн зан төлөвт суурилсан баталгаажуулалтын аргачлалууд, вэб аппликейшнуудын MFA хамгаалалтыг автомат шалгах систем зэрэг сэдвүүдийг гүнзгийрүүлэн судлах болно.

Энэхүү судалгаа нь MFA-ийн хамгаалалтыг улам боловсронгуй болгоход хувь нэмэр оруулж, вэб аппликейшнуудын аюулгүй байдлыг дээшлүүлэхэд чухал ач холбогдолтой юм.

ТАЛАРХАЛ

Энэхүү судалгааны ажлыг хийж гүйцэтгэхэд туслалцаа үзүүлсэн ШУТИС-ийн Мэдээлэл Холбоо Технологийн Сургуулийн Мэдээллийн Сүлжээ, Аюулгүй Байдлын Салбарын багш нартаа талархсанаа илэрхийлье.

НОМ ЗҮЙ

[1] Multi-Factor Authentication (MFA) Statistics You Need To Know In 2025. Available from https://expertinsights.com/insights/multi-factor-authentication-statistics/

[2] 2025 SME IT Trends Report Learn how IT is steering innovation and security in 2025. Grab Your Copy Toggle Search JumpCloud, Cloud Directory, Logo 2025 Multi-Factor Authentication (MFA) Statistics & Trends to Know. Available from https://jumpcloud.com/blog/multi-factor-authentication-statistics

[3] 17 Essential Multi-factor Authentication (MFA) Statistics [2023]. Available from https://www.zippia.com/advice/mfa-statistics/

[4] Ometov, A., Bezzateev, S., Mäkitalo, N., Andreev, S., Mikkonen, T., & Koucheryavy, Y. (2018). Multi-factor authentication: A survey. *Cryptography*, *2*(1), 1.

[5] Fassl, M., Gröber, L. T., & Krombholz, K. (2021, May). Exploring user-centered security design for usable authentication ceremonies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-15).

[6] Klivan, S., Höltervennhoff, S., Huaman, N., Krause, A., Simko, L., Acar, Y., & Fahl, S. (2023, November). " We've Disabled MFA for You": An Evaluation of the Security and Usability of Multi-Factor Authentication Recovery Deployments. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security* (pp. 3138-3152).

[7] Otta, S. P., Panda, S., Gupta, M., & Hota, C. (2023). A systematic survey of multi-factor authentication for cloud infrastructure. *Future Internet*, *15*(4), 146.

[8] Sinha, A., Shrivastava, G., & Kumar, P. (2019). A pattern-based multi-factor authentication system. *Scalable Computing: Practice and Experience*, *20*(1), 101-112.

Seeing the Unseen: Visualizing and Enriching Cyber attack

Nyamjargal Tulgabaatar[1], Ganbayar Uuganbayar[*], Adiyabat Enkhjargal[3], and Telmuun Tumnee[4]

[1]School of Information Technology and Electronics, National University of Mongolia, Ulaanbaatar, Mongolia,,
Email:24m1num0250@stud.num.edu.mn {ganbayar, adiyabat.e, telmuun}@num.edu.mn

***Abstract*—Cybersecurity remains a prevailing issue due to the increasing sophistication of attacks and the growing reliance on digital systems across all sectors. This research introduces a practical approach to visualizing and enriching cyber threat data based on real-time security alerts and Indicators of Compromise (IoCs). Using open-source tools, the system automatically parses alert data from IDS systems such as Suricata and Snort, extracts key elements (e.g., IP addresses, domains, CVEs), and enriches them with external threat intelligence sources, including MITRE ATT&CK and CVE databases, known exploit approaches. The enriched data is then visualized using graph visualization as a dynamic graph that maps relationships between entities such as threats, vulnerabilities, and attack techniques. This method provides security analysts and SOC teams with a clear, time-based view of attack flows, enabling faster investigation and improved response capabilities. The approach demonstrates how automated CTI correlation and graph-based visualization can enhance situational awareness and decision-making in cybersecurity operations.**

***Index Terms*—Cyber threat intelligence (CTI), Indicator of Compromise (IoC), Attack visualization**

I. Introduction

While technological advancement continues to drive innovation across industries, it has also introduced significant cybersecurity risks. In recent years, cyber threats have escalated in both frequency and sophistication, making them one of the most pressing global concerns. The World Economic Forum identifies cyber incidents and their risks as one of the top-ranked risks in terms of impact and likelihood, placing them alongside natural disasters and geopolitical conflict in the global risk landscape [1]. In parallel, the monetary loss of cyberattacks continues to rise; the IBM-Ponemon Institute reported that the average cost of a data breach increased from USD 3.86 million in 2020 to USD 4.88 million by 2024 [2].

Despite the evolving threat landscape, the most effective defense remains the ability to detect, analyze, and respond to attacks in a timely and informed manner [2]. Among these capabilities, Cyber Threat Intelligence (CTI) plays a pivotal role by providing actionable insights into threat actors, tactics, techniques, and procedures (TTPs). Accenture's recent findings indicate that organizations that effectively operationalize CTI can reduce incident response time by up to 40% and overall breach impact by over 50% [3]. However, one of the major limitations in leveraging CTI is the fragmented nature of threat feeds, which are often unstructured, dispersed across multiple sources, and difficult to correlate in real-time. This challenge is especially problematic for small and mid-sized organizations that lack the resources to integrate expensive threat intelligence platforms into their workflows.

From an analyst's perspective, the core challenges extend beyond collection. There is a need for automated enrichment of IoC data, contextual mapping to known threat models, and efficient presentation of complex relationships. The use of graph theory or visualization in cybersecurity has shown promise for modeling relational data, including entities such as IP addresses, vulnerabilities, malware, and threat actors. Prior works have applied graph-based methods to intrusion detection and attack path mapping, but most fall short when it comes to real-time enrichment and layered visualization of evolving attacks and the integration with CTI platforms [4], [5].

In this research, we address these limitations by integrating open-source intrusion detection systems (Suricata, Snort) with an automated threat enrichment and visualization framework. Alert data is parsed to extract Indicators of Compromise (e.g., IPs, CVEs, domains), which are then enriched through correlation with external intelligence sources such as MITRE ATT&CK, CVE databases, Exploit sources, and others. Finally, the enriched data is visualized using the customized Tulip graph platform, enabling dynamic modeling of attack progression. This approach supports better situational awareness, faster incident triage, and more informed decision-making for cybersecurity analysts and SOC teams.

The remainder of this paper unfolds as follows. In the next section, we review existing approaches and literature related to our work (*Section II*). In *Section III*, we specify the problems that we will solve with the proposed solution (*Section IV*). The conclusion and future work are illustrated in *Section V*.

II. Related work

Cyber threat detection, analysis, and intelligence sharing have led to the development of a wide range of tools and platforms, both commercial and open-source. Among the most widely used open platforms are MISP (Malware Information Sharing Platform), OpenCTI, and frameworks based on STIX/TAXII standards. These tools support structured data exchange and collaborative analysis. Additionally, the MITRE ATT&CK framework provides a comprehensive model of adversary tactics and techniques that many platforms integrate with for contextual enrichment.

While these platforms offer powerful features for intelligence professionals, their integration into automated analysis

workflows often requires substantial customization. For example, extracting specific Indicators of Compromise (IoCs) from threat feeds, correlating them with live alerts, and visualizing the relationships between data points often demand additional development efforts.

A. Intrusion Detection Systems and Log Analysis

Intrusion Detection Systems (IDS) are essential components in identifying and responding to cyber incidents. They monitor network and host activity to detect suspicious or unauthorized behavior, contributing directly to the protection of data confidentiality, integrity, and availability [6]. IDS generally operate through a pipeline of stages: data collection (e.g., packet capture, system logs), event analysis using detection algorithms, alert generation upon suspicious activity, and log storage for post-event investigation [7].

IDS solutions fall into two main categories:

- Network-based IDS (NIDS): Deployed at network entry points (routers, gateways), NIDS analyzes traffic patterns for signs of intrusion [8].
- Host-based IDS (HIDS): Installed on endpoints or servers, HIDS monitors internal system behavior, such as file changes or process anomalies [9].

In terms of detection methodology:

- Signature-based detection compares traffic against a database of known attack patterns. It is fast and reliable for known threats but fails to detect unknown or zero-day attacks unless signatures are constantly updated [10], [11].
- Anomaly-based detection uses statistical or machine learning models to flag deviations from normal behavior. While better at identifying novel attacks, it typically has higher false positive rates [12].

Modern IDS solutions leverage AI and cloud technologies to enhance detection accuracy, scalability, and automation. Artificial Intelligence-based IDS utilize machine learning and deep learning techniques to analyze network traffic, enabling the detection of complex and previously unseen cyber threats with improved accuracy and reduced false positives [19]. Cloud-based IDS leverage the scalability and flexibility of cloud computing to monitor and analyze network activities in real-time, offering efficient intrusion detection across distributed environments [20].

Open-source IDS such as Suricata offer high-performance alerting with structured output formats like JSON /eve.json in particular/, making them ideal for integration with visualization and analysis pipelines [8]. Suricata alerts include fields such as IP addresses, ports, protocols, timestamps, signature IDs, and classifications, providing a rich dataset for automated processing [13].

In our study, we developed a Python-based parser that ingests these alert logs, extracts relevant IoCs, and maps each alert as a graph node. Relationships between entities — such as attacker IPs, exploited CVEs, and target systems — are represented as edges. This structured representation enables a clearer understanding of attack flow and technique attribution.

B. Graph-Based Visualization

To enhance analytical clarity, we employed Tulip, an open-source graph visualization platform, to model and explore the relationships between threat entities [16], [17]. Unlike Maltego, which is primarily tailored for open-source intelligence gathering, and Gephi, which excels in static network visualizations, Tulip offers a flexible plugin architecture and supports dynamic graph analysis, making it ideal for custom cybersecurity applications. Its capability to manage hierarchical subgraphs and integrate various visualization techniques allows for in-depth exploration of complex network structures [21].

In our system, graph nodes represent entities such as IP addresses, CVE identifiers, attack types, MITRE techniques, domains, and assets. Edges represent relationships, for example, an IP launching an exploit using a specific technique on a target system.

Tulip's plugin architecture enabled further enhancements. We implemented a popup script that retrieves MITRE ATT&CK metadata for a given technique when a user interacts with a node. This real-time enrichment supports analysts in contextualizing alerts within the broader kill chain or attack campaign.

Graph visualization, especially when combined with time-based layouts and CTI overlays, has shown strong potential in identifying multi-stage threats, detecting patterns of lateral movement, and recognizing recurring attack paths. However, existing research often focuses on static graphs or requires manual data preparation. Our work addresses this gap by automating both graph construction and enrichment based on live IDS alert data.

C. CTI Integration from Open Sources

For threat enrichment, we integrated multiple open-source CTI feeds and services. These included:

- AlienVault OTX: For attacker IPs, domains, URLs, and malware hashes
- Google exploits
- VirusTotal, AbuseIPDB, Maltrail, Anyrun: For validating IoCs against threat reputations and blacklists
- MITRE ATT&CK: For linking techniques and tactics to observed behaviors
- CVE/NVD: For vulnerability descriptions and references
- other available sources

We developed scripts to fetch, normalize, and store CTI data in formats suitable for graph augmentation (JSON, CSV). For instance, a CVE observed in an alert can be mapped to its corresponding MITRE technique, which is then visualized as a connected node in the Tulip graph. In another case, attacker IPs observed in logs are matched against reputation services to determine known associations with threat actors or campaigns.

This enrichment process not only adds analytical depth but also enables real-time fusion of internal alerts with global threat context, helping analysts answer questions like '*Has this IP appeared in other campaigns*?', '*Is this vulnerability linked*

to a known APT group?' or '*Is there any related indicator like threat-actor, CVE, exploits out there*?'

III. Problem specification

Although intrusion detection systems (IDS) and open-source cyber threat intelligence (CTI) platforms are widely available, significant challenges remain in transforming raw alert data into actionable intelligence. Analysts often lack the tools to enrich Indicators of Compromise (IoCs) in real-time or to effectively visualize the relationships between observed threats, vulnerabilities, and adversary tactics. These limitations hinder both situational awareness and rapid incident response.

The following key questions guide this research:

1) How can raw IDS alert logs be automatically parsed and enriched with contextual threat intelligence in a timely and structured way?
2) In what ways can IoCs such as IP addresses, CVEs, and domain names be dynamically linked to public threat datasets (e.g., MITRE ATT&CK, CVE, OTX) to provide investigative depth?
3) Can a graph-based model be constructed using open-source tools (e.g., Tulip) to visualize multi-phase attacks and their relationships effectively?
4) How does such a visualization improve threat interpretation, communication, and decision-making for SOC teams and security analysts?

By addressing these questions, this study proposes a system that integrates IDS log parsing, threat enrichment, and dynamic graph visualization — bridging the gap between detection and insight.

IV. Proposed Solution

This section comprises the proposed solution that falls into the following sub-sections.

A. Process flow

To address the challenges of transforming low-level IDS alerts into actionable threat intelligence, this research proposes a multi-stage pipeline /as is shown in ıthat integrates log parsing, threat enrichment, and graph-based visualization. The architecture supports both real-time monitoring and offline analysis, enabling security analysts to correlate Indicators of Compromise (IoCs) with external threat intelligence and visualize the full attack lifecycle.

*1) **Information Gathering phase**:* The system ingests alert and event data from both real-time and offline sources to support continuous monitoring as well as retrospective forensic analysis.

Real-time monitoring: Alerts are captured live from intrusion detection systems (IDS) such as Suricata or Snort, which generate structured output in JSON format. These alerts are streamed or collected continuously from active network sensors.

Offline analysis: The system also supports the ingestion of previously collected data such as:

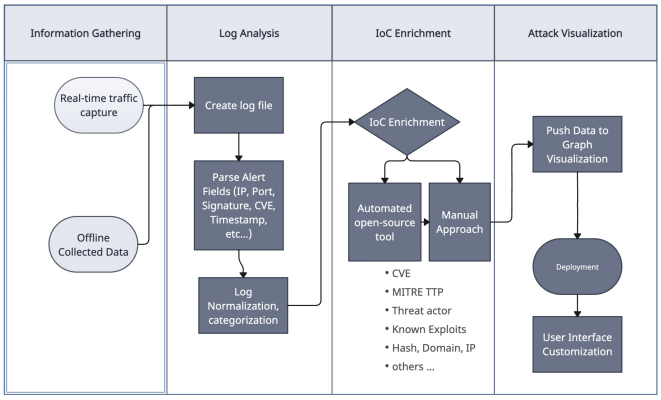


Fig. 1. Solution Flow

Packet capture files (.pcap) generated by network sniffers (e.g., tcpdump, Wireshark), which can be reprocessed through IDS engines to regenerate alerts.

Event logs are collected from network monitoring tools (e.g., Zeek), host detection agents, or security appliances.

Stored IDS alert logs such as Suricata's .eve.json, which can be directly parsed without requiring reprocessing.

This hybrid ingestion model allows the system to function in both live environments and investigative scenarios where retrospective log and traffic analysis is required.

*2) **Log Analysis phase**:* Collected alerts are processed using a custom Python parser that performs:

Log parsing: Extracting structured fields such as source/destination IPs, ports, protocols, timestamps, CVEs, and alert signatures.

Normalization and categorization: Events are transformed into a common schema, grouped by threat type, and mapped to relevant stages of the attack lifecycle. Where possible, events are matched to MITRE ATT&CK tactics and techniques to provide tactical context.

This preprocessing step ensures that diverse alerts are consistently formatted and ready for enrichment and visualization.

*3) **IoC Enrichment phase**:* The core of the system is the automated enrichment of IoCs, which adds critical context to otherwise raw indicators:

Automated enrichment: Using public CTI feeds and APIs (e.g., AlienVault OTX, AbuseIPDB, Virustotal, Anyrun, MITRE ATT&CK, CVE/NVD), the system queries threat intelligence data related to observed IPs, domains, CVEs, and file hashes.

Manual enrichment: Analysts may optionally annotate alerts with campaign-level or actor-specific information when automation is insufficient.

This stage enables dynamic threat correlation, turning atomic alert data into a meaningful representation of the adversary's tools, techniques, and procedures (TTPs).

*4) **Attack Visualization**:* To provide analysts with a clear view of how attacks unfold, enriched data is visualized using the Tulip graph platform:

Node-edge modeling: Entities such as IP addresses, CVEs, MITRE techniques, and domains are represented as nodes, with edges illustrating causal or relational links (e.g., "exploits," "targets," "originates from").

Timeline-based visualization: Time properties are used to animate the attack progression, allowing users to understand the sequence of events and lateral movement.

User interface customization: Interactive features, such as pop-up descriptions via MITRE API or filtering based on attack stages, are implemented to enhance usability.

This visual representation not only aids in threat analysis but also improves communication within security operations teams by presenting the threat in a structured and interpretable format.

B. Case studies

To evaluate the effectiveness of the proposed solution, a series of use cases were developed using both synthetic and real-world attack models. Each case illustrates the system's ability to process IDS alerts, enrich IoC data using external CTI sources, and generate an interpretable visual representation of the attack.

*1) **Case study 1: Pivot-Based Malware Attack**:* The first experiment focuses on malware analysis and the correlation of multiple IoC types—particularly file hashes, IP addresses, and domain names. This use case demonstrates how the system supports pivoting across different types of threat indicators to uncover broader relationships, such as command-and-control (C2) infrastructure and malware distribution points.

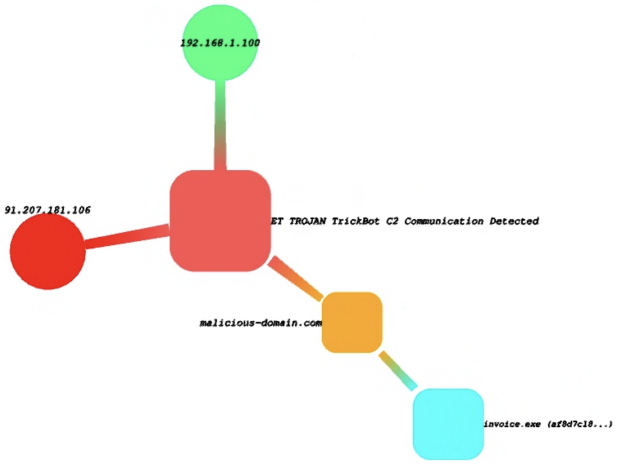


Fig. 2. Case 1 visualization

In this scenario, a malicious executable named invoice.exe was delivered to a host system (IP: 192.168.1.100). The file was identified as part of the TrickBot malware family—a well-known banking trojan with modular capabilities and built-in C2 communication. Upon execution, the infected system attempted to contact external domains and IPs associated with TrickBot's control infrastructure.

The analysis involved querying public CTI sources (e.g., VirusTotal, AlienVault OTX, AbuseIPDB) using the file hash, resulting in the identification of:

- Associated malware family: TrickBot
- Known C2 domains and IP addresses
- Threat actor tagging and attack campaign metadata

	TI	type
0	IP: 192.168.1.100 Malicious: 0 Reputation: 1	IP
1	IP: 91.207.181.106 Malicious: 10 Reputation: 0	IP
2	Domain: malicious-domain.com Malicious: 2	Domain
3	Hash: af8d7c18fe97415d7ede0cba166dce004ebbe94eacd1af6f61d8d3527b Malicious: 58 Names: Server.exe, bD0b.exe, unknown, bD0b.bin	File
4	Signature: ET TROJAN TrickBot C2 Communication Detected TI: TrickBot C2 communication detected by Suricata.	Signature

Fig. 3. CTI information

The visualization phase produced a graph model in Tulip that captured the flow and pivot points of the attack. Nodes represent the malware hash, target system, C2 servers, domains, and related IPs, while edges represent behavioral and attribution relationships. Enrichment data, such as malware classification and threat actor associations, were displayed as interactive metadata tied to specific nodes.

By linking alerts across hash, domain, and IP space, this case study highlights how malware-based investigations can be expanded into multi-dimensional threat analysis. The graph visualization enables analysts to pivot from a single artifact (e.g., a file hash) to a broader understanding of its network behavior and infrastructure linkages.

*2) **Case Study 2: Multiple Known Attacks on a Single Target**:* The second experiment models a scenario in which a single server is subjected to multiple distinct cyber-attacks. This case is designed to stress-test the entire CTI enrichment pipeline and demonstrate how the system handles the complexity of overlapping tactics, techniques, and vulnerabilities. It also serves to validate the clarity and scalability of the graph-based visualization model.

Attack Description: The targeted server (192.168.1.100) experienced three separate attack instances originating from different vectors:

- **WannaCry-like Exploit:** The first attack leveraging the SMB vulnerability (CVE-2017-0144) to propagate laterally and encrypt system files. It served as a control baseline for validating enrichment consistency.
- **BlueKeep RDP Exploit:** The second attack exploited CVE-2019-0708 (BlueKeep), targeting the Remote Desktop Protocol (RDP) to gain unauthorized access. Following exploitation, the attacker likely engaged in credential theft via Input Capture techniques, including potential keylogging to extract sensitive login information. This activity maps to MITRE techniques such as T1056.001 –

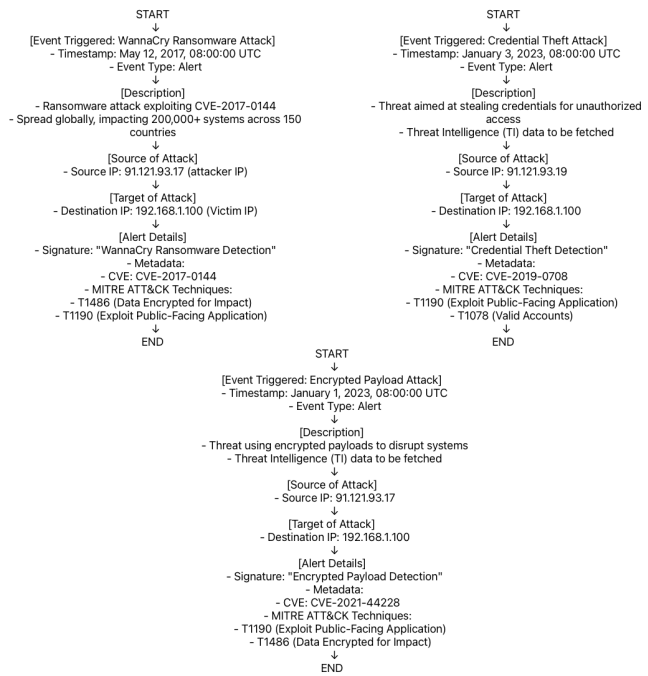


Fig. 4. Case 2 log file flowchart

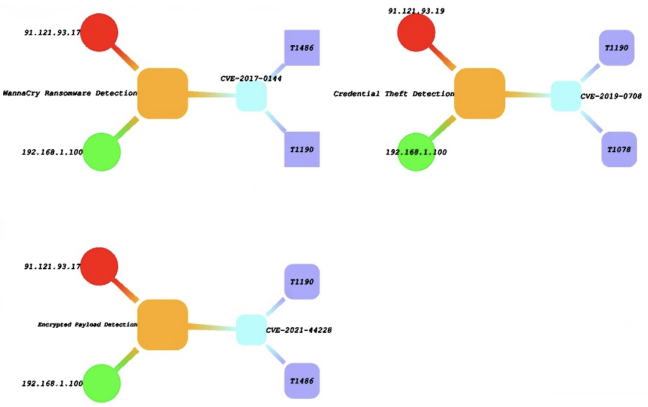


Fig. 5. Case 2 visualization

Keylogging and T1203 – Exploitation for Client Execution.

- **Log4Shell Attack:** The third attack involved the Log4Shell vulnerability (CVE-2021-44228), a critical flaw in the Apache Log4j logging library. The attacker sent specially crafted JNDI requests that triggered remote code execution (RCE), allowing arbitrary payloads to run on the target system. This incident potentially enabled further actions on objectives, such as data encryption or lateral movement.

After automated enrichment, the system generated a comprehensive threat graph using Tulip. Each attack sequence is visualized with its own trail of IoCs, CVEs, and techniques, while shared nodes (e.g., the victim IP) unify the attack context. The graph clearly displays the diversity of TTPs, severity of vulnerabilities, and intersections between campaigns.

This case demonstrates the system's ability to aggregate, enrich, and visualize multi-attack scenarios on a single asset. The layered CTI representation allows analysts to quickly assess which vulnerabilities were exploited, what attacker behaviors were observed, and how these behaviors correspond to known patterns in public threat databases.

*3) **Case study 3: Pivot-based complex attack**:* This final experiment demonstrates the system's ability to model and visualize a multi-stage, pivot-based attack that unfolds across multiple systems within a network. The case emphasizes advanced attacker behavior and the role of chained vulnerabilities, credentials, and lateral movement in progressing from reconnaissance to data exfiltration.

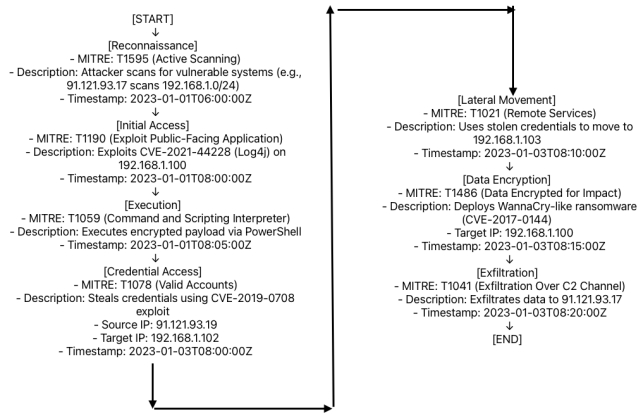


Fig. 6. Case 3 log file flow

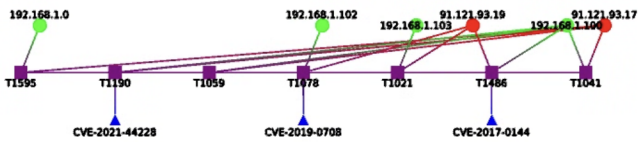


Fig. 7. Case 3 visualization

Initially, the automatically generated graph from the raw data was dense and difficult to interpret. The graph was subsequently refined by applying structured layout and filtering strategies in Tulip to emphasize the attack flow and associated MITRE ATT&CK techniques.

A total of seven distinct attack phases were observed:

- **Active Scanning** – Reconnaissance The attacker (IP: 91.121.93.17) scanned the 192.168.1.0/24 network range, searching for vulnerable systems or open ports. This phase corresponds to MITRE Technique T1595 and serves as the foundation for selecting viable targets.
- **Initial Access** – Exploit Public-Facing Application Using the Log4Shell vulnerability (CVE-2021-44228), the attacker exploited a publicly accessible Log4j service

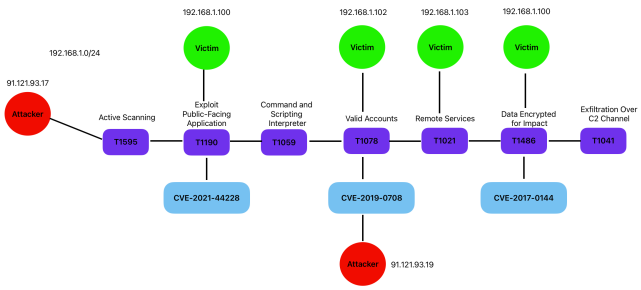


Fig. 8. Case 3 purified and rectified visualization

running on host 192.168.1.100, achieving initial remote code execution. This phase aligns with T1190 – Exploit Public-Facing Application.

- **Execution** – Command and Scripting Interpreter Post-compromise, the attacker executed a PowerShell-based payload, allowing further commands to be run remotely. This activity maps to T1059.001 – PowerShell under the Command and Control tactic.
- **Credential Access** – Valid Accounts Credentials were harvested from system 192.168.1.102 via exploitation of CVE-2019-0708 (BlueKeep). This enabled the attacker to move laterally with legitimate account access, aligning with T1078 – Valid Accounts.
- **Lateral Movement** – Remote Services Using stolen credentials, the attacker accessed additional systems (e.g., 192.168.1.103) over the network. This stage corresponds to T1021.001 – Remote Services: SMB/Windows Admin Shares.
- **Data Encryption** – Impact on Availability Ransomware was deployed to encrypt data on the compromised systems, rendering them inoperable. This maps to T1486 – Data Encrypted for Impact, a technique often used to coerce ransom payments.
- **Exfiltration** – Data Transfer to C2 Finally, stolen data—including user files and credentials—was transmitted to the attacker's external command-and-control (C2) server at 91.121.93.17. This phase corresponds to T1041 – Exfiltration Over C2 Channel.

The resulting visual graph displayed a step-by-step representation of the attack lifecycle. Each node was enriched with metadata from external CTI sources, including MITRE ATT&CK IDs, CVE references, and associated threat actor patterns. The graph clearly illustrated the flow from reconnaissance through to impact and exfiltration, enabling a holistic understanding of the threat.

This use case demonstrates the system's capacity to model complex, multi-vector intrusions, capture pivot behavior, and enhance investigation clarity through layered visualization. For security analysts, such clarity is critical for incident reconstruction, threat hunting, and both proactive and reactive defense strategy planning.

V. Conclusion and Future work

In this study, we proposed and implemented a comprehensive solution that automates the collection, enrichment, and visualization of cyber threat data based on intrusion detection alerts and CTI enrichment. By integrating open-source tools such as Suricata, CTI platforms (e.g. Virustotal, OTX, CVE, MITRE ATT&CK), and the Tulip graph visualization framework, we demonstrated how raw IoC data can be transformed into contextualized threat intelligence, visually represented as dynamic, time-aware graphs. Through three distinct case studies including known ransomware, pivot-based malware activity, multi-attack scenarios, and multi-stage pivot attacks, we validated the flexibility, clarity, and operational value of the proposed solution.

The approach enhances situational awareness by allowing analysts to quickly understand not just what was detected, but how different indicators and techniques relate to one another. It also supports better incident response by surfacing associated vulnerabilities, TTPs, and threat infrastructure in a single, navigable view. The ability to visually trace attacker behavior across different phases of the kill chain — from reconnaissance to exfiltration — provides critical support for SOC teams, incident responders, and threat hunters.

Future work will focus on automating exploit generation and simulation based on enriched IoCs, enabling proactive testing of known vulnerabilities in controlled environments. We also plan to develop AI models capable of predicting exploitability, enriching IoCs, and identifying threat relationships based on partial or observed indicators. Additionally, we aim to extend the visualization and analysis capabilities into collaborative, web-based platforms for real-time threat monitoring and response.

References

[1] World Economic Forum, "Global Risks Report 2024," Jan. 2024. [Online]. Available: https://www.weforum.org/reports/global-risks-report-2024/

[2] IBM Security, "Cost of a Data Breach Report 2024," IBM Corporation, 2024. [Online]. Available: https://www.ibm.com/reports/data-breach

[3] Accenture Security, "State of Cybersecurity Resilience 2021," Accenture, 2021. [Online]. Available: https://www.accenture.com/us-en/insights/security/state-cybersecurity-resilience-2021

[4] [D. Camacho et al., "Cyber threat intelligence using machine learning, graph analysis and threat sharing," *Computers & Security*, vol. 87, 2019.

[5] S. Noel et al., "Measuring the Cyber Threat Susceptibility of Enterprise Networks," *IEEE Security & Privacy*, vol. 10, no. 1, pp. 45–53, Jan. 2012.

[6] H. K. Idrissi and A. Kartit, "Network Intrusion Detection using Combined Deep Learning Models: Literature Survey and Future Research Directions," *IAENG International Journal of Computer Science*, vol. 51, pp. 998–1010, 2024.

[7] Scarfone, K., and Mell, P., "Guide to Intrusion Detection and Prevention Systems (IDPS)," NIST Special Publication 800-94, 2007.

[8] "Eve JSON Output," *Suricata 8.0.0-dev Documentation*, 2023. [Online]. Available: https://docs.suricata.io/en/latest/output/eve/eve-json-output.html

[9] Kemmerer, R. A., and Vigna, G., "Intrusion detection: a brief history and overview," IEEE Security Privacy, vol. 1, no. 1, pp. 27–30, Jan. 2003.

[10] Axelsson, S., "Intrusion detection systems: A survey and taxonomy," Technical Report No. 99-15, Chalmers University, 2000.

[11] Sommer, R., and Paxson, V., “Outside the closed world: On using machine learning for network intrusion detection,” in 2010 IEEE Symposium on Security and Privacy, pp. 305–316.
[12] Patcha, A., and Park, J. M., “An overview of anomaly detection techniques: Existing solutions and latest technological trends,” Computer Networks, vol. 51, no. 12, pp. 3448–3470, 2007.
[13] Suricata Documentation, “Suricata User Guide – JSON Output Format,” The OISF, 2023. [Online]. Available: https://docs.suricata.io/
[14] C. Wagner, F. Dulaunoy, A. Iklody, and G. Wagener, “MISP: The Design and Implementation of a Collaborative Threat Intelligence Sharing Platform,” in Proc. of the 2016 ACM Workshop on Information Sharing and Collaborative Security, pp. 49–56.
[15] MITRE ATTCK Framework, “Enterprise Matrix,” MITRE Corporation. [Online]. Available: https://attack.mitre.org/
[16] Tulip Team, “Tulip Graph Visualization Framework,” LaBRI / University of Bordeaux. [Online]. Available: https://tulip.labri.fr/
[17] D. Auber and P. Mary, "Tulip - A Huge Graph Visualization Framework," *SpringerLink*, 2023. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-18638-7_5
[18] Accenture, “State of Cybersecurity Resilience 2021,” Accenture Security, 2021. [Online]. Available: https://www.accenture.com/us-en/insights/security/state-cybersecurity-resilience-2021
[19] Sowmya, T. e a, Mary. (2023). A comprehensive review of AI based intrusion detection system. Measurement: Sensors. 28. 100827. [Online]. Available: https://doi.org/10.1016/j.measen.2023.100827
[20] Al-Ghuwairi, AR., Sharrab, Y., Al-Fraihat, D. et al. Intrusion detection in cloud computing based on time series anomalies utilizing machine learning. J Cloud Comp 12, 127 (2023).[Online]. Available: https://doi.org/10.1186/s13677-023-00491-x
[21] Auber, David Archambault, Daniel Bourqui, Romain Lambert, Antoine Mathiaut, Morgan Mary, Patrick Delest, Maylis Dubois, Jonathan Melançon, Guy. (2012). The Tulip 3 Framework: A Scalable Software Library for Information Visualization Applications Based on Relational Data.

Хүний биеийн 3D хөдөлгөөнийг дүрслэх мэдээллийг шахах арга

Өлзий-Ундралын Оргил
Машин оюуны лаборатори
Мэдээлэл, компьютерын ухааны тэнхим
МТЭС, МУИС
22m1num2873@stud.num.edu.mn

Ганболдын Амарсанаа
Машин оюуны лаборатори
Мэдээлэл, компьютерын ухааны тэнхим
МТЭС, МУИС
amarsanaag@num.edu.mn

***Хураангуй*—Хүний биеийн 3D хөдөлгөөнийг дүрслэх нь видео тоглоомоос бодит орчны симуляц зэрэг олон практик хэрэглээнд өргөн ашиглаж байна. Хөдөлгөөнийг өндөр чанартай, илүү нарийвчлалтай дүрслэх тусам мэдээллийн хэмжээ ихсэх боловч зарим офлайн хэрэглээнд мэдээллийн хэмжээг бага байх хэрэгцээ байдаг. Бид энэ ажлаар хүний биеийн 3D хөдөлгөөнийг чанарын алдагдалгүй шахаж, задлах шинэ аргыг танилцуулж байна. Энэ арга нь биеийн байрлал болон араг ясны 52 үеийн эргэлтийн өгөгдлийг тоймлох, төстэй мэдээллийг нэгтгэх замаар шахдаг. Үүнийг таван төрлийн 50 хөдөлгөөний хувьд туршихад мэдээллийг 1:14-ээс 1:961 хүртэл, дунджаар 1:44.9 харьцаагаар шахаж чадсан юм. Бүжгийн огцом хөдөлгөөн зэрэг нарийн төвөгтэй дүрслэлд шахалт харьцангуй сул байсан ч хөдөлгөөн бага болон зогсох хөдөлгөөн дээр маш сайн үр дүн үзүүлсэн. Энэ аргаар хөдөлгөөний мэдээллийг үр ашигтай ашиглах боломжтой, хэрэглээг өргөжүүлэхэд хувь нэмэр оруулна гэж дүгнэж байна.**

***Түлхүүр үгс*—Хүний биеийн 3D хөдөлгөөн, шахах алгоритм**

I. Удиртгал

Хүний биеийн 3D хөдөлгөөнийг дүрслэх нь видео тоглоом [1]–[9], кино [2], [4], [6], [8], [10], спорт [9], [10], анагаах ухаан [2], [10], симуляц [3], [4] зэрэг олон салбарт ихээхэн хэрэгцээтэй болж байна. Хөдөлгөөнийг бодит байдалд илүү ойр, чанартай дүрслэхэд мэдээллийн бүтцийг нарийвчлах, улмаар хэмжээг ихэсгэж байдаг. Энэ нь файлын хэмжээг ихэсгэж программын ажиллагааг удаашруулах, мэдээлэл дамжуулалт, хадгалуурыг нэмэгдүүлэх зэргээр зарим хэрэглээнд сөрөг нөлөөтэй. Тухайлбал, фитнесийн дасгал заадаг аппыг дасгалын хөдөлгөөнийг харуулсан видео бичлэг тоглуулах эсвэл 3D хөдөлгөөн тоглуулах маягаар хөгжүүлсэн байдаг. 3D хөдөлгөөн ашиглавал олон талаас нь харах, өндөр чанартайгаар томруулж жижгэрүүлж болно. Офлайн ажилладаг бол хөдөлгөөнийг апп дотроо хадгалах шаардлагатай. Өөр жишээ дурдвал бодит хугацаанд 3D хөдөлгөөний мэдээлэл дамжуулдаг AR/VR болон метаверс аппуудад ашиглаж болно. Иймээс хөдөлгөөний мэдээллийн хэмжээг чанарын алдагдалгүйгээр шахах асуудал тулгардаг.

3D хөдөлгөөний мэдээллийг шахахдаа хүний биеийн ясны бүтэц, үе мөчний байрлалыг Principal Component Analysis (PCA) [1], [3], [4], [6], [7], безиер муруй [1], [10], урвуу кинематик [1], [3], [7], [9], долгионд хувиргах [1], [3], [5], [9] болон бүлэглэх [1], [2], [8], [10] аргуудаар хувиргаж ашигласан судалгаа олон байна. Эдгээрээс хүний ясны үений эргэлтийн өнцгийн мэдээлэлд тулгуурласан аргууд цөөн, харьцангуй бага судлагдсан байна.

Энэхүү судалгааны зорилго нь хүний биеийн 3D хөдөлгөөний мэдээллийг үр дүнтэйгээр шахах, чанарын алдагдалгүй задлах аргыг боловсруулж турших юм. Бид хүний биеийн байрлал, араг ясны 52 үеийн эргэлтийн мэдээллийг тоймлох, биеийн байрлал эсвэл үеийн эргэлт нь өмнөхтэйгөө ижил тохиолдолд нэгтгэж илэрхийлэх, эдгээрийн завсрын утгуудаар илэрхийлж болохоор байвал хураангуйлах замаар мэдээллийг шахах аргыг санал болгож байна. Энэ ажлаар хүний биеийн хөдөлгөөний зогсох, явах, дасгал хийх, тулалдах, бүжиглэх зэрэг нийт таван төрлийн 50 хөдөлгөөн дээр туршиж үр дүнд нь 1:14-1:961 хүртэл харьцаатайгаар, дунджаар 1:44.9 харьцаагаар файлын хэмжээг шахаж чадсан.

Хөдөлгөөний мэдээллийг шахах аргуудын судалгаа болон хэрэглээний талаас хүний биеийн хөдөлгөөнийг ашигладаг гар утасны аппуудыг судалж өгүүллийн хоёрдугаар бүлэгт, шахах арга болон хэрэгжүүлсэн алгоритмын тухай гуравдугаар бүлэгт, дөрөвдүгээр бүлэгт хэрэгжүүлэлт болон туршилтын үр дүнг тус тус танилцуулаа.

II. Холбоотой ажлууд

Энэ бүлэгт хөдөлгөөний мэдээллийг шахах аргуудын талаар тайлбарлахаас гадна хүний биеийн 3D хөдөлгөөнийг ашигладаг хэрэглээний аппууд, тэдгээрт хэрэглэж буй хөдөлгөөний мэдээллийн хэмжээг товч танилцуулна.

A. Хөдөлгөөний мэдээлэл шахах аргууд

Хүний биеийн 3D хөдөлгөөнийг шахдаг олон аргууд байдаг. Хөдөлгөөнийг шугаман дэд орон зайн цуглуулгад дүрслэгдэх хэсгүүдэд хуваах байдлаар 1:31-1:35 харьцаатай шахдаг [1], хөдөлгөөнийг хэв загварт хуваан дугаарлаж давтагдсан хэв загваруудад дугаарыг нь ашиглах байдлаар 1:23-1:42 харьцаатай шахдаг [2], араг ясны хөдөлгөөний мэдээллийг долгионд хувиргаад дараа нь урвуу кинематик ашиглан бие болон орчны харилцан

үйлчлэлийг зохицуулах байдлаар 1:6.4-1:51 харьцаатай шахдаг [3] аргууд байна. Үүнээс гадна, хөдөлгөөний дараалалд байдаг орон зайн болон цаг хугацааны уялдаа холбоог ашиглах байдлаар 1:18-1:62 харьцаатай шахдаг [4], хөдөлгөөнийг мэдрэхэд нөлөөлж болох хамгийн чухал хүчин зүйлсийг харгалзан хөдөлгөөнийг мэдрэхүйгээр удирдах байдлаар 1:25-1:30 харьцаатай шахдаг [5], PCA суурилж шахдаг [6] арга бас бий. PCA нь олон хэмжээст мэдээллийг анхны мэдээллийн ихэнх хэсгийг хадгалж үндсэн цөөн хэмжээст болгон бууруулж өгдөг байна. Мөн кластерын арга болон үндсэн бүрэлдэхүүн хэсгүүдийн шинжилгээг ашиглан хөдөлгөөний мэдээллийг илэрхийлэх байдлаар 1:7.5-1:13 харьцаатай шахдаг [10], үндсэн геодезын шинжилгээ ашиглан араг ясны байрлалын авсаархан загварыг хөдөлгөөний дараалаар бүтээх байдлаар 1:18-1:182 харьцаатай шахдаг [7], хөдөлгөөнийг хэд хэдэн тэнцүү хэмжээтэй кластерт хуваж хөдөлгөөний мэдээллийг матриц хэлбэрээр төлөөлж кластер бүрийн хувьд альфа параллелограммын таамаглагчийг ашиглаж шахдаг [8], долгионы хувиргалт ба урагшлах кинематикийг хослуулсан ашигладаг болон урвуу кинематик дээр суурилсан таамаглалыг ашиглаж шахдаг [9] аргуудыг ч судалжээ.

Дээрх бүх аргуудын нийтлэг шинж нь хүний бие, үеийн байрлалын мэдээлэл тулгуурлаж төрөл бүрийн хувиргалт хийдэг байна. Бидний ажлын хувьд байрлалаас гадна үеийн эргэлт ашиглан хувиргалт хийж байгаагаараа өөр юм.

B. Хэрэглээ

Хүний биеийн 3D хөдөлгөөнийг ашигладаг бүх программууд дээр ашиглаж болно. Олон хөдөлгөөн ашигладаг байх тусмаа файлын хэмжээг бага байлгах хэрэгтэй. Олон хөдөлгөөн ашигладаг хамгийн түгээмэл программ гэвэл дасгал заадаг программ. Эсвэл тулааны урлаг, бүжиг заадаг гээд ямар ч байж болно. Гэхдээ энэ ажлаар дасгал заадаг программ дээр тайлбарлая. Дэлхийд маш олон дасгал заадаг программууд байгаа. Гэхдээ ихэнх программууд нь бичлэг эсвэл зураг ашиглаж заадаг. Нэг бол өөрийн дасгал хийж байгаа бичлэгээ ашиглаж эсвэл Gym Visual[1] сангийн төлбөртэй бичлэгийг ашиглаж заадаг.

Монголын бүс нутгаас Google Play Store болон App Store сангаас ойролцоогоор 100 гаруй аппуудаас хайж үзэхэд 3D хөдөлгөөн ашиглаж фитнес дасгал заадаг GymStreak[2] гэдэг нэг, тулааны урлаг заадаг Fighting Techniques Collection[3] гэдэг нэг, Krav Maga Secrets[4] гэдэг нэг болон Zhen Style Wing Chun[5] хөгжүүлэгчийн гаргасан дөрвөн апп гээд цөөн аппууд олдсон. Энэ аппууд бүгд 100MB-аас их багтаамжтай байсан. Онлайн үед бичлэг нь харагддаг дасгалын аппын хэмжээ ойролцоогоор 30MB хэмжээтэй байдаг. Эндээс харахад 3D хөдөлгөөн ашиглаж хийсэн аппуудын багтаамж нь бичлэг ашиглаж дасгал заадаг аппуудаас 70MB-аар их багтаамжтай байна. Энэ 70MB-ийг 3D хөдөлгөөний файлуудын хэмжээ гэж таамаглавал нэг хөдөлгөөнд харгалзах файлын хэмжээнээс үзэхэд энэ аппууд нь хөдөлгөөн шахаагүйгээр ашигласан гэж таамаглаж болохоор байна. Эндээс үзэхэд 3D хөдөлгөөн ашиглаад дасгал заадаг аппууд маш цөөн харин 3D шахсан хөдөлгөөн ашиглаад дасгал заадаг апп шалгасан аппууд дунд байгаагүй.

Энэ судалгааны ажилд ашигласан 50 хөдөлгөөний дундаж файлын хэмжээ нь 738.6KB байсан. Тэгвэл ийм мянган хөдөлгөөнтэй программ байвал зөвхөн хөдөлгөөний хэмжээ нь гэхэд ойролцоогоор 738.6MB болно. Олон 3D хөдөлгөөн ашиглавал файлын хэмжээ их болж байгаа учраас хөдөлгөөнүүдийг шахаж ашиглах нь энэ төрлийн аппуудад хэрэгтэй болж байна.

III. Хөдөлгөөний мэдээллийг шахах

Хүний биеийн 3D загварын риг буюу араг ясны бүтэц нь хүний араг ясны бүтцийн хураангуйлсан хувилбар гэж ойлгож болно. Цаашдаа нэг ясыг үе гэж нэрлээд явъя. Загварын риг нь мод бүтэцтэй байдаг. Зураг 1-ийн хувьд модны үндэс нь ташаа юм. Ташаанаасаа нуруу болон хоёр хөл рүү салаалж цаашаагаа үргэлжилж бүх үеүд руу холбогдоно. Энэ зурагт энгийн хүний ригийг харуулсан. Хүний загварыг хийхдээ хэдэн ч ястай хийж болно. Гэхдээ ихэнх тохиолдолд энэ энгийн ригтэй адилхан 52 үетэйгээр хийдэг.

A. Хөдөлгөөний мэдээлэл

3D хөдөлгөөн нь үргэлжлэх хугацаа, нийт фрэймийн тоо, үеүдийн нэр, үеүдийн бүтэц болон фрэймүүдийн дараалал гэх мэт зүйлсээс бүрдэнэ. Нийт фрэймийн тоогоор үргэлжлэх хугацааг илэрхийлж болох учраас үргэлжлэх хугацааг орхиж болно. Мөн үеүдийг тогтмол дараалаар хадгалбал үеүдийн нэр болон бүтцийг нь орхиж болно. Фрэйм нь тухайн агшинд өөрчлөгдсөн мэдээллүүдийг хадгалдаг. Харин бидний хувьд хүний хөдөлгөөний фрэйм нь фрэймийн дугаар, үеүдийн байрлал болон эргэлтийн мэдээллийг хэлнэ. Хүн биеэ хөдөлгөхдөө үеүдээ эргүүлж хөдөлгөдөг учраас зөвхөн үеүдийн эргэлтийг ашиглаж биеийг нь хөдөлгөж болно. Тийм учраас зөвхөн биеийн байрлал буюу ташааны байрлалыг хадгалаад бусад үеүдийн байрлалыг орхиж болно. Ингэвэл нэг фрэймд фрэймийн дугаар, ташааны байрлал болон өөрчлөгдсөн үеүдийн эргэлтийн мэдээллийг л хадгална гэсэн үг. Нийтэд нь үзвэл хүний хөдөлгөөний файлд нийт фрэймийн тоо болон фрэймүүдийн дараалал багтана гэж ойлгож болно.

Бидний зорилго бол хөдөлгөөний мэдээллийг аль болох чанарын алдагдалгүйгээр шахах юм. $[0, 40000[$ завсар дахь бүхэл тоог хоёр байтад хадгалж чадна мөн 3D координатыг гурван тоо учраас зургаан байтад хадгалж чадна гэж үзье. Байрлал болон эргэлтүүд нь 3D координатууд.

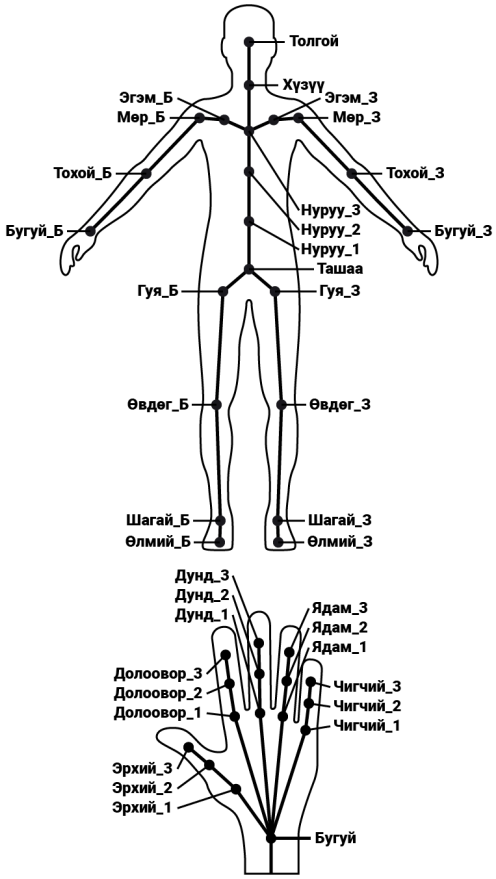
[1] https://gymvisual.com/

[2] https://www.gymstreak.com/

[3] https://play.google.com/store/apps/details?id=com.kaysoft63.ftc

[4] https://play.google.com/store/apps/details?id=com.DefaultCompany.Kravmaga

[5] https://play.google.com/store/apps/dev?id=5688888613330247431



Зураг 1: Хүний биеийн 3D загварын араг ясны бүтэц. Хар цэгүүд нь үеүдийг илтгэнэ.

Нийт фрэймийн тоог хоёр байтад хадгалъя. Фрэйм тус бүрийн хувьд фрэймийн дугаарыг хоёр байтад, ташааны байрлалыг зургаан байтад мөн үеүдийн эргэлтийг тус бүр зургаан байтад хадгалъя.

B. Алгоритм

Нэг байтад 0-255 хүртэл тоог хадгалж чадна. Энийг нэг ASCII тэмдэгт гэж үзэж болно. 1-200 хүртэлх ASCII тэмдэгтүүдийг 200 суурьтай тооны цифрүүд гэж үзье. $[0, 40000[$ завсар дахь n бүхэл тоог 200 суурьтай хоёр оронтой тоо буюу тэмдэгт мөр рүү хөрвүүлэхдээ тоогоо 200-д хуваагаад бүхэл хэсэг дээр нь нэгийг нэмээд тэмдэгт болгоод дараа нь бутархай хэсэг дээр нь нэгийг нэмээд тэмдэгт болгоод хооронд нь нийлүүлнэ. (1) томьёог $NumEncode$ гэж нэрлэе.

$$char(\lfloor n/200 \rfloor + 1) + char((n \bmod 200) + 1) \tag{1}$$

3D координатыг 200 суурьтай зургаан оронтой тоо руу хөрвүүлэхийн тулд x, y, z тус бүрийг 200 суурьтай хоёр оронтой тоо руу хөрвүүлээд нийлүүлнэ. Байрлал эсвэл эргэлт гэдгээсээ хамаараад өөр өөрөөр хөрвүүлнэ. Байрлалыг 200 суурьтай тоо руу хөрвүүлэх томьёог зөвхөн x дээр тайлбарлая. x нь $[-20, 20[$ завсарт байгаа гэж үзье. Эхлээд x-ийг мянгад үржээд дараа нь тоймлоод 20000-ийг нэмэхэд $[0, 40000[$ завсар дахь тоо руу хөрвүүлэгдэнэ дараа нь $NumEncode$ томьёогоор хөрвүүлнэ. Ингэж хөрвүүлбэл x-ийг таслалаас хойш гурван орноор тоймлоно гэсэн үг. Одоо x, y, z тус бүр дээр энэ томьёогоо хэрэглээд нийлүүлнэ. Байрлал дээр энэнээс том завсар ашиглахыг хүсвэл зургаан байт биш есөн байтаар илэрхийлж болно. (2) томьёог $PosEncode$ гэж нэрлэе.

$$\begin{aligned} &NumEncode(\lfloor x * 1000 \rceil + 20000) + \\ &NumEncode(\lfloor y * 1000 \rceil + 20000) + \\ &NumEncode(\lfloor z * 1000 \rceil + 20000) \end{aligned} \tag{2}$$

Эргэлтийн томьёо нь байрлалын томьёотой төстэй учраас товч тайлбарлая. x нь $[0, 360[$ завсарт байгаа гэж үзье. Эхлээд x-ийг зууд үржээд дараа нь тоймлоод 36000-д хуваагаад үлдэгдлийг аваад дараа нь $NumEncode$ томьёогоор хөрвүүлнэ. Ингэж хөрвүүлбэл x-ийг таслалаас хойш хоёр орноор тоймлоно гэсэн үг. Байрлалтай адил x, y, z тус бүр дээр томьёогоо хэрэглээд нийлүүлнэ. (3) томьёог $RotEncode$ гэж нэрлэе.

$$\begin{aligned} &NumEncode(\lfloor x * 100 \rceil \bmod 36000) + \\ &NumEncode(\lfloor y * 100 \rceil \bmod 36000) + \\ &NumEncode(\lfloor z * 100 \rceil \bmod 36000) \end{aligned} \tag{3}$$

Тоог хоёр байт руу харин байрлал болон эргэлтийг зургаан байт руу хөрвүүлж чаддаг боллоо. Одоо эсрэгээр нь 200 суурьтай хоёр байтын s тоог $[0, 40000[$ завсар дахь тоо руу мөн зургаан байтын тоог байрлал болон эргэлт рүү хөрвүүлэх томьёог үзье. 200 суурьтай хоёр байтын тооны эхний тэмдэгтээ бүхэл тоо болгоод нэгийг хасаад 200-д үржээд дараа нь хоёр дахь тэмдэгтээ бүхэл тоо болгоод нэгийг хасаад эхний тоон дээрээ нэмнэ. (4) томьёог $NumDecode$ гэж нэрлэе.

$$(s_0 - 1) * 200 + (s_1 - 1) \tag{4}$$

200 суурьтай зургаан байтын s тоогоо байрлал болон эргэлт рүү хөрвүүлэхийн тулд хоёр байтаар нь x, y, z гэж гурав хуваана. Байрлал руу хөрвүүлэх томьёог x дээр эхэлж тайлбарлая. x-ээ $NumDecode$ томьёогоор хөрвүүлээд 20000-ийг хасаад мянгад хуваана. x, y, z тус бүр дээр олоод 3D координат үүсгэнэ. (5) томьёог $PosDecode$ гэж нэрлэе.

$$\begin{aligned} &(x : (NumDecode(s_0 + s_1) - 20000)/1000, \\ &y : (NumDecode(s_2 + s_3) - 20000)/1000, \\ &z : (NumDecode(s_4 + s_5) - 20000)/1000) \end{aligned} \tag{5}$$

Эргэлт рүү хөрвүүлэх томьёог x дээр эхэлж тайлбарлая. x-ээ $NumDecode$ томьёогоор хөрвүүлээд зууд хуваана. x, y, z тус бүр дээр олоод 3D координат үүсгэнэ. (6) томьёог $RotDecode$ гэж нэрлэе.

$$\begin{aligned} &(x : NumDecode(s_0 + s_1)/100, \\ &y : NumDecode(s_2 + s_3)/100, \\ &z : NumDecode(s_4 + s_5)/100) \end{aligned} \tag{6}$$

Тэмдэгт мөрийг шахахдаа нэг тэмдэгт хоёр болон түүнээс дээш удаа дараалдаж орсон бол дараалалдсан тэмдэгтүүдийн эхнийхээс бусад тэмдэгтүүдийг $char(200+$дараалдаж орсон тоо$)$ тэмдэгтээр солино. Энэ функцийг $EncodeText$ гэж нэрлэе.

Шахсан тэмдэгт мөрийг задлахын тулд тэмдэгтийн тоон утга нь 202-оос их буюу тэнцүү байвал тухайн тэмдэгтийн өмнөх тэмдэгтийг (тухайн тэмдэгтийн тоон утга-201) удаа дараалуулсан тэмдэгт мөрөөр солино. Энэ функцийг $DecodeText$ гэж нэрлэе.

Хөдөлгөөнийг нэгдүгээр хувилбараар шахах нь фрэймийн тоогоо хөрвүүлээд дараа нь фрэйм тус бүрийн хувьд фрэймийн дугаар, биеийн байрлал болон үеүдийн эргэлтүүдийг бүгдийг нь хөрвүүлээд нийлүүлнэ. Дараа нь нийлүүлэн мэдээллээ $EncodeText$ функцээр шахна.

Алгоритм 1 Хөдөлгөөнийг нэгдүгээр хувилбараар шахах

```
procedure Encode1(хөдөлгөөн)
    s = NumEncode(хөдөлгөөн.фрэймийнТоо)
    for each фрэйм ∈ хөдөлгөөн.фрэймүүд do
        s += NumEncode(фрэйм.дугаар)
        s += PosEncode(фрэйм.байрлал)
        for each эргэлт ∈ фрэйм.эргэлтүүд do
            s += RotEncode(эргэлт)
        end for
    end for
    return EncodeText(s)
end procedure
```

Нэгдүгээр хувилбараар шахсан хөдөлгөөнийг задлахын тулд хийсэн үйлдлээ урвуугаар нь хийнэ. $DecodeText$ функцээр шахсан тэмдэгт мөрөө хөрвүүлнэ. Эхний хоёр байтаа хөрвүүлээд фрэймийн тоогоо авна дараа нь фрэйм тус бүрээр фрэймүүдийн мэдээллийг уншиж авна. Нэг фрэймийн хувьд эхний хоёр байт нь фрэймийн дугаар дараагийн зургаан байт нь биеийн байрлал дараагийн $52 * 6$ байт нь 52 үеүдийн эргэлтүүдийн мэдээллүүдийг хөрвүүлж авна. Алгоритм дээр ашиглагдах $Substr(s, startIndex, length)$ функц нь s тэмдэгт мөрийн $startIndex$-ээс эхэлсэн $length$ урттай дэд тэмдэгт мөрийг авна.

Хөдөлгөөнийг хоёрдугаар хувилбараар шахах нь нэгдүгээр хувилбарын сайжруулсан хувилбар юм. $char(201)$ тэмдэгтийг цаашдаа c гэж мөн зургаа дараалсан c тэмдэгттэй тэмдэгт мөрийг $c6$ гэж нэрлэе. Гол ялгаа нь Алгоритм 1-ийн тавдугаар мөрний байрлал, долоодугаар мөрний эргэлтийн хувьд өмнөх фрэймийн харгалзах утгатайгаа тэнцүү байвал c тэмдэгтээр солино.

Хоёрдугаар хувилбараар шахсан хөдөлгөөнийг задлах нь Алгоритм 2-той адилхан. Гол ялгаа нь хоёрдугаар мөрний s тэмдэгт мөрний c тэмдэгт бүрийг $c6$-аар солино. Долоодугаар мөрний байрлал, аравдугаар мөрний эргэлтийн хувьд утга нь $c6$ тэнцүү бол өмнөх фрэймийн харгалзах утгыг өгнө.

Хөдөлгөөнийг гуравдугаар хувилбараар шахах нь нэгдүгээр хувилбарын сайжруулсан хувилбар юм. Гол

Алгоритм 2 Нэгдүгээр хувилбараар шахсан хөдөлгөөнийг задлах

```
procedure Decode1(s)
    s = DecodeText(s)
    фрэймийнТоо = NumDecode(Substr(s,0,2))
    фрэймүүд = хоосон жагсаалт
    for each фрэйм ∈ s шахсан хөдөлгөөний хоёроос хойших байтуудыг фрэймүүдэд хуваасан хүснэгт do
        дугаар = NumDecode(Substr(фрэйм,0,2))
        байрлал = PosDecode(Substr(фрэйм,2,6))
        эргэлтүүд = хоосон жагсаалт
        for each эргэлт ∈ фрэймийн наймаас хойших байтуудыг эргэлтүүдэд хуваасан хүснэгт do
            эргэлт = RotDecode(эргэлт)
            эргэлтүүдэд эргэлтийг нэмнэ
        end for
        фрэйм = (дугаар, байрлал, эргэлтүүд)
        фрэймүүдэд фрэймийг нэмнэ
    end for
    хөдөлгөөн = (фрэймийнТоо, фрэймүүд)
    return хөдөлгөөн
end procedure
```

ялгаа нь Алгоритм 1-ийн тавдугаар мөрний байрлал, долоодугаар мөрний эргэлтийн хувьд тухайн завсрын захын хоёр фрэймүүдийн утгаар дундах фрэймүүдийг илэрхийлж чадахаар бол c тэмдэгтээр солино.

Завсар гэдэг нь биеийн байрлалууд эсвэл аль нэг үеүдийн эргэлтүүдийн фрэймүүдийн дурын хоёр завсрыг илэрхийлж байгаа. Гэхдээ бидний хувьд завсрын тоог хамгийн цөөн байхаар сонгоно. Байрлалыг 3D координат биш тоо гэж үзээд жишээгээр тайлбарлая. Нийт фрэймийн тоо нь найм харин байрлалын фрэйм бүрийн утгуудын хүснэгт нь $[10, 20, 30, 40, 60, 80, 100, 120]$ байг. Энэ хүснэгтийг хамгийн багадаа $[10, 20, 30, 40]$ болон $[40, 60, 80, 100, 120]$ хоёр завсраар илэрхийлж болохоор байна. $[10, 20, 30, 40]$ завсрын хувьд захын утгууд нь $10, 40$ харин дундах $20, 30$ утгуудыг арифметик прогресс ашиглаж үүсгэж болохоор байна. Бүтэн хүснэгтийг дундах утгуудыг c-ээр орлуулж бичвэл $[10, c, c, 40, c, c, c, 120]$ болно. Байрлал болон эргэлтийн хувьд саяны жишээтэй адилхан шугаман хамааралтай завсар байвал тухайн завсрын дундах утгыг c тэмдэгтээр солино. Эргэлтийн хувьд x, y, z нь өнцөг учраас 0-360°-ийн хувьд шугаман хамаарал нь тойрдог гэдгийг бодолцох хэрэгтэй.

Гуравдугаар хувилбараар шахсан хөдөлгөөнийг задлах нь Алгоритм 2-той адилхан. Гол ялгаа нь хоёрдугаар мөрний s тэмдэгт мөрний c тэмдэгт бүрийг $c6$-аар солино. Долоодугаар мөрний байрлал, аравдугаар мөрний эргэлтийн хувьд утга нь $c6$ тэнцүү бол харгалзах завсрын захын хоёр утгаараа илэрхийлнэ.

IV. Хэрэгжүүлэлт ба туршилт

Unity Game Engine дээр C# хэл дээр хэрэгжүүлэлт хийсэн. Туршилт хийсэн дүрээ болон энэ ажилд

Хөдөлгөөн	Утга	ФТ	ХХ	ФХ	Хэмжээ			Харьцаа			Хувь		
					Х1	Х2	Х3	Х1	Х2	Х3	Х1	Х2	Х3
Бүх	ХБ	16	0.5	297.2	5.2	2.4	**0.8**	1:57.0	1:122.1	**1:961.2**	1.8	0.8	**0.1**
	ДУ	181.5	6.0	739.3	55.3	44.4	**22.6**	1:17.4	1:23.7	**1:44.9**	5.8	4.2	**2.2**
	ХИ	843	28.1	2554.1	255.4	252.0	**161.1**	1:8.1	1:10.1	**1:14.3**	12.4	9.9	**7.0**
Зогсох	ХБ	250	8.3	768.7	77.0	29.0	**0.8**	1:12.1	1:26.5	**1:961.2**	8.3	3.8	**0.1**
	ДУ	308.5	10.3	1022.2	94.0	72.3	**9.8**	1:10.8	1:14.5	**1:109.4**	9.2	6.9	**0.9**
	ХИ	427	14.2	1404.2	132.7	119.8	**23.5**	1:8.1	1:11.7	**1:51.2**	12.4	8.5	**2.0**
Явах	ХБ	16	0.5	297.2	5.2	2.4	**2.4**	1:57.0	1:122.1	**1:122.9**	1.8	0.8	**0.8**
	ДУ	32.4	1.1	365.9	10.1	7.4	**5.8**	1:37.8	1:51.9	**1:65.3**	2.6	1.9	**1.5**
	ХИ	70	2.3	511.8	20.5	14.4	**10.0**	1:25.0	1:31.6	**1:45.5**	4.0	3.2	**2.2**
Дасгал хийх	ХБ	32	1.1	342.0	10.1	4.7	**3.4**	1:36.9	1:73.2	**1:101.9**	2.7	1.4	**1.0**
	ДУ	59.4	2.0	448.1	18.5	13.0	**8.4**	1:24.7	1:35.8	**1:55.9**	4.0	2.8	**1.8**
	ХИ	94	3.1	602.0	29.1	23.5	**19.6**	1:19.2	1:25.6	**1:29.4**	5.2	3.9	**3.4**
Тулалдах	ХБ	40	1.3	360.5	11.1	5.6	**4.5**	1:34.8	1:64.1	**1:80.7**	2.9	1.6	**1.2**
	ДУ	54.6	1.8	414.3	16.9	8.8	**7.3**	1:24.7	1:47.8	**1:58.0**	4.1	2.1	**1.7**
	ХИ	75	2.5	531.9	23.3	17.2	**13.7**	1:21.1	1:31.0	**1:38.7**	4.7	3.2	**2.6**
Бүжиглэх	ХБ	134	4.5	536.5	25.9	17.8	**16.9**	1:20.7	1:30.2	**1:31.7**	4.8	3.3	**3.2**
	ДУ	452.6	15.1	1446.0	137.0	120.4	**81.6**	1:11.4	1:13.4	**1:19.3**	8.8	7.4	**5.2**
	ХИ	843	28.1	2554.1	255.4	252.0	**161.1**	1:9.3	1:10.1	**1:14.3**	10.8	9.9	**7.0**
Нийт		9075	302.5	36965.2	2764.3	2220.6	**1128.4**	1:13.4	1:16.6	**1:32.8**	7.5	6.0	**3.1**

Хүснэгт I: Шахсан хөдөлгөөний үр дүнгийн дундаж, хамгийн бага, хамгийн их утгуудын хүснэгт. ХБ - Хамгийн бага, ДУ - Дундаж утга, ХИ - Хамгийн их, ФТ - Фрэймийн тоо, ХХ - Хөдөлгөөний хугацаа Секунд, ФХ - Файлын хэмжээ KB, Х1 - Нэгдүгээр хувилбар, Х2 - Хоёрдугаар хувилбар, Х3 - Гуравдугаар хувилбар, Хэмжээ - Шахсан файлын хэмжээ KB, Харьцаа - Нэг файлын хэмжээнд хэдэн шахсан файл багтах харьцаа, Хувь - Шахсан файлын хэмжээг файлын хэмжээнд харьцуулсан хувь.

ашигласан 50 хөдөлгөөнийг Mixamo[6] сангаас төлбөргүй татаж авсан. Mixamo сангаас хөдөлгөөнийг татахдаа дүрдээ Characters: Y Bot дүрийг сонгоод Download Settings тохируулахдаа Format: FBX Binary(.fbx), Skin: Without Skin, Frames per Second: 30, Keyframe Reduction: none байхаар тохируулж хөдөлгөөнийг татсан. Without Skin-ийг сонговол FBX файл дотроо дүрийн 3D загвараа оруулахгүйгээр зөвхөн хөдөлгөөний файлыг татна. Keyframe Reduction дээр none-ийг сонговол хөдөлгөөнийг шахахгүйгээр татна. Мөн Unity программаар Y Bot дүрийг ашиглаж хөдөлгөөнийг тоглуулж туршсан. 50 хөдөлгөөндөө зогсох, явах, дасгалын, тулаан, бүжгийн тус бүр арван хөдөлгөөнийг ашигласан.

Гурван хувилбарын алгоритмыг хэрэгжүүлэхдээ хөдөлгөөний фрэймүүдийн дугаарууд нь тэгээс эхлээд бүгд дараалдаж байвал дараалал нь тодорхой учраас бүх фрэймүүдийн дугаарыг орхиж хөдөлгөөнөө шахаад дараа нь хамгийн урд нь c тэмдэгт байрлуулсан. Шахсан хөдөлгөөнөө задлахдаа урд нь c тэмдэгт байвал фрэйм бүрээ уншихдаа эхний хоёр байтыг байхгүйгээр уншиж авдаг болгож өөрчилсөн. Хэрвээ бүх фрэймүүдийн дугаар нь дараалдаж орсон байвал файлын хэмжээ (Нийт фрэймийн тоо$*2-1$) байтаар буурна.

[6]https://www.mixamo.com/

А. Хэмжих арга

Шахсан хөдөлгөөн маань үндсэн хөдөлгөөнөөс хэр зөрүүтэй вэ гэдгийг хэмжихдээ өнцгийн зөрүүг ашигласан. Үеүдийн эргэлтүүдийг зургаан байтын 200 суурьтай тоо руу хөрвүүлэхдээ x, y, z тус бүрийг таслалаас хойш хоёр орны нарийвчлалтайгаар тоймлосон учраас x-ийн алдаа нь хамгийн ихдээ 0.005° байна харин x, y, z алдааг нийлүүлбэл 0.007° гарна. Хүний загвар маань мод бүтэцтэй байдаг ба энгийн ригтэй загварын хувьд гүн нь 11 байна. 11 үен дээр үе тус бүр дээр хамгийн их зөрүү болох 0.007°-ийн алдаа үүссэн гэж тооцвол алдаа нь хамгийн ихдээ 0.077° гарна. 0.077°-ийг 180°-тай харьцуулбал хамгийн ихдээ 0.0428% алдаатай гарна. Үеүдийн эргэлтийн алдааг хамгийн их алдаа болох 0.0428% тооцъё. Биеийн байрлалын алдаа нь таслалаас хойш гурван оронгоор тоймлож байгаа учраас x утгын зөрүү нь 0.0005 харин x, y, z нийлбэр зөрүү нь 0.0009 буюу 0.09мм гарна. Гэхдээ эргэлтийн 180° шиг харьцуулах тогтмол утга байхгүй учраас орлуулаад нэг метртэй харьцуулахад 0.09% гарна. 52 эргэлт нэг байрлал байгаа бөгөөд байрлалын нэг алдааны утга нь маш бага учраас нийт алдаагаа ойролцоогоор эргэлтийн алдаагаа тоймлоод 0.043% гэзд тооцъё. Харин бусад утгууд болох фрэймийн дугаар болон нийт фрэймийн тоог хөрвүүлэхэд манай тохиолдолд алдаа үүсэхгүй яагаад гэвэл фрэймийн тоо нь 40000-аас бага бүхэл тоо байна.

Нэг фрэймийн хувьд шахсан хөдөлгөөн маань үндсэн

хөдөлгөөнтэй адилхан байх магадлал нь 99.957% гарна. Нэгдүгээр хувилбар болон хоёрдугаар хувилбарууд үндсэн хөдөлгөөний фрэймүүдийг огт хураагаагүй учраас магадлал нь яг адилхан гэсэн үг. Харин гуравдугаар хувилбарын хувьд захын утгуудаас дундах утгуудыг гаргаж байгаа ч гэсэн дундах утгаа шалгахдаа ихдээ байрлалыг 0.0005 харин эргэлтийг 0.005°-ийн зөрүүтэй байвал $char(201)$ тэмдэгтээр сольж байгаа. Энэ нь байрлал болон эргэлтийг зургаан байтын 200 суурьтай тоо болгохтой адилхан тоймлож байгаа учраас нэг фрэймийн алдааны зөрүүтэй бүх фрэймийн алдааны зөрүү нь тэнцүү гэсэн үг.

Алдааны хэмжүүрээ байрлал дээр тулгуурлан хийвэл манай ажил үений эргэлт дээр тулгуурласан учраас эхлээд өнцгийг ойролцоогоор байрлал руу хөрвүүлэх хэрэгтэй. Хүний биеийн ясны уртыг дундаж эрэгтэйнхээр тооцвол гуя 48см, шилбэ 39см, дээд гар 34см, шуу 28см байна. Шуу болон дээд гарын нийлбэрийг гар харин гуя болон шилбэний нийлбэрийг хөл гэе. Хүний загварт ашиглаж байгаа ригний дөрвөн мөчнөөс бусад ясны бүтэц жинхэнэ ястай адил биш учраас ойролцоогоор бүтэн нурууг гартай, (хос ташаа, хос эгэм, хос мөрийг) хөлтэй, (хос ул, хүзүү, толгойг) хөлтэй, сарвууг хөлтэй адил урттай гэж тооцъё. 51 ясны дундаж хэмжээг тооцвол $((48+39)*6+(34+28)*3)/51$ ойролцоогоор 13.9см байна. Адил талт гурвалжны адил талыг 13.9см гэж үзээд хоорондох өнцгийг хамгийн их алдаа болох 0.007° гэж үзвэл гурвалжны суурь буюу үений хамгийн их алдаа нь ойролцоогоор 0.17мм гарна. Биеийн байрлал буюу 0.09мм нэмээд дунджийг тооцвол үе нь 52 учраас 0.168мм гарна. Энийг нэг метртэй харьцуулбал хамгийн их алдаа нь 0.168% харин адил байх магадлал нь 99.832% байна.

B. Туршилтын үр дүн

Хүснэгт I-ээс шахсан хөдөлгөөний үр дүнгийн дундаж, хамгийн бага, хамгийн их утгуудыг харж болно. Хүснэгт II-оос шахсан 50 хөдөлгөөн тус бүрийн дэлгэрэнгүй мэдээллийг харж болно. Зогсох болон бүжгийн хөдөлгөөнүүд нь файлын хэмжээ нь том бусад нь дундаж хөдөлгөөнүүд. Зогсох хөдөлгөөний хувьд шахсан хувь нь хамгийн бага харин бүжгийн хөдөлгөөний хувьд хамгийн их хувьтай байгаа. Хоёулаа файлын хэмжээ нь том ч гэсэн яагаад хувь нь нэг нь хамгийн бага нөгөө нь хамгийн их байгаа вэ гэвэл зогсох хөдөлгөөний хувьд фрэймүүдийн эргэлтийн шилжилт харьцангуй бага харин бүжгийн хөдөлгөөний хувьд фрэймүүдийн эргэлтийн шилжилт нь огцом буюу завсрын захын утгаар дундах утгыг илэрхийлэхэд хэцүү гэсэн үг. Харин бусад явах, дасгал, тулааны шахсан хөдөлгөөнүүдийн дундаж утгууд нь ойролцоо байгааг харж болно. Нэгдүгээр хувилбарыг хоёр болон гуравдугаар хувилбарууд сайжруулсан. Гуравдугаар хувилбар нь хоёрдугаар хувилбараас илүү байгааг хүснэгтээс харж болно.

V. Дүгнэлт

Энэхүү судалгааны ажлын хүрээнд хүний биеийн 3D хөдөлгөөний мэдээллийн хэмжээг чанарын алдагдалгүйгээр шахах алгоритмыг гурван хувилбараар боловсруулаа. Судалгааны үр дүнд боловсруулсан хувилбаруудаас хамгийн сайн нь хөдөлгөөний мэдээллийг 0.1%-7% харьцаагаар 1:14-1:961, дунджаар 2.2% харьцаагаар 1:44.9 хүртэл шахаж чадсан. Эдгээрээс нэг болон хоёрдугаар хувилбарыг мэдээллийг урсгаж (stream) дамжуулахад, харин гуравдугаар хувилбарыг мэдээллийг урсгалд нь бус нэг удаад унших үед ашиглавал илүү тохиромжтой болсон. Нэгдүгээр хувилбар нь фрэйм бүр нь тусдаа учраас мэдээлэл урсах үед одоогийн фрэймийн дугаарыг буюу хугацааг өөрчилсөн ч шууд тоглуулах боломжтой. Харин хоёрдугаар хувилбар нь өмнөх фрэймийн мэдээлэл хэрэгтэй учраас урсах үед хугацааг нь өөрчлөхгүй тохиолдолд ашиглавал тохиромжтой. Энэхүү алгоритмын гол санаа нь үеийн байрлал болон эргэлтийн 3D координатыг хүний нүдэнд анзаарагдахааргүйгээр тоймлож бүхэл тоонд хадгална. Хоёрдугаар арга нь өмнөх фрэймийн утгатай тэнцүү бол хураана харин гуравдугаар арга нь захын утгаар илэрхийлж болохоор бол хураана. Урсгал, онлайн эсвэл офлайн ямар нөхцөлд ашиглахаас хамаарч нэг, хоёр болон гуравдугаар аргаас сонгож ашиглаж болно. Бусад аргуудаас хэрэгжүүлэхэд хялбар бас эцсийн үр дүнтэй харьцуулбал шахалтын харьцаа нь бага байв.

Энэхүү ажлыг цааш сайжруулахын тулд хоорондоо ойролцоо фрэймүүдийг нэгтгэж фрэймийн тоогоо бууруулснаар файлын хэмжээгээ илүү шахаж болно. Одоогийн хэрэгжүүлсэн гуравдугаар хувилбарын хувьд дундах утгаа шалгахдаа зөрүүг нь илүү их байхаар шалгавал хөдөлгөөний чанар тодорхой хэмжээнд буурах ч гэсэн файлын хэмжээгээ илүү шахах боломжтой. Мөн $EncodeText$ болон $DecodeText$ функцдээ өөр шахах аргуудыг туршиж файлын хэмжээгээ өшөө бууруулж болно. Мөн хөдөлгөөн нь зүүн баруун талдаа адилхан бол зөвхөн нэг талын мэдээллийг хадгалаад файлын хэмжээгээ хоёр дахин шахаж болно.

3D загварын араг ясны бүтцийг мэдэж байгаа бол ямар ч хөдөлгөөн байсан энэ аргыг араг ясанд нь тохируулж өөрчилж ашиглаж шахаж болно. Амьтан, ургамал, робот, хүн хөдөлдөг юу ч байсан болно. Энэ ажилд байрлал дээр зөвхөн хүний биеийн байрлалыг ашигласан харин өөрсдөө өөрчилбөл аль үеүдийн байрлал хэрэгтэйгээс хамаараад хэдэн ч байрлалыг хадгалж болно. Мөн үений харьцааг (scale) 3D координатаар илэрхийлээд зургаан байтдаа хувиргаад ашиглаж болно.

Зүүлт

[1] O. Arikan, "Compression of motion capture databases," *ACM Trans. Graph.*, vol. 25, no. 3, p. 890–897, Jul. 2006. [Online]. Available: https://doi.org/10.1145/1141911.1141971

[2] Q. Gu, J. Peng, and Z. Deng, "Compression of human motion capture data using motion pattern indexing," vol. 28, no. 1, pp. 1–12, 2009.

[3] P. Beaudoin, P. Poulin, and M. van de Panne, "Adapting wavelet compression to human motion capture clips," pp. 313–318, 2007.

[4] G. Liu and L. McMillan, "Segment-based human motion compression," pp. 127–135, 2006.
[5] A. Firouzmanesh, I. Cheng, and A. Basu, "Perceptually guided fast compression of 3-d motion capture data," *IEEE Transactions on Multimedia*, vol. 13, no. 4, pp. 829–834, 2011.
[6] L. Váša and G. Brunnett, "Rate-distortion optimized compression of motion capture data," vol. 33, no. 2, pp. 283–292, 2014.
[7] M. Tournier, X. Wu, N. Courty, E. Arnaud, and L. Reveret, "Motion compression using principal geodesics analysis," vol. 28, no. 2, pp. 355–364, 2009.
[8] P. Wang, Z. Pan, M. Zhang, R. W. Lau, and H. Song, "The alpha parallelogram predictor: A lossless compression method for motion capture data," *Information Sciences*, vol. 232, pp. 1–10, 2013.
[9] S. Li, M. Okuda, and S.-i. Takahashi, "Compression of human motion animation using the reduction of interjoint correlation," *EURASIP Journal on Image and Video Processing*, vol. 2008, pp. 1–15, 2007.
[10] K. Zhou and F. Tian, "A clustering compression method for 3d human motion capture data," *International Journal of Computer Science and Network Security (IJCSNS)*, vol. 14, no. 8, p. 1, 2014.

Хавсралт. Хөдөлгөөн бүрийн үр дүн

Хөдөлгөөний нэр	ФТ	ХХ	ФХ	Хэмжээ			Харьцаа			Хувь		
				Х1	Х2	Х3	Х1	Х2	Х3	Х1	Х2	Х3
Зогсох	250	8.3	768.7	77.0	29.0	**0.8**	1:10.0	1:26.5	**1:961.2**	10.0	3.8	**0.1**
Бэлэн зогсох	300	10.0	986.7	81.6	75.7	**3.5**	1:12.1	1:13.0	**1:281.8**	8.3	7.7	**0.4**
Амьсгалж зогсох	298	9.9	1080.8	91.7	87.8	**14.1**	1:11.8	1:12.3	**1:76.7**	8.5	8.1	**1.3**
Залхуу зогсох	263	8.8	908.8	80.9	64.3	**4.9**	1:11.2	1:14.1	**1:187.1**	8.9	7.1	**0.5**
Хөлөө тулж зогсох	283	9.4	1026.1	87.0	75.8	**5.9**	1:11.8	1:13.5	**1:175.3**	8.5	7.4	**0.6**
Нинжа зогсох	275	9.2	961.0	84.6	48.4	**13.6**	1:11.4	1:19.9	**1:70.4**	8.8	5.0	**1.4**
Гэмтэлтэй зогсох	280	9.3	978.5	86.2	75.8	**12.5**	1:11.4	1:12.9	**1:78.2**	8.8	7.8	**1.3**
Хөгшин хүн зогсох	365	12.2	906.2	112.2	44.3	**6.2**	1:8.1	1:20.4	**1:146.5**	12.4	4.9	**0.7**
Дайчин зогсох	344	11.5	1201.3	105.8	102.5	**23.5**	1:11.4	1:11.7	**1:51.2**	8.8	8.5	**2.0**
Мутант зогсох	427	14.2	1404.2	132.7	119.8	**13.2**	1:10.6	1:11.7	**1:106.5**	9.4	8.5	**0.9**
Алхах	35	1.2	381.8	11.0	10.7	**6.9**	1:34.6	1:35.8	**1:55.0**	2.9	2.8	**1.8**
Шогших	25	0.8	347.3	8.0	6.5	**5.5**	1:43.5	1:53.5	**1:63.2**	2.3	1.9	**1.6**
Удаан гүйх	22	0.7	336.2	7.1	7.0	**5.4**	1:47.7	1:47.8	**1:62.3**	2.1	2.1	**1.6**
Гүйх	19	0.6	325.2	6.1	4.6	**4.5**	1:53.5	1:70.3	**1:71.9**	1.9	1.4	**1.4**
Хурдан гүйх	16	0.5	297.2	5.2	2.4	**2.4**	1:57.0	1:122.1	**1:122.9**	1.8	0.8	**0.8**
Бөхийж алхах	31	1.0	336.8	9.8	4.3	**3.8**	1:34.3	1:78.7	**1:87.9**	2.9	1.3	**1.1**
Гэтэж алхах	32	1.1	339.4	10.1	4.6	**4.0**	1:33.5	1:74.4	**1:84.8**	3.0	1.3	**1.2**
Мөлхөх	54	1.8	454.3	16.8	14.4	**10.0**	1:27.0	1:31.6	**1:45.5**	3.7	3.2	**2.2**
Доогуур мөлхөх	70	2.3	511.8	20.5	14.4	**9.8**	1:25.0	1:35.6	**1:52.1**	4.0	2.8	**1.9**
Мөлхөж гүйх	20	0.7	328.8	6.5	5.6	**5.5**	1:50.5	1:59.1	**1:59.6**	2.0	1.7	**1.7**
Суулт хийх	56	1.9	457.8	17.5	16.7	**7.8**	1:26.2	1:27.4	**1:59.0**	3.8	3.6	**1.7**
Голтой суулт хийх	68	2.3	477.3	21.2	13.8	**4.9**	1:22.6	1:34.6	**1:97.2**	4.4	2.9	**1.0**
Гэдэсний таталт хийх	67	2.2	502.3	20.8	19.5	**9.2**	1:24.1	1:25.8	**1:54.3**	4.1	3.9	**1.8**
Хэвтээд дугуй жийх	33	1.1	342.0	10.4	4.7	**3.4**	1:32.8	1:73.2	**1:101.9**	3.0	1.4	**1.0**
Суниах	46	1.5	377.9	14.4	6.7	**4.1**	1:26.2	1:56.0	**1:92.7**	3.8	1.8	**1.1**
Дөрвөн талдаа үсрэх	61	2.0	475.9	19.0	18.4	**16.2**	1:25.0	1:25.9	**1:29.4**	4.0	3.9	**3.4**
Суниагаад үсрэх	94	3.1	602.0	29.1	23.5	**19.6**	1:20.7	1:25.6	**1:30.7**	4.8	3.9	**3.3**
Гараа савж үсрэх	32	1.1	373.3	10.1	10.1	**8.1**	1:36.9	1:37.0	**1:46.1**	2.7	2.7	**2.2**
Гийр савах	59	2.0	408.5	18.2	7.2	**4.7**	1:22.5	1:56.7	**1:87.6**	4.4	1.8	**1.1**
Сумо таталт хийх	78	2.6	463.4	24.1	9.8	**6.0**	1:19.2	1:47.1	**1:77.1**	5.2	2.1	**1.3**
Шулуун цохих	40	1.3	360.5	12.6	5.6	**4.5**	1:28.7	1:64.1	**1:80.7**	3.5	1.6	**1.2**
Хоёр цохих	59	2.0	410.6	18.4	7.9	**6.8**	1:22.3	1:51.7	**1:60.5**	4.5	1.9	**1.7**
Хажуугаас цохих	65	2.2	426.4	20.2	8.8	**6.7**	1:21.1	1:48.2	**1:63.3**	4.7	2.1	**1.6**
Өлгөж цохих	40	1.3	387.2	11.1	8.7	**8.1**	1:34.8	1:44.6	**1:48.1**	2.9	2.2	**2.1**
Саваж цохих	60	2.0	413.2	18.7	8.2	**5.7**	1:22.1	1:50.1	**1:72.8**	4.5	2.0	**1.4**
Тохойгоор цохих	50	1.7	386.8	15.6	7.0	**5.5**	1:24.7	1:55.7	**1:70.5**	4.0	1.8	**1.4**
Доогуур өшиглөх	45	1.5	421.1	14.1	9.8	**9.2**	1:30.0	1:42.9	**1:45.9**	3.3	2.3	**2.2**
Урдаас өшиглөх	48	1.6	381.6	15.0	6.5	**5.7**	1:25.4	1:58.6	**1:66.8**	3.9	1.7	**1.5**
Жийх	64	2.1	423.8	19.9	8.6	**6.9**	1:21.3	1:49.0	**1:61.3**	4.7	2.0	**1.6**
Дээшээ өшиглөх	75	2.5	531.9	23.3	17.2	**13.7**	1:22.8	1:31.0	**1:38.7**	4.4	3.2	**2.6**
Хип хоп бүжиглэх	134	4.5	536.5	25.9	17.8	**16.9**	1:20.7	1:30.2	**1:31.7**	4.8	3.3	**3.2**
Робот бүжиглэх	463	15.4	1492.1	141.1	123.6	**86.3**	1:10.6	1:12.1	**1:17.3**	9.5	8.3	**5.8**
Чөлөөт бүжиглэх	146	4.9	686.3	44.7	43.3	**33.9**	1:15.4	1:15.8	**1:20.3**	6.5	6.3	**4.9**
Шафал бүжиглэх	225	7.5	904.4	70.1	42.5	**29.5**	1:12.9	1:21.3	**1:30.6**	7.8	4.7	**3.3**
Хаус бүжиглэх	641	21.4	1902.7	196.2	173.1	**97.5**	1:9.7	1:11.0	**1:19.5**	10.3	9.1	**5.1**
Ташаагаар бүжиглэх	283	9.4	971.3	87.1	49.3	**40.6**	1:11.2	1:19.7	**1:23.9**	9.0	5.1	**4.2**
Твэрк бүжиглэх	456	15.2	1507.8	140.1	136.6	**76.5**	1:10.8	1:11.0	**1:19.7**	9.3	9.1	**5.1**
Савж бүжиглэх	741	24.7	2115.8	227.5	202.5	**148.3**	1:9.3	1:10.5	**1:14.3**	10.8	9.6	**7.0**
Салса бүжиглэх	843	28.1	2554.1	255.4	252.0	**161.1**	1:10.0	1:10.1	**1:15.9**	10.0	9.9	**6.3**
Самба бүжиглэх	594	19.8	1789.5	181.7	163.5	**125.2**	1:9.8	1:10.9	**1:14.3**	10.2	9.1	**7.0**

Хүснэгт II: Бүх шахсан хөдөлгөөний үр дүнгийн хүснэгт. ФТ - Фрэймийн тоо, ХХ - Хөдөлгөөний хугацаа Секунд, ФХ - Файлын хэмжээ KB, Х1 - Нэгдүгээр хувилбар, Х2 - Хоёрдугаар хувилбар, Х3 - Гуравдугаар хувилбар, Хэмжээ - Шахсан файлын хэмжээ KB, Харьцаа - Нэг файлын хэмжээнд хэдэн шахсан файл багтах харьцаа, Хувь - Шахсан файлын хэмжээг файлын хэмжээнд харьцуулсан хувь.

Ерөнхий физикийн лабораторийн хичээлд сургалтын Moodle системийг ашигласан нь

Баярсайханы Алтанхүү
Физикийн тэнхим
Монгол Улсын Их Сургууль
Монгол улс, Улаанбаатар
altankhuu@num.edu.mn

Шаравын Мөнхжаргал
Физикийн тэнхим
Монгол Улсын Их Сургууль
Монгол улс, Улаанбаатар
munkhjargal@num.edu.mn

Самбуугийн Мөнхцэцэг
Физикийн тэнхим
Монгол Улсын Их Сургууль
Монгол улс, Улаанбаатар
munkhtsetseg_s@num.edu.mn

***Хураангуй* — Их дээд сургуулийн ерөнхий физикийн лабораторийн хичээлийн хүрээнд оюутны хийж гүйцэтгэсэн лабораторийн ажлыг хамгаалуулахдаа, хэмжилтийн тооцоог математик талаас нь алхам бүрээр нь нягтлан шалгах боломжтой нэмэлт тооцоолуурыг сургалт удирдах нээлттэй эх бүхий Moodle систем дээр Moodle STACK нэмэлт хэрэгслийн (plug-in) тусламжтайгаар боловсруулж, хичээл сургалтын үйл ажиллагаанд туршин ашиглав. Энэ нь мэдээллийн технологийн ололтыг сургалтад нэвтрүүлэн хэрэглэж буй дэвшилтэт нэгэн хэлбэр болно гэж үзэж байна.**

Түлхүүр үг — Moodle, STACK, сургалтын технологи, ерөнхий физик, физикийн сургалт, мэдээллийн технологи, лабораторийн хичээл

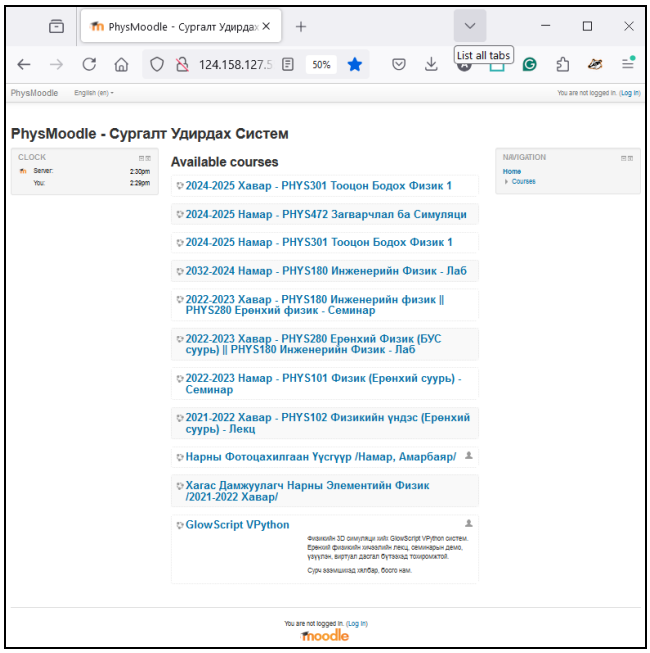
I. Удиртгал

Цар тахлын үед бүх нийтээр цахим сургалтад гэнэт шилжихэд, ерөнхий физикийн хичээлийн лабораторийн сургалтыг хэрхэн цахимжуулж интерактив болгох вэ гэдэг нь тулгамдсан асуудал болж, тухайн хичээлийг заадаг багш бүр өөр өөрийн аргаар үүнийг шийдсэн. Бидний хувьд, цахим сургалтын Moodle систем, түүний нэмэлт (third-party) "багаж" болох STACK хэрэгслийн тусламжтайгаар уг асуудлыг шийдэх гэж оролдсон арга туршлагаа энэхүү өгүүллээр дамжуулан товч хуваалцахыг зорилоо.

Moodle (Modular Object-Oriented Dynamic Learning Environment) нь хичээл сургалтын явцыг удирдан хөтлөх боломж олгодог, нээлттэй эх бүхий үнэгүй цахим талбар (програм хангамж) юм. Өнөөдөр дэлхий даяар олон зуун сая хэрэглэгч сургалтын болон бусад зориулалтаар уг системийг ашиглаж байгаа бөгөөд цаашид энэ тоо улам өсөх хандлагатай [1]. Moodle нь сүлжээний орчинд client-server дизайны дагуу ажилладаг. Веб сервер дээр Moodle-ийг байршуулж (moodle server), харин жирийн хэрэглэгчид онлайнаар веб хөтөч ашиглан системтэй холбогдоно (moodle clients). Хэрэглэгчийн компьютер дээр ямар нэг нэмэлт програм хангамж суулгах шаардлагагүй. Moodle нь Google Forms-той харьцуулахад олон дахин илүү боловсронгуй, сургалт удирдах хэрэгсэл.

Манай боловсролын байгууллагууд Moodle системийг багцаагаар 2009-2010 оны үеэс анх ашиглаж эхэлсэн бололтой [2]. Бидний хувьд интернетэд гарцтай Ubuntu Linux үйлдлийн систем бүхий виртуал сервер түрээслэн, Moodle системийг суулгаж туршилтын журмаар (албан бусаар) ажиллуулж байгаа (Зураг 1) [3].

Бид ерөнхий суурийн лабораторийн хичээлийг ковидын үед дараах байдлаар интерактив болгох гэж оролдсон. Нэгэнт танхимд лабораторийн ажлыг бодитоор хийлгэх боломжгүй тул хэмжилтийн бэлэн утгуудыг оюутанд өгөөд, оюутан зааврын дагуу зөвхөн тооцоогоо хийсэн. Хэмжилтийн тооцоогоо хэрхэн хийснээр нь тухай лабораторийн ажлын үнэлгээг нь өгсөн. Хэмжилтийн тооцоог нь онлайнаар шалгахдаа Moodle системийг ашигласан.



Зураг 1. Тэнхимийн албан бус Moodle систем.

Хэмжилтийн тооцоо гэдэг нь, тухайн ажлын хэмжилтийн алдааг үнэлэхийг хэлж буй. Хялбар тохиолдолд, зөвхөн шууд хэмжилтийн алдаа үнэлэхийн тулд, хэмжилтийн бүл утгын дунджийг олох, хэмжилт бүрийн шууд (абсолют) алдааг олох, бүл хэмжилтийн ерөнхий абсолют алдааг олох (хэмжилтийн тоо цөөн бол шууд дундажлан олох, хэмжилтийн тоо хангалттай олон бол стандарт хазайлт тооцон дундаж квадрат алдаа бодох), харьцангуй алдааг бодох гэсэн математик үйлдлийг дэс дараалан гүйцэтгэх шаардлагатай [4].

Moodle-ийн стандарт хэрэгслүүдийн тусламжтайгаар эдгээрийг гүйцэлдүүлэхэд учир дутагдалтай. Учир нь Moodle дээр шалгалт (quiz) авахад зориулагдсан дараах стандарт асуултын төрлүүд (standard question types) байдаг. Үүнд:

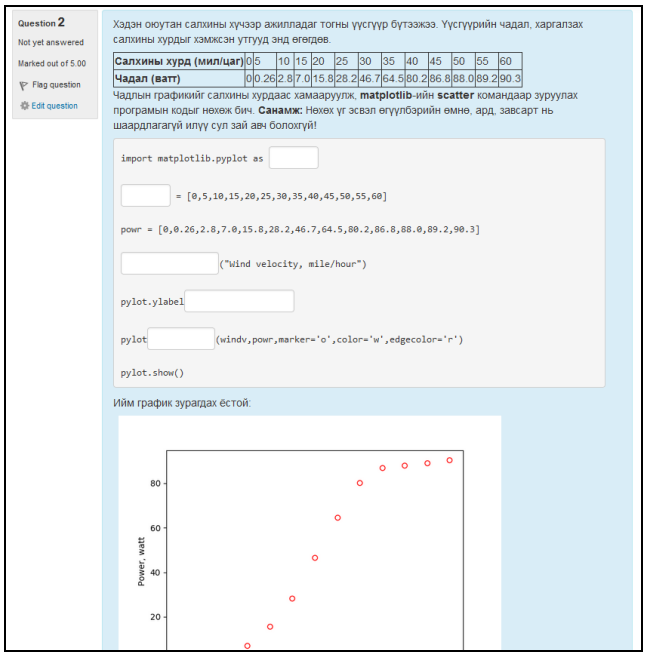
- True/False – Өгөгдсөн хоёр хариултын нэгийг сонгоно;
- Multiple choice – Өгөгдсөн олон хариултаас нэгийг эсвэл хэд хэдийг сонгоно;
- Matching – Жагсаалтаар өгөгдсөн асуулт ба хариултуудыг хооронд нь тааруулна;

- Calculated – Асуултын хариултыг бодож гаргаад тоон утга бичиж оруулна;
- Embedded question – Асуултынхаа текст дунд хариултаа бөглөж бичих зориулалттай. Бусад стандарт төрлийн асуултыг өөртөө агуулж болно

гэх мэт. Тухайлбал оюутан хариултаа бичиж оруулах боломжтой Calculated болон Embedded question type төрлүүдийг авч үзье. Calculated төрлийн асуулт нь энгийн арифметик илэрхийлэл, цөөн тооны элементар функцтэй л ажиллаж чадна. Оюутан зөвхөн тоон утга оруулах боломжтой. Оюутны оруулсан тоон утгыг зөв утгатай жишнэ (тодорхой нарийвчлалтайгаар). Харин Embedded question type нь зөвхөн текстэй ажиллана. Зөв хариултын текстүүдийг (тэмдэгтийн дараалыг) урьдчилан тодорхойлж өгсөн байх ёстой. Оюутан тоо оруулахад түүнийг нь тэмдэгтүүдийн дараалал гэж үзээд "зөв" тэмдэгтүүдийн дараалалтай жишнэ. Оюутан илэрхийлэл оруулахад түүнийг нь зөв илэрхийллийн текстэй жишнэ. Өөрөөр хэлбэл хариултыг шалгахын тулд ямар нэг математик тооцоо цаана нь хийгдэхгүй байна [1]. Зураг 2-т Embedded question type ашиглан үүсгэсэн ийм асуултын жишээг харуулав.

Гэтэл лабораторийн ажлынхаа хэмжилтийн тооцоог шалгуулах гэж буй оюутан зөвхөн тоон утга оруулахаас гадна зөв томъёогоо хэрэглэж байна уу гэдгээ шалгуулахын тулд математик илэрхийлэл юм уу тэгшитгэлийг аналитик хэлбэрээр оруулах шаардлага гарч болно. Дам хэмжилтийн алдаа бодож буй тохиолдолд жишээ нь уламжлал авах томъёо [4] оруулж болно. Дээр дурдсанчлан, энэ бүхнийг гүйцэлдүүлэхэд Moodle-ийн стандарт хэрэгслүүд хангалтгүй учир бид, тэдгээрээс илүү өргөн боломжтой **STACK** нэмэлт хэрэгслийг сонгон авч ажилдаа ашигласан.

STACK (System for Teaching and Assessment using a Computer algebra Kernel) нь, математик тооцооллын Maxima програм [5] дээр суурилсан автомат үнэлгээний хэрэгсэл (automatic assessment package) бөгөөд Moodle системд нэгдэн ажиллах боломжтой [6, 7] байдаг.



Зураг 2. Embedded төрлийн асуулт.

STACK нь, Moodle дээр байршуулсан хичээлийн хүрээнд үүсгэсэн шалгалтын асуултын нэг төрөл болж орж ирдэг. Тиймээс STACK төрлийн асуулт (STACK question type) хэмээн заримдаа нэрийднэ. STACK нь нээлттэй эх бүхий үнэгүй хэрэгсэл.

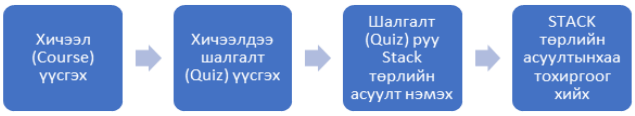
Оюутан тоон утга оруулсан тохиолдолд, STACK хэрэгсэл өөрөө харгалзах томъёог нь бодож, гаргасан утгыг нь оюутны оруулсан утгатай (тодорхой нарийвчлалтайгаар) жишнэ.

Оюутан томъёо бичиж оруулсан тохиолдолд STACK хэрэгсэл нь,

- өөрт буй томъёоны илэрхийллийг оюутны бичиж оруулсан илэрхийлэлтэй жишнэ;
- оюутны бичиж оруулсан томъёог өгөгдсөн параметруудынх нь хувьд бодож хариуг гаргаад өөрт буй томъёоны хариутай жишнэ.

Аль ч тохиолдолд бүх тооцоо үйлдэл нь веб сервер дээр ажиллаж буй Maxima тооцооны систем ("kernel") рүү илгээгдэн хийгдэж байгаа бөгөөд STACK-ийн интерфейс нь "front-end"-ийн үүрэг гүйцэтгэж буй хэрэг юм [5, 6, 7]. Бүрэн хэмжээний компьютерын алгебр систем дээр суурилсан гэдэг утгаараа STACK хэрэгсэл нь илүү хүчирхэг.

Уг хэрэгслийг ашиглахын тулд Moodle дээр гүйцэтгэх алхмуудыг Зураг 3-т схемчлэн харуулав.



Зураг 3. STACK хэрэгслийг ажилд бэлдэх бүдүүвч схем.

Энэ схемийн сүүлийн алхам (STACK төрлийн асуултын тохиргоог хийх) бол STACK-тооцоолуурыг ажилд бэлдэж буй үе шат юм.

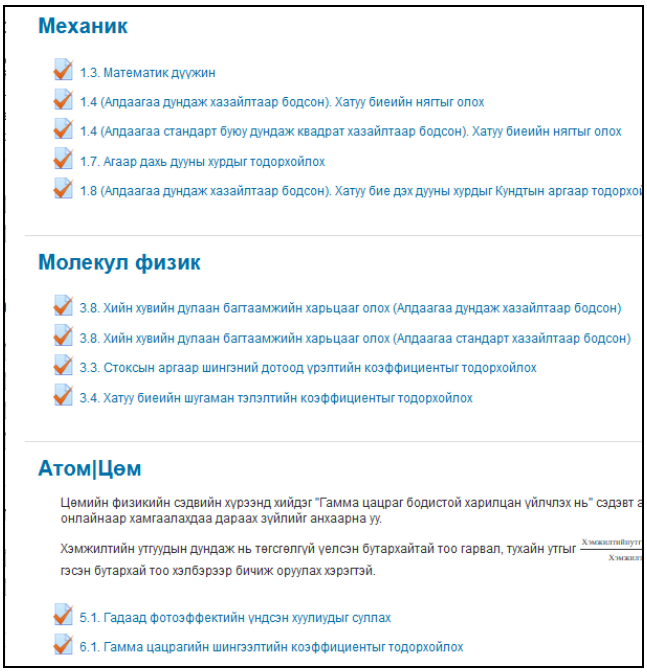
II. STACK-ТООЦООЛУУРЫГ ЕРӨНХИЙ ФИЗИКИЙН ЛАБОРАТОРИЙН ХИЧЭЭЛД АШИГЛАХ НЬ

Бид эхлээд ерөнхий физикийн лабораторийн ажлуудыг тодорхой тооны онол, практикийн асуултаас тогтох шалгалт (quiz) хэлбэрээр Moodle системд оруулж тавьсан (Зураг 4).

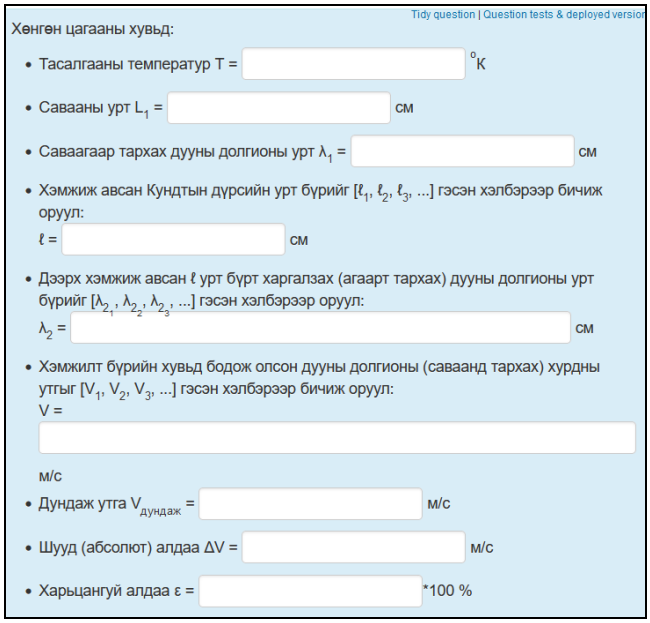
Эдгээрээс, "Механик" бүлгийн №1.8 "Агаарт дуу тархах хурдыг Кундтын аргаар тодорхойлох" сэдэвт ажлын [4] жишээ дээр STACK-тооцоолуурыг авч үзье. Энэ ажил (quiz) нийт 6 асуулттайгаас №1-ээс №4 нь онолын, №5 ба №6 нь STACK төрлийн асуултууд юм. Тухайлбал Асуулт №5 нь оюутанд ямар байдалтайгаар харагдахыг доорх Зураг 5-т үзүүллээ.

Өмнө дурдсанчлан, энэ асуултын тохиргоог хийнэ гэдэг нь STACK-тооцоолуур бэлдэж буй хэрэг юм. Тохиргооны дараалсан үе шатуудыг Зураг 6 болон 7-д харуулав.

Эхлээд бид Question variables буюу тооцоонд ашиглагдах өгөгдлүүдийг (физикийн тогтмолууд, параметрууд, тэдгээрийг холбосон томъёонууд, хэмжилтийн хувьсагчууд гэх мэт) тодорхойлно.

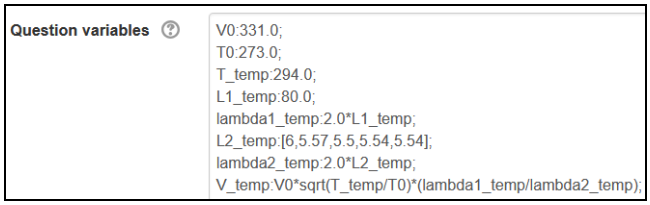


Зураг 4. Moodle системд байршуулсан лабораторийн хичээлүүд (quiz).



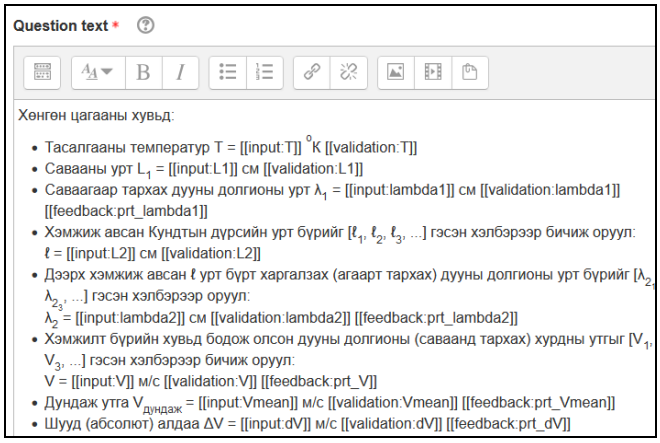
Зураг 5. “Механик” бүлгийн №1.8 ажлын 5-р асуулт (STACK төрөл).

№1.8 ажлын хувьд эдгээр нь агаарт дуу тархах хурд V_0, тасалгааны температур t, саваны урт L_1, саваагаар тархах долгионы урт λ_1, хэмжигдэх ёстой Кундтын дүрсийн урт l, түүгээр тодорхойлогдох долгионы урт λ_2, түүгээр тодорхойлогдох дууны хурд V гэсэн хэмжигдхүүнүүд бөгөөд анхны утга өгч “идэвхжүүлсэн” байна (Зураг 6).



Зураг 6. Тооцоолуурт (хэмжилтэд) ашиглагдах өгөгдлүүд.

Дараа нь асуултын текстийг (Question text) цаанаас тогтоосон форматын [6, 7] дагуу бэлдэнэ (Зураг 7). Энд [[Input: *]] гэж тодорхойлогдсон хэмжигдхүүнүүдийн утгыг гаднаас оруулна; [[validation: *]] гэж тодорхойлогдсон бол, гараас оруулсан утгыг STACK-ийн дүрмийн хувьд зөв бичигдсэн эсэхийг шалгана; [[feedback:prt_*]] гэж тодорхойлогдсон бол, тухайн хэмжигдхүүний утгыг Хариултын мод буюу Potential Response Tree (PRT) хэмээх граф ашиглан анализ хийнэ.

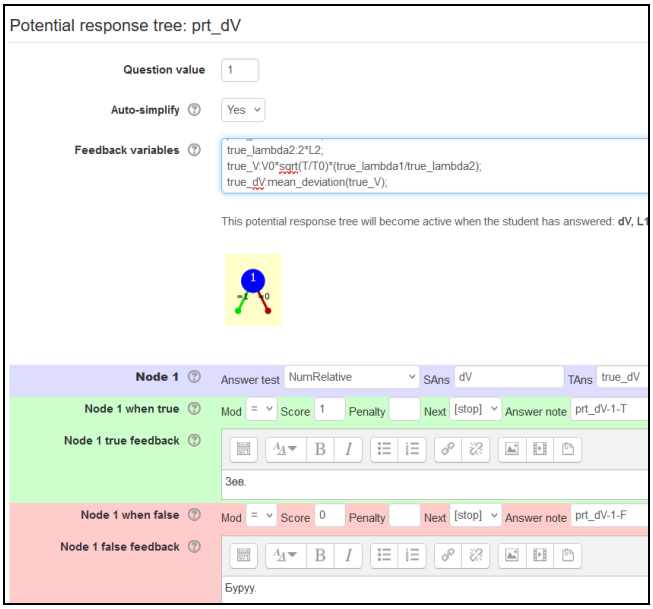


Зураг 7. Асуултын текстийг ингэж бэлдэнэ.

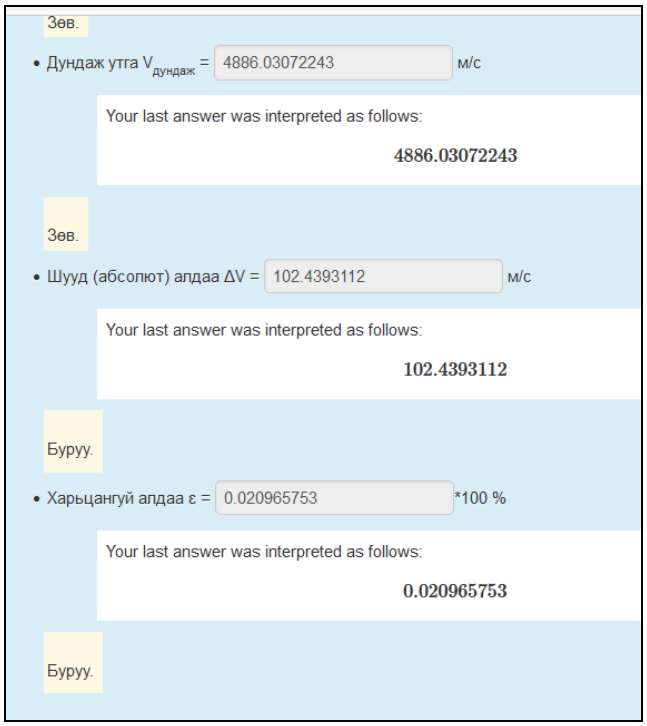
Жишээ нь дараах Зураг 8-д, №1.8 лабораторийн ажлын хэмжилтийн шууд алдаа болох ΔV хэмжигдхүүний утгыг боловсруулан үнэлэх хариултын модын тохиргоог үзүүлэв. Энэ нь ерөөсөө ганцхан зангилаа (Node 1) бүхий хялбар бүтэцтэй мод байна. Хариулт буруу бол 0, зөв бол 1 буюу бүтэн оноо өгнө.

Тухайн хариултын модонд харгалзах тохиолдлыг Зураг 9-д харуулжээ. Энд, оюутны өөрөө бодож олсон (шууд хэмжилтийн) абсолют ба харьцангуй алдаануудын хариу нь хэмжилтийн өгөгдлүүдтэйгээ таарахгүй буй учир тооцоолуур алдаа өгсөн байна.

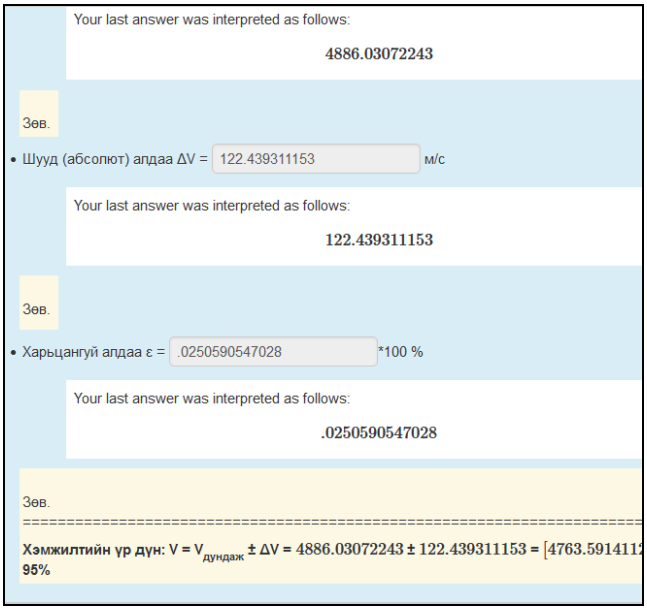
Хэрэв оюутан хэмжилтийн алдаагаа зөв бодож оруулбал тооцоолуур тохирох хариуг үзүүлнэ (Зураг 10).



Зураг 8. Дууны хурдны шууд алдааны (ΔV) хариултыг үнэлэх мод (хялбар тохиолдол)



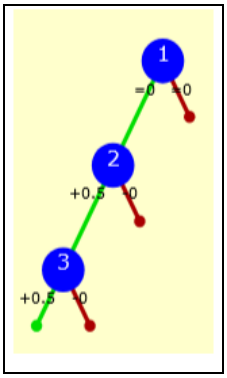
Зураг 9. Оюутан буруу утга оруулах тохиолдол.



Зураг 10. Оюутан зөв утга оруулах тохиолдол.

Хариултын модыг илүү боловсронгуй байдлаар үүсгэж бас болно. Зураг 11 дээр, оюутны тооцоолж гаргасан дууны хурдны шууд алдааны утгыг оруулахад зөв бурууг нь хэрхэн шалгах модыг харуулсан байна.

Дууны хурд V нь λ_1, λ_2 гэсэн параметруудаас хамаардаг. Харин λ_1, λ_2 нь шууд хэмжилтийн утгуудаас хамаарна. Хэрэв оюутан λ_1 юм уу λ_2-ыг (эсвэл хоёуланг нь) хэмжилтээс зөрсөн утга оруулсан байвал, дууны хурдны утгыг шалгах мод нь эхлээд (Node 1) параметрын оруулсан утгуудын хувьд V-г зөв бодсон эсэхийг шалгаад зөв бодсон тохиолдолд, одоо параметруудын утга нь хэмжилтээсээ зөрсөн эсэхийг үргэлжлүүлэн шалгана (Node 2 ба Node 3).



Зураг 11. Хариултын мод (боловсронгуй тохиолдол).

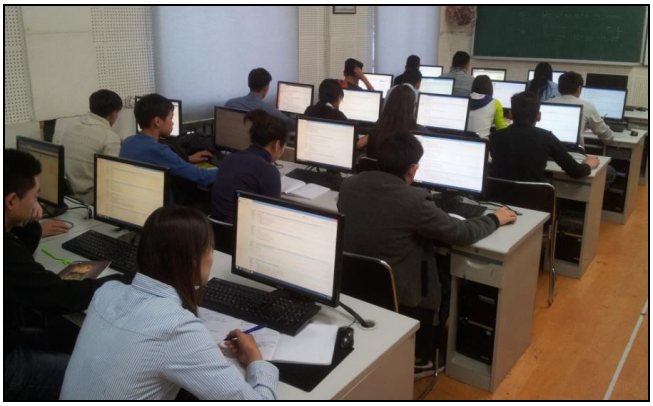
Ингэснээр тухайн асуултад оюутны хариулсан байдлаас хамааран бүтэн, хагас оноо эсвэл 0 оноо өгөх боломжтой болно. Тухайлбал, параметруудын оруулсан утгандаа дууны хурдны хэмжилтийн алдааг зөв тооцсон бөгөөд параметрууд нь өөрсдөө хэмжилтээсээ зөрөөгүй бол бүтэн оноо (0+0.5+0.5), параметруудын оруулсан утгандаа дууны хурдны хэмжилтийн алдааг зөв бодсон ч параметрууд нь өөрсдөө хэмжилтээсээ зөрсөн бол хагас оноо (0+0.5+0) өгнө гэх мэт.

III. ДҮГНЭЛТ

Гэхдээ STACK-тооцоолуур нь багшийн ажлыг бүрэн орлох зориулалтгүй гэдгийг тэмдэглэж хэлэх хэрэгтэй. Учир нь, манай ажлын хувьд тооцоолуур нь бүрэн дүүрэн ажиллагаатай (өөрөөр хэлбэл лабораторийн ажлын тооцоог физик талаас нь оновчтой шалган дүгнэх) хэрэгсэл биш, зөвхөн хэмжилтийн алдааны математик тооцоог нь шалгаж байгаа. Гагцхүү,

- цар тахлын үе шиг танхимаар орох боломжгүй үед лабораторийн хичээлийг онлайнаар өрнүүлж, хэмжилтийн тооцоог үнэлэх нэг арга зам нь байна;
- нэг багш олон оюутантай ажиллаж буй ачаалал ихтэй тохиолдолд оюутан нэг бүрийн хэмжилтийн тооцоог алхам бүрээр нь нэгт нэгэнгүй шалгахад учир дутагдалтай, цаг их авсан байдгийг хөнгөвчилж, шалгах процессын зарим хэсгийг автоматжуулах л зориулалттай.

Ингэхдээ, оюутны зүгээс хууран мэхлэх гэсэн аливаа оролдлогоос урьдчилан сэргийлэх арга хэмжээгээ багш өөрөө авсан байх ёстой. Тухайлбал, оюутнаас лабораторийн ажлын тайланг нь цаас хэлбэрээр давхар хүлээн авч (ковидын үед боломжгүй), тэнд буй хэмжилтийн утгууд болон тооцоо нь, энд онлайнаар оруулсан утгууд болон тооцоотой таарч буй эсэхийг (оюутан бусдаас хуулсан уу, эсвэл хийгээгүй хэмжилтийн тооцоог оруулсан уу гэх мэтийг шалгая гэвэл) хянах. Ковидын үе шиг, бүх үйл ажиллагаа 100 хувь онлайн хэлбэрт орсон үед, асуултад хариулж буй оюутан бүрийг веб камераар нь давхар хянах гэх мэт. Танхимын сургалттай үед, компьютерийн өрөөнд семинарын бүлгэмийг оруулж, шалгалт хэлбэрээр хамгаалуулж болно (Зураг 12).



Зураг 12. Оюутнууд STACK MOODLE систем дээр лабораторийн ажлаа онлайнаар хамгаалж буй нь

Эцэст нь дүгнэж хэлэхэд, Moodle STACK хэрэгсэл нь ерөнхий физикийн лабораторийн хичээлийн үнэлэх болон шалгах үйл ажиллагааг онлайнаар, хагас автомат горимоор өрнүүлэх боломжийг багшид олгож байна. Цар тахлын үед энэ нь нэн тохиромжтой байсан. Түүнчлэн ерөнхий суурийн хичээлийн хувьд нэгэн зэрэг олон оюутантай харьцаж буй үед энэ нь багшийн хөдөлмөрийг жаахан ч гэсэн хөнгөлөх боломж гэж харж байна.

Урьдчилж цаг товлоод, нэг талаас багш, нөгөө талаас оюутнууд онлайнаар (цахимаар) уулзаж, лабораторийн ажлын хамгаалалтыг явуулахад дөхөмтэй байх болов уу.

Сургалт удирдах Moodle системийг хичээлд ашигласнаар мэдээллийн технологийн ололтыг боловсролын салбарын идэвхтэй үйл ажиллагаанд хэрэглэж буй дэвшилтэт нэгэн хэлбэр болох юм.

НОМ ЗҮЙ

[1] Moodle documentation, https://moodle.org.

[2] Б. Уугангэрэл, “Цахим сургалтыг хөгжүүлж буй туршлага: АНУ, БНСУ болон Монголын зарим жишээн дээр”, Монголын хүн амын сэтгүүл, № (390) 22, pp. 52-55, 2013.

[3] PhysMoodle систем: http://124.158.127.55/moodle/

[4] Л. Дэмбэрэл, Физикийн Лабораторийн Практикум, 2 дахь хэвлэл, БитПресс, УБ, 2012.

[5] Maxima Computer Algebra System, https://maxima.sourceforge.io/.

[6] Stack Moodle plug-in, https://moodle.org/plugins/qtype_stack.

[7] STACK | Online assessment, https://stack-assessment.org/.

Чанарт төвлөрсөн Ажайл хөгжүүлэлтийн менежментийн арга болон түүний дэмжих орчин

Баттулгын Цэлмэг
Шинжлэх ухаан,инженерчлэлийн тэнхим
Шибаура Технологийн Их Сургууль
Токио, Япон
Ma22504@shibaura-it.ac.jp

Накашима Цуёоши
Шинжлэх ухаан, инженерчлэлийн тэнхим
Шибаура Технологийн Их Сургууль
Токио, Япон
tsnaka@shibaura-it.ac.jp

Хураангуй- Ажайл/Agile хөгжүүлэлт дэлхий дахинд өргөнөөр хэрэглэгдэж эхлээд байна. Гэвч нүдэнд шууд харагдах функционал шаардлагын хөгжүүлэлтэнд хэт анхаарсанаас, нүдэнд ил харагдахгүй богино хугацаанд сэргээгдэх, модуль бүтэцтэй байх гэх мэт чанарын шаардлагууд нь орхигдох хандлагатай байдаг. Энэхүү чанарын асуудлыг шийдвэрлэхийн тулд, бид бүхэн ISO/IEC 25000 бүлгийн чанарын моделд суурилан чанарын шаардлага, түүний хэрэгжүүлэлт мөн үнэлэлтийг дэмжих зориулалттай чанарын-дижитал ихэр(ЧДИ)-Quality Digital Twin гэх консепцийг хөгжүүлж байна. Энэхүү судалгаагаар, тэрхүү ЧДИ-ийг ашиглан, ажил хөгжүүлэлтийг дэмжих чанарын менежментийн аргыг санал болгон тухайн аргыг туршилтийн үр дүнг харуулахыг зорих болно.

Түлхүүр үг – ажайл хөгжүүлэлт, чанарын менежмент, төслийн менежмент, чанарын модел

I. Удиртгал

Нилээд өмнө нь систем, програм хангамжийн бүтээгдэхүүний хэрэглэгчид систем нь ажиллаж байвал түүний чанарт сэтгэл хангалуун байдаг. Гэсэн хэдий ч өнөө үед хэрэглэгчдийн шаардлага улам боловсронгуй, олон талт болж, системээс хэрэглэгчид бий болгож буй үнэ цэнийг нэмэгдүүлэхийг шаарддаг болсон байна. Үүний үр дүнд хэрэглэгчийн үнэ цэнийг хэрэгжүүлэхэд чиглэсэн **“agile”** хөгжүүлэлт илүү өргөнөөр хэрэглэгдэж байна.

Гэвч ажайл хөгжүүлэлтэнд, хэрэглэгчийн хувьд функционал шаардлага болон түүний хэрэгжүүлэлт нь илүү чухалд тооцогдон, найдвартай байх, аюулгүй байдал гэх мэт чанарын шаардлагууд нь хойш тавигдан улмаар менежментийн хувьд баланстай шийдвэр гаргаж чаддаггүй асуудал бий[1]. Мөн дизайны түвшинд чанарыг шингээх (Quality by design), функционал биш шаардлагын тестийн төлөвлөлт болон хэрэгжүүлэлт нь орхигдох хандлага байдаг[2]. Үүнээс шалтгаалан чанар болон зардлын харьцааг хадгалсан менежмент хийхэд хүндрэлтэй байдаг.

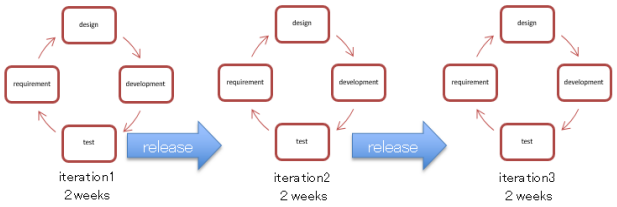
Бид чанарын менежментийн асуудлыг шийдвэрлэхийн тулд ЧДИ консепцийг хөгжүүлж байна[3]. Энэхүү консепц нь ISO/IEC 25000 (SQuaRE) бүлгийн чанарын загварт суурилан, **“agile”** хөгжүүлэлтэд чанарыг төлөвлөх, хэрэгжүүлэх, чанарын үнэлэхэд, дэмжлэг үзүүлдэг систем юм [4] .

Энэхүү судалгаанд бид ЧДИ системийг ашиглан чанарын менежментийн аргыг санал болгож, туршилт хийж үнэлгээ гаргана.

II. Ажайл хөгжүүлэлт ба сорилтууд

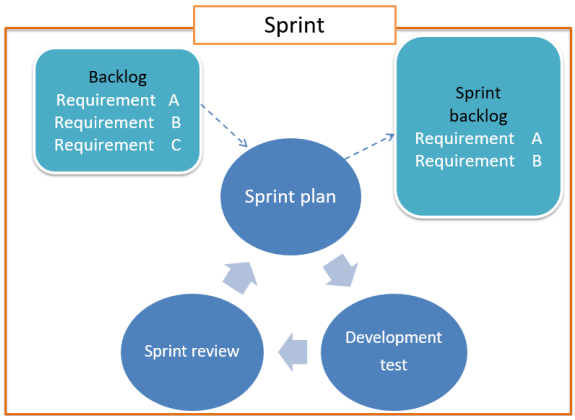
2.1 Ажайл хөгжүүлэлт

Ажайл хөгжүүлэлт нь, Зураг 1-д үзүүлсэнчлэн, тогтоосон богино хугацаанд программ хөгжүүлэлтийг хийж бүтээгдэхүүн гарган, хэрэглэгчийн сэтгэгдлийг хүлээн авч, дараагийн хөгжүүлэлтээрээ сайжруулалт хийх замаар цикл давталттай хөгжүүлдэ арга зүй юм[5].



Зураг 1 Ажайл хөгжүүлэлт

Ажайлын нэг онцлог нь давталт бүрт тодорхой хэмжээний нүдэнд харагдах бүтээгдэхүүн бэлэн болдог



Зураг 2 Спринтийн бүтэц

тул хэрэглэгчийн сэтгэл ханамжийг сонсож, бүтээгдэхүүнээ улам сайжруулах боломжийг олгодог.

2.2 Ажайл скрамм

Скрамм бол хамгийн түгээмэл ажайл хөгжүүлэлийн нэг арга зүй юм [5]. Скрамм-д нэг давталтыг спринт гэж нэрлэдэг бөгөөд спринт ихэвчлэн хоёроос дөрвөн долоо хоног үргэлжилдэг.

Функционал шаардлагуудыг хэрэглэгчийн түүх (user story) гэж нэрлэн цуглуулж боловсруулна. Тэрхүү жагсаалтыг баклог (backlog) гэж нэрлэдэг. Нэг спринтийг төлөвлөхдөө тухайн спринтэд хөгжүүлэх баклогуудийг сонгон, хамгийн чухал, чухал зэргээр нь эрэмбэлнэ. Үүнийг спринтийн баклог гэж нэрлэдэг. Түүний дараа спринтийн баклогийг таск болгон задалж хөгжүүлэлтийн багийн гишүүдэд хуваарилан, хөгжүүлэлтийн ажил эхэлдэг. Спринтийн төгсгөлд гарсан бүтээгдэхүүнийг хянаж үзэх замаар хэрэглэгчээс санал хүсэлтийг сонсоно. Үүнийг спринт хянах (sprint review) гэж нэрлэдэг. Тухайн спринтийн үр дүнд

Хүснэгт 1 Уламжлалт арга болон ажайлын чанар, зардал, хугацааны (Q.C.D) харьцуулалт

	Waterfall	Agile development
Quality (Q)	Constant (ideal)	Changeable
Delivery (T)	Constant Intermediate process settings	Number of times: undecided
Cost (C)	Changeable	Changeable

үндэслэн дараагийн спринтийг төлөвлөж эхэлнэ.

2.3 Ажайл хөгжүүлэлт дэх төслийн менежментийн сорилтууд

Төслийн менежментийн (TM) үндсэн зарчим нь төслийг төлөвлөх, хэрэгжүүлэх, хянах, түүнчлэн хамрах хүрээ, эрсдэлийг харгалзан үзэж төслийнхөө чанар, цаг хугацаа, зардлыг (Q.C.D) тэнцвэржүүлэх явдал юм. [6].

Хүснэгт 1-д уламжлалт хүрхрээ арга ба ажайлын чанарын (Q), цаг хугацаа (D), зардлын (C) шинж чанарыг харууллаа. Ажайл-д спринтийн буюу давталтын тоо хязгааргүй болович зардал, чанар нь хувьсах чадвартай байдаг. Иймд, уламжлалт хүрхрээ аргатай харьцуулан ажайл төслүүдийг удирдахад дараах асуудлуудыг хөндөж байна:

[Сорилт 1] Хүрэхийг хүсч байгаа чанар нь эхлээд тодорхойгүй, хувьсах шинж чанартай байдаг. Энэ нь төслийн хүргэх хугацаа, өртөг зардал, хүрэх ёстой чанарыг тэнцвэржүүлэхэд хэцүү болгодог.

[Сорилт 2] Функционал шаардлагыг л хөгжүүлэх нь нэн тэргүүнд тавигдах хандлагатай байдаг. Тиймээс чанарын шаардлага, функционал шаардлагын тэнцвэртэй шийдвэр гаргахад хэцүү байдаг.

III. Өмнөх судалгаа

Шийдвэр гаргагчид болон менежментийн багийг мэдээллээр хангаж дэмжлэг үзүүлэх зорилготой чанарын стратеги, хяналтын дашбоард систем (QaSD) хөгжүүлэгдсэн байдаг[7]. Зураг 3-т үзүүлснээр, QaSD нь спринтийн явцад чанарын үйл ажиллагааны үр дүнгийн мэдээллийг цуглуулж, Q-Rapids гэх чанарын модел ашиглан, чанарын явцын гүйцэтгэл, бүтээгдэхүүний чанарын явцыг харуулдаг. Гэвч бүтээгдэхүүний чанарыг, зөвхөн согогийн тоо төдийхнөөр тооцдог нь хангалттай биш. Илүү олон талт үзүүлэлтээр чанарыг хэмжих шаардлагатай байдаг.

Зураг 3 Чанарын дашбоард систем QaSD (出典[7])

Мөн Жозеф В. Йодер болон түүний багаас Чанарыг сайжруулахын тулд уламжлалт чанарын тухай ойлголтоос(QA) ажайлын чанар руу шилжихийн тулд дараах 8 ойлголтуудыг хэрэгжүүлэх хэрэгтэй хэмээн санал болгосон[8]:

- Чанарын инженерийг багтаа заавал оруулах,
- Зайлшгүй байх шаардлагатай чанарыг тодорхойл,
- Чанарын хэрэгжих замын зураг, (road map)
- Чанарын хэмжих нэгж тодорхойлох,
- Чанарын буух бүсийг тодорхойлох,
- Чанарын хэрэглэгчийн түүх(user story),
- Чанарын баклог(backlog),
- Чанарын дашбоард,

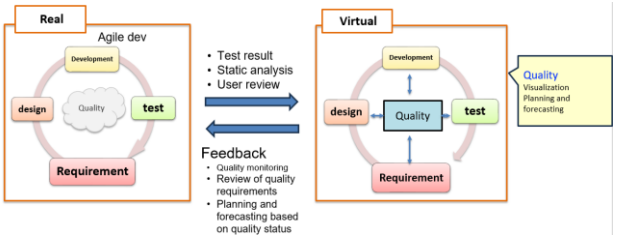
Гэвч эдгээр ойлголтын гол санаа нь чанарын өндөр мэдлэг ур чадвартай инженер ажайл багт зайлшгүй

шаардлагатай хэмээн үзсэн боловч бодит нөхцөл байдалд хэрэгжүүлэхэд хүндрэлтэй байдаг.

IV. Чанарын- дижитал ихэр систем

4.1 Системийн бүтэц

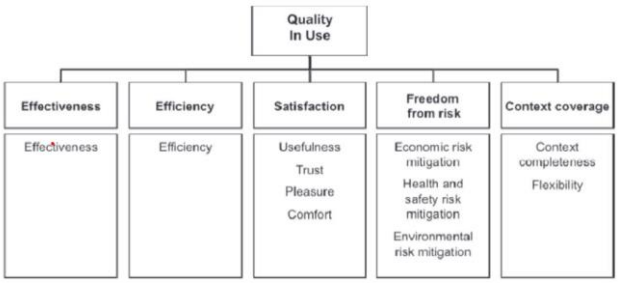
ЧДИ буюу /QualityDigitalTwin/ систем нь ажайл хөгжүүлэлт дэх чанарыг дүрслэн харуулах платформ бөгөөд уг дүрслэлд тулгуурлан хөгжүүлэлт болон менежментэд дэмжлэг үзүүлэх зорилготой. Зураг 4-т үзүүлсэнчлэн чанарын дижитал ихэр систем нь бодит ертөнцийн ажайл хөгжүүлэлтийн туршилт, статик шинжилгээ, хэрэглэгчийн үнэлгээний үр дүнг оруулж, систем автоматаар чанарын төлөвийн модел рүү буулгадаг. Энэ нь бүтээгдэхүүний чанарын өнөөгийн байдлыг харагдуулдаг. Чанарыг дүрслэн харуулснаар ажайл багууд чанарыг хянаж, чанарын шаардлагын гүйцэтгэлийг хянаж, төлөвлөлт, урьдчилсан таамаглалт зэргийг хийх боломжтой.



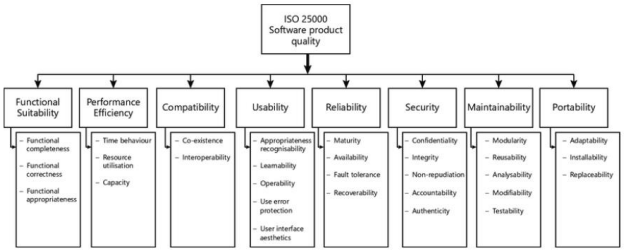
Зураг 4 Чанарын- дижитал ихэр системийн бүтэц

4.2 Чанарын төлөвийн модел

Энэхүү загвар нь ISO/IEC 25000 (SQuaRE) чанарын моделд тулгуурлан чанарын шаардлагыг тодорхойлж, тэдгээрийг хэрэгжүүлэх, үнэлэх аргуудыг **мод** бүтцээр илэрхийлдэг[9]. SQuaRE чанарын моделд програм хангамжийн чанарыг хэрэглэгчийн талаас нь авч үзвэл ашиглалтын чанар (Зураг 5), бүтээгдэхүүний өөрийнх нь үүднээс бүтээгдэхүүний чанар(Зураг 6) гэж тодорхойлогддог. Бүтээгдэхүүний чанар нь функционал бүрэн байдлыг хамт багтаадаг тул SQuaRE чанарын модел нь чанар мөн үйл ажиллагаанд чиглэсэн бүтэн



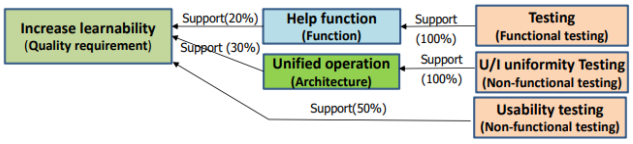
Зураг 5 Ашиглалтын чанарын модел(出典[10])



Зураг 6 Бүтээгдэхүүний чанарын модел(出典[10])

цогц модел юм.

Зураг 7-д чанарын төлөвийн модел ашиглан чанарын дэд шинж чанар болох модулийн задралын жишээг үзүүлэв. Чанарын төлөв байдлын моделд, чанарын шаардлагыг чанарын хэрэгжилт, тэдгээрийг дэмжих чанарын үйл ажиллагаа болгон задалдаг. Чанарын хэрэгжилт нь чанарыг хэрэгжүүлэх үйл ажиллагааг, чанарын үйл ажиллагаа нь чанарын амжилтыг баталгаажуулах үйл ажиллагааг илэрхийлдэг.



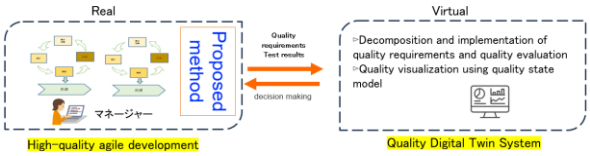
Зураг 7 Чанарын модел

Нэмж өгүүлэхэд чанарын тодорхой шаардлагыг хангасан чанарын хэрэгжилт, чанарын үйл ажиллагаанд хувь нэмэр оруулах түвшинг тогтоодог. Энэ нь чанарын хэрэгжилт, чанарын үйл ажиллагааны амжилтын түвшинг үнэлэх замаар чанарын шаардлагад хүрэх түвшинг доороос дээш тооцох боломжийг олгодог.

V. Санал болгож буй арга

5.1 ТМ-ын сорилтуудыг шийдвэрлэх арга зам

Энэхүү судалгаагаар бид 2.3-т тодорхойлсон ажайл хөгжлийн ТМ-ийн сорилтуудыг шийдвэрлэхийн тулд, ЧДИ платформ ашиглан, төслийн менежерүүдэд



Зураг 8 ЧДИ систем ашигласан чанарын менежмент

шийдвэр гаргахад нь туслах мэдээлэл болон чанарын удирдлагын аргыг санал болгож байна (Зураг 8)

5.2 Санал болгох арга

Энэ арга нь ажайл хөгжүүлэлтэнд өндөр чанарыг хангахад дэмжлэг үзүүлэх менежментийн арга юм. Энэ арга нь

A) спринт төлөвлөлтийн шатанд тухайн спринтэнд хөгжүүлэх чанарын шаардлагуудыг сонгон, тэдгээрийг хэрхэн хэрэгжүүлэх үйл ажиллагааг тодорхойлон, моделд буулгах,

B) спринт хянах шатанд(sprint review), тухайн бүтээгдэхүүний чанарын түвшинг шалган, менежментийн оновчтой шийдвэр гаргах зэргээс бүрдэнэ. Дэлгэрэнгүй:

A. Спринт төлөвлөлт: хөгжүүлэх чанарын шаардлага, хэрэгжүүлэх үйл ажиллагааг тодорхойлох

S1. Чанарын шаардлагыг чанарын хөгжүүлэлт・чанарын үнэлэх үйл ажиллагаа болгон хуваах задлах

Спринтийн төлөвлөлтийн үед ажайл баг нь тухайн спринтэд хөгжүүлэх чанарын шаардлагыг сонгон эдгээр шаардлагыг биелүүлэхийг тулд чанарыг хэрхэн хөгжүүлэх, чанарын хэрхэн үнэлэх үйл ажиллагааг сонгоно. Ингэхдээ чанарын моделийг ашиглан дараах хоёр хувилбараас чанарын шаардлагыг хэрхэн хэрэгжүүлэхээ сонгоно:

• Хөгжүүлэлтийн хувьд тодорхой загвар аргуудаас сонгох.

• Чанарын хэмжилт үнэлгээний хувьд статик шинжилгээ, функционал/функционал бус тест хэрэглэгчийн тойм дундаас тохирох аргыг сонгох

Дээрх чанарын шаардлагын задралыг ЧДИ систем

Зураг 9 Чанарын модел дэлгэц

ашиглан гүйцэтгэх боломжтой. Зураг 9-д систем дээрх жишээг үзүүлэв.

S2. чанарын хөгжүүлэлт・чанарын үнэлэх үйл ажиллагааг үүрэг даалгаварт задалж, түүний зардалыг тооцох

Ажайл багууд сонгосон үүрэг даалгавар бүрийн зардлыг (зарцуулагдах хүч чармайлт) тооцно. Чанарын дижитал ихэр системд ажлын зорилго, сонгосон чанарын хэрэгжилт, үйл ажиллагааны цар хүрээг тодорхойлох замаар зорилтыг тодорхойлдог. Зардлын тооцоог автоматаар тооцож харуулдаг.

S3. Хүний нөөцийн нөхцөл байдалтай уялдуулан чанарын хэрэгжүүлэлтийн төлөвлөгөөг гаргах

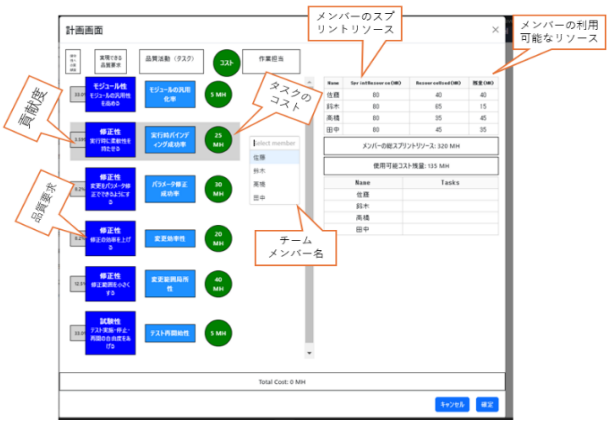
Ажайл баг спринтэд байгаа хүний нөөцийн байдлаа харгалзан үзэж спринтээр хийхийг хүсч буй чанарын шаардлага болон тэдгээрийн үнэ цэнийг (систем чанарт оруулах хувь нэмэр) тогтоох:

•Спринтээр хөгжүүлэх чанарын шаардлага

•Чанарын шаардлагыг хөгжүүлэлт хийхийн тулд гүйцэтгэх даалгаварууд

•багийн гишүүдэд даалгавар хуваарилах

Энэ алхамыг Зураг10-д харуулав. Зураг дээр хараахан хийгдээгүй ажлууд бөгөөд тэдгээрийн зүүн талд байгаа "Чанарын бодит шаардлага" нь даалгавруудыг амжилттай гүйцэтгэсэн тохиолдолд хүрч болох чанарын шаардлага юм. Цаашилбал, тухайн чанарын шаардлага нь хамаарах чанарын шинж чанарт оруулсан хувь нэмрийг зүүн талд харуулав. Даалгаврын нөөцийн хуваарилалтыг нөөцтэй гишүүдийг сонгох замаар хийж болно. Өөрөөр хэлбэл, аливаа ажилд нөөцийг



Зураг10 Төлөвлөлтийн дэлгэц

хуваарилах үед нөөц нь багасч, хариуд нь чанарын үнэ цэнийг бий болгодог.

Эдгээр гурван алхамын үр дүнд нь ажайл багууд спринт төлөвлөлтийн шатанд, хөгжүүлэлт хийхийг хүсч буй чанарын шаардлага болон түүний зардалын тооцоо хоёрын тэнцвэрийг хангасан төлөвлөлт хийх боломжтой.

B. Спринт хянах шатанд(sprint review)：чанарын хэрэгжилтийг шалган төслийн менежментийн шийдвэр гаргах

Спринтийн үеэр чанарын үнэлэх хэмжих үйл ажиллагаанаас туршилтын үр дүн болон бусад өгөгдлийг цуглуулж, чанарын төлөв байдлын модел ашиглан чанарын, шаардлагад хангах түвшинг шалгаж болно. Үүнээс гадна чанарын үзүүлэлт бүрийнх нь хүрэх түвшинг ЧДИ системийн "чанарын дашбоард" ашиглан шалгаж болно. Энэ нь ажайл багуудад спринт

хяналтын үеэр чанар, функционал шаардлагыг тэнцвэржүүлсэн ТМ шийдвэр гаргах боломжийг олгодог. Зураг 11

Зураг11 ЧДИ - Чанарын дашбоард

VI. Туршилт, үнэлгээ

6.1 Туршилтийн зорилго

ЧДИ систем ашиглан, санал болгож буй менежментийн аргыг чанар, зардлыг тэнцвэржүүлэхэд ТМ-д үр дүнтэй эсэхийг баталгаажуулахаар туршилтийг зохион байгуулсан.

6.2 Туршилтийн арга

Оролцогчид: SQuaRE Чанарын моделийн мэдлэгтэй магистрийн 5 оюутан,

Хэрэгжүүлэх дараалал:

(1) Оролцогчид ЧДИ системийг ашиглан, өгөгдсөн зохиомол нөхцөлд тулгуурлан, Ажайл хөгжүүлэлтийн спринтийн чанарын төлөвлөгөөг гаргах

(2) Чанарын менежменттэй холбоотой таван асуулттай Анкетанд хариулж, 1-5 оноогоор үнэлгээ хийсэн.

6.3 Туршилтын үр дүн

Хүснэгт 4-т туршилтийн үр дүнг харуулах бөгөөд 5 оролцогч дунджаар 20 минутийн хугацаанд гүйцэтгэсэн. Үр дүнгийн хүснэгт болон оролцогчдийн үлдээсэн сэтгэгдэлээс доорхи дүгнэлт гаргасан.

(1) Асуулт нэг: Чанарын менежментийн төлөвлөлт хийхэд хангалттай мэдээлэл авч чадсан эсэх

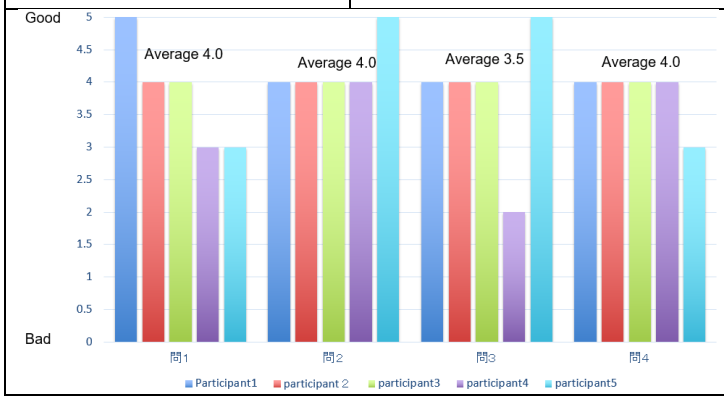
Дундаж үнэлгээ нь 4, нэлээн туслалцаа дэмжлэг болж чадсан гэсэн үнэлгээ гарсан. Оролцогчдийн өгсөн сэтгэгдэлд чанарын шаардлагыг хөгжүүлэхэд зарцуулах зардал, хүрэх үр дүн зэрэг мэдээллүүд нэг цогц учир чанарын төлөвлөлт хийхэд хялбар байв.

(2) Асуулт хоёр: Чанарын зардалын тооцоо тодорхой байх нь төслийн менежмент хийхэд үр дүнтэй эсэх

Дундаж үнэлгээ нь 4, менежментэд үр дүнтэй гэсэн үнэлгээ гарсан. Оролцогчдийн өгсөн сэтгэгдэлд тухайн нэгж чанарын шаардлагын эзлэх хувь зэрэг мэдээллүүд нь харахад маш ойлгомжтой байсан.

Хүснэгт2 Туршилтын үр дүн

Оролцогчид	5
Дундаж хугацаа	20минут



(3) Асуулт гурав: Чанар болон зардалын балансыг харж чадсан эсэх

Дундаж үнэлгээ нь 3.5, чанар болон зардалын балансыг харж чадсан боловч бусдаасаа бага үнэлгээтэй байна. Оролцогчдийн өгсөн сэтгэгдэлд чанар болон зардалын баланс гэдэг ойлголт нь тодорхойгүй.

(4) Асуулт дөрөв: Харагдах байдал нь ойлгомжтой байсан эсэх

Дундаж үнэлгээ нь 4. Харагдах байдал нь технологийн мэргэжлийн бус хүнд ч ойлгомжтой үнэлгээ гарсан. Зардал, тооцоолол нь харахад ойлгомжтой байв.

Дээрх үр дүнгээс үзэхэд, санал болгож байгаа чанарын менежментийн арга болон түүнийг ЧДИ систем дээрх хөгжүүлэлт нь ажайл дэх [Сорилт 1] -ийг шийдвэрлэх боломжтой гэдгийг харуулж байна.

VII. Дүгнэлт

Энэхүү судалгаанд бид ажайл хөгжүүлэлт дэх төслийн менежментийн сорилтуудийг онцлон гаргаж, ажайл төслийн менежментийг дэмжих тэр дундаа чанарт голчилсон менежментийн аргыг санал болгосон. ЧДИ систем дээр дэмжих хөгжүүлэлт хийн, туршилтыг хийсэн бөгөөд, үр дүн нь санал болгосон менежментийн арга нь чанар болон зардалын балансыг хадгалсан менежментийн шийдвэр гаргахад үр дүнтэй гэдэг нь харагдсан.

Цаашид хийгдэх судалгааны төлөвлөгөө

Ажайл хөгжүүлэлт дэлхий дахинд өргөн нэвтэрч буй тул манай улсад ч мөн адил сүүлийн жилүүдэд технологийн компаниуд, банкны салбарт ажайл

нэвтрүүлээд байгаа нь хүснэгт 3–аас харагдана.

Хүснэгт 3 Одоогоор ажайл хэрэглэдэг компани

Нэр	Ажайл хэрэглсэн жил	Технологийн салбарын ажилчдын тоо	Арга зүй (method)
Мобиком	4 (2021)	210 гаруй	Scrum
Хаан банк	4 (2021)	400 гаруй	Scrum
Ай си ти групп	3(2022)	30 гаруй	Scrum
Анд глобал	4(2021)	110 гаруй	Scrum

Дээрх хүснэгтээс үзэхэд Монгол улсын тэргүүлэх технологийн компаниуд ажайл аргыг 3-4 жилийн өмнө нэвтрүүлсэн байна.

Дижитал шилжилтийн гол зорилго нь иргэдийн амьдралын чанарыг дээшлүүлэхэд чиглэдэг билээ. Тиймээс компанийн бүтээгдэхүүн үйлчилгээ нь чанартай, хэрэглэгчдээ үнэ цэнтэй байх хэрэгтэй.

- Дотоодод чиглэсэн шаардлага: Монгол улсад дотооддоо нийлүүлж байгаа програм хангамжийн бүтээгдэхүүний чанарыг дээшлүүлэх,
- Гадагшаа чиглэсэн шаардлага: Манай компаниуд дэлхийн зах зээлд технологийн бусад компанитай өрсөлдөн амжилттай ажиллахын тулд бүтээгдэхүүний чанарыг дээшлүүлэх нь нэн чухал юм.

Ном зүй

[1] Colin Werner, et al. QaSD: Continuously Managing NFRs: Opportunities and Challenges in Practice (2021)

[2] アジャイル開発のプロジェクトマネジメントと品質マネジメント 58Q＆A で学ぶ (2021)

[3] 中島　毅：ユーザ主導の高品質なアジャイル開発を SQuaRE で支援する構想，IPSJ SIGSE211 (2022)

[4] 独立行政法人情報処理推進機構：つながる世界のソフトウェア品質ガイド（2015）

[5] 独立行政法人情報処理推進機構(IPA) .アジャイル開発の進め方, 2020.

[6] プロジェクトマネジメント BOK 知識体系ガイド第 7 版, 2021

[7] L.López a,*, M.Manzanoa, et al."QaSD: A Quality-aware Strategic Dashboard for supporting decision makers in Agile Software Development," (2020)

[8] Joseph W. Yoder , Rebecca Wirfs-Brock, Ademar Aguiar"QA to AQ：Patterns about transitioning from Quality Assurance to Agile Quality"(2017)

[9] 島田他：アジャイル開発における品質要求の展開と評価を支援する機能の提案と実装，電子情報通信学会 総合大会，2024 年 3 月.

[10] JIS X 25010:2013

[11] プロジェクトマネジメント BOK 知識体系ガイド第 6 版, 2017

[12] 福田怜哉：品質デジタルツインにおける SQuaRE モデルに基づく非機能テストカタログの提案とその利用を支援するシステムの実装

Монгол Улсын нийтийн өргөдлийн системд хийсэн үнэлгээ

Сайннямбуугийн Баясгалан
МУИС-ийн МТЭС-ийн Мэдээлэл, компьютерын ухааны тэнхимийн магистрант,УИХ-ын Тамгын газрын референт, Улаанбаатар хот, Монгол
bayasgalan.s@parliament.mn

Самбуугийн Уянга*
МУИС-ийн МТЭС-ийн Мэдээлэл, компьютерын ухааны тэнхимийн профессор, доктор (PhD), Улаанбаатар хот, Монгол,
uyanga@seas.num.edu.mn

Сайнбилэгийн Алунгоо
УИХ-ын Тамгын газрын ахлах референт, Улаанбаатар хот, Монгол,
alungoos@parliament.mn

Алтанцэцэгийн Мөнхзул
МУИС-ийн МТЭС-ийн Мэдээлэл, компьютерын ухааны тэнхимийн багш, Улаанбаатар хот, Монгол,
munkhzul.a@num.edu.mn

**Холбоо барих зохиогч*

Хураангуй - XXI зууны төрийн шинэчлэлийн нэг чухал хөдөлгөгч хүч нь мэдээлэл, харилцаа холбооны технологийн дэвшил юм. Өдгөө дэлхийн бүх улс цахим засаглалын системийг боловсронгуй болгох замаар төрийн үйлчилгээг иргэдэд илүү хүртээмжтэй, шуурхай, хяналттай, ил тод хүргэх боломжийг эрэлхийлж байна. Төрийн цахим үйлчилгээний гол зорилго нь иргэд байгууллагын хоорондын мэдээлэл солилцоог автоматжуулах, төрийн нээлттэй байдлыг хангах, иргэдийн оролцоог өргөжүүлэхэд чиглэдэг боловч үүнийг практикт нэвтрүүлэхэд багагүй бэрхшээл тулгарсаар байна. Тухайлбал, итгэлцэл, мэдээллийн аюулгүй байдал, системийн найдвартай байдал, хэрэглэгчийн ойлголт зэрэг хүчин зүйлс шууд болон шууд бусаар нөлөөлж байна. Энэхүү судалгаагаар UTAUT загва ашиглан мэдээллийн системийн хэрэглээнд хэрэглэгчийн зан төлөв хэрхэн нөлөөлж буйг Монгол Улсын парламентын нийтийн өргөдлийн судлан шинжилсэн болно. Судалгаанд 500 орчим түүвэр цуглуулж, чанарын хяналт хийсний дараа 300 итгэмжлэгдсэн хариултыг ашиглан бүтцийн тэгшитгэлийн загварчлал (SEM) хийсэн. Судалгааны үр дүнгээс үзэхэд технологийн ойлгомжтой байдал, төрд итгэх итгэл, мэдээллийн чанарт эерэг нөлөөтэй, харин хувийн мэдээллийн талаарх эргэлзээ нь хэрэглээг сааруулдаг болох нь харагдлаа.

Түлхүүр үг - Төрийн цахим үйлчилгээ, итгэлцэл, хувийн мэдээлэл, эрсдэл, технологийн хүлээн авалт, UTAUT, парламентын нийтийн өргөдлийн систем.

I. Удиртгал

Монгол Улс сүүлийн жилүүдэд цахим засаглалын бодлогоо эрчимтэй хэрэгжүүлж, бодит платформ болон системийн түвшинд нэвтрүүлээд байна. Бодлогын түвшинд "e-Mongolia" цахим системийн боломжид тулгуурлан төрийн бусад харилцааны хэлбэрийг цахимжуулах үйл ажиллагаа үе шаттайгаар хийгдэж байгаа бөгөөд үүний нэг тод жишээ нь парламентын нийтийн өргөдлийн цахим систем (petition.parliament.mn) юм. Энэхүү цахим систем нь иргэд төрд чиглэсэн санал, гомдол, санаачилга, дуу хоолойгоо парламентын түвшинд цахимаар илэрхийлэх боломжийг олгож буй онцлог систем юм.

Нийтийн өргөдлийн цахим систем нь ардчилал, олон нийтийн оролцооны түвшинийг илэрхийлэх хөгжлийн шинэ хүчин зүйл болон өргөжиж, парламентын ил тод байдал, иргэдийн оролцоог хангах үр дүнтэй хэрэгсэл болж байна. Энэхүү цахим систем нь уламжлалт өргөдөл гомдлын арга хэлбэрийг цахим орчинд шилжүүлж, иргэдийн санал бодол, шаардлага, дуу хоолойг хууль тогтоох байгууллагад хүргэх шууд суваг болж чаджээ. Гэхдээ улс орон бүрийн хөгжлийн түвшин, улс төрийн соёл, хууль эрх зүйн орчин, технологийн дэд бүтцээс хамааран нийтийн өргөдлийн цахим системийн хэрэглээ, нөлөө, хүлээн авалт нь ихээхэн ялгаатай байна [10].

Дэлхийн бусад орнуудын туршлагаас харахад нийтийн өргөдлийн цахим системийн амжилттай хэрэгжилт нь дараах хүчин зүйлстэй уялдаатай байна. Үүнд:

- Нийтийн өргөдлийг шийдвэрлэх үйл явц тодорхой, үр дүнтэй, хариуцлагатай байдал.
- Нийтийн өргөдлийг олон нийт ил тод хэлэлцэх, хянах боломж.
- Иргэний хувийн мэдээлэл хамгаалах бодлого, мэдээллийн технологийн найдвартай байдал.
- Нийтийн өргөдөл бодит үр нөлөөтэй, хууль эрх зүйн хүчин төгөлдөр байдал.
- Иргэд нийтийн өргөдлийг идэвхтэй ашиглах цахим соёл, дижитал бичиг үсэгт тайлагдсан байдал.

Монгол Улсын парламентын нийтийн өргөдлийн цахим систем нь олон улсын сайн туршлагатай харьцуулахад техникийн хувьд боломжтой ч иргэдийн итгэлцэл, ойлголт, оролцооны түвшин харьцангуй сул хэвээр байна. Түүнчлэн нийтийн өргөдлийн шийдвэрлэлтийн ил тод байдал, хариу арга хэмжээний механизм нарийвчлан тодорхойлогдоогүй байгаа нь системийн үр өгөөжид сөргөөр нөлөөлж байна. Харин олон улсын жишээнээс үзвэл нийтийн өргөдлийн цахим системийг зөвхөн бүртгэлийн хэрэгсэл бус, шийдвэр гаргалтад нөлөөлөх цахим ардчиллын суурь хэлбэр болгох нь төр-иргэний харилцаанд итгэлцэл бий болгоход чухал нөлөөтэй болохыг харуулж байна.

II. Цахим системийн үнэлгээний онол

Гэсэн хэдий ч төрийн цахим системүүдийн хэрэглээнд тулгамдаж буй нэг чухал асуудал бол иргэдийн хувийн мэдээлэлд хандах итгэлцэл, системийн аюулгүй байдлын хүлээлт, мэдээлэл алдагдах эрсдэл, мөн технологийг хүлээн авах сэтгэлзүй, туршлага юм. Олон улсын түвшинд энэ асуудлыг UTAUT, TAM, Trust in E-Government, Privacy Concerns, Perceived Risk зэрэг онолын хүрээнд судалж ирсэн хэдий ч хөгжиж буй орнууд, тэр дундаа Монголын Улсын парламентын нийтийн өргөдлийн

системийн хүрээнд иж бүрэн судалгаа өнөөдрийн байдлаар хийгдээгүй байна.

Энэхүү судалгаа нь UTAUT загварын уламжлалт хувьсагчид болох “Гүйцэтгэлийн хүлээлт” (Performance Expectancy), “Хялбар хэрэглээний хүлээлт” (Effort Expectancy), “Нийгмийн нөлөөлөл” (Social Influence), “Дэмжих нөхцөлүүд” (Facilitating Conditions)-ийг “Итгэлцэл” (Trust in Government & Technology), “Хувийн мэдээллийн талаарх санаа зовнил” (Privacy Concerns), “Мэдээллийн эрсдэлийн ойлголт” (Perceived Risk) зэрэг хувьсагчидтай уялдуулан нэгтгэж шалгуур үзүүлэлт болгож Монголын Улсын парламентын нийтийн өргөдлийн системийн хэрэглээний зан төлөвийг судалж, нөлөөллийг тодорхойлох зорилготой.

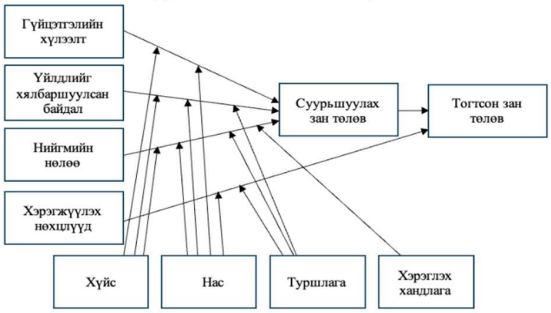
А.Онолын үндэслэл

Технологийн хүлээн авалтын нэгдсэн онол (Unified Theory of Acceptance and Use of Technology) нь хэрэглэгчид технологийг яагаад хүлээн авч, хэрэглэх зан төлөвт шилждэгийг тайлбарлахад өргөн ашиглагддаг онолын загвар юм [9]. Тус онол нь дараах найман онолыг нэгтгэсэн цогц загвар юм. Үүнд:

1. Үндэслэлтэй үйлдлийн онол (Theory of Reasoned Action, TRA).
2. Төлөвлөсөн зан үйлийн онол (Theory of Planned Behavior, TPB).
3. Технологийг хүлээн авах загварын онол (Technology Acceptance Model, TAM).
4. Мотивацийн загварын онол (Motivational Model, MM).
5. Компьютерын хэрэглээний загварын онол (Model of PC Utilization, MPCU).
6. Инновацын тархалтын онол (Innovation Diffusion Theory, IDT).
7. Нийгмийн танин мэдэхүйн онол (Social Cognitive Theory, SCT).
8. TAM болон TPB хосолсон онол.

Технологийн хүлээн авалтын нэгдсэн онол (цаашид “UTAUT” гэх) нь гүйцэтгэлийн хүлээлт, хялбар хэрэглээний хүлээлт, нийгмийн нөлөөлөл, дэмжих нөхцөлүүд гэсэн үндсэн дөрвөн хүчин зүйлээр хэрэглэгчийн ашиглах зорилго (behavioral intention) болон бодит хэрэглээ (use behavior)-г тайлбарладаг. UTAUT загварыг цахим засаглалын системд нэвтрүүлэх судалгаанд өргөнөөр ашигладаг. Тухайлбал цахим төрийн үйлчилгээний хэрэглээнд дээр дурдсан дөрвөн хүчин зүйл тус бүр нь статистикийн хувьд болон ач холбогдлын хувьд эерэг хамааралтай болохыг баталсан байна [8]

Зураг 1.UTAUT загвар [8]



“Гүйцэтгэлийн хүлээлт” нь тухайн системийн үр ашиг, үр нөлөөг хэрхэн ойлгож байгааг, харин “Хялбар хэрэглээний хүлээлт” нь хэрэглэгч системийг хэр төвөггүй, ойлгомжтой гэж үзэж байгааг тус тус илэрхийлдэг. Монголын нөхцөлд эдгээр хүчин зүйлс нь ялангуяа орон нутгийн иргэд, өндөр настнуудын хувьд илүү хүчтэй нөлөөтэй байж болох юм.

Цахим засаглалын хүрээнд “Итгэлцэл” нь хэрэглэгчдийн хүлээн авалт, системд бүртгүүлэх эсэх, мэдээлэл өгөх эсэх зэрэг зан төлөвийн шийдвэр гаргалтад шууд нөлөөлдөг [4]. Итгэлцлийн онолд хэрэглэгчийн төрд итгэх итгэл (Trust in Government) болон технологид итгэх итгэл (Trust in Technology) гэсэн хоёр үндсэн бүрэлдэхүүн хүчин зүйлсийг ялгаж үздэг. “Төрд итгэх итгэл” нь “төр засаг иргэдийн мэдээллийг хэр найдвартай хамгаалж чадах вэ гэсэн хүлээлт” байдаг бол “Технологид итгэх итгэл” нь “тухайн системийн техникийн найдвартай байдал, гүйцэтгэл, хамгаалалт, тогтвортой ажиллагаанд суурилсан итгэлцэл” байдаг. Иргэдийн хувьд эдгээр итгэлцлийн түвшин нь цахим системд хувийн мэдээлэл өгөх эсэх, бүртгүүлэх эсэх, ашиглах эсэхэд шууд нөлөөлдөг болох нь судалгаагаар батлагдсан. Тус судалгаагаар төрийн цахим системд итгэх иргэдийн итгэлцэл сул байх нь хэрэглээнд бодитоор нөлөөлж, хэрэглээг үлэмж хэмжээгээр бууруулж шууд хамааралтай болох нь батлагдсан [3].

Хэсэг судлаачдын боловсруулсан APCO (Antecedents privacy concerns outcomes) загвар нь хэрэглэгчийн хувийн мэдээлэлтэй холбоотой санаа зовнилыг хэмжихэд өргөн ашиглагддаг [6]. Уг загвараар хэрэглэгчийн хувийн мэдээллийг төр, байгууллага хэрхэн цуглуулж, ашиглаж, хадгалж буй талаарх хүлээлт, эргэлзээ зэргээс шалтгаалан төрийн цахим системийг хэрэглэх шийдвэрт хэрхэн нөлөөлж байгааг судалдаг.

Нийтийн өргөдлийн цахим систем нь иргэний нэр, регистрийн дугаар, хаяг, цахим гарын үсэг зэрэг хувийн мэдээллийг авдаг тул тухайн системийг хэрэглэх хязгаарлагч хүчин зүйл болдог. Өөрөөр хэлбэл хувийн мэдээллийн талаарх санаа зовнилыг хэрэглэгчид үүсгэдэг аж. Энэ нь ялангуяа мэдээллийн алдагдал, зөвшөөрөлгүй хандалт, гуравдагч этгээдийн ашиглалттай холбоотойгоор улам нэмэгддэг.

Мэдээллийн эрсдэлийн ойлголт нь цахим үйлчилгээ ашиглахад хэрэглэгчийн сэтгэлзүйд үүсдэг сөрөг үр дагаврын магадлал, түүнд хандах хүлээлтийг илэрхийлдэг [5]. Энэ нь санхүүгийн эрсдэл, мэдээлэл алдах эрсдэл, нэр хүндийн эрсдэл зэрэг хэлбэрээр илрэх боломжтой. Манай улсын хувьд мэдээллийн эрсдэлийн ойлголт нь цахим гарын үсэг, иргэний мэдээллийн нэгдсэн сан зэрэг бүртгэлийн мэдээллийг төрийн системд оруулах шаардлагатай үед хүчтэй яригдаж, хэлэлцүүлэг хийгддэг. Эрсдэл өндөр байна гэсэн ойлголт нь тухайн хэрэглэгчийг системээс татгалзахад хүргэх боломжтой ба энэ нь тухайн хэрэглэгчийн боловсролын түвшин, мэдээллийн аюулгүй байдлын ойлголт, өмнөх туршлагаас шууд хамаарах нь бий.

Б.Ижил төстэй судалгаануудын тойм

Итгэлцлийн онолын хүрээний ойлголтуудыг ялган ойлгох нь цахим засаглалын судалгаанд зайлшгүй чухал

шаардлагатай юм. Цахим засаглалын чиглэлээрх судалгаануудад төрд итгэх итгэл (Trust in Government - TG), системд итгэх итгэл (Trust in System - TPS) гэсэн хоёр ойлголтыг ялган авч үзсэн байдаг бөгөөд төрд итгэх итгэл нь системд итгэх итгэлд дам нөлөө үзүүлдэг болохыг онцолжээ [4]. Эдгээр судалгаануудад иргэн төрийг шударга, ил тод, хариуцлагатай гэж үздэг бол тухайн төрийн харьяа цахим системийг найдвартай гэж хүлээн зөвшөөрөх хандлага нэмэгддэг байна. “Төрд итгэх итгэл” (TG) нь улс төрийн орчин, хууль эрх зүйн тогтолцоо, засаглалын ил тод байдалтай холбоотой итгэлцлийг илэрхийлдэг бол “Системд итгэх итгэл” (TPS) нь системийн ажиллагаа, хэрэглээний хялбар байдал, мэдээллийн аюулгүй байдлын талаарх үнэлгээнд суурилдаг. Мичиганы Их сургуулийн судлаачид “Технологид итгэх итгэл” (Trust in Technology - TT)-ийг хэрэглэгчийн анхны итгэлцлийн онолыг дэвшүүлж, систем болон технологид хандах итгэлцлийг хэрэглээний зан төлөвт хэрхэн нөлөөлдгийг тогтоосон байна [7].

Хэрэглэгчийн хувийн мэдээлэл (Privacy Concerns - PC)-ийн талаарх санаа зовнил нь цахим засаглалын хэрэглээг хязгаарлагч сэтгэл зүйн хүчин зүйл болохыг олон судалгаа харуулж байна. APCO загвар нь цахим үйлчилгээний агуулгад хэрэглэгчийн мэдээллийг хэрхэн хадгалах, боловсруулах, хамгаалах талаар тодорхой дүрэм, стандарт хэрэгтэй болохыг онцолж байна. Монголын нөхцөлд мэдээллийн нууцлал, хамгаалалт, төрийн мэдээллийн ил тод байдалтай холбоотой асуудлууд идэвхтэй яригдаж байгаа нь энэхүү ойлголтыг судалгаанд тусгах шаардлагатайг нотолж байна.

Мэдээллийн эрсдэл (Perceived Risk - PR)-ийн талаарх хэрэглэгчийн ойлголт нь системд итгэх итгэлд шууд сөргөөр нөлөөлдөг. Энэ төрлийн олон улсын судалгаануудаас үзэхэд PR-ийг мэдээлэл алдагдах, систем гацах, гуравдагч этгээд мэдээлэл ашиглах эрсдэлүүдтэй холбоотойгоор хэрэглэгчийн дотоод сэтгэл зүйн хандлагыг илэрхийлдэг үзүүлэлт гэж ихэвчлэн үзсэн байна [7]. Судалгаагаар хэрэв хэрэглэгч тухайн системийг аюулгүй бус, мэдээлэл алдах магадлалтай гэж үзвэл түүнийг ашиглах хандлага эрс буурдаг. Тиймээс PR нь системийн хэрэглээний зан төлөвийг хэмжихэд хүчтэй хамааралтай хувьсагч болдог. Энэхүү ойлголт нь хэрэглэгчдийн өмнөх туршлага, мэдээллийн мэдлэг, технологийн соёл зэрэг хүчин зүйлтэй холбоотой байдаг. Иймээс төрийн цахим системийн дизайнд мэдээллийн хамгаалалтын баталгаа, тодорхой хяналтын механизм тусгах нь иргэдийн хэрэглээний итгэлцлийг нэмэгдүүлэх стратеги болно.

Дээр дурдсан онолын үзэл баримтлалуудыг нэгдсэн байдлаар ашиглаж нийтийн өргөдлийн цахим системийн хэрэглээнд нөлөөлж буй зан төлөвийн хүчин зүйлсийг бүх талаас нь судлах боломжтой юм. UTAUT загвар бол хэрэглэгчийн зан төлөвийг тайлбарлахад үр ашигтай, сайн баталгаажсан загвар бөгөөд гүйцэтгэлийн хүлээлт (PE), хүч чармайлтын хүлээлт (EE), дэмжих нөхцөл (FC), нийгмийн нөлөө (SI) зэрэг хувьсагчидтай. Харин итгэлцлийн онолын хувьсагчид болон хувийн мэдээллийн талаарх санаа зовнил, мэдээллийн эрсдэл зэрэг хувьсагчид нь хэрэглээний сэтгэл зүйг гүнзгий төвшинд тайлбарлаж, шийдвэр гаргалтад нөлөөлдөг. И.ймд энэхүү судалгаанд дээрх онолуудыг нэгтгэж Монгол Улсын парламентын нийтийн өргөдлийн цахим системийн хэрэглээг оновчтой дүгнэх, практик санал дэвшүүлэхэд чиглэсэн онол-практикийн холимог арга зүйг бүрдүүлж өгч байна.

III. Үнэлгээний загвар

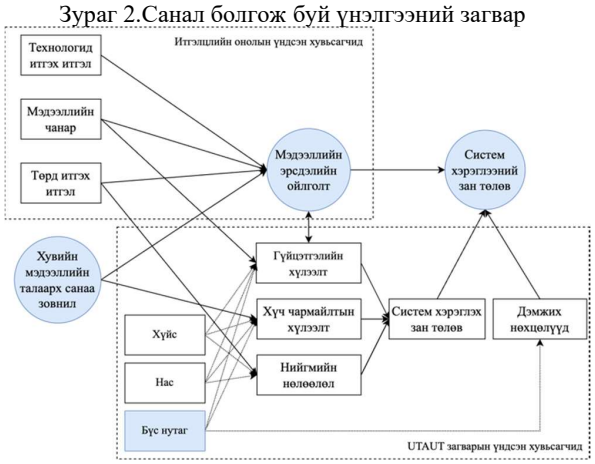
Энэхүү судалгаа нь тоон шинжилгээнд суурилсан, тайлбарлах шинж чанартай бөгөөд бүтцийн тэгшитгэлийн загварчлалыг ашиглан хувьсагч хоорондын хамаарлыг судалгаагаар тодорхойлсон. Бүтцийн тэгшитгэлийн загварчлал нь тоон судалгааны арга бөгөөд чанарын аргуудтай нэгтгэж болдог загварчлал юм. Уг загварчлалыг хувьсах хэмжигдэхүүний хоорондох учир шалтгааны холбоог харуулахад хэрэглэгддэг.

A. Судалгааны санал болгож буй үнэлгээний загвар

Бид судалгааны загвартаа UTAUT, итгэлцлийн онол, хувийн мэдээллийн хамгаалалт, мэдээллийн эрсдэлийн ойлголт зэрэг хэд хэдэн онолыг нэгтгэж, тухайн хувьсагчдын иргэдийн цахим систем ашиглах хандлагад үзүүлэх нөлөөллийг хэмжихэд чиглүүлсэн. Энэхүү судалгааны санал болгож буй үнэлгээний загвар нь бодит нөхцөлд судалгаа хийхээс өмнөх, тестлэх зорилготой онолын хүрээнд боловсруулагдсан болно. Загварын үндсэн хувьсагчийг дараах байдлаар тодорхойлов. Үүнд:

- Гүйцэтгэлийн хүлээлт (Performance Expectancy, PE);
- Хялбар хэрэглээний хүлээлт (Effort Expectancy, EE);
- Нийгмийн нөлөөлөл (Social Influence, SI);
- Дэмжих нөхцөлүүд (Facilitating Conditions, FC);
- Төрд итгэх итгэл (Trust in Government, TG);
- Технологид итгэх итгэл (Trust in Technology, TT);
- Мэдээллийн чанар (Information Quality, IQ);
- Хувийн мэдээллийн талаарх санаа зовнил (Privacy Concerns, PC);
- Мэдээллийн эрсдэлийн ойлголт (Perceived Risk, PR);
- Хэрэглэх хандлага (Behavioral Intention, BI).

Зураг 2.Санал болгож буй үнэлгээний загвар



Эх сурвалж: Судлаачийн боловсруулснаар

Судалгааны загвар нь хэрэглэгчийн технологид хандах хандлага, итгэлцэл, болгоомжлол, орчны дэмжлэг зэрэг олон хүчин зүйлийн шалтгаант хамаарлыг тусгасан бөгөөд Зураг2-т харуулсан.

Загварыг боловсруулахдаа хувьсагчдыг тайлбарлах (экзоген), тайлбарлагдах (эндоген), зуучлагч (mediator), тохируулагч (moderator) гэж ангилсан. Тайлбарлах хувьсагчид нь дараах байв: гүйцэтгэлийн хүлээлт (PE), хялбар хэрэглээний хүлээлт (EE), нийгмийн нөлөөлөл (SI), дэмжих нөхцөлүүд (FC), төрд итгэх итгэл (TG), технологид итгэх итгэл (TT), мэдээллийн чанар (IQ), хувийн мэдээллийн талаарх санаа зовнил (PC). Эдгээр хувьсагч нь мэдээллийн эрсдэлийн ойлголт (PR) болон хүлээн авах хандлагад (BI) шууд ба дам байдлаар нөлөөлнө гэж үзсэн. Тайлбарлагдах хувьсагчид нь системийн хэрэглээний зан төлөв (USE), хүлээн авах хандлага (BI) байсан бол, мэдээллийн эрсдэлийн ойлголт (PR) нь зуучлагч хувьсагчаар ажилласан. Харин нас, хүйс, оршин суугаа бүс нутгийн хувьсагчдыг тохируулагч хувьсагч байдлаар тусган, хэрэглэгчдийн сегментэд хэрхэн ялгаатай нөлөөлж буйг шинжлэв. Төрийн цахим үйлчилгээнд иргэний хувийн мэдээллийг хамгаалах загвар нь дараах нуугдмал хувьсагч (latent variables)-ыг агуулна.

Б.Боловсруулсан загварын судалгаа

Боловсруулсан загварын хувьсагч тус бүрийг 3–5 хэмжүүртэй Лайкертийн хэмжээс (1–5) ашиглан хэмжсэн. Судалгааны асуулгыг урьдчилан боловсруулж, Google forms дамжуулан авч, эхний байдлаар туршин шалгах зорилгоор 30 хэрэглэгчийн жишээн дээр туршсан. Туршилтаар асуултын ойлгомжтой байдал, логик дараалал, хариулах хугацааг хэмжсэн.

Асуулгыг дараах 3 хэсэгт хуваасан. Үүнд:

1. Хүн ам зүйн мэдээлэл: Нас, хүйс, боловсрол, оршин суугаа газар;
2. Онолын хувьсагчийг хэмжих асуулт;
3. Санал, дүгнэлтийн нээлттэй хариулт (optional).

Судалгааны ажлын хүрээнд 500 орчим түүвэр цуглуулж, чанарын хяналт хийсний дараа 300 итгэмжлэгдсэн хариултыг сонгосон [1][2]. Энэ нь SEM шинжилгээ хийхэд хангалттай хэмжээ бөгөөд олон хувьсагчийн өгөгдлийн шинжилгээний (Hair et al., 2010) зөвлөмжийн дагуу “latent variable” тутамд хамгийн багадаа 10 оролцогч шаардлагатай гэж тооцов. Уг өгөгдлийг SPSS болон SmartPLS 4.0 программ ашиглан дараах байдлаар боловсруулсан. Үүнд:

- SPSS: Хүн ам зүйн мэдээлэл, Cronbach’s Alpha, анхан шатны статистик шинжилгээ;
- SmartPLS 4.0: Баталгаажуулах хүчин зүйл (CFA), AVE, CR, path coefficients, R^2, Q^2, модельд тохирох байдал (model fit) зэрэг үзүүлэлтийг ашиглан SEM шинжилгээг гүйцэтгэсэн.

Асуулгыг дотоод найдвартай байдал (Cronbach's Alpha ≥ 0.70), нийлмэл найдвартай байдал (Composite Reliability ≥ 0.70), дундаж хувийн вариант (AVE ≥ 0.50) үзүүлэлтийн дагуу үнэлсэн.

Судалгааны загварын далд хувьсагч (latent variable)-ийн бүх үзүүлэлт нь босго үзүүлэлтийг хангаж, хүчин зүйл тус бүрийн ачаалал 0.70-аас их гарсан.

IV.СУДАЛГААНЫ ҮР ДҮН

А.Судалгааны түүвэр болон хүн ам зүйн мэдээлэл

Судалгааны асуулгад 30 орчим хэмжүүртэй хувьсагч байгаа тул хамгийн багадаа 300 оролцогч шаардлагатай гэж тооцож, 500 орчим түүвэр цуглуулж, чанарын хяналт хийсний дараа 300 итгэмжлэгдсэн хариултыг сонгосон. Судалгаанд хамрагдсан хэрэглэгчдийн насны хувьд 18–25 насныхан 32.7%, 26–35 насныхан 30.7%, 36–45 насныхан 20.7%, 46–55 насныхан 11.3%, 56-аас дээш насныхан 4.6% байв. Хүйсийн хувьд 50.7% нь эмэгтэй, 49.0% нь эрэгтэй, 0.3% нь тодорхойгүй хэмээн тэмдэглэжээ. Оршин суугаа газрын хувьд 71% нь Улаанбаатар хотын, 29% нь хөдөө, орон нутгийн харьяалалтай байв.

Үүнээс үзэхэд уг түүвэрлэлтэд Монгол Улсын хүн ам зүйн төлөөлөл бүрийг тодорхой хэмжээнд тусгасан байна. Иргэдийн боловсролын түвшин, технологи ашиглах дадал, төрийн цахим үйлчилгээний хэрэглээний байдал зэргийг асуулгаар давхар бүртгэсэн болно.

Б.Хувьсагчдын найдвартай байдал, хүчин зүйлсийн шинжилгээний үр дүн

SPSS болон SmartPLS 4.0 программын тусламжтайгаар хувьсагчдын дотоод найдвартай байдал (reliability) болон баталгаажуулах хүчин зүйлийн шинжилгээ (CFA)-г хийлээ.

Хүснэгт 1. Хувьсагчдын дотоод найдвартай байдал болон баталгаажуулах хүчин зүйлийн шинжилгээ

Хувьсагчийн нэр	Cronbach’s Alpha	Composite Reliability (CR)	AVE
PE	0.83	0.88	0.61
EE	0.79	0.85	0.59
SI	0.76	0.83	0.57
FC	0.81	0.86	0.62
TG	0.85	0.89	0.64
TT	0.87	0.91	0.66
IQ	0.83	0.88	0.63
PC	0.80	0.86	0.58
PR	0.77	0.84	0.55
BI	0.89	0.92	0.68

Эх сурвалж: Судлаачийн боловсруулснаар

Cronbach’s Alpha буюу найдвартай байдлын хувьсагчийн үзүүлэлтүүд бүгд 0.70-аас их гарсан. Хүчин төгөлдөр байдлын бүх үзүүлэлт (AVE) 0.50-аас их гарлаа.

В..Бүтцийн тэгшитгэлийн загварчлалын үр дүн

SmartPLS 4.0 ашиглан бүтцийн тэгшитгэлийн загварчлалыг боловсруулсан. Модельд нийт 10 далд хувьсагч (latent variable), 38 хэмжигдэхүйц хувьсагч (measurable variable) багтсан. Загварчлалын стандарт хэмжээг дараах үзүүлэлтээр тооцож, баталлаа. Үүнд:

- SRMR = 0.057 (≤ 0.08);

- NFI = 0.918 (≥ 0.90);
- R^2 (BI) = 0.684 Behavioral Intention хувьсагчийн тайлбарлагдах чадвар өндөр гарлаа.

Г.Таамаглалын шалгалт

Хүснэгт 2. Таамаглалын шалгалт

Таамаглал болон хувьсагчийн хамаарал	Коэффициент (β)	t-утга	p-утга	f^2 нөлөөний хэмжээ	Нөлөөний чиглэл	Таамаглалын статус
H1 TG → TPS	0.341	6.12	< 0.001	0.217 (дунд)	Эерэг	Батлагдсан
H2 TPS → BI	0.406	7.43	< 0.001	0.264 (дунд)	Эерэг	Батлагдсан
H3 PE → BI	0.191	3.85	< 0.001	-	Эерэг	Батлагдсан
H4 EE → BI	0.155	2.97	0.003	-	Эерэг	Батлагдсан
H5 PC → BI	– 0.264	5.68	< 0.001	0.141 (бага/дунд)	Сөрөг	Батлагдсан
H6 PR → TPS	– 0.188	4.33	< 0.001	-	Сөрөг	Батлагдсан
H7 FC → BI	0.142	2.51	0.012	-	Эерэг	Батлагдсан

Эх сурвалж: Судлаачийн боловсруулснаар

Таамаглалтын шалгалтын үр дүнд итгэлцэл, гүйцэтгэлийн хүлээлт, дэмжих нөхцөл нь хэрэглээний хандлагад эерэг, харин эрсдэл ба санаа зовнил нь сөрөг нөлөөтэй болох нь батлагдлаа.

H1: Төрд итгэх итгэл нийтийн өргөдлийн цахим системд итгэх итгэлд эерэг нөлөө үзүүлдэг. Таамаглал батлагдсан. Иргэдийн төрд итгэх итгэл нь нийтийн өргөдлийн цахим системд итгэх хандлагад эерэг, хүчтэй нөлөөтэй байна. Энэ нь Bélanger & Carter (2008) болон AlHadid et al. (2022) нарын судалгааны үр дүнтэй нийцэж байгаа бөгөөд төрийн итгэлцэл нь технологийн итгэлцэлд дам нөлөө үзүүлдэг болохыг харуулж байна.

H2: Нийтийн өргөдлийн цахим системд итгэх итгэл нийтийн өргөдлийн цахим системийг ашиглах зан төлөвт эерэг нөлөө үзүүлдэг. Таамаглал батлагдсан. Нийтийн өргөдлийн цахим системд итгэх итгэл нэмэгдэх тусам системийг хэрэглэх хандлага мэдэгдэхүйц нэмэгддэг байна. Энэ нь итгэлцлийн онолын гол зарчмуудыг дэмжиж байгаа бөгөөд иргэдийн итгэл бол системийн хэрэглээний үндсэн хөдөлгүүр гэдгийг нотолж байна.

H3: Гүйцэтгэлийн хүлээлт нийтийн өргөдлийн цахим системийг ашиглах зан төлөвт эерэг нөлөө үзүүлдэг. Таамаглал батлагдсан. Цахим систем нь үр ашигтай, үр дүнтэй байх хүлээлт нь нийтийн өргөдлийн цахим системийг ашиглах зан төлөвт эерэгээр нөлөөлж байна. Энэ нь UTAUT загварын үндсэн бүрэлдэхүүн болох performance expectancy-ийн нөлөөг Монголын контекстод баталж байна.

H4: Хүч чармайлтын хүлээлт нийтийн өргөдлийн цахим системийг ашиглах зан төлөвт эерэг нөлөө үзүүлдэг. Таамаглал батлагдсан. Нийтийн өргөдлийн цахим системийг хэрэглэхэд хялбар гэж үзэж буй иргэдийн хувьд хэрэглээний хандлага өндөр байна. Энэ нь технологийн хүлээн авалтын онолын чухал хүчин зүйлсийн нэг болох effort expectancy хүчинтэйг баталж байна.

H5: Хувийн мэдээллийн талаарх санаа зовнил нийтийн өргөдлийн цахим системийг ашиглах зан төлөвт сөрөг нөлөө үзүүлдэг. Таамаглал батлагдсан. Иргэдийн хувийн мэдээлэл алдагдах, буруугаар ашиглагдах болзошгүй байдалд санаа зовох нь системийг хэрэглэх сонирхлыг мэдэгдэхүйц бууруулж байна. Иймд энэхүү судалгааны үр дүн APCO загварын судалгааны үр дүнтэй нийцэж байна.

H6: Мэдээллийн эрсдэлийн ойлголт системд итгэх итгэлд сөргөөр нөлөөлдөг. Таамаглал батлагдсан. Нийтийн өргөдлийн цахим системийг мэдээллийн аюулгүй бус гэж үзэх нь системд итгэх итгэлийг бууруулж байна. Энэ нь PR хувьсагчийн итгэлцэлд үзүүлэх шууд сөрөг нөлөөг онцлон харуулж байна.

H7: FC нийтийн өргөдлийн цахим системийг ашиглах зан төлөвт эерэг нөлөө үзүүлдэг. Таамаглал батлагдсан. Системийг хэрэглэхэд шаардлагатай дэд бүтэц, дэмжлэг хангагдсан гэж үздэг иргэд нийтийн өргөдлийн цахим системийг ашиглах өндөр хандлагатай байна. Энэ нь UTAUT загварын facilitating conditions хувьсагчийн үр нөлөөг Монголын нөхцөлд бататгаж байна гэж үзнэ.

Бүх таамаглал статистикийн хувьд батлагдсан бөгөөд загварын таамагласан хамаарлууд нь онолын үндэслэлтэй, эмпирик баталгаатай байгааг харуулж байна.

Д.Хүн ам зүйн хувьсагчдын зохицуулах нөлөө

Энэхүү судалгаанд насны бүлэг, хүйс, бүс нутаг (Улаанбаатар/хөдөө) гэсэн гурван үндсэн хүн ам зүйн хувьсагчийг сонгож, судалгааны гол хамаарлуудад үзүүлэх зохицуулах нөлөөг шинжилсэн.

Насны бүлгийн зохицуулах нөлөөг шалгасан байдал:

- H2 хамаарал нь 18–25 насныханд илүү хүчтэй гарсан ($\beta = 0.487$, $p < 0.01$).
- H3 хамаарал нь 36–45 насныханд хамгийн өндөр байсан ($\beta = 0.274$, $p < 0.05$).

Залуучуудын хувьд итгэлцэлд суурилсан шийдвэр гаргалт давамгайлж байхад, нас тогтсон хэрэглэгчдийн хувьд ашиг тусын талаарх хүлээлт (performance expectancy) нь илүү нөлөөтэй байна. Энэ нь насны онцлогт тохируулсан цахим систем нэвтрүүлэлтийн ялгамжтай стратеги хэрэгжүүлэх шаардлагыг харуулж байна. Хүйсийн зохицуулах нөлөөг дараах хамаарлаар шалгахад:

- H5 хамаарал эмэгтэйчүүдийн дунд илүү хүчтэй ($\beta = -0.311$, $p < 0.01$).
- H4 хамаарал эрэгтэй хэрэглэгчдэд давамгайлсан ($\beta = 0.176$, $p < 0.05$).

Эмэгтэй оролцогчид хувийн мэдээллийн хамгаалалт дээр илүү мэдрэмтгий байсныг илтгэж байна. Харин эрэгтэй оролцогчдын хувьд системийг ашиглахад хялбар эсэх нь түлхүү нөлөөлж байна. Энэ нь жендэрийн хандлага, хүлээлтийн ялгааг бодлого

тодорхойлолтод тусгах шаардлагатайг харуулна. Бүс нутгийн зохицуулах нөлөөг дараах хамаарлаар шалгахад:

- H7 хамаарал хөдөө оршин суугчдын дунд илүү хүчтэй ($\beta = 0.223$, $p < 0.01$).
- H6 хамаарал хот суурин хэрэглэгчдэд илүү мэдрэгдсэн ($\beta = -0.215$, $p < 0.01$).

Хөдөө орон нутгийн хэрэглэгчдийн хувьд системийг дэмжих орчин, техникийн нөхцөл шийдвэр гаргалтад ихээхэн нөлөөлж байгаа бол, Улаанбаатар хотын иргэдийн хувьд мэдээллийн аюулгүй байдлын эрсдэл нь итгэлцлийг бууруулах гол хүчин зүйл болж байна. Энэ нь дэд бүтэц болон мэдээллийн аюулгүй байдлыг бүс нутгаар ялган оновчлох бодлогын үндэслэл болно.

Тохиолдолд үндэслэн хувьсагчдын хоорондын хамаарлыг бүхэлд нь авч үзвэл статистикийн хувьд хүн ам зүйн хувьсагчуудын зохицуулах нөлөөний илэрсэн байдлыг Хүснэгт 18-т нэгтгэн харууллаа.

Хүснэгт 3. Хувьсагчдын хоорондын хамаарал

Таамаглал	Зохицуулагч хувьсагч	Зохицуулах нөлөө илэрсэн бүлэг
H2	Нас	18–25 насныхан
H3	Нас	36–45 насныхан
H4	Хүйс	Эмэгтэй
H5	Хүйс	Эрэгтэй
H7	Бүс нутаг	Орон нутаг
H6	Бүс нутаг	Улаанбаатар

Эх сурвалж: Судлаачийн боловсруулснаар

Эндээс үзвэл нас, хүйс, байршил зэрэг хувьсагч нь EE → BI, PR → TT харилцаанд зохицуулах (moderating) нөлөө үзүүлж байна. Тухайлбал 26-35 насны хэрэглэгчдийн хувьд гүйцэтгэлийн хүлээлт илүү хүчтэй нөлөөтэй байсан бол хөдөө орон нутгийн хэрэглэгчдийн хувьд мэдээллийн эрсдэлийн ойлголт өндөр үзүүлэлттэй гарсан.

V. Дүгнэлт ба зөвлөмж

Энэхүү судалгаа нь Монгол Улсын парламентын нийтийн өргөдлийн системийн хэрэглээнд нөлөөлөх хүчин зүйлсийг UTAUT, итгэлцлийн онол, хувийн мэдээллийн хамгаалалт болон мэдээллийн эрсдэлийн ойлголтыг нэгтгэсэн концептуал загвараар тодорхойлон судалсан. Судалгаанд “Mock data” ашиглаж, бүтцийн тэгшитгэлийн загварчлал (SEM)-ын аргаар таамаглалыг шинжилж, үндсэн үр дүнг гаргасан. Судалгаан үр дүн нь төрийн байгууллагуудын цахим системийн хөгжүүлэлт, мэдээллийн аюулгүй байдлын бодлогыг сайжруулах, иргэдийн оролцоог нэмэгдүүлэхэд бодитой хувь нэмэр оруулах суурь мэдээлэл болох юм.

А. Дүгнэлт

- Нийтийн өргөдлийн цахим системийг хэрэглэх хандлагад итгэлцэл, ашиг тусын талаарх хүлээлт, хувийн мэдээллийн талаарх болгоомжлол голлон нөлөөлж байна.
- Төр болон технологид итгэх итгэл нь нийтийн өргөдлийн цахим системд итгэх итгэлд эерэг хамааралтай бөгөөд энэ нь хэрэглээг нэмэгдүүлдэг.
- Хувийн мэдээлэлтэй холбоотой эргэлзээ, мэдээллийн эрсдэл нь системд итгэх итгэлийг бууруулж, улмаар хэрэглээг сааруулдаг.
- UTAUT загварын гүйцэтгэлийн хүлээлт, хүчин чармайлтын хүлээлт, дэмжих нөхцөл зэрэг хүчин зүйлс мөн хандлагад эерэг нөлөөтэй байна.
- Хүн ам зүйн хувьсагчид (нас, хүйс, бүс нутаг) нь хэрэглээний хандлагад зохицуулах нөлөө үзүүлж буйг тогтоолоо.

Эдгээр үр дүн нь төрийн цахим системийн хөгжүүлэлтэд хэрэглэгчийн сэтгэл зүй, итгэлцэл, мэдээллийн аюулгүй байдлыг зэрэг авч үзэх шаардлагатайг баталгаажуулж байна.

Олон улсын ижил төстэй судалгаатай харьцуулан авч үзвэл дараах байдлаар онцлог судалгаа болсон. Үүнд:

- Судалгаагаар санал болгосон үнэлгээний загвар нь UTAUT, итгэлцлийн онол, мэдээллийн болгоомжлолын загварыг интеграцчилсан шинэ загвар юм.
- Үнэлгээнд зохицуулах хувьсагчийг оруулан сегментчлэлд суурилсан шинжилсэн нь ижил төстэй олон улсын судалгаануудаас давуу болно.
- Монгол Улсын цахим засаглал, ялангуяа нийтийн өргөдлийн цахим системийн хэрэглээг онолын түвшинд анх удаа иж бүрэн судалсан анхны эмпирик судалгаа болж байна.

Б.Зөвлөмж

1. Итгэлцлийг нэмэгдүүлэх бодлого боловсруулах, хэрэгжүүлэх: Төрийн цахим системд хадгалагдаж буй хувийн мэдээллийн хамгаалалтын бодлого, авч буй арга хэмжээ, холбогдох албан тушаалтны хариуцлагыг ил тод мэдээлж, иргэдийн хүлээлтийг зохицуулах;
2. Өгөгдлийн менежмент ба аюулгүй байдлын холбогдох арга хэмжээг авах: Монгол Улсын парламентын нийтийн системд аюулгүй байдлын шалгалтыг тогтмол хийж, эрсдэл үүсэх нөхцөлийг урьдчилан тооцоолж, хэрэгжүүлэх арга хэмжээний төлөвлөгөөг боловсруулахсзамаар эрсдэлийг багасгах;
3. Иргэдийн оролцоог дэмжих хөшүүргийг бий болгох: Иргэд нийтийн өргөдөл илгээх, оролцоог нэмэгдүүлэх урамшууллын бодлого эсвэл дэмжих механизмыг судлах;
4. MXXT-ийн сургалт, сурталчилгааг тогтмол хийх: Иргэдэд хувийн мэдээлэл, нууцлалын ойлголтыг нэмэгдүүлэх мэдээллийн боловсролын хөтөлбөр боловсруулж, сургалт, сурталчилгааг тогтмол хийх. Тухайлбал, орон нутгийн иргэд, ахмад настнууд зэрэг зорилтот бүлгүүдэд зориулсан сургалт, хэрэглээний зааварчилгааг боловсруулж, дижитал тэгш бус байдлыг бууруулах бодлого хэрэгжүүлэх;
5. Системийн хөгжүүлэлт, сайжруулалтыг тогтмол хийх: Хэрэглэгчийн гомдол, саналд тулгуурлан Монгол Улсын парламентын нийтийн өргөдлийн системийг тогтмол сайжруулах тогтолцоог бүрдүүлэх.

НОМ ЗҮЙ

[1] Монгол Улсын Их Хурлын Тамгын газар, “D-Parliament цахим системийн хэрэглэгчдийн оролцооны статистик мэдээлэл”, 2024, https://d.parliament.mn/

[2] Монгол Улсын Их Хурлын Тамгын газар, “Нийтийн өргөдөл, гомдлын цахим системийн хэрэглэгчдийн оролцооны статистик мэдээлэл”, 2024, https://petition.parliament.mn/

[3] AlHadid and Issam et al., “Predictors for E-Government Adoption of SANAD App Services Integrating UTAUT, TPB, TAM, Trust, and Perceived Risk”, International Journal of Environmental Research and Public Health (IJERPH), pp.1-26, July 2022

[4] F.Bélanger and L.Carter, “Trust and Risk in E-Government Adoption”, Journal of Strategic Information Systems, Vol.17, pp.165-176, 2008

[5] M.S.Featherman and P.A.Pavlou, “Predicting e-Services Adoption: A Perceived Risk Facets Perspective”, International Journal of Human-Computer Studies, Vol.59, pp.451-474, 2003

[6] J.A.Smith, “Evaluating the Contribution of Interpretive Phenomenological Analysis”, Health Psychology Review, Vol.5, pp.9-27, 2011

[7] T. Zhou, “Examining mobile banking user adoption from the perspectives of trust and flow experience,” Inf Technol Manag, vol. 13, no. 1, pp. 27–37, Mar. 2012, doi: 10.1007/s10799-011-0111-8

[8] Venkatesh, Thong, and Xu, “Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology”, MIS Quarterly, Vol. 36(1), pp.157-178, 2012, doi: 10.2307/41410412

[9] Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, “User Acceptance of Information Technology: Toward a Unified View”, MIS Quarterly, Vol.27(3), pp. 425–478, 2003, doi: 10.2307/30036540

[10] Weerakkody, R. El-Haddadeh, and S. Al-Shafi, “Exploring the complexities of e-government implementation and diffusion in a developing country: Some lessons from the State of Qatar”, Journal of Enterprise Information Management, vol. 24, no. 2, pp. 172–196, Feb. 2011, doi: 10.1108/174103911111062

Эрүүл Мэндийн Мэдээлэл зүй сургалтын хөтөлбөрийн шинэчлэлт

Пүрэвдолгор Лувсанцэрэн
Физик, мэдээлэл зүйн тэнхим
АШУҮИС, БАСургууль
Монгол Улс
purevdolgor@mnums.edu.mn

Одгэрэл Болдбаатар
Ерөнхий эрдмийн тэнхим
АШУҮИС, ССургууль
Монгол Улс
odgerel.b@mnums.edu.mn

Ахыт Тилуебай
Физик, мэдээлэл зүйн тэнхим
АШУҮИС, БАСургууль
Монгол улс
akhit@mnums.edu.mn

Жавзмаа Цэнд
Физик мэдээлэлзүйн тэнхим,
АШУҮИС, БАСургууль
Монгол улс, Улаанбаатар хот
javzmaa.ts@mnums.edu.mn

Баасандорж Чилхаасүрэн
Физик мэдээлэлзүйн тэнхим
АШУҮИС, БАСургууль
Монгол улс, Улаанбаатар хот
baasandorj@mnums.edu.mn

Нямдаваа Уугандаваа
Физик мэдээлэлзүйн тэнхим
БАС, АШУҮИС
Монгол улс
nyamdavaa.u@mnums.edu.mn

Ажнай Лувсан
Физик, мэдээлэл зүйн тэнхим
АШУҮИС, БАСургууль
Монгол Улс
ajnai@mnums.edu.mn

Оюунцэцэг Сандаг
Харилцааны ур чадварын тэнхим
АШУҮИС, БАСургууль
Монгол Улс
oyuntsetseg.sa@mnums.edu.mn

Энх-Үрэл Энхжаргал
Физик, мэдээлэл зүйн тэнхим
АШУҮИС, БАСургууль
Монгол Улс
enkhurel@mnums.edu.mn

***Хураангуй*—Эрүүл мэндийн мэдээллийн технологийн салбар нь анагаах ухаан болон эрүүл мэндийн их өгөгдлийг шинжлэх, тандах, загварчлах, мэдээллийн сан үүсгэхийн зэрэгцээ, мэдээллийн аюулгүй байдлыг хангаж, хууль эрх зүйн хүрээнд эрүүл мэндийн үйлчилгээ эрхлэгч, үйлчлүүлэгчдийн хооронд мэдээлэл солилцох цогц үйл ажиллагааг явуулдаг.**

Монгол улсын Эрүүл Мэндийн салбарт ажиллах Эрүүл мэндийн Мэдээлэл зүйн мэргэжилтнийг АШУҮИС-ийн Био-Анагаахын сургууль бэлтгэж байна. Төгсөгчдийн хөдөлмөрийн зах зээлд эзлэх байр суурийн судалгаагаар 85% мэргэжлээрээ ажиллаж байгаа бөгөөд эрүүл мэндийн салбарт Эрүүл мэндийн мэдээлэл зүйн мэргэжилтний хэрэгцээ шаардлага байна.

Энэхүү мэргэжилтэн бэлтгэх хөтөлбөр нь 2008, 2017, 2020 онуудад хөтөлбөр шинэчлэн боловсруулагдаж, 2018, 2020 онд БМИҮЗ-өөр хоёр удаа магадлан итгэмжлэгдсэн болно.

Боловсрол, Шинжлэх Ухааны сайдын А147 тоот тушаалын Дээд боловсролын сургалтын хөтөлбөрт тавих нийтлэг шаардлагын дагуу хийгдсэн хөтөлбөрийн шинэчлэлтийг судаллаа. Сургалтын хөтөлбөрийг 2020 оны хөтөлбөртэй харьцуулахад ерөнхий суурь 7.5%, мэргэжлийн суурь 5%, мэргэших хичээлүүдэд 7.5% өөрчлөлт хийжээ. Шинэчилсэн сургалтын хөтөлбөр нь олон улсын ижил төстэй Health Informatics сургалтын хөтөлбөртэй харьцуулан, мэдээллийн технологийн сургалтын хөтөлбөрийн зөвлөмжөөр (IT curriculum guide) боловсруулсан нь АШУҮИС-ийн алсын хараа, эрхэм зорилготой бүрэн нийцсэн эрүүл мэндийн салбарт мэдээллийн технологийн хэрэгцээ, шаардлагыг хангаж, хөгжүүлэх, хууль эрх зүйн хүрээнд эрүүл мэнд болон анагаах ухааны их өгөгдлийн мэдээллийн аюулгүй байдлыг хангах, боловсруулах, зохион байгуулах мэдлэг, чадвартай мэргэжилтэн бэлтгэх хөтөлбөр боллоо.

Түлхүүр үг—Эрүүл мэндийн Мэдээлэл зүй, сургалтын хөтөлбөр, шинэчлэлт

I. Удиртгал

Эрүүл мэндийн тусламж үйлчилгээний мэдээлэл нь өвчтөний эмчилгээ, оношилгооны төлөвлөлт, судалгаа, жор, эрүүл мэндийн эцсийн үр дүн гэх мэт мэдээллүүд нь ирээдүйд үзүүлэх тусламж үйлчилгээг сайжруулахад туслах үнэ цэнэтэй мэдээлэл юм. Энэхүү мэдээлэл нь

- *Өвчтөний тусламж үйлчилгээг сайжруулах:* Өвчтөний илүү их мэдээлэл нь өвчтөнийг илүү сайн ойлгож, тэдний үзүүлж буй тусламж үйлчилгээг сайжруулах боломжийг нэмэгдүүлнэ.
- *Сайжруулсан судалгаа:* Их өгөгдөл нь анагаах ухааны судлаачдад их хэмжээний өгөгдөл, мэдээлэл цуглуулах нь урьд өмнө байгаагүй өргөн боломжуудыг нээснээр хүний амь насыг аврах анагаах ухааны чухал нээлтүүдийг хийх боломжуудыг нэмэгдүүлнэ.
- *Ухаалаг эмчилгээний төлөвлөгөө:* Өвчтөний эмчилгээний төлөвлөгөөнд дүн шинжилгээ хийх нь дараагийн өвчтөн, үйлчлүүлэгчдэд илүү сайн эмчилгээний төлөвлөгөөг гаргана.
- *Өвчтөн болон эрүүл мэндийн үйлчилгээ үзүүлэгчдийн эрүүл мэндийн тусламж үйлчилгээний зардлыг бууруулах:* Эрүүл мэндийн тусламж үйлчилгээ нь их зардалтай бөгөөд их өгөгдөл нь зохих эмчилгээний төлөвлөгөөг тодорхойлох, нөөц боломжуудыг ухаалгаар хуваарилах, эрүүл мэндийн эрсдэлт асуудлуудыг үүсэхээс нь өмнө тодорхойлох замаар эрүүл мэндийн тусламж үйлчилгээ авах, үзүүлэх зардлыг бууруулах боломжийг санал болгодог гэх мэт эдгээр бүртгэл болон өгөгдлүүд нь эрүүл мэндийн салбарын төлөвлөлт, шийдвэр гаргалт, эрүүл мэндийн тусламж үйлчилгээний чанарыг сайжруулах нөхцөл бүрддэг байдаг.

Өөрөөр хэлбэл Эрүүл мэндийн салбарт Big Data Analytics (BDA) нэвтрүүлснээр өвчтөнийг эмчлэх болон эрүүл мэндийн менежментэд шинэ технологи ашиглах боломжийг олгоно [1].

Монгол улсын “Алсын хараа-2050” урт хугацааны хөгжлийн бодлого болон олон улсын эрүүл мэндийн холбогдох зорилтуудтай уялдуулан Монгол Улсын

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

цахим эрүүл мэндийн нөхцөл байдлын үнэлгээг НҮБ-ын Хүүхдийн сангийн тусламж, Азийн цахим эрүүл мэндийн мэдээллийн холбооны дэмжлэгтэйгээр Монгол Улсын Засгийн газар, төрийн байгууллагууд, хувийн хэвшлийн байгууллага, мэргэжилтэн, судлаачдын хамтаар хийж гүйцэтгэсэн байна. Энэхүү цахим эрүүл мэндийн нөхцөл байдлын үнэлгээний тайланд Эрүүл мэндийн газруудад мэдээлэл холбооны технологийн мэргэжилтэн хүрэлцээтэй биш, олон орон нутагт ганц мэргэжилтэн байдаг нь хангалтгүй аж. Зарим бүс нутаг мэдээлэл холбооны технологийн нэгжтэй боловч эрэлт хэрэгцээг хангахуйц том хэмжээтэй биш юм. Мэдээлэл холбооны технологийн мэргэжилтний ийм хүрэлцээгүй байдал цахим эрүүл мэндийг үр дүнтэй хэрэгжүүлэхэд томоохон саад тотгор учруулж байна [2].

1990 онд эрүүл мэндийн салбарт анхны компьютерийг нэвтрүүлэх үеэс л төрийн зүгээс цахим шийдлийг эрүүл мэндийн салбарт нэвтрүүлэх чиглэлээр идэвхтэй ажиллаж иржээ. Ийнхүү үндэс суурьиа тавьснаар 1997 онд анхны эрүүл мэндийн мэдээллийн систем болох H-Info-ийн анхны хувилбарыг нэвтрүүлэх зэрэг дараа дараагийн ажлуудыг хэрэгжүүлэх боломж бүрдсэн байна. 2003 он гэхэд “Эрүүл мэндийн мэдээллийн технологич” мэргэжлийн хөтөлбөрийг нээсэн нь Монгол Улс цахим эрүүл мэндийн чиглэлд өндөр ач холбогдол өгч байсныг илтгэж байна [3].

Үнэлгээний үр дүнд мөн эрүүл мэндийн байгууллагууд дахь цахим эрүүл мэндийн үйл ажиллагааг удирдан зохион байгуулах мэргэжилтнийг бэлдэх биоанагаахын эсвэл эрүүл мэндийн мэдээлэл зүй гэх мэт цахим эрүүл мэндийн ур чадварыг хөгжүүлэх сургалтын хөтөлбөрийг нэмэгдүүлэх шаардлагатай байгааг харуулж байна.

Дотоодын их сургуулиудад сургалтын хөтөлбөр нээх төлөвлөгөө байдаг, нэг их сургуульд эрүүл мэндийн мэдээллийн технологич мэргэжлийн хөтөлбөрийг хэрэгжүүлдэг боловч цахим эрүүл мэндийн мэргэжилтнүүдийн өсөн нэмэгдэж байгаа эрэлт хэрэгцээг хангахын тулд эдгээр хөтөлбөрийг хэд хэдэн сургуульд өргөжүүлэх хэрэгтэй [2].

Монгол Улсын газар зүйн байршил, дэд бүтцийн хөгжлийн харилцан жигд биш түвшин, эрүүл мэндийн тогтолцооны төвлөрлийг сааруулсан удирдлага, төвлөрсөн бус архитектурыг сонгон нэвтрүүлсэн цахим эрүүл мэнд зэргийг бодолцоход тус улсын засгийн газар, Эрүүл мэндийн яамны хувьд хамгийн зөв зам бол цахим эрүүл мэндийн стратегийг хөгжүүлэх явдал юм [2] .

Дэлхий нийтийн аж үйлдвэрийн дөрөвдүгээр хувьсгалын нөлөөгөөр эрүүл мэндийн салбарт зүйлсийн интернет, их өгөгдлийн шинжилгээ, блокчэйн, үүлэн тооцоолол, хиймэл оюун ухааны технологиуд нэвтэрч, өвчтөний тусламж үйлчилгээг сайжруулахын тулд дэвшилтэт тоон технологийг эрүүл мэндийн тусламж үйлчилгээний салбарт ашиглах Эрүүл мэндийн тусламж үйлчилгээний 4.0 (Healthcare 4.0) технологийг бий болгосон [4] [5].

Олон улсад Healthcare 4.0 технологийг эрүүл мэндийн бакалаврын сургалтын хөтөлбөрт хэрэгжүүлж байна. Оюутнуудад эрүүл мэндийн мэдээллийг цуглуулж, дамжуулах мэдрэгчтэй зүүдэг төхөөрөмж, алсаас хянах технологи бүхий эмнэлгийн зүйлсийн интернет (Internet of Medical Things (IoMT))[6], өвчтөнд үзүүлэх тусламж үйлчилгээг сайжруулах, өвчний дэгдэлтийг урьдчилан таамаглах их өгөгдлийн шинжилгээний аргууд [5], алсын зайнаас зөвлөгөө өгөх, тусламж үйлчилгээ үзүүлэх, үйлчилгээ багатай газруудад телемедицинийг ашиглах [7], эрүүл мэндийн бүртгэл мэдээллийг хадгалах, хандахад блокчейн болон үүлэнд суурилсан системийг ашиглах тухай [6] зааж байна.

Манай улсын эрүүл мэндийн тусламж үйлчилгээнд эдгээр технологиуд өргөнөөр нэвтэрч, эрүүл мэндийн мэргэжилтнүүд тэдэнтэй ажиллах хэрэгцээ, шаардлага үүсэж байгаа. Иймд АШУҮИС-ийн Био-Анагаахын сургуулийн Физик, Мэдээлэл зүйн тэнхим 2003 оноос Эрүүл мэндийн Мэдээлэл зүй мэргэжилтнийг бэлтгэж байна.

Эрүүл мэндийн Мэдээлэл зүйн мэргэжилтнийг бэлтгэх хэрэгцээ шаардлага, тэдний ур чадварууд нь Healthcare 4.0 технологийн хөгжилд олон талын хувь нэмэр үзүүлэх ач холбогдолтой бөгөөд эрүүл мэндийн үйлчилгээний чанарыг сайжруулж, үр ашгийг дэмжинэ. Жишээ нь эрүүл мэндийн их өгөгдлийг шинжлэх, хандлагыг тодорхойлох, шийдвэр гаргахад туслан, хиймэл оюун ухааныг хөгжүүлэхэд дэмжлэг үзүүлдэг [8]. Healthcare 4.0 технологийн аюулгүй байдлыг сайжруулж, зөвшөөрөлгүй хандалтаас сэргийлэхэд тусална [9]. Технологийн хөгжлийн үе шатанд шинэ болон өмнөх хувилбар бүхий системүүдийг хослуулан ашиглах боломжийг олгодог [10]. Зайнаас өгөгдөл дамжуулах, боловсруулах, хянах ур чадвараа ашиглан алслагдсан бүс нутагт эрүүл мэндийн тусламж үйлчилгээ үзүүлэхэд дэмжлэг үзүүлдэг [8]. IoT төхөөрөмжүүдээс мэдээлэл хүлээн авч, боловсруулан, эрүүл мэндийн тусламж үйлчилгээний чанарыг сайжруулахад тусалдаг [9].

АШУҮИС-ийн Био-Анагаахын суруулийн Эрүүл мэндийн Мэдээлэл зүйн мэргэжилтний цогц чадамжид олон улсын Healthcare 4.0 технологиудын мэдлэг, чадвар, хандлагыг зохих хэмжээгээр нэвтрүүлсэн болно. Энэ нь Эрүүл мэндийн Мэдээлэл зүйн мэргэжилтний ажлын байранд ажиллахад нь тус нэмэр болно.

Манай улсын хэмжээнд өнөөгийн байдлаар улсын II, III шатлалын эмнэлгүүдийн мэдээллийн технологийн 20 гаруй албанд Эрүүл Мэндийн Хөгжлийн Төвийн статистик мэдээллийн 2023 оны тайланд дурдсанаар 35465 орчим компьютер ажиллаж байна [11]. Эрүүл Мэндийн Сайдын “Мэдээллийн Технологийн талаар баримтлах журам батлах тухай” 140 тоот тушаалд 40 компьютер бүхий ажлын байранд мэдээллийн технологийн нэг ажилтан ажиллахаар заасан байдаг [12]. Эрүүл мэндийн хөгжлийн төвийн 2023 оны тайлангаас харахад улсын хэмжээнд одоогийн байдлаар 275 мэдээллийн технологийн мэргэжилтэн ажиллаж байгаа бөгөөд цаашид эрүүл мэндийн салбарт 886 мэргэжилтэн ажиллах хэрэгцээ, шаардлага харагдаж байна [13]. Иймээс Монгол Улсын эрүүл мэндийн салбарт Эрүүл мэндийн Мэдээлэл зүйн мэргэжилтний хэрэгцээ их байгаагаас энэхүү хөтөлбөрийг амжилттай хэрэгжүүлэх, хөтөлбөрийг орчин үеэтэй нийцээхээр шинэчлэн сайжруулсаар байгаа бөгөөд 2025 онд хийсэн шинэчлэлтийг судлахаар зорилоо.

Дээрх зорилгын хүрээнд хөтөлбөрийг үндэсний болон олон улсын стандарттай нийцүүлэн багшлах бүрэлдэхүүн, ажил олгогч зэрэг оролцогч талуудтай зөвшилцсөний үндсэн дээр Эрүүл мэндийн Мэдээлэл зүйн мэргэжилтэн бэлтгэх хөтөлбөрийн зорилгыг тодорхойлоод

- Монгол Улсын Эрүүл мэндийн салбарт мэргэжилтний хэрэгцээ болон ажлын байрны шаардлагад нийцлийг тодорхойлох,
- Төгсөгчид, ажил эрхлэлтийн судалгаа хийх,
- Хөтөлбөрийн бүтэц, хичээлүүдийг сонгох, хичээлийн дараалал, багц цагыг төлөвлөх гэсэн зорилтуудыг дэвшүүллээ.

II. Үндсэн хэсэг

A. *Эрүүл мэндийн Мэдээлэл зүйн хөтөлбөрийн хэрэгжилт*

Эрүүл мэндийн Мэдээлэл зүйн мэргэжилтэн бэлтгэх бакалаврын сургалтын хөтөлбөр нь эрүүл мэндийн салбарт мэдээллийн технологийн хэрэгцээг хангах, хөгжүүлэх, нутагшуулах, нэвтрүүлэх, эмнэлэг биологийн их өгөгдлийг шинжлэх, суурь судалгаанд гүйцэтгэгчээр ажиллах чадвартай мэргэжилтэн, судлаач бэлтгэх хөтөлбөр [14] бөгөөд 4 жилийн хугацаанд, өдрөөр 120 багц цаг судална [15, 16].

Одоо тус хөтөлбөрөөр I-IV дамжааны нийт 40 орчим оюутнууд суралцаж байгаа бөгөөд компьютерт суурилсан электроникийн лаборатори, орчин үеийн өндөр хүчин чадал бүхий сүлжээний орчинд компьютерийн 3 лабораториудад сургалт явагдаж байна. Энэхүү хөтөлбөрийг хэрэгжүүлэхэд АШУҮИС-ийн Био-Анагаахын Сургууль, Нийгмийн Эрүүл Мэндийн сургуулийн 13 тэнхимийн нийт 54 багш оролцож байна. Хөтөлбөрийг хэрэгжүүлэгч багш нар бүгд эрдмийн зэрэг, цолтой, анагаах ухааны боловсролын ба мэргэжлийн сургалтад хамрагдсан, багшлах эрхтэй байна. Нийт багш нарын 48% нь доктор, 52% нь магистрын зэрэгтэй, ажилласан жилээр нь авч үзвэл 5 хүртэл жил ажилласан 19%, 10 хүртэл жил ажилласан 29%, 15 жил ажилласан 21%, 20 хүртэл жил ажилласан 17%, 25 жил хүртэл ажилласан 8%, 30 хүртэл жил ажилласан 4%, 40 хүртэл жил ажилласан 2% байна. Эндээс үзвэл нийт багш нарын 50% дээш хувь 10 дээш ажилласан туршлагатай багш нар байна [14].

Хөтөлбөрт оролцогч багш нарын 6% нь профессор, 6% нь дэд профессор, 42% нь ахлах багш, 25% нь багш, 17% нь дадлагажигч багш, 4% нь мэдээлэл технологийн инженер, албаны дарга албан тушаалд ажиллаж байна. Эндээс үзвэл нийт багш нарын 50% дээш хувь нь ахлах багш, түүнээс дээш албан тушаалд ажиллаж байна.

Хөтөлбөрийг хэрэгжүүлэгч тэнхимийн хамт олон Өгөгдлийн сан, Өгөгдлийн бүтэц C#, Өгөгдлийн шинжилгээ R програмчлал, Анагаахын физик, Математик бодлогын хураамж, Компьютер програмчлал, Компьютерын практикум, Биоинформатик зэрэг мэдээллийн технологийн чиглэлээр олон ном, сурах бичиг, гарын авлага хэвлүүлж оюутан, эмч, мэргэжилтнүүдийн хүртээл болгожээ.

"Анагаахын физикийн лекцүүд" Зохиогчийн эрх №7952 [17], "Хэт авиан оношилгоо-эмчилгээ физик үндэс " Зохиогчийн эрх №12528 [18], Анагаахын физик дадлагын хичээлүүд "[19], "Анагаахын физик" сорилын эмхэтгэл Зохиогчийн эрх №12527 [20], Зохиогчийн эрх №15659 Компьютер програмчлал [21], Компьютерын програм хангамжийн төрлөөр Зохиогчийн эрх "Өсвөр үеийн сэтгэл зүйн үнэлгээ" №14421[22] тоот зохиогчийн эрхийн гэрчилгээ авсан байна.

Мөн хэрэгжүүлэгч тэнхимийн багш нар нийгэмд чиглэсэн үйл ажиллагаа явуулах зорилгоор 2021 онд Монголын Анагаахын физик биоанагаахын инженерчлэлийн нийгэмлэг шинээр байгуулан ажиллаж байна [23].

ШУТИС, МУИС-тай хамтран "Biomedical Applied Sciences and Engineering", "Монголын Мэдээллийн Технологи 2016", Биофизик, биоинформатикийн үндэсний эрдэм шинжилгээний бага хурлыг зохион байгуулсан [24] .

Сүүлийн жилүүдэд дотоодын болон гадаадын сэтгүүлүүдэд багш нар олон өгүүлэл хэвлүүлэн, илтгэл хэлэлцүүлсэн. Үүнд Mongolian Journal of Engineering and Applied Science сэтгүүл, Дээд боловсрол, American Journal of Networks and Communications сэтгүүлд, International Conference on Cloud Computing and Big Data (CloudCom-Asia), BMJ Health & Care Informatics Online, Mongolian Journal of Engineering and Applied Sciences, Advances in Artificial Intelligence and Machine Learning зэрэг 40 гаруй өгүүлэл хэвлүүлжээ.

B. *Хөтөлбөрийг ижил чиглэлийн гадаадын хөтөлбөрүүдтэй харьцуулсан байдал*

Бид хөтөлбөрөө Information technology curiculum guide 2017 зөвлөмжийн дагуу боловсруулсан бөгөөд сургалтын хөтөлбөрт мэдээлэл технологийн 47%, мэдээлэл технологийн бус 45%, математикийн 8% хичээлүүд байгаа нь Information technology curiculum guide 2017-д тусгагдсан зөвлөмжийн дагуу байна [25].

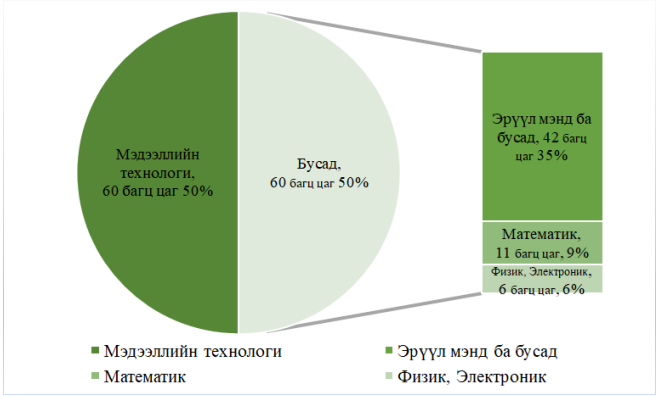


Fig. 1. Хөтөлбөрийн хичээлүүдийн эзлэх хувь

Гадаадын ижил төстэй их сургуулиудын дараах хөтөлбөрүүдтэй харьцуулан судлав. Үүнд:

- Австралийн Сидней их сургуулийн Bachelor of Advanced Computing/Bachelor of Science (Health) [26],
- Америкийн Нэгдсэн Улсын Oregon Tech их сургууль [27],
- Америкийн Нэгдсэн Улсын MontanaTech их сургууль [28]
- Этиопын Гондарийн их сургуулийн Health informatics [29]

Bachelor of Advanced Computing and Bachelor of Science (Health)

Хөтөлбөрөөр суралцах хугацаа 5 жил, 240 credit point байна. Эхний жилд математик, мэдээллийн технологийн программчлал, сүлжээ, өгөгдлийн сан зэрэг хичээлүүдийг судалдгаараа төстэй, сүүлийн 4 жилд тооцооллын болон компьютерийн шинжлэх ухаан, (Computer Science-Major), мэдээллийн систем (Information System-Major), программ хөгжүүлэлт (Software Development-Major), эрүүл мэнд, Health (Major) гэсэн чиглэлээс сонгон суралцана. Хөтөлбөрөөр төгсөгч нь программист, зөвлөх, байгууллагын эрүүл мэнд, тусгай хэрэгцээт, насжилтын менежмент, дэлхийн эрүүл мэндийн бодлогын шинжээч, эмнэлгийн менежмент, мэдээллийн үйлчилгээний менежмент, сэтгэцийн эрүүл мэнд ба аюулгүй байдал, программ хангамжийн инженер, веб хөгжүүлэлт ба менежмент, эмнэлгийн программ хангамж хөгжүүлэгч зэрэг ажлын байранд ажилладаг.

Америкийн Нэгдсэн Улсын Oregon Tech их сургууль

Хөтөлбөрөөр суралцах хугацаа 4 жил, 181 credit hours, I жилд ерөнхий суурь, мэргэжлийн суурь хичээл 45, II жилд мэргэжлийн суурь хичээл 46, III жилд мэргэжлийн хичээл 47, IV жилд мэргэжлийн хичээл, дадлага 43 багц цаг тус тус судална. Энэхүү хөтөлбөр нь эмнэлгийн тусламж үйлчилгээг сайжруулахын тулд биологи, компьютерийн шинжлэх ухаан, мэдээллийн шинжлэх ухааныг хослуулснаар онцлог юм. Төгсөгчид нь мэдээллийн менежмент, эрүүл мэндийн боловсрол, судалгаа, өвчний оношилгоо, эмчилгээ, өвчтөний хяналт, тайлан зэрэг эрүүл мэндийн тусламж үйлчилгээний бүхий л салбарт шаардлагатай компьютерийн системийг зохион бүтээх, хэрэгжүүлэх, дэмжих чадвартай болдог.

Healthcare Informatics, B.S, Этиопын Гондарийн их сургууль

Эрүүл мэндийн Мэдээлэл зүйн хичээлийн кредит оноо нь манай сургалтын хөтөлбөрт хамрагдсан хөтөлбөрүүдээс илүү өндөр байгаа нь төгсөх оюутнууд улсын эрүүл мэндийн системийн гүнзгий ойлголттой төслийн менежментийн болон техникийн ур чадварын өндөр түвшний мэдлэгтэй оюутан болно гэж үзжээ.

Төгсөгч нь тус улс орны эрүүл мэндийн мэдээллийн технологийн өсөн нэмэгдэж буй эрэлт хэрэгцээг хангах чадвартай, менежментийн болон техникийн албан тушаалд сайн бэлтгэгдсэн мэдээлэл зүйч болно гэж төлөвлөжээ.

Дээрх олон улсын хөтөлбөрүүдэд математик, эрүүл мэндийн мэдээллийн технологи, эрүүл мэндийн тусламж үйлчилгээний хүртээмж, аюулгүй байдал, өгөгдлийн тандалт, өгөгдлийн шинжилгээ зэрэг мэдээллийн технологийн хичээлүүд байгаа нь ижил байна.

Healthcare Informatics, B.S, MontanaTech их сургууль

Хөтөлбөрөөр суралцах хугацаа 4 жил, 120 багц цагийг судална. Эрүүл мэндийн тусламж үйлчилгээний мэдээллийн системийн менежмент, үнэлгээ, хэрэгжилт, хөгжүүлэлт болон дизайны олон төрлийн хичээлийг агуулахын зэрэгцээ удирдлагын шаардлагатай мэдлэг, чадварыг оюутнуудад өгдөг. Төгсөгч нь эмнэлэг, эрүүл мэндийн тусламж үйлчилгээний болон нийгмийн эрүүл мэндийн агентлагууд, мэдээллийн технологийн фирмүүд, эрдэм шинжилгээний хүрээлэнгүүд, даатгалын салбар зэрэг олон ажлын байранд ажилладаг.

С. Хөтөлбөрөөр төгсөгч, ажил олгогчдоос авсан судалгаа

Хөтөлбөрийг хэрэгжүүлэгч тэнхим, төгсөгчид, ажил олгогч байгууллагууд байнга холбоотой ажилладаг бөгөөд ажлын байрны шаардлага, нийцтэй байдал буюу мэргэжилтний мэдлэг чадварын судалгааг авсаар байна. Эдгээр төгсөгчдийн хөдөлмөрийн зах зээлд эзлэх байр суурийн судалгаагаар 85% нь ажлын байраар хангагдсан байна [30] бөгөөд 58% нь мэдээллийн технологийн мэргэжилтэн, 10% нь сүлжээний инженерийн албан тушаалд ажиллаж байна (Зураг 1).

Судалгаанд Эрүүл мэндийн Мэдээлэл зүйн хөтөлбөрөөр төгссөн нийт 49 төгсөгчөөс (17 эр, 32 эм) хөтөлбөрийг үнэлэх болон хөтөлбөрөөр эзэмшсэн ур чадварын үнэлгээний нийт 18 асуултаар үнэлэхэд ерөнхий дундаж **3.45** гарлаа.

Ажил олгогчийн сэтгэл ханамжийн судалгаанд 14 байгууллагын 14 ажил олгогчоос 10 асуумжаар судалгаа хийлээ.

TABLE I. ТӨГСӨГЧДИЙН СЭТГЭЛ ХАНАМЖИЙН ҮНЭЛГЭЭ

Асуулт	Дундаж
Та АШУҮИС-д эзэмшсэн мэдлэг, чадвар, хандлагад сэтгэл хангалуун байна уу?	3.31
Сургалт үр дүнтэй сайн эсэх	3.31
Хичээл хоорондын уялдаа	3.29
Сургалтын зохион байгуулалт	3.41
Оюутны ачаалал тохиромжтой эсэх	3.67
Сургалтын орчин хангалттай сайн байсан уу	4.00
Багшийн заах арга зүй, ур чадвар сайн байсан уу	3.61
Үнэлгээ бодитой байсан эсэх	3.73
Ном, сурах бичгийн хангамж сайн байсан эсэх	3.57
Суралцаж байх хугацаандаа олж авсан гадаад хэлний мэдлэг тань хангалттай байна уу?	2.55
Суралцаж байх хугацаандаа нийгмийн харилцааны чадварууд (багаар ажиллах, мэдээлэл солилцох, гадаад хэлээр харилцах) зэрэг мэдлэгийг хэр олж авсан бэ?	3.27
Дадлагаар ур чадвар мэдлэгээ сайжруулж чаддаг уу?	3.31
Дадлага хийх хугацаа хүрэлцдэг үү?	3.27
Судалж байсан хичээлүүдийн дотор сэдэв, агуулгын давхардал хэр зэрэг байсан вэ?	3.29
Сургалтын материаллаг бааз, лабораторийн тоног төхөөрөмжийн хангамж ямар байсан бэ?	3.65
Тухайн мэргэжлээр суралцаж төгссөндөө сэтгэл хангалуун байдаг уу?	3.43
Одоогийн ажилдаа сэтгэл хангалуун байдаг уу?	3.98
Мэргэжлийн хувьд өсөн дэвжих боломжууд нээлттэй байдаг уу?	3.51
Ерөнхий дундаж	**3.45**

Судалгаанд мэргэжилтний ажлын байрны шаардлага буюу мэргэжилтний мэдлэг чадварыг үнэлэхэд ерөнхий дундаж **4.08** гарлаа.

TABLE II. АЖИЛ ОЛГОГЧИЙН СЭТГЭЛ ХАНАМЖИЙН ҮНЭЛГЭЭ

Асуулт	Дундаж
АШУҮИС-ийн бэлтгэдэг ЭММЗ мэргэжилтэн бэлтгэдэг Мэдээллийн технологийн хөтөлбөрийн чанарт хэр сэтгэл хангалуун байдаг бэ?	3.86
Би АШУҮИС-ийн "Эрүүл мэндийн Мэдээлэл зүйн" мэргэжилтнийг үргэлжлүүлэн бэлтгэх шаардлагатай байна уу?	4.57
Манай төгсөгч ажлын байрны шаардлагыг хэрхэн хангаж байна вэ?	4.00
Мэргэжлээрээ ажиллахад түүний онолын мэдлэгийн түвшин хэр зэрэг шаардлага хангаж байна вэ?	3.93
Манай төгсөгч нь мэргэжлийн ур чадварын хувьд хэр зэрэг шаардлага хангаж байна вэ?	4.00
Манай төгсөгч нь орчин үеийн хөгжлийн чиг хандлагатай нийцүүлэн ажлаа гүйцэтгэж чаддаг уу?	4.07
Манай төгсөгч нь бусад сургуулийг төгссөн мэргэжилтнүүдтэй харьцуулахад мэдлэг, ур чадварын хувьд хэр зэрэг байгааг үнэлнэ үү.	4.14
Ерөнхий дундаж	**4.08**

Эрүүл Мэндийн Сайдын тушаалаар “Эмнэлгийн зарим мэргэжилтний болон эмнэлгийн бус зарим ажилтны ажлын байрны тодорхойлолтын үлгэрчилсэн загвар”-ыг батлах тухай тушаалын XVI хавсралтаар “Мэдээлэл технологийн мэргэжилтэний ажлын байрны тодорхойлолтын үлгэрчилсэн загвар”-ыг баталсан байдаг [31].

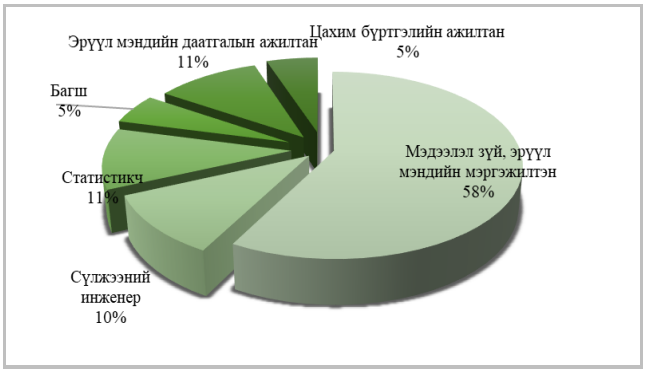


Fig. 2. Төгсөгчдийн ажил эрхлэлт

Одоогийн байдлаар төгсөгчдийн 23% нь төгсөлтийн дараах сургалтанд хамрагдсан ба үүнээс 11.6% нь магистрын зэргээ амжилттай хамгаалсан байна [32].

III. МАТЕРИАЛ, АРГА ЗҮЙ

Энэхүү судалгааны материалд Монгол Улсын эрүүл мэндийн хөгжлийн төвийн тайлан, олон улсын ижил төстэй сургалтын хөтөлбөрүүд, хөтөлбөрөөр төгсөгч болон ажил олгогчийн сэтгэл ханамжийн үнэлгээгээр үүссэн мэдээллийн сангууд орно.

Олон улсын ижил төстэй сургалтын хөтөлбөрүүдэд харьцуулсан шинжилгээ хийх болон төгсөгчдийн сэтгэл ханамжийн үнэлгээ болон бусад холбогдох тайлан мэдээлэлд үндэслэн хөтөлбөрийн шинэчлэлтийг хийлээ.

IV. СУДАЛГААНЫ ҮР ДҮН

A. Хөтөлбөрийн шинэчлэлт

Эрүүл мэндийн Мэдээлэл зүйн мэргэжилтэн бэлтгэх сургалтын хөтөлбөрийг 2003, 2005, 2008 2010, 2017, 2020, 2025 онуудад шинэчлэн баталжээ.

Шинэчилсэн 2025 оны сургалтын хөтөлбөр нийт багц цаг 120 багц цаг бөгөөд үүнээс

- ерөнхий суурь, сонгонтойгоо 34 багц цаг,
- мэргэжлийн суурь сонгонтойгоо 30 багц цаг,
- мэргэших сонгонтойгоо 46 багц цаг,
- сүлжээ ба эмнэлзүйн программ хангамжийн дадлага 3 багц цаг,
- өгөгдлийн удирдлага, программчлалын дадлага 3 багц цаг,
- диплом ажил 4 багц цаг байхаар төлөвлөгдсөн байна.

Эрүүл мэндийн Мэдээлэл зүйн сургалтын хөтөлбөрийг шинэчлэн сайжруулахдаа

- Боловсрол, Шинжлэх Ухааны сайдын 2024 оны А147 тоот тушаалын дагуу
- Information Technology guide зөвлөмж болон дотоодын мэдээллийн технологийн хөтөлбөрүүдтэй харьцуулсан судалгаа,
- Гадаад, дотоодын ижил төстэй мэргэжлийн хөтөлбөрүүдтэй харьцуулсан судалгаа,
- Ажил олгогчоос авсан асуумж судалгаа, хэлэлцүүлэг,
- Төгсөгчдөөс авсан асуумж судалгаа, хэлэлцүүлгийн үр дүнд шинэчлэв.

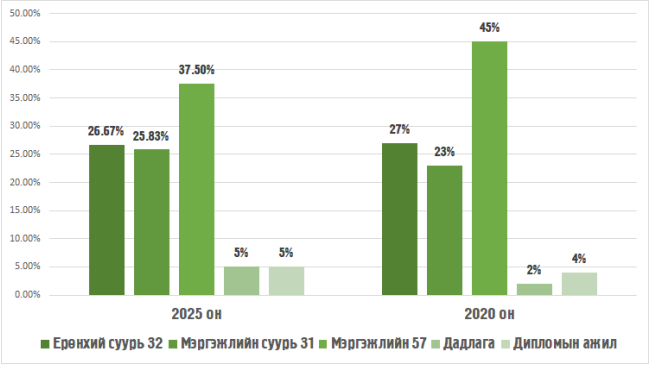


Fig. 3. Сургалтын хөтөлбөрийн харьцуулалт

Шинэчлэлтэд:

- Байгууллагын мэдээллийн технологийн хэрэгцээ, шаардлагыг хангах,
- Өвчний олон улсын ангилал, анатоми, физиологийн үндэс, анагаахын нэр томьёо судлах зэрэг эмнэлэг биологийн их өгөгдөл, түүнд хийх боловсруулалт, шинжилгээ, загварчлал хийх
- Эрүүл мэндийн тусламж үйлчилгээ, зайн эмчилгээ, өвчтөний цахим түүхийг сүлжээнд аюулгүй байдлыг ханган ажиллах, нууцлал, хамгаалалт хийх
- Эрүүл мэндийн байгууллагын мэдээлэл, цахим харилцаа, түүний тогтолцооны бүтэц зохион байгуулалтыг удирдах,

- Эрүүл мэндийн эдийн засгийн өртөг зардал тооцох, үр ашиг, нөлөөний дүн шинжилгээ хийх, дүгнэлт гаргах зэрэг онолын мэдлэг, чадвартай мэргэжилтэн бэлтгэхээр төлөвлөсөн байна.

B. Ерөнхий суурь хичээлийн өөрчлөлт

Эрүүл мэндийн Мэдээлэл зүйн хөтөлбөрийн ерөнхий суурь хичээлүүдийг Боловсрол, Шинжлэх Ухааны сайдын 2024 оны 3 сарын 20 өдрийн А/147 тушаалын "Дээд боловсролын бакалаврын мэргэжлийн чиглэл /хөтөлбөр/-д тавигдах нийтлэг шаардлага"-ын дагуу Байгалийн ухаан, Хүмүүнлэгийн ухаан, Нийгмийн ухаан, Амьдрах ухаан, Салбар чиглэлийн ерөнхий суурь гэсэн таван агуулгын хүрээнд боловсруулав.

C. Мэргэжлийн суурь хичээлийн өөрчлөлт

Шинэчилсэн хөтөлбөрийг өмнөх 2020 оны хөтөлбөртэй харьцуулахад дараах өөрчлөлтүүд хийгдсэн байна. Үүнд: *Нэр солигдсон*

- Анагаахын нэр томьёо,
- Компьютерын сүлжээ,

Суурь, нэр болон агуулга, багц цагийн өөрчлөлт

- Алгоритм, программчлал,
- Магадлалын онол, математик статистик,
- Тоон электроник,
- Веб систем хичээлүүдийн гэсэн гурван өөрчлөлт зэрэг хийгдсэн байна.

TABLE III. МЭРГЭЖЛИЙН СУУРИЙН ХИЧЭЭЛИЙН ЧИГЛЭЛ

Програмчлал (багц цаг)	
- Өгөгдлийн бүтэц	3
- Визуал программчлал	3
Өгөгдлийн сан (багц цаг)	
- Өгөгдлийн сангийн систем	3
- Магадлалын онол, математик статистик	3
Эрүүл мэнд (багц цаг)	
- Биохими ба биоөгөгдөл	3
- Анатоми	2
- Физиологи	2
Сүлжээ (багц цаг)	
- Компьютерын сүлжээ	3
- Тоон электроник	3
Вэб (багц цаг)	
- Веб систем	3

D. Мэргэших хичээлийн өөрчлөлт

Шинэ хөтөлбөрийн мэргэших хичээлүүдийг 2020 оны хөтөлбөртэй харьцуулахад дараах өөрчлөлтүүд хийгдсэн байна. Үүнд:

Шинээр нэмэгдсэн хичээл

- Биомэдрэгч,
- Судалгааны арга зүй

Суурь солигдсон

- Визуал программчлал,
- Үүлэн тооцоолол

Нэр солигдсон

- Үйлдлийн систем
- Эрүүл мэндийн мэдээллийн систем ба дизайн,
- Хиймэл оюун хичээл

Суурь, нэр солигдсон, багц цаг өөрчлөлт

- Эрүүл мэндийн өгөгдлийн сангийн технологи
- Зайн анагаах ухаан гэсэн гурван өөрчлөлт хийгдсэн.

TABLE IV. МЭРГЭШИХ ХИЧЭЭЛИЙН ЧИГЛЭЛ

Эрүүл мэндийн өгөгдөл боловсруулалт 50% (багц цаг)	
- Анагаахын нэр томьёо	2
- Эрүүл мэндийн өгөгдлийн сангийн технологи	3
- Хиймэл оюун	3
- Өгөгдлийн шинжилгээ ба тандалт	3
- Зайн анагаах ухаан	2
- Биоинформатик	2
- Дүрс боловсруулалт	2
- Биофизик	2
- Биологийн загварчлал, симуляци	2
- Биомэдрэгч	2
Мэдээллийн системийн удирдлага, зохион байгуулалт 30.43% (багц цаг)	
- Үйлдлийн систем	3
- Эрүүл мэндийн мэдээллийн систем ба дизайн	3
- Эрүүл мэндийн хууль, ёс зүй	2
- Эрүүл мэндийн цахим түүх	2
- Үүлэн тооцоолол	2
Компьютерийн сүлжээ 19.57% (багц цаг)	
- Мэдээллийн аюулгүй байдал	2
- Өгөгдлийн холбоо	2

2025 оны сургалтын хөтөлбөрийг

- Програмчлал,
- Өгөгдлийн сан,
- Эрүүл мэнд,
- Компьютерийн сүлжээ гэсэн чиглэлээр боловсруулжээ (Хүснэгт 3).

Хөтөлбөрийн нийт хичээлүүдээс Эрүүл мэндийн өгөгдөл боловсруулалт 18.3%, Мэдээллийн системийн удирдлага, зохион байгуулалт 12.5%, Компьютерийн сүлжээ 3.3% эзэлж байна.

Сургалтын хөтөлбөрт орсон нийт өөрчлөлтийг тоймлон үзүүлэв (Зураг 3).

Шинэчилсэн Эрүүл мэндийн Мэдээлэл зүйн хөтөлбөрийн мэргэжлийн хичээлүүдийг эрүүл мэндийн салбарт нэвтэрч байгаа мэдээллийн технологийн сүүлийн үеийн хөгжил, дэвшлийн чиг хандлага, эрүүл мэндийн мэдээллийн технологийн мэргэжлээр мэргэжилтэн бэлтгэдэг Австрали, АНУ, Этиоп зэрэг улсын хөтөлбөрүүдтэй харьцуулан, боловсролын магадлан итгэмжлэлийн үндэсний зөвлөлийн шинжээчид, мэдээллийн технологийн салбарын тэргүүлэх зэргийн эрдэмтэн судлаачдын зөвлөмжөөр Эрүүл мэндийн өгөгдөл боловсруулалт, Мэдээллийн системийн удирдлага, зохион байгуулалт, Компьютерийн сүлжээ гэсэн гурван чиглэлээр төлөвлөн боловсруулав (Хүснэгт 4).

V. ХЭЛЭЛЦҮҮЛЭГ

Дэлхийн нийт мэдээллийн 30 орчим хувийг эрүүл мэндийн салбар бүрдүүлдэг бөгөөд 2025 он гэхэд 36 хувь болж өсөх төлөвтэй байна. Тэр дундаа эрүүл мэндийн салбарт хүн амын өсөлтөөс гадна, хүний насжилттай холбоотойгоор мэдээлэл ихсэж байна.

Энэхүү мэдээлэлд үр дүнтэй дүн шинжилгээ хийж, ашиглах чадвар нь эрүүл мэндийн байгууллагуудын амжилттай үйл ажиллагаанд чухал үүрэгтэй.

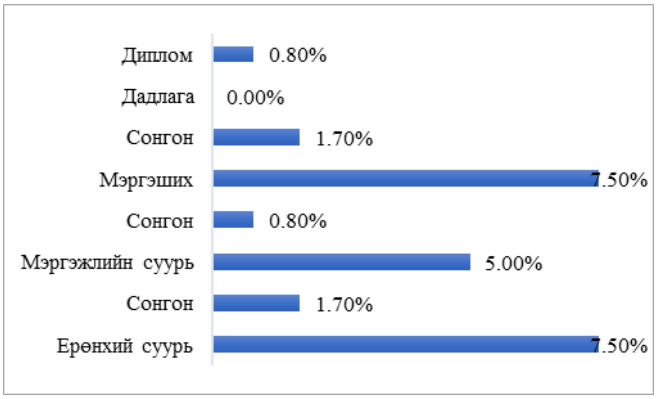


Fig. 4. Хөтөлбөрт орсон өөрчлөлт

Гэсэн хэдий ч Аркадиа саяхан гаргасан тайланд эрүүл мэндийн байгууллагуудын үүсгэсэн мэдээллийн 60 хүрэхгүй хувийг ухаалаг бизнесийн шийдвэр гаргахад ашигладаг болохыг тогтоожээ. Үүний зэрэгцээ эрүүл мэндийн салбарын удирдагчдын 93% нь тэдний гүйцэтгэлд чухал үүрэгтэй гэдэгтэй санал нэг байна. Гэхдээ аналитикт хөрөнгө оруулах нь том, жижиг эрүүл мэндийн байгууллагуудын нийтлэг сорилт хэвээр байна.

Энэхүү мэдээллийн тэсрэлт, анагаах ухааны салбарын технологийн дэвшил, нууц мэдээллийг нууцлах хэрэгцээ нь мэдээллийн шинжлэх ухаан, компьютерийн шинжлэх ухаан, эрүүл мэндийн тусламж үйлчилгээний уулзвар дээр ажилладаг эрүүл мэнд, эрүүл мэндийн мэдээллийн мэргэжилтнүүдийн хүчирхэг зах зээлийг бий болгосон.

Энэхүү Эрүүл мэндийн Мэдээлэл зүйн мэргэжлээр төгсөгчдийн 85% нь ажлын байраар хангагдсан байгаа нь "Хөгжлийн бодлого (2015-2024)"-ын шалгуур үзүүлэлт болох төгсөгчдийн ажил эрхлэлтийн хувь 85.1% гэсэн үзүүлэлтийг хангаж байна [33].

Мэргэжилтний ажлын байран дахь сул болон давуу талын судалгаанаас дүнгээс харахад ажил олгогч байгууллагууд мэргэжлийн ур чадвартайгаас гадна мэргэжлийн гадаад хэлтэй, багаар ажиллах чадвартай, харьцаа хандлага сайтай, ёс зүйтэй, ажлын ачаалал болон бие даан ажиллах ур чадвартай төгсөгчийг ажилд авах хүсэлтэй нь харагдаж байна. Иймээс сургалтын хөтөлбөрт гадаад хэлний хичээлийг сонгон судлах хичээлд оруулж, оронд нь академик англи хэл IELTS (ENG-IT11) болон мэргэжлийн англи хэл (PEN-IT12) гэсэн хоёр хичээлийг нэмж оруулсан. Мөн харьцаа хандлага сайтай, ёс зүйтэй мэргэжилтэн бэлтгэхэд чиглэсэн Хувь хүний манлайлал (LED-IT21) гэсэн хичээлийг сонгон судлахаар оруулсан болно.

Хөтөлбөрийн зах зээлийн хэрэгцээ, шаардлага нийцтэй байдлын кластер шинжилгээний дүнгээс харахад хичээл нь орчин үеийн хөгжлийн чиг хандлагатай хэр нийцдэг эсэх болон мэдлэг, ур чадварын хувьд бусад мэргэжлийн сургуулиудтай өрсөлдөх чадвартай эсэхэд анхаарах хэрэгтэй байна. Иймээс сургалтын агуулгыг сайжруулах, онол дадлагын хичээлийн харьцааг ойртуулах, багш нарын сурган заах ур чадварыг дээшлүүлэхэд зорилт тавин ажиллаж байна. Энэ ажлын хүрээнд зарим дадлагын хичээлийг семинарын хичээл болгоод гарын авлагуудыг гаргаад байна. Жишээ нь: Өгөгдлийн сангийн систем, анагаахын физик, биоинформатик хичээлүүдийн гарын авлага хэвлэгдэн гарч сургалтанд хэрэглэгдэж байна.

М. Солонго нар 2015 онд Эрүүл мэндийн Мэдээлэл зүйн хөтөлбөрөөр төгсөгчид болон ажил олгогчдоос төгсөгчдийг үнэлэх сэтгэл ханамжийн судалгааг хийсэн байна [34]. Энэхүү судалгаанд 51 ажил олгогч, хөтөлбөрөөр төгссөн 119 төгсөгч оролцсон байна. Энэ судалгаанд ажил олгогчид төгсөгчдийг үнэлсэн сэтгэл ханамжийн түвшин "дунджаас дээгүүр" буюу 3.71 гарсан байна. Энэхүү үр дүн нь манай судалгааны дүнтэй харьцуулахад 0.22%-аар бага байгаа нь харагдаж байна.

Мөн Эрүүл мэндийн Мэдээлэл зүйн хөтөлбөрөөр төгсөгчдийн ажил эрхлэлт 67% байна. Манай судалгаанд 84.5% буюу 17.5% илүү байна. Эндээс харахад мэдээлэл зүйн хөтөлбөрөөр төгсөгчид нь ажил олгогчдын ажлын байрны хэрэгцээ шаардлагыг хангасан байна. Иймээс ажил эрхлэлт нэмэгдсэн байна.

Цаашид эрүүл мэндийн мэдээлэл зүй сүүлийн жилүүдэд ихээхэн анхаарал хандуулж, 2027 он гэхэд албан тушаал 14.1%-аар өсөх төлөвтэй байна. Энэ нь үндсэн гурван чиг хандлагатай холбоотой

- Эрүүл мэндийн үйлчилгээ үзүүлэгчдийн дунд эрүүл мэндийн тусламж үйлчилгээний технологи, цахим эрүүл мэндийн бүртгэл (EHR) улам бүр нэмэгдэж байна.
- Ухаалаг төхөөрөмжүүдэд өгөгдөл цуглуулах, дамжуулах боломжийг олгодог зүйлсийн интернет (IoT) бий болж,
- Эмч болон бусад эрүүл мэндийн ажилтнуудад өвчтөнүүдийг алсаас эмчлэх боломжийг олгодог телемедициний хувьсал ихээр хийгдэж байна.

VI. ДҮГНЭЛТ

- Энэхүү хөтөлбөрөөр бэлтгэгдэн гарсан мэргэжилтнүүдийн ажлын байрны судалгаагаар 85% мэргэжлээрээ ажиллаж байгаа нь хөдөлмөрийн зах зээлийн хэрэгцээ шаардлагыг бүрэн хангаж байна.
- Хөтөлбөрөөр төгсөгчдийн сэтгэл ханамжийн үнэлгээний түвшин дунджаас дээгүүр 3.45 гарсан бөгөөд суралцаж байх хугацаандаа олж авсан гадаад хэлний мэдлэгтээ 2.55 гэсэн хамгийн бага үнэлжээ. Ажил олгогч төгсөгчийг үнэлэх мэргэжилтний мэдлэг чадварыг үнэлэхэд ерөнхий дундаж 4.08 гарлаа.
- Сургалтын хөтөлбөрийн мэргэших хичээлд Биомэдрэгч, Судалгааны арга зүй гэсэн хичээлүүд шинээр нэмэгдэж нийт өөрчлөлтийн 1.7%, суурь солигдсон 4.2%, нэр агуулга солигдсон 4.2%, суурь, нэр агуулга солигдсон, багц цагууд цуг өөрчлөгдсөн 5%, 1 багц цаг хасагдсан 0.8% зэрэг нийт 17.5% өөрчлөлт хийгдсэн байна.
- Шинэчилсэн сургалтын хөтөлбөрт эрүүл мэндийн өгөгдөл боловсруулалтын 50%, Мэдээллийн системийн удирдлага, зохион байгуулалтын 30%,

Компьютерийн сүлжээний чиглэлийн 20% хичээлүүд орлоо.

- Цаашид энэхүү хөтөлбөрийг амжилттай хэрэгжүүлэхэд мэргэжлийн багш нар олон улсад мэргэжил дээшлүүлэх болон сургалтын орчинг орчин хэрэгцээ шаардлагад нийцүүлэн шинэчлэх шаардлагатай.
- Мөн хөтөлбөрөөр төгсөгчдөд төгсөлтийн дараах сургалт болон мэргэжил дээшлүүлэх сургалтын хөтөлбөр хэрэгжүүлэх нь чухал болоод байгааг судалгаа харууллаа.

НОМ ЗҮЙ

[1] Ś. A. Batko K, "J Big Data," *The use of Big Data Analytics in healthcare,* vol. 9(1):3, 2022.

[2] Л. З. Амийра Хамид-Салдулкер, "Монгол улсын цахим эрүүл мэндийн нөхцөл байдлын үнэлгээний тайлан," 2024.

[3] "Харилцаа холбоо, мэдээллийн технологийн газар. Харилцаа холбоо, мэдээллийн технологийн онцлох үйл явдал. Улаанбаатар: НЭПКО ХХК; .", 2021.

[4] S. Paul *et al.*, "Industry 4.0 applications for medical/healthcare services," *Journal of Sensor and Actuator Networks,* vol. 10, no. 3, p. 43, 2021.

[5] I. Kickbusch, "Health promotion 4.0," vol. 34, ed: Oxford University Press, 2019, pp. 179-181.

[6] M. Osama *et al.*, "Internet of medical things and healthcare 4.0: Trends, requirements, challenges, and research directions," *Sensors,* vol. 23, no. 17, p. 7435, 2023.

[7] M. Ma *et al.*, "The need for digital health education among next-generation health workers in China: a cross-sectional survey on digital health education," *BMC Medical Education,* vol. 23, no. 1, p. 541, 2023.

[8] M. Toni and G. Mattia, "Healthcare in the 4.0 Era of Digitization," in *The Digital Healthcare Revolution: Towards Patient Centricity with Digitization, Service Innovation and Value Co-creation*: Springer, 2022, pp. 5-18.

[9] V. V. Popov, E. V. Kudryavtseva, N. Kumar Katiyar, A. Shishkin, S. I. Stepanov, and S. Goel, "Industry 4.0 and digitalisation in healthcare," *Materials,* vol. 15, no. 6, p. 2140, 2022.

[10] J. Chanchaichujit, A. Tan, F. Meng, and S. Eaimkhong, "Healthcare 4.0," *Springer Nature, Singapore. doi,* vol. 10, pp. 978-981, 2019.

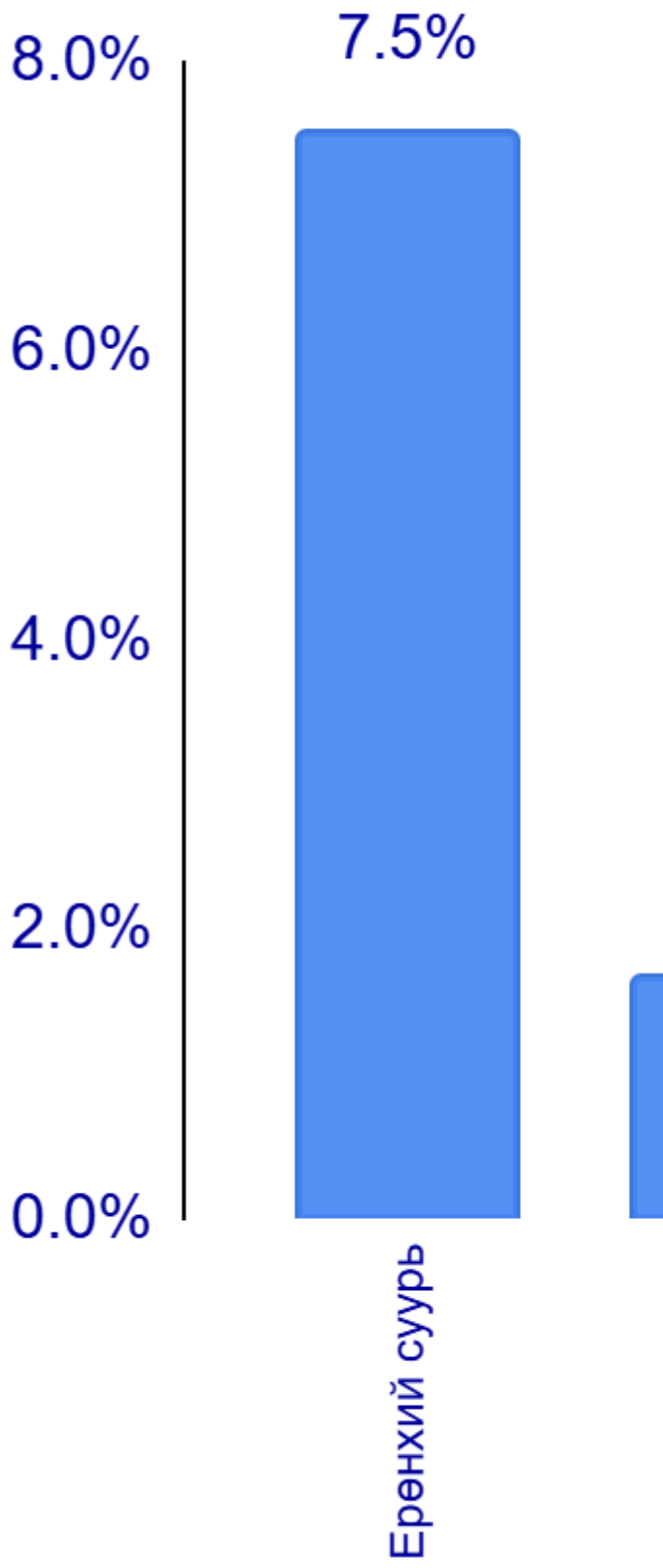
[11] "Эрүүл мэндийн статистикийн мэдээлэл боть - I," 2023.

[12] O. N. o. t. M. o. Health, *Procedures for information technology in the health sector (Responsibilities of healthcare organizations regarding information technology), Issue.* 2013.

[13] *Report, Center for Health Development*. 2019.

[14] "Хөтөлбөрийн танилцуулга," 2020 он.

[15] Б. АШУҮИС, "Эрүүл мэндийн Мэдээлэл зүйн сургалтын төлөвлөгөө," 2020 он.



Simplified graph neural network for node ranking

Amarsanaa Ganbaatar
School of Information Technology and Electronics
National University of Mongolia
Ulaanbaatar, Mongolia

Dorjnamjirmaa Badraa
School of Information Technology and Electronics
National University of Mongolia
Ulaanbaatar, Mongolia

Dalaijargal Purevsuren
School of Information Technology and Electronics
National University of Mongolia
Ulaanbaatar, Mongolia
dalaijargal@num.edu.mn

***Abstract*—Node ranking in graphs is a critical task in many applications such as social network analysis, recommendation systems, and infrastructure networks. Among various centrality measures, betweenness centrality is particularly important as it identifies the most influential nodes in terms of their control over information flow. Traditional algorithms for computing betweenness centrality often suffer from high computational complexity, especially in large-scale networks. In this paper, we investigate the impact of a simplified graph neural network model on node ranking based on betweenness centrality. Our results demonstrate that the proposed models offer a promising alternative to traditional methods and GNN-Bet model from Maurya et al [8], significantly improving computational efficiency while maintaining a competitive accuracy in node ranking.**

***Index Terms*—Betweenness centrality, graph neural networks (GNNs), node ranking**

I. Introduction

In the study of complex networks, understanding the importance of individual nodes within the network is crucial for a wide range of applications, from social network analysis and recommendation systems to transportation and infrastructure networks. One of the most widely used metrics for quantifying node importance is betweenness centrality. Betweenness centrality measures the extent to which a node lies on the shortest paths between other nodes, capturing its role in facilitating information flow across the network. Nodes with high betweenness centrality are considered critical for maintaining the connectivity and robustness of the network. However, computing betweenness centrality for large-scale networks is computationally expensive. The classic algorithm, based on the calculation of shortest paths between all pairs of nodes, has a time complexity of $\mathcal{O}(nm)$, making it impractical for large graphs, where n is the number of nodes, and m is the number of edges. As networks continue to grow in size and complexity, more efficient approaches are needed to estimate node importance without compromising accuracy.

Recent advancements in Graph Neural Networks (GNNs) offer a promising direction for addressing this challenge. GNNs have emerged as a powerful class of deep learning models for analyzing graph-structured data. Traditionally, GNNs have been applied to tasks such as node classification, link prediction, and graph classification, but their potential for node ranking tasks—such as ranking nodes by centrality measures—has recently gained attention. A key advantage of GNNs is their ability to aggregate information across multiple layers, enabling them to capture both local and global graph structures. In the case of betweenness centrality, nodes that mediate information flow across the graph (e.g., those involved in many shortest paths) will have strong connections with other important nodes. A well-trained GNN can learn to identify such nodes based on their interactions within the graph, and predict their centrality without the need for expensive path computations.

In this paper, we investigate the impact of simplified graph neural network model for node ranking based on betweenness centrality. Our results show that simplified GNN-based approach provides an efficient solution for node ranking, maintaining competitive accuracy while significantly reducing computational overhead.

The rest of the paper is structured as follows: Section 2 reviews related work in betweenness centrality computation, node ranking and graph neural network models. Section 3 describes the proposed methods, including the architecture of the GNN-based model and the training procedure. Section 4 presents experimental results and a comparative analysis of the proposed method against GNN-Bet method from [8] . Finally, Section 5 concludes the present work and outlines future directions for research.

II. Related works

A. Betweenness centrality computation

The computation of betweenness centrality has been a significant area of research due to its computational complexity, especially for large graphs. Brandes proposed an improved algorithm that generates a directed acyclic graph (DAG) using breadth-first search (BFS) to aggregate contributions of nodes on paths, which reduces the time complexity to $\mathcal{O}(|V||E|)$ but remains prohibitive for large graphs. To address this, two main approaches have been developed: distributed computation algorithms that extend the computation from a single machine to a cluster of high-performance machines, significantly reducing

computation time, and approximation methods that do not require exact calculations.

Several approximation algorithms have been proposed, such as the method by Brandes and Pich, which uses a subset of k starting nodes (pivots) to estimate betweenness centrality. However, this method can lead to overestimation for unimportant nodes near the pivot. Geisberger et al. [1] introduced a scaling function to mitigate this bias by adjusting the contributions from nearby pivots. Riondato et al. [2] proposed a different approach that samples a set of shortest paths between randomly selected source-target pairs to approximate betweenness centrality. Additionally, Borassi et al. [3] developed an adaptive sampling technique that provides faster computation of betweenness within a specified absolute error.

B. Graph neural network

Graph Neural Networks (GNNs) have gained significant attention in recent years for their ability to process graph-structured data and the inductive capabilities of GNNs, enabling models to generalize across different graph structures, as highlighted by various studies. Various models have been proposed to leverage the unique properties of graphs for different tasks. For example, Hamilton et al. [4] introduced a GNN framework that utilizes a message-passing mechanism to aggregate features from neighboring nodes, which is foundational to many GNN architectures. Additionally, the work by Kipf and Welling [5] on Graph Convolutional Networks (GCNs) has been influential, demonstrating how convolutional operations can be adapted to graph data, allowing for semi-supervised learning. GCNs stack layers of learned first-order spectral filters followed by a nonlinear activation function to learn graph representations. The authors have achieved state-of-the-art results in various application areas, including but not limited to citation networks ([5]), social networks ([6]), applied chemistry, natural language processing and computer vision. In particular, Wu et al. [7] argues the unnecessary complexity of GCN, developing a simplified GCN (SGCN) model by removing nonlinearities and collapsing multiple weight matrices into one. The performance of SGC rivals, if not surpasses, the performance of GCNs and state-of-the-art graph neural network models across a wide range of graph learning tasks. In addition, by pre-computing the fixed feature extractor S^K, the training time is reduced to a record low.

C. Node ranking

Node ranking in graphs is a critical task that helps to identify influential nodes based on their roles in the spread of information and connectivity. This process is essential for various applications, such as controlling the spread of infectious diseases and preventing misinformation on social networks. Recent advancements include the use of Graph Neural Networks (GNNs) for node ranking, which leverage feature aggregation from neighboring nodes to approximate centrality measures more efficiently. Maurya et al., [8] proposed GNN-based models, GNN-Bet and GNN-Close, specifically target the approximation of betweenness and closeness centrality, respectively, and have shown significant improvements in both performance and speed compared to traditional methods. These models are inductive, allowing them to be trained on one set of graphs and applied to others, which enhances their versatility in dynamic graph scenarios.

III. Methodology

In this study, we propose two simplified Graph Neural Network (GNN) models derived from the GNN-Bet model presented in the work of Maurya et al. [8] The primary modifications involve simplifying the architecture by removing nonlinear activation functions (ReLU) and multi-layer perceptron (MLP) layers. The simplified approaches aim to achieve computational efficiency while maintaining the accuracy of the node ranking based on the centrality of the connection.

A. Proposed models

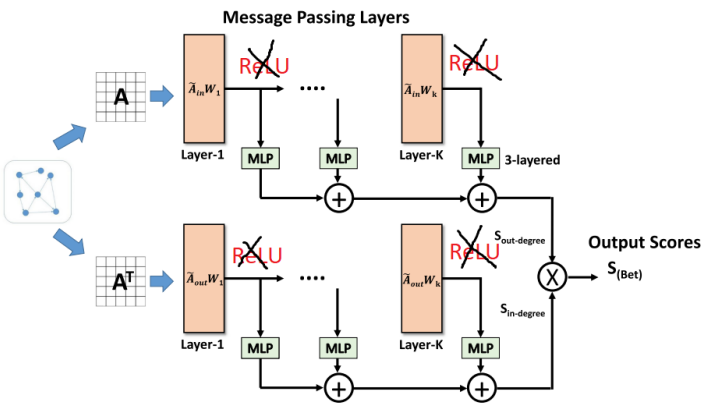


Fig. 1. An illustration of GNN-Linear-Bet model architecture.

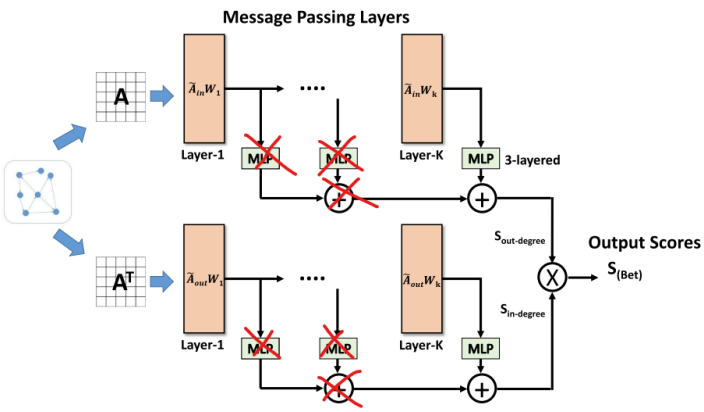


Fig. 2. An illustration of sGCN-Bet model architecture.

1) GNN-Linear-Bet model: The GNN-Linear-Bet model simplifies the original GNN-Bet architecture by eliminating the nonlinear activation function (ReLU). The linear transformation is maintained to capture essential graph information more efficiently.

Algorithm 1 GNN-Linear-Bet (Forward propagation)

Input: Directed Graph adjacency matrix A; depth K; weight matrices $W^{(k)}$

Output: Betweenness Centrality value vector, S_{Bet}

$\tilde{A}_{\text{out-degree}} \leftarrow \text{ModifyAdjacencyRow}(A)$
$\tilde{A}_{\text{in-degree}} \leftarrow \text{ModifyAdjacencyRow}(A^\top)$
for $k = 1, \ldots, K$ **do**
  $H^{(k)}_{\text{out-degree}} \leftarrow \tilde{A}_{\text{out-degree}} H^{(k-1)}_{\text{out-degree}} W^{(k)}$
  $H^{(k)}_{\text{in-degree}} \leftarrow \tilde{A}_{\text{in-degree}} H^{(k-1)}_{\text{in-degree}} W^{(k)}$
  $s^{(k)}_{\text{out-degree}} \leftarrow \text{MLP}\big(H^{(k)}_{\text{out-degree}}\big)$
  $s^{(k)}_{\text{in-degree}} \leftarrow \text{MLP}\big(H^{(k)}_{\text{in-degree}}\big)$
end for
$S_{\text{out-degree}} \leftarrow \sum_{k=1}^{K} |s^{(k)}_{\text{out-degree}}|$
$S_{\text{in-degree}} \leftarrow \sum_{k=1}^{K} |s^{(k)}_{\text{in-degree}}|$
$S_{\text{Bet}} \leftarrow S_{\text{out-degree}} \times S_{\text{in-degree}}$

2) sGCN-Bet model: The sGCN-Bet model further simplifies the architecture by removing both the nonlinear activation function (ReLU) and the intermediate scoring layers, utilizing only the final layer outputs to compute the betweenness centrality.

Algorithm 2 sGCN-Bet (Forward propagation)

Input: Directed Graph adjacency matrix A; depth K; weight matrices $W^{(k)}$

Output: Betweenness Centrality value vector, S_{Bet}

$\tilde{A}_{\text{out-degree}} \leftarrow \text{ModifyAdjacencyRow}(A)$
$\tilde{A}_{\text{in-degree}} \leftarrow \text{ModifyAdjacencyRow}(A^\top)$
for $k = 1, \ldots, K$ **do**
  $H^{(k)}_{\text{out-degree}} \leftarrow \tilde{A}_{\text{out-degree}} H^{(k-1)}_{\text{out-degree}} W^{(k)}$
  $H^{(k)}_{\text{in-degree}} \leftarrow \tilde{A}_{\text{in-degree}} H^{(k-1)}_{\text{in-degree}} W^{(k)}$
end for
$S_{\text{out-degree}} \leftarrow \text{MLP}\big(H^{(K)}_{\text{out-degree}}\big)$
$S_{\text{in-degree}} \leftarrow \text{MLP}\big(H^{(K)}_{\text{in-degree}}\big)$
$S_{\text{Bet}} \leftarrow S_{\text{out-degree}} \times S_{\text{in-degree}}$

B. Models analysis

To deeply analyze the impact of our architectural simplifications, we present a detailed evaluation inspired by the approach focusing on layer-wise and structural component analysis.

Impact of Non-linear Activation Removal: The primary modification in the GNN-Linear-Bet and sGCN-Bet models, compared to the baseline GNN-Bet, is the removal of the nonlinear activation function (ReLU). This simplification eliminates nonlinear transformations at each layer, directly propagating structural information. Empirical results show linearization improves computational efficiency without significantly compromising accuracy, suggesting a beneficial trade-off in complexity and performance.

Impact of Scoring Layer Removal (sGCN-Bet Model): The sGCN-Bet model further simplifies by omitting intermediate scoring (MLP) layers and directly computes node centrality scores using the final embedding layer. This significantly reduces computational complexity but makes the model more sensitive to network depth. Experiments indicate sGCN-Bet performs optimally at shallower depths, with noticeable performance degradation at higher layer counts.

Layer-wise analysis: Experimental results highlight that deeper models capture broader graph structures but may also amplify noise, particularly when simplifications are involved. While GNN-Bet and GNN-Linear-Bet remain stable with increasing depth, sGCN-Bet's performance peaks early, emphasizing the importance of carefully selecting network depth in simplified architectures.

Computational Efficiency: Our simplified models substantially reduce both training and inference times. Notably, sGCN-Bet exhibits the lowest computational cost due to minimal structural complexity. This computational advantage, paired with competitive accuracy, positions the simplified models as practical solutions for large-scale node-ranking tasks.

Mathematical Relationship with Simplified Graph Convolutional Network (SGCN): To further understand the relationship between our models and existing simplified GCN architectures, we provide a mathematical comparison with the original Simplified Graph Convolutional Network (SGCN) proposed by Wu et al. [7]:

The original SGCN defines propagation as:

$$H^{(k)} = \tilde{A}^k X W$$

where:
- $\tilde{A}$: normalized adjacency matrix
- X: initial node features
- W: learnable weight matrix
- k: propagation depth

Our models adapt this propagation to directional scenarios as follows:

GNN-Linear-Bet employs distinct linear transformations at each layer, formalized as:

$$H^{(k)}_{out} = \tilde{A}_{out} H^{(k-1)}_{out} W^{(k)}, \quad H^{(k)}_{in} = \tilde{A}_{in} H^{(k-1)}_{in} W^{(k)}$$

and computes betweenness centrality via multi-layer aggregation:

$$S_{Bet} = \left(\sum_{k=1}^{K} |MLP(H^{(k)}_{out})| \right) \times \left(\sum_{k=1}^{K} |MLP(H^{(k)}_{in})| \right)$$

sGCN-Bet closely follows the original SGCN linear propagation but applies directional adjacency matrices and a single final-layer scoring step:

$$H^{(k)}_{out} = \tilde{A}_{out} H^{(k-1)}_{out} W^{(k)}, \quad H^{(k)}_{in} = \tilde{A}_{in} H^{(k-1)}_{in} W^{(k)}$$

$$S_{Bet} = MLP(H^{(K)}_{out}) \times MLP(H^{(K)}_{in})$$

This comprehensive analysis underscores how our simplified GNN models relate mathematically and conceptually to the foundational SGCN framework, clearly illustrating the contributions and improvements made specifically for efficient, directed node-ranking tasks.

IV. Experimental Results

In this section, we conduct experiments and ablation studies to analyze performance and present comparisons between models.

A. Training setup

1) Hardware and software setup: All experiments were conducted on Intel Core i7-11700 CPU @ 2.50GHz machine with 8 cores, 24 GB RAM. Software frameworks used are Pytorch for the model implementation, NetworkX, NetworKit for graph generation and betweenness calculations.

2) Hyperparameters: The model size is 100000 that is the largest training/test graph size. For smaller graph inputs, the adjacency matrix of the input graph is placed along diagonal and the rest is padded with zeros. The training is performed for 5 epochs with Adam optimizer with a learning rate parameter set to 0.005. We have set the embedding dimension equal to 20 and dropout for MLP layer equal to 0.6.

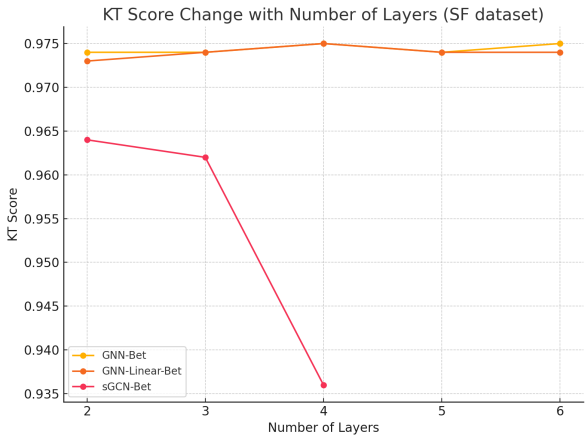
3) Training: We use a class of directed synthetic graphs, specifically Scale-free graphs, to train the model. The graph sizes range from 50,000 to 100,000 nodes. Fifteen graphs are generated with randomly assigned node numbers and generation parameters to introduce structural variations in the dataset. For training, five graphs are selected to generate 250 training samples by randomly permuting the node sequence in the adjacency matrices and their corresponding target centrality vectors. The model is trained on this training set and evaluated on 10 graphs from the test set, as well as real-world datasets from the Stanford Large Network Dataset (SNAP) website.

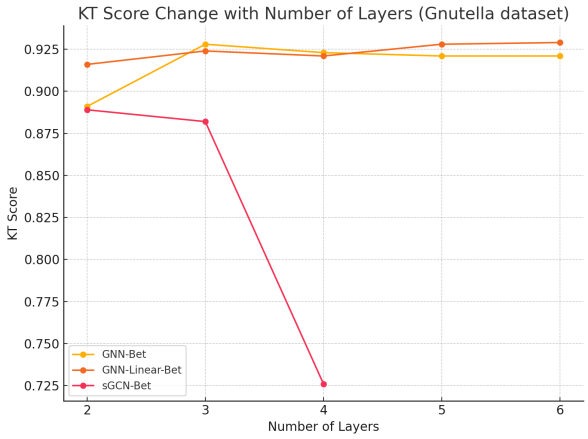
4) Real world datasets: The study uses two real-world datasets from SNAP for evaluation: **p2p-Gnutella31**, a peer-to-peer file-sharing network from Gnutella comprising 62,586 nodes and 147,892 directed edges (average degree approximately 2.36), characterized by a sparse, scale-free structure due to influential hubs and high dynamics from frequent host changes; and **soc-Slashdot0811**, a denser social network derived from interactions on the technology-focused website Slashdot.org, featuring 77,360 nodes and 905,468 directed edges (average degree approximately 11.7), notable for strong clustering, clear community structures, and interactions indicating user trust or distrust. Both datasets are employed to assess the betweenness centrality performance of simplified Graph Neural Network models, identifying influential hosts crucial for efficient communication in Gnutella, and influential users essential for information dissemination across Slashdot communities.

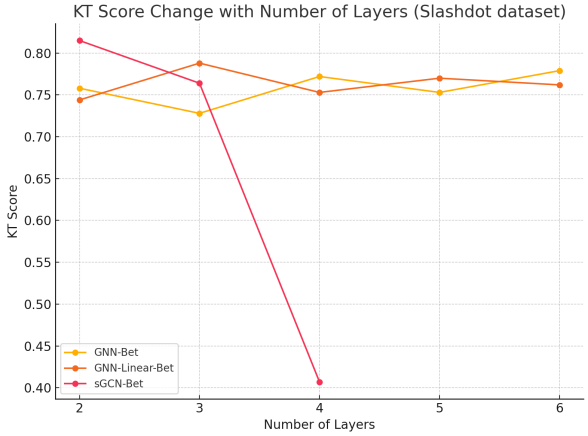
B. Comparisons

Initially, we compare the GNN-Bet, GNN-Linear-Bet, and sGCN-Bet models based on the number of layers for betweenness centrality. Then, for each corresponding number of layers, we evaluate our models for synthetic and real-world graphs using the KT score and computation time.

In the next figures shown that how does change in KT score with respect to number of layers for betweenness variant.







In the Table I, we compare the ranking performances and execution time comparison of our models GNN-Bet, GNN-Linear-Bet and sGCN-Bet for static graphs and some real-world graphs.

TABLE I
EXPERIMENTAL RESULTS

Model	Layer	Training Time (sec)	Testing Time (sec)			KT Scores		
			SF	Gnutella	Slashdot	SF	Gnutella	Slashdot
GNN-Bet	2	365.0	0.648	0.660	0.173	0.974	0.891	0.758
	3	534.9	1.04	0.125	0.284	0.974	0.928	0.728
	4	703.2	1.285	0.130	0.389	0.975	0.923	0.772
	5	903.9	1.659	0.168	0.459	0.974	0.921	0.753
	6	1093.2	1.821	0.215	0.536	0.975	0.921	0.779
GNN-Linear-Bet	2	349.8	0.674	0.079	0.169	0.973	0.916	0.744
	3	510.8	0.927	0.091	0.270	0.974	0.924	0.788
	4	674.5	1.169	0.120	0.323	0.975	0.921	0.753
	5	867.4	1.571	0.175	0.497	0.974	0.928	0.770
	6	1041.5	1.773	0.190	0.541	0.974	0.929	0.762
sGCN-Bet	2	215.1	0.450	0.042	0.201	0.964	0.889	0.815
	3	245.5	0.602	0.073	0.209	0.962	0.882	0.764
	4	290.1	0.701	0.069	0.281	0.936	0.726	0.407

V. CONCLUSIONS

In this paper, we explore the effectiveness of a simplified graph neural network model for approximating betweenness centrality. Using synthetic scale-free graphs, we train the proposed models to rank nodes based on their betweenness centrality scores. We compare the performance of several models, including GNN-Bet, GNN-Linear-Bet, and sGCN-Bet. The experimental results demonstrate that the simplified graph neural network models offer competitive performance while significantly reducing both training and testing times.

For future work, expanding the experiments to large-scale, real-world networks containing millions of nodes could further demonstrate the effectiveness and resilience of simplified graph neural network (GNN) architectures in practical, large-scale scenarios. Additionally, exploring the performance and adaptability of simplified GNN models on dynamic and evolving graphs would offer valuable insights into their capability to efficiently monitor changes in node centrality over time—especially relevant in fast-changing environments such as social media or financial transaction networks.

REFERENCES

[1] Robert Geisberger, Peter Sanders, and Dominik Schultes. 2008. Better approximation of betweenness centrality. In Proceedings of the Meeting on Algorithm Engineering & Expermiments. 90–100. Retrieved from http://dl.acm.org/citation. cfm?id=2791204.2791213

[2] Matteo Riondato and Evgenios M. Kornaropoulos. 2016. Fast approximation of betweenness centrality through sampling. Data Mining and Knowledge Discovery 30, 2 (2016), 438–475. DOI:https://doi.org/10.1007/s10618-015-0423-0

[3] Michele Borassi and Emanuele Natale. 2019. KADABRA is an Adaptive Algorithm for Betweenness via Random Approximation. ACM Journal of Experimental Algorithmics 24, 1 (2019), 1.2:1–1.2:35. DOI:https://doi.org/10.1145/ 3284359

[4] Will Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. In Proceedings of the 31st International Conference on Neural Information Processing Systems. I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Eds.), ACM, 1024–1034. Retrieved from http://papers.nips.cc/ paper/6703-inductive-representation-learning-on-large-graphs.pdf.

[5] Thomas N. Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. In International Conference on Learning Representation (ICLR'17).

[6] Deli Chen, Yankai Lin, Wei Li, Peng Li, Jie Zhou, and Xu Sun. 2019. Measuring and relieving the over-smoothing problem for graph neural networks from the topological view. Proceedings of AAAI conference on Artificial Intelligence 34 (2019), 3438–3445. Retrieved from http://arxiv.org/abs/1909.03211

[7] Felix Wu, Amauri H. Souza, Tianyi Zhang, Christopher Fifty, Tao Yu, and Kilian Q. Weinberger. 2019. Simplifying graph convolutional networks. In Proceedings of the 36th International Conference on Machine Learning. 6861–6871.

[8] Sunil Kumar Maurya, Xin Liu, and Tsuyoshi Murata. 2021. Graph Neural Networks for Fast Node Ranking Approximation. ACM Trans. Knowl. Discov. Data 15, 5, Article 78 (May 2021), 32 pages. https://doi.org/10.1145/3446217

[9] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, and M. Wang. Lightgcn: Simplifying and powering graph convolution network for recommendation. In International ACM SIGIR conference on research and development in Information Retrieval (SIGIR), pages 639–648, 2020.

Физикээс Сэдсэн Граф Нейрал Сүлжээний хувилбарууд, тэдгээрийн харьцуулалт

Д.Тэмүүжин
МУИС, МТЭС,
Мэдээлэл, Компьютерын Ухааны Тэнхим
temuujin.d@num.edu.mn

П.Далайжаргал
МУИС, МТЭС,
Мэдээлэл, Компьютерын Ухааны Тэнхим
dalaijargal@num.edu.mn

***Хураангуй*---Энэхүү судалгаанд бид Физикээс Сэдсэн Граф Нейрал Сүлжээ (PI-GNN)-ний гүйцэтгэлийг комбинаторик оновчлолын MaxCut бодлогыг шийдвэрлэхэд үнэлж, харьцуулав. Тодруулбал,Бид шинээр PI-GAT болон түүний сайжруулсан хувилбар болох PI-GATv2 загваруудыг санал болгож, өмнөх PI-GCN загвар болон PI-SAGE зэрэг аргачлалуудтай харьцуулан туршсан. Судалгаа нь шийдийн чанар (огтлолын хэмжээ) болон тооцооллын өртөг (сургалтын хугацаа, сургалтын үеийн тоо) хоорондын харилцан хамаарал, мөн хасалтын коэффициент болон суралцах хурд зэрэг хайперпараметрүүдийн эдгээр физикээс сэдсэн загваруудад үзүүлэх нөлөөг шинжилсэн. PI-GAT, PI-GATv2 болон PI-SAGE загварууд нь том графуудад илүү сайн үр дүн үзүүлж байна.**

***Түлхүүр үг*---Физикээс Сэдсэн Граф Нейрал Сүлжээ, MaxCut бодлого, Граф Конволюцийн Сүлжээ (GCN), Граф Анхаарлын Сүлжээ (GAT/GATv2), Граф Түүврийн Сүлжээ (GraphSAGE)**

I. Удиртгал

Комбинаторик оновчлолын бодлогууд компьютерын ухаан, инженерчлэл, сүлжээний дизайн, тоон биологи зэрэг олон салбарт чухал үүрэг гүйцэтгэдэг [1].Тухайлбал MaxCut нь цахилгаан хэлхээний дизайн, статистик физик зэрэгт хэрэглэгддэг [2]. MaxCut бодлогын хувьд өгөгдсөн графын оройнуудыг хоёр бүлэгт хуваахад, эдгээр хоёр бүлгийн хооронд холбогдсон ирмэгүүдийн нийт жин (эсвэл тоо) хамгийн их байлгахыг зорьдог оновчлолын бодлого юм. Уламжлалт аргаар MaxCut бодлогыг том графуудад шийдвэрлэх нь түүний NP-хард шинж чанараас (NP-hard nature) үүдэн хүндрэлтэй байдаг бөгөөд яс (exact) шийд олох нь ихэнх тохиолдолд боломжгүй байна [3].

Сүүлийн жилүүдэд чиглүүлэггүй сургалтын аргууд нь эдгээр асуудлыг үр дүнтэйгээр шийдвэрлэх шинэ боломжийг олгож байна. Үүнд Erdős Goes Neural [11], RUN-CSP [12], PI-GCN [9], PI-SAGE [10] зэрэг аргууд хөгжсөн байна. Erdős Goes Neural нь Erdős-ийн магадлалын аргыг ашиглан нейрал загваруудыг чиглүүлэггүй сургалтаар сургаж, шийдлийг тодорхой магадлалтайгаар баталгаажуулдаг [11]. RUN-CSP нь чиглүүлэггүй, рекуррент нейрал архитектурыг ашиглан хязгаарлалтыг хангах (constraint satisfaction) төрлийн бодлогуудыг шийдэх бөгөөд жижиг жишээн дээр сургасан загвараа том хэмжээний жишээнд амжилттай ашиглаж болдог [12]. PI-GCN нь статистик физикийн зарчим, гамилтонианы энергийн функцийг ашиглан физикийн харилцан үйлчлэлийг граф нейрал сүлжээнд шууд нэгтгэдэг. Жишээлбэл, энергийн функцийг алдааны функц эсвэл мессеж дамжуулах механизмд оруулах замаар оройнуудын хоорондын физик харилцан үйлчлэлийг загварчлахыг оролддог. Уг загварын хувьд алдааны функцэд оруулж загварчилсан байна [9]. Энэ нь оновчлолын бодлогын бүтцийг илүү сайн тусгаж, шийдлийн чанарыг сайжруулахад тусалдаг гэж үздэг. Энэ судалгаанд ашигласан суурь PI-GCN загвар нь статистик физикээс сэдэн гамилтонианы энергийн тэгшитгэлтэй тэнцүү тэгшитгэлийг дифференциалчилж алдааны функцийг тодорхойлсон. PI-GCN загварт GCN ашиглан хэрэгжүүлдэг бөгөөд энэ нь хөрш оройнуудаас мэдээлэл цуглуулахдаа хөршийн зэргийн нормчлол дээр суурилдаг. Энэ арга нь үр дүнтэй ч хөршүүдийн нөлөөг ялгах чадвар муу, ялангуяа бүтэц нь жигд бус графуудад сул талтай.

Граф Нейрал Сүлжээ (GNN) нь граф бүтэцтэй өгөгдлийг боловсруулах хүчирхэг арга болон хөгжиж байна [4]. Эдгээр сүлжээнүүд нь оройг төлөөлөх (node representations) хөршийн мэдээллийг нэгтгэх замаар суралцдаг. Гол архитектуруудад GCN [5], GAT [6], GATv2 [7], GraphSAGE [8] багтдаг бөгөөд эдгээр нь PI-GNN загваруудын суурьуд болж болно. Харгалзан тайлбарлавал, оройн шинэ төлөөллийг хөршүүдийн мэдээллийг жигд эсвэл жигнэсэн дунджаар нэгтгэдэг [5], анхаарлын механизмыг ашиглан хөршүүдийн мэдээлэлд ялгаатай жин оноож нэгтгэдэг [6], GAT-ийн ялгаатай жин оноох аргачлалыг илүү динамик болгосон [7], том графуудад зориулан хөршүүдээс түүвэрлэлт хийж, мэдээллийг нэгтгэх функц ашигладаг [8]. Энэхүү судалгаа нь MaxCut бодлогыг шийдвэрлэхэд зориулан шинээр боловсруулсан PI-GAT загварын гүйцэтгэлийг PI-GCN болон PI-SAGE суурь загваруудтай харьцуулан үнэлэхэд чиглэгдсэн. PI-GAT анхаарлын механизм оруулж ирснээр олон хөрштэй оройн хувьд аль оройн мэдээлэл чухал гэдгийг динамикаар суралцаж чадах тул том граф байх үед бусад загваруудаас ялгагдана. PI-SAGE түүврийн аргаар хөршийг сонгож авдаг энэ арга нь бага графд PI-GCN загвараас ялгарах зүйлгүй болох тул том графд шалгах хэрэгцээтэй. Загваруудын гүйцэтгэлийг харьцуулахын тулд төлөөллийн хэмжээсүүд ижил томъёонд захирагдах шаардлагатай гэж үзсэн. Хэдийгээр PI-GCN гэх мэт физикээс сэдсэн граф нейрал сүлжээний судалгаа хийгдсэн байгаа ч, графын анхаарал (PI-GAT), хөрш түүвэрлэлт (PI-SAGE) зэрэг өөр архитектуруудыг MaxCut зэрэг бодлогод зориулан гүйцэтгэлийг нь харьцуулсан судалгаа одоогоор дутмаг байна.

II. Арга Зүй

Бид MaxCut оновчлолын бодлогыг шийдвэрлэх зорилгоор хоёр шинэ GNN архитектурууд Физикээс Сэдсэн Графын Анхаарлын Сүлжээ (PI-GAT), (PI-GATv2) болон Физикээс Сэдсэн GraphSAGE (PI-SAGE)-ыг санал болгож байна.

GCN-ийг дараах хоёр хувилбараар орлуулсан:

Граф анхаарлын сүлжээ (GAT) нь хөршүүдийн төлөөлөл дээр үндэслэн өөр өөр жингүүдийг тооцоолдог анхаарлын механизмтай. Ингэснээр илүү чухал хөршүүдийг онцолж, энерги багасгах үйл явцыг илүү ухаалгаар удирдах боломжтой болдог. GAT архитектурт нь хөршийн анхаарлын коэффициентийг дараах томъёогоор тодорхойлдог [6]:

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(a^T[Wh_i||Wh_j]))}{\sum_{k\in\mathcal{N}(i)} \exp(\text{LeakyReLU}(a^T[Wh_i||Wh_k]))} \quad (1)$$

Энд h_i, h_j нь оройн төлөөллүүд, W нь суралцах жингүүд, a нь анхаарлын параметр бөгөөд $||$ нь конкатенаци (concatenate) үйлдэл юм.

GATv2-д хөрш оройн төлөөллүүдийг ашиглан илүү уян хатан байдлаар анхаарлын жинг тооцоолдог ба ингэснээр графын оройн хоорондох харилцан үйлчлэлийг илүү нарийн илэрхийлэх боломжийг олгодог [7].

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(a^T\text{LeakyReLU}\left(W[h_i||h_j]\right)\right)\right)}{\sum \exp\left(\text{LeakyReLU}\left(a^T\text{LeakyReLU}\left(W[h_i||h_k]\right)\right)\right)} \quad (2)$$

Энд h_i, h_j нь оройн төлөөллүүд, W нь суралцах жингүүд, a нь анхаарлын параметр бөгөөд $||$ нь конкатенаци (concatenate) үйлдэл юм.

GraphSAGE архитектур нь хөршүүдээс мэдээллийг санамсаргүйгээр түүвэрлэн дараах байдлаар нэгтгэдэг [8]:

$$h_i^{(l+1)} = \sigma\left(W \cdot \text{AGGREGATE}\left(h_i^{(l)}, h_j^{(l)}, \forall j \in \mathcal{S}_i\right)\right) \quad (3)$$

Энд $\mathcal{S}_i$ нь санамсаргүй сонгогдсон хөршүүд, W ба σ нь суралцах параметр болон активаци функц юм. Энэ арга нь том графуудад үр ашгийг нэмэгдүүлж, загварын ерөнхийлөх чадварыг сайжруулахад тусалдаг.

Эдгээр өөрчлөлтүүд нь PI-GCN-ийн гол санаа болох физикт суурилсан энергийг багасгах зарчмыг хадгалж үлдэнэ. Өөрөөр хэлбэл, загвар нь энергийн функц $L = p^T Qp$ хэлбэрээр өгөгдсөн зорилгыг алдааны функц болгон ашигладаг хэвээр үлдэнэ. Харин GAT болон GraphSAGE-г ашигласнаар мессеж дамжуулалтын илэрхийлэх чадвар болон ерөнхийшүүлэх чадвар нэмэгддэг. Бид эдгээр шинэ загваруудын (PI-GAT, PI-SAGE) гүйцэтгэлийг суурь PI-GCN загвартай харьцуулан үнэлнэ.

III. Туршилт

Туршилтад Gset цуглуулга [13]-аас доорхи графуудыг ашигласан.

A. Төлөөллийн хэмжээс өөрчлөгүй туршилт

PI-GCN загварт туршилтын үр дүнгийн хувьд төлөөллийн хэмжээсүүд болоод хайперпараметрүүдийг граф тус бүрт шилдэг байхаар үзүүлсэн [9]. Энэ нь дараагийн графд тооцоолоход хүндрэлтэй тул төлөөллийн хэмжээнүүд $d_0 = \sqrt{n}$, $d_1 = d_0/2$, $d_2 = 1$ байхаар туршсан. Энд хувилбар нь PI-GNN хувилбарууд, огтлолын хэмжээ нь сургалтын дүнд олдсон бодлогын шийдийн хариу, эпох нь сургалтын үе ба 100,000 хүрээгүй байгаа нь эрт зогссоныг илтгэнэ (early stopped). Туршсан гол хайперпараметрууд нь суралцах хурд (learning rate)-ийг 0.0001, 0.001, 0.002 болон хасалтын коэффициент (dropout)-ийг 0, 0.1, 0.2, 0.3 байхаар сонгосон.

Хүснэгт I
Gset цуглуулга дахь графуудын товч мэдээлэл, орой нь оройн тоог, ирмэг нь ирмэгийн тоог илтгэнэ

Граф	Орой	Ирмэг
G14	800	4,694
G15	800	4,661
G22	2,000	19,990
G49	3,000	6,000
G55	5,000	12,498

1) Хасалтын Нөлөө: Хасалтын коэффициент ашиглах (хасалтын коэффициент > 0) нь ерөнхийдөө илүү сайн шийд өгч байсан ч, сургалтын үеийн тоо (epoch) болон нийт хугацаа (time) мэдэгдэхүйц нэмэгдсэн. G22 граф дээр PI-GCN загвар хасалтын коэффициент 0.3 үед 24,360 эпох зарцуулан 11,025 огтлолын хэмжээнд хүрсэн бол, хасалтын коэффициент 0 үед 5,221 эпоход 10,804 огтлолын хэмжээнд хүрсэн. G55 граф дээр PI-GAT загвар хасалтын коэффициент 0.3 үед хамгийн их эпохын тоонд (100,000) хүрч 9,754 огтлолын хэмжээтэй байсан бол, хасалтын коэффициент 0 үед 2,281 эпоход огтлолын хэмжээ 8,467 хүрсэн. Хасалтын коэффициент ашиглаагүй үед загварууд илүү хурдан нийлэлт болж байсан ч, ихэнх тохиолдолд бага огтлолын хэмжээтэй шийд олсон. Энэ нь хасалтын коэффициент нь физикээс сэдсэн загваруудад ч шийдлийн орон зайг илүү өргөн хүрээнд судлах, локал минимумд гацахаас сэргийлэхэд тусалдаг боловч тооцооллын зардлыг нэмэгдүүлдэг болохыг харуулж байна.

Хүснэгт II
хасалт хийгээгүй, суралцах хурд 0.0001 байх туршилт

Хувилбар	Граф	Огтлолын хэмжээ	Эпох	Хугацаа
PI-GCN	G14	1,462	8,444	16.16
	G15	1,497	7,777	15.37
	G22	**10,804**	5,221	10.46
	G49	**5,454**	3,569	8.25
	G55	6,027	4,580	12.90
PI-GAT	G14	1,940	2,909	9.99
	G15	2,467	4,332	14.79
	G22	6,119	1,044	3.57
	G49	2,260	1,345	4.72
	G55	**8,467**	2,281	9.44
PI-SAGE	G14	**2,278**	5,189	13.96
	G15	**2,517**	5,496	14.28
	G22	9,936	3,004	8.34
	G49	4,188	3,163	8.98
	G55	7,990	2,246	7.96

2) Загварын Гүйцэтгэл ба Хуваалтын Чанар: Хамгийн өндөр огтлолын хэмжээнд хүрэхэд анхаарлаа хандуулбал

Хүснэгт III
хасалтын коэффициент 0.3, суралцах хурд 0.0001 байх туршилт

Хувилбар	Граф	Огтлолын хэмжээ	Эпох	Хугацаа
PI-GCN	G14	1,939	100,000	199.91
	G15	2,179	100,000	202.52
	G22	11,025	24,360	49.63
	G49	**5,508**	14,501	32.79
	G55	8,200	94,818	267.92
PI-GAT	G14	2,398	100,000	353.29
	G15	2,547	100,000	346.13
	G22	11,335	100,000	359.31
	G49	3,578	100,000	373.17
	G55	**9,754**	100,000	422.75
PI-SAGE	G14	**2,882**	100,000	280.71
	G15	**2,910**	100,000	281.03
	G22	**12,627**	100,000	293.04
	G49	5,424	100,000	305.40
	G55	9,709	88,892	333.51

Хүснэгт V
хасалтын коэффициент 0.3, суралцах хурд 0.002 байх туршилт

Хувилбар	Граф	Огтлолын хэмжээ	Эпох	Хугацаа
PI-GCN	G14	2,955	100,000	194.42
	G15	2,960	100,000	198.97
	G22	10,921	4,964	9.85
	G49	**5,498**	1,376	2.98
	G55	8,558	7,706	20.93
PI-GAT	G14	2,414	100,000	338.19
	G15	2,575	100,000	338.77
	G22	**13,014**	100,000	352.98
	G49	3,700	100,000	370.54
	G55	**9,864**	100,000	420.32
PI-SAGE	G14	**2,977**	100,000	272.27
	G15	**2,964**	100,000	279.03
	G22	12,748	100,000	288.63
	G49	5,408	7,371	21.86
	G55	9,709	8,417	30.54

(ялангуяа хасалтын коэффициент 0.3 үед), PI-GAT болон PI-SAGE загварууд том хэмжээний графууд (G22, G55) дээр ихэвчлэн сайн гүйцэтгэл үзүүлсэн бол, PI-GCN загвар G49 граф дээр өрсөлдөхүйц үр дүн үзүүлсэн.

G22 дээр PI-GAT (огтлолын хэмжээ: 13,046) болон PI-SAGE (огтлолын хэмжээ: 12,748) хамгийн өндөр шийд олсон ч ихэвчлэн хамгийн их эпохын тоо шаардагдсан. G55 дээр PI-GAT (огтлолын хэмжээ: 9,870) болон PI-SAGE (огтлолын хэмжээ: 9,709) хамгийн сайн үр дүнг үзүүлсэн. G49 дээр PI-GCN (хасалтын коэффициент=0.3, бага суралцах хурдтай) 5,508 огтлолын хэмжээтэй шийдэнд хүрч сайн гүйцэтгэлтэй байсан бөгөөд өндөр суралцах хурдтай үед (5,502, 5,498) бага зэрэг буурсан ч хурдан нийлэлт болсон. Эдгээр үр дүнгээс харахад, физикийн зарчмыг анхаарал болон түүвэрлэлт/нэгтгэлттэй хослуулсан PI-GAT болон PI-SAGE архитектур нь сургалтын урт хугацаа зөвшөөрөгдөх нөхцөлд PI-GCN загвараас илүү өндөр чанартай шийд олоход давуу талтай байна.

Хүснэгт IV
хасалтын коэффициент 0.3, суралцах хурд 0.001 байх туршилт

Хувилбар	Граф	Огтлолын хэмжээ	Эпох	Хугацаа
PI-GCN	G14	2,919	100,000	191.77
	G15	2,957	100,000	196.31
	G22	10,899	7,176	14.39
	G49	**5,502**	2,529	5.66
	G55	8,542	15,871	44.83
PI-GAT	G14	2,437	100,000	350.14
	G15	2,588	100,000	345.54
	G22	**13,046**	100,000	424.77
	G49	3,662	100,000	366.66
	G55	**9,870**	100,000	422.82
PI-SAGE	G14	**2,951**	100,000	271.10
	G15	**2,966**	100,000	276.19
	G22	12,726	100,000	282.48
	G49	5,350	12,262	36.63
	G55	9,709	25,114	92.98

3) Суралцах хурдны нөлөө: Суралцах хурд нь нийлэлтийн хурд болон эцсийн шийдлийн чанарт, ялангуяа хасалт өндөртэй үед (хасалтын коэффициент 0.3) чухал нөлөө үзүүлсэн.

G49 граф дээрх PI-GCN хасалтын коэффициент 0.3 үед загварын хувьд суралцах хурдыг 0.0001-ээс 0.002 болгон нэмэгдүүлэхэд эпохын тоо (14,501 -> 1,376) болон сургалтын хугацаа (32.79с -> 2.98с) эрс буурсан бөгөөд огтлолын хэмжээ-д (5,508 -> 5,498) маш бага бууралт ажиглагдсан. G55 граф дээрх PI-SAGE хасалтын коэффициент 0.3 үед загварын хувьд суралцах хурдыг 0.0001-ээс 0.002 болгон нэмэгдүүлэхэд эпохын тоо (88,892 -> 8,417) болон хугацаа (333.51с -> 30.54с) мэдэгдэхүйц буурч, огтлолын хэмжээ (9,709) хэвээр хадгалагдсан. Энэ нь хасалтын коэффициент ашиглах үед өндөр суралцах хурд нь PI-GCN болон PI-SAGE загваруудын хувьд нийлэлтийг хурдасгаж, сургалтын хурд болон шийдлийн чанарын хооронд илүү сайн буулт хийх боломжийг олгож болохыг харуулж байна. PI-GAT болон PI-SAGE нь ихэвчлэн том хэмжээтэй графууд (G22, G55) дээр сайн үр дүн үзүүлсэн бөгөөд энэ нь хөршүүдийн мэдээллийг илүү үр дүнтэй ашиглах чадвартай холбоотой байж болох юм. Ялангуяа PI-GAT анхаарлын механизм нь олон хөрштэй оройн мэдээллийг оновчтой ашигладаг нь том графууд дээр онцгой үр дүн үзүүлдэг болохыг харуулж байна. Туршилтын үр дүнгээс харахад PI-GAT болон PI-SAGE загварууд нь том хэмжээний графуудад огтлолын чанар өндөр байсан боловч тооцооллын зардал өндөртэй байсан. Эсрэгээр, PI-GNN загвар нь зарим граф дээр илүү хурдан нийлэлттэй байсан боловч ерөнхийдөө шийдлийн чанар бага байсан. Эдгээр загваруудыг ашиглахдаа тооцооллын хүчин чадал болон шийдлийн чанарын шаардлагад тохируулан хайперпараметрийн тохиргоог оновчлох хэрэгтэй.

B. PI-GAT ба PI-GATv2 загваруудын харьцуулалт

Бид PI-GAT болон түүний сайжруулсан хувилбар болох PI-GATv2 архитектуруудын гүйцэтгэлийг G22 граф дээр харьцуулсан. Энэхүү туршилтаар суралцах хурд болон хасалтын коэффициентийг янз бүрээр сонгон загваруудын нийлэлтийн хурд, шийдийн чанарыг үнэлсэн.

Туршилтын үр дүнгээс харахад PI-GATv2 загвар нь ихэнх тохиолдолд PI-GAT загвараас илүү өндөр огтлолын хэм-

Хүснэгт VI
PI-GAT ба PI-GATv2 загваруудын огтлолын хэмжээний харьцуулалт (G22 граф, 50,000 эпох)

Суралцах хурд	Хасалт	PI-GAT	PI-GATv2
0.001	0.1	9,759	**12,040**
0.001	0.2	9,987	**10,017**
0.001	0.3	11,314	**11,655**
0.0012	0.1	**11,073**	9,805
0.0012	0.2	9,979	**12,469**
0.0012	0.3	10,168	**11,594**
0.0013	0.1	9,816	**9,856**
0.0013	0.2	**9,997**	9,984
0.0013	0.3	10,975	**11,027**
0.0015	0.1	9,784	**12,226**
0.0015	0.2	11,053	**11,238**
0.0015	0.3	10,130	**11,644**

жээнд хүрч чадсан байна. Ялангуяа суралцах хурд болон хасалтын коэффициентийг оновчтой тохируулсан тохиолдолд PI-GATv2 загварын гүйцэтгэл мэдэгдэхүйц сайжирч байгааг харууллаа.

IV. Дүгнэлт

MaxCut оновчлолын бодлогод зориулан физикээс сэдсэн гурван GNN загвар – PI-GCN (суурь), PI-GAT, болон PI-SAGE–ийн гүйцэтгэлийг янз бүрийн граф болон хайперпараметрийн тохиргоон дор харьцуулан үнэлэв. Туршилтын үр дүн нь шийдийн чанар болон тооцооллын зардал хоёрын хооронд эдгээр загваруудын хувьд ч тодорхой харилцан буулт (trade-off) байгааг харууллаа.

Том хэмжээний графууд (G22, G55) дээр ихэвчлэн хасалтын коэффициент 0.3 үед PI-GAT болон PI-SAGE загваруудд нь суурь PI-GCN загвараас илүү өндөр огтлолын хэмжээтэй шийд олох хандлагатай байсан. Энэ нь анхаарлын механизм (PI-GAT) болон сайжруулсан түүвэрлэлт/нэгтгэлт (PI-SAGE) нь физикийн зарчимтай хослон MaxCut шийдлийн чанарыг сайжруулахад хувь нэмэр оруулж болохыг харуулж байна. Харин G49 граф дээр PI-GCN загвар сайн гүйцэтгэлтэй байсан нь загварын сонголт нь графын шинж чанараас хамаарч болохыг илтгэнэ.

Хамгийн сайн огтлолын хэмжээг авахын тулд PI-GAT эсвэл PI-SAGE загварыг хасалтын коэффициенттай, удаан хугацаанд сургах шаардлагатай байж болно. Хэрэв тооцооллын хурд чухал бол хасалтын коэффициентгүй сургах эсвэл PI-GCN загварыг өндөр суралцах хурдтай ашиглах нь сонголт байж болох ч огтлолын хэмжээ буурах магадлалтай.

Энэхүү судалгаа нь PI-GAT, PI-GATv2 болон PI-SAGE загваруудд нь MaxCut бодлогод ирээдүйтэй болохыг харуулж байна. Цаашид эдгээр загваруудын архитектурыг сайжруулах, физикийн зарчмыг илүү үр дүнтэй нэгтгэх арга замыг судлах нь чухал юм. Мөн өөр төрлийн оновчлолын бодлогуудад эдгээр PI загваруудын ерөнхий чадавхыг үнэлэхэд хэрэгцээтэй.

Ном зүй

[1] Ding, Chris HQ, et al. "A min-max cut algorithm for graph partitioning and data clustering." Proceedings 2001 IEEE international conference on data mining. IEEE, 2001.

[2] Commander, Clayton W. "Maximum cut problem, MaxCut." Encyclopedia of Optimization 2 (2009).

[3] Karp, Richard M. "Reducibility among combinatorial problems." In 50 Years of Integer Programming 1958-2008: from the Early Years to the State-of-the-Art, pp. 219-241. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[4] Wu, Zonghan, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S. Yu. "A comprehensive survey on graph neural networks." IEEE transactions on neural networks and learning systems 32, no. 1 (2020): 4-24.

[5] Scarselli, Franco, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. "The graph neural network model." IEEE transactions on neural networks 20, no. 1 (2008): 61-80.

[6] Veličković, Petar, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. "Graph attention networks." arXiv preprint arXiv:1710.10903 (2017).

[7] Brody, Shaked, Uri Alon, and Eran Yahav. "How attentive are graph attention networks?." arXiv preprint arXiv:2105.14491 (2021).

[8] Hamilton, Will, Zhitao Ying, and Jure Leskovec. "Inductive representation learning on large graphs." Advances in neural information processing systems 30 (2017).

[9] Schuetz, Martin JA, J. Kyle Brubaker, and Helmut G. Katzgraber. "Combinatorial optimization with physics-inspired graph neural networks." Nature Machine Intelligence 4, no. 4 (2022): 367-377.

[10] Schuetz, Martin JA, J. Kyle Brubaker, Zhihuai Zhu, and Helmut G. Katzgraber. "Graph coloring with physics-inspired graph neural networks." Physical Review Research 4, no. 4 (2022): 043131.

[11] Karalias, Nikolaos, and Andreas Loukas. "Erdos goes neural: an unsupervised learning framework for combinatorial optimization on graphs." Advances in Neural Information Processing Systems 33 (2020): 6659-6672.

[12] Toenshoff, Jan, Martin Ritzert, Hinrikus Wolf, and Martin Grohe. "Graph neural networks for maximum constraint satisfaction." Frontiers in artificial intelligence 3 (2021): 580607.

[13] Y. Ye, The Gset Dataset, URL (2003), https://web.stanford.edu/yyye/yyye/Gset/.

Төрийн цахим үйлчилгээг мэдлэгийн графт анхааралтай сүлжээгээр санал болгох туршилт

Мөнхжаргалын Билгүүндалай
МКУТ, МТЭС
Монгол Улсын Их Сургууль
Монгол Улс, Улаанбаатар хот
21b1num1559@stud.num.edu.mn

Гантулгын Сарнай
МКУТ, МТЭС
Монгол Улсын Их Сургууль
Монгол Улс, Улаанбаатар хот
sarnai.g@num.edu.mn

Ганболдын Амарсанаа
МКУТ, МТЭС
Монгол Улсын Их Сургууль
Монгол Улс, Улаанбаатар хот
amarsanaag@num.edu.mn

***Хураангуй*—Графт түшиглэсэн санал болгох системийн судалгаа эрчимжиж бараа бүтээгдэхүүн, сурталчилгаа, эсвэл кино, хөгжим зэрэг зүйлсийг санал болгоход ихэвчлэн нэвтрүүлж байна. Харин төрийн цахим үйлчилгээний санал болгох системийн судалгаа цөөн. Энэ ажлаар Монгол Улсын төрийн цахим үйлчилгээний портал E-Mongolia системийн өгөгдлөөс мэдлэгийн граф байгуулан, мэдлэгийн графт анхааралтай сүлжээ загвар хөгжүүлэн, олон төрлийн өгөгдлөөр мэдлэгийн графаар баяжуулан туршиж үнэлсэн. Туршилтын үр дүнд мэдлэгийн графыг үйлчилгээнд чухал нөлөөлөх агуулгаар баяжуулахад таамаглах хувийг нэмэгдүүлж байгааг олж илрүүллээ. Нийтдээ 50,000 хэрэглэгчийн үйлчилгээ авсан түүхийг ашиглан сургасан тус загварын recall@20, ndcg@20 үнэлгээ нь тус тус 0.3535, 0.3091 утгатай гарсан болно.**

***Түлхүүр үг*—мэдлэгийн граф, санал болгох систем, төрийн цахим үйлчилгээ, мэдлэгийн графт анхааралтай сүлжээ.**

I. Танилцуулга

Санал болгох системийг хөгжүүлэхэд хиймэл оюуныг ашиглах судалгаа эртнээс эхлэлтэй боловч ямар бүтээгдэхүүн, үйлчилгээг яагаад санал болгосныг тайлбарлах чадвар сул байдаг. Харин сүүлийн үед илүү нарийвчилсан үр дүнгээр санал болгохоос гадна тухайн хүнд яагаад тухайн зүйлсийг санал болгосныг тайлбарлах чадвартай болсон [1]. Өгөгдлийг мэдлэгийн графаар дүрслэх нь санал болгох системийн тайлбарлах чадварыг нэмэгдүүлэхээс гадна илүү үр дүнтэй ажилладаг болгосон [2]. Иймд мэдлэгийн графт түшиглэсэн санал болгох судалгаа сүүлийн жилүүдэд эрчимжиж байна [2]–[6].

Нейрон сүлжээний CNN, RNN, GNN, GAN зэрэг архитектурыг санал болгох системд хэрэглэж ном, кино, мэдээ, бараа бүтээгдэхүүн, хөгжим болон нийгмийн сүлжээ зэргийг хэрэглэгчийн сонирхолд нийцүүлэн санал болгодог [7], [8]. Гэсэн хэдий ч төрийн цахим үйлчилгээний салбарт санал болгох систем ашиглах судалгаа харьцангуй ховор байдаг [3], [4]. Төрийн цахим үйлчилгээ нь хэрэглэгчийн сонирхлоос илүүтэйгээр тухайн хэрэглэгчийн хэрэгцээ шаардлагад тулгуурладаг онцлогтой. Энэхүү судалгаагаар Монгол улсын төрийн цахим үйлчилгээний E-Mongolia порталын өгөгдлийн бүрдлийг ашиглан мэдлэгийн граф үүсгэн Мэдлэгийн графт анхааралтай сүлжээ (KGAT) загварыг туршиж цахим төрийн үйлчилгээг санал болгох боломжийг судаллаа. KGAT загвар нь хэрэглэгч болон үйлчилгээнүүдийн хоорондох өндөр эрэмбийн холбоо хамаарлыг граф анхаарлын сүлжээ (GAT) ашиглан оновчтой тооцоолдог [2], [9]. Уламжлалт хамтын шүүлтүүрт суурилсан аргууд зөвхөн хэрэглэгч-үйлчилгээ хоорондын шууд холбоог л авч үздэг [10] бол KGAT нь олон эрэмбийн хөрш оройгоор дамжин, илүү цогц холбоо хамаарлыг тусгаснаар илүү бодит таамаг гаргах боломжийг олгодог [2], [4]. Мөн өөрсдийн загварыг үнэлэхээр KGAT-ыг санал болгоход ашиглагдаж байсан Amazon-book, Last-FM, Yelp2018 өгөгдлүүд дээр туршиж үр дүнг харьцуулсан [2].

Манай судалгаанд E-Mongolia порталын 50,000 хэрэглэгчийн өмнө авсан үйлчилгээний түүхэн өгөгдлийг баяжуулан гурван өөр бүтэцтэй мэдлэгийн граф үүсгэж KGAT загварын гүйцэтгэлд графын бүтэц хэрхэн нөлөөлөхийг шинжилсэн. Уг гурван төрлийн графаар сургагдсан KGAT загваруудын хамгийн сайн үзүүлэлтүүд нь recall@20 болон NDCG@20 хэмжигдэхүүн дээр харгалзан 0.3535 ба 0.3091-д хүрч, Amazon-book, Last-FM, Yelp2018 өгөгдлөөс сургагдсан KGAT загвараас давсан үр дүнг гаргасан. Туршилтын үр дүнгээс харахад, мэдлэгийн граф дахь холбоо хамаарлыг төрөлжүүлж баяжуулах нь KGAT загварын гүйцэтгэлийг сайжруулж байв. Энэхүү судалгаа нь төрийн цахим үйлчилгээний мэдлэгийн граф болон санал болгох KGAT загварын боломжийг харуулж байна.

II. Мэдлэгийн граф

Мэдлэгийн граф нь бодит амьдрал дээрх объектууд болон тэдгээрийн хоорондын холбоо хамаарлыг *[толгой объект]—[холбоо хамаарал]→[сүүл объект]* гэсэн гуравласан бүтцээр илэрхийлдэг чиглэлтэй граф өгөгдлийн бүтэц юм [11]. Жишээлбэл, *[Бат]—[авсан_үйлчилгээ]→[Оршин суугаа газрын тодорхойлолт]* гэсэн гурвал нь иргэн Бат нь Оршин суугаа газрын тодорхойлолт үйлчилгээ авсан гэсэн баримтыг илэрхийлж байна. Мэдлэгийн графын орой нь объектыг, ирмэг нь объектуудын хоорондын холбоо хамаарлыг илэрхийлдэг.

A. Санал болгох загварын асуудал

Төрийн цахим үйлчилгээг санал болгох загварын гол зорилго нь хэрэглэгчид ирээдүйд шаардлагатай болох

магадлалтай үйлчилгээг урьдчилан таамаглаж санал болгох явдал юм. Мэдлэгийн графын үүднээс авч үзвэл энэхүү асуудал нь иргэн болон төрийн цахим үйлчилгээний хооронд үүсэж болох шинэ ирмэг буюу холбоо хамаарлыг илрүүлэх асуудал болж хувирна [2]–[4].

Энэ туршилтын ажлаар дараах хоёр судалгааны асуултад (CA) хариу олохыг зорьсон.

1) **CA1:** Мэдлэгийн граф дахь хэрэглэгч эсвэл үйлчилгээтэй холбогдсон ирмэгүүдийн төрөл нь санал болгох үр дүнд ялгаатай нөлөө үзүүлэх үү?
2) **CA2:** Мэдлэгийн граф дахь холбоо хамаарал буюу ирмэгийн төрөл нэмэгдэх тусам загварын санал болголтын үр дүн сайжрах уу?

B. Хоёр талт граф

Санал болгох системийн үндсэн суурь бүтцүүдийн нэг бол хоёр талт граф (bipartite graph) юм [2], [6]. Энэ нь хоёр төрлийн оройтой граф бөгөөд санал болгох системийн хувьд хэрэглэгч ба үйлчилгээ хоорондын хамаарлыг илэрхийлэхэд өргөнөөр ашиглагддаг. Ийм графын хувьд хэрэглэгчид зөвхөн үйлчилгээтэй, харин үйлчилгээнүүд зөвхөн хэрэглэгчидтэй холбогдох байдлаар хоёр төрлийн оройн хооронд ирмэг үүсгэдэг. Энэ бүтэц нь хэрэглэгчийн өмнө авсан үйлчилгээний түүхийг дүрслэхэд тохиромжтой. Гэсэн хэдий ч энэхүү граф нь хэрэглэгч болон үйлчилгээний нэмэлт шинж чанар, тэдгээрийн хоорондын нарийн хамаарлыг тусгах боломжгүй учир өндөр эрэмбийн холболт болон утга агуулгатай тайлбар шаардсан загваруудад хангалтгүй байдаг. Харин мэдлэгийн графт суурилсан санал болгох загварууд нь хоёр талт графын бүтэц дээр нэмэлт мэдээллийг холбон өргөтгөж, мэдлэгийн граф үүсгэх замаар энэхүү хязгаарлалтыг шийдвэрлэдэг [2].

C. Хамтын мэдлэгийн граф

Хамтын мэдлэгийн граф (collaborative knowledge graph) нь хэрэглэгчийн өмнө авсан үйлчилгээний түүхийг илэрхийлэх хоёр талт графыг түүний оройнуудын шинж чанарыг агуулах мэдлэгийн графтай нэгтгэн өргөтгөсөн нэг төрлийн граф бүтэц юм. Жишээлбэл, төрийн цахим үйлчилгээ санал болгох системийн хувьд хэрэглэгч үйлчилгээний хоёр талт граф дээр үйлчилгээний ангилал, хэрэглэгчийн насны бүлэг зэрэг олон төрлийн оройгоор баяжуулах боломжтой. Санал болгох системийн загварт хамтын мэдлэгийн графыг ашигласнаар хэрэглэгчийн нэг үйлчилгээг авсан мэдээллээр дамжуулан түүнтэй холбогдсон бусад үйлчилгээг санал болгох мөн тайлбарлах чадварыг нэмэгдүүлдэг [2]. Тухайлбал: (Бат - Оршин суугаа газрын тодорхойлолт - лавлагаа - НДШ төлөлтийн талаарх тодорхойлолт) гэсэн өндөр эрэмбийн холболт ашиглан *Бат*ад *НДШ төлөлтийн талаарх тодорхойлолт* гэсэн үйлчилгээг авсан үйлчилгээтэй нь ижил *лавлагаа* төрөлтэй учир санал болгож байна.

III. Цахим үйлчилгээг санал болгох загвар

E-Mongolia портал дахь төрийн цахим үйлчилгээ нь олон төрлийн шинж чанартай холбоотой бөгөөд энэхүү судалгаанд ангилал, үзүүлдэг байгууллага гэсэн шинж чанаруудыг авч үзсэн. Жишээлбэл, ”Оршин суугаа газрын тодорхойлолт” үйлчилгээ нь дараах шинжүүдтэй уялдана:

- үйлчилгээний төрөл - лавлагаа
- үзүүлдэг байгууллага - баг, хороо

Харин хэрэглэгчид нь хүйс, насны ангилал гэсэн хоёр төрлийн шинж чанаруудтай. Эдгээр өгөгдлийг гурвал хэлбэрээр илэрхийлсэн мэдлэгийн граф болон хэрэглэгчийн өмнө авсан үйлчилгээний түүхийг илэрхийлэх хоёр талт графуудыг нэгтгэн хамтын мэдлэгийн граф болгож боловсруулсан. Улмаар уг хамтын мэдлэгийн граф дээр түшиглэн шинээр үүсэж болзошгүй холбоо хамаарал буюу ирмэгийг таамаглах зорилгоор KGAT загварыг хөгжүүлсэн.

Мэдлэгийн графт түшиглэсэн санал болгох системийн загваруудыг ерөнхийд нь гурван төрлөөр ангилж болно [3], [4], [12]. Нэгдүгээрт, хамтын шүүлтүүрт суурилсан арга нь хэрэглэгчтэй ижил зан төлөв бүхий бусад хэрэглэгчийн өмнө нь авсан үйлчилгээний мэдээлэлд үндэслэн санал болгодог. Хоёрдугаарт, агуулгад суурилсан шүүлтүүр нь хэрэглэгчийн өмнө нь авсан үйлчилгээтэй ижил төстэй шинж чанартай үйлчилгээг санал болгодог [12]. Гуравдугаарт, холимог шүүлтүүр нь хоёр аргын давуу талыг хослуулан санал болголтын чанарыг сайжруулахыг зорьдог [3], [4], [12]. KGAT загвар нь хэрэглэгчид тохирох үйлчилгээг санал болгохдоо хэрэглэгчийн өмнө авсан үйлчилгээний түүхийг харгалзан ижил зан төлөв бүхий хэрэглэгчдийн сонголтыг ашиглахаас гадна, мэдлэгийн граф дээр үндэслэн хэрэглэгчийн өмнө сонгосон үйлчилгээнүүдтэй төстэй бусад үйлчилгээг илрүүлж, санал болгодог. Хоёр аргыг хослуулан ашигладаг тул холимог шүүлтүүр загварын төрөлд хамаарна [2].

A. Мэдлэгийн графт анхааралтай сүлжээ

Хамтын шүүлтүүрийн арга нь хэрэглэгч болон үйлчилгээ хоорондын харилцаанд тулгуурлан оновчтой таамаглал гаргах чадвартай боловч хэрэглэгч болон үйлчилгээний шинж чанар, холбоо хамаарлыг илэрхийлэх нэмэлт өгөгдлийг ашигладаггүй [2]–[4]. Үүнээс шалтгаалан хэрэглэгчийн өмнө нь авсан үйлчилгээний түүх хомс буюу шинэ хэрэглэгчид тохирох үйлчилгээг санал болгох, мөн шинэ үйлчилгээ нэмэгдэх үед энэ төрлийн загваруудын гүйцэтгэл буурах сул талтай байдаг. KGAT нь хамтын мэдлэгийн граф ашиглан хэрэглэгч болон үйлчилгээ хоорондын өндөр эрэмбийн холболтуудыг илрүүлж загварын санал болголтын чанарыг сайжруулах замаар дээрх сул талыг шийдвэрлэхээр боловсруулагдсан загвар юм. Энэхүү загвар нь мэдлэгийн граф дахь өндөр эрэмбийн холболт буюу олон орой дамжсан уялдааг анхаарлын механизмын тусламжтайгаар оновчтой тусгаж, хэрэглэгчийн сонирхолд илүү нийцсэн үйлчилгээ санал болгох боломжийг бүрдүүлдэг [2]. Энэхүү арга нь дараах үндсэн 3 давхаргад хуваагдана:

1) Эмбеддинг давхарга: Энэ давхарга нь хэрэглэгч, үйлчилгээ болон хамтын мэдлэгийн граф дахь бусад төрлийн оройнуудыг графын бүтцийг алдагдуулахгүйгээр

тоон вектор болгон буулган машин сургалтад тохиромжтой хэлбэрт оруулдаг. KGAT загвар нь хамтын мэдлэгийн графыг эмбеддинг хийхдээ холбоо хамаарал тус бүрийг ялгаатай орон зайд тусгаарлаж тооцоолдог TransR алгоритмыг ашигладаг [4], [13].

2) Анхааралд суурилсан эмбеддинг дамжуулалтын давхаргууд: Хамтын мэдлэгийн граф дахь оройнуудын эмбеддингийг өндөр эрэмбийн холболтын дагуу давтамжтайгаар дамжуулахын тулд граф конволюцын сүлжээний (GCN) архитектурт суурилдаг. Мөн граф анхаарлын сүлжээний зарчмыг ашигласнаар, дамжуулалтын үе шат бүрд хөрш оройнуудын жинг анхаарлын сүлжээгээр тооцоолж, тухайн холбоосын ач холбогдлыг тодорхойлдог. Анхааралд суурилсан эмбеддинг дамжуулалтын давхарга нь мэдээлэл дамжуулах, мэдлэгт суурилсан анхаарал тооцоолох, мэдээлэл нэгтгэх гэсэн гурван бүрэлдэхүүн хэсгээс бүрдэнэ [4].

- Мессеж дамжуулах: Энэ механизм нь оролтын хамтын мэдлэгийн граф дахь орой бүрийн эмбеддингийг сайжруулах зорилготой бөгөөд тухайн оройд холбогдсон гурвалуудын хүрээнд хөрш оройнуудын мэдээллийг дамжуулан нэгтгэдэг. Орой нь оролтын граф дахь хэд хэдэн гурвалын бүрэлдэхүүнд орсон байдаг. Эдгээр гурвалын хүрээнд тухайн орой нь толгой оройн үүрэг гүйцэтгэж, түүний сүүл оройнууд буюу хөрш оройнуудын мэдээллийг өөртөө хамааруулан дамжуулдаг. Ийнхүү орой бүр холбогдсон гурвалуудаар дамжуулан мэдээлэл хүлээн авч, өөрийн улам эмбеддингийг боловсронгуй болгодог.
- Мэдлэгт суурилсан анхаарал тооцоолох: Хамтын мэдлэгийн граф дахь хөрш орой бүр нь тухайн оройд ижил нөлөө үзүүлдэггүй. Энэ нь хөрш орой тухайн оройн ямар шинж чанарыг илэрхийлж буй болон ямар төрлийн холбоосоор холбогдсон зэрэг хүчин зүйлээс хамаардаг. Иймээс энэ бүрэлдэхүүн хэсэг нь граф анхаарлын сүлжээ зарчмыг ашигласан мэдлэгт суурилсан анхаарлын сүлжээг хэрэгжүүлдэг. Энэ механизмаар хөрш орой тус бүрийн оролтын оройд үзүүлэх нөлөөллийг тооцоолж, харгалзах жинг оноодог. Мессеж дамжуулахдаа эдгээр жинг нь харгалзан үзэж тохирох хэмжээний мэдээллийг уг оройд дамжуулна. Ингэснээр өндөр хамааралтай, ач холбогдол бүхий хөршүүд илүү их нөлөө үзүүлэх боломжтой болно. Жишээлбэл, хэрэв хэрэглэгч тогтмол лавлагаа төрлийн үйлчилгээг авдаг бол, тухайн хэрэглэгчийн хувьд лавлагаа гэсэн үйлчилгээний төрөлтэй холбоотой хамаарлыг илүү өндөр жинтэй гэж тооцно. Үүний үр дүнд хамаарал багатай холбоосуудаас илүү ач холбогдолтойгоор үнэлэгдэж, санал болголтын үйл явцад илүү үр дүнтэй нөлөө үзүүлнэ. Ингэснээр санал болгох үйл явц нь илүү хувь хүнд чиглэсэн санал болголтыг үүсгэх болно.
- Мэдээлэл нэгтгэх: Хөрш оройнууд мессеж дамжуулсны дараа эдгээр нэмэлт мэдээллийг оролтын оройн анхны эмбеддингтэй нэгтгэн шинэ эмбеддинг үүсгэх шаардлагатай. Энэ үйлдлийг мэдээлэл нэгтгэх бүрэлдэхүүн хэсэг гүйцэтгэнэ. Оролтын оройг хөршүүдийнх нь мэдээлэлтэй нэгтгэснээр илүү оновчтой, хэрэглэгчийн шаардлагад нийцсэн таамаглалыг гаргах боломжтой болдог.

Өндөр эрэмбийн дамжуулалт

Хамтын мэдлэгийн граф дахь нарийн уялдаа холбоо, шууд бус хамаарлыг илүү үр дүнтэйгээр санал болголтод тусгахын тулд KGAT загвар нь анхааралд суурилсан эмбеддинг дамжуулалтын давхаргыг давтамжтайгаар олон удаа гүйцэтгэх замаар өндөр эрэмбийн холболтыг үе шаттайгаар загварчилдаг. Эхний давхаргад тухайн оройн эмбеддинг нь түүний шууд хөршүүдийн мэдээллээр баяждаг бол дараагийн давхаргуудад өмнөх шатанд шинэчлэгдсэн эмбеддинг нь дараагийн эрэмбийн (2-hop, 3-hop гэх мэт) хөрш оройнуудын мэдээллийг тусган улам боловсронгуй болно. Энэ рекурсив бүтэц нь зөвхөн 1-р эрэмбийн (1-hop) шууд хөршүүдийн төдийгүй өндөр эрэмбийн (multi-hop) гүнзгий холболтоор дамжсан мэдээллийг үе шаттайгаар хүлээн авч, эмбеддингийг агуулга болон уялдаа холбооны хувьд илүү баялаг, илэрхийлэл сайтай болгох нөхцөлийг бүрдүүлдэг.

3) Таамаглах давхарга: Анхааралд суурилсан эмбеддинг дамжуулалтын давхаргуудыг гүйцэтгэсний дараа тухайн хэрэглэгч болон үйлчилгээг илэрхийлсэн оройн эмбеддинг илүү нарийн, боловсронгуй болж сайжирдаг [3], [4]. KGAT загвар нь бүх давхаргын эмбеддингүүдийн дараалаар нь нэг векторт нэгтгэн илэрхийлдэг. Үр дүнд гарсан векторoос хэрэглэгч болон үйлчилгээний тохирох оноог тооцож, санал болгох таамаглалыг гаргана. Энэ оноо нь тухайн хэрэглэгчийн хувьд уг үйлчилгээ хэр тохиромжтойг илэрхийлэх ба эцсийн санал болголтын зэрэглэл тогтооход ашиглагдана.

IV. Туршилт

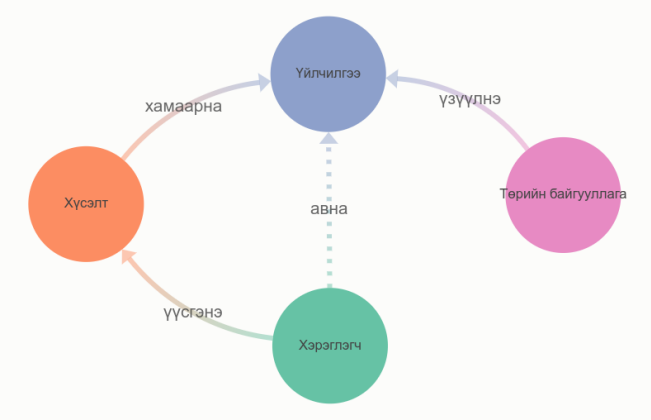
A. Өгөгдлийн бүрдэл

Нийтийн мэдээллийн ил тод байдлын тухай хуулийн 13 дугаар зүйлд заасны дагуу, судалгааны зорилгоор ашиглах өгөгдлийн хүсэлтийг И-Монгол Академи УТҮГ-т тавьж өгөгдлийг хүлээн авсан. Хүсэлтийн дагуу хэрэглэгч, үйлчилгээ, хэрэглэгчийн үйлчилгээний түүх, нэгж үйлчилгээний үйл явц, төрийн үйлчилгээний байгууллагын талаарх нийт таван төрлийн өгөгдлийг судалгаанд зориулан цуглуулсан. Цуглуулсан өгөгдөлд хувь хүнийг тодорхойлох мэдээллийг нэргүйжүүлсэн, мөн системд байгаа үйлчилгээ болон бусад хүснэгтэн өгөгдлийн хувьд системийн ID дугаарыг ч хуурмаг дугаараар хүлээн авсан болно. Иймд хэрэглэгчийн хувийн мэдээлэл цуглуулаагүй болно. Энэ өгөгдлийн суурь дээр Зураг 1-т үзүүлсэн мэдлэгийн графыг бүтээх боломжтой болсон.

Зураг 1 дахь үүсгэх боломжтой ерөнхий загвараас оройн төрөл, ирмэгийн төрөл, гурвалын хэмжээгээр ялгаатай

Хүснэгт I: E-Mongolia порталын өгөгдлийн бүрдэл

Нэр	Мөрийн тоо	Баганын тоо
Хэрэглэгч	50,000	4
Үйлчилгээ	2,620	7
Хэрэглэгчийн авсан үйлчилгээний түүх	2,163,544	7
Нэгж үйлчилгээний үйл явц	1,048,575	4
Төрийн үйлчилгээний байгууллага	656	5



Зураг 1: Мэдлэгийн графын загвар

гурван мэдлэгийн граф үүсгэн туршсан. Туршилтуудыг ашигласан мэдлэгийн графаар нь нэрлэн дараах байдлаар тодорхойлсон:

1) *Иргэн-үйлчилгээ + байгууллага*:
Энэ туршилтад ашигласан мэдлэгийн граф нь зөвхөн нэг төрлийн ирмэг болох *[Үйлчилгээ]—[үзүүлдэг]→[Байгууллага]* холбоог хамаарлыг агуулна.
2) *Иргэн-үйлчилгээ + насны бүлэг*:
Эхний туршилттай адил энэ туршилтад ашиглагдах мэдлэгийн граф нь нэг төрлийн ирмэгтэй. Энэ туршилтад үйлчилгээтэй бус хэрэглэгчтэй холбоотой *[Хэрэглэгч]—[хамаарна]→[Насны бүлэг]* ирмэгийг нэмэлт мэдээлэл болгон мэдлэгийн графт оруулсан. 50,000 хэрэглэгчийн насыг k-means алгоритмаар бүлгийн төвүүдийг олж 27-оос доош, 27→34, 35→41, 42→50, 51→60, 60-аас дээш гэсэн 6 насны бүлэгт хуваасан.
3) *Иргэн-үйлчилгээ + насны бүлэг + байгууллага*:
Сүүлийн туршилт нь өмнөх туршилтуудад илэрхийлэгдсэн бүх холбоо хамаарлуудыг агуулах буюу холбоо хамаарлын төрлийн тоо болон гурвалын тоогоор хамгийн их байна.

B. Үнэлэх аргачлал

Туршилтын үр дүнгээр санал болгох системийн загварын гүйцэтгэлийг үнэлж аль мэдлэгийн графын хувилбар илүү сайн үр дүн үзүүлж байгааг дараах хэмжигдэхүүнүүдээр харьцуулан үнэлсэн:

- Recall@K:
Загвар K ширхэг үйлчилгээ санал болгох үед тухайн хэрэглэгчид санал болгох ёстой байсан үйлчилгээнүүдээс хэд нь багтсан байгааг илэрхийлэх харьцаа. Энэ нь системийн хамрах чадварыг хэмждэг [14].
- Precision@K:
Загварын санал болгосон K ширхэг үйлчилгээний хэдэн хувь нь зөв байсан болохыг илэрхийлдэг. Энэ нь системийн нарийвчлалыг илтгэнэ [15].
- Hit@K:
Санал болгосон K үйлчилгээ дотор ядаж нэг зөв үйлчилгээ орсон эсэхийг шалгадаг үнэлгээний үзүүлэлт. Энэ нь систем хэрэглэгчид хэрэгтэй мэдээллийг ерөнхийд нь хүргэж чадаж байгааг илтгэнэ [10].
- NDCG@K:
Санал болгох ёстой үйлчилгээнүүдэд эрэмбэ тогтоож, K ширхэг үйлчилгээ санал болгох явцад тухайн эрэмбийн дагуу хэр оновчтой санал болгосныг хэмжих үнэлгээ. Зөв үйлчилгээнүүд санал болголтын эрэмбийн дээд хэсэгт байршсан эсэхийг тооцдог. Энэ нь зөвхөн зөв санал болгох чадавхыг бус, тэдгээрийг эрэмбэлэх оновчтой байдлыг харуулна [16].

C. Туршилтын бэлтгэл

1) Өгөгдлийг сургалт болон туршилтад хуваах: Уламжлалт машин сургалтын загварууд (жишээлбэл, объект илрүүлэлт зэрэг) сургалт болон туршилтын өгөгдлийг тодорхой харьцаагаар (ихэвчлэн 80:20, 70:30 гэх мэт) санамсаргүй байдлаар хуваах ашиглах нь түгээмэл бөгөөд өгөгдлийн тархалт ижил үед үр дүнтэй байдаг. Гэсэн хэдий ч хэрэглэгчийн хэрэгцээ, үйлчилгээний сонголт зэрэг цаг хугацаанаас хамаарсан өгөгдлийн хувьд энэхүү санамсаргүй хуваалтын арга нь тохиромжгүй байж болзошгүй юм. Мэдлэгийн графт суурилсан санал болгох системүүд нь хэрэглэгч ба үйлчилгээний хоорондын хамаарлыг хоёр талт граф хэлбэрээр загварчилдаг. Графын өгөгдлийг санамсаргүй байдлаар хуваах нь графын бүтцийг эвдэж, холбооснуудыг задлах замаар загварын сургалтын үр дүнд сөргөөр нөлөөлөх эрсдэлтэй. Иймээс графт суурилсан, цаг хугацаанаас хамаарсан загваруудын хувьд хэрэглэгч бүрийн үйлчилгээний түүхийг цаг хугацааны дараалалд үндэслэн хуваах нь илүү оновчтой шийдэл болдог. Энэ аргын хүрээнд хэрэглэгч бүрийн түүхэн өгөгдлийг тодорхой харьцаагаар хугацааны хувьд хоёр хэсэгт хуваах, эхний хэсгийг сургалт, үлдсэн хэсгийн туршилтад ашиглана. Сургалтын хугацаанд хамаарах графт суурилан сурсан загвар туршилтын хугацаанд хамаарах үйлчилгээг зөв таамаглан санал болгох чадварыг нь үнэлдэг. Судалгааны ажлын хүрээнд бид E-Mongolia порталын өгөгдлөөс үүсгэсэн хоёр талт графыг хугацааны хувьд 80:20 харьцаатайгаар хуваан ашигласан болно. Нөгөө талаар, мэдлэгийн граф нь хугацааны тухайн агшнаас үл хамааран тогтвортой бүтэцтэй тул сургалт болон туршилтын өгөгдөлд тусгайлан хуваагдах

шаардлагагүй. KGAT загвар нь оролтын хувьд хэрэглэгч-үйлчилгээний түүхийг илэрхийлсэн хоёр хэсэг граф (сургалтын ба туршилтын) болон нэмэлт мэдээллийг агуулсан мэдлэгийн графыг хамтад нь ашиглан, хамтын мэдлэгийн графыг үүсгэж загварыг сургадаг.

2) Параметр тохиргоо: E-Mongolia порталын өгөгдөлд суурилан төрийн цахим үйлчилгээ санал болгох загварыг KGAT загвараар сургахдаа үнэлэх болон KGAT өгүүллийн үр дүнтэй харьцуулах зорилгоор KGAT загварыг боловсруулсан судлаачдын Yelp2028, Amazon-book, Last-fm өгөгдлүүдэд суурилсан загвар сургахад ашигласан параметрийн тохиргоог ямар нэг өөрчлөлтгүйгээр байдлаар ашигласан. Үүнд: эмбеддинг хэмжээ 64, learning rate 0.0001, нийт 1000 давталт (epoch), node dropout 0.1, message dropout 0.1, анхаарлын сүлжээ болон мэдлэгийн графын embedding-г идэвхжүүлсэн байна.

D. Үр дүн

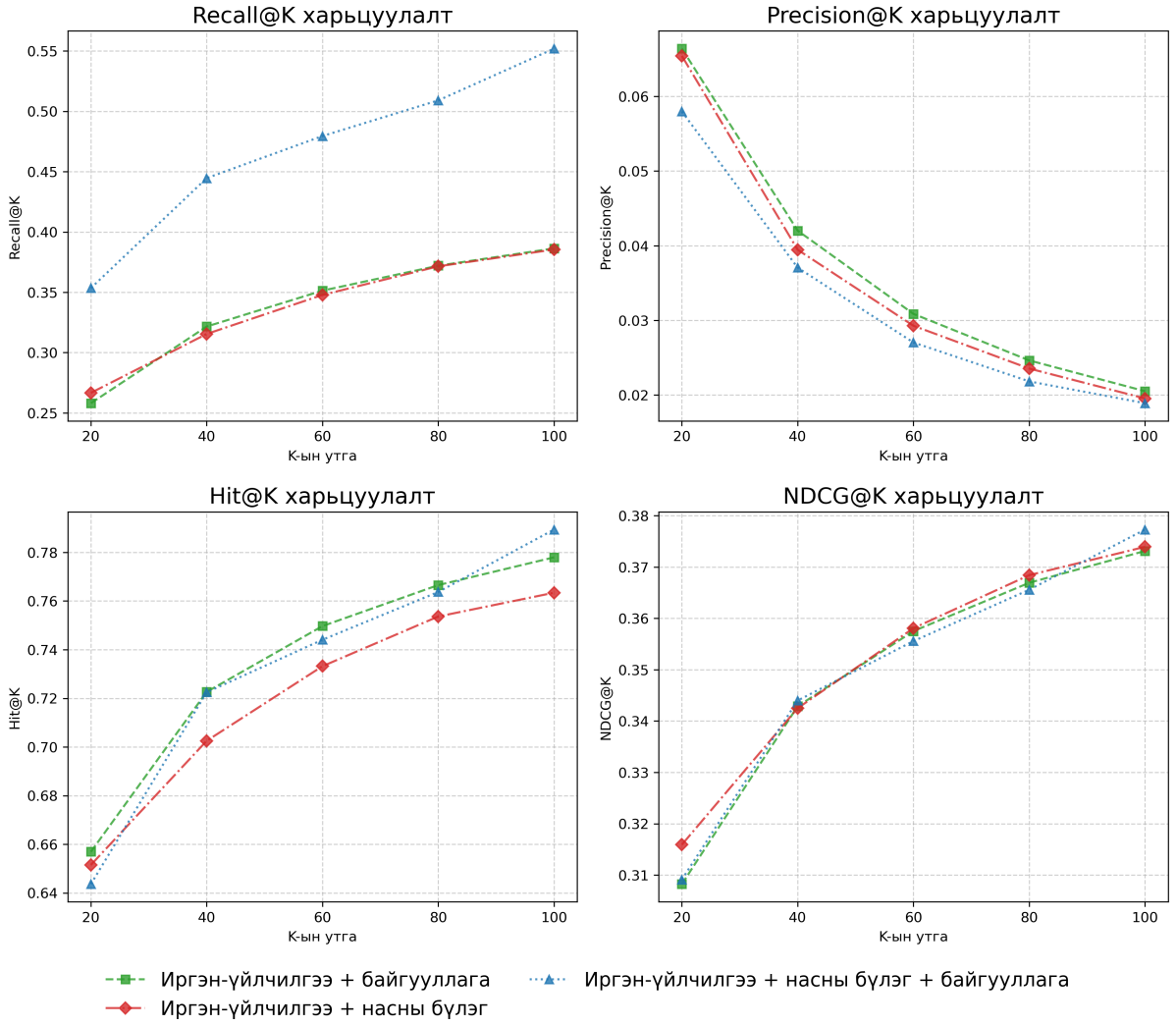
Хэмжигдэхүүнүүдийн k=20 байх үеийн гурван туршилтын үр дүнг Хүснэгт II-д харуулав.

Хүснэгт II: K=20 байх туршилтын үр дүн

Туршилт	Recall	Precision	Hit	NDCG
Иргэн-үйлчилгээ +байгууллага	0.2581	**0.0664**	**0.6569**	0.3082
Иргэн-үйлчилгээ +насны бүлэг	0.2666	0.0654	0.6515	**0.3159**
Иргэн-үйлчилгээ +насны бүлэг+байгууллага	**0.3535**	0.0580	0.6436	0.3091

CA1: Ялгаатай ирмэгийн төрлийн нөлөө

KGAT загвар нь хэрэглэгч ба үйлчилгээний хоорондын хоёр талт граф болон эдгээр оройнуудтай холбогдох нэмэлт мэдээллийг илэрхийлсэн мэдлэгийн графыг нэгтгэх замаар хамтын мэдлэгийн граф үүсгэн санал



Зураг 2: Туршилтын үр дүнгийн харьцуулалт

болгох загварыг сургадаг. Судалгааны эхний хоёр туршилтаар ялгаатай төрлийн ирмэг агуулсан мэдлэгийн графууд санал болгох үр дүнд хэрхэн нөлөөлж байгааг шинжлэв. Туршилтуудын үр дүнг харьцуулан үзэхэд ”иргэн-үйлчилгээ + байгууллага” бүтэцтэй граф нь *recall@k*, *precision@k*, *hit@k* хэмжигдэхүүнүүдийн хувьд илүү байсан бол ”иргэн-үйлчилгээ + насны бүлэг” граф нь *ndcg@k* хэмжигдэхүүний хувьд бага зэрэг илүү үр дүнтэй байсан. Энэ нь E-Mongolia порталын өгөгдөлд суурилсан KGAT загвараар сургасан санал болгох загварын хувьд ялгаатай ирмэгийн төрөл нь ялгаатай нөлөө үзүүлэх ба үйлчилгээ үзүүлдэг байгууллагын мэдээлэл нь хэрэглэгчийн насны бүлгээс илүү нөлөө үзүүлж байгааг харуулж байна. Энэ нь үйлчилгээг үзүүлдэг төрийн байгууллагын мэдээлэл нь тухайн үйлчилгээний ялгарах онцлог шинж чанарыг илүү нарийн илэрхийлж, KGAT загварын анхаарлын сүлжээ тухайн онцлогт төвлөрөх боломжийг нэмэгдүүлж байгаатай холбоотой. Харин хэрэглэгчийн насны бүлэг гэх мэт шинж чанарууд нь олон хэрэглэгчийн хувьд ижил байж болзошгүй тул ялгамж багатай, ерөнхий мэдээлэл болж хувирч, оновчтой таамаглал хийхэд харьцангуй бага хувь нэмэр оруулдаг.

CA2: Ирмэгийн төрөл нэмэгдэх үеийн нөлөө

Гурав дахь *Иргэн–үйлчилгээ + насны бүлэг + байгууллага* туршилт нь граф дахь оройнуудын хоорондын холбоо хамаарал буюу ирмэгийн төрөл мөн гурвалын тоогоор өмнөх туршилтуудаас илүү юм. Энэ туршилтад хоёр өөр төрлийн ирмэг (*[үйлчилгээ]—[үзүүлдэг]→[байгууллага]*, *[хэрэглэгч]—[хамаарна]→[насны бүлэг]*) бүхий мэдлэгийн графыг ашигласан бөгөөд, зөвхөн нэг төрлийн ирмэгтэй мэдлэгийн граф дээр суурилсан загваруудын гүйцэтгэлтэй харьцуулан, олон төрлийн ирмэг бүхий мэдлэгийн графыг ашиглах нь санал болгох загварын үр дүнд хэрхэн нөлөөлж байгааг судлах зорилготой. Судалгаанд ашиглагдсан мэдлэгийн графуудын хэмжээг Хүснэгт III-т харуулав.

Хүснэгт III: Туршилтуудын мэдлэгийн графын хэмжээ

Туршилт	Оройн тоо	Холбоосын төрөл	Гурвалын тоо
Иргэн-үйлчилгээ+байгууллага	3,227	1	2,620
Иргэн-үйлчилгээ+насны бүлэг	50,006	1	50,000
Иргэн-үйлчилгээ+насны бүлэг+байгууллага	53,068	2	52,620

Туршилтын үр дүнгээс харахад мэдлэгийн граф дахь холбоо хамаарал буюу ирмэгийн төрөл нэмэгдэхэд *recall@k* хэмжигдэхүүний утга өссөн нь KGAT загварыг ашигласан санал болгох системийн загварын санал болголтын хамрах хүрээ сайжирч байгааг харуулж байна. Гэсэн хэдий ч *precision@k* хэмжигдэхүүний хувьд зөвхөн нэг төрлийн ирмэгтэй графыг ашигласан загварууд өндөр нарийвчлалтай байсан нь ирмэгийн төрөл нэмэгдэхэд зөв санал болголтын эзлэх хувь бага зэрэг буурч болзошгүйг харуулж байна.

Өгөгдлийн харьцуулалт

KGAT загварыг боловсруулсан судлаачид Amazon-book, Last-FM, Yelp2018 гэсэн гурван нээлттэй өгөгдлөөс хамтын мэдлэгийн граф үүсгэж, өөрсдийн боловсруулсан загварын дагуу хиймэл оюуны санал болгох загваруудыг сурган recall@20, ndcg@20 хэмжигдэхүүнүүдээр үнэлж харьцуулсан [2]. Харин манай судалгаанд ижил KGAT загварыг ижил параметрийн тохиргоотойгоор E-Mongolia порталын өгөгдлөөс үүсгэсэн хамтын мэдлэгийн граф дээр сургаж, гүйцэтгэлийг нь дээрх нээлттэй эхийн өгөгдөлд суурилсан загваруудын үр дүнтэй харьцуулсан бөгөөд Хүснэгт IV-д үзүүлэв.

Хүснэгт IV: Өгөгдлүүдийн үр дүнгийн харьцуулалт

	Amazon-book	Last-FM	Yelp2018	E-Mongolia
recall@20	0.1489	0.0870	0.0712	**0.3535**
ndcg@20	0.1006	0.1325	0.0867	**0.3091**

Хүснэгтээс харахад E-Mongolia өгөгдөлд сургасан загварын ncdg@20 үзүүлэлт нь бусад өгөгдлийн санд суурилан сургасан загваруудтай харьцуулахад хамгийн өндөр байгаа нь санал болгосон үйлчилгээний эрэмбийг хэрэглэгчийн хэрэгцээнд нийцүүлэн оновчтой тодорхойлж байгааг харуулж байна. Мөн recall@20 үзүүлэлт нь өндөр байгаа нь хэрэглэгчид хамааралтай үйлчилгээнүүдийг хангалттай хэмжээнд санал болгож байгааг илтгэж байна. Эдгээр үр дүн нь E-Mongolia порталын өгөгдлөөс үүсгэсэн мэдлэгийн граф нь сайн холбогдсон, мэдээлэл сайн байгаагаас шалтгаалан KGAT загварын анхаарлын сүлжээ хэрэгжүүлэн өндөр эрэмбийн холболтуудыг илрүүлэхэд тохиромжтой болж загварын үр дүнд нөлөөлж байгааг харууллаа.

V. Холбоотой ажил

Туршилтад ашиглах загварыг сонгох явцад Мэдлэгийн графт анхааралтай сүлжээ (KGAT) аргаас гадна дараах загваруудыг судалсан. Эдгээр аргачлалууд нь тус бүр нь мэдлэгийн графт түшиглэсэн санал болгох загварын ялгаатай сул талыг шийдвэрлэхээр загварчлагдсан.

Мэдлэгийн графт түшиглэсэн гүн сүлжээ

Уламжлалт мэдээ санал болгох аргууд нь мэдээний гарчиг дахь объектууд болон агуулагдаж буй утга зүйг хангалттай хэмжээнд илрүүлж чаддаггүйгээс болж гарчгуудын хоорондын холбоог зөв тодорхойлох чадвар нь хязгаарлагддаг. Мэдлэгийн графт түшиглэсэн гүн сүлжээ арга (DKN) нь нэмэлт мэдлэгийн граф ашиглан мэдээний гарчиг дахь үгийн утга зүйн талаарх ойлголтыг сайжруулдаг агуулгад суурилсан гүн сургалтын арга юм. Энэ нь санал болгох системийн загварт мэдээний утга зүйг илүү сайн тусгаж мэдээ санал болголтын нарийвчлалыг нэмэгдүүлдэг [5].

Олон үүрэгт шинж чанарын сургалт

Мэдлэгийн графт түшиглэсэн санал болгох загвар сургах хамтын шүүлтүүрт суурилсан арга нь хэрэглэгч өөртэйгөө төстэй бусад хэрэглэгчдийн өмнө нь сонирхсон зүйлсийг сонирхох магадлалтай гэж үздэг. Гэвч энэхүү арга нь хүйтэн эхлэл буюу өгөгдөл хангалтгүй үед үр дүн багатай байх асуудалтай тулгардаг [3], [4], [6]. Энэхүү асуудлыг

шийдвэрлэхийн тулд олон үүрэгт шинж чанарын сургалт (MKR) арга боловсруулагдсан бөгөөд хэрэглэгч болон зүйлсийн түүх болон мэдлэгийн граф хоёроос нь сурах замаар дурдсан асуудлыг багасгах зорилготой [6].

Дээр дурдсан аргууд тус бүр мэдлэгийн графт суурилсан санал болгох системийн тодорхой сул талыг шийдвэрлэхэд чиглэгдсэн. DKN арга нь мэдээний утга зүйг илүү сайн тусгах замаар агуулгын ойлголтыг сайжруулж, гарчгуудын хоорондын утга зүйн холбоог гүн сүлжээгээр дамжуулан илүү нарийн тодорхойлохыг зорьдог бол MKR арга нь хүйтэн эхлэл, өгөгдөл хомс нөхцөлд хэрэглэгчийн болон зүйлсийн түүхийн мэдээллийг мэдлэгийн графтай хамтад нь сургах замаар гүйцэтгэлийг сайжруулдаг. Харин KGAT арга нь зөвхөн хэрэглэгч болон үйлчилгээ хоорондын шууд харилцан үйлчлэлд тулгуурлахаас илүүтэйгээр хамтын мэдлэгийн граф дахь өндөр эрэмбийн холболтуудыг граф анхаарлын механизмаар тооцож, холбогдох хөршүүдээс илүү ач холбогдолтой мэдээллийг шүүж авах боломжийг олгодог. Ийнхүү KGAT нь хэрэглэгч болон үйлчилгээний хоорондын олон түвшний, нарийн уялдааг тусган, хувь хүнд илүү нийцсэн санал болголтыг хийх боломжийг бүрдүүлдгээрээ онцлогтой.

VI. Дүгнэлт

Энэхүү судалгаагаар мэдлэгийн графт суурилсан санал болгох системийн загвар сургахад ашиглах мэдлэгийн графыг өргөтгөж, хэрэглэгч болон үйлчилгээ хоорондын уялдаа холбоог илүү нарийвчлалтай тайлбарлах тусам загварын гүйцэтгэл сайжирч, хэрэглэгчид тохирох үйлчилгээг илүү үр дүнтэй таамаглах боломж нэмэгдэж байгааг туршилтаар баталсан. Судалгааны үр дүн дэх хамгийн сайн гүйцэтгэлтэй KGAT загвар нь Recall@20 үзүүлэлт 0.3535 буюу хэрэглэгч бүрийн хувьд системийн санал болгосон 20 үйлчилгээ дотор тухайн хэрэглэгчид хамааралтай үйлчилгээний 35.35 хувийг дунджаар зөв таамаглаж байгааг илтгэнэ. Өөрөөр хэлбэл, хэрэглэгчийн ирээдүйд авах боломжтой бүх үйлчилгээнүүдээс тодорхой хэсгийг систем амжилттай санал болгож байгааг илэрхийлэх хэмжигдэхүүн юм. Мөн E-Mongolia порталын өгөгдлөөр хөгжүүлсэн KGAT санал болгох загварын хувьд мэдлэгийн граф нь хэрэглэгчээс илүүтэй үйлчилгээг тодотгосон нэмэлт мэдээллийг агуулах нь үр дүнд эерэг нөлөө үзүүлэх нь харагдсан. Цаашид санал болгох загварын сургалтад ашиглах мэдлэгийн графыг зөв холбоосоор баяжуулж загварын гүйцэтгэлийг улам сайжруулах боломжтой.

Талархал

Энэхүү судалгааны ажлыг хийж гүйцэтгэхэд шаардлагатай өгөгдлийг гаргаж өгч тусалсан, төр, хувийн хэвшлийн байгууллага болон их сургууль, эрдэм шинжилгээний байгууллагууд өгөгдөлд түшиглэсэн хамтын ажиллагааг дэмжсэн И-Монгол Академи УТҮГ-ын удирдлага, хамт олонд талархаж байна.

Зүүлт

[1] Y. Zhang, X. Chen *et al.*, "Explainable recommendation: A survey and new perspectives," *Foundations and Trends® in Information Retrieval*, vol. 14, no. 1, pp. 1–101, 2020.

[2] X. Wang, X. He, Y. Cao, M. Liu, and T.-S. Chua, "Kgat: Knowledge graph attention network for recommendation," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery Data Mining*, ser. KDD '19. ACM, Jul. 2019, p. 950–958. [Online]. Available: http://dx.doi.org/10.1145/3292500.3330989

[3] Q. Guo, F. Zhuang, C. Qin, H. Zhu, X. Xie, H. Xiong, and Q. He, " A Survey on Knowledge Graph-Based Recommender Systems ," *IEEE Transactions on Knowledge & Data Engineering*, vol. 34, no. 08, pp. 3549–3568, Aug. 2022. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/TKDE.2020.3028705

[4] J. Liu and L. Duan, "A survey on knowledge graph-based recommender systems," in *2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 2021, pp. 2450–2453.

[5] H. Wang, F. Zhang, X. Xie, and M. Guo, "Dkn: Deep knowledge-aware network for news recommendation," 2018. [Online]. Available: https://arxiv.org/abs/1801.08284

[6] H. Wang, F. Zhang, M. Zhao, W. Li, X. Xie, and M. Guo, "Multi-task feature learning for knowledge graph enhanced recommendation," 2019. [Online]. Available: https://arxiv.org/abs/1901.08907

[7] S. Zhang, L. Yao, A. Sun, and Y. Tay, "Deep learning based recommender system: A survey and new perspectives," *ACM computing surveys (CSUR)*, vol. 52, no. 1, pp. 1–38, 2019.

[8] S. Wu, F. Sun, W. Zhang, X. Xie, and B. Cui, "Graph neural networks in recommender systems: a survey," *ACM Computing Surveys*, vol. 55, no. 5, pp. 1–37, 2022.

[9] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," *arXiv preprint arXiv:1710.10903*, 2017.

[10] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," in *Proceedings of the 26th International Conference on World Wide Web*. ACM, 2017, pp. 173–182.

[11] G. Tamašauskaitundefined and P. Groth, "Defining a knowledge graph development process through a systematic review," *ACM Trans. Softw. Eng. Methodol.*, vol. 32, no. 1, Feb. 2023. [Online]. Available: https://doi.org/10.1145/3522586

[12] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE transactions on knowledge and data engineering*, vol. 17, no. 6, pp. 734–749, 2005.

[13] Y. Lin, Z. Liu, M. Sun, Y. Liu, and X. Zhu, "Learning entity and relation embeddings for knowledge graph completion," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 29, no. 1, 2015.

[14] P. Cremonesi, Y. Koren, and R. Turrin, "Performance of recommender algorithms on top-n recommendation tasks," in *Proceedings of the fourth ACM conference on Recommender systems*, 2010, pp. 39–46.

[15] G. Shani and A. Gunawardana, "Evaluating recommendation systems," *Recommender systems handbook*, pp. 257–297, 2011.

[16] K. Järvelin and J. Kekäläinen, "Cumulated gain-based evaluation of ir techniques," *ACM Transactions on Information Systems (TOIS)*, vol. 20, no. 4, pp. 422–446, 2002.

Өсөн Нэмэгдэх Стохастик Блок Загвар

Г.Гантулга
Мэдээллэл, Компьютерийн ухааны тэнхим
Мэдээлэл Технологи, Электроникийн Сургууль
МУИС
Улаанбаатар, Монгол улс
gantulga_g@num.edu.mn

П.Далайжаргал*
Мэдээллэл, Компьютерийн ухааны тэнхим
Мэдээлэл Технологи, Электроникийн Сургууль
МУИС
Улаанбаатар, Монгол улс
dalaijargal@num.edu.mn

***Хураангуй*---Энэхүү судалгаагаар бид бүлэг бүтэц дээр суурилсан Стохастик Блок Загвар ба преференциал холболтын (PA) механизмыг хослуулсан, Өсөн Нэмэгдэх Стохастик Блок загвар (GSBM) гэх цоо шинэ загварыг санал болгож байна. Уг загвар нь том хэмжээний нийгмийн болон мэдээллийн сүлжээнд түгээмэл ажиглагддаг бүлэг бүтэц болон хүнд сүүлтэй зэрэг тархалтыг үзүүлэх сүлжээг үр дүнтэйгээр байгуулах боломжийг олгоно. Сонгодог стохастик блок загвар (SBM) нь тогтсон оройн олонлогуудыг хооронд нь холбохдоо статик магадлалын дүрэм дээр суурилдаг бол, GSBM загвар нь сүлжээний динамик өсөлтийн механизмыг авч үздэгээрээ ялгаатай юм. Бид дэвшүүлж буй загвараа албан ёсоор тодорхойлж, хиймлээр үүсгэсэн сүлжээний шинж чанарыг аналитик байдлаар судалсан. Энэхүү эрлийз арга нь SBM загварын гол шинж болох бүлэг доторх нягт холболт болон бүлэг хоорондын сийрэг холболтыг хадгалахын зэрэгцээ, преференциал холболтын загваруудад түгээмэл ажиглагддаг нөлөө бүхий "төв зангилаануд"ыг (хабууд) аяндаа үүсгэж, улмаар бодит сүлжээнүүдэд түгээмэл байдаг шинжүүдийг хадгалах хиймэл сүлжээ байгуулдаг болохыг симуляци туршилтаар батлав. Эдгээр үр дүнгүүд нь GSBM загвар нь бүлэг илрүүлэх алгоритм, тархалтын динамик, мөн төвөгтэй бөгөөд тасралтгүй хувьсан өөрчлөгдөж буй сүлжээнүүдийн тогтвортой байдлыг судлахад илүү бодит, үр дүнтэй аргачлал болохыг харуулж байна.**

***Түлхүүр үг*---Нийгмийн сүлжээ, бүлэг бүтэц, преференциал холболт, стохастик блок загвар**

I. Удиртгал

Бодит нийгмийн сүлжээний шинж чанарууд бүхий хиймэл сүлжээ үүсгэх нь өнөөгийн тооцооллын нийгмийн шинжлэх ухаан, мэдээллийн тархалтын судалгаа болон бүлэг илрүүлэх, холбоосыг таамаглах зэрэг алгоритмуудын үнэлгээний чухал бүрэлдэхүүн хэсэг болж байна [1]--[3]. Шигүү, нутгийн кластерчлал, богино дундаж замын урт, зэрэг тархалтын хүнд сүүлтэй (heavy-tailed power law) байдал зэрэг бодит ертөнцийн сүлжээний гол шинж чанаруудыг тусгасан синтетик сүлжээг үүсгэснээр судлаачид хувийн, нууцлалтай эсвэл дуу чимээтэй (ноизи) өгөгдлөөс хараат бусаар таамаглалаа системтэй шалгаж, алгоритмуудыг илүү нарийн харьцуулан үнэлэх боломжтой болно [4]. Түүнчлэн хянах боломжтой хиймэл сүлжээ ашиглан бодлогын ойролцоо сүлжээний хувьсалд хэрхэн нөлөөлөх зэрэг "хэрэв ингэвэл яах вэ" гэх зохиомол нөхцөл байдлуудыг туршиж үзэх боломжийг олгож, том хэмжээний онлайн нийгэмлэгүүдийн үүсэл, динамикийн ар дахь механизмуудыг илүү гүнзгий ойлгоход тусалдаг [5], [6].

Сүлжээний бүлэг бүтцийг загварчлах хамгийн чухал аргуудын нэг бол бүлгүүдэд хуваагдсан сүлжээг дүрслэх зорилгоор анх санал болгосон Стохастик блок загвар (SBM) юм [8], [26]. Сонгодог SBM загварт зангилаануудыг урьдчилан тодорхойлсон бүлгүүдэд хуваах, ирмэг үүсэх магадлал нь зангилаануд аль бүлэгт хамаарахаас шууд хамаардаг. Энэ арга нь бүлэг дотор нягт холболт, бүлэг хоорондсийрэг холболт үүсгэж, бүлгийн бүтэц бүхий сүлжээг үр дүнтэй дүрсэлдэг. Гэвч уламжлалт SBM загвар нь зангилааны олонлогийг тогтмол, бүлгийн хуваарилалтыг статик гэж үздэг тул динамик өсөлт болон бодит нийгмийн сүлжээнд түгээмэл ажиглагддаг зэрэг тархалтын хүнд сүүлтэй байдлыг хангалттай загварчилж чаддаггүй [9], [10].

Үүний эсрэгээр, преференциал холболт (PA) загварууд нь сүлжээний өсөлтийг зангилаануудын зэрэгтэй шууд хамааралтайгаар шинэ зангилаануд холбогддог "баяжсан нь улам баяждаг" зарчмын дагуу дүрсэлдэг [12], [28]. Барабаши ба Альбертаас анх санал болгосон PA загвар нь зэрэг тархалт нь хүнд сүүлтэй (scale-free) байдаг, өндөр нөлөөтэй "төв зангилаануд" (хабууд) болон богино дундаж зам зэрэг бодит сүлжээний чухал шинж чанаруудыг аяндаа үүсгэдэг [2]. Гэвч сонгодог PA загварууд нь бүлэглэлийн бүтцийг тодорхой авч үздэггүй тул бодит нийгмийн болон мэдээллийн сүлжээнүүдэд нийтлэг байдаг зохион байгуулалттай бүлгийн бүтцийг сайн тусгаж чаддаггүй [10].

Сүүлийн үеийн зарим судалгаанууд нь SBM загварт зэрэг тархалтын ялгаатай байдал, бүлгийн давхар гишүүнчлэл, шаталсан бүтцийн зэрэг илүү бодит нарийн төвөгтэй байдлуудыг тусгах замаар дээрх загваруудын сул талыг нөхөхөөр ажиллаж байна. Жишээлбэл, холимог гишүүнчлэлтэй SBM (MMSBM) загвар нь зангилааг олон бүлэгт нэгэн зэрэг харьяалуулж болох байдлаар дүрсэлснээр нийгмийн болон хамтын ажиллагааны сүлжээнүүдэд загварчлах боломжийг нэмэгдүүлдэг [13]. Шаталсан SBM загвар нь бүлгүүдийг олон түвшний шатлал хэлбэрээр дүрсэлж, сүлжээний бүтцийг өөр өөр масштабуудад судлах боломж олгодог [14]. SBM-ийн нэг нөлөө бүхий өөрчлөлт болох зэрэг залруулгатай SBM (DCSBM) нь зангилаа тус бүрийн зэрэгтэй холбоотой параметр ашиглан хүнд сүүлтэй зэрэг тархалтыг илүү бодитойгоор тусгаж чаддаг [9].

Гэвч DCSBM нь зангилаа болон бүлгийн харьяаллыг статик гэж үзэн зэрэг залруулгыг зангилааны түвшинд хийдэг тул сүлжээний бодит өсөлтийн процессыг харуулж чаддаггүй. Харин бидний санал болгож буй Өсөн нэмэгдэх стохастик блок загвар (GSBM) нь зангилаануудын динамик ирэлт болон преференциал холболтыг тодорхой тусгаж, бүлгүүдийн хэмжээг зэрэг хуваарилалтаар бодитой өсгөж, төв зангилаануудыг аяндаа бий болгодгоороо ялгаатай юм [15]--[18].

Динамик SBM загварууд нь мөн сүлжээний хувьслыг дүрсэлдэг боловч ихэвчлэн тогтвортой зангилааны олонлогийн бүлгийн гишүүнчлэлийн шилжилт, өөрчлөлтийг хянахад төвлөрдөг [19], [20]. Харин бидний GSBM загвар нь сүлжээний өсөлт, шинэ зангилааны тасралтгүй ирэлтийг бүлгийн хэмжээний зэрэг хуваарилалт болон преференциал холболтоор удирддаг. Энэхүү ялгаа нь хурдацтай тэлж буй бодит нийгмийн сүлжээнүүдийн шинж чанарууд болох зэрэг тархалтын хүнд сүүл, төв зангилаануудыг илүү бодитойгоор тусгах боломж олгож байна. Иймээс манай GSBM нь бодит сүлжээний хэмжээ болон бүлгийн бүтцэд цаг хугацааны явцад гарах ихээхэн өөрчлөлтүүдийг загварчлахад илүү тохиромжтой юм.

Цаашид уг өгүүллийн бүтэц дараах байдлаар үргэлжилнэ. 2-р хэсэгт санал болгож буй GSBM загварыг дэлгэрэнгүй тайлбарлана. 3-р хэсэгт туршилтын үр дүнг танилцуулна. 4-р хэсэгт дүгнэлт болон цаашдын судалгааны чиглэлийг хэлэлцэнэ.

SBM ба PA загваруудын давуу талуудыг нэгтгэсэн манай GSBM загвар нь том хэмжээний нийгмийн болон мэдээллийн сүлжээнүүдийн бүтэц, динамикийг илүү бодитоор дүрслэх боломж олгоно.

II. Аргачлал

Энэ хэсэгт бид Өсөн нэмэгдэх стохастик блок загвар (GSBM)-ыг албан ёсоор танилцуулж, түүний преференциал холболтын (PA) механизмыг дэлгэрэнгүй тайлбарлана. Энэхүү загвар нь үндсэндээ дараах гурван үе шаттай хэрэгжинэ: (1) Эхлүүлэх үе шат, (2) Бүлэг доторх ирмэгүүдийг үүсгэх үе шат, (3) Бүлгүүдийн хоорондох ирмэгүүдийг үүсгэх үе шат. Алгоритм 1-д GSBM алгоритмын хуурмаг кодыг харуулав.

A. Эхлүүлэх үе шат (Initialization)

GSBM алгоритмын *generate-community* үйлдэл зэрэг тархалтыг ашиглан 1-ээс n оройнуудыг дугаараар нь дараалан k ширхэг бүлэгт хуваарилна. Хуваарилахдаа бүлэгт тус бүрд $n\frac{1}{i^{\alpha}}$ гэсэн зэрэг тархалтыг ашигласан. Энд i тухайн бүлгийн дугаар бөгөөд бүлэг хоорондын хэмжээний зөрөөг бага байлгах үүднээс $\alpha = 1.2$ гэж сонгосон. Бүлгийн хэмжээний нийлбэр $\sum |C_i| = n$ байхад анхаарах бөгөөд хамгийн жижиг бүлгийн хэмжээ ядаж m_2 байна. Оройн хуваарилалтыг хийсний дараа $C = \{C_1, C_2, \ldots, C_k\}$-г буцаана.

Algorithm 1 GSBM алгоритм

GSBM(n, k, m_1, m_2)
1: n оройтой G граф үүсгэ.
2: $G.E \leftarrow \{\emptyset\}$
3: $C \leftarrow$generate-community(n, k)
4: **for** $C_i \in C$
5: $S \leftarrow$initialize-community(G, C_i, m_2)
6: $C'_i \leftarrow C_i / S$
7: **for** $u \in C'_i$
8: **for** $v \in S$
9: $w[v] \leftarrow G.deg[v]/(2 \cdot |G.E_{C_i}|)$
10: $e \leftarrow$ 1-ээс $2m_2$ хооронд санамсаргүй сонгосон утга.
11: $R \leftarrow \{w$ жинтэй магадлалаар S олонлогоос сонгосон e ширхэг орой$\}$
12: **for** $v \in R$
13: $G.E \leftarrow G.E \cup \{(u, v)\}$
14: $G.deg[u], G.deg[v]$ тус бүрийг нэгээр нэмэгдүүлнэ
15: $S \leftarrow S \cup \{u\}$
16: **for** $C_i \in C$
17: $V' \leftarrow G.V / C_i$
18: **for** $u \in C_i$
19: **for** $v \in V'$
20: $w[v] = G.deg[v]/|G.E_{V'}|$
21: $e \leftarrow$ 0-ээс $2m_1$ хооронд санамсаргүй сонгосон утга.
22: $R \leftarrow \{w$ жингээр, V' олонлогоос санамсаргүйгээр сонгосон e ширхэг орой$\}$
23: **for** $v \in R$
24: $G.E \leftarrow G.E \cup \{(u, v)\}$
25: $G.deg[u], G.deg[v]$ тус бүрийг нэгээр нэмэгдүүлнэ
26: **return** G

B. Бүлэг доторх ирмэгүүдийг үүсгэх үе шат (Intra-block Edge Generation)

Энэ үе шатаар зөвхөн бүлэг доторх ирмэгүүдийг үүсгэн сүлжээнд нэмнэ. Эхлээд бүлэг ямар ч холбоосгүй байх тул эхний m_2 ширхэг оройг зэрэг тархалтыг ашиглан холбоно. Үүнийг хийхдээ *initialize-community* үйлдлээр хийх ба уг функц нь бүлгийн i-р оройг бусад $max(1, m_2/2^{i-1})$ ширхэг оройтой холбож ирмэг үүсгэнэ. Эхний m_2 орой тус бүр ядаж нэг ширхэг оройтой холбогдсон байна. Оройнуудыг хооронд нь холбохдоо оройн дугаараар, тухайн оройгоос хойно байрлах оройнуудыг сонгон авсан.

Эхний m_2 ширхэг орой холбогдсоны дараагаар оройнуудыг дараалаар нь сонгон авч орой тус бүрийн хувьд преференциал холболт зарчмаар өмнөх e ширхэг оройтой, жинтэй магадлалаар холбоно (Алгоритм 1, мөр 7--15). e-ийн утгыг 1-ээс $2m_2$ хооронд санамсаргүйгээр сонгосон бөгөөд энэ нь шинээр холбогдож буй оройнуудын зэрэг янз бүрийн тархацтай байхад нөлөөлнө. Магадлалын жинг тооцохдоо тухайн оройн зэргийг, бүлэгт байх ирмэгүүдийн

тоог 2-оор үржсэн үржвэрт хувааж жинг тооцсон.

C. Бүлгүүдийн хоорондох ирмэгүүдийг үүсгэх үе шат (Inter-block Edge Generation)

Сүлжээний орой бүр өөрийн агуулагдаж буй бүлгээсээ гадна ихдээ $2m_1$ ширхэг оройнуудтай холбогдоно (Алгоритм 1, мөр 16--25). Хэдэн ширхэг холбоос үүсгэхийг $e \in [0, 2m_1]$ утгаар санамсаргүйгээр сонгоно. Тухайн оройтой холбогдох e ширхэг оройг мөн преференциал холболтын зарчмаар үүсгэх бөгөөд холбогдох оройнуудыг мөн өмнөх алхамтай адил зарчмаар жинтэй магадлалаар сонгоно.

Бүлэг дотор болон бүлгүүдийн хоорондын преференциал холболтын механизмуудыг нэгтгэснээр GSBM алгоритм нь бүлэг бүрийн дотоод холболтыг нягт хадгалахын зэрэгцээ, нийт сүлжээний бүтцийг бүлгүүдийн хооронд ховор боловч хүчтэй холболтыг бий болгодог. Ингэснээр бодит ертөнцийн олон сүлжээнүүдэд ажиглагддаг бүлэглэсэн болон шаталсан (hierarchical) бүтцийг илүү бодитоор дүрслэх боломжийг олгодог.

III. Туршилт

Бид өөрсдийн боловсруулсан GSBM загварыг орчин үеийн алгоритмууд болох Barrabasi Albert (BA) [28], Erdos Reyni (ER) [27], болон уламжлалт SBM [26] аргуудтай харьцуулж туршив. Эдгээр аргуудаар сүлжээ үүсгэхдээ аль болохоор Facebook-ийн бодит өгөгдөлтэй адил байхаар үүсгэсэн. Оройн тоо, ирмэгийн тоо дундаж зэрэг нь facebook-тэй адил байдлаар үүсгэсэн. Бидний ашигласан бодит өгөгдөл болох *facebook* нь олон нийтэд өргөн танигдсан бөгөөд нийгмийн сүлжээний жишиг өгөгдөл болно. Уг графын оройнууд нь *facebook* хэрэглэгчид, ирмэгүүд нь тэдгээрийн хоорондын найзын холбоо юм [25]. Энэ өгөгдлийн FB (*facebook*) гэж товчлон, жишиг сүлжээ болгон үүсгэсэн синтетик сүлжээг уг сүлжээтэй харьцуулах болно.

Туршилтыг хийхдээ networkx [23] санг ашигласан уг санд хэрэгжүүлсэн SBM, BA, ER аргуудыг ашигласан.

A. Алгоритмуудын параметр тохиргоо

Уламжлалт SBM алгоритм сүлжээг үүсгэхдээ жишиг өгөгдлийн оройн тоо, дундаж зэрэгтэй адил сүлжээ үүсгэхийг зорьсон. Уг алгоритм нь оройн тоо n, бүлгүүдийн хэмжээ болон тоо k, бүлэг хоронд ирмэг үүсэх магадлалыг заасан P матриц авдаг. Бүлгийг хэмжээгээр GSBM-тэй яг адилхан зэргийн тархалтыг авч ашигласан. Харин P магадлалын матрицыг үүсгэхдээ томьёо 1, 2-ийг ашигласан. Хэрэв адил бүлгийн оройнуудын хооронд ирмэг үүсгэх бол томьёо 1 магадлалыг ашиглах ба ялгаатай бүлгүүдэд харьяалагдах оройнуудын хооронд ирмэг үүсгэх бол томьёо 2 магадлалыг ашиглана. Энд n_i, n_j нь харгалзан i-р бүлэг, j-р бүлгийн оройн тоо, k нь бүлгүүдийн тоо, d_{avg} нь дундаж зэрэг болно.

$$P_{i,i} = \frac{2 \cdot n_i \cdot d_{avg}}{n_i \cdot (n_i - 1) \cdot k} \tag{1}$$

$$P_{i,j} = \frac{(n_i + n_j) \cdot d_{avg}}{(n_i \cdot n_j) \cdot k} \tag{2}$$

Хүснэгт I
Үүсгэсэн сүлжээнүүдийн бүтцийн талаарх нийтлэг мэдээлэл

	FB	GSBM	SBM	BA	ER
Оройн тоо	4039	4039	4039	4039	4039
Ирмэгийн тоо	88234	87720	85633	84378	88084
Дундаж зэрэг	21.8	21.7	21.2	20.8	21.8
Диаметр	8	6	4	4	4
Дундаж ХБЗУ	3.7	3.0	2.6	2.5	2.6
Зэргийн нийцэл	0.063	0.026	-0.096	-0.006	-0.006

BA алгоритм сүлжээг үүсгэхдээ оройн тоо n, нэг оройн үүсгэж болох холболтын тоо p гэсэн параметрүүдийг жижиг өгөгдлийн оройн тоо болон дундаж зэргээр өгсөн.

Харин ER аргаар сүлжээг үүсгэхдээ оройн тоог n-ийг жижиг өгөгдөлтэй адилаар харин ирмэг үүсгэх магадлал p-ыг дараах томьёогоор өгсөн.

$$p = \frac{2 \cdot m}{n \cdot (n-1)}$$

Энд n, m нь харгалзан жишиг өгөгдлийн оройн тоо, ирмэгийн тоо болно. Уг магадлалын зорилго нь орой бүрийн үүсгэх ирмэгийн тоо нь жишиг өгөгдлийн дундаж зэрэгтэй ойролцоо байлгахын зорьсон.

GSBM алгоритм параметрээрээ оройн тоо n, бүлгийн тоо k, m_1 бүлэг хоронд үүсгэх ирмэгийн тоо, m_2 бүлэг дотроо үүсгэх ирмэгийн тоо гэсэн параметрүүдийг авдаг. Дундаж зэргийг жишиг графын дундаж зэрэгтэй адил байлгахын тулд $m_1 + m_2 \approx d_{avg}$ байхаар сонгон авсан. Туршилтад $m_1 = 1, m_2 = 22, k = 5$ утгыг ашигласан.

B. Үр дүн

Үүсгэсэн сүлжээг бүхлээр нь авч үзсэн шинжийг нийтлэг шинж гэх бөгөөд сүлжээнүүдийн нийтлэг шинжийн харьцуулалтыг Хүснэгт I-д харуулав.

Дундаж зэрэг нь нэг хүн дунджаар хэдэн хүнтэй холбоотой гэдгийг илтгэх бөгөөд жишиг сүлжээний дундаж зэрэг 21.8 байгаа нь уг сүлжээ харьцангуй шигүү сүлжээ гэдгийг илтгэж байна. Шигүү байх тусам тухайн сүлжээнд өвчин хурдан тарах эсвэл мэдээлэл хурдан түгэх боломж бүрддэг. Нийгмийн сүлжээнд өндөр зэрэгтэй байна гэдэг бол тухайн хүний өндөр нөлөөтэйг илтгэнэ [22]. Зохиомлоор үүсгэсэн сүлжээнүүдийн дундаж зэргийн үзүүлэлт бүгд FB сүлжээтэй ойрхон гарсан байгаа нь бидний FB сүлжээтэй ойролцоо сүлжээ байгуулах параметрүүд зөв гэдгийг илтгэж байна.

Диаметрийн хувьд жишиг сүлжээтэй хамгийн ойрхон гарсан нь GSBM байна. Диаметрийг бодохдоо бүх оройн хоорондох хамгийн богино замыг, тухайн замд оролцсон оройн тоогоор олж, хамгийн уртыг нь сонгон авсан. Нийгмийн сүлжээнд small-world нөлөөллөөр сүлжээнд байгаа ямар хоёр хүний хоронд 6 дотор багтах найзуудаар дамжин хүрэх боломжтой байдаг [24]. GSBM алгоритмын үүсгэсэн сүлжээний диаметр 6 гарсан нь бодит нийгмийн сүлжээтэй нийцэлтэй байгааг илтгэж байна. Бусад синтетик сүлжээнүүдийн диаметр 4 байгаа нь хэтэрхий компакт сүлжээ болсныг илтгэнэ.

Дундаж хамгийн богино замын урт (дундаж ХБЗУ) нь дараах томьёогоор илэрхийлэгдэнэ. Энд V нь бүх оройн олонлог, $\sigma_{u,v}$ нь u, v оройн хоорондох хамгийн богино замын уртыг илэрхийлнэ.

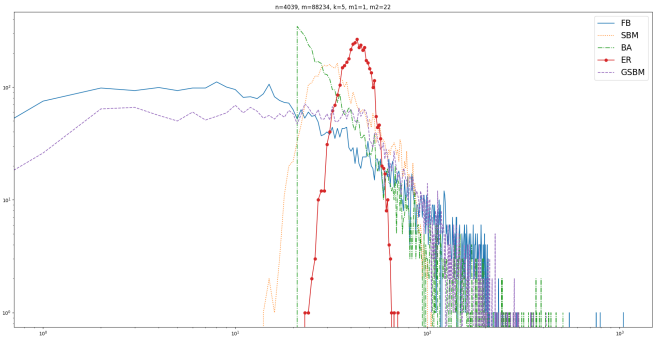
$$\text{дундаж ХБЗУ} = \frac{1}{n \cdot (n-1)} \cdot \sum_{u,v \in V} \sigma_{u,v}$$

Жишиг сүлжээний хувьд дундаж ХБЗУ нь 3.7 гарсан нь мэдээлэл дунджаар 3-аас 4-орой дамжин сүлжээг бүрхэж байгааг илтгэнэ. GSBM алгоритм 3 гаргаж жишиг сүлжээтэй хамгийн ойрхон үр дүн үзүүллээ. Харин бусад сүлжээний хувьд уг үзүүлэлт 2.5, 2.6 гарсан нь хэт компакт байгаар илэрхийлж байна.

Зэргийн нийцэл үзүүлэлт (degree assortavity) оройн зэргийн хувьд тухайн орой өөртэйгөө адилхан зэрэгтэй оройтой холбогдож байгаа эсэхийг харуулдаг. Уг үзүүлэлтийг $[-1, 1]$ завсар хэмждэг бөгөөд эерэг утга нь корреляц сайн байгааг илтгэх бөгөөд сөрөг утга нь тухайн орой өөртэйгөө адил зэрэгтэй оройтой холбогдоогүй гэдгийг илтгэдэг. Бодит нийгмийн сүлжээний хувьд энэ үзүүлэлт чухал бөгөөд баян хүн баян хүнтэйгээ найз, олон найзтай хүн олон найзтай хүнтэйгээ найз болох магадлалтай гэдгийг илтгэх болно. FB жишиг сүлжээний хувьд уг үзүүлэлт эерэг гарсан бөгөөд мөн адил GSBM эерэг хариу гаргасан нь бодит нийгмийн сүлжээний бүтэцтэй нийцтэй байна. GSBM-ийн үзүүлэлт FB-тэй харьцуулахад 0.04-өөр бага үзүүлэлт харуулав. Харин SBM алгоритмын хувьд харьцангуй өндөр утгаар корреляц багатай гарсан бол BA, ER хоёр хоёулаа 0-тэй ойрхон сөрөг утга гаргасан байна.

Оройн зэргийн тархалтыг судлах нь тухайн сүлжээ ямар шинж чанартай сүлжээ болох дүгнэлтийг гаргахад чухал нөлөөтэй. Нийгмийн сүлжээ болон бусад олон бодит сүлжээний зэргийн давтамж нь хүнд сүүлтэй зэргэн тархалтыг (heavy-tailed power law distribution) үзүүлдэг [21]. Энэ нь тухайн сүлжээнд цөөн тооны чухал нөлөөтэй тоглогчид оршин байгааг илтгэнэ. Зураг 1-д харьцуулж буй алгоритмуудаар үүсгэсэн сүлжээний зэргийн тархалтыг харуулав. Зургаас үзэхэд FB сүлжээ хамгийн хүнд сүүлийг харуулж байхад дараагаар нь BA, GSBM орж байна. BA, GSBM хоёрын сүүл хоорондоо ойрхон зөрж байх бөгөөд GSBM харин эхлэл хэсэгтээ FB сүлжээтэй илүү адил гэдэг нь харагдаж байна. FB, GSBM сүлжээнүүдэд 10-аас бага зэрэгтэй хэдэн зуун орой байна. Харин SBM, BA, ER сүлжээнүүдэд байх оройнуудын зэргийн тоо бүгд 10-ээс дээш байна. ER алгоритмын зэргийн тархалт нормаль хэлбэрийнх болох нь ажиглагдаж байгаа нь уг алгоритм бүх оройнуудад дундаж зэрэгтэй ойролцоо холболтуудыг үүсгэн ажилладгийг харуулав. SBM алгоритм дангаараа хүнд сүүл харуулж чадахгүй болох харагдах бөгөөд харин SBM-ийг PA-тэй хослуулсан хувилбар хүнд сүүл харуулж чадаж байгаа нь ажиглагдав.

Алгоритмуудаар үүсгэсэн сүлжээнүүдийг networkx сангийн spring layout-аар гаргасан байршлын дагуу зурж Зураг 2-г гаргасан. Зурагт FB сүлжээний олон бүлэгт нь хоорондоо ялгарч харагдаж байгаа бөгөөд зарим бүлгүүдийн хоронд ямарч шууд ирмэг байхгүй байгаа нь харагдана.



Зураг 1. Хэвтээ тэнхлэгийн дагуу зэргийн тоо, босоо тэнхлэгийн дагуу зэргийн давтамжийг харуулав.

Харин GSBM нь 5 ширхэг бүлэг бүтэц харуулж байгаа нь $k = 5$-тай холбоотой бөгөөд бүх бүлэг бүтцийн хоорондох шууд ирмэг байгаа нь SBM загварчлалтай холбоотой. SBM алгоритмын хувьд бүлэг бүтцээр загварчилсан боловч бүлгүүд хоорондоо ялгарч харагдаж чадахгүй байгаа нь бүлгүүд маш нягт холбоотойгоор үүсгэгдэж бүлгийн ялгаа алга болсонтой холбоотой. Харин BA, ER алгоритмуудын хувьд бүлэг бүтэц загварчлал ашигладаггүй бөгөөд үүсгэсэн сүлжээнүүдэд бүлэг бүтэц ялгарч харагдахгүй байна.

IV. Дүгнэлт

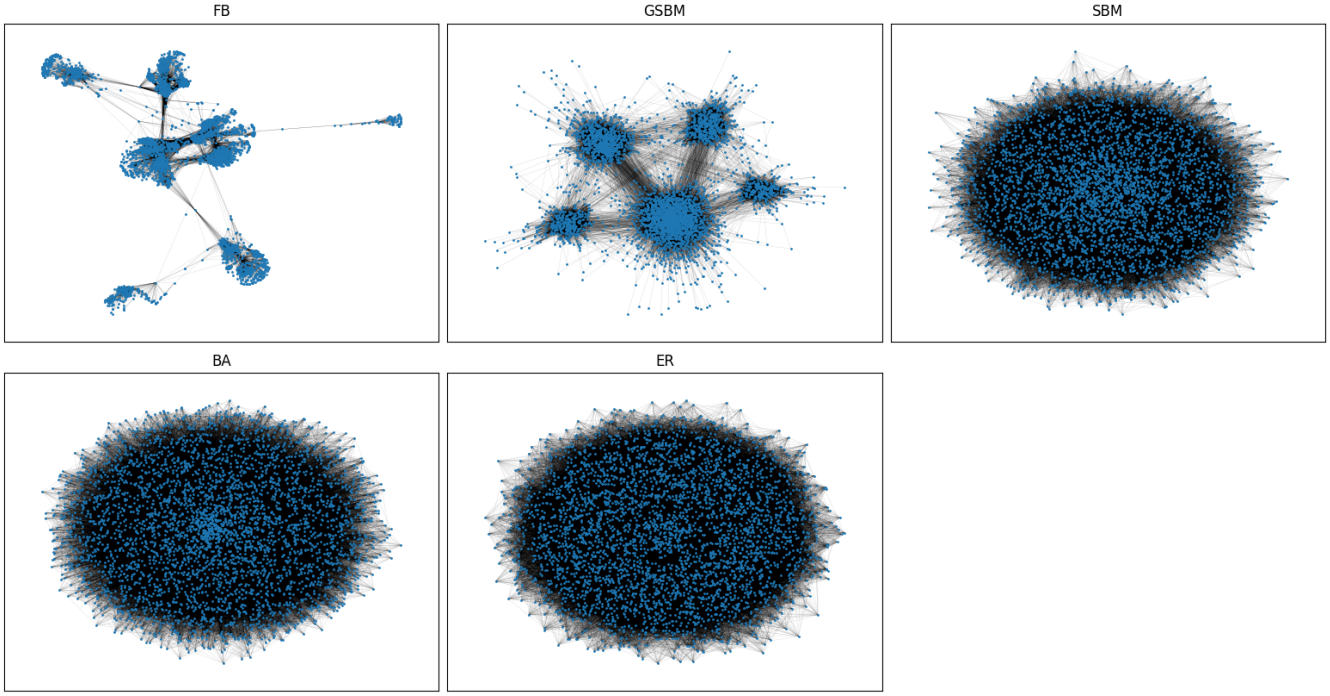
Уламжлалт SBM алгоритм болон PA механизмыг хослуулсан GSBM алгоритмыг зохиомжилж үр дүнг орчин үеийн аргуудтай харьцуулан туршив. Туршилтын үр дүнгээс үзэхэд GSBM алгоритм бодит өгөгдөлтэй ойролцоо бүтэцтэй синтетик сүлжээг үүсгэх чадвартай болохыг харуулав. GSBM алгоритмаар үүсгэсэн синтетик сүлжээний зэргийн давтамж нь бодит сүлжээнд ажиглагддаг хүнд сүүлтэй зэрэг тархалт харуулж байсан ба бүлэг бүтцийг хадгалж байгааг туршилтын өгөгдөл харуулав. GSBM алгоритмын үүсгэсэн сүлжээний дундаж зам нь BA, SBM, ER аргуудаас 0.4 их, бодит сүлжээнээсээ 0.7-оор бага ба зэргийн нийцэл нь бусад синтетик сүлжээнүүдээс хамгийн өндөр буюу 0.026 байна. Үүнээс үзэхэд GSBM алгоритмаар үүсгэсэн сүлжээ жишиг сүлжээтэй хамгийн ойрхон дөхсөн болох нь харагдаж байна.

Цаашдаа сайжруулах ажил гэвэл GSBM алгоритм бүх бүлгүүдийн хоорондох холболт үүсгэж байгаа нь бодит сүлжээтэй нийцэхгүй байна. FB сүлжээг ажиглавал зарим бүлэг нь бусад бүлгүүд дундаас зөвхөн ганц бүлэгтэй харилцаж байгаа нь харагдсан. Энэ шинжийг GSBM-д оруулж ирвэл бодит нийгмийн сүлжээтэй илүү ойр дөхсөн синтетик сүлжээ үүсгэж болох боломж харагдаж байна.

Ном зүй

Ашигласан ном

[1] M. E. J. Newman, *Networks: An Introduction*, Oxford Univ. Press, 2010.
[2] A.-L. Barabási, *Network Science*, Cambridge Univ. Press, 2016.
[3] D. Lusher, J. Koskinen, and G. Robins, *Exponential Random Graph Models for Social Networks*, Cambridge Univ. Press, 2012.



Зураг 2. FB сүлжээнээс бусад нь синтетикээр үүсгэсэн сүлжээ. Сүлжээг зурахдаа networkx санг ашигласан.

[4] A. Lancichinetti and S. Fortunato, ``Benchmarks for testing community detection algorithms," *Phys. Rev. E*, vol. 80, 016118, 2009.

[5] D. J. Watts, ``A simple model of global cascades on random networks," *Proc. Natl. Acad. Sci.*, vol. 99, no. 9, pp. 5766–5771, 2002.

[6] M. O. Jackson, *Social and Economic Networks*, Princeton Univ. Press, 2010.

[7] P. W. Holland, K. B. Laskey, and S. Leinhardt, ``Stochastic blockmodels: First steps," *Social Networks*, vol. 5, pp. 109–137, 1983.

[8] K. Nowicki and T. A. B. Snijders, ``Estimation and prediction for stochastic blockstructures," *J. Amer. Statist. Assoc.*, vol. 96, pp. 1077–1087, 2001.

[9] B. Karrer and M. E. J. Newman, ``Stochastic blockmodels and community structure in networks," *Phys. Rev. E*, vol. 83, 016107, 2011.

[10] M. E. J. Newman and A. Clauset, ``Structure and inference in annotated networks," *Nat. Commun.*, vol. 7, 11863, 2016.

[11] A.-L. Barabási and R. Albert, ``Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999.

[12] S. N. Dorogovtsev and J. F. F. Mendes, ``Evolution of networks," *Adv. Phys.*, vol. 51, pp. 1079–1187, 2002.

[13] E. M. Airoldi et al., ``Mixed membership stochastic blockmodels," *J. Mach. Learn. Res.*, vol. 9, pp. 1981–2014, 2008.

[14] T. P. Peixoto, ``Hierarchical block structures," *Phys. Rev. X*, vol. 4, 011047, 2014.

[15] P. Gopalan et al., ``Scalable inference of overlapping communities," *NeurIPS*, vol. 25, pp. 2249–2257, 2013.

[16] C. Lee and W. Wilkinson, ``A review of stochastic block models," *J. Mach. Learn. Res.*, vol. 22, pp. 1–50, 2021.

[17] A. Cerqueira et al., ``Consistent model selection for degree-corrected SBM," *arXiv:2302.03734*, 2023.

[18] G. Poux-Médard et al., ``Dynamic Mixed Membership SBM," *arXiv:2304.05894*, 2023.

[19] T. Yang et al., ``Detecting communities," *Mach. Learn.*, vol. 82, pp.157–189, 2011.

[20] K. Xu and A. Hero, ``Dynamic SBM," *IEEE J. Sel. Topics Signal Process.*, vol. 8, pp.552–562, 2014.

[21] David A Bader and Kamesh Madduri. *Parallel algorithms for evaluating centrality indices in real-world networks.* In 2006 International Conference on Parallel Processing (ICPP'06), pages 539--550. IEEE, 2006.

[22] Ta-Shun Cho and Hsin-Yu Shih. *Patent citation network analysis of core and emerging technologies in taiwan: 1997--2008.* Scientometrics, 89(3):795--811, 2011.

[23] Aric Hagberg, Pieter J Swart, and Daniel A Schult. *Exploring network structure, dynamics, and function using networkx.* Technical report, Los Alamos National Laboratory (LANL), Los Alamos, NM (United States), 2008.

[24] Marcelo Kuperman and Guillermo Abramson. *Small world effect in an epidemiological model.* Physical review letters, 86(13):2909, 2001.

[25] Jure Leskovec and Rok Sosič. *Snap: A general-purpose network analysis and graph-mining library.* ACM Transactions on Intelligent Systems and Technology (TIST), 8(1):1, 2016.

[26] Holland, P., Laskey, K. & Leinhardt, S. Stochastic blockmodels: First steps. *Social Networks*. **5**, 109-137 (1983)

[27] Erdős, P. & Rényi, A. On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*. **12**, 261-267 (1961)

[28] Barabási, A. & Albert, R. Emergence of scaling in random networks. *Science*. **286**, 509-512 (1999)

Машин сургалт ба граф суурьт хамтын шүүлтүүрийг ашиглан бүтээгдэхүүн санал болгох системийг сайжруулах нь

Энхтөрийн Сэргэлэнбаатар
Компьютерын Ухааны салбар
ШУТИС, МХТС
Улаанбаатар, Монгол
sergelenbaatare@gmail.com

Тэгшээгийн Түвшинсайхан
Мэдээллийн Технологийн салбар
ШУТИС, МХТС
Улаанбаатар, Монгол
J.DS24E051@must.edu.mn

Ишдоржийн Цэрэн-Онолт
Компьютерын Ухааны салбар
ШУТИС, МХТС
Улаанбаатар, Монгол
tseren-onolt@must.edu.mn

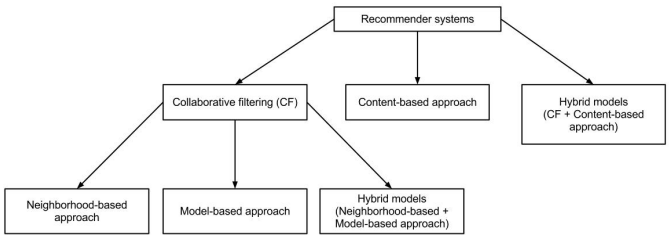
Батжаргалын Долгорсүрэн
Мэдээллийн Технологийн салбар
ШУТИС, МХТС
Улаанбаатар, Монгол
b.dolgorsuren@must.edu.mn

Хураангуй - Цахим худалдаа, олон төрлийн платформууд өсөн нэмэгдэж буй энэ үед хэрэглэгчийн сонирхолд нийцсэн бүтээгдэхүүн санал болгох нь хэрэглэгчийн сэтгэл ханамж болон борлуулалтыг нэмэгдүүлэх гол хүчин зүйл болж байна. Хамтын шүүлтүүрийн (Collaborative Filtering) аргууд нь хэрэглэгчдийн зан төлөв дээр үндэслэн ажилладаг ч өгөгдлийн ховор байдал (data sparsity), шинэ хэрэглэгч барааны танил бус байдал (cold start problem) зэрэг хүндрэлтэй тулгардаг. Энэхүү судалгаанд граф суурьт хамтын шүүлтүүр (Graph-based Collaborative Filtering) болон машин сургалтын (Machine Learning) арга зүйг ашиглан хослуулж авах бараг санал болгох, цаашлаад дэлгүүр тус бүр дэх барааны өрөлтийг оновчилж одоо байгаа борлуулалтын системийг илүү сайжруулах, ухаалаг болгохыг зорьсон юм. Хэрэглэгч болон бүтээгдэхүүнийг зангилаа (nodes), тэдгээрийн хоорондын харилцааг холбоос (edges) хэлбэрээр дүрсэлсэн, цаашлаад тухайн графыг өөрчлөн бүтээгдэхүүн-бүтээгдэхүүн графыг гаргаж авсан. Graph Neural Networks (GNN) зэрэг гүнзгий сургалтын аргыг ашиглан нарийвчилсан загвар боловсруулав. Судалгааны үр дүнд уламжлалт аргуудтай харьцуулахад илүү өндөр нарийвчлалтай, хэрэглэгчийн сонирхолд нийцсэн, бараа болон түүнтэй хослуулж авах боломжтой бараануудыг тодорхой хувийн нарийвчлалтай санал болгох системийг бий болгосон.

Түлхүүр үг—машин сургалт, граф суурилсан шүүлтүүр, хосолмол зөвлөмжийн систем

I. Удиртгал

Өнөөгийн мэдээлэл технологийн эрин үед, хэрэглэгчийн сонирхолд нийцсэн мэдээлэл, бүтээгдэхүүн санал болгох нь цахим платформуудын өрсөлдөх чадварыг тодорхойлогч гол хүчин зүйл болж байна [1]. Санал болгох систем (Recommendation Systems) нь хэрэглэгчийн сонирхолд тохирсон бүтээгдэхүүн, үйлчилгээг урьдчилан таамаглах зорилгоор ашиглагддаг мэдээллийн систем юм. Эдгээр системүүд нь хэрэглэгчийн өмнөх үйлдлүүд, үзсэн зүйлс, худалдан авалт, үнэлгээ зэрэг мэдээллийг ашиглан хэрэглэгчид хамгийн тохиромжтой, таалагдах бүтээгдэхүүнийг санал болгодог [2]. Энэ төрлийн системүүд нь collaborative filtering, content-based filtering, болон hybrid models гэх мэт аргачлалуудыг ашиглан илүү тодорхой, хувьчилсан санал өгөх боломжтой байдаг. Collaborative filtering нь бусад хэрэглэгчдийн үйлдэлд суурилсан санал болгох арга бөгөөд хэрэглэгчдийн хоорондын харилцаа, үйлдлүүдийг ашиглан зөвлөмжүүдийг гаргадаг. Content-based filtering нь бүтээгдэхүүний онцлог, шинж чанарт суурилдаг бөгөөд тухайн бүтээгдэхүүний тодорхойлолт болон хэрэглэгчийн өмнөх сонголтыг үндэслэн санал болгох боломжийг олгодог [3]. Харин hybrid models нь эдгээр аргуудыг хослуулан ашиглах бөгөөд, хоёр аргачлалын давуу талуудыг нэгтгэж илүү үр дүнтэй санал болгох боломжийг олгодог.



Зураг 1. Зөвлөмж өгөх систем дэх хамтын шүүлтүүр

Эдгээр системүүд нь хэрэглэгчийн өмнөх үйлдлүүдийг шинжилж, хамгийн тохиромжтой бүтээгдэхүүнийг санал болгохдоо ихэнхдээ machine learning болон data mining аргачлалуудыг ашигладаг. Энэхүү алгоритм нь хэрэглэгчийн сонирхол, зан төлөвийг цаг хугацаа өнгөрөх тусам илүү сайн таньж, сайжруулсан

таамаглалуудыг хийх боломжийг олгодог [4]. Тэгвэл бид энэхүү судалгааны ажлын хүрээнд collaborative filtering талаар дэлгэрэнгүй дурдах юм. Ихэнх санал болгох системүүд нь онлайн худалдааны сайтууд, урлагийн контент (кино, хөгжим), нийгмийн сүлжээг хөгжүүлэхэд чухал үүрэг гүйцэтгэдэг бөгөөд тэдгээр нь хэрэглэгчийн туршлага, хэрэгцээг хамгийн өндөр түвшинд нийцүүлэн үйл ажиллагаагаа зохион байгуулдаг. Ийм систем нь хэрэглэгчдэд илүү хялбар, тохиромжтой үйлчилгээ үзүүлэх зорилготой бөгөөд цахим худалдаа, контент зохиогчид болон өөр олон салбарт чухал үүрэгтэй болжээ. Хамтын шүүлтүүр (Collaborative filtering) нь санал болгох системийн голлох, өргөн хэрэглэгддэг аргуудын нэг бөгөөд хэрэглэгч болон бүтээгдэхүүний хоорондын харилцаанд үндэслэн таамаглал дэвшүүлдэг. Энэ арга нь хэрэглэгчдийн өмнөх үйлдлүүдийг, түүнчлэн ижил төстэй хэрэглэгчдийн зан төлөвийг үндэслэн ирээдүйд аль бүтээгдэхүүн эсвэл үйлчилгээг хэрэглэгчид илүү сонирхож, хэрэглэнэ гэж таамаглахад ашиглагддаг.

Гэвч, collaborative filtering нь хэд хэдэн хязгаарлалтад өртдөг, тухайлбал өгөгдлийн хомсдол (data sparsity) болон шинэ хэрэглэгч болон шинэ бүтээгдэхүүний асуудал (cold start problem) зэрэг асуудлууд үүсдэг [5]. Энэ нь тухайн системд хэрэглэгчид болон бүтээгдэхүүнүүдийн хоорондын хамтын ажиллагааг бүрэн дүүрэн тооцох боломжийг хязгаарлаж, санал болгох системийн үр ашигт сөргөөр нөлөөлдөг. Эдгээр хүндрэлүүдийг шийдвэрлэх зорилгоор, машин сургалтын (machine learning) аргуудыг хамтын шүүлтүүртэй хослуулан ашиглах нь үр дүнтэй арга зам болсон байна [6]. Түүнчлэн, хэрэглэгч-бүтээгдэхүүний харилцааг граф бүтэц (graph structure) хэлбэрээр дүрслэн боловсруулах чиг хандлага сүүлийн жилүүдэд эрчимтэй хөгжиж байгаа бөгөөд энэ нь санал болгох системийн гүйцэтгэлийг сайжруулах, илүү нарийвчилсан шийдэл гаргахад чухал нөлөө үзүүлж байна [7]. Граф суурьт хамтын шүүлтүүр (Graph-based collaborative filtering) нь хэрэглэгч болон бүтээгдэхүүнийг зангилаа (node), тэдгээрийн харилцааг холбоос (edge) хэлбэрээр илэрхийлдэг бөгөөд хэрэглэгчийн мэдээллийг гүнзгийрүүлэн судалж, орчны мэдээллийг ашиглах боломжийг олгодог. Түүнчлэн, граф нейрон сүлжээ (Graph Neural Network – GNN) зэрэг гүнзгий сургалтын аргууд нь шууд болон дам утгатай холбоосыг зэрэг тооцон, санал болгох системийн нарийвчлалыг сайжруулахад чухал нөлөө үзүүлдэг [8]. Иймд, энэхүү судалгааны ажлын хүрээнд граф суурьт хамтын шүүлтүүр болон машин сургалтын аргуудыг хослуулан, бүтээгдэхүүн санал болгох системийн гүйцэтгэлийг сайжруулах боломж, арга зүйг судлан боловсруулсан болно

II. Судлагдсан байдал

Санал болгох системийн хөгжил нь мэдээлэл технологийн салбарт хурдтай явцтайгаар өргөжиж байна. Хамтын шүүлтүүр (collaborative filtering) нь санал болгох системийн хамгийн өргөн хэрэглэгддэг аргуудын нэг бөгөөд хэрэглэгчид болон бүтээгдэхүүнүүдийн хоорондын харилцаа, зан төлөвийг шинжилж, ирээдүйн хэрэгцээг урьдчилан таамаглах боломжийг олгодог [9]. Тухайлбал, Хамтын шүүлтүүрийн анхны хувилбарууд нь хэрэглэгчид болон бүтээгдэхүүнүүдийн хоорондын үнэлгээний өгөгдлийг ашиглан санал болгох системийг бий болгосон байдаг [10]. Гэвч энэ арга нь өгөгдлийн хомсдол болон шинэ хэрэглэгчдийн асуудал (cold start problem) зэрэг хязгаарлалтад өртдөг. Иймд машин сургалтын аргууд нь хамтын шүүлтүүрийг сайжруулахад чухал үүрэг гүйцэтгэж байгаа бөгөөд хэрэглэгчдийн болон бүтээгдэхүүний хоорондын харилцааг илүү нарийвчилсан байдлаар таамаглах боломжийг олгодог. Тухайлбал, Matrix Factorization (MF) болон Singular Value Decomposition (SVD) зэрэг аргачлалууд нь хэрэглэгчийн зан төлөвийг илүү нарийвчилсан байдлаар шинжлэх боломжийг олгож байна [11]. Мөн Deep Learning (гүнзгий сургалт) арга нь хэрэглэгчийн зан төлөвийг илүү гүнзгий, автоматчилагдсан байдлаар судлах боломжийг олгодог [12]. Граф суурьт хамтын шүүлтүүрийн аргууд нь энэ чиглэлд шийдэл санал болгож байгаа бөгөөд хэрэглэгчийн болон бүтээгдэхүүний харилцааг граф бүтэц хэлбэрээр илэрхийлэх боломжийг олгодог. Граф нейрон сүлжээ (GNN) нь энэ хандлагыг нэмэгдүүлэн, хэрэглэгч болон бүтээгдэхүүний хоорондын харилцааг илүү нарийвчилсан хэмжээгээр тооцох боломжийг олгоно [13]. Граф суурьт хамтын шүүлтүүр нь улам бүр хүчирхэгжиж, олон төрлийн шинэ хэрэглэгчийн мэдээллийг хурдан боловсруулах, өгөгдлийн хомсдолыг багасгах зэрэг асуудлыг шийдвэрлэхэд тусалдаг. Эдгээр судалгаанууд нь хэрэглэгчийн зан төлөвийг илүү нарийвчилсан, үр дүнтэй судлах шаардлагыг тодорхойлсон бөгөөд энэ нь одоогийн санал болгох системүүдийн гүйцэтгэлийг сайжруулахад чухал үүрэг гүйцэтгэж байна. Мөн граф суурьт хамтын шүүлтүүрийн чиглэлээр олон судалгаа хийгдсэн. Sun нар нь олон граф давхаргыг ашигласан хамтын шүүлтүүрийн аргыг санал болгож, хэрэглэгч-бүтээгдэхүүний олон түвшний харилцааг илүү үр дүнтэйгээр ашигласан байна. Мөн Wang нар [15]нь мэдлэгийн граф ашиглан илүү утгатай (semantic) холболтуудыг системд тусгах замаар санал болголтын чанарыг сайжруулсан. Edges буюу ирмэгүүд дээр жинг өгч, хэрэглэгчийн зан төлөв болон итгэлцлийн түвшинг тусгах оролдлогууд ч хийгдсэн байна. Сүүлийн жилүүдэд гүнзгий сургалтад суурилсан нейрон санал болгох системүүд (Neural Recommendation Systems) нь улам бүр хүчирхэгжиж, хэрэглэгч-бүтээгдэхүүний харилцааг илүү ухаалаг загварчлах боломжийг олгож байна. NeuMF (Neural Matrix Factorization) загвар нь matrix factorization ба гүн нейрон сүлжээг хослуулснаар хэрэглэгчийн болон бүтээгдэхүүний латент шинжүүдийг илүү үр дүнтэй загварчилж чадсан [16]. Үүний дараагаар граф мэдээллийг нейрон аргаар ашигладаг NGCF (Neural Graph Collaborative Filtering) бо-

лон LightGCN (Light Graph Convolutional Network) загварууд боловсруулагдсан. NGCF нь графын бүтэц дэх хөршүүдийн мэдээллийг олон давхаргаар нэмж тусган, харилцан нөлөөллийг нарийвчлан тооцдог бол LightGCN нь энэхүү архитектурыг хялбаршуулан, параметр болон тооцооллын хувьд хөнгөн болгосон шийдэл юм [17] [18]. Эдгээр загварууд нь санал болголтын нарийвчлалыг сайжруулаад зогсохгүй, өгөгдлийн хомсдол, шинэ хэрэглэгчийн асуудлыг илүү үр дүнтэй шийдвэрлэх боломжийг бүрдүүлж байна.

III. Онолын хэсэг

Санал болгох системүүд (Recommender Systems) нь хэрэглэгчдэд хамаарал бүхий мэдээлэл, бүтээгдэхүүн, үйлчилгээг автоматаар санал болгох зорилготой ухаалаг системүүд юм. Эдгээр системүүд нь орчин үеийн цахим орчин дахь хэрэглэгч төвтэй үйлчилгээний гол тулгуур болж, цахим худалдаа (e-commerce), видео болон хөгжмийн стриминг платформууд (Netflix, Spotify гэх мэт), нийгмийн сүлжээ, мэдээ мэдээллийн сайтууд, боловсролын цахим платформ зэрэг өргөн хүрээний хэрэглээнд нэвтэрч, хэрэглэгчийн туршлагыг (user experience) дээд түвшинд хүргэхэд чухал үүрэг гүйцэтгэж байна [19].Санал болгох системүүд нь агуулга суурьт (content-based), хамтын шүүлтүүрт (collaborative filtering), эрлийз (hybrid) загварууд гэх мэт хэд хэдэн ангилалд хуваагддаг. Үүнээс хамгийн түгээмэл хэрэглэгддэг, хэрэглэгчдийн өмнөх зан төлөв, үнэлгээний өгөгдөл дээр үндэслэн ажилладаг нь хамтын шүүлтүүр (Collaborative Filtering, CF) юм. Тус арга нь хэрэглэгчдийн өнгөрсөн хандлага, харилцан үйлдлүүдийг дүн шинжилж, тэдний сонирхолд нийцсэн шинэ санал гаргахад чиглэдэг бөгөөд хэрэглэгчдийн ижил төстэй байдал болон бүтээгдэхүүнүүдийн ижил төстэй онцлогийг ашиглан таамаглал хийдэг.

A. Хамтын шүүлтүүрийг үндсэндээ хоёр үндсэн хэлбэрт ангилж болно:

- Хэрэглэгч суурьт хамтын шүүлтүүр (User-Based Collaborative Filtering): Энэ хандлага нь тухайн хэрэглэгчтэй ижил төстэй хэрэглэгчдийг тодорхойлох замаар санал болголтыг боловсруулдаг. Ижил төстэй хэрэглэгчид гэж ижил бүтээгдэхүүнд төстэй үнэлгээ өгсөн буюу төстэй хандлага үзүүлсэн хүмүүсийг ойлгоно. Ижил төстэй байдлыг хэмжихэд cosine similarity, Pearson correlation coefficient зэрэг статистик аргачлалуудыг ашигладаг. Энэхүү хандлага нь хялбар хэрэгжих боломжтой ч хэрэглэгчийн тоо нэмэгдэхийн хэрээр тооцооллын өртөг өндөрсдөг.
- Бүтээгдэхүүн суурьт хамтын шүүлтүүр (Item-Based Collaborative Filtering): Энэ нь хэрэглэгчийн өмнө нь өндөр үнэлгээ өгсөн эсвэл хэрэглэсэн бүтээгдэхүүнүүдтэй төстэй онцлог бүхий бусад бүтээгдэхүүнийг санал болгодог арга юм. Энэ арга нь бүтээгдэхүүнүүдийн хоорондын ижил төстэй байдлыг тодорхойлох замаар хэрэглэгчийн ирээдүйн сонголтыг урьдчилан таамаглах боломжийг олгодог. Харьцуулбал, хэрэглэгчийн суурьт аргын хувьд энэхүү арга нь илүү тогтвортой бөгөөд том хэмжээний өгөгдөл дээр үр дүнтэй ажилладаг.

Гэсэн хэдий ч хамтын шүүлтүүрийн эдгээр уламжлалт аргууд нь дараах хоёр томоохон хүндрэлтэй асуудалтай тулгардаг:

- Өгөгдлийн сийрэгжилт (Data Sparsity): Хэрэглэгчдийн зүгээс нийт боломжит бүтээгдэхүүнүүдийн маш бага хувьд л үнэлгээ өгсөн байдаг. Энэ нь хэрэглэгч, бүтээгдэхүүний матриц маш их хоосон утгатай болох шалтгаан болдог бөгөөд ижил төстэй байдлыг тооцоход хүндрэл үүсгэдэг.
- Хүйтэн эхлэл (Cold Start): Шинэ хэрэглэгч болон шинэ бүтээгдэхүүн нэвтэрч ирэхэд тэдгээрийн талаар хангалттай мэдээлэл байхгүй тул санал болгох систем тэдэнд тохирсон зөвлөмж гаргахад бэрхшээлтэй болдог.

Тэгвэл эдгээр асуудлуудыг шийдвэрлэх зорилгоор сүүлийн жилүүдэд өгөгдлийн бүтцийн шинэ хандлагуудыг судалж эхэлсэн бөгөөд тэдгээрийн нэг нь граф онолыг санал болгох системд ашиглах явдал юм.

B. Граф онол ба түүний санал болгох системд ашиглах нь

Граф онол (Graph Theory) нь математикт өгөгдөл дэх элементүүдийн хоорондын хамаарлыг илэрхийлэх хүчирхэг хэрэгсэл бөгөөд орчин үеийн өгөгдлийн шинжилгээ, хиймэл оюун ухаан, сүлжээний онол зэрэг олон салбарт өргөн хэрэглэгддэг. Граф нь орой (node буюу vertex) болон ирмэг (edge)-ээс бүрдэх ба эдгээр нь тухайн өгөгдлийн бүтэц, логик харилцааг илтгэдэг.Санал болгох системийн граф бүтэцтэй загварт дараах байдлаар өгөгдлийг дүрслэн харуулж болно:

- Орой (Nodes): Хэрэглэгч болон бүтээгдэхүүн бүрийг тус тусад нь орой хэлбэрээр төлөөлнө.
- Ирмэг (Edges): Хэрэглэгчийн бүтээгдэхүүнтэй хийсэн харилцан үйлдэл (жишээ нь үнэлгээ өгөх, худалдан авах, хандалт хийх гэх мэт) нь хэрэглэгч болон бүтээгдэхүүний хоорондуд үүссэн холбоо буюу ирмэгийг үүсгэнэ. Ирмэг нь жингийн утгатай байж болох ба тухайн харилцан үйлдлийн давтамж, чанар зэргийг илэрхийлж болно.

Энэхүү графын бүтэц нь хоёр хэсэгт граф (bipartite graph) буюу хоёр өөр төрлийн оройнуудын хоорондуд зөвхөн холбоо үүсгэж болох графын ангилалд хамаардаг. Энэ бүтэц нь зөвхөн шууд харилцаанд суурилсан төдийгүй хэрэглэгч болон бүтээгдэхүүний хоорондох шууд бус, олон шатлалт холбоо (multi-hop relationships)-г илрүүлэх, анализ хийх боломжийг олгодог [20]. Мөн графын хөрш оройн мэдээлэл болон бүтцийн мэдээллийг ашиглан уламжлалт матриц суурьт аргуудтай харьцуулахад илүү уян хатан, гүнзгий ойлголт бүхий төлөөлөл (representation) үүсгэх боломж бүрддэг. Зураг 2 дээр

санал болгох системийг графын хэлбэрээр дүрсэлсэн байна. Энэ граф дараах элементүүдээс бүрдэнэ.

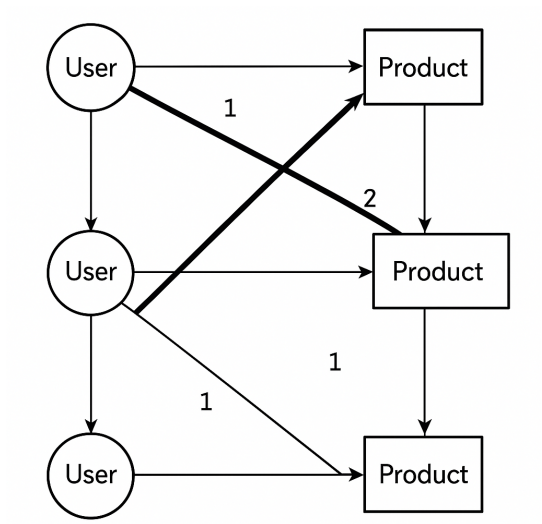
Оройнууд (Nodes)

- User A, User B, User C — эдгээр нь хэрэглэгчидийг илэрхийлж байна.
- Product 1, Product 2, Product 3 — эдгээр нь бүтээгдэхүүнүүдийг илэрхийлж байна.

Ирмэгүүд (Edges)

- Хэрэглэгч болон бүтээгдэхүүнийг холбосон сумтай шугамууд нь хэрэглэгчийн тухайн бүтээгдэхүүнтэй хийсэн харилцан үйлдлийг илэрхийлнэ.
  Жишээ нь:
- User A → Product 1 гэдэг нь "User A тухайн бүтээгдэхүүнтэй харилцсан (жишээ нь үзсэн, үнэлсэн, худалдан авсан)"гэсэн утгатай.
- Ирмэгийн зузаан нь тухайн харилцан үйлдлийн чанар эсвэл давтамжийг илэрхийлж болно. Илүү зузаан ирмэг = илүү их харилцан үйлдэл гэж ойлгож болно.

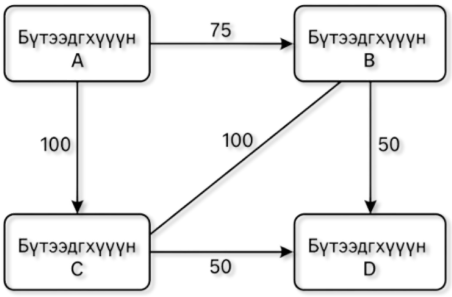
Ийм граф бүтэц нь санал болгох алгоритмд хэрэглэгч болон бүтээгдэхүүний хоорондын хамаарлыг тодорхойлох, ижил хэрэглэгчидтэй төстэй үйлдэл хийсэн бусад хэрэглэгчдийг олж, шинэ бүтээгдэхүүн санал болгоход ашиглагддаг.



Зураг 2. Санал болгох системийг графын хэлбэр

Эхний Зураг 2-т үзүүлсэн графаас харахад хэрэглэгчдийн хамгийн их сонгож хэрэглэдэг барааг тодорхойлох боломжтой. Энэ нь хэрэглэгчийн сонирхол, худалдан авалтын давтамжаас хамааран илэрч байгаа бөгөөд тухайн бүтээгдэхүүн нь бусад бараануудтай харьцуулахад илүү эрэлттэй, төвлөрөлтэй байна.

Зураг 3-т дахь графаас бүтээгдэхүүнүүдийн хоорондын хамт хэрэглэгдэх хандлага буюу хамаарал тодорхойлогдож байгаа бөгөөд энэ нь нэг бараа худалдаж авахад дагаад ямар бараа түүнтэй хамт авах магадлал



Зураг 3. Санал болгох системийг графын хэлбэр

өндөр байгааг харуулж байна. Тиймээс эдгээр хоёр графыг нэгтгэн ойлговол дараах дүгнэлтийг хийж болно.

Хэрэглэгчийн хамгийн их авдаг бараатай (жишээ нь: талх) хамгийн их хамааралтай, хамтдаа авдаг хандлагатай барааг (жишээ нь: масло) ойрхон байрлуулах буюу хамт санал болговол хэрэглэгчийн давхар худалдан авалтыг нэмэгдүүлэх боломжтой. Энэ нь дэлгүүрийн хувьд бүтээгдэхүүний байршлыг оновчтой төлөвлөх, борлуулалтыг нэмэгдүүлэх, хэрэглэгчдэд тохирсон бүтээгдэхүүнийг системээр санал болгох чухал стратеги болно.

Жишээлбэл, "Талхны лангууны хажууд масло байрлуулбал хэрэглэгч хоёр бүтээгдэхүүнийг хамтад нь авах магадлал өндөр"гэдэг нь хэрэглэгчийн худалдан авалтын зан төлөвт суурилсан ухаалаг шийдэл юм. Тэгэхээр бид бодит өгөгдөл дээрээс энэхүү зүй тогтол ажиглагдах эсэхийг батлахыг зорилоо.

C. Граф суурьт хамтын шүүлтүүр

Граф суурьт хамтын шүүлтүүр (Graph-based Collaborative Filtering, GCF) нь санал болгох системийг графын бүтэц дээр загварчилж, хэрэглэгч болон бүтээгдэхүүнүүдийн хоорондох шууд болон дам харилцааг илүү гүнзгий тусган авч үзэх аргуудын цогц юм. Энэ арга нь уламжлалт хамтын шүүлтүүрийн аргачлалуудыг гүнзгий сургалтын аргуудтай нэгтгэж, илүү оновчтой, ухаалаг төлөөлөл үүсгэх боломжийг нээж өгдөг. Сүүлийн жилүүдэд Граф нейрон сүлжээ (Graph Neural Networks, GNNs) нь граф өгөгдөлтэй ажиллах, нарийн төвөгтэй хамаарлыг илэрхийлэхэд өндөр үр дүн үзүүлж байгаа бөгөөд санал болгох системийн хөгжлийн шинэ чиг хандлага болоод байна. GNN-ийн тусламжтайгаар орой бүрийн хөрш оройн мэдээллийг аггрегац хийж, оройн шинэ төлөөлөл (embedding)-ийг үе шаттайгаар сурч авдаг. Үүний үр дүнд хэрэглэгч болон бүтээгдэхүүний илүү гүнзгий, утгатай онцлог шинжүүдийг тусгасан векторуудыг гарган авах боломжтой болдог. Түгээмэл хэрэглэгддэг GNN суурьтай алгоритмууд

- Graph Convolutional Networks (GCNs): Графын хөрш оройноос жигнэсэн дундаж мэдээллийг цуг-

луулж, олон давхаргад дамжуулан оройн төлөөллийг сайжруулдаг. Энэ арга нь графын бүтцийг convolution үйлдлээр илэрхийлдэг бөгөөд сүлжээний гүн болох тусам илүү холын хөршүүдийн мэдээллийг сурч авдаг.

- GraphSAGE: Том хэмжээний граф дээр ажиллахад зориулагдсан бөгөөд хөршүүдээс мэдээлэл аггрегац хийхдээ sampling буюу дээж авах арга хэрэглэдэг. Мөн хөршүүдийн мэдээллийг нэгтгэхдээ mean, LSTM, pooling гэх мэт янз бүрийн функцүүд ашигладаг.
- PinSage: Pinterest компанийн санал болгох системд зориулан хөгжүүлсэн энэхүү загвар нь граф бүтэц болон бүтээгдэхүүний визуал онцлогийг хослуулж embedding үүсгэдэг. Энэ нь графд суурилсан гүнзгий сургалтыг жинхэнэ хэрэглээнд ашиглаж болдгийг харуулсан анхны жишээнүүдийн нэг юм.

Эдгээр аргачлалууд нь зөвхөн хэрэглэгчийн өнгөрсөн зан төлөвийг тусгаад зогсохгүй, графын бүтцийн мэдээллийг ашиглан дам холбоос бүхий хамаарлыг тодорхойлж, хэрэглэгчийн ирээдүйн хандлагыг илүү оновчтойгоор урьдчилан таамаглахад тусалдаг.

D. Комбинацийн томъёо (C(n, k)) болон хамтын шүүлтүүр (CF)-ийг хэрхэн холбох вэ?

E. Комбинацийн томъёо (C(n, k)):

Комбинацийн томъёо нь олон зүйлийг тодорхой тоогоор хослуулах боломжийг олгодог математик томъёо юм. Энэ нь n зүйлээс k зүйл сонгох тооны хослолыг илэрхийлдэг. Жишээлбэл, гурван бараа (A, B, C)-г хоёр хоёроор нь хослуулахад комбинацийн томъёо ашиглана. Хэрвээ бид $C(n, k)$ гэж бичвэл, энэ нь дараах томъёогоор тооцогдож байна:

$$C(n, k) = \frac{n!}{k!(n-k)!} \tag{1}$$

- n — нийт элементийн тоо (жишээ нь, барааны тоо)
- k — сонгох элементийн тоо (жишээ нь, хослуулах бараа)
- ! — факториал (тоо хэдэн удаа буурахыг илэрхийлдэг)

F. Жишээ:

Хэрвээ $n = 3$ бөгөөд $k = 2$ бол:

$$C(3, 2) = \frac{3!}{2!(3-2)!} = \frac{3 \times 2 \times 1}{2 \times 1 \times 1} = 3 \tag{2}$$

Энэ нь гурван бараа (A, B, C)-г хоёр хослох боломжийн тоо юм. Нийтдээ 3 хослол үүснэ: $\{(A, B); (A, C); (B, C)\}$.

G. Комбинацийн томъёо болон хамтын шүүлтүүрийг холбох:

- Хэрэглэгчийн сонирхлын хослолуудыг илрүүлэх: Хамтын шүүлтүүрийн алгоритмд хэрэглэгчид болон бүтээгдэхүүнүүдийн хооронд олон төрлийн хослол үүсгэдэг. Жишээ нь, хэрэглэгч A бүтээгдэхүүн X болон Y-г хамт худалдан авсан бол, энэ нь хоёр бараа хоорондын хослол болж, эдгээр хослолыг ашиглан шинэ бүтээгдэхүүн санал болгож болно. Энэ тохиолдолд, $C(2, 2)$-г ашиглан 2 барааг хослуулах тоог тооцоолж, эдгээрийг хамтран зөвлөмж болгож болно.
- Хэрэглэгчийн зан төлөвийг анализ хийх: Хэрэглэгчид бараа эсвэл үйлчилгээг хамтран хэрэглэж байгаа тохиолдолд эдгээр хослолуудыг комбинацийн томъёо ашиглан тооцоолж, хослолуудыг зөвлөснөөр хэрэглэгчийн сонирхлыг илүү нарийвчлан тодорхойлж, хамгийн тохиромжтой бүтээгдэхүүнүүдийг санал болгож чадна.
- Санал болгох системд комбинацийн томъёог ашиглах: Хэрэглэгчид болон бүтээгдэхүүнүүдийн хоорондох хослолуудын тоог тооцоолох нь хамтын шүүлтүүрийн хамгийн чухал хэсэг юм. Хэрвээ хэрэглэгч A бүтээгдэхүүн X болон Y-г хамт үзсэн бол энэ нь эдгээр хоёр бүтээгдэхүүний хоорондох хослол бөгөөд дараа нь $C(2, 2)$ томъёог ашиглан хослол үүсгэж, эдгээрийг хамт санал болгох боломжтой.
- Граф суурьт алгоритмд комбинацийн ашиглалт: Граф суурьт хамтын шүүлтүүрт комбинацийн хослолуудыг ашиглан хэрэглэгч болон бүтээгдэхүүн хоорондын харилцааг нарийвчлан тодорхойлох боломжтой. Графт хэрэглэгчийн үйлдэл бүрийг хослон, эдгээрийн харилцааг өргөжүүлэн нарийн үзүүлэлтүүдийг тодорхойлж, зөвлөмжийн чанарыг сайжруулах боломжтой.
- Комбинацийн томъёо (C(n, k)) нь хамтын шүүлтүүрийн систем (CF)-д хэрэглэгч болон бүтээгдэхүүний хоорондын хослолуудын тоог тооцоолох, эдгээр хослолыг ашиглан хамгийн тохиромжтой санал болгох стратегиудыг үүсгэхэд хэрэглэгдэнэ.
- Хэрэглэгчийн зан төлөв, бүтээгдэхүүний хослолуудыг комбинацийн томъёо ашиглан нарийвчлан тооцоолж, хамгийн сайн зөвлөмжүүдийг гаргах боломжтой.

Энэ нь CF-ийн алгоритмуудад хэрэглэгчийн сонирхлыг илүү нарийвчлан ойлгох, хамгийн тохиромжтой бүтээгдэхүүнийг санал болгоход чухал үүрэг гүйцэтгэдэг.

- Борлуулалт нэмэгдүүлэх: Зөвлөмжийн системүүдийг компанийн хамгийн гол зорилго нь орлого олох явдал юм. Зөвлөмжийн систем нь хэрэглэгчдэд хамгийн тохиромжтой бүтээгдэхүүнийг санал болгосноор борлуулалтыг нэмэгдүүлдэг. Үүний зэрэгцээ, хэрэглэгчдийн идэвхтэй оролцоо болон вэбсайт дээрх сессийн хугацаа ч нэмэгдэж, хэрэглэгчид улам илүү оролцох боломжтой болно. Энэ нь цахим худалдааны платформуудад хэрэглэгчийн худалдан авалтын амжилттай түвшинг нэмэгдүүлэхэд нөлөөлдөг.
- Системийн ачааллыг бууруулах: Зөвлөмжийн сис-

тем нь хэрэглэгч бүрт хамгийн тохиромжтой бүтээгдэхүүнүүдийг шүүн сонгох замаар системийн ачааллыг бууруулж, хэрэглэгчдэд илүү хялбар, хурдан туршлага олгоно. Энэ нь зөвлөмжийн системийн гүйцэтгэлийг сайжруулахад тусалж, урт хугацаанд өртгийг бууруулахад туслана. Өгөгдөлд тулгуурласан шийдвэр гаргалт нь системийн ачааллыг бууруулж, илүү үр ашигтай үйл ажиллагааг баталгаажуулдаг.

- Хэрэглэгчийн оролцоо ба сэтгэл ханамжийг нэмэгдүүлэх: Зөвлөмжийн системүүд нь хэрэглэгчид төгс тохирсон, хувийн сонирхолд нийцсэн бүтээгдэхүүнийг тасралтгүй санал болгосноор хэрэглэгчдийн оролцоог үргэлжлүүлэн нэмэгдүүлдэг. Түүнчлэн, зөвлөмжийн системүүд нь хэрэглэгчийн туршлагаас илүү үр дүнтэй ашиглах боломжийг олгохоор, хаягдсан болон хэрэгцээгүй контентийг багасган, холбоотой контентыг дэмжих замаар хэрэглэгчийн сэтгэл ханамжийг нэмэгдүүлдэг.

IV. Туршилтын хэсэг

A. Сургалтын орчин

Бид туршилтыг Lenovo 21KUA004CD загварын зөөврийн компьютерт хийсэн. Intel(R) Core(TM) Ultra 5 125H 1.20 GHz хүчин чадалтай процессортой бөгөөд 32.0 GB RAM (31.5 GB ашиглах боломжтой) суурилагдсан. Үйлдлийн систем нь Windows 11 Pro (Version 24H2, OS Build: 26100.3476), 64-битийн бүтэцтэй.

B. Өгөгдөл бэлдэлт

Энэхүү судалгаанд бид Наран супермаркетын Наран Плейс салбарын 2025 оны 1-р сарын 1-нээс 3-р сарын 31-ний өдрийг хүртэлх хугацааны борлуулалтын мэдээллийг ашиглав. Нийт 120,231 мөр өгөгдөлд үндэслэн анализ хийсэн. Хэрэглэгчид нэг удаагийн худалдан авалтаар хамтад нь худалдан авсан бүтээгдэхүүнүүдийн хослолыг тодорхойлохын тулд $C(n, 2)$ (хослолын) томъёог ашиглаж, бүтээгдэхүүнүүдийг хоёр хоёрын хос болгон хөрвүүлэв. Үүссэн хос бүрийг графын ирмэг (edge) хэлбэрээр бүртгэж, давтамжийн үндсэн дээр жин (weight) өгч тэмдэглэв. Ингэснээр бүтээгдэхүүнүүдийн хоорондын хамаарлыг илүү нарийвчлалтай тусгах боломжтой болсон.

- ProductID: Гол бүтээгдэхүүний дугаар
- CopurchasedProductID: Тэр бүтээгдэхүүнтэй хамт хамгийн олон удаа худалдан авсан бүтээгдэхүүний дугаар

Өгөгдөл цэвэрлэх явцад давхардсан хослолуудыг нэгтгэн, хамгийн өндөр давтамжтай хослолуудыг хадгалсан. Бүтээгдэхүүн бүрт дахин давтагдахгүй ID оноож, графын зангилаа (node) хэлбэрт хөрвүүлэв. Дараа нь өгөгдлийг 80:20 харьцаагаар сургалт болон тестийн багцад хуваасан. Манайх сургалтанд ашиглсан өгөгдлийг .csv форматаар хадгалж судлаачид болон сонирхогчдыг чөлөөтэй ашиглах боломжийг бүрдүүлэх үүднээс уг дата өгөгдлийг GitHub [21] платформ дээр олон нийтэд нээлттэй байдлаар байршуулан хуваалцсан болно.

C. Үнэлгээний загвар

Бүтээгдэхүүн санал болгох системийн гүйцэтгэлийг үнэлэх нь хэрэглэгчийн туршлага, системийн үр ашиг, зөв таамаглалын чанарыг сайжруулахад чухал үүрэгтэй. Машин сургалтын аргууд, тухайлбал collaborative filtering ашиглаж буй үед системийн нарийвчлал, хариулах чадвар, болон нийт гүйцэтгэлийг үнэлэхийн тулд precision, recall, F1-score зэрэг гүйцэтгэлийн үзүүлэлтүүдийг өргөн ашигладаг.

- Recall нь хэрэглэгчийн өмнөх хандлага, сонирхлыг үндэслэн системийн санал болгосон бүтээгдэхүүнүүд дотор хэрэглэгчийн жинхэнэ сонирхолтой байсан бүтээгдэхүүн хэдийг зөв илрүүлсэн болохыг илтгэнэ. Энэ нь систем хэрэглэгчийн сонирхлыг бүрэн хамарч чадаж буй эсэхийг хэмжих үзүүлэлт юм.
- Precision нь системийн санал болгосон бүх бүтээгдэхүүнүүдээс хэд нь үнэхээр хэрэглэгчид тохирсон байсан бэ гэдгийг харуулдаг. Хэрэв систем олон буруу буюу хамааралгүй бүтээгдэхүүн санал болговол precision буурна.
- F1 score нь precision ба recall хоёрын тэнцвэрийг илэрхийлдэг нийлмэл үзүүлэлт бөгөөд системийн нийт гүйцэтгэлийг илүү бодитойгоор үнэлэх боломжийг олгодог. Precision болон recall-ын аль нэг нь хэт өндөр, нөгөө нь бага байх нөхцөлд F1 score нь дундаж утгаар буурдаг тул системийн ерөнхий чадварыг тэнцвэртэйгээр илтгэнэ.

Эдгээр үнэлгээний аргачлалууд нь хамтын шүүлтүүрийг ашиглан бүтээгдэхүүн санал болгох системийн чанар, оновчтой байдал, хэрэглэгчийн оролцооны чанарыг сайжруулахад чиглэгддэг. Ялангуяа граф суурьт хамтын шүүлтүүрийн загварууд нь хэрэглэгч болон бүтээгдэхүүний хоорондын харилцааг граф бүтэц ашиглан илэрхийлж, санал болголтыг оновчтойгоор гүйцэтгэх боломжийг олгодог.

Үнэлгээний үзүүлэлтүүдийн математик томъёо дараах байдалтай байна:

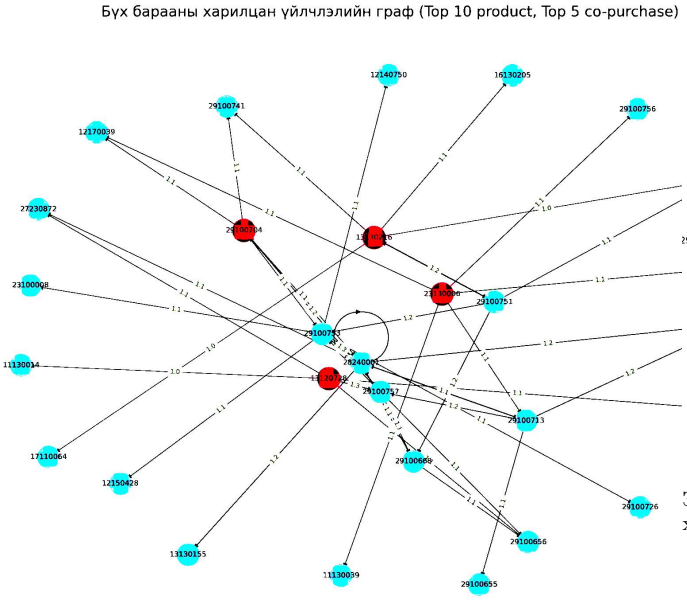
- Precision (Нарийвчлал):

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

  Хаана:
  - TP (True Positive): Зөв санал болгосон, үнэхээр сонирхсон бүтээгдэхүүн
  - FP (False Positive): Буруу санал болгосон, хэрэглэгч сонирхоогүй бүтээгдэхүүн
- Recall:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

  Хаана:



Зураг 4. Хамгийн их худалдан авалт хийгдсэн 10 бүтээгдэхүүн болоод тэдгээртэй хамт зарагдах бүтээгдэхүүнүүд

  - TP: Зөв санал болгосон, хэрэглэгч сонирхсон бүтээгдэхүүн
  - FN (False Negative): Систем санал болгоогүй ч хэрэглэгч сонирхсон бүтээгдэхүүн
- F1 Score (Тэнцвэржүүлсэн үзүүлэлт):

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

Жишээ: Хэрэглэгчид 5 бүтээгдэхүүн санал болгосон гэж үзье. Түүнээс 3 нь үнэхээр сонирхсон бүтээгдэхүүн байв. Гэвч хэрэглэгчийн сонирхсон нийт бүтээгдэхүүний тоо 6 байсан бол:

$$\text{Precision} = \frac{3}{5} = 0.6 \tag{6}$$

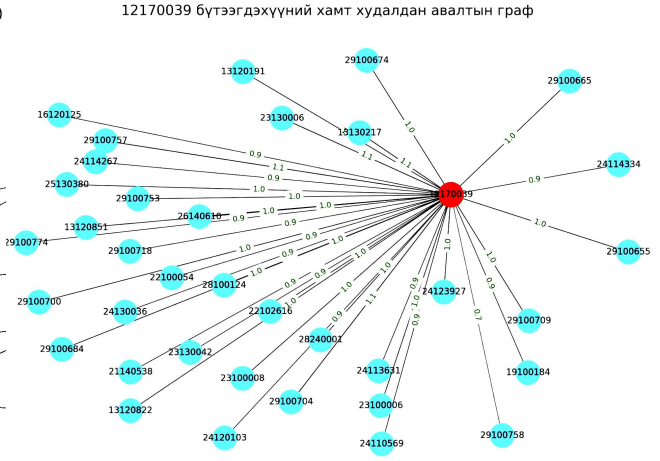
$$\text{Recall} = \frac{3}{6} = 0.5, \tag{7}$$

$$F_1 = 2 \cdot \frac{0.6 \cdot 0.5}{0.6 + 0.5} = 0.545 \tag{8}$$

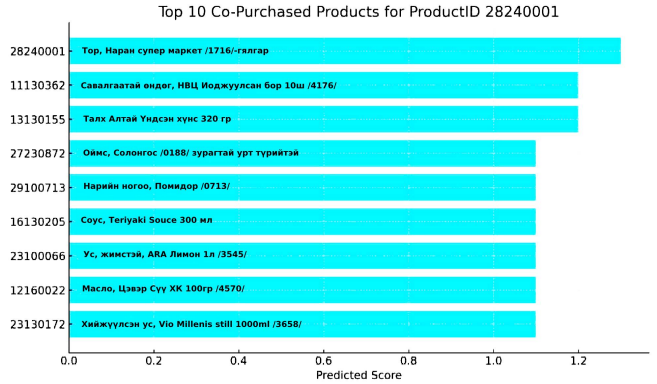
Иймд Precision, Recall, F1 score зэрэг үнэлгээний үзүүлэлтүүдийг хамтын шүүлтүүр ашиглан системийн гүйцэтгэлийг нарийвчлан хэмжих, сайжруулах, хэрэглэгчдэд илүү оновчтой бүтээгдэхүүн санал болгоход хэрэглэж болно.

Энд ProductID = 12170039-ийн хамт худалдан авалтын харилцааг графаар дүрслэн харууллаа:

- Улаан зангилаа — гол бүтээгдэхүүн (12170039).
- Цэнхэр зангилаа — холбогдсон бусад бүтээгдэхүүнүүд.



Зураг 5. 12170039 дугаартай Сүү 3.2%, Миний монголын 1л хамт худалдан авах өндөр магадлалтай бараануд



Зураг 6. 28240001 дугаартай TOP болон түүнтэй хамт зарагддаг бүтээгдэхүүн

- Ногоон тоо — ML.NET-ийн таамагласан хамт худалдан авах оноо (score).

Зураг 6-өөс харвал өндөг, талх, нарийн ногоо, терайки соус, ус, масло зэргийг хамт авбал өндөр магадлалтайгаар тор бичүүлдэг гэсэн дүгнэлт гаргаж болно. Мөн бид Соёрхол alc 36% 500gr /3298/ архийг хүмүүс ямар бараатай хамтатган авдаг болохыг сонирхож үзсэн бөгөөд дараах барааны жагсаалт гарч байна. Үүнд:

1) Самар, Cheetan honey roasted 150gr /3998/
2) Эрүүл мэндийн бэлдмэл,Цэнэг VITOS /0253/
3) Бохь, Dirol ledyanay /0812/
4) Ундаа, Coca cola 1,93л /5035/
5) Ундаа, Coca Cola 900мл/5882/
6) Цэвэр ус, Bonaqua 500мл

V. Дүгнэлт

Энэхүү судалгаагаар граф суурьт хамтын шүүлтүүр болон машин сургалтын аргуудыг хослуулан, хэрэглэгчийн зан төлөвт нийцсэн, нарийвчлал өндөртэй бүтээгдэхүүн санал болголтын системийг боловсруулллаа.

Уламжлалт хамтын шүүлтүүрийн аргачлалд тулгардаг өгөгдлийн ховор байдал болон шинэ хэрэглэгч, бүтээгдэхүүний мэдээллийн дутмаг байдлыг графын бүтэц болон гүнзгий сургалтын аргуудаар шийдвэрлэх боломжтойг нотолж, илүү ухаалаг, оновчтой борлуулалтын систем бүтээх үндэс тавигдлаа. Судалгааны үр дүн нь зөвхөн хэрэглэгчийн сэтгэл ханамжийг нэмэгдүүлэхээс гадна, дэлгүүрийн дотоод бүтээгдэхүүн өрөлт, санал болголтод стратегийн давуу тал бий болгох боломжтойг хэд хэдэн жишээн дээр харууллаа.

Ашигласан ном

[1] Adomavicius, G., Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering, 17(6), 734–749.

[2] Bobadilla, J., Serradilla, F., Bernal, J. (2010). A framework for collaborative filtering recommender systems. Expert Systems with Applications, 38(12), 14609–14623.

[3] Lops, P., De Gemmis, M., Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In F. Ricci, L. Rokach, B. Shapira, P. B. Kantor (Eds.), Recommender Systems Handbook (pp. 73–105). Springer.

[4] Burke, R. (2002). Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction, 12(4), 331–370.

[5] Bobadilla, J., Ortega, F., Hernando, A., Gutiérrez, A. (2013). Recommender systems survey. Knowledge-Based Systems, 46, 109–132.

[6] Zhang, S., Yao, L., Sun, A., Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. ACM Computing Surveys (CSUR), 52(1), 1–38.

[7] Wu, L., Sun, P., Fu, Y., Hong, R., Wang, X. (2020). A survey on graph neural networks for recommender systems. arXiv preprint arXiv:2011.02260.

[8] Zhang, S., Yao, L., Sun, A., Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. ACM Computing Surveys (CSUR), 52(1), 1–38

[9] Koren, Y., Bell, R., Volinsky, C. (2009). Matrix factorization techniques for recommender systems. Computer, 42(8), 30–37.

[10] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J. (1994). GroupLens: An open architecture for collaborative filtering of netnews. Proceedings of the 1994 ACM conference on Computer supported cooperative work, 175–186.

[11] Sarwar, B., Karypis, G., Konstan, J., Riedl, J. (2000). Application of dimensionality reduction in recommender system — a case study. WEBKDD.

[12] Zhang, S., Yao, L., Sun, A., Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. ACM Computing Surveys (CSUR), 52(1), 1–38.

[13] Wu, L., Sun, P., Zhang, Y., Tan, T. (2020). A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems, 32(1), 4–24.

[14] Sun, F., Liu, J., Wu, J., Pei, C., Lin, X., Ou, W., Jiang, P. (2020). Graph neural networks for social recommendation. The Web Conference (WWW).

[15] He, X., Liao, L., Zhang, H., Nie, L., Hu, X., Chua, T. S. (2017). Neural collaborative filtering. Proceedings of the 26th International Conference on World Wide Web, 173–182.

[16] Wang, X., He, X., Wang, M., Feng, F., Chua, T. S. (2019). Neural graph collaborative filtering. Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, 165–174.

[17] Wang, X., He, X., Wang, M., Feng, F., Chua, T. S. (2019). Neural graph collaborative filtering. Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, 165–174.

[18] Wang, H., Zhang, F., Xie, X., Guo, M. (2019). DKN: Deep knowledge-aware network for news recommendation. Proceedings of the 2018 World Wide Web Conference, 1835–1844.

[19] Ricci, F., Rokach, L., Shapira, B., Kantor, P. B. (Eds.). (2011). Recommender systems handbook. Springer.

[20] Su, X., Khoshgoftaar, T. M. (2009). A survey of collaborative filtering techniques. Advances in Artificial Intelligence, 2009, 421425.

[21] Enkhbaatar, S. (2025). $\text{sales}_data[Sourcecode].GitHub.https://github.com/Sergelen1005/sales_data$

5G NSA сүлжээн дэх чадлын үр ашигтай нөөц хуваарилалт

Ванданцэрэнгийн Нарантуяа
Мэдээллийн технологийн тэнхим
АБДС, ҮБХИС
Улаанбаатар, Монгол
vnarantuya0123@gmail.com

Найманнарангийн Чулуунбанди
Холбооны салбар
МХТС, ШУТИС
Улаанбаатар, Монгол
chuka@must.edu.mn

Ламжавын Эрдэнэбаяр
Холбооны салбар
МХТС, ШУТИС
Улаанбаатар, Монгол
erdenebayar.l@must.edu.mn

***Хураангуй*— Утасгүй холбооны сүлжээний хурдацтай өсөлт, ялангуяа 5G бие даасан бус (NSA) суурилуулалт нь спектрийн үр ашиг, интерференц, чадлын оновчлолыг сайжруулахын тулд олон цэгийн хүртэмжит дэвшилтэт техникийг шаардаж байна [1]. Rate-Splitting Multiple Access (RSMA) нь хэрэглэгчийн өгөгдлийг нийтлэг болон хувийн бүрэлдэхүүн хэсгүүдэд үр дүнтэй хуваах замаар ердийн Ортогональ бус олон цэгийн хүртэмжит арга(NOMA)-ын ирээдүйтэй хувилбар болж гарч ирсэн [2]. Энэхүү баримт бичиг нь хурдыг хуваах харьцаа болон чадлын хуваарилалтын хамтарсан оновчлолын загварыг боловсруулах замаар уруудах шугамын RSMA систем дэх нөөц хуваарилалтын асуудлыг авч үзсэн. Бид сонголт, кроссовер, мутаци зэрэг хувьслын үйлдлүүдээр дамжуулан оновчтой шийдлүүдийг давталттайгаар судалдаг Генетик Алгоритм (GA)-д суурилсан аргыг санал болгож байна. Үр дүн нь уламжлалт аргуудтай харьцуулахад чадлын үр ашиг 25%, спектрийн үр ашиг 20% сайжирсан байна. Нэмж дурдахад Жайны шударга байдлын индекс 0.9-тэй ойролцоо хэвээр байгаа нь хэрэглэгчдийн дунд нөөцийн тэгш хуваарилалтыг баталгаажуулдаг.**

Түлхүүр үгс—Спектрийн үр ашиг, Хурд хуваах харьцаа, интерференцийн менежмент, шударга байдал

I. Оршил

Утасгүй холбооны хэрэглээний эрчимтэй өсөлт, ялангуяа 4G-ээс 5G болон түүнээс цааш шилжих үед үүсэх хэт найдвартай, бага хоцролттой холбоо (URLLC), олон тооны хэрэглэгчийг нэгэн зэрэг холбох, спектрийн үр ашигийг нэмэгдүүлэх зэрэг эрэлт хэрэгцээг хангахын тулд олон цэгийн хүртэмжит дэвшилтэт схемүүдийг боловсруулах шаардлагатай болж байна [1,2]. Уламжлалт ортогональ олон цэгийн хүртэмжит (OMA) систем нь спектрийн нөөцийг хэрэглэгч бүрд тус тусад нь хуваарилдаг бөгөөд энэ нь спектрийг үр ашиггүй ашиглахад хүргэдэг. Энэхүү хязгаарлалтыг шийдвэрлэх зорилгоор ортогональ бус олон цэгийн хүртэмжит арга(NOMA)-ыг нэвтрүүлсэн бөгөөд энэ нь олон хэрэглэгчийн дохиог нягтруулга хийж, ижил давтамж болон хугацааны нөөцийг хуваалцах боломжийг олгодог. Гэвч NOMA нь хэрэглэгч хоорондын интерференцийн асуудалтай бөгөөд төгс бус сувгийн нөхцөлд дараалсан интерференцийг арилгах (SIC) төвөгтэй арга техникийг шаарддаг тул системийн гүйцэтгэл буурах хязгаарлалттай байдаг [3,4].

Rate-Splitting Multiple Access (RSMA) нь NOMA-аас илүү уян хатан, бат бөх олон цэгийн хүртэмжит аргачлал болж гарч ирсэн бөгөөд хэрэглэгчдийн мессежийг нийтлэг болон хувийн хэсэгт хуваах замаар сайжруулсан интерефернцийн менежментийг санал болгодог [5]. Энэхүү мессежийг хуваах механизм нь нэг хэсгийг олон хэрэглэгч зэрэг тайлах боломжийг олгодог (нийтийн урсгал) бол үлдсэн хэсгийг хэрэглэгч бүр тусдаа (хувийн урсгал) нууцаар тайлах боломжтой байдаг. Энэ эрлийз арга нь спектрийн үр ашигийг нэмэгдүүлж, хэрэглэгчдийн дунд шударга байдлыг сайжруулан, сувгийн тооцооллын алдааны эсрэг NOMA-тай харьцуулахад илүү бат бөх байдлыг хангадаг [6,7]. RSMA-д суурилсан системүүдийн гол зорилтын нэг нь хамгийн бага SINR болон хэрэглэгчдийн шударга байдлын шаардлагыг хангахын зэрэгцээ чадлын үр ашигийг нэмэгдүүлэхийн тулд хурд хуваах харьцаа болон чадлын хуваарилалтыг оновчтой болгох явдал юм [8]. Хурд хуваах харьцаа нь нийтлэг болон хувийн мессежүүдэд хуваарилагдах дамжуулалтын чадлыг тодорхойлдог. Энэ харьцааг буруу сонгосноор системийн дамжуулах чадвар буурч, хэрэглэгч хоорондын интерференц нэмэгддэг учраас 5G-ийн бие даасан бус (NSA) сүлжээн дэх практик хэрэглээнд динамик дасан зохицох чадвар нь онцгой чухал юм [9,10].

Чадлын хуваарилалт нь RSMA-ийн гүйцэтгэлийг сайжруулахад чадлын үр ашиг болон хэрэглэгчдийн шударга байдлыг тэнцвэржүүлэх чухал үүрэгтэй байдаг. Бааз станцын нийт чадлыг нийтлэг болон хувийн урсгалд оновчтой хуваарилах замаар бүх хэрэглэгчийн хамгийн бага SINR шаардлагыг хангахын зэрэгцээ спектрийн ашиглалтыг хамгийн өндөр түвшинд хүргэх ёстой [11]. RSMA-д суурилсан чадлын хуваарилалтыг оновчтой болгоход уламжлалт програмчлал, water-filling алгоритм зэрэг аргуудыг ашиглаж ирсэн боловч бодит сүлжээн дэх асуудлуудыг шийдвэрлэхэд хүндрэлтэй тулгарч байна. Эдгээр тооцооллын сорилтуудыг даван туулахын тулд сүүлийн үед RSMA системийн чадал болон хурд хуваарилах асуудалд метахевристик оновчлолын арга, тэр дундаа Генетикийн алгоритмууд (GA)-ыг өргөнөөр судлах болсон. GA нь шалгаруулах, кроссовер, мутаци зэрэг хувьслын үйлдлүүдийг ашиглан шийдлийг алхам алхмаар боловсронгуй, бат бөх бөгөөд дасан зохицох чадвартай оновчлолын тогтолцоог бүрдүүлдэг. Иймээс GA нь детерминист аргуудаас ялгаатай нь шийдлийн нарийн төвөгтэй хайлтын орон зайд үр ашигтай ажиллаж чаддаг тул бодит цагийн 5G NSA сүлжээнд үүсэх RSMA-ийн нөөц хуваарилалтын асуудлуудад илүү тохиромжтой юм.

Энэ өгүүлэлд RSMA (Rate-Splitting Multiple Access)-д суурилсан утасгүй сүлжээний нөөц хуваарилалтын асуудлыг генетик алгоритм (GA)-аар оновчтой болгох аргыг санал болгож байна. RSMA нь нийт дохиог common stream ба private stream гэсэн хоёр хэсэгт хуваадаг. Common stream нь бүх хэрэглэгчдэд зориулагдсан мэдээллийг дамжуулдаг бол private stream нь зөвхөн тухайн хэрэглэгчийн хувийн мэдээллийг агуулдаг. Дамжуулах тал (TX)-д энэхүү хуваагдлыг ашиглан илүү уян хатан нөөц хуваарилалт хийх боломжтой бөгөөд хүлээн авах тал (RX)-д decoding процессыг тодорхой дарааллын дагуу гүйцэтгэнэ.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Энэхүү судалгааны гол зорилго нь RSMA систем дэх хурд хуваах харьцаа болон чадлын хуваарилалтыг хамтад нь оновчтой болгох GA-д суурилсан оновчлолын системийг боловсруулахад оршино. Санал болгож буй аргыг 5G NSA-ийн бодит орчинд үнэлж, чадлын үр ашиг болон спектрийн үр ашгийн хувьд уламжлалт NOMA болон Simulated Annealing (SA)-д суурилсан аргуудаас илүү үр дүнтэй болохыг харуулсан.

II. Холбогдох ажлууд

5G Non-Standalone (NSA) сүлжээ бий болсноор нөөцийн үр ашигтай хуваарилалт нь спектрийн өндөр үр ашиг, бага хоцролт, чадлын үр ашгийг хангах судалгааны чухал чиглэл болж байна [12]. Бие даасан (SA) 5G-ээс ялгаатай нь NSA 5G нь одоо байгаа 4G LTE дэд бүтцийг ашиглахын зэрэгцээ, 5G New Radio (NR) технологийг нэвтрүүлэн гүйцэтгэлийг сайжруулдаг. Энэхүү хосолсон сүлжээний архитектур нь спектрийн хуваарилалт, чадлын хуваарилалт, хэрэглэгчдийн шударга байдал болон интерференцийн менежмент зэрэгт шинэ сорилтуудыг бий болгож байна [13]. Маш олон судалгаанууд 5G NSA-ийн нөөц хуваарилалтыг оновчтой болгох арга техникүүдийг санал болгож байгаа бөгөөд үүнд машин сургалт (ML)-д суурилсан шийдэл, тоглоомын онолд суурилсан оновчлол, Генетик алгоритм (GA), бөөмийн бөөгнөрлийн оновчлол (Particle Swarm Optimization, PSO) зэрэг метахевристик аргууд, мөн бататгах сургалт (Reinforcement Learning, RL) зэрэг хиймэл оюун ухаанд суурилсан аргууд орж байна [14].

Спектрийн хуваарилалт нь 5G NSA сүлжээний үр ашгийг нэмэгдүүлэхэд чухал үүрэгтэй. Урьдчилан тодорхойлсон зурвасын өргөнийг хэрэглэгч бүрд тогтмол хуваарилах зэрэг хэд хэдэн аргыг судалсан бөгөөд эдгээр нь динамик орчинд үр ашиггүй байдлыг бий болгож байсан [15]. Тиймээс LTE болон 5G NR-д спектрийн үр ашгийг дээшлүүлэх зорилгоор Динамик спектр хуваалцах (DSS)-ыг нэвтрүүлсэн бөгөөд энэ нь спектрийг нөхцөл байдалд тохируулан динамикаар хуваарилах боломжийг олгодог [16]. Сүүлийн үед Reinforcement Learning (RL) болон Deep Q-Networks (DQN) зэрэг хиймэл оюун ухаанд суурилсан аргууд нь спектрийн менежментийн өндөр үр дүнг харуулж, спектрийг оновчтой таамаглах, хуваарилах боломжийг олгож байна [17]. Чадал хуваарилалт нь 5G NSA сүлжээний гүйцэтгэл болон чадал зарцуулалтад шууд нөлөөлдөг. Уламжлалт Water-filling algorithm зэрэг аргууд нь сувгийн нөхцөлд тулгуурлан чадлыг оновчтой хуваарилдаг тул MIMO системд илүү тохиромжтой байдаг [18]. Харин тоглоомын онолд суурилсан аргууд нь Нэшийн тэнцвэрийн зарчмыг ашиглан олон хэрэглэгчтэй орчинд чадлын шударга хуваарилалтыг баталгаажуулдаг [19]. Мөн сүүлийн үед мэдрэлийн сүлжээ (Neural Network, NN)-д суурилсан гүнзгий сургалтын аргуудыг чадлын үр ашгийг дээшлүүлэх зорилгоор өргөнөөр судалж байна [20]. Хэрэглэгчийн хуваарь нь 5G NSA-ийн нөөц хуваарилалтын бас нэг чухал тал бөгөөд үйлчилгээний чанар (QoS), нөөцийн шударга хуваарилалтыг хангахад чиглэдэг. Уламжлалт Round Robin арга нь бүх хэрэглэгчдэд ижил хугацааг хуваарилдаг ч сүлжээний ачаалал өндөр үед үр ашиггүй байдаг [21]. Харин Пропорциональ шударга хуваарилалт (Proportional Fair Scheduling, PFS) зэрэг илүү дэвшилтэт аргууд нь сувгийн нөхцөл байдал, хэрэглэгчийн шаардлагад тулгуурлан динамикаар нөөц хуваарилдаг [22]. Сүүлийн үеийн судалгаагаар Reinforcement Learning-д суурилсан хуваарийн алгоритмууд нь хэрэглэгчийн бодит хугацааны эрэлт хэрэгцээнд дасан зохицож, сүлжээний ерөнхий үр ашгийг сайжруулдаг болохыг харуулсан байна [23].

Оновчлолын техникүүд нь 5G NSA сүлжээнд нөөц хуваарилалтын үр ашгийг нэмэгдүүлэхэд чухал үүрэгтэй. Генетикийн алгоритмууд (GA) нь төвөгтэй, шугаман бус оновчлолын асуудлуудыг үр дүнтэй шийдвэрлэж, бараг оновчтой шийдлийг олох чадвартай тул спектр болон чадлын хуваарилалтыг сайжруулахад өргөнөөр ашиглагдаж байна. GA-г олон хэрэглэгчийн чадлын хуваарилалт, цацраг үүсгэх оновчлол, сүлжээний зүсэлт зэрэг төрөл бүрийн асуудлуудад хэрэглэсэн бөгөөд уламжлалт аргуудаас илүү үр дүнтэй болохыг харуулсан [24]. Өөр нэг ирээдүйтэй арга бол бодит хугацааны сүлжээний нөхцөлд дасан зохицох нөөц хуваарилалт хийх боломжтой Reinforcement Learning юм. Түүний дэвшилтэт хэлбэр болох Deep Reinforcement Learning нь хүрээлэн буй орчны хариу үйлдлээс тасралтгүй суралцах замаар нэг төрлийн бус (heterogeneous) сүлжээнд нөөцийг оновчтой хуваарилахад онцгой үр дүн үзүүлсэн [25]. Мөн Multi-Agent Reinforcement Learning аргачлалыг олон үүрийн 5G NSA сүлжээнд чадал болон спектрийн нөөцийг хамтран хуваарилахад судалсан бөгөөд энэ нь системийн ерөнхий гүйцэтгэл болон шударга байдлыг сайжруулахад үр дүнтэй байсан. Түүнчлэн, тоглоомын онолд суурилсан загваруудыг хэрэглэгчдийн шударга өрсөлдөөнийг хангах, сүлжээний гүйцэтгэлийг сайжруулахад ашигладаг. Нэшийн тэнцвэрт суурилсан загварууд нь олон хэрэглэгчдийн дунд чадал болон зурвасын өргөнийг шударгаар, үр ашигтай хуваарилахад тусалдаг. Эвслийн тоглоомын загварууд нь спектрийн үр ашгийг дээшлүүлж, интерференцийг багасгахын тулд 5G NSA сүлжээний динамик спектр хуваарилалтыг дэмжих зорилгоор сүлжээний зүсэлтэнд ашиглагдаж байна.

Эвристикт суурилсан аргуудаас RSMA-д чадлын хуваарилалт болон хурд хуваах харьцааг оновчтой болгоход Генетикийн алгоритм (GA) болон Бөөмийн бөөгнөрлийн оновчлол (Particle Swarm Optimization, PSO)-ыг өргөн хэрэглэдэг. [26]-д хийсэн судалгаанд RSMA-ийн дасан зохицох чадлын хуваарилалтад GA-д суурилсан аргачлалыг санал болгосон бөгөөд энэ нь бүрэн хайлтын аргуудтай (exhaustive search) харьцуулахад тооцооллын төвөгтэй байдлыг бууруулж, бараг оновчтой шийдлийг олох боломжийг харуулсан. Өөр нэг судалгаанд [27] хэрэглэгчдийн холбоо (user association) болон чадлын хяналтыг (power control) зохицуулахад PSO-г ашигласан бөгөөд үүгээр нэг төрлийн бус сүлжээн дэх хэрэглэгчдийн холболтын хурд болон чадлын үр ашгийг нэмэгдүүлсэн. Мөн сүлжээний зүсэлт (network slicing) болон програм хангамжаар тодорхойлсон сүлжээ (Software-defined networking, SDN)-г 5G сүлжээнд динамик нөөц хуваарилалтын үр дүнтэй шийдлийг санал болгож байна. Сүлжээний зүсэлт нь операторуудад өөр өөр хэрэглээний шаардлагад нийцүүлэн виртуалчлагдсан сүлжээний сегмент үүсгэх боломжийг олгодог бөгөөд ингэснээр үйлчилгээний чанарыг (QoS) сайжруулдаг [28]. SDN болон сүлжээний зүсэлтийг хослуулснаар төвлөрсөн удирдлагын тогтолцоо бий болж, бодит хугацааны хөдөлгөөний хяналт болон ухаалаг хуваарилалтыг хэрэгжүүлэх боломжтой болсон [29]. Түүнчлэн SDN-ийг хиймэл оюун ухаанд суурилсан нөөц хуваарилалтын аргуудтай нэгтгэх

нь сүлжээний ашиглалт болон үйлчилгээний найдвартай байдлыг эрс сайжруулсан болохыг судалгаагаар баталсан [30,31]. Эдгээр дэвшил гарч байгаа хэдий ч 5G NSA сүлжээнд RSMA-д суурилсан нөөц хуваарилалтыг бодит цагт хэрэгжүүлэхэд тулгардаг хэд хэдэн асуудал бий. Тухайлбал, тооцооллын хүндрэлтэй байдал болон бодит хугацааны дасан зохицох чадвар хоёрын хоорондын зохистой харьцааг олох шаардлагатай. Олон оновчлолын аргууд нь их хэмжээний тооцооллын нөөц шаардаг тул том хэмжээний 5G сүлжээнд бодит цагийн хэрэглээнд тохиромжгүй байдаг. Үүнээс гадна үүр хоорондын интерференц, сувгийн төлөв байдлын мэдээллийн төгс бус байдал (CSI), техник хангамжийн хязгаарлалт зэрэг нь нөөцийн хуваарилалтын алгоритмын гүйцэтгэлийг бууруулдаг хүчин зүйлс юм [32]. Тиймээс ирээдүйн судалгаагаар тооцооллын үр ашиг болон оновчтой гүйцэтгэлийн тэнцвэрийг хангахын тулд хиймэл оюун ухаанд суурилсан сургалтын загваруудыг уламжлалт оновчлолын техниктэй хослуулсан хосолсон аргуудад анхаарах шаардлагатай.

Одоогийн судалгааны үр дүнгээс үзэхэд 5G NSA сүлжээнд зориулсан нөөц хуваарилалтын олон аргачлалууд судлагдсан бөгөөд эдгээрээс RSMA нь уламжлалт NOMA-гийн хувилбар болгон өндөр потенциалтай болох нь тогтоогдож байна. Ялангуяа Генетикийн алгоритм (GA), Reinforcement Learning, сүлжээний зүсэлт зэргийг нэгтгэсэн аргачлал нь чадлын үр ашиг, спектрийн үр ашиг, хэрэглэгчдийн шударга байдлыг сайжруулах боломжийг харуулсан. Гэхдээ цаашдын судалгаагаар RSMA-д суурилсан нөөц хуваарилалтыг бодит хугацааны хэрэглээнд өргөтгөх боломжийг баталгаажуулж, практик байршуулалтын сорилтуудыг шийдвэрлэх шаардлагатай байна.

III. Системийн загвар

A. *RSMA системийн тойм*

Энэхүү судалгаанд авч үзсэн Rate-Splitting Multiple Access системийн загвар нь олон антенаар тоноглогдсон нэг бааз станц нь хуваалцсан давтамжийн зурвасаар олон нэг антентай хэрэглэгчдэд үйлчилдэг 5G NSA сүлжээнд зориулагдсан. RSMA нь уламжлалт олон цэгийн хүртэмжит аргуудаас ялгаатай нь RSMA нь спектрийн үр ашгийг дээшлүүлж, интерференцийг үр дүнтэй удирдахын тулд дамжуулагдсан өгөгдлийг нийтлэг болон хувийн урсгалд динамик байдлаар хуваадаг.

Санал болгож буй системд хэрэглэгч бүр нийтлэг болон хувийн мессежээс бүрдсэн давхардсан дохиог хүлээн авдаг. Нийтлэг мессежийг бүх хэрэглэгчид тайлдаг бол хувийн мессежийг тухайн хэрэглэгч тус бүр дангаар нь тайлдаг. Эдгээр бүрэлдэхүүн хэсгүүдэд чадлын хуваарилалт болон хурдыг хуваах харьцаа нь системийн ерөнхий гүйцэтгэлд шууд нөлөөлдөг. Судалгаанд GA-г ашиглан RSMA-ийн дамжуулах хүч болон stream хуваарилалтыг оновчтой болгох зорилго тавьсан. GA нь олон шийдлийн дундаас хамгийн оновчтойг сонгох, хувьслын зарчмаар сайжруулах давуу талтай. Энэхүү аргачлалын хүрээнд RSMA-ийн системийн загвар, сувгийн нөхцөл, хэрэглэгчийн тоо зэрэг үзүүлэлтүүдийг харгалзан (1)-(13) томъёог дахин боловсруулсан. Тухайлбал, [конкрет өөрчлөлтүүдийг оруулна] гэх мэтээр GA-ийн хувьслын процесс томъёонд хэрхэн нөлөөлснийг тусгасан.

B. *Дохионы загвар*

i-р хэрэглэгчийн хүлээн авсан дохиог дараах байдлаар илэрхийлж болно.

$$y_i = h_i^H x + n_i \tag{1}$$

Энд:

- h_i^H нь BS ба i хэрэглэгчийн сувгийн өсгөлтийн векторыг илэрхийлнэ,
- x нь BS-ээс нэвтрүүлсэн дохио
- $n_i \approx CN\left(0, \sigma^2\right)$ нь σ2 варианстай хүлээн авагч дээрх Гауссын цагаан шуугианыг (AWGN) илэрхийлнэ..

Нэвтрүүлсэн дохио x нь нийтлэг болон хувийн мессежүүдийн шугаман хослол юм:

$$x = w_c s_c + \sum_{i=1}^{N} w_i s_i \tag{2}$$

Энд:

- sc нь цацраг үүсгэх вектор wc бүхий нийтлэг мессеж юм,
- si нь цацраг үүсгэх вектор wi бүхий i хэрэглэгчийн хувийн мессеж юм,

Дамжуулах нийт чадлыг $P_{total} \leqslant P_{max}$-аар хязгаарласан. нөлөөлдөг.

C. *Хурд хуваах, код тайлах дараалал*

Хэрэглэгч бүр хувийн мессежийг дуу чимээ гэж үзэх замаар эхлээд нийтлэг мессежийг тайлдаг. Хэрэглэгч ii дээрх нийтлэг мессежийг тайлах боломжтой хурдыг өгсөн:

$$R_{c,i} = \log_2\left(1 + \frac{\left|h_i^H w_c\right|^2}{\sum_{j=1}^{N}\left|h_i^H w_j\right|^2 + \sigma^2}\right) \tag{3}$$

Нийтлэг мессежийг бүх хэрэглэгчид амжилттай тайлах ёстой бөгөөд ингэснээр нийтлэг зурваст хуваарилагдсан нийт хувь хэмжээ болно:

$$R_c = \min_{i \in N} R_{c,i} \tag{4}$$

Нийтлэг мессежийг тайлж, дараалсан интерференцг арилгах (SIC) -аар хассаны дараа хэрэглэгч бүр өөрийн хувийн мессежийг хүрч болох хурдаар тайлдаг.

$$R_i = \log_2\left(1 + \frac{\left|h_i^H w_i\right|^2}{\sum_{j \neq i}\left|h_i^H w_j\right|^2 + \sigma^2}\right) \tag{5}$$

Ингээд системийн нийт нийлбэр хурд (дамжуулах хурд) нь дараах байдлаар тодорхойлогдоно.

$$R_{sum} = R_c + \sum_{i=1}^{N} R_i \tag{6}$$

Томъёо 3, 5-д логарифм функц ашигласан нь Shannon-ийн мэдээллийн онолын суурь томъёололд үндэслэсэн. Энэ нь өгөгдлийг битээр хэмжих үед суурь 2 логарифм ашиглах шаардлагатай болгодог. Зарим судалгаанд лог функцгүйгээр илэрхийлсэн нь оновчлолын загварын ялгаатай байдлаас шалтгаалдаг.

D. RSMA дахь чадлын хуваарилалт

Нийт дамжуулах чадлыг нийтлэг болон хувийн урсгалуудын хооронд хуваарилах шаардлагатай.

$$P_{total} = P_c + \sum_{i=1}^{N} P_i \tag{7}$$

Энд Pc нь нийтлэг урсгалд хуваарилагдсан чадал, Pi нь хэрэглэгчийн ii хувийн урсгалд хуваарилагдсан чадал юм. Оновчлолын асуудал нь үйлчилгээний чанарын (QoS) хязгаарлалтыг хангахын зэрэгцээ чадлын үр ашгийг дээд зэргээр нэмэгдүүлэх чадлын оновчтой хуваарилалтыг тодорхойлоход чиглэгддэг.

RSMA дахь нөөц хуваарилалтын үндсэн зорилго нь шударга байдлын хязгаарлалтыг хангахын зэрэгцээ системийн чадлын үр ашгийг (EE) нэмэгдүүлэхийн тулд хурд хуваах харьцаа α ба чадлын хуваарилалтыг шийдвэрлэх явдал юм. Шийдвэрлэх асуудлыг дараах байдлаар томъёолж болно:

$$\max_{\alpha, P_c, P_i} EE = \frac{R_{sum}}{P_{total}} \tag{8}$$

RSMA нь дараах хязгаарлалтуудыг хангана:

- Чадлын хязгаарлалт

$$P_c + \sum_{i=1}^{N} P_i \le P_{\max} \tag{9}$$

- Хамгийн бага хурдны хязгаарлалт:

$$R_i \ge R_{\min}, \quad \forall i \in \{1, 2, ..., N\} \tag{10}$$

- Шударга байдлын хязгаарлалт: Хуваарилалт нь бүх хэрэглэгчдийн дунд шударга байж, нөөц хуваарилалтын тэнцвэрт байдлыг хангах ёстой.

$$J = \frac{\left(\sum_{i=1}^{N} R_i\right)^2}{N \sum_{i=1}^{N} R_i^2} \tag{11}$$

Энд J нь Жэйний шударга байдлын индекс бөгөөд үүнийг нэмэгдүүлэх шаардлагатай.

- Үйлчилгээний чанарын (QoS) хязгаарлалт: Нэг хэрэглэгчдэд хэт их хуваарилалт хийхээс сэргийлэхийн тулд хэрэглэгч бүрт хуваарилсан хувь хэмжээ нь урьдчилан тодорхойлсон дээд хязгаараас хэтрэхгүй байх ёстой.

$$R_i \le R_{\max}, \quad \forall i \in \{1, 2, ..., N\} \tag{12}$$

- Интерференцийн удирдлага: Сүлжээний тогтвортой байдлыг хангахын тулд бүх хувийн урсгалын интерференцийн нийлбэр нь босго утгаас хэтрэхгүй байх ёстой.

$$\sum_{j \ne i} \left| h_i^H w_j \right|^2 \le I_{\max}, \qquad \forall i \in \{1, 2, ..., N\} \tag{13}$$

Санал болгож буй системийн загвар нь чадлын хуваарилалт болон хурд хуваах харьцааг оновчтой болгосноор 5G NSA сүлжээнд RSMA нь илүү ашигтай. Энэхүү томъёолол нь чадлын хязгаарлалт, шударга байдал, үйлчилгээний хамгийн бага шаардлагууд зэрэг бодит хязгаарлалтуудыг авч үздэг. Дараагийн хэсэгт асуудлын томъёоллыг танилцуулж, RSMA-г идэвхжүүлсэн 5G NSA сүлжээнд нөөцийн оновчтой хуваарилалтад хүрэх оновчлолын аргуудыг хэлэлцэнэ.

IV. Нөөц хуваарилалтын асуудал

RSMA-д суурилсан 5G NSA сүлжээнд хэрэглэгч ба сувгийн хуваарилалт нь спектрийн үр ашгийг нэмэгдүүлэх, шударга байдлыг хангахад чухал үүрэгтэй. Хүчтэй болон сул хэрэглэгчдийг сувгийн үр ашгийн дагуу хослуулдаг NOMA системээс ялгаатай нь, RSMA нь чадлын хуваарилалт болон хурд хуваах харьцааг динамикаар тохируулснаар хэрэглэгчийн сонголтыг оновчтой болгодог. Энэхүү сонголтын үйл явц нь интерференцийн түвшин, дохионы чанар, өгөгдлийн хурдны шаардлагыг харгалзан, хэрэглэгчдийг тодорхой дэд сувгуудад хуваарилах хосолсон оновчлолын асуудлаар томьёологддог.

Хэрэглэгчийн сувгийн оновчтой хуваарилалтад хүрэхийн тулд Генетик Алгоритм (GA) дээр суурилсан оновчлолын тогтолцоог ашигладаг бөгөөд энэ нь хуваарилалтыг дараах байдлаар дахин сайжруулдаг.

- Санамсаргүй байдлаар хэрэглэгчийн сувгийн даалгаврыг эхлүүлэх.
- Спектрийн үр ашиг, шударга байдалд үндэслэн фитнессийн функцийг үнэлэх.
- Сонголт, огтлолцол (crossover), мутаци (mutation) үйлдлүүдийг гүйцэтгэж, илүү сайн хуваарилалтыг хөгжүүлэх.
- Оновчтой эсвэл оновчтой хуваарилалтад ойртохыг баталгаажуулах.

GA-д суурилсан хэрэглэгчийн хуваарилалтыг нэгтгэснээр RSMA системүүд нь нөөцийг динамикаар хуваарилж, дамжуулах чадварыг нэмэгдүүлэхийн

зэрэгцээ олон хэрэглэгчийн орчинд хэрэглэгчийн шударга байдлыг хадгалах боломжтой.

A. *Нийтлэг ба хувийн урсгалын чадлын хуваарилалт*

RSMA дахь чадлын хуваарилалт нь ердийн олон цэгийн хүртэмжит схемээс эрс ялгаатай бөгөөд энэ нь хэрэглэгч бүрийн нийтлэг болон хувийн мессежийн бүрэлдэхүүн хэсгүүдийн хооронд чадалийг хуваарилдаг. Оновчлолын зорилго нь хэрэглэгч бүр өөрийн хамгийн бага SINR шаардлагыг хангасан байхын зэрэгцээ системийн чадлын үр ашгийг нэмэгдүүлэх явдал юм.

Бааз станц (БС) дээрх нийт дамжуулах чадлыг дараах байдлаар хуваарилна:

$$P_{total} = P_c + \sum_{i=1}^{N} P_i \tag{14}$$

Энд:

- Компьютер нь нийтлэг урсгалд хуваарилагдсан чадал (бүх хэрэглэгчдийн кодыг тайлсан),
- Pi нь i хэрэглэгчийн хувийн урсгалд хуваарилагдсан чадал юм.

GA-д суурилсан оновчлолын арга нь дасан зохицох чадлын хуваарилалтын стратегийг баталгаажуулдаг бөгөөд үүнд:

- Өндөр интерференцтэй хэрэглэгчид нийтлэг урсгалын чадлын илүү их хэсгийг авдаг.
- Хувийн урсгалын хүчийг хувь хүний SINR нөхцөл дээр үндэслэн хуваарилдаг.
- Оновчлолын процесс нь чадлын хэрэглээг багасгахын зэрэгцээ damжуулах чадварыг дээд зэргээр нэмэгдүүлдэг.

Энэхүү дасан зохицох хуваарилалт нь NOMA-тай харьцуулахад илүү сайн интерференцийн удирдлагыг хангадаг бөгөөд чадлын хуваарилалт нь хэрэглэгчийн тэргүүлэх түвшний түвшинд тулгуурладаг.

B. *Хурд хуваах харьцааг оновчтой болгох*

RSMA-ийн гол сорилтуудын нэг бол дохионы кодыг оновчтой тайлах, интерференцийн удирдлагыг хангахын тулд хурд хуваах харьцааг оновчтой болгох явдал юм. Хурд хуваах харьцаа α нь damжуулагдсан өгөгдлийг нийтлэг болон хувийн бүрэлдэхүүн хэсгүүдэд хэрхэн хуваахыг тодорхойлдог:

$$R_i = \alpha_i R_c + (1-\alpha_i) R_i^{private} \tag{15}$$

Энд:

- α_i нь нийтлэг урсгалд хуваарилагдсан нийт ханшийн хувийг илэрхийлнэ.
- R_c нь нийтлэг мессежийн хурд юм.
- $R_i^{private}$ нь i хэрэглэгчийн хувийн мессежийн хурд юм.

Оновчлолын асуудал нь шударга байдлыг хангахын зэрэгцээ нийлбэр хурдыг хамгийн их байлгах α_i -ийн оновчтой утгуудыг тодорхойлох явдал юм. GA-д суурилсан сайжруулалт нь янз бүрийн хуваах харьцааг давтаж, тэдгээрийн нөлөөллийг үнэлдэг:

- Спектрийн үр ашиг – интерференцийг бууруулахаас хэрэглэгчдэд ашиг тусаа өгөх.
- Чадлын үр ашиг – урсгал damжих чадлын хэрэглээг зохицуулах.
- Шударга байдал – хэрэглэгчдийн дунд өгөгдлийн хурдыг тэгш байлгах.

RSMA дахь нөөц хуваарилалтын асуудал нь гурван бүрэлдэхүүн хэсгээс бүрдэнэ.

- Хэрэглэгч ба сувгийн хуваарилалт: GA ашиглан хэрэглэгчдийг сувгууд руу динамикаар холбох.
- Чадлын хуваарилалт: Нийтлэг болон хувийн урсгалын хооронд чадлын дасан зохицох хуваарилалт.
- Хурд хуваах харьцааг оновчтой болгох: Спектрийн үр ашиг ба шударга байдлыг тэнцвэржүүлэхийн тулд параметрийг нарийвчлан тохируулах.

Энэхүү судалгаа нь генетикийн алгоритмд суурилсан аргыг ашигласнаар 5G NSA сүлжээн дэх RSMA системүүд нь уламжлалт NOMA болон OFDMA техникүүдээс илүү чадлын үр ашиг, илүү шударга байдал, спектрийн өндөр үр ашигтай байдлыг хангадаг. Дараагийн хэсэгт санал болгож буй оновчлолын хүрээ, түүний хэрэгжилтийг танилцуулна.

V. Симуляцийн үр дүн

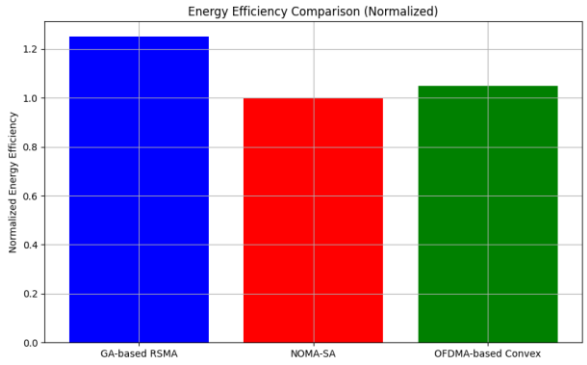
5G NSA дахь RSMA-д зориулсан GA-д суурилсан нөөц хуваарилалтын гүйцэтгэлийг үнэлэхийн тулд олон хэрэглэгчийн уруудах шугамын системийг ашиглан өргөн хүрээний симуляци хийсэн. Симуляци нь 5G сүлжээний орчинд хийгдсэн бөгөөд хэрэглэгчид нь random location-тэй, MIMO антенны систем ашигласан. Симуляцийн параметрүүдийг 5G сүлжээний хувьд тохируулсан:

Хүснэгт I. Тооцооллын параметр

№	*Параметр*	*Утга*
1	Зөөгч давтамж	3.5 GHz
2	Системийн зурвасын өргөн	100 MHz
3	Бааз станцын антен	8 нэвтрүүлэх антен
4	Хэрэглэгчийн тоо	10 single-antenna users
5	Шуугианы чадлын нягт	-174 dBm/Hz
6	Хамгийн их нэвтрүүлэх чадал	40 dBm
7	Замын алдагдлын загвар	3GPP Urban Macro-cell (UMa) model
8	Замхралтын загвар	Релейн замхралт
9	GA параметр	Population size = 50, Generations = 200, Crossover rate = 0.8, Mutation rate = 0.1

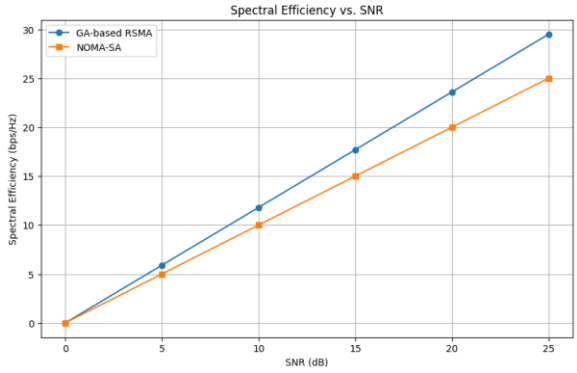
Санал болгож буй нөөц хуваарилалтын схемийг олон үзүүлэлтээр иж бүрэн үнэлгээ хийсэн.. Энергийн үр ашгийг (EE) нэгж чадлын хэрэглээнд ногдох нийт

нийлбэрээр хэмжсэн нь системийн чадлын үр ашгийн талаарх ойлголтыг өгдөг. Спектрийн үр ашгийг (SE) нэгж зурвасын өргөнд ногдох нийт дамжуулах чадвараар шинжилсэн нь системийн өгөгдөл дамжуулах чадварыг дээд зэргээр харуулсан. Шударга байдлын индексийг (FI) Жайны шударга байдлын индексийг ашиглан үнэлж, хэрэглэгчдийн дунд нөөцийн тэгш хуваарилалтыг баталгаажуулсан. Нэмж дурдахад, тооцооллын төвөгтэй байдлыг гүйцэтгэх хугацаа болон нийлэмж (convergence)-ийн шинж чанараар үнэлсэн бөгөөд энэ нь санал болгож буй алгоритмыг бодит хугацааны орчинд ашиглах боломжийн талаар ойлголт өгдөг.



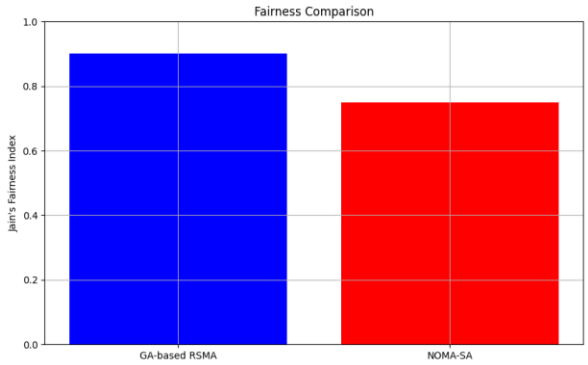
Зураг. 1. Чадлын үр ашгийн харьцуулалт (Нормчилсан)

Симуляцийн үр дүн нь чадлын үр ашиг нь 5G NSA сүлжээнд чухал хүчин зүйл болохыг харуулсан. Зураг 1-д үзүүлснээр, санал болгож буй GA-д суурилсан RSMA тогтолцоо нь NOMA-SA схемтэй харьцуулахад EE-д дунджаар 25%, OFDMA-д суурилсан гүдгэр оновчлолын аргуудтай харьцуулахад 20% сайжирсан. Энэхүү сайжруулалт нь GA-ийн чадлын хуваарилалт, хурдыг хуваах харьцааг оновчтой оновчтой болгож, гүйцэтгэл болон чадлын хэрэглээний хооронд илүү сайн тохирдогтой холбоотой юм.



Зураг. 2. Спектрийн үр ашиг vs SNR

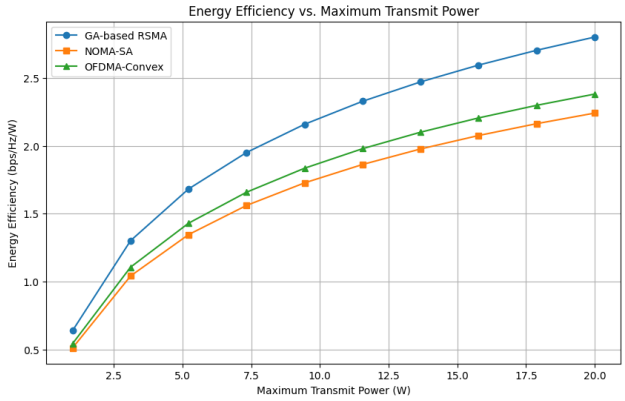
Зураг 2-т үзүүлсэн спектрийн үр ашиг нь дохио-шуугианы харьцаа (SNR) янз бүрийн нөхцөлд NOMA-д суурилсан хандлагыг гүйцэтгэсэн GA-д суурилсан RSMA хуваарилалтын схемийн давуу талыг цаашид дэмждэг. Үр дүн нь NOMA-SA-тай харьцуулахад спектрийн үр ашиг 18%-иар өссөнийг харуулж байгаа нь оновчтой интерференцийн удирдлага болон хурд хуваах дасан зохицох чадварын давуу талыг онцлон харуулж байна.



Зураг. 3. Шударга байдлын харьцуулалт

Олон хэрэглэгчийн сүлжээнд нөөцийн хуваарилалтыг шударга байлгах, ялангуяа бүх хэрэглэгчдэд үйлчилгээний чанар (QoS) шаардлагуудыг хангахад тулгарч буй гол сорилт хэвээр байна. Зураг 3-т үзүүлсэн шударга байдлын үнэлгээний үр дүнгээс харахад GA-д суурилсан RSMA схем нь Жайны шударга байдлын индексийг 0.9-д ойртуулж, ачаалал ихтэй нөхцөлд шударга байдал алдагдсныг харуулсан NOMA-SA-аас давж гарсныг харуулж байна. Эдгээр үр дүнгээс санал болгож буй арга нь хэрэглэгчдийн дунд нөөц хуваарилалтыг үр дүнтэй тэнцвэржүүлж, чадлын мужийн NOMA техникүүдэд ихэвчлэн үүсдэг хэт тэгш бус байдлаас сэргийлж байгааг харуулж байна.

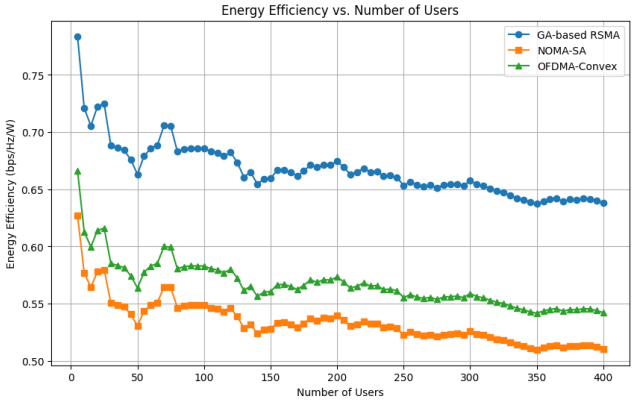
Симуляцийн үр дүнгээс гарсан гол дүгнэлтүүд нь GA-д суурилсан RSMA систем нь NOMA-SA-тай харьцуулахад энергийн үр ашгийг 25%-иар нэмэгдүүлж, чадлын оновчлолын өндөр чадавхийг харуулж байгааг илтгэж байна. Спектрийн үр ашгийг NOMA-д суурилсан шийдлүүдээс 18%-иар дээшлүүлсэн нь нөлөөллийн удирдлага болон динамик нөөц хуваарилалтад RSMA-ийн үр нөлөөг улам баталгаажуулж байна. Шударга байдлын шинжилгээ нь RSMA нь сүлжээний янз бүрийн нөхцөлд Жайны шударга байдлын индексийг 0.9-д ойртуулж, нөөцийн тэгш хуваарилалтыг хангах чадварыг онцолж байна. Эцэст нь, тооцооллын техник, эдийн засгийн шинжилгээний үр дүнгээс үзэхэд, GA-д суурилсан арга нь SA-д суурилсан аргуудтай харьцуулахад гүйцэтгэлийн хугацааг 30%-иар багасгаж, бодит хугацааны сүлжээг оновчтой болгох практик хувилбар болж байгааг баталж байна.



Зураг. 4. Чадлын үр ашиг ба хамгийн их дамжуулах чадал

Зураг 4-т үзүүлсэнчлэн, GA-д суурилсан RSMA оновчлолын арга нь NOMA-SA болон OFDMA-д

суурилсан гүдгэр оновчлолын аргуудтай харьцуулахад дамжуулах хамгийн дээд чадлын янз бүрийн түвшинд чадлын үр ашгийг (bps/Hz/W) тогтмол хангадаг. Тодруулбал, GA-д суурилсан RSMA арга нь NOMA-SA-тай харьцуулахад дунджаар 25%, OFDMA-д суурилсан гүдгэр оновчлолоос 15% орчим сайжирдаг. Эдгээр үр дүн нь системийн гүйцэтгэлийг нэмэгдүүлэхийн тулд байгаа чадлын нөөцийг үр ашигтай ашиглах генетикийн алгоритмын үр нөлөөг тодорхой харуулж байна.



Зураг. 5. Энергийн үр ашиг ба Хэрэглэгчийн тоо

Хэрэглэгчдийн тоо нэмэгдэхийн хэрээр чадлын үр ашгийн (bps/Hz/W) үзүүлэлтээр гүйцэтгэлийн харьцуулалтыг Зураг 5-д үзүүлэв. GA-д суурилсан RSMA алгоритм нь харьцуулсан аргуудаас тогтвортой давуу байдлаа харуулж, ачаалал ихтэй нөхцөлд ч чадлын үр ашгийг мэдэгдэхүйц өндөр байлгадаг. Тодруулбал, GA-д суурилсан RSMA нь NOMA-SA-тай харьцуулахад ойролцоогоор 20%-иар илүү сайн гүйцэтгэлтэй бөгөөд OFDMA-д суурилсан гүдгэр аргуудаас ойролцоогоор 10%-иар сайжирч янз бүрийн хэрэглэгчийн нягтралтай байдаг. Энэхүү гүйцэтгэлийн тогтвортой байдал нь GA-ийн үр ашгийг мэдэгдэхүйц бууруулалгүйгээр өсөн нэмэгдэж буй сүлжээний төвөгтэй байдлыг үр дүнтэй удирдах чадварыг онцолж байна.

Дүгнэлт

Энэхүү судалгаагаар 5G NSA сүлжээн дэх Rate-Splitting Multiple Access (RSMA)-д зориулан генетикийн алгоритм (GA)-д суурилсан шинэ нөөц хуваарилалтын тогтолцоог боловсруулсан. Судалгааны үр дүнгээс харахад санал болгож буй GA-д суурилсан RSMA аргачлал нь уламжлалт болон Simulated Annealing (SA)-д суурилсан NOMA схемүүдтэй харьцуулахад чадлын үр ашгийг 25%-иар сайжруулж байгааг харуулсан. RSMA нь хурд хуваах механизм болон интерференцийн оновчтой удирдлагын үр дүнд уламжлалт чадлын мужийн NOMA болон OFDMA аргаас 18%-иар илүү өндөр спектрийн үр ашгийг үзүүлсэн байна. Түүнчлэн, GA-д суурилсан оновчлолын аргаар хэрэглэгчдийн дунд нөөцийн шударга хуваарилалтыг баталгаажуулсан бөгөөд Жайны шударга байдлын индекс нь одоо байгаа аргуудаас илүү гарч, ойролцоогоор 0.9-ийн өндөр утгыг хадгалсан.

Туршилтын үр дүнг дүрслэхдээ Зураг 1-д GA-RSMA, NOMA-SA, OFDM Convex 2 гэсэн гурван механизмийг харьцуулсан. Манай санал болгож буй GA-RSMA нь бусдаас 20–25% илүү bps/Hz/W үзүүлэлттэй байгаа нь дамжуулах хүчний оновчлол, stream хуваарилалт сайжирсан байгаатай холбоотой. Энэхүү үр дүнг Зураг 1-ийн томъёо болон тайлбар хэсэгт нэгтгэн үзүүлсэн. Зураг 2, 3-д зөвхөн хоёр механизм харьцуулагдсан нь [тухайн нөхцөл байдлын шалтгааныг бичих], Зураг 4-д үзүүлсэн үр дүн нь дамжуулах хүчний хувьд үнэмшилтэй боловч хүлээн авах талын параметр тусгагдаагүй байсан тул сайжруулалт хийсэн. Зураг 5-д хэрэглэгчийн тоо нэмэгдэхэд үзүүлэлт хэрхэн өөрчлөгдөж байгааг харуулсан боловч өмнөх хэсэгт хэрэглэгчийн тоог дурдаагүй байсан учир параметрт нэмж тусгасан. RSMA-д хэрэглэгчийн тоо нь resource allocation-ийн үр дүнд шууд нөлөөлдөг чухал хүчин зүйл тул GA-RSMA загвар хэрэглэгчийн тоонд илүү мэдрэмжтэй хариу үйлдэл үзүүлж байгааг тодотгосон.

Эдгээр ирээдүйтэй үр дүнг харуулж байгаа ч судалгааг цаашид илүү боловсронгуй болгохын тулд хэд хэдэн чиглэлээр нэмэлт судалгаа хийх шаардлагатай. Тухайлбал, бодит хугацааны сүлжээний нөхцлөөс динамик байдлаар суралцаж чаддаг дасан зохицох оновчлолын системийг бий болгохын тулд GA-г Reinforcement Learning эсвэл гүнзгий сургалт (Deep Learning, DL)-тай хослуулсан хиймэл оюун ухаанд (Artificial Intelligence, AI) суурилсан хосолсон аргуудыг судлах шаардлагатай. Мөн санал болгож буй GA-д суурилсан RSMA аргачлалыг олон үүрийн сүлжээнд өргөтгөн үүр хоорондын интерференцийн асуудлыг шийдвэрлэх нь сүлжээний хэмжээнд нөөцийн оновчтой хуваарилалтыг хангах боломжтой юм. Үүнээс гадна програм хангамжаар тодорхойлогдсон радио (SDR) эсвэл 5G туршилтын дэд бүтцийг ашиглан бодит орчинд туршиж, баталгаажуулснаар уг аргачлалын практик үр дүнг баталгаажуулж, бодит хэрэглээний боломжийн талаар нарийвчилсан ойлголтыг олж авах болно.

Дүгнэж хэлэхэд, энэ судалгаагаар GA-д суурилсан RSMA-ийг оновчтой хослуулснаар 5G NSA сүлжээний үр ашиг, шударга байдлыг хангаж, уламжлалт олон цэгийн хүртэмжит схемүүдээс илүү гэдгийг харуулсан. 6G болон ирээдүйн сүлжээнүүдийн эрэлт хэрэгцээ өсөн нэмэгдэхийн хэрээр RSMA нь интерференцийн удирдлага, спектрийн үр ашиг, QoS баталгааг хангах бат бөх шийдэл болж, ирээдүйн олон цэгийн хүртэмжит схемийн хөгжүүлэлтэд чухал үүрэг гүйцэтгэх төлөвтэй байна.

Ашигласан материал

[1] [1] Xudong Li, Rugui Yao, Theodoros A. Tsiftsis, Alexandros-Apostolos A. Boulogeorgos, Secure and Green Rate-Splitting Multiple Access Integrated Sensing and Communications, 2502.13699 19 Feb 2025

[2] Hameed. I, Camana. M.R, Sejan. M.A.S Song, H.K, Enhancing Uplink Communication in Wireless Powered Communication Networks Through Rate-Splitting Multiple Access and Joint Resource Optimization. Mathematics 2025, 13,888. https://doi.org/10.3390/math13050888

[3] Sundar Aditya, Yong Jin Daniel Kim, David Vargas, David Redgate, Onur Dizdar, Neil Bhushan, Xinze Lyu, Sibo Zhang, Stephen Wang and Bruno Clerckx, Rate-Splitting Multiple Access for 6G: Prototypes, Experimental Results and Link/System level Simulations, 2502.09283 17 Feb 2025

[4] Harvanek, M.; Bolcek, J.; Kufa, J.; Polak, L.; Simka, M.; Marsalek; R. Survey on 5G Physical Layer Security Threats and Countermeasures, Sensors 2024, 24, https://doi.org/10.3390/s24175523

[5] Mao, Y., Clerckx, B., & Shi, Y. (2022). Energy Efficiency of Rate-Splitting Multiple Access in 5G Multi-Cell Networks. *IEEE Transactions on Wireless Communications*, 21(5), 3361-3375.

[6] Mao, Y., Clerckx, B., & Li, V. O. K. (2021). Rate-Splitting Multiple Access for Downlink Communication Systems: Bridging, Generalizing, and Outperforming SDMA and NOMA. EURASIP Journal on Wireless Communications and Networking, 2021(1), 1-54.

[7] Liu, Y., et al. (2018). Rate-splitting multiple access for 5G and beyond. IEEE Communications Magazine, 56(11), 70-75.

[8] Joudeh, H., & Clerckx, B. (2016). Sum-rate maximization for linearly precoded downlink multiuser MISO systems with partial CSIT: A rate-splitting approach. IEEE Transactions on Communications, 64(11), 4847-4861.

[9] Dai, L., Wang, B., Ding, Z., Wang, Z., Chen, S., & Hanzo, L. (2018). A Survey of Non-Orthogonal Multiple Access for 5G. IEEE Communications Surveys & Tutorials, 20(3), 2294-2323.

[10] Mao, Y., & Clerckx, B. (2020). Beyond dirty paper coding for multi-antenna broadcast channel with partial CSIT: A rate-splitting approach. IEEE Transactions on Communications, 68(11), 6775-6791.

[11] B. Clerckx, H. Joudeh, C. Hao, M. Dai and B. Rassouli, "Rate splitting for MIMO wireless networks: a promising PHY-layer strategy for LTE evolution," in *IEEE Communications Magazine*, vol. 54, no. 5, pp. 98-105, May 2016, doi: 10.1109/MCOM.2016.7470942.

[12] Shi, L., et al. (2021). Optimization of NOMA systems with simulated annealing. IEEE Access, 9, 43312-43320.

[13] Tang, J., So, D. K. C., Zhao, N., Shojaeifard, A., & Wong, K. K. (2019). Energy efficiency optimization with SWIPT in MIMO broadcast channels for Internet of Things. IEEE Internet of Things Journal, 6(3), 4337-4347.

[14] Andrews, J. G., et al. (2014). What will 5G be? IEEE Journal on Selected Areas in Communications, 32(6), 1065-1082.

[15] Sachs, J., et al. (2018). 5G Radio Network Design for Ultra-Reliable Low-Latency Communication. IEEE Network, 32(2), 24-31.

[16] Elsayed, M., Erol-Kantarci, M. (2019). AI-enabled Future Wireless Networks: Challenges, Opportunities, and Open Issues. IEEE Vehicular Technology Magazine, 14(3), 70-77.

[17] Han, Z., et al. (2012). Game Theory in Wireless and Communication Networks: Theory, Models, and Applications. Cambridge University Press. https://doi.org/10.1017/CBO9780511895043

[18] Holland, J. H. (1992). Genetic algorithms. Scientific American, 267(1), 66-73.

[19] Li, R., et al. (2018). Intelligent 5G: When Cellular Networks Meet Artificial Intelligence. IEEE Wireless Communications, 24(5), 175-183.

[20] Zhao, N., et al. (2019). Spectrum Sharing for 5G Networks: A Survey and Perspectives. IEEE Communications Surveys & Tutorials, 21(1), 111-133.

[21] Qualcomm Technologies, Inc. (2020). Dynamic Spectrum Sharing (DSS) in 5G NR. White Paper.

[22] Luong, N. C., et al. (2019). Applications of Deep Reinforcement Learning in Communications and Networking: A Survey. IEEE Communications Surveys & Tutorials, 21(4), 3133-3174.

[23] Tse, D., & Viswanath, P. (2005). Fundamentals of Wireless Communication. Cambridge University Press.

[24] Buzzi, S., et al. (2016). A Survey of Energy-Efficient Techniques for 5G Networks and Challenges Ahead. IEEE Journal on Selected Areas in Communications, 34(4), 697-709.

[25] Sun, Y., et al. (2018). Learning to Optimize Wireless Resource Allocation: Deep Reinforcement Learning. IEEE Wireless Communications, 25(3), 31-37.

[26] Wang, X., Chen, Y., & Tian, L. (2020). Energy-efficient resource allocation in 5G networks using genetic algorithm. IEEE Transactions on Green Communications and Networking, 4(2), 362-371.

[27] Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. Proceedings of ICNN'95 - International Conference on Neural Networks, 4, 1942-1948.

[28] Foukas, X., et al. (2017). Network Slicing in 5G: Survey and Challenges. IEEE Communications Magazine, 55(5), 94-100.

[29] Kreutz, D., et al. (2015). Software-Defined Networking: A Comprehensive Survey. Proceedings of the IEEE, 103(1), 14-76.

[30] Li, Z., Zhao, J., & Wang, X. (2021). Artificial Intelligence Empowered Resource Allocation in Software-Defined Networks: A Comprehensive Survey. IEEE Communications Surveys & Tutorials, *23(4), 2357-2389. https://*doi.org/10.1109/COMST.2021.3076430

[31] Ye, H., Li, G. Y., & Juang, B. H. F. (2019). Deep Reinforcement Learning Based Resource Allocation for V2V Communications. IEEE Transactions on Vehicular Technology, 68(4), 3163-3173.

[32] Sanguinetti, L., Björnson, E., & Hoydis, J. (2019). Toward Massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination. IEEE Transactions on Communications, 67(1), 573-589.

Холбооны зориулалттай цацаргагч фидерээс энерги хуримтлуулах тородойл загвар

Батболд ДЭЛГЭРМӨРӨН[1], Батбаатар ЗАГАРСҮРЭН[1], Гантас ЭНХЗУЛ[1], Дашбаатар АНУДАРЬ[1], Баярсайхан ПҮРЭВЦЭРЭН[1a], Пүрэвдорж УУГАНБАЯР[1]

[1]Холбооны салбар, Мэдээлэл, Холбооны Технологийн Сургууль
Шинжлэх Ухаан Технологийн Их Сургууль
Улаанбаатар, Монгол улс
J.RC24E004@must.edu.mn , B222300004@must.edu.mn, B222300502@must.edu.mn, B222300007@must.edu.mn, purevtseren_b@must.edu.mn[1a], p.uuganbayar@must.edu.mn

***Хураангуй* — Энэ өгүүлэлд гүний уурхайн цацаргагч фидерээс тородойл ороомог ашиглан RF энергийг цуглуулж, тогтмол гүйдэл болгон хувирган супер конденсаторт хуримтлуулан бага чадлын мэдрэгчийг тэжээх шийдлийн хөгжүүлэлтийг танилцуулж байна. Цацаргагч фидер нь уурхайн хонгил дахь харилцаа холбоог хангах зориулалттай бөгөөд аюулгүй байдлын мэдээллийг үр ашигтай дамжуулдаг. Уг орчинд ашиглагддаг мэдрэгч төхөөрөмжүүд ихэвчлэн батарейгаар ажилладаг тул тогтмол солих шаардлага нь хүндрэл, зардал үүсгэдэг. Иймээс уг судалгаа нь батарей ашиглахгүйгээр мэдрэгчийг тогтвортой, найдвартай тэжээх боломжийг бүрдүүлэхэд чиглэгдэж байна.**

***Түлхүүр үг*—** радио давтамжийн, индуктор

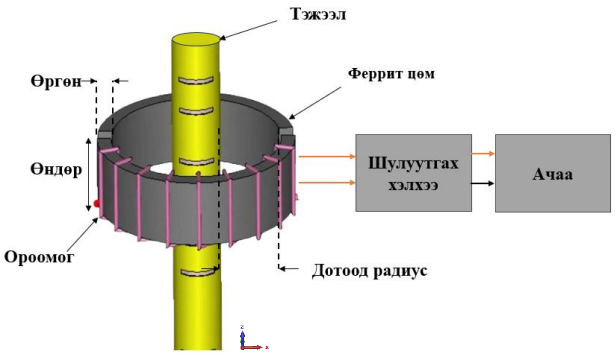
I. Удиртгал

Орчин үеийн утасгүй холбооны системүүд нь гүний уурхай, метро, хонгил зэрэг хүнд нөхцөлд аюулгүй ажиллагааг хангахад чухал үүрэгтэй. Эдгээр орчинд ашиглагддаг температур, чийгшил, галын мэдрэгчүүд нь ихэвчлэн батарейгаар ажилладаг тул тогтмол солих шаардлагатай байдаг. Энэ нь засвар үйлчилгээний зардлыг нэмэгдүүлдэг. Иймээс сул алдагдаж буй RF энергийг цуглуулж дахин ашиглах нь баттерей ашиглахгүй, удаан хугацаанд ажиллах мэдрэгч хөгжүүлэх боломжийг олгодог. Уг судалгаагаар цацаргагч фидерийн радио долгионоос феррит цөм бүхий тородойл ороомгоор энерги хурааж, тогтмол хүчдэл болгон хувирган супер конденсаторт хуримтлуулах технологийг танилцуулсан. Өмнө нь энерги хуримтлуулах технологийг өндөр хүчдэлийн шугамд ашиглаж байсан ч тэдгээр нь 50–60 Гц давтамжтай, мегаваттын түвшний чадалтай байсан нь энэ судалгаанаас эрс ялгаатай. Манай тохиолдолд UHF буюу 0.4–0.6 ГГц-ийн давтамжтай, -17 дБм чадалтай дохионоос энерги хуримтлуулж байгаа нь онцлог юм. Ийм төрлийн бага чадлын энерги хуримтлуулагчийг хөгжүүлснээр микроконтроллер болон мэдрэгч төхөөрөмжүүдийг батарейгүйгээр удаан хугацаанд тасралтгүй ажиллуулах нөхцөлийг бүрдүүлнэ. Мөн гүний уурхайн 1300 метрийн гүнд орших битүү хонгилд нар, салхи, чичиргээ зэрэг бусад төрлийн энерги эх үүсвэрүүд тохиромжгүй тул радио долгионоос энерги цуглуулах нь хамгийн боломжит шийдэл болж байна.

II. Цахилгаан соронзон энерги хураагуурын бүтэц, ажиллагааны зарчим

Цахилгаан соронзон энерги хуримтлуулагч нь бусад төрлийн энерги хуримтлуулагч төхөөрөмжүүдтэй харьцуулахад хэд хэдэн давуу талтай. Тухайлбал, цаг агаарын нөхцлөөс хамаарахгүй, байнгын засвар үйлчилгээ шаардахгүй, хамгийн гол нь дамжуулах шугамд гүйдэл байгаа л бол тасралтгүй ажиллаж чадна. Тиймээс ийм төрлийн систем нь гүний уурхайн хонгил мэтийн алслагдсан болон хүрэхэд хүндрэлтэй, хязгаарлагдмал орон зайн орчинд байрлуулахад хамгийн тохиромжтой шийдэл болдог.

Феррит цөмийн U хэлбэрийн хийц нь зарчмын хувьд трансформатор шиг ажиллаж төрөл бүрийн дамжуулах шугамаас энерги хураах боломжтой. Энэ тохиолдолд дамжуулах шугам нь анхдагч ороомгийн үүргийг гүйцэтгэж, харин феррит цөм дээр байрласан ороомгууд нь хоёрдогч ороомог болж, орчны соронзон орны нөлөөгөөр Фарадейн индукцийн хуулийн дагуу хуйларсан соронзон орон нь хувьсах гүйдэл үүсгэж хэрэглэгч дээрх хүчдэлийн уналтыг бий болгох боломжтой [2]. Ингэснээр дамжуулах шугамын гүйдлээс эрчим хүч хураах боломжтой болдог. Доорх Зураг 1-т феррит цөм ашигласан цахилгаан соронзон энерги хураах системийг харуулсан байна.



Зураг 1. Энерги хуримтлуулах зориулалттай феррит цөмийн бүтэц, хэрэглээний диаграмм

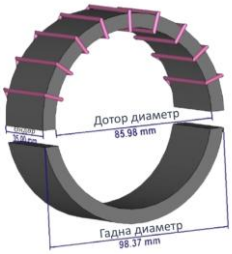
Индукц ороомог болон феррит цөм ашиглан цахилгаан соронзон ороомгоос үүссэн хувьсах хүчдэлийг хүлээн авч түүнийг LC (индукц-багтаамж) резонансын

цахилгаан хэлхээгээр дамжин гүүрэн шулуутгагч ашиглан тогтмол гүйдэл, хүчдэл болгон хувиргадаг.

Энэхүү арга нь уламжлалт индуктив холболт дээр тулгуурладаг боловч, түүн дээр LC резонансны хэлхээг нэмснээр резонансны давтамж дээр илүү үр ашигтайгаар хүчдэл хураан авах боломжыг бүрдүүлдэг. Ийнхүү боловсруулагдсан тогтмол гүйдэл нь хэт бага чадлын төхөөрөмжүүд жишээлбэл: (мэдрүүр, Юмсын интернэт IoT гэх мэт)-ийг тэжээхэд ашиглагдана. Мөн энэхүү гүйдлийг супер конденсаторт хуримтлуулж, цаашид бага чадлын хэрэглэгчийг тогтвортойгоор тэжээлээр хангах боломжтой болно.

III. Индуктив ороомгийн загварчлал, шулуутгах хэлхээ

Тородойл ороомгийн хувьд феррит цөмийн дотор диаметр, гадна диаметр, ферритийн өргөн, өндөр, ороодсын тоо зэрэг хэд хэдэн хэмжээсээс бүрддэг. (Зураг 2-д үзүүлсэн). Эдгээр хэмжээсүүд нь энерги хуримтлуулагч төхөөрөмжийн гаргаж чадах эрчим хүчний хэмжээнд шууд нөлөөлдөг. Өөрөөр хэлбэл, феррит цөмийн хэлбэр хэмжээ өөрчлөгдөх тусам хуримтлагдах энергийн хэмжээ мөн өөрчлөгддөг. Тиймээс эдгээр геометрийн үзүүлэлтүүдийг зөв сонгох нь төхөөрөмжийн үр ашигт чухал ач холбогдолтой юм.



Зураг 2. Феррит цөм

A. Феррит цөмийн тооцоолол

Соронзон орны энерги хуримтлуулагч төхөөрөмжүүд нь Фарадейн хууль болон Амперийн хуулийн зарчим дээр үндэслэн ажилладаг. Соронзон орны хүчлэг (H) нь дамжуулж буй гүйдлийн хэмжээнээс хамаарч, гүйдэл ихсэхэд соронзон орны хүчлэг пропорционалиар өсдөг. Феррит цөмийн хувьд соронзон орны хүчлэгийг дараах тэгшитгэл 1-ээр тооцоолно [6] [7].

$$H = \frac{i}{2\pi r} \tag{1}$$

Доорх илэрхийлэл нь феррит цөмийн соронзон индукц (L)-ийг тодорхойлон соронзон орон үүсгэх, хуримтлуулах чадварыг үнэлдэг бөгөөд энэхүү томьёоноос феррит цөмийн параметрүүдийг тооцоолж болно.

Индукц (L)-ийг дараах томьёогоор тодорхойлно:

$$L = N^2\mu(\frac{Ae}{Le}) \tag{2}$$

Феррит цөмийн соронзон урсгал дамжих дундаж замын урт Le [6]:

$$\text{Le} = \frac{\pi(D_0 - D_i)}{ln\left(\frac{D_0}{D_i}\right)} \tag{3}$$

Феррит цөмийн хөндлөн огтлолын талбай Ae:

$$\text{Ae} = \pi(R_o^2 - R_i^2) = 2\text{rH} \tag{4}$$

Феррит цөм болон цацаргагч фидерийг CST STUDIO SUITE программ ашиглан загварчилсан.

Б. Резонансны тохируулга

Харилцаа холбооны зориулалттай цацаргагч фидер кабелиас UHF давтамжын зурваст хувьсах соронзон орон үүсдэг. Энэхүү өндөр давтамжтай соронзон орон нь ойр орших индукцийн ороомогт индукцийн хүчдэл үүсгэх бөгөөд үүнийг Фарадейгийн индукцийн хууль тодорхойлдог.

ХҮСНЭГТ 1. Феррит цөмийн параметр утгууд

	Феррит цөмийн хэмжээ параметр		43 x 50 x 35
Феррит цөмийн материал	феррит	Дотор радиус (R_i)	43 мм
		Гадна радиус (R_0)	50 мм
		Өндөр (H)	35 мм
	зэс	Утасны радиус (Wr)	1 мм
		Ороодсын тоо (N)	9

Иймд индукцийн хүчдэлээс эрчим хүч хураах үр ашгийг нэмэгдүүлэхийн тулд LC хэлхээг зохион байгуулж, резонансны давтамжийг тухайн UHF дохионы давтамжид тохируулна. Ингэснээр тухайн давтамжид хариу үйлдэл хамгийн ихтэй болгодог [3]. Резонанс давтамжийг f_0 дараах томъёогоор тодорхойлно.

$$f_0 = \frac{1}{2\pi * \sqrt{\frac{L}{C}}} \tag{5}$$

Резонанс нь ороомог болон конденсаторын реактив эсэргүүцлүүд тухайн давтамжид тэнцэлд орсноор, хэлхээ тухайн давтамж дээр хамгийн их гүйдэл дамжуулах чадвартай болдог. Энэ үед энергийн шингээлт ихэсч, хариу үйлдэл хамгийн хүчтэй илэрдэг. Резонансын үед хэлхээний гүйцэтгэлийг тодорхойлох бас нэг чухал параметр бол чанарын хүчин зүйл (Q) юм. Энэ нь дараах томъёогоор тодорхойлогдоно:

$$Q = \frac{1}{R} * \sqrt{\frac{L}{C}} \tag{6}$$

Q өндөр байх тусам энерги алдагдал багатай, резонанс зурвасын өргөн нарийсан давтамжийн хариу үйлдэл нэмэгднэ.

В. Шулуутгах хэсэг

Резонанс LC хэлхээнээс үүссэн хувьсах хүчдэлийг гүүрэн шулуутгагч (rectifier circuit) ашиглан тогтмол гүйдэл (DC) болгон хувиргадаг. Ингэхдээ нийтлэг хэрэглэгддэг аргуудын нэг нь гүүрэн шулуутгагч (bridge rectifier) бөгөөд энэ нь хувьсах хүчдэлийг хоёр туйлтай хэлбэрээр хүлээн авч, хоёр чиглэлд дамжиж буй гүйдлийг зөвхөн нэг чиглэлд урсгах замаар бүтээн үеийн тогтмол гүйдэл үүсгэдэг. Бид симуляцын орчинд хүчдэлийг тооцооллоор шулуутган харах боломжтой бөгөөд практикт Шотки диод гэх мэт элементэд үндэслэсэн бага чадлын өндөр үр ашигтай бэлэн интеграл шулуутгах хэлхээг шууд ашиглах боломжтой юм [4].

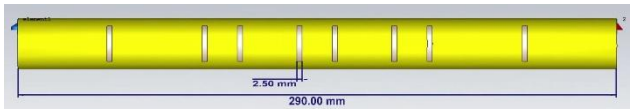
Г. Нам давтамжын шүүр ба ачаа

Шулуутгагчийн гаралтыг тогтворжуулахын тулд LC нам давтамжын шүүрийг (LC low-pass filter) ашиглана. Энэ шүүр нь зэрэгцээ холбогдсон ороомог болон конденсаторааас бүтдэг ба өндөр давтамжийн хэлбэлзэл, хүчдэлийн шуугианыг дарж тогтмол гүйдлийн чанарыг сайжруулдаг.

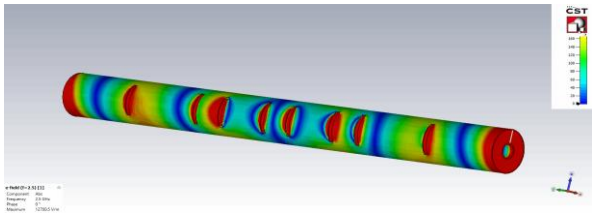
Тогтмол хүчдэл нь тогтворжсон тул үүнийг бага чадлийн тэжээлийн удирдлага ашиглан супер конденсаторт цэнэглэж бага чадлийн мэдрүүр болон микроконтроллер ажиллуулах боломжтой [5].

IV. ЦАЦАРГАГЧ ФИДЕР БОЛОН ФЕРРИТ ЦӨМИЙН ҮР ДҮН

Судалгааны ажлын анхны зорилгын дагуу UHF давтамжын зурваст ажиллах цацаргагч фидер болон феррит цөмийг CST програм дээр тус тусад загварчилсан.



Зураг 3. Цацаргагч фидерийн CST загварчлал

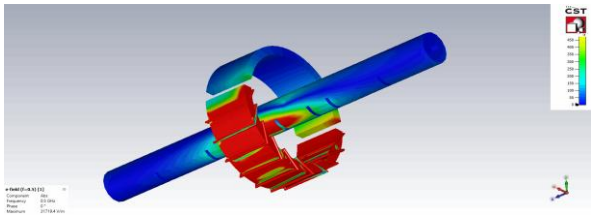


Зураг 4. Цацаргагч фидерт үүсч буй цахилгаан орон (e-field)

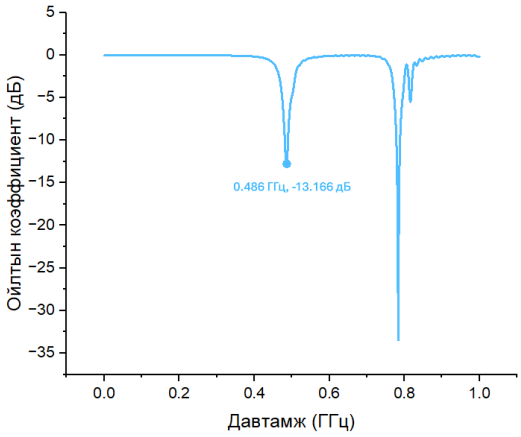
Цацаргагч фидер нь (TE11) горимтой бөгөөд Зураг 5-д үзүүлэв.

Зураг 5. Цацаргагч фидерийн TE11 горим

Цацаргагч фидер болон феррит цөм хэрхэн харилцан нөлөөлөлд орж байгааг Зураг 6-с харж болно. Энэ тохиолдолд шугамын тэжээлийн дохиог цацаргагч фидерт өгөөд феррит цөмд индукцлэлийн үзэгдлээр гүйдэл үүсч байгаа бөгөөд феррит цөмд үүссэн энергийг цааш шулуутган хэрэглэх боломжтой юм.

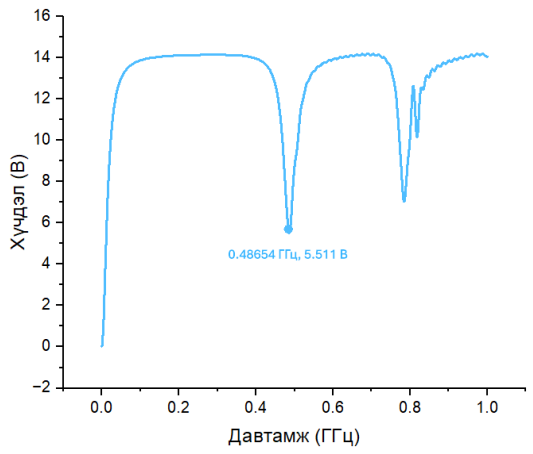


Зураг 6. Феррит цөмтэй цацаргагч фидерийн цахилгаан орны тархалт (e-field)



Зураг 7. Ойлтын коэффициентийн график

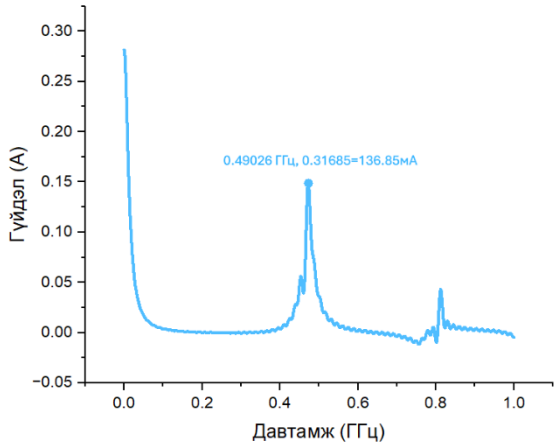
Зураг 7-д феррит цөмийн ойлтын коеффициентийг харууллаа. Феррит цөм нь 0.486 ГГц давтамж дээр зохицсон резонанстай байгаа нь гүний уурхайд хэрэглэдэг 0.4-0.6 ГГц датамжийн зурвастай давхцаж байна. Симуляц хийхдээ цацаргагч фидерийг гүний уурхайн холбоонд байгаа хэмээн төсөөлж 0.486 ГГц давтамжтай дохиогоор тэжээсэн байгаа учир цацаргагч фидер болон феррит цөмийн ажиллаж байгаа давтамжууд ижил болж тэдгээрийн хооронд явагдаж байгаа индукцлэлийн үр ашиг нэмэгдэж байгаа юм.



Зураг 8. Хүчдэлийн симуляцийн график

Зураг 8-д санал болгож буй загвараас үүсч буй хүчдэлийн хэмжээг давтамжаас хамааруулан харуулж буй симуляцын үр дүнг харууллаа. Мургуйгаас харахад резонансны давтамж дээр 5.511 В хүчдэл гарган авах боломжтой байгааг харуулж байна.

Зураг 9-д санал болгож буй загвараас үүсч буй гүйдлийн симуляцын үр дүнг харуулж байгаа бөгөөд резонансны давтамж дээр ойролцоогоор 137 мА гүйдэл гүйж байгаа нь харагдаж байна. Гүйдлийн энэ утга нь резонансны бус бусад давтамжууд дээрх гүйдлээс илүү байгааг Зураг 9-с харж болно. Иймд хүчдэл, гүйдлийн хэмжээг харилцан баланстай хадгалахын тулд энерги хуримтлуулагч нь заавал тодорхой резонансны давтамж дээр төлөвлөгдөн ажиллах шаардлагатай нь харагдаж байна.



Зураг 9. Гүйдлийн симуляцийн график

Зураг 8 болон Зураг 9 -д харуулсан хүчдэл, гүйдлийн симуляцын үр дүнгээс нэгтгээд харахад санал болгож буй тородойл бүтэцтэй энерги хуримтлуулагч нь 0.754 Вт чадал гаргаж байгаа юм.

V. ДҮГНЭЛТ

Практик хэрэглээнд, бага оврын IoT мэдрүүрүүд нь ойролцоогоор 0.5-5 мВт чадал хэрэглэдэг бөгөөд эдгээрийн хэрэглээний гүйдэл нь 30 мА орчим, хүчдэл нь 3.3 В эсвэл 5 В байдаг. Энэ нөхцөлд санал болгосон 0.754 Вт чадалтай энерги хуримтлуулагчийн загвар нь гүний уурхайн хонгилд байрлах цацаргагч фидерийн холбооны системээс үүсэж буй дохионоос энерги хуримтлуулах зарчмаар үйл ажиллагаа явуулах боломжтойг харуулж байна. Иймд уг энерги хуримтлуулагчийг тухайн орчинд амжилттай хэрэглэх бүрэн боломжтой. Цаашид, резонансын чансааг нэмэгдүүлж, энерги хуримтлуулагчийн чадлыг илүү оновчтой, өндөр түвшинд сайжруулах боломжтой юм.

АШИГЛАСАН МАТЕРИАЛ

[1]. M. K. Patel et al., "IoT-Based Wireless Communication Systems for Mine Safety Monitoring," IEEE Internet of Things Journal, vol. 8, no. 15, pp. 12145-12159, 2021. DOI: 10.1109/JIOT.2021.3087456.

[2]. Griffiths, D. J. (2017). *Introduction to Electrodynamics* (4th Edition). Pearson Education.

[3]. Wang, Z., & Zhang, Y. (2016). "Design of a Magnetic Resonant Wireless Power Transfer System for Low Power Devices." *IEEE Transactions on Power Electronics*, 31(7), 4805–4814.

[4]. Pressman, A. I., Billings, K., & Morey, T. (2009). *Switching power supply design* (3rd ed.). McGraw-Hill.

[5]. Rashid, M. H. (2013). *Power electronics: Circuits, devices and applications* (4th ed.). Pearson.

[6]. Numerical analysis of the harvester having toroidal structure and examination of the application result e-ISSN: 2618-575X

[7]. Optimization design of toroidal core for magnetic energy harvesting near power line by considering saturation effect Bumjin Park,1 Dongwook Kim,1 Jaehyoun

Био-анагаахын радио долгион дээр элэгний хорт хавдрыг илрүүлэх мэдрүүр антены судалгаа

Хүрэлсүх ЕСӨН-ЭРДЭНЭ[1], Батбаатар ЗАГАРСҮРЭН[1], Баярсайхан ПҮРЭВЦЭРЭН[1], Пүрэвдорж УУГАНБАЯР[1a]
[1]Холбооны салбар, Мэдээлэл, Холбооны Технологийн Сургууль
Шинжлэх Ухаан Технологийн Их Сургууль
Улаанбаатар, Монгол улс
yesunerdenek@gmail.com[1], zagarsuren0109@gmail.com[1], purevtseren_b@must.edu.mn[1], p.uuganbayar@must.edu.mn[1a]

Хураангуй— Элэгний хорт хавдрыг бичил долгионы техникийг ашиглаж илрүүлэх бага оврын урвуу F антены загварыг танилцуулаа. Антен нь 401-406 МГц-ийн эмнэлгийн төхөөрөмж (Medradio)-ийн зурваст загварчлагдсан. Санал болгож буй антены ойлтын коэффициентын симуляцын үр дүнгээс харахад резонансны давтамжийн шилжилтээр хорт хавдартай биеийг эрүүл биеэс ялгаж мэдрэхээс гадна хавдрын өөр өөр үе шатуудад антены ойлтын коэффициентын утга -27.4 дБ-с -29дБ хооронд өөрчлөгдөж байгаа учир тус антены тусламжтайгаар хорт хавдрын үе шатуудыг ялгаж чадаж байна.

Түлхүүр үг—Элэгний эсийн хорт хавдар, хавтгай урвуу F антен

I. Оршил

Элэгний эсийн хорт хавдар (HCC) нь дэлхий даяар тохиолддог тав дахь хамгийн түгээмэл неоплазм (хэвийн бус эдийн өсөлт) бөгөөд зарим бүс нутагт хорт хавдартай холбоотой нас баралтын үндсэн шалтгаан болдог [1]. Элэгний хорт хавдарт холангиокарцинома (cholangiocarcinoma), ангиосаркома (angiosarcoma), элэгний эсийн хорт хавдар (hepatocellular carcinoma) гэсэн төрлүүд байдаг ба эдгээрээс хамгийн түгээмэл нь элэгний эсийн хорт хавдар юм [2]. Эдгээр төрлийн хорт хавдрууд нь бүх насанд тохиолддог бөгөөд хүний биед хөнөөлтэй тул эрт илрүүлэх шаардлагатай байдаг. Элэгний эсийн хорт хавдрыг цусны дээжийн шинжилгээ, компьютер томограф (CT), соронзон резонансын дүрслэл (MRI), хэт авианы дүрслэл (US) зэрэг аргуудаар оношилж болно [3]. Мөн элэгний биопсигоор тухайн хүний элэгнээс дээж авч микроскопоор ажиглах, туршилт зэргийг гүйцэтгэж хавдрыг оношилно. Эдгээр бүх хавдрыг илрүүлэх аргууд нь хүний биеийн бүхий л хэсэгт гүйцэтгэгдэх боломжтой ба нарийн шинжилгээ хийгддэг. Гэвч эдгээр нь өртөг өндөртэй, олон удаа шинжилгээнд орох боломжгүй байдаг сул талуудтай.

Сүүлийн жилүүдэд антенаар хорт хавдар илрүүлэх аргууд түгээмэл судлагдаж байна [5], [6]. Бичил долгионы дүрслэл, давтамжийн шилжилт зэрэг аргууд нь хямд, нарийвчлал сайн олон удаа шинжилгээнд хамрагдах боломжийг үүсгэдэг нь гол давуу тал болж байна. Мэдрэгч антен нь давтамжийн шилжилт, цацрагийн өөрчлөлт, тусгай шингээлтийн харьцаа (SAR) зэрэг шинж чанарын өөрчлөлтөөр температур, даралт, глюкоз, чийг зэрэг мэдрэхэд хүнд зүйлсийг мэдэрдэг [7]. Энэхүү өгүүлэлд элэгний хорт хавдрыг илрүүлэхэд зориулсан хавтгай урвуу F антеныг 404 МГц-ийн давтамж дээр загварчилсан. Антеныг хүний биеийн модель Густаво дээр 2 мм зайд байрлуулж үзүүлэлтүүдийг харууллаа. Ингэхдээ хадрын 4 үе шат дээр тус тусад нь симуляцийг хийж гүйцэтгэлээ.

II. Элэгний биеийн загвар ба элэгний хорт хавдар

A. Элэгний хорт хавдар

Дэлхийн эрүүл мэндийн байгууллага (WHO) мэдээлснээр хорт хавдар нь хэвийн бус эсийн өсөлтөөс үүсдэг.

Барселона клиникийн элэгний хорт хавдрын системээс (BCLC) тогтоосон стандартаар элэгний эсийн хавдрын үе шатыг тодорхойлдог. Элэгний хорт хавдрын үе шатуудад дараах зүйлс орно.

- I үе: Элгэнд 2 см-ээс бага хэмжээтэй хавдрын эсүүд үүссэн байна. Энэ үед цусны шинжилгээнд хавдаргүй гэж гардаг.
- II үе: Элэгний эдэд 5 см ба түүнээс бага хэмжээтэй ганц хавдартай эсвэл 3 см-ээс бага хэмжээтэй нэгээс олон хавдрын голомттой. Хавдар цусаар дамжин судсанд тархсан байж магадгүй. Иймээс цусны шинжилгээнд илэрнэ.
- III үе: Энэ үе шатанд нэгээс олон хавдар тархсан ба /эсвэл 5 см-ээс дээш хэмжээтэй хавдартай байж болно/. Хавдар лимфийн зангилаа, том судас эсвэл өөр эрхтэн рүү тархсан /метастаз/ байж магадгүй.
- IV үе: Хорт хавдар нь таны уушги, яс, тунгалгийн булчирхай зэрэг биеийн бусад хэсэгт тархсан.

Зураг.1-т элэгний хорт хавдрын үе шатуудыг үзүүлэв.

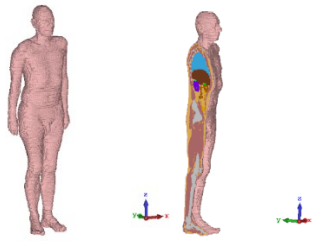
Зураг. 1 Элэгний хорт хавдрын үе шат

Энэхүү судалгаанд хүний биед тохиолддог хамгийн түгээмэл хавдар болох элэгний эсийн хавдрыг авч үзсэн. [12]-т буй эрэгтэй 65 настай элэгний эсийн хорт хавдартай өвчтөний мэдээллийг ашигласан.

B. Элэгний загвар

Антен элэгний эс хоорондох үзүүлэлтийг үндэслэн нотлохын тулд CST Studio Suite программ хангамжид зураг.2-т буй хүний биеийн модель болох Густавог ашигласан. Густаво нь CST Studio suite программ хангамжид нээлттэй байдаг. Эдгээр эдийг үүсгэхдээ хүний биед бодитоор тохиолддог диэлектрик шинж

чанаруудыг ашигласан [8]. Симуляцийг гүйцэтгэхдээ моделийн хэвлийн дээд хэсгийг сонгон авч симуляцийг гүйцэтгэсэн.



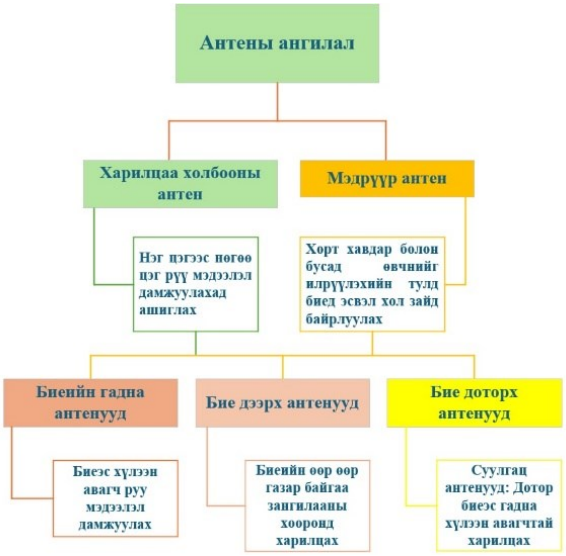
Зураг. 2 Хүний биеийн 3D модель Густаво

Хүснэгт I. ЭЛЭГНИЙ ЗАГВАРТ АШИГЛАСАН ДИЭЛЕКТРИК ШИНЖ ЧАНАР

Эд / Шинж чанар	*Арьс*	*Өөх*	*Булчин*	*Элэг*	*Хавдрын эс*
Нэвтрүүлэх чадвар	46.7	11.6	57.1	51.2	59.89
Цахилгаан дамжуулалт [С/м]	0.69	0.08	0.79	0.65	0.65
Нягт [кг/м^3]	1109	911	1090.4	1078.7	~
Дулааны дамжуулалт [Вт/°С/м]	0.37	0.21	0.49	0.51	~
Дулааны багтаамж [кЖ/К/кг]	3390.5	2348.3	3421.2	3540.2	~

III. АНТЕНЫ ЗАГВАРЧЛАЛ

Хавдрыг илрүүлэх антенууд нь голчлон хавдрын эсийг мэдрэх болон дүрслэл хийх зорилготой бүтээгддэг. Нөгөө талаас холбооны антенуудыг эмнэлгийн төхөөрөмж, мэдрэгч, суулгац эсвэл зүүж болох төхөөрөмжүүдийн хоорондох утасгүй холболт хийх зорилгоор био-анагаахын хэрэглээнд ашигладаг [4]. Зураг.3-т эдгээрийн төрлийг харуулав.



Зураг. 3 Био-Анагаахын антены төрлүүд

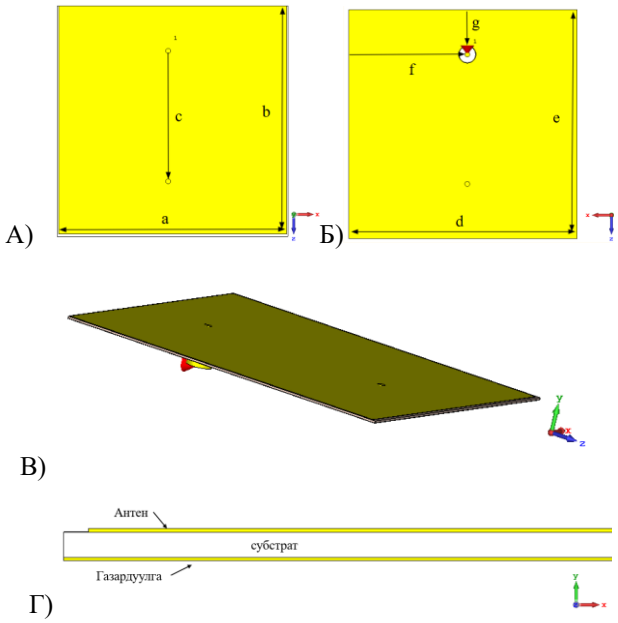
Тиймээс антен жижгэрүүлэхийн тулд өндөр диэлектрик тогтмолтой материал ашиглах, хагас долгионы антен, слот арга гэх мэт антен жижгэрүүлэх аргуудыг ашигласан [9].

Хавдрыг илрүүлэхэд илүү бага давтамжийн зурвасыг ашиглахыг илүүд үздэг. Учир нь нам давтамжууд нь эд эсэд илүү амархан нэвтэрч чаддаг гэдгийг нэвтрэлтийн гүний тэгшитгэлд дурдсан байдаг [10]. Антен бүтээхэд долгионы уртаас хамаарч антеныг загварчилдаг. Ихэвчлэн антенууд нь долгионы уртын хагастай тэнцүү хэмжээтэй байдаг [11]. Иймд антен нь хэмжээний хувьд том болох тул антены хэмжээг багасгах маш том асуудал үүсдэг.

Хүснэгт II. ХАВТГАЙ УРВУУ F АНТЕНЫ ТЭМДЭГЛЭЭ БА ХЭЖЭЭС

Тэмдэглэгээ	Хэмжээ(мм)
Антены урт, a	45.5
Антены өргөн, b	45.5
Тэжээл слотын хоорондох зай, c	46
Газардуулгын урт, d	46
Газардуулгын өргөн, e	26
Тэжээлийн урт, f	22
Тэжээлийн өргөн, g	9

Иймээс хавтгай урвуу F антеныг 401-406 МГц-ийн эмнэлгийн төхөөрөмж (Medradio)-ийн зурваст загварчилсан. Дээд болон доод хязгаар давтамжийг 0 МГц-ээс 800 МГц гэж сонгосон. Антен газардуулга хоорондох материалыг Rogers 6010 гэсэн материалаар сонгож симуляци хийсэн. Материалын зузаан 0.254 мм, зэсийн зузаан 0.035 мм.

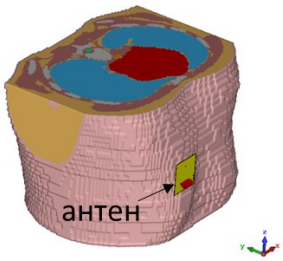


Зураг. 4 Хавтгай урвуу F антен (а- нүүр, б- ар тал, в- 3D, г- хэвтээ)

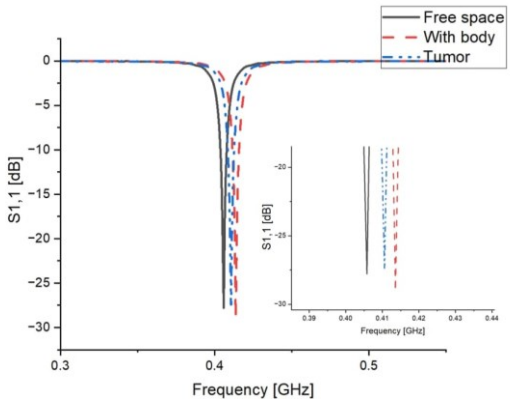
Хавтгай урвуу F антены 3D загварыг зураг.4-т үзүүлэв. Мөн хүснэгт. II нь хавтгай урвуу F антены тэмдэглэгээ ба хэмжээсийг харуулав.

IV. Симуляцийн үр дүн

Антен жижгэрүүлэх аргуудыг ашиглаж элэгний хорт хавдар илрүүлэхэд зориулсан антеныг 16 дахин жижгэрүүлсэн. Хүний биеийн модель антены симуляцийг CST Studio Suite программын тусламжтай хийж гүйцэтгэв.



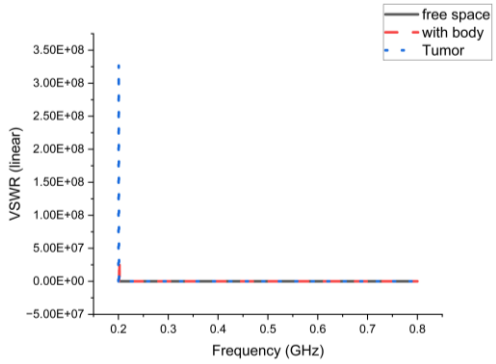
Зураг. 5 Хүний биеийн цээжин хэсэгт антен байрлах байрлал



Зураг. 6 Антены ойлтын коэффициент

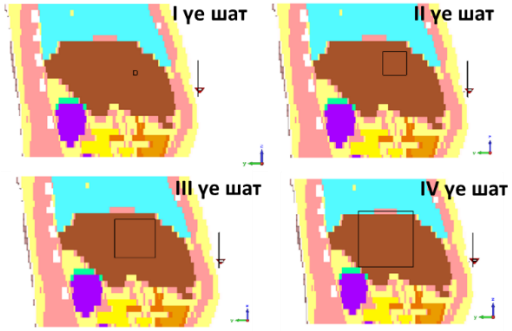
Зураг.4-т санал болгож буй антеныг хүний биеийн элэг байрладаг хэсэгт хэрхэн байрлуулахыг харуулж байна. Антен нь хүний биед 2 мм зайтай байрлаж хавдрын шинжилгээг хийх ба энгийн дүрс оношилгоотой харьцуулахад хавдрын эсийг хавдрын эхэн үе шатанд нь мэдрэх чадвартайгаа онцлог юм. Өөрөөр хэлбэл маш бага хэмжээтэй хавдрын үүслийг цаг алдалгүй илрүүлэх боломжтой.

Зураг.6-т антеныг чөлөөт орон зайд байрлуулсан, эрүүл хүний биед байрлуулсан, элэгний хорт хавдартай биед байрлуулсан зэрэг 3 ялгаатай нөхцөлд түүний ойлтын коэффициент хэрхэн өөрчлөгдөж, давтамж шилжиж байгааг үзүүлэв. Эндээс хавдар байгаа үед резонансны давтамж шилжиж антен мэдрүүр байдлаар амжилттай ажиллаж байгааг харж болно.



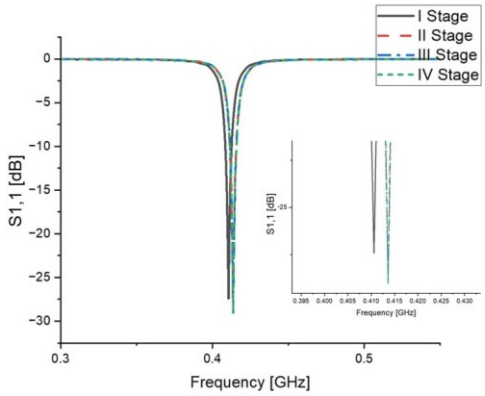
Зураг. 7 Ялгаатай нөхцөлд антены VSWR

Зураг 7-оос харахад антеныг чөлөөт орон зайд байрлуулсан, эрүүл хүний биед байрлуулсан, элэгний хавдартай биед байрлуулсан зэрэг 3 ялгаатай нөхцөлд VSWR үзүүлэлт тогтвортой байгаа бөгөөд эндээс бид гадны физик шалтгааны улмаас антены мэдрэх чадвар өөрчлөгдөхгүй тогтвортой байж чадна гэдгийг баталж болно.



Зураг. 8 Хавдрын 4 шат

Зураг.8-т элэгний хавдрын I, II, III, IV шат дахь хэмжээний дагуу элгэнд хорт хавдрыг байршуулав. Эцсийн шатанд бусад эрхтэн лүү үсэрхийлсэн байдлаар байршуулсан. Хорт хавдрын тархалтыг дөрвөлжин хүрээгээр тэмдэглэн харуулсан бөгөөд симуляци хийх үед дөрвөлжин хүрээн доторх хэсгийг хавдар гэж үзээд ялгаатай диэлектрик нэвтрэх чадвартай материалаар дүүргэсэн. Ийнхүү хавдрын ялгаатай үе шатуудыг загварчилснаар санал болгож буй антен хавдарт хэрхэн мэдрэг байгааг шинжлэх боломжтой болж байгаа юм.

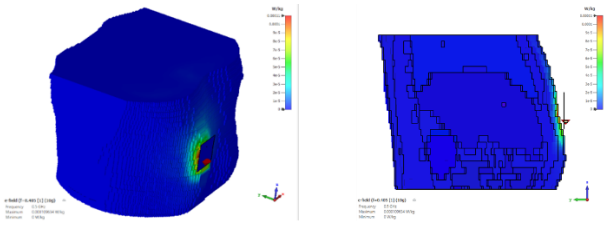


A)

Б)

Зураг. 9 Хавдрын 4 үе шат дээрх ойлтын коэффициентын өөрчлөлт

Үүний дараа хорт хавдрын үе шатуудад антены ойлтын коэффициент хэрхэн өөрчлөгдөж буйг зураг.9а,б-д үзүүлэв. Зураг 9б-ээс харахад хавдрын өөр өөр үе шатуудад антены ойлтын коэффициентын утга -27.4 дБ-29 дБ хооронд өөрчлөгдөж байгаа учир тус антены тусламжтайгаар хорт хавдрын үе шатуудыг ялгах боломжтой гэж үзэж байна.



А) Б)

Зураг. 10 Тусгай шингээлтийн харьцаа

Хүний биеийн тусгай шингээлтийн харьцааны (SAR) стандарт утга нь 1.6 Вт/кг байдаг бол санал болгож буй антены хувьд Зураг 10-т үзүүлсэн симуляцийн үр дүнгээс харахад 0.00011 Вт/кг байгаа нь энэ антен нь хүний биед аюултай нөлөөгүй гэдгийг баталж байна.

Хүснэгт III. НӨХЦӨЛ БҮР ДЭЭРХ ПАРАМЕТР ҮЗҮҮЛЭЛТҮҮД

Параметер	**Чөлөөт орон зай**	**Хүний бие**	**Хавдар**
S_{11}(дБ)	-28.13	-10.09	-10.08
Өсгөлт (дБ)	1.27	1.25	1.23
VSWR	1.08	1.93	1.91

Хүснэгт IV. ТӨСӨӨТЭЙ ЗАГВАРУУДТАЙ ХАРЬЦУУЛСАН БАЙДАЛ

Ном зүй	S_{11}(дБ)	Хэмжээ (мм)	Ажлын давтамж (ГГц)
[5]	-27.48	20.5×14×1.6	3.35
[6]	-27	137 × 109 ×109	0.58-1.6
Энэ ажил	-27.78	46 × 46× 0.324	0.403

V. Дүгнэлт

Энэхүү өгүүлэлд урвуу F хэлбэрийн антеныг 401-406МГц зурваст элэгний хорт хавдрыг эрт илрүүлэх оношилгоонд зориулж загварчиллаа. Санал болгож буй антены ойлтын коэффициентын симуляцын үр дүнгээс харахад резонансны давтамжийн шилжилтээр хавдартай биеийг эрүүл биеэс ялгаж мэдрэхээс гадна хавдрын өөр өөр үе шатуудад антены ойлтын коэффициентийн утга -27.4 дБ-с -29дБ хооронд өөрчлөгдөж байгаа учир тус антены тусламжтайгаар хорт хавдрын үе шатуудыг ялгах боломжтой гэж үзэж байна.

Санал болгож буй антены гол давуу тал нь UHF био-радио давтамжийн зурваст буюу харьцангуй нам давтамжид антены геометр хэмжээг хангалттай бууруулж 46мм×46мм×0.324мм болгож чадсан явдал юм.

Ном зүй

[1] J. Bruix, L. Boix, M. Sala, and J. M. Llovet, “Focus on hepatocellular carcinoma,” *Focus*, vol. 5, no. 3, pp. 201–219, Mar. 2004.

[2] J. Baj, L. Brylinski, F. Wolinski, M. Granat, K. Kostelecka, and P. Duda, “Biomarkers and genetic markers of hepatocellular carcinoma and cholangiocarcinoma—What do we already know,” *Cancers*, vol. 14, no. 6, p. 1493, 2022. [Online]. Available: https://doi.org/10.3390/cancers14061493

[3] Yu, X.L., Li, A.H., Jiang, T.A., Chen, M.H., Zhao, B.Z., Zhou, X.D. and Wang, J.R., 2007. Comparison of contrast enhanced ultrasound and contrast enhanced CT or MRI in monitoring percutaneous thermal ablation procedure in patients with hepatocellular carcinoma: a multicenter study in China. Ultrasound in medicine & biology, 33(11), pp.1736-1749.

[4] A. R. Chishti, M. A. Qureshi, M. N. Abbasi, S. S. Iqbal, and A. M. Algarni, “Advances in antenna-based techniques for detection and monitoring of critical chronic diseases: A comprehensive review,” *IEEE Access*, published Sept. 18, 2023.

[5] R. Islam, F. Mahbub, S. A. Al-Nahium, and M. I. Hossain Rabbi, “Design and performance analysis of monopole antenna for early-stage hepatocellular carcinoma (HCC) in human liver,” in *Proc. IEEE Annu. Information Technology Conf.*, 2021.

[6] S. A. Razaeieh, A. Brankovic, A. S. Janani, B. Mohammed, A. Darvazenban, A. Zamani, G. A. Macdonald, and A. M. Abbosh, “Wearable electromagnetic belt for steatotic liver detection using multivariate energy statistics,” *IEEE Access*, published Nov. 3, 2020.

[7] M. E. Gharbi, R. Fernandez-Garcia, S. Ahyoud, and I. Gil, “A review of flexible wearable antenna sensors: Design, fabrication methods, and applications,” *Materials*, published Aug. 27, 2020..

[8] IT’IS Foundation. Tissue Properties Database. Version 4.2 [Online]. Available: https://itis.swiss/virtual-population/tissue-pSroperties/database/dielectric-properties/

[9] N. A. Malik “ Design of implantable antennas for biomedical applications” Ph.D dissertation, Univ. of Bedfordshire pp. 10-12, 2022.

[10] S. I. Alekseev and M. C. Ziskin, ‘‘Human skin permittivity determined by millimeter wave reflection measurements,’’ Bioelectromagnetics, vol. 28, no. 5, pp. 331–339, 2007

[11] Antenna thoery “The Dipole antenna” [Online]. Available: https://www.antenna-theory.com/antennas/dipole.php

[12] A. Peyman, B. Kos, M. Djokic, B. Trotovsek, C. Limbaeck-Stokin, G. Sersa, and D. Miklavcic “ Variation in dielectric properties due to pathological changes in human liver” Bioelectromagnetics, vol. 36, pp. 603-612, 2015

Дрон унагаах төхөөрөмжид зориулсан өндөр үзүүлэлттэй антены хөгжүүлэлт

Ганбаатар ДӨЛМАНДАХ[1] , Баярсайхан ПҮРЭВЦЭРЭН[1] , Пүрэвдорж УУГАНБАЯР[1a]

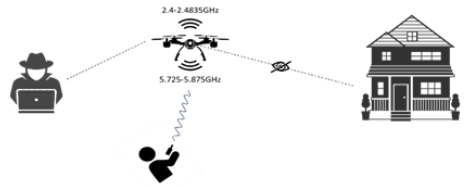
Монгол Улсын Шинжлэх Ухаан Технологийн Их Сургууль,
Мэдээлэл холбоо технологийн сургууль Холбооны салбар
gdolmandah@gmail.com[1], purevtseren_b@must.edu.mn[1], p.uuganbayar@must.edu.mn[1a]

Хураангуй--Дрон эсэргүүцэх төхөөрөмжид зориулсан нарийн чиглэлтэй антенны судалгаануудад ихэвчлэн Яаги-Уда төрлийн антен ашигласан байдаг бол энэхүү симуляцийн судалгаанд өндөр үр ашигтай мушгиа (Helix) антены сайжруулсан загварыг танилцуулж байна. Уг мушгиа антен нь дрон үйлдвэрлэлийн стандарт дамтамжууд болох 2.4 ГГц болон 5.8 ГГц-ийн хос давтамжийн зурваст ажиллаж дроны удирдлагын дохиог сарниулах юм. Ингэхдээ тухайн дроны антен ямар төрлийн туйлширалтай(шугаман, дугуй) вэ гэдгээс хамаарахгүйгээр мушгиа антен нь дугуй туйлширалтай ажиллах юм. Антены өсгөлтийг нэмэгдүүлж, чиглэлийн диаграммын үндсэн чиглэлийн өнцгийг нарийсгахын тулд антенд нэмэлтээр чиглүүлэгч линз нэмж, үндсэн цацаргагч болох мушгиа вибраторыг нарийссан (Tapered) хэлбэртэй болгож өөрчилсөн. Туршилтын үр дүнд уг антен нь 2.4 ГГц болон 5.8 ГГц давтамжийн зурваст ажиллаж 5.8 ГГц-ийн зурваст 15.5 dBi өсгөлттэй, чиглэлийн диаграммын хагас чадлын өргөн нь 24.1⁰ болсон нь ижил зориулалттай антенуудтай харьцуулахад сайжруулсан үр дүн болж чадсан.

Түлхүүр үг: 2.4 ГГц Helix antenna, 5.8 ГГц Tapered antenna, Narrow HPBW

I. Оршил

Сүүлийн үед дрон хэрэглэгчдийн тоо хурдацтай нэмэгдэж, энгийн иргэдээс эхлээд цэргийн зориулалт хүртэлх олон салбар, хэрэглээнд өргөн ашиглагдах болсон. Гэвч зарим нэг хараа хяналтгүй бусдын эрхэнд халдсан дроны үйл ажиллагаанд хязгаарлалт тавих технологи хөгжүүлэх шаардлага нэмэгдэж байна. Тухайлбал хилийн орчмын бүсэд хил зөрчиж дрон нисгэж зураг бичлэг хийх зэрэг нь хил хамгаалах байгууллагынхны зайлшгүй анхаарах асуудал болоод байна. Өмнө нь дронуудыг радио жаммер (саатуулах төхөөрөмж) ашиглан эсэргүүцэх талаар олон төрлийн судалгаанууд хийгдэж байсан [1]. Гэвч ихэнх ийм төрлийн дрон эсэргүүцэх төхөөрөмжүүд зөөврийн бус байдаг. Зөөврийн дрон эсэргүүцэх төхөөрөмж нь дроны аюулд яаралтай хариу арга хэмжээ авах хэрэгтэй болдог. Иймд радио давтамжийн (RF) бууг ашиглан дроны ажиллагааг алдагдуулах арга нь илүү тохиромжтой юм. Дрон эсэргүүцэх төхөөрөмжийн радио долгионыг цацаргах антен нь хэд хэдэн давтамжийн зурвасуудыг хамрах ёстой.



Зураг.1 Ажиллагааны зарчим

Дронууд нь өөрийн байршлыг тодорхойлоход Байршил Тогтоох Систем (GPS-Global Positioning System) L1 зурвас (1562.92–1587.92 МГц)-ыг ашигладаг [2, 3]. Үйлдвэрлэл Шинжлэх ухаан, Эмнэлгийн зориулалтын (ISM-industrial, scientific, and medical) 2.4 ГГц (2.4–2.4835 ГГц) болон 5.8 ГГц (5.725–5.875 ГГц) зурвасууд нь дрон болон алсын удирдлагатай төхөөрөмжүүдийн утасгүй удирлагын дохиог дамжуулахад хэрэглэгддэг [4,5]. Мөн дронуудаас удирдлагын төхөөрөмж рүү зураг, бичлэг дамжуулахад дээрх зурвасууд ашиглагддаг. Эдгээр зурвасуудад ажиллах чадамжтай сайжруулсан антеныг загварчилснаар дроны ажиллагааг саатуулах хямд, зөөврийн төхөөрөмжийг хөгжүүлэх боломжтой юм. Зах зээл дээр ийм төрлийн төхөөрөмжүүд худалдаалагддаг хэдий ч үнэ маш өндөр байдаг бөгөөд ийм төрлийн төхөөрөмжийн үзүүлэлтийг сайжруулахын тулд антены өсгөлт, чиглэлийг сайжруулах ажил судлаачдын дунд сорилт хэвээр байна.

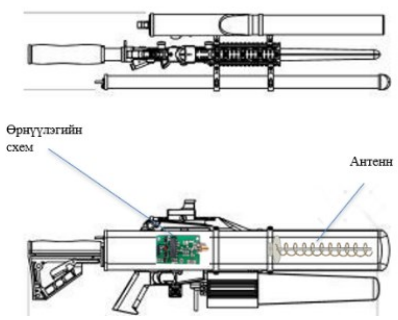
Дрон саатуулах системийн гол бүрэлдэхүүн хэсэг болох антен нь өндөр өсгөлттэй, хамгийн гол нь нарийн чиглэлийн диаграммтай, тогтвортой ажиллах чадамжтай байх хэрэгтэй. Тэгвэл “мушгиа” антен нь тойрог туйлширалтай, бага алдагдалтай, авсаархан хэмжээтэй тул дрон саатуулах хэрэглээнд хамгийн тохиромжтой шийдэл болох юм.

II. Төхөөрөмжийн бүтэц

Радио долгионоор дрон саатуулах төхөөрөмжийг нэгэн төрлийн долгионы буу гэж үзэж болох бөгөөд одоогийн байдлаар судлагдаж буй ихэнх төхөөрөмжүүд дараах хэлбэртэй болно. Энэхүү төхөөрөмж нь хэд хэдэн үндсэн хэсгээс бүрдэх бөгөөд уг бүтцийг Зураг.2-т үзүүлэв. Бидний хөгжүүлж байгаа антенныг суурилуулах төхөөрөмж нь Зураг 2-т үзүүлсэнчлэн дараах хэсгүүдээс бүрдэх юм.

1. Антен нь 2.4 ГГц болон 5.8 ГГц хос ажлын давтамжтай мушгиа хэлбэрийн өндөр өсгөлттэй нарийн чиглэлтэй байх ёстой.
2. Дохио өрнүүлэх хэлхээ нь тухайн давтамжийн зурваст үр дүнтэй ажиллах чадвартай, өндөр үр ашигтай хэлхээ байдаг.
3. Гадна их бие нь 3D дизайны тусламжтай гаргасан загварыг сайжруулсан хуванцар материалаар хэвлэсэн бүтэцтэй байна.

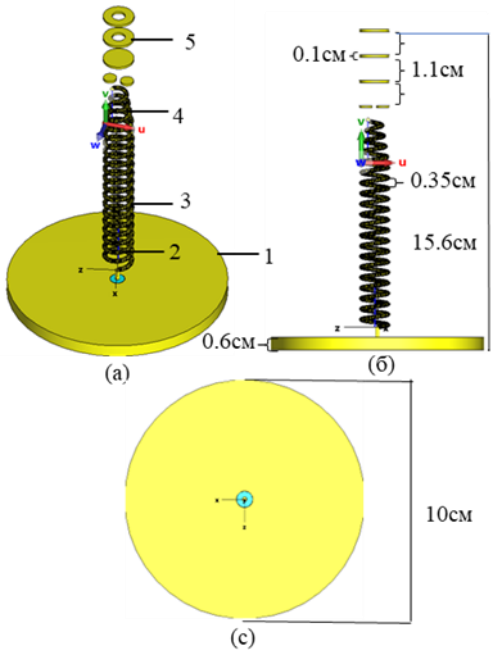
Антены үндсэн цацаргагч элемент болох мушгиа цацаргагчийг хийх зэс утасны диаметр 2мм, антены урт 90.87мм, алхамт хоорондын зай 32 мм, ороодосны тоо 15 байхыг тооцоолсон. Чиглүүлэгч линзний радиус 0.6мм бөгөөд энэ нь антенаас ирж буй долгионыг хэлбэржүүлэн цуглуулах үүрэгтэй.



Зураг.2 Анти-дрон төхөөрөмжийн бүтэц

III. АНТЕНЫ БҮТЭЦ

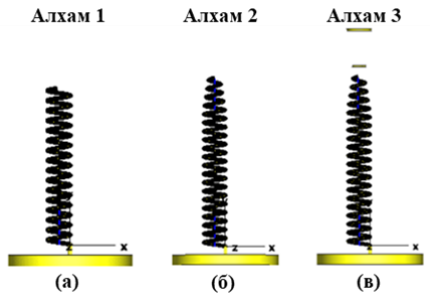
Зураг3-т санал болгож буй антены бүтэц, геометр хэмжээсүүдийг дэлгэрэнгүй харуулллаа.



Зураг.3. 5.8 ГГц Мушгиа сайжруулсан антены загвар

Антен нь ойлгогч (1), тэжээлийн хэсэг (2), антен (3), нарийссан (4), линз (5) зэргээс бүрдэнэ.

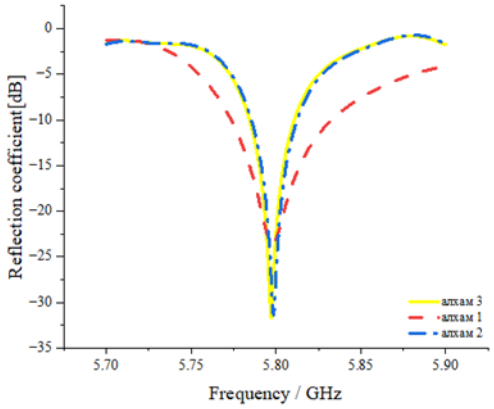
Зураг.3-т үзүүлсэн мушгиа антеныг 3 алхамтай сайжруулсан. Ингэхдээ CST Studio Suite симуляцийн программ дээр гүйцэтгэсэн.



Зураг 4 Антены хөгжүүлэлтийн үе шат

Зураг.4.а.б,в-д антены хөгжүүлэлтийн үе шатуудыг харуулж байгаа бөгөөд алхам 1-д энгийн мушгиа антены бүтэц, алхам 2-т нарийссан мушгиа антены бүтэц, алхам 3-т чиглүүлэх линзтэй нарийссан мушгиа антены бүтэц гэсэн дараалаар хөгжүүлэлт хийсэн.

Зураг 5-д хөгжүүлэлтийн үе шатуудад антены ойлтын коеффициент хэрхэн өөрчлөгдөж байгааг харуулж байгаа бөгөөд дээрх 3 бүтцийн хувьд аль аль зорилтот давтамж дээр зохицсон үр дүнтэй байна. Алхам 3-т хөгжүүлсэн линзэн чиглүүлэлгчтэй нарийссан мушгиа антены хувьд ойлтын коеффициент ажлын давтамж дээр хамгийн бага байгааг харж болно.



Зураг.5. Ойлтын коэффициент

1. МУШГИА АНТЕНЫ СИМУЛЯЦИЙН ХАРЬЦУУЛАЛТ

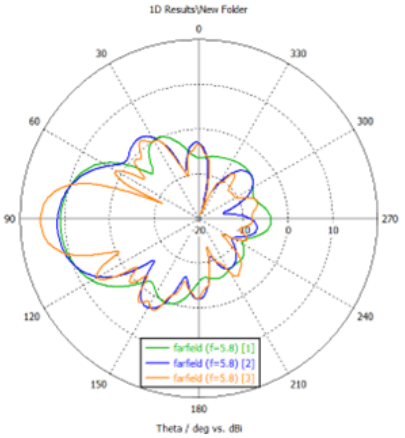
	Front/Back ratio (дБ)	Өсгөлт (дБ)	Цацрагийн өргөн (HPBW)
Алхам 1	14.5	8	50.9º
Алхам 2	23	11.6	41.7º
Алхам 3	23.5	15	23.8º

Хүснэгт.1-т харуулснаар дрон ажлын давтамж дээр мушгиа антены үзүүлэлтийг сайжруулан ажиллуулах боломжтойг харуулж байна. Алхам 3 буюу сүүлийн санал болгож буй чиглүүлэх линзтэй загварын хувьд чиглэлийн диаграммын үндсэн болон арын дэлбээний харьцаа (Front to back ratio) 23.5 дБ болж сайжирсан байгаа нь антен үндсэн цаца ёстой чиглэлдээ энергийг илүү сайн төвлөрүүлж цацаж, хажуугийн болон арын чиглэлд үр ашигтгүй энерг цацаргахгүй, цацаргалтын ашигт үйлийн коеффициент нь дээшилсэн гэдгийг баталж байна. Мөн антены өсгөлт өмнөх загваруудтайгаа харьцуулахад 15 дБ болж өссөн байна. Антенны үндсэн дэлбээны хагас чадлын цацрагийн өргөн (HPBW) нь 23.8º болж нарийссан үзүүлэлтийг харахад чиглэлтэй антены хувьд илүү нарийн чиглэлээр энерги цацаргаж дрон эсэргүүцэх хэрэглээнд нэн тохиромжтой болж хөгжсөнийг харуулж байна.

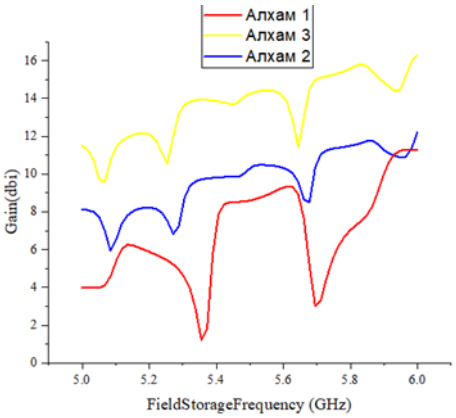
Зураг 6-д антены хөгжүүлэлтийн үе шатууд дахь антены чиглэлийн диаграмын симуляцын үр дүнг харьцуулан харуулсан байна. Улбар шар өнгөөр алхам 3-т хөгжүүлсэн антены чиглэлийн диаграммыг харуулсан бөгөөд алхам-1 болон алхам-2 ийн загваруудтай харьцуулахад илүү нарийн чиглэлд (23.8º) энергийг төвлөрүүлэн цацах чадвартай болсон нь харагдаж байна.

Зураг 7-д антены хөгжүүлэлтийн үе шатуудыг антены өсгөлт-давтамжын муруйгаар харьцаалан харуулсан ба санал болгож буй алхам 3-т хөгжүүлсэн

антен нь өмнөх загваруудаас илүү өндөр өсгөлтийг давтөмжийн зурвасын туршид үзүүлж байна.



Зураг.6. Чиглэлийн диаграмм



Зураг 7 Антены өсгөлт (3 алхам)

Санал болгож буй антены зурвасын өргөний тооцоололыг дараах томьёогоор тооцоолсон ба 40.15% -ийн зурвасын өргөнтэй байгаа нь хэрэглээний шаардалагад нийцэж байна.

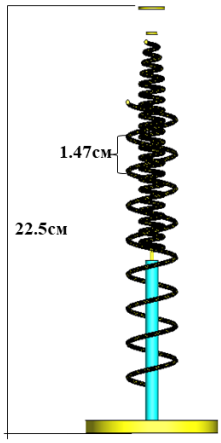
$$FBW(Fractional\ BandWidth) = \frac{\Delta f}{f_c} * 100\% \quad \Delta f = (f_{high} - f_{low}) \ (1)$$

FBW= 40.15%

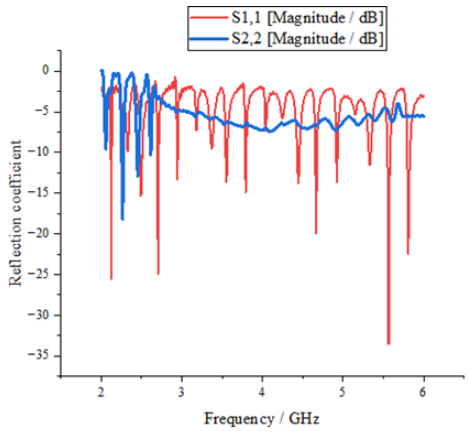
Мөн дрон эсэргүүцэх антеныг хос зурваст (dualband) ажиллуулах нь үндсэн зорилгын нэг бөгөөд үүнийг хийхийн тулд алхам 3-т хөгжүүлсэн 5.8 ГГц-ийн давтамжийн зурваст ажиллах антеныг 2.4 ГГц давтамж дээр ажиллах энгийн мушгиа антентай бүтцийн хувьд нэгтгэж тэжээлийн 2 порттой байхаар зохицуулсан бөгөөд үүний загварыг Зураг.8-д харуулав. Нийлмэл антены хувьд ойлгогч элементээс чиглүүлэч хүртэлх антены нийт урт 20.4 см бөгөөд хоорондоо хэрхэн байрлаж ажиллахыг зургаас харж болно.

Зураг 9-д хос зурваст ажиллах нийлмэл бүтэцтэй 2 порттой антены ойлтын коеффициентуудын симуляцын үр дүнг порт тус бүрээр харуулсан ба 2.4 ГГц болон 5.8 ГГц давтамж дээр ажиллах антенны хэсгүүд тус бүрдээ тухайн ажлын давтамж дээрээ зохицон ажиллах боломжтой байна.

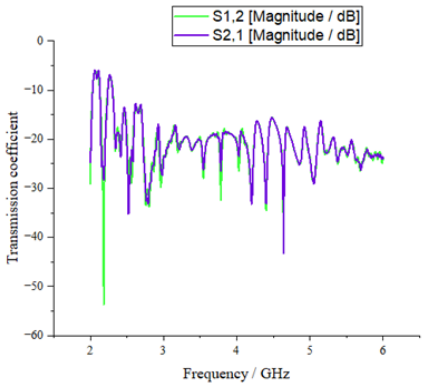
Зураг 10-т нийлмэл бүтэцтэй хос зурвастай мушгиа антены портууд хоорондын дамжууллын коеффициентыг харуулсан ба симуляцын үр дүнгээс харахад 2.4 ГГц болон 5.8 ГГц давтамжын зурвас дээрх бүтцүүд хоорондын дамжууллын коефициент -15дБ буюу түүнээс бага байгаа нь эдгээр антенууд хоорондоо харицан нөлөөлөлгүй нэг бүтэц болон ажиллаж чадаж байгаагийн баталгаа юм.



Зураг.8 Хосолсон зурвасын мушгиа антены загвар



Зураг.9 Хосолсон зурвасын Мушгиа антены загвар



Зураг.10 Хосолсон зурвасын мушгиа антены дамжууллын коэффициент

IV. Дүгнэлт

Ихэнх дрон эсэргүүцэх төхөөрөмжид Яаги-Уда антен ашигладаг бол энэ судалгаагаар үүний оронд мушгиа төрлийн антеныг нарийссан болон линзэн бүтэцтэй болгож антены параметрүүд болох өсгөлт, чиглэлийн диаграммын үзүүлэлтүүдийг сайжруулснаар ийм төрлийн антеныг хэрэглэх боломжтойг харууллаа. Цаашид санал болгож буй антены загварыг дрон эсэргүүцэх төхөөрөмжил ашиглан үйлдвэрлэн турших боломжтой.

V. Ном зүй:

[1] S. Park, H. T. Kim, S. Lee, H. Joo, and H. Kim, “Survey on anti-Drone systems: Components, designs, and challenges,” IEEE Access, vol. 9, pp. 42635–42659, 2021, doi: 10.1109/ACCESS.2021.3065926.

[2] E. Horton and P. Ranganathan, “Development of a GPS spoofing apparatus to attack a DJI matrice 100 quadcopter,” J. Glob. Position. Syst., vol. 16, no. 9, Jul. 2018, doi: 10.1186/s41445-018-0018-3.

[3] A. D. B. A. Rahman, K. A. Ghani, N. H. H. Khamis, and A. R. M. Sidek, “Unmanned aerial vehicle (UAV) GPS jamming test by using software defined radio (SDR) platform,” J. Phys.: Conf. Ser., vol. 1793, no. 1, pp. 12–60, 2021, doi: 10.1088/1742-6596/1793/1/012060.

[4] Z. Cui, C. Briso-Rodríguez, K. Guan, Z. Zhong, and F. Quitin, “Multi frequency air-to-ground channel measurements and analysis for UAV communication systems,” IEEE Access, vol. 8, pp. 110565–110574, 2020, doi: 10.1109/ACCESS.2020.2999659. [5] M. Ghamari, P. Rangel, M. Mehrubeoglu, G. S. Tewolde, and R. S. Sherratt, Access, “Unmanned aerial vehicle communications for civil applications: A review,” IEEE vol. 10, 10.1109/ACCESS.2022.3208571.

5G Сүлжээний Хуваалт Хоорондын Хэрэглэгч Баталгаажуулалт

Г.Манлайбаатар, Н.Угтахбаяр, Д.Баттулга*, Б.Өсөхбаяр
Электроник, Холбооны Инженерчлэлийн Тэнхим
Мэдээллийн Технологи,Электроникийн Сургууль, Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
22m1num0283@stud.num.edu.mn, battulgad@num.edu.mn

Товчлол*—5G сүлжээнд олон шинэ технологи багтдаг бөгөөд тэдгээр нь хэрэглээг дэмжих зорилготой юм. Нөгөө талаас эдгээр шинэ технологи нь сүлжээг улам динамик болгодог онцлогтой. Хатуу дэд бүтэц дээр олон хуурмаг сүлжээ үүсгэх технологийг 5G сүлжээнд нэвтрүүлэн 'хуваалт' (slice) гэж нэрлэсэн. Хуваалтын хувьд уламжлалт утастай сүлжээнийхээс ялгаатай тал нь хэрэглэгчийн хөдөлгөөнт байдал юм. Өөрөөр хэлбэл хэрэглэгч хуурмаг сүлжээ хооронд шилжих хэндовр үйлдэл хийх боломжтой. Технологи болон хэрэглээний хувьд өргөн боломж олгож байгаа ч хуваалт хооронд шилжихэд хэрэглэгчийн баталгаажуулах шаардлага гарч ирнэ. Энэхүү өгүүллээр бид хуваалт хооронд шилжихэд хэрэглэгчийн баталгаажуулах механизмыг санал болгож байна. Бидний санал болгож буй механизм нь сүлжээний ачаалал болон баталгаажуулалт хийх хоцрогдол зэргийг багасгах зорилгоор төвлөрсөн бус захын тооцоолол ашигладаг. Механизмын гүйцэтгэлийн үзүүлэлтийг 5G технологид ашиглаж буйтай харьцуулан дүн шинжилгээ хийсэн.

Түлхүүр үг—Сүлжээний архитектур, сүлжээний хуваалт, хэрэглэгч баталгаажуулалт, хэндовр, 5G

I. Эхлэл

Тав дахь үеийн мобайл сүлжээний технологи 5G -д сүлжээний функцийн виртуалчлал (NFV), программаар тодорхойлогддог сүлжээ (SDN) технологиудыг ашиглан хуурмаг сүлжээ үүсгэж боломжтой болсонг slice гэж нэрлэдэг [1]. Сүлжээний хуваалт (slice) -ийг хатуу дэд бүтэц (радио хандалтын сүлжээ багтана)-ийг үйлчилгээ, хэрэглэгч зэрэг буюу хэрэглээ талаас, аюулгүй байдал, хамрах хүрээ зэргээс хамааруулсан технологи талаас харгалзан олон сүлжээний хуваалт үүсгэх боломжтой юм. Жишээ нь мобайл сүлжээний оператор компани нэгэн уурхайд зориулан хамрах хүрээг үүсгэх буюу олон үүр дамнасан хуурмаг сүлжээг үүсгэн, хүнд механизмуудын сенсоруудаас оффист байрлах серверт өгөгдөл дамжуулах сүлжээний хуваалт үүсгэж болно. Энэ тохиолдолд уурхайд байрлах сенсорууд нь оператор компаний энгийн хэрэглэгчээс тусдаа дэд сүлжээтэй болно.

Уламжлалт мобайл сүлжээнд цөм сүлжээ болон радио хандалтын сүлжээ нь төхөөрөмж хангамжуудтай шууд уялдаатай байх ба зөвхөн операторын үйлчилгээнүүдийг гүйцэтгэхээр зохион байгуулагддаг [2-3]. Энэ нь үйлчилгээнүүд, төхөөрөмжүүдийг удирдахад хялбар боловч өртөг өндөр байдаг. Мөн нэг операторын төхөөрөмжүүд, үйлчилгээнүүд учир хэрэглэгчийн баталгаажуулалт эхний удаа хийгдэнэ [3,5]. Хэндовр нь мобайл сүлжээний гол функцүүдийн нэг бөгөөд уламжлалт байдлаар хөндлөн буюу нэг радио хандалтын сүлжээний үүр хооронд, босоо буюу өөр өөр технологи хооронд гэж ашиглагддаг. Харин 5G сүлжээнд хуваалт үүсгэснээр хэрэглэгч сүлжээний үйлчилгээг авахын тулд байрлаж буй хуваалтруу шилжих буюу хэндовр хийх болно. Мөн хэрэглэгчийг баталгаажуулах үйлдлийг сүлжээний хуваалт хооронд хэндовр хийхэд дахин хийх шаардлагатай [5]. Жишээ нь: өмнөх дурьдсан байдлаар сүлжээнд хуваалт үүсгэсэн ба хэрвээ энгийн хэрэглэгч (сүлжээний администратор) уурхайн сенсоруудын тохиргоог өөрчлөх болвол тухайн сүлжээний хуваалтруу нэвтэрч орно. Энэ нь мобайл операторын энгийн хэрэглэгчдийн сүлжээнээс уурхайн сенсоруудын хуваалтруу хэндоврийг баталгаажуулалтын хамт хийнэ гэсэн үг юм.

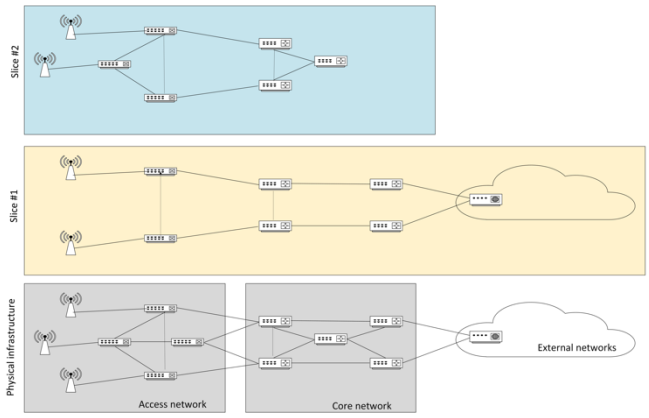
Мобайл оператор компани өртөг өндөр хатуу дэд бүтэц дээрээ хуваалт үүсгэн бусдад ашиглуулснаар нөөцийн ашиглалт сайжрах, хөрөнгө оруулалтаа нөхөх зэрэг эерэг талтай юм [5,6]. Уламжлалт байдлаар өөр сүлжээний үйлчилгээг хэрэглэгч авахын тулд операторын сүлжээний цөм, интернэт хоёроор дамжин өнгөрч холбогддог. Харин хуваалт хийх, хуваалтыг шууд гадаад сүлжээтэй холбох боломжтой болгосноор хатуу дэд бүтцийг хуваан ашиглах боломж улам нэмэгдэнэ.

Мобайл сүлжээний хэрэглэгчийн баталгаажуулах уламжлалт механизмууд нь зөвхөн нэг удаад хийгддэг, төвлөрсөн бүтцээр ажилладаг, хуваалт тус бүрийн онцлогт тааруулж өөрчлөх боломжгүйгээрээ сул талтай [5-7]. Мөн хэт олон хэрэглэгч нэгэн зэрэг баталгаажуулалт хийвэл удаашрах магадлалтай байдаг. Мөн хуваалтын нэг зорилго болох хэрэглэгч оператрын бус үйлчилгээг гадаад сүлжээнээс авах тохиолдолд нэмэлт баталгаажуулалт хийх хэрэгтэй болно.

Уламжлалт механизмуудыг сүлжээний хуваалтуудын хооронд шууд хэрэглэвэл хэрэглэгчийн төхөөрөмж хэндовр хийх бүртээ сүлжээний цөмд байрлах сервертэй баталгаажуулалт хийнэ. Бид энэхүү асуудлыг шийдэх зорилгоор хуваалт хооронд хэндовр хийх үед хамт хийгдэх хөндлөн, тархсан, сүлжээний захад хийгдэх хэрэглэгчийн баталгаажуулалтын механизмыг санал болгоно. Мөн бидний санал болгож буй механизм нь хэрэглэгч мобайл операторын хуваалт хийсэн баз станцаас шууд гадаад сүлжээрүү холбогдох тохиолдолд илүү тохиромжтой юм.

Энэхүү өгүүллийн бүтэц нь хоёрдугаар бүлэгт 5G сүлжээ, сүлжээний хуваалт болон хэрэглэгчийн баталгаажуулалтын талаар онол, гуравдугаар бүлэгт сүлжээний бүтцийн талаар тайлбар, дөрөвдүгээр бүлэгт өөрсдийн механизмын дэлгэрэнгүй тайлбар, тавдугаар бүлэгт бусад хэрэглэгчийн баталгаажуулалтын механизмуудтай харьцуулалт хийсэн тухай, төгсгөлийн тавдугаар бүлэгт дүгнэлтийг багтаасан.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE



Зураг 1. 5G сүлжээний хуваалт

II. 5G СҮЛЖЭЭ, ХУВААЛТ, БАТАЛГААЖУУЛАЛТ

A. *5G сүлжээ ба хуваалт*

5G нь үүрэн телефоны тав дахь үеийн технологи бөгөөд хэрэглэгчдэд илүү хурдан, хоцрогдол маш бага, найдвартай мобайл дата сүлжээний үйлчилгээ үзүүлэх зорилготой, IP суурьтай сүлжээ юм [1-4].

5G сүлжээний цөм нь Зураг 1А-д үзүүлсэн олон үйлчилгээнээс бүрдэнэ [7].

- *Unified Data Repository (UDR):* Хэрэглэгч болон сүлжээний талаарх өгөгдөл хадгална. Сүлжээний функцүүдэд нэвтрэх боломжийг олгодог.
- *Unified Data Management (UDM):* нь хэрэглэгчийн тухайн датаг удирдах ба баталгаажуулалтын итгэмжлэх байдлаар оролцоно.
- *Access and Mobility Management (AMF):* Хэрэглэгчийн бүртгэл, хөдөлгөөнт байдлын удирдлага болон баталгаажуулалтыг удирдан явуулна.
- *Session Management Function (SMF):* сейшн, IP хаяглалт, чиглүүлэлт зэргийг удирддаг.
- *User Plane Function (UPF):* Хэрэглэгчийн өгөгдлийн урсгалыг чиглүүлэх, дамжуулах гарцын үүргийг гүйцэтгэнэ.
- *Network Slice Selection Function (NSSF):* Сүлжээний хуваалтуудаас тодорхой хэрэглээнд зориулан сонгоно.
- *Authentication Server Function (AUSF):* Сүлжээнд нэвтрэх аюулгүй байдлыг хангахын тулд хэрэглэгчдийг баталгаажуулдаг.
- *Network Repository Function (NRF):* Сүлжээний функцүүдэд шаардлагатай мэдээллийг хадгалж, үйлчилгээг эрж олох, хоорондын холбоог дэмждэг.

B. *Хэрэглэгчийн баталгаажуулалт*

Хэрэглэгчийн баталгаажуулалт нь зөвхөн зөвшөөрөгдсөн хэрэглэгчид болон төхөөрөмжүүдийг 5G сүлжээнд нэвтрүүлэх чухал үйл явц юм. Нөгөө талаас сүлжээний нууцлал, хамгаалалтыг хангах зорилготой. 5G сүлжээнд дараах алхмуудыг хийж гүйцэтгэнэ [8]. Үүнд:

1) Бүртгэл: Хэрэглэгч (User) нь AMF-д бүртгүүлэх хүсэлт илгээнэ.

2) Хүсэлт: AMF нь хэрэглэгчийн хүсэлтийг AUSF-д дамжуулна. Цаашлаад UDR-ээс хэрэглэгчийн итгэмжлэлийн талаарх мэдээллийг авна.

3) Хариу: AUSF нь хэрэглэгчрүү асуулт илгээх ба хэрэглэгч нь SIM дээрх нууцлалын түлхүүр ашиглан тооцоолол хийдэг.

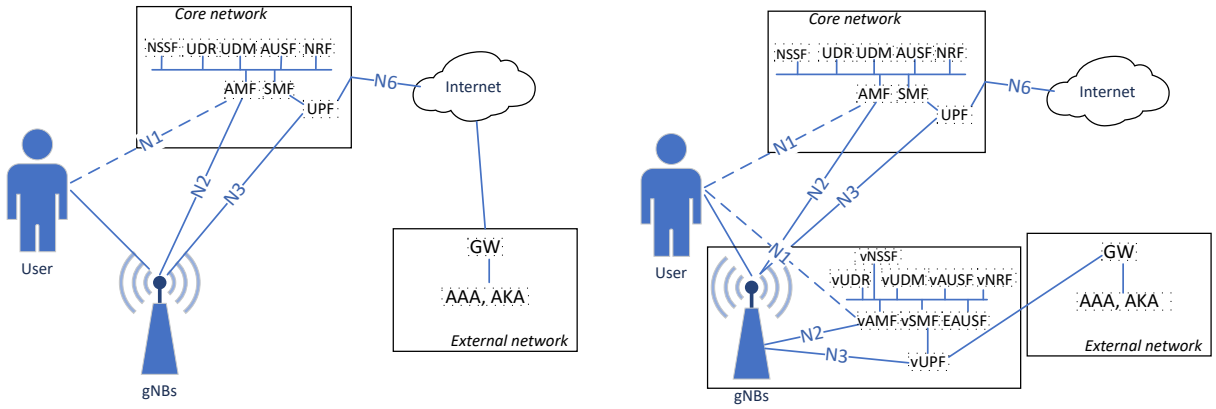
4) Баталгаажуулах: Хэрэглэгчээс ирсэн тооцооллыг агуулсан хариуг AUSF баталгаажуулна. Хэрэв зөв бол хэрэглэгчийг баталгаажуулсан төлөвт шилжүүлнэ.

5) Хандах: Хэрэглэгчид сүлжээнд хандах эрх олгож, тохиргоо хийх нэгж функцүүдрүү шилжүүлнэ.

5G сүлжээнд 3-р үйлдлийг гүйцэтгэхдээ 5G-AKA гэсэн шифрлэлт болон нууцлал хамгаалалтын механизм ашигладаг. Энэ механизм нь хурдан, аюулгүй, хэрэглэгчийн баталгаажуулалтыг хийдэг стандарт арга юм.

III. СҮЛЖЭЭНИЙ БҮТЭЦ

Энэ бүлэгт бид 5G сүлжээнд хуваалт хийн ашиглаж байгаа хувилбарыг танилцуулах болно. Өөрөөр хэлбэл дараах архитектур бүхий сүлжээг байгуулж түүндээ тохирсон хэрэглэгч баталгаажуулах механизм санал болгож байна. 5G сүлжээний баз станц дээр хуваалт хийж,



А) 5G Сүлжээний архитектур

Б) 5G сүлжээний хуваалт хийсний дараах архитектур

Зураг 2. 5G сүлжээний үйлчилгээ ба архитектур, өөрчлөлт

хэрэглэгчийг гадаад сүлжээтэй шууд холбогдож байгааг Зураг 1Б-д үзүүлэв. Мобайл операторын сүлжээний цөмд 5G -ийн үндсэн бүх функц байрлаж энгийн хэрэглэгчдэд үйлчилнэ [7]. Харин хуваалтын хэрэглэгч сүлжээнд анхлан холбогдох үед үндсэн функцүүд холболтыг хийнэ. Үүний дараа баз станц дээр хуулбарлан хуурмагаар байруулсан 5G үйлчилгээнүүд рүү хэрэглэгчийн мэдээллийг илгээж өгөгдөл дамжуулах зам болон чиглүүлэлтийг гүйцэтгэнэ.

Энэхүү 5G сүлжээний хуваалт нь хэрэглэгч баз станц дээрх хуваалтаар дамжин шууд гадаад сүлжээ буюу гурав дахь байгууллагын үйлчилгээтэй холбогдох юм. Дараах алхмуудын дагуу хэрэглэгч сүлжээнд анхлан холбогдоно. Үүнд:

1) Хэрэглэгч эхний удаад сүлжээнд нэвтрэхдээ баз станцаар дамжин үндсэн функцүүдтэй холбогдож хэрэглэгчийн бүртгэл, баталгаажуулалт, холболтын бүх тохиргоог хийж гүйцэтгэнэ.

2) Хэрэглэгч амжилттай нэвтэрсний дараа үндсэн NRF функц шилжих боломжтой хуваалтын тохиргоог хийх ба баз станц дээрх хуулбар NRF-т хуваалттай холбоотой бүх үйлдлийг шилжүүлж тохируулна.

3) Ингэснээр хэрэглэгч, хуваалтууд хооронд хэндовр хийх, баталгаажуулалт болон чиглүүлэлт зэрэг үйлдлийг баз станц дээрх хуулбар функцүүдтэй хийнэ.

4) Хэрэглэгч үндсэн сүлжээнээс хуваалтруу шилжинэ. Үүнд: хуваалт сонгох, хэрэглэгч баталгаажуулалт, хандовр гэсэн үйлдлүүдийг хийж гүйцэтгэнэ.

IV. Баталгаажуулалт ба Хуваалт Хоорондын Хэндовр

Бидний гол зорилго бол хөнгөн-хурдан ажиллах хэрэглэгч баталгаажуулалттайгаар 5G сүлжээний хуваалт хооронд шилжих юм.

A. Баталгаажуулалт хийх алхмуд

Хэрэглэгчийн баталгаажуулалтад дараах дөрвөн нэгж оролцоно: 1) идэвхтэй хэрэглэгч, 2) сүлжээний хуваалт үүсгэсэн баз станц, 3) оператор, 4) гадаад сүлжээ.

Хэрэглэгч үндсэн сүлжээнээс хуваалтруу шилжих буюу хэндовр хийхэд дараах алхмуудыг хийж гүйцэтгэнэ. Үүнд:

1) Хэрэглэгч үндсэн сүлжээнд нэвтрэхэд SIM-д агуулагдах түлхүүр ашиглан 5G -ийн үндсэн функүүдээр баталгаажуулалт хийнэ.

2) Үндсэн функцүүдээс хуулбарууд руу хэрэглэгчийн мэдээллийг өгч хадгалуулна. Үүнд: хэрэглэгчийн баталгаажуулалт, өгөгдөл замын нууцлал, хэндоврын удирдлага багтана.

3) Хэрэглэгч боломжит хуваалтуудаас сонгоно. Үүнд тусгайлсан тохиргоо хийхгүй. Учир нь хэрэглэгч урьдчилан тохируулсан хуваалтруу шилжинэ.

4) Сонгосон хуваалтруу шилжих хүсэлтийг баз станцад илгээнэ.

5) Баз станц хэрэглэгчийн хүсэлтийг шалгаад баталгаажуулалт хийнэ.

6) Гадаад сүлжээний гарц руу хэрэглэгч хэндовр хийж байгаа тухай мэдээллийг баталгаажуулалтын хамт илгээнэ.

7) Гарцын төхөөрөмж хэрэглэгчийг өөрийн дотоод серверийн тусламжтай баталгаажуулж хариуг баз станцруу илгээнэ.

8) Хэрэв баталгаажуулалт амжилттай болвол хэрэглэгч рүү хэндоврыг эхлүүлэх комманд өгнө.

9) Хэрэглэгч хэндоврыг гүйцэтгэж өгөгдлийн замыг шинэчилснээр хуваалт руу шилжиж дуусна.

B. Баз станц дахь нэмэлт функц

Бид баз станц дээр NRF, AMF, AUSF функцүүдээс гадна Гадаад Сүлжээний Баталгаажуулалт (EAUSF) -ийг байгуулсан. Энэ функц нь дараах алхмуудыг хийж гүйцэтгэнэ. Үүнд:

1) Мобайл операторын үндсэн функцүүдээс хуваалт руу шилжих хэрэглэгчийн мэдээллийг хүлээн авна. Үүнд хэрэглэгчтэй мэдээлэл солилцоход ашиглах тэгш хэмт түлхүүр багтана.

2) Гадаад сүлжээний гарц руу хэрэглэгчийн баталгаажуулахтай холбоотой тохиргоо хийнэ. Үүнд нийтийн болон нэг удаагийн түлхүүр багтана.

3) Хэрэглэгчээс нэвтрэх хүсэлт ирэхэд тэгш хэмт түлхүүр ашиглан нийтийн түлхүүр, нэг удаагийн түлхүүрийг хариу илгээнэ.

4) Хэрэглэгчээс нийтийн түлхүүрээр нууцалсан сейшн ID-ийг хүлээн авч гадаад сүлжээний гарц руу илгээнэ.

5) Гадаад сүлжээнд хэрэглэгчийн сейшн ID-г нэг удаагийн түлхүүрийн тусламжтай бататгаад хариу илгээнэ. Ингэхдээ хувийн түлхүүрээр нууцлал тайлна. Энэ хариунд мөн үйлчилгээнүүдтэй холбогдох тохиргоонуудыг багтаана.

6) Баз станц ирсэн тохиргооны дагуу хэрэглэгчийн хэндоврыг удирдан дуусгана.

7) Хэрэглэгч баз станц, баз станцаас гадаад сүлжээрүү гэсэн өгөгдлийн замыг эцэслэж үндсэн сүлжээнээс сална.

Хэрэглэгч одоо байгаа хуваалтад байхгүй үйлчилгээг авах хэрэгцээ үүсвэл шилжих шаардлагатай болно. Ингэхдээ эхлээд өөрийн холбогдсон байгаа баз станцруу хүсэлт илгээнэ. Ингэснээр хуваалт хоорондын хэндовр хийх юм.

C. Хуваалт хоорондын хэндовр

Өмнөх дэд бүлгүүдэд дурьдсаны дагуу хэрэглэгч үндсэн сүлжээнээс хуваалтруу хэндовр хийнэ. Энэ бүлэгт бид хуваалт хооронд хэндовр хийх алхмуудыг тодорхойлно. Хэрэглэгч нэг хуваалтаас боломжит өөр хуваалт руу хэндовр хийх дараах хувилбарууд байж болно. Үүнд:

a) Одоогийн болон сонгосон хуваалт нэг гадаад сүлжээ рүү чиглэсэн

Хэрэглэгч хэндовр хийх хүсэлтийг орох хуваалтын мэдээллийн хамт баз станц руу илгээнэ. Энэ нь IV-A дахь хэндоврын дөрөв дэх алхамтай адил юм.

5) Баз станц хэрэглэгчийн хүсэлтийг хүлээн аваад өмнө хийсэн баталгаажуулалт байгаа эсэхийг шалгана. Баталгаажуулалт байгаа учир дахин батлах шаардлагагүй.

6) Гадаад сүлжээний гарц руу хэрэглэгч хуваалт солисон тухай мэдээллийг илгээнэ.

7) Хэрэглэгчийн холболтын шаардлагатай тохиргоог өөрчилнө. Жишээ нь хуваалтын гарц, ID зэрэг болно.

b) Одоогийн болон сонгосон хуваалт өөр өөр гадаад сүлжээрүү чиглэсэн

Хэрэглэгч хэндовр хийх хүсэлтийг орох хуваалтын мэдээллийн хамт баз станцруу илгээнэ (IV-A -ийн дөрөв дэх алхам). Мөн

5) Баз станц гадаад сүлжээний гарц руу хэрэглэгч нэвтэрч орох хүсэлт илгээнэ.

6) Гадаад сүлжээний гарц баталгаажуулалт хийсний дараа хариу ирүүлнэ.

7) Хэрэв зөвшөөрсөн хариу ирсэн бол хэрэглэгчийн холболтын шаардлагатай тохиргоог өөрчилнө. Жишээ нь хуваалтын гарц, ID, гарц зэрэг болно.

c) Хуваалт нь өөр өөр баз станцад байрласан

Хэрэглэгч хэндовр хийх хүсэлтийг орох хуваалтын мэдээллийн хамт баз станц руу илгээнэ. Баз станц шилжин орох хуваалтын байрлаж буй баз станц руу хүсэлтийг илгээнэ. Энэ үйлдэл нь IV-A -ийн дөрөв дэх алхмын оронд хийгдэх бөгөөд бусад 5-9 -р алхмуудыг адилхан хийж гүйцэтгэнэ.

d) Хуваалт нь өөр өөр мобайл операторт байрласан

Энэ тохиолдолд хэрэглэгч баз станцдаа сонгосон хуваалтын мэдээллийг хүсэлтийн хамт илгээнэ. Үүний дараа IV-A -ийн бүх алхмыг эхнээс нь хийнэ.

V. Хэрэглэгчийн баталгаажуулалтын механизмуудтай харьцуулалт

Энэ бүлэгт бид санал болгож буй механизмаа 3GPP-ийн нэгэн механизм [9]-тай харьцуулсан үр дүнг танилцуулах болно. Сонгож авсан механизм нь хэрэглэгч мобайл операторын сүлжээнд нэвтрэх үед сим карт дахь мэдээллийг ашиглан баталгаажуулалт хийнэ. Харин хуваалтруу шилжихэд нь ижил аргаар баталгаажуулалт хийнэ. Хэрэглэгч гадаад сүлжээний үйлчилгээрүү хандахад операторын сүлжээний цөмөөр дамжин аппликейшн түвшинд баталгаажуулалт хийдэг. Энэ тохиолдолд операторын үүрэг оролцоо хэвээр байна.

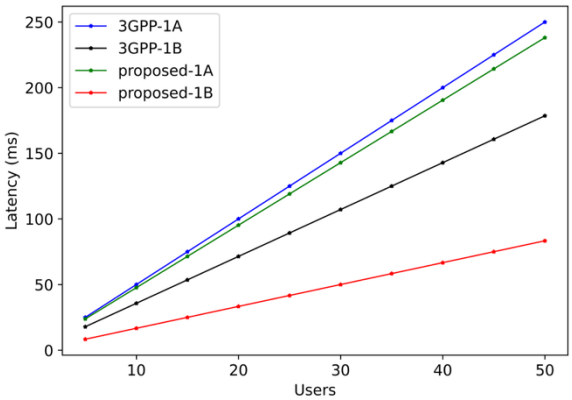
Хоёр механизмыг Зураг 2-д үзүүлсэн хоёр архитектур тус бүрд туршилтын орчинд ажиллуулж туршсан. Мөн хэрэглэгчийг хугацааны давтамжтайгаар хуваалт хооронд хэндовр хийхээр тохируулсан. 3GPP TS 33.501-ийн дагуу 5G -ийн аюулгүй байдал, түүний процедурыг тохируулсан [10]. Туршилтад бид стандарт механизмыг стандарт архитектурт ажиллаж байгааг *3GPP-1A*, стандарт механизмыг өөрчилсөн архитектурт ажиллаж байгааг *3GPP-1B*, санал болгож буй механизмыг стандарт архитектурт ажиллаж байгааг *proposed-1A*, санал болгож буй механизмыг өөрчилсөн архитектурт ажиллаж байгааг *proposed-1B* гэж нэрлэв.

Зураг 3-д бид хэрэглэгч сүлжээнд анхлан нэвтрэх, хуваалт руу шилжих үе дэх хугацааны хоцрогдлыг үзүүлсэн. Ингэхдээ сүлжээнд нэгэн зэрэг нэвтэрч буй хэрэглэгчийн тоог, хуваалт руу нэвтэрч дуусах хүртэлх хугацаа буюу хоцрогдолтой хэрхэн уялдаж байгааг харуулав. *3GPP-1A*, *3GPP-1B, proposed-1A* гэсэн гурван хувилбар нь хэрэглэгчийн тоотой шууд хамааралтайгаар хоцрогдол ихсэж байгааг зургаас харж болно. Харин *proposed-1B* буюу өөрчилсөн архитектуртай үед хоцрогдол мэдэгдхүйц бага өөрчлөгдсөнг харж болно. Жишээ нь хэрэглэгчийн тоо 10 байхад хоцрогдол *3GPP-1A, 3GPP-1B, proposed-1A* нь ойролцоогоор 45-50 ms байгаа бол *proposed-1B* нь 24 ms байна. Харин хэрэглэгчийн тоо 50 болоход *proposed-1B* нь 76 ms буюу бусад хувилбараас дариу 120-200 ms -ээр бага байна. Нэмж дурьдахад хэрэглэгчийн тоо өсөх тусам *proposed-1A* нь нөгөө *3GPP-1A, 3GPP-1B* -аас хоцрогдол багасаж байна.

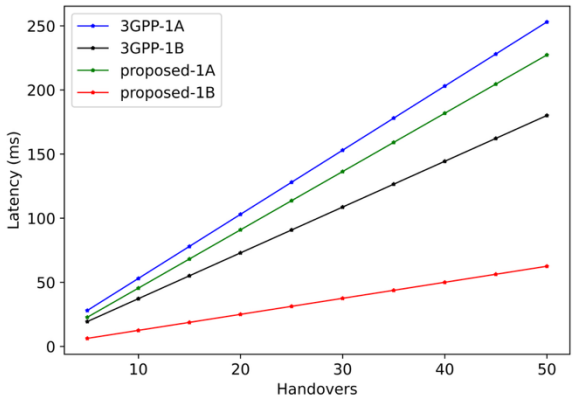
Зураг 4-д бид хэрэглэгч, хуваалт хооронд хэндовр хийх үеийн хугацааны хоцрогдлыг үзүүлэв. Ингэхдээ сүлжээнд нэгэн зэрэг нэвтэрч буй хэрэглэгчийн тоог, хуваалт руу нэвтэрч дуусах хүртэлх хугацаа буюу хоцрогдолтой хэрхэн уялдаж байгааг харуулав. *3GPP-1A, 3GPP-1B, proposed-1A* гэсэн гурван хувилбар нь хэрэглэгчийн тоотой шууд хамааралтайгаар хоцрогдол ихсэж байгааг зургаас харж болно. Харин *proposed-1B* буюу өөрчилсөн архитектуртай үед хоцрогдол мэдэгдхүйц өөрлөгдсөнг харж болно. Жишээ нь хэрэглэгчийн тоо 10 байхад хоцрогдол ойролцоогоор *3GPP-1A* нь 50 ms, *3GPP-1B* нь 47 ms, *proposed-1A* нь 49 ms байгаа бол *proposed-1B* нь 15 ms байна. Харин хэрэглэгчийн тоо 50 болоход *proposed-1B* нь 56 ms буюу бусад хувилбараас дариу 120-200 ms -ээр бага байна.

Хүснэгт 1-д бид *proposed-1B* -ийн үр дүнг дэлгэрүүлж үзүүлэв. Ингэхдээ хоцрогдлыг IV-C дэд бүлэгт тодорхойлсон дөрвөн төрлийн хэндовроор ялгасан. Хэрэглэгчийн тоог 50 байх үеийн хоцрогдлыг авч үзсэн болно.

Харьцуулалтын үр дүнгээс харвал санал болгож буй механизм нь өөрчилсөн сүлжээний архитектурт тохиромжтой буюу хоцрогдол багатай ажиллаж байна. Харин стандарт архитектурт ажиллахад хангалттай сайн үр дүн үзүүлсэнгүй. Энэ нь сүлжээний архитектураас шууд хамаарч байна.



Зураг 3. Хэрэглэгч сүлжээнд нэвтрэх үеийн хугацааны хоцрогдол



Зураг 4. Хуваалт хоорондын хэндоврын хоцрогдол

ХҮСНЭГТ 1. ХЭНДОВРИЙН ТӨРӨЛ БА ХОЦРОГДОЛ

Хоцрогдол	***A хувилбар***	***B хувилбар***	***C хувилбар***	***D хувилбар***
1A архитектур	72 ms	81 ms	78 ms	110 ms
1B архитектур	60 ms	68 ms	66 ms	105 ms

VI. Дүгнэлт

Энэхүү судалгааны ажлаараа бид 5G сүлжээнд хуваалт үүсгэсэн үед хэрэглэгчийн баталгаажуулах, хуваалт хооронд хэндовр хийх үйлдлүүдийг нэгтэсэн механизмыг танилцуулав. Бидний механизм нь хэрэглэгч баз станц дээрээс шууд гадаад сүлжээ рүү ханддаг байхаар сүлжээний архитектурт өөрчлөлт оруулснаараа онцлогтой юм. 3GPP байгууллагаас гаргасан стандарт механизмтай харьцуулалт хийсэн үр дүн нь өөрчлөгдсөн архитектуртай үед бидний механизм сайн ажиллаж, энэ төрлийн хэрэглээнд чухал давуу тал буй болгосон байна. Цаашид санал болгосон механизмд ашиглах түлхүүрийн менежмент, нууцлалын алгоритмын хөгжүүлэлт, хурдан хөнгөн ажиллах талаар механизмыг сайжруулах болно. Нөгөө талаас мобайл операторын ашиглаж буй технологийн зөрүүтэй байдлаас хамаарч хэрэглэгч баталгаажуулалтын механизм төвөгтэй болж магадгүй юм. Энэхүү ажлыг өргөжүүлж хэрэглээнд нэвтрүүлснээр мобайл операторууд хоорондоо дэд бүтцийг хуваалцаж ашиглахаас гадна үйлдвэр, уурхай зэрэг газарзүйн хувьд алсмагдал байршилд мобайл операторын баз станцыг ашиглан утасгүй сүлжээ байгуулж дотоод хэрэглээгээ шийдэхэд нэг түлхэц болно.

Ашигласан материал

[1] LEI., Wan, A. C. Soong, Liu Jianghua, Wu Yong, Brian Classon, Weimin Xiao, David Mazzarese, Zhao Yang, and Tony Saboorian. 5G system design. Cham: Springer International Publishing, 2021.

[2] Pirinen, Pekka. "A brief overview of 5G research activities." In 1st international conference on 5G for ubiquitous connectivity, pp. 17-22. IEEE, 2014.

[3] Attar, Hani, Haitham Issa, Jafar Ababneh, Mahdi Abbasi, Ahmed AA Solyman, Mohammad Khosravi, and Ramy Said Agieb. "5G system overview for ongoing smart applications: structure, requirements, and specifications." Computational intelligence and Neuroscience 2022, no. 1 (2022): 2476841.

[4] Rodriguez, Jonathan. Fundamentals of 5G mobile networks. John Wiley & Sons, 2015.

[5] Prasad, Anand R., Sivabalan Arumugam, and Alf Zugenmaier. "3GPP 5G security." Journal of ICT Standardization 6, no. 1-2 (2018): 137-158.

[6] Zhang, Shunliang. "An overview of network slicing for 5G." IEEE Wireless Communications 26, no. 3 (2019): 111-117.

[7] Abdelwahab, Sherif, Bechir Hamdaoui, Mohsen Guizani, and Taieb Znati. "Network function virtualization in 5G." IEEE Communications Magazine 54, no. 4 (2016): 84-91.

[8] Olimid, Ruxandra F., and Gianfranco Nencioni. "5G network slicing: A security overview." Ieee Access 8 (2020): 99999-100009.

[9] Huang, Xiaoting, Takahito Yoshizawa, and Sheeba Backia Mary Baskaran. "Authentication mechanisms in the 5G system." Journal of ICT Standardization 9, no. 2 (2021): 61-78.

[10] Kunz, Andreas, and Xiaowei Zhang. "New 3GPP security features in 5G phase 1." In 2018 IEEE Conference on Standards for Communications and Networking (CSCN), pp. 1-6. IEEE, 2018.

Албан хэрэг хөтлөлтийн мэдээллийн систем нэвтрүүлснээр гарах үр ашиг

Баярмагнай Даваахүү
Мэдээлэл, компьютерын ухааны тэнхим МУИС, Мэдээллийн технологи, электроникийн сургууль
Монгол Улс, Улаанбаатар хот
23m1num0439@stud.num.edu.mn

Самбуу Уянга*
Мэдээлэл, компьютерын ухааны тэнхим МУИС, Мэдээллийн технологи, электроникийн сургууль
Монгол Улс, Улаанбаатар хот
uyanga@seas.num.edu.mn

Алтанцэцэг Мөнхзул
Мэдээлэл, компьютерын ухааны тэнхим МУИС, Мэдээллийн технологи, электроникийн сургууль
Монгол Улс, Улаанбаатар хот
munkhzul.a@num.edu.mn

* *Холбоо барих зохиогч*

Хураангуй — Байгууллага, бизнесийн үйл ажиллагаанд мэдээллийн системийг нэвтрүүлснээр өдөр тутмын үйл ажиллагааг автоматжуулах, бүх шатны шийдвэр гаргалтыг шуурхай болгох төдийгүй эдийн засгийн хувьд үр ашигтай байдаг. Мэдээллийн системийг байгууллагын үйл ажиллагаанд нэвтрүүлснээр зардал хэмнэсний үр ашиг, сайжруулагдсан мэдээлэл болон үйлчилгээний үр ашиг гардаг. Энэхүү өгүүлэлд мэдээллийн системийн эдийн засгийн үр ашгийг уламжлалт цаасан суурьт албан хэрэг хөтлөлтийн үйл ажиллагааг мэдээллийн систем ашиглан сайжруулснаар гарах эдийн засгийн хэмнэлт, үр ашгийг архив, албан бичиг хөтлөлтийн системийн жишээн дээр харуулахаар зорьсон.

Түлхүүр үг — мэдээллийн систем, архив, цахимжуулалт, үйл ажиллагааны нөлөөлөл, албан хэрэг хөтлөлт, эдийн засгийн үр ашиг

I. ОРШИЛ

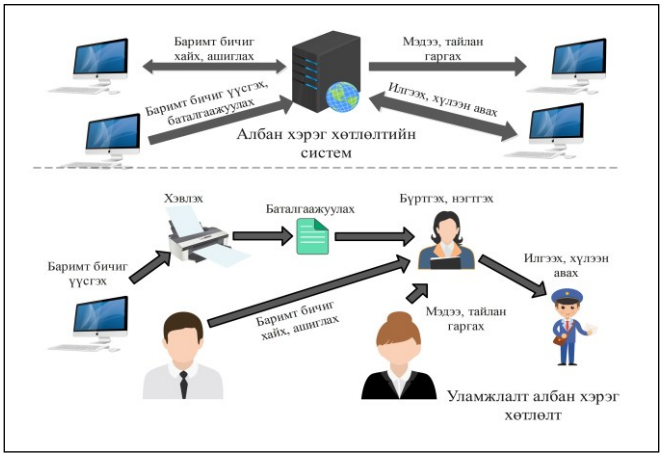
Монгол Улсын цахим засаглалыг Нэгдсэн Үндэстний Байгууллагын Тогтвортой хөгжлийн зорилтын тайланд суурь түвшнийг 0.5824 (2018 он) хэмээн үнэлж, 2020 оны байдлаар 0.6497, 2022 оны байдлаар 0.7209, 2024 оны байдлаар 0.8457 (2021, 2023 онуудад үнэлээгүй) гэж гаргасан байна. 2025 оны зорилгыг 0.6358 гэж таамагласан нь 2024 онд 0.2099 нэгжээр давсан байгаа нь манай улсын цахим засаглалын индекс тогтвортой нэмэгдэж, төлөвлөснөөс илүү гарсныг илтгэж байна [1].

Цахим баримт бичгийн удирдлага, зохион байгуулалтыг шинжлэх ухааны байдлаар хэрэгжүүлж, буруу арга барилаас үүдсэн ихээхэн хохирлыг арилгахын тулд баримт бичгийн удирдлагын онол, судалгаа болон хэрэгжилтийг сайжруулах нь нэн чухал үүрэгтэй байна [3]. Жишээлбэл, Европын Цөмийн Судалгааны Байгууллага (CERN) нь Цахим өгөгдлийн удирдлагын EDMS системийг амжилттай нэвтрүүлж, байгууллагын цахим баримт бичгийн боловсруулалтын EDH системээр тус байгууллагын ажиллагсад, хүрээлэнгийн удирдлагын нэгж хоорондын бүх харилцааны үүргийг гүйцэтгэдэг. Тухайлбал, баримт бичигт гарын үсэг зурах, ажилтантай холбоотой асуудлууд (гэрээ үзэглэх, дахин үнэлгээ хийх, амралт гэх мэт), мэргэжил дээшлүүлэх сургалтад бүртгүүлэх, шалгалт өгөх зэрэг багтдаг. Уг систем өдөрт 2,000 баримт бичиг боловсруулах боломжтой [2].

Монгол Улсад албан хэрэг хөтлөлтийн үйл ажиллагааг Архив, албан хэрэг хөтлөлтийн тухай хууль (2020), Албан хэрэг хөтлөлтийн нийтлэг журам (2024), Албан хэрэг хөтлөлтийн стандарт (2021)-ын дагуу хэрэгжүүлдэг. Архив, албан хэрэг хөтлөлтийн тухай хууль (2020)-ын дагуу төрийн бүх шатны байгууллагууд нь албан хэрэг хөтлөлтөд цахим тогтолцоог нэвтрүүлэх үүрэгтэй бөгөөд төрийн 5,574 байгууллага (ҮСХ, 2024) 15 төрлийн мэдээллийн системийг албан бичиг боловсруулах үйл ажиллагаанд ашиглаж, Нийтийн мэдээллийн ил тод байдлын тухай хууль (2021)-ын дагуу албан бичиг солилцооны дундын www.digidoc.gov.mn системээр албан бичгийг солилцдог [4]. Баримт бичгийн MNS 5140-2021 стандартаар захирамжлалын, зохион байгуулалт, мэдээлэл лавлагааны баримт бичиг гэсэн 3 ангиллаар 31 нэр төрлийн баримт бичгийг зохион бүрдүүлэхэд тавигдах шаардлагыг тогтоосон ба эдгээрээс албан бичгийг мэдээллийн систем ашиглан боловсруулж, дамжуулдаг. Албан бичиг солилцооны дундын системд 2025 оны 3 дугаар сарын 25-ны өдрийн байдлаар 15 систем, төрийн 4,724 байгууллага холбогдож, нийт 526,107 албан бичгийг илгээж, 562,902 албан бичгийг хүлээн авчээ. Ингэхдээ 751,117 ширхэг .pdf файлыг дамжуулж, 797,864 API хүсэлтэд хариу өгчээ [5].

Мэдээллийн систем ашиглан албан хэрэг хөтлөлтийг зохион байгуулах болон уламжлалт цаасан суурьт баримтаар албан хэрэг хөтлөлтийн үйл ажиллагааг явуулах ажлын урсгалыг Зураг 1-д харуулав.

Байгууллагын архивын үйл ажиллагааны нийтлэг журам (2021)-д байгууллага бүр тухайн жилд хөтөлж дуусан баримт бичгээр архивын хадгаламжийн нэгж үүсгэн байгууллагын архивт шилжүүлэх үүрэгтэй гэж заасан. Архивын нэг нэгж нь олон төрлийн баримт бичгээс бүрдэх бөгөөд нэг баримт бичиг дунджаар 205 орчим хуудастай байдаг [6] ба түүнийг баталгаажуулахад баримт бичгийн төслийг 2 удаа, эх хувийг 2 удаа гаргасан байна [7]. Байгууллагын архивт хадгалах баримт бичгийн 35% нь байнга хадгалах, 60% нь түр хадгалах, 5% нь 70 жил хадгалах баримт бичиг байгаагаас 8% нь албан бичиг байдаг аж [6].



Зураг 1. Албан хэрэг хөтлөлтийн уламжлалт арга болон мэдээллийн систем ашиглалт

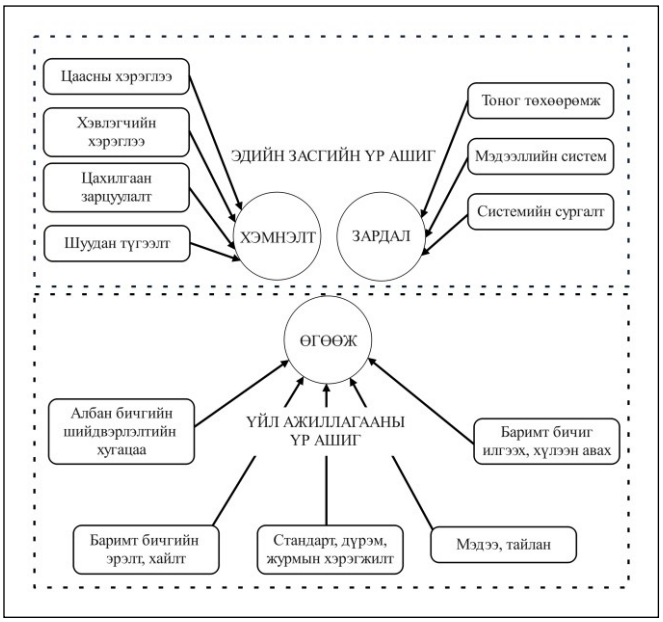
XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Архив, албан хэрэг хөтлөлтийн тухай хуулийн дагуу байнга хадгалах баримтыг заавал цаасан хувь, цахим хувийг хамт үйлдэж, түр болон 70 жил хадгалах баримтыг цаасан хэлбэрт үйлдэхгүй байж болохоор хуульчилсан нь цаасан баримтын хэрэглээг хязгаарлаж өгсөн эрх зүйн зохицуулалт болсон.

II. АРГА ЗҮЙ

Энэхүү судалгаанд харьцуулсан тоон шинжилгээний аргыг ашигласан бөгөөд байгалийн болон нийгмийн шинжлэх ухаанд өргөнөөр ашиглагдаж аливаа юмсын үзэгдэл, шинж чанар, үр дүнг тайлбарлах, тодорхойлоход өргөнөөр ашигладаг арга юм. Харьцуулсан шинжилгээг тоон, чанарын болон фаззи гэж ангилж болно. Тоон харьцуулсан шинжилгээ нь хувьсагч төвтэй, чанарын харьцуулсан шинжилгээ нь үзэгдэл төвтэй, харин фаззи харьцуулсан шинжилгээ нь цуглуулга төвтэй байдаг [8].

Албан хэрэг хөтлөлтийн системийн үйл ажиллагааны нөлөөллийг ашиглагч байгууллагад үзүүлж буй эдийн засгийн хэмнэлт, үйл ажиллагааны үр ашиг гэсэн бүрэлдэхүүнээр тооцсон. Эдийн засгийн хэмнэлтийг хэмнэлт ба зардал, үйл ажиллагааны үр ашгийг байгууллагын үйл ажиллагаанд гарсан үр өгөөж гэсэн загвараар (Зураг 2) судалсан. Судалгааны өгөгдлийг өмнөх судалгааны ажлууд, төрийн архивын тоон мэдээ, Төрийн худалдан авах ажиллагааны www.tender.gov.mn систем зэрэг эх сурвалжаас цуглуулсан болно.



Зураг 2. Судалгааны загвар

III. ҮР ДҮН

А. Эдийн засгийн хэмнэлт

Эдийн засгийн хэмнэлтийг цаас, хэвлэх төхөөрөмж, шуудан, харилцаа, цахилгааны зардлыг хэрхэн хэмнэж буйгаар тооцож, зардлыг мэдээллийн системийг хөгжүүлэх болон ашиглалтын гэж ангилан тооцоолсон. Цаасны хэрэглээ буурснаар байгалийн нөөцийн хэмнэлт өгөх дүнг харьцуулан судаллаа.

ХҮСНЭГТ 1. ТӨРИЙН БАЙГУУЛЛАГЫН БАРИМТ БИЧГИЙН БҮТЭЦ, ХУУДАСНЫ ТОО

Хадгалалт	Эзлэх хувь	Нэг жилд ашиглах хуудас цаас	
Байнга хадгалах баримт	35%	35%x177x205x4	50,799
Түр хадгалах баримт	60%	60%x177x205x4	87,084
70 жил хадгалах баримт	5%	5%x177x205x4	7,257

1) Цаасны хэмнэлт

Монгол Улсын Архивын ерөнхий газрын www.edoc.mojha.gov.mn, Нийслэлийн Архивын газрын www.edoc.ulaanbaatar.mn системд 2017-2024 онд боловсруулсан албан бичгийн өгөгдлөөс харахад 1 албан бичгийг боловсруулж баталгаажуулахад төсөл болон эх хувийг дунджаар тус бүр 2 хувь үйлддэг. Архивын ерөнхий газрын www.e-document.gov.mn системд 2000-2022 онд бүртгэлтэй 4,396 байгууллагын нэг жилд боловсруулсан хадгаламжийн нэгжийн дундаж тоо 177, нэг хадгаламжийн нэгж дунджаар 205 хуудастай байна.

Нэг байгууллага нэг жилд баримт бичиг үүсгэхэд ашиглах цаасны хэмжээг Хүснэгт 1-д харуулсны дагуу 145,140 хуудас гэж тооцвол мэдээллийн системээр үүний 8% болох албан бичгийг үүсгэдэг гэж үзэж 11,611 албан бичгийг хасвал 133,529 хуудас болно. Үүнийг 2024 оны байдлаар ҮСХ-ны тайланд дурдагдсан төрийн 5,574 байгууллагаар авч үзвэл 744,290,646 хуудас буюу ойролцоогоор А4 хэмжээтэй 1,488,581 ширхэг боодол бичгийн цаас болно.

Худалдан авах ажиллагааны www.tender.gov.mn системээс 2024-2025 онд А4 хэмжээтэй бичгийн цаас дангаар худалдан авахаар зарласан 8 тендерийг жишээ болгон авч үзвэл, А4 хэмжээтэй 80г/м2 нягтралтай 500 хуудас буюу нэг боодол цаасны дундаж үнэ 17,375.69 төгрөг байна. Иймд улсын хэмжээнд хэрэглэж байгаа 1,488,581 боодол цаасыг дунджаар 17,375.69 төгрөгийн үнэтэйгээр тооцоход ойролцоогоор 25,865,122,000 төгрөг болох юм (Хүснэгт 2).

ХҮСНЭГТ 2. А4 ХЭМЖЭЭТЭЙ 1 БООДОЛ БИЧГИЙН ЦААСНЫ ҮНИЙН ТООЦООЛОЛ

Он	Тендер шалгаруулалтын дугаар	Тоо	Нийт үнэ (төгрөг)	Нэгжийн үнэ (төгрөг)
2025	БЗСБЧДЭХАШТШТГ/20250202004	1,760	30,000,000	17,045
2025	НГЗБА/20250202013	2,455	43,000,000	17,515
2025	ЗБДС/20250102002	9,235	157,000,000	17,001
2025	ЦЕГ/20250202002	1,500	30,000,000	20,000
2025	ЦЕГСАТА/20250202002	2,665	40,000,000	15,009
2024	ДААТҮТӨХК/20240202120	1,320	28,564,800	21,640
2024	ИНЕГ/20240202035	3,000	46,000,000	15,333
2024	ЦДҮС/20240202167	1,885	29,145,000	15,462
	Дундаж үнэ			17,375.69

ХҮСНЭГТ 3. ЦААСНЫ ХЭРЭГЛЭЭГ МОД, УС, ЦАХИЛГААН, ХҮЧИЛТӨРӨГЧӨӨР ИЛЭРХИЙЛСЭН БАЙДАЛ (ЖИЛЭЭР)

Хэрэглээ	Нэгж
Хэмнэсэн цаас	744,290,646 хуудас, 59,543 тонн
Хэмнэсэн мод	142,903 тонн, 1,012,231 ширхэг мод
Хэмнэсэн ус	5,954,300 тонн
Хэмнэсэн цахилгаан	452,526 мега ватт
Хүчилтөрөгч	52,636,012 хүн/өдөр

Цаасны үйлдвэрлэл нь дэлхийн агаар, усны бохирдол, хог хаягдал, хүлэмжийн хий ялгаралтад ихээхэн нөлөөлдөг. Энэ нь байгалийн нөөцийн хэт хэрэглээ, ой модны устгал, агаар, ус, газрын бохирдолд ноцтойгоор нөлөөлж, байгаль орчинд ихээхэн хүндрэлийг үүсгэж байна [10].

Нэг тонн 1 дүгээр зэргийн цаас боловсруулахад мод 2.4 тонн, ус 100 тонн, цахилгаан 7.6 мега ватт зарцуулдаг. Нэг насанд хүрсэн мод нь өдөрт 52 хүний амьсгалах хүчилтөрөгчийн хэрэгцээг хангах чадалтай бөгөөд 1 тонн цаас гаргаж авахад 17 ийм мод ашигладаг [11]. Үүнийг энэ судалгааны үр дүнд гарсан цаасны хэрэглээгээр Хүснэгт 3-т харуулав.

Цаасан баримт нь унших явцад тэмдэглэх, олон баримтыг зэрэг унших зэрэг уян хатан шинжтэй, орон зайн зохион байгуулалтад онцгой давуу талтай байж, мэдээлэл солилцох, хэлэлцэх, хамтын ажиллагааны орчинд архивын үүрэг гүйцэтгэдэг. Гэвч мэдээллийн систем нэвтрүүлснээр байгууллагад олон давуу тал бий болдог. Тухайлбал, ажлын бүтээмж, гүйцэтгэлийг нэмэгдүүлэх, зардлыг бууруулах, ашиг орлогыг өсгөх, байгаль орчинд ээлтэй байх зэрэг орно [9].

2) Хэвлэх төхөөрөмжтэй холбоотой зардлын хэмнэлт

Хэвлэх төхөөрөмжийн хорны хэрэглээ болон үнийг Төрийн худалдан авах ажиллагааны www.tender.gov.mn системээс цуглуулж, нэг хорны үнэ болон хэвлэх А4 хэмжээтэй хуудас цаасны тоог гаргаж, Хүснэгт 6-д үнийн тооцооллыг харуулав. Төрийн 5,574 байгууллагын албан хэрэг хөтлөлтийн шатанд нэг жилд хэрэглэдэг 744,290,646 хуудас цаасыг хэвлэх хорны хэмжээ, үнийг Хүснэгт 4-өөр тооцоход 471,966 хор шаардагдах ба үүнийг мөнгөн дүнгээр илэрхийлбэл 16,501,578,521 төгрөг болно.

3) Цахилгаан эрчим хүчний хэмнэлт

А4 хэмжээтэй нэг нүүр цаасыг хэвлэхэд 50 ватт цахилгаан зарцуулдаг [12]. Монгол Улсад 2025 оны байдлаар мөрдөгдөж байгаа цахилгааны тарифыг Эрчим хүчний зохицуулах хорооны 2024 оны 11 дүгээр сарын 01-ний өдрийн 632 дугаар тогтоолын дагуу байгууллагын хэрэглээг 300 киловаттаас их тохиолдолд 285 төгрөгөөр тооцохоор заасан. Иймд нэг жилд хэрэглэдэг 744,290,646 хуудас баримтыг хэвлэхэд 37,214,532 киловатт цахилгаан зарцуулах буюу тарифт үнээр 10,606,141,705 төгрөг болно.

4) Шуудан, түгээлтийн хэмнэлт

Монгол шуудан ХК-ийн 2025 оны 1 дүгээр сарын 24-ний өдрийн шинэчилсэн үнийн тарифаар аймаг, хот доторх бичиг захидлын шуудан хүргэлтийн үнэ 1,000-7,700 төгрөг, хот хоорондын тариф

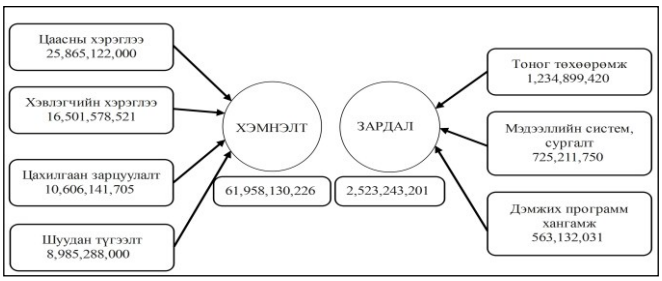
ХҮСНЭГТ 4. ХЭВЛЭГЧИЙН ХОРНЫ ҮНИЙН ТООЦООЛОЛ

Он	Тендер шалгаруулалтын дугаар	Тоо	Нийт үнэ	Нэгжийн үнэ	Хэвлэх цаасны тоо
2024	БГХУДИХАШ ШТГ/20240202008	1,286	62,000,000	48,212	1,654
2024	АТҮТ/ 20240202033	2,776	57,000,000	20,533	-
2024	МШ/ 20240202070	1,375	43,000,000	31,273	1,500
2024	ИНЕГ/ 20240202036	979	39,000,000	39,837	-
Дундаж				34,963.49	1,577

1,500-9,900 төгрөг байна (1 албан бичгийн жинг 200 грамм хүртэлхээр тооцов). Хүргэлт нь энгийн, бүртгэлтэй, буухиа гэсэн 3 төрөлтэй байх ба төрлөөс хамаарч хот дотор 1-2 хоног, хот хоорондын 3-5 хоног байдаг [13]. Нэг байгууллагад нэг жилд үүсдэг 145,140 хуудас баримт бичгийн 8% болох албан бичгийг бусдад дамжуулдаг ба ингэхдээ цаасан 1 хувийг шуудангаар илгээдэг гэж тооцвол 2,902 хуудас албан бичгийг илгээдэг тооцоо гарч байна. Нийслэлийн архивын газрын www.edoc.ulaanbaatar.mn системийн нэг албан бичгийн дундаж хуудасны тоо 1.8 байдгаар жишиж тооцвол нэг байгууллагын нэг жилд илгээдэг албан бичгийн тоо 1,612 болно. Улсын хэмжээнд нийт 8,985,288 албан бичгийг шуудангаар солилцдог гэж тооцвол шуудан хүргэлтийн хамгийн бага дүн 1,000 төгрөгөөр тооцвол 8,985,288,000 төгрөг болно.

5) Мэдээллийн систем ашиглахад үүсэх анхны зардал

Энэхүү судалгаа нь Монгол Улсын Засгийн газар, Бүгд Найрамдах Солонгос Улс хоорондын ерөнхий хэлэлцээрийн хүрээнд Архивын ерөнхий газарт хэрэгжиж буй МОН-16 төслийн Албан хэрэг хөтлөлтийн системийг суурь болгон ашиглаж байгаа тул 2023 онд зарлагдсан олон улсын тендерийн баримт бичигт тусгагдсан мэдээллийг ашиглан шаардагдах тоног төхөөрөмж, мэдээллийн систем, сургалт, ашиглалт зэргийг төрийн нийт байгууллагын хүрээнд авч үзсэн болно. Уг төслийн хүрээнд сервер, галт ханыг параллель байхаар тооцож, тус бүр 2 төхөөрөмж худалдан авахаар төлөвлөсөн. Албан хэрэг хөтлөлтийн системийн хөгжүүлэлт болон нэвтрүүлэх сургалтыг багцалж, 213,000 ам.доллар байхаар гэрээлсэн бол системийг ажиллуулах тоног төхөөрөмж 362,698 ам.доллар, дэмжин ажиллах программ хангамж 165,396 ам.долларoор (Монгол банкны сүүлийн нэг жилийн дундаж ханшаар 1 ам.долларыг 3,404.75 төгрөгөөр тооцов) тус тус тооцож, эдийн засгийн үр ашгийг төгрөгөөр илэрхийлэв (Зураг 3).



Зураг 3. Эдийн засгийн үр ашгийн харьцуулалт (төгрөгөөр)

Б. Нийгмийн нөлөөлөл

Мэдээллийн систем нь байгууллага, салбарын ажил үүргийг автоматжуулж, хялбарчлах үүрэгтэй. Энэ нь байгууллага, салбарын үйл ажиллагааны үр ашгийг ихээр нэмэгдүүлэхэд хувь нэмрээ оруулдаг. Гар ажиллагаатай, давтамж ихтэй ажлуудыг автомат үйл ажиллагаагаар солих нь тодорхой ажлыг гүйцэтгэхэд шаардагдах хугацаа, хүчин чармайлтыг үлэмж багасгадаг нь зөвхөн хүний алдаа гарах магадлалыг бууруулах төдийгүй ажилтнуудыг илүү үнэ цэнэ бүтээх, инновац, стратегийн өсөлтөд чиглэсэн үйл ажиллагаандаа цаг гаргах боломжийг бүрдүүлдэг [14]. Түүнчлэн мэдээллийн систем нь өгөгдлийг нэгтгэх, шинжлэх өргөн боломжийг олгож, бизнесийн үйл явцыг оновчтой болгоход маш чухал үүрэгтэй. Жишээлбэл, өөр өөр хэлтсийн мэдээллийг нэгтгэн боловсруулах нь үйл ажиллагаагаа оновчтой эрхлэн явуулахад туслахын зэрэгцээ байгууллагын хэв маяг, хандлагыг тодорхойлж, мэдээлэлд суурилсан шийдвэр гаргах боломжийг нэмэгдүүлдэг. Удирдлагын зүгээс бодит хугацааны мэдээлэлд үндэслэн оновчтой, шуурхай зохион байгуулалтыг хийж, нөөцийн хуваарилалтыг оновчтой хийх, үйл ажиллагааны гүйцэтгэлийг бүхэлд нь сайжруулах боломжтой болдог [15].

6) Албан бичгийн шийдвэрлэлтийн хугацаа

Баримт бичгийн шийдвэрлэлтийн хугацааг яаралтай шийдвэрлэх асуудлыг ажлын 3 хоногт, баримт бичигт дурдсан бол тэр хугацаанд, хууль тогтоомжоор заасан бол тэр хугацаанд, эдгээрээс бусад тохиолдолд ажлын 7 өдөр байхаар Архив, албан хэрэг хөтлөлтийн тухай хуульд заасан.

2014-2016 оны байдлаар мэдээллийн систем ашигладаггүй байх үед Нийслэлийн нутгийн захиргааны байгууллагуудын албан бичгийн шийдвэрлэлийн дундаж хугацаа 15.5 хоног, харин 2023 оны байдлаар мэдээллийн систем ашиглан 4.34 хоногт, 2024 онд 5.02 хоногт шийдвэрлэснээр Албан хэрэг хөтлөлтийн www.e-doc.ulaanbaatar.mn системд бүртгэгджээ.

7) Системийн хэрэглээ

Баримт бичгийн эрэлт, хайлт, загвар ашиглалт, дамжуулалт, хүлээн авах зэргийг тооцоолохдоо систем ажиллаж буй веб сервер (WEB), аппликэйшн сервер (WAS), өгөгдлийн сангийн сервер (DB)-ийн хариу өгсөн хугацааг сүүлийн нэг жилийн байдлаар авч үзлээ (Зураг 4).

Зураг 4. WEB, WAS, DB серверийн RTT (round trip time)

WEB сервер хэрэглэгчээс хүсэлт хүлээн авч боловсруулаад WAS серверт дамжуулж, хариу илгээх хугацааг WEB RTT, WAS сервер WEB серверээс хүсэлт хүлээн авч боловсруулаад DB серверт дамжуулж, хариу хүлээн авах хугацааг WAS RTT, DB сервер WAS серверээс ирсэн хүсэлтийг боловсруулах хугацааг DB RTT гэж тус тус тэмдэглэв. Зураг 4-өөс WEB RTT=1.9 ms, WAS RTT=1.9 ms , DB RTT=1.8 ms байгааг харж болно. Эдгээрийн нийлбэр нь 5.6 ms гарч байгаа нь хэрэглэгчийн хүсэлтийг хүлээн авснаас буцаж хариу өгөх хүртэлх хугацаа юм.

8) Стандарт, дүрэм, журмын хэрэгжилт

Албан хэрэг хөтлөлтийн мэдээллийн систем ашигласнаар хууль, дүрэм, журам, заавраар тогтоосон ажлын урсгалыг бүрэн хэрэгжүүлэх, хянах боломжтой. Энэ зорилгоор ашигладаг www.edoc.ulaanbaatar.mn, www.gas.e-document.gov.mn, www.edoc.*.mn, www.erp.e-mongolia.mn систем нь албан бичиг боловсруулалттай холбоотой хууль, стандарт, журмын хэрэгжилтийг хангаж хэвлэмэл хуудас боловсруулах батлуулах модультай. Иймд хэрэглэгчийн мэдлэгээс үл хамааран систем нь стандарт, дүрэм, журмын хэрэгжилтийг хангаж ажилладаг.

IV. Дүгнэлт

Энэхүү судалгаа нь албан хэрэг хөтлөлтийн системийн байгууллагад үзүүлэх нөлөөллийг эдийн засгийн болон нийгмийн гэсэн хоёр бүлэгт хуваан шинжилж, тэдгээрийн онцлогийг тодорхой жишээ, баримтаар харуулахыг зорьсон.

Өдгөө манай улсад ашиглаж байгаа албан хэрэг хөтлөлтийн системүүд нь албан бичиг боловсруулж, солилцох үүрэгтэй байгаа нь албан хэрэг хөтлөлтийн шатанд үүсдэг 31 нэр төрлийн баримтыг бүрэн хамарч чадахгүй байна. Албан бичиг нь дунджаар байгууллагад үүсэж буй баримт бичгийн 8%-ийг эзэлж, үлдсэн 92%-ийг цаасан хэлбэрээр боловсруулж, баталгаажуулдаг. Цаасан хэлбэрээр боловсруулж, ашиглаж байгаа баримт бичигт нэг жилд А4 хэмжээтэй, 80г/м2 нягттай 744,290,646 хуудас бичгийн цаас зарцуулдаг ба ийм хэмжээний цаасыг үйлдвэрлэхэд 1,012,231 ширхэг мод, 5,954,300 тонн ус, 452,526 мегаватт цахилгаан зарцуулдаг. Модны тухайд, 1,012,231 ширхэг мод нь Монгол Улсын 3,544,835 иргэний ойролцоогоор 15 өдөр амьсгалах хүчилтөрөгчийг ялгаруулах боломжтой ба нэг м2 газарт нэг мод ургадаг гэж үзвэл 101.2 га газрыг ойжуулах боломжтой зэргээс харахад цаасны хэрэглээ нь байгалийн асар их нөөц ашигладаг нь харагдаж байна. Иймд албан хэрэг хөтлөлтийн үйл ажиллагааг цахимжуулах, мэдээллийн системийг нэвтрүүлэх нь нийгэм, эдийн засгийн өндөр ач холбогдолтой.

Бидний судалгаагаар албан хэрэг хөтлөлтийн системийг бүрэн ашигласнаар 2,523,243,201 төгрөгийн зардал гарахаар, харин уламжлалт цаасан баримт бичгээр ашиглахад 61,958,130,226 төгрөг зарцуулахаар байна. Үүнээс үзэхэд мэдээллийн систем ашиглах нь уламжлалт цаасан баримт ашигласнаас даруй 24.5 дахин бага зардал шаардахаас гадна баримт бичгийг богино хугацаанд хайж олох, стандарт, дүрэм, журмын хэрэгжилтийг бүрэн хангах, баримт бичгийн шийдвэрлэлтийн хугацааг хурдасгах, төрөөс иргэдэд үзүүлэх үйлчилгээний хурдыг нэмэгдүүлэх, төрийн залгамж халаа тасрахгүй байх зэрэг олон давуу тал, эерэг нөлөөтэй байна.

НОМ ЗҮЙ

[1] Монгол Улсын тогтвортой хөгжлийн зорилтын тайлан, http://sdg.gov.mn/Home/Page/SDV_Indicator2050, 2024.

[2] I. Alexandrov *et al*., “Development and implementation of electronic document management system "EDMS Dubna" at JINR”, Procedia Computer Science, 145 (2018) 248–253 253 6, 2018

[3] Jin H.,Cheng W., Jie M., Mingxin L., Yingchun W., Jin S., “Research on Electronic Document Management System Based onCloud Computing”, Computers, Materials & Continua. DOI:10.32604/cmc.2021.014371, 2021.

[4] Davaakhuu B., Uyanga S., Oyun-Erdene N., Azzaya B., “Solution for enhancing the information system of the archives sector In mongolia”, 2024.

[5] Албан бичиг солилцооны дундын систем, http://www.digidoc.gov.mn/, 2025.

[6] Төрийн архивын нэгдсэн систем, https://www.e-document.gov.mn, 2025.

[7] Албан хэрэг хөтлөлтийн систем, https://edoc.mojha.gov.mn, 2025.

[8] Sa’ei, Ali, Comparative research method: Quantitative, historical and fuzzy analysis. Tehran. Iran. pp. 10-50, 2013.

[9] Obeidat, M. A. Empirical analysis for the factors affecting realization of paperless office. International Journal of Economics, Commerce and Management, 3(6), 773-792, 2015.

[10] Shah, I. A., Amjed, S., Alhamar Alkathiri, N. The economics of paper consumption in offices. Journal of Business Economics and Management, 20(1:43-62). https://doi.org/10.3846/jbem.2019.6809, 2019.

[11] Ahmad Tutus, Mustafa Cicekler. Waste Paper Recycling: Contributions to Giresun and Turkey Economies, https://www.researchgate.net/publication/328927430_Waste_Paper_Recycling_Contributions_to_Giresun_and_Turkey_Economies, 2018.

[12] Прогрессив харилцааны холбоо (APC), https://www.apc.org/sites/default/files/SustainableITtips5_0.pdf

[13] Монгол шуудан ХК, https://www.mongolpost.mn/, 2025.

[14] B. J. Ai. A. A.-A. Ali, “Operational Efficiency and The Adoption of Accounting Information System (AIS): A Comprehensive Review of The Banking Sectors,” 2020.

[15] Y. K. Surajit Bag , Lincoln C. Wood , Lei Xu d , Pavitra Dhamija , “Big data analytics as an operational excellence approach to enhance sustainable supply chain performance”.

(K-means) боловсролын салбарын оновчтой шийдвэр гаргалтад хэрэглэх

Г.Батдэлгэр
МУИС, МТЭС,
Хэрэглээний Математикийн Тэнхим
20M6NUM0103@stud.num.edu.mn

Г.Баттөр
МУИС, МТЭС,
Хэрэглээний Математикийн Тэнхим
battur@num.edu.mn

С.Батбилэг
МУИС, МТЭС,
Хэрэглээний Математикийн Тэнхим
batbileg_su@num.edu.mn

Т.Цэрэннадмид
МУИС, МТЭС,
Мэдээлэл Компьютерын Ухаан Тэнхим
tserennadmid@num.edu.mn

К.Есенбек
МУИС, МТЭС,
Мэдээлэл Компьютерын Ухаан Тэнхим
esenbek@num.edu.mn

***Хураангуй*---K-means арга нь машин сургалтын аргууд дотроос хамгийн өргөн хэрэглэгддэг машин сургалтын кластерчлах аргад тооцогддог. Уг арга нь өгөгдлүүдийг ижил төстэй байдалд (хүчин зүйлүүдэд тулгуурлан) үндэслэн кластеруудад хуваадаг арга юм. Хүчин зүйлүүдийн тоо хэдий чинээ их байх тусам кластерлах процесс улам төвөгтэй болдог. Тиймээс хүчин зүйлүүдийн хэмжээсийг үндсэн бүрэлдэхүүн хэсгийн шинжилгээг (PCA) хийн бууруулах, тооцоллыг хялбаршуулах боломжоотой.Бидэнд ирүүлсэн оролцогч талуудаас авсан санал асуулгын үр дүнг ашиглан жинг тооцож гаргасан.Бид энэхүү судалгааны ажлаараа K-means аргын онолыг судалж, боловсролын салбарын өгөгдөлд дүн шинжилгээ хийж, боловсролын салбарт оновчтой шийдвэр гаргалтыг дэмжих зорилгоор өгөгдлийг хэд хэдэн хүчин зүйлүүдээр нь кластерчлал хийсэн. Гарсан үр дүнгээ харуулах платформ хөгжүүлсэн.**

***Түлхүүр үг*---K-means, PCA (Үндсэн бүрэлдэхүүн хэсгийн шинжилгээ) , Боловсрол, Өгөгдөл**

I. Танилцуулга

Их өгөгдлийн эрин үед боловсролын байгууллагууд их хэмжээний өгөгдөл дээр дүн шинжилгээ хийж, өгөгдөл олборлолтын аргуудыг өргөнөөр ашиглах болжээ. Эдгээр аргуудын нэг нь болох K-means алгоритм нь боловсролын салбар дах хэв танилт, шийдвэр гаргалтыг дэмжих чухал аргын нэг болсон.K-means арга нь машин сургалтын алгоритм бөгөөд өгөгдлийг к ширхэг кластер болгон ангилж, өгөгдлийн цэг кластер руу хуваарилдаг. Энэхүү арга нь энгийн, үр ашигтай, их өгөгдөлтэй ажиллахад тохиромжтой байдлаараа онцлогтой.

K-means алгоритм нь боловсролын салбартай хамааралтай судалгаанд олон талт хэрэглээтэй бөгөөд оюутны сурлагын гүйцэтгэлийг шинжлэх, суралцах зан төлвийг тодорхойлох, сургалтын хөтөлбөр боловсруулах зэрэгт өргөн хэрэглэгддэг.Жишээлбэл, Liu [1] боловсролын үнэлгээний өгөгдөлд үндэслэн кластерчлалын аргачлалаар загвар боловсруулж, суралцахад тулгарч буй бэрхшээлүүдийг 95.6 хувь нарийвчлалтай илрүүлэн, оюутны хөгжлийн менежментэд бодит дэмжлэг үзүүлсэн байна. Үүнтэй адил Al-Badareen [2] оюутны гүйцэтгэлийн өгөгдлийг кластерчлах замаар мэдээллийн технологийн аргуудыг нийгмийн болон боловсролын судалгаанд амжилттай ашигласан.

Оюутны сурлагын амжилтыг урьдчилан таамаглах судалгааны хүрээнд Oyelade нар [3] кластерчлах алгоритмыг оюутны академик ахиц дэвшлийг хянах хэрэгсэл болгон ашигласан нь дээд боловсролыг төлөвлөгч түвшний хүмүүст чухал гарц болохыг харуулсан. Shovon болон Haque [4] кластерчлах алгоритмыг шийдвэрийн модтой хослуулан оюутны голч дүнг урьдчилан тооцоолох аргыг боловсруулсан нь эрт шатны дэмжлэг үзүүлэх боломжийг бүрдүүлсэн.Кластерийн бодлого нь оюутны сэтгэн бодох арга барил болон суралцах зан төлөвийг судлахад үр дүнтэй хэрэглэгдэж байна. Zhang [7] их хэмжээний өгөгдөлд суурилан багшлах чанарыг хянах, сайжируулах аргачлалд K-means алгоритмыг бусад аргуудтай хослуулан боловсруулсан. Balovsyak нар [8] K-means болон fuzzy logic-ийг нэгтгэн боловсруулсан програм хангамж ашиглан оюутны амжилтыг үнэлж, ангиллын нарийвчлалыг сайжируулсан. Мөн K-means алгоритмыг академик хандлага, гүйцэтгэлийн хандлагыг дүгнэхэд хэрэглэсэн. Agha нар [9] Mehran Их Сургуулийн оюутны амжилтыг шинжилж, оюутануудын гүйцэтгэлийн чиг хандлагыг их өгөгдөлд суурилсан судалсан.Эдгээр судалгаанууд нь K-means кластерчлалыг боловсролын судалгаанд үзүүлэх нөлөө, уг аргыг ашиглах боломжийг харуулж байгаа бөгөөд оюутны гүйцэтгэл, тэдний урам зориг, оролцоог олон талаас дүгнэн шинжилж, боловсролын чанарыг сайжруулахад суурь мэдээлэлүүдийг өгч байна. Цаашлаад оюутны мэдээллийг янз бүрийн үзүүлэлтээр нь үр дүнтэй бүлэглэх замаар боловсролын байгууллагуудад сургалтын хандлага, агуулгыг илүү оновчтой төлөвлөх боломжийг олгож байна.Бид энэхүү судалгааны хүрээнд Монгол улсын бүх төрийн өмчийн болон хувийн ерөнхий боловсролын сургуулиудын (ЕБС) болон сургуулийн өмнөх боловсролын байгууллагуудын (СӨБ) өгөгдлийг ашиглан тэдгээрийн нийтлэг өгөгдлийн шинж чанарт тулгуурлан дүн шинжилгээ хийж, оновчтой кластерын тоог тодорхойлж өгөгдлийн шинж чанарт 7 хүчин зүйлд тулгуурлан K-means алгоритмаар класстерчлал хийсэн. Мөн тухайн

класстерчлалын үр дүнг симуляци хийх программ хангамж хөгжүүлсэн. Уг судалгааны ажил нь дараах хэсгүүдээс бүрдэнэ: 2-р хэсэгт K-means алгоритмын онолын судалгааг авч үзнэ. 3-р хэсэгт өгөгдлийн шинж чанарыг танилцуулна. 4-р хэсэгт класстерчлалын үр дүнг хэлэлцэн дүгнэнэ. 5-р хэсэгт судалгааны ажлын дүгнэлтийг хийнэ.

II. K-means арга

K-means арга нь хэв таних, өгөгдөл олборлолтын салбарт өргөнөөр ашиглагддаг машин сургалтын салбартын алгоритм. Энэ арга нь хоорондоо төстэй шинж чанар бүхий өгөгдлийн к тооны кластер болгон ангилдаг. Энэ аргыг анх MacQueen[10] 1967 онд нээсэн бөгөөд энэхүү арга нь хэрэглэхэд хялбар, тооцооллыг хурдан хийх боломжийг олгодог учраас өгөгдлийн шинжилгээнд өргөн хэрэглэгддэг үндсэн алгоримт болсон.

K-means алгоритм нь дараах үндсэн алхамуудыг биелүүлдэг:

1) Оноох алхам: K ширхэг төв цэгийг (centroid) санамсаргүй байдлаар сонгож эхлүүлнэ.
2) Ангилах алхам: Өгөгдлийн цэг бүрийг хамгийн ойр байрлалтай төв цэгт хамааруулна. Өөрөөр хэлбэл K цэгт ойр байгаагаар нь өгөгдлийг кластерт хуваана.
3) Шинэчлэх алхам: Кластер бүрийн шинэ төв цэгийг тухайн кластер бүх цэгийн дундаж координатоор тооцно.Кластерийн төвийг дахин тооцсоны дараа өгөгдлийн утгуудыг дахин кластерд хуваарлана.

K-means алгоритм нь дараах үндсэн алхмуудыг агуулдаг:

1) **Давталт:** Ангилалт болон шинэчлэлтийг зардлын функц хамгийн бага утгатай болтол эсвэл өмнөх давталтаас өөрчлөгдөхгүй болтол давтана.

$$\text{argmin} \sum_{i=1}^{K} \sum_{x \in S_i} \|x - \mu_i\|^2 \tag{1}$$

Энд μ_i нь S_i кластерын төв цэг, x нь тухайн кластер доторх өгөгдлийн цэгийг илэрхийлнэ. Кластеруд нь тойрог хэлбэртэй, ялгагдахуйц бөгөөд ойролцоо хэмжээтэй үед K-means арга нь өндөр үр дүнтэй ажилладаг. Үндсэн давуу талууд нь их хэмжээний өгөгдөлд тохиромжтой, хэрэгжүүлэхэд хялбар, хурдан тооцоологддог. Гэвч тус алгоритм нь кластерийн тоо болох к-г урьдчилан тооцоолох шаардлагатай. Тооцооллын нарийвчлал нь эхний төв цэгийн байрлалаас шууд хамааралтай байдаг. Кластеруудын төв буюу анхны дөхөлтийн цэгийн сонголтоос хамаараад зорилгын функцийн утга өөрчлөгдөж байгаа.Локал минимум шийдүүд нь тухайн нэг кластерд нэмэгдэж байгаа.

III. Боловсролын өгөгдлийг кластер шинжилгээнд тохируулах нь

Энэхүү ажлаар бид ЕБС, СӨБ-ын байгууллагуудын өгөгдлийг нийт 7 шинж чанар бүхий (хүчин зүйлүүд) дээр тулгуурлан K-means кластерчлалын аргыг хэрэглэн K=3, K=4, K=5 гэсэн кластеруудад хуваасан. Мөн тухайн кластеруудад орсон элементүүдийг (сургууль, цэцэрлэгүүд) төв цэгээс алслагдсан зайн тусламжтайгаар эрэмбэлсэн. Эдгээр кластер шинжилгээнүүдийн үр дүнг харуулах программ хангамж нь MATLAB болон Python программчлалын хэлд тулгуурлан хөгжүүлэгдсэн. Бид MATLAB R2017b программыг ашиглан туршилт хийсэн. MATLAB програмын kmeans функцийг тооцоололдоо ашиглав. Энэхүү функц нь heuristic төрлийн K-means++ алгоритм дээр тулгуурладаг тул локаль шийдийг илүү хурдан олдогоороо давуу талтай. Өгөгдлийн тохиргоог байгууллагын үнэлгээнд код оноож, сургууль бүрд аймаг сумын код(Бүсчлэлийн код), өмчийн хэлбэрийн код, сургуулийн түвшний кодыг тус тус оноосон.Мөн Улаанбаатараас тухайн аймгийн төв хүртлэх зай, аймгийн төвөөс алслагдсан байдал, тухайн аймаг сумын (одоогоор тухайн сумдын хувьд дата олдоогүй байгаа тул аймгийн хэмжээнд) хөгжлийн индекс зэргийг оноов. Багануудыг мэдээлэл өгөх байдлаар нь байгууллагын үндсэн мэдээлэл, байгууллагын багшийн үндсэн мэдээлэл, байгууллагын удирдах ажилтан, багш нарын нарийвчилсан, ажилтны зэрэг мэдээлэл, байгууллагын орчин /хичээлийн байр, дотуур байр, хүчин чадал, интернет гэх мэт/, байгууллагын болон багшийн гүйцэтгэлийн үнэлгээ, хичээл төрөл бүрээр заадаг багш нарын мэдээлэл, багш нарын эрдмийн болон боловсролын зэрэг, ажилласан жил, мэргэшлийн зэрэг мэдээллийг хүйс бүрээр нь ангилалд багтаан хуваасан.

Үр дүнд нь ЕБС-ийн хувьд 871(сургууль) мөртэй, 314(үзүүлэлтүүд) багана бүхий, СӨБ-ын байгууллагын хувьд 1409(цэцэрлэг) мөртэй, 105(үзүүлэлтүүд) багана бүхий матрицыг байгуулсан. Уг матрицыг K-means аргыг хэрэгжүүлэхэд шууд ашигласан болно. Эхлээд бид өгөгдлөө кластерлахдаа ижил жинтэй гэж тооцоод кластерлаж, туршиж зүгшрүүлсэн болно.

Энэхүү аргын үндсэн зорилго нь N тооны ажиглалтын өгөгдлийг m тооны шинж чанарыг харгалзан үзэж, K ширхэг кластер болгон ангилах бөгөөд энэ нь өгөгдлийн цэгүүдийг кластерын төв цэгээс алслагдсан зайг ашиглан бүлэглэх зарчимд үндэслэнэ.

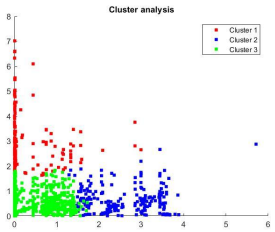
IV. Боловсролын өгөгдлийг кластер шинжилгээний үр дүн

A. Ерөнхий боловсролын сургуулиудын (ЕБС) өгөгдлийн шинжилгээний үр дүн

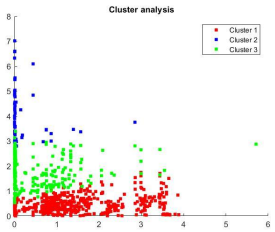
ЕБС-ийн хувьд нийт 871 ажиглалтын утгуудад буюу өгөгдлүүдэд кластерчлал хийсэн. Туршилтын явцаас харахад Газарзүйн байршил, Суралцагчдын тоо, Нийт багш+ажилчид, Хичээлийн байрны хүчин чадал, Сургуулийн ээлжийн коэффициент, Интернет үзүүлэлт, Байгууллагын гүйцэтгэлийн үнэлгээ зэрэг 7 хүчин зүйлүүд дээр тулгуурлан K=3, K=4, K=5 кластеруудад хуваах нь тохиромжтой байсан. Эхлээд дээрх 7 хүчин зүйлүүд нэг ижил жинтэй гэж үзээд кластерчлал хийв.

Мөн бид туршилтыг дээрх 7 хүчин зүйлүүд дээр өөр өөрсдийн жинтэй үед хийсэн. ЕБС, СӨБ-ын 87 оролцогч хүнээс /багш, сургалтын менежер, захирал/ санал асуулга авч, эдгээр санал асуулгын үр дүнд ЕБС-ийг газар зүйн байршил, нийт багш+ажилчид, суралцагчдын тоо, хичээлийн байрны хүчин чадал, сургуулийн ээлжийн коэффициент, интернет үзүүлэлт, байгууллагын үнэлгээ 7 хүчин зүйлээр

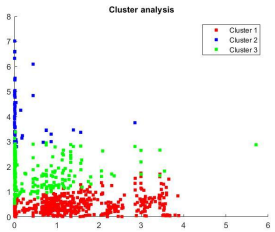
нь, СӨБ-ын хувьд газарзүйн байршил, суралцагчдын тоо, нийт багш+ажилчид, хичээлийн байрны хүчин чадал, цэцэрлэгийн ээлжийн коэффицент 6 хүчин зүйлийг гол болгон авч AHP (Analytical Hierarchy Process) аналитик шаталсан процесс аргыг хэрэглэж жинг нь бодохдоо ач холбогдлын түвшинг үнэлж тооцсон.Асуулгын үр дүнгээр газарзүйн байршил-5, суралцагчдын тоо-10, нийт багш+ажилчид-9, хичээлийн байрны хүчин чадал-8, сургуулийн ээлжийн коэффициент-8, интернет үзүүлэлт-6, байгууллагын үнэлгээ 7 гэсэн жинтэй гарсан. Энэ жинг ашиглан 7 үзүүлэлтүүдийг нэгтгэн 1 вектор болгон хувиргаж, зорилгын функцийн утгыг тооцон кластеруудад хуваасан болно. Тооцоог K=3, K=4, K=5 үед хийж туршсан. Үүний үр дүнг Зураг 1-4-д харуулав. Туршилтын үр дүнг эдгээр зургуудаас харахад кластеруудыг нэмэх бүрд хуваагдсан кластеруудын дунд давхардал нь ихсэж байна.



Зураг 1. Боловсролын өгөгдөл дээрх кластер шинжилгээний үр дүн ЕБС (K=3 үед)



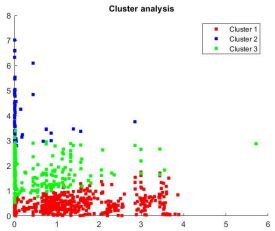
Зураг 2. Боловсролын өгөгдөл дээрх кластер шинжилгээний үр дүн ЕБС (K=3 үед)



Зураг 3. Боловсролын өгөгдөл дээрх кластер шинжилгээний үр дүн ЕБС (K=4 үед)

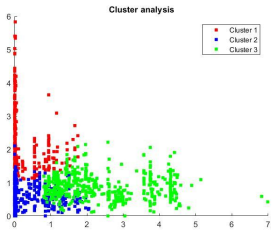
B. Сургуулийн өмнөх боловсролын байгууллага (СӨБ)

СӨБ-ын хувьд 1409 ажиглалтын утга буюу элементүүд байгаа. СӨБ-ын байгууллагуудын хувьд ЕБС-ийн хувьд

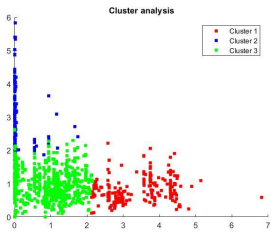


Зураг 4. Боловсролын өгөгдөл дээрх кластер шинжилгээний үр дүн ЕБС (K=5 үед)

хийсэнтэй адил аргачлалаар кластер хуваалтыг хийв. СӨБ-ын байгууллагын хувьд Газарзүйн байршил, Суралцагчдын тоо, Нийт багш+ажилчид, хичээлийн байрны хүчин чадал, Цэцэрлэгийн ээлжийн коэффициент, Байгууллагын гүйцэтгэлийн үнэлгээ зэрэг 6 төрлийн хүчин зүйлүүд дээр эхлээд дээрх үзүүлэлтүүдийг ижил жинтэй гэж үзэн K=3,K=4,K=5 гэсэн класстеруудад хуваасан туршилтуудыг хийв.Туршилтын үр дүнг Зураг 5-8-д харуулав. Тайлбар: Ээлжийн коэффициент гэдэг нь хүүхдийн тоог хичээлийн байрны багтаамид харьцуулсан харьцаагаар тодорхойлогдож байгаа. Жишээлбэл 40 хүүхдийн багтаамжтай цэцэрлэгд 50 хүүхэд СӨБ-д хамрагдаж байвал (хүүхдийн тоо)/(хичээлийн байрны багтаамж)=50/40=1.25 нягтаршилтай.

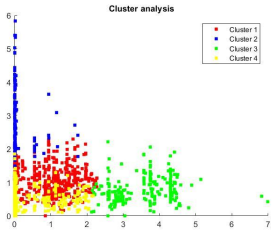


Зураг 5. Боловсролын өгөгдөл дээрх кластер шинжилгээний үр дүн СӨБ (K=3 үед)

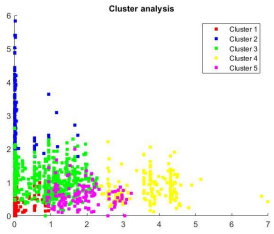


Зураг 6. Боловсролын өгөгдөл дээрх кластер шинжилгээний үр дүн СӨБ (K=3 үед)

Туршилтын үр дүнг Зураг 5-8-аас харахад кластеруудын тоог нэмэх бүрд хуваагдсан кластеруудын дунд давхардал ихсэж байгаа нь тодорхой харагдаж байна. Эндээс дүгнэхэд бидний судалгаанд ашиглагдсан өгөгдлүүдийн хувьд кластерийн тоо K=3 эсвэл K=4 байх нь илүү тохиромжтой харагдаж байна.



Зураг 7. Боловсролын өгөгдөл дээрх кластер шинжилгээний үр дүн СӨБ (K=4 үед)



Зураг 8. Боловсролын өгөгдөл дээрх кластер шинжилгээний үр дүн СӨБ (K=5 үед)

V. Дүгнэлт

Орчин үед мэдээллийн технологийн салбар хурдацтай хөгжиж статистик, математик загварчлалын арга онолд тулгуурлсан их хэмжээний өгөгдөл боловсруулах, анализ хийх, ашиглахтай холбоотой машин сургалт, хиймэл оюуны арга алгоритмууд судлаачдын анхаарлыг зүй ёсоор татаж байна.Бид энэхүү ажлаараа Монгол улсын хэмжээнд буй ЕБС болон СӨБ-ын байгууллагуудын өгөгдөлд тулгуурлан K-means алгоритмыг ашиглан уг өгөгдлийн шинж чанарын 7 хүчин зүйлд үндэслэн кластерчлал хийж ангиллаа. Туршилтын үр дүнгээс харахад кластеруудын тоо цөөн байх нь кластерчлалын үр дүнд эерэгээр нөлөөлж байна. Энэхүү кластерчлалыг хийснээр боловсролын салбарын хүртээмжтэй хөгжилд оновчтой шийдвэр гаргалтын дэмжих зорилготой.Боловсролын хүрээлэн буюу захиалагчийн зүгээс гол 7 үзүүлэлтүүд дээрээ жин бодож тооцох талаар санал ирсэн. Боловсрол Судлалын Хүрээлэнгийн нэгэн чухал ажил нь ЕБС, СӨБ-уудад очиж үнэлгээ хийх,ажлын явцад нь дүгнэлт хийх, кластерласны дараагийн ажил нь тухайн нэг кластерд байгаа ЕБС, СӨБ-уудад очиж санал зөвлөмж өгөх, кластерд байгаа сургууль цэцэрлэгүүдээр явах төсөв нь бага байдаг яавал хүргэх талаар бодлого бодох нь дараагийн шатны ажил юм.Боловсролын хүрээлэн тухайн үеийн зорилгоосоо хамааруулж шинж чанарыг нь сонгож кластерчлах боломжтой. Backend программчлал нь Python Django фреймворк, Front-End программчлал нь Next JS, өгөгдлийн сан MongoDB ашигласан болно.

Ном зүй

[1] R. Liu, "Data Analysis of Educational Evaluation Using K-means Clustering Method," Computational Intelligence and Neuroscience, vol. 2022, Article ID 3762431, 2022.

[2] M. A. Al-Badareen, "Clustering Student Performance Data Using K-means Algorithms," Journal of Computing and Informatics, vol. 13, no. 1, pp. 1–11, 2019.

[3] O. J. Oyelade, O. O. Oladipupo, and I. C. Obagbuwa, "Application of K-means Clustering Algorithm for Prediction of Students Academic Performance," arXiv preprint, arXiv:1002.2425, 2010.

[4] Md. H. I. Shovon and M. Haque, "An Approach of Improving Students Academic Performance by Using K-means Clustering Algorithm and Decision Tree," arXiv preprint, arXiv:1211.6340, 2012.

[5] O. R. Battaglia, A. M. Di Paola, and S. Fazio, "K-means Clustering to Study How Student Reasoning Lines Can Be Modified by a Learning Activity Based on the Use of a Scientific Model," Eurasia Journal of Mathematics, Science and Technology Education, vol. 12, no. 4, pp. 733–747, 2016.

[6] M. A. Al-Badareen, "Detecting Learning Patterns in Tertiary Education Using K-means Clustering," Information, vol. 13, no. 2, p. 94, 2022.

[7] Y. Zhang, "Big Data Technology for Teaching Quality Monitoring and Improvement Method," Journal of Innovation and Knowledge, vol. 6, no. 4, pp. 234–245, 2021.

[8] S. Balovsyak, O. Derevyanchuk, H. Kravchenko, Y. Ushenko, and Z. Hu, "Clustering Students According to their Academic Achievement Using Fuzzy Logic," arXiv preprint, arXiv:2312.10047, 2023. [Online].

[9] D. Agha, A. F. Meghji, and S. Bhatti, "Clusters of Success: Unpacking Academic Trends with K-means Clustering in Education," VFAST Transactions on Software Engineering, vol. 11, no. 4, pp. 1–17, 2023.

[10] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, no. 14, pp. 281–297, University of California Press, 1967. https://web.stanford.edu/yyye/yyye/Gset/.

Нарны теоретик потенциалд үүлнээс үүсэх нөлөөллийг IEEE-ийн загвараар тооцоолж үнэлэх нь

1st Отгонбаярын Уянга
Мэдээллийн технологи
Хүрээ МХТДС
Монгол, Улаанбаатар
uyanga.o@huree.edu.mn

2nd Wooyoung Song
Эмнэлгийн тоног төхөөрөмжийн ашиглалт
Хүрээ МХТДС
Монгол, Улаанбаатар
songwy54@gmail.com

***Хураангуй*—Дэлхий нийтээр сэргээгдэх эрчим хүчний хэрэглээг нэмэгдүүлэх хүчин зүтгэл улам бүр өргөжиж байгаа нь уур амьсгалын өөрчлөлтийн сөрөг нөлөөг бууруулах чухал алхам болж байна. Нарны эрчим хүч нь сэргээгдэх эрчим хүчний тогтвортой хөгжлийн чухал хэсэг болохын хувьд дэлхийн олон улс оронд эрчимтэй хөгжүүлж буй салбар юм. Гэвч нарны эрчим хүчний үйлдвэрлэл нь цаг агаарын нөхцөл байдал, ялангуяа үүлшилтээс ихээхэн хамааралтай бөгөөд энэ нь тухайн газар нутгийн цаг агаарын онцлог нарны теоретик потенциалд шууд нөлөө үзүүлдэг.**

Монгол Улс нь цэлмэг тэнгэртэй, жилд дунджаар 250-300 нартай өдөртэй, нарны цацрагийн эрчим нь 1200-1600 кВт.цаг/м² хүрдэг. Тиймээс дэлхийн нарны потенциал өндөр орнуудад багтдаг. Гэсэн хэдий ч Монгол орны газар зүйн онцлог нь эх газрын эрс тэс уур амьсгал, бүсчилсэн нутаг дэвсгэр, уулархаг болон тал хээрийн ялгаатай газрын гадаргуутай, мөн улирлын хувьд үүлшилтийн ялгаатай төлөв байдалтай байдаг нь нарны эрчим хүчний потенциалыг үнэлэхэд нарийн тооцооллыг шаардаг.

Энэхүү судалгааны ажлаар Монгол Улсын Улаанбаатар хотын онцлогт тохируулан IEEE (Institute of Electrical and Electronics Engineers) 1565-2021-ийн загвараар үүлшилтийн нөлөөллийг тооцоолж, нарны теоретик потенциалыг үнэлэх аргачлалыг хөгжүүлсэн. Судалгаанд Улаанбаатар хотод оршин суугаа айл өрхүүдийн цахилгаан хэрэглээг нарны эрчим хүчээр хангах оновчтой шийдэлд тохиромжтой тооцоолол боловсруулж, айлуудын боломжит хэмжээнд байршихад тохиромжтой 20 м² хэмжээнд нарны теоретик потенциалыг тооцоолов.

Түлхүүр үг— нарны эрчим хүч, үүлшилтийн нөлөөлөл, IEEE загвар, нарны теоретик потенциал, Улаанбаатар хот.

I. Удиртгал

Дэлхийн хэмжээнд тогтвортой хөгжил, ногоон эдийн засгийн бодлогын хүрээнд сэргээгдэх эрчим хүчний хэрэглээг нэмэгдүүлэх хандлага нэмэгдэж байна. Үүний нэг тод жишээ нь нарны эрчим хүчний технологийн хөгжил, өртгийн бууралт бөгөөд дэлхийн олон улс орнууд нарны эрчим хүчний хэрэглээг төрийн бодлогоор дэмжиж, нүүрстөрөгчийн ялгарлыг бууруулах зорилтуудаа хангахыг эрмэлзэж байна [1, 2].

Монгол улсын хувьд нарны эрчим хүчний потенциалтай орон юм. Жилд дунджаар 250-300 нартай өдөртэй, нарны цацрагийн эрчим нь 1200-1600 кВт·цаг/м² хүрдэг гэсэн үзүүлэлт нь дэлхийн бусад орнуудтай харьцуулахад нарны эрчим хүчийг ашиглахад онцгой тохиромжтой нөхцөл болж өгдөг [3, 4]. Ялангуяа Улаанбаатар хот нь өндөрлөг бүс нутагт байрлаж, жилийн дийлэнх хугацаанд нартай хэрнээ үүлтэй байдгаараа онцлогтой.

Сүүлийн жилүүдийн судалгаанаас харахад Улаанбаатар хотын үүлшилтийн байдал нь нарны эрчим хүчний ашиглалтад тодорхой хэмжээний нөлөө үзүүлж болзошгүй нь харагдаж байна. Б.Батзориг, М.Лхагвасүрэн нарын 2022 онд хийсэн судалгаанаас [4] харахад хот дахь үүлшилтийн байдал улирлаас хамаарч харилцан адилгүй байдаг бөгөөд энэ нь нарны цацрагийн эрчмийг ихээхэн бууруулж байгааг тогтоожээ.

Нөгөө талаас, Улаанбаатар хотын хүн амын өсөлт, хотжилтын үйл явц хурдацтай явагдаж байгаа өнөө үед эрчим хүчний найдвартай хангамжийн асуудлыг шийдвэрлэх шаардлага тулгараад байна. Нийслэлийн Статистикийн газрын хүн ам зүйн судалгаанаас 2024 оны мэдээллээр [5] Улаанбаатар хотод 1,640,781 хүн амьдарч байгаагаас 715,121 хүн (нийт хүн амын 43.6%) гэр хорооло́лд оршин сууж байна. Гэр хорооллын айл өрхүүдийн эрчим хүчний хэрэглээний горим, найдвартай ажиллагаа, байгаль орчинд үзүүлэх нөлөөлөл нь нийслэлийн тулгамдсан асуудлуудын нэг хэвээр байна.

С. Түвшинжаргал, Д. Сайнбаяр нарын судалгаагаар [5] гэр хорооллын айл өрхүүдийн цахилгаан хэрэглээ жил бүр 3-5%-иар өсөж байгаа бөгөөд энэ нь цахилгаан эрчим хүчний хангамжид шинэ асуудлуудыг бий болгож байна. Мөн Ц. Доржпалам, Г. Лхагважав нарын судалгаа [6] хотын агаарын бохирдлыг бууруулахын тулд гэр хорооллын энергийн хэрэглээний шилжилтийг хийх шаардлагатайг тэмдэглэжээ.

Сүүлийн жилүүдэд нарны цахилгаан үүсгүүрийн технологи хурдацтай хөгжиж, үйлдвэрлэлийн өртөг буурч, үр ашиг нэмэгдсэнээр гэр хорооллын айл өрхүүдэд хүртээмжтэй болох боломж нээгдэж байна [7, 8]. Гэвч Монгол улсын нөхцөлд, ялангуяа Улаанбаатар хотын үүлшилтийн онцлогийг тооцсон нарны теоретик потенциалын судалгаа дутагдалтай байна.

Иймд Улаанбаатар хотын үүлшилтийн төлөв байдлыг судалж, нарны эрчим хүчний тереотик потенциалд үзүүлэх нөлөөллийг тодорхойлохыг зорьсон.

Улаанбаатар хотын үүлшилтийн байдлыг жил, сар, өдрөөр нь судалж, нарны цацрагт үзүүлэх нөлөөллийг тооцоолох, үүлшилт-нарны эрчим хүчний харилцан хамаарлын загварыг боловсруулах.

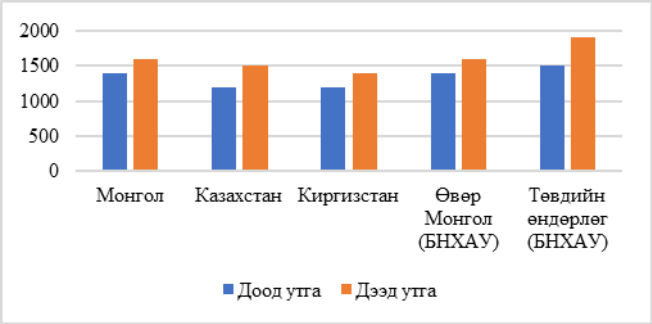
Монголын цаг уурын нөхцөлд нарны эрчим хүчний потенциалыг тооцоолох, үүлшилтийн нөлөөгөөр бодит нарны эрчим хүчний потенциалыг гаргах.

Судалгааны үр дүнгээр Монгол улсын нарны эрчим хүчний бодлого боловсруулалт, гэр хорооллын нарны эрчим хүчний төслүүдийн хэрэгжилт, иргэдийн эрчим хүчний хангамжийн найдвартай байдал, байгаль орчны асуудалд чухал хувь нэмэр оруулах юм.

II. НАРНЫ ЭРЧИМ ХҮЧНИЙ ПОТЕНЦИАЛЫН ХАРЬЦУУЛСАН СУДАЛГАА

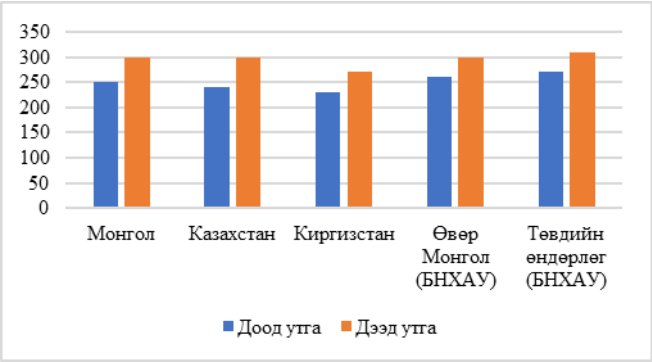
A. Улс орнуудын нарны эрчим хүчний потенциалын харьцуулалт

Монгол Улсын нарны эрчим хүчний потенциалыг олон улсын түвшинд үнэлэхийн тулд төстэй газарзүйн байршил, уур амьсгалтай орнуудтай харьцуулсан судалгааг хийв. Судалгааны үр дүнг зураг 1, 2-т дэлгэрэнгүй харуулав.



Зураг 1. Жилийн дундаж нарны тусгалын ын харьцуулалт[1,8]

Жилийн дундаж нарны тусгалын хувьд Монгол (1400 - 1600 кВтц/м²) Казахстан (1200-1500 кВтц/м²), Киргизстан (1200-1400 кВтц/м²) зэрэг хөрш орнуудаас давуу байгаа нь онцлог юм.



Зураг 2. Жилийн нарлаг өдрийн харьцуулалт [1,8]

Зураг 2-оос харахад Монгол Улс нь нарны эрчим хүчний потенциалын хувьд олон улсын түвшинд өндөр үзүүлэлттэй байна. Ялангуяа жилийн нарлаг өдрийн тоо (250-300) нь харьцуулсан бүх орнуудтай дүйцэхүйц, зарим тохиолдолд илүү байна.

Зураг 3. Нарны эрчим хүчний хүчин чадлын өсөлт

IRENA-гийн 2023 оны тайлангаас [3] харахад бүс нутгийн орнууд дахь нарны эрчим хүчний хүчин чадлын өсөлтийн хувьд Монгол Улс (32%) нь Киргизстан (27%)-аас өндөр боловч Казахстан (41%), Өвөр Монгол (52%), Төвдийн өндөрлөг (48%)-ийн өсөлтөөс доогуур байна. Энэ нь Монгол Улсын нарны эрчим хүчийг хөгжүүлэх боломж, потенциал өндөр хэвээр байгааг харуулж байна.

Дэлхийн нарны атласын мэдээллээр [8] Монгол Улсын нутаг дэвсгэр нь нарны цацрагийн эрчмийн хувьд дэлхийн топ 25% бүс нутагт багтаж байгаа нь онцлог юм.

B. Гол хотуудын нарны эрчим хүчний потенциалын харьцуулалт

Сонгосон улс орнуудаас Улаанбаатар хоттой уур амьсгал төстэй томоохон хотуудыг сонгон нарны эрчим хүчний потенциалыг харьцуулан судлав.

ХҮСНЭГТ 1. ХОТУУДЫН ХАРЬЦУУЛСАН ДҮН ШИНЖИЛГЭЭ [1,4, 9]

Хот	**Жилийн дундаж нарны тусгал (кВтц/м²)**	**Нарлаг өдөр (хоног)**	**Жилийн дундаж үүлшилт**	**Үр ашигт нөлөөлөх онцлог хүчин зүйлс**
Улаанбаатар (Монгол)	1,400 - 1,500	260 - 280	44-50%	Нарлаг хоног олон, өндөрлөг, агаарын чанар хөдөө орон нутагт сайн
Астана (Казахстан)	1,350 - 1,450	260 - 280	50-55%	Нарны тусгал төстэй, бага зэрэг нам газар, өвөл цас их
Денвер (АНУ, Төв)	1,600 - 1,800	280 - 300	35-40%	Тусгал өндөр ч, хүчилтөрөгч багатай орчин, панелийн температур өндөр
Улаанцав (Өвөр Монгол, Хятад)	1,500 - 1,600	270 - 290	40-45%	Нөөц өндөр, үйлдвэржсэн бүс, агаарын бохирдол өндөр
Лхас (Төвдийн өндөрлөг)	1,700 - 1,850	290 - 310	25-30%	Маш өндөр, агаар тунгалаг, температурын хэлбэлзэл их

Хүснэгт 1-ээс харахад Улаанбаатар хот нь нарны тусгалын хувьд Астана (1350-1450 кВтц/м^2) хоттой төстэй боловч Лхас (1700-1850 кВтц/м^2), Денвер (1,600-1,800 кВтц/м^2) хотуудаас доогуур байна.

Хамгийн бага үүлшилттэй газар нь Лхас (25-30%) байгаа бөгөөд энэ нь өндөр уулын бүс, цэвэр агаартай, нарлаг өдрийн тоо их байдагтай холбоотой. Харин Астана хот хамгийн өндөр үүлшилттэй (50-55%) газар болж байна.

Улаанбаатар хотын хувьд жилийн дундаж үүлшилт 45-50% гэж тооцсон нь дундаж түвшинд орж байна. Энэ мэдээлэл нь нарны эрчим хүч ашиглах төлөвлөлтөд чухал үзүүлэлт болох юм.

Судалгаанаас харахад Улаанбаатар хотын нарны эрчим хүчний потенциалд хамгийн их нөлөөлөх хүчин зүйл нь улирлын хэлбэлзэл, үүлшилтийн төлөв байдал юм. Үүнээс үүлшилтийн төлвийг аван тооцоолох аргачлалыг тооцов.

III. Тооцоолох арга аргачлал

Нарны гэрэл үүлнээс нэвтрэн ирэх харьцаа [10]

$$TF = 1 - 0.75 \times C \qquad (1)$$

Жишээ нь: 50% үүл→ $TF = 1 - 0.75 \times 0.5 = 0.625$ болно.

$$H_m = H_{0m} \times TF \qquad (2)$$

Энд:

H_m - тухайн сарын үүлшилтийг тооцсон нарны цацрагийн дундаж хэмжээ (кВт.цаг/м²/өдөр)

H_{0m} - агаар мандлын гадна хүрэх нарны цацраг (атмосфергүй орчинд)

TF - агаар мандлаар нэвтрэх фактор, үүнд үүлшилт голлох нөлөөтэй.

IEEE 1565-2021 стандарт дахь нарны потенциалын үндсэн суурь ойлголт $P_{теоретик} = H_m \times D_m$ байдаг.

Иймд, үүлшилтийн томьёог ашиглан сарын нарны цацрагийн хэмжээг (H_m) тооцоолж, түүнийгээ дараах 3-р томьёонд орлуулж , нарны системийн жилийн гаралтыг ($P_{Теоретик}$) тооцоолж болно.

$$P_{Теоретик} = A \times \eta \times PR \times \sum_{m=1}^{12} H_m \times D_m \qquad (3)$$

D_m - сар дахь өдрийн тоо (Жишээ нь: 1-р сар = 31, 2-р сар = 28 гэх мэт)

$A=20м^2$ - панелийн эсвэл ашиглах талбайн хэмжээ (м²)

$\eta=0.18$ - панелийн үр ашиг (жишээ нь 0.15…0.20)

$PR=0.75$ - системийн бусад алдагдлыг (албум, кабель, инвертор г.м.) харгалзан үзэх коэффициент (0.75…0.85)

E_m – сар бүрийн нэг м²-д хуримтлагдах энерги (кВт·цаг/м²)

$P_{Теоритик}$ – 20 м² талбайд тухайн сард хуримтлагдах энерги (кВт·цаг)

Эхлээд үүлшилтийн мэдээллийг ашиглан нарны цацрагийн бодит хэмжээг тооцоолно:

$$H_m = H_{0m} \times (1 - 0.75 \times C_m) \qquad (4)$$

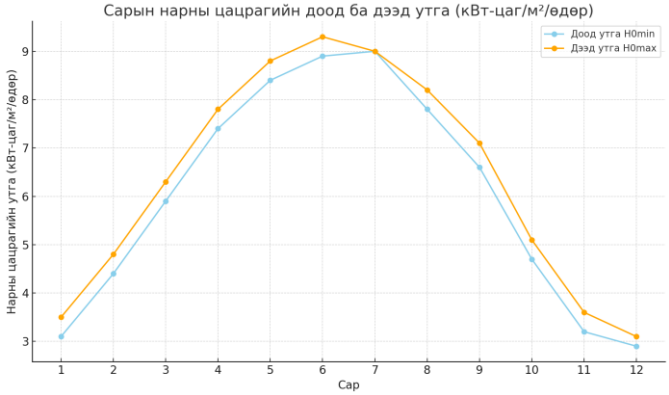
Энд:

- H_m - тухайн сарын үүлшилтийг тооцсон нарны цацрагийн дундаж хэмжээ (кВт.цаг/м²/өдөр)
- H_{0m} - тухайн сарын атмосферийн гадна талд хүрэх нарны цацраг (тухайн газар зүйн байршил, сар, өнцөг зэргээс хамаарна)
- C_m - тухайн сарын үүлшилтийн дундаж хувь (0-1 хүртэл утгатай)

Дараа нь тооцоолсон H_m утгыг томьёо 3-т орлуулж тооцоолно.Үүлшилтийн нөлөөг тооцсон нарны цацрагийн хэмжээг тодорхойлох:

$$P_{Теоретик} = A \times \eta \times PR \times \sum_{m=1}^{12} [H_{0m}(1 - 0.75 \times C_m)] \times Dm \qquad (5)$$

Монгол орны хувьд H_{0m} буюу атмосферийн гадна тал дахь нарны цацрагийн хэмжээ нь газар зүйн байршил, өндөршил, сар болон жилийн улирлаас хамаарч өөр өөр утгатай байдаг. Зураг 4-т Улаанбаатар хотын H_{0m} утгуудыг харуулав.

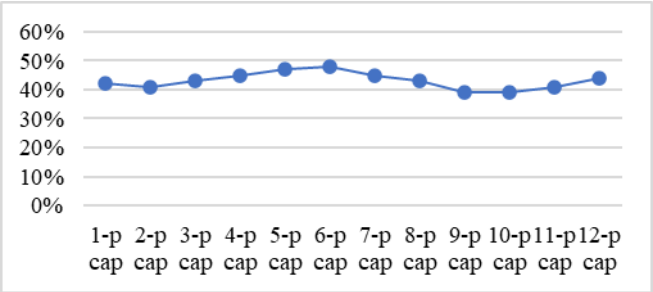


Зураг 4. Улаанбаатар хотын атмосферийн гадна тал дахь нарны цацрагийн сарын дундаж (H_{0m})

IV. Тооцоолсон үр дүн

Улаанбаатар хотын үүлшилтийн дундаж хувийг 7 жилийн (2017-2025.04) өдөр тутмын үүшилтийн дундаж [11] утгуудаар график байгуулав.

Улаанбаатар хотод үүлсээр бүрхэгдсэн тэнгэрийн дундаж хувь нь жилийн туршид улирлын бага зэргийн өөрчлөлттэй байна.



Зураг 4. Сүүлийн 7 жилийн сар сарын дундаж

Улаанбаатар хотод жилийн хамгийн цэлмэг үе наймдугаар сарын 20-ноос эхэлж, 3 сар үргэлжилж, арван нэгдүгээр сарын 28-нд дуусна. Улаанбаатар хотын жилийн хамгийн цэлмэг сар бол 10-р сар бөгөөд энэ

хугацаанд дунджаар тэнгэр цэлмэг, эсвэл багавтар үүлтэй, нийт 61% нь бүрхэг байна. Жилийн үүлэрхэг үе нь 11-р сарын 28-нд эхэлж, 7 -аас 8 сарын хооронд үргэлжилж, 8-р сарын 20-нд дуусна. Улаанбаатар хотод жилийн хамгийн үүлэрхэг сар бол 6-р сар бөгөөд энэ хугацаанд тэнгэр бүрхэг буюу нийт цагийн 48% нь бүрхэг байдаг.

Томёо 2-т сар болгоны үүлшилтийн хэмжээг оруулан тухайн сарын үүлшилтийг тооцсон нарны цацрагийн дундаж хэмжээ (кВт.цаг/м²/өдөр)-г сар болгоноор дараах хүснэгтэнд гаргав.

ХҮСНЭГТ 2. УЛААНБААТАР ХОТЫН НАРНЫ ЦАЦРАГИЙН ДУНДАЖ ХЭМЖЭЭ

Сар	Үүлшилтийн дундаж хувь (%)	TF	H_{0m}	H_m
Нэгдүгээр сар	42%	0.69	3.30	2.26
Хоёрдугаар сар	41%	0.69	4.60	3.19
Гуравдугаар сар	43%	0.68	6.10	4.13
Дөрөвдүгээр сар	45%	0.66	7.60	5.04
Тавдугаар сар	47%	0.65	8.60	5.57
Зургадугаар сар	48%	0.64	9.10	5.82
Долдугаар сар	45%	0.66	8.80	5.83
Наймдугаар сар	43%	0.68	8.00	5.42
Есдүгээр сар	39%	0.71	6.60	4.67
Аравдугаар сар	39%	0.71	4.90	3.47
Арван нэгдүгээр сар	41%	0.69	3.50	2.42
Арван хоёрдугаар сар	44%	0.67	2.90	1.94
Дундаж хувь	43%	0.680	6.167	4.19

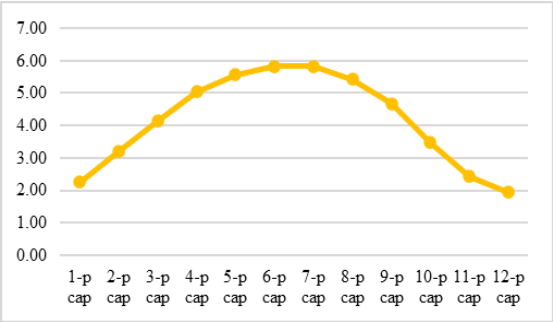
Тооцооллоос харахад жилийн дундаж нарны цацрагийн хэмжээ атмосферийн гадна талд ~6.80 кВт·цаг/м²/өдөр байхад үүлшилтийн нөлөөгөөр ~4.19 кВт·цаг/м²/өдөр болж буурсан байна.

Үүлшилтийн нөлөөгөөр буурах хувь жилийн дундажаар 38.38% байна.

Сар бүрийн H_m утгууд нь нарны эрчим хүчний системийн тооцоололд ашиглагдаж, жилийн цахилгаан эрчим хүчний гаралтыг тооцоолоход хэрэглэгдэнэ.

Жилийн дундаж үүлшилт (0.43) болон нарны нэвтрүүлэх фактор (0.680) нь нарны эрчим хүчний системийн ерөнхий үр ашгийг тодорхойлдог чухал үзүүлэлт болно.

Дээрх тооцоолол нь Улаанбаатар хотын хувьд нарны эрчим хүчний системийн онолын тооцоолол хийхэд ашиглагдах H_m утгуудыг гаргав.



Зураг 5. Сар болгоны нарны эрчим хүчний хэмжээ

Жилийн илүү эрчимтэй үе нь 4-р сараас 8-р сарын 9-ний хооронд 3-4 сар үргэлжилдэг.

ХҮСНЭГТ 3. УЛААНБААТАР ХОТЫН САР БҮРИЙН НАРНЫ ЭРЧИМ ХҮЧНИЙ ПОТЕНЦИАЛЫН ТООЦООЛОЛ

Сар	H_m (кВт·цаг/м²·өдөр)	D_m (өдөр)	PТереотик
1-р сар	2.26	31	189.16
2-р сар	3.19	29	249.78
3-р сар	4.13	31	345.68
4-р сар	5.04	30	408.24
5-р сар	5.57	31	466.21
6-р сар	5.82	30	471.42
7-р сар	5.83	31	487.97
8-р сар	5.42	31	453.65
9-р сар	4.67	30	378.27
10-р сар	3.47	31	290.44
11-р сар	2.42	30	196.02
12-р сар	1.94	31	162.38
Нийт	49.76	2024 он	4099.22

Жилийн нийлбэр: 4099.22 кВт·цаг

20м² талбайтай нарны панел суурилуулсан тохиолдолд жилийн турш нийт 4099.22 кВт·цаг эрчим хүч үйлдвэрлэх потенциалтай байна.

Хүснэгт 5-д Монгол орны сар бүрийн нарны эрчим хүчний теоретик потенциал болон бодитоор ашиглах боломжит хэмжээг харьцуулж байна. Теоретик потенциал (P_m) нь цаг агаарын нөхцөл бүрэн тааламжтай байх үеийн потенциал бол бодит потенциал нь ашиглалтын коэффициент 0.135-аар утга авсан. Зуны саруудад (5-8 сар) дундаж эрчим хүчний потенциал хамгийн өндөр байгаа бөгөөд 7-р сард 487.97 кВт·цаг болж байна.

V. ДҮГНЭЛТ

Улаанбаатар хотын үүлшилтийн байдал нь нарны эрчим хүчний ашиглалтад тодорхой хэмжээнд сөрөг нөлөө үзүүлж байгааг судалгааны үр дүн харуулж

байна. 2017–2025 оны хооронд хийсэн судалгаагаар хотын дундаж үүлшилт 43% байгаа нь нарны цацрагийн ойролцоогоор 32%-ийг хаадаг болохыг тогтоосон. Монгол Улс нарны эрчим хүчний теоретик потенциалаараа дэлхийд дээгүүрт ордог ч, Улаанбаатар хотын хувьд үүлшилтийн нөлөөгөөр онолын потенциалын зөвхөн ~18% нь бодитоор ашиглагдаж болохыг тооцоолсон. Зуны саруудад (5–8 сар) үүлшилт хамгийн өндөр буюу 43–48% байдаг бол намрын эхэн саруудад (9–10 сар) хамгийн бага буюу 39% байна.

20 м² талбай бүхий нарны панелын жилийн теоретик үйлдвэрлэл 4099.22 кВт·цаг боловч, үүлшилт болон төхөөрөмжийн үр ашгийг тооцсон бодит гарах боломжтой эрчим хүч ойролцоогоор 737.86 кВт·цаг болохоор тооцоологдсон нь анхаарал татахуйц үзүүлэлт юм.

Энэхүү судалгааны үр дүн нь нарны эрчим хүчний салбар дахь бодлого төлөвлөлт, гэр хорооллын бүсэд хэрэгжих сэргээгдэх эрчим хүчний төслүүд, иргэдийн эрчим хүчний хүртээмж, байгаль орчны тогтвортой байдалд чухал ач холбогдолтой юм.

Талархал

Бидний судалгааны үр дүнг эрдэм шинжилгээний орчинд толилуулах үнэт боломжийг бүрдүүлсэн "Монголын мэдээллийн технологи 2025" эрдэм шинжилгээний хурлын зохион байгуулах хороо болон удирдах зөвлөлд чин сэтгэлийн талархал илэрхийлье.

Мөн уг хурлын эрдэм шинжилгээ, шүүмжийн түвшинг баталгаажуулж, шинжлэх ухааны манлайллаараа мэргэжлийн чиг баримжаа олгосон эрхэм хүндэт эрдэмтэн багш нарт гүн талархал дэвшүүлж байна.

Ном зүй / Reference

[1] International Renewable Energy Agency (IRENA), "Renewable Energy Statistics 2023", Abu Dhabi, 2023.
[2] International Energy Agency (IEA), "Renewables 2023 - Analysis and Forecasts to 2028", Paris, 2023.
[3] Д. Намхайням, Г. Содномдорж, "Монгол орны нарны эрчим хүчний потенциал, ашиглалт", Шинжлэх ухаан, технологийн сэтгүүл, №25, х. 45-58, 2021.
[4] Б. Батзориг, М. Лхагвасүрэн, "Улаанбаатар хотын үүлшилтийн динамик ба нарны цацрагийн хэмжилт", Монголын Шинжлэх Ухааны Академийн бүтээл, №42, х. 112-125, 2022.
[5] С. Түвшинжаргал, Д. Сайнбаяр, "Гэр хорооллын цахилгаан хэрэглээний судалгаа", Эрчим хүчний эдийн засаг, №8, х. 76-89, 2023.
[6] Ц. Доржпалам, Г. Лхагважав, "Улаанбаатар хотын агаарын бохирдол ба сэргээгдэх эрчим хүчний хэрэглээний хамаарал", Экологи ба тогтвортой хөгжил, №15, х. 211-228, 2022. *В. Бага хурлын илтгэлүүд*
[7] Б. Билгүүн, Ж. Энхсайхан, "Нарны эрчим хүчний системийн загварчлал – Улаанбаатар хотын жишээн дээр", Тогтвортой хөгжлийн эрчим хүчний бага хурал, Улаанбаатар хот, Монгол Улс, х. 245-250, 2022.
[8] Г. Батгэрэл, Д. Даваасүрэн, "Нарны эрчим хүчний системийн Монголд тохиромжтой бүтэц, төрлүүд", Инженерчлэлийн технологи, №31, х. 105-119, 2022.
[9] Дэлхийн нарны атлас, "Монгол улсын нарны эрчим хүчний зураглал", www.solaratlas.org, хандсан огноо: 2024.03.20.
[10] Ангерстром, А. (1924). "Solar and terrestrial radiation." Quarterly Journal of the Royal Meteorological Society, 50(210), pp. 121-126.
[11] Weatherspark, "Average Weather in Ulan Bator, Mongolia: Year-Round", https://weatherspark.com/y/117604/Average-Weather-in-Ulan-Bator-Mongolia-Year-Round#Sections-Clouds, хандсан огноо: 2025.04.01.

Хиймэл оюунд ашиглах сүлжээний халдлагын сан үүсгэх

1st П.Баясгалан
МУИС, МТЭСургууль Улаанбаатар, Монгол улс
bayasgalan.p@num.edu.mn

2nd Г.Манлайбаатар
Хил хамгаалах ерөнхий газар Улаанбаатар, Монгол улс

5rd Д.Баттулга
МУИС, МТЭСургууль Улаанбаатар, Монгол улс

4rd Н.Угтахбаяр*
МУИС, МТЭСургууль Улаанбаатар, Монгол улс
ugtakhbayar@num.edu.mn

5nd Б.Өсөхбаяр
МУИС, МТЭСургууль Улаанбаатар, Монгол улс

Хураангуй: Кибер аюулгүй байдлыг хангахад сүлжээний халдлага илрүүлэх системүүдийн үр нөлөөг сайжруулах нь чухал асуудал хэвээр байсаар байна. Энэ чиглэлийн судлаачдын ашиглаж буй KDD99, NSL-KDD99 гэх мэт сангуудад агуулагдаж буй халдлагын мэдээлэл нь харьцангуй хуучирсан, шинэ төрлийн халдлагын мэдээлэл агуулаагүй зэрэг нь халдлага илрүүлэх системийн халдлага илрүүлэх үр дүнг сайжруулах боломжийг тодорхой түвшинд хязгаарлаж байгаа юм. Энэхүү судалгааны ажлаар машин сургалт, хиймэл оюун ашигласан сүлжээний халдлага илрүүлэх системийг хөгжүүлэхэд шаардлагатай шинэ төрлийн халдлагын мэдээлэл агуулсан бодит мэдээллийн санг боловсруулахад чиглэгдсэн. Судалгааны ажлын хүрээнд боловсруулсан топологи бүхий сүлжээний орчинд халдлага гүйцэтгэн сүлжээний урсгалыг цуглуулж тус санг бүтээсэн. Бид энэ удаагийн судалгааны ажлаараа вэбийн халдлага илрүүлэх сан үүсгэхээр зорин ажилласан ба нийт 2545563 мөр, 19 онцлог бүхий сан үүсгэж машин сургалтын алгоритмуудыг хялбараар ашиглах боломжтой байхаар тооцож arff файл хэлбэрээр үүсгэсэн. Уг санг ашиглан судлаачид судалгааны үр дүнгээ баталгаажуулах, шинэ халдлага бүхий сан ашиглаж үр дүнгээ сайжруулах боломжийг бүрдүүлж байна.

Түлхүүр үг: Кибер аюулгүй байдал, сүлжээний халдлага илрүүлэх систем, KDD99, машин сургалт, хиймэл оюун

I. Удиртгал

Кибер аюулгүй байдал нь орчин үеийн сүлжээний системүүдийн хамгийн чухал хэсэг юм. Сүлжээний халдлагын давтамж болон эрсдэл нь жилээс жилд өсөн нэмэгдэж байгаа нь сүлжээний халдлага илрүүлэх системүүдийг (IDS) оновчтой хөгжүүлэх шаардлагыг нэмэгдүүлж байна. Эдгээр системүүд нь сүлжээний аюулгүй байдлыг хангах, сүлжээний халдлагыг шуурхай илрүүлэх, хариу арга хэмжээ авахад чухал үүрэг гүйцэтгэдэг [9]. Гэсэн хэдий ч, сүлжээний халдлага илрүүлэх үйл явц нь технологийн хурд, халдлагыг хиймэл оюун ашиглан гүйцэтгэх болсон зэрэг олон шалтгаанаас үүдэн төдийлөн амжилттай байхгүй байна. Энэ нь нэг талаасаа шинэ төрлийн халдлагын илрүүлэлт, сүлжээний урсгалын анализ хийхэд шаардлагатай өгөгдлийн хомсдолтой байдал, ашиглаж буй сангуудын хуучирсан байх зэрэгтэй шууд хамааралтай байна.

Халдлага илрүүлэх системүүд нь signature-д суурилсан болон гажигт суурилсан гэж хоёр ангилж болно. Сигнатурт суурилсан IDS нь танил аюулуудыг илрүүлэхэд үр дүнтэй боловч шинэ төрлийн халдлагуудыг илрүүлэхэд амжилтгүй байдаг. Харин гажигт суурилсан IDS нь шинэ төрлийн халдлагуудыг илрүүлэх боломжтой ч, энэ нь өндөр хуурамч сөрөг илрүүлэлттэй байдаг. Гэвч шинэ төрлийн халдлагууд эдгээр системийг хууран мэхлэх боломжтой байдаг. Иймд уламжлалт халдлагын мэдээллийн сангууд, тухайлбал KDD99 болон NSL-KDD99, нь олон төрлийн шинэ төрлийн халдлагыг илрүүлэхэд дутагдалтай байдаг. Мөн эдгээр сангуудад байгаа өгөгдлүүд нь ихэвчлэн хуучирсан бөгөөд бодит цагийн сүлжээний халдлагуудыг илрүүлэхэд тусгайлан тохирсон мэдээллийн дутагдалтай байдаг. [10]

Бид уг судалгааны ажлын хүрээнд халдлага илрүүлэх системийг үнэлэх (benchmarking) боломжтой систем ашиглан халдлагын мэдээллийг цуглуулах замаар сүүлийн үеийн халдлагын мэдээлэл агуулсан сан үүсгэхийг зорин ажиллав. Уг санг үүсгэснээр халдлага илрүүлэх системийн танилт, үр дүн сайжрах, судлаачид шинэ төрлийн судалгаа хийх боломжийг бүрдүүлэх зэрэг олон давуу тал бий болно гэж үзэж байна. Уг санг үүсгэхдээ халдлагын мэдээлэлд бусад мэдээлэл орохгүй, үр дүнд нөлөөлөхгүй байх зорилгоор холболтын мэдээллийг агуулсан байх, үндсэн сүлжээнээс бүрэн

тусгаарлагдсан байх, интернэт орчны урсгал дамжихгүй байх зэрэг нөхцөлүүдийг тооцон ажилласан.

II. Судлагдсан байдал

Сүүлийн жилүүдэд сигнатурт суурилсан халдлага илрүүлэх системтэй хослуулан гажигт суурилсан халдлага илрүүлэх системийг түгээмэл ашиглаж байгаагаас гадна машин сургалт, хиймэл оюуны хэрэгсэл ашигласан системүүд цөөнгүй бий болсон. Үүнтэй уялдаатайгаар машин сургалт, хиймэл оюунд ашиглах боломжтой сангуудыг шинэчлэх, сайжруулах шаардлага бий болсноос гадна шинэ төрлийн халдлага бүхий санг үүсгэх нь энэ чиглэлийн судалгааны ажилд чухал үүрэг гүйцэтгэх юм. Халдлага илрүүлэх системийн сан үүсгэх ажил цөөнгүй хийгдэж байгаагаа хэдий ч сүүлийн үеийн судалгааны эх сурвалжуудад ашиглаж буй сангийн дийлэнх нь нийтэд нээлттэй биш эсвэл хуучирсан байна. Н.Угтахбаяр нарын судлаачид [11] ажлаар NUM15 сан үүсгэж халдлага илрүүлэх системийн үр дүнг сайжруулсан хэдий ч тус сан нь нийтэд нээлттэй биш байна. Нийтэд нээлттэй сангийн нэг болох RoEduNet-SIMARGL2021 нь бодит академик сүлжээнээс цуглуулсан өгөгдлөөс үүссэн сан бөгөөд халдлага болон хэвийн урсгалын мэдээллийг цуглуулан 44 онцлог шинж бүхий санг үүсгэсэн байна. Уг санг үүсгэхдээ судлаачид өгөгдлийн тэнцвэргүй байдал (imbalanced) дээр анхаарлаа хандуулжээ [1]. Энгийн сүлжээний урсгалаас гадна IOT, роботын сүлжээнд ашиглагдах сангууд үүсгэсэн судалгааг судлаачид мөн хийж байгаа ба үүний нэг нь ROSPaCe сан ба Robot operating system 2 (ROS2)-д суурилсан системээс цуглуулсан бодит урсгалын мэдээлэл ашигласан бөгөөд гол онцлог нь халдлага болон энгийн урсгалын мэдээллийг цаг хугацааны дараалалтайгаар бүртгэсэнд оршиж байна. Судлаачид уг ажлаа халдлага илрүүлэх систем халдлагыг хурдан хугацаанд илрүүлэх боломжтой байх гэдэгт үндэслэн үүсгэжээ cite2.

Судалгааны ажлуудын дийлэнх нь KDD99, NSL-KDD дээр хийгдсэн байна. [3] ажилд судлаачид convolutional neural networks (CNN) алгоритм ашиглан халдлагыг илрүүлэх загвар боловсруулсан ба үр дүн сайн гарчээ. Үүнээс гадна [4] ажилд судлаачид XGBoost алгоритм ашигласан судалгааны үр дүнд халдлагыг 99.3 хувийн нарийвчлалтай таньсан үр дүнг гаргажээ. Дээрхээс гадна судлаачид синтетик сан үүсгэх ажлыг гүйцэтгэж байгаа ба Салех болон бусад судлаачид [5] ажилдаа зохиомол өгөгдөл үүсгэх CTGAN (Conditional Tabular Generative Adversarial Network)-ийг ашиглан халдлага илрүүлэх системийн сургалтад зориулагдсан синтетик сан үүсгэсэн нь бодит өгөгдөлтэй шинж чанарын хувьд ижил гарсан байна. Пилип болон бусад судлаачид [6] өгүүллийн хүрээнд халдлага илрүүлэх системийн судалгаанд ашиглагдаж буй 48 сангийн шинж чанар, давуу болон дутагдалтай талуудыг тодорхойлох зорилготой ажлыг гүйцэтгэсэн. Мөн Аркан болон бусад судлаачид [7] ажлын хүрээнд NetFlow ашиглан онцлог шинжийг баяжуулснаар халдлага илрүүлэх системийн үр дүнг сайжруулсан байна.

III. Судалгааны аргачлал

Судалгааны хүрээнд бид халдлагын мэдээллээс бусад мэдээлэл дамжих боломжгүй байхаар сүлжээний орчинг зохион байгуулсан. Ийм туршилтын орчинг бэлдэхдээ "ARP", "IPv6"гэх мэт сүлжээгээр тогтмол хугацаанд дамжуулал хийдэг протоколуудыг хаасан. Бэлэн болсон сүлжээний орчиндоо Pytbull системийг ашиглан шийн төрлийн халдлага хийж, үүнийг сүлжээний халдлага илрүүлэх систем болох Zeek суулгасан төхөөрөмж дээгүүр дамжиж зорилтот төхөөрөмж дээр очно. Zeek халдлаг илрүүлэх систем нь болж буй бүх процесийг лог болгон хадгалдаг тул халдлагын мэдээлэл агуулсан сан нь нарийвчлалтай, найдвартай байдлыг тодорхойлох зорилгоор zeek халдлага илүүлэх системийг ашигласан. Цуглуулсан сан хэр оновчтой болохыг тодорхойлоход дөхөм байх зорилгоор Weka программын тусламжтайгаар arff файл хэлбэрээр үүсгэсэн.

A. Pytbull

Pytbull нь халдлага илрүүлэх, эсэргүүцэх систем (IDS/IPS)-ийг шалгах боломжтой нээлтэй эхийн фрэймворк юм. Энэ нь 11 модулиас бүрдэх бөгөөд 278 төрлийн халдлага хийх замаар системд шалгалт хийх боломжтой, 11 модуль нь тус бүрдээ өөр төрлийн халдлага хийдэг. Pytbull ашиглан туршилтын сүлжээний орчинд хортой урсгал үүсгэхэд ашигласан. Үүнийг ашиглахад

B. Zeek

Zeek халдлага илрүүлэх систем нь сүлжээний аюулгүй байдлын хяналт, халдлага илрүүлэх, болон халдлагын мэдээлэлд анализ хийх зориулалттайгаас гадна өргөтгөх боломжтой хэрэгсэл юм. Бид энэхүү судалгааны ажлын хүрээнд Pytbull-аар гүйцэтгэсэн халдлагыг Zeek илрүүлэх эсэхийг тодорхойлохоос гадна цуглуулж буй урсгалын найдвартай байдлыг шалгахыг зорьсон. Үүнээс гадна урсгалын мэдээллийг Zeek ажиллаж буй машин дээр цуглуулсан нь өгөгдлийг баталгаажуулахад хугацааны алдагдалгүй, зөрүүгүй, баталгаажуулах боломжтой болгосон. Zeek-ийг нэмэлт хөгжүүлэлт хийх замаар олон хувилбараар ашиглах боломжтой бөгөөд анализ хийх дүрэм, процессуудыг автоматжуулах боломжтой. [8]

C. Weka

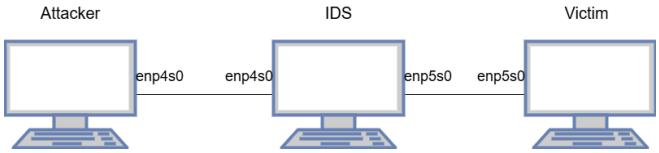
Weka түүл нь "Waikato Environment for Knowledge Analysis"гэсэн үгний товчлол бөгөөд өгөгдлийн олборлолт болон машин сургалтын нэг түүл бөгөөд төрөл бүрийн алгоритмууд, өгөгдлийн урьдчилан боловсруулалт хийх хэрэгслүүдийг багтаасан иж бүрэн программ хангам юм. WEKA нь машин сургалтын олон төрлийн алгоритмуудыг агуулдаг. Мөн өгөгдлийг урьдчилан

Хүснэгт I
Төхөөрөмжийн техник үзүүлэлт

№	Төхөөрөмжийн нэр	CPU	RAM	HARD	Үйлдлийн систем	Interface
1	Mini computer (attacker)	Intel Atom(TM) C3558R@2,40Ghz	16Gb DDr4	256GB ssd	Ubuntu 22.04.5-Server	5ш 1gb lan, 4ш 10gb fiber
2	Mini computer (zeek)	Intel Atom(TM) C3558R@2,40Ghz	16Gb DDr4	256GB ssd	Ubuntu 22.04.5-Server	5ш 1gb lan, 4ш 10gb fiber
3	Mini computer (Victim)	Intel Atom(TM) C3558R@2,40Ghz	16Gb DDr4	256GB ssd	Ubuntu 22.04.5-Server	5ш 1gb lan, 4ш 10gb fiber

боловсруулалт хийх хэрэгслүүд, жишээлбэл дискретизаци, сонголт хийх гэх мэт олон төрлийн хэрэгслүүдтэй, суралцсан загвар болон түүний гүйцэтгэлийг визуалжуулж, үр дүнг үнэлэх боломжийг олгодог.

Судалгааны ажлын хүрээнд ашигласан топологид CPU: C3558R, RAM: 16Gb, Network interface: 1Gbps үзүүлэлт бүхий 3 эмбеддэд төхөөрөмж ашигласан бөгөөд халдлагын явцад төхөөрөмжийн нөөц (resource) хүрэлцэж буй эсэхийг хянаж өгөгдлийн алдагдал гарахгүй байхад анхааран ажилласан.



Зураг 1. Туршилтийн сүлжээний топологи

Цуглуулсан өгөгдлөөс халдлага илрүүлэх системд илүү нөлөө бүхий онцлог шинжийг тодорхойлох, стандарт алгоритмууд бидний үүсгэсэн санг ашиглан халдлага илрүүлж буй хувийг тооцох зэрэг ажлыг Weka программ ашиглан гүйцэтгэх боломжтой хувилбараар үүсгэсэн бөгөөд одоогийн байдлаар холболтын мэдээлэл, вэбийн мэдээлэл, файлын мэдээлэл тус бүрд нь уг туршилтыг хийх боломжтой байхаар үүсгэсэн.

Үр дүн

Уг судалгааны ажлаар Зураг1-т харуулсан бодит сүлжээний орчин үүсгэж халдлага хийж халдлагын болон холболтын мэдээллийг цуглуулан 2 сая гаруй мөр бүхий сан үүсгэж судлаачид, аюулгүй байдлын мэргэжилтнүүд ашиглах боломжтой байхаар нийтэд нээлттэй байршуулав [1]. Бидний туршилтын үрд дүнд бий болсон халдлагын сан нь үндсэн гурван хэсгээс бүрдэх бөгөөд тэдгээр нь холболтын мэдээлэл, вэб урсгалын мэдээлэл, файлын мэдээлэл байна. Уг мэдээллийг хамтад нь ашиглаж халдлага илрүүлснээр халдлагыг илүү нарийвчлалтайгаар илрүүлэх боломжтой гэсэн дүгнэлтэд хүрсэн. Сангийн хувьд:

- Холболтын мэдээлэл - 287893kb хэмжээтэй, 2368003 мөр
- Вэбийн мэдээлэл - 21677kb хэмжээтэй, 101750 мөр
- Файлын мэдээлэл - 18032KB хэмжээтэй, 75010 мөр

[1] https://github.com/Bayasgalan02/Data-set

зэрэг мэдээллийг агуулсан. Уг сангийн хувьд бүх мэдээлэл нь халдлагын мэдээлэл бөгөөд үүнийг энгийн урсгалаар баяжуулан supervised алгоритмд ашиглах, одоо байгаа санд cluster, CNN зэрэг аргачлалууд ашиглан гажигт суурилсан халдлага илрүүлэх систем үүсгэх, судлаачид өөрсдийн боловсруулсан гажигт суурилсан халдлага илрүүлэх системийн үр дүнг баталгаажуулах зэрэг олон төрлөөр ашиглах боломжтой.

IV. Дүгнэлт

Энэхүү судалгааны ажлын хүрээнд халдлага илрүүлэх системийн судалгаа хийдэг судлаачид, аюулгүй байдлын мэргэжилтнүүд ашиглах боломжтой нээлттэй сан үүсгэх ажлыг эхлүүлж вэбд суурилсан халдлагын санг үүсгэлээ. Уг ажлыг гүйцэтгэхдээ Pytbull, Zeek, Weka зэрэг хэрэгслүүдийг ашиглав. Энэхүү судалгааны ажлыг үргэлжлүүлэн вэбээс бусад төрлийн халдлага илрүүлэх боломжтой сан үүсгэнэ. Уг ажил нь зөвхөн халдлагын мэдээллээс гадна холболтын мэдээллийг агуулж буйгаараа бусад сангаас илүү онцлогтой болсон. Мөн уг ажлын хүрээнд холболтын болон файлын мэдээллийг бусад сангуудтай нэгтгэн халдлага илрүүлэх нэг сан үүсгэх ажлыг эхлүүлээд байна.

Ашигласан ном

[1] Mihailescu ME, Mihai D, Carabas M, Komisarek M, Pawlicki M, Hołubowicz W, Kozik R. The Proposition and Evaluation of the RoEduNet-SIMARGL2021 Network Intrusion Detection Dataset. Sensors (Basel). 2021 Jun 24;21(13):4319. doi: 10.3390/s21134319. PMID: 34202616; PMCID: PMC8272217.

[2] Puccetti T, Nardi S, Cinquilli C, Zoppi T, Ceccarelli A. ROSPaCe: Intrusion Detection Dataset for a ROS2-Based Cyber-Physical System and IoT Networks. Sci Data. 2024 May 10;11(1):481. doi: 10.1038/s41597-024-03311-2. PMID: 38729994; PMCID: PMC11087584.

[3] Yalei Ding and Yuqing Zhai. 2018. Intrusion Detection System for NSL-KDD Dataset Using Convolutional Neural Networks. In Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence (CSAI '18). Association for Computing Machinery, New York, NY, USA, 81–85. https://doi.org/10.1145/3297156.3297230

[4] Amit Singh, Jay Prakash, Gaurav Kumar, Praphula Kumar Jain, and Loknath Sai Ambati. 2024. Intrusion Detection System: A Comparative Study of Machine Learning-Based IDS. J. Database Manage. 35, 1 (Jan 2024), 1–25. https://doi.org/10.4018/JDM.338276

[5] Alabdulwahab, S.; Kim, Y.-T.; Seo, A.; Son, Y. Generating Synthetic Dataset for ML-Based IDS Using CTGAN and Feature Selection to Protect Smart IoT Environments. Appl. Sci. 2023, 13, 10951. https://doi.org/10.3390/app131910951

[6] Bönninghausen, P., Uetz, R. and Henze, M., 2024, August. Introducing a comprehensive, continuous, and collaborative survey of intrusion detection datasets. In Proceedings of the 17th Cyber Security Experimentation and Test Workshop (pp. 34-40).

[7] Aouatif Arqane, Omar Boutkhoum, Hicham Boukhriss, and Abdelmajid El Moutaouakkil. 2021. A Review of Intrusion Detection Systems: Datasets and machine learning methods. In Proceedings of the 4th International Conference on Networking, Information Systems amp; Security NISS 21. Association for Computing Machinery, New York, NY, USA, Article 7, 1–6. https://doi.org/10.1145/3454127.3456576

[8] https://docs.zeek.org/en/master

[9] AbdulRaheem, M., Oladipo, I.D., Imoize, A.L. et al. Machine learning assisted snort and zeek in detecting DDoS attacks in software-defined networking. Int. j. inf. tecnol. 16, 1627–1643 (2024). https://doi.org/10.1007/s41870-023-01469-3

[10] D. Sudyana et al., "Improving Generalization of ML-Based IDS With Lifecycle-Based Dataset, Auto-Learning Features, and Deep Learning,"in IEEE Transactions on Machine Learning in Communications and Networking, vol. 2, pp. 645-662, 2024

[11] Ugtakhbayar, N., Usukhbayar, B., Baigaltugs, S. (2020). A Hybrid Model for Anomaly-Based Intrusion Detection System. In: Pan, JS., Li, J., Tsai, PW., Jain, L. (eds) Advances in Intelligent Information Hiding and Multimedia Signal Processing. Smart Innovation, Systems and Technologies, vol 157. Springer, Singapore. https://doi.org/10.1007/978-981-13-9710-3_44

Монгол Улс дахь Нээлттэй Эхийн Сангийн Аюулгүй Байдал, Хэрэглээ

1st Индра Өлзийбадрах
Электроник холбооны, инженерчлэлийн тэнхим
МТЭС, МУИС
Улаанбаатар, Монгол Улс
indra.ulziibadrakh@gmail.com

2nd Ганбаяр Ууганбаяр*
Электроник холбооны, инженерчлэлийн тэнхим
МТЭС, МУИС
Улаанбаатар, Монгол Улс
ganbayar@num.edu.mn

3rd Угтахбаяр Найдансүрэн
Электроник холбооны, инженерчлэлийн тэнхим
МТЭС, МУИС
Улаанбаатар, Монгол Улс
ugtakhbayar@num.edu.mn

Хураангуй - Нээлттэй эхийн сан нь өнөөгийн программ хангамж хөгжүүлэлтийн салшгүй бүрэлдэхүүн хэсэг бөгөөд түүний аюулгүй байдлын эрсдэл, эрх зүйн орчин нь олон улсад анхаарал татсан асуудал болоод байна. Энэхүү судалгаагаар Монгол Улсад нээлттэй эхийн сангийн хэрэглээ, аюулгүй байдлын талаарх туршлага, программ хангамжийн аюулгүй хөгжүүлэлтийн хэрэгжилт, тулгамдаж буй асуудлыг тодорхойлохыг зорив. Уг судалгааны ажлын хүрээнд тоон болон чанарын судалгааны аргачлалыг ашиглан хөгжүүлэгч, удирдлагын түвшинд авч үзэв. Бид энэхүү судалгааныхаа үр дүнд тулгуурлан нээлттэй эхийн сангийн аюулгүй байдлын талаарх мэдлэгийг нэмэгдүүлэх, хууль эрх зүйн орчныг боловсронгуй болгох, аюулгүй хөгжүүлэлтийн соёлыг төлөвшүүлэх шаардлагатай гэсэн дүгнэлтэд хүрсэн.

Түлхүүр үгс - Программ хангамж хөгжүүлэлт, Нээлттэй эхийн сан, Кибер аюулгүй байдал

I. Удиртгал

Программ хангамжийн хөгжилд нээлттэй эхийн сангуудын үүрэг сүүлийн хэдэн арван жилд эрс өссөн. Орчин үеийн нийтлэг программ хангамжийн 80-90% хүртэлх эх код нь нээлттэй эх дээр суурилсан сан ашиглаж буйг судлаачид онцолдог [1]. Нээлттэй эхийн сангууд өргөн хэрэглэгддэг болсон нь хөгжүүлэлтийн зардлыг бууруулж, хөгжүүлэлтийн хурдыг нэмэгдүүлсэн давуу талтай ч нөгөө талаасаа аюулгүй байдлын эрсдэлийг бий болгодог. Хөгжүүлэгчид нээлттэй эхийн сан нь олон талын оролцоонд суурилан гардаг учир аюулгүй гэж үздэг. Гэсэн хэдий ч бодит байдал дээр нээлттэй эхийн сан дээр суурилан бүтээгдсэн программууд дунджаар 7 эмзэг байдалтай байдаг ба үүнд Remote Code Execution (RCE), Cross-Site Scripting (XSS), SQL Injection, Insecure Deserialization, Broken Authentication, Sensitive Data Exposure, Using Components with Known Vulnerabilities зэрэг орно [2]. Эдгээр программуудын 44% нь ноцтой эмзэг байдал агуулдаг нь томоохон кибер халдлага үүсгэх эрсдэлтэй байдаг [2].

Нээлттэй эхийн сангийн аюулгүй байдлын асуудал 2000 оноос эхлэн анхаарал татаж эхэлсэн. Веб аппликэйшний аюулгүй байдлын “OWASP Top 10” жагсаалтад 2013 оноос “Using Known Vulnerable Components” A9 гэсэн ангилал орсон нь энэ чиглэлийн эрсдэлийг олон улсад хүлээн зөвшөөрснийг илтгэнэ [3]. Тухайлбал, 2014 онд илэрсэн Heartbleed хэмээх OpenSSL криптографийн сангийн ноцтой эмзэг байдал нь интернэтийн орчинд өргөн хүрээтэй нөлөөлж, нээлттэй эхийн сангуудын аюулгүй байдлын сул тал ямар их хор хөнөөл учруулж болохыг тод харуулсан [4]. “Heartbleed” халдлагын үр дүнд байгууллагын нууц мэдээлэл алдагдан, яаралтай засвар хийгдсэн хэдий ч олон систем удаан хугацааны турш хамгаалалтгүй хэвээр үлдсэн [4]. Энэхүү эмзэг байдал нь нээлттэй эхийн чухал сангуудыг тогтмол хянаж, шинэчлэх зайлшгүй шаардлага байгааг тодотгон харуулж, улмаар 2014 онд Linux Foundation-ийг “Core Infrastructure Initiative” хэмээх санхүүжилтийн хөтөлбөрийг эхлүүлэхэд хүргэсэн юм. Үүний дараа 2017 онд Equifax байгууллагын 143 сая хэрэглэгчийн мэдээлэл алдагдсан халдлага гарсан бөгөөд уг халдлагын үндэс нь Apache Struts хэмээх нээлттэй эхийн веб фреймвөүкийн (framework) цаг тухайд нь сайжруулалт хийгдээгүй эмзэг байдал (CVE-2017-5638) байсан нь илэрсэн [5].

Дээрх жишээнүүд нь нээлттэй эхийн сангийн аюулгүй байдлыг хангахгүй орхих нь хор уршигтайг ойлгуулсан тул “аюулгүй программ хөгжүүлэлт” хэмээх ойлголтод тусгайлан гуравдагч талын сангуудын эмзэг байдлыг удирдах үйл явцыг багтаах шаардлагатай болсон. Өөрөөр хэлбэл, программ хөгжүүлэлтийн амьдралын циклд (SDLC) гаднын сангийн эмзэг байдлыг илрүүлэх, засварлах, шинэчлэх алхмуудыг урьдчилан төлөвлөх нь зайлшгүй хэсэг болжээ [6]. 2010 оны дунд үеэс хакерууд ч энэ сул талыг идэвхтэй ашиглаж эхэлсэн. Тухайлбал, өргөн хэрэглээтэй нээлттэй эхийн сангуудын эх кодод зориуд хортой код оруулж, бусад оролцогч талууд татаж авч, ашиглахыг хүлээх тактик дэлгэрсэн бөгөөд үүний нэг жишээ нь 2018 онд npm сангийн event-stream багцад гарсан халдлага юм. Ийнхүү нээлттэй эхийн программ хангамжийн экосистемд гарч буй ноцтой халдлагууд, үүнээс үүдэн халдлагад өртөх эрсдэл нэмэгдэж байгаа нь тус асуудлыг академик судалгаа болон программ хангамж хөгжүүлэлтийн практикт нэн чухал ач холбогдолтойгоор судлах үндэслэл болж байна.

Монгол Улсын хувьд сүүлийн жилүүдэд кибер аюулгүй байдлыг хангах эрх зүйн орчныг бүрдүүлэхээр хэд хэдэн алхам хийсэн. Тухайлбал, 2021 онд УИХ-аас “Кибер аюулгүй байдлын тухай хууль”-ийг баталж, кибер орчин дахь мэдээллийн нууцлал, бүрэн бүтэн байдал, хүртээмжийг хамгаалах эрх зүйн үндсийг тогтоосон [7]. Мөн 2023 оны 1-р сард Монгол Улсын кибер аюулгүй байдлын үндэсний стратегийг баталж, төр, иргэн, хуулийн этгээдийн мэдээллийн аюулгүй байдлыг хангах чиглэлээр зорилтуудыг тодорхойлсон байна [8].

Иймд энэхүү судалгааны зорилго нь Монгол Улсад нээлттэй эхийн сангуудын хэрэглээтэй холбоотой

аюулгүй байдлын мэдлэг, туршлага, хэрэглэгдэж буй арга хэрэгсэл, бодлогын хэрэгжилтийг судлахад оршино.

II. СУДЛАГДСАН БАЙДАЛ

Нээлттэй эхийн сангийн эмзэг байдал, аюулгүй хөгжүүлэлтийн чиглэлээр сүүлийн жилүүдэд олон улсын түвшинд эрчимтэй судалгаа хийгдэж, шийдлүүдийг гаргаж байна.

A. Судалгааны хүрээнд гарч буй асуудлууд

Synopsys компанийн CyRC (Cybersecurity Research Center)-с жил бүр гаргадаг "Open Source Security and Risk Analysis" (OSSRA) тайланд дурдсанаар [9]:

- 2023 онд судалгаанд хамрагдсан 1481 программын 84% нь нээлттэй эхийн сангийн мэдэгдэж буй эмзэг байдал агуулж байсан бөгөөд тэдгээрийн 48% нь Remote Code Execution.
- Кодын сангийн 91 хувь нь хуучирсан, засвар хийгдээгүй нээлттэй эхийн сан агуулж байсан.
- Автомашин, логистик, эрчим хүч гэх мэт салбаруудад нийт кодын 70–78 хувийг нээлттэй эхийн сан эзэлж байсан, эдгээрийн дийлэнх нь эмзэг байдал илэрсэн ч засаагүй орхигдсон.

Sonatype компанийн судлаачид 2018 онд программ хангамжийн нийлүүлэлтийн экосистемд 300 тэрбум гаруй нээлттэй эхийн сангийн бүрэлдэхүүн татагдсан мэдээлэлд дүн шинжилгээ хийхэд [10]:

- Татагдсан бүрэлдэхүүний "найм тутмын нэг" нь CVE (Common Vulnerabilities and Exposures)-д бүртгэгдсэн эмзэг байдлыг агуулсан байжээ.
- Нээлттэй эхийн эмзэг байдлын тоо 1 жилийн дотор 120%-иар өсөж, хакеруудын эмзэг байдлыг ашиглах хугацаа 45 хоногоос 3 хоног болж эрс багассан нь тогтоогдсон.

2022 онд Snyk болон Linux Foundation хамтран 550 программ хангамж хөгжүүлэгчдээс авсан судалгаагаар [11]:

- Оролцогчдын 51% нь ямар ч бодлогогүй байна гэсэн хариу гарсан.
- Нэг төсөлд дунджаар 49 эмзэг байдал, 80 орчим хамааралтай сан бүртгэгдэж, эдгээрийн 40% нь дам хамаарлын сангуудад(transitive dependencies) нуугдсан байдаг гэсэн дүгнэлт нь онцгой анхаарал татсан асуудал болсон.
- 33% нь CI/CD орчиндоо SAST, SCA зэрэг хэрэгслийг нэвтрүүлсэн гэж хариулсан.

2021 онд European Union Agency for Cybersecurity (ENISA)-аас гаргасан тайланд нээлттэй эхийн төслүүдэд аюулгүй байдлыг хангах бүтэцтэй бодлого, нөөц дутмаг байдгаас эмзэг байдал урт хугацаанд хадгалагдаж үлддэг нь нийтлэг үзэгдэл гэж тэмдэглэсэн [12]. Түүнчлэн GitHub-ийн Octoverse 2020 тайланд хөгжүүлэгчдийн 94% нь нээлттэй эхийн сангийн код дахин ашигладаг ч 30%-иас бага хувь нь аюулгүй байдлыг шалгадаг гэсэн үр дүн гарчээ [13].

B. Хийгдэж буй ажлууд

Түүнчлэн Open-Source Security Foundation (OpenSSF) нээлттэй эхийн хамгаалалтыг сайжруулах 10 гаруй чиглэлд хөрөнгө оруулалт хийх төлөвлөгөө гаргасан. 2022 онд Log4Shell-н эмзэг байдал илэрсний дараагаар эдгээр санаачилгад нийт $150 сая ам.долларын санхүүжилт шаардлагатайг тооцоолон, Google, Microsoft, Intel, Amazon зэрэг компаниудын дэмжлэгтэйгээр хэрэгжих уг төлөвлөгөө нь SBOM хэрэгжилт, найдвартай програмчлалын хэл ашиглах, мөн нээлттэй эхийн аюулгүй байдлын талаарх мэдлэгийг сайжруулахад чиглэж байна [14]. Мөн аюулгүй байдлын олон улсын экспертүүд нээлттэй эхийн программ хангамжийг улс орнуудын стратегийн дэд бүтэцтэй дүйцэхүйц ач холбогдолтой болсон хэмээн дүгнэж байна. Тухайлбал, криптографийн нэрт судлаач Брюс Шнайер нээлттэй эхийн программ хангамжийг олон нийтийн баялаг хэмээн үзэж, зам, гүүртэй адил стандарт, арчилгаа, хяналт шаардлагатай гэж онцолсон нь энэ чиглэлд хандлага өөрчлөгдөж буйг тод илэрхийлж байна [15]. Германы засгийн газар ч нээлттэй эхийг нийтийн баялагт тооцон дэд бүтцийн нэг хэсэгт тооцох бодлого баримталж эхэлсэн [15].

Олон улсын жишгээр АНУ нь 2023 оны Үндэсний кибер аюулгүй байдлын стратегидаа Нээлттэй эхийн программ хангамжийн аюулгүй байдлын санаачилга (OS3I)-ыг тусгаж [16], нээлттэй эхийн сангуудын бүрэлдэхүүний жагсаалт (SBOM) гаргахыг худалдан авалтын журамд заасан бол [17], Европын холбоо "Кибер тогтвортой байдлын акт" [18] нэртэй зохицуулалтаар программ хангамжийн аюулгүй байдлын наад захын стандартыг хуульчилж эхэлсэн.

III. НЭЭЛТТЭЙ ЭХИЙН АЮУЛГҮЙ БАЙДЛЫН СУДАЛГАА

Энэ бүлэгт Монгол Улсад нээлттэй эхийн сангийн аюулгүй байдлын хэрэглээ, мэдлэг, туршлагын бодит нөхцөл байдлыг тодорхойлох зорилготойгоор хийсэн тоон ба чанарын судалгаа, түүний үр дүнг танилцуулж байна.

Судалгаанд хосолсон аргыг (mixed-method) ашигласан. Тоон судалгаагаар 70 программ хангамжийн хөгжүүлэгч, мэдээллийн аюулгүй байдлын мэргэжилтнүүдээс санал асуулга авч, чанарын судалгаагаар төр, хувийн хэвшлийн байгууллагын 8 удирдах болон гүйцэтгэх түвшний ажилтантай ярилцлага хийв. Үр дүнг сэдэвчилсэн байдлаар нэгтгэн дүгнэж, сайжруулах чиглэлийг боловсруулсан.

Тоон судалгаа

A. Судалгааны аргачлал ба өгөгдөл цуглуулалт:

Судалгааг Google Forms ашиглан 25 асуултаас бүрдэх цахим санал асуулгаар зохион байгуулсан. Оролцогчид боловсрол, банк, харилцаа холбоо, төрийн үйлчилгээ, программ хөгжүүлэлтийн салбарыг төлөөлж байгаа ба судалгаа 14 хоногийн хугацаанд явагдсан.

B. Өгөгдөл боловсруулах ба шинжилгээний арга:

Цуглуулсан өгөгдлийг статистикийн аргаар боловсруулж, хувь хэмжээ, хариултын бүлэглэлээр дүн шинжилгээ хийв. Сэдэвчилсэн байдлаар үр дүнг нэгтгэн боловсруулсан.

С. Үндсэн үр дүн:

1. **Нээлттэй эхийн сангийн хэрэглээний түвшин:** Судалгаанд оролцогчдын 64% нь нээлттэй эхийг тогтмол ашигладаг, зөвхөн 11% нь огт хэрэглэдэггүй гэж хариулсан.
2. **Аюулгүй байдлын талаарх мэдлэг, ойлголт, туршлага:** Оролцогчдын 79% нь аюулгүй код бичих (Secure Coding Practices) чиглэлээр туршлагагүй, харин 30% нь байгууллагад нь огт сургалт зохион байгуулдаггүй гэжээ.
3. **Secure SDLC хэрэгжилт:** Аюулгүй программ хангамж хөгжүүлэлтийн мөчлөг (Secure SDLC)-ийн хэрэгжилтийн талаарх судалгаанд оролцогчдын 31% нь байгууллагад нь холбогдох аргачлал огт хэрэгждэггүй гэж хариулсан. DevSecOps зарчим нэвтрүүлсэн байгууллага 16% байсан бол автоматжуулсан тест, эрсдэлийн загварчлал зэрэг нарийвчилсан аргыг ашигладаг нь 10%-иас бага байв.
4. **Эмзэг байдлын удирдлага, шалгалт, хариу арга хэмжээний практик:** Нээлттэй эхийн сангийн аюулгүй байдлыг хэрхэн хянаж, асуудалд хэрхэн хариу үйлдэл үзүүлдэг талаар хийсэн судалгаагаар олон байгууллага урьдчилан сэргийлэх хяналт сул, гол төлөв асуудал үүссэний дараа арга хэмжээ авдаг хандлагатай байна. Оролцогчдын 15% нь эмзэг байдлыг тогтмол урьдчилан шалгадаг гэж хариулсан бол 13% нь огт шалгадаггүй. Нээлттэй эхийн санг ашиглах үеийн аюулгүй байдлын арга хэмжээ авах талаарх асуултад ~35% нь "тодорхой арга хэмжээ авдаггүй", 28.75% нь "мэргэжилтэнд даатгадаг" гэжээ. Харин 25% нь өмнөх эмзэг байдлын мэдээллийг шалгадаг ба GitHub, CVE, NVD зэрэг албан ёсны болон албан бус эх сурвалж ашигладаг. Мөн оролцогчдын 12.5% орчим нь Dependency-Check, Snyk, Dependabot зэрэг автоматжуулсан хэрэгсэл ашигладаг гэжээ. 15% нь нээлттэй эхийн кодыг гаргасан талын аюулгүй байдлын бодлогод бүрэн найдаж ашигладаг байна.
5. **Нээлттэй асуултуудын дүн шинжилгээ:** Нээлттэй асуултаар оролцогчдоос Монгол Улсад нээлттэй эхийн сангийн аюулгүй байдалд анхаарал хандуулахгүй байгаа гол шалтгаануудыг тодруулахад, дараах хүчин зүйлс хамгийн их давтагдсан. Үүнд:

- Мэдлэг, ойлголт дутмаг – Оролцогчдын 60% нь нээлттэй эхийн сангийн аюулгүй байдлын талаарх ерөнхий мэдлэг тааруу байгааг онцолсон.
- Хөгжүүлэлтийн хугацаа давчуу – Программ хангамж хөгжүүлэх хугацаа шахуу байдаг тул аюулгүй байдлыг нэгдүгээрт тавих боломжгүй болдог хэмээн тэмдэглэсэн.
- Бодлого, журам дутмаг – 32.5% нь компаниудад нээлттэй эхийн сангийн зохицуулалт байхгүй,
- Санхүүгийн хязгаарлалт – 30% орчим оролцогч төсөв, санхүүгийн боломж муугаас шалтгаалж аюулгүй байдлыг орхигдуулдаг гэж үзжээ.
- Хүний нөөцийн хомсдол – 26% нь тусгайлан энэ чиглэлээр ажиллах IT ажилтан, аюулгүй байдлын шинжээч хангалтгүй гэжээ. Олон байгууллагад аюулгүй байдлын зориулалтын баг байхаас илүүтэй, хөгжүүлэгчдэд энэхүү чиг үүргийг давхар хариуцуулж байна.
- Reactive хандлага – Оролцогчдын 24% нь ерөнхийдөө байгууллагууд "асуудал гарсны дараа л аюулгүй байдалд анхаардаг" хандлагатайг онцолсон. Өөрөөр хэлбэл, proactive бус reactive арга барил түгээмэл байна.
- "Нээлттэй эх бол аюулгүй" гэсэн эндүүрэл – 22.5% нь нээлттэй эхийг анхнаасаа туршигдаж батлагдсан, олон хүн оролцдог учраас аюулгүй гэж хүлээж авдгаас болж анхаарал сул байдаг гэжээ.
- Ойлголтын зөрүү – 21% нь "программ хангамжийн аюулгүй байдал ба нээлттэй эхийн сангийн аюулгүй байдал хоёрын ялгааг сайн ойлгодоггүй" гэсэн нь байгууллагын удирдлага болон хөгжүүлэлтийн багийн гишүүд эдгээр ойлголтыг тусад нь авч үзэхгүй байх, төөрөгдөлд орж буруу үнэлэх, улмаар нээлттэй эхийн аюулгүй байдлын ач холбогдлыг бууруулах хандлага түгээмэл байж болзошгүйг харуулж байна.
- Бизнесийн ач холбогдол багатай гэж үздэг – 18.8% нь компаниуд нээлттэй эхийн сангийн аюулгүй байдлыг бизнесийн амжилт, үр дүнтэй холбоотой чухал үзүүлэлт биш гэж хардаг гэж дүгнэсэн.

Нээлттэй эхийн сангийн аюулгүй байдлаас үүдэлтэй нийтлэг асуудалд оролцогчдын тал орчим (48.75%) нь "Асуудал тулгарч байгаагүй" гэж хариулсан. Гэсэн хэдий ч энэ нь тухайн байгууллагад аюулгүй байдлын зөрчил бодитоор байхгүйг бүрэн батлах үндэслэл болохгүй. Учир нь уг бүлгийн оролцогчдын дийлэнх нь нээлттэй эхийн сангийн аюулгүй байдлыг урьдчилан шалгах тодорхой арга хэмжээ авдаггүй (жишээ нь: эмзэг байдлын автоматжуулсан шинжилгээ (SAST/DAST), Dependency Scanning (Snyk, Dependabot), код шалгалт (Code Review), эсвэл Software Bill of Materials (SBOM) боловсруулах), мөн "аюулгүй байдлын албан ёсны бодлого, журам байгууллагад нь байхгүй" гэж хариулсан байна. Үүнтэй адилаар зарим хариултад логик зөрүү илэрч байсан нь аюулгүй байдлын талаарх ойлголт, мэдлэгийн төвшин байгууллага дотор нэгдмэл бус, тогтолцоогүй байгааг харуулж байна.

Чанарын судалгаа

А. Судалгааны зорилго:

Чанарын судалгааны зорилго нь Монгол Улсад аюулгүй программ хангамж хөгжүүлэлт болон нээлттэй эхийн сангийн хэрэглээ, аюулгүй байдлын талаарх мэргэжлийн практик, сорилт, хандлагыг тодорхойлох явдал байв. Судалгаа нь мэдээллийн технологийн

салбарт ажилладаг, олон жилийн туршлагатай удирдах болон гүйцэтгэх түвшний мэргэжилтнүүдийн бодит туршлага, үнэлэмжид тулгуурласан.

B. Судалгааны аргачлал ба өгөгдөл цуглуулалт:

Нийт 9 мэргэжилтэн ярилцлага өгсөн бөгөөд тэдгээр нь төрийн болон хувийн хэвшлийн байгууллагад ажилладаг ахлах инженер, хөгжүүлэгч, мэдээллийн аюулгүй байдлын чиглэлээр ажилладаг мэргэжилтнүүд байв. Ярилцлагууд биечлэн уулзаж, урьдчилан бэлтгэсэн асуултын дагуу ярилцаж, дунджаар 30–60 минут үргэлжилсэн.

C. Өгөгдөл боловсруулах ба шинжилгээний арга:

Ярилцлагын үр дүнг сэдэвчилсэн шинжилгээ (thematic analysis) хийж, аюулгүй хөгжүүлэлт, нээлттэй эхийн сангийн аюулгүй байдал, хүний нөөц, эрх зүйн орчны талаарх хандлага, хэрэгжилтийг үнэлсэн.

D. Үндсэн үр дүн:

1. **Аюулгүй хөгжүүлэлтийн хэрэгжилт:** Судалгаанд оролцогчдын ихэнх нь SDLC-ийн бүх үе шатанд аюулгүй байдлыг тусган хэрэгжүүлдэггүй, харин хөгжүүлэлтийн дараах шатанд авч үздэг хандлагатай байна. Зарим байгууллага олон улсын стандарт мөрддөг ч нийтлэг практик биш байна. Нийтлэг практик гэдэг нь ISO/IEC 27001 (Мэдээллийн аюулгүй байдлын стандарт), OWASP Top 10-ийн заавар, NIST-ийн Secure Software Development Framework (SSDF) зэрэг олон улсад хүлээн зөвшөөрөгдсөн аргачлалуудыг хэлнэ. Эдгээр нь Монгол Улсад хангалттай түгээгдээгүй байна.
2. **Нээлттэй эхийн сангийн бодлого:** Нээлттэй эхийн сангийн хэрэглээ өндөр түвшинд байгаа ч сонголт, хяналт, баталгаажуулалтын үйл явц нь олон байгууллагад албан ёсны бодлогогүй, инженерийн хувь хүний шийдвэрт суурилсан байдал давамгайлж байна.
3. **Аудит, шалгалт:** Тогтмол аюулгүй байдлын шалгалт, эрсдэлийн хяналт хэрэгжүүлдэг байгууллагын тоо цөөн байна. Зарим байгууллага систем бүрэн ашиглалтад ороогүй гэдэг шалтгаанаар аудит хийх шаардлагагүй гэж үздэг.
4. **Хүний нөөцийн бэлэн байдал:** Аюулгүй байдлын чиглэлээр мэргэшсэн тусгайлсан албан тушаалтантай байгууллага ховор бөгөөд хөгжүүлэгчид аюулгүй байдлыг давхар хариуцдаг.
5. **Хууль, бодлогын орчин:** Эрх зүйн үндэс бүрдсэн ч нээлттэй эхийн сангийн хэрэглээнд чиглэсэн нарийвчилсан заавар, хэрэгжилтийн шалгуур дутагдаж байна.

Үүнээс дүгнэвэл Монгол Улсад программ хангамжийн аюулгүй хөгжүүлэлтийн түвшин харилцан адилгүй бөгөөд:

- Туршлагатай, олон нийтэд чиглэсэн үйлчилгээ үзүүлдэг байгууллагууд илүү системтэй, зохион байгуулалттай, ISO болон хууль журмын шаардлагыг дагаж мөрддөг бол
- Бусад байгууллагууд хувь хүний чадвар, дотоод туршлагад суурилсан, стандартгүй хандлага давамгайлж байна.

IV. МОНГОЛ УЛСАД НЭЭЛТТЭЙ ЭХТЭЙ ХОЛБООТОЙ ТУЛГАМДСАН АСУУДЛУУД, ШИЙДВЭРЛЭХ САНАЛ

Энэхүү судалгааны хүрээнд нээлттэй эхийн аюулгүй байдлын талаарх тоон болон чанарын судалгаанд үндэслэн дараах асуудлуудыг тодорхойлов. Үүнд:

A. Бодлого, стандартын хомсдол:

Нээлттэй эхийн санг аюулгүй ашиглах талаар салбарын түвшинд мөрдөх стандарт, бодлогын баримт бичиг дутмаг байна. Төрийн зүгээс нээлттэй эхийн программ хангамжийг ашиглах, үнэлэх, сонгох, шинэчлэхэд баримтлах нэгдсэн зааварчилгаа гараагүй нь байгууллагууд өөр өөрсдийн ойлголт, боломжоор асуудлыг шийдэхэд хүргэж байна.

B. Мэргэжилтний хомсдол ба технологийн нөөцийн дутагдал:

Монгол Улсад кибер аюулгүй байдлын хүний нөөцийн хомсдол нь байгаа нь программ хангамж хөгжүүлэлт, нээлттэй эхийн сангийн аюулгүй байдалд шууд нөлөөтэй байна.

C. Лицензийн зөрчил ба эрх зүйн эрсдэл:

Нээлттэй эхийн санг “үнэгүй” гэсэн ойлголтоор шууд ашиглаж, лицензийн нөхцөлийг үл тоомсорлох хандлага түгээмэл байна. Энэ нь программ хангамж, улмаар байгууллагыг эрсдэлд оруулах боломжтой.

Дээр дурдсан асуудлууд нь Монгол Улсад нээлттэй эхийн сангийн аюулгүй байдлын талаарх зохицуулалт, хэрэгжилт тал дээр анхаарах шаардлагатай олон сул талыг илэрхийлж байна. Гэсэн хэдий ч дэлхий нийтийн чиг хандлагыг судалж, сайн туршлагыг нутагшуулах замаар эдгээрийг шийдвэрлэх боломжтой.

Шийдвэрлэх санал

Монгол Улс олон улсын чиг хандлагаас суралцаж, нээлттэй эхийн сангийн аюулгүй байдлыг дээшлүүлэхэд тулгарч буй сорилтуудыг харгалзан дараах практик арга хэмжээг хэрэгжүүлэхийг санал болгож байна:

A. Үндэсний гарын авлага, стандарт гаргах:

Нээлттэй эхийн санг сонгох, ашиглах, шинэчлэх, аудит хийх, лицензийн нийцлийг шалгах, эмзэг байдлын хариу арга хэмжээ авах үйл явцыг зохицуулсан нэгдсэн стандарт, гарын авлага боловсруулах. Уг стандарт нь олон улсын OWASP OpenSAMM, NIST SSDF зэрэг

аюулгүй хөгжүүлэлтийн санамжуудтай уялдсан байвал тохиромжтой.

B. Нийлүүлэлтийн сүлжээний ил тод байдлыг нэмэгдүүлэх:

АНУ-ын Ерөнхийлөгчийн 14028 тоот зарлигийн хүрээнд [2021, Executive Order 14028 – Improving the Nation's Cybersecurity] SBOM-ыг бүх төрийн болон нийлүүлэгч байгууллагуудад шаардлагатай бүрэлдэхүүн гэж тодорхойлсон [19]. Үүний дагуу NIST болон CISA зэрэг агентлагууд SBOM-тай холбоотой практик аргачлал, хэрэгжүүлэлтийн гарын авлага боловсруулсан. SBOM нь программ хангамжийн бүрэлдэхүүнийг бүртгэж, эмзэг байдлыг хянахад тусалдаг.

Монгол Улсад ч энэхүү жишгээр, төрийн болон стратегийн мэдээллийн системүүдэд SBOM хөтлөх, шинэ эмзэг байдал үүсэхэд системийн нөлөөллийг тогтоох, хурдан засварлах зохицуулалт хийх боломжтой. Dependency-Track, Synopsys Black Duck зэрэг автоматжуулсан хэрэгслүүдийг ашиглан автоматаар эмзэг байдлыг илрүүлэх зэрэг нэвтрүүлэх боломжийг судлах нь зүйтэй.

D. Хүний нөөцийг чадавхжуулах, мэргэшүүлэх:

Их дээд сургуулиудад Secure Coding, Open-Source Security зэрэг хичээл оруулах, олон улсад хүлээн зөвшөөрөгдсөн сертификат (жишээ нь, (ISC)² CSSLP) олгох хөтөлбөрийг нэвтрүүлэхийг санал болгож байна. Тухайлбал, Coursera, edX зэрэг платформтой хамтран хөнгөлөлттэй сургалт явуулах, эсвэл OpenSSF-ийн нээлттэй сургалтын материалыг (OpenSSF Education, OpenSSF Training) Монгол хэл дээр орчуулан хөгжүүлэгчдэд түгээх боломжтой.

E. Нээлттэй эхийн оролцоог нэмэгдүүлж, тогтмол сайжруулах систем бий болгох:

Гадаадын программ хангамжид түшиглэхээс гадна, Монголын хөгжүүлэгчид чухал нээлттэй төслүүдэд идэвхтэй хувь нэмэр оруулах орчныг бүрдүүлэх. Өргөн хэрэглэгддэг нээлттэй эхийн сангууд (жишээ нь, кирилл хөрвүүлэлт) дээр үндэслэн fork үүсгэж, сайжруулалт, хамгаалалтын шинэчлэл хийх дотоод баг бүрдүүлэх. Ингэснээр алдаа гарах үед дотоод мэргэжилтнүүд хурдан хариу өгөх чадвартай болох юм. Энэ нь урт хугацаанд аюулгүй байдлыг хангах суурь болно.

F. Лиценз, эрхзүйн мэдлэг түгээх:

Хөгжүүлэгчид болон эрхзүйн ажилтнуудад зориулсан нээлттэй эхийн лицензийн сургалт, гарын авлага боловсруулахыг дэмжих. Ялангуяа GPL, Apache, MIT, MPL зэрэг гол лицензийн ялгаа, эрх зүйн үүргийг тайлбарласан монгол хэл дээрх ойлгомжтой материал (ном, товхимол, цахим контент) бэлтгэх шаардлагатай. Мөн компанийн хуульч, эрх зүйн зөвлөхүүдэд зориулсан семинар зохион байгуулж, гэрээ боловсруулахдаа лицензийн нөхцөлийг зөв тусгах мэдлэг олгох нь чухал. Лицензийн зөрчил нь зохиогчийн эрхийн маргаан үүсгэх эрсдэлтэйг бодит жишээгээр ойлгуулах нь зүйтэй.

V. ДҮГНЭЛТ

Монгол Улсад нээлттэй эхийн санг өргөнөөр хэрэглэгддэг хэдий ч аюулгүй байдлыг хангах соёл, мэдлэг, стандартчилал сул, зарим тохиолдолд бүр байхгүй байгааг энэхүү судалгаа харууллаа. Судалгаанд оролцогчдын 35%-аас илүү нь эмзэг байдлыг шалгахад тогтмол арга хэмжээ авдаггүй гэж хариулсан нь SAST, DAST зэрэг хэрэгслийг нэвтрүүлэх, аюулгүй код бичилтийн соёлыг төлөвшүүлэх шаардлагатайг онцоллоо.

Тулгарч буй гол асуудлууд:

- Мэдлэг, мэдээллийн хомс байдал;
- Мэргэшсэн хүний нөөцийн дутагдал;
- Цаг хугацаа, санхүүгийн хязгаарлалт;
- Хууль, бодлогын нарийвчилсан зохицуулалтын дутмаг байдал зэрэгт оршиж байна.

Иймд дараах түлхүүр чиглэлүүдийг хэрэгжүүлэх нь зүйтэй:

1. Нээлттэй эхийн сангийн хэрэглээ, аюулгүй байдлыг зохицуулах OWASP, NIST-д суурилсан үндэсний стандарт, гарын авлага боловсруулах;
2. SBOM, автоматжуулсан хэрэгсэл ашиглан нийлүүлэлтийн сүлжээний аюулгүй байдлыг сайжруулах;
3. Хүний нөөцийг чадавхжуулахад чиглэсэн сургалт, гэрчилгээжүүлэлтийн тогтолцоог бий болгох;
4. Салбар хоорондын хамтын ажиллагаа, нээлттэй мэдлэгийн орчныг дэмжих арга хэмжээг хэрэгжүүлэх.

Эцэст нь дүгнэхэд, "аюулгүй хөгжүүлэлт" болон "нээлттэй эхийн сангийн аюулгүй байдал" нь цахим эринд амжилттай орших, өрсөлдөх чадвартай байх үндэс болохын сацуу, Монгол Улсын кибер орчны тогтвортой байдал, хөгжлийн итгэлцлийг хадгалах стратегийн хүчин зүйл гэдгийг онцлон тэмдэглэх нь зүйтэй.

VI. НОМ ЗҮЙ

[1] Nagle, F. & Hoffman, J. "The Hidden Vulnerabilities of Open Source Software," Harvard Business School, 2020 https://www.library.hbs.edu/working-knowledge/the-hidden-vulnerabilities-of-open-source-software

[2] Veracode, "The Dangers of Open Source Risk," Veracode Security Resources. [Online]. Available: https://www.veracode.com/security/dangers-open-source-risk

[3] OWASP Foundation, "OWASP Top 10," MIT CSAIL Course Reading, 2014. [Online]. Available: https://css.csail.mit.edu/6.858/2014/readings/owasp-top-10.pdf

[4] D. Raywood, "Ten years of Heartbleed: Lessons learned," SC Media, 2024. [Online]. Available: https://www.scworld.com/analysis/ten-years-of-heartbleed-lessons-learned

[5] F. Bals, "Equifax, Apache Struts, and CVE-2017-5638 vulnerability," Black Duck Blog, Sep. 14, 2017. [Online]. Available: https://www.blackduck.com/blog/equifax-apache-struts-vulnerability-cve-2017-5638.html

[6] M. Vizard, "Survey Uncovers Depth of Open Source Software Insecurity," DevOps.com, June 21, 2022. [Online]. Available: https://devops.com/survey-uncovers-depth-of-open-source-software-insecurity

[7] “Кибер аюулгүй байдлын тухай хууль,” Legalinfo.mn, 2021. [Online]. Available: https://legalinfo.mn/mn/detail?lawId=16390365491061

[8] Монгол Улсын Цахим хөгжил, харилцаа холбооны яам, “Кибер аюулгүй байдлын үндэсний стратеги (2023),” MDDIC.gov.mn, Jan. 2023. [Online]. Available: https://mddic.gov.mn/mn/2023/01/11900

[9] A. Venkat, “At least one open source vulnerability found in 84% of code bases: Report,” CSO Online, Feb. 23, 2023. [Online]. Available: https://www.csoonline.com/article/574607/at-least-one-open-source-vulnerability-found-in-84-of-code-bases-report.html

[10] T. Spring, “2018 Has Been Open Season on Open Source Supply Chains,” Threatpost, Sep. 26, 2018. [Online]. Available: https://threatpost.com/2018-has-been-open-season-on-open-source-supply-chains/137726

[11] M. Vizard, "Survey Uncovers Depth of Open Source Software Insecurity," DevOps.com, Jun. 21, 2022. [Online]. Available: https://devops.com/survey-uncovers-depth-of-open-source-software-insecurity/

[12] ENISA, “Open Source Software Security: Risk, compliance and assurance,” European Union Agency for Cybersecurity, 2021. [Online]. Available: https://www.enisa.europa.eu/publications

[13] GitHub, "The State of the Octoverse 2020," GitHub Octoverse Reports, 2020. [Online]. Available: https://octoverse.github.com/

[14] R. Lemos, “Big Tech Pledges Millions to Secure Open Source,” IEEE Computer Society, May 2022. [Online]. Available: https://www.computer.org/publications/tech-news/trends/big-tech-pledges-money-to-open-source

[15] B. Schneier, “Securing Open-Source Software,” Schneier on Security, Jul. 27, 2022. [Online]. Available: https://www.schneier.com/blog/archives/2022/07/securing-open-source-software.html

[16] J. Gregory, “What is the Open-Source Software Security Initiative (OS3I)?”, IBM Think Blog, Feb. 26, 2024. [Online]. Available: https://www.ibm.com/think/news/open-source-software-security-initiative

[17] Cybersecurity and Infrastructure Security Agency, "Software Bill of Materials (SBOM)," CISA. [Online]. Available: https://www.cisa.gov/sbom

[18] European Commission, “Cyber Resilience Act,” Digital Strategy, Mar. 6, 2025. [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/cyber-resilience-act

[19] J. R. Biden, “Executive Order 14028: Improving the Nation’s Cybersecurity,” Federal Register, vol. 86, no. 93, pp. 26633–26647, May 17, 2021. [Online]. Available: https://www.federalregister.gov/documents/2021/05/17/2021-10460/improving-the-nations-cybersecurity

Нийгмийн сүлжээн дэх стандарт бус галиглагдсан бичвэрийг Кирилл бичвэрт хөрвүүлэх загварын судалгаа

Бямбадоржийн Золзаяа [1*], Соном-Очирын Өлзийбаяр [1], Аюушийн Алтангэрэл [1], Энхбаярын Мөнхсүх [1], Kim Hyunchul[2]

[1]Мэдээллийн Технологийн тэнхим, Шинжлэх Ухаан Технологийн Их Сургууль, Улаанбаатар, Монгол
[2]Солонгос улс, Чонан, Сангмён Их Сургууль, Програм хангамжийн тэнхим

Email: b.zolzaya@must.edu.mn, ulziibayar@must.edu.mn, a.altangerel@must.edu.mn, munkhsukh@must.edu.mn, hkim@smu.ac.kr

Хураангуй— Сүүлийн жилүүдэд олон нийтийн мэдээллийн хэрэгсэл, ялангуяа сошиал медиа нь хүмүүсийн өдөр тутмын амьдралын салшгүй хэсэг болж, нийгмийн харилцаа, мэдээлэл түгээх шинэ орон зайг бүрдүүлж байна. Сошиал медиа нь нийгэм, эдийн засгийн асуудлыг шинжлэх, бодлого боловсруулах, бизнесийн шийдвэр гаргалтыг дэмжих баялаг текст өгөгдлийг агуулдаг ч эдгээр нь ихэвчлэн албан бус хэлбэртэй, шуугиантай, түүнчлэн Латин үсгээр галигласан байдаг. Энэ нь хэл зүйн боловсруулалтыг хүндрүүлж, мэдээллийн ойлголцолд төдийгүй өгөгдөлд суурилсан дүн шинжилгээ, автомат шийдвэр гаргалтад ч саад болдог. Энэхүү судалгаагаар Латин үсгээр галиглагдсан, шуугиантай Монгол үгсийг стандарт Кирилл үгрүү хөрвүүлэх асуудлыг шийдвэрлэх зорилгоор үгийн түвшний болон тэмдэгтийн түвшний Seq2Seq загварууд, статистик галиглах загвар (TM), эдгээрийн хослол бүхий гибрид загваруудын гүйцэтгэлийг харьцуулан судалсан. Үүнээс гадна хөрвүүлэлтийн нарийвчлалыг нэмэгдүүлэх зорилгоор толь бичигт суурилсан шалгах, edit distance ашиглан засварлах, мөн BERT хэлний загвар дээр суурилсан контекст ойлголттой дахин эрэмбэлэх ажиллагааг нэгтгэсэн сайжруулагч модулийг боловсруулж хэрэгжүүлсэн. Туршилтын үр дүнгээс харахад сайжруулагч модулийг ашигласнаар бүх загварын гүйцэтгэл мэдэгдэхүйц сайжирсан. Анхны гүйцэтгэлээр TM загвар 80.52%, Seq2Seq загвар 74.86%, TM+Seq2Seq гибрид загвар 83.88%-ийн нарийвчлалтай байсан бол сайжруулагч модулийг нэмж хэрэгжүүлснээр эдгээр нь тус тус 84.90%, 81.59%, 86.53% болж өссөн байна. Сургалтын өгөгдлийн багцад агуулагдсан үгсийн хувьд бүх загвар 94%-иас дээш нарийвчлалтай байсан бол агуулагдаагүй буюу шинэ үгсийн хувьд TM загвар 71.90%, Seq2Seq загвар 65.56%, TM+Seq2Seq гибрид загвар 75.50%-ийн гүйцэтгэлтэй байсан нь сайжруулагч модуль шуугиантай, өгөгдлийн хомсдолтой нөхцөлд илүү үр дүнтэй ажиллаж байгааг харуулж байна. Судалгааны үр дүнгээс дүгнэхэд, боловсруулсан сайжруулагч модулийг хөрвүүлэлтийн аль ч загварын гүйцэтгэлийг сайжруулах нэмэлт давхарга болгон ашиглах боломжтой бөгөөд хэлний дүрэм болон гүн сургалтын хосолсон арга нь Монгол хэлний шуугиантай галиглагдсан үгийг стандарт хэлбэрт хөрвүүлэхэд үр дүнтэй шийдэл болж байна.

Түлхүүр үгс—Text normalization, noisy text, transliterated text, neural network, seq2seq model, transliteration model.

I. Оршил

Сүүлийн жилүүдэд олон нийтийн мэдээллийн хэрэгсэл нь бичгийн хэлийг хамгийн их ашигладаг платформ болсон. Эдгээр нийгмийн сүлжээнүүд дээр хуваалцсан мэдээлэл, өгөгдөл нь ихээхэн үнэ цэнэтэй хэдий ч, дүрмийн бус нийтлэлүүд, товчлолууд, эможи зэрэг нь түгээмэл байдаг тул шууд ашиглахад хүндрэл учруулдаг. Түүнчлэн, олон нийтийн сүлжээнд Латин үсэг ашиглан галиглагдсан бичвэрүүд нь түгээмэл бөгөөд эдгээр текстүүдийг эх хэлний боловсруулалтад ашиглахад бэрхшээл учруулдаг. Иймээс галиглагдсан бичвэрийг хэвийн болгох буюу Кирилл үсэгтэй бичвэр рүү хөрвүүлэх нь эх хэлний боловсруулалтын хамгийн том сорилтуудын нэг болсон юм.

Энэ чиглэлийн судалгааны арга зүй нь дүрэмд суурилсан хандлага, статистикийн арга, мэдрэлийн сүлжээ, гибрид арга зэргийг хамардаг. Эдгээр аргуудыг ашигласан судалгаануудыг дурьдвал, [1] судалгаанд Малай шуугиантай текстийг хэвийн болгоход дүрэмд суурилсан аргыг буюу урьдчилан тодорхойлсон дүрмүүдийг ашигласан бол, [2] судалгаанд SMS англи текстийг стандарт хэлбэрт хөрвүүлэхийн тулд статистикийн машин орчуулгын (СМО) загварыг санал болгосон. Мөн түүнчлэн, [3] судалгаанд шуугиан ихтэй бичигдсэн англи текстийг хэвийн болгох оновчтой нэр дэвшигчдийг сонгох зургаан аргыг хэлний загвар ашиглан хоёр үе шаттай хяналтгүй системийг боловсруулсан. Сүүлийн жилүүдэд мэдрэлийн сүлжээнд суурилсан аргууд, ялангуяа sequence-to-sequence (seq2seq) арга нь текстийг хэвийн болгоход гайхалтай ахиц дэвшлийг харуулсан. Энэ аргыг [4-8] судалгаанд анх танилцуулсан бөгөөд үгийн болон тэмдэгтийн түвшинд seq2seq загваруудыг хэрэгжүүлсэн. [9] энэхүү загварыг текстийг хэвийн болгох даалгаварт ашигласан ба шуугиан ихтэй япон текстийг хэвийн болгосон. Тэд өгөгдлийг нэмэгдүүлэхийн тулд урьдчилан тодорхойлсон дүрмүүд бүхий синтетик өгөгдлийн багцыг үүсгэж, өөрсдийн мэдрэлийн сүлжээний загварыг дүрэмд суурилсан аргуудтай харьцуулсан. Үүний нэгэн адил, [10] судалгаанд үгэнд суурилсан seq2seq аргыг ашиглан Казахстаны интернетийн сегментээс казах хэл дээрх хэрэглэгчийн үүсгэсэн контентыг хэвийн болгоход анхаарлаа хандуулсан. Эдгээр судалгаанууд нь өөр өөр хэл дээрх шуугиантай текстийг хэвийн болгох ажлыг

гүйцэтгэхэд seq2seq загваруудын үр нөлөөг онцолж байна.

Үгийн болон тэмдэгтийн түвшний seq2seq аргууд нь өөрийн давуу болон хязгаарлагдмал талуудтай байсан. [11]-д дурьдсанчлан, үгийн түвшний seq2seq арга нь сургалтын өгөгдлийн багцад агуулагдсан үгсийг хэвийн болгоход сайн хэдий ч, агуулагдаагүй үгсийг хэвийн болгоход бэрхшээлтэй байсан. Үүний эсрэгээр, тэмдэгтийн түвшний seq2seq арга нь сургалтын өгөгдлийн багцад агуулагдаагүй үгсийг эсвэл шуугиантай үгсийг илүү сайн зохицуулдаг боловч тооцоллын хувьд илүү их нөөц шаарддаг. Иймд эдгээр асуудлыг шийдвэрлэхийн тулд үг болон тэмдэгтийн түвшний seq2seq загваруудыг хослуулсан гибрид аргыг боловсруулсан. Энэхүү гибрид арга нь үгийн түвшний seq2seq загварыг ашиглах ба, сургалтын өгөгдлийн багцад агуулагдаагүй үгтэй тулгарах үед тэмдэгтийн түвшний seq2seq загвар руу шилжиж, үгийг хэвийн болгосноор гүйцэтгэлийг сайжруулсан. Нэмж дурдахад, [12] судалгаанд тэмдэгтийн түвшний seq2seq-г машин сургалтын арга техниктэй хослуулж шуугиантай текстийг хэвийн болгоход анхаарлаа хандуулсан. Эдгээр судалгаанууд нь нийгмийн сүлжээн дэх шуугиан ихтэй галиглагдсан текстийг хэвийн болгох сорилтыг даван туулах боломжтойг харуулж байна.

Манай судалгааны зорилго нь шуугиантай, Латин үсгээр галиглагдсан текстийг Кирилл үсгээр бичсэн стандарт хэлбэрт бага хэмжээний өгөгдлийн тусламжтайгаар үр дүнтэйгээр хэвийн болгох арга санал болгох явдал юм. Үүний хүрээнд бид статистик арга болох галиглах загвар (transliteration model-TM), seq2seq загвар болон эдгээрийн хослол бүхий аргуудын гүйцэтгэлийг харьцуулан судалж, хамгийн үр нөлөөтэй аргыг тодорхойлсон. Мөн үүн дээр нэмэлтээр засварлах зай (edit distance-ED) болон BERT загварын хослол бүхий сайжруулагч модулийг хэрэгжүүлснээр үр дүнг улам сайжруулсан.

II. Өгөгдлийн сан

Бид нийгмийн сүлжээнээс шуугиантай, Латин үсгээр галиглагдсан бодит өгөгдлийг цуглуулан, өгүүлбэр бүрийг үг үгээр нь салгаж, тэдгээрийн харгалзах Кирилл үсгийн стандарт хувилбарыг гар аргаар үүсгэсэн. Сургалтын өгөгдлийн багцын статистик мэдээллийг Хүснэгт I-д үзүүлэв. Уг өгөгдлийг тэмдэгтийн болон үгийн түвшний seq2seq загваруудыг сургахад ашигласан. Нэмж дурдахад, бид Монгол хэлний толь бичгийн өгөгдлөөс [13] нийт 7680 Кирилл үгийг түүвэрлэн авч, тэдгээрийн стандарт галиглалыг MNS 5217:2012 стандартын [14] дагуу автоматаар хөрвүүлэн үүсгэж, ийнхүү бүрдүүлсэн хос үг бүхий өгөгдлийг TM загварын сургалтад ашигласан. Үгийн сонголтыг хийхдээ толь бичгийн бүтэц нь цагаан толгойн үсгийн дараалаар зохион байгуулагддаг онцлогт тулгуурлан, Монгол цагаан толгойн бүх үсгийг төлөөлөх зарчмаар түүвэрлэлт хийж, хэлний төлөөлөл, өгөгдлийн тэнцвэрт байдлыг хангахыг зорьсон болно. Энэхүү өгөгдлийн багцын дэлгэрэнгүй мэдээллийг Хүснэгт II-д харуулав. Мөн нийгмийн сүлжээнээс цуглуулсан шуугиантай, галиглагдсан 200 өгүүлбэрээс бүрдсэн тестийн өгөгдлийг Хүснэгт III-д тусгасан. Манай тестийн өгөгдлийн 58% нь сургалтын өгөгдлийн багцад багтсан үгсээс бүрддэг бол үлдсэн 42% нь багтаагүй буюу шинэ үгс байна. Цаашид сургалтын өгөгдөлд орсон үгсийг in-vocabulary (IV), харин агуулагдаагүй үгсийг out-of-vocabulary (OOV) гэж тэмдэглэх болно.

Хүснэгт I. Сургалтын өгөгдлийн сангийн статистик

Тайлбар	*Оролт*	*Гаралт*
Нийт үгийн тоо	24073	25023
Ялгаатай үгийн тоо	14197	12030
Үгийн хамгийн их урт	21	22

Хүснэгт II. ТМ загварыг сургахад ашигласан өгөгдлийн багц

Тайлбар	*Стандарт галиглагдсан үгс* **MNS 5217:2012**	*Кирилл үгс*
Ялгаатай үгийн тоо	7680	7680

Хүснэгт III. Тестийн өгөгдлийн сангийн статистик

Тайлбар	*Input*
Нийт үг	1663
Ялгаатай үгийн тоо	1178
IV үгийн тоо	970 (58%)
OOV үгийн тоо	694 (42%)

III. Арга зүй

Бид шуугиантай текстийг хэвийн болгохын тулд үгийн түвшний болон тэмдэгтийн түвшний seq2seq загваруудыг тус тус хэрэгжүүлсэн. Үүнээс гадна, статистикийн ТМ загварыг ашиглан шуугиантай текстийг стандарт хэлбэрт хөрвүүлэхийг зорьсон.

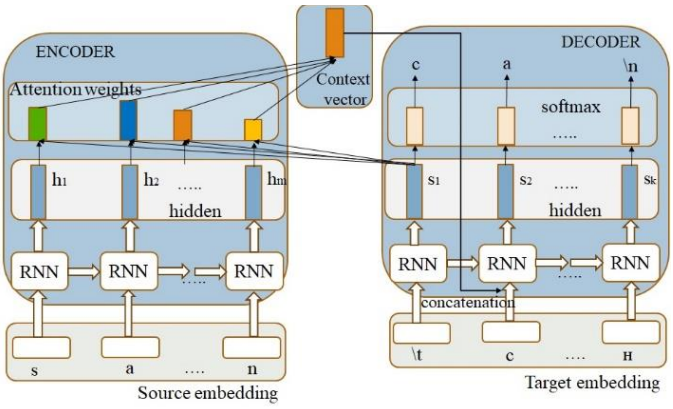
A. *Sequence to sequence загвар*

Шуугиантай, галигласан текстийг стандарт хэлбэрт хөрвүүлэхийн тулд бид [7]-д үндэслэн Зураг 1-т үзүүлсэн seq2seq загварыг боловсруулсан. Энэхүү загвар нь хэрэглэгчийн оруулсан текстийг Кирилл үсэгт суурилсан стандарт хувилбарт шилжүүлэхэд зориулагдсан бөгөөд хоёр үндсэн хэсгээс бүрдэнэ: энкодер болон декодер, эдгээр нь хоёулаа давтагддаг мэдрэлийн сүлжээ (RNN) юм.

Энкодер нь оролтын дараалыг боловсруулж, түүний мэдээллийг тогтмол урттай контекст вектор болгон шахдаг. Энэ контекст вектор нь декодерыг эхлүүлэх үүрэг гүйцэтгэдэг бөгөөд декодер нь тухайн оролтод тохирох хэвийн хэлбэрийг дараалаан үүсгэнэ. Энкодер нь оролтын дараалын элемент бүрд харгалзах далд төлөв (hidden state)-үүдийг үүсгэж, эдгээрийг декодер ашиглан гаралтыг үүсгэхдээ хэрэглэдэг.

Декодер нь тухайн алхмын өмнөх далд төлөв болон энкодерийн бүх далд төлөвүүдтэй харьцуулж, Бахданаугийн анхаарлын онооны функц [7] (Bahdanau alignment score function)-ийг ашиглан анхаарлын оноо (alignment score)-г тооцоолдог. Энэ нь анхаарлын механизм (attention mechanism)-ыг хэрэгжүүлэх боломжийг олгодог.

Үгийн болон тэмдэгтийн түвшний seq2seq загваруудыг TensorFlow [15] ашиглан хэрэгжүүлсэн ба дараах гиперпараметрүүдийг тохируулж өгсөн: энкодер болон декодерт тус бүр 512 далд нэгжтэй нэг GRU (Gated Recurrent Unit) давхарга, мөн 250 хэмжээтэй embedding вектор. Загваруудыг Adaptive Moment Estimation (Adam) оновчлогч ашиглан 0.0002 сургалтын хурд (learning rate) болон 32-ийн багцын хэмжээгээр (batch size) сургаж боловсруулсан.



Зураг 1. Seq2Seq загварын архитектур

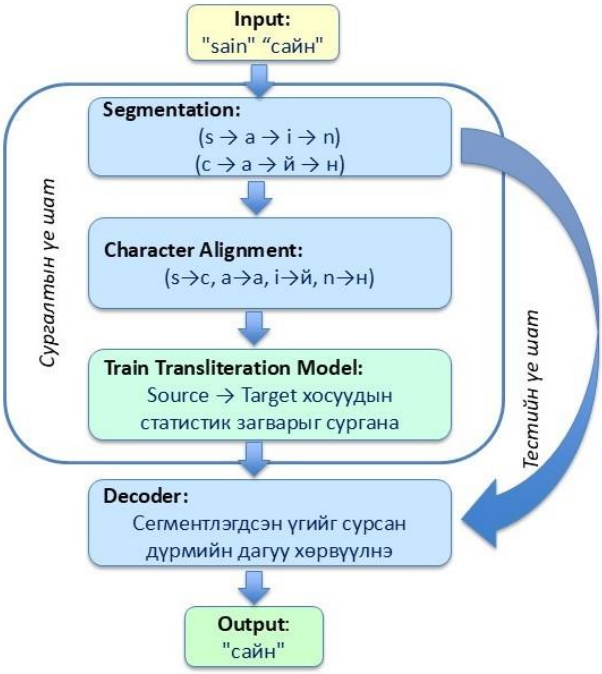
B. *Галиглах загвар*

Энэхүү судалгаанд галиглагдсан үгсийг стандарт Кирилл бичигт хөрвүүлэх галиглах загварыг статистик машин орчуулгын аргачлал дээр тулгуурлан боловсруулсан. Энэхүү загвар нь сургалтын өгөгдөлд байхгүй шинэ үгсийг ч авиа зүйн хувьд оновчтой хөрвүүлэх боломжтой.

Загварыг сургах үйл явц хэд хэдэн үндсэн үе шаттай. Эхний шатанд галиглагдсан бичвэрийг үсэг тус бүр эсвэл богино хэмжээний тэмдэгтийн дараалал болгон сегментчилнэ. Үүний дараа эх үсэг болон зорилтот хэлний үсгийн хооронд уялдаа тогтоох буюу харгалзуулах үйлдэл хийгдэнэ. Энэхүү уялдааг GIZA++ болон EM алгоритмууд ашиглан автоматаар тооцоолж, Латин үсгүүдийг Кирилл үсгүүдтэй харгалзуулдаг. Ингэснээр олон янзын тэмдэгт хослолуудыг тохирох авиатай холбох боломж бүрдэнэ.

Харгалзуулах үйлдлийн үр дүнд үүссэн *source-target* тэмдэгтийн хослолын давтамж, нөхцөлт магадлалд $P(target|source)$ үндэслэн статистик хөрвүүлэлтийн загвар суралцдаг ба хөрвүүлэлтийн хамгийн өндөр магадлалтай дүрмүүдийг автоматаар тодорхойлно. Энэхүү загвар нь *phrase-based statistical machine translation* (СМО) аргын нэгэн хувилбар бөгөөд үгсийн дотоод авиа зүйг баримталж ажилладаг.

Загварыг сургаж дууссаны дараа шинэ Латин үсэг ашиглан галиглагдсан бичвэрүүд орж ирэхэд өмнө тогтоогдсон дүрэм, магадлалуудын үндсэн дээр хөрвүүлэлт хийгдэнэ. Тухайн оролтын сегментүүдийг хамгийн тохиромжтой Кирилл тэмдэгтүүдэд хөрвүүлж, эцсийн гаралтыг бүрдүүлдэг. Энэ нь авиа зүйн хувьд оновчтой, орчуулгын чанар сайтай үр дүнг гаргах нөхцөлийг бүрдүүлдэг бөгөөд бага нөөцтэй нөхцөлд ч үр ашигтайгаар ажилладаг шийдэл юм. Энэүү загварын бүтэц, үйл ажиллагааны зарчмыг Зураг 2-т харуулав.



Зураг 2. ТМ загварын бүтэц болон ажиллах зарчим

IV. Туршилт, үр дүн

A. *Загваруудын харьцуулалт*

Бид тэгшитгэл (1)-д үзүүлсэн үгийн түвшний нарийвчлалын дагуу бүх загварыг үнэлсэн.

$$AccWord = \frac{\text{Зөв нормалчилсан үгийн тоо}}{\text{Бүх үгийн тоо}} \quad (1)$$

Туршилтын үр дүнг Хүснэгт IV-т үзүүлэв. Манай тестийн өгөгдөлд OOV үгс их хэмжээгээр орсон тул үгийн түвшний seq2seq загвар эдгээр үгсийг хэвийн болгоход хүндрэлтэй байж, харьцангуй бага нарийвчлал (47.50%) үзүүлсэн. Үүний эсрэгээр, OOV үгсийн хувьд илүү тэсвэртэй байх чадвартай тэмдэгтийн түвшний seq2seq болон ТМ загваруудыг ашигласан бөгөөд эдгээр нь харгалзан 74.86% болон 80.52% нарийвчлалтайгаар хамгийн өндөр гүйцэтгэл үзүүлсэн. Ялангуяа ТМ загвар нь маш бага хэмжээний өгөгдөл дээр сурсан ч хамгийн өндөр нарийвчлал үзүүлсэн нь, өгөгдлийн нөөц хязгаарлагдмал орчинд үр дүнтэй шийдэл болохыг харуулсан ба хэлний дүрэмд суурилсан хандлага нь чухал үүрэг гүйцэтгэж байгааг харуулж байна. Дараагийн алхамд, бид үгийн түвшний seq2seq загварын OOV үгтэй холбоотой асуудлыг шийдвэрлэх зорилгоор түүнийг тэмдэгтийн түвшний seq2seq болон ТМ загваруудтай хослуулсан. Хосолсон эдгээр аргууд нь үгийн түвшний seq2seq загварын гүйцэтгэлийг сайжруулсан ч, тус бүрээрээ ашигласан тэмдэгтийн түвшний seq2seq болон ТМ загварын гүйцэтгэлийг давж чадаагүй.

ХҮСНЭГТ IV. ЯЛГААТАЙ ЗАГВАРУУДЫГ ТЕСТИЙН ӨГӨГДӨЛ ДЭЭР ТУРШСАН ҮР ДҮН

#	*Загварууд*	*Үгийн түвшний нарийвчлал*
1	Word-level seq2seq	47.50%
2	Char-level seq2seq	**74.86%**
3	ТМ	**80.52%**
4	Word-level seq2seq+ТМ	73.12%
5	Word-level seq2seq+char level seq2seq	67.11%

Бид мөн бүх загваруудын IV болон OOV үгсийг хэвийн болгох чадварыг Хүснэгт V-д дэлгэрэнгүй үнэлж үзсэн. Үгийн түвшний seq2seq загвар нь IV үгс дээр сайн нарийвчлалтай байсан ч OOV үгсийг боловсруулах боломжгүй байсан. Тэмдэгтийн түвшний seq2seq загвар нь IV үгсийн хувьд маш өндөр нарийвчлалтай, харин OOV үгсийн нарийвчлал харьцангуй бага боловч тодорхой хэмжээнд стандарт хэлбэрт хөрвүүлэх боломжтой байгааг туршилтын үр дүнгээс харж болно. TM загвар нь IV болон OOV үгс дээр харьцангуй сайн гүйцэтгэлтэй байлаа. Энэ нь өгөгдлийн хомсдол болон шуугиантай галиглалыг хэвийн хэлбэрт шилжүүлэхэд илүү үр дүнтэй гэдгийг харуулсан. Хосолсон загварууд (Word-level seq2seq + TM, Word-level seq2seq + Char-level seq2seq) нь анхны үгийн түвшний seq2seq загварын OOV үгсийг боловсруулах боломжийг олгож, гүйцэтгэлийг сайжруулсан ч тэмдэгтийн түвшний seq2seq загвар болон TM загваруудыг дангаар нь ашигласан гүйцэтгэлийг давж чадаагүй.

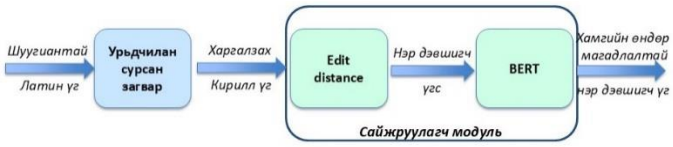
ХҮСНЭГТ V. ТЕСТИЙН ӨГӨГДӨЛ ДЭХ IV БА OOV ҮГСИЙН НАРИЙВЧЛАЛ

#	Загварууд	Үгийн түвшний нарийвчлал	
		IV үг	OOV үг
1	Word-level seq2seq	81.44%	-
2	Char-level seq2seq	**94.74%**	**47.04%**
3	TM	**94.12%**	**61.47%**
4	Word-level seq2seq+TM	81.44%	61.47%
5	Word-level seq2seq+char level seq2seq	81.23%	47.04%

B. *Сайжруулагч модуль*

Бидний хэрэгжүүлж буй сайжруулагч загварын үйл ажиллагааг Зураг 3-д дэлгэрэнгүй үзүүлсэн байна. Энэхүү сайжруулагч модуль нь урьдчилан сурсан загварын гаралтыг стандарт Кирилл үгс агуулсан толь бичгийн үгнүүдтэй харьцуулж, хэрэв толь бичигт байгаа бол тухайн үгийг зөв хөрвүүлэгдсэн гэж үздэг. Харин толь бичигт байхгүй үгсийг загварын буруу хөрвүүлэлт гэж тооцон, эдгээр үгсийг *Levenshtein edit distance* аргаар 1-2 тэмдэгтээр засварлаж, толь бичгийн үгтэй хамгийн ойр тохирох нэр дэвшигч үгсийг гаргадаг. Хамгийн оновчтой нэр дэвшигчийг сонгохын тулд бид тухайн үг өгүүлбэрт контекстийн хувьд хэр тохирохыг үнэлэх зорилгоор BERT загварын *next token scoring* аргыг ашигласан. Өөрөөр хэлбэл, нэр дэвшигч үгсийг тухайн байрлалын өмнөх бүх контекстэд үндэслэн үнэлж, эдгээрийн дараа ямар үг орох нь хамгийн өндөр магадлалтай байгааг BERT загвар тооцоолно. Бид монгол хэлний текст өгөгдөл дээр урьдчилан сургагдсан, *uncased* (том, жижиг үсгийн ялгаа гаргадаггүй) төрлийн BERT-base загварыг [16] ашигласан. Бидний санал болгож буй энэхүү сайжруулагч модулийн тусламжтайгаар загварын гүйцэтгэлийг үр дүнтэй сайжруулж, түүний нарийвчлал болон найдвартай байдлыг нэмэгдүүлэх боломжийг олгодог.

Өмнөх туршилтын үр дүнд үндэслэн, бид хамгийн өндөр гүйцэтгэлтэй байсан үгийн түвшний Seq2Seq болон TM загварууд дээр энэхүү сайжруулагч модулийг хэрэгжүүлэхээр төлөвлөж, эдгээр загваруудын хөрвүүлэлтийн нарийвчлалыг цаашид нэмэгдүүлэх боломжийг судалсан.



Зураг 3. Сайжруулагч модулийн үйл ажиллагааны бүдүүвч

Сайжруулагч модулийг TM буюу дүрмийн үндэслэлтэй хөрвүүлэлтийн арга, Seq2Seq загвар болон эдгээрийг хослуулсан гибрид арга дээр тус тус хэрэглэснээр, бүх тохиолдолд хөрвүүлэлтийн гүйцэтгэл мэдэгдэхүйц сайжирч байгааг Хүснэгт VI-аас харж болно. Тухайлбал, зөвхөн TM загвар ашигласан тохиолдолд 80.52% байсан гүйцэтгэл сайжруулагч модулийг нэмж хэрэгжүүлснээр 84.90% болж өссөн байна. Үүнтэй адилаар Seq2Seq загварын хувьд 74.86%-иас 81.59% болж, харин гибрид загварын хувьд хамгийн өндөр үзүүлэлттэйгээр 83.88%-иас 86.53% болж сайжирсан нь сайжруулагч модуль нь аль ч төрлийн загварын гүйцэтгэлд эерэг нөлөө үзүүлж байгааг харуулж байна.

Мөн IV болон OOV үгсийн тусдаа нарийвчилсан үзүүлэлтүүдийг авч үзвэл, сайжруулагч модулийг ашигласнаар IV үгс дээр бүх загварын хувьд 94%-иас дээш гүйцэтгэлтэй байгаа нь үндсэн хөрвүүлэлтийн үр дүнг хадгалах чадвартайг илтгэж байна. Харин хамгийн их сорилт үүсгэдэг OOV үгс дээр үзүүлэлт илүү ялгаатай харагдаж байна. TM загварын хувьд OOV үг дээр 71.90%, Seq2Seq загварын хувьд 65.56% байгаа бол эдгээрийг хослуулсан TM+Seq2Seq гибрид загварын хувьд 75.50%-д хүрсэн нь сайжруулагч модуль нь OOV үгсийн зөв хөрвүүлэлтийг илүү үр дүнтэйгээр сайжруулж байгааг баталж байна.

Эдгээр үр дүнгээс харахад, сайжруулагч модулийг нэмэлт давхарга болгон ашигласнаар зөвхөн үндсэн загварын гүйцэтгэлийг нэмэгдүүлэхээс гадна, мэдлэгт суурилсан дүрэм болон гүн сургалтын аргуудын сул талуудыг харилцан нөхөх, хөрвүүлэлтийн нарийвчлалыг бүх түвшинд (IV ба OOV үгс дээр) сайжруулах боломжтойг харуулж байна. Тиймээс уг модуль нь олон төрлийн хөрвүүлэлтийн системд ерөнхийлсөн байдлаар ашиглагдах, өргөтгөх боломжтой үр дүнтэй шийдэл болохыг дүгнэж болно.

ХҮСНЭГТ VI. САЙЖРУУЛАГЧ МОДУЛИЙГ ХЭРЭГЖҮҮЛСНИЙ ДАРААХ ТЕСТИЙН ӨГӨГДӨЛ ДЭЭРХ ҮР ДҮН

#	Загварууд	Ерөнхий нарийвчлал		Үгийн түвшний нарийвчлал	
		CM-гүй	CM-тай	IV үг	OOV үг
1	TM + CM*	80.52%	84.90%	94.22%	71.90%
2	Seq2Seq+ CM	74.86%	81.59%	94.74%	65.56%
3	TM+Seq2Seq+CM	**83.88%**	**86.53%**	**94.42%**	**75.50%**

CM - сайжруулагч модуль*

V. ДҮГНЭЛТ

Энэхүү судалгаанд бид үг болон тэмдэгтийн түвшний Seq2Seq, дүрмийн үндэслэлтэй TM, мөн тэдгээрийн хослолыг ашигласан гибрид хөрвүүлэлтийн загваруудыг боловсруулан харьцуулж судаллаа. Судалгааны үр дүнд, OOV үгс орсон нөхцөлд тэмдэгтийн түвшний Seq2Seq болон TM загварууд нь илүү үр дүнтэй байгааг баталлаа. Ялангуяа, TM загвар нь өгөгдлийн хэмжээ бага байсан ч хамгийн өндөр гүйцэтгэл үзүүлж, дүрмийн мэдлэгт суурилсан арга нь шуугиантай галиглагдсан текстийг хөрвүүлэхэд онцгой үүрэг гүйцэтгэж байгааг харуулсан.

Мөн бид сайжруулагч модулийг эдгээр загваруудад нэмэлт давхарга болгон хэрэгжүүлснээр хөрвүүлэлтийн нарийвчлал, найдвартай байдлыг бүх түвшинд (IV болон OOV үгс) мэдэгдэхүйц нэмэгдүүлж чадсан. Сайжруулагч модуль нь толь бичиг, edit distance, BERT загвар зэрэг аргуудыг хослуулсан бөгөөд үндсэн загваруудын гүйцэтгэлд мэдэгдэхүйц эерэг нөлөө үзүүлсэн.

Үр дүнгээс дүгнэхэд, сайжруулагч модулийг ерөнхийлсөн байдлаар олон төрлийн хөрвүүлэлтийн системд ашиглах боломжтой бөгөөд үгийн болон тэмдэгтийн түвшний хөрвүүлэлтийн сул талыг нөхөж, нарийвчлалыг цогцоор нь сайжруулах үр дүнтэй шийдэл болохыг харуулж байна.

МЭДЭЭЛЭЛ

Энэхүү судалгааг KOICA (Солонгосын Олон Улсын Хамтын Ажиллагааны Байгууллага)-аас "Монгол дахь ШУТИС-ийн Мэдээлэл, харилцаа холбооны технологийн сургуулийн чадавхийг бэхжүүлэх төсөл"-өөр (Гэрээ P2019- 00124) дэмжин ажиллав.

АШИГЛАСАН МАТЕРИАЛ

[1] S. Tiun and S. N. A. Noor Ariffin, "Rule-based Text Normalization for Malay Social Media Texts," Int. J. Adv. Comput. Sci. Appl., vol. 11, pp. 156-162, Nov. 2020. DOI: 10.14569/IJACSA.2020.0111021.

[2] A. Aw, M. Zhang, J. Xiao, and J. Su, "A phrase-based statistical model for SMS text normalization," in Proc. COLING/ACL 2006 Main Conf. Poster Sessions, Sydney, July 2006, pp. 33-40.

[3] M. A. Saloot, N. Idris, and A. Aw, "Noisy text normalization using an enhanced language model," in Proc. Int. Conf. Artif. Intell. Pattern Recognit., Kuala Lumpur, Malaysia, 2014.

[4] N. Kalchbrenner and P. Blunsom, "Recurrent continuous translation models," in *Proc. 2013 Conf. Empirical Methods Natural Lang. Process.*, Seattle, WA, USA, 2013, pp. 1700–1709.

[5] K. Cho, B. van Merriënboer, Ç. Gülçehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proc. 2014 Conf. Empirical Methods Natural Lang. Process.*, Doha, Qatar, 2014, pp. 1724–1734.

[6] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, Montreal, Canada, 2014, pp. 3104–3112.

[7] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in *Proc. 3rd Int. Conf. Learn. Represent.*, San Diego, CA, USA, 2015.

[8] M. Luong, H. Pham, and C. D. Manning, "Effective approaches to attention-based neural machine translation," in *Proc. 2015 Conf. Empirical Methods Natural Lang. Process.*, Lisbon, Portugal, 2015, pp. 1412–1421.

[9] T. Ikeda, H. Shindo, and Y. Matsumoto, "Japanese text normalization with encoder-decoder model," in Proc. 2nd Workshop on Noisy User-generated Text (W-NUT), Osaka, Japan, Dec. 2016.

[10] Z. Kozhirbayev and Z. Yessenbayev, "Kazakh text normalization using machine translation approaches," CEUR Workshop Proc., vol. 2780, pp. 115-122, 2020.

[11] I. Lourentzou, K. Manghnani, and C. X. Zhai, "Adapting sequence to sequence models for text normalization in social media," in Proc. 13th Int. AAAI Conf. Web Social Media, Association for the Advancement of Artificial Intelligence, Apr. 2019.

[12] Z. Byambadorj, R. Nishimura, A. Ayush, and N. Kitaoka, "Normalization of Transliterated Words Using Seq2Seq Model with Spell Checker," in Proc. Conference Name, Mar. 2020.

[13] https://mn.wiktionary.org

[14] Mongolian Agency for Standardization and Metrology, “MNS 5217:2012 - Rules of Mongolian Script Spelling,” Mongolian national standard, 2012.

[15] M. Abadi et al., "TensorFlow: A system for large-scale machine learning," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)*, 2016, pp. 265–283.

[16] Tugstugi, “BERT Base Mongolian Uncased,” Pre-trained model. Accessed [2025.03.22]. [Online]. Available: https://huggingface.co/tugstugi/bert-base-mongolian-uncased

Balancing Retrieval and Forgetting: A Comparative Analysis of RAG and Machine Unlearning Hallucination For Hallucination

Taivanbat Bayarsaikhan

National University of Mongolia

School of Information Technology and Electronics

22B1NUM6769

Altangerel Chagnaa

National University of Mongolia

School of Information Technology and Electronics

altangerel@num.edu.mn

Abstract- **This paper explores the interplay between information retrieval and intentional forgetting in large language models, specifically comparing GPT-4o augmented with Retrieval-Augmented Generation (RAG) to a Mongolian BERT model. Using data from the National University of Mongolia Open Data repository, we evaluate the effectiveness of top-k retrieval mechanisms with k values of 50 and 75 and analyze appropriate unlearning methodologies for the Mongolian BERT model. Our experiments demonstrate that while RAG enhances contextual relevance in GPT-4o, the Mongolian BERT model benefits most from a gradient-ascent unlearning approach that respects the morphological complexity of the Mongolian language. Results suggest that cultural and linguistically adapted models require specialized approaches to both retrieval and forgetting mechanisms, with implications for developing more responsive and ethically aligned AI systems in diverse linguistic contexts.**

Keywords-Retrieval-Augmented Generation, Machine Unlearning, Mongolian BERT, GPT-4o, Top-k Retrieval, Gradient Ascent

I. INTRODUCTION

The tension between remembering and forgetting represents a fundamental challenge in artificial intelligence systems. As large language models (LLMs) continue to advance, two competing imperatives have emerged: enhancing models with improved retrieval capabilities for greater contextual awareness and developing robust unlearning mechanisms to selectively remove information when necessary for privacy, accuracy, or ethical considerations.

Retrieval-Augmented Generation (RAG) has emerged as a powerful approach to enhance LLMs by providing them with access to external knowledge repositories, enabling more accurate, up-to-date, and verifiable outputs. Conversely, machine unlearning techniques seek to selectively remove specific information from trained models—a capability increasingly important considering regulatory frameworks like the "right to be forgotten" and the need to eliminate outdated or erroneous information.

This research examines these dual aspects through a comparative analysis of two distinct approaches:

A. GPT-4o enhanced with RAG using top-retrieval (k=50, k=75) against a private knowledge base from the National University of Mongolia - an award-winning approach that secured first place in a hackathon jointly organized by XacBank and the National University of Mongolia

B. A Mongolian BERT model with various unlearning algorithms applied to the same dataset

The Mongolian language context provides a particularly interesting case study due to its agglutinative morphology and unique cultural references, presenting challenges for both retrieval precision and targeted unlearning. Through this comparative analysis, we aim to identify optimal approaches for balancing retrieval enhancement and

forgetting capabilities in language models, with particular attention to linguistically diverse applications.

II. BACKGROUND AND RELATED WORK

A. *Retrieval-Augmented Generation (RAG)*

Retrieval-Augmented Generation (Lewis et al., 2020) combines neural retrievers with sequence-to-sequence models to enhance generation by accessing relevant documents from an external knowledge base. Originally conceived as an end-to-end trainable model, RAG has evolved into a flexible framework that can be integrated with various pre-trained language models (Gao et al., 2023).

The top-k retrieval mechanism limits the retrieved documents to the k most relevant documents determined by a similarity metric. This approach represents a crucial design choice that balances computational efficiency with information comprehensiveness. Recent work by Shi et al. (2023) demonstrates that appropriate calibration of k values significantly impacts performance across different tasks and domains.

B. *Machine Unlearning*

Machine unlearning, introduced by Cao & Yang (2015), aims to remove the influence of specific training examples from a trained model without complete retraining. Several approaches have emerged:

- SISA (Sharded, Isolated, Sliced, Aggregated) (Bourtoule et al., 2021): Divides training data into shards, trains models independently on each shard, and aggregates predictions, enabling efficient unlearning by retraining only affected shards.
- Gradient Ascent Unlearning (Graves et al., 2021): Updates model parameters in the direction that increases loss on targeted examples, effectively teaching the model to "forget" specific data.
- Amnesiac Unlearning (Graves et al., 2022): Identifies and removes the effects of specific examples by leveraging influence functions to approximate parameter changes.
- Approximate Unlearning (Sekhari et al., 2021): Employs privacy mechanisms like differential privacy to limit the influence of any single training example.

C. *Mongolian BERT and Language Modeling*

The Mongolian language presents unique challenges for NLP due to its agglutinative morphology, vowel harmony, and Cyrillic script adaptation. Mongolian BERT (Erdenebileg et al., 2022) represents a significant advancement in Mongolian language processing, adapting the BERT architecture to the language's specific characteristics.

Prior work on Mongolian language models has highlighted challenges including word segmentation complexity, resource limitations, and domain adaptation issues (Gunchinish et al., 2020). These challenges extend to both retrieval precision and unlearning effectiveness, as morphological variations can complicate exact matching and parameter influence tracing.

III. METHODOLOGY

A. *Datasets*

We utilized the National University of Mongolia Open Data repository, which contains approximately 1.2 million Mongolian text documents spanning academic articles, research papers, educational materials, and cultural texts. This dataset was processed to create:

1. A retrieval corpus for the RAG implementation with GPT-4o
2. A training dataset for fine-tuning the Mongolian BERT model
3. A set of 5,000 documents designated for unlearning experiments

The dataset was pre-processed using specialized Mongolian language tokenization tools, with attention to stemming and morphological analysis to improve retrieval and unlearning precision.

B. *GPT-4o with RAG Implementation*

Our implementation integrated GPT-4o with a retrieval mechanism using the following architecture, which won first place in a prestigious hackathon organized by XacBank in collaboration with the National University of Mongolia. The same architecture was subsequently adapted and applied to healthcare data as part of a datathon organized by the Ministry of Health of Mongolia:

1. *Document Indexing:* The Mongolian corpus was embedded using multilingual sentence transformers fine-tuned on Mongolian text pairs.
2. *Vector Database:* Embeddings were stored in a FAISS vector database optimized for top-k retrieval.
3. *Query Processing*: User queries were embedded using the same embedding model and used to retrieve the top-k most similar documents.
4. *Augmented Generation:* Retrieved documents were concatenated with the original query and provided to GPT-4o as context.

We experimented with k=50 and k=75 for the top-k retrieval to evaluate the impact of retrieval breadth on generation quality.

C. *Mongolian BERT Configuration*

We utilized a pre-trained Mongolian BERT model (Erdenebileg et al., 2022) and further fine-tuned it on our dataset. The model specifications included:

- 12 transformer layers
- 768 hidden dimensions
- 12 attention heads
- 4M parameters
- Vocabulary size of 32,000 tokens with specialized Mongolian subword units

Fine-tuning was performed using a masked language modeling objective with additional next sentence prediction on the National University of Mongolia dataset.

D. *Unlearning Algorithms Implementation*

We implemented and evaluated four unlearning approaches for the Mongolian BERT model:

1. *SISA (Sharded, Isolated, Sliced, Aggregated):*

- Divided the training data into 20 shards
- Trained independent models on each shard
- Implemented an aggregation layer for combined predictions

2. *Gradient Ascent Unlearning:*

- Identified parameters most influenced by target documents
- Applied targeted gradient updates to increase loss on forgotten examples
- Implemented regularization to preserve general performance

3. *Approximate Differential Privacy-Based Unlearning:*

- Implemented a DP-SGD training regime
- Leveraged the privacy properties to limit memorization
- Applied selective noise addition to targeted parameter subspaces

4. Amnesiac Unlearning:

- Tracked and stored gradient updates during training
- Implemented selective reversal of gradient updates from forgotten examples
- Applied fine-tuning on retained data to restore performance

E. Evaluation Metrics

We evaluated both systems using the following metrics:

1. *Retrieval Precision and Recall:* Measured the relevance of retrieved documents
2. *Generation Quality:* BLEU, ROUGE, and BERTScore against reference answers
3. *Unlearning Effectiveness*: Measured using membership inference attack success rates
4. *Unlearning Efficiency*: Computational cost and time required for unlearning
5. *Retention of Unaffected Knowledge*: Performance on a test set excluding unlearning targets

IV. RESULT

A. *RAG Performance with Different k Values*

The performance of GPT-4o with RAG varied significantly based on the k value used in retrieval:

TABLE I. Performance of GPT-4o with different top-k values

Metric	GPT-4o	GPT-4o + RAG (k=50)	GPT-4o + RAG (k=75)

	(no RAG)		
BLEU Score	0.342	0.418	0.427
ROUGE-L	0.389	0.456	0.461
BERTScore	0.712	0.768	0.773
Factual Acuracy(%)	0.712	0.768	0.773
Domain Relevance (%)	76.3	89.2	90.1
	68.7	87.4	88.2

The results demonstrate that RAG significantly improved performance across all metrics, with k=75 offering marginally better results than k=50. However, the performance gain from k=50 to k=75 was relatively small (approximately 2-3% improvement) compared to the computational cost increase (approximately 40% longer inference time).

The most substantial improvements were observed in domain-specific knowledge questions and recent factual information queries, where the base GPT-4o model showed limitations due to its training cutoff date and general knowledge scope.

Metric	SISA	Gradient Ascent	Approximate DP	Amnesiac
Unlearning Success Rate (%)	92.4	96.8	87.3	91.2
Model Performance Retention (%)	89.7	94.2	92.8	88.3
Computational Efficiency (hours)	24.5	8.3	12.7	16.2
Membership Inference Defense (%)	88.6	94.7	92.5	86.9
Morphological Variation Handling (%)	76.8	92.3	71.2	81.4

B. *Unlearning Algorithm Performance on Mongolian BERT*

We evaluated the four unlearning algorithms on the Mongolian BERT model with the following results:

TABLE II. Unlearning Algorithm Performance on Mongolian BERT

Gradient Ascent Unlearning demonstrated the best overall performance, particularly in handling morphological variations—a critical factor for the agglutinative Mongolian language. The approach showed superior unlearning success rates while maintaining strong performance on unaffected knowledge.

SISA, while effective, required substantially more computational resources due to the need to maintain multiple model shards. The Approximate DP method provided strong privacy guarantees but struggled with morphological variations in the language, which affected precise unlearning.

C. *Comparative Analysis: RAG vs. Unlearning Approaches*

Figure 1: Performance Comparison Across Different Tasks

[Note: Figure would show a radar chart comparing performance dimensions of both approaches]

The comparative analysis revealed complementary strengths:

1. *GPT-4o with RAG excelled in:*

- Retrieval of specific cultural and academic information
- Integration of domain knowledge with general reasoning
- Handling of ambiguous queries through contextual enhancement

2. *Mongolian BERT with Gradient Ascent Unlearning excelled in:*

- Precise removal of targeted information
- Preservation of grammatical and morphological knowledge
- Efficient adaptation to changing information requirements

V. DISCUSSION

A. *Optimal RAG Configuration for Mongolian Language Applications*

Our results indicate that for Mongolian language applications, RAG with k=50 represents an optimal balance between retrieval comprehensiveness and computational efficiency.

The limited additional benefit of k=75 suggests diminishing returns beyond certain retrieval breadth.

The key factors influencing RAG performance in this context were:

1. *Embedding Quality:* The multilingual embedding model required specific fine-tuning to capture Mongolian semantic relationships effectively
2. *Document Granularity:* Chunking documents into semantically coherent segments improved retrieval precision
3. *Context Integration*: GPT-4o demonstrated stronger ability to incorporate retrieved context compared to smaller models tested in preliminary experiments

B. *Gradient Ascent as Optimal Unlearning Approach for Mongolian BERT*

The superior performance of Gradient Ascent Unlearning for Mongolian BERT can be attributed to several factors:

1. *Parameter Specificity*: The approach allowed for targeted modification of parameters most associated with specific examples
2. *Morphological Robustness*: Gradient Ascent demonstrated better handling of morphological variations of the same concept
3. *Computational Efficiency:* The approach required significantly less computational resources than SISA while achieving better results
4. *Performance Preservation:* Careful regularization during unlearning helped maintain model performance on unaffected knowledge

C. *Balancing Retrieval and Forgetting*

Our research highlights the complementary nature of enhanced retrieval and selective forgetting. While RAG provides flexibility by keeping knowledge external to the model, unlearning offers a pathway to modify knowledge embedded within model parameters.

The ideal approach appears to be context-dependent:

1. *For rapidly changing information domains*: RAG provides advantage through easily updated external knowledge bases
2. *For privacy-sensitive applications:* Unlearning offers stronger guarantees of information removal
3. *For resource-constrained environments:* Gradient-based unlearning may be more practical than maintaining complex retrieval infrastructure

VI. CONCLUSION

This research demonstrates that balancing retrieval enhancement and selective forgetting capabilities is essential for developing responsible and effective language models. Our comparative analysis of GPT-4o with RAG and Mongolian BERT with unlearning algorithms reveals that:

1. RAG significantly enhances domain-specific performance of general models like GPT-4o, with k=50 offering an optimal efficiency-performance balance in the Mongolian language context.
2. Gradient Ascent Unlearning emerges as the most effective approach for the Mongolian BERT model, offering superior performance in handling the language's morphological complexity while maintaining general knowledge.
3. The choice between external retrieval and parameter modification through unlearning should be guided by application-specific requirements regarding information freshness, privacy guarantees, and computational constraints.

ACKNOWLEDGMENTS

We thank the National University of Mongolia for providing the Open Data repository used in this research. We also extend our gratitude to XacBank for co-organizing the hackathon where our RAG implementation was recognized with the first-place award, and to the Ministry of Health of Mongolia for organizing the datathon that allowed us to further validate our approach with healthcare data.

REFERENCES

[1] Bourtoule, L., Chandrasekaran, V., Choquette-Choo, C. A., Jia, H., Travers, A., Zhang, B., ... & Papernot, N. , Machine unlearning. In IEEE Symposium and Security and Privacy, 2021.

[2] Y. &. Y. J. Cao,). Towards making systems forget with machine unlearning. In 2015 IEEE Symposium on Security and Privacy (pp. 463-480)., 2015.

[3] L. M. X. J. &. C. J. Gao, "Precise zero-shot dense retrieval without relevance labels. In

Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing.," 2023.

[4] A. N. S. B. G. &. M. D. Graves, " Amnesiac machine learning. arXiv preprint arXiv:2010.10981," 2021.

[5] S. B. E. &. S. D. Gunchinish, " Challenges in Mongolian text processing. In Proceedings of the 12th Language Resources and Evaluation Conference .," 2020, pp. 6214-6221.

[6] P. P. E. P. A. P. F. K. V. G. N. .. &. K. D. Lewis, " Retrieval-augmented generation for knowledge-intensive NLP tasks. In Advances in Neural Information Processing Systems," 2020, pp. 9459-9474.

[7] A. A. J. K. G. &. S. A. T. Sekhari, " Remember what you want to forget: Algorithms for machine unlearning. In Advances in Neural Information Processing Systems," 2021, pp. 15076-15087.

[8] W. Z. R. L. H. &. H. X. Shi, "The impact of retrieval k on RAG-based query systems. In Findings of the Association for Computational Linguistics," 2023.

[9] N. P. H. &. Z. X. Erdenebileg, MonBERT: Mongolian language model using BERT architecture. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (pp. 7333-7338)., 2022.

Flickr30k-MN: Англи-Монгол зургийн тайлбарын өгөгдөл, орчуулгын үнэлгээ

Очирбатын Билгүүн
Мэдээлэл, компьютерийн ухааны тэнхим
Монгол улсын их сургууль
Улаанбаатар, Монгол улс
bilguun.o@num.edu.mn

Чагнаагийн Алтангэрэл
Мэдээлэл, компьютерийн ухааны тэнхим
Монгол улсын их сургууль
Улаанбаатар, Монгол улс
altangerel@num.edu.mn

***Хураангуй*— Энэхүү судалгааны ажлаар зургийг Монгол хэлээр ойлгох, хайх, тайлбарлах машин сургалтын загвар хөгжүүлэхэд шаардлагатай өгөгдөл бэлтгэхээр зорьсон. Бид Flickr30k хэмээх 31 мянган зурагтай, зураг тус бүр 5 англи тайлбартай өгөгдлийг 1. машин орчуулга буюу Google Translate, 2. мультимодал LLM (хэлний том загвар) болох Gemini 2.0 Flash-н тусламжтай монгол хэл рүү автоматаар орчуулсан. Орчуулгын чанарыг үнэлэх зорилгоор хүмүүсийн дунд веб аппликейшнд суурилсан судалгаа явуулж, оролцогчдоос орчуулгын аль хувилбар нь тухайн зургийг илүү оновчтой тайлбарласан байгааг нь сонгохыг хүссэн. Судалгаанд оролцогчдын нийт саналын 63.8% нь хоёрдугаар хувилбар болох LLM ашигласан орчуулгыг илүүд үзсэн бол 36.2% нь машин орчуулгын системээр орчуулсан тайлбарыг сонгосон. Энэхүү үр дүнгээс зургийн тайлбарыг монгол хэл рүү орчуулахад мультимодал LLM нь хүнд илүү ойлгомжтой, чанартай үүсгэж байгааг харж болохоор байна. Flickr30k өгөгдлийн зураг болон англи тайлбараас бидний автомат орчуулгын аргаар үүсгэсэн монгол тайлбар нь цаашдын монгол хэл дээрх зураг, бичвэрийн судалгаа, зураг ойлгох, хайлт хийх хиймэл оюун, машин сургалтын загвар хөгжүүлэхэд чухал ач холбогдолтой нөөц болно гэж үзэж байна.**

Түлхүүр үг— Монгол хэл, англи хэл, зургийн тайлбар, машин орчуулга, орчуулгын үнэлгээ, Google Translate, Gemini, LLM, хэлний том загвар, Flickr30k

I. Удиртгал

Сүүлийн жилүүдэд хиймэл оюун ухаан, тэр дундаа компьютерын хараа (computer vision) болон хэл боловсруулалт (natural language processing)-ын салбарт асар хурдацтай хөгжил дэвшил гарч байна. Эдгээр салбарын нэг чухал огтлолцол бол зураг тайлбарлах (image captioning) буюу компьютерт зургийн агуулгыг ойлгуулж, түүнийг хүний хэлээр тайлбарлах чадварыг бий болгох явдал юм. Зураг тайлбарлах технологи нь харааны бэрхшээлтэй иргэдэд мэдээлэл хүргэх, зурган контентын хайлтыг сайжруулах, роботын орчин тойрноо ойлгох чадварыг нэмэгдүүлэх зэрэг олон төрлийн ач холбогдолтой.

Өндөр чанартай зураг тайлбарлах загвар бүтээхийн тулд их хэмжээний, сайн чанарын "зураг-тайлбар" хосолсон бүхий өгөгдөл зайлшгүй шаардлагатай байдаг. Англи хэл болон бусад өргөн хэрэглэгддэг хэлнүүдийн хувьд COCO (Common Objects in Context) [1], Flickr30k [2] зэрэг томоохон хэмжээний, хүний оролцоотой цуглуулсан өгөгдөл өргөн ашиглагдаж байна. Харин Монгол хэл шиг харьцангуй бага нөөцтэй хэлнүүдийн хувьд ийм төрлийн мэдээллийн сан хомс байгаа нь тухайн хэл дээрх зураг тайлбарлах технологийн хөгжилд томоохон сорилт болж байна.

Энэхүү асуудлыг шийдвэрлэх нэг арга зам бол англи хэл дээрх өгөгдлийг зорилтот хэл рүү орчуулах явдал юм. Гэвч машин орчуулгын чанар нь ашиглаж буй арга, хэлний онцлогоос хамааран харилцан адилгүй байдаг. Ялангуяа Монгол хэлний үгийн баялаг сан, өвөрмөц хэл зүйн бүтэц нь автомат орчуулгын үед алдаа гарах, утгын зөрүү үүсгэх, хэллэг нь хиймэл сонсогдох зэрэг асуудлуудыг үүсгэдэг.

Иймд бид энэхүү судалгааны ажлаар дамжуулан Монгол хэл дээрх зураг тайлбарын хосолсон өгөгдлийг бүрдүүлэхийг зорьж, англи хэл дээрх Flickr30k өгөгдлийн 31 мянган зургийн тайлбарыг хоёр өөр автомат орчуулгын аргаар Монгол хэл рүү орчуулж, аль аргаар орчуулсан орчуулга нь илүү чанартай болохыг тодорхойлох зорилгоор хүний оролцоотой үнэлгээ хийсэн. Уг үнэлгээг веб аппликейшн ашиглан зохион байгуулж, оролцогчдоос аль орчуулга нь зургийг илүү сайн тодорхойлсон, хэр бодит сонсогдож байгаа, алдаагүй эсэхийг харгалзан сонголт хийхийг хүссэн.

II. Холбогдох судалгаа

Зураг ойлгох нь компьютерын хараа ба хэл боловсруулалтын салбарыг холбосон идэвхтэй судалгааны чиглэл юм. Сүүлийн үед зургаас тайлбар үүсгэх, зураг бичвэр хоёрын утга зүйн харьцуулалт хийх судалгаануудаас дурьдвал CLIP[3], BLIP[4] зэрэг машин сургалтын загварууд англи хэл дээр хийгдсэн байна. Энэ чиглэлийн судалгаа нь ихэвчлэн томоохон хэмжээний, тэмдэглэл хийгдсэн өгөгдөлд тулгуурладаг. Хамгийн өргөн ашиглагддаг өгөгдлийн хувьд:

COCO (Common Objects in Context) [1]: Өдөр тутмын орчин дахь объектуудыг харуулсан 330,000 гаруй зурагтай, зураг бүр 5 тайлбартай. Энэ нь зураг таних, сегментчлэх, тайлбарлах зэрэг олон төрлийн даалгаварт стандарт болсон.

Flickr30k [2]: Flickr сайтаас цуглуулсан 31,783 зурагтай, зураг бүр 5 тайлбар буюу нийт ялгаатай 158,915 тайлбартай. Зургууд нь олон төрлийн, бодит амьдралын агшнуудыг харуулдаг бөгөөд үүнийг бид үндсэн өгөгдлөөр сонгон авсан.

Visual Genome [5]: Зургийн бүсүүд, объектууд, тэдгээрийн хоорондын хамаарал, шинж чанаруудын талаарх дэлгэрэнгүй мэдээллийг агуулсан илүү баялаг бүтэцтэй өгөгдөл.

Эдгээр өгөгдлүүд нь голчлон англи хэл дээр бөгөөд бусад хэл дээрх өгөгдөл үүсгэх нэг түгээмэл арга бол машин орчуулгын тусламжтайгаар орчуулах явдал юм. Жишээлбэл, COCO болон Flickr30k өгөгдлийг герман хэл рүү орчуулж, тухайн хэл дээрх загваруудыг сургах, үнэлэхэд ашигласан судалгаанууд[6] бий.

Гэсэн хэдий ч машин орчуулга нь төгс биш бөгөөд орчуулгын чанар нь загварын архитектур, сургалтын өгөгдөл, хэлний онцлогоос ихээхэн хамаардаг. Сүүлийн үеийн нейрон сүлжээн, attention механизмд суурилсан машин орчуулгын (Neural Machine Translation - NMT) [7] загварууд нь уламжлалт статистик машин сургалтын (Statistical Machine Translation - SMT) загваруудаас илүү сайн үр дүн үзүүлж байгаа ч, хэлний хослол болон өгөгдлийн хэмжээнээс хамааран чанар нь харилцан адилгүй байдаг. Мөн сүүлийн үед хэлний том загвар (Large Language Model - LLM) гарч ирснээр заавал машин сургалтын загвар сургахгүйгээр LLM-г зааварчлах байдлаар машин орчуулга хийх хандлага нэмэгдсээр байна[8].

Монгол хэл дээрх зураг ойлгох чиглэлийн судалгаа харьцангуй эхэн шатандаа явж байгаа бөгөөд түүнд шаардлагатай томоохон өгөгдөл байхгүй байна. Зарим судалгааны хүрээнд англи хэл дээрх өгөгдлийг Google Translate зэрэг олон нийтэд нээлттэй орчуулгын хэрэгслүүдээр орчуулахыг оролдсон байж болох ч, орчуулгын чанарыг системтэйгээр үнэлж, сайжруулах, өөр өөр орчуулгын аргуудыг харьцуулсан судалгаа ховор байна.

III. Аргачлал

Судалгааны хүрээнд бид Flickr30k өгөгдлийн англи тайлбарыг машин орчуулгын Google Translate болон мультимодал хэлний том загвар болох Google Gemini 2.0 Flash[9]-н тусламжтай орчуулга хийсэн.

A. Англи тайлбарыг монгол руу орчуулах

Орчуулгын хоёр хувилбар үүсгэсэн ба эхний хувилбарыг Google Translate API-р орчуулж гаргасан. Google Translate орчуулгын үр дүнд бүх тайлбар асуудалгүй орчуулагдсан.

Хоёр дахь хувилбарыг Gemini 2.0 Flash загварт орчуулгын зааварчилгаа өгөх байдлаар англи тайлбар болон зургаас харгалзах монгол тайлбарыг үүсгүүлж авсан. LLM-р хүссэн даалгавраа гүйцэтгүүлэхийн тулд зааварчилгаа буюу prompt бичиж өгөх шаардлагатай бөгөөд энэ ажлын хүрээнд хэрэглэсэн prompt-г доор үзүүлэв (<> тэмдгээр хувьсах өгөгдлийг тэмдэглэв):

System: You are a native Mongolian speaker with a keen understanding of cultural nuances. You will receive an image and five English captions. Your mission is to create five authentic and natural-sounding Mongolian captions for the image, capturing the essence of the English captions while reflecting typical Mongolian phrasing. Output the captions in the same order as the English, one per line, hyphen at the beginning.

User: Write 5 captions for this image in Mongolian.
English captions:
<5 англи тайлбарыг зураасаар (-) эхлүүлэн жагсаана>
<хавсралт: тайлбарт харгалзах зураг>

Мөн LLM бичвэр үүсгэхдээ temperature параметрээр дараагийн үгийг санамсаргүй байдлаар сонгох хэмжээг тодорхойлдог бөгөөд энэ ажлаар тухайн параметрийн утгаг 0.5-аар авав.

LLM-н үр дүнд зааварчилгаанд өгсөн зураг болон англи тайлбаруудад харгалзах 5 ширхэг монгол тайлбар тус бүр нэг мөрөнд, зураасаар (-) эхэлсэн жагсаалт байдлаар гарч ирнэ. Энэ үйлдлийг бүх зураг тайлбарт давтан хэрэгжүүлж, нийт 158,740 тайлбар (31,744 зураг) бүрэн үүссэн[1]. Орчуулж чадаагүй/асуудалтай үлдсэн 175 тайлбарын 170 нь “prohibited content” алдаа, харин үлдсэн 5 нь таван ялгаатай зургийн нэг, нэг тайлбар бөгөөд шалгаж үзэхэд үндсэн өгөгдөл доторх англи тайлбар нь буруу, зурагтай тохиромжгүй байсан. Gemini загвар нь зураг болон бичвэр аль алиныг нь харж байгаа тул тэрхүү алдааг нь зассан тайлбар нэмж бичсэн байна. Засаж бичсэн таван тохиолдол нь англи болон Google Translate-р орчуулсан тайлбараас зөрсөн боловч зурагтай нийцэлтэй байсан. Гэхдээ энэхүү судалгааны хүрээнд тэдгээр тайлбаруудыг түр тооцож үзэхгүй алгассан болно.

Зураг 1. Flickr30k өгөгдлийн жишээ. 41-р зургийн 2-р тайлбар: **Англи тайлбар:** “*A man leans into a car to talk to the driver , as a man on a bicycle looks on .*”, **Google Translate-н орчуулсан тайлбар:** “*Унадаг дугуй унасан хүн машин руугаа тонгойж жолоочтой ярилцаж байна .*”, **Gemini-н үүсгэсэн тайлбар:** “*Нэг эрэгтэй жолоочтой ярилцахаар машинд тонгойж байгаа харагдаж байна, хажууд нь дугуйтай хүн ажиглаж байна.*”

B. Орчуулгын чанарын судалгаа

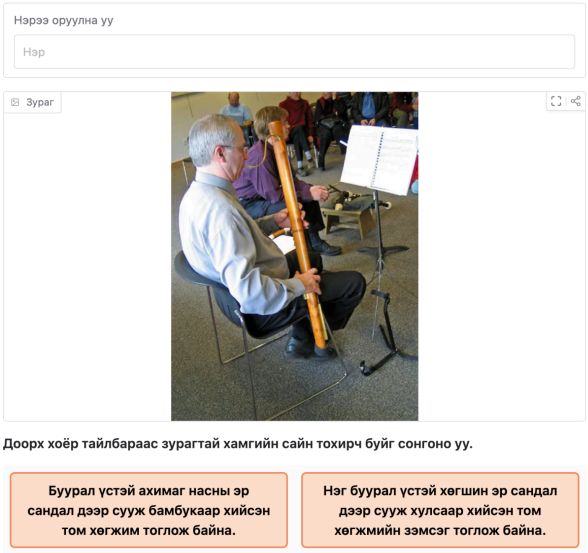
Google Translate болон Gemini загвараар үүсгэсэн орчуулгын хоёр хувилбарын чанарыг шалгах зорилгоор хүмүүсийн дунд веб аппликейшнд суурилсан судалгаа явуулсан[2]. Судалгаа авахдаа орчуулгын хоёр хувилбарыг зургийн хамт үзүүлж, “Доорх хоёр тайлбараас зурагтай хамгийн сайн тохирч буйг сонгоно уу.” асуултыг асуусан. Судалгаа явуулсан вебийг зураг 2-т үзүүлэв.

Үнэлгээний хариултын чанарт нөлөөлөх нөлөөллийг бага байлгах үүднээс тайлбарыг аль аргаар орчуулсныг хэлэхгүй, сонголтын байрлалаас хамаарах хамаарлыг багасгах зорилгоор хоёр тайлбарын дараалыг санамсаргүй солих буюу заримдаа эхний тайлбар нь Google Translate, заримдаа хоёр дахь тайлбар нь Google Translate-р орчуулсан хувилбар гарч ирэхээр тохируулсан. Мөн нэг хэрэглэгчид нэг зургийн өөр тайлбар аль болох үзүүлэхгүй байлгахын тулд зураг гарч ирэх дараалыг мөн санамсаргүй байхаар, нэг тайлбарт дор хаяж нэг санал авсан бол “in_progress”, 3 санал

---

[1] Flickr30k-MN өгөгдөл: https://huggingface.co/datasets/bilguun/flickr30k-mn
[2] Flickr30k-MN өгөгдөлд санал авах веб аппликейшн: https://huggingface.co/spaces/bilguun/flickr30k-mn

авсан бол төлөвийг “done” болгож, санал нэмж авахгүй байхаар тус тус тохируулсан.



Зураг 2. Орчуулгын үнэлгээ хийлгэх веб аппликейшн.

IV. Үр дүн

Судалгаанд 98 хүн оролцсон бөгөөд нийт 746 зургийн 2,246 ялгаатай тайлбарт 5,138 санал авсан. Саналын тоон үзүүлэлтийг доор жагсаавал:

- Хувилбарын сонголт:
  - Google Translate (GT): 1,784 (39%)
  - Gemini (Gem.): 2,769 (61%)
- Ижил тайлбар: 50
- Веб дээрх тайлбарын байрлал:
  - GT – Gem.: 865
  - Gem. – GT: 1,381
- Төлөв:
  - done: 1,216
  - in_progress: 1,030

Судалгааны үр дүн тооцохдоо санал авсан тайлбарууд дундаас аль аль орчуулгын арга нь ижил тайлбартай эсвэл 2 санал аваад, санал нь тэнцсэн тайлбаруудыг оролцуулалгүй тооцсон. Шүүлт хийсний дараа 2,024 тайлбар үлдсэн. Түүнээс олонх санал авсан хувилбарыг сонгоход 36.2% (733) нь Google Translate-р орчуулсан хувилбар, 63.8% (1,291) нь Gemini-н үүсгэсэн хувилбар сонгогдсон.

V. Дүгнэлт

Орчуулгын чанарын үнэлгээгээр дан бичвэрээс орчуулга хийдэг машин орчуулгын Google Translate системтэй харьцуулахад ерөнхий зориулалтын, мультимодал хэлний том загвар болох Gemini 2.0 Flash нь зураг болон англи тайлбараас илүү чанартай, зургийг мөн харж байгаа тул утга зүйн хувьд илүү ойлгомжтой орчуулга хийснийг судалгаанд оролцогчдын олонх нь буюу ~63.6% нь үзжээ. LLM, ялангуяа Vision LLM-үүд өргөн хүрээний зураг, бичвэрийн өгөгдөл дээр суралцсанаар төрөл бүрийн даалгаврыг биелүүлэх чадвартай болж, цэвэр машин орчуулгын зорилгоор бүтсэн системээс давуу үр дүн үзүүлэх боломжтойг эндээс харж болохоор байна.

Тиймээс зураг, бичвэртэй холбоотой, монгол хэл дээрх өгөгдөл багатай салбарт холбогдолтой гадаад хэл дээрх өгөгдлөөс орчуулах, бэлтгэх боломж байгааг эндээс мөн харж болохоор байна.

Цаашид энэхүү судалгааны ажлын хүрээнд үүсгэсэн зураг, тайлбар хосолсон өгөгдөлд суурилан зургаас тайлбар үүсгэх, бичвэрийн утгад тохирох зураг хайх гэх мэт машин сургалтын загвар, мөн англи, монгол зэрэгцээ өгөгдлөөр машин орчуулгын систем үүсгэх судалгаа, шинжилгээний ажил цааш үргэлжлүүлэн хийх боломжтой.

VI. Талархал

Энэхүү судалгааны ажлын хүрээнд хэрэглэсэн Flickr30k өгөгдөл бэлдэх ажилд оролцсон МУИС, МТЭС-ийн компьютерийн ухааны бакалаврын төгсөгч Г.Саруулд талархал илэрхийлье.

Ном зүй

[1] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, “Microsoft COCO: Common objects in context,” in Proceedings of the European Conference on Computer Vision (ECCV), 2014, pp. 740–755.

[2] P. Young, A. Lai, M. Hodosh, and J. Hockenmaier, “From image descriptions to visual denotations: New similarity metrics for semantic inference over event descriptions,” in Transactions of the Association for Computational Linguistics, vol. 2, pp. 67–78, 2014.

[3] A. Radford et al., "Learning Transferable Visual Models From Natural Language Supervision," CoRR, vol. abs/2103-00020, 2021. [Online]. Available: https://arxiv.org/abs/2103.00020

[4] J. Li, D. Li, C. Xiong, and S. Hoi, "BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation," in ICML, 2022.

[5] R. Krishna, Y. Zhu, O. Groth, J. Johnson, K. Hata, J. Kravitz, et al., "Visual genome: Connecting language and vision using crowdsourced dense image annotations," in International Journal of Computer Vision, vol. 123, no. 1, pp. 32–73, 2017.

[6] D. Elliott, S. Frank, K. Simaan, and L. Specia, “Multi30k: Multilingual English-German Image Descriptions,” in Proceedings of the 5th Workshop on Vision and Language, pp. 70–74 , 2016.

[7] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), 2017, pp. 6000–6010.

[8] M. Enis and M. Hopkins, “From LLM to NMT: Advancing Low-Resource Machine Translation with Claude,” arXiv preprint arXiv:2404.13813, 2024. [Online]. Available: https://arxiv.org/abs/2404.13813.

[9] "Gemini 2.0 Flash," Google AI, [Online]. Available: https://gemini.google.com. [Accessed: 2025.03.27].

УХААЛАГ ХӨТӨЧ

Мөнхбаяр Түмэн-Аюуш (нэгдүгээр зохиогч)[1], Чулуунжин Мичидгоо (хоёрдугаар зохиогч)[2],
Сүхбаатар Батмөнх (гуравдугаар зохиогч)[3], Соном-Очир Өлзийбаяр (дөрөвдүгээр зохиогч)[4],
Монгол улс, Улаанбаатар, ШУТИС, Мэдээлэл, холбооны технологийн сургууль, Мэдээллийн технологийн салбар
B210930048@must.edu.mn , B210920035@must.edu.mn, B210930011@must.edu.mn, ulziibayar@must.edu.mn

***Хураангуй* — Мэдээллийн технологийн хурдацтай хөгжлийн энэ үед харааны бэрхшээлтэй иргэд шинэ дэвшлийг хүртэн, өөрт тохирсон технологийн шийдлээр хангагдах нь нэн чухал болоод байна. Үндэсний статистикийн хорооны 2023 оны мэдээллээр Монгол Улсад 12,053 харааны бэрхшээлтэй иргэн бүртгэгдсэн бөгөөд тэдний хувьд өдөр тутмын энгийн үйлдлүүдийг бие даан гүйцэтгэхэд хүндрэлтэй, зарим тохиолдолд аюултай байдаг. Тиймээс энэхүү судалгааны ажлаар Монгол орны нөхцөлд нийцсэн, харааны бэрхшээлтэй иргэдийн бие даасан байдал болон аюулгүй зорчих орчныг бүрдүүлэх зорилготой ухаалаг хөтөч системийг хөгжүүлсэн. Тус систем нь гар утасны аппликейшн хэлбэртэй бөгөөд хэрэглэгчийн дуу хоолойг таньж, очихыг хүссэн газрыг тодорхойлсны дараа GPS ашиглан байршлыг тогтоож, чиглэл заах мэдээллийг өгдөг. Хэрэглэгч нь замчилгааг аудио хэлбэрээр хүлээн авч, гар утасны камерын тусламжтайгаар орчны объектыг таньж, тэдгээрийн зайг тооцоолох боломжтой. Ингэснээр нэмэлт төхөөрөмж шаардлагагүй, хэрэглэгчдэд илүү хялбар бөгөөд хүртээмжтэй шийдэл болж байна. Системийг хөгжүүлэхдээ машин сургалтын YOLO загварыг ашиглаж, Монгол орны нөхцөлд тохирсон бодит замын бичлэгүүд дээр туршилт хийсэн. Үүний үр дүнд объект илрүүлэлтийн нарийвчлал 98.58%, зай тодорхойлолт 98.36%-тай гарсан. Мөн Text-to-Speech функц нь хэрэглэгчдэд мэдээллийг ойлгомжтой, тод хэлбэрээр хүргэх боломжийг бүрдүүлж, 98%-ийн гүйцэтгэлтэй үнэлэгдсэн. Цаашид уг системд GPS функцийн өргөтгөл, асран хамгаалагч руу байршлын мэдээлэл дамжуулах, IOS болон Android системд зэрэг ажиллах боломжийг бүрдүүлж, харааны бэрхшээлтэй иргэдэд зориулсан үнэ төлбөргүй, олон нийтэд нээлттэй хоёр хувилбар болгон өргөжүүлэхээр зорьж байна. Энэхүү судалгаа нь Монгол Улсын нөхцөлд үндэслэсэн, бодит хэрэгцээг хангасан хүртээмжтэй технологийн шийдэл болох бүрэн боломжтойг харуулж байна.**

Түлхүүр үг* — *машин сургалт, object distance estimation, object detection, харааны бэрхшээлтэй иргэн, аудиогоор мэдээлэх, дуу хоолойн удирдлага, газрын зураг

I. УДИРТГАЛ

Ухаалаг хөтөч систем нь харааны бэрхшээлтэй иргэдийн өдөр тутмын хэрэгцээнд нийцсэн, аюулгүй зорчих боломжийг бүрдүүлэх зорилготой технологийн шийдэл юм. Дэлхийн Эрүүл Мэндийн Байгууллагын (ДЭМБ)-ын мэдээлснээр, дэлхий даяар 246 сая хүн харааны бэрхшээлтэй байгаагийн 36 сая нь хараагүй, 210 сая нь хараа муутай байдаг [1]. Монгол Улсад ойролцоогоор 11,400 орчим харааны бэрхшээлтэй хүн амьдардаг бөгөөд үүний 3 мянга орчим нь Улаанбаатар хотод оршин суудаг [2]. Эдгээр иргэд өдөр тутамдаа цагаан таяг ашиглан зорчдог ч тэдний аюулгүй байдал, зорчих нөхцөл бүрэн хангагдаагүй байна. Иймээс, ухаалаг хөтөч системийг боловсруулсан бөгөөд энэ нь орчин үеийн технологийн дэвшлийг ашиглан харааны бэрхшээлтэй иргэдийн амьдралын чанарыг дээшлүүлэх чухал шийдэл болж байна. Тийм ч учраас судлаачид энэ төрлийн судалгааг сүүлийн жилүүдэд маш эрчимтэй өрнүүлж байна. Тэдгээр судалгаануудыг ашигласан технологи, арга аргачлалаас хамааруулан дүрсийн боловсруулалтад суурилсан арга, төхөөрөмжид суурилсан арга болон эдгээр аргын хосолсон арга гэсэн гурван бүлэгт хуваах үздэг. Бид эдгээр аргуудаас хамгийн өртөг багатай боловч, үр дүн сайтай арга болох дүрсийн боловсруулалтад суурилсан аргаар хийхээр зорьсон. Объект таних олон төрлийн алгоритм байдгаас бид хэд хэдэн шалгуур болон судалгаанд үндэслэн өөрсдийн шаардлагад хамгийн тохиромжтой алгоритмыг сонгосон. Тус алгоритмуудын ялгаа нь хурд, нарийвчлал, нарийн төвөгтэй байдлын харьцаанд оршдог бөгөөд хэрэглээний өөр өөр болон тооцооллын хязгаарлалттай уялддаг. Бид сүлжээний дизайн, алдааны хувь, нэгж хуудас дахь фреймүүдийн тоо, дундаж үзүүлэлтүүдийг нарийвчлан судалсны үндсэн дээр YOLO буюу дүрс боловсруулах моделийг сонгож, түүнд тохирсон хэрэглээний нөхцөл болон өөрсдийн шаардлагыг харгалзан үзсэн.

Энэхүү судалгаанд санал болгож буй систем нь аудио мэдээлэлд суурилан орчны талаар мэдээлэл дамжуулах, мөн хэрэглэгчээс шууд дуу хоолойгоор удирдлага авч замчлах шийдлийг танилцуулж байна. Уг систем нь машин сургалт, зургийн боловсруулалт, аудио технологи зэрэг дэвшилтэт аргачлалыг багтаасан. Камераар дамжуулан объектын зай, байршил, төрөл зэрэг шинж чанаруудыг тооцоолж, дүн шинжилгээ хийдэг. Зургийн боловсруулалтын салбарын өндөр хөгжлийг илүү шинэлэг байдлаар ашиглаж, орчны асуудлыг шийдвэрлэхэд чиглүүлсэн нь энэ системийн онцлог юм. Тус систем нь харааны бэрхшээлтэй хүмүүст бодит цагийн аудио мэдээлэл дамжуулж, тэднийг

хүрээлэн буй орчинтой холбогдох боломжийг нэмэгдүүлнэ.

Иймд, ухаалаг хөтөч систем нь дээр дурдсан асуудлуудыг шийдвэрлэж, орчин үеийн технологитой хослуулан аюулгүй, үр ашигтай зорчих орчныг бүрдүүлэх чухал ач холбогдолтой.

Эцэст нь, бид бодит хугацаанд объект таних, мэдээлэх системийн тасралтгүй ахиц дэвшлийг тодорхойлохын зэрэгцээ цаашдын судалгаа, хөгжүүлэлтийн боломжит чиглэлүүдийг судалсан. Мөн системийн онцлог, давуу талуудыг бусад ижил төстэй судалгаа, шийдлүүдтэй харьцуулан тайлбарласнаар бидний санал болгож буй шийдэл ялгарч буйг тодотгосон.

1-Р ХҮСНЭГТ. СУДАЛГААНЫ АЖЛУУДЫН ХАРЬЦУУЛАЛТ

Судалгааны ажил	[3]	[4]
Модель	CNN болон LSTM	YOLO v5
Аудио	Тийм	Үгүй
Сул тал	Архитектур дурдаагүй ба 5000 зурагтай өгөгдлийн багцыг ашигласан ба олон функцийг дэмжихгүй.	Доторх зургууд нь маш хязгаарлагдмал, цөөхөн (8000 зурагтай зөвхөн 16 доторх объектыг илрүүлэх боломжтой)

Эдгээр судалгааны ажлууд нь манай системийн үйл ажиллагаатай төстэй бөгөөд тус судалгаанд үндэслэн өөрийн системд сайжруулалт хийх шаардлагатай гол асуудлуудыг илрүүлэхэд тус болсон.

Амазонын амжилтын шалтгаан нь өнөөгийн сошиал хэрэглэгчдийн хамгийн их эрэлт хэрэгцээтэй байгаа онцлог нь яриагаар болон дуу хоолойн удирдлагатай харилцаж чаддагт оршдог [15]. Үүнээс үзэхэд хөгжлийн бэрхшээлтэй иргэд гэлтгүй жирийн хэрэглэгчдэд хүртэл дуу хоолойн удирдлага ямар нэмэртэй, мөн ажиллагдаж байгааг харж болно.

Эндээс үзэхэд эдгээр хязгаарлалтуудыг үндэслэн бидний судалгааны ажилд YOLO моделийг ашиглаж, илүү өргөн хүрээний объект таних, бодит хугацаанд мэдээллийг аудио хэлбэрээр дамжуулах боломжийг оруулснаар хэрэглэгчдэд илүү хүртээмжтэй, өргөн цар хүрээтэй системийг санал болгож байна. Мөн түүнчлэн дан ганц объект илрүүлэн, уг объектыг тодорхойлохын зэрэгцээ тухайн объектын орон зайн байршил, хөдөлгөөний чиглэл, хурд зэрэг нэмэлт мэдээллийг харааны бэршээлтэй хүнд мэдээлсэнээр бодит орчноо илүү сайн ойлгож, шаардлагатай шийдвэр гаргахад туслах юм.

Baidu компани Baidu IHome-г, Алибаба ч Tmall гаргасан. Эдгээр компаниудаас гадна ухаалаг дуу хоолойг туслахтай бүтээгдэхүүн гаргасан Хятадын өөр олон амжилттай компани байдаг[16].

Үүнтэй адил хэрэглэгчтэй бодит хугацаанд дуу хоолойгоор харилцах боломжтой болж, хэрэглэгч очихыг хүссэн газрынхаа мэдээллийг дуу хоолойгоор оруулна. Аппликэйшн нь хэрэглэгчийн өгсөн мэдээллийг боловсруулан, очих газрын байршлыг тодорхойлсны үндсэн дээр дуу хоолойгоор чиглэл заах, замын талаарх мэдээллийг бодит хугацаанд өгч эхэлнэ. Ингэснээр хэрэглэгч гар болон нүдний оролцоогүйгээр, зөвхөн дуу хоолойн тусламжтайгаар зорьсон газартаа хүрэх боломжтой болно. Энэ нь харааны бэрхшээлтэй болон хөдөлгөөнд хязгаарлалттай хүмүүст асар их дөхөмтэй байхаас гадна, жолоо барьж яваа үед ч аюулгүй байдлыг хангах үр дүнтэй шийдэл болох юм.

II. ХЭРЭГЛЭГЧИЙН СУДАЛГАА

Харааны бэрхшээлтэй хүмүүс харааны хагас бэрхшээлтэй эсвэл хараа бүрэн алдсан байдаг. Энэ нь тэдний өдөр тутмын амьдралд ихээхэн нөлөөлж, хөдөлгөөн, ажил гүйцэтгэх чадвар, амьдралын чанарт нь нөлөөлсөөр байгаа. Ийм хөгжлийн бэрхшээлтэй хүмүүсийн дийлэнх нь 50-иас дээш насныхан байдаг.

Харааны эмгэгийн олон шалтгаан бий. Булчингийн доройтол мөн зарим генетикийн эмгэгүүд нь харааг муутгадаг. Нүд, толгой болон бусад төрлийн гэмтэл нь нүдний өвчинд хүргэдэг бөгөөд энэ нь эмгэгийг үүсгэдэг. Нүдийг гэмтээж, харааны эмгэгт хүргэдэг бусад өвчинд чихрийн шижингийн ретинопати, катаракт (cataracts), глауком (glaucoma), насжилттай холбоотой доройтол орно. Хорт бодист өртөх, тод дэлгэц, хурц гэрлийг удаан хугацаагаар үзэх, контакт линзийг удаан хугацаагаар ашиглах зэрэг нь хараа бүрэн алдагдахад хүргэдэг. Энэ бол хэзээ ч алга болох

асуудал биш бөгөөд хараа муудсан, хараагүй болсон ч эргэн тойрон, урд байгаа саад тотгорыг таних систем нэн шаардлагатай байгаа билээ.

Харааны бэрхшээлтэй иргэдийн хувьд нийтийн тээврийн болон явган зорчих орчны хүртээмж, аюулгүй байдал нь өдөр тутмын үйл ажиллагаанд оролцох, нийгмийн харилцаанд идэвхтэй оролцох боломжийг ноцтойгоор хязгаарлаж байна. Монгол Улсад нийтийн тээврийн хэрэгслүүд зогсоол бүр дээрээ тогтмол зогсдоггүй, чиглэлийн талаарх мэдээлэл нь дуу хоолой эсвэл аудио зарлалаар хүрдэггүй нь харааны бэрхшээлтэй иргэдийн хувьд зорчиход хүндрэл үүсгэж байна. Үүнээс гадна явган хүний зам дагуу байрлах гэрэлтүүлгийн шон, модны мөчир, сурталчилгааны самбар, зам дээр байрлуулсан унадаг дугуйнууд зэрэг нь тэдний хөдөлгөөнд саад учруулж, гэмтэл бэртэлд хүргэх эрсдэлтэй.

Монголын Хараагүйчүүдийн Үндэсний Холбооны (МХҮХ) Ерөнхийлөгч Б.Ганзоригийн тэмдэглэснээр дээрх асуудлуудыг холбогдох байгууллагад удаа дараа уламжилсан ч дорвитой шийдвэрлэлт хийгдээгүй хэвээр байна [2]. Энэ нь харааны бэрхшээлтэй иргэдийн эрх тэгш амьдрах, орчиндоо бие даан оролцох нөхцөлийг алдагдуулж буйг харуулж байна. Харааны бэрхшээлтэй иргэдийн үндсэн хэрэгцээ, шаардлага нь дараах байдлаар тодорхойлогддог [5]:

- *Орчны мэдээллийг аудио хэлбэрээр шууд авах,*
- *Ойртож буй объектын зайг тодорхойлох,*
- *Объектын хөдөлгөөний чиглэл, хурдыг мэдээлэх.*

Эдгээр шаардлагыг биелүүлэхэд ухаалаг төхөөрөмж болон гар утасны нэгдсэн системийн оролцоо чухал юм. Судалгаагаар харааны бэрхшээлтэй олон иргэн өдөр тутмын зорчих хөдөлгөөндөө зөвхөн цагаан таяг ашигладаг ч, зөвхөн энэ хэрэгслээр хангалттай аюулгүй байдал бүрдэхгүй байгаа нь тогтоогдсон. Манай судалгаа нь энэхүү нөхцөл байдлыг сайжруулж, гар утасны ухаалаг шийдлээр дамжуулан харааны бэрхшээлтэй иргэдийг орчинтой нь илүү үр дүнтэй холбож, мэдээллийг дуу хоолойгоор өгч, аюулгүй зорчих боломжоор хангахад чиглэж байна. Уг систем нь таягтай хослон ажиллах ба цаашлаад таяггүйгээр бие даан зорчих боломжийг бүрдүүлэх суурь нөхцөлийг бий болгох боломжтой.

Үүнээс гадна Монгол орны хот төлөвлөлт, дэд бүтцийн стандарт хангалтгүй, замын хөдөлгөөний зохион байгуулалт нь харааны бэрхшээлтэй иргэдэд төдийгүй бүх иргэдэд эрсдэл дагуулж буй өнөөгийн нөхцөл байдал нь энэхүү төслийн хэрэгцээ шаардлагыг улам нэмэгдүүлж байна.

III. СУДЛАГДСАН БАЙДАЛ

Өмнөх судалгаануудад Convolutional Neural Networks (CNN) буюу Мэдрэлийн сүлжээ гэгддэг технологийг ашиглан объект илрүүлж, таних системүүдийг бүтээхэд бага зардалтай ч давуу талтай гэж үзсэн байдаг. Гэвч эдгээр судалгаанд хэд хэдэн сул талууд илэрсэн. Тухайлбал, [3] CNN болон LSTM ашиглан хязгаарлагдмал зураг дээрх объектыг танин, аудио хэлбэр болгон хувиргах туршилтыг хийсэн байсан. Энэ систем нь тооцооллын хурд сайтай байсан боловч бодит хугацаанд үйлдлийг дэмждэггүй байсан нь бодит орчинд хэрэглэх боломжийг хязгаарласан.

2020 онд хийгдсэн бас нэгэн судалгаа [4] YOLOv5 ашиглан харааны бэрхшээлтэй хүмүүсийн хөдөлгөөнд дэмжлэг үзүүлэх зорилгоор орчны объект илрүүлэх системийг боловсруулсан. YOLOv5-ийн тусламжтайгаар илрүүлэх хурд сайжирсан хэдий ч уг систем нь зөвхөн хязгаарлагдмал өгөгдлийн сантай (8000 зураг) бөгөөд 16 төрлийн объектыг илрүүлэх боломжтой байсан тул хэрэглээний цар хүрээ багатай хэвээр байв. Үүнээс гадна уг систем нь илэрсэн объектын талаар дуу хоолойн мэдээлэл дамжуулах боломжгүй байсан нь хэрэглэгчдэд шууд мэдээлэл хүргэх тал дээр дутагдалтай хэвээр байлаа.

Олон төрлийн дараах шийдлүүд нь хараагүй хүмүүст туслахын тулд компьютерийн тусламжтай технологийг ашигласан. Жишээлбэл, TapTapSee гар утасны програм [11] нь хараа бэрхшээлтэй хэрэглэгчдийн 10 секундын дотор авсан зургийг танихын тулд компьютер, олон сүлжээг ашигладаг. Blindsight's Text Detective [12] нь камерт авсан зургуудаас текстийг таних, уншихад тэмдэгт таних (OCR) аргыг ашигладаг. Baidu саяхан DuLight төслийн [14] демо видеог үзүүлж, зургийн дүрслэл, хүмүүс, мөнгөн тэмдэгт, бараа бүтээгдэхүүн, явган хүний гарцын дохиог таних тухай ойлголтуудыг харуулсан. Гэсэн хэдий ч эдгээр нь Монгол улсын нөхцөлд хараагүй иргэдэд шууд хүргэхэд дутмаг байна.

Системийг хөгжүүлэхдээ дэлхий дахинд болон Монголд ямар байдлаар судлагдсан, хэрэгжсэн талаар холбогдох дараах судалгааг хийсэн. Объектыг таних, орчиндоо дасан зохицох зэрэг үйл явц нь харааны бэрхшээлтэй иргэдэд олон сорилтыг бий болгодог бөгөөд эдгээр асуудлыг шийдвэрлэхийн тулд судлаачид компьютер болон хиймэл оюуны (AI) алгоритмуудыг өргөнөөр ашиглах болсон. Харааны бэрхшээлтэй хүмүүст зориулсан объект илрүүлэх системүүдийн гол зорилго нь илрүүлэлтийн хурд, нарийвчлалыг сайжруулж, олон төрлийн объектыг бодит цагийн горимд зөв таньж илрүүлэхэд чиглэгддэг.

A. Судалгааны ажил

Энэхүү судалгаанд [6] YOL0v4 болон Resnet101 загваруудыг хослуулж, олон объект таних нарийвчлалтай системийг хөгжүүлэхэд анхаарсан. YOL0v4 нь объектыг илрүүлэх өндөр үр дүнтэй загвар бөгөөд хурд, нарийвчлалын үзүүлэлтээрээ нэр хүндтэй бол Resnet101 нь гүний сургалтын нейрон сүлжээний архитектур ашиглан илүү нарийн таних чадварыг бий болгосон. Судалгааны үр дүнд гарсан YOL0v4_Resnet101-ийн өөрчилсөн хувилбар нь 4000 туршилт дээр 96.34%-ийн нарийвчлал, 0.073%-ийн алдааны үзүүлэлттэй байсан бөгөөд энэ нь харааны бэрхшээлтэй хүмүүст өдөр тутмын амьдралд тулгардаг объект таних асуудлыг шийдвэрлэхэд өндөр ач холбогдолтой болсон.

- Тус судалгаагаар дээрх объект илрүүлэх загварыг текстээс яриа руу хөрвүүлэх технологитой хослуулсан бөгөөд энэ нь харааны бэрхшээлтэй хүмүүст хүрээлэн буй орчноо бодит цагийн хоолойгоор дамжуулан ойлгож, мэдрэхэд тусалсан. Бодит цагийн туршилтууд болон хэрэглэгчийн судалгааны үр дүнд YOL0v4_Resnet101 интеграцийн үр дүн, бодит байдал дээр ажиллах чадвар нь батлагдсан бөгөөд энэ нь харааны бэрхшээлтэй хүмүүст эргэн тойрны орчиндоо илүү дасан зохицох боломжийг нээж өгсөн.

B. Судалгааны ажил

Сүүлийн жилүүдэд текстээс яриа үүсгэх системүүд (TTS) нь хиймэл оюун ухаан болон гүний сургалтын (deep learning) аргачлалуудын тусламжтайгаар хурдацтай хөгжиж, хүн-компьютерийн харилцааны асуудлыг шийдэх чухал хэрэгсэл болж байна. TTS системүүд нь янз бүрийн хэлэнд зориулсан загваруудыг агуулдаг бөгөөд үүнийг сайжруулахад харилцааны хиймэл оюун ухаан, чатбот, intelligent speech interaction зэрэг олон төрлийн хэрэглээний хувилбарууд хөгжүүлэгдсээр байна. Энэ нь олон хэлний онцлогийг харгалзан үзэж, хэрэглэгчийн хувьд илүү дуу хоолойн уншлагыг бий болгоход тусалж байгаа юм.

- Дээрх судалгаанд TTS системийг хөгжүүлэхэд тулгарч буй гол бэрхшээлүүдийг судалж, тэдгээрийг шийдвэрлэхэд ашиглагдаж буй аргачлалуудыг тоймлон багтаасан. Санах ой, машин сургалт, гүний сургалтын загваруудыг ашиглах нь TTS системийн чанарыг сайжруулж, текстээс яриа үүсгэх нарийвчлал, хурдыг нэмэгдүүлэх, хэлний суурьтай загваруудын үгийн алдаа болон санах ойд зарцуулах хэрэглээг бууруулахад онцгой үр дүнтэй болохыг харуулсан.

2-Р ХҮСНЭГТ. АШИГЛАСАН ТЕХНОЛОГИУД

Судалгааны ажил	**Аргачлал**
[1]	❖ таних хурд(rate), ❖ нарийвчлал (accuracy), ❖ TTS (текстээс аудио), ❖ Санах (recall), ❖ Зэрэг чанарын хэмжүүрүүд.
[6]	❖ YOL0v4_Resnet101 ❖ MS COCO dataset

C. Судалгааны ажил

Судалгааны ажлуудын нэг болох "Third Eye" хэмээх энэ аргачлал нь YOLOv5 загвар болон текстээс яриа үүсгэх (TTS) технологийг ашиглан илүү өндөр нарийвчлалтай объект илрүүлж, хараагүй хүмүүст бодит цагийн мэдээлэл дамжуулах системийг хөгжүүлэх судалгаа юм.

Third Eye систем нь нийт 95 төрлийн объектыг таних чадвартай бөгөөд үүнд 15 төрлийн тусгайлсан объектыг агуулсан dataset болон MS COCO 2017 dataset-ийг ашигласан. Загварын илрүүлсэн объектыг текст болгон хувиргаж, дараа нь pyttsx3 болон gTTS зэрэг Python программчлалын хэл дээрх TTS сангуудын тусламжтайгаар аудио формат руу

хөрвүүлсэн. Эдгээр TTS сангуудыг харьцуулж туршсанаар pyttsx3 нь бодит цагийн горимд илүү өндөр хурдтай ажиллаж, илүү найдвартай үр дүн гаргасан нь харааны бэрхшээлтэй хүмүүст сонсголын хариу өгөх хугацааг багасгаж, мэдээллийг хурдан дамжуулахад онцгой үүрэгтэй болохыг баталсан.

- Аудио хэлбэрээр үүсгэх үед pyttsx3 нь ямар ч интернэт холболт шаарддаггүй, харин нөгөө талаас gTTS нь байнгын интернэт холболт шаарддаг болохыг ажигласан. Аудио файл үүсгэхийн тулд gTTS нь Google-ийн серверүүд рүү текст илгээж, дараа нь буцааж өгдөг. Тиймээс pyttsx3-ийг gTTS-ээс илүү ашигтай гэж үзсэн. Тус судалгааны ажлаас манай баг нь аудио бүтээхэд зарцуулах цаг хугацаа, саатал, lib, сүлжээний холболт зэргийг мөн харгалзан үзэх ёстойг анхаарсан. Учир нь бид муу сүлжээ гэх мэт гаднын хүчин зүйлийн нөлөөнд автахгүй бүтээхийг хүсч байна.

D. Судалгааны ажил

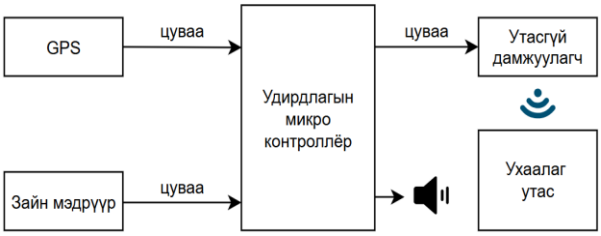
Дэлхий олон улсад хэрэглэж буй цагаан таягийг Монгол улсын нөхцөл байдалд зориулсан, ажиллагааны шинэлэг загвартай, ухаалаг болгох зорилгоор техникийн шийдэл гаргаж, програмчилсан байсан. [10] Уг таяг нь :

- Ультра - соник : Таяг дээр байрлах сенсор нь хэрэглэгчийн мөрнөөс дээш байрласан саадыг мэдэрч, хэрэглэгчийн програмд мэдээлнэ.
- GPS модуль : Асран хамгаалагч болон төөрсөн тохиолдолд товчлуур дарна өөрийн одоогийн байршлыг заасан мессежийг илгээнэ.
- Bluetooth : 2G дэмждэг СИМ суулгаснаар хаягийн мэдээлэл илгээгдэнэ.

Тухайн таягийн үзүүлэлт :

- Таягийн үнэ 2019 оны ханшаар 131,300 төгрөг болсон.
- Чийгийн мэдрүүр нь усны гүнийг мэдэрдэг
- Монгол хэл дээр ажилдаг

Хэрэглэгчийн хариу : Өмнө өмнөх жилүүдэд ухаалаг таяг хийсэн олон оюутнууд байсан ба зөвхөн өмнө нь байгаа саадыг мэдрэх л үүрэгтэй байсан. Үүнийг цаашлаад сайжруулах хэрэгтэй байгаа.



1-р зураг. Системийн ерөнхий бүтэц

E. Судалгааны ажил

Хиймэл оюун ухааны дуу хоолойгоор хэрэглэгчдэд туслахад дүн шинжилгээ хийж, Хятад болон гадаадын загваруудын ялгааг харьцуулж, өдөр тутмын амьдралдаа нэгтгэх талаар тус судалгааны ажилд тусгасан байсан.

- China Telecom Terminal Research, туршилтын төвийн судалгаагаар дуу хоолойн туслахуудыг ашиглахад хэрэглэгчийн мэдлэг 95.6%, ашиглалтын түвшин 58.6% байна [14].

Тус судалгааны ажлаас дуу хоолойн туслах ашиглах бол үүнд хэрхэн хиймэл оюун ашиглаж болох мөн яаж электрон төхөөрөмжтэй холбогдож болохыг тусгасан байсан.

3-Р ХҮСНЭГТ. СУДАЛГААНЫ АЖЛУУДЫН ХАРЬЦУУЛАЛТ

Судалгааны ажил	**Гол ажиллагаа**
[17]	❖ AI шийдлүүд ❖ Электрон төхөөрөмжтэй холбогдох ❖ Өөр өөр функциональ сонголттой
[18]	❖ Машин сургалт болон TTS ❖ Автомат хоолой танигч (Automatic Speech Recognition) ❖ Natural Language Understanding, Generation (NLU, NLG)

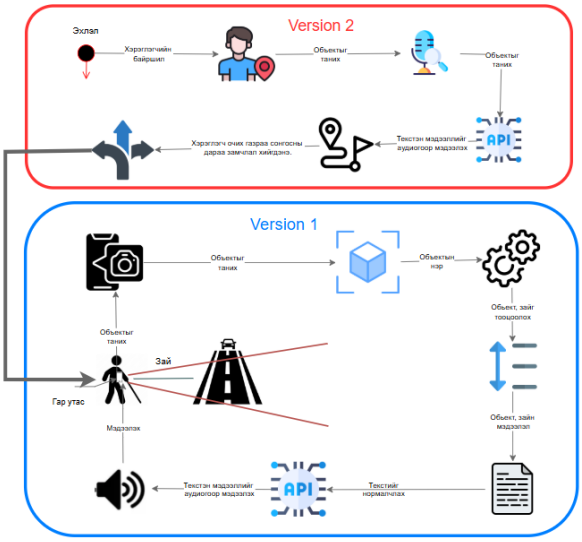
IV. ТӨСЛИЙН ХЭСЭГ

Төслийн үйл ажиллагаа: Төслийн үйл ажиллагаа нь харааны бэрхшээлтэй иргэдийн аюулгүй, бие даасан байдлыг хангахад чиглэсэн бөгөөд ухаалаг хөтөч системийг хөгжүүлэх, турших, хэрэглэгчдийн хэрэгцээнд нийцүүлэх ажлуудаас бүрдэнэ. Төслийн эхний шатанд Монгол Улсын нөхцөлд тохирсон хэд хэдэн объект таних, зайг тооцоолох технологиудыг судалж, YOLOv5 алгоритмыг ашиглан үндсэн загварыг боловсруулсан.

2-р зураг. Системийн ерөнхий ажиллагаа

Дараагийн шатанд системийн чанарыг сайжруулахын тулд туршиж үзэх, алдааг илрүүлэх, нарийвчлан хөгжүүлэх ажлууд хийгдэж байна.

Үүний зэрэгцээ, хэрэглэгчийн санал хүсэлтэд үндэслэн уг системийг текстэн мэдээллийг яриа руу хөрвүүлэх функцээр баяжуулж, харааны бэрхшээлтэй иргэдийн өдөр тутмын амьдралд туслах, хэрэглэхэд хялбар, үнэ төлбөргүй программ болгохоор ажиллаж байна.



3-р зураг. Системийн ерөнхий бүтэц

Дээрх зураг нь харааны бэрхшээлтэй иргэдэд зориулсан ухаалаг хөтөч системийн үйл ажиллагааны үндсэн үе шатуудыг харуулсан.

1. Гар утасны камер ашиглах: Хэрэглэгч гар утасны камерыг ашиглан зам болон орчны зураглалыг авдаг. Камераар дамжуулан тухайн орчны дүрслэлийг системд илгээдэг.
2. Объект таних: Камераар авсан дүрслэлийг систем анализ хийж, объект таних алгоритм ашиглан зам дээрх гол объектуудыг танина.
3. Объектын нэр, зайг тооцоолох: Систем таньсан объект бүрийн нэр болон хэрэглэгчээс хэр зайтай байгааг тооцоолж, тодорхойлсон мэдээллийг боловсруулахад бэлтгэдэг.
4. Объект болон зайны мэдээллийг дамжуулах: Тодорхойлогдсон объектын нэр, зайг API ашиглан дараагийн шат руу дамжуулна.
5. Текстийг нормалчлах: Дамжуулсан мэдээллийг текст хэлбэрт шилжүүлж, аудио мэдээлэлд бэлтгэнэ.
6. Текстэн мэдээллийг аудио болгох: Объектын нэр, зай зэрэг текстэн мэдээллийг Text-to-Speech (текстийг ярианд хөрвүүлэх) технологи ашиглан аудио болгон хөрвүүлж, хэрэглэгчид сонсголын хэлбэрээр хүргэнэ.

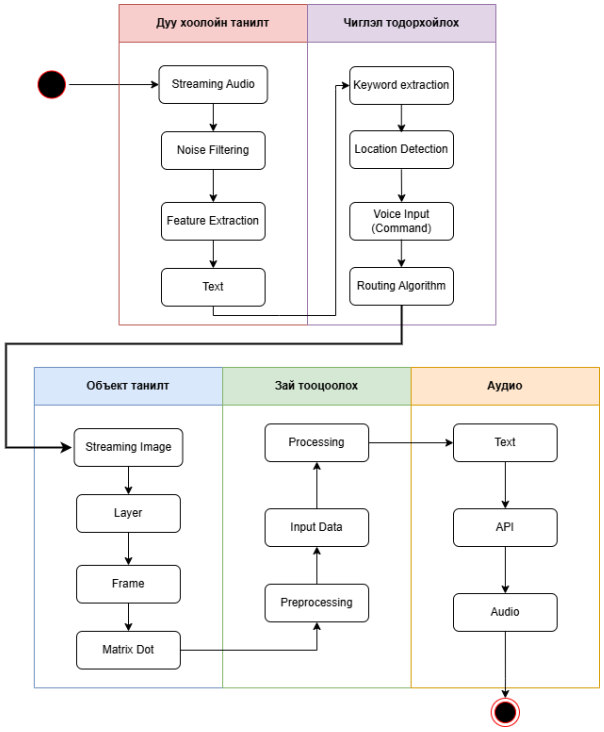
Энэ бүхэн нь харааны бэрхшээлтэй иргэдийг замын хөдөлгөөнд аюулгүй зорчих, орчны саад тотгорыг таних, тэдгээрээс зайлсхийхэд туслах ухаалаг хөтөч системийг бүрдүүлж байна.

4-Р ХҮСНЭГТ. ТӨСЛИЙН ДАВУУ БОЛОН СУЛ ТАЛ

Давуу тал	● Offline орчинд ажиллах боломжтой. ● Үнэгүй ● Ашиглахад хялбар ● Объект, зайн мэдээллийг аудио хэлбэрээр мэдээлэх ● Харааны бэрхшээлтэй иргэдийн бие даасан байдлыг дэмжинэ. ● Өргөтгөх боломжтой.
Сул тал	● Утасны ашиглалт, баттерей (зай) хурдан дуусах магадлалтай. ● Орчны хамааралтай танихгүй байх. ● Системийг цаг агаарын таагүй нөхцөлд ашиглахад бэрхшээлтэй байж болзошгүй.

V. МОДУЛЬ

Манай систем нь дуу хоолойн танилт, чиглэл тодорхойлох, объект танилт, зай тооцоолох, аудио гэсэн гэсэн таван үндсэн модулийг агуулсан бөгөөд тус бүрдээ дараах үйл явцуудаар дамжуулан харааны бэрхшээлтэй иргэдэд хүрэх аудио мэдээллийг боловсруулна.



4-р зураг. Модулиудын холбоо хамаарал

VI. ДУУ ХООЛОЙН ТАНИЛТ

Дуу хоолойн танилт нь хэрэглэгчийн өгсөн ярианг системд ойлгомжтой команд болон хувиргах үйл явц юм. Энэхүү процесс нь дараах үндсэн алхмуудаас бүрддэг.

Streaming Audio: Хэрэглэгчийн дуу хоолойг бодит хугацаанд тасралтгүй хүлээн авна.

Noise Filtering: Орчны дуу чимээг шүүж, зөвхөн хэрэглэгчийн ярианг онцлон ялгана. Энэ нь хэлбэр, аялга зөрүүтэй орчны нөхцөлд системийн ойлгох чадварыг нэмэгдүүлдэг.

Feature Extraction: Ярианы онцлог шинж чанарууд (pitch, tone, frequency) зэргийг ялгаж, дараагийн боловсруулалтад зориулан онцлог тэмдэгтүүд болгон хувиргана.

Text Conversion (Speech-to-Text): Танилт хийсэн дуу хоолойг бичвэр болгон хөрвүүлдэг. Үүнд STT (Speech-to-Text) систем ашиглагдана.

VII. ЧИГЛЭЛ ТОДОРХОЙЛОХ

Чиглэл тодорхойлох нь хэрэглэгчийн өгсөн ярианаас чиглэлийн командыг ялгаж, зорьсон байршлыг тогтоон, тухайн газарт хүрэх замыг тооцоолох үе шат юм.

Keyword Extraction (KE): Дуу хоолойноос гаргаж авсан бичвэрээс “баруун”, “зүүн”, “зах”, “эмийн сан” зэрэг чухал түлхүүр үгсийг ялгаж авна.

Location Detection (LD): Түлхүүр үг болон агуулгыг ашиглан газрын зураг дээрээс хэрэглэгчийн зорьж буй байршлыг тодорхойлно.

Voice Input (Command Analysis): Хэрэглэгчийн өгсөн командын утга агуулгыг гүнзгийрүүлэн задлан шинжилж, системд хэрэгжүүлэх боломжтой хэлбэрт оруулна (жишээ: “Зүүн эргээд, эмийн сан хүр”).

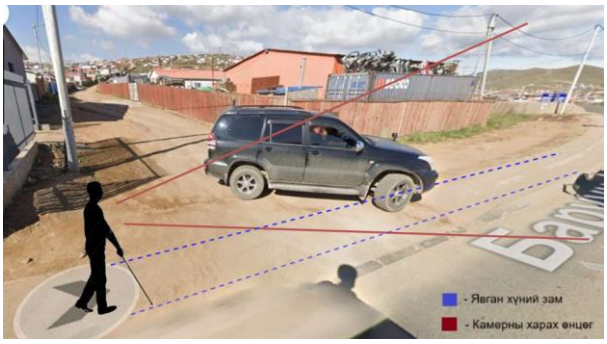
Routing Algorithm: Хэрэглэгчийн байршлаас зорьсон газарт хүрэх хамгийн богино, аюулгүй чиглэлийг тооцоолно. Алгоритм нь замын мэдээлэл, хөдөлгөөн, объектуудаас зайлсхийх боломж зэрэг олон хүчин зүйлийг харгалзан үздэг.

VIII. ОБЪЕКТ ТАНИЛТ

YOLO-ийн ажлын урсгалын диаграм (workflow) YOLO framework-ийг ерөнхийд нь дөрвөн үе шатанд хуваадаг:

1. *Боловсруулалт:* Зурган дээрх объектуудын харьцааг гаргахын оролтын дүрсийг 416 x 416 пиксел гэх мэт тогтмол хэмжээтэй болгож өөрчилж хадгалдаг.
2. *Мэдрэлийн сүлжээ технологи (CNN):* Энэ үе шатанд DCNN нь зураг дээрх объектуудын байгаа байдал, байршлын талаар таамаглал гаргахад ашиглагддаг. DCNN нь объектыг таних, тэдгээрийн хязгаарлах хэсгийг урьдчилан таамаглахад зургийг өгөгдлийн багц ашиглагддаг.
3. *Хувиарлалт:* Энэ алхамд дүрсийг хуваадаг бөгөөд хуваалт нь зургийн тодорхой хэсэгт байгаа объектуудын байгаа байдал, байршлыг урьдчилан таамаглах үүрэгтэй. YOLO нь олон тооны хязгаарлах хэсгийг гаргаж, хэсэг бүрд байгаа объектуудын ангиллын магадлалыг тодорхойлдог.
4. *Объект таних:* Дараа нь эдгээр таамаглалыг ашиглан шүүж, давхцаж буй танилтыг хасаж эцсийн илрүүлэлтийн үр дүнг гаргана. Эцэст нь, тухайн хэсэгт илрүүлсэн объектын нэр, зайг хамт гаргадаг байна. Бүх үйл явц нь фрейм бүрийн хувьд давтагддаг бөгөөд энэ нь YOLO-д бодит цагийн объект илрүүлэх боломжийг олгодог.

Мөн хэд хэдэн загвар (model) байдаг ба хувилбар бүр нь ялгаатай функцүүд агуулж болдог.



5-р зураг. Объект танилтын харагдац

4-Р ХҮСНЭГТ. ОБЪЕКТ ТАНИХ МОДЕЛИУДЫН ХАРЬЦУУЛАЛТ

Модель	Гүйцэтгэл (%)	FPS			Өгөгдлийн сан
		Tesla T4	1660 Ti	Jetson Nano	
YOLOv3	**54.3**	80	21	8	MS COCO 2017
YOLOv5s	37.6	**100**	**28**	**15**	MS COCO 2017
MobileNet-SSD V2	33.7	94	26	15	MS COCO 2017

IX. ЗАЙ ТООЦООЛОХ

Object distance estimation нь аливаа объект болон тухайн объект руу харж буй гар утасны камер хоёрын хоорондын зайг тодорхойлох аргачлал юм. Энэ аргыг компьютерын хараа, робот техник, зураг боловсруулах, автомат жолоодлого, виртуал болон нэмэлт бодит байдал гэх мэт салбаруудад өргөнөөр ашигладаг.

- Дүрсний тооцоололт – Объектын өнцөг болон зайг тодорхойлохдоо хоёр өөр байрлал эсвэл өнцгөөс авсан мэдээллийг ашигладаг. Энэ аргыг лазер болон stereo камерын мэдрэгчтэй хамт ашигладаг.
- Стерео – Stereo камер ашиглан нэг объектыг хоёр өөр өнцгөөс дүрслээд, хоёр дүрсний хоорондын ялгааг (disparity) ашиглан зайг тооцдог.
- Тооцооллын арга (monocular method) – Нэг камерын дүрслэлийг ашиглан объектын том, жижиг, тод, бүдгэрэл зэрэг шинж чанараар зайг тооцоолдог.
- Илрүүлэлт болон хүрээг тооцоолох (LiDAR) – Лидар технологи ашиглан лазерын тусламжтайгаар гэрлийг объектоос цацаж, буцаж ирэх хугацааг хэмжин зайг тодорхойлдог.
- Таних болон сургах загвар (TensorFlow) – Гэрэл эсвэл дууны дохио ашиглан объектоос ойртон очоод буцаж ирэх хугацааг хэмжих замаар зайг тооцдог.

Эдгээр нь object distance estimation буюу зайг тодорхойлох аргуудын хамгийн түгээмэл хэрэглэгддэг аргачлалууд юм.

Объектын зайг тооцоолох томьёо[7]:

$$object\ distance = \frac{focal\ length\ *\ real\ height\ of\ the\ object\ *\ image\ height}{object\ height\ *\ sensor\ height}$$

- focal length - Камерын линзний фокусын урт, буюу линзний төвөөс мэдрэгч хүртэлх зай.
- real height of the object - Бодит объектын өндөр буюу тухайн объектын жинхэнэ өндөр хэмжээс.
- image height - Мэдрэгч дээр дүрслэгдсэн объектын дүрсний өндөр.
- object height - Дүрсэн дэх объектын өндөр.
- sensor height - Камерын мэдрэгчийн өндөр буюу мэдрэгчийн физик хэмжээсний нэг параметр.

Энэхүү томьёо нь тухайн объект мэдрэгч дээр хэр том хэмжээтэй дүрслэгдсэн, мөн мэдрэгчийн хэмжээ, линзний фокусын урт зэрэг хүчин зүйлсийг ашиглан объектын бодит зайг тооцоолдог.

6-р зураг. Объектын зайг тооцоолох

X. АУДИО

TTS нь тухайн объектыг таньсан тохиолдолд дараагийн алхам буюу зорилтот хэрэглэгчид хэрхэн мэдэгдэх, мэдээлэхэд ашиглагдах арга техник юм. Тухайн шүүж гаргаж ирсэн мэдээллийг хэрэглэгчийн сонголтоор дамжуулан сэргийлүүлнэ.

Илрүүлж байгаа объект нь эхний ээлжид хязгаарлагдмал байгаа ба тухайн объектыг нэр болон зайн мэдээллийг мэдээлнэ.

Ярих, бичих нь хүмүүсийн текст эсвэл аман хэлбэрээр ашигладаг харилцааны үндсэн хэлбэр юм. Гэсэн хэдий ч ингэх чадваргүй хүмүүс эсвэл амархан уншиж, тайлж чаддаггүй хэл, үг байдаг. Үүний үр дүнд өгөгдсөн текстээс хүний хэлийг задлан шинжилж, ойлгож, үүсгэж чадах системийг зохион бүтээх нь чухал юм. Тиймээс, үүн дээр ажиллахын тулд яриа, хэл боловсруулах энэхүү судалгааны ажил нь компьютерт хүний хэлээр бичсэн эсвэл хэлсэн мэдээллийг ойлгох боломжийг олгоно.

Үндсэндээ машинд хүний хэлийг боловсруулах талаарх тооцооллын загварыг боловсруулдаг бөгөөд энэ нь машинаар хүний яриаг зохиомлоор бий болгож байна гэсэн үг юм. TTS нь бичгийн текстийг яриа болгон хувиргах механизм юм. Энэ арга нь текстийн шинжилгээ (text analysis), ярианы долгион (speech waveform) үүсгэх гэсэн хоёр үндсэн үе ажиллагаатай. Сүүлийн хэдэн жилийн хугацаанд Text-to-Speech технологи нь хиймэл оюун ухаан, машин сургалт зэрэг төрөл бүрийн технологийн тусламжтайгаар хамгийн сайн хувилбар нь гарсаар байгаа. Машин сургалтын технологийг ашиглан TTS дахь яриаг нь хүнтэй төстэй байдлаар дүрслэхэд хувь нэмрээ оруулсаар байгаа. Зах зээл дээр шинэ программууд нэмэгдэж байгаа тул ярианы хэлбэрийн хэрэгцээ улам бүр нэмэгдсээр байгаа.

7-р зураг. Text - to - speech

XI. ХАРЬЦУУЛАЛТ

YOLOv5 загварыг ашиглан объект илрүүлэлт хийхдээ, сургалтын хугацааг хэмнэх зорилгоор урьдчилан сургагдсан параметрүүдийг ашигладаг. YOLOv5 нь PyTorch фреймворк дээр хөгжүүлэгдсэн бөгөөд Google Colab-ийн үнэгүй GPU ашиглан сургалтыг хурдан хийж болдог. MS-COCO зэрэг өгөгдлийн санг Google Drive-д байршуулан Colab-т mount хийж, сургалтын өгөгдөлд обьектуудын координатыг агуулсан байх шаардлагатай.

YOLOv5-д сургалт хийхийн тулд дараах үндсэн тохиргооны файлууд шаардлагатай:

1. **.yaml файл** – сургалтын болон баталгаажуулах өгөгдлийн замуудыг багтаана.
2. **Data.yaml файл** – замын мэдээллийг тодорхойлно.
3. **Загварын тохиргоо** – YOLOv5-н загварын хэмжээнүүд (s, m, l, x) сонгох ба эдгээр нь тухайн моделийн нарийвчлалыг тодорхойлно.

Сургалтын явцад загвар нь batch size, сургалтын хурд зэрэг параметрүүдийг ашиглан суралцана. Сургалтын дараа, загварыг туршилтын зургууд дээр ажиллуулж, нарийвчлалыг хэмжихийн тулд дундаж нарийвчлал mAP, FPS зэрэг үзүүлэлтээр үнэлнэ. Тиймээс дээрх 4-р хүснэгтээс харахад mAP болон FPS үзүүлэлтээс нь хамааралтай YOLOv5 моделийг сонгон, сургасан.

YOLO фреймворк хөгжихийн хэрээр хурд болон нарийвчлалын хоорондын тохиргоо улам нарийссан. YOLOv4 болон YOLOv5 зэрэг загваруудад сүлжээний шинэ арга, өгөгдлийн өргөтгөлийн сайжруулсан аргууд, сургалтын оновчтой стратеги зэрэг шинэлэг өөрчлөлтүүдийг нэвтрүүлсэн нь загварын бодит цагийн гүйцэтгэлийг огцом бууруулахгүйгээр нарийвчлалын өндөр ахиц гаргах боломжийг олгосон. YOLOv5-ээс эхлэн бүх албан ёсны YOLO загварууд нь хурд болон нарийвчлалын хоорондох тохиргоог нарийн тааруулж, тухайн хэрэглээний шаардлага, техник хангамжийн боломжид тохирсон загварын янз бүрийн хэмжээсүүдийг санал болгодог болсон. Тухайлбал, эдгээр хувилбарууд нь захын төхөөрөмжүүдэд зориулагдсан хөнгөн жинтэй загваруудыг санал болгодог бөгөөд эдгээр нь нарийвчлалыг бууруулахын зэрэгцээ тооцооллын нарийн төвөгтэй байдлыг багасгаж, боловсруулалтын хурдыг нэмэгдүүлдэг.

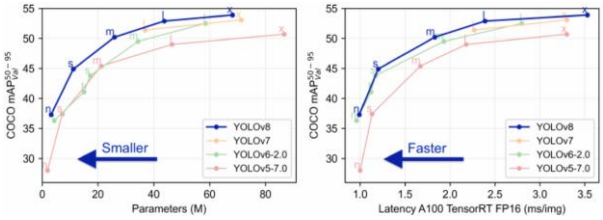


График 1. Гүйцэтгэлийн харьцуулалт

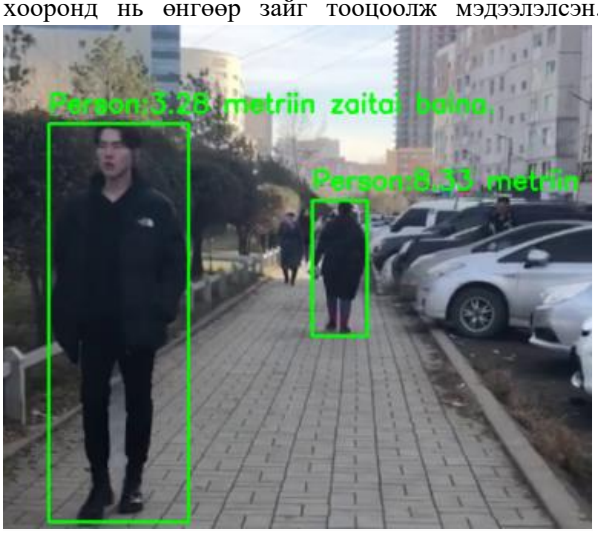
График 1 -т YOLOv5-аас YOLOv8 хүртэлх загваруудын хэмжээсүүдийг харьцуулсан. Энэ график нь нь янз бүрийн YOLO загваруудын нарийн төвөгтэй байдал болон гүйцэтгэлийг харьцуулсан шинжилгээг харуулдаг. Зүүн талын графикт параметрийн тоо (сая нэгжээр) болон COCO баталгаажуулалтын өгөгдлийн багц дээрх дундаж нарийвчлалын утгыг (mAP) IOU босго 50-95-ын хооронд харьцуулсан байна. Энэ нь параметрийн тоо нэмэгдэхийн хэрээр загварын нарийвчлал сайжирдаг чиг хандлагыг тодорхой харуулдаг. Загвар бүрийг nano (n), small (s), medium (m), large (l), extra-large (x) гэсэн хэмжээстүүдээр тэмдэглэсэн.

XII. ХЭРЭГЖҮҮЛЭЛТ

Ухаалаг хөтөч системийн туршилтыг явган хүний зам дээрх тодорхой объектуудыг таних болон объектуудын зайн мэдээллийг аудио хэлбэрээр гаргах зорилгоор туршиж үзсэн. Туршилт нь гурван үндсэн хэсгээс бүрдсэн бөгөөд үүнд объект таних, объектын зайг тооцоолох, мэдээллийг аудио хэлбэрээр дамжуулах туршилтуудыг хийсэн.

8-р зураг. Системийн туршилтын хэсэг

Хэрэгжүүлэх явцад нийт 80 гаруй бодит цагийн бичлэг хийж хэрэгжүүлсэн ба энэ нь объектыг таниж хооронд нь өнгөөр зайг тооцоолж мэдээлэлсэн.



Аудио мэдээ : Хүн 3 метрийн зайд явж байна. Та болгоомжтой явна уу

Дараах байдлаар туршилтуудыг хийж манай системийн одоогийн ажиллагааны хувийг гаргаж ирсэн. Мөн цаашдын анхаарах зүйлсийг танихад тус болсон.

XIII. ҮР ДҮН БОЛОН ИРЭЭДҮЙ

Туршилтын явцад хийгдсэн үйл явц, ажиллагааг нарийвчлан хянаж, алдааны хувиуд дээр үндэслэн харьцуулалтыг хийсэн. Энэ харьцуулалтаас цаашид анхаарах шаардлагатай зүйлс болон ирээдүйн хөгжүүлэлтэд хэрэгжүүлэх шаардлагатай шинэ санаанууд тодорхой болсон юм.

Туршилтын харьцуулалтыг хийхдээ хэд хэдэн үндсэн атрибутуудыг сонгон авч, тэдгээрийг хувиар тооцоолон гаргаж, үр дүнг график хэлбэрээр харуулсан болно. Туршилт нийт 20 бичлэг дээр хийгдэж, тэдгээрээс дараах үндсэн харьцуулалтуудыг гаргаж ирсэн.

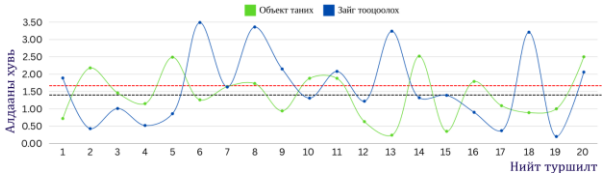


График 2. Харьцуулалт

A. *Ногоон → Объект таних*

Объект таних хурд нь туршилтын голлох үзүүлэлтүүдийн нэг бөгөөд энэхүү атрибутыг хугацаагаар хэмжин, үр дүнг хувиар илэрхийлсэн графикаар харуулсан болно.

Объект таних хурд нь системийн гүйцэтгэлийг тодорхойлж, үр дүнгийн бодит байдалд шууд нөлөөлдөг тул бидний судалгааны чухал хэмжигдэхүүн байлаа.

B. Улаан → Зайг тооцоолох

Зай тооцоолох үзүүлэлт нь хэрэглэгчдэд зөв, бодитой мэдээлэл хүргэхэд нэн шаардлагатай нэгж юм. Зайг зөв тооцоолох чадвар нь системийн найдвартай байдал болон бодит байдлыг баталгаажуулдаг бөгөөд энэ нь хэрэглэгчдийн ашиглах туршлага, үр дүнд шууд нөлөөлдөг. Туршилтын явцад гарсан алдааг илрүүлж, анализ хийснээр алдааны хэмжээг бууруулах, цаашдын хөгжүүлэлтийн суурь нөхцөлийг бий болгосон болно.

Эдгээр үзүүлэлтүүд дээр тулгуурлан хийсэн харьцуулалтууд нь судалгааны үр дүнгийн чанарыг дээшлүүлэх, үр ашигтай хөгжүүлэлтийн чиглэл тодорхойлоход онцгой ач холбогдолтой байлаа.

	Min (%)	Max (%)	Avg (%)	Accuracy (%)
Объект таних	0.24	2.52	1.41	98.58
Зайг тооцоолох	0.2	3.49	1.63	98.36

9-р зураг. Харьцуулалтын хүснэгт

B. Аудио мэдээлэх

Ашиглагдсан API : Бичвэрээс яриа руу хөрвүүлэх, бичвэрийг аудио файл болгон хөрвүүлэхдээ ашигласан. Хөрвөсөн аудио файлын төрөл нь audio/x-wav байна.
Авсан параметрээр хоолойгоо сонгож, speed болон pitch параметрээр ярианы хурд болон өнгийг тааруулсан.
Эдгээр туршилт болон харьцуулалтын үр дүнд манай систем эхний шаардлагын дагуу мэдээлэх чадвар нь 98% дүнтэй гэж үзэж байна.

Бид энэхүү шийдлийг 5-10 жилдээ дэлхийд гаргах боломжтой гэж үзээд эдийн засгийн судалгаа мөн хийсэн. Ирээдүйд бид илүү хялбарчилсан хэрэглээнд суурьлаж нүдний шил дээр камерыг байрлуулах төлөвлөгөөтэй байгаа. Энэ ажлыг ерөнхий зураг дараах байдалтай байна.

Тухайн шилний үзүүлэлт :

- Энэхүү шил ойролцоогоор 110,000 - 150,000 төгрөгөөр бүтнэ.
- Баттерей болон камер гэсэн эд ангитай.
- Монгол хэл дээр ажилдаг.
- Аудио мэдээг хүргэхдээ чихэвчийг ажиглах төсөөлөлтэй байгаа

XIV. ДҮГНЭЛТ

Энэхүү судалгаагаар бид гар утасны камерыг ашиглан орчны объектуудыг таньж, тэдгээрийн зайг тооцоолон, хэрэглэгчид мэдээллийг аудио хэлбэрээр дамжуулах ухаалаг хөтөч системийг хөгжүүлэхэд чиглэсэн цогц аргачлалыг боловсруулаа. Судалгааны явцад системийн гүйцэтгэлийг үнэлэх туршилтуудыг гүйцэтгэх замаар объект таних нарийвчлал, зай тооцоолох чадвар, дуу хоолойн мэдээлэл дамжуулах чанарыг оновчтой үнэлж, цаашдын хөгжүүлэлтэд шаардлагатай шинэ санаа, сайжруулалтын чиглэлүүдийг тодорхойлсон болно.

YOLO загварт суурилсан объект таних алгоритм нь Монгол Улсын явган зорчих орчинд нийтлэг тохиолддог саад, объектуудыг 98.58%-ийн нарийвчлалтайгаар илрүүлсэн нь энэхүү системийн бодит хэрэгцээг хангах үндсэн боломжийг баталгаажуулж байна. Зай тооцоолох модуль 98.36%-ийн нарийвчлалтай ажилласан нь орчныг аюулгүй байдлаар үнэлэхэд чухал дэмжлэг үзүүлж буйг харуулж байна. Мөн аудио мэдээлэл дамжуулахад ашигласан API нь дууны өнгө, хурдыг тохируулах боломжтой байснаар мэдээлэл хүргэлтийн үр ашиг 98%-д хүрч, хэрэглэгчийн ойлголт, хүлээн авалтад эерэг нөлөө үзүүлсэн.

Системийн хоёр дахь хувилбарт дуу хоолойн удирдлагыг амжилттай нэвтрүүлсэн нь хэрэглэгч өөрийн очих газрыг дуу хоолойгоор зааж, систем тухайн байршлыг таньж чиглэл өгөх боломжийг бүрдүүлж байна. Энэ нь харааны бэрхшээлтэй хэрэглэгчдэд бие даан зорчих, үйлчилгээ авах орчныг бүрдүүлж буйгаараа нийгмийн хүртээмж, цахим оролцооны шинэ жишгийг тогтоож байна.

Түүнчлэн судалгааны үр дүнг харааны бэрхшээлтэй хүмүүст зориулсан олон улсын нэр хүндтэй "Seeing AI" системтэй харьцуулан үнэлэхэд, манай систем үндсэн функцүүдийг хангаж чадсан ч гүнзгий тайлбар, нүүр таних, бичвэр унших зэрэг өргөтгөсөн боломжуудаар хязгаарлагдаж байгааг тогтоолоо. Гэсэн хэдий ч манай систем нь монгол хэл дээр суурилсан анхны дуу хоолойгоор удирддаг ухаалаг хөтөч шийдэл болж, үндэсний нөхцөлд чиглэсэн инновацийн шийдлийн эхлэл болж чадсан.

Эдгээр үр дүн нь харааны бэрхшээлтэй иргэдийн амьдрах орчны хүртээмжийг нэмэгдүүлэх, технологийн тусламжтайгаар тэгш оролцоог хангах бодит боломж байгааг нотлон харуулж байна. Судалгаанд тулгуурласан дүн шинжилгээ нь төслийг илүү өргөн хүрээнд хөгжүүлж, нээлттэй өгөгдөл, дуу хоолойн процессинг, гүн сургалтын аргуудыг ашиглан Монгол Улсын нөхцөлд нийцсэн, уян хатан, өргөтгөх боломжтой туслах технологийг бий болгох суурь нөхцөлийг бүрдүүлж өгч байна..

ТАЛАРХАЛ

Энэхүү төслийг хийж гүйцэтгэхэд туслалцаа үзүүлж, их зүйлийг зааж, чиглүүлж өгсөн удирдагч багш Доктор (Ph.D) С.Өлзийбаяр, зөвлөх багш нартаа баярлалаа.

Ном зүй

[1] Kuriakose, B.; Shrestha, R.; Sandnes, F.E. DeepNAVI: A deep learning based smartphone navigation assistant for people with visual impairments. Expert Syst. Appl. 2023

[2] https://mpress.mn/p/2765 - Хараагүй иргэдэд тулгамддаг гол асуудал бол дэд бүтэц

[3] Building_A_Voice_Based_Image_Caption_Generator_with_Deep_Learning, 2021он

[4] Indoor objects detection and recognition for an ict mobility assistance of visually impaired people, 2020.

[5] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[6] Enhancing Object Detection for VIPs Using YOLOv4_Resnet101 and Text-to-Speech Conversion Model

Published online : 2 August 2023

Authors : Academic Editors: Christos Troussas, Cleo Sgouropoulou, Akrivi Krouska, Ioannis Voyiatzis and Athanasios Voulodimos

[7] OpenCV: How-to calculate distance between camera and object using image? - Stack Overflow

[8] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[9] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[10] Харааны бэрхшээлтэй иргэдэд зориулсан таяг Author : Б. Билгүүн 2019 он

[11] Joseph Redmon болон Anelia Angelova, RealTime Grasp Detection Using Convolutional Neural Networks (ICRA), 2015.

[12] A. Quattoni, болон A.Torralba. Recognizing Indoor Scenes. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009.

[13] Tadas Naltrusaitis, Peter Robison, болон LouisPhileppe Morency, 3D Constrained Local Model for Rigid and Non-Rigid Facial Tracking (CVPR), 2012.

[14] Li S.W., Cheng G.F., He S.W., Zhang D. (2019) Research and Trend Analysis of Voice Assistant Capability Evaluation

[15] The development and future of mobile phone intelligent voice assistants" [J]. Communication World

[16] "Status and Prospect of Intelligent Voice Products" [J]. Guangdong Communication Technology

[17] Research on the Development of Voice Assistants in the Era of Artificial Intelligence, 2023

[18] Voice assistants in private households: a conceptual framework for future research in an interdisciplinary field, 2023

Гүн сургалтын арга ашиглан зурагнаас ой модыг илрүүлж тоолох нь

Т. Оюунбаатар
Мэдээлэл компьютерын ухааны тэнхим
Мэдээллийн технологи, электроникийн сургууль
Монгол улсын их сургууль
Улаанбаатар, Монгол
t.oyunbaatar@gmail.com

Т.Цэрэннадмид
Мэдээлэл компьютерын ухааны тэнхим
Мэдээллийн технологи, электроникийн сургууль
Монгол улсын их сургууль
Улаанбаатар, Монгол
tserennadmid@seas.num.edu.mn

Х.Оюундолгор
Мэдээлэл компьютерын ухааны тэнхим
Мэдээллийн технологи, электроникийн сургууль
Монгол улсын их сургууль
Улаанбаатар, Монгол
oyundolgor@seas.num.edu.mn

Хураангуй – **Монгол орны ойн нөөцийг судалж тогтоох, биологийн олон янз байдлыг хамгаалах, ойн менежментийг хийх зэрэг олон үйл ажиллагаанд ойн тооллого чухал үүрэгтэй. Сүүлийн жилүүдэд дэвшилтэт технологи ашиглан асуудлыг шийдвэрлэх боломжууд нээгдэж байна. Тухайлбал, дроноор авсан зургийг гүн сургалтын загвараар боловсруулан мод илрүүлэх, тоолох аргачлалыг түлхүү ашиглах болсон нь өртөг хямд, нарийвчлал өндөртэй шийдэл болж байна. Энэхүү судалгааны ажилд бид дроноор авсан зурагнаас гүн сургалтын YOLOv5 загварыг ашиглан болц гүйцсэн моддыг илрүүлэх, тоолох туршилт судалгааг гүйцэтгэнэ. Бидний дэвшүүлж буй аргачлал нь уламжлалт аргачлалтай харьцуулахад болц гүйцсэн ой модыг бүхэлд нь тоолох боломжийг олгох бөгөөд байгаль орчны талаарх нарийвчилсан мэдээлэл цуглуулахад өндөр ач холбогдолтой болохыг энэхүү судалгааны үр дүнгээр харуулахыг зорьлоо.**

Түлхүүр үг – Ой мод, ойн тооллого, мод илрүүлэх, YOLOv5, дроны зураг

I. УДИРТГАЛ

Ой мод нь экосистемийн тэнцвэрийг хадгалах, нүүрстөрөгчийн хийг шингээж, агаарын чанарыг сайжруулахад онцгой үүрэг гүйцэтгэдэг. Монгол Улсын хувьд ойн нөөцийг үнэлэх, менежмент хийхдээ ойн тооллогыг уламжлалт аргад тулгуурлан гүйцэтгэдэг тул өртөг зардал өндөртэй, цаг хугацаа их шаарддаг, мөн ойг бүхэлд нь хамарсан тооллого хийхэд хязгаарлагдмал байдаг. Иймээс ой модны талаарх иж бүрэн, нарийвчилсан мэдээлэл авахад хүндрэл үүсч, бодитой төлөвлөлт, шийдвэр гаргахад сөрөг нөлөө үзүүлдэг. Тухайлбал, түүвэрлэлтэд суурилсан уламжлалт аргын хувьд ойн санг бүхэлд нь хамруулахад учир дутагдалтай байдаг учраас үр дүнтэй шийдвэр гаргахад хүндрэл учруулдаг [1].

Сүүлийн жилүүдэд дэвшилтэт технологийн тусламжтай энэ асуудлыг шийдвэрлэх шинэ арга замууд нээгдэж байна. Жишээ нь лидар (LiDAR), дроны агаарын зураг, гүн сургалтын алгоритмуудыг ашиглан ойн мэдээллийг автоматаар боловсруулах судалгаа идэвхтэй хөгжиж байна.

Лидар (LiDAR) технологийн хувьд модны тархалт, бүтцийг агаараас тандан судлах, модны гурван хэмжээст бүтцийг илрүүлэх өргөн боломжийг олгох боловч өртөг зардал өндөр, техник боловсруулалт их шаарддаг зэрэг байдлаас шалтгаалан өргөн хэрэглээнд ашиглахад хязгаарлагдмал байна [2]. Харин лидартай харьцуулахад дроны зураг нь өртөг хямд тул энэхүү судалгааны ажилд бид дроноор авсан зурагнаас гүн сургалтын загварыг ашиглан болц гүйцсэн моддыг илрүүлэх, тоолох туршилт судалгааг гүйцэтгэнэ.

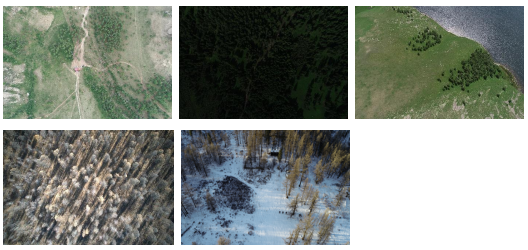
Зурагнаас объектыг илрүүлэх гүн сургалтын олон төрлийн аргууд ашиглагдаж байна. Тухайлбал, YOLOv5 загвар нь бодит цагт өндөр үр дүнтэй ажиллах чадвартай тул энэ аргыг ойн тооллогыг автоматжуулахад шинэ боломжийг нээж байна. Сүүлийн үеийн судалгаануудаас харахад Юань [3] YOLOv4 загварыг ашиглан модны суулгацын өндрийг хэмжиж, тоолох системийг боловсруулсан бол Дима Мохаррам [4] YOLOv5 загварыг ашиглан суулгац таних, ангилах нарийвчлалыг сайжруулсан. Мьянмарын нөхцөлд Htet, K.S. [5] Mask R-CNN алгоритмыг ашиглан далдуу модны ангилал, тооллогыг 80–95% нарийвчлалтай хийсэн. Парк, Жун нар YOLOv5 загвар ашиглан агаарын зурагнаас хотод тарьсан моддыг тоолох судалгааг хийсэн бөгөөд танилтын дундаж нарийвчлал (mAP) нь 0.663 байсан [6]. Харин манай орны хувьд энэ чиглэлийн судалгаа хомс байна.

Энэхүү судалгааны ажлын зорилго нь дроноор авсан агаарын зурагнаас YOLOv5s ба YOLOv5m зэрэг гүн сургалтын загваруудыг ашиглан нас бие гүйцсэн модыг илрүүлэх, тоолох системийг Монгол орны нөхцөлд боловсруулах явдал юм.

II. АРГА ЗҮЙ

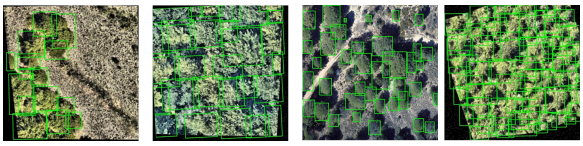
А. Өгөгдөл

Энэхүү судалгаанд Монгол орны 5 өөр байршил, 5 өөр цаг хугацааны дроноор авсан зургууд болох Тэрэлжид байрлах Арьяабалын хийд орчмын ойн зураг, Өвөрхангай аймгийн Бат-Өлзий сумд байрлах Төвхөн хийдийн орчмын ойн зураг, Уянга сумд байрлах Хүйсийн найман нуурын эрэг дэх ойн зураг, Төв аймгийн Батсүмбэр сумд байрлах МУИС-ийн Ой судлалын баазын орчмын ойн зураг (намар), Төв аймгийн Гүнжийн сүм орчмын ойн зураг (өвөл) буюу нийт 222 ширхэг зургийг цуглуулан ашигласан. Модны харагдах байдал бүх улиралд өөр өөр байх тул шим тэжээл сайтай үе болох зун болон намрын 158 зургуудыг сонгон авсан.



Зураг 1. Дроноор авсан Монгол орны ойн зураг

Мөн Roboflow [7] нээлттэй сангаас нэмэлтээр 6761 ширхэг модны өгөгдлийг бүтнээр нь болон түүвэрлэн ашигласан.



Зураг 2. Нээлттэй өгөгдөл

В. Өгөгдлийн нэмэгдэл (Image augmentation):

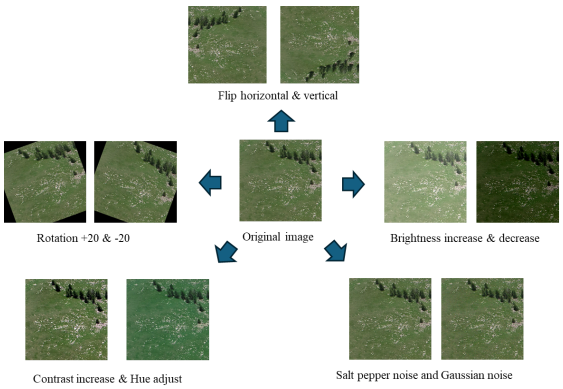
Хэт тохируулгыг багасгахын тулд зургийг эргүүлэх эсвэл бага зэрэг өөрчлөх замаар өгөгдлийн багцын олон янз байдлыг нэмэгдүүлэхийн тулд

- Геометрийн хувиргалт: Хэвтээ/босоо толин буулгалт (flip), 90°-180° эргүүлэлт (rotation).
- Гэрэл, өнгөний өөрчлөлт: Тодосголт (brightness), контраст (contrast), hue.
- Шуугиан нэмэх: давс буюу перцийн шуугиан (Salt & Pepper Noise), Гауссын шуугиан (Gaussian Noise).
- Бусад: Ойртуулах/Холдуулах (Zoom in/Zoom out)

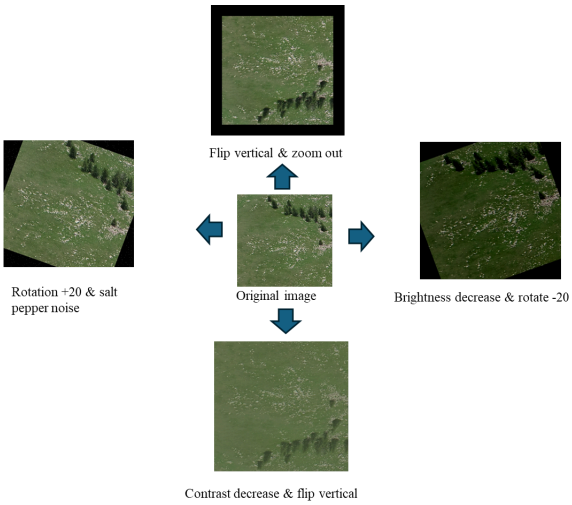
зэрэг аргуудыг ашигласан бөгөөд дараах давуу талуудтай. Үүнд:

- ➢ Санамсаргүй эргүүлэх: Дроны янз бүрийн өнцөг, чиглэлийг дуурайлгана.
- ➢ Томруулах/жижигрүүлэх: Хол, ойр харааны өөрчлөлтийг дуурайлгана.
- ➢ Толин буулгалт (Хэвтээ/Босоо): Камераас үүдэлтэй толин тусгал болон тэгш хэмийг тооцоолно.
- ➢ Гэрлийг тодруулах: Гэрэлтүүлгийн өөрчлөлт, сүүдэр болон өдрийн цагийг тооцоолно.
- ➢ Зургийн нягтралыг өөрчлөх: Цаг агаарын нөхцөл байдал болон зургийн чанарын өөрчлөлтийг тооцоолно.
- ➢ Гауссын шуугиан: Дроны мэдрэгчээс гарах шуугианыг дуурайлгана.
- ➢ Бүдэгрүүлэх (хөдөлгөөний бүдэгрэл): Дроны хөдөлгөөнөөс үүдэлтэй нөлөөг тооцоолно.
- ➢ Өнгөний хэлбэлзэл: Улирлын өөрчлөлт, ургамлын өнгөний ялгааг дуурайлгана.

Дээрх аргуудыг ашиглан өгөгдлөө хэрхэн баяжуулсан үр дүнг Зураг 3,4-т харуулав.



Зураг 3. Image augmentation 1

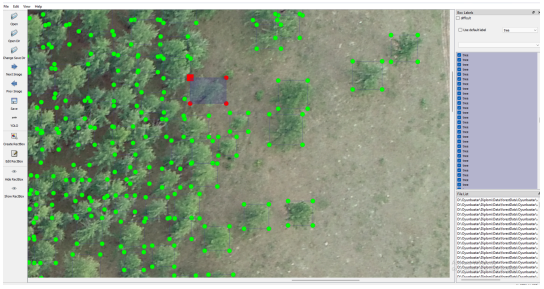


Зураг 4. Image augmentation 2 *(санамсаргүй байдлаар өгөгдлөө нэмэгдүүлсэн байдал.)*

С. Тэмдэглэл хийх(Annotation):

Чиглүүлэгтэй машин сургалтын загварт зориулан сургалтын өгөгдлийн хувьд модны байрлал, хил хязгаарын тэмдэглэгээг гараар тэмдэглэх шаардлагатай болсон тул бид LabelImg tool-ийн

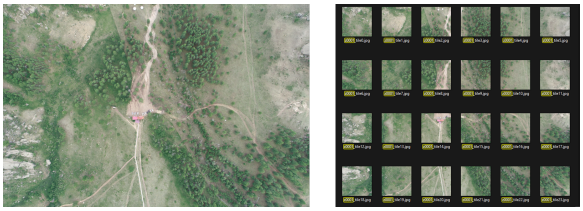
тусламжтай хязгаарлагч хайрцаг (bounding box) ашиглан моднуудын тэмдэглэгээг гараар хийсэн.



Зураг 5. LabelImg-р хийсэн тэмдэглэл

D. Зургийг хэсэгчлэн хуваах (Image Tiling):

Зурагнаас овор хэмжээ багатай объектыг илрүүлэхэд зургийн нягтршил чухал үүрэгтэй. Бидний дроноор (DJI Phantom 4 Pro) цуглуулж авсан зургууд 5472x3648 пикселийн нягтралшилтай байсан тул зургийг хэсэгчлэн хуваах аргыг ашиглан зургуудын чанарыг алдагдуулахгүйгээр 912x912 хэмжээтэй байхаар хэсэгчлэн хуваасан.



Зураг 6. Image Tiling

E. Өгөгдөл хуваалт(Data split):

Нэмэгдүүлсэн өгөгдөл болон нээлттэй өгөгдлийн нийлбэр нь 1333 + 6761 = 8094 болсон бөгөөд үүнээс:

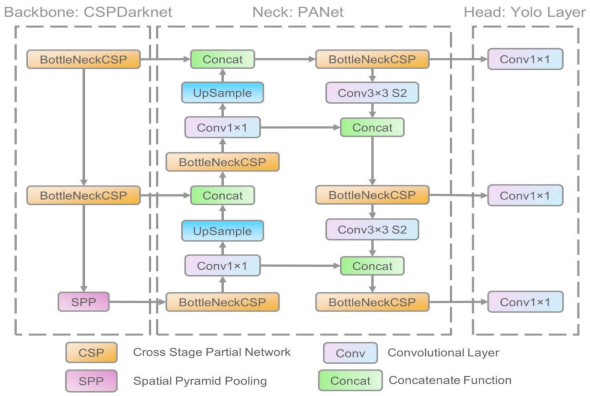
(a) 80% (train)-1066, 10% (valid)-133, 10% (test)-134

(b) 80% (train)-6474, 10% (valid)-809, 10% (test)-811

байхаар өгөгдлийг хуваасан.

F. YOLOv5 model

YOLO буюу You Only Look Once [8] загвар нь объект илрүүлэх нэг шатат загвар (single stage object detector) юм. Зураг 7-т харуулснаар YOLOv5 загвар нь CSP DarkNet53 суурийн сүлжээг ашиглах ба өмнөх загваруудтай харьцуулахад тооцооллын зардлыг бууруулж, илрүүлэлтийн хурдыг дээшлүүлсэн [10]. Энэхүү архитектур нь давхаргын хооронд градиентын мэдээллийг хуваалцах замаар давхардсан тооцооллыг арилгадгаараа онцлогтой.



Зураг 7. YOLOv5 architecture [9]

Мөн загварын "хүзүү" хэсэгт PA-Net (Path Aggregation Network)-ийг ашиглан FPN-ийн нэг чиглэлийн мэдээлэл дамжуулалтын дутууг нөхөх шийдлийг оруулсан бөгөөд өөр өөр масштабын хувьд ч шинж чанарыг интеграцалж өгдөг.

YOLOv5 нь хамгийн жижгээс хамгийн том хүртэлх 5 өөр хувилбартай. Тухайлбал, YOLOv5n (хамгийн жижиг), YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x (хамгийн том). Хувилбар томрох тусам нарийвчлал нэмэгдэх боловч хурд нь буурдаг [11].

G. Үнэлгээний метрик (Evaluation metrics)

mAP: (mean Average Precision) Дундаж нарийвчлал

train/box_loss: Сургалтын явцад (Bounding Box) илүү нарийвчлалтай тодорхойлоход гарч байгаа алдааг илэрхийлнэ.

train/obj_loss: Загвар обьект байгаа эсэхийг хэрхэн зөв тодорхойлж байгааг харуулдаг

metrics/precision: Нарийвчлал нь зөв илэрхийлсэн эерэг таамаглалуудын нийт эерэг таамаглалд эзлэх хувь юм.

metrics/recall: Энэ нь бодит эерэгүүдээс зөв илэрхийлсэн эерэгүүдийн харьцааг илэрхийлнэ.

val/box_loss: train/box_loss төстэй боловч баталгаажуулалтын өгөгдөлд Bounding Box гарч байгаа алдааг үнэлнэ.

mAP_0.5: IoU(Intersection over Union) 0.5 босго, объект илрүүлэх загварт зориулсан нийтлэг үнэлгээний хэмжүүр юм.

mAP_0.5:0.95: 0.5-0.95 (0.05 алхамаар) хооронд дундажлагдах бөгөөд гүйцэтгэлийн хамгийн хатуу цогц хэмжүүрийг өгдөг тул ихэвчлэн "эцсийн нарийвчлал" гэж тооцогддог.

III. ҮР ДҮН БА ХЭЛЭЛЦҮҮЛЭГ

Туршилтыг CPU: I7-10700 @ 2.90 GHz, RAM: 32GB, GPU: NVIDIA GeForce RTX2060 техникийн үзүүлэлт бүхий компьютер дээр явуулав.

Онцлог ихтэй зургаан туршилтын үр дүнгийн графикийг Хүснэгт 1-д харуулав.

Exp43 - Зургийн нягтрал нь 1920x1920(resize) байх үед сургасан. Зургийн нягтрал өндөр байсан тул сургалт хийхэд 12.9 цаг шаардагдсан. Нарийвчлал нь (mAP) 44.8%.

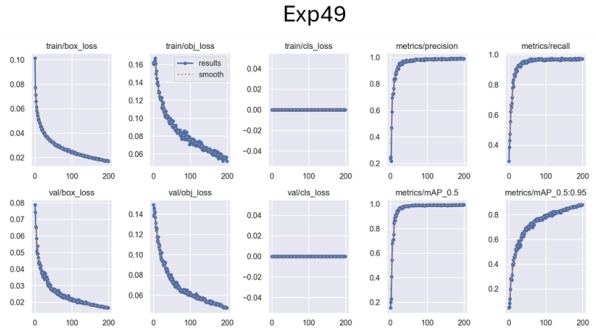
Exp45 - Зургийн нягтралыг 640x640 болгон сургасан. Сургалтын хугацаа хурдан боловч нарийвчлал нь 26.9%.

Exp46 - YOLOv5m дунд загвар байсан тул нарийвчлал 34.6% буюу бага зэрэг сайжирсан.

Exp48 - Зургийн нягтралыг 1280x1280 болгож, хуваалт хийсэн зургуудаа нэмж сургасан. Нарийвчлал нь 57.3%.

Exp49 - Дан өөрийн өгөгдлөө хуваан (tile) зургийн чанарыг алдагдуулахгүйгээр, 1066 зургийн нэмэгдэл (augmentation) сургахад нарийвчлал нь 87.9% болсон.

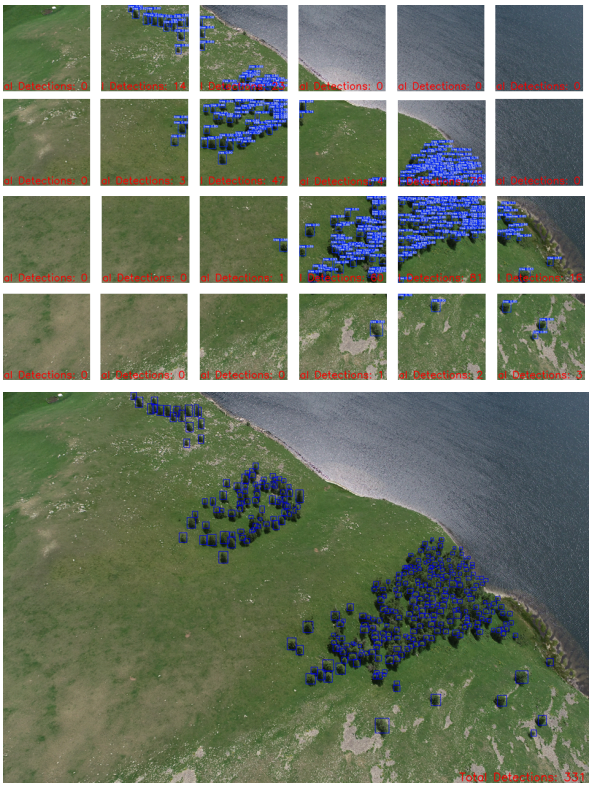
Exp53 – Бодит өгөгдлөө нээлттэй өгөгдөлтэй нийлүүлэн сургасан. 6474 өгөгдлөөр сургахад нарийвчлал нь 94.2% болсон.



Зураг 8. Туршилт 49*(Exp49)* үр дүнгийн график

Эхний 2 туршилт дээр зургийн нягтралуудыг 1920x1920, 640x640 байх үед сургалтыг хийхэд зарцуулах хугацаа жижиг хэсгүүдэд хуваан сургалт хийх хугацаанаас 30 дахин их байсан. Сүүлийн 2 туршилтыг өгөгдлийг нэмэгдүүлэх (augmentation) аргыг ашиглан явуулсан учраас эхний 4 туршилтаас хангалттай сайн нарийвчлалтай байсан. Өгөгдлийн тоо ихсэх, зургийн нягтрал эсвэл YOLO загвар томрох тусам нарийвчлал сайжрах боловч, дагаад сургалтын хугацаа болон хурд удааширч байсан. Тиймээс зургийг хэсэгчлэн хуваах (image tiling) аргаар зургийн нягтралыг хамгийн бага буюу 640x640 пиксель байх үед хурдан, үр дүн сайн байсан.

Зураг 9. Эхний зураг зургийн нэмэгдэл (*augmentation*) хийхээс өмнө, 2 дахь зураг зургийн нэмэгдэл хийсний дараах үр дүн



Зураг 10. Нэгтгэсний дараах эцсийн үр дүн

IV. ДҮГНЭЛТ

Энэхүү судалгааны ажилд бид дроноор авсан зурагнаас гүн сургалтын YOLOv5 загварыг ашиглан болц гүйцсэн моддыг илрүүлэх, тоолох туршилт судалгааг гүйцэтгэсэн. Судалгааны үр дүнгээс харахад дроноор авсан зурагнаас бид болц гүйцсэн моддыг 87.9 - 94.2 хувийн нарийвчлалтай илрүүлсэн. Зургийг хэсэгчлэн хуваах болон augmentation хийх аргуудыг оновчтой хослуулах нь загварын үр дүнг 30 хувиар сайжруулсан.

Хүснэгт 1. Туршилтын үр дүн

Туршилт	Epoch	Obj_ loss	Preci sion	Recall	mAP 0.95	Total_ epoch	Img_size	Num_img	Type Yolo	Train time
exp43	188	46.518	90.719	53.48	44.828	200	1920x1920	364	small	12.950 hours
exp45	170	14.788	64.976	33.24	26.91	200	640x640	364	small	0.460 hours
exp46	189	13.053	79.53	41.04	34.593	200	640x640	364	medium	0.660 hours
exp48	118	18.734	86.354	75.69	57.342	200	1280x1280	445	medium	4.605 hours
exp49	191	4.7655	98.946	97.2	87.93	200	640x640	1066	medium	1.764 hours
exp53	159	6.6692	99.68	99.06	94.243	200	640x640	6474	small	7.711 hours.

Бидний дэвшүүлж буй энэхүү аргачлал нь уламжлалт аргачлалтай харьцуулахад болц гүйцсэн ойн моддыг бүхэлд нь тоолох боломжийг олгох бөгөөд байгаль орчны талаарх нарийвчилсан мэдээлэл цуглуулахад өндөр ач холбогдолтой болохыг харуулж байна. Цаашид дэвшилтэт технологид суурилсан энэ мэт шийдлүүдийг хэрэглээнд нэвтрүүлэх, зурагнаас объект илрүүлэх бусад аргууд тухайлбал, Mask-RCNN, EfficientDet зэрэг бусад архитектуртай харьцуулан судлах өргөн боломжтой юм. Ирээдүйд ойн нягт байдал, моддыг төрөл зүйлээр ангилах, сегментчилэх аргыг хэрэгжүүлэх, мөн 3 хэмжээст цэгэн өгөгдөлтэй (LiDAR) интеграцчилсан системийг боловсруулснаар илрүүлэлтийн нарийвчлалыг улам боловсронгуй болгох боломжтой юм.

НОМ ЗҮЙ

[1] Олон зорилтот ойн тооллогын хээрийн хэмжилтийн аргачлал 2014 он

[2] Park, C.W. and Jung, H.S., 2022. Detection of urban trees using YOLOv5 from aerial images. Korean Journal of Remote Sensing, 38(6_2), pp.1633-1641.

[3] Yuan, X., Li, D., Sun, P., Wang, G. and Ma, Y., 2022. Real-Time Counting and Height Measurement of Nursery Seedlings Based on Ghostnet–YoloV4 Network and Binocular Vision Technology. *Forests*, *13*(9), p.1459. https://www.mdpi.com/1999-4907/13/9/1459.

[4] Moharram, D., Yuan, X. and Li, D., 2023. Tree seedlings detection and counting using a deep learning algorithm. *Applied Sciences*, *13*(2), p.895. https://www.mdpi.com/2076-3417/13/2/895.

[5] Htet, K.S.; Sein, M.M. Toddy Palm Trees Classification and Counting Using Drone Video: Retuning Hyperparameter Mask-RCNN. In Proceedings of the 2021 7th International Conference on Control, Automation and Robotics (ICCAR), Singapore, 23–26 April 2021; pp. 196–200.

[6] Park, C.W. and Jung, H.S., 2022. Detection of urban trees using YOLOv5 from aerial images. Korean Journal of Remote Sensing, 38(6 2), pp.1633-1641.

[7] https://universe.roboflow.com/csi-sgqu2/trees-counting/dataset/2/images

[8] https://iq.opengenus.org/yolov5/

[9] Xu, R., H. Lin, K. Lu, L. Cao, and Y. Liu, 2021. A forest fire detection system based on ensemble learning, Forests, 12(2): 217. https://doi.org/10.3390/f12020217

[10] Wang, C. et al. "CSPNet: A New Backbone that can Enhance Learning Capability of CNN", arXiv:1911.11929, 2021.

[11] https://github.com/ultralytics/yolov5

[12] https://blog.roboflow.com/yolov5-improvements-and-evaluation/

Микроскопын дүрснээс Enterobius vermicularis-ийн өндөгийг гүн сургалт ашиглан илрүүлэх нь

Тэгшээгийн Түвшинсайхан
Мэдээллийн технологийн салбар
ШУТИС, МХТС
Улаанбаатар, Монгол
tuvshinsaikhan@gmail.com

Доржсүрэнгийн Тэмүүлэн
Биологийн тэнхим
АШУҮИС,
Био-Анагаахын сургууль
Улаанбаатар, Монгол
temuulen@mnums.edu.mn

Наранцогтын Гиймаа
Биологийн тэнхим
АШУҮИС,
Био-Анагаахын сургууль
Улаанбаатар, Монгол
giimaa.n@mnums.edu.mn

Батжаргалын Долгорсүрэн
Мэдээллийн технологийн салбар
ШУТИС, МХТС
Улаанбаатар, Монгол
b.dolgorsuren@must.edu.mn

Хураангуй - Монгол улсын хүн амд шимэгчдээр үүсгэгдэх халдвараас хамгийн өргөн тархалттай нь Enterobius vermicularis (E.vermicularis) буюу хүн амын дундах халдварын 80 орчим хувийг дангаараа эзэлдэг. Дэлхий нийтэд E.vermicularis (Хялгасан хорхой)-ийг илрүүлэхэд Tape scotch аргыг өргөн ашигладаг. Гэсэн хэдий ч энэхүү арга нь шинжилгээ хийж буй хүнээс их цаг хугацаа, чадвар шаарддаг, олон шинжилгээ хийх үед лабораторийн ажилтнуудад ачаалал үүсгэн, ажлын стресст өртүүлэх, улмаар шинжилгээний хариу буруу тавигдах магадлалд хүргэдэг.

Иймд You Only Look Once(YOLO) гүн сургалтын загварыг ашиглан E.vermicularis-ийн өндгийг автоматаар илрүүлж, тоолох аргыг боловсруулах нь олон хүний шинжилгээг богино хугацаанд мэдрэг болон өвөрмөц чанар өндөртэй хийж, хүний оролцоог багасган, хариуг хурдан гаргаж өгөх боломж бүрдэж байна. Сургалтад ашигласан датасет буюу микроскопоос гарган авсан зурган өгөгдөл тус бүрд хялгасан хорхойн өндөгнүүд, наалдамхай туузаас үүссэн хийн бөмбөлөг, хувцасны утас, бусад паразитын үүсгэгч зэрэг хуурамч элементүүд (artifacts) багтаав.

YOLO8 загвар нь 92.5%-ийн дундаж нарийвчлалтай (mAP) E.vermicularis-ийн өндгийг илрүүлж, тоолж байгаа нь уламжлалт оношилгооны аргуудтай харьцуулахад мэдрэг болон өвөрмөц чанар илүү болох нь батлагдлаа. Бидний боловсруулсан хиймэл оюунд тулгуурласан энэхүү аргачлал нь бодит цаг хугацаанд өндөр гүйцэтгэлтэй оношилгоо хийж, халдварын экстенсив болон интенсив чанарыг тогтоох боломж олгож байна.

Түлхүүр үг: цагаан хорхой, хялгасан хорхой, yolo8, гүн сургалт, дүрс боловсруулалт

I. Удиртгал хэсэг

E. vermicularis-аар үүсгэгддэг "Хялгасан хорхойтох" өвчин нь манай орны хүн амын дунд ялангуяа цэцэрлэгийн насны хүүхдүүдэд өргөн тархсан бөгөөд энэхүү өвчин үүсгэгчийг Монголчууд "Цагаан хорхой" хэмээн нэрлэдэг. Хүний бүдүүн болон мухар гэдсэнд тодорхой хугацаанд шимэгчлэн амьдарч буй эр, эм E. vermicularis нь эвцэлдэн, улмаар эмэгчин хорхой өндгөө гаргахаар бүдүүн гэдсээр дамжин шулуун гэдэсний амсарт гарч ирээд, өөрөөсөө наалдамхай шингэн ялгаруулж, халдварлагчдын ануст өндгөө наалдуулдаг нь халдвар тараах үндсэн шалтгаан болдог. Иймд халдварлагчдад бие тавгүйтэх, загатнах шинж тэмдэг илэрхий мэдрэгддэг. Мөн нойр муудах, хоолны дуршил буурах, хэвлийгээр өвдөх зэрэг шинж тэмдгүүд илэрдэг [1]. E.vermicularis-ийн өндөг нь 50-60 микрометр(μm) урт, 20-30 μm өргөн тул нүдэнд харагддаггүй. Иймд хялгасан хорхойн өндөг халдвартай хүмүүсийн бохирлогдсон гар, хувцас, ор дэрний хэрэглэл, тоглоом, хоол хүнсээр дамжиж хүнээс хүнд халдварлахаас гадна тухайн халдвартай хүн өөрөө дахин авдаг. E. vermicularis-ийн халдварлалтын түвшин халдвар тархах магадлал [3] нь халдвартай хүнээс гадагшилж байгаа өндөгний тоо, халдвар тархсан орчны эрүүл ахуйгаас хамаадаг. Одоогийн байдлаар манай оронд E.vermicularis өндгийг илрүүлэхэд хошногоны амсраас арчдас авч, микроскопоор шинжилдэг арга нь уламжлал болон ашиглагддаг. БНСУ E.vermicularis-ийн халдвар илрүүлэх Tape scotch аргыг [25] 1975 оноос өргөнөөр ашиглаж ирсэн бөгөөд Монгол улсад энэхүү аргыг анх 2008 оноос ашиглаж байгаа ч, Эрүүл мэндийн салбарт өргөн хэрэглэдэггүй байсаар л байна. Хэдийгээр Tape scotch арга нь Монгол улсад уламжлал болгон ашигладаг арчдасын аргаас мэдрэг чанар өндөртэй [4], тул давуу талтай ч, бас л шинжилгээ хийж буй хүнээс их цаг хугацаа, чадвар шаарддаг, олон хүний шинжилгээ хийх үед лабораторийн ажилтнуудад ачаалал их үүсгэн, ажлын стресст өртүүлэх зэрэг хүчин зүйл нь хэвээр байна. Иймд бидний судалгааны зорилго нь E. vermicularis-ийг илрүүлэх Tape scotch-ийн шинжилгээг микроскопоор шалгах явцад харах талбай дахь өндөгний дүрсийг YOLO [5] заг-

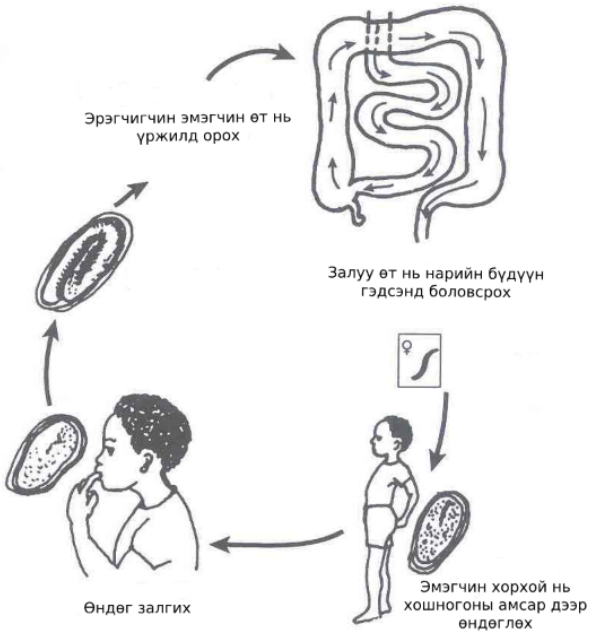
варыг ашиглан автоматаар илрүүлж, шинжилгээний дээж дэх өндгийг тоолох загвар боловсруулах явдал юм. YOLO8 нь зургийг бүхэлд нь нэгэн зэрэг боловсруулж, объектын байршил болон ангиллыг нарийн тогтоож, илрүүлэлтийн хурд, нарийвчлалыг сайжруулах [6] боломжийг олгодог. Энэ системийг хөгжүүлснээр E. vermicularis-ийн өндөг илрүүлэх олон хүний шинжилгээг богино хугацаанд мэдрэг болон өвөрмөц чанар өндөртэй хийж, хүний оролцоог багасган, хариуг хурдан гаргаж өгөх боломж бүрдэж байна. Цаашилбал шинжилгээний өртөг ч багасах боломжтой гэж үзэж байна.

II. Судлагдсан байдал

Олон улсын судалгаагаар [6] АНУ-д жилд 40 сая орчим хүн E. vermicularis-ийн халдвар өртдөг гэж тооцоологддог бөгөөд үүний дийлэнх нь 5–14 насны хүүхдүүд байна. Европын орнууд болон Азийн зарим бүс нутагт ч [7] мөн адил өндөр тархалттай байгаа нь харьцангуй түгээмэл, илүү анхаарал шаардсан халдварт тооцогддог. Монголд хийгдсэн судалгаагаар [8] ч E. vermicularis-ийн халдвар хамгийн түгээмэл байгааг харуулсан. Д.Ганболд (1990) нарын судалгаанд Улаанбаатар хотын хүүхдийн дунд халдварын тархалт 40.2% байсан бол үүнээс 82.5%-ийг 4–7 насныхан эзэлж байжээ. Мөн Н.Гиймаа (2008) нарын судалгаагаар [9] энтеробиозын тархалт 46.55±0.13% байжээ. Тус судлаачид бага насны хүүхдүүдийн дунд ариун цэврийн дадал бүрэн хэвшээгүй байгаа нь бие биедээ халдвар тараах боломж олгож байна хэмээн дүгнэжээ. Өндөр хөгжилтэй орнуудын E. vermicularis-ийн оношилгоонд хамгийн өргөн хэрэглэгддэг арга нь Грэхэмийн наалдуулагч туузны арга буюу (Scotch tape test) юм. Энэхүү арга [10] нь энгийн, хямд үнэтэй боловч шинжлүүлэгчдээс сорьц зөв аваагүй тохиолдолд оношилгооны мэдрэг чанар багасаж, хуурамч сөрөг хариу гарах магадлалтай. Иймээс нэг шинжлүүлэгчээс 14 хоногийн зайтай 3–5 удаа дараалан сорьц авахыг зөвлөдөг ч энэ нь лабораторийн ажилтны ачааллыг нэмэгдүүлж, оношилгооны хариу гарах хугацааг сунгадаг болох нь [11] батлагджээ. Мөн микроскопийн сорьцод агаарын бөмбөлөг, хувцасны утас, ялгадасны үлдэгдэл зэрэг хольц их байх нь оношилгоог будлиулах, өндөгтэй андуурах эрсдэл нэмэгддэг [12] гэж зарим судлаачид үзэж байна. Сүүлийн жилүүдэд энэхүү асуудлыг шийдвэрлэх үүднээс хиймэл оюун ухаан болон объект илрүүлэлтийн гүн сургалтын арга болох YOLO зэрэг загваруудыг ашиглан автомат оношилгооны аргыг хөгжүүлэх оролдлогууд хийгдэж эхэлж байна. YOLOv8 [13] загвар нь хурд, нарийвчлал сайтай тул микроскопийн зурагт хялгасан хорхойн өндгийг өндөр нарийвчлалтайгаар илрүүлж, онош тавих шинэ боломж олгож байгаа юм. Хиймэл оюун машин сургалтын аргуудыг оношилгоонд ашиглах нь [14] хүний алдаа, цаг хугацааны хязгаарлалтыг багасгах давуу талтай бөгөөд оношилгооны үр дүнг сайжруулахад чухал хувь нэмэр оруулах юм.

III. Онолын хэсэг

E. vermicularis-аар халдварлагчдын 70–90 хувийг 5–12 насны хүүхдүүд эзэлдэг [15] гэсэн судалгаа хийгджээ. Нас бие гүйцсэн эмэгчин хорхой 8–13 мм урттай бөгөөд шөнө хүнийг унтаж байх үед хошного орчмоор гадагшилж 10,000–15,000 ширхэг өндөг гаргадаг. Зураг 1- д уг өвчний амьдралын мөчлөгийг харууллаа. Өндөг нь наалдамхай гадаргуутай учраас арьсанд амархан наалддаг бөгөөд хэдхэн цагийн дотор халдварлах чадвартай болдог. Халдварлагчдад хамгийн түгээмэл шинж тэмдэг бол шөнийн цагаар хошного загатнах бөгөөд энэ нь халдварыг тараах гол зам юм. Хүн маажих үед өндөг нь хумс, гар, хувцас болон ор дэрний даавуугаар дамжин бусад хүмүүст халдварладаг. E. vermicularis-ийн өндөг нь агаараар тархаж, хүн амьсгалах үед залгигдан халдвар тархах боломжтой хэмээн зарим судлаачид өгүүлжээ. Жишээлбэл, АНУ-ын Өвчний Хяналт, Урьдчилан Сэргийлэх Төвийн мэдээлснээр [16], "Өндөг нь жижиг хэмжээтэй тул агаарт хөөрч, амьсгалаар залгигдах боломжтой"гэжээ. Мөн UpToDate эмнэлзүйн мэдээллийн санд, "Өндөг нь агаарт хөөрч, амьсгалаар залгигдан халдвар үүсгэж болно"гэж тэмдэглэсэн байна. Халдвартай хүүхдүүд гэр бүл болон хамт олондоо халдвараа тараадаг.



Зураг 1. Цагаан хорхойн амьдралын мөчлөг

A. Паразит судлалд ML ба DL-ийн хөгжил

Сүүлийн жилүүдэд машин сургалт (ML) болон гүн сургалт (DL)-ын технологиудын хурдацтай хөгжил нь Анагаах ухааны оношилгоонд тэр дундаа шимэгчдийн халдварыг илрүүлэхэд томоохон дэвшил гаргаж байна. ML-ийн аргачлалуудыг ашиглан Plasmodium төрлийн

үүсгэгч [18] [19], мөн гэдэсний гельминтүүдийг [20] өндөр үр дүнтэй илрүүлсэн судалгааны өгүүллүүд байна. Түүнчлэн, конволюцийн нейрон сүлжээнүүд (CNNs) нь өтгөн дэх паразитын өндгийг мэдрэг чанар өндөртэй илрүүлэхэд амжилттай ашиглагдаж байна [21]. Цистик эхинококкозын оношилгоонд ML-ийн сүүлийн үеийн аргуудыг нэвтрүүлснээр оношийн нарийвчлалыг нэмэгдүүлж, онош тавих эмчийн гаргах алдааг бууруулах боломжийг тодорхойлсон судалгаанууд [22] хийгдсэн байна. Эдгээр ололт амжилт нь хиймэл оюун ухааны тусламжтайгаар паразитын илрүүлэлтийг уламжлалт микроскопын шинжилгээнээс илүү хурдан, найдвартай болгох, улмаар оношилгооны үйл явцыг хөнгөвчилж, үр ашгийг мэдэгдэхүйц нэмэгдүүлэх бололцоог харуулж байна

B. E.vermicularis өндөг илрүүлэхэд ML ашиглах нь

Бид YOLOv5, YOLOv8 болон YOLOv12 загваруудыг харьцуулан хамгийн мэдрэг чанар өндөр үр дүн гаргасан гүн сургалтын автоматжуулсан илрүүлэх системийг санал болгож байна. Энэхүү системийн гол зорилго нь шимэгчийн халдварын оношилгоог сайжруулахад оршиж, E.vermicularis-ийн өндөг болон түүнтэй төстэй бусад артефактуудыг мэдрэг ялган таньж, лабораторийн эмч нарын ачааллыг бууруулж, алдааг багасгах явдал юм. Мөн энэ систем нь үр ашигтай, нарийн мэдрэмжтэй, өргөн хүрээтэй хэрэглээнд тохиромжтой шийдэл болж, оношлогооны процессыг хурдтай, нарийвчилсан болгох зорилготой.

C. ML загварын сонголт: YOLO хувилбарууд

Энэхүү судалгааны даалгаварт зориулан бид YOLOv5, YOLOv8, болон YOLOv12 загваруудыг ашиглсан. Загваруудын дунд YOLOv8 нь хамгийн сайн үр дүнг үзүүлсэн бөгөөд илрүүлэлтийн хурд болон нарийвчлалаараа хамгийн өндөр үзүүлэлттэй байсан.

- YOLOv5 нь хурдтай бөгөөд энгийн орчинд сайн ажилладаг. Энэ загвар нь хурдтай боловч нарийвчлал нь зарим нарийн төвөгтэй орчинд хязгаарлагдмал байв.
- YOLOv8 нь YOLOv5-ийн сайжруулсан хувилбар бөгөөд том өгөгдөл болон нарийн төвөгтэй объектуудтай илүү нарийвчилсан гүйцэтгэлтэй. Энэ загвар нь хамгийн сайн үр дүнг гаргаж, нарийвчлал болон илрүүлэлтийн хурдны тэнцвэрийг хангаж чадсан.
- YOLOv12 нь хамгийн сүүлийн үеийн технологийг ашигладаг ч, том архитектуртай байсан тул их хэмжээний нөөц шаарддаг. Энэ нь өндөр гүйцэтгэлтэй боловч илүү их хүчин чадал бүхий компьютер хэрэгтэй болсон. Үүнтэй холбоотойгоор YOLOv8 загвар нь тухайн орчинд хамгийн сайн үр дүнг үзүүлсэн бөгөөд нарийвчлал болон хурдны тэнцвэртэй хялбар тохиргоо олсон.

YOLOv12 нь микроскопийн нарийн төвөгтэй орчныг боловсруулахад хамгийн өндөр үр дүнтэй байсан ч, шаардлагатай нөөцийн хэмжээ нь түгээмэл хэрэглээний хувьд сөрөг нөлөө үзүүлж байв. Харин YOLOv8 нь хамгийн үр дүнтэй, нарийвчилсан илрүүлэлт хийж, лабораторийн эмч нарын ачааллыг бууруулж, алдааг багасгах боломжийг олгосон.

Үзүүлэлт	Утга
Алгоритм	YOLOv4
Нарийвчлал (Precision)	89.31%
Мэдрэмж (Recall)	95.83%
F1-үнэлгээ	92.46%
IoU (дундаж)	–
Тайлбар	Бодит цагийн илрүүлэлтэд тохиромжтой
Алгоритм	Xception (CNN)
Нарийвчлал (Precision)	99%
Мэдрэмж (Recall)	99%
F1-үнэлгээ	99%
IoU (дундаж)	0.6136
Тайлбар	Өндөр нарийвчлал, сайн локализаци
Алгоритм	YOLOv4-Tiny
Нарийвчлал (Precision)	96.25%
Мэдрэмж (Recall)	95.08%
F1-үнэлгээ	95.66%
IoU (дундаж)	–
Тайлбар	Хөнгөн, хурдан ажиллагаатай
Алгоритм	ResNet (CNN)
Нарийвчлал (Precision)	–
Мэдрэмж (Recall)	–
F1-үнэлгээ	–
IoU (дундаж)	55.48%
Тайлбар	Дунд зэргийн үр дүнтэй

Хүснэгт I
Гүн сургалтын алгоритмуудын E. vermicularis болон бусад шимэгчийг илрүүлэх үр дүнгийн дэлгэрэнгүй жагсаалт

IV. Туршилтын хэсэг

A. Өгөгдөл бэлдэх аргачлал

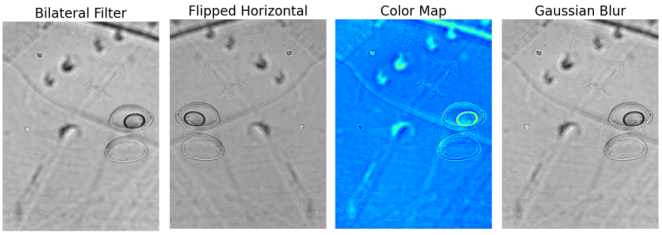
Бид АШУҮИС-ийн Биологийн тэнхимийн багш нарын судалгаагаар E.vermicularis-ийн халдвартай хэмээн оношлогдсон 5 хүний шинжилгээний дээжийг ашиглан, датасетээ бэлтгэсэн. Дээж тус бүрт 4, 30, 36, 41, 68 тооны E. vermicularis-ийн өндөг байсан. Мөн энэхүү зурган өгөгдөл нь E. vermicularis-ийн өндгүүд, бусад шимэгчүүд, ялгадас, мөн Грэхэмийн наалдуулагч туузын хий зэрэг объектуудыг агуулж байсан. Эдгээр зургийг машиндаа сургаснаар бид шимэгчдийг мэдрэг болон өвөрмөц чанар өндөртэй аргаар оношилж, улмаар өндгийг тоолуулж халдварлагчдын дата мэдээллийг файл болгон хадгаласан нь цаашид экстенсив болон интенсив чанарыг судлах боломж олгож байгаагаараа давуу талтай. Бидний судалгааны нэг шинэлэг тал

нь мөн шинжилгээний дээжнээс шинэ өгөгдлийн багц (dataset creation) гаргаж авсан явдал юм.

B. Өгөгдөл ихэсгэх аргачлал

Хиймэл оюуны загвар нь оношилгооны мэдрэг чанар өндөртэй байхын тулд олон хүний шинжилгээний дээж, түүнээс гаргаж авсан дата мэдээлэл хангалттай байх зайлшгүй шаардлагатай. Дата мэдээлэл дутмаг байвал, хэчнээн дэвшилтэт модел ашигласан ч оношилгооны мэдрэг чанар өндөр гарах боломжгүй. Тиймээс бид нийт 5 хүний шинжилгээний дээжид хамаарах 179 ширхэг E. vermicularis өндөгний өгөгдөл цуглуулсан. Энэ өгөгдлийн хэмжээ хангалттүй байсан тул өгөгдөл ихэсгэх (augmentation) аргыг ашиглан өгөгдлийн хэмжээг 1970 болгон өсгөсөн. Өгөгдөл ихэсгэх үйл явцыг Зураг 3-д харуулсан. Өгөгдлийг ихэсгэхдээ нийт 10 төрлийн аргыг ашигласан бөгөөд үүнд: bilateral filter, flipped horizontal, color map, gaussian blur, болон бусад зургаан төрлийн дүрс боловсруулалтын техникүүд багтсан. Эдгээр аргууд нь өгөгдлийн олон янз байдлыг нэмэгдүүлж, загварын сургалтад чухал үүрэг гүйцэтгэсэн. Өгөгдөл ихэсгэх шалтгаануудыг дараах байдлаар тайлбарлав: Үүнд:

- Overfitting-ээс зайлсхийх: Загварын хэт тохиргоо буюу overfitting нь сургалтын өгөгдлөөс хэт хамааралтай болж, шинэ, хараахан үзээгүй өгөгдлүүд дээр муу гүйцэтгэл үзүүлэх нөхцөл үүсгэдэг. Өгөгдөл нэмэгдүүлэх нь загварын сургалтанд олон төрлийн өгөгдөл оруулснаар энэ асуудлыг шийддэг.
- Ерөнхийлөн боловсруулсан загвар: Загвар илүү олон төрлийн өгөгдөлтэй харьцаж, шинэ өгөгдөл дээр илүү сайн үр дүн гаргадаг.
- Өгөгдлийн хэмжээ нэмэгдэх: Бага хэмжээний өгөгдөлтэй байх үед загвар сургалтаа дуусгахад хүндрэлтэй болдог. Augmentation техник ашигласнаар өгөгдлийн хэмжээ ихсэж, сургалт үр дүнтэй болно.
- Загварын бат бөх байдал: Загвар нь янз бүрийн хувилбаруудтай танилцаж, дасан зохицох чадвартай болдог, ингэснээр шинэ өгөгдөл дээр илүү сайн үр дүн гаргах боломжтой.



Зураг 2. Өгөгдөл нэмэгдүүлсэн байдал

C. Гиперпараметрүүдийн тохиргоо

YOLO-ийн гиперпараметрүүд нь алгоритмын сургалт, гүйцэтгэлд маш их нөлөө үзүүлдэг. Эдгээр гиперпараметрүүдийг тохируулах нь загварын үр дүнг оновчтой болгох, хурдыг нэмэгдүүлэх, болон нарийвчлалыг сайжруулахад чухал үүрэгтэй. Гиперпараметрүүдийг тохируулах нь туршилт, алдаа засах, мөн тооцооллын хугацаа, боломжийн нөөцийг хамгийн ихээр ашиглахад чиглэгддэг.

1) Сургалтын хурдыг сайн тохируулах нь загварын тогтвортой байдал, хурд, чанарт нөлөөлнө.
2) Туршилт хийхдээ сургалтын хурдад бага, дунд, өндөр зэрэг хэд хэдэн утгуудыг туршиж үздэг.

D. Гиперпараметрүүдийн зарим чухал жишээ:

1) Learning Rate: Сургалтын хурд нь загварын жингийн шинэчлэлтийн хэмжээг тодорхойлдог хамгийн чухал гиперпараметр юм. Хэрэв сургалтын хурд хэт өндөр бол загвар хурдан суралцаж эхэлнэ, гэхдээ тухайн боломжит шийдлийг хэтэрхий хурдан илрүүлж, хамгийн сайн үр дүнд хүрэхгүй байж болзошгүй. Хэрэв сургалтын хурд хэт бага байвал загвар удаан суралцаж, шаардлагатай үр дүнд хүрэхэд их хугацаа шаардлагатай.

V. Туршилтын Үнэлгээ

E. vermicularis -ийн өндөг илрүүлэх зорилгоор YOLOv8 суурилсан автомат системийн туршилтын үнэлгээ нь загварын гүйцэтгэл, нарийвчлал болон бодит орчинд хэрэглэгдэх боломжийг шалгах хэд хэдэн чухал алхмаас бүрдэнэ. Дараах хэсгүүд нь туршилтын тохиргоо, үнэлгээний шалгуур үзүүлэлтүүд, мөн туршилтын үеийн үр дүнгүүдийг нарийвчлан тайлбарлаж байна

A. Өгөгдөл

E.vermicularis -ийн өндөг болон түгээмэл гадны биетүүдийн баталгаажсан микроскопийн зургууд ашиглан өргөн хүрээтэй мэдээллийн санг үүсгэжээ. Энэхүү мэдээллийн сан нь клиник лабораторийн орчинд Scotch tape тестийн аргаар цуглуулсан зургуудаас бүрдэж, лабораторийн нөхцөлд

- Эерэг дээжүүд: E.vermicularis -ийн нэг эсвэл түүнээс олон өндөгнүүдийг агуулсан дээж
- Сөрөг дээжүүд: E.vermicularis -ийн нэг ч өндөг агуулаагүй дээж

Мэдээллийн санг сургалт, баталгаажуулалт, туршилтын бүлгүүдэд хуваах, сургалтанд 70%, баталгаажуулалтанд 15%, туршилтад 15% нь тус тус зарцуулсан. Өгөгдлийн санг илүү сайн хэрэглэхийн тулд эргүүлэх, эргэх, томруулах болон Гауссын шуугиан нэмэх зэрэг өгөгдлийн нэмэгдэл аргуудыг ашигласан.

B. Загварын Сургалт ба Гиперпараметрүүдийг Тохируулах

YOLO загварыг бэлдсэн мэдээллийн сан дээр дамжуулан сургалт хийсэн бөгөөд сургалтыг урьдчилан сургасан жингүүдийг ашиглан дамжуулан явуулсан. Сургалтын явцад гиперпараметрүүдийг (суралцах хурд, багцын хэмжээ, сургалтын давтамж) грид

Гиперпараметр	Тохиргооны утга
learning rate	0.001
batch size	16
epoch	500
image size	640x640 пиксел

Хүснэгт II
Гиперпараметрийн тохиргоо

хайлтын(grid search) арга ашиглан тохируулсан. Грид хайлт нь машин сурахад гиперпараметр тааруулахад ашигладаг уламжлалт арга бөгөөд хамгийн сайн загварыг олохын тулд өгөгдсөн гипер-параметрийн утгуудын хослол бүрийг сайтар туршиж үздэг. Грид хайлтын үр дүнгээс хамгийн сайн суралцах чадвартай загварын гиперпараметрүүдын тохиргоо нь дараах Хүснэгт V-Вд үзүүлсэн байдлаар гарсан.

C. Үнэлгээний Шалгуур Үзүүлэлтүүд

Загварын гүйцэтгэлийг объект илрүүлгийн стандарт үзүүлэлтүүдээр үнэлсэн бөгөөд эдгээр нь микроскопийн зургийн өгөгдлийг зөв илрүүлж, байршуулах чадварыг харуулах чухал үзүүлэлтүүд юм. Дараах шалгуур үзүүлэлтүүдийг бид ашиглан загварын гүйцэтгэлээ үнэллээ. Үүнд:

- mAP: Энэ нь загварын нийт гүйцэтгэлийг нарийвчлал ба тохирол дээр үндэслэн хэмжиж, үр дүнгүүдийг нэгтгэсэн үзүүлэлт юм
- Precision:Загвар E.vermicularis-ийн өндөгнүүдийг зөв илрүүлсэн бүх тохиолдлуудын зөв таамагласан хувь.
- Recall: Загвар Шинжилгээний дээж дэх E. vermicularis-ийн өндөгнүүдийг бүхэлд нь илрүүлэх чадвар буюу өндөгнүүдийг хэр мэдрэг чанар өндөртэй илрүүлж буйг илтгэдэг үзүүлэлт юм.
- F1: Нарийвчлал ба тохирлын дундаж утга буюу тэнцвэртэй үнэлгээг гаргах боломжийг олгодог.
- Хуурамч Эерэг ба Хуурамч Сөрөг: Загварын хуурамч эерэг ба хуурамч сөрөг тохиолдлуудыг тооцоолж, E. vermicularis -ийн өндөгнүүдийг хэр зөв ялган таньж, орхигдуулахгүй тоолж байгааг илтгэнэ.

D. Туршилтын үр дүн

YOLO8v загвар нь E.vermicularis -ийн өндөгнүүдийг гадны биетүүдээс ялган илрүүлж, амжилттай тоолж, интенсив чанарыг тооцоолсон. Туршилтын үр дүнгүүдийг дараах байдлаар хураангуйлан орууллаа. Үүнд:

- mAP: 92.5%
- Нарийвчлал: 91.8%
- Recall: 93.2%
- F1: 92.5%
- Хуурамч Эерэг: 3.5%
- Хуурамч Сөрөг: 4.2%

Эдгээр туршилтын үр дүн нь машин сургалтын энэхүү загвар нь E. vermicularis -ийн өндгийг илрүүлэх мэдрэг болон өвөрмөц чанар өндөр (хуурамч эерэг сөрөг утга бага) болохыг нотолж байна.

Зураг 3. Туршилтын үр дүн

E. Харьцуулсан Шинжилгээ

Сүүлийн жилүүдэд хиймэл оюун ухаанд (AI) суурилсан аргачлалуудыг паразитын халдвар, тэр дундаа Plasmodium (хумхаа үүсгэгч)-ыг илрүүлэхэд амжилттай ашиглаж эхэлжээ. Эдгээр шинэчлэлүүд нь уламжлалт микроскопын аргыг орлох буюу нэмэлт хэрэгсэл болгон ашиглах боломжийг нээж байна.

Тухайлбал, 2025 онд Carles Rubio Maturana нарын судлаачид [26] нь хямд өртөгтэй, роботжуулсан микроскоп болон AI-д суурилсан дүрс боловсруулалтын хэрэгслийг ашиглан хумхааг автоматаар оношлох системийг боловсруулсан байна. Энэхүү судалгаа нь Испанийн Барселона хотын Vall d'Hebron-ийн олон улсын эрүүл мэндийн нэгжид хийгдсэн бөгөөд 54 ширхэг Giemsa-гаар будсан зузаан цусны түрхэцийн дээжийг ашиглан Plasmodium паразитыг илрүүлэхэд AI системийн үр нөлөөг үнэлсэн байна.

Мөн 2023 онд Liu нарын судлаачид [27] нь YOLOv5 болон Transformer загваруудыг хослуулан AIDMAN хэмээх AI-д суурилсан обьект илрүүлэх системийг боловсруулсан. Энэхүү систем нь нимгэн цусны түрхэцийн зургууд дээр Plasmodium паразитыг илрүүлэхэд 98.44%-ийн нарийвчлалтай ажилласан бөгөөд энэ нь туршлагатай микроскопчдын үр дүнтэй ойролцоо байжээ.

Эдгээр судалгаанууд нь AI технологийг ашиглан паразитын халдварыг илрүүлэх шинэ аргачлалуудыг хөгжүүлж, уламжлалт микроскопын аргыг сайжруулах боломжийг харуулж байна. Цаашид эдгээр технологийг E. vermicularis зэрэг бусад паразитын халдварыг илрүүлэхэд хэрхэн ашиглах талаар судалгаа хийх нь зүйтэй

юм.

Манай судалгаанд ашигласан YOLOv8 загварт суурилсан автомат илрүүлгийн систем нь Enterobius vermicularis-ийн өндгийг 93.2%-ийн мэдрэг чанар, 96.5%-ийн өвөрмөц чанар-тайгаар, зурган дээрх объектыг дунджаар 0.12 секундэд илрүүлж, тоолох боломжтой байгааг тогтоосон. Энэ нь зөвхөн оношлогооны өндөр нарийвчлалтай төдийгүй, богино хугацаанд олон тооны дээжийг боловсруулж чадах хурдаараа урьдчилан сэргийлэх шинжилгээний өргөн хүрээтэй хэрэглээнд өндөр үр ашигтай болохыг харуулж байна.

Уламжлалт микроскопын аргууд нь туршлагатай мэргэжилтнээс хамааралтай, цаг хугацаа шаардсан процесс байдаг бол манай систем нь энэхүү процессийг автоматжуулснаар шинжилгээний хугацааг богиносгож, хүний нөөцийн ачааллыг бууруулах, улмаар олон нийтийг хамарсан скрининг шинжилгээг илүү үр дүнтэй зохион байгуулах боломжийг бүрдүүлж байна. Иймд тус систем нь хиймэл оюун ухаанд суурилсан, практикт шууд хэрэглэх боломжтой, амьдралд ойр шийдэл болж байгаагаараа онцлог юм.

VI. Дүгнэлт

1) Хүүхдийн хялгасан хорхойтох өвчнийг оношлоход YOLO8v суурилсан автомат илрүүлгийн аргын мэдрэг чанар 93.2%, өвөрмөц чанар 96.5% байна.
2) YOLO8v загвар нь зурган дээрх объект (E. vermicularis-ийн өндөг)-ийг дунджаар 0.12 секундэд илрүүлж, тоолж байгаа нь олон хүнийг хамруулсан урьдчилан сэргийлэх шинжилгээний хариуг богино хугацаанд мэдрэг чанар өндөртэй гаргах боломжтойг харуулж байна.
3) Хүүхдийн хялгасан хорхойтох өвчний оношилгоог автоматжуулснаар олон хүнийг хамруулсан урьдчилан сэргийлэх шинжилгээний үед дээж шинжилж буй лабораторийн мэргэжилтнүүдийн ажлыг хөнгөвчилнө.

Ашигласан ном

[1] Burkhart, C. G., & Burkhart, C. N. (2005). Assessment of frequency, transmission, and genitourinary complications of enterobiasis (pinworms). International Journal of Dermatology, 44(10), 837–840.

[2] Centers for Disease Control and Prevention. (2022). Parasites - Enterobiasis (Pinworm Infection).

[3] Cook, G. C. (1994). Enterobius vermicularis infection. Gut, 35(9), 1159–1162.

[4] Fan, C.-K., et al. (2019). Global status of human enterobiasis and its health implications. Acta Tropica, 199, 105106.

[5] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

[6] Garcia, L. S. (2016). Diagnostic Medical Parasitology (6th ed.). ASM Press.

[7] Giymaa, N. (2008). [Enterobiasis infection rate study]. Parasitology Research Reports.

[8] ENTEROBIUS VERMIVULARIS-ИЙН ОНОШИЛГООНЫ АСУУДАЛД

[9] Улаанбаатар хотын цэцэрлэгийн насны хүүхдүүдэд наалдуулагч туузын аргаар энтербиозыг илрүүлсэн дүн

[10] Ganbold, D. (1990). [Helminthic infection in Ulaanbaatar city population]. Mongolian Journal of Parasitology.

[11] Jocher, G., et al. (2023). YOLOv8 Documentation.

[12] Rahman, T., et al. (2022). Deep learning-based parasite detection in medical imaging: A survey. Biomedical Signal Processing and Control, 71, 103158.

[13] Shin, H. C., et al. (2020). Medical image analysis with deep learning. Annual Review of Biomedical Engineering, 22, 221–248.

[14] Takizawa, M., et al. (2021). Misdiagnosis due to non-parasitic artifacts in microscopic stool examinations. Journal of Infection and Chemotherapy, 27(1), 123–128.

[15] Batbayar, B., & Batsukh, U. (2021). Study on the prevalence of Enterobius vermicularis among school-aged children in Arkhangai province. Mongolian Journal of Parasitology and Public Health, 3(2), 45–52.

[16] Fan, C. K., Chuang, T. W., Huang, Y. C., Yin, A. W., Chou, C. M., Hsu, Y. T., Kios, R., Hsu, S. L., Wang, Y. T., Wu, M. S., Lin, J. W., Briand, K., & Tu, C. Y. (2019). Enterobius vermicularis infection: prevalence and risk factors among preschool children in kindergarten in the capital area, Republic of the Marshall Islands. BMC Infectious Diseases, 19(1), 536.

[17] Centers for Disease Control and Prevention. (2024). Enterobiasis (Pinworm). In CDC Yellow Book 2024: Health Information for International Travel

[18] Miah, M. S., Rahman, M. M., & Razu, M. H. (2024). Real-time object detection in medical imaging using YOLO models for kidney stone detection. European Journal of Computer Science and Information Technology, 12(5), 15–25

[19] Burgess, I. F., et al. (2005). Prevalence and transmission of Enterobius vermicularis in primary school children in England. Journal of Public Health, 27(2), 172–174.

[20] Dong, H., et al. (2017). Malaria parasite detection using machine learning. Journal of Parasitology, 104(6), 45-59.

[21] Yang, L., et al. (2020). Detection of Plasmodium species using deep learning. Journal of Clinical Microbiology, 58(12), 1023-1035.

[22] Gomes, P. E., et al. (2021). Intestinal helminth detection using machine learning methods. Parasites and Vectors, 14(1), 53-60.

[23] Oliveira, S. A., et al. (2019). Convolutional neural networks for detecting parasitic eggs in stool samples. International Journal of Parasitology, 49(4), 375-380.

[24] Tegshee, T., et al. (2023). A study on staging cystic echinococcosis using machine learning methods. Journal of Parasitology Research, 15(2), 45-53.

[25] Cho, S. Y., & Kang, S. Y. (1975). Significance of Scotch-tape anal swab technique in diagnosis of Enterobius vermicularis infection. Korean J Parasitol, 13(2), 102-114.

[26] Rubio Maturana, C., de Oliveira, A. D., Zarzuela, F., Mediavilla, A., Martínez-Vallejo, P., Silgado, A., Goterris, L., Muixí, M., Abelló, A., Veiga, A., López-Codina, D., Sulleiro, E., Sayrol, E., & Joseph-Munné, J. (2025). Evaluation of an Artificial Intelligence-Based Tool and a Universal Low-Cost Robotized Microscope for the Automated Diagnosis of Malaria. International Journal of Environmental Research and Public Health, 22(1), 47. https://doi.org/10.3390/ijerph22010047

[27] Liu R, Liu T, Dan T, Yang S, Li Y, Luo B, Zhuang Y, Fan X, Zhang X, Cai H, Teng Y. AIDMAN: An AI-based object detection system for malaria diagnosis from smartphone thin-blood-smear images. Patterns (N Y). 2023 Aug 3;4(9):100806. doi: 10.1016/j.patter.2023.100806. PMID: 37720337; PMCID: PMC10499858.

Бодит цаг хугацаанд дамжуурга дээрх цементийг тоолох систем

Цолмонгийн Энхзаяа
Мэдээллийн Технологи,
Компьютерын Ухааны Тэнхим
МУИС, МТЭС
Монгол улс, Улаанбаатар
enkhzayatsolmon@gmail.com

Батсуурийн Сувдаа
Мэдээллийн Технологи,
Компьютерын Ухааны Тэнхим
МУИС, МТЭС
Монгол улс, Улаанбаатар
suvdaa@num.edu.mn

Ганболдын Ууганбаяр
Мэдээллийн Технологи,
Компьютерын Ухааны Тэнхим
МУИС, МТЭС
Монгол улс, Улаанбаатар
uuganbayar.g@num.edu.mn

Абстракт*— Үйлдвэрлэлийн салбарт туузан дамжуурга дээрх бүтээгдэхүүнийг бодит цаг хугацаанд үнэн зөв тоолох нь үйл ажиллагааны үр ашиг, бараа материалын менежментэд чухал ач холбогдолтой юм. Уламжлалт тоолох аргууд нь ихэвчлэн гар ажиллагаатай байдаг ба хөдөлмөрийн бүтээмж бага мөн автоматжуулалтад нэгтгэх боломж муу байдаг. Эдгээр асуудлыг шийдвэрлэхийн тулд энэ судалгаагаар компьютерын харааны техникт суурилсан автомат тоолох системийг санал болгож байна. Санал болгож буй систем нь хөдөлж буй туузан дамжуурга дээрх ууттай цементийг дээр байрлуулсан камерыг ашиглан автоматаар таньж тоолох юм. Хоцрогдолгүй бодит цагт нарийвчлалтай илрүүлэхийн тулд гүн сургалтын YOLO-ийг ашиглан ууттай цементийг илрүүлж, тоолоход ашигласан. Энэхүү загварыг үйлдвэрлэлийн орчинд найдвартай байдлыг хангах үүднээс янз бүрийн чиглэл, хүрээлэн буй орчны нөхцөлд цементийн зургийг агуулсан өгөгдлийн багцыг үүсгэн, түүн дээр сургасан. Боловсруулалтын дараах үе шатанд илрүүлсэн цементийг дагах, тоолсон. Системийг туршилтын өгөгдөл дээр илрүүлэлтийн дундаж нарийвчлал нь 98.6% (0.5 давхцал дээрх) байна. Уг системийн туршилтын үр дүн нь гар ажиллагаатай холбоотой үйл ажиллагааг бууруулж, үйл ажиллагааны үр ашгийг дээшлүүлж, үйлдвэрлэлийн орчинд өргөнөөр нэвтрүүлэхэд тохиромжтойг харуулж байна.

Түлхүүр үг—Компьютер Хараа; Цемент Илрүүлэлт; Автоматжуулалт;

I. Оршил

Орчин үеийн үйлдвэрлэлийн үйл ажиллагаанд автоматжуулалт нь үр ашгийг дээшлүүлэх, хүний алдааг бууруулах чухал бүрэлдэхүүн хэсэг болж байна. Автоматжуулалт чухал үүрэг гүйцэтгэдэг нэг гол талбар бол туузан дамжуулагч системээр дамжиж буй бүтээгдэхүүнийг бодит цаг хугацаанд тоолж, хянах явдал юм. Үнэн зөв тоолох нь бараа материалын менежмент, үйл явцыг хянахад чухал ач холбогдолтой юм [1]. Механик мэдрэгч эсвэл гараар тоолох гэх мэт уламжлалт аргууд нь объектын бөглөрөл эсвэл хувьсах дамжуургын хурд зэрэг хүчин зүйлээс шалтгаалж алдаа гарах хандлагатай байдаг.

Машины хараа болон гүнзгий сургалтын сүүлийн үеийн дэвшил нь эдгээр үйл явцыг автоматжуулах шинэ боломжийг нээж өгсөн. Эдгээр дэвшлийн дотроос YOLO (You only Look one) зэрэг бодит цагийн объект илрүүлэх загварууд нь хоцрогдол багатайгаар объектыг өндөр нарийвчлалтайгаар илрүүлж, ангилах чадвараараа үйлдвэрлэлийн янз бүрийн хэрэглээнд ашиглагдаж байна. Сүүлийн үеийн YOLOv8 нь хурд, нарийвчлалын хувьд улам сайжирч, бодит цагийн боловсруулалт шаардлаг үйлдвэрлэлийн ажлуудад онцгой тохиромжтой болсон [2].

Цементийн үйлдвэрлэлийн салбарт дамжуургын системийг үйлдвэрлэлийн шугамын янз бүрийн үе шатанд цементийн уут тээвэрлэхэд ихэвчлэн ашигладаг. Логистикийн хяналт, чанарын хяналтын аль алинд нь дамжуургаар өнгөрөх цементийг нарийн тоолох шаардлагатай. Гэсэн хэдий ч үйлдвэрлэлийн орчин дахь үйл ажиллагааны хүндрэлүүд, тухайлбал дамжуургын хурд, хувьсах гэрэлтүүлгийн нөхцөл, давхацсан эсвэл хэсэгчлэн бөөгнөрсөн ууттай цемент зэрэг нь уламжлалт аргуудыг нарийвчлал, найдвартай илрүүлэхэд хүндрэл учруулдаг.

Эдгээр сорилтыг шийдвэрлэхийн тулд энэхүү судалгаагаар YOLOv8 болон компьютер харааND суурилсан цементийн уут тоолох бодит цагийн системийг танилцуулж байна. Туузан дамжуурга дээр суурилуулсан өндөр нарийвчлалтай камерыг ашиглан систем нь YOLOv8 объект илрүүлэх загвараар видео урсгалаас цементийг илрүүлэн тоолдог. Энэхүү загвар нь янз бүрийн чиглэлтэй цемент, хүрээлэн буй орчны өөр өөр гэрэлтүүлгийн нөхцөлд ууттай цементийг илрүүлэх, тоолоход сургагдсан. Гүнзгий сургалтыг бодит цагийн дүрс боловсруулалттай нэгтгэснээр үйлдвэрлэлийн нөхцөлд тоолох үйл явцыг автоматжуулах найдвартай, өргөтгөх боломжтой аргыг санал болгож байна.

Энэхүү ажлын үндсэн хувь нэмэр нь дараах байдалтай байна.

- Өндөр хурдны үйлдвэрлэлийн орчинд ашиглахад оновчтой YOLOv8 олон олон хувилбарт туршилтыг хийж, туршилтын дүнд тулгуурлан цементийн уут тоолох бодит цагийн системийг хөгжүүлсэн.
- Загварын найдвартай байдлыг сайжруулахын тулд янз бүрийн гэрэлтүүлгийн нөхцөл, чиглэл,

давхцалын үеийн ууттай цементийн зургийг агуулсан өөрсдийн өгөгдлийн багцыг бий болсон.

- Бодит үйлдвэрлэлийн орчинд системийн гүйцэтгэлийг туршиж үр ашигтайг харуулсан.

Энэхүү өгүүллийн үлдсэн хэсгийг дараах байдлаар зохион байгуулав: II хэсэгт үйлдвэрлэлийн орчинд объект илрүүлэх, тоолох чиглэлээр холбогдох ажлуудыг авч үзнэ. III хэсэгт архитектур, өгөгдөл цуглуулах, сургалтын үйл явц зэрэг санал болгож буй системийн дэлгэрэнгүйг оруулсан болно. IV хэсэгт туршилтын үр дүн болон гүйцэтгэлийн үнэлгээг харуулав. Эцэст нь системийн цаашдын сайжруулалт болон боломжит хэрэглээний талаар хэлэлцэх асуудлыг дэвшүүлсэн.

II. СУДЛАГДСАН БАЙДАЛ

Бодит цагийн объект илрүүлэх, тоолох салбарт сүүлийн жилүүдэд, ялангуяа гүнзгий сургалт, компьютер харааны технологи бий болсноор ихээхэн ахиц дэвшил гарч байна. Үйлдвэрлэлийн эхэн үед туузан дамжуулагч дээрх объектыг тоолохдоо инфраред эсвэл лазер дээр суурилсан систем гэх мэт механик мэдрэгчийг ашиглан голчлон ашигладаг байсан. Эдгээр аргуудыг хэрэгжүүлэхэд харьцангуй хялбар боловч тэдгээр нь хэд хэдэн хязгаарлалттай байдаг, ялангуяа объектууд давхцах үед мэдрэх боломжгүй болдог. Жишээлбэл,ууттай цементүүд зэрэгцэн байрласан үед мэдрэгч нь буруу тоолж болно.

Balkrishna нар. (2018) автоматжуулсан савлагааны шугам дахь туузан дамжуурга дээрх объектуудыг тоолоход зайн мэдрэгчийг ашиглах талаар судалжээ [3]. Хяналттай орчинд үр дүнтэй байсан ч өндөр хурдны дамжуулагч эсвэл давхцаж буй объекттой холбоотой хувилбаруудад алдаа гарах хандлагатай байсан нь илүү боловсронгуй шийдлүүдийн хэрэгцээг онцолж байв.

Компьютер харааны системүүд хөгжихийн хэрээр ирмэгийг илрүүлэх, ирмэгийн шинжилгээ, Хьюгийн хувиргалт зэрэг дүрс боловсруулах аргуудыг дамжуурга дээр суурилсан объект илрүүлэх, тоолоход ашигласан. Эдгээр аргууд нь объектыг геометрийн шинж чанарт нь тулгуурлан илрүүлэхэд үр дүнтэй байсан боловч хувьсах гэрэлтүүлэг, объектын бөглөрөл эсвэл жигд бус хэлбэр бүхий орчинд таних чадвар нь муу байдаг.

Тухайлбал, Zhang (2019), Yi Zhang(2020) нар үйлдвэрлэлийн шугамын чанарын хяналтад дүрс боловсруулахад суурилсан аргыг хэрэгжүүлсэн[4][5]. Тэдний систем нь дамжуурга дээрх объектуудыг илрүүлж чанарыг нь шалгахын тулд дүрсийг сегментчилэх уламжлалт аргыг ашигласан. Гэсэн хэдий ч системийн гүйцэтгэл нь гэрэлтүүлгийн тогтвортой нөхцөл, тодорхой объектын хил хязгаараас ихээхэн хамааралтай байсан тул өндөр хурдны үйлдвэрлэлийн хэрэглээнд тохиромжгүй болсон[6].

Гүнзгий суралцах болон хувиралт мэдрэлийн сүлжээ (CNN) бий болсноор YOLO(You Only Look Once)[7], Faster R-CNN[8], SSD (Single Shot Detector)[9] зэрэг загварууд нь объект илрүүлэх салбарт хувьсгал хийсэн. Эдгээр загварууд нь өгөгдлөөс шууд объектын шинж чанарыг судлах замаар хатуу загварчилдаг уламжлалт аргуудаас хамаагүй давуу бөгөөд өөр өөр орчин, объектын олон шинж чанарыг нэгтгэх боломжийг олгосон.

Тэр дундаа YOLO-ийн гүйцэтгэл нь бусад мэдрэлийн сүлжээнд суурилсан объект илрүүлэгчээс нарийвчлал, хурдны хувьд илүү байсан нь тогтоогдсон. Тус загвар нь зургийн хэмжээтэй харьцангуй жижиг объектыг илрүүлэхэд тааруу үр дүн үзүүлдэг. Энэхүү судалгааны хүрээнд цементийн хэмжээ нь жижиг бус тул YOLO загвар нь энэ тохиолдолд сайн ажиллана гэж үзсэн. Мөн (YOLO) нь нэг мэдрэлийн сүлжээг бүтэн зурагт хэрэглэж, дүрсийг бүс болгон хуваах, бүс тус бүрийн хязгаарлах хайрцаг, магадлалыг урьдчилан таамагладаг бөгөөд үүгээрээ Fast-RCNN-ээс хурдан байдаг.

YOLOv3 нь динамик орчинд объектыг үнэн зөв илрүүлэх чадварыг харуулсан бөгөөд энэ нь туузан дамжуулагчийн хяналт зэрэг үйлдвэрлэлийн хэрэглээнд тохиромжтой. YOLOv4 ба YOLOv5 зэрэг дараагийн хувилбарууд нь ялангуяа нөөц хязгаарлагдмал орчинд илрүүлэх нарийвчлал, хурдыг сайжруулснаар энэхүү архитектурыг улам сайжруулсан.

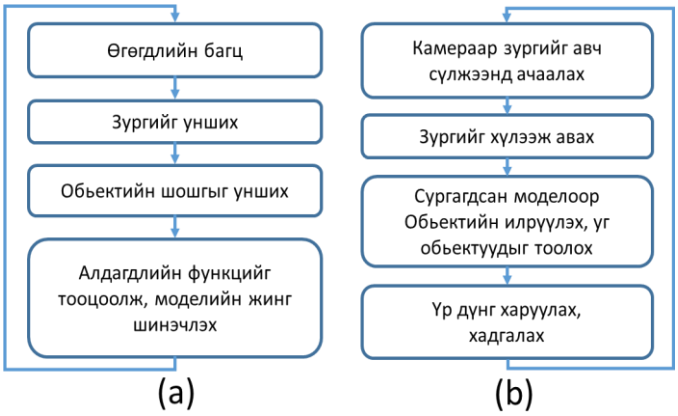
Варна [10] бодит цагт дамжуурга дээрх электроникийн эд ангийг илрүүлэх болон ангилахад YOLOv4-ийн хэрэглэж өндөр үр дүнг харуулсан. Хэдийгээр электроникийн эд ангид төвлөрч байгаа ч тэдний ажил нь үйлдвэрлэлийн хэрэглээнд хамааралтай бөгөөд YOLO нь объектын жигд бус хэлбэр, хувьсах зайтай орчинд ч бодит цаг хугацаанд объектыг үнэн зөв тоолж чаддаг болохыг харуулж байна. Эдгээр ойлголтыг дамжуурга дээрх цементийн уут тоолох асуудалд шилжүүлэх боломжтой.

III. СУДАЛГААНЫ МАТЕРИАЛ АРГА ЗҮЙ

Санал болгож буй систем нь өндөр нягтралтай камер болон YOLOv8 объект илрүүлэх загвар ашиглан туузан дамжуулагч дээрх ууттай цементийг автоматаар тоолоход зориулагдсан. Энэхүү систем нь цемент бөөгнөрсөн, гэрэлтүүлгийн янз бүрийн нөхцөл зэрэгт дамжуургаар урсах цементийг алдагдалгүй тоолоход тулгарч буй бэрхшээлийг даван туулахын тулд компьютерын хараа, гүнзгий сургалт, бодит цагийн боловсруулалтын чадварыг нэгтгэдэг. Системийн бүтэц нь дараах үндсэн бүрэлдэхүүн хэсгүүдээс бүрдэнэ.

- Камерын тохиргоо ба өгөгдөл цуглуулалт
- Объект илрүүлэх (YOLOv8) загварын сургалт
- Өмнөх болон дараах дэд боловсруулалт
- Системийн гүйцэтгэлийн хэмжүүр ба үнэлгээ

Энэхүү систем нь камер, тооцоолох хэрэгсэл болох компьютер, үр дүнг харуулах дэлгэц, тэдгээрийн холбох сүлжээний төхөөрөмж, дамжуурга зэргээс бүрдэнэ. Системийн ажлын урсгалыг Зураг 1-д үзүүлэв.



1-р зураг. (a) Сургалтын үеийн ажлын диаграм, (b) Суурилагдсан системийн ажлын диаграм

Камер нь ууттай цемент дамжин өнгөрдөг туузан дамжуургын дээрээс доошоо тодорхой харагдахаар байрлуулсан. Камерын хувьд өөрөө таталт болон олон нягтралтайгаар дүрсээ илгээх боломжтой камерыг сонгож авсан. Өгөгдлөө цуглуулахын тулд MAK цементийн үйлдвэрийн авто ачлагын хэсэгт гаралтын дамжуурга дээр суурилуулсан камерыг ашигласан. Өгөгдөл бүрхэг өдөр, хурц нартай, камерын эсрэг чиглэлд хурц нартай, бүрэнхий болон шөнийн гэрэлтэй үед өгөгдлийг видео хэлбэрээр цуглуулсан. Уг видеоноос фрейм бүрээр салган авч модел сургалтандаа ашигласан. Нийт 2500 орчим зургийг шошголох багаж[9] ашиглан гараар шошголосон. Сургах өгөгдөл 2057 зураг, баталгаажуулах өгөгдөл 224 зураг, туршилтын өгөгдөл 301 болно. Цуглуулсан өгөгдлийн жишээг Зураг 2-д үзүүлэв.

Системийн цөм нь YOLOv8 объект илрүүлэх загвар бөгөөд камерт авсан фрейм бүр дэх цементийг илрүүлэх, тоолох үүрэгтэй. YOLO дэвшилтэт хувилбар болох YOLOv8 нь хурд, нарийвчлал, уян хатан байдлын хувьд сайжирсан. YOLOv8 загварыг янз бүрийн цементийн чиглэл, гэрэлтүүлгийн нөхцөл, цементийн давхцал зэргийг агуулсан өөрсдийн өгөгдлийн багц дээр сургасан. Энэхүү өгөгдлийн багцыг бодит дамжуургын орчноос цуглуулснаар загвар нь бодит хувилбарт сайн тохирсон.

2-р зураг. Бүрхэг өдөр(зүүн дээд), камерын эсрэг чиглэлд хурц нартай (баруун дээд), шөнийн гэрэлтэй(зүүн доод), шөнийн гэрэлтэй цемент арын нүүр (баруун доод)

YOLOv8 нь YOLOv8n (нано), YOLOv8s (жижиг), YOLOv8m (дунд), YOLOv8l (том), YOLOv8x (хэт том) гэсэн таван том хувилбарыг санал болгосон. Бид өөрсдийн тооцооллын төхөөрөмж болон хурдыг шалгахын тулд YOLOv8n, YOLOv8s, YOLOv8m загваруудыг сонгон туршилтыг гүйцэтгэсэн. Бид тус загваруудыг сургахдаа AdamW (lr = 0.01, momentum = 0.94) оновчлогчийг ашиглаж, параметрүүдийг дараах байдлаар тохируулсан сан. Сургалт болон баталгаажуулалт, шалгалтын туршилтыг гүйцэтгэсэн ба зургийн хэмжээг 640 болгож тохируулсан. YOLOv8 загварыг NVIDIA RTX 3070 (8192MB) дээр CUDA дэмжлэгтэй Ultralytics YOLOv8.3.15, Python-3.12.2, torch-2.4.0 + cu118 ашиглан хэрэгжүүлсэн. Сургалтын үйл явц нь 30 epoch ба багц бүрийн хэмжээг 16 гэж тохируулсан.

Камераас (1200x800цэг) орж ирэх зургийг сургасан загварт оруулахын өмнө системийн хурдыг сайжруулахын тулд зургийн нарийвчлалыг бууруулж 640x640 цэг болгосон. YOLOv8 загвар нь зураг тус бүр дэх цементийн уутыг илрүүлсний дараа уутны тоог үнэн зөв тоолохын тулд дараах боловсруулалтыг хийдэг. Цементийн уут бүрийг зөвхөн нэг удаа тоолохын одоогийн фреймд илэрсэн цементийг дараа дараагийн фреймд олж дагах шаардлага гардаг. Үүнийг гүйцэтгэхийн тулд Ultralytics-д агуулагдсан мөрдөх алгоритм болох BoT-SORT [10]-ийг ашигласан.

Цемент илрүүлэлтийн гүйцэтгэлийг гурван хэмжүүр ашиглан хийсэн: (precision, Recall, mean Average Precision). Эдгээр хэмжүүрүүд нь загварын гүйцэтгэл болон өгөгдсөн өгөгдлийн багц доторх объект эсвэл тохиолдлуудыг зөв тодорхойлж, ангилах чадварыг үнэлэхэд чухал ач холбогдолтой юм. Итгэлийн босго 50% (mAP50) нарийвчлалын дундаж утгыг мөн өргөн ашигладаг тул үнэлгээндээ ашигласан.

IV. Үр дүн ба Хэлэцүүлэг

A. Үр дүн

Сургалтын явцад 30 epochs-оос дээш мэдэгдэхүйц ахиц гараагүй тул сургалтын үйл явцыг 30 хүртэл хийсэн. yolov8-ийн n,s,m гурван загвар дээр илрүүлэлтийн үр дүнг харьцуулсан ба уг гурван моделууд нь бидний хэрэглээнд хангалттай гэж үзсэн тул бусад моделуудыг туршилтад оруулаагүй. Сургагдсан загварын гүйцэтгэлийн хэмжүүрүүд болох precision, Recall, mean Average Precision (mAP50)-ийг модел тус бүрд Хүснэгт 1-д үзүүлэв. Илрүүлэлтийн дүнгээс харахад гурван моделийн илрүүлэлт нь маш ойролцоо байлаа. Моделийн сонголтыг хийхдээ туршилтын өгөгдөл дээрх F1 үнэлгээг үндэслэн сонгосон ба YOLOv8s нь өндөр байна. Илрүүлэлтийн үр дүнгээс нэг анхаарал татсан нь YOLOv8m модел дээр илрүүлэлт хамгийн бага байгаа явдал бөгөөд үүнд дараах хүчин зүйлээс шалтгаалан бага модел нь өгөгдөл дээр өндөр танилт үзүүлсэн гэж үзэж болохоор байна.

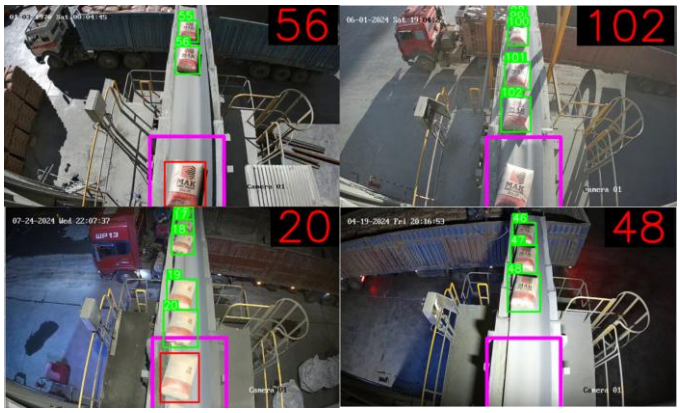
1. Уг судалгааны өгөгдлийн онцлог нь хялбар бөгөөд тогтвортой шинжтэй. Учир нь камер нь тогтмол байрлалд байрласан тул объектын хэмжээ нь хувьсахгүй.

2. Цемент нь дамжуурга дээр урсах тул зургийн төвд байрласан.

Иймд цаашид бид YOLOv8m-ийг ашиглах бол өгөгдлийг тоо болон шинж чанарыг нэмэгдүүлэх шаардлагатай гэж үзэж байна. Үр дүнгээс харахад YOLOv8s нь хангалттай сайн үр дүнг үзүүлсэн тул уг ажлын хүрээнд ашиглагдах бүрэн боломжтой гэж үзэж байна.

Хүснэгт 1. Цементийн өгөгдлийг багцуудад дахь гүйцэтгэлийн үзүүлэлтүүд.

Модел	Өгөгдөл	Precision	Recall	mAP50	F1
YOLOv8n (нано)	Сургалт	0.967	0.984	0.994	0.975
	Баталгаажуулалт	0.980	0.984	0.994	0.981
	Туршилт	0.984	0.988	0.994	0.985
YOLOv8s (жижиг)	Сургалт	0.980	0.977	0.994	0.978
	Баталгаажуулалт	0.988	0.977	0.993	0.982
	Туршилт	0.981	0.992	0.994	0.986
YOLOv8m (дунд)	Сургалт	0.977	0.984	0.993	0.980
	Баталгаажуулалт	0.976	0.984	0.993	0.980
	Туршилт	0.980	0.985	0.994	0.984



3-р зураг. Тоололтын үр дүнгийн зураг. Улаан дөрвөлжин нь илэрсэн ч тоологдоогүй илэрц. Ногоон дөрвөлжин нь тоологдсон илэрц. Ягаан дөрвөлжин нь илэрцийг тоолох хязгаарыг илтгэнэ.

Хүснэгт-1 дээрх үр дүнгээс харахад уг системийн цементийг илрүүлэлт нь хангалттай өндөр байгааг харж болно. Үүний шалтгаан нь дамжуургаар урсан цементийн хэмжээ болон өнгө төдийлөн өөрчлөгдөхгүйд байгаа гэж үзэж байна.

Илэрсэн цемент бүрийг бид дараалсан фрейм бүрд мөрдөж зураг-3 дээрх үзүүлсэнчлэн ягаан дөрвөлжин дотроос илэрцийн төв цэг гарахад тоолохоор програмчилсан. Мөн бид бодит үйлдвэрийн ачилтын хэсэг дээр автомашинд ууттай цементийг уг системийг ашиглан тоолж шалгасан ба алдагдалгүй тоолж чадсан. Тоололтын үр дүнгийн жишээг зураг-3 дээр харууллаа.

Сургалтын явцад орчин ихээхэн өөрчлөгдөх үед илрүүлэлтийн нарийвчлал ихээхэн буурч байсан. Жишээлбэл Зураг-2 баруун дээд буланд буй зураг шиг камерын урдаас гялбасан үеийн зураг сургалтын өгөгдөлд байгаагүй ба бодит орчинд туршихад уг өгөгдөл дээр алдаж байсан. Иймд уг өгөгдлийг нэмж загварыг сургаснаар нарийвчлал илүү сайжирсан. Энэ мэт шинээр орчин солигдох буюу өгөгдлийн багц дахь өгөгдлөөс эрс ялгаатайгаар орчин өөрлөгдөх бүрд тухайн орчны өгөгдлийг нэмж сургах шаардлагатай гарч болзошгүй.

B. Ажиллах хурдны туршилтын дүн

Илрүүлэлтийн үзүүлэлтийн хувьд гурван модел нь мэдэгдэхүйц ялгааг үзүүлээгүй. Ажиллах хугацааны туршилтыг гүйцэтгэсэн ба харьцуулалтад yolov8-ийн n,s,m гурван загвар харьцуулсан. Модел бүрийг график карт дээр болон төв процессор дээр ажиллуулж тус бүрийн нэг зураг дээр танилт хийхэд шаардагдаж буй хугацааг хэмжив. Үр дүнг Хүснэгт 2-т үзүүлэв. Илрүүлэлтийн болон хурдны туршилтаас харахад YOLOv8n, YOLOv8s загварууд нь хангалттай илрүүлэлт болон бидний хэрэглээний бодит цагийн хэмжүүрт нийцэж байна гэж үзлээ. Иймд бид цаашид YOLOv8s загварын сайжруулсан хувилбарыг хэрэглэхээр шийдвэрлэлээ.

Хүснэгт 2. Моделиудын ажиллах хурдны харьцуулалт.

Модел	Төв процессор	График карт
YOLOv8n (нано)	28мс	7мс
YOLOv8s (жижиг)	72мс	9мс
YOLOv8m (дунд)	170мс	18мс

C. Хэлэлцүүлэг ба цаашдын судалгааны чиглэл

Орчин өөрчлөгдөж одоогийн загвар алдах үед өгөгдлийг цуглуулж, загварыг дахин сайжруулах процессыг автоматжуулах шаардлага гарч байгаа ба үүнийг шийдэх боломжуудыг илэрхийлж байна.

Одоогоор зөвхөн тоолох үйлдлийг автоматжуулсан ба цаашид дамжуургаар урсан цементийн бүрэн бүтэн байдлыг хянах хяналтын хэсгийг нэмж хийх, түүнийг гүйцэтгэхэд хэрхэн тооцоолох төхөөрөмжид ачаалал багатайгаар шийдвэрлэх шийдлийг илэрхийлж байна.

Ашигласан ном, хэвлэл:

[1] Yusoff, Mohd Shahrizan, et al. "Image Processing Recognition Technique for Quality Control (Conveyor System)." Technological Frontiers and Sustainable Innovations. Cham: Springer Nature Switzerland, 2024. 15-22.

[2] Varghese, Rejin, and M. Sambath. "Yolov8: A novel object detection algorithm with enhanced performance and robustness." 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS). IEEE, 2024.

[3] Balkrishna, Ms Tapase Snehal, and S. A. Pardeshi. "Model design and simulation of automatic counting machine using proximity sensor." JournalNX—Multidisciplinary Peer Reviewed J (2018).

[4] Zhang, Xuanni. "Online defect detection method of array optical sensor based on watershed image processing algorithm." Multimedia Tools and Applications (2019): 1-15.

[5] Yi Zhang, Mengbo You, Tasuku Miyoshi, Takuya Akashi, A Vision System for Canning with Fish Sensing Using Rule-based Matching and Segmentation, IEEJ Transactions on Electrical and Electronic Engineering, Vol.15, No.6, pp.956-964, 2020.

[6] Kaur, Ravpreet, and Sarbjeet Singh. "A comprehensive review of object detection with deep learning." Digital Signal Processing 132 (2023): 103812.

[7] Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.

[8] Ren, Shaoqing, et al. "Faster R-CNN: Towards real-time object detection with region proposal networks." IEEE transactions on pattern analysis and machine intelligence 39.6 (2016): 1137-1149.

[9] Liu, Wei, et al. "Ssd: Single shot multibox detector." Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016.

[10] Varna, Dainius, and Vytautas Abromavičius. "A system for a real-time electronic component detection and classification on a conveyor belt." Applied Sciences 12.11 (2022): 5608.

[11] Open Data Annotation Platform: https://www.cvat.ai/

[12] Aharon, Nir, Roy Orfaig, and Ben-Zion Bobrovsky. "BoT-SORT: Robust associations multi-pedestrian tracking." *arXiv preprint arXiv:2206.14651* (2022).

Веб Камер Ашиглан Нүдний Хөдөлгөөнөөр Виртуал Аяллыг Удирдах

Рэнцэндорж Жавхлан*
Мэдээллийн Технологи,
Компьютерийн Ухааны Тэнхим
Мэдээллийн Технологи,
Электроникийн Сургууль, Монгол
Улсын Их Сургууль
Улаанбаатар, Монгол Улс
*r.javkhlan@num.edu.mn

Намсрайдорж Мөнхцэцэг
Мэдээллийн Технологи,
Компьютерийн Ухааны Тэнхим
Мэдээллийн Технологи,
Электроникийн Сургууль, Монгол
Улсын Их Сургууль
Улаанбаатар, Монгол Улс
munkhtsetseg_n@num.edu.mn

Намсрай Оюун-Эрдэнэ **
Мэдээллийн Технологи,
Компьютерийн Ухааны Тэнхим
Мэдээллийн Технологи,
Электроникийн Сургууль, Монгол
Улсын Их Сургууль
Улаанбаатар, Монгол Улс
**oyunerdene@num.edu.mn

Эрдэнэбаатар Цэцэгжаргал
Мэдээллийн Технологи,
Компьютерийн Ухааны Тэнхим
Мэдээллийн Технологи,
Электроникийн Сургууль, Монгол
Улсын Их Сургууль
Улаанбаатар, Монгол Улс
tsetsegjargal.e@num.edu.mn

Цоодол Лхамролом
Мэдээллийн Технологи,
Компьютерийн Ухааны Тэнхим
Мэдээллийн Технологи,
Электроникийн Сургууль, Монгол
Улсын Их Сургууль
Улаанбаатар, Монгол Улс
lkhamrolom@num.edu.mn

Хураангуй — Виртуал бодит байдал (VR) технологи нь боловсрол, зугаа цэнгэл, соёлын өвийг сурталчлахад илүү гүнзгий, бодит мэдрэмжийг хэрэглэгчдэд олгодог хүчирхэг хэрэгсэл болон хөгжиж байна. Гэвч өнөөдөр ашиглагдаж буй виртуал аяллууд ихэвчлэн гара ар эсвэл удирдлага ашиглан хийгддэг учир хэрэглэгчийн оролцоо, мэдрэмжийг хязгаарладаг. Энэхүү судалгаагаар веб камерын тусламжтайгаар хүний нүдний хөдөлгөөнийг бодит цаг хугацаанд хянаж, ямар ч нэмэлт мэдрэгч төхөөрөмж ашиглахгүйгээр виртуал аяллыг илүү бодитой, хэрэглэгчид таламжтай хэлбэрээр удирдах шинэлэг аргыг танилцууллаа. Энэхүү систем нь хүний нүдний харц болон физиологийн бусад үзүүлэлтүүдийг 360 градусын виртуал орчинтой холбон, хэрэглэгчийн анхаарал төвлөрөлд нийцүүлэн орчныг өөрчилдөг. Энэхүү аргыг олон түвшний дүрслэл, нүдний хөдөлгөөний загварчлал, анхааралд тулгуурласан динамик тохируулгаар баяжуулсан. Туршилтын үр дүнгээр веб камер ашигласан энэхүү технологи нь уламжлалт аргуудаас илүү хүртээмжтэй, хэрэглэгчийн танин мэдэхүйн ачааллыг багасгаж, илүү гүнзгий мэдрэмжийг төрүүлсэн нь тогтоогдсон. Иймд энэхүү судалгаагаар веб камерт суурилсан харааны хяналтын систем нь VR төхөөрөмж ашиглах боломжгүй эсвэл зардал өндөртэй орчинд илүү тохиромжтой, үр ашигтай шийдэл болохыг онцоллоо.

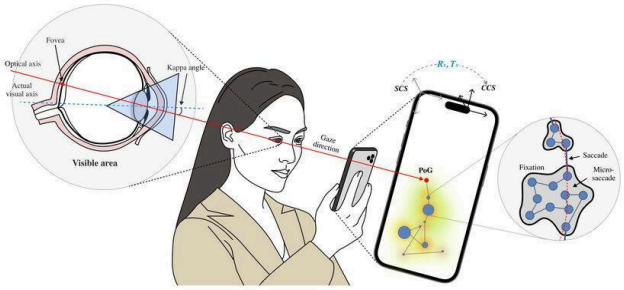
Түлхүүр үг — Виртуал бодит байдал, хараа хянах, веб камер, виртуал аялал, нүдээр удирдах, хүн-компьютерийн харилцаа

I. Танилцуулга

Виртуал бодит байдлын (VR) төхөөрөмжүүд музей, соёлын өв, боловсролын орчинд виртуал аялал хийх өргөн боломжийг олгож буй боловч өртөг өндөр, олон нийтэд хүртээмжгүй байна. Харин энгийн веб камер ашиглан хүний нүдний хөдөлгөөнийг хянаж, навигац болон харилцааг хангах шинэ боломж гарч ирж байна. Орчин үеийн компьютерийн харааны технологи ашигласнаар жирийн веб камераар ч хэрэглэгчийн харааны чиглэлийг маш нарийвчлалтай тодорхойлох боломжтой болсон нь энэхүү технологийг илүү хүртээмжтэй болгож байна. Энэхүү судалгаагаар веб камерын тусламжтайгаар нүдний хөдөлгөөнөөр виртуал орчныг удирдах аргыг VR төхөөрөмжтэй харьцуулж, физиологийн онолын үндэс, хэрэгжүүлэлтийн техник шийдлүүд болон боловсрол, соёлын өв, аялал жуулчлал, хүртээмжтэй байдлын бодит жишээг танилцуулж, хэрэглэгчийн оролцоог нэмэгдүүлэх хямд өртөгтэй шийдлийн давуу талуудыг тодорхойллоо.

II. Хүний нүдний хөдөлгөөн ба харааны физиологи

Хүний нүдний хөдөлгөөн ба харааны физиологи суурилсан интерфэйсийг бүтээхийн тулд хүний нүдний физиологи, хөдөлгөөнийг нарийн ойлгох шаардлагатай. Хүний харааны системийн төв хэсэг нь фовеад байрладаг бөгөөд хүн аливаа зүйлийг тод харахын тулд фовеагаар төвлөрүүлдэг. Нүдний хөдөлгөөний үндсэн төрлүүдэд дараах зүйлс орно:



Зураг 1. Хүний нүдний хөдөлгөөн ба харааны физиологи [11]

Гол үйлдлүүд нь:

- Фиксаци (Fixations): Тодорхой нэг зүйл дээр анхаарлаа төвлөрүүлж (~0-400 миллисекунд үргэлжилдэг), хараагаа тогтворжуулах үйл явц бөгөөд энэ үед хүн мэдээллийг хүлээн авч, боловсруулдаг. [2].
- Саккад (Saccades): Нүдний хурдан шилжилт бөгөөд хараг нэг фиксацаас дараагийн фиксаци руу

шилжүүлдэг хөдөлгөөн юм. Энэхүү хөдөлгөөн нь шинэ мэдээллийг харааны төвд оруулдаг. [2].

- Зөөлөн дагах хөдөлгөөн (Smooth pursuit): Хөдөлж буй объектыг хараагаар мөрдөхөд ашигладаг удаан, тасралтгүй хөдөлгөөн юм [2].
- Хүүхэн харааны өргөсөлт (Pupil Dilation): Хүүхэн харааны хэмжээ өөрчлөгдөх нь танин мэдэхүйн ачаалал, сэтгэл зүйн байдлын үзүүлэлт болж, гэрэл, сэтгэл хөдлөлд мэдрэг байдаг [2].
- Нүд цавчих (Blink Rate ба Eye Closure): Нүдний ядаргаа, анхаарлын түвшинг илэрхийлэх чухал үзүүлэлт юм. Энэхүү физиологийн үзүүлэлтүүд нь хэрэглэгчийн анхаарал, сэтгэл зүйн байдлын талаар мэдээлэл өгч, харааND суурилсан интерфэйсийн загварыг илүү боловсронгуй болгох боломжийг олгодог [2].

Хүний эдгээр байгалийн нүдний хөдөлгөөнүүд нь хэрэглэгчийн анхаарлын төвлөрөл, танин мэдэхүйн байдал, эрүүл мэндийн талаар үнэ цэнтэй мэдээллийг өгдөг. Жишээлбэл, хэрэглэгчийн харааны хэв маягаас анхаарал төвлөрч байгаа эсвэл анхаарал сарниж байгаа эсэхийг тодорхойлж болно. Удаан хугацааны фиксаци болон огцом саккадууд нь хэрэглэгчийн анхаарал төвлөрөлт эсвэл сатаарлыг илтгэдэг. Мөн эмнэлгийн салбарт хэвийн бус нүдний хөдөлгөөн (тогтворгүй фиксаци эсвэл өөрчлөгдсөн саккадын динамик гэх мэт) нь мэдрэлийн өвчин, ядралтын шинж тэмдэг байж болдгийг тогтоосон. Пупиллометр (хүүхэн харааны хэмжилт) нь сэтгэлзүйд сонирхол болон сэтгэл хөдлөлийг хэмжихэд ашиглагддаг бол, нүд цавчилтын шинжилгээ нь ядралт болон танин мэдэхүйн ачааллыг илрүүлэх боломжтой. Энэ бүх шинжлэх ухааны үндэслэлүүд нь веб камерт суурилсан нүдний хөдөлгөөний системийг хэрэглэгчийн интерфэйс загварт ашиглах баттай суурийг бий болгож байна.

III. Веб камерт суурилсан нүдний хөдөлгөөний хяналт: Боломж ба Нарийвчлал

Анхны нүдний хөдөлгөөн хянах системүүд өндөр нарийвчлалтай үр дүн гаргахын тулд тусгай зориулалтын хэт улаан туяаны (IR) төхөөрөмж ашигладаг байсан бол өнөө үед компьютер харааны технологийн хөгжлийн үр дүнд энгийн веб камераар хүний нүдний хөдөлгөөнийг хянах боломжтой болжээ [8].

А. Нарийвчлалын судалгаанууд:

Веб камерт суурилсан хараа хянах технологийн бодит боломжийг хэд хэдэн судалгаагаар баталсан байна. Тухайлбал, Nascimento нарын судалгаагаар веб камерт суурилсан нүд хянах систем нь арилжааны IR төхөөрөмжтэй харьцуулж болохуйц (~11.58% алдаа) нарийвчлалыг (~11.29% алдаа) үзүүлсэн бөгөөд энэхүү нарийвчлалын ялгаа статистикийн хувьд ач холбогдолгүй болохыг тогтоожээ [3]. Уг судалгаагаар веб камер нь IR төхөөрөмжөөс бага зэрэг илүү хувьсамтгай, бага нарийвчлалтай боловч практик хэрэглээнд энэ ялгаа төдийлөн ач холбогдолгүй гэж үзсэн байна. Судлаачдын дүгнэснээр веб камер дээр суурилсан хараа хянах технологи нь бодит орчинд, ялангуяа сургалтын танхимд харааны зан төлөвийг хянахад IR төхөөрөмжтэй бараг ижил үр дүн үзүүлэх чадвартай болохыг тогтоосон.

Өөр нэгэн судалгаагаар Liu нар хэдэн арван оролцогчийг хамруулан веб камер болон өндөр түвшний SMI RED 250mobile нэртэй IR төхөөрөмжийг уншлагын даалгавар дээр харьцуулжээ. Уг судалгаагаар веб камерын төхөөрөмж нь нүдний хөдөлгөөний үйл явдал (саккад, фиксаци гэх мэт)-ыг илрүүлэхэд арилжааны төхөөрөмжөөс ялгарахгүй өндөр нарийвчлалтай болохыг тогтоосон. Хэдийгээр веб камер нь харааны байрлалыг нарийн тодорхойлох чадварын хувьд арай сул байсан ч ерөнхийдөө нүдний хөдөлгөөний дараалал, үйл явдлыг бүрэн хангалттай илрүүлж чадсан байна. Энэхүү үр дүнгээс веб камерын тохиргоо нь үнэтэй тусгай төхөөрөмжийг бүрэн орлох боломжтойг нотолжээ.

Б. Бодит хэрэглээний баталгаа:

Веб камерт суурилсан хараа хянах технологи нь зөвхөн лабораторийн нөхцөлд ч биш, бодит хэрэглээний орчинд ч үр дүнтэй гэдгээ харуулсан. Жишээлбэл, Hutt нарын 100 гаруй хэрэглэгчийг хамруулсан онлайн уншлагын туршилтаар WebGazer нэртэй веб камерын системийг ашиглан хэрэглэгчийн анхаарал сарних зэрэг танин мэдэхүйн далд төлөвүүдийг амжилттай илрүүлжээ. Зөвхөн нэг удаагийн тохируулгын дараагаар хоёр өөр туршилтын бүлэгт уг систем нь хэрэглэгчид анхааралгүй байх үеийг өндөр нарийвчлалтайгаар тодорхойлж чадсан байна [8]. Энэ нь веб хөтөч дээр суурилсан энэхүү систем лабораторийн хамгийн сайн нөхцөлгүй орчинд ч хэрэглэгчийн анхаарал төвлөрөлтийг өндөр түвшинд тодорхойлох боломжтойг харуулсан юм. Үүнтэй ижил төстэйгээр боловсролын салбарт веб камерт суурилсан хараа хянах системүүд нь сурагчдын анхаарлын төвлөрлийг хэмжих, онлайн шалгалтын явцад хуулах зэргийг илрүүлэхэд үр дүнтэй болохыг харуулсан хэд хэдэн судалгаа бий [3], [4].

Эдгээр бүх судалгаанаас харахад веб камерт суурилсан нүдний хөдөлгөөн хянах технологи нь хүн-компьютерийн харилцааны олон төрлийн хэрэглээнд техникийн хувьд хэрэглэх боломжтой, хангалттай нарийвчлалтай байна. Орчин үеийн алгоритмууд нь IR төхөөрөмжөөс бага чанартай (хэт улаан туяагүй энгийн видео дүрс) дүрсийг ашигладаг ч хэдхэн градусын алдаатайгаар харааны байрлалыг өндөр түвшинд тодорхойлох чадвартай болсон [7]. Тусгай зориулалтын IR төхөөрөмжүүд нь харааны байрлалыг 0.5°-1.0°-ийн алдаатайгаар тодорхойлдог бол веб камер дээр суурилсан системүүд нь ердийн дэлгэц дээр хэдхэн сантиметрийн нарийвчлалтай ажиллаж байна. Энэ нь интерфейсийн удирдлага, хэрэглэгчийн судалгаа, боловсролын адаптив систем зэрэгт бүрэн хангалттай үзүүлэлт юм.

Хамгийн чухал нь веб камерын шийдлүүд нь бараг ямар ч нэмэлт зардалгүйгээр өндөр үр дүн үзүүлэх боломжтой. Нүд хянах чиглэлийн нэгэн симпозиумын илтгэлд дурдсанаар "Веб камерт суурилсан хараа хянах технологи нь IR төхөөрөмжүүдийн хямд өртөгтэй, бодит хувилбар болж чадна" гэжээ [7]. Хэдийгээр веб камерын системүүд нь заримдаа гэрэлтүүлэг, толгойны хөдөлгөөн зэргээс хамаарч мэдрэмтгий байдаг ч компьютерийн хараа болон машин сургалтын технологийн дэвшлийн үр дүнд энэхүү дутагдлууд нь улам багасаж, үр дүнтэй шүүлтүүр, тохируулгын тусламжтайгаар интерактив аппликейшнийг удирдахад шаардлагатай гүйцэтгэлийг хангах боломжтой болж байна. Энэ нь веб камерт суурилсан харааны удирдлага нь өргөн хүрээнд

ашиглагдах, хүртээмжтэй, үр ашигтай шийдэл болох боломжтойг харуулж байна.

IV. Веб камер ашиглан хараа хянах технологийг VR төхөөрөмжтэй харьцуулах нь

VR төхөөрөмжүүд ихэвчлэн гүн мэдрэгч (IMU), гүний камерууд, заримдаа дотор суурилагдсан нүд хянах төхөөрөмжийг ашиглан хэрэглэгчдэд илүү бодит мэдрэмж бүхий туршлага олгодог. Гэвч веб камер дээр суурилсан нүдний хөдөлгөөн хянах технологи болон VR төхөөрөмж дээр суурилсан технологийн хооронд өртөг, хүртээмж, нууцлал, үр дүнгийн хувьд томоохон ялгаа бий.

Төхөөрөмжийн өртөг: Жирийн веб камер нь ихэнх зөөврийн компьютер, таблет, гар утсанд суурилагдсан байдаг тул үнэ төлбөргүй эсвэл нэмэлтээр худалдан авахад ойролцоогоор 20-50 долларын үнэтэй байдаг. Харин нүдний хөдөлгөөнийг хянах чадвартай VR төхөөрөмжүүд нь тусгай зориулалтын өндөр өртөгтэй төхөөрөмж бөгөөд зөвхөн нүд хянах модулийг нэмэхэд л хэдэн зуун ам.долларaас (ойролцоогоор 200-1700 доллар хүртэл) хэдэн мянган ам.долларын өртөгтэй болдог (WorldViz VR Budgeting Guidelines, 2025). Тухайлбал, Vive Pro Eye зэрэг дээд түвшний мэргэжлийн VR төхөөрөмжүүд 1000 долларoос илүү үнэтэй байдаг. Иймээс веб камерт суурилсан шийдэл нь өртгийг эрс бууруулж, байгууллага болон сургуулиуд нэмэлт тоног төхөөрөмж худалдан авахгүйгээр өргөнөөр хэрэглэх боломжийг олгодог [5].

Хэрэглэгчийн хүртээмж: Веб камер нь нийтлэг бөгөөд дэлхий даяар тэрбум тэрбумаар тоологдох хэрэглэгчдийн төхөөрөмжүүдэд суурилагдсан байдаг. Харин VR төхөөрөмжийн эзэмшил харьцангуй бага, жишээлбэл АНУ-д л гэхэд насанд хүрэгчдийн зөвхөн 19% нь VR-ийг ашиглаж үзсэн байдаг (Virtual Reality Statistics and Facts, 2025). Энэ нь VR шаарддаг виртуал аяллыг бүтээхэд хэрэглэгчдийн хүрээ хязгаарлагдана гэсэн үг юм. Үүнтэй харьцуулахад интернет холболттой ямар ч төхөөрөмжтэй хүн веб камер ашиглан виртуал орчныг хялбархан удирдах боломжтой байдаг. Үүнээс гадна, веб камер ашиглан нүдний хөдөлгөөн хянах технологи нь хэрэглэгчдэд илүү хялбар, энгийн бөгөөд ердөө богино хугацааны тохируулга л шаарддаг. VR төхөөрөмжүүд харьцангуй хүнд, хэрэглэгчдэд эвгүй мэдрэмж төрүүлж, удаан хугацааны хэрэглээнд толгой эргэх зовиур үүсгэдэг. Иймд залуу хүүхдүүд, ахмад настнууд, хөгжлийн бэрхшээлтэй хүмүүс зэрэг бүлгүүдэд веб камер ашиглах нь илүү хүртээмжтэй, тохиромжтой байдаг [6].

Нууцлалын асуудал: Нүдний хөдөлгөөний мэдээлэл нь маш эмзэг бөгөөд хэрэглэгчийн хувийн мэдээллийг (таних тэмдэг, нас хүйс, сэтгэл хөдлөл, эрүүл мэндийн үзүүлэлт г.м) агуулж байдаг. Зарим судалгаагаар нүдний түүхий мэдээлэл нь биометрийн танигч, танин мэдэхүйн онцлог болон эрүүл мэндийн байдлыг илэрхийлэх чадвартай болохыг харуулсан (What Does Your Gaze Reveal About You? [2]). VR төхөөрөмжүүдийн хувьд ийм мэдээллийг интернетээр дамжуулан үүлэн дээр хадгалах нь түгээмэл тул мэдээлэл алдагдах эрсдэл өндөр. Харин веб камер дээр суурилсан системүүд нь бүх мэдээллийг хэрэглэгчийн төхөөрөмж дээр боловсруулж, сервер рүү ямар ч дүрс бичлэг эсвэл мэдээлэл илгээдэггүй тул хэрэглэгчийн нууцлалыг илүү сайн хамгаалдаг (WebGazer.js). Ийм хэрэглэгч талын боловсруулалт хийдэг арга нь хэрэглэгчдэд өөрсдийн мэдээллээ хянах бүрэн боломж олгодог. Иймд веб камер дээр суурилсан системүүд нь мэдээллийн нууцлалыг илүү сайн хангадаг бөгөөд VR төхөөрөмжүүдийн хувьд үүлэнд боловсруулалт хийдэг тул илүү их болгоомжлол шаарддаг [7].

Үр дүн ба нарийвчлал: VR төхөөрөмжүүд нь зөвхөн нүдний хөдөлгөөн төдийгүй толгойн байрлал болон гарын удирдлагыг өндөр нарийвчлалтайгаар тодорхойлдог. Энэ нь хэрэглэгчдэд байгалийн хөдөлгөөнөөр виртуал орчныг хянах, өндөр түвшний дүрслэлтэй сонголт хийх боломж олгодог. Харин веб камер дээр суурилсан системүүд зөвхөн дэлгэцийн 2D орчинд нүдний харааны координатыг хянадаг бөгөөд толгойн 3D хөдөлгөөнийг бүрэн тодорхойлох боломжгүй байдаг (зарим системүүд веб камерын тусламжтай толгойн байрлалыг бага зэрэг тооцоолдог). Энэ нь VR төхөөрөмжүүд илүү гүнзгий, бодит мэдрэмжийг бий болгож, хэрэглэгч байгалийн толгой хөдөлгөөнөөр виртуал орчныг удирдах боломж олгодог гэсэн үг юм. Нарийвчлалын хувьд тусгай зориулалтын VR төхөөрөмжүүд веб камераас илүү байдаг ч веб камерын нарийвчлал нь виртуал аяллыг удирдахад хангалттай бөгөөд бага зэргийн нарийвчлалын алдаа нь практик хэрэглээнд бараг мэдэгдэхгүй байдгийг судалгаагаар нотолсон [3],[4],[1].

Иймд веб камер нь бага зэргийн нарийвчлалын дутагдалтай ч олон төрлийн харилцаа болон интерактив хэрэглээнд хангалттай үр дүн үзүүлдэг.

Виртуал аяллын хувьд VR төхөөрөмж нь хэрэглэгчийг 360° орчинд бүрэн гүнзгий оруулдаг бол веб камер дээр суурилсан систем нь ердийн дэлгэц дээр харааны чиглэлээрээ зураг эргүүлэх, объектуудыг томруулж харуулах зэргээр илүү хүртээмжтэй аргачлал ашигладаг. Энэ нь хэрэглэгчийн талын тоног төхөөрөмжийн шаардлагыг эрс бууруулж, илүү олон хүнд хүрэх боломжийг олгодог. Иймд боловсролын болон соёлын өвийн салбарт веб камер дээр суурилсан системийг хэрэглэх нь хэрэглэгчийн тоог нэмэгдүүлж, хүртээмжийг сайжруулах чухал ач холбогдолтой юм. Энэ нь илүү олон хүний сонирхлыг татах, хэрэглэгчийн оролцоог нэмэгдүүлэх үр дүнтэй шийдэл болно.

V. Веб технологи (JavaScript ба HTML) ашиглан хэрэгжүүлэх нь

Веб камер ашиглан нүдний хөдөлгөөнийг хянах технологийн хамгийн давуу талуудын нэг нь тусгай програм хангамж эсвэл нэмэлт төхөөрөмж шаардалгүйгээр стандарт веб технологиор хэрэгжүүлж болдог юм. Орчин үеийн веб хөтөчүүдэд камер ашиглах боломж (camera access), дүрсийг бодит хугацаанд боловсруулах (real-time video processing), JavaScript дээр харааны байрлал тогтоох алгоритм (gaze estimation algorithms) ажиллуулахад шаардлагатай бүх нөхцөл боломж бүрдсэн байдаг [8].

A. Нүдний хөдөлгөөн таних:

Энэ чиглэлээр хамгийн их танигдсан жишээ бол WebGazer.js бөгөөд энэ нь ямар ч веб камерыг шууд нүдний хөдөлгөөн хянах төхөөрөмж болгон хувиргадаг JavaScript сан юм. WebGazer нь “веб хөтөч дээр

суурилсан нүдний хөдөлгөөнийг хянах технологийг хүртээмжтэй болгох” зорилгоор бүтээгдсэн бөгөөд хэдхэн мөр код ашиглан дурын веб хуудсанд суулгаж, хэрэглэгчийн харааны байрлалыг дэлгэцэн дээрх (x, y) координатаар бодит хугацаанд авах боломжийг олгодог. WebGazer нь хэрэглэгчийн нүдний хөдөлгөөнийг дүрсний урсгалаас илрүүлж, нүдний онцлог шинжүүдийг дэлгэц дээрх байрлалтай холбохын тулд компьютерийн харааны технологийг ашигладаг. Мөн энэхүү систем нь өөрийгөө автоматаар тохируулах (self-calibration) боломжтой бөгөөд хэрэглэгчийн хулганаар дарах, текст унших гэх мэт энгийн үйлдлийг ажиглан, гараар нарийн тохируулга хийх шаардлагагүйгээр аажмаар тохируулгаа хийдэг энэ санааг анх Papoutsaki-гийн докторын ажилд бүрэн тодорхойлсон [12]. Энэ бүх үйл явц нь HTML5-ийн API (жишээлбэл, getUserMedia болон <canvas> ашиглан дүрс боловсруулалт хийх) тусламжтайгаар хэрэглэгчийн веб хөтөч дотор л ажилладаг тул мэдээлэл сервер рүү илгээгддэггүй бөгөөд хэрэглэгчийн хувийн мэдээллийн аюулгүй байдлыг бүрэн хангаж чаддаг. Энэ нь веб хуудсан дээр өндөр хурдтайгаар харааны байрлалыг тодорхойлж, хэрэглэгчийн интерфэйсийг сааталгүйгээр удирдах боломж олгодог.

JavaScript ашиглан харааны хяналтын системийг шинээр бүтээхэд тодорхой хэмжээний сорилттой боловч орчин үеийн хэрэгслүүдийн тусламжтайгаар хэрэгжүүлэх боломжтой болсон. Ерөнхийдөө дараах алхмуудыг хийдэг:

- Веб камераас видео фрэймүүдийг авах. Вебкамерын дүрсийг аваад фрейм тус бүр дээр нүүрний болон нүдний байршлыг илрүүлэх.
- Дүрс дотроос нүүр болон нүдийг тодорхойлох. Нүдний хэсгийн шинж чанарыг (сормуусны үзүүр, хүүхэн харааны байршил гэх мэт) гаргаж авах.
- Нүдний онцлог шинжүүдийг ялган авах (хүүхэн харааны төв, нүдний булангийн тусгал гэх мэт). Эдгээр шинж чанарыг дэлгэцэн дээрх харааны цэг рүү хувиргах ML эсвэл регрессийн загвар ашиглах.
- Олж авсан мэдээллийг машин сургалтын загвар ашиглан дэлгэц дээрх харааны координаттай холбох. Хэрэглэгчид богино калибровк хийлгэх (нүдээрээ тодорхой цэгүүдийг харах) замаар нарийвчлалыг өсгөх.

Хэт улаан туяаны тусгай гэрэлтүүлэг байхгүй нөхцөлд веб камерын системүүд нүдний харагдах байдал дээр үндэслэсэн аргуудыг (appearance-based methods) ихэвчлэн ашигладаг. Энэ нь нүдний дүрсний онцлог шинж (нүдний хэлбэр, нүдний цагаан доторх солонгон бүрхэвчийн байрлал г.м.) ба харааны чиглэлийн хоорондын хамаарлыг тодорхойлдог арга юм. Ийм системийг эхний ээлжинд хэрэглэгч дэлгэц дээрх тодорхой цэгүүдийг харах замаар тохируулж өгдөг бөгөөд цаашид тухайн хэрэглэгчийн харааны байрлалыг дэлгэц дээр хэдхэн сантиметрийн нарийвчлалтайгаар тодорхойлох боломжтой болдог.

Б. Нүд, зовхи, толгойн жест-контроллёр:

Мөн JavaScript дээр суурилсан машин сургалтын (ML) загваруудыг ашиглан хэрэгжилтийг илүү хялбар болгож чаддаг. Жишээлбэл, TensorFlow.js эсвэл MediaPipe ашиглан нүүрний 3D загварчлал (face mesh models) хийж, нүдний өнцөг, хүүхэн харааг тодорхойлон харааны чиглэлийг бодит хугацаанд тооцоолох боломжтой.

Орчин үеийн **face-landmarks-detection** (TensorFlow.js) модель нь 468 нүүрний зангилаа + 10 ирисийн нэмэлт цэгийг бодит цагт (~30 FPS) илрүүлдэг бөгөөд нүд анивчих, толгой гилжийн хөдөлгөөн зэргийг прагматик босго-шүүлтээр ялган танихад хангалттай [13]. Бид дараах 7 жестийг (Ж-1 ~ Ж-7) тодорхойлж, 360° виртуал аяллын командтай рүүл хийлээ.

Жест	**Илрүүлэх онцлох зангилаа (фэйс-мэш ID)**	**Тооцоолох метрик**	**Навигац үйлдэл**
Ж-1 Нүдний хараа (gaze)	Iris (468–477) + нүдний булан 33, 133, 362, 263	Iris-центроид ы → дэлгэц (x,y)	Pan / эргүүлэх
Ж-2 Зовхи нээлт/хаалт (blink ratio)	FACEMESH_LEFT_EYE, FACEMESH_RIGHT_EYE босоо болон хэвтээ зай	EAR < 0.20 (т < 120 мс)	Pause / Resume
Ж-3 Баруун нүдээ анивчих (wink-R)	$O_{(R)}$ EAR < 0.20, $O_{(L)}$ EAR ≥ 0.25	Ялгаатай EAR	Hotspot → Next
Ж-4 Зүүн нүдээ анивчих (wink-L)	$O_{(L)}$ EAR < 0.20, $O_{(R)}$ ≥ 0.25		Hotspot → Prev
Ж-5 Толгой дээш (pitch > +10°)	Нүүр-box тэнхлэг (solvePnP)	Pitch °	FOV Zoom-in
Ж-6 Толгой доош (pitch < –10°)			FOV Zoom-out
Ж-7 Толгой зүүн/ баруун (yaw ± 15°)	Yaw °	нүдний хөдөлгөөндөө ашиглагдах эргэлт	

Хүснэгт 1. Толгойн болон нүдний хөдөлгөөнөөр хэрхэн виртуал аялал удирдах арга зүй

Хүний зориудаар нүдээ анивчих болон байгалийн жамаар анивчих үйлдлүүдийг ялгахдаа EAR (Eye Aspect Ratio)-ийн босго 0.2 нь MediaPipe-ийн blink detection-д хэрэглэгддэг жишиг [14] прагматик байдлаар 3 кадр дараалан хангагдсан нөхцөлийг тооцон үзэж ялгаатай тооцон “анивчилт” гэж шийднэ.

Apple компани 2024 онд iPhone болон iPad дээр нүүрний урд камер ашиглан хэрэглэгчийн интерфэйсийг нүдний хөдөлгөөнөөр удирдах "Eye Tracking" функцийг нэвтрүүлсэн бөгөөд энэ нь нэмэлт тоног төхөөрөмж

шаардлагагүйгээр, зөвхөн секундийн хугацаанд тохиргоо хийж, бүх боловсруулалтыг төхөөрөмж дээрээ хийдэг байна. Энэ нь веб камерын нүд хянах системийн зарчмаас ялгаагүй бөгөөд орчин үеийн мобайл төхөөрөмж дээр ч маш үр дүнтэй хэрэгжиж байгааг харуулж байна.

Өөр нэгэн бодит веб хэрэгжилт бол SearchGazer бөгөөд энэ нь WebGazer-ын тусгай хувилбар бөгөөд хайлтын системийн үр дүнг харааны хяналтаар тодорхойлж, хэрэглэгч аль хайлтын үр дүнг уншиж байгааг таних боломжийг олгодог. Эдгээр төслүүд нь энгийн хэрэглэгчийн төхөөрөмж дээр веб хөтөч дотор секундэд 15-30 кадрын хурдтайгаар харааны хяналтыг бодитоор хэрэгжүүлж болдгийг харуулсан.

Хэрэв хөгжүүлэгчид ямар нэгэн тусгай сан ашиглахгүйгээр веб технологи ашиглан анхан шатны нүд хянах системийг хийхийг хүсвэл HTML5-ийн Canvas API ашиглан камерын дүрсийг боловсруулж болно. Тухайлбал, дүрсийг саарал өнгөнд шилжүүлэн, хүүхэн харааг хар дугуй хэсэг байдлаар тодорхойлж болно. Мөн WebAssembly-ийг ашиглан OpenCV зэрэг дүрс боловсруулалтын сангуудыг веб хөтөч дотор хурдан ажиллуулах боломжтой бөгөөд WebGL эсвэл WebGPU ашиглан JavaScript дээр дүрс боловсруулалтыг хурдасгах боломжтой.

Дүгнэж хэлэхэд веб технологийн орчин нь нүдний хөдөлгөөн хянах системийг шууд веб хөтөч дээр бодит хугацаанд ажиллуулахад хангалттай боловсронгуй болсон. Иймээс веб хөгжүүлэгчид ердөө HTML, JavaScript технологийг ашиглан хэрэглэгчийн хараагаар удирдах боломжтой виртуал аялал зэрэг олон төрлийн интерактив системийг бүтээх бүрэн боломжтой болж байна. Энэ нь веб хөтөч дээр суурилсан системийг өргөн хүрээнд хэрэглэх үүд хаалгыг нээж байгаа юм.

VI. Хэрэглээний бодит жишээнүүд

A. Боловсрол ба цахим сургалт

Веб камер ашиглан нүдний хөдөлгөөнийг хянах технологийг хамгийн үр дүнтэй ашиглаж болох салбаруудын нэг бол боловсрол юм. Онлайн сургалтын үед багш нар сурагчдын анхаарал, хичээлдээ төвлөрч буй байдлыг бодитоор хэмжихэд хүндрэлтэй тулгардаг. Энэ тохиолдолд нүдний хөдөлгөөнийг хянах технологи нь сурагч дэлгэцийн яг аль хэсэгт, хэрхэн анхаарлаа төвлөрүүлж буй талаар бодит өгөгдөл цуглуулж өгдөг. Энгийн веб камер ашиглан цахим сургалтын платформ нь сурагчдын анхаарал суларч, самуурч буй мөчүүдийг шууд илрүүлэх боломжтой болно. Жишээлбэл, хэрэв сурагч хичээлийн бичлэгээс харцаа байнга өөр тийш чиглүүлж байвал систем автоматаар сурагчийн анхаарал сарнисан гэж үзээд интерактив асуулт эсвэл сануулга өгч, дахин оролцоог нь идэвхжүүлж болно. Судалгаагаар энэ боломжтойг нотолсон бөгөөд Hutt болон бусад судлаачдын туршилтаар веб камерын тусламжтайгаар сурагчдын анхаарлын сарнилтыг автоматаар илрүүлээд зогсохгүй, ойлголтын түвшин сул байх магадлалтай сурагчдыг урьдчилан таамаглаж чадсан байна [1].

Мөн энэ технологийн бас нэг боловсролын хэрэглээ нь сурагчдын унших зан төлөв, судалгааны арга барилыг шинжлэх юм. Танхимын хичээлийн явцад сурагчид комик унших үед веб камер ашиглан ямар хэсэгт, хэр удаан харж байгааг тодорхойлсон судалгаа нь багш нарт сурах материалын ямар хэсэг сурагчдын анхаарлыг татаж, ямар хэсэг ойлгоход хүндрэлтэй байгааг тодорхойлоход тусалсан байдаг. Веб камер ашигласан сургалтын орчин дахь бусад судалгааны тоймоор сурагчдын оролцооны түвшин, тэмдэглэл хийх стратеги, бүр онлайн шалгалт өгөх үед хуулж байгаа эсэхийг хүртэл тодорхойлох боломжтойг харуулсан. Жишээлбэл, онлайн шалгалтын үеэр сурагч байнга дэлгэцнээсээ хол эсвэл өөр тийшээ харвал сэжигтэй үйлдэл гэж үзээд хяналтын системд мэдээлэл өгдөг.

Веб камер дээр суурилсан нүдний хөдөлгөөн хянах технологи нь зах зээл дээрх бусад технологиудаас хүний амьдралд илүү ойр учраас жирийн анги танхимд хүртэл нэвтрэх боломжтой юм. Хөдөлгөөний бэрхшээлтэй сурагчид энэхүү технологийг ашиглан зөвхөн нүдний хөдөлгөөнөөрөө боловсролын тоглоом болон бусад сургалтын програм хангамжийг ашиглах боломжтой болно. Мөн эрүүл сурагчдад ч илүү сонирхолтой, интерактив сургалтын арга барилыг санал болгодог. Жишээлбэл, виртуал лабораторид аливаа тоног төхөөрөмжийг харцаараа сонгомогц түүний талаарх мэдээлэл гарч ирэх, цахим ном унших явцад сурагчийн харц нэг газар удаан зогсвол текстийг автоматаар томруулах, уншихыг нь хөнгөвчлөх гэх мэт боломжуудыг санал болгож болно.

Хураангуйлан дүгнэвэл, боловсролын салбар веб камер ашигласан нүдний хөдөлгөөн хянах системээс ихээхэн ашиг тус хүртэх боломжтой бөгөөд энэ нь өмнө нь зөвхөн тусгай лабораторийн орчинд л хийгддэг байсан анхаарлын түвшин, оролцоог хэмжих үйл ажиллагааг өдөр тутмын сургалтын үйл явцад авчрах боломжийг бүрдүүлж байна. Цахим сургалтын орчин улам өргөжиж байгаа өнөө үед уг технологи нь суралцах үйл явцыг илүү хувь хүнд чиглүүлж, сурагчдын анхаарал болон оролцоог нэмэгдүүлэх үр дүнтэй хэрэгсэл болох нь гарцаагүй юм.

B. Виртуал Музей ба Соёлын Өв

Ковид-19 цар тахлын дараагаас музей, соёлын өвийн байгууллагууд дэлхийн үзэгчдэд хүрэхийн тулд виртуал аялалд ихээхэн анхаарал хандуулж байна. Ийм нөхцөлд веб камерын тусламжтай нүдний хөдөлгөөн хянах технологийг ашигласнаар виртуал музейн аяллыг илүү сонирхолтой, хэрэглэгчдэд таатай болгож, музейн удирдлагад үнэ цэнтэй дата өгдөг. Дьюк Их сургуулийн урлагийн түүхийн ангид сурагчид веб камер ашиглан виртуал музейн интерфэйсийн үнэлгээ хийж үзэхэд багш нь веб камерын тусламжтай нүдний хөдөлгөөн хянах технологийг “маш нарийн бөгөөд танин мэдэхүйн нөлөөлөл, оюуны төвлөрөл гэх мэт бодит мэдээллийг цуглуулдаг өндөр үр дүнтэй хэрэгсэл” гэж тодорхойлсон байдаг. Энэхүү судалгааны үр дүнд музейн дижитал загвар, контентын үр нөлөөг сайжруулахад ямар хэсгүүдийг хэрэглэгчид илүү анхаарч үздэг, аль хэсэгт мэдээлэл дутмаг байгаа зэргийг тодорхойлж чаджээ.

Хэрэглэгчийн хувьд энэхүү технологи нь виртуал музейгээр аялахад хулгана ашиглан зураг гүйлгэхийн оронд зүгээр л харах чиглэлээрээ галерей дотор чөлөөтэй шилжих боломжийг олгодог. Хэрэглэгчийн

сонирхсон үзмэр дээр хэсэг хугацаанд харцаа төвлөрүүлэхэд тухайн үзмэрийн талаарх дэлгэрэнгүй мэдээлэл автоматаар гарч ирнэ. Энэ нь яг л бодит музейгээр явж байгаа мэт хэрэглэгчийн сониуч зан, анхааралд тулгуурлан аяллыг хөтөлдөг.

Мөн виртуал музей, соёлын өвийн виртуал аялалд нүдний хөдөлгөөнийг ашигласнаар өгүүлэмжийг илүү сонирхолтой болгож болно. Тухайлбал, виртуал хөтөч (хиймэл оюунтай аватар) нь хэрэглэгчийн анхаарч буй үзмэрийн онцлог дээр суурилан тайлбараа өөрчлөх боломжтой. Хэрэглэгч нэг зураг эсвэл үзмэрийн жижиг деталиас харцаа салгахгүй байвал хөтөч тухайн деталь дээр нэмэлт мэдээлэл өгч болно. Харин хэрэглэгчийн анхаарал сарниж, өөр зүйл рүү шилжиж байвал хөтөч илүү товч тайлбар өгөх эсвэл асуулт асууж хэрэглэгчийг буцаан татаж болно.

Музейн байгууллагуудын хувьд веб камер ашигласан технологи нь бодит үзэсгэлэнд тусгай шил зүүлгэж хийгддэг нүдний хөдөлгөөний судалгаанаас ч илүү үр дүнтэй байж болно. Виртуал үзэгчид аль хэсгийг илүү анзаарч, анхаарч байгааг тодорхойлох нь музейн дижитал болон бодит үзэсгэлэнгийн загварыг сайжруулах үнэ цэнтэй өгөгдөл болж чадна.

Эцэст нь, веб камерын тусламжтайгаар нүдний хөдөлгөөн хянах технологиор баяжуулсан виртуал музейн аялал нь хэрэглэгчдэд илүү сонирхолтой, интуитив мэдрэмжийг өгч, музейн байгууллагуудад хэрэглэгчдийн зан төлөвийг ойлгох, виртуал контентыг сайжруулахад чухал түлхэц болж байна.

C. Виртуал аялал ба алсын зайнаас судлах боломж

Музейн хүрээнээс цаашлаад аялал жуулчлалын салбар ч гэсэн веб камер ашиглан нүдний хөдөлгөөнөөр удирдах технологийг виртуал аялал, судалгааны чиглэлээр ашиглах боломжтой болж байна. Виртуал аялал гэдэг нь 360 градусын зураг, видеог ашиглан хот, байгалийн гайхамшигт газрууд, түүхэн дурсгалт газруудаар цахимаар зочлохыг хэлдэг бөгөөд энэ нь хүмүүс аялах боломжгүй эсвэл аялахыг хүссэн газраа урьдчилан судлах боломж олгодог. Ихэнхдээ хэрэглэгчид ийм контентыг хулгана, мэдрэгчтэй дэлгэц, эсвэл VR төхөөрөмжөөр удирддаг байсан бол веб камер дээр суурилсан нүдний хөдөлгөөн хянах технологи нь дундын шийдлийг санал болгож байна: энэ нь хулганаас илүү байгалийн бөгөөд хялбар, мөн бүрэн VR төхөөрөмж шаардахгүйгээр энгийн камерын тусламжтайгаар ажиллах боломжийг бүрдүүлдэг.

Жишээлбэл, эртний балгас, ЮНЕСКО-гийн дэлхийн өвд бүртгэгдсэн газруудаар виртуал аялал хийхдээ веб камерын тусламжтай хэрэглэгчийн харцыг удирдлага болгон ашиглах боломжтой. Виртуал сүмийн тааз руу харвал фреск зураг харагдахуйцаар дээшээ харах өнцгийг автоматаар өөрчилж, баруун тийш харахад дараагийн танхим руу шилжих байдлаар орчныг удирдана. Энэ нь хэрэглэгчид ердөө л харцаараа виртуал орчныг “эргүүлж харах” бодит мэдрэмжийг өгч чадна. Виртуал орчинд урагшлахын тулд тусгай цэгүүдэд харцаа түр хугацаанд төвлөрүүлснээр дараагийн байршил руу шилжих үйлдэл хийж болох бөгөөд энэ аргачлал нь VR орчинд хэдийн туршигдаж, үр дүнгээ өгсөн байдаг. Ингэснээр хэрэглэгчид ширээний ард суусан хэвээр, яг л бодит орчинд биеэрээ явж байгаа мэт дэлгэц дээрх орчныг харцаараа удирдан аялах боломжтой болно.

Энэхүү технологи нь эрүүл мэндийн асуудалтай, санхүүгийн боломжгүй, эсвэл өөр ямар нэг шалтгаанаар бодитоор аялах боломжгүй хүмүүст асар их ач холбогдолтой. Энэ нь дэлхийн өнцөг булан бүрд байгаа гайхамшигт газруудыг виртуал орчинд мэдрэх боломжийг хүн бүрт нээж өгдөг. Аялал жуулчлалын байгууллага, боловсролын салбарын хувьд энэ нь илүү өргөн хүрээний үзэгчдийг агуулгатай холбож, утга учиртайгаар харилцах боломжийг бүрдүүлдэг. Үүнээс гадна музейн контентын жишээтэй ижил, виртуал аялал хийж буй хэрэглэгчийн харцны мэдээллээс тухайн контентийн аль хэсэг илүү сонирхол татаж байгааг тодорхойлж, аяллын танилцуулга, маркетингийн арга барилыг улам сайжруулахад ашиглах боломжтой юм.

Үүнээс гадна алсын зайны аялал, хүртээмжтэй байдлыг сайжруулах зорилгоор нүдний хөдөлгөөний удирдлагыг телепрезенс (алсын удирдлагатай) роботтой нэгтгэх шинэлэг шийдлүүд ч бий болж байна. Зарим музей, аялал жуулчлалын газрууд өөрийн биеэр ирж чадахгүй хүмүүсийг зориулсан алсын зайнаас удирдах роботыг туршиж эхэлсэн бөгөөд эдгээр роботуудыг хэрэглэгчид интернетээр дамжуулан хулгана, гар эсвэл джойстикоор удирддаг байсан юм. Харин одоо веб камерын тусламжтайгаар харцаараа удирдах боломжтой интерфэйсийг роботтой хослуулснаар гар хөдөлгөх боломжгүй, гэрээсээ гарах боломжгүй хүмүүс хүртэл роботоор дамжуулан музей эсвэл аяллын бүсэд өөрийн хүслээр аялах боломжтой болж байна. Роботын камерын дүрс хэрэглэгчийн дэлгэц дээр харагдаж байгаа үед тухайн үзмэр эсвэл объектыг харснаар роботыг тэр чиглэлд явуулах боломжтой бөгөөд энэ нь гарын хөдөлгөөний асуудалтай хүмүүсийн хувьд аяллын туршлагыг илүү хялбар болгож өгдөг.

Нэгэн жишээ дурдахад, дөрвөн мөч нь саажсан хэрэглэгч алсын зайнаас телепрезенс роботын тусламжтайгаар музей үзэж чадсан бодит тохиолдол байдаг бөгөөд уг технологийг веб камерын тусламжтай нүдний хөдөлгөөнөөр баяжуулснаар удирдлагыг илүү энгийн болгож, хүртээмжийг нэмэгдүүлж чадна.

Эцэст нь веб камер ашиглан нүдний хөдөлгөөнөөр удирдах виртуал аялал нь интернетэд холбогдсон хэн бүхэнд хүртээмжтэй, идэвхтэй бөгөөд сонирхолтой аяллын туршлагыг бий болгож чадна. Энэхүү технологи нь сургууль (сурагчид нүдээрээ түүхэн газруудаар аялах боломжтой), ахмад настны асрамжийн газар (ахмад хүмүүс төрсөн нутгаараа аялах боломжтой) болон аялал жуулчлалын маркетинг зэрэг олон төрлийн хэрэглээнд тохиромжтой юм. Ингэхдээ хэрэглэгчээс нэмэлт төхөөрөмж шаардахгүйгээр, зөвхөн энгийн веб камер ашиглахад л хангалттай.

D. Хүртээмж ба Тусгай Хэрэгцээт Хүмүүст

Веб камер ашиглан нүдний хөдөлгөөн хянах технологийн хамгийн том давуу тал нь хөгжлийн бэрхшээлтэй хүмүүсийн хүртээмжийг нэмэгдүүлэхэд чиглэдэг. Гар болон хулгана ашиглах боломжгүй хүмүүс (ALS, тархины гэмтэл, хүнд хэлбэрийн артрит гэх мэт) веб камер ашиглан компьютерийн интерфэйсийг зөвхөн нүдний хөдөлгөөнөөрөө удирдах боломжтой болно.

Apple компани 2024 онд iPhone болон iPad дээр нүдээр удирдах боломжтой функц нэвтрүүлснээр хэрэглэгчид өөрийн төхөөрөмж дээр ямар ч нэмэлт тоног төхөөрөмж шаардлагагүйгээр интерфэйсийг удирдах боломжтой болжээ. Энэ нь веб камер дээр суурилсан нүд хянах технологийн бодит хэрэглээний хамгийн тод жишээ болж байна.

Энэхүү технологи нь хөгжлийн бэрхшээлтэй сурагчдад боловсролд тэгш оролцох боломж олгохын зэрэгцээ гар нь түр хугацаагаар гэмтсэн эсвэл хэвтэрт байгаа хэрэглэгчдэд ч ашиглахад тохиромжтой. Ийнхүү веб камерын тусламжтайгаар хүртээмжтэй байдлыг эрс нэмэгдүүлж байна.

VII. Дүгнэлт

Веб камерт суурилсан нүдний хөдөлгөөнийг хянах технологи нь VR төхөөрөмжүүд болон тусгай зориулалтын нүд хянах төхөөрөмжүүдийн үр ашигтай бөгөөд практик хувилбар болон хүч түрэн гарч ирлээ. Энэхүү технологи нь хаа сайгүй тархсан энгийн веб камерууд болон орчин үеийн компьютер харааны технологийг ашиглан, өмнө нь зөвхөн тусгай төхөөрөмжөөр л хэрэгждэг байсан нүдний хөдөлгөөний хяналтыг олон нийтэд хүртээмжтэй болгож байна [1],[3],[4]. Манай энэхүү судалгааны тоймд дурдсанаар, уг технологи нь хэрэглэгчийн харааны байршлыг өндөр нарийвчлалтайгаар тодорхойлж, хэрэглэгчийн танин мэдэхүйн төлөв байдлыг хүртэл урьдчилан таамаглах боломжтойг шинжлэх ухааны үндэслэлээр баталж байна. Мөн хүний физиологийн олон төрлийн үзүүлэлт, тухайлбал саккад хөдөлгөөн, хүүхэн харааны хэмжээний өөрчлөлт зэрэг нь хэрэглэгчийн анхаарал, сэтгэл зүйн байдлыг нарийвчлан хэмжих боломжийг олгож, хэрэглэгчид төвлөрсөн интерактив интерфэйс бүтээх боломжийг бүрдүүлж байгаа юм .

Веб камерын энэхүү шийдэл нь VR төхөөрөмжүүдтэй харьцуулахад өртөг зардал багатай, хүртээмж өндөр тул хүн бүрийн хэрэглэх боломжтой байдлаар харааны хяналтын системийг өргөн хүрээнд түгээх боломжийг бий болгож байна. Үүний үр дүнд сургууль, музей, гэр бүлийн орчинд хүртэл энэхүү технологийг өндөр зардал гаргалгүйгээр ашиглах боломж бүрдэж байгаа бөгөөд VR төхөөрөмжүүдийн гүн мэдрэмжийг бүрэн орлохгүй ч илүү олон хүнд виртуал орчинтой харилцах боломжийг олгодог юм. Түүнчлэн веб камер ашигласнаар хувийн мэдээллийн аюулгүй байдлыг илүү сайн хамгаалж (төхөөрөмж дээрх боловсруулалт хийдэг), нүдний хөдөлгөөнөөр удирдах боломжтой тул гар удирдлага болон толгойн төхөөрөмж ашиглаж чадахгүй хүмүүсийн хүртээмжийг ч мөн нэмэгдүүлж байна.

Цэвэр веб технологи дээр суурилсан энэхүү шийдэл нь өнөөдөр хэрэглээнд нэвтрэхэд бүрэн бэлэн болсон бөгөөд WebGazer зэрэг сангууд нь ямар ч нэмэлт програм суулгалгүйгээр шууд веб хөтөч дээр бодит хугацаанд ажиллаж байгаагаараа үүнийг баталж байна [8]. Үүний үр дүнд виртуал аялал, цахим сургалт, хүртээмжтэй байдлыг дэмжих аппликэйшн бүтээхэд веб хөгжүүлэгчид харааны хяналтын функцийг хялбархан нэвтрүүлэх боломжтой болсон юм. Цаашдаа энэхүү технологи улам боловсронгуй болж, мэдрэгчтэй болон дуут интерфэйсүүд шиг түгээмэл хэрэглэх технологи болж хөгжих бүрэн боломжтой.

Бидний дээр дурдсан хэрэглээний жишээнүүд нь энэхүү технологийн олон салбар дахь хувьсгах чадварыг харуулж байна. Боловсролын салбарт багш нар сурагчдын анхаарал төвлөрөлтийг шууд тодорхойлж, сурагчид зөвхөн нүдний хөдөлгөөнөөр программ хангамжийг удирдах боломжтой болж байна [1]. Соёлын өв болон аялал жуулчлалын салбарт хэрэглэгчид виртуал орчныг нүдний хөдөлгөөнөөр илүү бодитой судалж, байгууллагууд ч хэрэглэгчдийн анхаарлын хэв маягт үндэслэн контентоо сайжруулах боломжтой болж байна. Хөгжлийн бэрхшээлтэй хүмүүсийн хувьд энэ технологи нь компьютерийн интерфэйсийг ашиглах, бусадтай харилцах боломжийг олгосноор амьдралыг өөрчлөх хэмжээний нөлөө үзүүлж байна.

Мэдээж энэхүү технологид шийдвэрлэх ёстой тодорхой сорилтууд байсаар байна. Гэрэлтүүлэг, нүдний хэлбэр, харааны шил гэх мэт хүчин зүйлсээс шалтгаалсан нарийвчлалын өөрчлөлтүүд нь цаашдын судалгаа, хөгжүүлэлтийн явцад шийдвэрлэх асуудал бөгөөд интерфэйсийн загварыг хүний нүдийг ядраахгүйгээр ээлтэй байдлаар хийх шаардлагатай. Гэсэн ч эдгээр асуудлууд шийдвэрлэх боломжтой бөгөөд цаашдын технологийн хөгжлөөр улам сайжрах нь тодорхой юм.

Эцэст нь веб камерт суурилсан нүдний хөдөлгөөн хянах технологи нь виртуал орчинд хэрэглэгч төвтэй, энгийн бөгөөд илүү хүний байгалийн мэдрэмжид ойр харилцааны шийдлийг бий болгож байна. Энэ нь бидний хамгийн байгалийн заагч төхөөрөмж болох “нүдийг” ашиглан интерфэйсийн дизайн болон интерактив харилцаанд хувьсгал авчирч байгаа бөгөөд энэхүү технологи хөгжихийн хэрээр бид өмнөхөөсөө илүү нарийвчлалтай, шинэлэг шийдлүүдтэй танилцах боломжтой болно. Төгсгөлд нь уг технологи нь бидний виртуал орчинтой харилцах арга барилыг үндсээр нь өөрчилж, хэрэглэгчдэд ямар ч нэмэлт төхөөрөмж шаардалгүйгээр зөвхөн харцаараа дижитал ертөнцийг удирдах боломжийг олгож байна. Энэ нь интерфэйсийн загварчлалын ертөнцийг шинэ түвшинд гаргаж, хэрэглэгчдийн нүдийг шууд утгаар нь “нээж” байгаа юм.

Ном зүй

[1] S. Hutt, A. Y. Wong, A. Papoutsaki, R. S. Baker, J. I. Gold, and C. Mills, “Webcam-based eye tracking to detect mind wandering and comprehension errors,” Behavior Research Methods, vol. 56, no. 1, pp. 1–17, Jan. 2024, doi: 10.3758/s13428-022-02040-x.

[2] J. L. Kröger, O. H.-M. Lutz, and F. Müller, “What Does Your Gaze Reveal About You? On the Privacy Implications of Eye Tracking,” in Privacy and Identity Management. Data for Better Living: AI and Privacy (Privacy and Identity 2019), IFIP AICT, vol. 576, pp. 226–241. Cham, Switzerland: Springer, 2020. doi: 10.1007/978-3-030-42504-3_15.

[3] J. A. Nascimento et al., “Assessing webcam-based eye-tracking during comic reading in the classroom: a feasibility study,” Einstein (São Paulo), vol. 23, eAO0911, 2025. doi: 10.31744/einstein_journal/2025AO0911.

[4] Y. Liu, Z. Lin, H. Wang, Z. Liu, S. Cai, Z. Zheng, Y. Zhou, and X. Zhang, “An eye tracker based on webcam and its preliminary application evaluation in Chinese reading tests,” Biomedical Signal Processing and Control, vol. 71, Art. 103521, Mar. 2022, doi: 10.1016/j.bspc.2022.103521.

[5] WorldViz, “2025 WorldViz VR Budgeting Guidelines for Scientific VR Labs,” WorldViz Blog, Feb. 25, 2025. [Online]. Available:

https://www.worldviz.com/post/worldviz-vr-budgeting-guidelines-for-scientific-vr-labs

[6] Market.us, "Virtual Reality Statistics and Facts (2025)," Market.us Scoop, Feb. 2025. [Online]. Available: https://scoop.market.us/virtual-reality-statistics/

[7] K. Wisiecka, K. Krejtz, I. Krejtz, D. Sromek, A. Cellary, B. Lewandowska, and A. T. Duchowski, "Comparison of Webcam and Remote Eye Tracking," in Proc. 2022 Symposium on Eye Tracking Research and Applications (ETRA '22), Seattle, WA, USA, 8–11 Jun. 2022, Article 32, pp. 1–7, doi: 10.1145/3517031.3529615.

[8] A. Papoutsaki, J. Laskey, and J. Huang, "WebGazer.js: Democratizing Webcam Eye Tracking on the Browser." Brown University, 2016. [Online]. Available: https://webgazer.cs.brown.edu

[9] M. Forte, "When You View the Sistine Chapel, Where Does Your Eye Go? An Art History Class Used Digital Tools to Find Out," Duke University Voices, Dec. 14, 2020. [Online]. Available: https://today.duke.edu/2020/12/when-you-view-sistine-chapel-where-does-your-eye-go-art-history-class-used-digital-tools-find

[10] Apple Inc., "Apple announces new accessibility features, including Eye Tracking," Press Release, Cupertino, CA, USA, May 15, 2024. [Online]. Available: https://www.apple.com/newsroom/2024/05/apple-announces-new-accessibility-features-including-eye-tracking/

[11] Y. Lei, S. He, M. Khamis, and J. Ye, "An end-to-end review of gaze estimation and its interactive applications on handheld mobile devices," ACM Computing Surveys, vol. 56, no. 2, Art. 34, pp. 34:1–34:38, Feb. 2024, doi: 10.1145/3606947.

[12] A. Papoutsaki, "*Democratizing Eye Tracking*," Ph.D. dissertation, Dept. Comput. Sci., Brown Univ., Providence, RI, USA, May 2018.

[13] T. Blaze et al., "face-landmarks-detection: Accurate facial geometry and iris tracking in TensorFlow.js," GitHub repository, 2025. [Online] https://github.com/tensorflow/tfjs-models/tree/master/face-landmarks-detection

[14] Google MediaPipe, "Facemesh eye-aspect-ratio blink heuristic," GitHub Issue #2313, 2024. [Online].

[15] A. Kedia, "Blink-detection with TensorFlow.js face-landmarks," GitHub, 2025. [Online] https://github.com/theankurkedia/blink-detection

Монтгомери үржүүлэгчийн дизайныг FPGA дээр хэрэгжүүлэх нь

Болдбаярын Намуунаа
Электроник, холбооны инженерчлэлийн тэнхим
МУИС, МТЭС
Улаанбаатар хот, Монгол улс
namuna.boldbayar@gmail.com

Батсүхийн Алтанцоож
Электроник, холбооны инженерчлэлийн тэнхим
МУИС, МТЭС
Улаанбаатар хот, Монгол улс
altantsooj.b@num.edu.mn

***Хураангуй*—Монтгомери алгоритм нь криптографын хэрэглээнд өргөн ашиглагддаг хурдан модуль үржүүлэлтийн арга юм. Энэхүү өгүүлэл нь Монтгомери алгоритмын дижит-цуваа хэрэгжилтийг том бүхэл тоон дээр судалсан. Carry-save нэмэгч дээр суурилсан үр ашигтай дижит-цуваа Монтгомери модуль үржүүлэгчийн архитектурын загвар гаргаж, түүний талбай эзлэлтийг тооцоолсон. Энэхүү архитектур нь pipelined carry-select нэмэгчийг ашиглан (1) модуль үржүүлэлтийн carry-save векторуудыг нэгтгэн нэмэх, (2) хэрэв шаардлагатай бол модулиар хасах хэмээх эцсийн хоёр даалгаврыг гүйцэтгэнэ. Мөн дурын δ дижит хэмжээ ба θ модульд хэрэгжүүлэх боломжтой.Монтгомери алгоритмд хэрэглэгдэх $\psi = -\theta^{-1} \bmod 2^{\delta}$ урьдчилсан тооцооллын битүүдийн логик томьёог $\delta \leq 8$ үед танилцуулсан. Эцэст нь Virtex 7 FPGA дээр хэрэгжүүлэн үр дүнг харуулсан.**

***Түлхүүр үгс*—Montgomery modular multiplication, Montgomery reduction, carry-save addition, carry-select addition**

I. Удиртгал

Өнөө үед өгөгдлийн харилцаа болон интернэт үйлчилгээ хурдацтай хөгжихийн хэрээр мэдээлэл дамжуулах үйл явц дахь нууцлал, аюулгүй байдлыг хамгаалах үндсэн шаардлага өндөрсөж байна [1]. Ихэнх системүүд ийм төрлийн аюулгүй байдлын үйлчилгээг криптографийн аргуудад тулгуурлан хангадаг. Жишээлбэл, Rivest, Shamir, Adelman (RSA) зэрэг түгээмэл ашиглагддаг алгоритмууд нь олон нийтийн түлхүүр (public key) үүсгэх аргаар криптографийн үйлдлүүдийг гүйцэтгэдэг. RSA алгоритм нь хоёр түлхүүр (public/private) ашигладаг асимметрик төрлийн криптограф бөгөөд нийтийн түлхүүрийг нийтээр ашиглах боломжтой бол хувийн түлхүүрийг зөвхөн цөөн хэрэглэгч эзэмшиж, хамгаалдаг [2], [3]. RSA алгоритм нь өргөн цар хүрээтэй модуль үржүүлэлтийг олон удаа гүйцэтгэх шаардлагатай байдаг. Аюулгүй байдлын түвшнийг хангах үүднээс түлхүүрийн уртыг ихэвчлэн 1024 битээр сонгодог тул томоохон битийн хэмжээтэй өгөгдлийг өгөгдлийн түвшинд шууд боловсруулахад заримдаа техник хангамжийн түвшний боловсруулалт, дэмжлэг нэн шаардлагатай болдог [4]. Иймээс бага эрчим хүч зарцуулж, өндөр гүйцэтгэл, хурд, бага хоцролттой нөхцөлд том бүхэл тоонуудын модуль үйлдлийг шуурхай гүйцэтгэх үр ашигтай архитектур зайлшгүй хэрэгтэй болдог. Энэ хэрэгцээнд Монтгомери модуль үржүүлэлтийн алгоритм нь өндөр хурд, бага хоцролтын ажиллагааг санал болгодог. Уламжлалт модуль үржүүлэлт нь хуваах үйлдэл шаардаг бөгөөд энэ нь өртөг өндөр, гүйцэтгэлийн хувьд удаан байдаг. Монтгомери алгоритм нь өндөр зардал бүхий хуваагчийг үр дүнтэй үржүүлэлт, бит шилжүүлэлтийн үйлдлүүдээр солин өндөр хурдтай модуль үржүүлэлт хийдэг арга юм.

II. Судлагдсан байдал

Техник хангамж (hardware) - ийн түвшинд үр ашигтай хэрэгжилтийг үгээр (word-wise) болон битээр (bit-wise) хэрэгжүүлэх гэсэн үндсэн 2 ангилалд хуваадаг. Өнөө үеийн FPGA (Field-Programmable Gate Array) дээр тусгай зориулалтын үржүүлэгчүүдийг үгээр (bit-wise) хэрэгжүүлэх загварт ашигладаг ба энэ нь уламжлалт загвартай харьцуулахад илүү хурдтай. Операндуудыг хэд хэдэн хэсэгт хуваж тус бүрийг тусгай үржүүлэгчүүдээр үржүүлсэнээр бага хугацаанд гүйцэтгэх шаардлагатай хэрэглээнд илүү үр дүнтэй болгосон. Судлаачид 64×64 битийн процессортой үржүүлэгчийн санааг талбайн хэмжээ эзлэлтийн хувьд үр ашигтай хэрэгжилтийг санал болгож байв [5]. 2013 онд Barreto–Naehrig (BN) муруй дээр хэрэгжих модуль үржүүлэгчийн концепцийг танилцуулсан ба тусгай анхны тоог ашиглан FPGA-ийн Digital Signal Processor (DSP) блокт тохируулан хуваах үйлдлийг хэрэгжүүлж байсан [6]. Үүнээс гадна Монтгомери алгоритмыг өндөр хурдтай, хямд өртөгтэй байдлаар хэрэгжүүлэхийн тулд carry-save adder (CSA) хэлбэрийг ашигласан [7].

Судлаачид битээр ажилладаг өндөр хурдтай Монтгомери үржүүлэгчийн архитектур боловсруулсан ба энэ нь хоёртын (binary) үржүүлэлтийг high-radix бутархай үржүүлэлттэй хослуулан гүйцэтгэдэг байв [8]. Цуваа 0 бит бүхий үржүүлэлтүүдийг нэг клок циклд гүйцэтгэх боломжтой болсон. Зарим судалгааны ажилд 512 битийн операнд нэмэлт болон 256 битийн үржүүлэлтийг 64 битийн carry chain болон 188 MHz давтамжтай soft-core multiplier ашиглан нэг клок циклд гүйцэтгэжээ [9].

FPGA дээр IP core ашиглан 512 ба 256 битийн үржүүлэлтийг каскад хэлбэрээр гүйцэтгэх замаар 50%-иар илүү үр дүнтэй хэрэгжүүлэлтүүдийг танилцуулсан байна. Зарим судлаачид ECC-д зориулсан MMM архитектурын хэрэгжилтийг хялбарчлах зорилгоор кодын алдааны хамгаалалтыг ашигласан бөгөөд бүрэн 256 битийн бүтэн тооны үржүүлэлтийг 16 битийн модуль дээр каскад хэлбэрээр хэрэгжүүлжээ [10].

Karatsuba алгоритмыг ашиглан операндуудыг 2 хэсэгт хуваж, үржүүлэгчийн зай эзлэлтийг бууруулж, клок циклийн тоог багасгасан схемүүдийг загварчилсан [11], [12]. Зарим судлаачид зэрэгцээ холбогдсон модуль үржүүлэгч (parallel interleaved modular multiplier) - ийн дизайн загварчилсан. Уг архитектурын дагуу операнд нь алгоритмд өгөгдсөн даалгаврыг дөрвөн зэрэгцээ боловсруулалтын элементийн тусламжтай гүйцэтгэдэг.

III. Монтгомери үржүүлэлт

a,b хоёр бүхэл тооны монтгомери үржвэрийг (1)-ээр тодорхойлно.

$$r = ab2^{-\delta} \bmod \theta. \tag{1}$$

Энэ үйлдлийг урьдчилсан тооцоолол $\theta^{-1} \bmod 2^{-\delta}$ ашиглан хуваах үйлдэл хийхгүйгээр гүйцэтгэж болно.

A. Монтгомери талбар

Монтгомери үржүүлэлтэд a, b, r утгууд нь $\tilde{a}, \tilde{b}, \tilde{r}$ гэсэн хэд хэдэн бүхэл тооны үлдэгдлийн төлөөллүүд байдаг. Үүнийг илэрхийлбэл:

$$a = \tilde{a}2^{\delta} \bmod \theta, \quad b = \tilde{b}2^{\delta} \bmod \theta, \quad r = \tilde{r}2^{\delta} \bmod \theta.$$

Энэхүү тусгай төлөөллийг *Монтгомери талбар* (Montgomery domain) гэж нэрлэдэг. Тухайн нөхцөлд Монтгомери үржүүлэлт нь тэгшитгэл (2) харьцааг хангана:

$$\begin{aligned} r &= ab2^{-\delta} \bmod \theta \\ \tilde{r}2^{\delta} \bmod \theta &= (\tilde{a}2^{\delta} \bmod \theta)(\tilde{b}2^{\delta} \bmod \theta)2^{-\delta} \bmod \theta \\ \tilde{r} &= \tilde{a}\tilde{b} \bmod \theta \end{aligned} \tag{2}$$

Эндээс харахад $r = ab2^{-\delta} \bmod \theta$ нь $\tilde{a}\tilde{b} \bmod \theta$ гэсэн үйлдлийн Монтгомери талбар дахь тооцоолол гэж ойлгож болно.

B. Монтгомери багасгалт

Монтгомери үржүүлэлт $r = ab2^{-\delta} \bmod \theta$-ийг, эхлээд үржүүлэх шат ($u = ab$), дараа нь багасгах шат ($r = u2^{-\delta} \bmod \theta$) гэсэн хоёр үе шатад хуваана. Монтгомери модуль багасгалтыг, өөрөөр хэлбэл $r = u2^{-\delta} \bmod \theta$-ийг (3)-аар хуваах үйлдэлгүйгээр тооцоолж болно:

$$r = u2^{-\delta} \bmod \theta \;\equiv\; \hat{r} \;=\; \frac{u+\theta q}{2^{\delta}} \tag{3}$$

Энд $q = u\psi \bmod 2^{\delta}$ нь урьдчилсан тооцоолол $\psi = -\theta^{-1} \bmod 2^{\delta}$-тэй холбоотой параметр юм. Энэхүү модуль багасгалтын логик үндэслэл нь дараах зарчим дээр суурилдаг.

θ нь сондгой тул $\gcd(\theta, 2^{\delta}) = 1$ байна. Ингэснээр $2^{-\delta} \bmod \theta$ болон $\theta^{-1} \bmod 2^{\delta}$ гэсэн хоёр бүхэл тоо орших ба тэдгээр нь

$$2^{\delta}(2^{-\delta} \bmod \theta) + \theta(\theta^{-1} \bmod 2^{\delta}) = 1$$

гэсэн нөхцлийг хангана. Урьдчилсан тооцоолол нь

$$\psi = -\theta^{-1} \bmod 2^{\delta}$$

Ингэснээр

$$2^{-\delta} \bmod \theta \;=\; \frac{1+\theta\psi}{2^{\delta}}.$$

Улмаар Монтгомери модуль багасгалтыг (4) тэгшитгэлээр тооцоолж болно.

$$\begin{aligned} r \;&= u2^{-\delta} \bmod \theta \equiv \frac{u+u\theta\psi}{2^{\delta}} \\ &\equiv \frac{u+\theta(u\psi \bmod 2^{\delta})}{2^{\delta}} \\ &\equiv \frac{u+\theta q}{2^{\delta}} \end{aligned} \tag{4}$$

Энд $q = u\psi \bmod 2^{\delta}$.

C. Монтгомери үржвэрийн тооцоолол

(1)-д өгсөн багасгалтыг ашигласнаар, Монтгомери модуль үржүүлэлт $r = ab2^{-\delta} \bmod \theta$-г хуваах үйлдэлгүйгээр (5) ойролцоолон тооцоолж болно:

$$r = ab2^{-\delta} \bmod \theta \;\equiv\; \hat{r} = \frac{ab+\theta q}{2^{\delta}} \tag{5}$$

Энд $q = ab\psi \bmod 2^{\delta}$ нь урьдчилсан тооцоолол $\psi = -\theta^{-1} \bmod 2^{\delta}$-оос хамаардаг параметр юм.

Хэрэв $a < a_{\sup}$ болон $b < 2^{\delta}$ байвал, бүхэл тоо δ-ийн хувьд Монтгомери үржүүлэлтийг тооцоолбол:

$$r = ab2^{-\delta} \bmod \theta \;\equiv\; \hat{r} = \frac{ab+\theta q}{2^{\delta}} \;<\; a_{\sup} + \theta \tag{6}$$

Энд мөн $q = ab\psi \bmod 2^{\delta}$, $\psi = -\theta^{-1} \bmod 2^{\delta}$. $\hat{r}$ нь зөвхөн $a_{\sup}$-ээс бага байх ёстой тул зарим тохиолдолд $\hat{r}$-ээс θ-г нэг удаа хасах шаардлагатай байж болно. Ийм үед:

$$r = ab2^{-\delta} \bmod \theta \;\equiv\; \hat{r}' = \hat{r} - \varepsilon \cdot \theta < a_{\sup} \tag{7}$$

энд ε нь (8) байна:

$$\varepsilon = \begin{cases} 1, & \text{if } \hat{r} \geq \theta \\ 0, & \text{if } \hat{r} < \theta \end{cases} \tag{8}$$

Бодит байдал дээр a, b болон модуль θ нь $n \geq \lceil \log_2 \theta \rceil$ битээр илэрхийлдэг. Мөн дээд хязгаар $a_{\sup}$ нь θ-аас багагүй байх ёстой.

$$a, b, \theta < 2^{n}, \quad a < a_{\sup} \in [0, 2^{n}]$$

Ийм нөхцөлд бүхэл тоо n-ийн хувьд Монтгомери үржүүлэлт нь (9) хэлбэртэй байна:

$$r = ab2^{-n} \bmod \theta \;\equiv\; \hat{r} = \frac{ab+\theta q}{2^{n}} < a_{\sup} + \theta \tag{9}$$

Энд $q = ab\psi \bmod 2^n$, $\psi = -\theta^{-1} \bmod 2^n$ байна.

Эцсийн үр дүн $\hat{r}$-г $a_{\sup}$-аас доош буулгахын тулд θ-г хасах шаардлагатай нөхцөл үүсч болох бөгөөд (10) илэрхийлнэ:

$$r = ab2^{-n} \bmod \theta \equiv \hat{r}' = \hat{r} - \varepsilon \cdot \theta < a_{\sup} \tag{10}$$

Энд ε мөн (8) хэлбэртэй байна.

D. Практик дээрх хоёр тохиолдол

1. $a_{\sup} = 2^n$
2. $0 < a_{\sup} = 2^n$

1-р тохиолдол: Бүрэн арифметик: Энэ тохиолдолд (9) тэгшитгэлийг хувиргавал:

$$r = ab2^{-n} \bmod \theta \equiv \hat{r} = \frac{ab + \theta q}{2^n} < 2^n$$

Энд $q = ab\psi \bmod 2^n$ ба $\psi = -\theta^{-1} \bmod 2^n$. Эцсийн багасгалтын дараа (10) үр дүн $\hat{r}'$ нь Монтгомери үржүүлэлттэй тэнцүү байна:

$$r = ab2^{-n} \bmod \theta \equiv \begin{cases} \hat{r} - \theta, & \text{if } \hat{r} \geq \theta \\ \hat{r}, & \text{if } \hat{r} < \theta \end{cases} \tag{11}$$

2-р тохиолдол: Дутуу арифметик (incomplete arithmetic): Энэ тохиолдолд (9) тэгшитгэлийг хувиргавал:

$$r = ab2^{-n} \bmod \theta \equiv \hat{r} = \frac{ab + \theta q}{2^n} < 2^n + \theta \tag{12}$$

Мөн $q = ab\psi \bmod 2^n$, $\psi = -\theta^{-1} \bmod 2^n$. (10) багасгалтын дараа $\hat{r}'$ нь Монтгомери үржүүлэлтийн утгатай тэнцүү боловч түүнээс их байж болох бөгөөд:

$$r = ab2^{-n} \bmod \theta \equiv \begin{cases} \hat{r} - \theta, & \text{if } \hat{r} \geq 2^n \\ \hat{r}, & \text{if } \hat{r} < 2^n \end{cases} \tag{13}$$

Эцсийн үр дүн $\hat{r}' < 2^n$ болон $\hat{r}' < \theta$ гэсэн нөхцөл заавал биелэх албагүй. Үнэн хэрэгтээ $\hat{r}'$ нь зөвхөн n битийн хэмжээтэй, Монтгомери үржүүлэлтийн утгатай тэнцүү утга юм. Дутуу арифметикийн зорилго нь оролт болон үр дүнг $[0, 2^n)$ мужид барихад оршдог.

IV. Зэрэгцээ Монтгомери үржүүлэлт (Interleaved Modular Multiplication)

Алгоритм 1 нь Монтгомери үржүүлэлтийг $r = ab2^{-n} \bmod \theta$ хэлбэрээр хэрэгжүүлнэ (9), (10)-р тэгшитгэлийг ашиглана:

Algorithm 1

$v = ab$, энд $b = (b_{n-1}, \ldots, b_0)_2$
$u = \dfrac{v + \theta q}{2^n}$, энд $q = v\psi \bmod 2^n$
if $u \geq a_{\sup}$ **then**
  $r = u - \theta$
else
  $r = u$
end if

Энд $\psi = -\theta^{-1} \bmod 2^n$, $a < a_{\sup} \in [0, 2^n]$. Алгоритм 2 дагуу a нь b-ийн δ биттэй хэсэг бүртэй үржиж, үр дүнг давталт доторх алхмуудад багасна. Үйлдлүүдийг зэрэгцүүлэн хэрэгжүүлж болно.

Algorithm 2

$u = 0$
for $i = 0$ to $n/\delta - 1$ **do**
  $v = u + a\beta$, энд $\beta = (b_{i+\delta-1}, \ldots, b_i)_2$
  $u = \dfrac{v + \theta q}{2^\delta}$, $q = v\psi \bmod 2^\delta$
end for
if $u \geq a_{\sup}$ **then**
  $r = u - \theta$
else
  $r = u$
end if

A. $\psi = -\theta^{-1} \bmod 2^\delta$ утгыг хурдан тооцоолох нь ($\delta \leq 8$)

Урьдчилсан тооцоолол болох ψ нь θ-г ашиглан бодогдож, регистр рүү ачаалж байж модуль үржүүлэлт гүйцэтгэнэ. Иймд дараах Буулийн(Boolean) функцүүдийг хэрэгжүүлнэ. Эдгээр нь θ-гийн битүүд дээр үндэслэн ψ-ийн битүүдийг 0-оос 7 хүртэл олно:

$$\begin{aligned}
\psi_0 &= 1 \\
\psi_1 &= \theta_1, \\
\psi_2 &= \theta_2 \\
\psi_3 &= \theta_1 \oplus \theta_2 \oplus \overline{\theta_3} \\
\psi_4 &= (\theta_1 \oplus \theta_3)(\theta_2 \oplus \theta_3) \oplus \overline{\theta_4} \\
\psi_5 &= (\theta_1 \oplus \theta_3)(\theta_2 \oplus \theta_3)(\theta_3 \oplus \theta_4) \oplus \theta_1 \oplus \theta_4 \oplus \theta_5 \\
\psi_6 &= (\theta_1 \oplus \theta_5)(\theta_2 \oplus \theta_5)(\theta_3 \oplus \theta_5)(\theta_4 \oplus \theta_5) \oplus \\
&\quad (\theta_1 \oplus \theta_4)(\theta_2 \oplus \theta_4)(\theta_3 \oplus \theta_5) \oplus \\
&\quad (\theta_1 \oplus \theta_5)(\theta_2 \oplus \theta_6) \oplus \theta_6 \\
\psi_7 &= (\theta_1 \oplus \theta_6)(\theta_2 \oplus \theta_6)(\theta_3 \oplus \theta_6)(\theta_4 \oplus \theta_6)(\theta_5 \oplus \theta_6) \oplus \\
&\quad (\theta_1 \oplus \theta_6)(\theta_2 \oplus \theta_6)(\theta_3 \oplus \theta_6) \oplus (\theta_1 \oplus \theta_2)(\theta_4 \oplus \theta_6) \oplus \theta_7
\end{aligned} \tag{14}$$

Эдгээр логик функцүүдийг техник хангамжийн түвшинд хялбар хэрэгжүүлж болох бөгөөд $\delta \leq 8$ бүхий дижит-цуваа (digit-serial) үржүүлэгчид ашиглахад тохиромжтой.

Жишээ нь:

Хэрэв $\theta \bmod 2^8 = 185$ гэж үзвэл:

$$\theta = (1\,0\,1\,1\,1\,0\,0\,1)_2 = 185$$

Дээрх Boolean функцүүдийг ашигласнаар:

$$\psi = (0\,1\,1\,1\,0\,1\,1\,1)_2 = 119$$

болно. Баталгаа:

$$\theta \cdot \psi = 185 \cdot 119 \equiv -1 \bmod 2^8$$

V. Санал болгосон модуль үржвэрийн загвар

Алгоритм 3 нь битийн түвшний зэрэгцээ (interleaved) Монтгомери модуль үржүүлэлтийг хэрэгжүүлдэг. Энэ алгоритм дахь операндууд болон модуль θ нь бүгд n битийн урттай.

$$a < a_{\sup} = 2^n, \quad b < 2^n, \quad \theta < 2^n$$

Теорем 1-ийн дагуу, алгоритм 3 дотор ашиглагдах завсрын хувьсагч u нь $u < a_{\sup} + \theta = 2^n + \theta$ хязгаарлалтыг хангана:

Иймд u нь $(n+1)$ битийн хувьсагч байна. Мөн $a < 2^n \quad \beta < 2^\delta, \quad a\beta < 2^{n+\delta}$ харьцааг анхаарах хэрэгтэй:

Үүнээс үндэслэн, алгоритм 3 дээрх v нь:

$$v = a\beta + u < 2^{n+\delta} + 2^{n+1} \tag{15}$$

Тэгвэл v нь $(n+\delta+1)$ битийн хувьсагч байна.

Algorithm 3 Digit-Serial Montgomery Modular Multiplication for a δ-Bit Digit Size

Оролт: n-bit integers a, b, θ
Гаралт: $r = ab2^{-n} \bmod \theta$
Precomputation: $\psi \leftarrow -\theta^{-1} \bmod 2^\delta$
$u[n:0] \leftarrow 0$
for $i = 0$ **to** $n/\delta - 1$ **do**
  3a $\beta \leftarrow b[\,\delta(i+1)-1 : \delta i\,]$,
    $v[n+\delta:0] \leftarrow a \times \beta + u[n:0]$
  3b $q \leftarrow \psi \times v[\,\delta-1:0\,] \bmod 2^\delta$
    $u[n:0] \leftarrow \big(\theta q + v[n+\delta:0]\big)/2^\delta$
end for
if $u[n:0] \geq 2^n$ **then**
  $r \leftarrow u[n:0] - \theta$
else
  $r \leftarrow u[n-1:0]$
end if

Алгоритм 4 нь Алгоритм 3-тай төстэй бөгөөд ялгаа нь u болон v хувьсагчуудыг carry-save хэлбэрээр хадгалдагт оршино. Зураг 1 нь Алгоритм 4-г дүрслэн үзүүлдэг. Доорх гол ажиглалтуудыг Зураг 1-ээс дүгнэж болно. Үүнд:

1) Эхний аккумлятор нь u-г carry-save хэлбэрээр аван, Алгоритм 4-ийн 3a алхамд $v = a\beta + u$ утгыг мөн carry-save хэлбэрээр тооцоолдог.
2) v-г $(n+\delta-1)$ битийн хадгалалт (save vector) болон $(n+\delta)$ битийн дамжуулалт (carry vector) хэлбэрээр илэрхийлж болно. Учир нь томьёо 15-оос үзэхэд $v < 2^{n+\delta} + 2^{n+1}$. Гэхдээ carry vector болон save vector-ийн хамгийн бага зэрэгтэй δ битүүдийг нийлүүлэн $q = v\psi \bmod 2^\delta$ тооцоололд ашиглана.
3) Хоёр дахь аккумлятор нь v-г илэрхийлсэн carry-save векторуудыг авч, Алгоритм 4-ийн 3b алхамд $\theta q + v$-г тооцоолно. Энэ үед хамгийн бага зэрэгтэй δ битүүд нь үргэлж 0 байдаг тул тэдгээрийг устгаж, дараах байдлаар u-г авна:

$$u = \frac{\theta q + v}{2^\delta}$$

Algorithm 4 Digit-Serial Montgomery Modular Multiplication With Carry-Save Addition for a δ-Bit Digit Size

Оролт: n-bit integers a, b, θ
Гаралт: $r = ab2^{-n} \bmod \theta$
Precomputation: $\psi \leftarrow -\theta^{-1} \bmod 2^\delta$
$c^{(u)}[n:0] = 0, \quad s^{(v)}[n-1:0] = 0$
for $i = 0$ **to** $n/\delta - 1$ **do**
  3a $\beta \leftarrow b[\,\delta(i+1)-1 : \delta i\,]$,
    $\big(c^{(v)}[n+\delta-1:\delta], s^{(v)}[n+\delta-2:0]\big) = a \times \beta + c^{(u)}[n:0] + s^{(u)}[n-1:0]$
  3b $q = \psi \times s^{(v)}[\delta-1:0] \bmod 2^\delta$,
    $\big(c^{(u)}[n:0], s^{(u)}[n-1:0]\big) = \theta \times q + c^{(v)}[n+\delta-1:\delta] + s^{(v)}[n+\delta-2:0]\big)/2^\delta$
end for
$t[n:0] = c^{(u)}[n:0] + s^{(u)}[n-1:0]$,
if $t[n:0] \geq 2^n$ **then**
  $r = t[n:0] - \theta$
else
  $r = t[n-1:0]$
end if

A. Гүйцэтгэлийн талбайн хуваарилалт

$$\begin{aligned}&(2n+\delta)T_{FF} + (2n\delta + \delta^2/2 - 3\delta/2)T_{AND}\\ &+ (2n\delta + \delta^2/2 - 5\delta/2 + 1)T_{FA} + (3\delta - 2)T_{HA}\end{aligned} \tag{16}$$

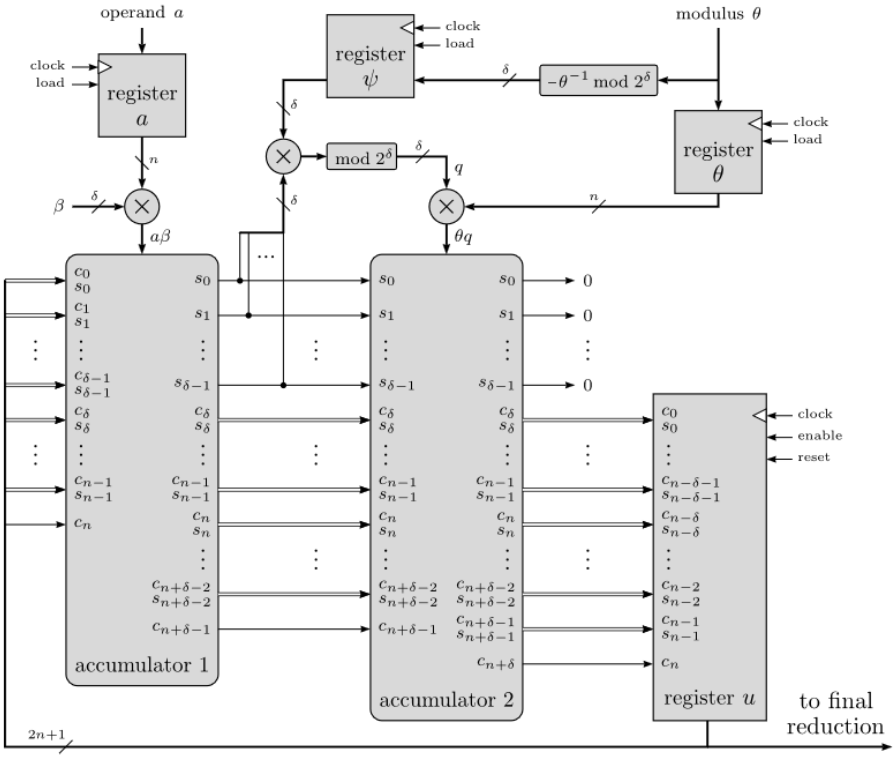
- $2n+\delta$ ширхэг флип-флоп
- $2n\delta + \delta^2/2 - 3\delta/2$ ширхэг AND гейт
- $2n\delta + \delta^2/2 - 5\delta/2 + 1$ ширхэг бүтэн нэмэгч
- $3\delta - 2$ хагас нэмэгч

Эдгээр тооцоололд үржигдэхүүн ба модулийг хадгалах санах ойг тооцоогүй болно. Мөн $\psi = -\theta^{-1} \bmod 2^\delta$-г тооцоолохыг орхигдуулсан ба энэ нь бага нөлөөтэй, хэрэгжилтээс шалтгаалах тул хялбарчлан авч үзсэн.

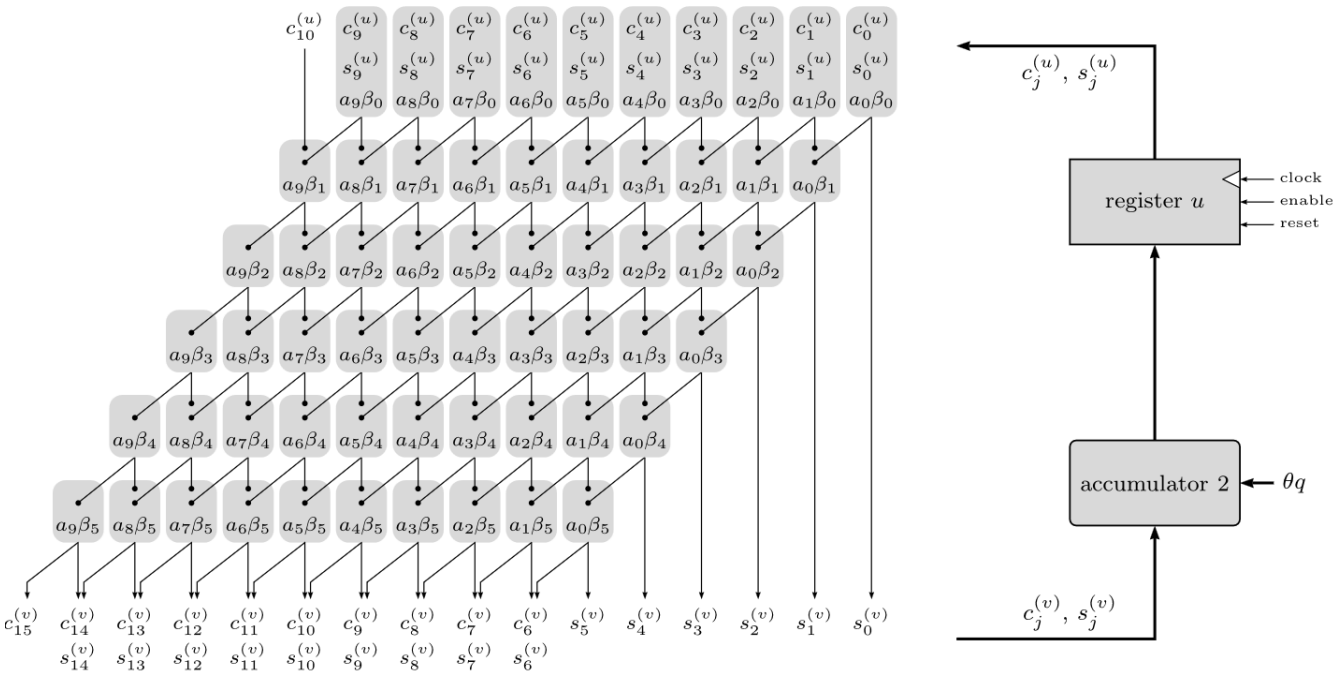
Нарийвчилсан задлал:

1) $c_i^{(u)}, s_i^{(u)}$ -ийг хадгалахад $2n+1$ flip-flop хэрэгтэй. ψ-г хадгалахад δ флип-флоп хэрэгтэй ч $\psi_0 = 1$ тул нэгийг хасаж болно.
2) Эхний аккумлятор $a\beta$-г бодох $n\delta$ ширхэг AND гейт, $n\delta$ ширхэг нэмэгч хэрэгтэй. Үүний $\delta - 2$ нь хагас нэмэгч байдаг.
3) q_i-г тооцоход $\delta(\delta+1)/2$ AND гейт, $(\delta-2)(\delta-1)/2$ ширхэг бүтэн нэмэгч, $\delta - 1$ ширхэг хагас нэмэгч шаардлагатай. Гэхдээ δ AND гейтийг $\psi_0 = 1$ гэдгийг ашиглан хэмнэх боломжтой.
4) Хоёр дахь аккумлятор δ AND гейт, $n(\delta-2) + 2(n+\delta-1)$ нэмэгч хэрэгтэй. Үүнээс $\delta + 2$ нь хагас нэмэгч байна. Мөн дараах давхар битүүд илүүц тул хасаж болно.

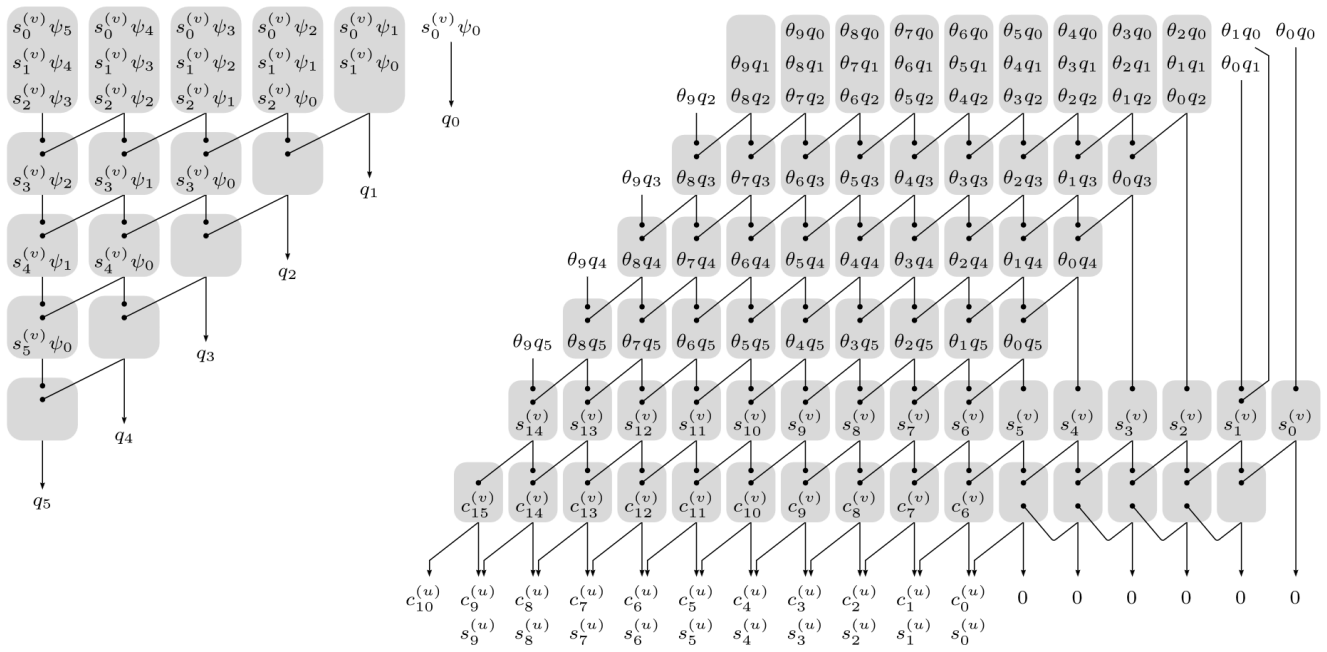
$$s_0^{(v)} + \theta_0 q_0 = s_0^{(v)} + q_0 \Rightarrow 2s_0^{(v)} \text{ (энд } \psi_0 = \theta_0 = 1)$$



Зураг 1. Дижит-цуваа бүхий Монтгомери үржигчийн хэлхээ



Зураг 2. $\delta = 6$ бит дижитийн хэмжээтэй, n = 10 бит операндын 3а) алхам дахь $a\beta$ хэсэгчилсэн үржвэрийн хэлбэр.



Зураг 3. Дижитийн хэмжээ $\delta = 6$ бит, операторын хэмжээ $n = 10$ бит үед 3b) алхам дахь q-ийн тооцоолол болон θq хэлбэр.

B. Хугацааны нарийвчлал

Үржүүлэгчийн critical path-ийн хугацааг тодорхойлбол:

$$\text{Хугацаа} = \begin{cases} 3T_{\text{AND}} + T_{\text{HA}} + (\delta+4)T_{\text{FA}}, & \delta > 2 \\ 3T_{\text{AND}} + T_{\text{HA}} + 4T_{\text{FA}}, & \delta = 2 \end{cases} \quad (17)$$

Энд:

- T_{AND} — AND гейтийн хугацааны хоцролт
- T_{FA} — Бүтэн нэмэгчийн хугацааны хоцролт
- T_{HA} — Хагас нэмэгчийн хугацааны хоцролт

Томоохон хугацааны хоцрогдол дараах байдлаар гардаг:

1) $s_i^{(0)}$-г эхний аккумляторт тооцоолохдоо:
$T_{\text{AND}} + (i+1)T_{\text{FA}}, \quad i < \delta$
2) q_i-г $s_i^{(v)}$-с тооцоолохдоо:
$T_{\text{AND}} + T_{\text{FA}} + T_{\text{HA}}, \quad 2 \le i < \delta$
3) θq_i-г хоёр дахь аккумляторт нэмэхдээ:
$T_{\text{AND}} + (\delta - i + 2)T_{\text{FA}}, \quad 2 \le i < \delta$

C. Эцсийн багасгалт

Алгоритм 4-ийн 3-р алхамд $(n+1)$ битийн дараах нэмэлт хийгдэнэ:

$$t = (c^{(u)}[n]\,c^{(u)}[n-1]\ldots c^{(u)}[0])_2 + (0\,s^{(u)}[n-1]\,s^{(u)}[n-2]\ldots s^{(u)}[0])_2 \quad (18)$$

Гаралт $r = t$ байх ба хэрэв $t[n] = 0$ бол шууд $r = t$, эс бөгөөс дараах хасалтыг гүйцэтгэнэ:

$$r = t - \theta = (t[n]\,t[n-1]\ldots t[0])_2 - (\theta[n-1]\ldots\theta[0])_2 \quad (19)$$

Өгөгдлүүдийг өмнөх алхмуудад carry-save хэлбэрээр хадгалсан нь томоохон carry propagation саатлаас зайлсхийх давуу талтай. Ингэснээр өмнөх алхмуудын critical path хоцролт нь зөвхөн δ-тай шугаман хамааралтай болно. Энэ нь $n \gg 1$ тохиолдолд маш ашигтай.

Харамсалтай нь, эцсийн нэмэлт ба хасалтууд $n+1$ битийн хэмжээтэй тул бүтэн нэмэгч ба borrow propagation шаардаг. Тиймээс энэ нэмэлт, хасалтыг CSA (Carry-Save Adder) ашиглан олон клок циклд pipeline хэлбэрээр гүйцэтгэх шаардлагатай.

1024 битийн жишээ

1024 битийн операндуудыг 168 ба 172 битийн блокуудад хуваах ripple-carry нэмэгч ашиглан гүйцэтгэнэ:

$$t[1024:0] = C[1024:0] + S[1023:0]$$
$$tt[1023:0] = t[1023:0] + \bar{\theta}[1023:0]$$

Энд $\bar{\theta}[1023:0]$ нь сөрөг модулийн утга. Гурван клок циклд нэмэлт, хасалт дараах байдлаар хийнэ:

1) $t[511:0]$ ба carry $mmc1$
2) $t[1024:512]$, $tt[511:0]$, borrow $mmb1$
3) $tt[1023:512]$
4) $r = \begin{cases} tt[1023:0], t[1024] = 1 \\ t[1023:0], t[1024] \neq 1 \end{cases}$

Талбайн хуваарилалт:

Алгоритм 4-ын 3-р алхамд шаардагдах ойролцоо талбай:

$$4nA_{\text{FA}} + 2nA_{\text{REG}} + 3nA_{\text{MUX2}} \quad (20)$$

Энд:

- A_{FA} — Бүтэн нэмэгчийн талбай
- A_{REG} — Флип-флоп эсвэл регистрийн талбай
- A_{MUX2} — 2 ороллтой мультиплексорын талбай

D. Хугацааны хоцрогдол:

Carry-save нэмэгчийн хугацааны хоцрогдлыг pipeline-оор бууруулах хэрэгтэй. 3-р алхмын нийт хугацааны хоцрогдол:

$$\mathcal{P}+2 \tag{21}$$

Энд $\mathcal{P}$ нь pipeline шатны тоо. $(\mathcal{P}+1)$ клок циклд сүүлийн хасалт гүйцэтгэгдэж, $(\mathcal{P}+2)$ клок циклд аль хариуг сонгох мультиплексор ажиллана.

VI. Гүйцэтгэлийн харьцуулалт

Хүснэгт I нь $\delta = 1$ битийн дижит хэмжээтэй зарим Монтгомери үржүүлэгчийн талбайн хуваарилалтыг харуулав. Мөн $\delta = 2, 4, 8$ битийн хувьд санал болгож буй үржүүлэгчийн талбайн хэрэглээг (16) болон (20)-оос авсан тооцооллууд дээр нэмээд регистр a болон θ-г хадгалах санах ойтой хамт нийт талбайг харуулдаг.

1. $\delta = 1$ битийн үржүүлэгчдийн онцлог

$\delta = 1$ битийн үржүүлэгчүүд хурдан нэмэгч шаардаггүй ба операнд болон үр дүнгээ үргэлж carry-save хэлбэрээр хадгалдаг. Эцсийн хасалтыг clever scheme ашиглан хийдэг. Үүний тулд:

$$\hat{r} = \frac{ab+\theta q}{2^{n+2}} \equiv ab\,2^{-(n+2)} \bmod \theta \tag{22}$$

байх илүү том хуваагч ашиглана. Иймд бүх операнд болон үр дүнг 2θ-оор хязгаарлаж, цаашдаа дараалан үржүүлэлт хийх боломжтой болно.

2. $\delta > 1$ үед санал болгосон арга

Санал болгосон $\delta > 1$ битийн үржүүлэгч нь илүү том талбай ашигладаг ч нийт гүйцэтгэл, нийт цаг хугацааны хувьд илүү давуу талтай. Гүйцэтгэлийн хугацааг (17) ба (21)-аас:

$$\frac{n}{\delta} + \mathcal{P}_\delta + 2 \approx \frac{n}{\delta} \tag{23}$$

гэж тооцно. Иймд нийт гүйцэтгэлийн хугацаа:

$$(\delta+4)\frac{n}{\delta}T_{\mathrm{FA}} + \frac{n}{\delta}(3T_{\mathrm{AND}} + T_{\mathrm{HA}}) \tag{24}$$

байдаг. $\mathcal{P}_\delta$ бол pipeline шатны тоо ба ихэвчлэн бага тоо байна.

3. FPGA дээрх хэрэгжилтийн үр дүн

Хүснэгт III-т Xilinx Vivado дээр хэрэгжүүлсэн бодит үр дүнгүүдийн харьцуулалтыг харуулав:

- $\delta = 2$ үед хамгийн бага area-time product гарсан.
- $\delta = 4$ нь ч мөн [16], [17]-аас бага area-time product-той.
- $\delta = 2$ загвар нь флип-флоп багатай мөртлөө slice багатай (учир нь [16], [17] олон комбинацийн логик агуулдаг).
- FPGA slice нь dedicated carry chain-тай тул хурдан нэмэлт гүйцэтгэнэ.

4. Бусад дижит хэмжээтэй үр дүн ба дүгнэлт

Хэрэв $\delta = 8$ бол area-time product бага зэрэг ихэсдэг. Гэвч нийт гүйцэтгэл сайжирсан хэвээр байна. Хуучин Virtex 2 төхөөрөмжүүд дээр ихэнх загварууд туршигдсан байдаг. Бусад загварууд зөвхөн тодорхой модуль утга эсвэл дижит хэмжээнд зориулсан байдаг тул энэхүү нийтлэг хэрэгжүүлэлттэй харьцуулахад сул байна.

5. Аюулгүй байдал

Санал болгосон үржүүлэгч нь мөн:

- Side-channel халдлагад тэсвэртэй.
- ECC, RSA зэрэг криптографийн практикт шууд ашиглах боломжтой.
- Хамгийн чухал нь нийт гүйцэтгэлийн хугацаа бага.

VII. Дүгнэлт

Энэхүү өгүүлэлд Монтгомери алгоритм болон түүний дижит-цуваа тооцооллыг нарийвчлан судалсан. Урьдчилсан тооцоолол болох $\psi = -\theta^{-1} \bmod 2^\delta$ нь (14)-д өгөгдсөн бөгөөд энэ нь $\delta \leq 8$ үед хүчинтэй.

Монтгомери алгоритмд зориулсан ерөнхий дижит-цуваа архитектур боловсруулсан. Энэ архитектурыг дурын θ модуль болон δ хэмжээний дижитэд хэрэгжүүлэх боломжтой. Үржүүлэгчийн критик замын саатлыг бууруулахын тулд хуримтлагдсан утгыг carry-save хэлбэрээр хадгалсан. Гэвч эцсийн үр дүнг гаргахын тулд нэмэлт ба хасалт зайлшгүй шаардлагатай байдаг.

Эдгээр нэмэлт болон хасалтыг CSA (Carry-Save Adder)-аар хэрхэн гүйцэтгэдгийг дэлгэрэнгүй тайлбарласан. FPGA хэрэгжилтэд CSA нь тун тохиромжтой, учир нь FPGA нь томоохон нэмэлтийг хурдан гүйцэтгэх dedicated carry chain-тэй байдаг. CSA стратеги нь том нэмэлтүүдийг жижиг хэсэгт хуваан, хэрэгтэй үедээ хурдан нэмэх боломжийг олгодог.

Санал болгож буй дижит-цуваа үржүүлэгчийн нийт гүйцэтгэлийн хугацаа нь уламжлалт бит-цуваа үржүүлэгчээс хамаагүй бага. Гэхдээ энэ сайжруулалт нь илүү их талбай ашигладаг. FPGA хэрэгжилтийн үр дүнгээс харахад хамгийн бага area–time product нь $\delta = 2$ үед гарсан.

Санал болгосон $\delta = 2$ архитектур нь уламжлалт бит-цуваа үржүүлэлтээс area–time үзүүлэлтээр илүү сайн байна. Мөн $\delta = 4$ үед ч уламжлалттай ижил түвшинд гүйцэтгэлтэй байсан.

Хүснэгт I
Ялгаатай хэрэгжүүлэлтүүдийн талбайн тооцоолол

Multiplier	Digit size δ	Area Requirement without Fast Adders	Additional Area for Fast Adders
[16] (5 to 2 adders)	1	$3n\ A_{\mathrm{FA}}+7n\ A_{\mathrm{REG}}+3n\ A_{\mathrm{AND}}$	—
[16] (4 to 2 adders)	1	$2n\ A_{\mathrm{FA}}+9n\ A_{\mathrm{REG}}+2n\ A_{\mathrm{MUX}_4}$	—
[17]	1	$2n\ A_{\mathrm{FA}}+9n\ A_{\mathrm{REG}}+3n\ A_{\mathrm{AND}}+n\ A_{\mathrm{MUX}_2}+n\ A_{\mathrm{MUX}_3}$	—
Proposed	2	$(4n-2)\ A_{\mathrm{FA}}+(4n+2)\ A_{\mathrm{REG}}+(4n-1)\ A_{\mathrm{AND}}+4\ A_{\mathrm{HA}}$	$4n\ A_{\mathrm{FA}}+2n\ A_{\mathrm{REG}}+3n\ A_{\mathrm{MUX}_2}$
Proposed	4	$(8n-1)\ A_{\mathrm{FA}}+(4n+4)\ A_{\mathrm{REG}}+(8n+2)\ A_{\mathrm{AND}}+10\ A_{\mathrm{HA}}$	$4n\ A_{\mathrm{FA}}+2n\ A_{\mathrm{REG}}+3n\ A_{\mathrm{MUX}_2}$
Proposed	8	$(16n+13)\ A_{\mathrm{FA}}+(4n+8)\ A_{\mathrm{REG}}+(16n+20)\ A_{\mathrm{AND}}+22\ A_{\mathrm{HA}}$	$4n\ A_{\mathrm{FA}}+2n\ A_{\mathrm{REG}}+3n\ A_{\mathrm{MUX}_2}$

Хүснэгт II
Ялгаатай хэрэгжүүлэлтүүдийн хугацааны тооцоолол

Multiplier	Digit size δ	Critical Path Delay	Cycles	Total Computation Time
[16] (5 to 2 adders)	1	$3T_{\mathrm{FA}}+T_{\mathrm{AND}}+2T_{\mathrm{XOR}}$	$n+1$	$(3n+3)T_{\mathrm{FA}}+(n+1)(T_{\mathrm{AND}}+2T_{\mathrm{XOR}})$
[16] (4 to 2 adders)	1	$2T_{\mathrm{FA}}+T_{\mathrm{AND}}+2T_{\mathrm{XOR}}+T_{\mathrm{MUX}_4}$	$n+2$	$(2n+4)T_{\mathrm{FA}}+(n+2)(T_{\mathrm{AND}}+2T_{\mathrm{XOR}}+T_{\mathrm{MUX}_4})$
[17]	1	$2T_{\mathrm{FA}}+T_{\mathrm{AND}}+T_{\mathrm{MUX}_2}$	$n+4$	$(2n+8)T_{\mathrm{FA}}+(n+4)(T_{\mathrm{AND}}+T_{\mathrm{MUX}_2})$
Proposed†	2	$4T_{\mathrm{FA}}+3T_{\mathrm{AND}}+T_{\mathrm{HA}}$	$\frac{n}{2}+P_2+2$	$(2n+4P_2+8)T_{\mathrm{FA}}+\left(\frac{n}{2}+P_2+2\right)(3T_{\mathrm{AND}}+T_{\mathrm{HA}})$

Хүснэгт III
FPGA хэрэгжүүлэлтийн талбай болон хугацааны харьцуулалт, $n=1024$

Multiplier	Digit size δ	slice	LUT	FF	LUT%	FF%	Multiplier	Adder	Period (ns)	Time (μs)	Area× time (slice×ms)
[16]	1	2981	9297	9564	–	–	1025	1025	2.26	2.32	6.91
	1	3078	9277	13570	–	–	1028	1028	2.04	2.10	6.45
[17]	2	2504	9304	7492	23%	42%	512	530	2.23	1.18	2.96
Proposed	4	5836	22625	7533	14%	41%	256	264	3.83	1.01	5.90
	8	9677	34904	7427	13%	41%	128	132	6.88	0.91	8.79

Ашигласан ном

[1] Kuang, Shiann-Rong, et al. "Energy-efficient high throughput Montgomery modular multipliers for RSA cryptosystems."Very Large Scale Integration (VLSI) Systems, IEEE Transactions on 21.11 (2013): 1999-2009

[2] Knezevic, Miroslav, Frederik Vercauteren, and Ingrid Verbauwhede. "Faster interleaved modular multiplication based on Barrett and Montgomery reduction methods. "Computers, IEEE Transactions on 59.12 (2010): 1715-1721.

[3] I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[4] R. L. Rivest, A. Shamir, and L. Adleman, “A method for obtaining digital signature and public-key cryptosystems,” Commun. ACM, vol. 21, no. 2, pp. 120–126, Feb. 1978.

[5] Mondal, A., Ghosh, S., Das, A., Chowdhury, D. R. (2012).Efficient FPGA implementation of Montgomery multiplier using DSP blocks. In Progress in VLSI Design and Test (pp. 370-372). Springer, Berlin, Heidelberg.

[6] Brinci, R., Khmiri, W., Mbarek, M., Rabaâ, A. B., Bouallegue, A.,Chekir, F. (2013). Efficient Multiplier for pairings over Barreto-Naehrig Curves on Virtex-6 FPGA. IACR Cryptol. ePrint Arch.,2013, 5

[7] Kuang, S. R., Wu, K. Y., Lu, R. Y. (2015). Low-cost high-performance VLSI architecture for Montgomery modular multiplication. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 24(2), 434-443

[8] Rezai, A., Keshavarzi, P. (2014). High-throughput modularmultiplication and exponentiation algorithms using multibit-scan–multibit-shift technique. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 23(9), 1710-1719.

[9] Javeed, K., Wang, X. (2014, July). Efficient Montgomery multiplier for pairing and elliptic curve-based cryptography. In 2014 9th International Symposium on Communication Systems,Networks Digital Sign (CSNDSP) (pp. 255-260). IEEE.

[10] McIvor, C. J., McLoone, M., McCanny, J. V. (2006). Hardware Elliptic Curve Cryptographic Processor Over rm GF (p). IEEE Transactions on Circuits and Systems I: Regular Papers, 53(9), 1946-1957.

[11] Gong, Y., Li, S. (2010, March). High-throughput FPGA implementation of 256-bit Montgomery modular multiplier. In 2010 Second International Workshop on Education Technology and Computer Science (Vol. 3, pp. 173-176). IEEE.

[12] Ghosh, S., Verbauwhede, I., Roychowdhury, D. (2012, May). Core-based architecture to speed up optimal ate pairing on the FPGA platform. In International Conference on Pairing-Based Cryptography (pp. 141-159). Springer, Berlin, Heidelberg.

[13] Javeed, K., Wang,(2017). Low latency flexible FPGA implementation of point multiplication on elliptic curves over GF(p). International Journal of Circuit Theory and Applications,45(2), 214-228.

[14] Ors, S. B., Batina, L., Preneel, B., Vandewalle, J. (2003, June).Hardware implementation of an elliptic curve processor over GF(p). In Proceedings IEEE International Conference on Application-Specific Systems, Architectures, and Processors. ASAP 2003 (pp. 433-443). IEEE.

[15] C. McIvor, M. McLoone, and J. V. McCanny, “Modified Montgomery modular multiplication and RSA exponentiation techniques,” IEE Proc.Comput. Digit. Techn., vol. 151, no. 6, pp. 402–408, Nov. 2004.

[16] M. D. Shieh, J. H. Chen, H. H. Wu, and W. C. Lin, “A new modular exponentiation architecture for efficient design of RSA cryptosystem,” IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 16, no. 9, pp. 1151–1161, Sep. 2008.

[17] C. D. Walter, “Montgomery exponentiation needs no final subtractions,” Electron. Lett., vol. 35, no. 21, pp. 1831–1832, Oct. 1999.

PQC McEliece алгоритмд зориулсан эрэмбэлэх алгоритмын хэрэгжүүлэлт

Баатарсүрэнгийн Энхзаяа
Электроник, Холбооны Инженерчлэлийн тэнхим
Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол Улс
enhzaya4503@gmail.com

Батсүхийн Алтанцоож
Электроник, Холбооны Инженерчлэлийн тэнхим
Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол Улс
altantsooj.b@num.edu.mn

***Хураангуй*—Энэхүү судалгаанд Classic McEliece алгоритмын түлхүүр үүсгэх шатанд шаардагддаг тогтмол хугацааны эрэмбэлэх процессыг хурдасгах шинэ төрлийн техник хангамжийн архитектур бүхий алгоритмыг санал болгож байна. McEliece алгоритм нь өндөр аюулгүй байдалтай боловч түлхүүр үүсгэх хурд бага, олон нийтийн түлхүүрийн хэмжээ том байдаг нь бодит хэрэглээнд саад болдог. Энэхүү ажлын хүрээнд Нэгтгэн эрэмбэлэх (Merge sort) алгоритмыг дөрвөн элементийг зэрэгцээ боловсруулах аргаар сайжруулж, харьцуулагчдыг бүрэн ашиглан, гэдрэг холбоогүй (feedback-less) бүтэц хэмнэлттэй шийдлийг боловсруулсан. Үр дүнгээс харахад санал болгож буй архитектур нь дамжуулах чадвар, нөөцийн хэрэглээ болон гүйцэтгэлийн хувьд өмнөх судалгаануудаас давуу болохыг харуулсан бөгөөд санах ойн хэрэглээг хэмнэж, өндөр давтамжтай ажиллагааг дэмжиж байна. Судалгааны үр дүн нь McEliece алгоритмын хурдасгуур системийг өндөр хүчин чадал бүхий платформ (жишээ нь, өгөгдлийн төвүүд) дээр хэрэгжүүлэх, түүнчлэн NIST-ийн пост-квант криптографын стандартчиллын дараагийн шатанд үнэлэх үндэс суурийг бүрдүүлж чадна.**

***Түлхүүр үгс*—PQC, McEliece, Эрэмбэлэх алгоритм, NTRU.**

I. Удиртгал

McEliece алгоритм нь 40 жилийн турш судлагдсан, аюулгүй байдлын нарийн шинжилгээ хийсэн Goppa код дээр суурилсан алгоритм бөгөөд энэ нь өнөөгийн криптографийн салбарт чухал байр суурь эзэлдэг. 2022 оны 7-р сард McEliece алгоритм нь National Institute of Standards and Technology (NIST)-ийн PQC (Post-Quantum Cryptography) стандартчилалд 4-р шатанд Түлхүүр Хамгаалах Механизм (KEM)-ын финалист болсон [1]. NIST-ийн 3-р шатны бичиг баримтад "Классик McEliece нь олон нийтийн түлхүүрийн хэмжээ нь маш том бөгөөд түлхүүр үүсгэх процесс нь харьцангуй удаан. Энэ нь олон нийтлэг нөхцөлд Классик McEliece-ийг хэрэглэх боломжгүй болгох магадлалтай гэж дурдсан байдаг. Тиймээс, McEliece алгоритм нь түлхүүр үүсгэх хурд болон олон нийтийн түлхүүрийн хэмжээ нь том байгаагаас шалтгаалан шууд стандарчилалд орж чадахгүй боловч, 4-р шатанд судалгаагаа үргэлжлүүлэх болно. Манай судалгааны гол зорилго бол эдгээр асуудлыг шийдвэрлэх явдал юм.

Сертификацийн байгууллага эсвэл олон нийтийн түлхүүрийн дэд бүтцийн хувьд дамжуулах чадвар өндөр, олон тооны хослол түлхүүр үүсгэх үйл явц шаардлагатай. Тиймээс манай судалгаанд дамжуулах чадвар, санах ойн үр ашиг (дамжуулах чадвар/талбай) нь хамгийн чухал юм. Одоогийн гүйцэтгэлийн шинжилгээнд үндэслэн, түлхүүр үүсгэх нь өгөгдлийг нууцлах болон нууцлалыг тайлахаас хоёр дахин их цикл зарцуулдаг гэж дурдсан [2]. Тиймээс, түлхүүр үүсгэх үйл явц нь McEliece алгоритмын гол асуудал болж, түүний өргөн хэрэглээг хязгаарлаж байна.

Өнөөгийн байдлаар, олон төрлийн McEliece-ийн алгоритм нь олон төрлийн программ хангамжийн платформ [2]–[5] эсвэл техник хангамжийн платформ дээр [6]–[9] хэрэгжүүлэгдсэн байна. [6] нь McEliece-ийг техник хангамж дээр хамгийн анх хэрэгжүүлсэн бөгөөд хэд хэдэн модуль бүхий дизайн нь дараа нь NIST-ийн 3-р шатны McEliece-ийн тодорхойлолтуудтай нийцээгүй байсан, үүнд жижгэрүүлэх боломжгүй олон гишүүнтүүдийг үүсгэх, псевдо-санамсаргүй тоо үүсгэгч (PRNG), нууц түлхүүрийн зохион байгуулалт зэрэг байсан. [7] ажилд техник хангамжийн түлхүүр үүсгэгчийн дизайныг танилцуулсан бөгөөд үүнд цуваа эрэмбэлэх, adaptive FFT (The Fast Fourier Transform), Карацуба алгоритмд суурилсан олон гишүүнтүүдийг үржүүлэх модуль гэх мэт багтсан. Мөн, том хэмжээний $GF(2)$-ийн Гаусс устгах(Gauss elimination) болон $GF(2^m)$-ийн матрицын эргүүлэлт нь [10]-д суурилсан дизайныг ашигласан. [8] нь энэ дизайныг өргөжүүлэн схемүүдийг хэрэгжүүлэх (нууцлах, тайлах) боломжийг олгосон. Мэдээллийг нууцлахад 6000 клок, тайлахад 14000 клок шаардлагатай байв. Үүний эсрэгээр, түлхүүр үүсгэхийн тулд нэг оролдлого тутамд 966000 клок шаардлагатай бөгөөд амжилттай түлхүүр үүсгэхэд дунджаар 3.5 оролдлого шаардлагатай байдаг. [11] судалгааны ажил нь амжилтгүй оролдлогын хугацааг багасгах эрт эргэж буцах дизайныг танилцуулсан. [9] нь өндөр түвшний синтезийн аргыг ашиглан McEliece-ийг техник хангамж-программ хангамж хослуулсан дизайнаар хурдасгахыг оролдсон. Гэвч энэ аргачлалын хязгаарлалтуудын улмаас гүйцэтгэл нь бүтэн техник хангамж хэрэгжүүлэлт [7] -тэй харьцуулахад дутаг-

далтай байлаа. Түлхүүр үүсгэх модулийн хувьд [8] -д $GF(2)$-ийн Гаусс устгах нь түлхүүр үүсгэхэд зарцуулагдсан клокын 76.4%-ийг эзэлдэг бол тогтмол хугацааны эрэмбэлэх клок 15.2%-ийг эзэлдэг. Энэ нь амжилттай ажиллах тохиолдолд юм. Тиймээс тогтмол хугацааны эрэмбэлэх алгоритм нь түлхүүр үүсгэхэд хурдасгах гол үүргийг гүйцэтгэдэг.

Энэхүү судалгаанд, бид шинэ төрлийн тогтмол хугацааны эрэмбэлэх алгоритмыг танилцуулж байна, энэ нь өмнөх судалгаануудад авч үзсэнээс илт цөөн харьцуулах үйлдлийг ашиглан, түүнд зориулан хэд хэдэн янзын түвшний схемчилсэн дизайныг боловсруулсан.

II. Эрэмбэлэх алгоритмууд

A. Одоогийн тогтмол хугацаанд эрэмбэлэгчийн техник хангамжийн хэрэгжүүлэлт

Эрэмбэлэх функцууд нь олон PQC схемд шаардлагатай байдаг, үүнд NTRU [12] болон McEliece [13] орно, учир нь эдгээр нь өгөгдсөн санамсаргүй массивын бүрэн эрэмбийг тогтмол хугацаанд олж авахад тусалдаг. Тиймээс эрэмбэлэгчийн дизайн нь PQC схемүүдэд санамсаргүй алдааны байрлалыг оруулахад зайлшгүй шаардлагатай хэсэг болдог. Мөн криптографийн техник хангамж өндөр хурдны үзүүлэлттэй байх шаардлагатайгаас гадна тогтмол хугацааны шинж чанарыг анхаарч үзэх хэрэгтэй. Учир нь хугацааны халдлагуудын (time attack) эрсдлийг багасгахад тусалдаг.

[8] нь нэг л харьцуулалт ашиглан скаляр Нэгтгэн эрэмбэлэгчийг танилцуулсан. Гэвч энэ нь гүйцэтгэлийн хурданд хангалттай сайн байгаагүй. [14] нь FIFO (First in Firt out) санах ойд суурилсан параллел Нэгтгэн эрэмбэлэгчийг санал болгосон. Энэ дизайнд, $log_2(l)$ шат нь l урттай массивыг эрэмблэхийн тулд уялдуулан холбоно, шат бүрт хоёр харьцуулалт байдаг бөгөөд хамгийн сүүлийн шат нь ялгаатай. Тус бүрийн харьцуулалтын оролт нь хоёр эрэмбэлэгдсэн жагсаалт бөгөөд гаралтын нэг жагсаалт нь хоёр дахин урт байна. Мөн тус бүрийн оролтын жагсаалт нь тусдаа санах ойд хадгалагддаг. Бүх хадгалах орон зайд хүчинтэй өгөгдөл агуулаагүй байдаг тул санах ойд хэрэгцээгүй өгөгдөл ихээр үүсгэдэг. [14] нь l элементийг эрэмблэхийн тулд ойролцоогоор $3l$ клок шаарддаг. Харьцуулалтын тоо нь $2 \times log_2(l) - 1$ гэж тооцогддог. Мөн элементийн хувьд санах ойн шаардлага нь $4(1+2+4+\cdots+2log_2 l - 2) + 2(2log_2 l - 1) \approx 3l$ гэж тооцогддог. [15] нь хоёр дахин өөрчлөгдсөн Bitonic эрэмбэлэгчид суурилсан тогтмол хугацааны дизайныг хэрэгжүүлсэн бөгөөд энэ нь гэдрэг холбооны механизмд үндэслэсэн. Энэ дизайн нь ASIC хэрэгжилтэд тохиромжтой боловч FPGA(Field-Programmable Gate Array) дээр урт үргэлжилсэн замтай (critical path) байсан тул хэрэгжүүлэх боломжгүй.

B. Уламжлалт эрэмбэлэгчийн дизайнууд

[16]–[21] нь Нэгтгэх-мод хэлбэрийн [21] дизайн нь тогтмол хугацааны биш байсан. Эрэмбэлэгчийн дизайнд, FIFO болон санах ойд хандах байдлууд нь мөн тогтмол хугацааны гүйцэтгэлтэй байдаггүй [17]. [20]-ын санаанаас сэдэвлэн тогтмол хугацаанд эрэмбэлэгчийг танилцуулсан бөгөөд энэ нь тогтмол санах ойд хандах ба эрэмбэлэгчийн хэрэгслийн шаардлага багатай юм.

C. Эрэмбэлэх алгоритмуудын харьцуулалт

Эрэмбэлэх алгоритмууд нь гүйцэтгэлийн шинж чанараараа ялгардаг бөгөөд тэдгээрийн зарим нь тодорхой нөхцөлд илүү давуу талтайгаар хэрэглэгддэг. Тухайлбал, Merge Sort, Selection Sort, Radix Sort, болон Count Sort алгоритмууд нь өгөгдлийн бүтэц, хэмжээ ямар байхаас үл хамааран тогтвортой ажиллагаатай байх буюу тогтмол хугацаанд гүйцэтгэх боломжтой гэдгээрээ онцлог юм. Энэ нь алгоритмыг их хэмжээний өгөгдөл дээр үр дүнтэй ажиллуулах нөхцөлийг бүрдүүлдэг.

Нөгөөтэйгүүр, өгөгдөл дэх ижил утгатай элементүүдийн анхны дараалыг хадгалах шаардлагатай тохиолдолд тогтвортой (stable) эрэмбэлэлт чухал ач холбогдолтой. Энэ төрлийн алгоритмуудад Merge Sort, Insertion Sort, Bubble Sort, болон Radix Sort багтдаг. Эдгээр нь өгөгдлийн бүртгэл, ангилал зэрэгт өргөн хэрэглэгддэг.

Мөн зарим алгоритмууд нь бага хэмжээний санах ой (memory footprint) ашигладаг давуу талтай байдаг ба үүнд мөн Merge Sort, Insertion Sort, Bubble Sort, болон Radix Sort хамаардаг. Энэ нь хязгаарлагдмал нөөц бүхий орчинд алгоритм хэрэгжүүлэхэд үр ашиг-тай болгодог.

Түүнчлэн, параллел боловсруулалт хийхэд тохиромжтой алгоритмуудад Merge Sort болон Quick Sort багтдаг. Эдгээр нь өгөгдлийг хэсэгчлэн боловсруулж, олон цөмт процессорын боломжийг бүрэн ашиглах нөхцөлийг бүрдүүлдэг тул өндөр хурдтай системүүдэд тохиромжтой алгоритмуудын нэг юм.

Доорх харьцуулалтуудын үр дүнд тулгуурлан Нэгтгэн эрэмбэлэх (Merge sort) сонгон хэрэгжүүлсэн.

Sort algorithm	Time complexity			Space complexity	Sort stability
	Best case	Average case	Worst case	Worst case	Stable
Bubble sort	$O(n)$	$O(n^2)$	$O(n^2)$	$O(1)$	stable
Selection sort	$O(n^2)$	$O(n^2)$	$O(n^2)$	$O(1)$	unstable
Insertion sort	$O(n)$	$O(n^2)$	$O(n^2)$	$O(1)$	stable
Merge sort	$O(n \log n)$	$O(n \log n)$	$O(n \log n)$	$O(n)$	**stable**
Heap sort	$O(n \log n)$	$O(n \log n)$	$O(n \log n)$	$O(n)$	unstable
Quick sort	$O(n \log n)$	$O(n \log n)$	$O(n^2)$	$O(n \log n)$	unstable
Radix sort	$O(nk)$	$O(nk)$	$O(nk)$	$O(nk)$	stable
Count sort	$O(n+k)$	$O(n+k)$	$O(n+k)$	$O(nk)$	stable
Bucket sort	$O(n+k)$	$O(n+k)$	$O(n^2)$	$O(nk)$	stable

Зураг 1. Эрэмбэлэх алгоритмын харьцуулалт.

III. Хэрэгжүүлэлт

Санал болгож буй сайжруулсан тогтмол хугацаанд эрэмбэлэгч нь өмнөх судалгаануудтай адил нэгтгэн эрэмбэлэх алгоритмд суурилсан. [8]-д дурдсан цуваа хэрэгжүүлэлтээс ялгаатай нь, энд дөрвөн элементийг зэрэгцээ эрэмбэлэх хэрэгжүүлэлт ашигласнаар үйл явцыг

хурдасгасан. FIFO-д суурилсан зэрэгцээ эрэмбэлэгч [14]-тай харьцуулахад, энэхүү загварт зөвхөн жижиг хэмжээтэй FIFO ашигласан бөгөөд харьцуулагчдыг бүрэн ашигласан.

Пост-квант криптографт пермутацийн матриц (permutation matrix) байршлын санамсаргүй матриц үүсгэх зорилгоор хэдэн зуу, мянган элементийг эрэмбэлэх шаардлагатай тул санал болгож буй энэхүү загвар нь том хэмжээтэй сүлжээнүүдийг ашигладаг [20]-аас илүү тохиромжтой.

Эрэмбэлэх төхөөрөмжийн гол давуу талууд:

1) **Харьцуулагчдын бүрэн ашиглалт:** Бүх харьцуулагчид эхэлмэгц идэвхтэй ажиллаж, давталт хоорондын шилжилтийн үед ч бүрэн ачаалалтай байна.
2) **Цөөн харьцуулагчид ашиглан өндөр гүйцэтгэл:** Шинэ төрлийн эрэмбэлэх механизмын тусламжтайгаар өмнөх ажлуудаас илүү үр ашигтай болсон.
3) **Жижиг хэмжээтэй FIFO-д суурилсан холбогч:** Санах ойн хандалт, тооцоолох хэрэглээг тэнцвэржүүлэн, тогтмол хугацааны шинж чанарыг хангана.
4) **Гэдрэг холбоогүй (feedback-less) бүтэц:** Эрэмбэлэх явц нь дан нэг чиглэлд буюу нэгэнт эрэмбэлэгдсэн өгөгдлийг дахин буцааж оруулах гэдрэг холбоогүй учир.

A. Эрэмбэлэх үйл явц ба бүтэц

Эрэмбэлэх үйл явц нь 2^m элементийг m давталтад хуваана. i-дахь давталтад 2^i элементийн бүлгүүдийг эрэмбэлэгдсэн дэд массив гэж үзнэ. Давталт бүрт эдгээр дэд массивууд хоорондоо нэгтгэгдэж, 2^{i+1} элемент бүхий урт массив үүснэ. Эцэст нь, m давталтын дараа бүрэн эрэмбэлэгдсэн 2^m элементтэй массив гарна.

Санал болгож буй зэрэгцээ эрэмбэлэгч нь дөрвөн элементийг зэрэг унших, бичих боломжтой бөгөөд дараах гурван үндсэн хэсэгтэй:

- **Ping-pong өгөгдлийн санах ой:** Унших/бичих үйлдлийг тасралтгүй хийхэд тусалдаг санах ой.
- **FIFO-д суурилсан холбогч:** FIFO-A болон FIFO-B нь тус бүр дөрвөн нэг-элементийн гүнтэй дэд-FIFO-уудаас бүрддэг. Бүх дэд FIFO-ууд бичих хаягийг хуваалцдаг.
- **Сондгой-тэгш эрэмбэлэх сүлжээ:** Дөрвөн элементийг зэрэг эрэмбэлэх чадвартай сүлжээ. Эрэмбэлэх сүлжээнд FIFO-A болон FIFO-B-ийн дээд 8 элемент оролцоно. Харьцуулагчдын үр дүнгээс шалтгаалан dequeue логик аль 4 элементийг pop хийхийг шийднэ. FIFO-A/B-ийн элементүүдийг ашиглан эрэмбэлэгдсэн 4-элементийг гарцын санах ойд бичнэ. Давталт бүрийн дараа оролт болон гарцын санах ойнуудын үүрэг солигддог.

B. Харьцуулалт ба баталгаа

Харьцуулалтад дараах функцийг ашиглана:

$$f(k) = a_{i+k \bmod 4} - b_{j+3-k \bmod 4}, \quad k \in [0,3]$$

Хэрэв $f(k) \leq 0$ бол $a_{i+k \bmod 4}$-г pop хийж хэрэглэнэ. Энэ функц монотоноор өсдөг.

Дараагийн клокт FIFO санах ойд байгаа элементүүд мөн эрэмбэлэгдсэн хэвээр байна. Ингэснээр үргэлж хамгийн бага 4 элемент гарч ирэх юм.

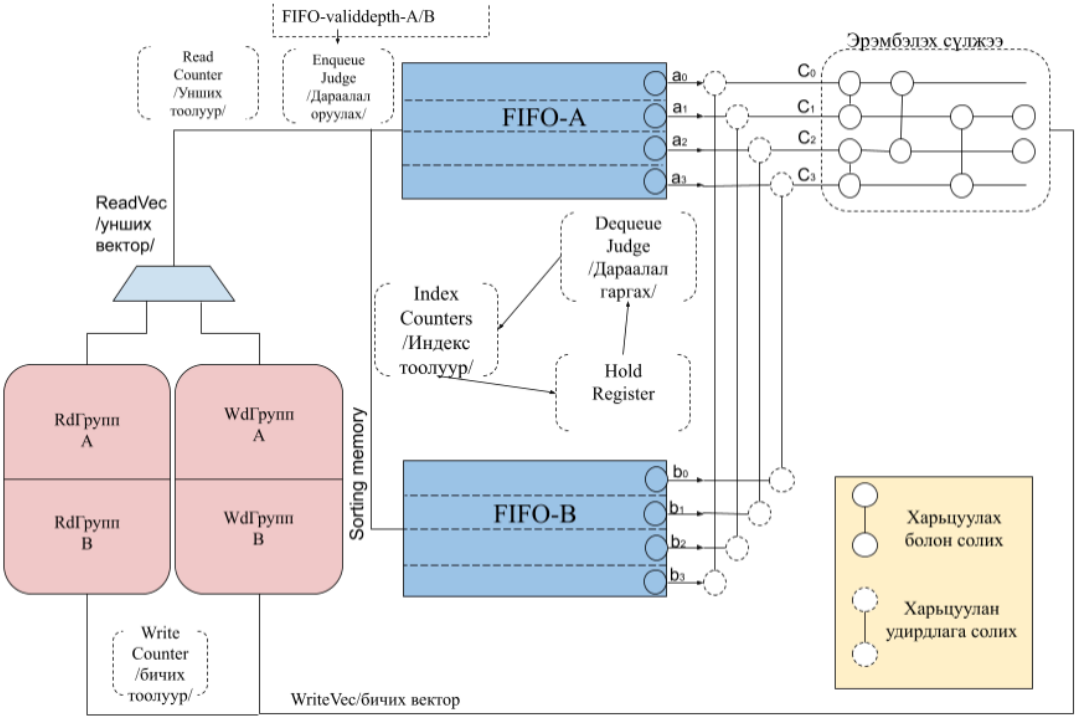
Унших хаягуудыг унших тоолуур үүсгэдэг. Энэхүү тоолуур нь A болон B бүлгүүдийн хаягийг удирддаг бөгөөд унших үйл явцын үед харгалзан 1-ээр нэмэгддэг. Харин бичих тоолуур нь вектор, бүлэг болон итерац тус бүрийн бичих индексүүдийг бие даан удирдаж, A болон B бүлгүүдийн бичих хаягуудыг гаралтын санах ойд ээлжлэн үүсгэнэ. Дараагийн давталтад уншигдах бүлгүүдийн өгөгдөл нь урьдчилан бичигдсэн байдаг тул давталтын шилжилтийн үе шатанд нөөцийн бүрэн ашиглалт (full-utilization) хангагдана. Нэг клокт FIFO-A-ийн дээд 4 элемент болох a_0, a_1, a_2, a_3, мөн FIFO-B-ийн b_0, b_1, b_2, b_3 нь тодорхой харьцааг хангаж байна гэж үзье. Энд i болон j нь $[0,3]$ мужид багтах бүхэл тоонууд юм. Хэрвээ энэ клокт k_1 ширхэг элемент FIFO-A-гаас, k_2 ширхэг элемент FIFO-B-гаас гарсан бол дараах дараалалтай:

$$a_i, a_{i+1 \bmod 4}, \ldots, a_{i+k_1-1 \bmod 4}$$
$$b_j, b_{j+1 \bmod 4}, \ldots, b_{j+k_2-1 \bmod 4}$$

Энэ нөхцөл нь бүлгүүдийн харьцуулалтын анхны клокт автоматаар хангагддаг бөгөөд энэ үед $i = j = 0$ байна. Үүнээс гадна, гаралтын вектор c_i нь үлдсэн бүх элементүүдээс хамгийн бага 4 элементийг агуулсан байдаг болохыг баталж болно. Мөн дээрх тэгшитгэлд дараагийн клокт ч хүчин төгөлдөр хэвээр хадгалагдана. Эдгээрийн топ заагчдыг дараагийн байрлал руу шилжүүлнэ. Үүний үр дүнд дараагийн клокт топод гарч ирэх үлдсэн элементүүд болох $a_{(i+k_1) \bmod 4}$, $a_{(i+1+k_1) \bmod 4}$, $\ldots$, $a_{(i+3+k_1) \bmod 4}$ нь мөн адил эрэмбэлэгдсэн хэвээр байх бөгөөд эдгээр нь өмнө pop хийгдсэн a_i-аас их байна. Энэ нь оролтын дараалал анхнаасаа эрэмбэлэгдсэн байсантай холбоотой. Мөн дараагийн клокт FIFO-A дахь үлдсэн хамгийн бага элемент болох $a_{(i+k_1) \bmod 4}$ нь FIFO-B-ийн хамгийн сүүлд гарсан элемент $b_{(j+k_2-1) \bmod 4}$-оос их байна гэдэг нь харьцуулалтын үр дүнгээр харагдана. Үүнтэй ижил нөхцөл FIFO-B-д ч мөн хамаарна. Иймд тухайн клокын гаралтын вектор c_i нь үргэлж хамгийн бага 4 элементээс бүрдэж байгааг баталж болно.

FIFO-A болон FIFO-B хоосорсон үед holding register идэвхжиж, дараагийн бүлгийн элементүүдийг pop хийхээс сэргийлнэ. Хэрэв аль аль нь хоосорч байвал holding register идэвхгүй болж, дараагийн бүлэг элементэд шилжинэ.

FIFO-ийн гүн нь 10, босго нь 5 байна. Pop $\rightarrow$ гүн шинэчлэх $\rightarrow$ эрэмбэлэх $\rightarrow$ унших $\rightarrow$ enqueue хийх гэсэн



Зураг 2. Санал болгож буй архитектур

логик зам нь нийт 4 клок шаардана. Иймд сүлжээ ажиллаж эхлэх үед FIFO бүрт дор хаяж 5 элемент байна.

C. Нөөцийн ашиглалт ба дахин эхлүүлэх нөхцөл

Энэ загварт 32-битийн 9 ширхэг харьцуулагчид ашиглагддаг. Үүн дээр 12/13 битийн "satellite" тогтмолууд нь санамсаргүй урьдчилсан утгуудтай хамт дамждаг. Хэрэв хоёр урьдчилсан утга тэнцүү бол харьцуулалт хийгч үүнийг илрүүлж, системийн хянагч руу дахин эхлүүлэх дохио илгээдэг.

Эрэмбэлэх клокын хугацаа:

$$(m-2) \times \frac{2^m}{4}$$

Нэмэлт хадгалах багтаамж нь зөвхөн 80 коэффициент хадгалах бөгөөд BRAM бус зөвхөн флип-флопуудаар хангагдсан. Энэ нь санах ойн хэрэглээг ихээхэн бууруулсан.

D. FIFO-ийн гүний хяналт ба нөхөн сэлбэлт

FIFO-A болон FIFO-B-ийн гүнийг тус тусын гурав-дугаар sub-FIFO (FIFO-A3, FIFO-B3)-д агуулагдаж буй хүчинтэй элементүүдийн тоогоор хянадаг. Энэ нь тухайн багана дахь өгөгдөл тасралтгүй рор хийгдсэн эсэхийг илтгэнэ.

Enqueue логик нь гүний мэдээллийг ашиглан тухайн FIFO-ийн хүчинтэй гүн босгоноос доош орсон эсэхийг шалгадаг. Хэрвээ тийм бол Groups A эсвэл B-гээс уншиж авах дараагийн векторын тухай шийдвэрийг гаргана. Энэхүү эрэмбэлэгч нь өгөгдлийг жигд, тогтмол хурдаар боловсруулж, FIFO-нуудын аль нэг нь хэт хоосрох эсвэл хэт дүүрэхээс сэргийлдэг.

Enqueue хийх хүртэл дараах 5 шат дамждаг бөгөөд нийт 4 клок шаарддаг:

1) Рор хийх
2) Хүчинтэй гүнийг шинэчлэх
3) Enqueue хийх эсэхийг шүүх
4) Унших үйлдэл
5) Эцсийн enqueue

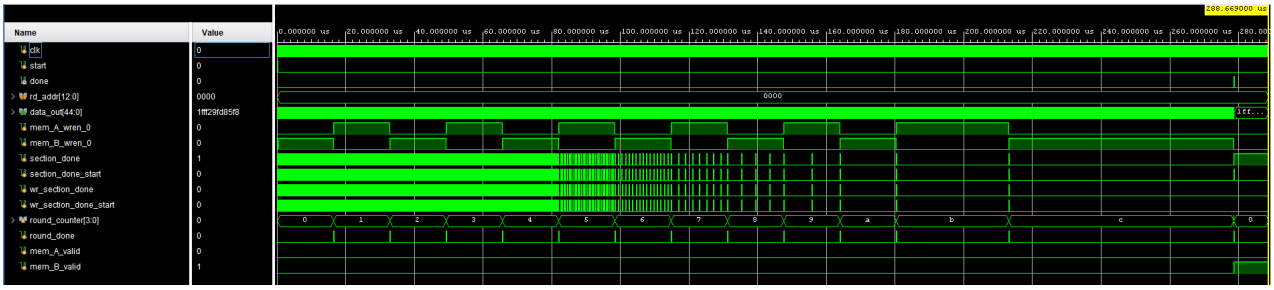
Энэхүү санал болгосон тогтмол хугацаат эрэмбэлэгчийг NTRU зэрэг пост-квант криптографын алгоритмуудын техник хангамжийн хэрэгжүүлэлтэд мөн ашиглаж болно. Хэрвээ тухайн алгоритмд байршилд санамсаргүй алдаа үүсгэхийн тулд эрэмбэлэлт хэрэглэгддэг бол энэхүү архитектур шууд ашиглагдах боломжтой юм.

Хүснэгт I
Хэрэгжүүлэлтийн харьцуулсан үр дүн

	LUTs	FFs	BRAMs	Cycles	Frequency
[8]	583	411	20	147505	250M
[14]	2533	1589	33	26646	250M
Proposed	2510	3887	20	24598	300M

IV. Дүгнэлт

Classic McEliece алгоритмын хамгийн их цаг зарцуулдаг модулиудын нэг болох эрэмбэлэх үйл явцыг



Зураг 3. Хэрэгжүүлэлтийн үр дүн

хурдасгах үүднээс шинэ төрлийн тогтмол хугацааны, параллел хэрэгжүүлэлт бүхий архитектурын загварыг боловсруулаа. Энэ загвар нь өндөр гүйцэтгэлтэйгээр харьцуулагчдыг бүрэн ачаалалтай ажиллуулж, FIFO-д суурилсан холбогчийн тусламжтайгаар эрэмбэлэх болон санах ойн ажиллагааг тэнцвэржүүлдэг. Шинэ архитектур нь зөвхөн 9 ширхэг харьцуулалт ашиглан эрэмбэлэх сүлжээг бүрдүүлж, 80-хан коэффициент хадгалах шаардлагатайгаар нийт хадгалах багтаамжийг ихээхэн бууруулсан. Санал болгож буй шийдэл нь зөвхөн McEliece алгоритмд бус, NTRU зэрэг постквант криптографын бусад алгоритмуудад ч ашиглагдах боломжтой юм. Иймд энэхүү судалгаа нь PQC-ийн техник хангамжийн хэрэгжилтийн ирээдүйн судалгаа, стандартчилалд чухал хувь нэмэр оруулна гэж үзэж байна.

Ашигласан ном

[1] D. Cooper, G. Alagic, D. Apon, and et al., "Status report on the second round of the nist post-quantum cryptography standardization process," NIST, Technical Report, 2022.

[2] D. J. Bernstein, T. Chou, and P. Schwabe, "Mcbits: Fast constant-time code-based cryptography," in *Cryptographic Hardware and Embedded Systems - CHES 2013*, G. Bertoni and J.-S. Coron, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 250–272.

[3] T. Chou, "Mcbits revisited," in *Cryptographic Hardware and Embedded Systems – CHES 2017*, W. Fischer and N. Homma, Eds. Cham: Springer International Publishing, 2017, pp. 213–231.

[4] J. Roth, E. Karatsiolis, and J. Kr"amer, "Classic mceliece implementation with low memory footprint," in *Smart Card Research and Advanced Applications*, P.-Y. Liardet and N. Mentens, Eds. Cham: Springer International Publishing, 2021, pp. 34–49.

[5] M.-S. Chen and T. Chou, "Classic mceliece on the arm cortex-m4," *IACR Transactions on Cryptographic Hardware and Embedded Systems*, vol. 2021, no. 3, pp. 125–148, Jul 2021.

[6] A. Shoufan, T. Wink, H. G. Molter, S. A. Huss, and E. Kohnert, "A novel cryptoprocessor architecture for the mceliece public-key cryptosystem," *IEEE Transactions on Computers*, vol. 59, no. 11, pp. 1533–1546, 2010.

[7] W. Wang, J. Szefer, and R. Niederhagen, "Fpga-based key generator for the niederreiter cryptosystem using binary goppa codes," in *Cryptographic Hardware and Embedded Systems – CHES 2017*, W. Fischer and N. Homma, Eds. Cham: Springer International Publishing, 2017, pp. 253–274.

[8] ——, "Fpga-based niederreiter cryptosystem using binary goppa codes," in *PQCrypto*, vol. 10786, 2018.

[9] V. Kostalabros, J. Ribes-Gonz'alez, O. Farràs, M. Moret'o, and C. Hernandez, "Hls-based hw/sw co-design of the post-quantum classic mceliece cryptosystem," in *2021 31st International Conference on Field-Programmable Logic and Applications (FPL)*, 2021, pp. 52–59.

[10] W. Wang, J. Szefer, and R. Niederhagen, "Solving large systems of linear equations over gf(2) on fpgas," in *2016 International Conference on ReConFigurable Computing and FPGAs (ReConFig)*, 2016, pp. 1–7.

[11] P.-J. Chen, T. Chou, S. Deshpande, N. Lahr, R. Niederhagen, J. Szefer, and W. Wang, "Complete and improved fpga implementation of classic mceliece," in *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer, 2022.

[12] J. Hoffstein, C. Chen, O. Danba *et al.*, "NTRU: Algorithm specifications and supporting documentation," NIST PQC Project, Technical report, 2020. [Online]. Available: https://ntru.org/

[13] D. J. Bernstein, M. R. Albrecht *et al.*, "Classic McEliece: Conservative code-based cryptography," NIST PQC Project, Technical report, 2020. [Online]. Available: https://classic.mceliece.org/nist/mceliece-20201010.pdf

[14] K. Mohajerani, V. B. Dang, and K. Gaj, "High-speed hardware architectures and fair fpga benchmarking of crystals-kyber, ntru, and saber," in *Third PQC Standardization Conference*, 2021.

[15] Y. Zhu, W. Zhu, M. Zhu, C. Li, C. Deng, C. Chen, S. Yin, S. Yin, S. Wei, and L. Liu, "A 28nm 48kops 3.4µj/op agile crypto-processor for post-quantum cryptography on multi-mathematical problems," in *2022 IEEE International Solid-State Circuits Conference (ISSCC)*, vol. 65, 2022, pp. 514–516.

[16] J. Casper and K. Olukotun, "Hardware acceleration of database operations," in *Proceedings of the 2014 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA '14)*. New York, NY, USA: ACM, 2014, pp. 151–160.

[17] W. Song, D. Koch, M. Luj'an, and J. Garside, "Parallel hardware merge sorter," in *2016 IEEE 24th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, 2016, pp. 95–102.

[18] S. Mashimo, T. V. Chu, and K. Kise, "High-performance hardware merge sorter," in *2017 IEEE 25th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, 2017, pp. 1–8.

[19] M. Saitoh, E. A. Elsayed, T. V. Chu, S. Mashimo, and K. Kise, "A high-performance and cost-effective hardware merge sorter without feedback datapath," in *2018 IEEE 26th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, 2018, pp. 197–204.

[20] P. Papaphilippou, C. Brooks, and W. Luk, "Flims: Fast lightweight merge sorter," in *2018 International Conference on Field-Programmable Technology (FPT)*, 2018, pp. 78–85.

[21] N. Samardzic, W. Qiao, V. Aggarwal, M.-C. F. Chang, and J. Cong, "Bonsai: High-performance adaptive merge tree sorting," in *2020 ACM/IEEE 47th Annual International Symposium on Computer Architecture (ISCA)*, 2020, pp. 282–294.

90нм технологи ашигласан LDO-ийн дизайн болон хэрэгжүүлэлт

Цогтгэрэлийн Энхтуяа
Электроник, Холбооны Инженерчлэлийн тэнхим
Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол Улс
enkhh0605@gmail.com

Батсүхийн Алтанцоож
Электроник, Холбооны Инженерчлэлийн тэнхим
Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол Улс
altantsooj.b@num.edu.mn

***Хураангуй*—Энэхүү өгүүлэлд 90нм CMOS технологид суурилсан, 1.2V гаралтын хүчдэлийг үүсгэх чадвартай Low Drop-Out (LDO) зохицуулагчийн бүтэц, загварчлал, хэрэгжүүлэлтийг танилцууллаа. Судалгааны гол зорилго нь бага чадал зарцуулалттай IoT болон зөөврийн төхөөрөмжид зориулсан өндөр үр ашигтай, тогтвортой ажиллагаатай LDO зохицуулагчийг зохион бүтээхэд оршино. LDO нь PMOS төрлийн pass transistor, хоёр шатлалт CMOS алдааны өсгөгч, буцах холбооны сүлжээ зэргээс бүрдсэн. Cadence Virtuoso орчинд хийсэн симуляци, туршилтын үр дүнгээр загвар нь 1.8–1.2V оролтын хэлбэлзэл дунд 1.2V тогтвортой гаралт үүсгэж, зөөврийн хэрэглээнд тохиромжтой болохыг харуулсан.**

***Түлхүүр үгс*—LDO, FVF, Pass transistor, Capacitorless LDO.**

I. Удиртгал

Систем-он-чип (SoC) технологи өндөр түвшинд нэгтгэгдэж, олон төрлийн үйлдлийг нэг чип дээр гүйцэтгэх болсонтой уялдан тэжээлийн үүсгүүрийн зохицуулалтыг үр ашигтай гүйцэтгэх шаардлага улам бүр нэмэгдэж байна. Энэ хүрээнд хүчдэлийн бага уналттай тогтворжуулагч (Low Dropout Regulator, LDO) нь өндөр үр ашигтай, дуу чимээ багатай, тогтвортой гаралтын хүчдэлтэй байж чаддаг учир чухал үүрэг гүйцэтгэж байна. LDO нь хурдан хариу үйлдэл үзүүлдэг, гаралтын хүчдэлийн хэлбэлзэл бага байдаг тул аналог болон RF хэлхээнүүдийн тэжээлд өргөн ашиглагддаг.

Гэвч CMOS технологи 90нм түвшинд шилжихийн хэрээр LDO зохиомж нь олон сорилттой тулгардаг. Нэгдүгээрт, хангамжийн хүчдэл багассанаар аналог удирдлагын хэлхээнд ажиллах хүчдэлийн нөөц буурдаг. Хоёрдугаарт, транзисторуудын дотоод өсгөлт буурч, гэдрэг холбооны тогтвортой байдал болон тэжээлийн үүсгүүрийн хэлбэлзэлд тэсвэр багасдаг. Мөн үүнтэй зэрэгцэн шүүрлийн гүйдэл (leakage current) ихэсч, үйлдвэрлэлийн процесс нөлөөлөх нь ихэсдэг.

Ийм нөхцөлд ч 90нм CMOS технологи нь өртөг багатай, эрчим хүчний хэмнэлттэй, жижиг төхөөрөмжүүд (жишээ нь IoT төхөөрөмж, мэдрэгчийн интерфейс, зүүдэг төхөөрөмжүүд)–д тохиромжтой хэвээр байна. Тиймээс эдгээр хэрэглээнд зориулсан бага чадал зарцуулдаг, найдвартай ажиллагаатай, талбайн хувьд хэмнэлттэй LDO зохиомж нь өндөр ач холбогдолтой юм.

Сүүлийн жилүүдэд 90нм CMOS технологид суурилсан хүчдэлийн бага уналттай тогтворжуулагчийн (LDO) дизайны талаар олон судалгааны ажил хийгдсэн байна. Эдгээр судалгаанууд нь LDO-ийн гүйцэтгэлийг сайжруулах, ялангуяа эрчим хүчний хэмнэлт, талбайн үр ашиг, түргэн хариу үйлдэл зэрэгт анхаарлаа хандуулсан.

2021 онд [1] нь систем-он-чип (SoC) хэрэглээнд зориулсан хурдан хариу үйлдэлтэй CMOS LDO зохиомжийг танилцуулсан. Энэхүү зохиомж нь Flipped Voltage Follower (FVF) архитектурыг ашиглан түргэн шилжилтийн хариу үйлдлийг сайжруулсан бөгөөд 90нм CMOS технологид хэрэгжүүлсэн байна.

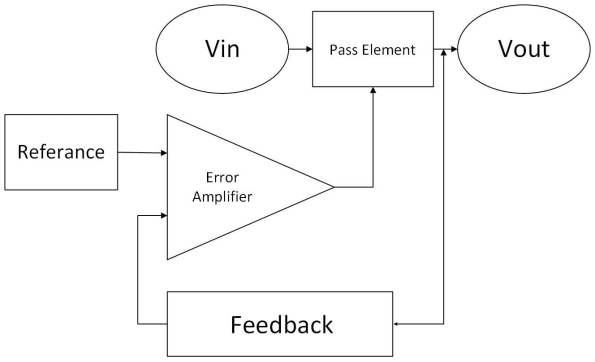
2022 онд [2] бага чадал, өндөр талбайн үр ашигтай NMOS LDO зохиомжийг танилцуулсан. Энэхүү зохиомж нь хурдан адаптив тэжээлийг ашиглаж, 0.005 mm^2 талбай эзэлдэг бөгөөд $2.91\mu A$-ийн нам гүйдэлтэй ажилладаг.

2023 онд [3] зөөврийн төхөөрөмжүүдэд зориулсан гурван шатлалт CMOS LDO зохиомжийг боловсруулсан. Энэхүү зохиомж нь 1.8V гаралтын хүчдэлийг 100mV уналттайгаар зохицуулж, $13.8\mu A$-ийн нам гүйдэлтэй, $314\mu m \times 150\mu m$талбай эзэлдэг. Энэ нь зөөврийн хэрэглээнд тохиромжтой болохыг харуулж байна.

Эдгээр судалгаанууд нь 90нм CMOS технологид суурилсан LDO зохиомжийн чиг хандлагыг тодорхойлж, эрчим хүчний хэмнэлт, талбайн үр ашиг, түргэн хариу үйлдэл зэрэг үзүүлэлтүүдийг сайжруулахад чиглэгдсэн байна. Ирээдүйд эдгээр чиглэлээр нэмэлт судалгаа, хөгжүүлэлт хийх нь чухал юм.

II. LDO-ийн дизайн

Хүчдэлийн бага уналттай тогтворжуулагч (Low Dropout Regulator, LDO) нь систем-он-чип (SoC), мэдрэгчийн интерфейс, био-эмнэлгийн төхөөрөмж зэрэг эрчим хүчний хэмнэлтийн өндөр шаардлагатай хэрэглээнд чухал үүрэгтэй. Энэхүү хэсэгт 90нм CMOS технологид суурилсан LDO-ийн дизайн, архитектурын сонголт, симуляци болон хэрэгжүүлэлтийн үе шатуудыг дэлгэрэнгүй танилцуулна.



Зураг 1. LDO ерөнхий дизайн

A. LDO-ийн Ерөнхий Бүтэц

Стандарт LDO нь дараах үндсэн блокуудаас бүрдэнэ:

- **Error amplifier** – Оролтын хүчдэл V_{ref} болон буурруулсан гаралтын хүчдэлийг харьцуулан ялгааг өсгөдөг.
- **Pass transistor** – Гаралтын гүйдлийг хянах үүрэгтэй (ихэвчлэн PMOS/NMOS).
- **Feedback network** – Гаралтын хүчдлийг буцаан error amplifier руу оруулдаг.
- **Reference voltage source** – Bandgap Reference хэлбэрээр хэрэглэгддэг.

LDO-ийн гаралтын хүчдэлийг дараах томьёогоор тодорхойлно:

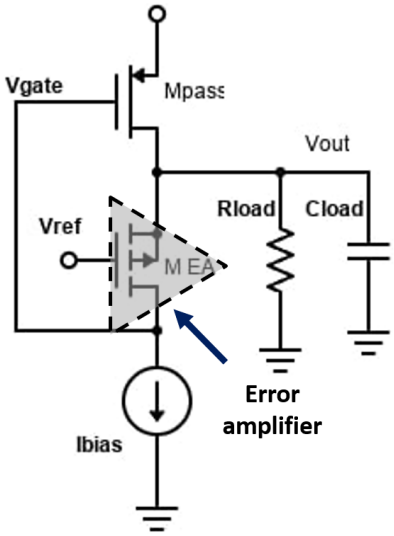
$$V_{\text{out}} = V_{\text{ref}}\left(1 + \frac{R_1}{R_2}\right)$$

B. Архитектурын Хувилбарууд

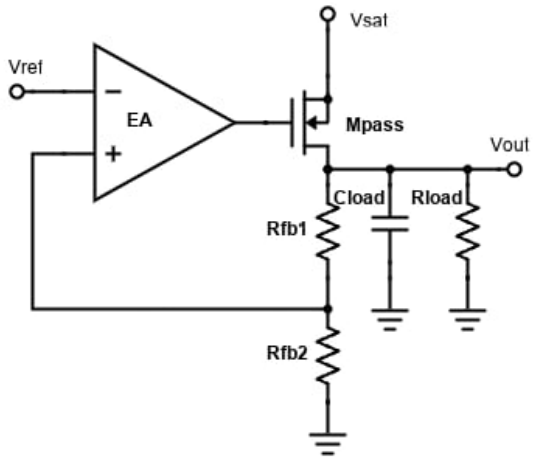
1) Flipped Voltage Follower (FVF): FVF архитектур нь pass transistor-ийн сурвалжийг өсгөгчөөр удирдаж хурдан хариу үйлдэлтэй үзүүлдэг [4]–[7]. Ажиллагаа тогтвортой бөгөөл симметрик хүчдэлийн удирдлагатай зэрэг давуу талуудтай. Ачаалал багатай үед хязгаарлагдмал, гаралтын хүчдлэл оролтын хүчдэлээс их хамаардаг зэрэг сул талуудтай. Үүний дизайныг Зураг 2-д үзүүлэв.

2) Pass Transistor: Pass transistor бүхий LDO нь оролтын болон гаралтын хүчдэл хоорондын ялгааг хянах үүрэгтэй PMOS эсвэл NMOS транзистор ашигладаг [8]–[10]. Энэ нь төрөлийн дизайн нь хүчдэлийн маш бага уналттай, өндөр үр ашигтай, тогтвортой ажиллагаатай бөгөөд ачааллын хариу үйлчлэл хурдантай зэрэг давуу талуудтай. Архитектур нь нарийн төвөгтэй тул тохиргоо хийхэд хэцүү сул талтай. Зураг 3-д хэлхээг үзүүлэв.

3) Capacitor-less LDO: Capacitor-less LDO нь гадны багтаамжгүй ажиллах зориулалттай бөгөөд эрчим хүчний хэрэглээг бага байлгах, том талбайг хэмнэх боломжийг олгодог. Энэ нь ихэвчлэн Miller Compensation болон Advanced Frequency Compensation аргачлалуудыг ашигладаг [11]–[13]. Гадны багтаамжгүй тул талбай бага эзэлдэг, нэмэлт багтаамж шаардахгүйгээр ажиллах



Зураг 2. FVF LDO

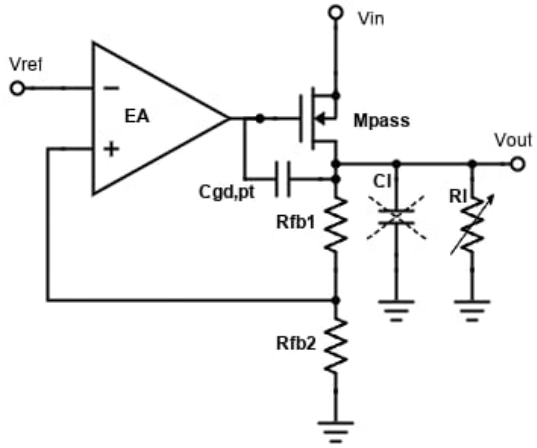


Зураг 3. Pass transistor LDO

давуу талуудтай. Хүчдэлийн уналт өндөр, зарим тохиолдолд ажиллагаа тогтворгүй болон өндөр нарийвчлалтай хэрэглээнд тохиромжгүй зэрэг сул талуудтай. Хялбар хэлхээг Зураг 4-д харуулав.

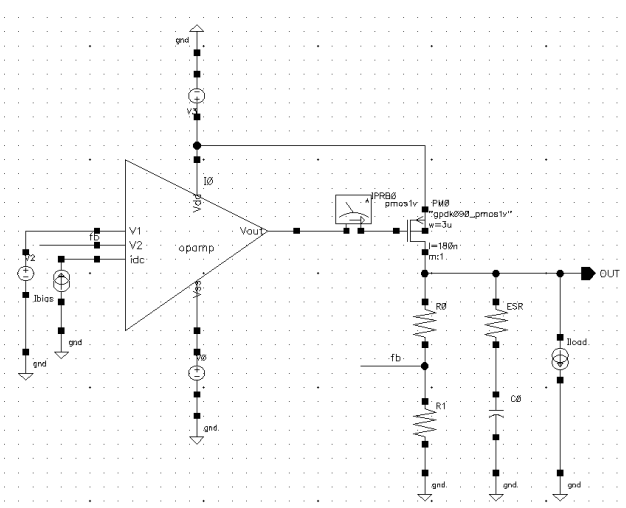
III. Хэрэгжүүлэлт

Энэхүү судалгааны ажилд CMOS технологид тулгуурлан, 1.2V гаралтын хүчдэлтэй, бүрэн CMOS бүтэцтэй LDO-ийн дизайныг боловсруулж, Cadence Virtuoso программын орчинд хэрэгжүүлэв. Загварын бүх үзүүлэлтүүд хангагдсан эсэхийг шалгахын тулд бид хэлхээг



Зураг 4. Capacitorless LDO

стандарт LDO-ийн ерөнхий бүтцийн дагуу угсарсныг Зураг 5-д үзүүлэв.



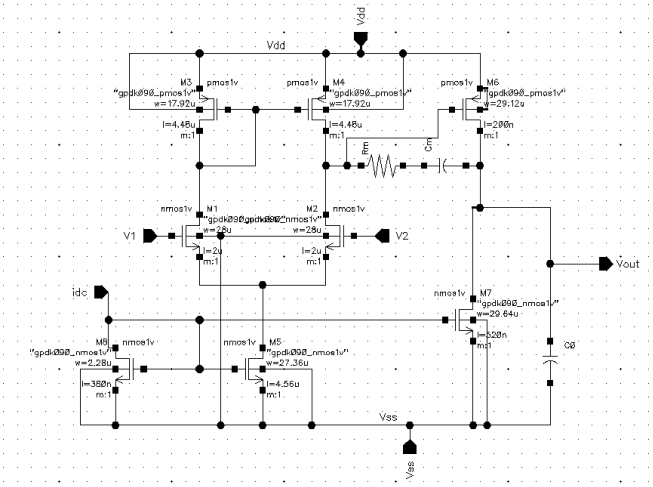
Зураг 5. Virtuoso программ дээрх LDO-ийн хэлхээ

LDO нь Error amplifier, Pass transistor, Feedback network, Reference voltage source гэсэн үндсэн блокуудаас бүрдэнэ.

A. Error amplifier

LDO дизайны энэ хэсэгт хоёр шатлалт CMOS үйлдлийн өсгөгч (two-stage operational amplifier) ашигласан ба үйлдлийн өсгөгчийн хэрэгжүүлэлтийн хэлхээг Зураг 6-д харуулав.

CMOS үйлдлийн өсгөгчийн хоёр үе шатыг оновчтой болгохын тулд онолын тооцооллуудыг хийсэн. Мөн транзисторын хэмжээ параметрүүдэд хэрхэн нөлөөлж буйд үндэслэн туршилтууд хийснээр транзисторуудын хэмжээг тодорхойлсныг Хүснэгт I -д оруулав.



Зураг 6. Хоёр шатлалт CMOS үйлдлийн өсгөгчийн хэрэгжүүлэлтийн хэлхээ

Хүснэгт I
Транзисторуудын хэмжээ

Транзисторын дугаар	**W** (μm)	**L** (μm)
M1	28.00	2.000
M2	28.00	2.000
M3	17.92	4.480
M4	17.92	4.480
M5	27.36	4.560
M6	29.12	0.200
M7	29.64	0.520
M8	2.28	0.380

B. Pass transistor

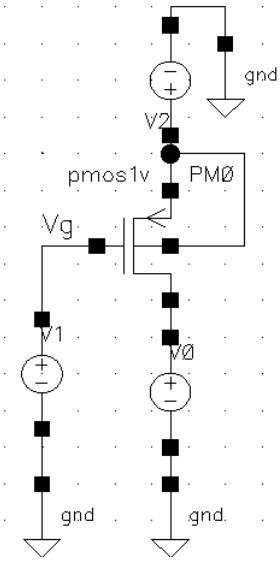
NMOS pass transistor нь ажиллахын тулд гейт дээрх хүчдэлээс хамааран өндөр тэжээлийн хүчдэл (VDD) шаардлаг. Үүнийг хангахын тулд Charge pump хэлхээг нэмэлтээр ашигладаг. Энэ нь гаралтын шуугианы түвшинг нэмэгдүүлдэг тул PMOS pass transistor-ийг сонгон ашиглав.Транзисторын параметрийн хэмжээг тодорхойлохын тулд Зураг 7 -д үзүүлснээр туршин тооцоолов. Тэжээлийн хүчдэл 1.8В, транзисторын суваг дээрх гаралтын хүчдэл 1.2В байна. Транзисторын Vgs хүчдэл нь гүйдлийг удирдахад 400мВ гэж үзээд PMOS pass transistor-ын W ба L нь тус тус 500нм ба 100нм байна.

IV. Үр дүн

Энэ хэсэгт хэрэгжүүлэлт хийсэн дизайны ажиллагаг Cadence Virtuoso программ дээр Spectre симуляцийн арга ашиглан туршилтууд хийн шалгав.

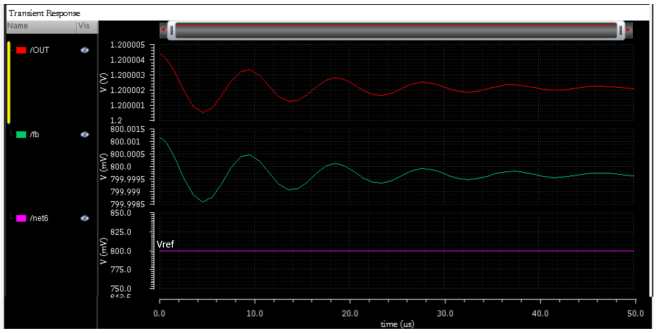
A. Шилжилтийн анализ (transient analysis)

Ачааллын өөрчлөлтөд гаралтын хүчдэл хэрхэн хариу үйлдэл үзүүлж байгааг шилжилтийн(transient) анализаар үнэлэв. Ачааллын гүйдлийн хязгаар 0 мА – 50 мА-ийн хооронд хэлбэлзсэн нөхцөлд симуляци хийгдсэн.



Зураг 7. PMOS pass transistor -ийн параметрийн утга тохируулах

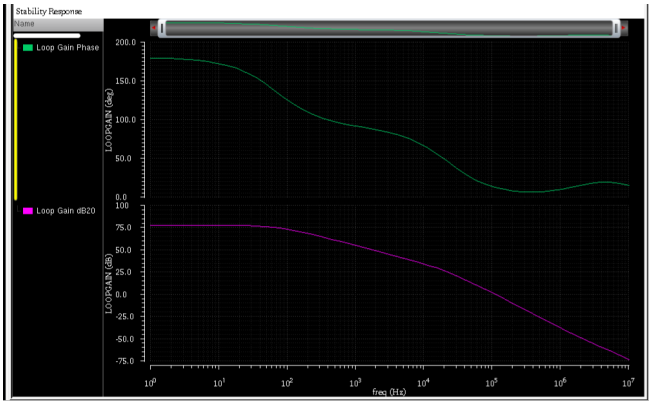
Зураг 4.1-д ачааллын гүйдлийн өөрчлөлтөд LDO-ийн гаралтын хүчдэлийн хариу үйлчлэлийг харуулсан Transient Response график байна. Гаралтын хүчдэл (V_{OUT}, улаан өнгөтэй) 1.2 В орчимд тогтвортой хадгалагдаж, хэлбэлзэл 10 мВ-ээс хэтрээгүй байна. Энэ нь хэрэгжүүлсэн LDO тогтворжуулагчийн дизайн нь ачааллын нөхцөлд тогтвортойгоор ажиллаж байгааг баталгаажуулж байна. Мөн уг үр дүнгээс харахад тогтворжих хугацаа нь 25-30uc байгаа нь харагдаж байна.



Зураг 8. LDO тогтворжуулагчийн Transient Response: Гаралтын хүчдэл (V_{OUT}, улаан), V_{fb} (ногоон), болон $V_{NET6=ref}$ (ягаан) сигналуудын хугацааны хамаарал.

B. Тогтворжилтын анализ (stability analysis)

Тогтворжуулагчийн тогтвортой байдлыг тогтворжилтын анализ(stability analysis) хийн шалгасан ба давтамжийн хариу үйлчлэлийг Зураг 9-д харуулав. Энэ анализын үед давтамжийн хязгаарыг 1Гц-ээс 10ГГц байхаар тохируулсан. Уг график дээр давталтын өсгөлт (loopgain)-ийн хэмжээ (dB-ээр) болон фазын (хэмээр) давтамжаас хамаарах хамаарлыг харуулж байна. Фазын хязгаар (phase margin)-ыг (PM, фазын хязгаар)



Зураг 9. LDO тогтворжуулагчийн Bode диаграмм: Loop Gain (ягаан) ба Phase Margin (ногоон) давтамжийн хамаарал.

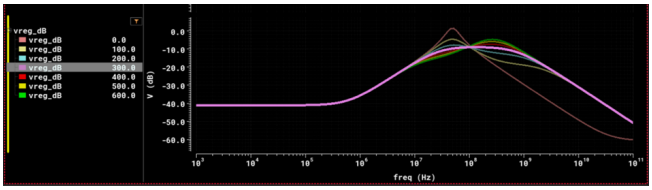
давталтын өсгөлтийн хэмжээ 0 dB байх үеийн (f_{gc}, өсгөлтийн огтлолын давтамж) фазаар тодорхойлно. Графикаас харахад $f_{gc} \approx 200\,\text{kHz}$ ба энэ давтамж дахь фаз ойролцоогоор $-45°$ байна. Иймд, фазын хязгаар нь $\text{PM} = 180° + (-45°) = 135°$ байна. 135°-ийн фазын хязгаар нь LDO хэлхээний тогтворжилт өндөр байна гэдгийг илтгэнэ. Өсгөлтийн хязгаар(gain margin)-ыг (GM, өсгөлтийн хязгаар) давталтын фаз $-180°$ байх үеийн (f_{pc}, фазын огтлолын давтамж) давталтын өсгөлтийн хэмжээгээр тодорхойлно. Графикт үзүүлсэн давтамжийн муж дотор фаз $-180°$-д хүрэхгүй байна. Тиймээс, өсгөлтийн хязгаар нь үр дүнтэйгээр хязгааргүй бөгөөд энэ нь мөн тогтворжилт сайтайг харуулж байна.

C. PSRR (AC анализ)

Зураг10 нь PSRR (Power Supply Rejection Ratio), давтамжийн хамаарлыг харуулж байна. Энэ график нь 10^3 Гц-ээс 10^{11} Гц хүртэлх давтамжийн хүрээнд бөгөөд долоон өөр ачааллын эсэргүүцлийн утга (0.0, 100.0, 200.0-600.0) дээр үндэслэн анализ хийгдсэн болно.

- Бага давтамжид (10^3 Гц орчимд) PSRR ойролцоогоор -40 dB-аас эхэлнэ. Эндээс харахад PSRR нь ачааллын утгаас бага хамаарч байна.
- Давтамж 10^7 Гц орчимд хүрэхэд PSRR-ийн утга хамгийн муу буюу -20 dB болж өснө. Энэ нь анализийн үеийн дунд давтамжид тэжээлийн хэлбэлзэл гаралтын хүчдэлд илүү их нөлөөлж буйг харуулж байна. Эсэргүүцлийн утгуудын хувьд 0.0 Ω-аас 600.0 Ω хүртэлх өөрчлөлт PSRR-ийн утгад бага зэрэг нөлөөлж байна.
- 10^9 Гц-ээс эхлэн PSRR буурч эхэлж, -40 dB орчимд тогтворжин, дараа нь 10^{11} Гц-д -60 dB хүрч байна. Энэ нь өндөр давтамжид хэлхээний тэжээлийн хэлбэлзлийг дарах чадвар сайжирч буйг илтгэнэ. Эсэргүүцлийн утгуудын нөлөө энд илүү тодорхой харагдаж, их утгатай резисторууд (600.0 Ω+) арай илүү сайн гүйцэтгэл үзүүлж байна.

$$\mathrm{PSRR}^{-1}(s) = \frac{V_{\mathrm{REG}}(s)}{V_{\mathrm{DD}}(s)} = \frac{H_{\mathrm{supply}}(s)}{(1 + L_{\mathrm{OL}}(s))} \tag{1}$$



Зураг 10. PSRR, давтамжийн хамаарал

V. Дүгнэлт

Энэхүү судалгаагаар 90нм CMOS технологид тулгуурласан бага чадлын хэрэглээнд зориулагдсан, 1.2V гаралттай, PMOS pass transistor ашигласан LDO зохицуулагчийг амжилттай загварчилж хэрэгжүүллээ. Хоёр шатлалт CMOS өсгөгч болон гадны багтаамж шаардлагагүй архитектурын сонголт нь хэлхээг илүү энгийн, найдвартай болгосон. Симуляцийн үр дүнгээр LDO зохицуулагч нь хүчдэлийн уналт багатай, хурдан хариу үйлдэлтэй бөгөөд IoT болон зөөврийн төхөөрөмжийн шаардлагад нийцэж байгааг нотоллоо. Цаашид хүчин чадал, үр ашгийг сайжруулах чиглэлээр өргөтгөсөн архитектурын судалгаа хийх нь зүйтэй.

Хүснэгт II
Гүйцэтгэлийн харьцуулалт

Parameter	This Work	[14]	[15]	[16]
Technology	**90nm**	130nm	90nm	90nm
Supply Voltage (V)	**1.5–1.8**	1.2	1.8	1.2–1.8
Output Voltage (V)	**1.2**	1.0	1.5	1.2
Quiescent Current (μA)	**35**	6	30	40
Dropout Voltage (mV)	**150**	32.06	100	150
Max Load Current (mA)	**100**	100	150	200
Power Consumption (μW)	**48**	7.2	36	48
Line Regulation (%/V)	**0.44**	0.02	0.1	1
Load Regulation (μV/mA)	**340**	0.5	1	4
Output Variation (mV)	**12.4**	10	15	24
PSRR (dB @ 50 kHz)	**–51.2**	–60 @ 2.4 MHz	–55	–50
Settling Time (μs)	**25**	2	1.5	100
Efficiency (%)	**93.3**	83.3	93.8	91.7
Output Cap (μF)	**1.0**	0	0.05	1

Ашигласан ном

[1] H. Wang, "Design of a fast response cmos low dropout regulator for system-on-chip applications," Master's thesis, Nanyang Technological University, Singapore, 2021.

[2] D. Zhang, J. Guo, and J. Xu, "A low-power high area-efficiency nmos ldo with fast adaptive bias," *Integration*, vol. 88, pp. 1–8, Nov. 2022.

[3] J. A. Serrano-Reyes, J. A. D. Luna-Domínguez, and L. T. Aguilera-Cortés, "A three-stage cmos ldo with high power and dynamic performance for portable devices," *Electronics*, vol. 12, no. 22, p. 4638, Nov. 2023.

[4] J. Lee, H. Kim, and B. Kim, "A 0.6-v low-dropout regulator with flipped voltage follower for ultra-low-power applications," *IEEE Transactions on Power Electronics*, vol. 29, no. 9, pp. 4701–4707, Sept. 2014.

[5] S. S. R. Reddy, S. B. K. V. P., and S. Chandra, "Design of low-power low-dropout regulator with flipped voltage follower architecture for soc applications," in *Proc. IEEE 8th Int. Conf. Electronics and Communication Systems (ICECS)*, Coimbatore, India, 2019, pp. 276–281.

[6] A. H. Amiri, M. Shams, and K. L. Lee, "A high-performance low dropout regulator with flipped voltage follower architecture," *Microelectronics Journal*, vol. 65, pp. 93–101, Jan. 2017.

[7] S. D. K. P. V. Kumar, R. Prabhu, and S. V. N. Krishna, "Flipped voltage follower (fvf) based low dropout regulator with enhanced load regulation," in *Proc. 2018 Int. Conf. Communication and Signal Processing (ICCSP)*, Chennai, India, 2018, pp. 0759–0763.

[8] J. Lee, H. Kim, and B. Kim, "A 0.6-v low-dropout regulator with pass transistor for ultra-low-power applications," *IEEE Transactions on Power Electronics*, vol. 32, no. 6, pp. 4741–4747, June 2017.

[9] S. Choi, J. Kim, and S. Lee, "A pass-transistor low-dropout regulator with enhanced load regulation and fast transient response," *Microelectronics Journal*, vol. 72, pp. 56–63, March 2018.

[10] R. Gupta, M. Kumar, and V. R. V. Srinivas, "A high-performance low dropout regulator using pass-transistor architecture," in *Proc. IEEE 2020 Int. Conf. on Electronics and Communication Systems (ICECS)*, Chennai, India, 2020, pp. 278–282.

[11] S. Choi, J. Kim, and S. Lee, "A capacitor-less low dropout regulator with high power supply rejection ratio for portable applications," *IEEE Transactions on Power Electronics*, vol. 32, no. 5, pp. 3523–3530, May 2017.

[12] L. Yang, W. Li, and X. Liu, "Design of a high-performance capacitor-less low dropout regulator for ultra-low-power applications," *Microelectronics Journal*, vol. 92, pp. 70–77, July 2021.

[13] X. Liu, Y. Zhang, and Z. Wang, "A capacitor-less low dropout regulator with high efficiency for iot devices," in *Proc. IEEE 2019 International Symposium on Circuits and Systems (ISCAS)*, Sapporo, Japan, 2019, pp. 2586–2589.

[14] L. Chen, K. Zhang, and Y. Liu, "A capacitor-less ldo with ultra-low dropout voltage in 130nm cmos," *Journal of Analog Integrated Circuits and Signal Processing*, vol. 107, no. 2, pp. 345–352, Apr 2021.

[15] M. Kumar and S. Gupta, "High-efficiency low dropout regulator in 90nm cmos for portable devices," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2022, pp. 1234–1238.

[16] J. Choi and H. Kim, "Design of a high-psrr fast-response ldo in 90nm cmos for iot devices," *Analog Integrated Circuits and Signal Processing*, vol. 110, pp. 115–122, 2022.

Хүний хоолойн өнгөнд суурилсан уушгины хатгалгааг оношлох судалгаа

Дорждэрэмийн Балжинням[1]
Мэдээллийн технологи
ШУТИС, МХТС
Монгол улс, Улаанбаатар
B210930837@must.edu.mn

Бямбадоржийн Золзаяа[2]
Мэдээллийн технологи
ШУТИС, МХТС
Монгол улс, Улаанбаатар
b.zolzaya@must.edu.mn

***Хураангуй*: Техник технологийн хурдацтай хөгжил нь бидний амьдралд өргөн цар хүрээтэй нөлөө үзүүлж, ялангуяа эрүүл мэндийн салбарт томоохон ахиц дэвшлийг авчирч байна. Энэхүү судалгааны ажлаар хиймэл оюуны технологийг ашиглан хүний дуу хоолойн өнгөнд тулгуурлан уушгины хатгалгаа өвчнийг эрт үе шатанд нь оношлох боломжийг судалсан. Хөгжүүлсэн систем нь хэрэглэгч хаанаас ч, хэзээ ч өөрийн хоолойгоор дамжуулан оношлуулж, өвчний талаарх зөвлөмж, мэдээллийг хялбар байдлаар авах боломжийг бүрдүүлдэг. Энэ нь хэрэглэгчдэд зөвхөн уушгины хатгалгаа төдийгүй амьсгалын замын өвчнүүдийн эрсдэлээ урьдчилан мэдэх, шаардлагатай арга хэмжээг цаг алдалгүй авах нөхцөлийг бүрдүүлж, цаашлаад цаг хугацаа болон санхүүгийн зардлыг хэмнэх бодит боломжийг олгож буй юм. Бид уг судалгаанд RF, CNN, SVM, KNN зэрэг аргуудыг эрүүл болон өвчний шинж тэмдэгтэй нийт 1310 аудио өгөгдлийг, мөн өгөгдлийн багц дах хатгалгаа өвчтэй хамааралтай нэмэлт 15 аттрибутуудаас хамаарлыг үнэлэн сонгож туршсан бөгөөд хамгийн өндөр гүйцэтгэлтэйгээр RF-ийн хувьд 95% нарийвчлалтай гарсан юм.**

Түлхүүр үг— RF- Random forest, CNN- convolutional neural network, KNN- k-nearest neighbors, SVM – Support vector machine, MFCC

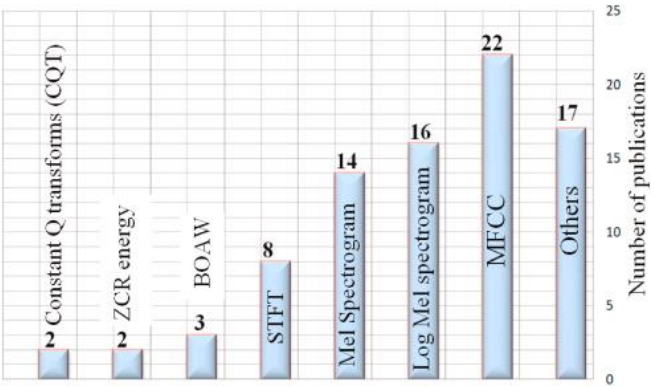
I. Оршил

Манай улсын эрүүл мэндийн салбар олон сорилт, хүндрэлтэй тулгарсаар байна. Тэдгээрийн нэг чухал асуудал нь иргэд эмнэлгийн тусламж үйлчилгээ авахын тулд удаан хугацаанд хүлээх шаардлагатай байдал юм. Жишээлбэл, уушгины хатгалгаа өвчнийг оношлуулахын тулд үзлэгийн цаг авахад дунджаар 7–14 хоног хүлээх тохиолдол түгээмэл бөгөөд үзлэгт орохдоо дахин уушгины зураг авахуулах, эмчид үзүүлэх, эмчилгээний зөвлөгөө авах зэрэг олон шат дамжлагыг дамжих шаардлагатай байдаг. Энэ нь иргэдийн хувьд цаг хугацаа, эдийн засгийн хувьд ихээхэн дарамт үүсгэж байна. Уушгины хатгалгаа нь зөвхөн Монгол Улсад төдийгүй дэлхий нийтэд [1] хүний амь насанд аюул учруулдаг амьсгалын замын ноцтой өвчнүүдийн нэг юм. Энэ өвчин нь уушгины үрэвсэл үүсгэж, амьсгалын бэрхшээл, сул дорой байдал, хоолойн сөөнгөтөлт, цээжээр хатгуулах зэрэг шинж тэмдгүүдээр илэрдэг. Иймд уг өвчнийг аль болох эрт үед нь илрүүлж, оношлох нь эмчилгээний үр дүнг нэмэгдүүлж, хүндрэлээс сэргийлэхэд чухал ач холбогдолтой юм.

Сүүлийн жилүүдэд хүний дуу хоолой, ханиалга, амьсгалын авиа, digital стетоскопоор бичигдсэн уушгины дуу чимээ [3] зэрэг аудио өгөгдөлд суурилан уушгины болон амьсгалын замын өвчнийг эрт үед нь оношлох чиглэлээр хиймэл оюун ухаанд суурилсан судалгаанууд эрчимтэй хөгжиж байна [4]. Машин сургалтын аргуудыг ашиглан аудио өгөгдөлд тулгуурласан оношлогооны системүүдийг хөгжүүлэх нь эмнэлгийн нөөц, цаг хугацааг хэмнэх, эрүүл мэндийн үйлчилгээний хүртээмжийг нэмэгдүүлэх зэрэг томоохон давуу талуудыг бий болгож байна.

Энэхүү судалгаанд бид хиймэл оюун ухааны аргуудыг ашиглан хүний дуу хоолойн өнгөнд суурилж, уушгины хатгалгааг эрт үед нь оношлох боломжийг судалж, хэрэгжүүлснээр эрүүл мэндийн үйлчилгээний хүртээмжийг нэмэгдүүлэх, эмнэлгийн байгууллагын ачааллыг бууруулах, иргэдийн цаг хугацаа болон зардлыг хэмнэхэд хувь нэмэр оруулах явдал юм.

Ижил төстэй судалгаануудын тойм судалгаанд [5] уушгины болон амьсгалын замын өвчнүүдийг дуу хоолойгоор оношлоход хэрэглэдэг янз бүрийн шинж чанар гаргаж авах (feature extraction) аргуудын хэрэглээг судалсан байна. Зураг 1-т харуулснаар MFCC (Mel-frequency cepstral coefficients) хамгийн түгээмэл хэрэглэгддэг арга бөгөөд тухайн өгүүллийн хүрээнд авч судалсан нийт судалгааны 25.5%-д хэрэглэгдсэн байна.



Зураг 1. Шинж чанаруудыг ялгах аргуудын хэрэглээний харьцуулалт

Энэхүү судалгаанд мөн шинж чанар гаргаж авах аргуудыг зөв сонгох нь уушги болон амьсгалын замын өвчнүүдийг дуу хоолойгоор ангилах, оношлоход маш чухал үүрэгтэй гэдгийг тэд тогтоосон байна. Түүнчлэн уг судалгаа аудио өгөгдөл дээр суурилсан оношлогооны аргуудыг голчлон хоёр чиглэлд ангилсан байна. Үүнд:

1. **Онцлог шинж дээр суурилсан уламжлалт машин сургалтын аргууд** – Хүний дуу хоолойн акустик онцлог шинжүүд болох MFCC, pitch, jitter, shimmer зэргийг гарган авч, эдгээрт тулгуурлан Random Forest, SVM зэрэг машин сургалтын загваруудыг сургадаг.

2. **Спектрограм дээр суурилсан гүн сургалтын аргууд** – Аудио өгөгдлийг Mel spectrogram

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

эсвэл raw waveform хэлбэрээр хувирган, түүнийг Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), эсвэл Transformer зэрэг гүн сургалтын сүлжээнд оруулан шууд дүгнэлт гаргах аргууд өргөн хэрэглэгдэж байна.

Дээрх хоёр чиглэлийн аль алинд нь аудио өгөгдлийн урьдчилсан боловсруулалт чухал үүрэгтэй бөгөөд шуугиан арилгах, дууны хэмжээ жигдлэх, хэсэгчлэх, онцлог шинжүүдийг гарган авах зэрэг үндсэн алхмуудыг багтаадаг. Өөрөөр хэлбэл амьсгалын замын өвчний оношлогоо хийхэд урьдчилсан боловсруулалтаас ангилал хүртэлх үе шатыг дамждаг. 1) *Урьдчилсан боловсруулалт*, энэ нь шаардлагагүй дуу шуугианыг арилгаж, дуу өгөгдлийг цаашдын шинжилгээнд бэлтгэх зорилготой бөгөөд аудио шүүлтүүр болон шуугиан багасгах техникийг ашигладаг. Түүнээс гадна цэвэрлэгдсэн өгөгдлөөс онцлог шинж чанарыг гаргаж авах шаардлагатай. 2) *Ангилал*, энэ нь гаргаж авсан онцлог шинжүүдийг ашиглан дууг янз бүрийн өвчний ангилалд хуваарилах үйл явц юм.

Иймд энэхүү судалгаагаар хүний дуу хоолойн өнгөнөөс уушгины хатгалгааны эрсдэлийг эрт шатанд оношлох боломж, ашиглах онцлог шинжүүд, тохиромжтой урьдчилсан боловсруулалтын арга зүй, машин сургалтын загваруудын гүйцэтгэлийг нарийвчлан судалж, нэвтрүүлэх боломжийг тодорхойлох зорилготой юм. Судалгааны хүрээнд уушгины хатгалгааг дууны өгөгдлөөс таамаглах зорилгоор машин сургалт болон гүн сургалтын дараах дөрвөн төрлийн загварыг сонгон ашигласан. Эдгээр загваруудыг сонгохдоо аудио өгөгдлийн онцлог болон шинж чанарт тулгуурласан үндэслэлийг харгалзсан болно. Үүнд:

- **Convolutional Neural Network (CNN)**: MFCC шинжүүд дээр орон зайн хамаарлыг харгалзан автоматаар онцлог шинжийг суралцдаг чадвартай. Дууны өгөгдлийн нарийн бүтэц, үл мэдэгдэх хамаарлуудыг илүү сайн илрүүлэх боломжтой тул судалгаанд оруулсан.

- **K-Nearest Neighbors (KNN)**: Бага хэмжээний өгөгдөл дээр үр дүн сайтай ажилладаг, мөн шинэ жишээг ойролцоох хөршүүдийн мэдээлэлд тулгуурлан ангилдаг энгийн боловч үр дүнтэй арга юм. Судалгаанд ашигласан өгөгдлийн хэмжээ хязгаарлагдмал байсан тул энэхүү загварыг сонгосон.

- **Support Vector Machine (SVM)**: Өндөр хэмжээст өгөгдөлд сайн тохирдог бөгөөд MFCC зэрэг олон хэмжээст шинжүүдийг ангилахад үр ашигтай. Мөн дууны онцлог шинжүүд nonlinear (шулуун бус) хамааралтай байдаг тул kernel trick ашиглан илүү нарийн загварчлах боломжийг олгодог.

- **Random Forest (RF)**: Шийдвэрийн модны ансамбль алгоритм болох RF нь шуугиантай, чанарын ялгаатай аудио өгөгдөлд өндөр тэсвэртэйгээр ажилладаг. Түүнчлэн чухал шинжүүдийг ялгах чадвартай учраас өгөгдлийн онцлог шинжүүдийг тайлбарлахад чухал үүрэгтэй байв.

Иймд дээрх дөрвөн загварыг сонгосноор дууны өгөгдлийн олон хэмжээст, шуугиантай, nonlinear шинжүүдийг үр дүнтэйгээр боловсруулж, уушгины хатгалгааг илүү өндөр нарийвчлалтайгаар таамаглах боломж бүрдсэн.

Үр дүнгээрээ хүний дуу хоолойн онцлог шинж чанарт тулгуурлан уушгины хатгалгааг эрт шатанд оношлох хамгийн оновчтой загварыг илрүүлж, түүнийг хэрэглэгчийн веб интерфэйстэй холбож, онош дээр үндэслэн зөвлөгөө зөвлөмж гаргана.

II. АШИГЛАСАН АРГА

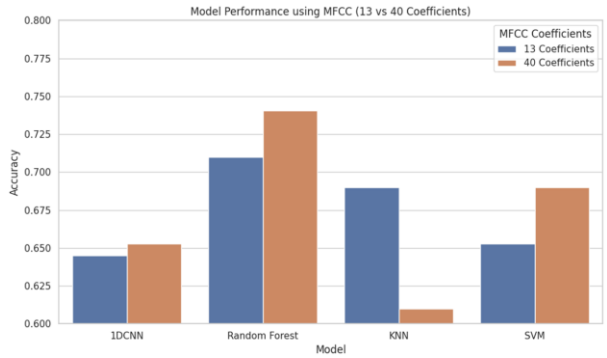
3.1 Өгөгдлийн багц

Судалгаанд нээлттэй эх сурвалжаас [6] зөвхөн судалгаа шинжилгээнд ашиглаж болохуйц нийт 1310 аудио өгөгдөл ашигласан ба үүнээс 453 нь уушгины хатгалгаа өвчтэй хүний, 857 нь эрүүл хүний дуу хоолойн бичлэг байв. Өгөгдлийн багц нь тухайн хүний дуу хоолойноос гадна уушгины хатгалгаа өвчинтэй холбоотой шинж чанаруудыг илтгэх 15 төрлийн аттрибут агуулсан. Эдгээр аттрибут нь нас, хүйс, шинж тэмдэг зэрэг дуу авианы болон мета мэдээллийг багтаадаг.

3.2 Өгөгдлийн урьдчилсан боловсруулалт

3.2.1 Шинж чанарыг ялгах (Feature extraction)

Аудио өгөгдлөөс чанартай, онцлог шинжүүдийг ялган авах нь загварын гүйцэтгэлийг шийддэг чухал алхам юм. Бид өөрсдийн судалгаанд MFCC аргын тохиргоонуудыг туршиж, хамгийн оновчтой параметрүүдийг тодорхойлохыг зорьсон. MFCC нь хүний сонсголын онцлогт тохирсон давтамжийн мэдээллийг илэрхийлдэг бөгөөд энэхүү аргаар аудио сигналын онцлог шинжүүдийг олон хэмжээст орон зайд хувиргаж өгдөг. Судалгаанд MFCC-ийн коэффицентийн тоог 13-40 хооронд өөрчилж туршилт хийсэн. Үүний дүнд, 13 болон 40 коэффицентээр задласан хувилбарууд дээр загварууд онцгой сайн үр дүн үзүүлсэн. Энэ нь Зураг 2-т харуулсан байна. Ялангуяа, 40 MFCC коэффицент ашигласан загварууд илүү нарийн ялгалт хийж, өндөр гүйцэтгэлтэй байсан нь онцлог шинжийн нарийвчлал нь эцсийн үр дүнд шууд нөлөөлдөг болохыг харуулж байна.



Зураг 2. Коэффициент харьцуулалт

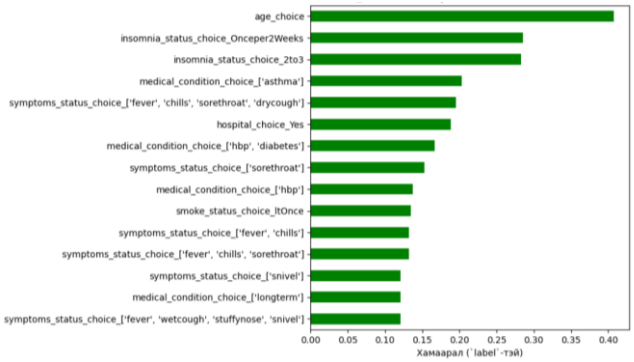
3.2.2 Өндөр хамааралтай аттрибутуудыг сонгох

Өмнөх шатанд зөвхөн аудио өгөгдөлд тулгуурлан MFCC онцлог шинжүүдийг сонгон ашигласан бол уушгины хатгалгаа өвчинг шинж тэмдэгүүд дээр нь үндэслэн илрүүлэх боломжийг [7] судалгаанаас авч, энэ удаа өгөгдлийн багцад багтсан уушгины хатгалгаатай холбоотой Хүснэгт 1-д харуулснаар

Шинж тэмдэг	Тайлбар
insomnia_status_choice['since2weeks']	2 долоо хоногийн хугацаанд нойргүйдэлтсэн
insomnia_status_choice['now']	Одоогийн үед нойргүйдэлт байгаа
medical_condition_choice['fever', 'chills', 'sorethroat']	Халуурах, дагжин чичрэх, хоолой өвдөх шинжүүд
hospital_choice['yes']	Эмнэлэгт очсон шинж тэмдэгтэй
medical_condition_choice['diabetes']	Чихрийн шижинтэй гэсэн түүхтэй
medical_condition_choice['heart']	Зүрхний өвчтэй
symptoms_status_choice['fever', 'chills', 'sorethroat']	Халуурах, дагжин чичрэх, хоолой өвдөх шинжүүдийг хамтад нь илэрсэн
symptoms_status_choice['fever', 'chills']	Халуурах байгаа, дагжин чичрэх шинж тэмдэгтэй
symptoms_status_choice['fever', 'coughwet', 'sneeze']	Халуурах байгаа, нойтон ханиалгах, найтаах шинж тэмдэгтэй
age_choice	Хэрэглэгчийн насны мэдээлэл
medical_condition_choice['asthma']	Астма өвчтэй эсэх
medical_condition_choice['hbp']	Өндөр даралттай эсэх (High Blood Pressure)
medical_condition_choice['others']	Бусад оношлогдсон өвчний түүхтэй эсэх
symptoms_status_choice['fingernem']	Хуруу мэдэгүй болох шинж тэмдэг байгаа эсэх

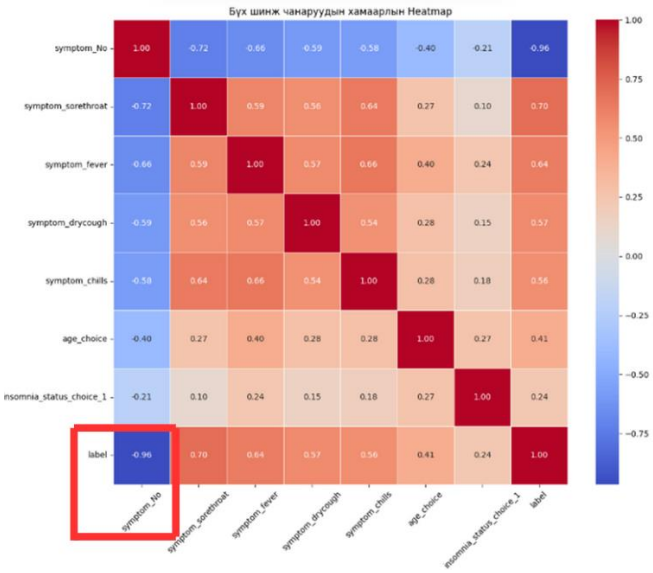
Хүснэгт-1. Аттрибутуудын тайлбар

Эдгээр нэмэлт аттрибутуудын хамаарлыг үнэлэн (Зураг 3), сургалтын өгөгдөлд нэмж ашиглах боломжийг судалсан. Үүний тулд корреляцийн шинжилгээ хийж, MFCC-ээр гаргасан 40 онцлог шинжүүдтэй хослуулан сургалт хийж загварын гүйцэтгэлд хэрхэн нөлөөлж буйг харьцуулсан болно.



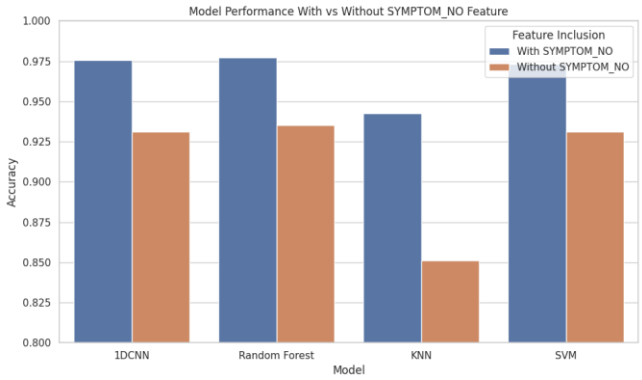
Зураг 3. Өндөр хамаарал бүхий аттрибутууд

Зорилтот атрибуттай энэхүү нэмэлт аттрибутууд хэр хамааралтайг илэрхийлэх хамаарлын матрицыг Зураг 4-т харуулсан ба шинжилгээний үр дүнд symptom_No буюу "шинж тэмдэг илрээгүй" гэсэн багана нь зорилтот атрибуттай 0.96-ийн шугаман хамаарал үзүүлж байв. Энэ нь тухайн хүн өвчтэй байсан ч өөрийгөө шинж тэмдэггүй хэмээн бүртгүүлсэн тохиолдолд, уг аттрибутын улмаас систем автоматаар "эрүүл" гэж ангилах өндөр магадлалтай болохыг харуулж байна. Энэ нь class imbalance буюу өгөгдлийн тэнцвэргүй байдалтай уялдан, загварын гүйцэтгэлд сөргөөр нөлөөлөх эрсдэлийг бий болгох эрсдэлтэйг харуулж байна.



Зураг 4. Аттрибутуудын хамаарлын матриц

Иймд энэхүү аттрибутыг сургалтын өгөгдлөөс хасч, түүний оролцоогүйгээр сургаж загварын гүйцэтгэлд үзүүлэх нөлөөг судалсан. Туршилтын дүнгээс харахад уг аттрибутыг хасахад (Зураг 5) загварын нарийвчлал 4–9%-иар буурсан ч, хуурмаг оношлох (false positive/false negative) эрсдэл багассанаар итгэмжтэй, бодит нөхцөлд илүү аюулгүй загвар үүсгэх боломжтой болсон.



Зураг 5. Шинж тэмдэг илрээгүй баганыг хассан үеийн загварын гүйцэтгэлийн харьцуулалт

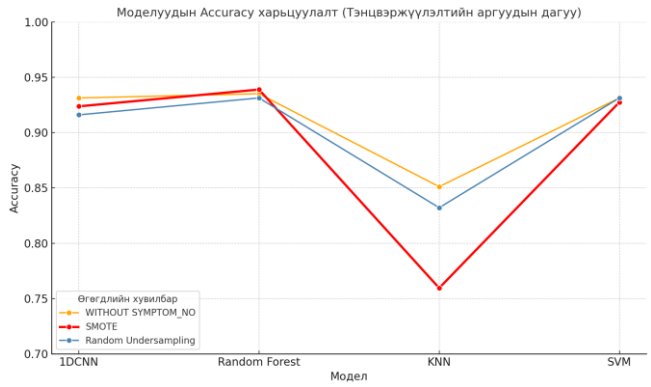
3.2.3 Өгөгдөл тэнцвэржүүлэлт

Өмнөх урьдчилсан боловсруулалтын шатанд өгөгдлийн тэнцвэргүй байдал (class imbalance) нь сургалтын загварын гүйцэтгэлд сөргөөр нөлөөлж, хуурмаг ангилал (false prediction) үүсгэж байгааг анзаарсан. Иймд өгөгдлийг тэнцвэржүүлэх зорилгоор өгөгдлийн багц болон аудио өгөгдөлд тус тусад нь тохирох аргуудыг туршин хэрэглэсэн. Өгөгдлийн багцын хувьд хоёр төрлийн тэнцвэржүүлэх аргыг туршсан.

- **Random undersampling** аргыг ашиглан өгөгдлийн дийлэнх ангиллыг бууруулахад, сургалтын өгөгдөл багассанаар моделийн гүйцэтгэл унаж, сөрөг нөлөө үзүүлсэн.

- Харин **SMOTE (Synthetic Minority Oversampling Technique)** аргыг хэрэглэн ховор ангилалд хиймэл өгөгдөл үүсгэснээр загварын гүйцэтгэл тогтвортой өсөлттэй байсан. Энэхүү арга нь feature-пространц дотор шинэ өгөгдлийг бодит өгөгдлийн орчны дагуу үүсгэх тул ангиллын ялгаа бууруулахгүйгээр өгөгдлийн тэнцвэрийг хангаж чадсан.

SMOTE-н үр дүнд загваруудын өвчтэй хүнийг зөв ангилах нь өсөж эхэлсэн ба харин Random undersampling хийж өгөгдлийг бууруулах нь загварын гүйцэтгэлд бүх талдаа муу үзүүлэлт үзүүлсэн. Үр дүнг Зураг 6-д харуулав.



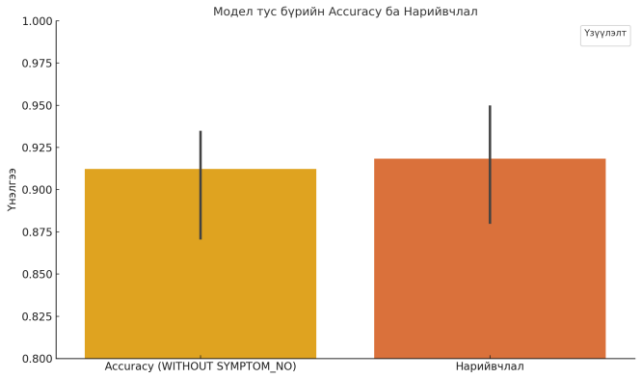
Зураг-6.SMOTE, UNDERSAMPLING

Аудио өгөгдлийн хувьд зөвхөн “bad” ангилал буюу өвчтэй хүний аудио өгөгдөлд rate=0.8 тохируулгатай augmentation хийсэн. Өөрөөр хэлбэл, тухайн аудио файлуудыг 20% удаашруулж, сургалтын өгөгдлийн 80%-д энэ өөрчлөлтийг хийснээр өгөгдлийг хиймлээр өсгөсөн. Ингэснээр:

- Өвчтэй хүнийг зөв ангилах магадлал (recall) нь 0.4-өөс 0.8 хүртэл өссөн.

- Моделийн ерөнхий гүйцэтгэл 2% хүртэл сайжруулалт гарсан.

Эдгээр туршилтууд нь тэнцвэргүй өгөгдөлтэй ажиллахад бидний хэрэглэсэн SMOTE болон аудио файлын хурдыг өөрчилж ховор ангиллын өгөгдлийн хэмжээг нэмэгдүүлсэн нь эдгээр аргууд үр дүнтэй болохыг баталж, практикт хэрэглэх боломжтойг харуулж байна. Нийт загваруудын дундаж өсөлтын үр дүнг Зураг 7-д харуулав.



Зураг 7. Нийт загваруудын дундаж өсөлтыг харуулав

3.3 Ашигласан загваруудыг ба сургалтын процесс

Өгөгдөл боловсруулалт, онцлог шинжийн ялгалт, тэнцвэржүүлэлтийн дараа бид хэд хэдэн машин сургалтын болон гүн сургалтын загваруудыг туршин харьцуулалт хийсэн. Судалгаанд дараах загваруудыг ашигласан:

- **Random Forest (RF)** – олон шийдвэрийн модыг ансамбл байдлаар ашиглаж, өгөгдлийг тогтвортой ангилах чадвартай. Өгөгдлийн тэнцвэргүй байдлын үед ч гүйцэтгэл өндөр байдаг учир нь олон янзын bootstrap өгөгдлөөр сурсан олон модноос нийлмэл шийдвэр гаргадаг.

- **Support Vector Machine (SVM)** – өгөгдлийг хамгийн оновчтой гипер хавтгайгаар ангилдаг, дуу авианы онцлог шинж дээр сайн гүйцэтгэл үзүүлдэг.

- **K-Nearest Neighbors (KNN)** – энгийн боловч хүчтэй статистик арга, ижил төстэй шинжүүд дээр суурилан ангилдаг.

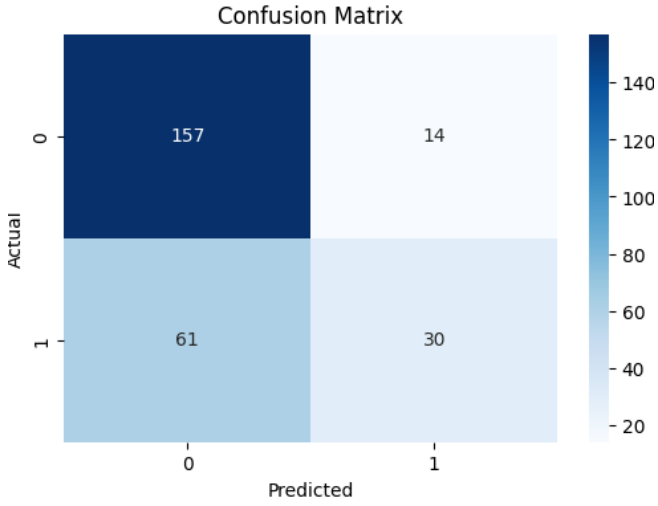
- **Convolutional Neural Network (CNN)** – дуу хоолойн өнгө, хүч, агаарын урсгалын хэлбэлзэл зэрэг аудио шинжүүдийг динамикаар боловсруулан ангилж, уушгины хатгалгааг оношлох зорилгоор хэрэглэгдсэн. MFCC feature-үүдийг ашигласнаар шууд аудио сигнал бус, сонсголын мэдрэмжтэй ойролцоо хэмжүүрүүдийг CNN-д оруулж сургасан.

CNN загварт early stopping, dropout, batch normalization зэрэг техникийг ашиглан хэт тохируулагдахаас (overfitting) сэргийлсэн. Загваруудыг TensorFlow, Keras, Scikit-learn зэрэг сангуудыг ашиглан

боловсруулж, librosa санг аудио өгөгдлийн шинж чанарыг гаргахад ашигласан.

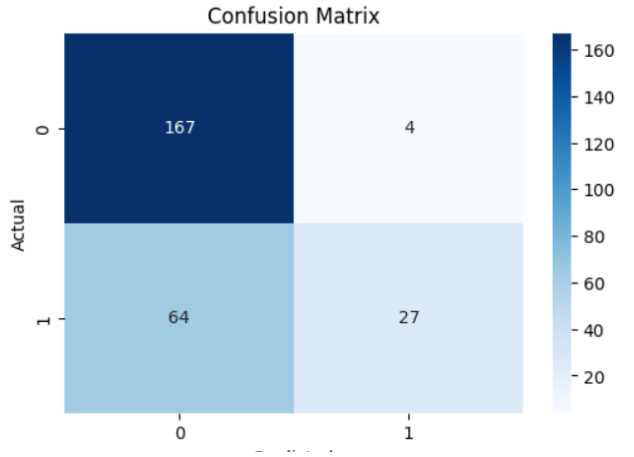
III. ҮР ДҮН

Бид уг сонгож авсан загварууд дээрээ өөрсдийн өгөгдлийн багц дээр туршиж дараах үр дүнгүүдийг гаргалаа. RF загварын гүйцэтгэлийг төөрөгдлийн матриц ашиглан Зураг 8-д харуулав.



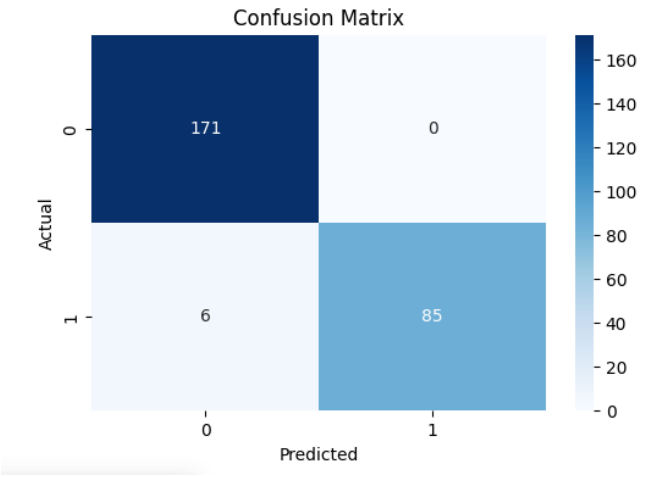
Зураг 8. MFCC 13 коэффициенттэй зөвхөн аудио өгөгдлийн хувьд

40 коэффициентээр буюу онцлог шинжийн нарийвчлалыг ихээр задалвал RF загварын хувьд тохиромжтой, гүйцэтгэл сайн байсан ба энэхүү үр дүнг Зураг 9-т харуулав.

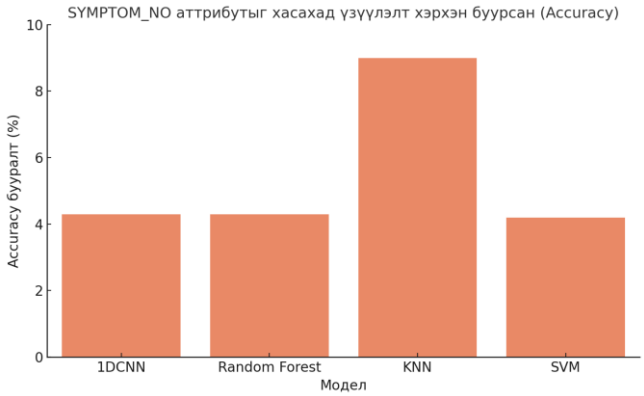


Зураг 9. 40 коэффициенттэй үед гүйцэтгэл

Өндөр хамаарал бүхий аттрибутуудыг сонгосноор загварын гүйцэтгэл (Зураг 10) нэмэгдэж байсан хэдий ч, Symptom_No аттрибутын хувьд хэт өндөр хамааралтай байснаар загвар тухайн шинжид хэт найдах хандлагатай болж бодит нөхцөлд буруу ангилалт хийх эрсдэл нэмэгдэж байв. Symptom_No аттрибутыг устгасны дараах бүх загваруудын гүйцэтгэл 4-9% буурсан (Зураг 11).

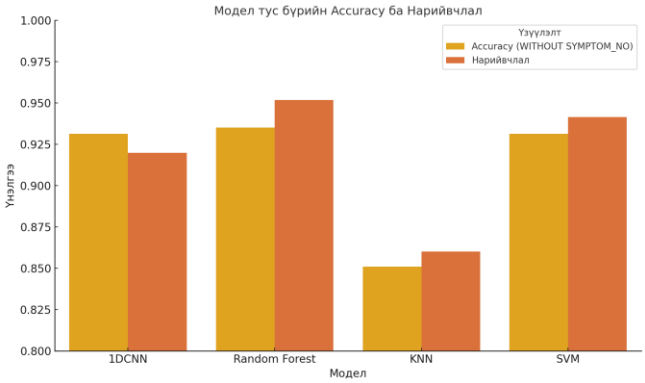


Зураг-10. Хамаарал бүхий аттрибутуудтай тохиолдлын үр дүн



Зураг-11. Symptom_No аттрибутгүйгээр загварын гүйцэтгэл 4-9% хүртэл буурав

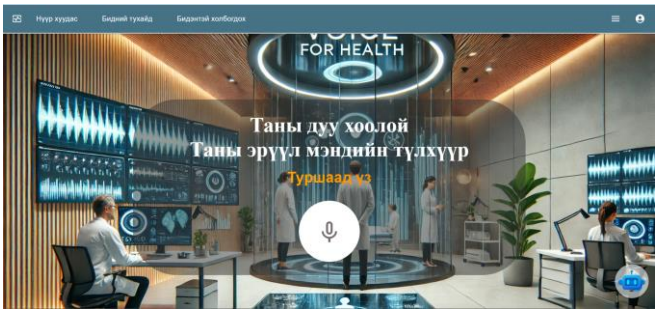
Уг аттрибутгүйгээр үргэлжлүүлэн оновчлолыг нэмэх нь бодит нөхцөлд оновчтой гэж үзэн нэмэлт туршилтуудыг хийв. Өгөгдлийн тэнцвэржүүлэх аргуудыг (SMOTE, UNDERSAMPLING, DATA AUGMENTATION) туршиж өгөгдлийг нэмэгдүүлж өгснөөр ангиллын гүйцэтгэлд эерэг нөлөө үзүүлж байв. Энэхүү үр дүн Зураг 12-т үзүүлэв.



Зураг-12.Загваруудын харьцуулалт

Судалгааны туршилтын шатанд урьдчилсан боловсруулалт хийгдсэн өгөгдлийг ашиглан нийт дөрвөн төрлийн загварыг сургаж, тэдгээрийн гүйцэтгэлийг харьцуулан шинжилсэн. Туршилтуудыг ижил өгөгдлийн хуваарилалт, ижил гиперпараметрийн тохиргоо дор явуулсан.

Ийнхүү сонгогдсон хамгийн үр дүнтэй загварыг бодит нөхцөлд турших зорилгоор веб интерфэйс боловсруулж, системд холбосон болно (Зураг 13).



Зураг-13. Веб интерфэйс

ДҮГНЭЛТ

Уушгины хатгалгаа өвчнийг эрт шатанд нь илрүүлэхэд хүний дуу хоолойн мэдээллийг ашиглах боломжтой эсэх, мөн Монголын эрүүл мэндийн салбарт тулгарч буй хүндрэлүүд — тухайлбал эмнэлгийн ачаалал, цаг авахад зарцуулдаг хугацаа зэргийг хиймэл оюуны тусламжтайгаар хэрхэн бууруулах талаар бид энэхүү судалгаанд анхаарч ажиллалаа. Судалгааны зорилго нь иргэдэд өндөр нарийвчлалтай, оновчтой оношилгоо, зөвлөгөөг дуу хоолойн өгөгдөлд тулгуурлан өгөх боломжийг тогтоох явдал байв. Хийсэн судалгаа, шинжилгээний үр дүнгээс харахад хэрвээ хэрэглэгч өөрийн мэдээллийг үнэн зөв, нягт нямбай оруулсан тохиолдолд хөгжүүлсэн систем нь өндөр нарийвчлалтай гүйцэтгэл үзүүлэх боломжтой байна.

Судалгаанд нийт 1310 аудио өгөгдөл (өвчтэй болон эрүүл хүмүүсийн ханиалгын дуу) ашигласан бөгөөд өгөгдлийг augmentation хийж, нийт өгөгдлийн 80%-д хиймэл аудио үүсгэн өгөгдлийг тэнцвэржүүлсэн. Үлдсэн 20%-ийг тестийн зориулалтаар ашиглаж, KNN, SVM, CNN, RF зэрэг загваруудыг туршсан. Эдгээрийн дундаас Random Forest (RF) загвар хамгийн өндөр нарийвчлалтай буюу 0.95-тай гарч, өвчтэй хүнийг зөв таних recall 0.8 хүртэл өссөн үзүүлэлттэй байлаа.

Мөн шинжилгээнд ашигласан шинж тэмдгийн аттрибутууд (metadata) нь загварын гүйцэтгэлд мэдэгдэхүйц эерэг нөлөө үзүүлж байсныг тогтоосон. Иймд хэрэглэгч талын өгөгдлийн үнэн бодит байдал нь системийн нарийвчлал, оновчтой шийдвэр гаргалтад хамгийн чухал хүчин зүйл болохыг харуулж байна.

НОМ ЗҮЙ

[1] Pneumonia Detection: An Efficient Approach Using Deep Learning. Saniya P.M., et al. (2024)

[2] Respiratory Diseases Diagnosis Using Audio Analysis and Artificial Intelligence: A Systematic Review . Kapetanidis., et al. (2024)

[3] AI Based Diagnosis of Pneumonia . B Vidhya., M Nikhil Madhav., et al. (2022)

[4] Deep Learning Approaches For Lung Disease Detection Through Voice Analysis . Dr.Y.Sowjanya Kumari., et al. (2024)

[5] Lung disease recognition methods using audio-based analysis with machine learning . Ahmad H.Sabry., Omar I. Dallal Bashi , et al(2024)

[6] Sound-Dr: Reliable Sound Dataset and Baseline Artificial Intelligence System for Respiratory Illnesses. Truong V, Hoang., et al(2022)

[7] The use of ai systems in diagnosis of pneumonia via signs and symptoms: Systematic review ., Katy Stokes ., et al .(2022)

[8] Our world in Data , Pneumonia., Bernadeta Dadonaite and Max Roser., (2024)

Object Pose Estimation for Random Bin Picking Using Synthetic Data

Munkhtulga Byambaa
Department of Electronics and Communication Engineering
School of Information Technology and Electronics
Ulaanbaatar, Mongolia
https://orcid.org/0000-0002-3560-9583

Lodoiravsal Choimaa
Department of Electronics and Communication Engineering
School of Information Technology and Electronics
Ulaanbaatar, Mongolia
https://orcid.org/0000-0002-1773-1059

Gou Koutaki
Faculty of Engineering
Kumamoto University
Kumamoto City, Japan
https://orcid.org/0000-0002-3414-1085

***Abstract*— With the rapid growth of industrial automation, estimating the 3D pose of objects has become a central challenge in random bin picking—particularly in cluttered environments with occlusions and a variety of object types. One major hurdle in developing and testing such systems is the reliance on industrial robot arms, which can cost up to $100,000 and are often unsuitable for research labs due to safety, space, and movement constraints. To overcome this, we built an open-source, 3D-printed robotic arm with six degrees of freedom. We integrated it into the ROS MoveIt framework using a URDF model, enabling inverse kinematics and motion planning. Our system uses a deep neural network trained entirely on synthetic data for 3D object pose estimation. This approach allows for fast and scalable generation of large, labeled training datasets. The network outputs belief maps and vector fields from input images; peaks are detected in the belief maps, keypoints are assigned via the vector fields, and the final object pose is computed using the Perspective-n-Point (PnP) algorithm. Using this method, we demonstrate a low-cost, real-time random bin-picking system that accurately estimates the poses of known assembly line objects. Project link: https://github.com/bmunkhtulga/rbx1-ros**

Keywords— Robotics, object pose estimation, synthetic data, random bin picking

I. INTRODUCTION

Recognizing objects and estimating their poses in 3D has a wide range of applications in robotic tasks. For instance, recognizing objects' 3D location and orientation is important for robot manipulation. It is also useful in human-robot interaction tasks such as learning from demonstration. However, the problem is challenging due to the variety of objects in the real world. They have different 3D shapes, and their appearances on images are affected by lighting conditions, clutter in the scene, and occlusions between objects.

There are two common approaches for 3D object pose estimation.

- CAD-based object pose estimation. The problem of 3D object pose estimation is tackled by matching feature points between 3D models and images [9, 12, 13]. However, these methods require that there are rich textures on the objects in order to detect feature points for matching. As a result, they are unable to handle texture-less objects. For template-based methods [10, 11], occlusions significantly reduce the recognition performance.
- CNN-based object pose estimation. While deep neural networks have been successfully applied to the problem of object detection in 2D [1, 2, 3], they have only recently begun to be applied to 3D object detection and pose estimation [4, 5, 6]. Recently, several deep learning-based approaches have appeared that operate directly on RGB images [4, 5, 6, 14], showing improved performance with respect to difficulties such as occlusion. Of these, BB8 [4] and PoseCNN [5] first segment the objects and then estimate the object pose, whereas the approaches of Tekin et al. [6] and Kehl et al. [14] rely on the single-shot networks of YOLO and SSD, respectively. All of these networks were trained on real data, thus limiting their ability to generalize to new scenarios (e.g., lighting conditions).

In this work, our goal is to estimate the 3D pose of known objects in the clutter from a single RGB image in real time for the purpose of grasping such objects with a robot manipulator. Unlike 2D object detection, it is difficult and time-consuming to manually label data for 3D detection.

Synthetic data is a promising alternative for training such deep neural networks, capable of generating an almost unlimited amount of pre-labeled training data with little effort. Synthetic data comes with its own problems, however. Chief among these is the reality gap, that is, the fact that networks trained on synthetic data usually do not perform well on real data without additional fine-tuning or other tricks.

A recently proposed solution to this problem is domain randomization (DR) [7], in which the training data is randomized in non-realistic ways so that, at test time, real data appears to the network as simply another variation. We

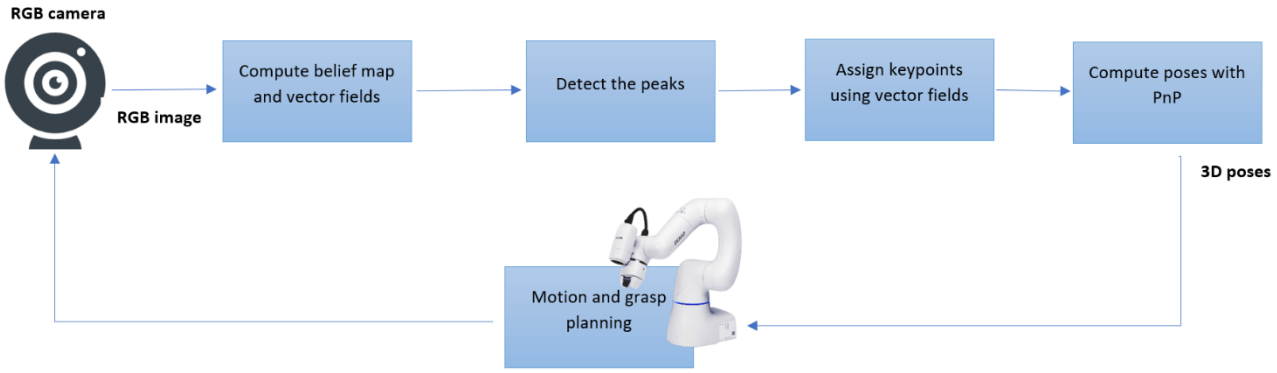


Figure 1. Overview of random bin picking system

think that a simple combination of DR data with photorealistic data yields sufficient variation and complexity to train a deep neural network that is then able to operate on real data without any fine-tuning.

II. SYSTEM MODEL

The flowchart of the proposed system is shown in Fig. 1. Our approach can be divided into five stages: (a) compute the belief map and vector fields, (b) detect the peaks of the belief map, (c) for each centroid assign keypoints using vector fields, (d) standard perspective-n-point (PnP) algorithm, and (e) robotic manipulation.

A. *Network architecture*

We used the DOPE (Deep Object Pose Estimation) for the network [17]. The DOPE network is a one-shot, fully convolutional deep neural network that detects key points using a multistage architecture. The feedforward network inputs an RGB image of size w × h × 3 and branches to produce two different outputs: belief maps and vector fields. There are nine belief maps, one for each of the projected 8 vertices of the 3D bounding boxes and one for the centroids. Similarly, there are eight vector fields indicating the direction from each of the 8 vertices to the corresponding centroid, similar to [5], to enable the detection of multiple instances of the same type of object. (In experiments, w = 640, h = 480.)

B. *Robot Arm Design*

For this work, we have used two kinds of robot arms to compare our newly 3D-printed robot arm's grasping accuracy. Industrial robot arms can cost up to a hundred thousand dollars. Also, it is unsafe to experiment in a research laboratory due to constraints of movements and workspace range. So, we made open source 3D-printed robot arm from scratch. It is called RBX1 Remix, which has 6 DOF and 60 cm reachability. The robot arm is shown in Fig. 2; it took us approximately 12 days or 115 hours to print. Fig. 2 shows 3D-printed parts and assembled robot. Denavit-Hartenberg parameters are given in Table 1. Using the table, the homogeneous transform matrix of RBX1 can be obtained which describes the relation between the end-effector coordinate system and the global coordinate system.

(a) 3D printed parts. *(b) After assembly*

Figure 2. RBX1 Remix open source 6 DOF robot arm

Table 1. RBX1 D-H parameters

i	a_i	α_i	d_i
1	0	0	L1
2	0	π/2	0
3	L2	0	0
4	0	π/2	L3
5	0	- π/2	0
6	0	π/2	L4

The dynamic equation of the robot manipulator with n-DOF is as follows,

$$D(q)\ddot{q} + C(q,\dot{q})\dot{q} + g(q) = \tau \qquad (1)$$

where $q \in R^n$ is the coordinate vector of the robot joint, $D(q) \in R^{n\times n}$ is the inertia matrix, $C(q,\dot{q}) \in R^{n\times n}$ is Coriolis matrix, $g(q) \in R^n$is the gravity vector, and $\tau \in R^n$ is the generalized input vector [11].

After loading the URDF model into ROS MoveIt, we solved inverse kinematics. We did motion planning, as shown in Fig. 3. Below, we can see motion trajectories being planned from an upright position (left) to a random valid position (right).

The second robot arm is Denso COBOTTA. COBOTTA is a 6 DOF and human collaborative robot arm with 34 cm reachability. It's smaller than RBX1 Remix but has much more capability. To evaluate the accuracy and performance

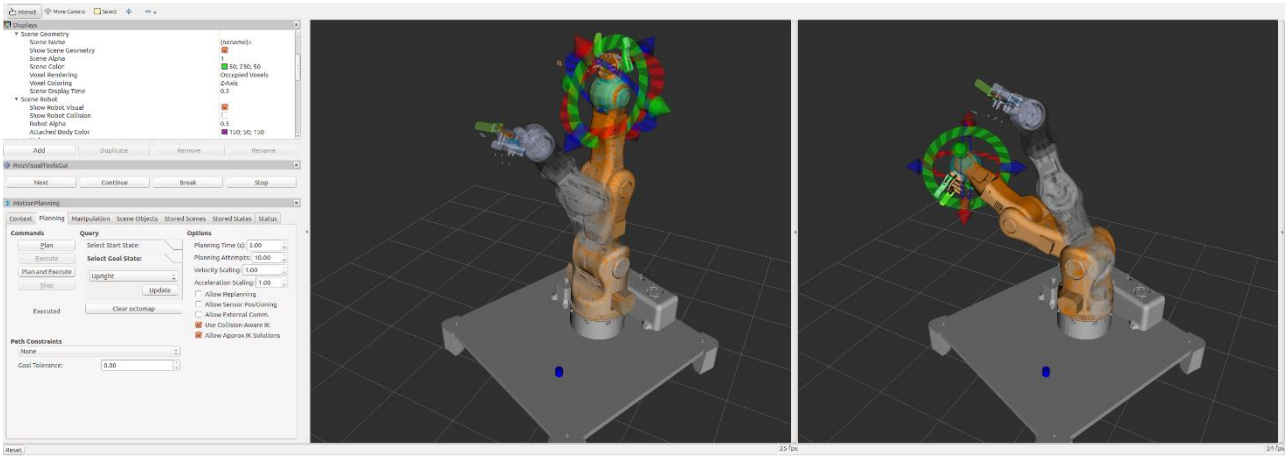


Figure 3. Motion trajectories in ROS MoveIt

of our 3D-printed robotic arm, we conducted a comparative analysis using a commercial industrial robot.

C. *Robot Coordinate Calibration*

We used the eye-to-hand camera system, as shown in Fig. 4, to capture the scene images under the camera coordinates. Therefore, to make the robot arm adopt the RGB camera information, the relationship between the camera coordinate system and the robot arm's position has to be determined. For that matter, we used the pinhole model. where (X, Y, Z) are the coordinates of a 3D point in the world coordinate space, (u, v) are the coordinates of the projection point in pixels, A is a camera matrix, or a matrix of intrinsic parameters (c_x, c_y) is a principal point that is usually at the image center and f_x, f_y are the focal lengths expressed in pixel units.

$$s\ m' = A[R|t]M'$$

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

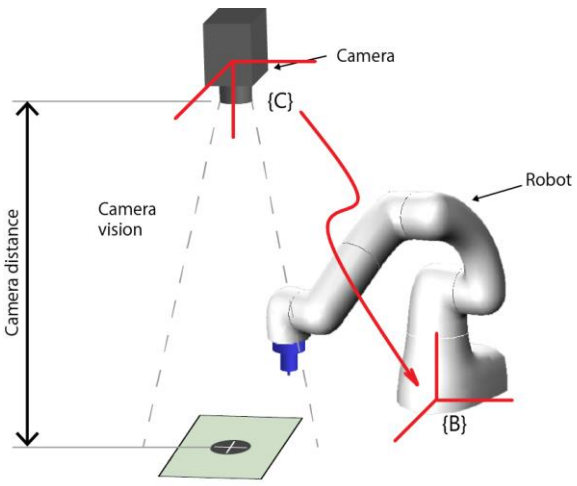


Figure 4. Coordinate relationship between the camera and COBOTTA

III. EXPERIMENTAL RESULTS

First, we did preliminary experiments with household objects to estimate the poses of assembly line objects. In this experiment, we used the YCB dataset to train the network.

A. *Training*

For training, we chose two objects from the YCB dataset (banana (011) and tomato soup can (005)). We used ~48k domain-randomized image frames mixed with ~12k photorealistic images. For data augmentation, Gaussian noise (σ = 2.0), random contrast (σ = 0.2) and random brightness (σ = 0.2) were added. The DOPE network was implemented using PyTorch v1.0 [15]. The VGG-19 feature extractions were taken from publicly available trained weights in *torchvision* open models. The networks were trained for 60 epochs with a batchsize of 32. Adam [16] was used as the optimizer with a learning rate set at 0.0001. The system was trained with NVIDIA V100 GPU 16GB. Some examples of images of training data are shown in Fig. 5.

Figure 5. Some example images of training data. (Left: RGB, Depth, Segmentation)

B. *Evaluation metrics*

For evaluating 3D pose estimation, the Average Distance (ADD) metric is commonly used [17], [18], [19]. It measures the mean 3D Euclidean distance between the model points transformed by the predicted pose and those transformed by the ground truth pose. Specifically, given the ground truth rotation **R** and translation **T**, and the estimated rotation $\tilde{\mathbf{R}}$ and translation $\tilde{\mathbf{T}}$, ADD calculates the average distance between corresponding 3D points after applying both transformations.:

$$\text{ADD} = \frac{1}{m}\sum_{x\in \text{M}}\left\|(\mathbf{Rx}+\mathbf{T})-(\tilde{\mathbf{R}}\mathbf{x}+\tilde{\mathbf{T}})\right\|, \qquad (2)$$

where M denotes the set of 3D model points, and m is the number of points. A 6D pose estimate is considered accurate if the computed average distance falls below a specified threshold.

C. *Results*

Given our focus on robotic manipulation, we restricted our evaluation to images of YCB objects, which are also the primary focus of the state-of-the-art 6-DoF pose estimation method, PoseCNN [5]. We compare our approach against PoseCNN using the YCB-Video dataset. Notably, PoseCNN has been shown to outperform the single-shot method proposed by Tekin et al. [6], which in turn surpasses the more complex architectures of BB8 [4] and SSD-6D [26] on standard benchmarks such as LINEMOD [23] and OccludedLINEMOD [30].

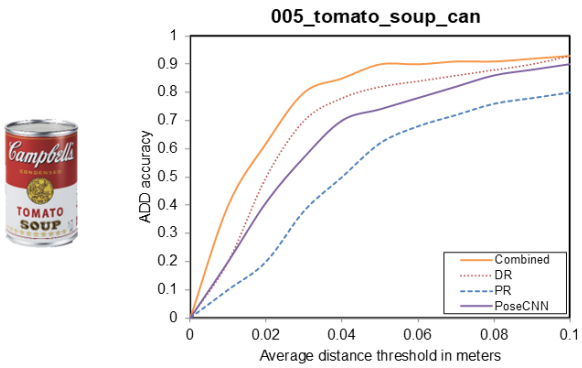


Figure 6. Accuracy-threshold curve for our approach compared with PoseCNN [5] on the tomato soup can (005) of the YCB dataset. The results display different variants of our approach trained exclusively on domain-randomized data (DR), solely on photorealistic data (PR), and on a combination of both (Ours (combined)). The X-axis represents the accuracy threshold in meters, while the Y-axis indicates the ADD accuracy.

Fig. 6 and Table. 2 present the accuracy-threshold curves for the tomato soup can (005) object, evaluated using the ADD metric, comparing our method with PoseCNN. We report results for three variants of our approach, each trained on a different dataset: domain-randomized data only (~48k synthetic images), photorealistic synthetic data only (~12k images), and a combination of both (~60k images). The results demonstrate that our lightweight network architecture, trained exclusively on synthetic data, achieves performance comparable to the state-of-the-art PoseCNN method. Fig. 7 shows examples of the 6D pose estimation results on the two objects in the YCB dataset.

Table 2. The area under the accuracy-threshold curve for 3D pose evaluation on the YCB dataset.

Objects	PoseCNN	DR	PR	DR+PR
005_tomato_soup_can	66.0	63.22	52.73	**75.69**
011_banana	72.3	69.50	58.39	**82.51**
Mean	69.1	66.3	55.5	**79.1**

D. *Robotic manipulation*

In our concluding experiment, we tested the effectiveness of our method by grasping real objects using a robotic arm. For this experiment, we used an RBX1 3D-printed robot arm, a Denso COBOTTA robot with a custom-made wide gripper for parallel-jaw grasping, and a Logitech C920 camera fixed above. The picking system setup is shown in Fig. 4. Two objects from the YCB dataset were used to evaluate grasping. The objects were placed in front of the robot at different locations and poses. Grasping was considered successful if the robot could pick it up and place objects in the box beside the robot. We conducted 15 trials for every object for each robot arm to test the proposed method, totaling 60 attempts. The results are presented in Table. 3. The grasping success rate for RBX1 is 83.3% which demonstrates that the 3D-printed robot arm is robust enough for real-world robotic manipulation.

Table 3. Results of grasping experiment (successful rate [%]). 15 trials for each object using each robot arm, totaling 60 attempts. The grasping success rate for RBX1 is 83.3% which demonstrates that the 3D-printed robot arm is robust enough for real-world robotic manipulation.

Object	RBX1	Cobotta
005_tomato_soup_can	76.6	83.3
011_banana	90.0	96.6
mean	83.3	89.9

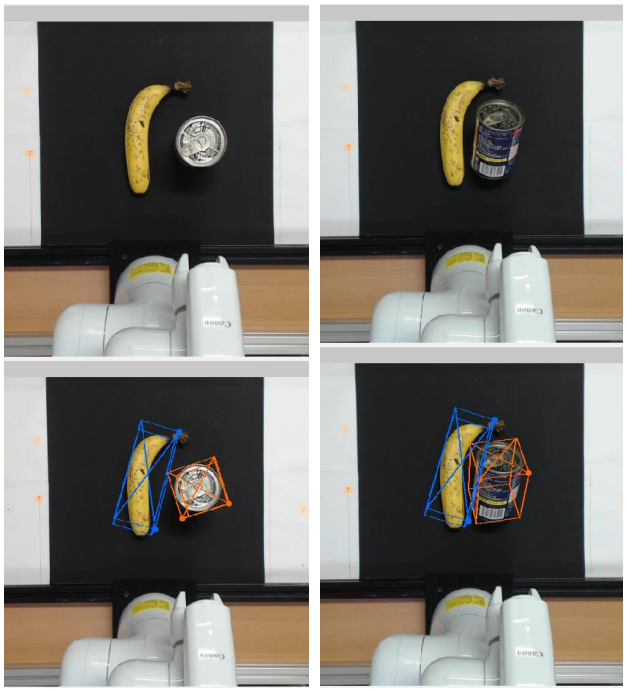


Figure 7. Examples of 6D object pose estimation results on the two objects of the YCB dataset.

IV. CONCLUSION AND FUTURE WORK

We have presented a system for detecting and estimating the 3D pose of known objects using a deep network with synthetic data. The network employs multiple stages to refine ambiguous estimates of the 2D locations of projected vertices of each object's 3D bounding cuboid. These points are then used to predict the final pose using PnP, assuming known camera intrinsic and object dimensions. We have shown that a network trained only on synthetic data can estimate poses with sufficient accuracy for robotic manipulation.

For robotic manipulation experiments, we utilized two different robotic arms: a 3D-printed robotic arm and a commercial industrial robotic arm. After conducting 30 trials with each, the 3D-printed arm (RBX1) achieved a grasping success rate of 83.3%, indicating that it is sufficiently robust for real-world robotic manipulation tasks.

Further research should be aimed at increasing the number of objects, handling symmetry, and incorporating closed-loop refinement to increase grasp success. Most importantly, we will try to tackle transparent and metallic objects that are not getting attention. Metallic objects are the most common objects in assembly lines; also, due to the characteristics of the material, extreme lighting conditions affect much more than other materials.

REFERENCES

[1] J. Dai, Y. Li, K. He, and J. Sun. R-FCN: Object detection via region-based fully convolutional networks. In NIPS, 2016.

[2] J. Redmon and A. Farhadi. YOLO9000: Better, faster, stronger. In CVPR, 2017.

[3] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg. SSD: Single shot multibox detector. In ECCV, 2016.

[4] M. Rad and V. Lepetit. BB8: A scalable, accurate, robust to partial occlusion method for predicting the 3D poses of challenging objects without using depth. In ICCV, 2017.

[5] Y. Xiang, T. Schmidt, V. Narayanan, and D. Fox. PoseCNN: A convolutional neural network for 6D object pose estimation in cluttered scenes. In RSS, 2018.

[6] B. Tekin, S. N. Sinha, and P. Fua. Real-time seamless single shot 6D object pose prediction. In CVPR, 2018.

[7] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel. Domain randomization for transferring deep neural networks from simulation to the real world. In IROS, 2017.

[8] V. Lepetit, F. Moreno-Noguer, and P. Fua. EPnP: An accurate O(n) solution to the PnP problem. International Journal Computer Vision, 81(2), 2009

[9] J. Hill and W. Park, "Real time control of a robot with a mobile camera", Proceedings of the 9th International Symposium on Industrial Robots, pp. 233–246. 1979

[10] A. Sanderson and L. Weiss, "Image-based visual servo control using relational graph error signals", Proc. IEEE, pp. 1074–1077, 1980

[11] D. Kragic and H.I. Christensen, "Survey on visual servoing for manipulation", Royal Institute of Technology, 2002

[12] E. Malis, "Survey of vision-based robot control", presented at the Eur. Naval Ship Des. Captain Comput. IV Forum, 2002.

[13] F. Chaumette and S.Hutchinson, "Visual servo control, Part 1: Basic approaches", IEEE Robotics and Automation Magazine, vol 13, no. 4, pp. 82-90, December 2006

[14] C. Lazăr and A. Burlacu, "Image-Based Visual Servoing for Manipulation Via Predictive Control–A Survey of Some Results", Memoirs of the Scientific Sections of the Romanian Academy 39. pp 71-81, 2016

[15] SM Baek, KS Lim, WG Han, and TY Kuc, "Dynamic Visual Servoing of Robot Manipulator", The Transaction of KIEE, D49.1 pp. 41-47, 2000 (in Korean)

[16] Hashimoto, Koichi, et al. "Manipulator control with image-based visual servo." Robotics and Automation, 1991. Proceedings., 1991 IEEE International Conference on. IEEE, 1991.

[17] Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards real-time object detection with region proposal networks. In: Advances in Neural Information Processing Systems (NIPS). (2015)

[18] Mellado, N., Aiger, D., Mitra, N.J.: Super 4pcs fast global pointcloud registration via smart indexing. Computer Graphics Forum 33(5) (2014) 205–215

[19] S. Bae, E. Kim, S. Yang, J. Park and T. Kuc, "A dynamic visual servoing of robot manipulator with eye-in-hand camera," *2018 International Conference on Electronics, Information, and Communication (ICEIC)*, Honolulu, HI, 2018, pp. 1-4.

[20] Carlos Hernandez, Mukunda Bharatheesha, Wilson Ko, Hans Gaiser, Jethro Tan, Kanter van Deurzen, Maarten de Vries, Bas Van Mil, Jeff van Egmond, Ruben Burger, Mihai Morariu, Jihong Ju, Xander Gerrmann, Ronald Ensing, Jan Van Frankenhuyzen, and Martijn Wisse, "Team Delft's Robot Winner of the Amazon Picking Challenge 2016"

[21] Alvaro Collet, Manuel Martinez, and Siddhartha S Srinivasa. The MOPED framework: Object recognition and pose estimation for manipulation. International Journal of Robotics Research (IJRR), 30(10):1284–1306, 2011.

[22] Stefan Hinterstoisser, Cedric Cagniart, Slobodan Ilic, Peter Sturm, Nassir Navab, Pascal Fua, and Vincent Lepetit. Gradient response maps for real-time detection of textureless objects. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 34(5):876–888, 2012.

[23] Stefan Hinterstoisser, Vincent Lepetit, Slobodan Ilic, Stefan Holzer, Gary Bradski, Kurt Konolige, and Nassir Navab. Model based training, detection and pose

estimation of texture-less 3D objects in heavily cluttered scenes. In Asian Conference on Computer Vision (ACCV), pages 548–562, 2012.

[24] David G Lowe. Object recognition from local scale-invariant features. In IEEE International Conference on Computer Vision (ICCV), volume 2, pages 1150–1157, 1999.

[25] Fred Rothganger, Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce. 3D object modeling and recognition using local affine-invariant image descriptors and Multiview spatial constraints. International Journal of Computer Vision (IJCV), 66(3):231–259, 2006.

[26] W. Kehl, F. Manhardt, F. Tombari, S. Ilic, and N. Navab. SSD-6D: Making RGB-based 3D detection and 6D pose estimation great again. In ICCV, pages 1521–1529, 2017.

[27] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer. Automatic differentiation in PyTorch. In NIPS Autodiff Workshop, 2017.

[28] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In ICLR, 2015.

[29] Jonathan Tremblay, Thang To, Balakumar Sundaralingam, Yu Xiang, Dieter Fox, Stan Birchfield. Deep Object Pose Estimation for Semantic Robotic Grasping of Household Objects. In CoRL, 2018

[30] E. Brachmann, A. Krull, F. Michel, S. Gumhold, J. Shotton, and C. Rother. Learning 6D object pose estimation using 3D object coordinates. In ECCV, 2014.

Хуулийн хэлний магадлалт биграмм загвар

Г.Чинбаяр
Бүртгэл, Программ хангамжийн тэнхим
Отгонтэнгэр Их Сургууль
swe24e103@otu.edu.mn

Ц.Ганбат
Бүртгэл, Программ хангамжийн тэнхим
Отгонтэнгэр Их Сургууль
0000-0003-0736-052X

***Хураангуй*— Эрх зүйн мэдээллийн нэгдсэн системээс Монгол улсын одоо хүчин төгөлдөр мөрдөгдөж буй хуулиудаас 6 хууль сонгож магадлалд суурилсан хуулийн униграмм ба биграмм загварыг танилцуулна. Хууль санаачлагч, хуулийн төсөл боловсруулагч, хуульч, шүүгч, өмгөөлөгч, иргэд хуулиас хайлт хийх, хуулийн төслийг боловсруулахад ашиглаж болно.**

***Түлхүүр үг*— цахим сан, магадлал, токен, зөвлөмж үг, алгоритм, униграмм**

I. Судлагдсан байдал

Хэлний цахим сан байгуулсан хэдэн судалгаа байдаг. 2011 онд МУИС-ийн Ж.Пүрэв, Ч.Алтангэрэл нарын бүрдүүлсэн хөмрөг уран зохиол, хууль, хэвлэл, сонингийн үгтэй [1], 2012 онд Krylov хүний гаргасан General Corpus of Modern Mongolian Language (GCML) роман, шүлэг, нууц товчоо, сонингийн үгтэй, Өвөр Монголын Давхарбаяр, Баярмэнд нарын Corpus of Modern Written Mongolian (CMWM) 10 сонин, уран зохиолын боть ном, улс төрийн боть ном, ЕБС-ын сурах бичгийн үгтэй, A.Kilgarriff монголын веб хуудсанд буй үгсээс тус тус бүрдсэн байсан.

2017 онд хэл шинжлэл талаас Д.Боролзой, М.Энхжаргал нарын сонингийн үгийн давтамжийн судалгаа [2], 2022 онд Т.Батбямба 5-р ангийн монгол хэлний сурах бичгийн үгийн давтамжийн судалгаа, 2020 онд Ж.Ган-Эрдэнэ, Б.Сэр-Од нар Боловсролын үнэлгээний Төвийн даалгаврын сангийн судалгаа [3], 2012 онд А.Хүдэр Zipf-ын хуулийн тогтмолын судалгаа [4], 2023 онд Чой Сон Сү үндсэн хуулийн бичвэрийн судалгаа [5] тус тус хийжээ.

II. Хуулийн хэлний униграмм цахим сан байгуулсан нь

Эрх зүйн мэдээллийн нэгдсэн системээс цуглуулсан үгс бол тоо, текст агуулсан боловсрогдоогүй, түүхий өгөгдөл юм. Харин хуулийн үгийн сан судалгааны тодорхой зорилгын дагуу компьютероор боловсруулагдсан, тэмдэглээ хийгдсэн, шинж чанар, холбоо хамаарлыг тодорхойлсон цаашид хэрэглээнд хэрэглэж болохуйц хэлний бичвэрийн цахим сан юм.

Эрх зүйн мэдээллийн нэгдсэн систем нь 2024 оны 01 сарын 08 өдрийн байдлаар Монгол улсын 934 хууль байдгаас Авлигын эсрэг хууль, Хөдөлмөрийн хууль, Хөдөлмөрийн аюулгүй байдал, эрүүл ахуйн тухай хууль, Боловсролын ерөнхий хууль, Дээд боловсролын тухай, Сургуулийн өмнөх болон ерөнхий боловсролын тухай хууль [6] гэсэн 6 хуулийг сонгон авч хуулийн униграмм, биграмм үгийн цахим санг байгуулсан.

Хуулийн бичвэрийн зарим онцлогийг дурдвал өгүүлбэрийн эхлэл том үсгээр эхэлж араас жижиг үсэг, тоо, текст хаалт (“) агуулсан. Өгүүлбэрийн төгсгөлүүд нэг цэг (.), дараалсан 3 цэг (…), асуултын тэмдэг (?), анхаарлын тэмдэг (!), цэгтэй таслал (;) төгссөн байв. Өгүүлбэр дотор дээр доор (:) цэгүүд агуулсан байв. Байгуулсан цахим сангийнхаа статистикийг хүснэгт 1-д харуулав.

Хүснэгт 1. Цахим сангийн статистик

№	Хууль	Нийт токен	Нийт үг	Ялгаатай үг	Өгүүлбэрийн тоо
1	Авлигын эсрэг	5675	5535	1354	325
2	Хөдөлмөрийн	20341	19874	2540	1219
3	Хөдөлмөрийн аюулгүй байдал, эрүүл ахуйн тухай	5837	5758	1198	358
4	Боловсролын ерөнхий	9241	9071	1650	562
5	Дээд боловсролын	4562	4428	1127	325
6	Сургуулийн өмнөх болон ерөнхий боловсролын тухай	4183	4046	1007	275
	Нийт	49839	48712	8876	

Хамгийн олон токентэй хууль нь Хөдөлмөрийн тухай хууль. Нийт 91 токентой байв.

Хуулийн униграмм үгийн цахим санг байгуулахад эдгээр 6 хууль давтагдсан тоогоор 48712 үгтэй, эдгээрийг 4284 үгээр бичсэн нь тогтоогдлоо. Хүснэгт 2-т цахим сангаас хамгийн их давтамжтай эхний болон төгсгөлийн үгсийг агуулсан хүснэгтийг харуулав.

Хүснэгт 2. Цахим сан

ug	davtamj
боловсролын	802
хөдөлмөрийн	794
ажил	670
болон	544
сургалтын	502
заасан	496
хуулийн	486
энэ	482
зүйл	365
ажлын	351
байгууллага	338
хууль	327
ахуйн	305
эрүүл	295
үйл	294
хамтын	270

ug	davtamj
дэвшүүлж	1
орохыг	1
бэлтгэгдсэн	1
үнэлгээгээр	1
хандлагын	1
туслахад	1
төгсөхөд	1
үсгэн	1
суудаг	1
төхөөрөмжөөр	1
хариуцаж	1
ахиулах	1
тандалтын	1
эрт	1
оношилгоонд	1
давтамжтай	1

Хууль бүр үгтэй, ялгаатай үгтэй. Хуулийн цахим сангийн үгээс хууль бүр хичнээн ялгаатай үгээр бичигдсэнг хуулийн үгийн баяжилт гэж нэрлэвэл доорх шинжтэй байлаа.

Хүснэгт 3. Хуулиудын үгийн баяжилт

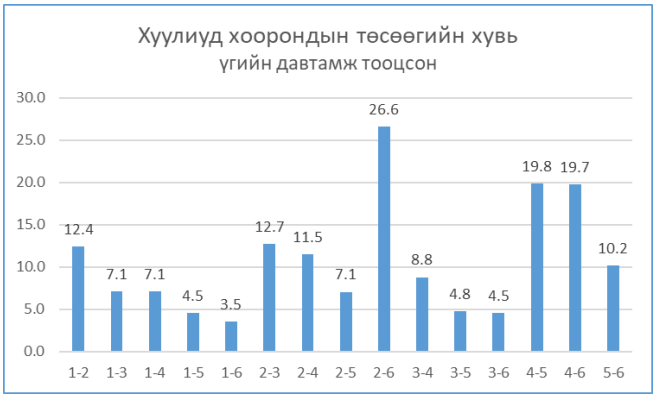
Хууль	Ялгаатай үг	Үгийн баяжилт
Авлигын эсрэг	1354	0.316
Хөдөлмөрийн	2540	0.592
Хөдөлмөрийн аюулгүй байдал, эрүүл ахуйн тухай	1198	0.279
Боловсролын ерөнхий	1650	0.385
Дээд боловсролын	1127	0.263
Сургуулийн өмнөх болон ерөнхий боловсролын тухай	1007	0.235
Сангийн нийт ялгаатай үгийн тоо	4284	
Баяжилтын дундаж утга	1479	0.35

Хуулийн үгийн цахим санд 2 үсэгтэй 26 үг, 3 үсэгтэй 112 үг гэх мэт 17 үсэгтэй үг 7, 18 үсэгтэй ганц үг тааралдлаа. Бүрэлдэхүүнтэйгээр, баталгаажуулалтыг, баталгаажуулалтын, баталгаажуулахаар, хэрэглэгдхүүнийг, хөдөлмөрлүүлэхийг зэрэг үгүүд хамгийн их үсэгтэй үг байлаа.

Хуулийн цахим санг шахах бол Heap, Zipf Mandelbrot-ын хуулийн тогтмолуудыг хуулийн хэлэнд мөн тооцон гаргав. Хуулийн хэлэнд Зипфын хуулиар бодсон давтамжийн хамаарлын томьёо цахим сан дээр тоолсонгоос дунджаар 9.6 үгийн зөрүүтэй байлаа. Харин Zipf-Mandelbrot хуулийн тогтмолуудыг хуулийн хэлэнд тооцоолоход a=1, b=-3482, 0 үед цахим сангаас давтамжийн зөрүү (-956.84, 4966.86) хооронд гарсан. Дундажаар 1.218 үгээр зөрүүтэй тооцоолж байна. Өөрөөр хэлбэл Зипфын томьёогоор бодсонтой харьцуулахад 9 үгээр илүү нарийвчлалтай бодсон байна.

A=2, b=-2160.78 үед цахим сангаас зөрөх зөрүүнүүдийн дундажуудыг авч үзвэл 1.673 байлаа. Хуулиудын төсөөг косинус төсөөгийн томьёогоор тооцоолоход хуулиуд хоорондын төсөөний хувь дараах байдалтай байлаа.

Зураг 1. Хуулиудын төсөөгийн хувь



Үүнд:

1 Авлигын эсрэг хууль
2 Хөдөлмөрийн хууль
3 Хөдөлмөрийн аюулгүй байдал, эрүүл ахуйн тухай
4 Боловсролын ерөнхий хууль
5 Дээд боловсролын тухай
6 Сургуулийн өмнөх болон ерөнхий боловсролын тухай хууль

III. Хуулийн хэлний магадлалт загвар

Хууль бүрт дэхь заалтуудаас үгийн дараа ямар үг ордогоор хуулийн хэлний магадлал загварыг дараах томьёогоор гаргаж болно.

$$P(d) = \frac{L_d!}{\mathrm{tf}_{t_1,d}!\mathrm{tf}_{t_2,d}!\cdots\mathrm{tf}_{t_M,d}!} P(t_1)^{\mathrm{tf}_{t_1,d}} P(t_2)^{\mathrm{tf}_{t_2,d}} \cdots P(t_M)^{\mathrm{tf}_{t_M,d}}$$

Үүнд tf нь ti үг d заалтын i-р үгэнд орох давтамж. Тухайлбал хэлний биграмм загвар 4 үгтэй өгүүлбэрт [7].

$$P_{\mathrm{bi}}(t_1 t_2 t_3 t_4) = P(t_1)P(t_2|t_1)P(t_3|t_2)P(t_4|t_3)$$

Үгийн биграмм сангаас хамгийн их давтагдаж орсон хоёр үгийг буурах дараалаар эхний үгсүүдийг харуулбал

Хүснэгт 4. Хуулийн хэлний биграмм сан

ug1	ug2	freq	prob
энэ	хуулийн	365	0.0852
боловсролын	сургалтын	262	0.0612
дугаар	зүйл	235	0.0549
ажил	олгогч	199	0.0465
эрүүл	ахуйн	172	0.0401
ерөнхий	боловсролын	167	0.0390
хөдөлмөрийн	аюулгүй	164	0.0383
дээд	боловсролын	162	0.0378
авлигатай	тэмцэх	143	0.0334
аж	ахуйн	129	0.0301
асуудал	эрхэлсэн	121	0.0282
сургалтын	байгууллагын	116	0.0271
дүгээр	зүйл	112	0.0261
хуульд	заасан	110	0.0257
хөдөлмөр	эрхлэлтийн	95	0.0222
заасны	дагуу	89	0.0208
тэмцэх	газрын	88	0.0205

IV. Хуулийн бичвэрт үгийн зөвлөмж өгөх загвар

Хэрэглэгч хуулийн бичвэрийн эхний үгийг бичих явцад униграмм загвар буюу үгийн давтамжийн сангаас хамгийн их давтагдаж орсон үгсийг давтамжаар эрэмбэлж харуулна. Үгийг бичиж дуусгамагц эхний үгтэй хосолж орох үгсийг биграмм загвараас эрэмбэлж харуулна. Ийм аргаар хуулийн бичвэр бичнэ.
Тухайлбал: “Энэ” гэдэг үгийг сонгоё. “Энэ” үгийг агуулсан хамгийн их давтагдсан хос үг “энэ хуулийн” болно. Одоо “хуулийн” үгийн агуулсан хамгийн их давтагдсан хос үг “хуулийн зорилт” … гэх мэт

Хамгийн их давтагдаж орсон эхний үгсийг автоматаар сонгоод байвал дараах бичвэрийг гарган авч байна:
“Энэ хуулийн зорилт нь боловсролын сургалтын байгууллагын үйл ажиллагааг зохион байгуулах хөдөлмөрийн аюулгүй байдлыг хангах арга зүйн байгууллага болон бусад бүрэн эрхийг хязгаарлах үндэслэл болохгүй”

Униграмм, биграмм загвар ашиглан хуулийн бичвэр бичихэд санал болгох зөвлөмж үгсийн дараалал харуулснаар хурдан бичих боломж олгохоос гадна алдаагүй зөв үгийг сонгох боломж олгоно. Тухайлбал 8 үгтэй 60 тэмдэгт мөртэй бичвэр бичихэд компьютерын 60 товч дарж байхад магадлалт загвараар 25 товч дарж байна. Өөрөөр хэлбэл зөвлөмж үгтэй магадлал загвар ашиглавал 2.4 дахин цөөн товч дарж байна.
Магадлалт загвар ашиглан хуулийн төсөл боловсруулах явцад бичсэн үгийн араас зөвлөмж үгүүдийг эрэмбэлж жагсаалтыг харуулах, тэдгээрээс сонгох боломж олгох, үг бичих явцад зөв бичгийн дүрмээр бичигдсэн зөв үгийн жагсаалт харуулах, сонгоход хэрэглэгдэнэ.

V. Дүгнэлт

1. Монгол улсын 930 хуулиас 6 хууль сонгож 4284 хуулийн ялгаатай үгтэй үгийн сан байгуулсан. Өөрөөр хэлбэл хуулийн цуглуулга нийт 49839 тэмдэгтүүдийн дараалалтай байгаагаас хуулиуд 4284 үг ашиглан 48712 үгээр бичигджээ. Дунджаар нэг хуулийг 1480 хуулийн үг ашигладаг байна.
2. Хуулийн бичвэрт хуулийн үгийн сангаас ашигласан баяжилтын хувийг гаргасан. Эндээс хамгийн их хуулийн хэллэг ашигласан хуулиар Хөдөлмөрийн хууль, хамгийн баяжилт муутай нь сургуулийн өмнөх ба ерөнхий боловсролын хууль байна.
3. Хууль дундын үгийн тоо, хувийг гаргаснаар хуулиудыг нэгтгэх, салгах боломжтой эсэхийг харж болмоор байна. Боловсролын хууль, дээд боловсролын хууль хоорондоо хамгийн их дундын үгтэй байна. Харин Хөдөлмөрийн хууль, сургуулийн өмнөх ба ерөнхий боловсролын хууль хоорондоо дундын үг багатай тул нэгтгэх боломжтгүй гэж харагдаж байна. Шинэ хуулийг бусадтай нэгтгэх, бусдаас ялгаатай болохыг тодорхойлох боломжтой.
4. Неар-ын хуулийн тогтмолыг хуулийн бичвэр дээрээс тооцоолсоноор шинэ боловсруулсан хууль хичнээн хуулийн хэллэгтэй болохыг урьдчилан таах боломжтой боллоо
5. Хуулиудын төсөөтэй байдлыг cosine томьёогоор тооцоолсны үр дүнд Хөдөлмөрийн хууль, сургуулийн өмнөх ба ерөнхий боловсролын хуулиуд хоорондоо (26.6%) хамгийн их төсөөтэйгээр тооцоологдон гарч байна. Хамгийн хоорондоо төсөө багатай хуулиудаар Авлигын эсрэг хууль, сургуулийн өмнөх ба ерөнхий боловсролын хууль байна. Эндээс шинэ хуулийг батлагдсан хуулиудаас хэр ялгаатайг тодорхойлоход хэрэглэж болохоор байна.
6. Хуулийн өгүүлбэрийн магадлалт загвар гаргаснаар хуулийг бичих үед үгийн араас дараагийн бичигдэх үгийн зөвлөмжийг гаргаж өгөх боломжтой. Үгийн араас ордоггүй үг бичихэд анхааруулга хэвлэхээр зохиомжлож бас болно. Мөн эрх зүйн мэдээллийн системээс хамаатай заалтыг хайхад магадлалт загварыг ашиглаж болно.

Ном зүй

[1] Purev J, Altangerel Ch. Language resources for Mongolian. Conference on Human Language Technology for Development, Alexandria, Egypt. 2011

[2] Боролзой.Д, Энхжаргал.М. Монгол хэлний үгийн давтамжийн толины арга зүйн асуудалд. Хэл зохиол судлал. 2018

[3] Ган-Эрдэнэ.Ж, Сэр-Од.Б. Даалгаврын сан дээр тулгуурлан Heaps-ын хуулийн тогтмолуудыг үнэлэх нь. ММТ 2020 эмхэтгэл. 2020

[4] Хүдэр.А. Монгол текст дээрх Цифтийн хуулийн тогтмолын үнэлгээ. ШУТИС-ийн эрдэм шинжилгээний эмхэтгэл. №16 2012

[5] Чой Сон Сү, Хөмрөг хэл шинжлэлийн аргаар Монгол улсын үндсэн хуулийн бичвэрийг судалсан нь. 2023 он. Диссертаци

[6] Эрх зүйн мэдээллийн нэгдсэн систем www.legalingo.mn

[7] Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, Introduction to Information Retrieval, ISBN: 0521865719 Cambridge University Press. 2008

3D,360 технологи дээр суурилсан сургалтийн цахим контент боловсруулалт

[1]Б.Батшагай,
[1]ШУТИС Мэдээлэл, холбооны технологийн сургууль
[1]e-mail: batshagai@must.edu.mn

[2]М.Жавхлан,
[2]Этүгэн их сургууль НАСС
[2] e-mail: javhlan.m@etugen.edu.mn

[3]Ц.Дэвжрэх
[3]ШУТИС Мэдээлэл, холбооны технологийн сургууль
[3] e-mail: wbhlietue@gmail.com

***Хураангуй* - Бид энэ өгүүлэлд өөрсдийн судлан хөгжүүлж буй нүүр ам судлалын салбарт 360 панорама зураг, видео, 3D интерактив сургалтын цахим веб контент бүхий иммерсив VR (Virtual Reality) виртуал симуляцид суурилсан сургалтын цахим хэрэглэгдэхүүн, түүний хөгжүүлэлт болон шийдлийн талаар тусгалаа. Хөгжүүлж буй цахим контент нь нүүр ам судлалын салбарт шүд авах болон шүдний эмнэлгийн орчны виртуал загварчлалд зориулагдсан. Энэхүү орчныг 360 панорама зураг, видео ашиглан бодит эмнэлгийн орчинд суурилан бүтээж, мөн 3D технологид суурилсан виртуал загваруудыг боловсруулсан. 360 панорама зураг, 3D анимашин бүхий интерактив цахим контентоор бодит эмнэлгийн болон сургалтын орчин, хичээлийн онол, зааварчилгаа, төхөөрөмж ба багаж хэрэгслийн танилцуулгыг багтаасан. Харин иммерсив VR цахим симуляцийн тусламжтайгаар шүд авах процессыг виртуал орчинд дүрслэн үзүүлсэн. Тус цахим контентыг ашиглан эмнэлгийн болон боловсролын бусад шинжлэх ухааны салбарт 360 панорама зураг, 3D интерактив цахим контент болон иммерсив VR виртуал симуляцийг хэрхэн үр дүнтэй ашиглаж болохыг санал болгож, энэхүү өгүүлэлд дэлгэрэнгүй танилцуулж байна.**

Түлхүүр үг: (keywords). Мэдээллийн технологи, Виртуал бодит байдал, виртуал орчин, 360 зураг видео, веб туур, 3D интерактив анимейшн

УДИРТГАЛ

Мэдээллийн технологи хөгжихийн хэрээр түүний хэрэглээний цар хүрээ тэлж, технологийн ололт амжилтууд бидний өдөр тутмын амьдралд улам бүр нэвтэрч байна. Мөн сошиал медиа, маркетинг, виртуал энтертайнмент зэрэг чиглэлээр олон төрлийн хэрэглээ, технологиуд түгэн дэлгэрсээр байна. Үүний нэг жишээ нь 360° зураг, видео ашиглан виртуал аялал болон виртуал маркетингийг өргөнөөр хийх болсон явдал юм.

360 технологийг ашиглан VR орчин үүсгэх, VR контент хөгжүүлэхэд VR болон MR технологиудыг мөн өргөнөөр ашиглаж байна. Бидний хувьд 360° панорама бүхий веб туур боловсруулах технологийг ашиглан сургалтын цахим контент хөгжүүлэхээр зорин ажиллаж байна. Хөгжүүлж буй МООС цахим контентын онцлог нь веб дээр суурилсан 360° туур, 360° видео зураг, 3D интерактив анимейшн ашигласан цахим веб контент юм.

Энэхүү контент нь манай улсын төдийгүй дэлхийн улс орнуудад 360° болон 3D технологид суурилсан анхдагч сургалтын цахим веб контент гэдгээрээ онцлогтой. Мөн 3D технологид суурилсан иммерсив VR виртуал симуляци нь манай улсын Нүүр Ам Судлалын салбарт анх удаа боловсруулж буй сургалтын хэрэглэгдэхүүний нэг гэж үзэж байна.

Бөөнөөр үйлдвэрлэл хөгжихийн хэрээр өртөг буурч, хэрэглээ нэмэгдсэн бөгөөд Insta360, GoPro, Ricoh, Kandao, Samsung Gear зэрэг олон төрлийн 360 камер худалдаанд гарсан. Мөн виртуал бодит байдлын (VR - Virtual Reality) төхөөрөмж болох дэлгэцтэй (HMD) VR төхөөрөмжүүдийн хөгжлийн үр дүнд хэрэглэгчид виртуал бодит байдлыг илүү хүртээмжтэй ашиглах боломжтой болсон. Oculus Rift, HTC Vive, Samsung VR Gear, Microsoft HoloLens, Sony-ийн PlayStation VR зэрэг олон төрлийн VR төхөөрөмжүүд бий болж, орчин үед телевизор, компьютерын дэлгэцийг харж үзэх биш, харин хэрэглэгч өөрөө виртуал орчинд орж, бодит мэдрэмжийг авах боломж нээгдээд байна.

Энэ нь виртуал орчныг илүү бодит мэт мэдрэх, виртуал аялал хийх боломжийг олгож байгаа юм. Виртуал бодит байдал (VR) гэдэг нь программ хангамжийн тусламжтайгаар бүтээгдсэн хиймэл орчин бөгөөд үүнд 360° зураг, видео ашигласнаар хиймэл орчныг илүү бодитой харагдуулах боломж бүрддэг.

Өнөөдөр дэлхий нийт, түүнчлэн манай улсын хувьд виртуал орчин үүсгэх, судлах, сургалт, судалгаа, үзвэр үйлчилгээ, маркетинг, сурталчилгаанд өргөнөөр ашиглаж байна [1]. Иймээс энэхүү өгүүлэлд виртуал орчны тухай ойлголт, түүний техник технологи, программ хангамж, виртуал орчинг хялбар үүсгэх аргачлал болон виртуал орчинд 3D модель ашиглах зэрэг өөрсдийн судлан хөгжүүлж буй контентын технологийн талаар тусгасан болно.

Виртуал бодит байдал (VR) гэсэн ойлголт 1992 оноос яригдаж эхэлсэн боловч тухайн үед өндөр өртөгтэй, төвөгтэй технологи байсан тул төдийлөн хөгжөөгүй. Гэхдээ виртуал орчны санаа анх 1838 онд Чарльз Вийтстоун хүний нүд хоёр өөр өнцгөөс объектийг харж, тархи үүнийг нэгтгэн гүн гүнзгий мэдрэмж үүсгэдэг болохыг судалж, стереоскопыг бүтээсэнтэй холбогддог. Түүний бүтээсэн стереоскопоор дамжуулан хоёр талаас авсан зургийг нэгтгэн үзэхэд хүн гүн гүнзгий орон зайн сэтгэгдэл авдаг байжээ.

Мөн 1860-аад онд 360° градусаар дүрсэлсэн панорамик ханын зурган урлаг үүссэн нь виртуал бодит байдлын үндэс суурийг тавьсан гэж үздэг. 1970 онд "хиймэл бодит байдал" (artificial reality) гэсэн нэр томьёо шинээр үүсч байв [2]. Өнөө үед Meta Quest, Samsung Gear, HTC VIVE, OSVR зэрэг VR сенсор төхөөрөмжүүд хэрэглээнд хурдтай нэвтэрч байна. Facebook 360 videos, YouTube 360 videos зэрэг платформууд ч хэрэглэгчдэд виртуал реалити технологийг өдөр тутамдаа мэдрэх боломжийг олгож байна. Энэхүү виртуал реалити (VR) болон нэмэгдсэн бодит байдлын (AR) технологийн энтертайнмент хэрэглээг боловсрол, шинжлэх ухаан зэрэг олон салбарт үр дүнтэй ашиглах асар их боломж байгааг харуулж байна [3].

ШУТИС-ийн хувьд зайн болон цахим сургалтын хэрэглэгдэхүүнийг 2000 оны эхэн үеэс хөгжүүлж эхэлсэн [4]. Мөн төрийн зүгээс цахим сургалтыг дэмжих хэд хэдэн

үндэсний хөтөлбөр боловсруулан хэрэгжүүлсэн. Тухайлбал, 2002 онд "Зайн сургалтын үндэсний хөтөлбөр", 2005-2012 онд "Цахим Монгол үндэсний хөтөлбөр", 2012-2016 онд "Нэг Монгол санаачилга" зэрэг ажлуудыг нэрлэж болно [5].Цахим сургалт нь дараагийн хөгжлийн төвшин рүү шилжиж тоон (дижитал), нээлттэй боловсролын зарчмаар технологид суурилсан олон чиглэлээр хөгжих төлөв ажиглагдаж байна.

ҮНДСЭН ХЭСЭГ

2010 оноос эхлэн виртуал лабораторийн боловсруулалт, хөгжүүлэлтийн чиглэлээр тасралтгүй судалгаа хийж эхэлсэн бөгөөд туршилтын үр дүнд хоёр хэмжээст онлайн болон оффлайн хэлбэрээр ажиллах, хичээлд хэрэглэх боломжтой интерактив цахим хэрэглэгдэхүүн, лабораториудыг бүтээж туршсан. Тухайн үед 67 дугаар тушаалаар (магистрын түвшний хичээлүүдийг е-хичээл болгон боловсруулах тухай) 260 орчим цахим хичээл болон 10 орчим виртуал лаборатори боловсруулсан байна [6].

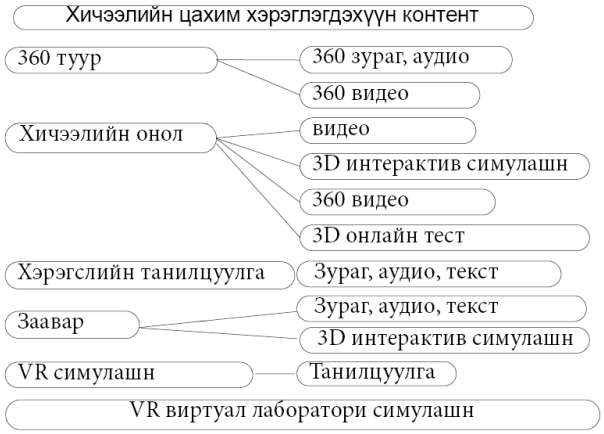
Үүнээс хойш VR технологид суурилсан цахим хичээлийн хэрэглэгдэхүүн, виртуал лабораторийн хөгжүүлэлтийг цаашид үргэлжлүүлэн хийсэн [1]. Харин энэ удаад Нүүр Ам Судлалын салбарт VR технологид суурилсан сургалтын хэрэглэгдэхүүнийг амжилттай боловсруулж дуусгалаа.

Тус ажлын хүрээнд дараах гурван төрлийн цахим контент болон симуляциудыг бүтээв. Үүнд:

- VR Headset (Meta Quest 3) төхөөрөмж дээр ажиллах виртуал бодит симуляци;
- 3D интерактив, 360° виртуал орчин бүхий вэб хуудас буюу цахим контент;
- Meta Quest 3 төхөөрөмжид суурилсан VR виртуал туур.

Эдгээр гурван хэлбэрийн цахим контент, симуляци, виртуал туурыг боловсруулан бүтээсэн болно.

Аливаа виртуал лаборатори контент боловсруулахад тухайн контентын зохиол, ажлын дараалал маш чухал үүрэгтэй. Энэхүү дараалын дагуу файл бүрдүүлэлт хийж, шүд авах процессын видео, аудио, зураглалыг бэлтгэн, 3D бүтээн байгуулалт хийж, мөн 3D VR орчинтой төстэй бодит шүдний эмнэлэгт 360° зураг авалтыг гүйцэтгэсэн. Ийнхүү VR болон 360° виртуал контентуудыг цогцоор нь бүтээж боловсруулсан болно.



Зураг 1. Цахим контенын агуулга, файлын бүтэц

VR, AR, MR, XR технологи нь энтертаймент салбарын шинэ талбар болоод байна. Интертаймент байдлыг боловсрол, бизнес маркетинг болон бусад шинжлэх ухааны салбарт өргөнөөр ашиглаж эхлээд байгаа ба асар их боломжийг агуулж буй салбар. 3D виртуал орчин нь бодит зураглалыг бүрэн гаргах нь дутмаг байдаг учраас сүүлийн жилүүдэд бодит 360 зураг, видео ашиглах хандлаг ажиглагдаж байна. VR болон виртуал орчныг 3 үндсэн ангиллыг ашигладаг:

- Иммерсивбус
- Хагасиммерсив
- Бүрэн иммерсив

360 видео болон зургыг сүүлийн үед хагас болон бүтэн иммерсив VR технологи дээр суурилсан маркетин суртчилгаан, үзвэр үйлчилгээ, виртуал аялаль музей зэргийг хийж байна.

Зураг 2. 360 зураг

360 градусын видео/зураг гэж юу вэ?

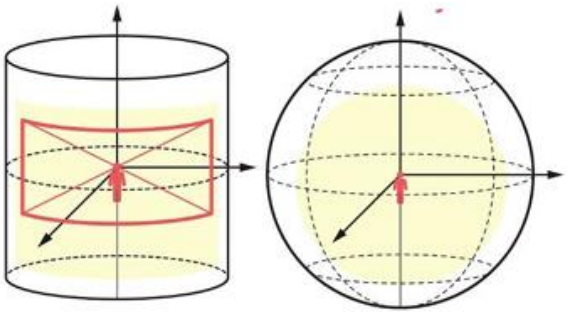
360 градусын видео болон зураг нь бүх чиглэлд бичигдсэн байдаг бөгөөд хэрэглэгчид тухайн орчныг 360 градусаар эргүүлж, харах боломжийг олгодог.

360 зураг, видео нь ихэвчлэн дэлгэмэл (equirectangular) хэлбэртэй файл байдлаар хадгалагддаг бөгөөд боловсруулалтын явцад хүний харах өнцөгт тохируулан CrystalView технологийг ашиглан тусгай сегментчилсэн бөмбөрцөг хэлбэрт шилжүүлдэг [7].

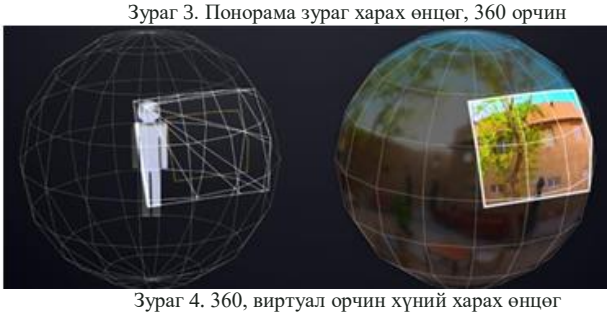
Өөрөөр хэлбэл, хэрэглэгч бөмбөлөг хэлбэртэй орчинд дотор нь орсон мэт өнцгөөс харж чадна. Энэ нь хоёр янзаар харагдах боломжтой:

- Энгийн дэлгэц дээр 360 градусаар эргүүлэн үзэх
- VR төхөөрөмж (жишээ нь: Meta Quest зэрэг VR headset) ашиглан бүрэн дүрслэлтэй, илүү бодит мэдрэмжтэй үзэх

VR headset-ээр үзэх тохиолдолд хэрэглэгч тухайн орчинд бүрэн автагдаж, илүү бодит мэдрэмж авдгаараа онцлогтой..



Зураг 3. Понорама зураг харах өнцөг, 360 орчин



Зураг 4. 360, виртуал орчин хүний харах өнцөг

VR орчин

Аливаа VR бүтээхэд виртуал орчин маш чухал үүрэгтэй байдаг. Виртуал орчин гэдэг нь бидний нүдээр харж буй объектууд болон орчныг хэлнэ. Түүнчлэн

"виртуал бодит байдал (VR)" болон "виртуал орчин" гэсэн ойлголтыг ялгаж ойлгох шаардлагатай. Виртуал орчин нь бодитоор хүрэлцэх боломжгүй буюу тодорхой алслалтай орчныг дүрслэхэд зориулагддаг бөгөөд үүнийг 360 градусын дүрс ашиглан бүтээх нь хамгийн оновчтой шийдэл болдог.

VR орчин бүтээх үндсэн хоёр арга зам бий:

- 3D график программ ашиглан бүрэн орчныг бүтээх
- 360° зураг, видео ашиглан орчин үүсгэх

Эдгээр аргыг сонгохдоо тухайн VR орчин ямар зориулалттай вэ гэдгээ сайтар анхаарах шаардлагатай.

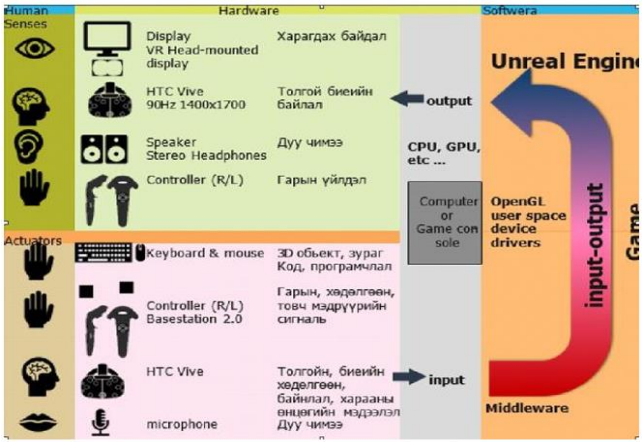
- 3D графикийг кино, анимэйшн, тоглоом, симуляци зэрэг интерактив VR орчин бүрдүүлэхэд түлхүү ашигладаг.
- Харин 360° зураг, видеог реклам, маркетинг, сошиал медиа, виртуал аялал зэрэг статик VR орчин бүтээхэд голчлон ашиглаж байна [3].

Бидний хөгжүүлж буй контентын онцлог нь дээрх хоёр аргыг хослуулан бүрэн ашигласан явдал юм.

Тодруулбал:

- Хичээлийн онол, агуулгыг 360° бодит виртуал орчинд бүтээсэн;
- Шүд авах симуляци хэсгийг 3D виртуал бодит VR орчинд ажиллахаар боловсруулсан.

360° болон 3D виртуал орчныг бүрэн иммерсив (immersive) хэлбэрээр буюу виртуал бодит байдлын VR орчинд харуулахын тулд дараах интерфейс болон төхөөрөмжүүд шаардлагатай байдаг.



Зураг 5. VR бодит байдалд орох интерфейс

Иммерсив VR орчинд ороход headset буюу HMD дэлгэц бүхий төхөөрөмж, чимээ хөдөлгөөн мэдрэх камер, гарын controller зэрэг төхөөрөмжийн тусламжтай өгөгдлийг системд оруулан боловсруулан программчиллаж буцаан HMD болон бусад төхөөрөмжид гаралтыг илгээх хэлбэрээр хүний харах болон бусад мэдрэхүйд хүрснээр VR орчинд орон үйлдэл хийх боломжийг олгоно[8].

Өөрөөр хэлбэл

1. Хүн /мэдрэх, үйлдэл илгээх/
2. Техник хангамж /Өгөгдлийн ороолт гаралтын сигнал/
3. Программ хангамж /программчлал, боловсруулалт/ гэсэн үндсэн гурван хэсгээс бүрдэх юм.

Цахим котентийн боловсруулалт

Технологийн шийдэл нь:

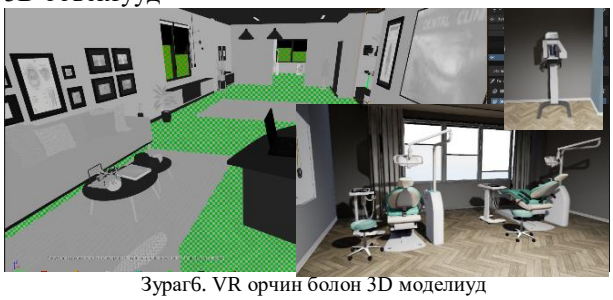
- 3D, 360 веб технологи дээр тулгуурлан хийсэн ба веб framework болон скрипт Three.js сан зэргийг ашиглан 360, 3D интерактив анимашн бүхий интертаймен веб контент хийсэн.
- VR орчин болон 3D модель хийхэд blender, maya, зэрэг программ ашигласан ба тоглоомын хөдөлгүүр Unity3D Game engine технологи ашиглан VR симулашн хөгжүүлэлт хийсэн.

Виртуал орчин хөгжүүлэлт

Виртуал орчны хөгжүүлэлтийн хүрээнд 3D зураглал болон объектуудыг бүрэн зурж байгуулсан бөгөөд мөн 3D анимашнуудыг боловсруулан хийсэн.

VR орчин болон веб контентийн 3D анимашн хөгжүүлэхэд дараах үндсэн бүрдэл хэсгүүд орсон:

- 3D байгуулалтын орчин
- 3D объектууд



Зураг6. VR орчин болон 3D моделиуд

Цахим хичээлийг хоёр үндсэн хэлбэрээр хийхээр төлөвлөсний дагуу

- VR орчин туур хэлбэрээр
- 3D 360 веб цахим контент гэсэн 2 хэлбэрээр хийсэн болно.

Зураг 7. 360 цахим контент

Зураг 8. 360 цахим контент 3D анимашн

VR цахим контент

3D 360 веб бүхий цахим контентыг мөн VR виртуал орчин headset дээр тулгуурлан тухайн орчин бодитоор орох үйлдэл хийх эмнэлгийн бодитод орчиг VR-аар үзүүлэх зорилгоор Meta Quest3 тохируулан виртуал туур хийсэн.

Зураг 9. Oculos quest3 дээр цахим контент харагдах байдал

3D 360 цахим контент тестийн хэсэг

Цахим контентын хүрээнд тухайн хичээлийн агуулгад тохирсон тестийн хэсгийг боловсруулсан. Энэхүү тест нь энгийн асуулт хариултаас ялгарах онцлогтой бөгөөд 360° виртуал лабораторийн орчин болон 3D объектын зургуудыг ашиглан сонирхолтой, идэвхтэй оролцоотой болгох шийдлийг тусгасан. Мөн тестийн хэсгийг 360° бичлэг, зураг, аудио, 3D анимашнаар баяжуулах боломжтой интерфейстэйгээр бүтээсэн..

Зураг 10. 3D 360 цахим контент тестийн хэсэг

Зураг 11. 3D VR шүд авах симулашн

VR виртуал бодит байдалд суурилсан шүд авах симуляц

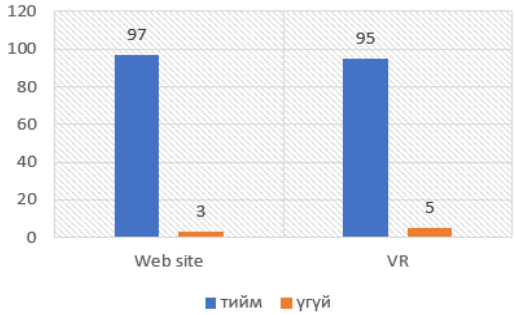
VR виртуал бодит байдлын төхөөрөмж (headset) дээр суурилсан шүд авах симуляцыг Unity3D game engine технологийг ашиглан хөгжүүлсэн. Энэхүү симуляци нь эмнэлгийн виртуал орчинд Hand Tracking технологи ашиглан интерактив үйлдэлийг бодит хэлбэрээр хийх мэдрэмжийг хэрэглэгчдэд төрүүлэх симуляцыг хөгжүүлэх явцад 3D интерактив үйлдлүүдийг (action) нэмэн, бодит орчныг аль болох ойртуулсан 3D орчин г загварчилсан. Тус ажилд 3D модель, орчин, объектоос гадна action, animation, character, assets зэрэг бүх элементүүдийг бүтээж нийтдээ 1000 гаруй мөр код бичиж гүйцэтгэсэн болохыг онцлох нь зүйтэй.

Энэхүү VR симуляц лаборатори нь хүний шүд авах процессыг виртуал орчинд бүрэн гүйцэд загварчилсан бөгөөд цаашид хөгжүүлэлтийг үргэлжлүүлэх суурь нөхцөл болох программчлалын үндэс, объектын орчин бүрэн хангагдсан байна..

Туршилт

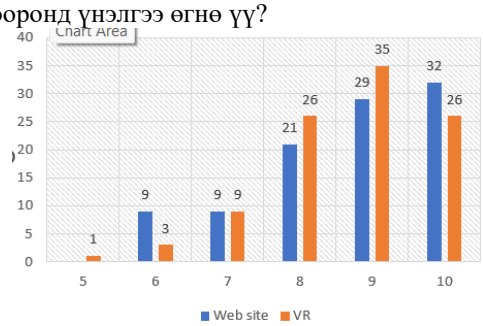
1. 360, 3D Web цахим контентоос 3-р курсийн эмэгтэй 82 эрэгтэй 18 нийт 100 оюутнаас асуумж судалгаа
2. VR буюу (виртуал бодит байдал)-ыг симуляц дээр шүд урлахуй материал судлал хичээлийн 3-р курсийн 71эмэгтэй 29 эрэгтэй нийт 100 оюутнаас асуумж судалгаа авахад гарсан судалгааны үр дүн:

Web site-аар дахин зочилох хүсэл төрсөн үү?
VR-ыг хэрэглэж үзэхэд таалагдсан эсэх?



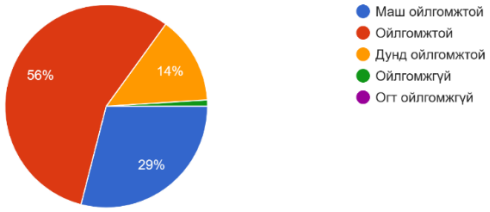
Зураг 12. 3 дугаар курсын 100 оюутаны асуумж

1-10 хооронд үнэлгээ өгнө үү?



Зураг 13. 1-10 оноогоор үнэлхэд 5 дээш үнэлгээ өгсөн

Web site дээрх тест ойлгомжтой байсан уу?

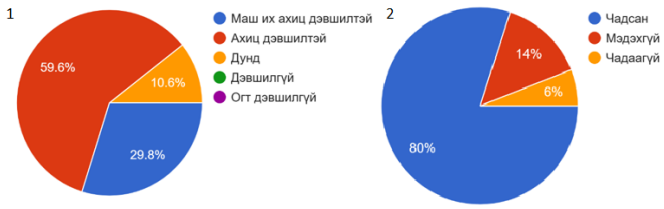


Зураг 14. 3D, 360 тестийн асуумж

1. Өмнө нь VR-ын тухай ойлголттой байсан уу?
2. VR-ыг анх хэрэглэж үзэхэд ямар сэтгэгдэл төрсөн бэ?

Зураг 15. VR хэрэглээний асуумж

1. VR(виртуал бодит байдал)-ыг хичээлтэй холбож зааснаар сурлагад ахиц дэвшил гарах уу?
2. Эмнэлгийн нөхцөлдөх практик ойлголтыг виртуал ертөнцөөр дамжуулан авч чадсан уу?



Зураг 16. VR сургалтад хэрэглэх асуумж

3D, 360 цахим контент VR симулашн хөгжүүлэлтийг сайжруулах зорилгоор авсан зарим асуумжийн үр дүнг дээр харуулсан ба сайжруулалт хөгжүүлэлтийг үргэлжлүүлэн хийсээр байна. Мөн одоогоор hand tracking технологийг ашиглан гарын control ашиглахгүй headset дээр виртуалаар объект барих нь илүү иммерсив мэдрэмж төрүүлж байна.

ДҮГНЭЛТ

Сургалтын цахим хэрэглэгдэхүүн, контент, виртуал лаборатори нь орон зай, цаг хугацаа, үнэ өртөг, суралцах, боломж боловсролын хүртээмжтэй байдал зэрэг маш олон талын хэрэглээ ач холбогдолтой. Цахим хэрэглэгдэхүүн виртуал лаборатори симулашн боловсруулахад олон шийдэлтэй байдаг технологийн шинэчлэл түүний ололт амжилттай уялдан шинэ шийдлүүд дээр тулгуурлан бүрэн иммерсив хэлбэрийн virtual reality цахим контент хийсэн нь бидний хувьд шинэлэг тал ач холбогдол нь оршиж байна. Ирээдүйд 5G болон бусад технологийн шийдлүүд болон WebVR, 3D, VR, AR технологиуд нь интертаймэнт 3D хэлбэрийн цахим хичээл, виртуал лаборатори, виртуал аялал, цахим хэрэглэгдэхүүнийг хөгжүүлж олон нийтийн хэрэглээнд өргөнөөр нэвтрүүлэх боломж харагдаж байна. Мөн 3D, VR хэлбэрийн энэ ажил нь цаг хугацаа болоод хүн боловсон хүчин, тоног төхөөрөмж, үнэ өртөг, технологийн шийдэл ихээр шаардана. Харьцангуй сүүлийн жилүүдэд технологи, төхөөрөмжийн үнэ дэлхийн зах зээл дээр буурч байгаа бөгөөд headset ийн үнэ хувьд дунджаар 500$ байна. Хэдийгээр үндсэн хөгжүүлэлт хийх программ хангамж хөдөлгүүр үнэгүй байгаач VR library, development environment зэргийг ашиглан engine хийх 3D объект боловсруулах зэрэг өртөгтэй цаг хугацаа их шаардсан ажил болж байна. Технологийн хөгжилтэй зэрэгцэн цахим хэрэглэгдэхүүн виртуал лаборатори хөгжүүлэлтийн орчин технологи өөрчлөгдөж байгаатай хөл нийлүүлэн шинэ шийдлүүдийг илэрхийлэн энэ хүү виртуал цахим контент боловсруулсан.

Учир нь тус контент нь 360 VR орчин болон 3D объект 3D анимашн ашиглан хийсэн ба энэ нь одоогоор Монгол улсад хийдээгүй бөгөөд дэлхийн бусад оронд эхлэн ашиглагдаж буй вебийн төрөл бөгөөд энэ технологийг ашигласнаар платформ хамааралгүй веб хөтөч дээр суурилан шүдний бүтцийг 3D харах боломжийг оюутан суралцагчдад олгож байгаа гэдгээрээ маш онцлог бөгөөд VR-виртуал бодит орчноо game engine технологи ашиглан хийж суралцах чиг хандлага, виртуалаар туршилт судалгаа дадлага хийх зэрэг олон талын боломжийг олгосон.

Туршилтийн вебийг http://202.70.34.27:8089/ хаяг дээр байшуулсан ба та одоогоор зөвхөн компьютер дээр үзэхийг анхаарна уу/ хөгжүүлэлт үргэлжилж байгаа/!

Бидний хувьд цахим хэрэглэгдэхүүний мөн чанар нь сургалтын агуулгыг баяжуулах, хичээлийн зорилго зорилтыг оюутан суралцагч нарт бүрэн хүргэхэд чиглэсэн байх ёстой гэсэн үндсэн шаардлагаар хэрэглэгдэхүүнийг боловсруулсаар ирлээ.

НОМ ЗҮЙ

[1]. Б.Батшагай “ИММЕРСИВ VR-Virtual Reality ЦАХИМ ЛАБОРАТОРИ ХӨГЖҮҮЛЭЛТ” ММТ2024.

[2]. Steven M. LaValle “VIRTUAL REALITY” Cambridge University Press, 2023.

[3]. Б.Батшагай, "360^0 зураг, видеоний хэрэглээ ба виртуал аялал,"ММТ 2023.

[4]. Т.Уранчимэг, Б.Батшагай, Б.Цогтсайхан “Цахим хэлбэрийн сургалтын технологийн зарим асуудал” Эрдэм шинжилгээний бүтээл 2010, 10, pp. 138-143.

[5]. Б.Очирбат "Цахим сургалт боловсролын хөгжлийн чиг хандлага," Зууны мэдээ 2020.

[6]. Б.Батшагай, Т.Уранчимэг "Цахим хэрэглэгдэхүүнүүдийг боловсруулж, инженерийн сургалтад нэвтрүүлсэн замнал, дүгнэлт," Нээлттэй боловсрол , 2021.

[7]. .https://www.insta360.com/product/insta360-pro2

[8]. Б.Батшагай, Т.Урагчимэг, "Virtual лабораторийн хөгжүүлэлт," ММТ 2022.

Ангиллын зарим аргаар өсвөр насныхны сэтгэл хөдлөл, зан үйлийг хүйсээс хамааруулан судлах

Батнаст Ганчөдөр
Физик мэдээлэлзүйн тэнхим, БАС
АШУҮИС
Монгол улс, Улаанбаатар хот
NastG9811@gmail.com

Баярмаа Ваанчиндорж
Сэтгэцийн эрүүл мэндийн үндэсний төв
Монгол улс, Улаанбаатар хот
vbayarmaa41@gmail.com

Ажнай Лувсан
Физик мэдээлэлзүйн тэнхим, БАС
АШУҮИС
Монгол улс, Улаанбаатар хот
ajnai@mnums.edu.mn

Булгансүрэн Дагвасамбуу
Физик мэдээлэлзүйн тэнхим, БАС
АШУҮИС
Монгол улс, Улаанбаатар хот
enkhurel@mnums.edu.mn

Ахыт Тилуебай
Физик мэдээлэлзүйн тэнхим, БАС
АШУҮИС
Монгол улс, Улаанбаатар хот
akhit@mnums.edu.mn

Нямдаваа Уугандаваа
Физик мэдээлэлзүйн тэнхим, БАС
АШУҮИС
Монгол улс, Улаанбаатар хот
nyamdavaa.u@mnums.edu.mn

Баасандорж Чилхаасүрэн
Физик мэдээлэлзүйн тэнхим, БАС
АШУҮИС
Монгол улс, Улаанбаатар хот
baasandorj@mnums.edu.mn

Жавзмаа Цэнд
Физик мэдээлэлзүйн тэнхим, БАС
АШУҮИС
Монгол улс, Улаанбаатар хот
javzmaa.ts@mnums.edu.mn

Пүрэвдолгор Лувсанцэрэн
Физик мэдээлэлзүйн тэнхим, БАС
АШУҮИС
Монгол улс, Улаанбаатар хот
purevdolgor@mnums.edu.mn

Энх-Үрэл Энхжаргал
Физик мэдээлэлзүйн тэнхим, БАС
АШУҮИС
Монгол улс, Улаанбаатар хот
enkhurel@mnums.edu.mn

***Хураангуй* —Олон улсад өсвөр үеийнхний сэтгэцийн эрүүл мэндийг SDQ асуулгаар үнэлэн, өгөгдөл тандалтын аргаар таамагласан судалгааны ажлууд нилээд бий болж байна. Монгол улсын сэтгэцийн эрүүл мэндийн салбарт цугларсан их өгөгдөлд өгөгдөл тандалтын аргаар тооцоолсон судалгааны ажил хомс байгаа нь энэхүү судалгааны ажлыг гүйцэтгэх үндэслэл болж байна. Ангиллын зарим аргаар өсвөр насныхны сэтгэл хөдлөл, зан үйлийг хүйсээс нь хамааруулан судлахыг зорилоо.**

СЭМҮТ-тэй хамтран Говь-Алтай аймгийн 6-12 дугаар ангийн 3764 сурагч, тэдний асран хамгаалагч, ангийн багшаас "Сэтгэл хөдлөл, зан үйлийн байдлыг илрүүлэх асуумж" (Strengths and Difficulties Questionnaire - SDQ)-аар өгөгдөл цуглуулан, сурагч, асран хамгаалагч, ангийн багш тус бүрийн сан үүсгэсэн.

Сурагчийн санг эрэгтэй, эмэгтэй хүйсээр нь хуваан, C50, Бейс, Дүрэмд үндэслэсэн аргуудаар тэдний сэтгэл хөдлөлийн дотоод мэдрэмж болон зан үйл, хэт хөдөлгөөнтөхийн гадаад илрэлээр зааг, хэвийн бус төлөвийг таамаглах загварууд үүсгэн, гүйцэтгэлийг үнэлэх хэмжүүрээдээр харьцуулан, өндөр үзүүлэлтэй загвараар үр дүнг илэрхийлсэн.

Түлхүүр үг— сурагч, сэтгэцийн эрүүл мэнд, өгөгдлийн тандалт, загвар, нарийвчлал

I. Удиртгал (Style: *Heading 1*)

Дэлхийн хэмжээнд 10-19 насныхны 13% сэтгэцийн эмгэгтэй байна. 2020 оны байдлаар 19 хүртэлх насны 1.289.587 хүүхэд байгаа нь 2015 оныхоос 0.9%-иар өссөн үзүүлэлттэй [1]. 2013 оны өсвөр үеийнхний судалгаагаар эрүүл саруул 60.5%, сэтгэл хөдлөл, зан үйлийн асуудалтай 30.5%, эмгэгтэй 9% байна. Говь-Алтай аймагт 2018-2019 онд хийсэн судалгаагаар 58.6% эрүүл, 36.1% асуудалтай, 5.2% эмгэгтэй байна [2]. Олон улсад өсвөр үеийнхний сэтгэцийн эрүүл мэндийг SDQ асуулгаар үнэлэн, хиймэл оюуны аргаар таамагласан судалгааны ажлууд нилээд бий болж байна. Монгол улсын сэтгэцийн эрүүл мэндийн салбарт цугларсан их өгөгдөлд хиймэл оюун, машин сургалтын аргаар тооцоолсон судалгааны ажил хомс байгаа нь энэхүү судалгааны ажлыг гүйцэтгэх үндэслэл болж байна. Иймд их өгөгдлийг боловсруулах өгөгдөл тагнадалтын ангиллын зарим аргаар өсвөр насныхны сэтгэл хөдлөл, зан үйлийг хүйсээс нь хамааруулан судлахыг зорилоо.

II. Материал, арга зүй

A. *SDQ асуумж*

Олон улсад өргөн хэрэглэгддэг ерөнхий боловсролын сургуульд суурилсан өсвөр насныхны сэтгэл хөдлөл, зан үйлийг үнэлэх "Сэтгэл хөдлөл, зан үйлийн байдлыг илрүүлэх асуумж" (Strengths and Difficulties Questionnaire - SDQ) нь сэтгэл хөдлөл, зан үйл, хэт хөдөлгөөнтөх, үе тэнгийнхний харилцааны бүлэг бүрээр 5 асуулт, асуулт бүр үнэн биш 0, заримдаа үнэн 1, яг үнэн 2 оноо бүхий 3 шалгуур, нийт 25 асуулттай болно. SDQ асуумж нь хүүхэд, асран хамгаалагч, ангийн багшийн хувилбартай бөгөөд асуултууд нь ижил байна. Сурагчдын дотоод илрэлийг нь сэтгэл хөдлөл, үе тэнгийнхний харилцааны 10, гадаад илрэлийг зан үйл, хэт хөдөлгөөнтөхийн 10 асуумжаар тодорхойлдог [2, 3].

B. *Сурагчийн сан*

Эрүүл Мэндийн Яамны Анагаах Ухааны Ёс Зүйн Хяналтын Хорооны 2022 оны 262 тоот тогтоолоор SDQ асуумжаар судалгааг хэрэгжүүлэх ёс зүйн зөвшөөрөл авсан. СЭМҮТ-тэй хамтран Говь-Алтай аймгийн 6-12 дугаар ангийн 3764 сурагч, асран хамгаалагч, ангийн багшаас SDQ асуумжаар судалгаа авч, сурагч, асран хамгаалагч, багшийн сан үүсгэсэн.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Сурагчийн сан нь хүйс, бүлэг тус бүрээр 5, нийт 21 аттрибуттай. Бүлэг бүрийн аттрибутудын онооны нийлбэрээр нийт оноог тооцсон. Нийт оноо нь сурагчийн сэтгэл хөдлөл, зан үйлийн хэвийн, зааг, хэвийн бус төлөвийг илэрхийлэх класс аттрибут болно. Нийт оноо 0-15 бол хүүхэд Сэтгэл хөдлөл, зан үйлийн хэвийн илрэлтэй, эрүүл, 16-19 үед Сэтгэл хөдлөл, зан үйлийн зааг илрэлтэй буюу "Сэтгэцийн тулгамдсан асуудалтай, 20-40 үед Сэтгэл хөдлөл, зан үйлийн хэвийн бус илрэлтэй буюу Сэтгэцийн тулгамдсан асуудалтай эмгэг үүсэнийг илэрхийлнэ. Класс аттрибутын хэвийн утга 48% (1813), зааг төлөв 28% (1070), хэвийн бус төлөв 23% (881) байна. Иймд сурагчийн сангаас зөвхөн зааг, хэвийн бус төлөвийн 1951 мөрийг сонгон, эрэгтэй 903, эмэгтэй 1048 бичлэгт хуваан, тэдний дотоод, гадаад илрэлээр зааг, хэвийн бус төлөвийг таамаглана.

C. *Ангилал*

Сурагч, багш, асран хамгаалагчийн сангийн аттрибутууд 0-2 ангилсан хэлбэртэй тул ангиллын аргуудыг тохиромжтой гэж үзсэн.

Эрэгтэй, эмэгтэй сурагчийн сангийн дотоод, гадаад илрэлийн аттрибут бүрийг класс аттрибутаас хамааруулан хи-квадрат тестээр (Chi-Square Test) тооцон, сонгосон. Өгөгдөл тандалтын ангиллын C50, Бейс, Дүрэмд үндэслэсэн аргаар тус хувьсагчдаас сэтгэл хөдлөл, зан үйлийн зааг, хэвийн бус төлөвийг ангилах загварууд үүсгэсэн. Сан тус бүрийг хөндлөн баталгаажуулалтаар санамсаргүйгээр k=10 бүлэгт хуваан, сургалт, туршилтыг k удаа хийнэ. I дугаар давталт бүрт i дахь багцыг туршилтын багцаар нөөцөлж, үлдсэнийг дээр дурдсан ангиллын аргаар сурган загварууд үүсгэсэн [4].

C5.0 алгоритмаар гадаад болон дотоод илрэлийн аттрибутыг класс аттрибутаар ангилах мэдээллийн тоо хэмжээг тооцон, хамгийн их мэдээллийн тоо хэмжээтэй аттрибутаар хуваан, мод загвар [5, 6], Бейсийн теоромоор класс аттрибутаар таамаглах загвар [7], Дүрэмд суурилсан RIPPER алгоритмаар IF урьдчилсан нөхцөлд атрибут бүрийг логик AND холбоосоор холбон THEN үр дагавараар класс аттрибутын төлөвийг таамаглан, дүрэмд суурилсан ангилагч тус тус үүсгэсэн [4].

D. *Загварын гүйцэтгэлийг үнэлэх хэмжүүрүүд*

Эрэгтэй, эмэгтэй сурагчийн санг хөндлөн баталгаажуулалтаар 10 хуваан тооцоолсон загваруудыг тэдгээрийн зөв ангилсан эерэг (TP) ба сөрөг багц (TN), буруу ангилсан эерэг багц FP, сөрөг багцын (FP) тоо бүхий алдааны матрицуудыг ашиглан зөв ангилсан туршилтын багцын эзлэх хувь **нарийвчлал**, эерэг багцын хувь **мэдрэмж**, сөрөг багцын хувь **өвөрмөц байдал** зэрэг гүйцэтгэлийн хэмжүүрээр үнэлэх болно.

III. ТУРШИЛТ, ҮР ДҮН

Судалааны ажлаар үүсгэсэн сурагчийн сангийн дотоод болон гадаад илрэлийн хүчин зүйл буюу аттрибутуудын тайлбарыг Хүснэгт1-д үзүүлэв. Аттрибутууд ангилсан хэлбэртэй байна.

Хүснэгт I. Сурагчийн сангийн аттрибутууд

Төрөл		Аттрибутууд		Үнэн биш	Заримдаа үнэн	Яг үнэн
Дотоод илрэл	Сэтгэл хөдлөл	S3	Ихэнхдээ толгой өвддөг	0	1	2
		S8	Олон зүйлд санаа зовнидог, түгшдэг	0	1	2
		S13	Ихэнхдээ аз жаргалгүй, сэтгэлээр унадаг.	0	1	2
		S16	Шинэ орчинд байхаараа сандарч түгшдэг	0	1	2
		S24	Олон зүйлээс айдаг	0	1	2
	Үе тэнгийнхэний харилцаа	S6	Ихэнхдээ ганцаараа тоглодог	0	1	2
		S11	Нэг л дотны сайн найзтай	2	1	0
		S14	Ерөнхийдөө бусад хүүхдүүдэд таалагддаг	2	1	0
		S19	Бусад хүүхдүүд дарамталдаг	0	1	2
		S23	Үеийн хүүхдүүдтэй байхаасаа илүү том хүмүүстэй байх амар	0	1	2
Гадаад илрэл	Зан үйл	S5	Дүрсхийн уурлаж, ердийн байдлаа алддаг	0	1	2
		S7	Ихэнхдээ дуулгавартай байдаг	2	1	0
		S12	Байнга зодолддог, бусдыг дарамталдаг	0	1	2
		S18	Байнга худлаа ярьдаг, хууран мэхэлдэг	0	1	2
		S22	Сургууль, гэр зэрэг газраас ямар нэгэн зүйлийг зөвшөөрөлгүй авч байсан	0	1	2
	Хэт хөдөлгөөнтөх	S2	Хөдөлгөөнтэй болохоор нэг газраа удаан байж чаддаггүй	0	1	2
		S10	Байнгын хөдөлгөөнтэй, тогтворгүй байдаг	0	1	2
		S15	Анхаарал амархан сарнидаг	0	1	2
		S21	Ямар нэгэн үйлдэл хийхийн өмнө сайтар тунгаан боддог	2	1	0

		S25	Эхэлсэн ажлаа дуусгадаг	2	1	0
Класс аттрибут		B6	Бүлэг тус бүрийн нийлбэр оноогоор нийт оноо тодорхойлогдоно	Хэвийн - 0	Зааг - 1	Хэвийн бус - 2
				0-15	16-19	20<

Нийт 20 аттрибуттай, зааг, хэвийн бус утгатай класс аттрибуттай, 1951 мөртэй сурагчийн санг хүйсээр нь салган, R программын train() функцыг ашглан [8, 9] хөндлөн баталгаажуулалтаар k=10 бүлэгт хуваан, ангиллын аргуудаар загварууд үүсгэн, confusionMatrix() функцаар гүйцэтгэлийг үнэлэх хэмжүүрүүдийг тооцоолон үнэлсэн.

Хи-квадрат тестээр аттрибутуудыг класс аттрибутаас хамааруулан R программын vcd сангийн assocstats() функцаар тооцоход эрэгтэй сурагчдын дотоод сэтгэл хөдлөл мэдрэмжийн s3, s13, s24, s6, s19, s23, гадаад илрэлийн s5, s12, s18, s22, s2, s10 үзүүлэлтүүд дундаж хамааралтай байна. Эмэгтэй сурагчдын хувьд дотоод илрэлийн s3, s8, s13, s16, s24, s6, s19, s23, гадаад илрэлийн s5, s12, s18, s2, s10, s15 хувьсагчид дундаж хамааралтай гарсан. Иймд хамааралтай хувьсагчдаар класс аттрибутаас хамаарах загварууд үүсгэсэн.

Хүснэгт II-д эрэгтэй сурагчийн санд тооцсон сэтгэл хөдлөл, зан үйлийн загваруудын нарийвчлал, мэдрэмж, онцлог байдал зэрэг гүйцэтгэлийг үнэлэх хэмжүүрүүдийг харуулав. Эрэгтэй сурагчдын дотоод, гадаад илрэлийн үзүүлэлтээр Бейсийн загварын нарийвчлал 0.01-0.03, мэдрэмж 0.03-0.05, өвөрмөц чанар -0.01-0.02 хооронд, гадаад илрэлийн хувьд нарийвчлал 0.02-0.03, мэдрэмж 0.01-0.03, өвөрмөц байдал 0.04-0.05 хооронд их байна.

Хүснэгт II. Эрэгтэй сурагчдын сэтгэл хөдлөл, зан үйлийн загварын гүйцэтгэлийг үнэлэх хэмжүүрүүд

Арга				Алдааны матриц		Гүйцэтгэлийн хэмжүүр		
				Таамагласан утга		Accuracy	Sensitivity	Specificity
				1	2			
Дотоод илрэл	C50	Бодит утга	1	43	15	0.71	0.75	0.67
			2	14	28			
	Bayes	Бодит утга	1	43	11	0.74	0.79	0.69
			2	15	32			
	RIPPER	Бодит утга	1	45	14	0.73	0.76	0.68
			2	13	29			
Гадаад илрэл	C50	Бодит утга	1	48	15	0.75	0.76	0.75
			2	9	28			
	Bayes	Бодит утга	1	49	13	0.79	0.79	0.78
			2	8	30			
	RIPPER	Бодит утга	1	46	13	0.76	0.78	0.74
			2	11	30			

Хүснэгт III-д эмэгтэй сурагчийн санд тооцсон загваруудын гүйцэтгэлийг үнэлэх хэмжүүрүүдийг харуулав. Эмэгтэй сурагчдын дотоод, гадаад илрэлийн үзүүлэлтээр Бейсийн загварын нарийвчлал 0.05, мэдрэмж 0.02-0.04, өвөрмөц чанар 0.08-0.09 хооронд, гадаад илрэлийн хувьд нарийвчлал 0.01-0.03, өвөрмөц чанар 0.03-0.07 хооронд их байна.

Хүснэгт III. Эмэгтэй сурагчдын сэтгэл хөдлөл, зан үйлийн загварын гүйцэтгэлийн хэмжүүрүүд

Арга				Алдааны матриц		Гүйцэтгэлийн хэмжүүр		
				Таамагласан утга		Accuracy	Sensitivity	Specificity
				1	2			
Дотоод илрэл	C50	Бодит утга	1	40	16	0.71	0.72	0.71
			2	13	31			
	Bayes	Бодит утга	1	44	14	0.77	0.76	0.78
			2	9	33			
	RIPPER	Бодит утга	1	38	14	0.71	0.73	0.69
			2	15	33			
Гадаад илрэл	C50	Бодит утга	1	39	15	0.71	0.72	0.69
			2	14	32			
	Bayes	Бодит утга	1	41	16	0.72	0.72	0.73
			2	12	31			

	RIPPER	Бодит утга	1	36	14	0.69	0.73	0.66
			2	17	34			

Бейсийн аргаар тооцсон эрэгтэй, эмэгтэй сурагчдын сэтгэл хөдлөл, зан үйлийн загварын гүйцэтгэлийн хэмжүүрүүдийн үр дүн өндөр гарсан тул үр дүнг нь мэдлэг болгон илэрхийллээ (ХүснэгтIV-ХүснэгтVII).

Эрэгтэй сурагчдад сэтгэл хөдлөлийн ихэнхдээ аз жаргалгүй, сэтгэлээр унах (S13), олон зүйлээс айх (S24), үе тэнгийнхэний харилцааны ихэнхдээ ганцаараа тоглодог (S6), бусад хүүхдүүд дарамталдаг (S19) зэрэг дотоод сэтгэл хөдлөлийн мэдрэмжээс хэвийн бус илрэл үүсэх магадлалтай (Хүснэгт IV).

Хүснэгт IV. Эрэгтэй сурагчдын дотоод сэтгэл хөдлөлийн мэдрэмжээс хамаарах сэтгэл хөдлөл, зан үйлийн ерөнхий төлөв

Үзүүлэлт	Үнэн биш - 0, заримдаа үнэн - 1, яг үнэн - 2								
	0	1	2	0	1	2	0	1	2
Сэтгэл хөдлөл	s3			s13			s24		
Зааг - 1	0.56	0.38	0.06	0.47	0.49	0.04	0.36	0.55	0.09
Хэвийн бус - 2	0.34	0.51	0.15	0.22	0.57	0.20	0.18	0.56	0.26
Үе тэнгийнхэн	s6			s19			s23		
Зааг - 1	0.48	0.42	0.10	0.48	0.44	0.08	0.33	0.59	0.09
Хэвийн бус - 2	0.23	0.55	0.22	0.22	0.60	0.19	0.16	0.61	0.23

Хүснэгт5-д харуулсанаар эмэгтэй сурагчдад дотоод илрэл сэтгэл хөдлөлийн ихэнхдээ толгой өвдөх (s3), олон зүйлд санаа зовних (s8), сэтгэлээр унах (S13), үе тэнгийнхэний харилцааны ганцаараа тоглох (S6), томчуудтай байх амар (S23) зэрэг шинж тэмдгээс хэвийн бус, шинэ орчинд сандрах (s16), айх (s24) зэрэг үзүүлэлтээс сэтгэл хөдлөл, зан үйлийн зааг төлөв үүсэх магадлалтай.

Эмэгтэй сурагчдад сэтгэл хөдлөлийн дотоод мэдрэмж илүү илэрч, эмгэг үүсэх төлөвтэй байна.

Эрэгтэй, эмэгтэй сурагчдад сэтгэл хөдлөлийн сэтгэлээр унах (S13), үе тэнгийнхэний харилцааны ганцаараа тоглох (S6), томчуудтай байх амар (S23) зэрэг дотоод сэтгэл хөдлөлийн үзүүлэлтээс хэвийн бус төлөв эмгэг үүсэх хандлагатай.

Хүснэгт V. Эмэгтэй сурагчдын дотоод сэтгэл хөдлөлийн мэдрэмжийн сэтгэл хөдлөл, зан үйлийн ерөнхий төлөв

Үзүүлэлт	Үнэн биш - 0, заримдаа үнэн - 1, яг үнэн – 2														
	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
Сэтгэл хөдлөл	s3			s8			s13			s16			s24		
Зааг - 1	0.39	0.51	0.10	0.05	0.53	0.42	0.34	0.58	0.09	0.16	0.58	0.26	0.19	0.53	0.28
Хэвийн бус - 2	0.18	0.56	0.26	0.01	0.34	0.64	0.12	0.60	0.28	0.08	0.42	0.51	0.06	0.43	0.52
Үе тэнгийнхэн	s6			s19			s23								
Зааг - 1	0.47	0.40	0.13	0.69	0.28	0.03	0.45	0.48	0.07						
Хэвийн бус - 2	0.23	0.51	0.26	0.42	0.46	0.12	0.27	0.55	0.17						

ХүснэгтVI-д харулсанаар эрэгтэй сурагчдад гадаад илрэл зан үйлийн дүрсхийн уурлаж, ердийн байдлаа алдах (s5), худлаа ярих, хууран мэхлэх (s18), хэт хөдөлгөөнтөхийн байнгын хөдөлгөөнтэй, тогтворгүй байх (s10) зэрэг шинж тэмдгээс хэвийн бус төлөв буюу эмгэг үүсэх төлөвтэй.

Хэт хөдөлгөөнтөхийн нэг газраа удаан байж чаддаггүй (s2), анхаарал нь амархан сарних (s15) зэрэг шинж тэмдгээс сэтгэл хөдлөл, зан үйлийн зааг төлөв үүсэх магадлалтай байна (Хүснэгт VI).

Хүснэгт VI. Эрэгтэй сурагчдын гадаад илрэлээс хамаарах сэтгэл хөдлөл, зан үйлийн ерөнхий төлөв

Үзүүлэлт	Үнэн биш - 0, заримдаа үнэн - 1, яг үнэн - 2											
	0	1	2	0	1	2	0	1	2	0	1	2
Зан үйл	s5			s12			s18			s22		
Зааг - 1	0.34	0.56	0.10	0.66	0.31	0.03	0.43	0.51	0.06	0.67	0.29	0.03
Хэвийн бус - 2	0.13	0.60	0.27	0.40	0.49	0.11	0.18	0.69	0.25	0.45	0.43	0.12
Хэт хөдөлгөөнтөх	s2			s10			s15					
Зааг - 1	0.20	0.64	0.16	0.43	0.50	0.07	0.12	0.66	0.23			

Хэвийн бус - 2	0.09	0.63	0.28	0.20	0.56	0.24	0.07	0.50	0.43	

Хүснэгт VII-д үзүүлсэнээр эмэгтэй сурагчдад гадаад илрэл зан үйлийн заримдаа худлаа ярих, хууран мэхлэх (s18) шинж тэмдгээс хэвийн бус төлөв эмгэг, дүрсхийн уурлаж, ердийн байдлаа алдах (s5) үед зааг буюу сэтгэл хөдлөл, зан үйлийн тулгамдсан асуудал үүсэж болзошгүй.

Хэт хөдөлгөөнтөхийн нэг газраа удаан байж чаддаггүй (s2), анхаарал нь амархан сарних (s15) зэрэг шинж тэмдгээс сэтгэл хөдлөл,зан үйлийн зааг төлөв үүсэх магадлалтай байна (Хүснэгт VII).

Хүснэгт VII. Эмэгтэй сурагчдын гадаад илрэлээс хамаарах сэтгэл хөдлөл, зан үйлийн ерөнхий төлөв

Үзүүлэлт	Үнэн биш - 0, заримдаа үнэн - 1, яг үнэн - 2								
	0	1	2	0	1	2	0	1	2
Зан үйл	s5			s12			s18		
Зааг - 1	0.23	0.67	0.11	0.81	0.18	0.01	0.49	0.43	0.08
Хэвийн бус - 2	0.10	0.55	0.35	0.62	0.33	0.05	0.28	0.54	0.18
Хэт хөдөлгөөнтөх	s2			s10			s15		
Зааг - 1	0.22	0.67	0.12	0.56	0.40	0.04	0.16	0.61	0.23
Хэвийн бус - 2	0.12	0.62	0.26	0.37	0.50	0.13	0.05	0.50	0.44

Эрэгтэй сурагчдад тэдний гадаад илрэлийн шинж тэмдэгээс шалтгаалан зааг, эмгэг төлөв үүсэх төлөвтэй.

Эрэгтэй, эмэгтэй сурагчдад гадаад илрэл болох зан үйлийн заримдаа худлаа ярих, хууран мэхлэх (s18) шинж тэмдгээс эмгэг, дүрсхийн уурлаж, ердийн байдлаа алдах (s5) үзүүлэлтээс зааг төлөв үүсэх хандлагатай.

IV. ХЭЛЦЭМЖ

Сурагчийн сангаас зөвхөн сэтгэл хөдлөлийн шинж чанарыг сонгон хөндлөн баталгаажуулалтаар 10 хуваан ангиллын аргуудыг туршихад Бейсийн загварын гүйцэтгэлийн хэмжүүрүүд 0.02-0.03 хооронд илүү, статистик ач холбогдлын туршилтаар тэд хоорондоо ялгаатай байв [10, 11].

Ye тэнгийхний харилцаа, хэт хөдөлгөөнтөхийн үзүүлэлтийг нэмэн ангиллын аргуудыг туршихад C50 алгоритмаар тооцсон мод загварын гүйцэтгэлийн хэмжүүрүүд 0.01-0.02 хооронд өндөр, загварууд хоорондоо статистик ач холбогдлын ялгаагүй гарсан. C50 алгоритмаар тооцсон загвараар хэт олон салаатай мод үүссэн тул тайлбарлахад хялбар, гүйцэтгэлийн хэмжүүрүүдийн зөрүү бага Бейсийн загвараар үр дүнг илэрхийлсэн [12].

Монгол Улсын сэтгэцийн эрүүл мэндийн салбарт өгөгдөл тандалтын ангиллын аргаар хийгдсэн судалгааны ажлууд хомс байна. Үүнээс дурдвал:
Улаанбаатар хот, Говь-Алтай аймагт 2017-2018 оны хооронд 7-12 насны 1789 хүүхдийг SDQ асуумжаар үнэлэн, логистик регрессээр шинжлэхэд эрэгтэй хүүхдүүд (AOR 1.64, 95% итгэлцүүр 1.29–2.10), ээжийн боловсролын түвшин бага (AOR 1.89, 95% итгэлцүүр 1.16–3.05), нойр бага (AOR 1.41, 95% итгэлцүүр 1.10–1.80), дасгал хийхгүй (AOR 1.31, 95% итгэлцүүр 1.03–1.67), дэлгэцийн хэрэглээ (AOR 1.53, 95% итгэлцүүр 1.20–1.94) зэргээс шалтгаалан сэтгэцийн эрүүл мэндийн асуудалтай байх магадлалтай гэж дүгнэсэн. Уг судалгаагаар хамрагдсан хүүхдүүдийн тоо 2 дахин бага, ашигласан арга нь өөр байна [13].

СЭМҮТөв 2018 онд SDQ асуумжийн эцэг эх, багш, хүүхэд хувилбарыг хот, хөдөөгийн ЕБС-ийн 11-18 насны 2250 хүүхдээс авч, сэтгэл хөдлөл, зан үйлийн эмгэгийн тархалтыг агшингийн аргаар, проспектив когорт аргаар 2013 оноос 2017 он хүртэл 82 хүүхдийг дагаж судлан эмгэгт өртөх эрсдэлт хүчин зүйлсийг судалсан. Охидод дотоод сэтгэл хөдлөлийн мэдрэмж, хөвгүүдэд зан үйл, хэт хөдөлгөөнтөх гадаад илрэл давамгайлдаг нь тэдний хүйсний онцлогтой холбоотой хэмээн дүгнэсэн байна [2].

Улаанбаатар хот болон Сэлэнгэ аймгийн 3000 гаруй хүүхдийг 2015-2017 оны хооронд SDQ асуумжаар үнэлж, нөлөөлөх хүчин зүйлсийг логистик регрессээр шинжлэхэд өрхийн эдийн засгийн байдал (β=−0.25, p<0.01), эцэг эхийн боловсрол (β=−0.18, p<0.05), болон хангалтгүй хооллолт (β=0.30, p<0.001) зэрэг хувьсагчид статистик ач холбогдолтой гарсан. Хүүхдүүдийн сэтгэцийн эрүүл мэндийг таамагласан Support Vector Machines (SVM) загварын нарийвчлал 78%, Precision 74%, Recall 70% буюу манай судалгааны үр дүнтэй ойролцоо байна [14]. Өсвөр үеийнхний сэтгэцийн эрүүл мэндийг үнэлэх SDQ асуумж, арга ашиглалт, загварын гүйцэтгэлийн хэмжүүрүүдийн үр дүн ойролцоо, регресс шинжилгээгээр нөлөөлөх хүчин зүйлийг таамагласанаараа ялгаатай байна.

V. ДҮГНЭЛТ

Сурагчийн санг хүйсээр нь хуваан, C50, бейс, дүрэмд үндэслэсэн аргаар дотоод болон гадаад илрэлээс хамаарах сэтгэл хөдлөл, зан үйлийн загваруудыг тооцоход бейсийн аргаар тооцсон загварын гүйцэтгэлийг үнэлэх хэмжүүр нарийвчлал 0.01-0.05, мэдрэмж 0.02-0.05, онцлог байдал 0.02-0.07 хооронд өндөр байна. Энэ нь өгөгдлийн сангийн мөр болон аттрибутуудын тоог ихэсгэж, багасгахад Бейсийн арга сайн ажиллаж байгаа тул үр дүнг илэрхийлсэн.

Сурагчдад сэтгэл хөдлөлийн сэтгэлээр унах, үе тэнгийнхэний харилцааны ганцаараа тоглох, томчуудтай байхад амар зэрэг дотоод илрэлээс

шалтгаалан хэвийн бус төлөв эмгэг үүсэх хандлагатай байна.

Зан үйлийн заримдаа худлаа ярих, хууран мэхлэх шинж тэмдгээс эмгэг, дүрсхийн уурлан, ердийн байдлаа алдах зэрэг гадаад илрэлээс зааг төлөв үүсэх хандлагатай.

Судалгааны үр дүнгээр гадаад илрэлийн шинж тэмдэг эрэгтэй сурагчдад, дотоод илрэлийн шинж тэмдэг эмэгтэй сурагчдад илүү байгаа нь дотоодын сэтгэцийн эрүүл мэндийн судалгаатай үр дүнтэй таарч байгаа ьулнь бидний судалгааны ажил амжилттай хэрэгжсэнийг харуулж байна.

Талархал

Энэхүү SDQ асуумжийг боловсруулсан хамтран гүйцэтгэгч СЭМҮТөвийн судлаач АУ-ны доктор В.Баярмаад гүнээ талархаж байна. Говь-Алтай аймгийн 3764 сурагч, тэдний асран хамгаалагч, ангийн багшид судалгааг амжилттай зохион байгуулахад тусалсанд талархал илэрхийлж байна. Судалгаагаар олж авсан өгөгдөлд боловсруулалт хийхэд хамтран ажилласан АШУҮИС-ийн Био-Анагаахын сургуулийн Физик, Мэдээлэл Зүйн хамт олондоо талархлаа.

Ном зүй

[1] UNICEF, "ДЭЛХИЙН ХҮҮХДИЙН АЖ БАЙДАЛ 2021," 2021. [Online]. Available: https://www.unicef.org/mongolia/media/3811/file/SOWC-2021-executive-summary-mn.pdf.

[2] В. Баярмаа, "Өсвөр насныхны сэтгэл хөдлөл, зан үйлийн судалгаа," Ph.D, 2019.

[3] R. Goodman, "The Strengths and Difficulties Questionnaire (SDQ)," in *Innovations in clinical practice: Focus on children & adolescents*: Professional Resource Press/Professional Resource Exchange, Inc, 2003, pp. 109-111.

[4] W. I. D. Mining, "Data mining: Concepts and techniques," *Morgan Kaufinann*, vol. 10, pp. 559-569, 2006.

[5] C. Yobero, "Determining Creditworthiness for Loan Applications Using C5. 0 Decision Trees," *RPubs by RStudio*, 2018.

[6] M. Balamurugan and S. Kannan, "Performance analysis of cart and C5. 0 using sampling techniques," in *2016 IEEE International Conference on Advances in Computer Applications (ICACA)*, 2016: IEEE, pp. 72-75.

[7] K. M. Leung, "Naive bayesian classifier," *Polytechnic University Department of Computer Science/Finance and Risk Engineering*, vol. 2007, pp. 123-156, 2007.

[8] M. Kuhn *et al.*, "Package 'caret'," *The R Journal*, vol. 223, no. 7, 2020.

[9] D. Meyer *et al.*, "Package 'e1071'," *The R Journal*, 2019.

[10] A. Tilyeubai *et al.*, "Significance Statistical Test Analysis on Classification Models of Adolescent's Emotional Problems," *Adv. Artif. Intell. Mach. Learn.*, vol. 3, no. 4, pp. 1743-1757, 2023.

[11] A. Tileubai *et al.*, "Study of decision tree algorithms: effects of air pollution on under five mortality in Ulaanbaatar," *BMJ Health & Care Informatics*, vol. 30, no. 1, 2023.

[12] Г. Батнаст, Т.Ахыт, Ц.Жавзмаа, У.Нямдаваа, and В.Баярмаа, "Өсвөр үеийнхний сэтгэл хөдлөл, зан үйлийн машин сургалтын загваруудыг харьцуулсан дүн," *Эрүүл мэндийн шинжлэх ухаан сэтгүүл,* 2025.

[13] A. Aoki *et al.*, "Socioeconomic and lifestyle factors associated with mental health problems among Mongolian elementary school children," *Social psychiatry and psychiatric epidemiology*, pp. 1-13, 2022.

[14] A. Dagvadorj *et al.*, "Prevalence and determinants of mental health problems among children in Mongolia: a population-based birth cohort," *Global Epidemiology*, vol. 1, p. 100011, 2019.

Analysis of Traffic Congestion in UB city, Mongolia using CNN and Agent-Based Modelling

Erdenetuya Erdenebileg
Department of telecommunications engineering
School of Information Communication Technology, MUST
Ulaanbaatar, Mongolia
erdenetuya.e@must.edu.mn

Altanzaya Tumen
Information security division
E-Mongolia academy
Ulaanbaatar, Mongolia
altanzaya@ema.gov.mn

***Abstract*—Urban traffic congestion poses significant challenges to the sustainable development of cities worldwide, and Ulaanbaatar, Mongolia, is no exception. With more than half of the population residing in Ulaanbaatar, the capital city, urbanization issues such as traffic jams, road accidents, and air pollution have recently reached extreme levels, significantly affecting residents' lives. This paper conducts a comprehensive network analysis to understand the dynamics of traffic congestion in Ulaanbaatar. To determine traffic patterns and congestion points, vehicles on the road were detected from satellite images using AI techniques, which were then integrated with GIS-based road network data and network analysis methods. Furthermore, the research included route analysis to find optimal routes, determine paths from accident sites to the nearest facilities, and identify service points. GIS mapping and network modeling were further integrated into Agent-Based Models to examine the complex interactions within the city's road network through dynamic modeling. This paper concludes that road congestion results from factors such as exceeding the traffic capacity of the road system, improper and inefficient road usage, ineffective traffic control, inadequacies in the transportation system, and a lack of awareness about road capacity.**

Keywords—GIS, satellite imaginary, CNN & Agent-Based Modelling, Analysis of Traffic congestion

I. Introduction

Ulaanbaatar, the capital city of Mongolia, has experienced rapid urban growth in recent years. This growth has led to significant increases in population and vehicle numbers, resulting in severe traffic congestion. The city's road infrastructure has struggled to keep pace with this rapid expansion, leading to daily traffic jams that impact residents' quality of life and the city's overall efficiency. The general road network of Ulaanbaatar city was designed to accommodate a population of 349,000 and 10,044 cars, which accounted for only 2.9% of the population owning cars. The city's maximum population capacity was estimated to be between 400,000 and 500,000 [1].

The road network of Ulaanbaatar spans 1,181.5 kilometers, with 711.3 kilometers consisting of major and minor roads, and 470.2 kilometers of residential roads [2]. Traffic congestion is caused by several factors, including the limited capacity of the road system compared to the heavy traffic volume. Furthermore, the city's underdeveloped public transportation system forces many residents to rely on private cars for travel. Understanding traffic flow is essential for developing effective solutions to manage and reduce congestion[3][5]. The aim of this paper is to analyze the current state of traffic congestion in Ulaanbaatar and support the development of strategies to improve traffic flow. Table 1 shows the basic road information of Ulaanbaatar city.

TABLE I. Basic road information of UB city

Name	Length of road	Total area	Road size per km^2	Total population
Ulaanbaatar	1181.5 km	470.4 km^2 (181.63 sq mi)	0.19	1,672,627

A. Literature review

Traffic congestion is a major issue in Ulaanbaatar, the capital of Mongolia, impacting numerous residents every day. This review explores researchers' findings on the causes of traffic congestion, its effects, and potential strategies for managing and reducing it in Ulaanbaatar.

Researchers use various methods to study traffic patterns in Ulaanbaatar [4] [6] [7] [8] [9] [10] [13]:

a) *Traffic Simulation Models:* These models help predict where and when traffic will be bad. They can identify which areas are the most congested.
b) *GPS Data Analysis:* Data from GPS devices in cars provides real-time information about traffic conditions.
c) *Remote Sensing and GIS:* Satellite images and Geographic Information Systems (GIS) are used to map traffic congestion and monitor changes over time.

Various strategies, such as enhancing public transportation and optimizing traffic signals, can help alleviate congestion. Learning from the experiences of other cities may offer valuable insights for Ulaanbaatar. Continued research and practical solutions are crucial to effectively tackle this issue.

B. Network analysis

Network analysis is a method of studying how different elements are connected and interact with each other. It involves examining a system of interconnected points (nodes) and the links (edges) between them. This type of analysis is applied in various fields, such as transportation, social sciences, and biology, to understand the structure and behavior of complex systems. Geographic Information System (GIS) software is used to determine the quickest or shortest routes between locations. Attribute data from the road network is collected to develop a database of the road network and identify the optimal route. The results of the analysis include detailed directions for traveling along the chosen route [12][14].

Planning the routes and schedules of city vehicles helps citizens reduce stress caused by traffic jams, save time otherwise wasted on congested roads, and ensures vehicles operate according to their schedules without delays.

There are three types of Network Analysis layers [11]:

1) *A Network Dataset Layer.* A Network Dataset Layer is a comprehensive digital representation of a network within a Geographic Information System (GIS). It includes all the elements that define how different points (nodes) are connected by pathways (edges). This dataset is essential for performing various types of network analysis, such as finding the shortest path, analyzing traffic flow, and determining service areas.

2) *A Route Layer.* A Route Layer in a Geographic Information System (GIS) is a specialized layer used to represent the optimal path or series of paths between specified points within a network. This layer is integral to network analysis, enabling users to determine the most efficient routes for transportation, logistics, and navigation.

3) *Closest Facility Analysis.* This is a spatial analysis technique used in Geographic Information Systems (GIS) to determine the nearest facility or point of interest from a specific location or a set of locations within a network. The analysis identifies the closest facilities based on defined criteria, such as distance, travel time, or cost. This technique is highly valuable across various fields, including emergency response planning, retail site selection, and urban development.

II. METHODLOGY

The objectives of this research are to use satellite imagery for traffic data generation, carrying facility and shortest path analysis, road congestion condition assessment, and traffic simulation to evaluate the optimal cost. The software used in research includes [15] [16] [17]:

- ArcGIS 10
- ArcMap ArcGIS Pro
- QGIS
- AnyLogic simulation software

This research is supported by both primary and secondary data to achieve specific objectives and address the research questions.

1) *Primary Data.* Google satellite imagery of Ulaanbaatar city from the year 2023 serves as the primary geographic data source. The sensor details are as follows:
   - Satellite Sensor Name: Google Satellite
   - Company: Maxar Technologies
   - Resolution: 15m
   - Bands: Blue, Green, Red
2) *Secondary Data.* Secondary vector data collection involved digitizing vector objects from Google satellite imagery to create road layers for the Ulaanbaatar city area.

The methodology of this research can be summarized as follows: it aims to manage high-resolution satellite images of research areas, particularly those with high demand. High-resolution satellite images were used as geographic references, with the dataset set to WGS 1984 and the projection set to UTM 48N. Two-dimensional editors in "ArcGIS Pro" and "ArcMap" were used to map main roads, secondary roads, vehicles in traffic on highways, and vehicles parked in parking lots. All relevant linear features were digitized, and after digitization, attribute data were assigned to the following categories: Road, Street Name, Start Elevation, Mid Elevation, End Elevation, TF Speed, FT Speed, TF Minutes, FT Minutes, and Meters. Network analysis tasks were performed to determine the optimal path and the shortest route between selected points. Additionally, to realistically represent traffic direction and road density on highways, an Agent-Based Simulation Model was created using data from traffic lights and main roads in Ulaanbaatar. Five cars were used as agents in the simulation.

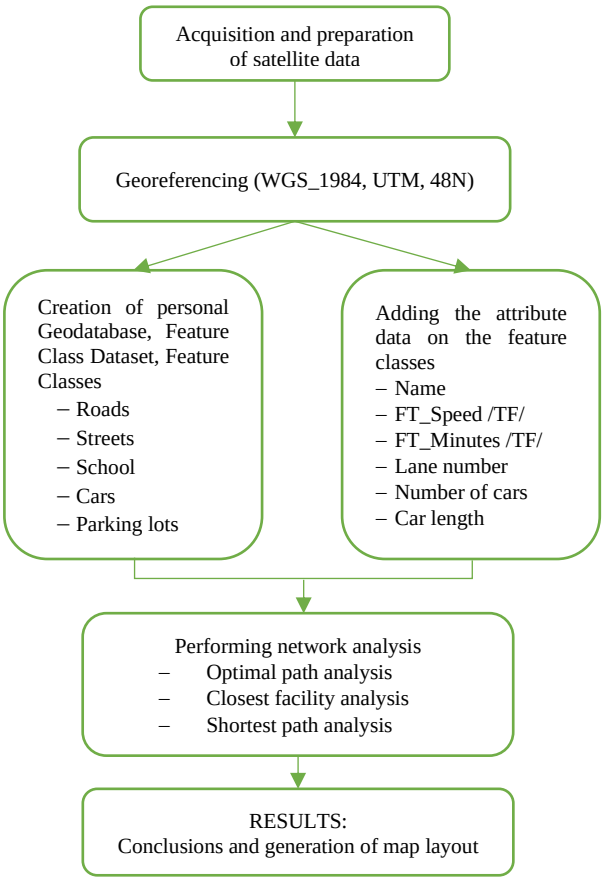


Fig. 1. Flowchart of Digitization

A. *Database preparation*

The ArcGIS Pro system was used to manage geographic data in our research [15]. At the beginning of the research, we created and prepared a geographic database containing my research data under the name "Network_Analysis_UB." Within this geographic database, a feature dataset named "Transportation" was created using the UTM projection with the WGS 1984 datum, set to zone 48. All feature classes within the dataset share the same coordinate system. These feature classes represent geographic features and their attributes as points, lines, and polygons, and may include markers and dimension values. Within the feature dataset, a feature class was created to define the data type of specific geographic features. This class includes additional information such as geometric properties, which serve as the main attribute table for further research. Since this attribute table plays a critical role in future analyses, it was created with careful attention to accuracy and completeness.

B. *Digitization of satellite imagery*

main_road_Layer1 | minor_road_Layer | narnii_zam_Layer | bridge_Layer

Field: Add | Calculate | Selection: Select By Attributes | Zoom To | Switch | Clear | Delete | Copy

OBJECTID *	SHAPE *	name	prefix	prefix_type	sufix	sufix_type	FT_speed	TF_speed	FT_minutes	TF_minutes	oneway	SHAPE_Length	cars	TF_cars	FT_cars	length_of_car	road_density	lane
13	Polyline	minor road	circus	circus	circus inersect left	intersect	30	<Null>	<Null>	<Null>	TF	62.282277	3	0	3	4.5	0.650328	3
14	Polyline	minor road	circus intersect	intersect	narnii zam	narnii zam	35	30	21.20632	21.20632	-	636.696805	38	12	26	4.5	0.537575	2
15	Polyline	minor road	peace bridge	bridge	narnii zam	narnii zam	40	35	10.95056	10.95056	-	383.768455	28	8	20	4.5	0.6575	2
16	Polyline	minor road	peace bridge left	bridge	narnii zam	narnii zam	35	35	9.575146	9.575146	-	335.381784	4	2	2	4.5	0.107421	2
17	Polyline	minor road	milaya	intersection	milaya right	intersection	30	<Null>	<Null>	<Null>	TF	74.408457	7	0	7	4.5	0.84863	2
18	Polyline	minor road	milaya right	intersection	120 bus stop	bus stop	30	<Null>	<Null>	<Null>	TF	165.844484	8	0	8	4.5	0.434179	2
19	Polyline	minor road	milaya right	intersect	naadam center	intersect	30	30	35.76831	35.76831	-	1024.199632	89	46	43	4.5	1.566658	4
20	Polyline	minor road	tunnel	tunnel	narnii zam	narnii zam	30	35	7.21651	7.21651	-	252.745768	18	5	13	4.5	0.641386	2
21	Polyline	tunnel	tunnel	tunnel	narnii guur	narnii guur	35	35	52.4521	52.4521	-	1841.566765	100	23	77	4.5	0.490243	2
22	Polyline	19 goliin zam	narnii guur	bridge	peace bridge roundabout	roundabout	35	45	31.93131	31.93131	-	1441.684282	155	98	57	4.5	1.941668	4
23	Polyline	tovchoonii zam	rockmon building	office	narnii zam	narnii guur	40	45	10.50158	10.50158	-	184.014301	92	38	54	4.5	8.999995	4
24	Polyline	ikh mongol zam	440 residence	apartment	ikh mongol	bar	45	40	6.90246	6.90246	-	277.121461	32	9	23	4.5	1.043106	2
26	Polyline	ikh mongol	ikh mongol	bar	circus intersection	intersection	45	<Null>	2.32646	2.32646	-	93.170901	2	0	2	4.5	0.193427	2
27	Polyline	minor road	SBD 1 khoroo	khoroo	TDB peace bridge	bank	40	35	8.073393	8.073393	-	283.081037	15	5	10	4.5	0.47776	2
30	Polyline	ard khan bank zam	ard khan bank	bank	ard khan bank front	bank	30	<Null>	<Null>	<Null>	TF	314.855894	25	0	25	4.5	0.358457	1
31	Polyline	jur ur zam	jur ur	bakery	barilagchdiin talbai	square	30	<Null>	<Null>	<Null>	TF	169.116025	7	0	7	4.5	0.186308	1
32	Polyline	mungun zaviya ariin zam	mungun zaviya	jewellery shop	department store	shopping mall	30	<Null>	<Null>	<Null>	TF	388.484058	19	0	19	4.5	0.441189	2
33	Polyline	department store ariin z...	adidas	sport shop	yves rocher	skin care shop	30	<Null>	<Null>	<Null>	TF	374.187964	30	0	30	4.5	0.72239	2
34	Polyline	urt tsagaan to ih delguur	urt tsagaan front inters...	shopping street	department store left	shopping mall	30	30	8.211186	8.211186	-	246.375773	25	3	22	4.5	0.913388	2
35	Polyline	mungun zaviya left	mungun zaviya lest side	jewellery shop	mungun zaviya back road	shopping mall	30	<Null>	<Null>	<Null>	TF	44.393169	2	0	2	4.5	0.202817	1
36	Polyline	ard state bank front to...	ard state bank	bank	mungun zaviya back road	jewellery shop	35	<Null>	<Null>	<Null>	TF	171.513424	3	0	3	4.5	0.078937	1
37	Polyline	mongol bank to montel	mongol bank left road	bank	gmobile	operator company	35	<Null>	<Null>	<Null>	TF	696.973029	40	0	40	4.5	0.258836	1
38	Polyline	ministry of MDDC back...	the mongolian stock e...	company	MDDC	ministry	35	30	2.878186	2.878186	-	86.620973	10	4	6	4.5	2.084646	4
40	Polyline	ard to gmobile road	gmobile	operator company	ard bus station left	bus station	45	30	2.540021	2.540021	-	76.483366	3	2	1	4.5	0.354328	2
41	Polyline	mongol bank to zoos go...	Mongol bank intersect...	bank	zoos goyol intersection	shop	45	30	7.800751	7.800751	-	234.062578	55	33	22	4.5	4.230362	4
42	Polyline	tedy road	tedy	mobile shop	urt tsagaan	shopping street	35	<Null>	<Null>	<Null>	TF	267.709217	80	0	80	4.5	4.034753	3
43	Polyline	naiman sharga road	naiman sharga	currancy exchange	holiday inn	hotel	35	<Null>	<Null>	<Null>	TF	273.457765	17	0	17	4.5	0.55959	2
44	Polyline	huuhdiin 100 dotor zam	circus back road	circus	az khur	book store	30	35	27.14867	27.14867	-	951.921795	73	21	52	4.5	0.691431	2
45	Polyline	huuhdiin 100 dotor zam	huuhdiin 100 bair	apartment	avilgatai temtseh gazar	company	30	35	10.439	10.439	-	366.213563	0	0	0	4.5	0	2
46	Polyline	huuhdiin 100 dotor zam	huuhdiin 100 bair	apartment	ayanchin outlet	shop	35	30	4.184598	4.184598	-	125.815649	6	2	4	4.5	0.430149	2
47	Polyline	UB ih delguuriin left road	the bull	hot pot restaurant	peace mall	shopping mall	35	30	10.44299	10.44299	-	313.345738	3	1	2	4.5	0.086182	2
48	Polyline	fountain right road	khaan buuz	restaurant	irish pub	lounge	30	<Null>	<Null>	<Null>	TF	283.06648	5	0	5	4.5	0.158999	2
49	Polyline	fountain left road	circus back intersection	intersection	department store	shopping mall	30	<Null>	<Null>	<Null>	TF	277.35444	11	0	11	4.5	0.357002	2

Fig. 2. Attribute table of Minor roads

Digitization of satellite images is the process of selecting the necessary image features and converting them into digital vector data for research purposes [15].

Steps:

1. Open ArcGIS Pro and add the geodatabase "Network_Analysis_UB."
2. Add the base layer, using high-resolution imagery.
3. Start editing using the Editor Toolbar and digitize linear, polygon, and point features, such as roads, streets, and facilities.

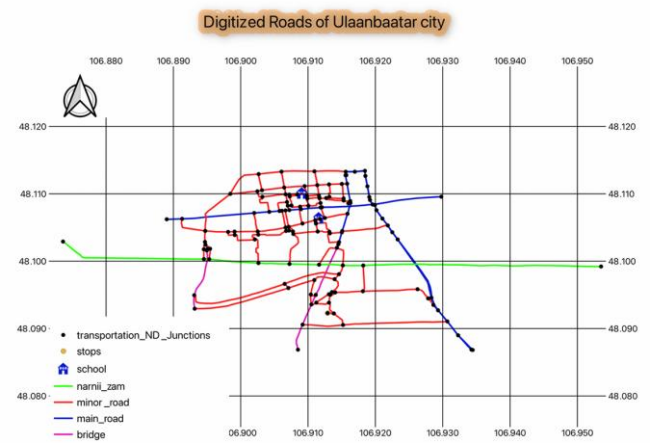


Fig. 3. Digitized roads

We added the necessary values to the digitized road network attribute table for my research, including FT_Speed, FT_Minutes, TF_Speed, TF_Minutes, Lane, FT_Cars, TF_Cars, total number of cars, and length_of_cars. This enhancement simplifies calculations and enables the extraction of more detailed information [Fig. 2].

C. *Extraction of vehicle from road using CNN models*

The Deep CNN YOLOv7 object detection algorithm is used to detect the locations of vehicles on the road and extract volumetric information. This technology has applications in traffic management and analysis, parking lot utilization, urban planning, and more. In the research, we used high-resolution satellite images of our research area and applied a deep learning model to extract data on the number of cars on the road and in parking areas, as well as their volume.

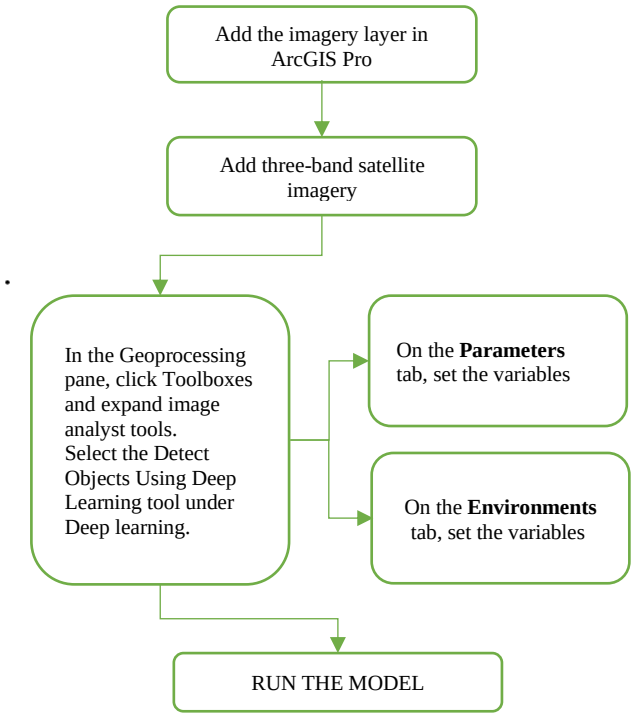


Fig. 4. Flowchart of CNN model

In figure 4, the Deep CNN YOLOv7 model demonstrates how object detection is used to identify cars on the road.

D. Results of CNN modelling

The Deep CNN YOLOv7 model is impressive when it comes to object detection, especially for identifying cars on the road. It builds on previous YOLO versions by enhancing accuracy and speed, making real-time detection more efficient. In practical applications, this technology is often used for traffic monitoring, autonomous driving, and safety analytics.

Fig. 5. It illustrates cars detected on the road using the Deep CNN Yolo-7 Object Detection algorithm.

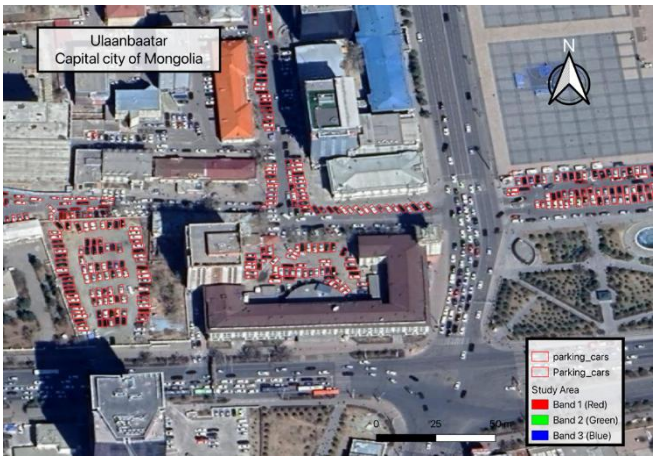


Fig. 6. It illustrates the Cars on Parking lot using the Deep CNN Yolo-7 Object Detection algorithm.

Fig. 7. It illustrates all the Cars on Road and Parking lot using the Deep CNN Yolo-7 Object Detection algorithm.

III. Simulation Model and Results

A. Modelling for validation

Although various definitions of Agent-Based Modeling (ABM) exist in the literature, from a practical application perspective, ABM can be simply defined as a decentralized, individual-centric approach to model design, as opposed to a system-level approach. When designing an Agent-Based Model, the modeler identifies the active agents—these agents can be people, companies, projects, assets, vehicles, cities, animals, ships, products, etc.—and defines their behavior, including key drivers, reactions, and history. The modeler then places them in a specific environment, establishes possible connections, and runs the simulation. The global (system-level) behavior emerges as a result of the interactions among many individual agents [9]. An Agent-Based Model is a computational model used to simulate the actions and interactions of autonomous agents in order to understand system behavior and the factors governing its outcomes.

In our research, vehicles on roads and in parking lots were detected using deep learning models. After detection, an Agent-Based Model was employed for throughput analysis, including generating congestion and traffic jam statistics. The model was also used to simulate traffic flow in Ulaanbaatar city, identifying directions with higher traffic concentration, analyzing traffic density, and demonstrating how vehicles slow down in response to traffic lights and congestion. To make the road network map more realistic, high-resolution satellite images are layered onto the digitized road network. Lane and direction information, along with certain turn restrictions at intersections, are incorporated to enhance accuracy.

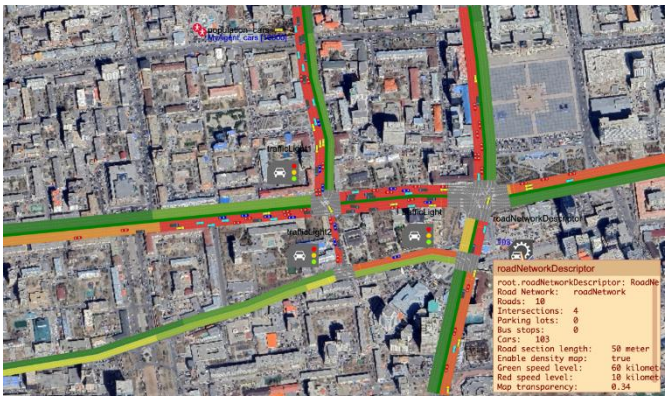


Fig. 8. Simulation results of a section of road

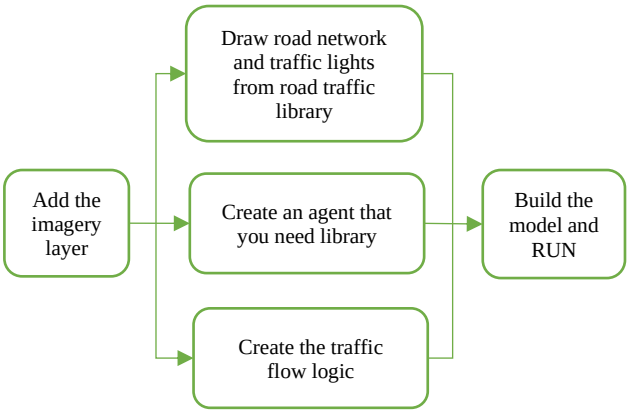


Fig. 9. Flowchart of Agent-Based Simulation Modelling

B. *Simulation parameters*

In the following tables, we present details of the Agent-Based Modeling used in our research.

TABLE II. DETAILS OF MODELLING

Name	Value
General	
Model time units	seconds
System Dynamics solver	
Differention Equations Method	Euler
Algebraic Equations Method	Modified Newton
Mixed Equations Method	RK45+Newton
Absolute accuracy	1.0E-5
Time accuracy	1.0E-5
Relative accuracy	1.0E-5
Fixed time step	0.001
Advanced	
Java package name	finalWork
File Name	/Users/altanzayatumen/Models/Network_Analysis_Altanzaya_final Work_1/Network_Analysis_Altanzaya_finalWork_1.alp

TABLE III. AGENT TYPE AND DETAILS

Name	Value
Agent in flowcharts	
Use in flowcharts as	Agent
Dimensions and movement	
Speed	(10 : MPS)
Rotate animation towards movement	true
Rotate vertically as well (along Z-axis)	false
Space and network	
Space Type	Continuous
Dynamic: Width	500
Dynamic: Height	500
Dynamic: z Height	0
Layout Type	User-defined
Layout Type Apply On Startup	true
Network type	User-defined
Network Type Apply On Startup	true
Enable steps	false
Advanced Java	
Generic	false
Advanced	
Logging	true
Auto-create datasets	true
AOC_DATASETS_UPDATE_TIME_PR OPERTIES	- Recurring Event Properties
Limit the number of data samples	false

TABLE IV. AGENT PARAMETERS

Name	Value
Arrivals defined by	self.RATE
Arrival rate	1000
Set agent parameters from DB	false
Limited number of arrivals	false
Appears	self.ON_ROAD
Road	road5
Enters	self.FORWARD_LANE
Random lane	true
New car	new Agent()
Length	4
Initial speed	60
Preferred speed	60
Max acceleration	1.8
Max deceleration	4.2
Add cars to:	false

TABLE V. PARAMETERS OF SIMUALTION EXPERIMENTS

Name	Value
General	
Bypass Initial Simulation Screen	true
Maximum available memory	512
Agent type	Main
Model time	
Execution mode	Real time with scale
Real time scale	1.0
Stop option	Never
Initial time	0.0
Initial date	Tue May 14 00:00:00 GMT 2024
Randomness	
Random Number Generation Type	Fixed seed (reproducible simulation runs)
Seed value	1
Window	
Title	test_1 : Simulation
Enable zoom and panning	true
Enable developer panel	true
Show developer panel on start	false
Advanced	
Load root from snapshot	false

C. *Network analysis and results*

Network analysis methodologies, incorporating techniques such as graph theory, spatial analysis, and computational algorithms, are used to study and interpret complex patterns in traffic networks. In the context of analyzing congestion in Ulaanbaatar, these methods help identify critical congestion points, assess traffic flow efficiency, and develop targeted strategies to alleviate congestion by understanding the underlying causes and dynamics of traffic movement in the city.

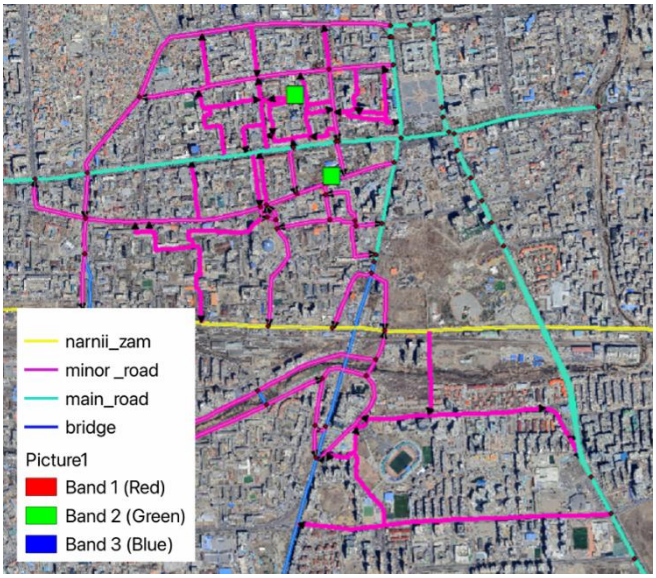


Fig. 10. Finding the optimal path processing

To obtain this result, the optimal path was determined by selecting two points—the route from home to school—and initiating the analysis.

Although multiple optional routes exist between the two selected points, the final route is typically chosen based on the fastest and most uninterrupted path. Additionally, travel time and distance are also evaluated.

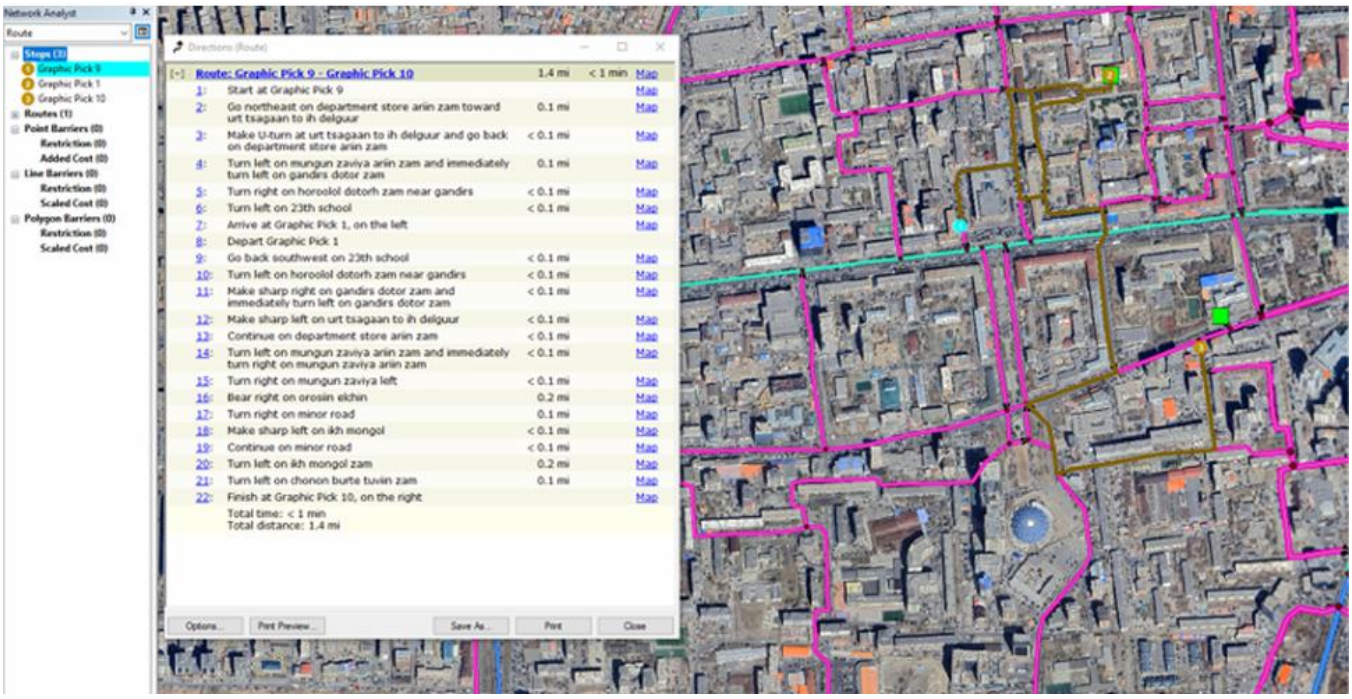


Fig. 11. Optimal paths of simulation

IV. Conclusion

In this paper, we aimed to identify the dynamic factors contributing to traffic congestion in Ulaanbaatar city. To achieve this, we developed an agent-based model using satellite imagery. Within this model, we employed a Convolutional Neural Network (CNN), specifically the YOLOv7 algorithm, to perform car detection. This allowed us to identify vehicles on roads and in parking areas [fig 5, 6], enabling the determination of the most densely congested areas of the road network [fig 7]. Furthermore, figure 8 presents the simulation results (green speed level of 60km/hr, red speed level of 10 km/hr & map transparency 0.34) for a road segment comprising 10 roads, 4 intersections, no parking lots or bus stops, and a total of 103 vehicles.

We also determined the shortest possible route between two selected locations based on road choices and vehicle density on those roads [fig 10, 11]. This represents the most efficient route in terms of travel time (total time < 1min, total distance 1.4 mi) to reach the desired destination. Figure 11 displays the optimal paths generated by the simulation, illustrating the route progression from Step 1 to Step 22. In terms of the detection process, the satellite image used in the research was too large, so it was divided into multiple smaller sections. The detection errors were evaluated to be minimal, not exceeding 20%.

In conclusion, applying quantitative network analysis to Ulaanbaatar's traffic congestion provides valuable insights into the city's transportation dynamics. By systematically analyzing the network, this methodology uncovers the root causes of congestion and helps develop effective strategies for improving traffic flow and reducing delays, ultimately contributing to a more efficient and livable urban environment.

References

[1] Davaa, S. (2017). Rapid Urbanization in Ulaanbaatar and Its Impact on Traffic Congestion. Journal of Urban Development, 23(2), 45-59.

[2] Batmunkh, D., Bayarsaikhan, T., & Narantuya, B. (2019). Increased Vehicle Ownership and Its Effects on Urban Traffic in Ulaanbaatar. Transportation Research Review, 18(3), 102-117.

[3] Boldbaatar, B. (2021). Infrastructure Challenges in Ulaanbaatar's Road Network. International Journal of Urban Infrastructure, 12(1), 33-48.

[4] Jargalsaikhan, N. (2018). Public Transportation Issues in Ulaanbaatar and Their Role in Traffic Congestion. Asian Journal of Transport and Infrastructure, 11(4), 87-99.

[5] Asian Development Bank. (2020). Impact of Traffic Congestion on Travel Times in Ulaanbaatar. ADB Urban Studies, 15(2), 20-32.

[6] Munkhbayar, L., Enkhbat, S., & Bat-Erdene, G. (2019). The Environmental Impact of Traffic Congestion in Ulaanbaatar. Journal of Environmental Health, 21(3), 56-68.

[7] World Bank. (2018). Economic Costs of Traffic Congestion in Ulaanbaatar. World Bank Urban Policy Reports, 22(5), 70-85.

[8] Enkhbayar, U. (2020). Traffic Simulation Models for Congestion Analysis in Ulaanbaatar. Journal of Traffic Management, 10(2), 44-58.

[9] Tserendulam, M. (2021). Utilizing GPS Data for Real-Time Traffic Analysis in Ulaanbaatar. International Journal of Smart Transportation Systems, 9(3), 99-112.

[10] Batzorig, S., Amarbayasgalan, T., & Bayarjargal, D. (2019). Remote Sensing and GIS Applications in Traffic Congestion Monitoring. Geospatial Journal, 14(4), 120-135.

[11] Enkhbaatar, B. (2019). Optimizing Traffic Signal Timings to Improve Flow at Ulaanbaatar Intersections. Traffic Engineering and Control, 17(1), 23-39.

[12] Batbayar, A. (2020). Enhancing Public Transportation to Alleviate Traffic Congestion in Ulaanbaatar. Urban Transport Solutions, 13(2), 55-69.

[13] Asian Development Bank. (2018). Infrastructure Development for Traffic Congestion Reduction in Ulaanbaatar. ADB Infrastructure Reports, 14(1), 40-53.

[14] Jargalsaikhan, N. (2018). Strategies for Traffic Demand Management in Ulaanbaatar. International Journal of Urban Planning and Policy, 16(3), 88-101.

[15] Esri. (n.d.). ArcGIS Pro. In Esri. Retrieved from https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview.

[16] Esri. (n.d.). ArcMap. In Esri. Retrieved from https://www.esri.com/en-us/arcgis/products/arcgis-desktop/overview.

[17] AnyLogic. (n.d.). AnyLogic: Multimethod Simulation Software for Business Applications. In AnyLogic. Retrieved from https://www.anylogic.com

ХИЙМЭЛ ОЮУН УХААНД СУУРИЛСАН RetCAD ПРОГРАММААР ЧИХРИЙН ШИЖИНГИЙН РЕТИНОПАТИЙН ЭРТ ИЛРҮҮЛЭГ ХИЙСЭН ҮР ДҮН

Пүрэвдорж, Баясгалан
Нүд Судлалын Тэнхим
АШУҮШИС, Анагаах Ухааны Сургууль
Монгол, Улаанбаатар хот
Bayasgalan.p@mnums.edu.mn

Эрдэнэбаатар, Түмэнжаргал
Нүд Судлалын Тасаг
Окулопластик эмнэлэг
Монгол, Улаанбаатар хот
Tumenjargal.e@mnums.edu.mn

Түлхүүр үг— Хиймэл Оюун ухаан, Сохрол, Чихрийн шижин, Нүдний уг, Эрсдэлт хүчин зүйл

УДИРТГАЛ

Чихрийн шижинтэй (ЧШ) хүмүүсийн тоо 1980 онд 108 сая байсан бол 2021 онд 537 сая болж өссөн. Цаашид чихрийн шижингийн тархалт 2030 он гэхэд 643 сая, 2045 он 783 саяд хүрнэ гэж тооцоолсон байна[1]. Дэлхийн хэмжээнд сохрол, сул хараа үүсгэх тэргүүлэх шалтгаануудын нэгд чихрийн шижингийн ретинопати орж байна.[2] Эрт илрүүлэг хийх нь уг эмгэгийн хүндрэлээс сэргийлэх хамгийн гол менежмент юм.[3] Манай оронд одоогоор нүд судлалын салбарт Чихрийн шижингийн ретинопати (ЧШР) оношилгоонд хиймэл оюун ухааны ашиглан хийсэн судалгааны ажил бүртгэгдээгүй байгаа нь бидний судалгааны ажлын үндэслэл боллоо.

Зорилго

Чихрийн шижингийн ретинопатийн эрт илрүүлэгд хиймэл оюун ухаанд суурилсан RetCAD программ хангамжийн оношилгооны нарийвчлалыг судлах.

Зорилт

RetCAD программ хангамж ашиглан хийсэн нүдний угийн өнгөт зургийн дүгнэлтийг харьцуулан судлах.

ЧШР-ийн ялгаатай үе шатуудыг тодорхойлоход RetCAD хиймэл оюун ухааны программын мэдрэг, өвөрмөц чанарыг үнэлэх.

Судалгааны шинэлэг тал

“RetCAD” нь convolutional neural networks (CNN) буюу нягтралтай мэдрэлийн сүлжээнд суурилсан гүн гүнзгий сургалтын загвар ашиглан, өнгөт нүдний угийн зургийг боловсруулж, чихрийн шижингийн ретинопати болон насжилттай холбоотой шар толбоны эмгэгийг өндөр нарийвчлалтайгаар илрүүлдэг. Бидний судалгаа нь эрүүл мэндийн салбарт өсөн нэмэгдэж буй хиймэл оюун ухааны салбарт хувь нэмрээ оруулж, ЧШР-ийн оношилгооны нарийвчлал, үр ашгийг дээшлүүлэх, уг дэвшилтэт технологийг Монгол хүнд судалж байгаа нь бидний судалгааны ажлын шинэлэг тал болно.

Материал, арга зүй

Судалгаанд нийт 19-82 насны 359 үйлчлүүлэгчийг тохиолдлын болон харьцангуй эрүүл бүлэгт санамсаргүй түүврийн аргаар сонгон хамрууллаа. Судалгаанд зайлшгүй шаардлагатай түүврийн хэмжээг тооцохдоо зураг 1-р харуулсан томьёог ашигласан.

$$n_{\min} = DEFF \frac{NZ^2 p(1-p)}{(N-1)e^2 + Z^2 p(1-p)}$$

n – судлагдах нэгжийн тоо
p – Судалж буй үзүүлэлтийн таамагласан тархалт (33%)
z – Стандарт хэвийн тархалтын онолын утга (1.92)
N – Эх олонлогийн хэмжээ (5000)
e – Алдааны хязгаар (0.05)

Fig. 1. Түүврийн хэмжээг олох томьё

Бүх үйлчлүүлэгчдэд нүдний эмчийн бүрэн үзлэг хийн чихрийн шижингийн ретинопатийн үе шат, хүндрэлийн зэргийг тодорхойлсон дүгнэлтийг хиймэл оюун ухааны хамгийн сүүлийн үеийн загвар RetCAD программын тоон хүснэгт 1-т харуулсны дагуу үнэлгээтэй харьцуулан агшингийн загвараар судаллаа.

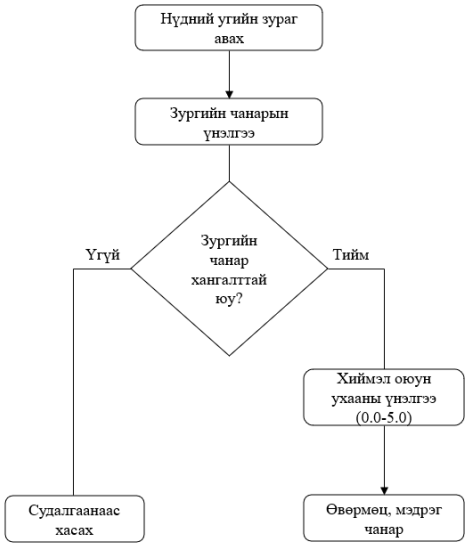


Fig. 2. RetCAD хиймэл оюун ухааны программын оношилгооны үйл явцын блок-схем

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

TABLE I. ОЛОН УЛСЫН НҮД СУДЛАЛЫН КОНСУЛААС ГАРГАСАН ЧШР-ИЙН АНГИЛАЛ БОЛОН RETCAD ПРОГРАММЫН ҮНЭЛГЭЭ

Ret- CAD үнэлгээ	**ЧШР-ийн ангилал**	**Нүдний угт илрэх өөрчлөлтүүд**
0-0.9	Илэрхий ЧШР байхгүй	Эмгэг өөрчлөлтгүй
1-1.9	ЧШР-ийн хөнгөн зэрэг	Зөвхөн бичил цүлхэнгүүд
2-2.9	ЧШР-ийн дунд зэрэг	Бичил цүлхэнгүүд ба бусад шинжүүд (цэгэн ба толбон цус харвалтууд, хатуу экссудатууд, хөвөн төст толбонууд) гэхдээ эдийн ургалтгүй ЧШР-ийн хүнд зэргээс бага
3-3.9	ЧШР-ийн хүнд зэрэг	Эдийн ургалтгүй ЧШР-ийн дунд зэрэг дээр илэрсэн шинжүүд, нэмээд дараах шинжүүдийн аль нэг нь: Торлогийн цус харвалт (квадрат тус бүрд ≥20) Илэрхий вений калибрын өөрчлөлт (2 квадратад) Торлогийн доторх бичил судасны гаж хөгжил (1 квадратад) Эдийн ургалттай ЧШР-ийн шинж байхгүй
≥4.0	Эдийн ургалттай ЧШР	Эдийн ургалтгүй ЧШР-ийн хүнд хэлбэр дээр илэрсэн шинжүүд, нэмээд дараах шинжүүдийн нэг нь эсвэл түүнээс олон: Шинэ судас Шилэнцэр/торлогийн өмнөх цус харвалт

Хиймэл оюун ухааны үнэлгээ нь зураг 2-т харуулсны дагуу нүдний угт илэрсэн өөрчлөлтүүдийн байршил зүй, тоон хэмжээ, нягтаршил, хэлбэрт үндэслэн 5.0 хүртэлх үнэлгээнүүдийг гаргана.

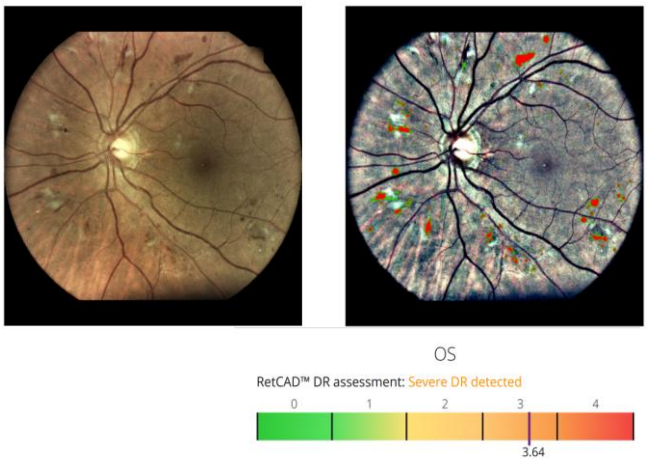


Fig. 3. Чихрийн шижингийн хүнд зэргийн ретинопати зүүн нүдний угийн өнгөт зураг, хиймэл оюун ухааны үнэлгээ

Үр дүн

Судалгаанд хамрагсдын дундаж нас 39.41±4.50 жил, ЧШ-тэй хүмүүсийн 40% нь эдийн ургалтгүй ЧШР-ийн хөнгөн зэрэг, 20% нь эдийн ургалтгүй ЧШР-ийн дунд зэрэг, 30.9% нь эдийн ургалтгүй ЧШР-ийн хүнд зэрэг, 9.1% нь эдийн ургалттай ЧШР оноштой байсан. Нүдний эмчийн дүгнэлт ба хиймэл оюун ухааны программын нийцэл хөнгөн зэргийн үед 93.5%, дунд зэргийн үед 87.5%, хүнд зэргийн үед 95.0%, эдийн ургалттай үед 100% байв. Нүдний угийн өнгөт зургийн үнэлгээнд RetCAD хиймэл оюун ухаан программ хангамж нь мэдрэг чанар 95.0%, өвөрмөц чанар 96.8%байв.

Дүгнэлт

RetCAD хиймэл оюун ухааны программ нь чихрийн шижингийн ретинопатийг богино хугацаанд олон хүнд мэдрэг, өвөрмөц чанар өндөртэй, харьцангуй алдаа багатай эрт илрүүлэг, оношилгоог хийж байна.

Хэлцэмж

RetCAD IDx-DR (Digital Diagnostics) 2018 онд АНУ-ын Хүнс, эмийн чанарын хяналтын газар /FDA/ зөвшөөрөгдсөн чихрийн шижингийн ретинопатийг илрүүлэх хиймэл оюун ухааны алгоритм. Мидриатик бус топкон нүдний угийн камер ашиглан авсан угийн өнгөт зургуудыг үүлэн сервер рүү илгээж үнэлгээ хийгддэг. Дунд ЭҮЧШР болон түүнээс дээш хэмжээний ЧШР-г илрүүлэхэд зориулагдсан. Мэдрэг чанар болон өвөрмөц чанар 89.5% ба 87.4%. Үүлэнд суурилсан хиймэл оюун ухааны сул тал нь мэдээллийн аюулгүй байдалд нэмэлт баталгаа шаардлагатай байдаг.[4]

EYEART(Eyenuk) 2020 онд АНУ-ын Хүнс, эмийн чанарын хяналтын газар /FDA/ зөвшөөрөгдсөн Чихрийн шижингийн ретинопатийг илрүүлэх хиймэл оюун ухааны алгоритм. Мидриатик бус канон нүдний угийн камераар авсан угийн өнгөт зургуудыг үүлэн сервер рүү илгээж үнэлгээ хийгддэг. Дунд ЭҮЧШР болон түүнээс дээш хэмжээний ЧШР-г илрүүлэхэд зориулагдсан. Мэдрэг чанар болон өвөрмөц чанар 94% ба 92%. Үүлэн системийн давуу тал нь адил нүдний угийн камераар авагдсан өнгөт зургуудад байршил хамаарахгүй үнэлгээ хийх боломжтой.[5]

Эдийн засгийн үр нөлөөний хувьд хиймэл оюун ухааныг ашиглан автомат үнэлгээг богино хугацаанд үр дүнтэй, богино хугацаанд хийж боломжийг бүрдүүлж байна. Иймд хиймэл оюун ухаанд суурилсан эрт илрүүлгийн арга нь эрүүл мэндийн ачааллыг бууруулж, ЧШР-ийн хүндрэлүүдийг үүсэх, даамжрахаас сэргийлэн үйлчлүүлэгчийн эмчилгээний зардлыг бууруулахад чухал үр нөлөөтэй болохыг харуулж байна.

Ном зүй

[1] Kumar A et al. Prevalence of diabetes in India: A review of IDF Diabetes Atlas 10th edition. Curr Diabetes Rev. 2023.

[2] Bourne et al. Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. The Lancet global health. 2021.

[3] Wong TY et al. Guidelines on diabetic eye care: the international council of ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings. Ophthalmology. 2018.

[4] Grzybowski A, Brona P. Analysis and comparison of two artificial intelligence diabetic retinopathy screening algorithms in a pilot study: IDx-DR and retinalyze. Journal of Clinical Medicine. 2021.

[5] Vought R, Vought V, Shah M, Szirth B, Bhagat N. EyeArt artificial intelligence analysis of diabetic retinopathy in retinal screening events. International Ophthalmology. 2023.d

ПОСТ КВАНТ КРИПТОГРАФИЙГ ЭМБЭДДЭД СИСТЕМД ХЭРЭГЖҮҮЛЭХ НЬ

Мөнхтөрийн МӨНХБАГАНА
Мэдээллийн Сүлжээ, Аюулгүй Байдлын салбар
ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль
Улаанбаатар, Монгол
baganaa237@gmail.com

Нэмэхбаярын ДАВХАРБАЯР
Мэдээллийн Сүлжээ, Аюулгүй Байдлын салбар
ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль
Улаанбаатар, Монгол
davkharbayr05@gmail.com

Дэмчигсүрэнгийн ГАНБААТАР
Мэдээллийн Сүлжээ, Аюулгүй Байдлын салбар
ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль
Улаанбаатар, Монгол
ganbaatar040617@gmail.com

Бат-Эрдэнийн МӨНХБАЯР
Мэдээллийн Сүлжээ, Аюулгүй Байдлын салбарын эрхлэгч, доктор(Ph.D), дэд профессор
Улаанбаатар, Монгол
munkhbayar.b@must.edu.mn

***Хураангуй*— Квант компьютерын хурдацтай хөгжил нь өнөөгийн криптографийн системүүдэд ноцтой аюул учруулж байна. RSA, ECC гэх мэт өргөн хэрэглэгддэг нийтийн түлхүүрт криптограф нь Шорын алгоритмд эвдэрнэ. Учир нь Шорын алгоритм нь квант компьютер дээр ажиллах үед бүхэл тоог үржигдэхүүнд задлах болон дискрет логарифмын бодлогыг экспоненциал бус, олон гишүүнт хугацаанд бодох чадвартай. Үүнийг шийдвэрлэхийн тулд пост-квант криптограф буюу квант компьютерт тэсвэртэй алгоритмууд хурдацтай хөгжүүлэгдэж байна. Судалгааны тайлангуудын тооцооллоор 2030 он гэхэд зүйлсийн интернэт (Internet of Things) төхөөрөмжийн тоо 29 тэрбумд хүрэх төлөвтэй бөгөөд үүний 98 орчим хувийг эмбэддэд системүүд эзлэх хандлагатай байна. Энэхүү судалгааны хүрээнд эмбэддэд системийн гүйцэтгэл, эрчим хүчний зарцуулалтад тохиромжтой Kyber-1024 пост-квант криптографийн алгоритмыг LWC-AES алгоритмтай хосолсон загвар ашиглан хэрэгжүүлсэн системийг боловсруулсан. Туршилтын явцад орчны температурыг хэмжихийн тулд DHT11 мэдрэгчийг ашиглана. Хэмжигдсэн мэдээллийг Raspberry Pi 4 SBC (Single Board Computer) дээр боловсруулж, LWS-AES алгоритмаар шифрлэнэ. Түлхүүр солилцоог хангахын тулд Kyber-1024 пост-квант алгоритмыг ашиглах бөгөөд шифрлэгдсэн өгөгдлийг дотоод сүлжээгээр сервер (зөөврийн компьютер) руу илгээнэ. Сервер хүлээн авсан мэдээллийг Kyber-1024 алгоритмаар солилцсон түлхүүр ашиглан тайлж, шифрлэлтээс гаргасан температурын утгыг дотоод сүлжээнд холбогдсон хэрэглэгчдэд веб интерфейсээр дамжуулан харуулна. Энэхүү судалгаа нь пост-квант криптографийг эмбэддэд систем дээр үр ашигтай хэрэгжүүлэх боломжийг тодорхойлж, бодит системүүдэд нэвтрүүлэхэд үнэтэй хувь нэмэр оруулах болно.**

***Түлхүүр үг*— Криптограф, Квант компьютер, Эмбэддэд систем, Кибер аюулгүй байдал.**

I. Удиртгал

1998 оноос эхлэлтэй квант компьютерын хурдацтай хөгжил нь Криптографийн салбарт өөрчлөлт авчирч байна. Шор болон Гроверийн алгоритмууд нь одоогийн өргөн хэрэглэгдэж буй асимметрик криптографийн алгоритмууд болох RSA, ECC зэрэг CNSA 1.0 стандартад багтах алгоритмуудыг эвдэх чадвартай [1].

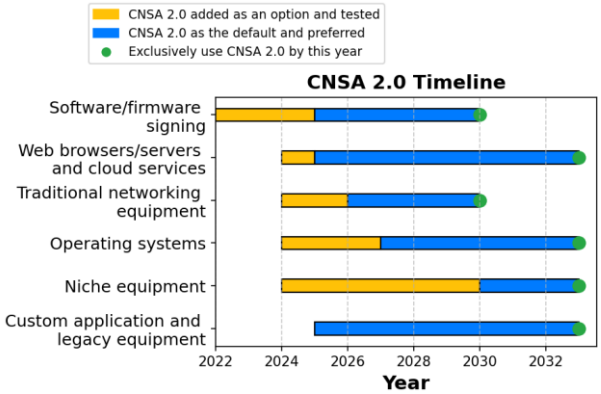
Криптограф нь мэдээллийн нууцлал, бүрэн бүтэн байдал, баталгаажуулалт зэрэг аюулгүй байдлын чухал шаардлагуудыг хангаж, цахим орчны найдвартай ажиллагааг дэмждэг. Банк, цахим худалдаа, харилцаа холбоо, аж үйлдвэрийн автоматжуулалт зэрэг олон чухал салбарт өргөнөөр ашиглагддаг бөгөөд мэдээллийн технологи болон кибер аюулгүй байдал дахь үндсэн технологийн нэг юм. 2015 онд National Security Agency (NSA) нь CNSA 1.0 криптограф стандартыг National Service Scheme (NSS)-ийг хамгаалах зорилгоор баталсан. АНУ-ын Үндэсний Стандарт ба Технологийн Хүрээлэн (NIST)-ийн тодорхойлсноор, тухайн стандарт нь квант тооцооллоос хамгаалагдсан шинэ алгоритмууд боловсруулагдах хүртэл ашиглагдах болно. Үүнээс өмнө Suite A болон Suite B криптографийн стандартууд ашиглагдаж байсан[1].

NIST нь мэдээллийн аюулгүй байдалд ойртож буй аюулыг таслан зогсоохын тулд 2016 оноос эхлэн квант компьютерын халдлагаас хамгаалах шинэ нийтийн түлхүүрт криптографийн стандарт боловсруулах ажлыг эхлүүлсэн нээлттэй сонгон шалгаруулалт явуулсан [2].

Сонгон шалгаруулалт дараах үе шатуудаас бүрдэнэ:

- Анхны шалгаруулалт (2016-2017): Шинэ алгоритмуудын саналыг хүлээн авч эхэлсэн.
- 1-р шат (2017-2019): Олон улсын Криптографийн судлаачид нийт 69 алгоритм санал болгосон.
- 2-р шат (2019-2020): 26 алгоритм сонгогдож, аюулгүй байдал болон гүйцэтгэлийн үнэлгээ хийгдсэн.
- 3-р шат (2020-2022): Сүүлийн шатанд 7 алгоритм шалгарсан.
- Албан ёсны сонголт (2022-2025): Олон улсын стандарт руу шилжүүлэх ажил хийгдсэн.

NSA нь NIST-ээс стандартчилахаар сонгосон алгоритмуудаас сонгон, 2024 оны 12-р сард CNSA 2.0 стандартыг баталсан. Үүнтэй холбоотойгоор, Зураг 1-т заасан хугацаанд Suite B болон CNSA 1.0 стандартын алгоритмуудыг ашиглаж буй бүх системүүд CNSA 2.0 руу шилжих шаардлагатай гэж заасан [3].



Зураг 1. CNSA2.0 Timeline

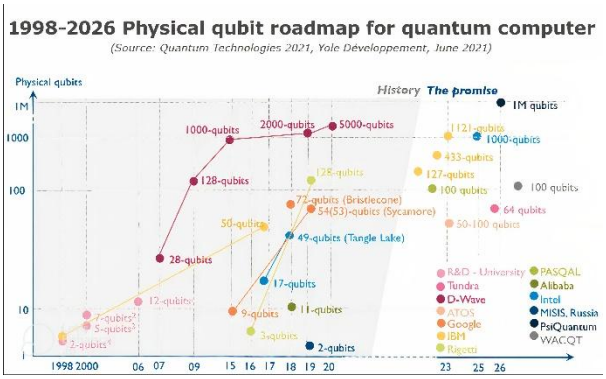
A. *Квант компьютер*

Квант компьютер нь бөөмсийн харилцан үйлчлэл, квант механикийн зарчмуудыг ашиглан мэдээлэл боловсруулах өндөр хүчин чадалтай тооцоолох төхөөрөмж юм. Энэ нь физикийн онолын нэг салбар болох квант механикт тулгуурладаг. [4]. Уламжлалт компьютероос ялгаатай нь мэдээллийг боловсруулах нэгж болох битийн оронд кубитийг ашигладаг. Кубит нь нэгэн зэрэг 0 болон 1-ийн төлөвт орших боломжтой ба суперпозиции (Superposition), квант орооцолдоо (Entanglement), квант интерференц (Quantum Interference), квант хэмжилт (Quantum Measurement) зэрэг квант механикийн шинж чанаруудыг ашиглан маш хурдан, үр ашигтай тооцоолол хийх боломжийг олгодог. [4].

Квант компьютерын гол онцлогууд:

- Суперпозиции – Кубитууд нь superposition зарчмаар ажилладаг бөгөөд нэгэн зэрэг 0 ба 1-ийн төлөвт байж болно. Энэ нь квант компьютерыг олон тооцооллыг зэрэг хийх боломжийг олгодог.
- Орооцолдоо – Хоёр ба түүнээс дээш кубитууд хоорондоо холбогдож, нэг кубитын төлөв нь нөгөө кубитын төлөвөөс шууд хамааралтай болох үзэгдэл юм. Энэ нь квант компьютерыг маш хурдан мэдээлэл боловсруулах боломжийг олгодог.
- Квант интерференц – Кубитуудын төлөвүүд хоорондоо харилцан үйлчлэлцэж, зөвхөн зөв хариултуудыг нэмэгдүүлж, буруу хариултуудыг хүчингүй болгодог. Энэ нь квант компьютерын үр ашгийг нэмэгдүүлдэг.
- Квант хэмжилт – Кубитийн төлөвийг классик 0 эсвэл 1 төлөв болгон хөрвүүлэх явц. Хэмжилт хийгдэсний дараа кубитийн суперпозиции алдагдаж зөвхөн нэг төлөв үлдэнэ[4].

Квант компьютерын хөгжил нь Криптографийн стандарт алгоритмуудад аюул учруулах ба 2035 он гэхэд RSA-2048 шифрлэлтийн алгоритмыг бүрэн эвдэлнэ[4]. Квант компьютерын хөгжлийн хурдацтай өсөлт болон ирээдүйд хүрэх боломжит түвшинг Зураг 2-т харуулав.



Зураг 2. 1998-2026 оны хооронд квант компьютерын физик кубитийн хөгжлийн замнал[5].

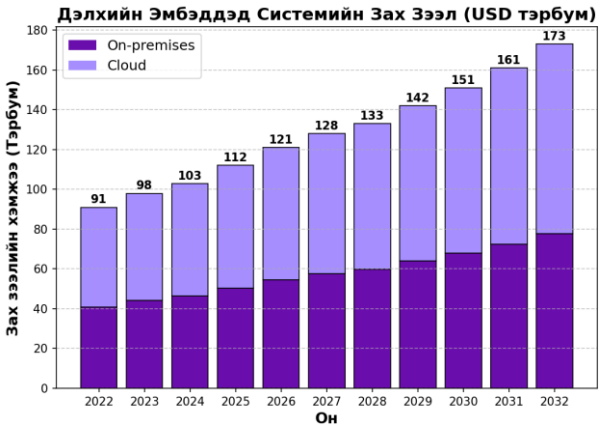
B. *Эмбэддэд систем*

Эмбэддэд Систем нь микропроцессор эсвэл микроконтроллер дээр суурилсан, тодорхой нэг үүрэг гүйцэтгэх зориулалттай, техник хангамж болон программ хангамжийг хослуулсан тооцооллын систем юм[6]. Эмбэддэд систем нь ерөнхий зориулалтын компьютертой харьцуулахад тодорхой нэг үүрэг гүйцэтгэх тусгай зориулалттай бөгөөд ихэвчлэн жижиг хэмжээтэй, хөнгөн жинтэй, бага эрчим хүч хэрэглэдэг тул батарейгаар ажиллах боломжтой. Мөн зарим системүүд бодит цагийн хариу үйлдэл хийх шаардлагатай байдаг ба ихэвчлэн ROM санах ой, Флэш санах ой ашигладаг, энгийн үйлдлүүдийг гүйцэтгэхээр программчилсан тул өндөр хүчин чадал шаарддаггүй өгөгдөл цуглуулах, дамжуулах үүрэгтэй ажилладаг.

Эмбэддэд системийн үйлдвэрлэл хурдацтай өссөөр байгаа бөгөөд үүнийг :

- Хиймэл оюун ухаан (Artificial Intelligence)
- Машин сургалт (Machine learning)
- Зүйлсийн интернэт (Internet of Things)

зэрэг технологиудын хөгжүүлэлт улам бүр хурдасгаж, үйлдвэрлэгдсэн бүх микропроцессорын 98 орчим хувь нь эмбэддэд системд ашиглагдаж байна [7]. Зураг 3-т үзүүлснээр эмбэддэд системийн зах зээлийн үнэлгээ 2022 онд 91 тэрбум доллар байсан бөгөөд 2032 он гэхэд 173.4 тэрбум гаруй доллар болж өсөх төлөвтэй байна[7].



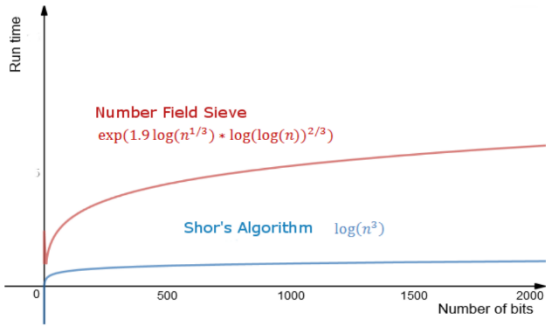
Зураг 3. Дэлхийн эмбэддэд системийн зах зээлийн өсөлт[7].

Урт хугацааны ашиглалттай эмбэддэд систем нь 30-аас дээш жилийн настай тул ирээдүйд хангалттай хүчирхэг квант компьютер бий болох үед ч ашиглагдаж байх магадлалтай. Ингэснээр "Harvest Now, Decrypt Later" буюу одоо цуглуул, дараа нь тайл гэх эрсдэл үүснэ. Одоогоор тайлах боломжгүй ч, ирээдүйд квант компьютер ашиглан шифрийг задлах боломжтой тул эсрэг талууд шифрлэгдсэн мэдээллийг урьдчилан цуглуулж, хадгалах арга ашиглаж болно.

C. Шорын алгоритм

1994 онд Белл Лабс-ын судлаач Питер Шор квант компьютер дээр олон оронтой тоог олон гишүүнт хугацаанд ($O(n^a)$; $1 \leq a$) задлах алгоритмыг боловсруулсан ба Шорын алгоритм (Shor's Algorithm) хэмээн нэрлэгддэг.

Уламжлалт компьютер олон оронтой тоог (жишээлбэл, 2048 битийн урттай тоо) олон гишүүнт хугацаанд задалж чадахгүй. Одоогоор хамгийн хурдан нийтийн алгоритм нь NFS (Number Field Sieve) ба экспоненциал (жишээлбэл, $O(2^n)$) хугацаанд ажилладаг[8].



Зураг 4. Том тоог задлах хугацаа : NFS болон Shor's алгоритм[8].

Шорын алгоритмын үндсэн санаа нь квант Фурьегийн хувиргалт (Quantum Fourier Transform, QFT) ашиглан үе олох процесс юм[8].

Шорын алгоритм нь нийт 11 алхамтай ба үндсэндээ QFT ашиглан r болон x тоог дараах томьёогоор тооцоолно:

$$\frac{1}{\sqrt{||A||}} \sum_{a' \epsilon A} \frac{1}{\sqrt{q}} \sum_{c=0}^{q-1} |c, k\rangle * e^{2\pi i a' c/q} \quad (1)$$

Сүүлчийн алхам нь дараах томьёогоор хийгддэг бөгөөд үүний үр дүнд том тооны анхны хуваагчийг олох боломжтой. Хамгийн их ерөнхий хуваагч (Greatest Common Divisor) тооцоолол нь уламжлалт компьютер дээр маш хурдан хийгддэг. [8].

$$gcd\left(x^{r/2}+1, n\right) \; and \; \gcd\left(x^{r/2}, -1n\right) \quad (2)$$

Хүснэгт 1. Квант алгоритмууд.

Алгоритм	**Он**	**Үүрэг, хэрэглээ**	**Ашиглагдаг салбар**
Deutsch-Jozsa algorithm	1992	Функцийн шинж чанарыг нэг хэмжилтээр тодорхойлох	Квантын мэдээллийн онол
Simon's algorithm	1994	Шорын алгоритмын үндэс болсон, экспоненциал давуу талтай	Криптограф, квантын онол
Shor's algorithm	1994	Анхны тооны задрал	Криптограф, аюулгүй байдал
Quantum Fourier Transform, QFT	1994	Шорын алгоритмын гол бүрэлдэхүүн хэсэг, Фурье хувиргалт хийх	Сигнал боловсруулалт, квант симуляци
Grover's algorithm	1996	O(√N) хугацаанд хайлт хийх	Өгөгдөл хайлт, Хиймэл оюун ухаан, Криптограф
Narrow-Hassidim-Lloyd algorithm	2009	Шугаман тэгшитгэлийн системийг квант аргаар шийдэх	Физик, Машин сургалт, Санхүү
Quantum Game Theory	1999–2000	Квант стратеги, Nash equilibrium загварчлал	Эдийн засаг, тоглоомын онол
Quantum Machine Learning, QML	2013-2017	Квант алгоритмуудыг машин сургалтад ашиглах (QAOA, VQE гэх мэт)	Хиймэл оюун ухаан, Өгөгдөл шинжилгээ, Биоинформатик

Шорын алгоритм нь квант компьютерын цорын ганц хэрэглээ биш юм. Хүснэгт 1-т 1992 оноос эхлэн хөгжүүлэгдсэн алгоритмуудыг ашиглагдах салбарын хамт харуулав. Криптографийн салбарт Гроверийн

алгоритм (Grover's Algorithm) ч бас чухал нөлөө үзүүлдэг бөгөөд n урттай дараалаас дурын элементийг $O(\sqrt{n})$ хугацаанд хайлт хийн олох давуу талтай[8].

D. *Пост квант криптограф*

NIST нь криптографийн салбарт үүсээд буй асуудлыг шийдэхийн тулд шинэ алгоритмуудыг сонгон шалгаруулсан бөгөөд CNSA 1.0 стандарт дахь нийтийн түлхүүрт алгоритмуудыг үндсэн дөрвөн шалгуурыг давсан пост-квант алгоритмуудаар шинэчлэн, CNSA 2.0 стандартад тусгасан. Хүснэгт 2-т шалгуурыг давсан алгоритмуудын мэдээллийг харуулав[9].

Хүснэгт 2. Сонгон шалгаруултад шалгарсан Пост-квант алгоритмууд.

Алгоритм	Хэрэглээ	Математик суурь	Төлөв
CRYSTALS -Kyber	Түлхүүр солилцоо/KEM/	Лэттис - суурьт	Стандартчилаг -дасан
CRYSTALS -Dilithium	Дижитал гарын үсэг /DSA/	Лэттис - суурьт	Стандартчилаг -дасан
FALCON	Дижитал гарын үсэг /DSA/	Лэттис - суурьт	Стандартчилаг -дасан
SPHINCS+	Дижитал гарын үсэг /DSA/	Хэш - суурьт	Стандартчилаг -дасан
BIKE	Түлхүүр солилцоо/KEM/	Код - суурьт (QC-MDPC)	Шинэ шатны судалгаа
Classic McEliece	Түлхүүр солилцоо/KEM/	Код - суурьт	Шинэ шатны судалгаа
HQC	Түлхүүр солилцоо/KEM/	Код - суурьт	Шинэ шатны судалгаа

Эдгээр алгоритмуудыг сонгохдоо дараах хүчин зүйлсийг үнэлсэн :

• Аюулгүй байдал – Квант болон уламжлалт компьютерт хоёуланд нь тэсвэртэй байх.

• Гүйцэтгэл – Мэдээллийг шифрлэх, тайлах хурд өндөр байх.

• Нийцтэй байдал – Өнөөгийн мэдээллийн системд хялбархан интеграци хийх боломжтой байх.

• Туршилт, баталгаажуулалт – олон улсын судлаачдын нээлттэй үнэлгээ[9].

Хүснэгт 3. CNSA 1.0 стандарт[1].

Алгоритм	Функц	Стандарт	Параметр
Advanced Encryption Standard (AES)	Блок шифрлэлт	FIPS Pub 197	256 битийн түлхүүр
Elliptic Curve Diffie-Hellman (ECDH) Key Exchange	Түлхүүр солилцох нийтийн түлхүүрийн алгоритм	NIST SP 800-56A	P-384 муруй
Elliptical Curve Digital Signature Algorithm (ECDSA)	Цахим гарын үсгийн нийтийн түлхүүрийн алгоритм	FIPS Pub 186-4	P-384 муруй
Secure Hash Algorithm (SHA)	Хэш алгоритм	FIPS Pub 180-4	SHA-384 хэш
Diffie-Hellman (DH) түлхүүр солилцоо	Түлхүүр солилцооны алгоритм	IETF RFC 3526	Хамгийн багадаа 3072-битийн модуль
Rivest-Shamir Adleman (RSA)	Түлхүүр үүсгэх алгоритм	NIST SP 800-56B rev 1	Хамгийн багадаа 3072-битийн модуль
Rivest-Shamir Adleman (RSA)	Цахим гарын үсэгт хэрэглэгддэг нийтийн түлхүүрийн алгоритм	FIPS PUB 186-4	Хамгийн багадаа 3072-битийн модуль

2015 онд батлагдсан CNSA 1.0 стандартад ашиглагдсан криптографийн алгоритмуудыг хүснэгт 3-т харуулав. Энэхүү стандарт нь асимметрик болон симметрик криптографийн хамгийн өндөр аюулгүй байдлыг хангах зорилготой[1].

Асимметрик криптографийн хувьд:

- CNSA 1.0 стандарт нь RSA болон ECC алгоритмуудыг ашигладаг.
- Эдгээр алгоритмууд анхны тооны задрал болон дискрет логарифмын математикийн бодлогод тулгуурладаг.

Квант компьютерын Шорын алгоритм нь RSA болон ECC-г хурдан хугацаанд эвдэх чадвартай учраас эдгээр алгоритмууд найдвартай бус гэж тооцогдох болсон. Иймд, 2024 оны 12 сард CNSA 2.0 стандартад пост-квант криптограф (PQC) буюу квантын халдлагад тэсвэртэй алгоритмууд багтааж, аюулгүй байдлын шинэ шаардлагад нийцүүлэн сайжруулсан[10].

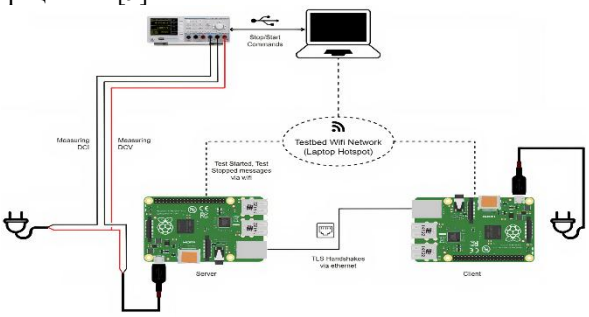
Хүснэгт 4. CNSA 2.0 стандарт[1].

Алгоритм	Функц	Стандарт	Параметр
Advanced Encryption Standard (AES)	Симметрик блок шифрлэлтийн алгоритм	FIPS PUB 197	256 бит түлхүүр
ML-KEM (CRYSTALS-Kyber)	Түлхүүр солилцох асимметрик алгоритм	FIPS PUB 203	ML-KEM-1024
ML-DSA (CRYSTALS-Dilithium)	Цахим гарын үсгийн асимметрик алгоритм	FIPS PUB 204	ML-DSA-87
Secure Hash Algorithm (SHA)	Хэш алгоритм	FIPS PUB 180-4	SHA-384, SHA-512
Leighton-Micali Signature (LMS)	Цахим гарын үсгийн асимметрик алгоритм	NIST SP 800-208	LMS SHA-256/192
Xtended Merkle Signature Scheme (XMSS)	Цахим гарын үсгийн асимметрик алгоритм	NIST SP 800-208	*
Secure Hash Algorithm 3 (SHA3)	Хэш алгоритм	FIPS PUB 202	SHA3-388, SHA3-512

II. ӨМНӨ НЬ СУДЛАГДСАН АЖИЛ

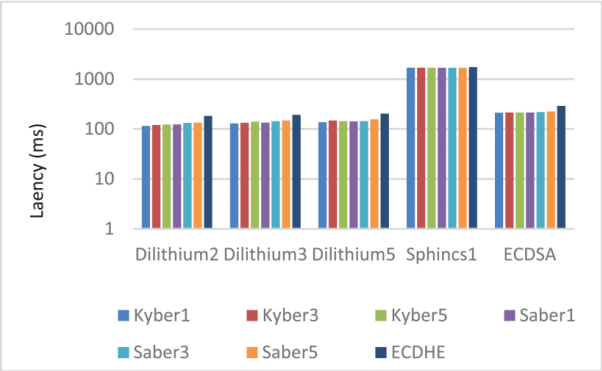
Эмбэддэд системийн тооцоолол болон эрчим хүчний зарцуулалтыг үнэлэхэд техник, төхөөрөмж хангалтгүй байсан тул IEEE-ээс 2024 онд нийтлэсэн "*Evaluation of performance, energy, and computation costs of quantum-attack resilient encryption algorithms for embedded devices. IEEE Access, 12, 8791-8805.*" судалгааны ажлаас иш татсан болно.

Энэхүү судалгааны туршилтын орчин нь хоёр Raspberry Pi 3b+ төхөөрөмжийг Cat5e Ethernet кабелаар шууд холбосон. Нэг төхөөрөмж нь сервер, дараагийн төхөөрөмж нь клиентийн үүрэг гүйцэтгэнэ[9].



Зураг 5. Raspberry Pi pico W үйлчлүүлэгчийн цахилгааны хэмжилтийн схем[9].

Raspberry Pi клиент тест орчны үр дүн пост-квант алгоритмуудын гүйцэтгэлийг нарийн харьцуулах зорилгоор түлхүүрийн энкапсуляци механизмууд болон цахим гарын үсгийн схемүүдийн бүх боломжит хослолуудыг хэрэгжүүлж, үнэлгээг хийсэн байсан. Тус бүрийн тохиолдолд латенси, санах ойн хэрэглээ, эрчим хүчний зарцуулалт зэргийг тооцоолсон хэмжилтээс иш татав.



Зураг 6. Raspberry Pi 3b (Клиент) дээрх TLS хэндшэйк латенси [9].

5-р хүснэгт нь судалж буй түлхүүр хүлээж авах алгоритмуудтай холбоотой латенси талаар илүү дэлгэрэнгүй харуулсан. Үүнд эдгээр схемүүдийн TLS протоколын тооцооллын латенси -тай холбоотой мэдээлэл багтана. Эдгээр тоонууд нь Kyber 1 нь шинэчлэгдсэн нийтийн түлхүүрийн алгоритмуудаас хамгийн сайн гүйцэтгэлтэй бөгөөд Saber5 нь хамгийн удаан хүлээгдэлтэй болохыг харуулж байна. Хүснэгтэд пост-квантум нь KEM-ийн аюулгүй байдлын түвшин нэмэгдүүлэх нь латенсын өсөлтөд хүргэдэг бөгөөд энэ нь аюулгүй байдлын өндөр түвшний схемүүдийн хувьд тооцооллын зардал илүү өндөр байх хандлагатай байдаг [9]. Kyber болон Saber-г харьцуулахад Kyber нь өөрийн алгоритмыг ажиллуулахад богино хугацаа зарцуулдаг болохыг харуулж байна.

Хүснэгт 5. TLS-д ашиглагдах KEM Алгоритмуудын харьцуулалт (латенси)[9].

Алгоритм	Түлхүүрийн энкапсуляци (Клиент) (мс)	Түлхүүрийн декапсуляци (Сервер) (мс)	Түлхүүр үүсгэх (Сервер) (мс)
Kyber1	2.68	2.4	1.57
Kyber3	4.52	3.93	2.78
Kyber5	6.88	5.44	4.15
Saber1	6.62	6.12	3.34
Saber3	11.99	10.45	6.52
Saber5	14.99	15.97	10.98

Хүснэгт 6. Нөөц хязгаарлагдмал төхөөрөмжүүд[9].

Development Board	MCU (Processor)	Freq. (MHz)	SRAM (KB)	Flash (KB)
Raspberry Pi 3B+	ARM Cortex-A53 (Quad-core)	1400	ARM Cortex-A53 (Quad-core)	microSD-based storage
Arduino Nano 33 IoT	SAMD21	48	32	256
Arduino Nano RP2040 Connect	RP2040 (2 ARM Cortex-M0+)	125	264	16000
Arduino Nano 33 BLE Sense	nRF52840 (ARM Cortex-M4)	64	256	1000
Raspberry Pi Pico W	RP2040 (2 ARM Cortex-M0+)	125	264	16000
ESP32-S2 DevKit	ESP32-S2 (Xtensa LX7)	240	512	384
STM32-F4 Series	STM32F4 (ARM Cortex-M4)	180	384	512-2056
STM32-F7 Series	STM32F7 (ARM Cortex-M4/M7)	216	64-248	256-512

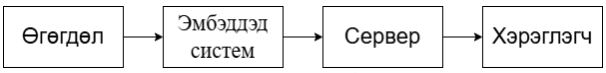
Эдгээр төхөөрөмжүүд нь тооцооллын чадвар болон нөөцийн хувьд бага ч, өргөжиж буй эмбэддэд системд өргөн ашиглагддаг. 2021 онд RP2040 процессор үйлдвэрлэлийн эхний жилдээ 10 сая орчим үйлдвэрлэгдсэн амжилтаас илэрхийлэгдэж байна. RP2040 нь зөвхөн 264KB SRAM агуулдаг бөгөөд 16MB флэш хүртэл дэмжиж, эрэлт ихтэй байдаг[9].

ARM Cortex M0+ нь хамгийн хязгаарлагдмал, эрчим хүчний хэмнэлттэй 32-бит процессор бөгөөд төхөөрөмжүүдийн доод түвшний хэрэгцээнд өргөн хэрэглэгддэг. Илүү хүчирхэг Cortex-M4, M7 зэрэг процессоруудын оронд энэ төрлийн процессорыг ашиглах нь туршилтын орчинд илүү тохиромжтой.

Дээрх судалгаанаас дүгнэхэд Raspberry Pi Pico W ба Arduino Nano RP2040 Connect нь сүлжээний чадвартай, том флэш сантай хөгжүүлэлтийн самбарууд юм. Pico W нь OpenOCD, GDB ашиглан SWD-ээр гадаад дебаг хийхэд хялбар, харин Arduino нь pad холболттой тул интеграци хийхэд төвөгтэй. Иймээс, хязгаарлагдмал MCU, бага өртөг, сайтар баримтжуулалт, хялбар программчлал зэргээс шалтгаалан Raspberry 4-ийг хэрэгжүүлэлтийн ажилдаа сонгосон.

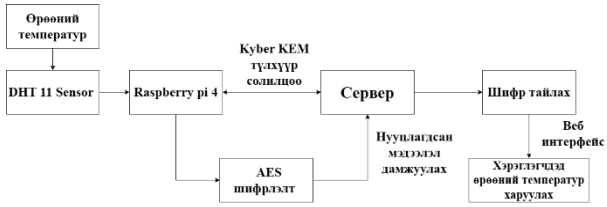
III. САНАЛ БОЛГОХ МЕХАНИЗМ

Судалгааны хүрээнд хэрэгжүүлж буй систем нь эх өгөгдлийг суурилагдсан систем буюу эмбэддэд системийн тусламжтай нууцлан мэдээллийг сервер (зөөврийн компьютер) рүү дамжуулах юм. Сервер өгөгдлийн нууцлалыг тайлан цааш хэрэглэгчдэд илгээх үндсэн үйлдлийг гүйцэтгэнэ.



Зураг 7. Ажиллагааны бүтцийн схем

Орчны температур болон чийгшлийн мэдээллийг хэмжихэд DHT-11 төрлийн мэдрэгчийг ашиглана. Уг мэдрэгч нь секунд тутамд хэмжилт хийж, мэдээлэл дамжуулах боломжтой бөгөөд судалгаа, туршилтад өргөнөөр ашиглагддаг. Эмбэддэд системийн сонголтыг 6-р хүснэгт болон [9] судалгааны ажлын үр дүнд тулгуурлан хийж, Raspberry Pi 4 SBC ашигласан. Эмбэддэд систем нь температур болон чийгшлийн мэдээллийг Lightweight Cryptography AES (LWC-AES) алгоритмаар шифрлэнэ.



Зураг 8. Мэдээлэл нууцлалтай дамжуулах бүтцийн схем

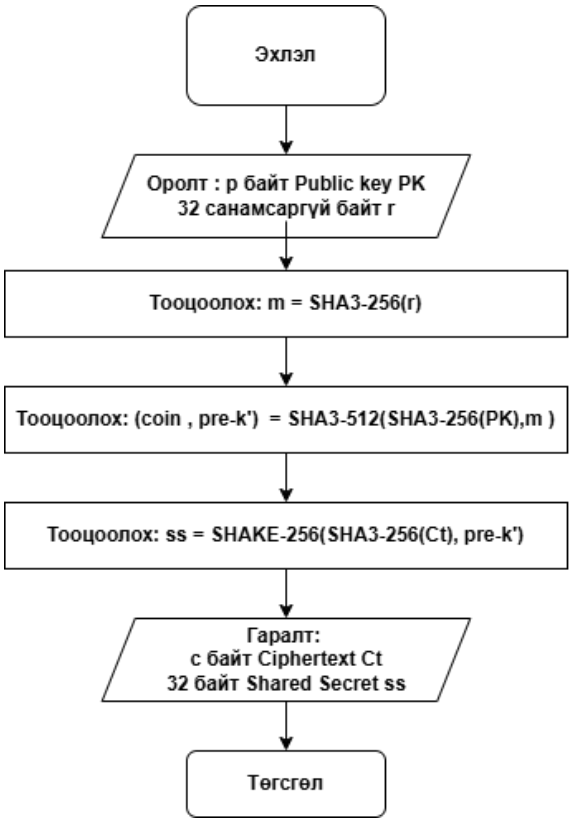
Хүснэгт 7.Техникийн үзүүлэлт

Үзүүлэлт	DHT11	Raspberry Pi 4	HP pavilion
Температур хэмжих хүрээ	0°C - 50°C, ±2°C		
Чийгшил хэмжих хүрээ	20% - 90%, ±5% RH		
Мэдээлэл дамжуулах хурд	~1Hz (Секунд 1 удаа)		
Процессор	-	Cortex-A72 ~ 1.5GHz	Intel/ AMD
RAM	-	8GB LP DDR4	8GB
Санах ойн хэмжээ	-	microSD карт	SSD 256 GB
USB порт	-	2x USB 3.0	USB 2.0
Wi-Fi & Bluetooth	-	Wi-Fi 802.11ac, Bluetooth 5.0	Wi-Fi 802.11ac
Энергийн хэрэглээ	~2.5mA (idling)	5V /3A(USB-C)	120 W

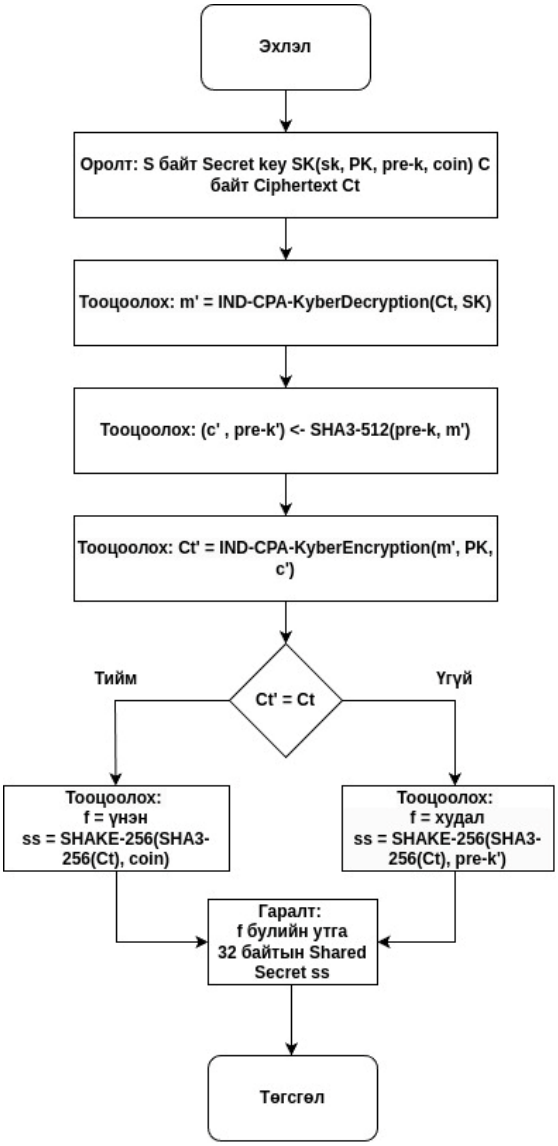
LWC-AES нь хязгаарлагдмал нөөцтэй төхөөрөмжүүдэд зориулан оновчлогдсон хөнгөн жингийн криптографийн хувилбар бөгөөд Advanced Encryption Standard (AES) алгоритмыг бага хэмжээний санах ой, бага эрчим хүч зарцуулдаг төхөөрөмжүүдэд ашиглах боломжтой болгосон хувилбар юм. LWC-AES нь эмбэддэд төхөөрөмжүүд, суурилагдсан системүүд, ухаалаг картууд зэрэг тооцооллын нөөц хязгаарлагдмал орчинд өгөгдлийг найдвартай шифрлэхэд ашиглагддаг.

Шифрлэгдсэн өгөгдлийг сервер рүү дамжуулах явцад пост-квант криптографийн Kyber-1024 алгоритмыг хэрэгжүүлж, дундын түлхүүр солилцох механизм ашиглан өгөгдлийн бүрэн бүтэн байдлыг хангана. Сервер нь хүлээн авсан өгөгдлийг өөрийн түлхүүрийг ашиглан тайлж, дотоод сүлжээнд холбогдсон хэрэглэгчдэд веб интерфейсээр дамжуулан мэдээллийг үзүүлнэ. 7-р зурагт хэрэгжүүлсэн системийн бүдүүвчийг харуулав. 7-р хүснэгтэд хэрэгжүүлэлтэд ашиглагдсан төхөөрөмжүүдийн техникийн үзүүлэлтийг харуулав.

Мэдээллийг нууцалсан LWC-AES алгоритмын түлхүүрийг дамжуулах явцад Kyber-1024 алгоритмаар солилцох бөгөөд 4-р зурагт харагдах алхмаар буюу IND-CPA-KyberDecryption функцээр шифрлэгдсэн өгөгдлийг тайлж, SHA3-512 хэш функцээр баталгаажуулалт хийсний дараа IND-CPA-KyberEncryption алгоритмаар дахин шифрлэн эх өгөгдөлтэй тохирч буй эсэхийг шалгана. Хэрэв өгөгдөл зөв тайлагдсан бол SHAKE-256 хэш функцээр 32 байтын Shared Secret үүсгэж, мэдээллийг аюулгүй байдлаар дамжуулна.

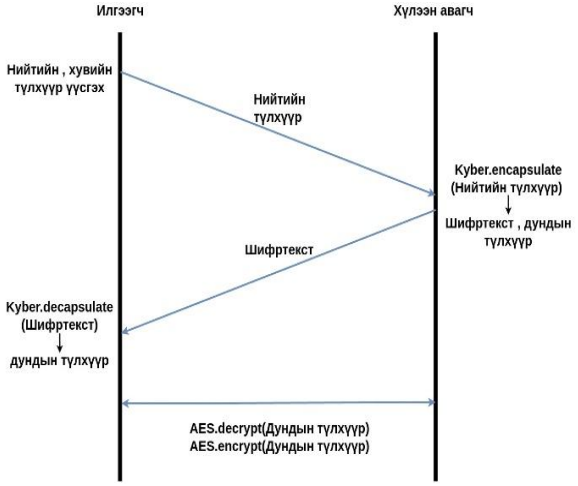


Зураг 9. Kyber-1024 алгоритм энкапсуляци схем.



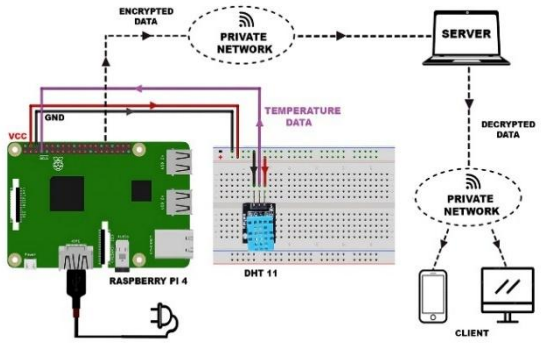
Зураг 10. Kyber-1024 алгоритм декапсуляци схем.

Илгээгч буюу Rasberry Pi 4 температур болон чийгшлийн мэдээллийг LWC-AES алгоритмаар өгөгдлийг шифрлэсэн ба түлхүүрийг зураг 11-т харагдах байдлаар буюу Kyber-1024 алгоритмаар нийтийн болон хувийн түлхүүр үүсгэж, нийтийн түлхүүрийг хүлээн авагч буюу сервер руу илгээх бөгөөд сервер уг түлхүүрийг ашиглан дундын түлхүүр болон нууцласан текст үүсгэж Raspberry Pi 4 рүү буцаах ба нууцалсан текстыг декапсуляци хийж дундын түлхүүрийг аюулгүй дамжуулна.

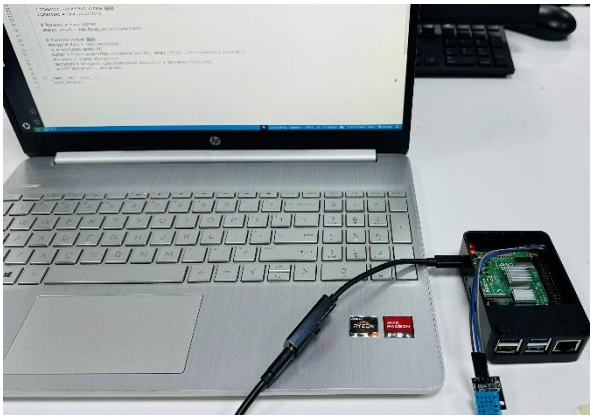


Зураг 11. PQC ашиглан мэдээллийг нууцлалтай дамжуулах хосолсон схем

Зураг 12-т хэрэгжүүлэлтийн системийн холболтын бүдүүвчийг үзүүлсэн бөгөөд энэхүү систем нь дотоод сүлжээнд холбогдсон Raspberry Pi 4 болон серверийн аюулгүй байдлыг ханган, температур, чийгшлийн өгөгдлийг LWC-AES алгоритмаар шифрлэн дамжуулж, сервер дээр боловсруулсны дараа хэрэглэгчдэд мэдээллийг найдвартай хүргэх болно.



Зураг 12. Холболтын бүдүүвч зураг.



Зураг 13. Туршилтын орчин

```python
def send_encrypted_data(payload):
    """Send encrypted data to the server."""
    try:
        response = requests.post(
            f"{SERVER_URL}/send_data",
            json=payload,
            headers={'Content-Type': 'application/json'}
        )
        response.raise_for_status()
        print(f"Server response: {response.status_code} - {response.text}")
        return True
    except Exception as e:
        print(f"Error sending data: {e}")
        return False
```

Зураг 14. Илгээгч талын код

Сервер нь хүлээн авсан нууцлагдсан мэдээллийг урьдчилан солилцсон криптографийн түлхүүрийг ашиглан зураг 14-т дүрслэгдсэн кодын дагуу өгөгдлийг тайлах бөгөөд туршилтын орчны бүтэц, тохиргоог зураг 13-т үзүүлсэн болно. Аюулгүй дамжуулагдсан мэдээллийг сервер нь нэг сүлжээнд байрлах хэрэглэгчдэд зураг 15-т өгөгдсөн програмчлалын хэрэгжилтийн дагуу веб интерфейсээр хүргэх ба үүний хэрэгжилтэд Пайтон програмчлалын хэлний Flask фрэймворк болон HTML бичиглэлийг ашигласан болно.

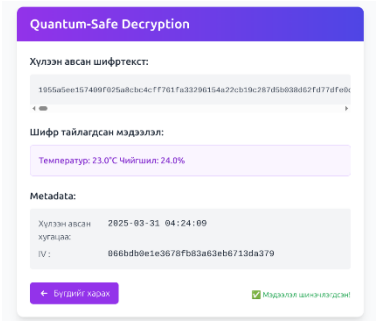
```python
from flask import Flask, render_template, request, jsonify
from oqs import KeyEncapsulation  # type: ignore
from cryptography.hazmat.primitives.ciphers import Cipher, algorithms, modes
from cryptography.hazmat.backends import default_backend
from datetime import datetime

app = Flask(__name__)

# Kyber1024 алгоритмыг ашиглан түлхүүрийн хос үүсгэнэ
kem = KeyEncapsulation("Kyber1024")
public_key = kem.generate_keypair()  # Нийтийн түлхүүр
private_key = kem.export_secret_key()  # Хувийн түлхүүр

decrypted_messages = []  # Шифрийг тайлсан мессежүүдийг хадгалах жагсаалт

@app.route('/')
def index():
    """Үндсэн хуудасыг харуулах."""
    return render_template('results.html')

@app.route('/get_public_key', methods=['GET'])
def get_key():
    """Нийтийн түлхүүрийг буцаах."""
    return public_key.hex()

@app.route('/send_data', methods=['POST'])
def receive_data():
    """Шифрлэгдсэн өгөгдөл хүлээн авч, тайлж хадгалах."""
    data = request.json
    ciphertext = bytes.fromhex(data['ciphertext'])  # Шифрлэгдсэн түлхүүр
    iv = bytes.fromhex(data['iv'])  # Эхний вектор
    encrypted = bytes.fromhex(data['encrypted'])  # Шифрлэгдсэн мессеж

    # Нууц түлхүүрийг тайлж авах
    shared_secret = kem.decap_secret(ciphertext)

    # AES шифрлэлтийг тайлах
    cipher = Cipher(algorithms.AES(shared_secret), modes.CFB(iv), backend=default_backend())
    decryptor = cipher.decryptor()
    decrypted = decryptor.update(encrypted) + decryptor.finalize()

    # Тайлсан мессежийг хадгалах
    decrypted_str = decrypted.decode()
```

Зураг 15. Веб код

Хэрэглэгчдэд харагдах веб сайтын үндсэн ямар ч мэдээлэл хүлээж аваагүй үед зураг 16-т харагдах хэлбэрээр байна.

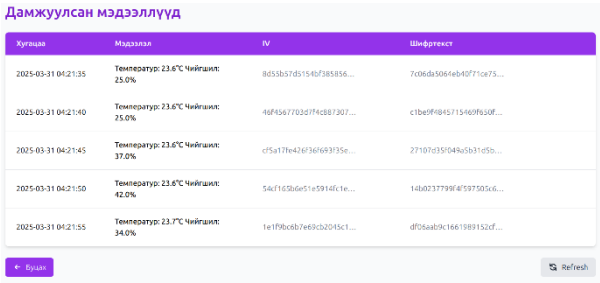
Зураг 16. Холболт тогтоох үе

Сервер орчны температурыг Raspbery Pi 4 - өөс хүлээн аваад нууцлалыг тайлан зураг 17-т харагдах үр дүнг хэрэглэгчид харуулна.



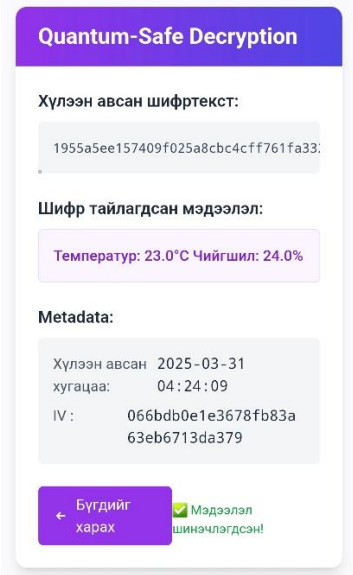
Зураг 17. Мэдээлэл ирсэн байдал

Тогтмол хугацаанд хэрэглэгч рүү температур, чийгшлийн мэдээллийг дамжуулах бөгөөд хэрэглэгч хүссэн үедээ өмнө нь ирсэн мэдээллүүдийг шалгах боломжтойг зураг 18-т харуулав.



Зураг 18. Цугларсан мэдээлэл харагдах байдал

Хэрэглэгч гар утаснаасаа вебээр температурын мэдээллийг хүлээж авах боломжтойг Зураг 19-т харуулав.



Зураг 19. Гар утаснаас холболт хийх

IV. ДҮГНЭЛТ

Хязгаарлагдмал нөөцтэй эмбэддэд системүүд дээр пост-квант криптографийн алгоритмуудыг хэрэгжүүлэх гүйцэтгэл, эрчим хүчний хэрэглээ, тооцооллын зардлын талаарх хийгдсэн судалгаанд дүн шинжилгээг хийлээ. Туршилтын үр дүнгээс Kyber-1024 нь өндөр түвшний аюулгүй байдлыг хадгалж, гүйцэтгэл болон эрчим хүчний зарцуулалтын хувьд эмбэддэд системд тохиромжтой нь харагдсан тул өөрсдийн системд хэрэглэв.

Raspberry Pi 4 төхөөрөмж дээр Kyber-1024 , LWC-AES алгоритмуудыг хослуулан ашиглан CNSA 2.0 стандартын алгоритмыг бодит орчинд туршсан.

Цаашид пост-квант криптографт суурилуулсан системүүдэд илүү үр ашигтай хэрэгжүүлэх арга замуудыг судлах нь чухал бөгөөд энэ чиглэлээр нэмэлт оновчлол, алгоритмын сайжруулалт хийх шаардлагатай байна.

НОМ ЗҮЙ

[1] Commercial National Security Algorithm (CNSA) Suite, MFS U/00/814670-15.

[2] Bush G. National security strategy of the United States. – White House, 1991.

[3] The Commercial National Security Algorithm Suite 2.0 and Quantum Computing FAQ, U/OO/194427-22 | PP-24-4014 | December 2024 Ver. 2.1

[4] Khan, S., Krishnamoorthy, P., Goswami, M., Rakhimjonovna, F. M., Mohammed, S. A., & Menaga, D. (2024). Quantum Computing And Its Implications For Cybersecurity: A Comprehensive Review Of Emerging Threats And Defenses. *Nanotechnology Perceptions*, *20*, S13.

[5] QUANTUM TECHNOLOGIES 2021 Market & Technology Report - June 2021

[6] Heath, S., 2002. Embedded systemc design. Elsevier.

[7] Lutkevich, B. (2020, December) 'Embedded Систем', IoT Agenda.

[8] Hayward, M. (2008). Quantum computing and shor's algorithm. *Sydney: Macquarie University Mathematics Department*, *1*.

[9] Halak, B., Gibson, T., Henley, M., Botea, C. B., Heath, B., & Khan, S. (2024). Evaluation of performance, energy, and computation costs of quantum-attack resilient encryption algorithms for embedded devices. *IEEE Access*, *12*, 8791-8805.

[10] National Institute of Standards and Technology (NIST). (2024) 'NIST Releases First 3 Finalized Post-Quantum Encryption Standart

Монгол бичгийн хэвлэмэл текстийг машин сургалт болон гүн сургалтаар таних нь

Даваацэрэнгийн Цэцэнбилэг
Мэдээллийн технологи,
электроникийн сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
22b1num1282@stud.num.edu.nm

Энхтөрийн Чойжилням
Мэдээллийн технологи,
электроникийн сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
21b1num0871@stud.num.edu.mn

Батсуурийн Сувдаа*
(холбоо барих зохиогч)
Мэдээллийн технологи,
электроникийн сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
suvdaa@num.edu.mn

Хураангуй— Хэл нь улс үндэстнийг бусдаас ялгаж байгаа нэгэн гол шинж болдог. Монгол хэл бичиг нь монгол үндэстний оюуны соёлын үнэт өв, үндэсний аюулгүй байдал, тусгаар тогтнолын нэгэн баталгаа мөн гэж үндсэн хуульд тусгасан байдаг. Уламжлалт монгол бичиг нь дээрээс доош холбогдсон бичиглэлтэй бөгөөд стандарт болон гар бичмэл гэсэн 2 төрөлтэй байдаг ба үсэг нь үгийн эхэнд, дунд, адагт өөр өөр зурлагаас үүсдэг өвөрмөц шинж чанартай бичиг. Энэхүү судалгааны ажлаар хэвлэмэл монгол бичгийн зураг дээр урьдчилсан боловсруулалтын аргууд ашиглан хэсэгчлэн хуваалт хийсэн. Үүссэн дүрсүүдийг 55 багц гэж үзэн гүн сургалтын арга болон машин сургалтын аргуудыг харьцуулан танилт хийсэн.

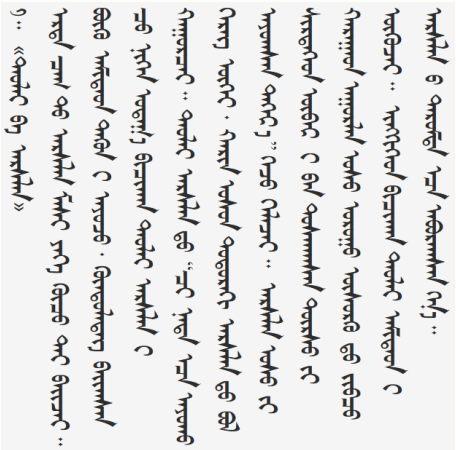
Түлхүүр үг— Уламжлалт монгол бичиг, дүрс боловсруулалт, тэмдэгт таних, машин сургалт, гүн сургалт, зураг ангилал (Image Classification)

I. УДИРТГАЛ

Монголчуудын дунд “уйгуржин монгол бичиг”, “монгол бичиг”, “худам монгол үсэг” анх хэдийд, хаанаас авсан тухай сүүлийн үед судлаачдын олонх нь зөвшөөрч буйгаар, эрдэмтэн Ц.Шагдарсүрэнгийн судалсанчлан VI-VII зууны үед Монгол, Уйгур хоёр одоогийн Тажикийн дээд өвөг Согд нараас нэгэн зэрэг бичиг үсэг авсан байж болох хэмээсэн байдаг. Монголчууд согд гаралтай бичгийг авч, өөрийн хэлэнд тохируулан боловсруулж хэрэглэсэн нь эртнээс эдүгээ хүртэл 10 зуунаас наашгүй гэж үзэх үндэстэй юм. Тийм ч учраас бид монгол бичгээ мянган жилийн түүхтэй хэмээн ярьж буй.[1] Монгол бичгээ сурч эзэмших, судлан сурталчлах орчин үеийн техник технологи, шинжлэх ухааны ололт амжилтуудыг ашиглан түгээн дэлгэрүүлэх, хамгаалах, өвлүүлэх, хөгжүүлэх нь Монгол хүн бүрийн бичиг соёлын өв сангийн өмнө хүлээж байгаа эрхэм үүрэг.[2] Монгол Улсын Засгийн газраас төрийн албан хэргийг хос бичгээр хөтлөн явуулах цаашдаа үндэсний бичгийг түгээн дэлгэрүүлж хэрэглээ, судалгаа, орчныг нэмэгдүүлэх зорилттой “Монгол бичгийн үндэсний хөтөлбөр III” хөтөлбөрийг энэ оны эхнээс хэрэгжүүлэн эхэлж байна. Тиймээс урт хугацааны туршид үүсэж бий болсон асар их хэмжээний монгол бичгийн өгөгдлийг боловсруулах хэрэгцээ шаардлага тулгарч байна. Энэхүү боловсруулалтад кирилл бичгээс үндэсний бичигт хөрвүүлэх, монгол бичгээс кирилл бичигт хөрвүүлэх цаашлаад хэвлэмэл бичиг, гар бичмэл олон төрлийн монгол бичгийн хэвүүдийг таних гэх мэт боловсруулалтууд орно.

Монгол бичиг нь өөрийн гэсэн онцлог шинж чанартай. Нэгдүгээрт нийлмэл бүтэцтэй буюу үсгүүд нийлж үг бүтээхэд үсэг бүр холбогдож үндсэн нурууг үүсгэдэг. Удаах онцлог нь дээрээс доош чиглэлтэй бичигддэг ба бичиглэлийн төрөл дармал, бичмэл гэсэн 2 төрөл байна. Цагаан толгой нь 29 үсэгтэй үүнд 7 эгшиг, 14 гийгүүлэгч, 8 галиг үсэг орно . Үсэг бүр нь үгийн эхэнд дунд адагт орж байгаагаас хамааран өөр өөр зурлагаар бичигддэг. Мөн энэхүү онцлог нь олон ястан өөр өөрийнхөө нутгийн аялгууг бие даан хадгалж хөгжүүлэх боломжийг хангаж өгсөн байдаг. [5]

Судалгааны ажлын зорилго нь хэвлэмэл байгаа өргөн фонттой монгол бичгийн фонт болон зурлагыг өндөр хувьтай таних. Үүнийхээ хүрээнд монгол бичгийн хэвлэмэл зураг дээр компьютер харааны дүрс боловсруулалтын энгийн аргууд ашиглан багана, үг, зурлага болгож хэсэгчлэн хуваах урьдчилан боловсруулалт хийсний дараа машин сургалт болон гүн сургалтын аргууд ашиглан туршилтууд хийж 90%-аас дээш үзүүлэлттэй танилтын үр дүнгүүд гаргаж чадсан. Доорх зурагд (Зур.1) өргөн фонттой монгол бичгийн жишээг харууллаа.



Зур. 1. Судалгаанд ашигласан үндсэн гурван фонтын нэг болох Mongolian script фонттой бичиг.

Энэхүү судалгааны ажил нь дараах бүтэцтэй. Судалгааны ажлын II хэсэгт сэдэвтэй ижил төстэй, холбоо хамааралтай өмнө хийгдсэн судалгааны ажлуудыг тоймлон, III хэсэгт судалгааны ажлын үндсэн аргачлал арга алгоритм, судалгаанд ашигласан өгөгдлийн багцын тухай, IV хэсэгт туршилтын үр дүнг танилцуулж эцэст нь энэхүү ажлыг дүгнэнэ.

II. ХОЛБООТОЙ АЖИЛУУД

Монгол бичгийн хэвлэмэл текстийн таниулахтай холбоотой судалгааны ажлууд олон улсын болон дотоодын түвшинд хийгдсэн байдаг. Эдгээрээс бидний ажилтай төстэй, мөн харьцуулан үзэх боломжтой зарим ажлыг дурдвал:

Traditional Mongolian Script Recognition Using Primitives and Template Matching Methods[7] судалгаанд Монгол бичгийн тэмдэгтүүдийг хэсэгчилсэн дүрс (primitive) болон загвар таарууулах (template matching) аргуудыг ашигласан. Энэхүү арга нь дүрсийн үндсэн хэлбэрийг урьдчилан тодорхойлж, өгөгдлийн сангийн доторх загвартай харьцуулах замаар танилт хийдэг. Энэ нь хэвүүдийг урьдчилан боловсруулах, байрлалд суурилсан бүтэц болон histogram мэдээлэлд тулгуурлан segmentation хийх давуу талтай боловч, үсгийн хэмжээ, байрлал, фонтоос хамаарч үр дүн нь буурах магадлалтай. Харин бид энэхүү аргаас ялгаатай нь гүн сургалтын аргуудыг ашигласнаар урьдчилсан өгөгдөл бэлдэх шаардлагагүй ба онцлог ялган таних Convolutional Neural Network ашигласан. CNN нь өөрөө онцлог шинжүүдийг гаргаж тооцдог илүү уян хатан, олон талт бүтэцтэй өгөгдөл дээр үр дүн сайтай ажилладаг. Мөн энэхүү судалгааны нуруу бүтэц ба histogram ашиглан segmentation хийх үйл явц нь бидний судалгааны pre-processing үеийн зурлага салгах үйл явцын гол тулгуур бүтэц болж өгсөн.

Machine-Printed Traditional Mongolian Characters Recognition Using BP Neural Networks[8] судалгаанд арын пропагацитай нейрон сүлжээ (Backpropagation Neural Network) ашигласан нь харьцангуй энгийн бүтэцтэй, гүн багатай, хэрэгжүүлэхэд хялбар загваруудын нэг юм. Гэвч BP нейрон сүлжээ нь дүрсийн нарийн онцлог ялган таних байдал сул бөгөөд их хэмжээний өгөгдөл дээр overfitting үүсэх эрсдэлтэй. Бидний судалгаанд хэрэглэсэн олон давхаргат Convolutional Neural Network (CNN) нь илүү олон давхарга бүхий бүтэцтэй, өгөгдлөөс илүү баялаг онцлог шинжийг тооцож чаддаг тул ялгаатай нөхцөлүүдэд илүү найдвартай үр дүн үзүүлдэг.

Convolutional Neural Network for Machine-Printed Traditional Mongolian Font Recognition[9] судалгаанд бидэнтэй адил CNN ашигласан боловч тус судалгаа нь үгний фонтын ангилалд төвлөрсөн. Тус ажлын давуу тал нь segmentation хийх шаардлагагүйгээр шууд үгний зураг дээр ангилал хийх боломжтой бөгөөд font-based OCR системд тохиромжтой. Харин бидний судалгаанд тухайн үг, өгүүлбэрийн доторх тэмдэгтүүдийг нэг бүрчлэн ялган таньж, танилтын нарийвчлал, уян хатан байдалд илүү ач холбогдол өгсөн. Иймд бид segment хийж таних MLP болон CNN хосолсон аргуудыг ашигласан бөгөөд энэ нь үг доторх олон янзын дүрс, байрлал, хэмжээтэй тэмдэгтүүдийг онцгой сайн ялгаж, гүйцэтгэлийг нэмэгдүүлэх давуу талтай.

A Survey on Arabic Character Recognition[10] судалгаанд араб бичгийн танилтын асуудлыг тоймлон судалсан ба түүний бүтэц, бичих зарчим нь Монгол бичигтэй ижил төстэй (үсгүүдийн залгаа байдал, байрлалаас хамаарч хэлбэр нь өөрчлөгддөг) тул энэхүү судалгаанд дурдсан segmentation аргууд, baseline буюу "нуруу"-д суурилсан бүтэц танилтын ойлголтыг бид өөрсдийн судалгаанд урьдчилсан боловсруулалтын үед ашигласан. Гэсэн ч араб бичгийн танилтын ихэнх судалгаа нь LSTM, attention механизмуудыг ашигласан sequence-to-sequence загварт төвлөрдөг бөгөөд бидний судалгаа тэмдэгт төвтэй, фонт болон дүрсийн нарийвчилсан ангилал хийхэд илүү чиглэсэн онцлогтой.

III. АРГАЧЛАЛ

Судалгааны энэхүү хэсэгт зургаас монгол бичгийн зурлага бүрээр салгасан аргачлал, архитектур бүтэц болон танилтад ашигласан машин сургалт, гүн сургалтын аргуудыг танилцуулах болно.

3.1. Өгөгдлийн багц(dataset) бэлтгэх архитектур

A. Оролтын зургийн тухай

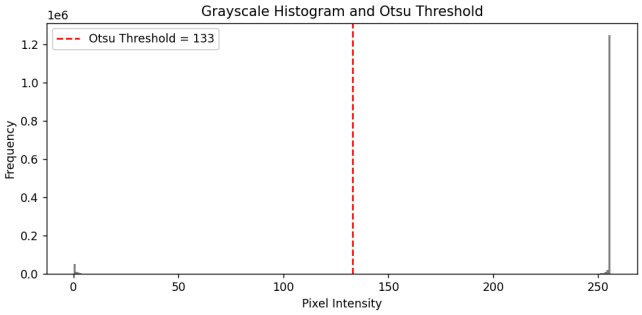
Оролтын зураг болгон ашиглах өгөгдлөө цахим орчинд нээлттэй веб хуудаснаас авсан бөгөөд онцлох гол шинж нь өргөн хэвтэй фонтын бичиг сонгосон. Нарийвчлан үзвэл нурууны өргөн 5 эсвэл 6 пиксел байх 3 өөр төрлийн дармал бичвэрүүд сонгон авсан. Оролтын бүх зургийг цахим орчноос бэлтгэсэн тул ямар нэгэн хазайлт өнцгийн зөрүү үүсэхгүй давуу талтай.

(1) (2) (3)

Зур. 2. Хүмүүн бичиг, Mongolian Script, Mongol Script фонтууд

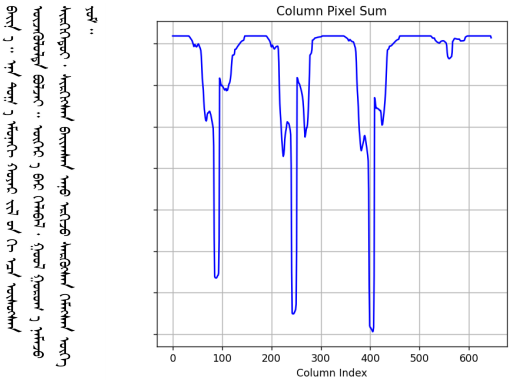
B. Баганаар хэсэгчлэн салгах

Баганаар хэсэгчлэн салгах үйлдэл хийхдээ зургаа саарал буюу нэг сувагтай зураг (grayscale image) болгож уншуулсны дараа саарал түвшний гистограммаас хамгийн тохиромжтой босгыг автоматаар тодорхойлж, дүрсийг хоёртын хэлбэрт үр дүнтэй хувиргах боломжийг олгодог Otsu-ийн арга [3] ашиглан хоёртын зураг болгосон (binary image).



Зур. 3. Otsu-ийн аргаар босго тодорхойлж байгаа нь

Үүний дараагаар баганын дагуу пикселийн утгуудын нийлбэрийн график үүсгэсэн.



Зур. 4. Эхний зургийн баганын дагуу нийлбэрийн график

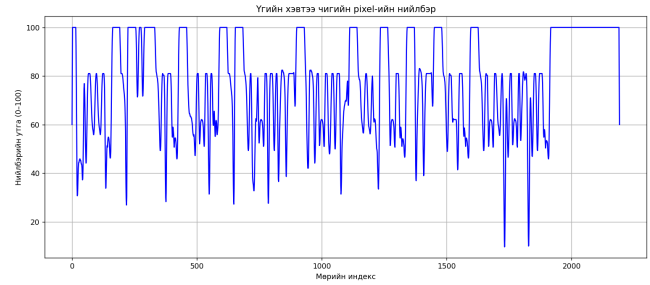
Ерөнхий байдлаар мөрүүдийн зааг мэдэгдэж байгаа тул таслалт хийх цэгүүдийн жагсаалт үүсгэсэн. Энэхүү жагсаалтыг тодорхой нэг а утга олоод 0-ээс а хүртэлх утгатай байснаа түүнээс дээш утгатай болж байгаа хэсгийн баганын утга эхлэх цэг тэрхүү мужийн утгаас өндөр утгатай байсны дараа уг муж руу орж байгаа хэсгийн баганын утгыг эцсийн цэс гэж үзнэ. Уг мужийг тодорхойлох а утгыг дараах байдлаар тооцоолсон. Эцэст нь эхлэх цэгээс төгсгөлийн цэгийн хооронд байгаа утгуудыг ашиглан шинэ зураг үүсгэснээр багана бүр таслагдан тусдаа зураг болно.

а = хамгийн_өндөр_утга (баганын_нийлбэр) * 0.001

```
0 <= baganii_utga[i] < a  and baganii_utga[i+1] > a) or
  (baganii_utga[i] > a and  a > baganii_utga[i+1] >= 0
```

C. *Үгээр хэсэгчлэн салгах*

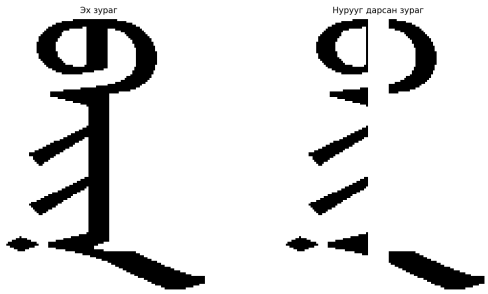
Энэ нь багана салгах үе хэрэгжүүлсэн аргийг давтан хэрэглэж байгаа бөгөөд энэ удаад мөрийн дагуух утгууд дээр тооцоолол хийж байгаа.



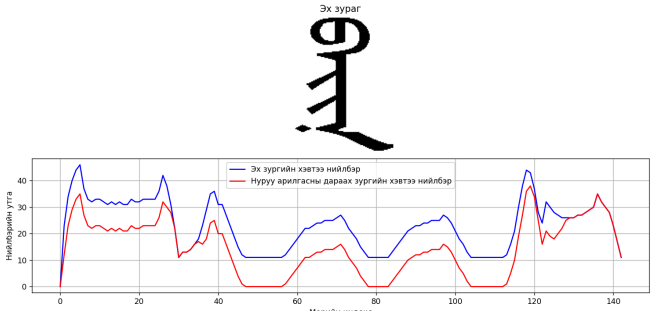
Зур. 5. Otsu-ийн аргаар босго тодорхойлж байгаа нь

D. *Нуруу арчих арга ашиглан зурлага салгах*

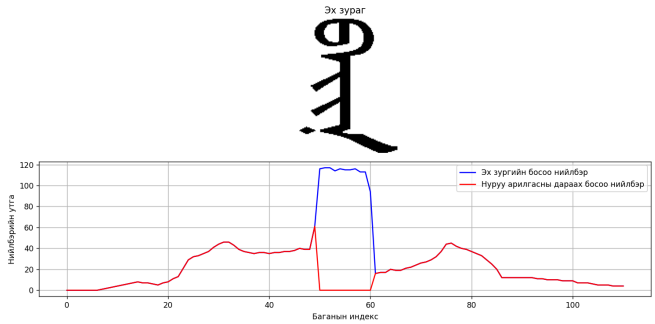
Баганын дагуу пикселийн нийлбэрүүдийг олоод тэндээс а=хамгийн_өндөр_утга (баганын_нийлбэр) * 0.8 байх а утгыг босго утга болгон авсан. Хэрвээ босго утгаас өндөр байвал түүнийг нуруу агуулж буй багана гэж үзээд баганын бүх пикселд 255 гэсэн утга өгч цагаан болгоно. Үүний дараа мөрийн дагуу пикселийн нийлбэрүүдийг олж өмнөх аргуудтай адилаар тасралтын цэгийн утгуудыг аваад таслалтууд хийсэн.



Зур. 6. Нурууг дарахаас өмнөх болон нурууг арилгасан зураг



Зур. 7. Хэвтээ чиглэлд пикселүүдийн нийлбэрийг өмнөх болон дараах үетэй харьцуулан харуулсан график.



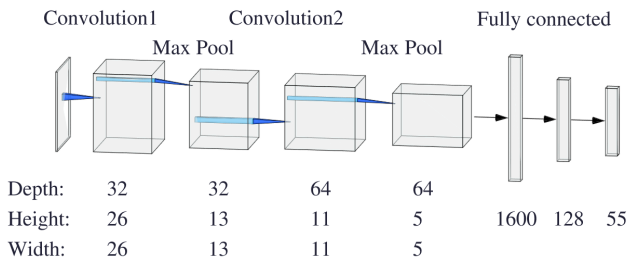
Зур. 8. Босоо чиглэлд пикселүүдийн нийлбэрийг өмнөх болон дараах үетэй харьцуулан харуулсан график.

Нурууг арилгах үйлдэл хийсний дараа зураг 6 дээр харагдаж байгаачлан 0-тэй тэнцүү болж байгаа хэсгүүдэд эх зураг дээр таслалт хийнэ. Үр дүнд нь анх орж ирсэн зургийг зурлага бүрээр нь амжилттай салгана.

E. *Ангилал үүсгэх*

Нийт 3700 гаруй өгөгдөл бий болгосон бөгөөд үүнийгээ ангилал бүрд салгаж хийхдээ хэв шинжээр нь адилтган жишиж давхардсан хэсгийг тооцдог зураг боловсруулалтын аргуудын нэг болох загвар тааруулах (template matching) арга[4] ашиглан ойролцоох нийт өгөгдлийн 20%-ийг ангилж чадсан ба үлдсэн 80%-ийн өгөгдлийг гар аргаар хавтас бүр лүү файлуудыг зөөх замаар шаардлагатай байршилд нь оруулах ялгалтыг хийсэн.

3.2 *Гүн сургалтын архитектур*



Зур. 9. Гүн сургалтын архитектур

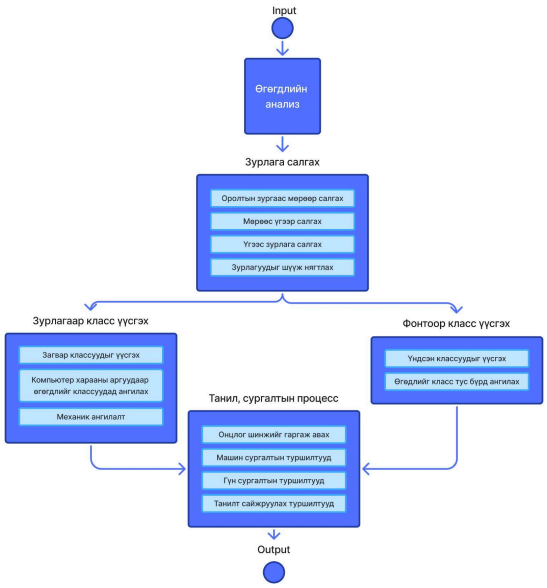
Convolutional Neural Networks (CNN) нь дүрс таних, объект илрүүлэх, загварын танилт хийх зэрэг чиглэлд өргөн хэрэглэгддэг гүн сургалтын алгоритмуудын нэг юм. CNN-ийн архитектур нь конволюцийн давхарга (convolutional layers), пуллинг давхарга (pooling layers), болон бүрэн холболттой давхаргуудаас (fully connected layers) бүрддэг бөгөөд энэ нь зурагт мэдээллийг автоматаар боловсруулан, онцлог шинж чанарыг таньж авахад тусалдаг. CNN-ийн хамгийн гол давуу тал нь өөрчлөлт, эргүүлэлттэй байдлыг (translation invariance) нөхөх боломжтой бөгөөд энэ нь түүний үндсэн давуу тал юм. Энэ бүтэц нь бүртгэх, таних процессыг улам хялбар болгож, эх өгөгдлийн мэдээллээс илүү их онцлог шинжийг бий болгож чаддаг. [6]

Конволюцийн Давхарга (Convolutional Layer): Зургийн онцлог шинж чанаруудыг таньж авахад ашигладаг. Энэ давхарга нь өнгө, текстур, хээг ойлгож, дараагийн процессуудад илүү нарийвчилсан шинж чанаруудыг ашигладаг. Энэхүү хэсэгт жижиг хэмжээний өөр өөр матрицуудаар үржүүлэхэд онцлог шинжийг агуулсан давхаргууд бий болгодог. Бид судалгааны ажилдаа анх бэлдсэн өгөгдлөө 28x28 хэмжээтэй болгоод эхний конволюцийн давхаргаар оруулахдаа 3x3 хэмжээтэй 32 кернэл гүйлгэж 26x26 хэмжээтэй 32 давхарга бүхий зураг үүсгэсэн.

Пуллинг Давхарга (Pooling Layer): Зургийн хэмжээг багасгаж, хэрэгтэй мэдээллүүдийг хадгалах үүрэгтэй. Энэ нь тооцоолох хурдыг нэмэгдүүлэх болон илүү сайн онцлог шинж чанаруудыг хадгалахад тусалдаг. Аливаа утгын хэмжээг багасгах уг аргаас Max pooling (2 , 2) буюу 2x2 хэмжээтэй матрицаас хамгийн өндөр байгаа утгыг хадгалаад 2 нүдээр үсрээд дараагийн матрицаас хамгийн их утгыг авах замаар явж байна.

Бүрэн Холболттой Давхарга (Fully Connected Layer): Энэ давхарга нь хамгийн сүүлийн үеийн шинж чанаруудыг боловсруулж, эцсийн ангиллыг гаргаж ирдэг. Хоёр удаа конволюцийн давхарга, пуллинг давхаргаар дамжсаны дараа бидний анх орж ирсэн эх зураг 64 давхаргатай 5x5 хэмжээтэй зураг болсон байгаа. Үүнийг сүүлчийн давхарга луу шилжүүлэхдээ 64x5x5 урттай вектор болгосон. Дараагийн шатанд 128 болгож эцсийн үр дүнд классын тоотой тэнцүү матриц үүсгэсэн.

Нэмэлтээр машин сургалтын хэд хэдэн арга алгоритмаар туршиж үзсэн бөгөөд судалгааны ажлын IV хэсэгт дэлгэрэнгүй тайлбарлах болно. Дараах схемээр судалгааны ажлын ерөнхий архитектурыг харууллаа.



Зур. 10. Судалгааны ажлын ерөнхий архитектур

Энэхүү алга алгоритмуудаар хэвлэмэл монгол бичгийг зурлагаар салгах, цаашлаад танихад тулгарсан хэд хэдэн асуудал бий. Баганаар салгалт хийх үед үгийн эхэнд ордог га үеийн давхар дусал үндсэн зурлагаас тусдаа байрлалтай байсан тул тухайн баганад шилбэ ба шилбэний төрлийн зурлага ороогүй бол тэрхүү дуслын хэсгийг тусдаа багана гэж үзэн таслах тохиолдол гарсан. Хэдийгээ нурууг арилгаж хооронд нь цагаан зай үүсгэсэн боловч үгийн ар хэсэгт зурагддаг гэзэг, эвэр гэх мэт зурлагууд үргэлжлэн удаах зурлагатай нийлж байгаа мэт болж таслалт хийхэд алдаа үүсгэж байсан.

IV. ТУРШИЛТЫН ҮР ДҮН

Судалгааны энэ хэсэгт дээр дурдсан аргачлалын дагуу зургаас монгол бичгийн зурлага бүрээр салгасан өгөгдөл дээр машин сургалт, гүн сургалтын аргуудыг ашиглан танилтыг сайжруулах туршилтуудын үр дүн болон үүсгэсэн өгөгдлийн багцын талаар дэлгэрэнгүй мэдээлэл харуулна.

4.1. Өгөгдлийн багцын дэлгэрэнгүй

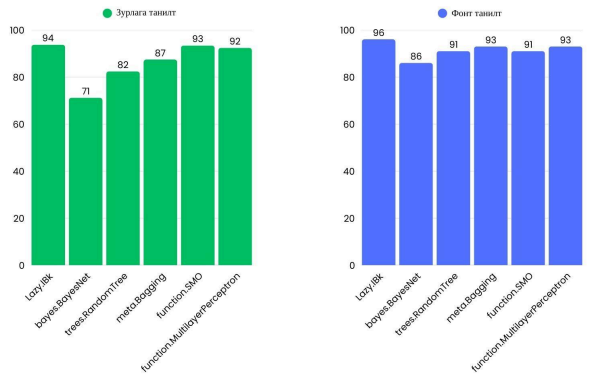
Монгол бичгийн цагаан толгойд нэг үсэг олон зурлагаас бүтдэг цаашлаад нэг зурлага нэлээн хэдэн үсгийг илэрхийлдэг тул ангилал хийхдээ ихээхэн анхаарах хэрэгтэй. Гэхдээ бид судалгааны ажлынхаа хүрээнд зөвхөн өндөр хувьтай танилт хийхэд анхаарах тул зурлагын илэрхийлэх илэрхийлэмжийг төдийлөн чухалчилж үзэхгүй. Харин компьютерт хэрхэн уншигдах вэ түүнийг хэрхэн боловсруулах вэ гэдэг дээр голчлон анхаарсан. Хэдийгээр монгол бичиг 7 эгшигтэй боловч үгийн эхэнд 5 өөр хэлбэрээр үгийн дунд 3 өөр хэлбэрээр үгийн адагт мөн 3 өөр хэлбэрээр бичигддэг. Харин гийгүүлэгчид эгшиг авсан үед болон эгшигжээгүй үед тус тус өөр бичигддэг.

Дээрх онцлог шинж чанаруудыг харгалзан үзээд монгол бичгийн нийт цагаан толгойн зурлагыг 55 багц (class) болгон дараах байдлаар ангилсан. Бусад судалгааны ажлуудаас өөр байгаа онцлог шинж нь ба, бо, фа, фо гэх мэт үгийн аль ч хэсэгт төстэй зурлагатай боловч дээр эсвэл доор нь нурууны хэсэг байгааг үл анхааран үеүдийг эхэнд дунд адагт ялгаагүй гэж үзсэн нь давуу тал болно.

№	Нэршил	Зурлага	Тэмдэг лэгээ	№	Нэршил	Зурлага	Тэмдэг лэгээ
1	Титэм		ti	29	Гэдэст нум		ba
2	Шүд		su	30	Гэдэст нум шилбэтэй		bi
3	Шилбэ		si	31	Давхар гэдэст нум		bo
4	Гэдэс		ge	32	Гэдэст нум орхицтой		bs
5	Нум		ia	33	Ятгар зартиг		ha
6	Гэдэстэй нум		gn	34	Ятгар зартиг давхар дусалтай		ga
7	Сүүл		sl	35	Богино шилбэт нум		ha
8	Цацлага		ts	36	Богино шилбэт нум шилбэтэй		hi
9	Эвэртэй титэм		le	37	Богино шилбэт нум гэдэстэй		gu
10	Эвэртэй шүд		ld	38	Эх энхлэг		gj
11	Эвэр сүүл		ld	39	Дусалтай агшилга		ga
12	Гэзэгтэй титэм		me	40	агшилга		ha
13	Гэзэг		md	41	Сэрээ		tc
14	Гэзэгтэй охир сүүл		ma	42			zj
15	Дусалтай титэм		ne	43	Итгэр шилбэ		ye
16	Дусал шүд		nd	44	Итгэр шилбэ		yd
17	Дусал сүүл		na	45	Матгар шилбэ		vd
18	Завж		se	46	Чагттай гэдэст нум		fa
19	Завж		sd	47	Чагттай гэдэст нум		fd
20	Завж тегсгөл		sa	48	Чагттай давхар гэдэст нум		fo
21	Давхар дусалт завж		we	49	Далбан чихт нум		pa
22	Давхар дусалт завж		wd	50	Далбан чихт нум шилбэтэй		pi
23	Өргөстэй шилбэ		rd	51	Далбан чихт нум гэдэстэй		po
24	Өргөстэй нуман шилбэ		ra	52	Далбан чихт нум орхицтой		ps
25	Гогцоот титэм		de	53	Сэртэн чихт нум		ka
26	Гогцоот шилбэ		dd	54	Сэртэн чихт нум		kd
27	таслал		du	55	Сэртэн чихт нум орхицтой		ka
28	Этгэр сүүл		es				

Зур. 11. Өгөгдлийн ангиллын загвар нийт 55 class үүсгэсэн.

4.2. Машин сургалт болон гүн сургалтын туршилтууд



Зур. 12. Машин сургалтын зурлага болон фонт танилтын үр дүн

Зураг 12-д машин сургалтын 6 өөр аргааг зурлага болон фонт таних танилтын хувийг хооронд нь харьцуулан харуулсан байна. Эндээс харахад ихэнх арга нь фонт танилт дээр илүү сайн ажиллаж байгаа ч хоёрдугаар арга зурлага танилт дээр бага зэрэг муу ажиллаж байгааг харж болно. Дараах хүснэгт 1, 2-д өмнөх зураг дээр харуулсан танилтын үр дүнгүүдийн зөв таньсан тоо, зөв таньсан хувь, буруу таньсан тоо, буруу таньсан хувийг тус тус дэлгэрэнгүй харуулсан.

№	Recognition algorithm	Correctly Classified Instances		Incorrectly Classified Instances	
		Number	Percent	Number	Percent
1	Lazy.IBk	3462	93.72%	232	6.28%
2	bayes.BayesNet	2630	71.20%	1064	28.20%
3	trees.RandomTree	3043	82.38%	651	17.62%
4	meta.Bagging	3230	87.44%	464	12.56%
5	function.SMO	3449	93.36 %	245	6.63 %
6	function.MultilayerPerceptron	3413	92.39%	281	7.61%

Хүснэгт. 1. Зурлага танилтын үр дүн

№	Recognition algorithm	Correctly Classified Instances		Incorrectly Classified Instances	
		Number	Percent	Number	Percent
1	Lazy.IBk	3378	96.05%	139	3.95%
2	bayes.BayesNet	3022	85.93%	495	14.07%
3	trees.RandomTree	3193	90.79%	324	9.21%
4	meta.Bagging	3276	93.15%	241	6.85%
5	functions.SMO	3188	90.65%	329	9.35%
6	functions.Multilayer Perceptron	3257	92.61%	260	7.39%

Хүснэгт. 2. Фонт танилтын үр дүн

Дээр туршсан машин сургалтын 6 аргаа товч байдлаар тайлбарлавал:

Lazy.IBk (Instance-based k-NN)[11]:

Энэ нь k-хамгийн ойрын хөрш (k-Nearest Neighbors) аргыг ашигладаг. Шинэ өгөгдлийг шууд сургалтын өгөгдөлтэй харьцуулж ангилдаг бөгөөд хурдан тохиргоо шаарддаггүй ч их хэмжээний өгөгдөлд удаан ажилладаг.

Bayes.BayesNet (Bayesian Network)[11]:

Магадлалд суурилсан загвар бөгөөд хувьсагчдын хоорондын хамаарлыг график байдлаар тодорхойлдог.

Trees.RandomTree:

Санамсаргүйгээр бүтэцлэгдсэн шийдвэрийн мод (decision tree) бөгөөд өгөгдлийг салаалж ангилан үр дүнг тооцдог.

Meta.Bagging (Bootstrap Aggregating)[11]:

Олон Decision Tree-г нэгтгэж, тогтворжсон үр дүн гаргадаг ensemble арга.

Function.SMO (Sequential Minimal Optimization)[11]:

SMO (Support Vector Machine) нь ангиллын хооронд хамгийн оновчтой хил заагийг тогтоох замаар өгөгдлийг жигд ялгаж ангилдаг.

Multilayer Perceptron (MLP)[11]:

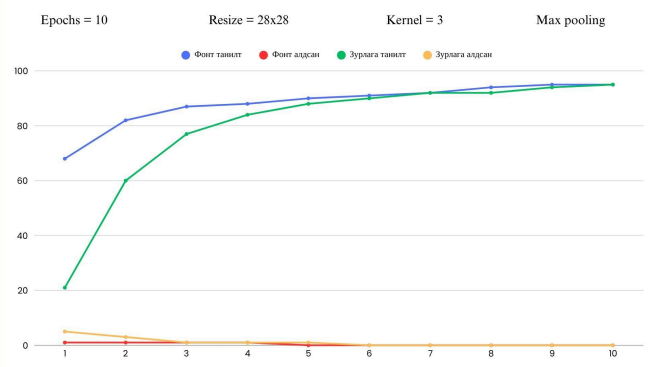
Энэ нь олон давхаргатай хиймэл мэдрэлийн сүлжээний бүтэц дээр суурилсан ба нарийн төвөгтэй онцлог бүхий дүрс, фонтын өгөгдлөөс хамаарлыг илрүүлж, тогтвортой, өндөр нарийвчлалтай үр дүн гаргадаг.

Машин сургалтын хэсгийн дүгнэлт:

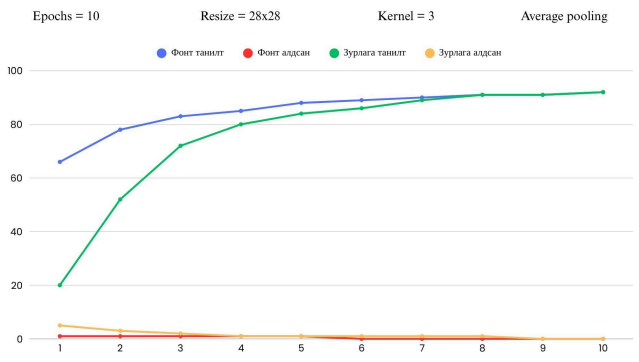
Зурлага танилт: Lazy.IBk арга нь хамгийн өндөр (93.72%) нарийвчлалтайгаар зурлагыг таньсан нь k-NN алгоритм дүрсийн хэлбэрийг сайн ялгаж чаддагтай холбоотой. Харин BayesNet 71.20% нарийвчлалтай байсан нь шугаман бус дүрслэлийг ялгаж танихад хүндрэлтэй байгааг харуулж байна.

Фонт танилт: Lazy.IBk, meta.Bagging, болон MultilayerPerceptron бүгд 93-96% өндөр нарийвчлалтайгаар фонтуудыг ялган таньж байна. Энэ нь эдгээр загварууд нарийн онцлог ялгахдаа сайн, гүн сургалт, ensemble аргыг ашигладагтай холбоотой.

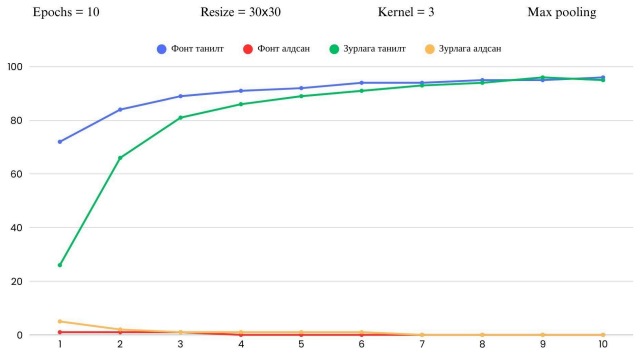
Харин энэ хэсэгд гүн сургалтын үндсэн CNN буюу Convolutional neural network аргынхаа сургалтын давтамж (epoch), оролтын зургийн хэмжээ (resize), шүүлтүүрийн хэмжээ (kernel), багасгал хийх арга (poolong) утгуудыг тус тус өөрчлөн туршиж аль аль утга дээр хамгийн өндөр танилттай ажиллахыг тогтоох зорилготойгоор хийсэн туршилтуудыг зураг 13-аас зураг 22-д харуулсан байна.



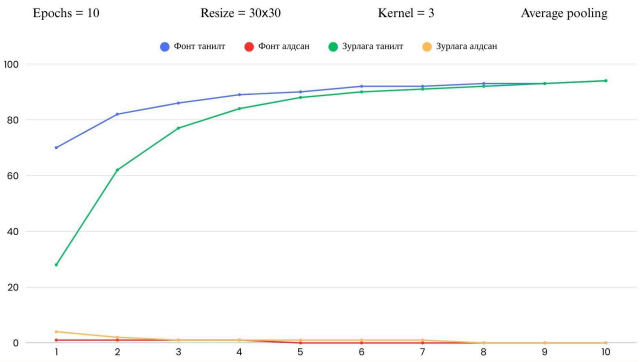
Зур. 13. Туршилт . 1.



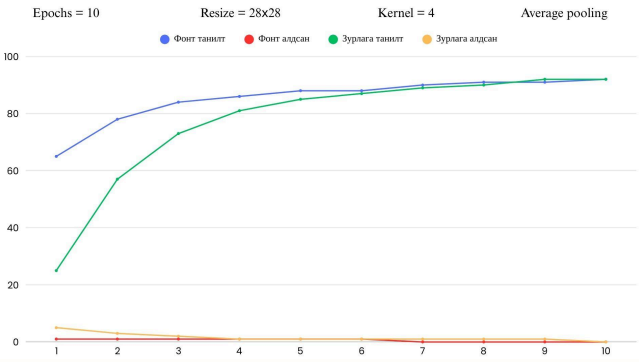
Зур. 14. Туршилт. 2.



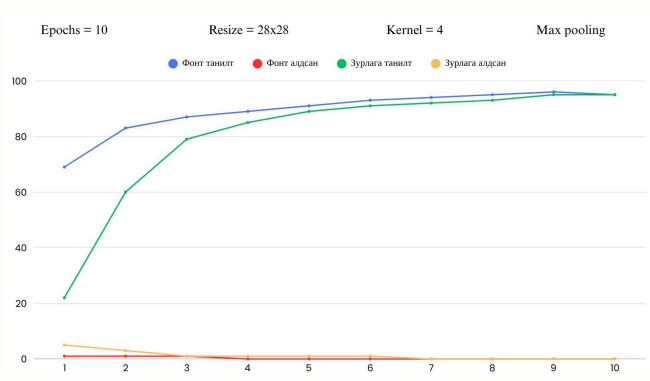
Зур. 15. Туршилт 3.



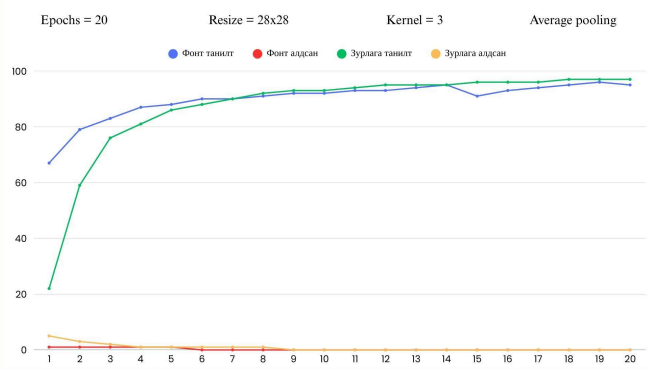
Зур. 16. Туршилт. 4.



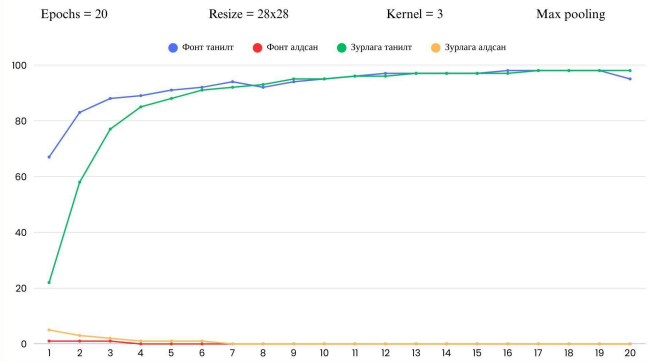
Зур. 17. Туршилт. 5.



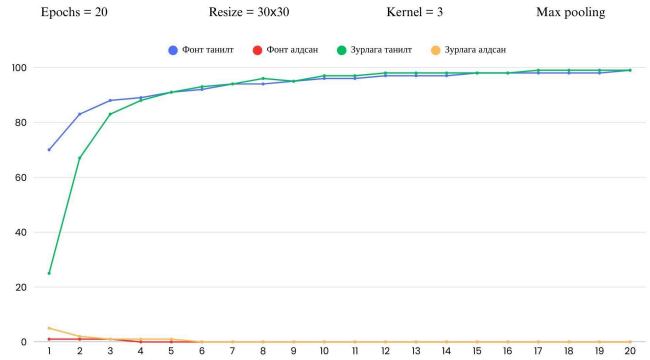
Зур. 18. Туршилт 6.



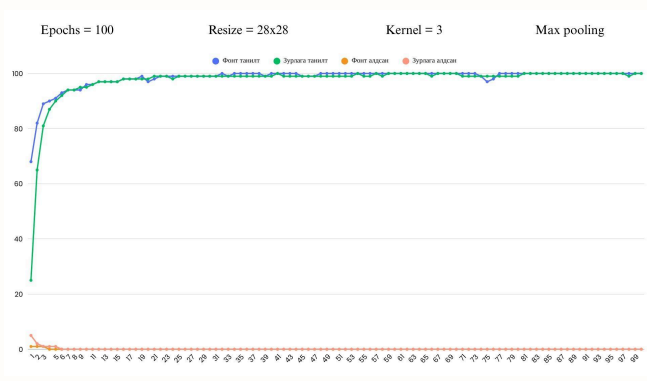
Зур. 19. Туршилт 7.



Зур. 20. Туршилт 8.

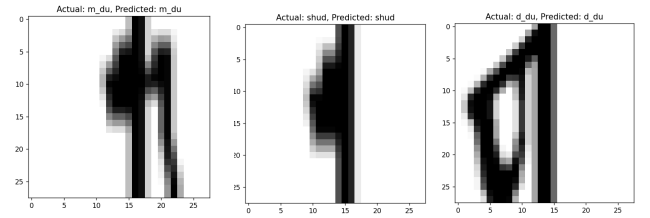


Зур. 21 Туршилт 9.

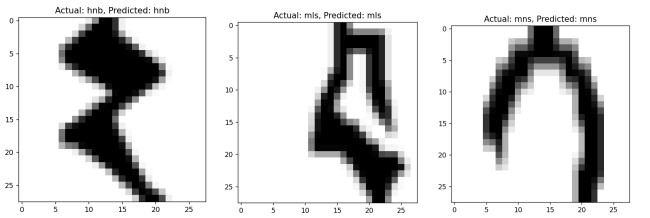


Зур. 22. Туршилт 10.

Дээрх туршилтуудын үр дүнд давталтын тоо (epoch) 7 болоход танилтын хувь 90%-аас дээш болж байсан ба давталтын тоог 20 дээш болгоход 98%-аас дээш хувьтай үр дүн гарсан. Мөн сургалтад оруулах зургийн хэмжээ том байх тусам сургалт удаж үр дүн илүү өндөр гарч байсан. Харин шүүлтүүрийн хэмжээ ихсэх тусам алдаа нэмэгдсэн ба 2x2 хэмжээтэй матрицаас хамгийн ихийг сонгож аваад 2 нүдээр алгасан явах арга буюу max pooling арга илүү өндөр танилттай болгосон. Эдгээр туршилтуудаас танилтын хувь хамгийн өндөртэй нь Туршилт 9-д зурлага танилт 98.81%, фонт танилт дээр 98.58% зэрэг үзүүлэлтэй байсан ба Туршилт 10-ын үр дүнд зурлага танилт 99.73% фонт танилт 100.00% буюу бүрэн танилт хийсэн.



Зур. 23. Үр дүнгийн жишээ (дунд м, шүд, дунд д)



Зур. 24. Үр дүнгийн жишээ (Хүмүүн бичиг, Mongolian script, Mongol Script)

V. ДҮГНЭЛТ

Бид энэхүү судалгаандаа дижитал хэлбэрийн хэвлэмэл монгол бичгийн зураг дээр дүрс боловсруулалтын аргууд ашиглан хэсэгчлэн хуваж машин сургалт болон гүний сургалтаар фонт болон зурлага танилт хийх хялбар аргыг танилцуулаа. Өгөгдөл дээр анхан шатны боловсруулалт хийхдээ багана, мөрийн пикселүүдийн утга дээр анализ хийсэн бөгөөд гүн сургалтын CNN(Convolutional neural network) арга ашиглан 95% аас дээш танилттай үр дүнгүүд гаргасан. Цаашдын судалгаанд танигдсан тэмдэгтүүдийг нэгтгэх тэмдэгт мөр үүсгэх өөр бусад гаралтын дараах боловсруулалтуудыг хийх шаардлагатай.

VI. НОМ ЗҮЙ

[1] Ц. Шагдарсүрэн, “Монголчуудын үсэг бичгийн товчоон”, Bibliotheca Mongolica, 2001

[2] Монгол Улсын Засгийн газар. (2020). Монгол бичгийн үндэсний хөтөлбөр III (Засгийн газрын тогтоол №96). Улаанбаатар хот.

[3] Nobuyuki Otsu, “A Threshold Selection Method from Gray-Level Histograms”, IEEE Transactions on Systems, Man, and Cybernetics, January 1979

[4] Timothy Malche, “What is Template Matching? An Introduction”, roboflow, October 2024

[5] Ц. Шагдарсүрэн, “Туурга тусгаар Монгол Улсын төрд үндэсний бичиг нь үгүйлэгдэж байна” МХББНҮХолбооны албан ёсны сайт, Нэгдүгээр сар 2006

[6] Rikiya Yamashita1,2 & Mizuho Nishio1,3 & Richard Kinh Gian Do2 & Kaori Togashi1, “Convolutional neural networks: an overview and application in radiology”, Insights into Imaging, June 2018

[7] Byambasuren Ivanov, Uuganbaatar Dulamragchaa, Markhamet Musa, Otgonnaran Ochirbat, “Recognition of traditional Mongolian script using primitives and template matching methods”. International Journal of Scientific and Engineering Research, February 2019.

[8] Hongxi Wei, Guanglai Gao, “Machine-Printed Traditional Mongolian Characters Recognition Using BP Neural Networks”, January 2010

[9] Hongxi Wei, Ya Wen, Weiyuan Wang, Guanglai Gao, “Convolutional Neural Network for Machine-Printed Traditional Mongolian Font Recognition: 25th International Conference, ICONIP 2018, Siem Reap, Cambodia, December 13–16, 2018, Proceedings, Part V”

[10] Ahmed Lawgali, “A Survey on Arabic Character Recognition”, February 2015

[11] "Weka Primer – Classifiers," Weka Wiki, University of Waikato. [Online]. Available: Primer - Weka Wiki. Apr. 6, 2025.

Уул уурхайн салбарын кибер аюулгүй байдлын төлөвшлийг тодорхойлох судалгаа

Хүүхээгийн Мөнхтөр
Монголын алт “МАК” ХХК
Монгол Улс, Улаанбаатар хот
m23md1209@ufe.edu.mn

Дэмчигийн Болормаа
Ахисан Түвшний сургууль
СЭЗИС
Монгол Улс, Улаанбаатар хот
dbolormaa@ufe.edu.mn

***Хураангуй*— Монгол улсын Уул уурхайн салбар нь ДНБ-ий 25 гаруй хувь, улсын нэгдсэн төсвийн орлогын 30 гаруй хувь, гадаадын шууд хөрөнгө оруулалтын 73 орчим хувь, нийт экспортын 92 гаруй хувийг дангаар бүрдүүлдэг эдийн засгийн тулгуур салбар юм (miningmongolia, 2024). Манай улсын эдийн засгийн тогтвортой байдалтай шууд хамааралтай тус салбар нь харилцаа холбооны технологид (ХХТ) улам бүр тулгуурладаг болсон нь кибер халдлагад өртөх эрсдэлийг нэмэгдүүлж байна. Энэхүү судалгааны ажил нь С2М2 төлөвшлийн загварт суурилсан кибер аюулгүй байдлын (КАБ) 10 домайн буюу хүрээгээр уул уурхайн салбарын КАБ-ын төлөвшлийн өнөөгийн түвшинг тодорхойлж, тухайн түвшинг ахиулах иж бүрэн стратеги боловсруулж, тасралтгүй сайжруулан боломжийг олгох саналыг дэвшүүлэх зорилготой юм. Судалгааны үр дүнгээс харахад уул уурхайн компаниудын КАБ-ын төлөвшлийн Т1 түвшний шаардлагуудыг 37 хувиар, Т2 түвшний шаардлагуудыг 28 хувиар тус тус хангаж, салбарын ерөнхий дүнгээрээ түвшин Т1-д буюу кибер аюулгүй байдлын үндсэн бодлого, дүрэм, үйл явцыг хэрэгжүүлж эхэлж байгаа түвшинд байна.**

Түлхүүр үг*— *Уул уурхайн салбар, кибер аюулгүй байдал (КАБ), кибер аюулгүй байдлын төлөвшлийн түвшин, төлөвшлийн С2М2 загвар

I. Удиртгал

2020 онд EY Global Information Security Survey (GISS) байгууллагын уул уурхайн салбарын кибер аюулгүй байдлын судалгаанд оролцогчдын 71 хувь нь кибер халдлагад өртөх нь ихэссэн гэж мэдээллэсэн бол уул уурхай, металлын салбарын удирдах ажилтнуудын 55 хувь нь өсөн нэмэгдэж буй аюул заналхийллийг удирдах чадвартаа гүнээ санаа зовж байгаагаа илэрхийлжээ [1]. Кибер довтолгооны давтамж нэмэгдэж байгаа энэ үед уул уурхайн салбарын өсөн нэмэгдэж буй дайралтаас өөрийгөө хамгаалах бэлтгэлгүй хэвээр байгааг олон шинжээч анхааруулж байна. Purple Security компанийн кибер аюулгүй байдлын мэргэжилтнүүдийн 2023 онд гаргасан тайлангаас харахад уул уурхай, металлын компаниудын 54% нь кибер аюулгүй байдлын осолд өртсөн бол 40% нь гадны аюулын тоо өмнөх жилээс огцом өссөнийг тэмдэглэжээ [2].

Монгол Улс 2021 онд кибер аюулгүй байдлын хуулийг баталж, хууль эрх зүйн, техникийн, байгууллагын бүтэц болон бусад хүчин зүйлсийг бий болгож эхэлснээр 2024 оны Дэлхийн кибер аюулгүй байдлын индекс (Global Cybersecurity Index)-т 3-р түвшний “Бэхжиж буй” ангилалд орж түвшингөө ахиулсан ч тэдгээрийг бүрэн хэрэгжүүлэх тал дээр учир дутагдалтай байна [3].

Уул уурхайн төрийн болон хувийн хэвшлийн томоохон топ аж ахуй нэгжүүд нь уурхайн ухаалаг хяналтын систем Pitram, Minestar системүүдийг нэвтрүүлж үйлдвэрлэлийн үйл ажиллагааны зардлыг бууруулж, тоног төхөөрөмжийн ашиглалтын хугацааг нэмэгдүүлсэн. Мөн ERP, Office 365, Пүүний 24/7 систем, AVG технологи, хиймэл оюун ухааны систем, 5G сүлжээнд суурилсан ухаалаг боомтын системүүдийг өргөн ашигладаг. Уул уурхайн тоног төхөөрөмжүүд зүйлсийн интернэт холбогдож /IoT/, цаашид хиймэл оюун ухаанд суурилсан уул уурхайн тоног төхөөрөмжүүд бизнесийн үйл ажиллагаанд нэвтэрч эхэлсэн. Эдгээр технологи нь бүгд сүлжээнд холбогдож, цаашид клауд орчинд холбогдох хэрэгцээ шаардлага өсөн нэмэгдэж байгаа үед кибер халдлага зөрчлөөс хэрхэн сэргийлэх, түүнд бэлэн байгаа эсэхийг кибер аюулгүй байдлын төлөвшлийн загвар дээр тулгуурлан төлөвшлийн үе шатыг тодорхойлох, дараагийн түвшинд хүрэхэд авч хэрэгжүүлэх гол арга хэмжээг олж илрүүлэх зайлшгүй шаардлага урган гарч ирж байна.

Хэдийгээр Монгол улсын хэмжээнд уул уурхай салбарт 2024 оны 3 улирлын байдлаар 1196 аж ахуй нэгж үйл ажиллагаа явуулж байгаа боловч 20 хүрэхгүй орчим хувь нь дээр дурьдсан үйлдвэрлэлийн системүүдийг ашиглаж байна.

II. Онолын судалгаа

A. Кибер халдлага ба аюулгүй байдал

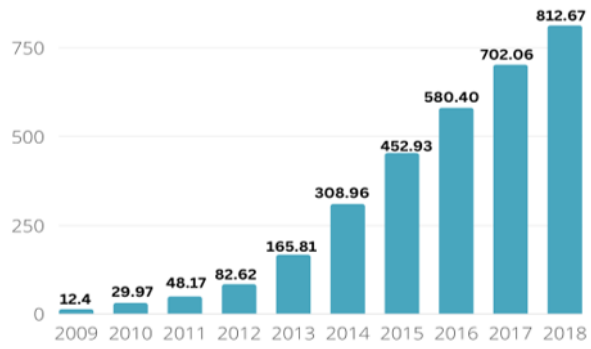
1940-өөд онд дэлхийн анхны дижитал компьютер нэвтэрснээс хойш кибер халдлагын асуудал үүсэн бий болсон. Сүлжээ, компьютерын систем, дижитал төхөөрөмжид зөвшөөрөлгүй нэвтрэх замаар өгөгдөл, программ болон бусад хөрөнгийг хулгайлах, мэдээллийг тараах, өөрчлөх, идэвхгүй болгох, устгах зорилготой аливаа оролдлогыг кибер халдлага гэнэ [4]. Энэхүү халдлагын эсрэг авдаг арга хэмжээг кибер аюулгүй байдал гэх бөгөөд өнөөгийн дижитал ертөнцөд байгууллагын хөрөнгө, үйл ажиллагаа, нэр хүндийг хамгаалахад чухал ач холбогдолтой юм.

Монгол Улсын их хурлын 2021 онд батлагдсан Кибер аюулгүй байдлын тухай хуульд "Кибер аюулгүй байдал гэж кибер орчинд мэдээллийн бүрэн бүтэн, нууцлагдсан, хүртээмжтэй байдал хангагдсан байхыг хэлнэ” гэж тодорхойлсон [5] Үүнд:

- Мэдээллийн нууцлал: Мэдээллийг зөвшөөрөлгүй хандалтаас хамгаалах.
- Мэдээллийн бүрэн бүтэн байдал: Мэдээллийг зөвшөөрөлгүй өөрчлөлтөөс хамгаалах.
- Мэдээллийн хүртээмжтэй байдал: Мэдээллийг зөвшөөрөгдсөн хэрэглэгчдэд шаардлагатай үед нь ашиглах боломжийг хангах гэсэн үг юм.

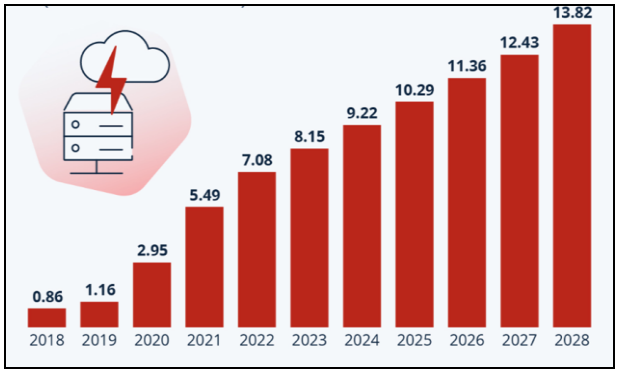
Кибер халдлагын тоо 2009 оноос эрс нэмэгдэж байгааг (Зураг 1) судалгааны байгууллагууд өөрсдийн тайлангуудад тэмдэглэсэн [6].

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE



Зур. 1. Кибер халдлагын өсөлтын хурд

Кибер халдлага өсөж байгаатай холбоотойгоор түүний жилд учруулах хор хөнөөл нэмэгдсээр, Statista Market Insight-ын байгууллагын тооцоолсноор кибер гэмт хэргийн дэлхийн өртөг ойрын дөрвөн жилд өсч, 2024 онд 9,22 их наяд доллар байсан бол 2028 он гэхэд 13,82 их наяд ам.доллар болж өсөх төлөвтэй байна. (Зур. 2).



Зур. 2. Кибер гэмт хэргийн жилд учруулж буй хохирол [7]

Кибер эрсдлийн орчны хамрах хүрээ, боловсронгуй байдал, учруулж буй хор хохирол ийнхүү өссөөр байгаа тул байгууллагын КАБ-ын чадавхийг тасралтгүй сайжруулах асуудал байгууллагын мэдээллийн технологийн үр дүнтэй менежментийн хувьд нэн чухал болоод байна.

B. Кибер аюулгүй байдлын төлөвшил

Байгууллагууд үйл ажиллагааны тасралтгүй байдал, дүрэм журмыг дагаж мөрдөхийн зэрэгцээ аюул заналаас үр дүнтэй хамгаалж, эрсдлийг бууруулж, хөгжиж буй кибер аюулгүй байдлын орчинд дасан зохицож чаддаг байх нь чухал. Байгууллагын КАБ-ын үйл ажиллагаа, бодлого, үйл явцын хөгжлийн түвшин, үр нөлөөг КАБ-ын төлөвшлөөр (Cybersecurity maturity) тодорхойлдог. Энэ нь тухайн байгууллага өөрийн тогтоосон аюулгүй байдлын тогтолцоонд тулгуурлан кибер аюулаас хэр зэрэг идэвхтэй хамгаалж урьдчилан сэргийлж, илрүүлж, хариу арга хэмжээ авч байгаа, Мөн бизнесийн тасралтгүй байдал, хууль журмыг дагаж мөрдөхийн зэрэгцээ аюул заналхийлэлд хэр дасан зохицох чадвартай байгааг нь хэмждэг [7].

Кибер аюулгүй байдлын төлөвшлийн загвар (CMM[1]) нь бүх төрлийн байгууллага, салбарын онцлог хэрэгцээнд нийцүүлэн уян хатан байхаар бүтээгдсэн, байгууллагуудад КАБ-ын байр сууриа үнэлж, сайжруулахад тусалдаг загвар юм. Энэхүү загвар нь КАБ-ын сул талуудыг тодорхойлох, үйл ажиллагааг эрэмбэлэх, кибер аюулаас хамгаалах ахиц дэвшлийг хэмжихэд тусалдаг.

CMM загвар нь ихэвчлэн хэд хэдэн түвшинд хуваагддаг бөгөөд тус бүр нь КАБ-ын төлөвшлийн өөр өөр түвшин илэрхийлдэг. Түвшинүүдийг үндсэн, дунд, ахисан гэж тодорхойлдог:

- *Суурь түвшин нь* КАБ-ын чиглэлээр шинээр ажиллаж байгаа эсвэл КАБ-ын арга хэмжээг хараахан хэрэгжүүлж эхлээгүй байгаа байгууллагуудын эхлэх цэг юм.
- *Дунд түвшин нь* КАБ-ын зарим арга хэмжээг хэрэгжүүлсэн боловч сайжруулах шаардлагатай байгаа байгууллагуудыг тодорхойлдог.
- *Ахисан түвшин* нь КАБ-ын цогц арга хэмжээг хэрэгжүүлсэн, кибер аюулгүй байдлын байр сууриа тасралтгүй сайжруулж буй байгууллагуудыг тодорхойлдог.

C. КАБ-ын төлөвшлийн загварууд ба тэдгээрийн ялгаа

Олон улсад КАБ-ын төлөвшлийн хэд хэдэн түгээмэл хэрэглэгддэг загварууд байдаг бөгөөд байгууллагын салбар, зорилго, дагаж мөрдөх шаардлага зэргээс шалтгаалан загвар бүр өөрийн давуу талуудтай. (Хүснэгт I). Хамгийн түгээмэл ашиглагддаг загваруудад дараах загварууд ордог:

ХҮСНЭГТ I. КАБ-ЫН ТӨЛӨВШЛИЙН ЗАГВАРУУД

Загвар	Зорилго	Зориулалт	Бүтэц	Тохиромж
C2M2	КАБ-ын төлөвшлийг үнэлэх, сайжруулах	Чухал дэд бүтэц, илүү өргөн салбарууд	10 хүрээ, 4 төлөвшлийн түвшин	Төрөл бүрийн салбарын байгууллагууд
NIST CSF	Эрсдэлийн удирдлага, кибер аюулгүй байдлын эрсдэлийг бууруулах	Уян хатан, өөрчлөх боломжтой, эрсдэлд суурилсан арга	5 үндсэн чиг үүрэг, хэрэгжүүлэх шатууд	Бүх төрлийн байгууллагууд, ялангуяа зохицуулалттай салбарууд
CMMC	DoD кибер аюулгүй байдлын стандартыг хангадаг	CUI-г хариуцдаг батлан хамгаалахын гүйцэтгэгчид	Тодорхой практик туршлага бүхий төлөвшлийн 5 түвшин	АНУ-ын батлан хамгаалахын гэрээт гүйцэтгэгчид болон ханган нийлүүлэгчид

[a] Судлаачийн боловсруулснаар

1. *Cybersecurity Capability Maturity Model (C2M2)* - КАБ-ын чадавхын төлөвшлийн загвар [8]
2. *NIST Cybersecurity Framework (CSF)* - кибер аюулгүй байдлын хүрээ [9]
3. *Cybersecurity Maturity Model Certification (CMMC)* - КАБ-ын төлөвшлийн загвар гэрчилгээ [10]

[1] CMM – Capability Maturity Model

Судалгаануудаас [8, 9, 10] дүгнэхэд C2M2 нь тусгай зохицуулалтын шаардлагад анхаарлаа хандуулалгүйгээр өргөн хүрээнд өөрсдийн КАБ-аа үнэлж, сайжруулах хүсэлтэй байгууллагуудад хамгийн тохиромжтой. Харин NIST CSF нь эрсдэлийн удирдлагад анхаарлаа төвлөрүүлэхийн зэрэгцээ КАБ-ын туршлагыг хэрэгжүүлэх уян хатан байдлыг эрэлхийлдэг байгууллагуудад тохиромжтой. CMMC нь АНУ-ын Батлан хамгаалах яамнаас (DoD) батлан хамгаалах үйлдвэрлэлийн баазад тусгайлан боловсруулсан загвар бөгөөд АНУ-ын Батлан хамгаалах яамтай хамтран ажиллах гэрээлэгчдээс шаардаг гэрчилгээжүүлэлтэд төвлөрсөн загвар ба тодорхой төлөвшлийн түвшин, аюулгүй байдлын практик туршлагын шаардлагуудыг агуулсан загвар юм.

Зарим шалгуур үзүүлэлтүүд нь уул уурхайн техникийн асуудлуудыг шууд тусгахгүй байж болох талтай ч эрчим хүчний салбарт ашиглагддаг, онц чухал дэд бүтэц болон хүнд үйлдвэрүүдэд зориулагдсан КАБ-ын төлөвшлийн C2M2 загвар нь бусад загвараас илүү сайн үйл ажиллагаа нь OT (Operational Technology)-д тулгуурладаг уул уурхайн салбарт бусад төлөвшлийн загваруудаас илүү сайн тохирохоор байна. C2M2 загвар нь эрсдэлийн удирдлага, ослын хариу арга хэмжээ болон бусад гол салбаруудыг хамарсан цогц хяналттай, төлөвшлийн түвшин нь ахиц дэвшлийг хянах, зорилгоо тодорхойлох боломжийг олгодог давуу талтай юм.

III. СУДАЛГААНЫ ДИЗАЙН БА АРГА ЗҮЙ

Энэхүү судалгааны ажлаар дараахь асуултуудад хариулахыг зорьсон. Үүнд:

- Уул уурхайн салбарын КАБ ямар түвшинд байна вэ?
- Салбарын КАБ-н өнөөгийн түвшинг ахиулахын тулд ямар арга хэмжээ авах хэрэгтэй вэ?

Эмпирик судалгаанд уул уурхайн салбарын эх олонлогоос үйлдвэрлэлийн үндсэн үйл ажиллагааны томоохон дараах мэдээллийн системүүдийг амжилттай нэвтрүүлсэн тус *топ таван компанийг* сонгон авсан. Үүнд:

1. Дэлхийд тэргүүлэгч уурхайн хяналтын систем болох Питрам системийг үйл ажиллагаандаа нэвтрүүлсэн дөрвөн уурхай Монгол Улсын 2023 оны топ 10 аж ахуй нэгжид шалгарсан.
2. Жолоочгүй автомат удирдлага бүхий AGV тээвэр, 5G сүлжээ зэрэг өндөр технологиудыг нэвтрүүлсэн нэг ухаалаг боомт

Дээрх 5 компаний *мэдээллийн технологи хариуцсан удирдах ажилтнуудаар* судалгааны түүврээ бүрдүүлсэн.

Судалгааны өгөгдөл цуглуулах арга хэрэгслийн хувьд АНУ-ын эрчим хүчний яамнаас гаргасан КАБ-ын төлөвшлийг үнэлэх *C2M2 загварын 2.1 хувилбарыг* сонгосон. C2M2 загвар нь дөрвөн түвшинд (T) байгууллагын КАБ-ын байдлыг тодорхойлдог. Үүнд:

- *T0 (No practices implemented)* түвшин нь КАБ-ын бодлого, процессыг хараахан хэрэгжүүлээгүй байгааг илтгэх бөгөөд төлөвшлийн түвшинг сайжруулахын тулд дээд түвшнүүдтэй холбоотой дүрмүүдийг хэрэгжүүлж эхлэх суурь цэг болдог.
- *T1 (Basic)* нь КАБ-ын үндсэн бодлого, дүрмийг хэрэгжүүлж эхэлсэн байдаг боловч эдгээр нь бүрэн ажиллагаатай биш эсвэл тогтмол дагаж мөрддөггүйг илтгэнэ.
- *T2 (Intermediate)* нь аюулгүй байдлын хяналтуудыг хэрэгжүүлсэн бөгөөд дүрмийг тогтмол хянах, шинэчлэхээс гадна эрсдэлийн удирдлагад анхаардгийг харуулна.
- *T3 (Advanced)* нь Өндөр түвшинд КАБ-ын бодлого, дүрмийг хэрэгжүүлдэг бөгөөд үүнд аюулыг идэвхтэй илрүүлэх, хариу арга хэмжээ авах чадвартай, тасралтгүй сайжруулах үйл явц нь үйл ажиллагаанд бүрэн нэвтэрсэн байдгийг илтгэдэг.

C2M2 загварын доорх 10 домайныг агуулдаг:

1. Asset - Хөрөнгийн аюулгүй байдал:
2. Threat - Аюул занал
3. Risk- Эрсдэл
4. Access- Хандалтын эрх
5. Situation-КАБ-ын нөхцөл байдлыг удирдах
6. Response - Кибер халдлагын үед хариу үзүүлэх
7. Third parties - Гуравдагч талын аюулгүй байдал
8. Workforce-Хүний нөөц
9. Architecture - Архитектурын аюулгүй байдал
10. Program – Хөтөлбөр

C2M2 загварын 2.1 хувилбар нь 10 домайны нийт гурван түвшний 356 хяналт буюу шалгуурыг агуулдаг. Үйлдвэрлэлийн мэдээллийн системүүдийг нэвтрүүлсэн уул уурхайн топ компаниуд нь 2021 оноос л КАБ-ын хууль, олон улсын холбогдох стандарт, фрэймворкуудын дагуу КАБ-ын туршлагуудыг нэвтрүүлж эхэлсэн тул C2M2 загварын *T3 (Advanced)* түвшинд хүрээгүй байгаа гэж үзэж судалгааны өгөгдөл цуглуулах хэрэгсэлд тус түвшний шалгууруудыг хамруулаагүй. Судалгааны түүвэрт хамрагдсан компаниудын КАБ-ын төлөвшлийн өнөөгийн түвшинг тодорхойлохын тулд T1 ба T2-ын харгалзах нийт 222 шалгууруудыг дэд шалгууруудаар нь ерөнхийлөн бүлэглэн нэгтгэж уул уурхайн компаниудад зориулсан 143 асуулт болгон Google Form-д оруулж өгөгдөл цуглуулав [11].

Google Form дээрх анкетын асуулгуудыг бүрэн бөглүүлж үнэн, бодит мэдээллийг цуглуулахын тулд түүвэрт хамрагдсан компаниудын холбогдох мэдээллийн технологийн удирдах ажилтнуудад Үндэсний Аудитын Газрын аудитор Ж.Чинбатаар дамжуулан хүсэлт тавьсан. Үүний үр дүнд нөхцөл байдлын судалгаанд хамрагдсан нийт 5 компанийхан бүгд асуултад бүрэн хариулсан байсан.

IV. СУДАЛГААНЫ ӨГӨГДӨЛ БОЛОВСРУУЛАЛТ, ҮР ДҮН

Судалгааны асуумжид дараахь 10 домайн буюу зорилтод харгалзах ерөнхий асуултуудын хүрээнд нийт 143 асуултад уул уурхайн топ таван компаний мэдээллийн технологи хариуцсан удирдах ажилтнуудаас хариулт

авч, Excel програмыг ашиглан өгөгдөлд боловсруулалт хийж, үр дүнг графикаар илэрхийлэв.

1. ***Asset - Хөрөнгийн аюулгүй байдлын*** *хамрах хүрээ*

Энэ домайн нь уурхайн бүх төрлийн мэдээллийн систем, сүлжээ, төхөөрөмж, програм хангамж, өгөгдөл зэрэг хөрөнгийг аюулгүй байлгах зорилготой.

Асуулт 1: Танай байгууллага дээр хөрөнгийн аюулгүй дараах бодлогуудаас аль нь хэрэгждэг вэ?

2. ***Threat - Аюул занал:***

КАБ-ын аюулыг тодорхойлох, үнэлэх, удирдах практик, үйл явцыг хамардаг. Энэ домэйн нь байгууллагын мэдээллийн технологи (МТ) болон үйл ажиллагааны технологийн (ОТ) хөрөнгөд нөлөөлж болох боломжит аюулыг ойлгож, бууруулахад чиглэгддэг.

Асуулт 2: Танай байгууллага дээр кибер аюул заналаас хамгаалах дараах бодлогуудаас аль нь хэрэгждэг вэ?

3. ***Risk- Эрсдэл:***

Байгууллагын болон кибер эрсдэлийг тодорхойлох, тодорхойлсон эрсдэлүүдийн нөлөөлөл, магадлалыг үнэлэх, эрсдэлд хариу арга хэмжээ авах болон эрсдэлийн удирдлагын хөтөлбөрийг байнга хянах, үнэлэх үйл ажиллагаануудыг багтаана.

Асуулт 3 : Танай байгууллага дээр кибер аюулгүй байдлын эрсдэлээс хамгаалах дараах бодлогуудаас аль нь хэрэгждэг вэ?

4. ***Access- Хандалтын эрх***

Энэ домайн нь байгууллагын мэдээлэл болон системүүдэд хандах эрхийг удирдах, хянах, байнга шалгах, цуцлах үйл явцыг агуулна.

Асуулт 4: Танай байгууллагад дараах хандах эрхийн бодлого журам хэрэгждэг үү?

5. ***Situation-Аюулгүй байдлын нөхцөл байдлыг удирдах***

Энэ зорилт нь байгууллагын КАБ-ын нөхцөл байдлыг байнга хянах, үнэлэх, нөхцөл байдлын мэдрэмжийг хадгалах, мэдээлэх, удирдлагыг сайжруулах үйл явц юм.

Асуулт 5: Кибер аюулгүй байдлын нөхцөл байдлыг удирдах дараах бодлогуудаас аль нь танай байгууллагад хэрэгждэг вэ?

6. ***Response - Кибер халдлагын үед хариу үзүүлэх***

Энэ зорилт нь байгууллагын кибер аюулгүй байдлын үйл явдлуудыг тодорхойлох, хянах, удирдах, ослуудыг удирдах, нь байгууллагын үйл ажиллагааг тасралтгүй үргэлжлүүлэх, ослын дараа үйл ажиллагааг сэргээн босгох үйл явцыг хамардаг.

Асуулт 6: Кибер халдлагын үед хариу үзүүлэх дараах бодлогуудаас аль нь хэрэгждэг вэ?

7. ***Third parties - Гуравдагч талын аюулгүй байдал***

Энэ зорилт нь байгууллагын нийлүүлэгчид болон бусад гуравдагч талуудаас үүдэлтэй кибер эрсдэлийг тодорхойлох, үнэлэх, эрсдэлд хариу арга хэмжээ авах, хянах үйл явцыг агуулдаг.

Асуулт 7: Гуравдагч талын аюулгүй байдлыг хангах дараах бодлогуудаас аль нь хэрэгждэг вэ?

8. ***Workforce-Хүний нөөц***

Энэ домайнд байгууллагын КАБ-ын хэрэгцээг хангахуйц ажиллах хүчнийг төлөвлөх, удирдах, ажилтнуудын кибер аюулгүй байдлын ур чадвар, мэдлэгийг тасралтгүй сайжруулах, үнэлэх, сайжруулах арга хэмжээнүүдийг хэрэгжүүлэх үйл явцууд хамаарна.

Асуулт 8: Кибер аюулгүй байдлыг хангах хүний нөөцийн бодлогуудаас аль нь хэрэгждэг вэ?

9. ***Architecture - Архитекторын аюулгүй байдал***

Энэ зорилт нь байгууллагын кибер аюулгүй байдлын архитектурыг төлөвлөх, хэрэгжүүлэх, удирдах, архитектурыг хууль, дүрэм журам, стандартуудтай нийцүүлэх үйл явц юм.

Асуулт 9: Кибер аюулгүй байдлын архитектурын бодлогуудаас аль бодлого нь хэрэгждэг вэ?

10. ***Program – Хөтөлбөр***

Тус домайн нь байгууллагын кибер аюулгүй байдлын хөтөлбөрийг төлөвлөх, хэрэгжүүлэх, удирдах, хууль, дүрэм журам, стандартуудтай нийцүүлэх үйл явц юм.

Асуулт 10: Кибер аюулгүй байдлыг хангах хөтөлбөрийн бодлогуудаас аль нь хэрэгждэг вэ?

Түүвэрт хамрагдсан уул уурхайн компаниудын КАБ-ын төлөвшлийн төвшинг тодорхойлдог C2M2 загварын шалгууруудыг хэрэгжүүлсэн байдлыг компани тус бүрээр болон салбарын хэмжээнд Хүснэгт II-д нэгтгэн харуулав.

Судалгааны үр дүнгээс харахад түүвэрт хамрагдсан компаниудын төлөвшлийн Т1 түвшний шалгууруудын хэрэгжилт нь Зур.3-т харуулсанчлан 15 хувиас 65 хувийн хооронд хэлбэлзэж байна.

Зур. 3. Уул уурхайн топ компаниуын КАБ-ын төлөвшлийн Т1 түвшний хэрэгжилт

Домайнуудаас харахад компаниуд байгууллагын мэдээллийн технологи (МТ) болон үйл ажиллагааны технологийн (ОТ) хөрөнгөд нөлөөлж болох боломжит аюулыг ойлгож, бууруулах чиглэлээр хамгийн их анхаарал тавьж байгаа (65%) бол харин нийлүүлэгчид болон бусад гуравдагч талуудаас үүдэлтэй кибер

эрсдэлийг тодорхойлох, үнэлэх, эрсдэлд хариу арга хэмжээ авах, хянах үйл ажиллагааг хамгийн сул хэрэгжүүлж байгааг (15%) судалгааны үр дүн (Зур.1) харуулж байна.

ХҮСНЭГТ II. УУЛ УУРХАЙН ТОП КОМПАНИУДЫН КАБ-ЫН ТӨЛӨВШЛИЙН ТҮВШНИЙ ШАЛГУУРЫГ ХАНГАСАН БАЙДАЛ

№	Аюулгүй байдлын хүрээ	Төлөвшлийн түвшин	Харгалзах шалгуурууд	Түүвэрт хамрагдсан уул уурхайн компаниудын хэрэгжүүлж буй шалгуурын тоо						Салбарын хэмжээний хэрэгжилт
				A	B	C	D	E	Нийт	
1	ASSET-Хөрөнгийн аюулгүй байдал	T1	4	4	3	0	4	1	12	60%
		T2	8	5	7	0	8	2	22	55%
2	THREAT- Аюул занал	T1	4	4	4	3	1	1	13	65%
		T2	9	7	4	3	5	0	19	42%
3	RISK - Эрсдэл	T1	4	2	1	1	0	0	4	20%
		T2	12	8	6	2	4	0	20	33%
4	ACCESS- Хандалтын эрх	T1	7	2	4	2	2	1	11	31%
		T2	11	2	1	1	1	0	5	9%
5	SITUATION-Аюулгүй байдлын нөхцөл байдлыг удирдах	T1	3	3	1	1	0	0	5	33%
		T2	4	1	0	0	2	0	3	25%
6	RESPONSE – Кибер халдлагын үед хариу үзүүлэх	T1	6	2	0	2	1	1	6	20%
		T2	18	8	0	4	5	0	17	19%
7	THIRD PARTIES - Гуравдагч талын аюулгүй байдал	T1	4	0	1	2	0	0	3	15%
		T2	7	1	1	1	2	1	6	17%
8	WORKFORCE-Хүний нөөц	T1	7	5	6	4	1	1	17	49%
		T2	4	0	4	0	3	0	7	35%
9	ARCHITECTURE- Архитектерын аюулгүй байдал.	T1	5	3	4	4	0	1	12	48%
		T2	17	3	9	7	5	0	24	28%
10	PROGRAM - Хөтөлбөр	T1	2	0	1	1	0	1	3	30%
		T2	7	1	0	4	2	0	7	20%

Зур.4-т харуулсанчлан уул уурхайн компаниуд КАБ-ын төлөвшлийн Т2 түвшний шалгуурууд 9 хувиас 55 хувийн хооронд хэрэгжүүлж байна. Судалгааны үр дүнгээс харахад компаниуд уурхайн бүх төрлийн мэдээллийн систем, сүлжээ, төхөөрөмж, програм хангамж, өгөгдөл зэрэг хөрөнгийг аюулгүй байлгахад (55%) үйл ажиллагаа түлхүү хэрэгжүүлж байна. Харин компаниуд байгууллагын мэдээлэл болон системүүдэд хандах эрхийг удирдах, хянах, байнга шалгах, цуцлах үйл явцыг тогтмол тууштай хэрэгжүүлэхгүй (9%) байна.

Зур. 4. Уул уурхайн топ компаниуын КАБ-ын төлөвшлийн Т2 түвшний хэрэгжилт

Төлөвшлийн түвшний нэгдсэн дүнгээр авч үзвэл уул уурхайн салбарын компаниуд КАБ-ын төлөвшлийн Т1 түвшний хяналтуудыг 37 хувиар харин Т2 түвшний хяналтуудыг 28 хувиар л хэрэгжүүлж байгааг судалгааны үр дүн харуулж байна (Зур.5).

Зур. 5. Уул уурхайн салбарын КАБ-ын төлөвшлийн түвшин

V. СУДАЛГААНЫ ДҮГНЭЛТ

АНУ-ын Эрчим хүчний яамнаас гаргасан C2M2 загварын шалгуурууудыг ашиглан уул уурхайн салбарын КАБ-ын төлөвшлийн түвшинг тодорхойлох судалгааны үр дүнгээс дараахь дүгнэлтүүдийг хийж байна.

Төлөвшлийн Т1 ба Т2 түвшний хяналтуудын хэрэгжилтийн үр дүнгээс харахад уул уурхайн компаниудын КАБ-ын Т1 түвшин дээр хамгийн өндөр хэрэгжилттэй гарсан домайнууд болох *"Asset -*

Хөрөнгийн аюулгүй байдал" болон *"Threat - Аюул занал"* гэсэн шалгуур Т2 түвшний шалгуур дээр мөн өндөр хувьтай гарсан байна. Энэ нь компаниуд МТ болон ОТ хөрөнгийн аюулгүй байдлыг ойлгож, эрсдэлийг бууруулах талаар нэлээд анхаарч байгааг нотолж байна.

Харин Т1 ба Т2 түвшин дээр хоёулан дээр нь хамгийн сул хэрэгжиж байгаа шалгуурууд нь *"Third parties - Гуравдагч талын аюулгүй байдал",* "Response – Кибер халдлагын үед хариу үзүүлэх" гэсэн домайнд хамаарах шалгуурууд байна.

Төлөвшлийн Т1 болон Т2 түвшний шалгууруудын хэрэгжилт харьцангуй доогуур байгаа нь байгууллагуудын кибер халдлагад хариу үзүүлэх, эрсдэлийг үнэлэх, гуравдагч талуудтай холбоотой аюулгүй байдалд анхаарах хэрэгцээтэйг харуулж байна.

Судалгааны нэгдсэн үр дүнгээс харахад уул уурхайн салбарын компаниуд КАБ-ын төлөвшлийн Т1 түвшний хяналтуудыг 37%-иар, Т2 түвшний хяналтуудыг 28%-иар хэрэгжүүлж байна. Энэ нь уул уурхайн салбар нь КАБ-ын Т1 түвшин болон түүнээс дээших түвшнүүдтэй холбоотой бодлого, дүрэм, үйл явцыг хэрэгжүүлж эхлэх суурриа бүрдүүлж эхэлснийг харуулж байна.

VI. Санал

Уул уурхайн салбарын компаниудын КАБ-ын төлөвшлийн төвшин одоогоор дунд зэргийн бөгөөд төслийн хэрэгжилт хангалттай сайн биш байна. Гэхдээ компаниуд өөрсдийн КАБ-ын төлөвшлийг сайжруулах, эрсдэлд хариу арга хэмжээ авах, хамгаалалтын механизмуудыг боловсронгуй болгох боломжтой. Энэ нь салбарын мэдээллийн аюулгүй байдлыг нэмэгдүүлэхэд чухал үүрэг гүйцэтгэх болно.

- Салбарын одоогийн КАБ-ын төлөвшлийн түвшинг ахиулахын тулд компаниуд бүх 10 домайны хяналтууд буюу шалгуурууд жигд анхаарч, олон улсын сайн туршлагад тулгуурлсан удирдлагын хөтөлбөрүүд болон үйл явцын оновчтой зохион байгуулалтыг хэрэгжүүлэх хэрэгтэй.
- Гуравдагч талуудтай холбоотой аюулгүй байдлын хяналтыг сайжруулах, мөн тэдгээртэй хийж буй гэрээ, хамтын ажиллагааныхаа аюулгүй байдлын стандартыг тодорхойлох хэрэгтэй. Гуравдагч талуудын үйл ажиллагааг хянах, үнэлэх, хариу арга хэмжээ авах үйл явцыг илүү боловсронгуй болгох;
- Байгууллагын мэдээлэл болон системд хандах эрхийг удирдах, хянах, шалгах үйл явцыг илүү сайн боловсронгуй болгож, эрхийн хяналтын системийг тогтмол шалгах хэрэгтэй. Эдгээр үйл явц нь мэдээллийн аюулгүй байдлын чухал хэсэг юм.
- Системийн хандалтын эрхийг сайжруулах:
- Байгууллагын мэдээлэл болон системд хандах эрхийг удирдах, хянах, шалгах үйл явцыг илүү сайн боловсронгуй болгож, эрхийн хяналтын системийг тогтмол шалгах хэрэгтэй. Эдгээр үйл явц нь мэдээллийн аюулгүй байдлын чухал хэсэг юм.
- Эрсдэл болон хариу арга хэмжээний аргачлалыг бодитойгоор сайжруулах.
- Кибер халдлагаас урьдчилан сэргийлэх, сэргийлэх арга хэмжээг сайжруулж, тогтмол шинэчлэх, төгөлдөршүүлэх, халдлага болсон тохиолдолд хурдан, үр дүнтэй хариу үзүүлэх үйл ажиллагааг боловсронгуй болгох хэрэгтэй. Хариу арга хэмжээ авах үйл ажиллагааг компанийн нийт ажилчдын ойлголт болон сургалтаар дэмжих шаардлагатай.
- Нөхцөл байдлын удирдлагын хөтөлбөрүүдийг сайжруулж, салбарын удирдлага, стратегийн төлөвлөлтөд кибер аюулгүй байдлын хөтөлбөрийг тодорхойлон, системийн аюулгүй байдлын төлөвшлийг тогтмол хянах хэрэгтэй.

НОМ ЗҮЙ

[1] Mitchell P. & Soh C. (2020). Cybersecurity In Energy and resources. Retrieved from EY: https://www.ey.com/en_gl/industries/energy-resources/mining-metals-cybersecurity

[2] PurpleSec. (2024). The Ultimate List Of Cybersecurity Stats Data, & Trends. Retrieved from PurpleSec: https://purplesec.us/resources/cybersecurity-statistics/

[3] КХЗТУТ. (2024, 09 18). Монгол Улс Дэлхийн кибер аюулгүй байдлын индекст ахиц гаргалаа. Retrieved from Кибер халдлага зөрчилтэй тэмцэх Үндэсний төв (National CSIRT): https://ncsirt.gov.mn/a/24

[4] IBM. (2022). What is a cyberattack.

[5] Кибер аюулгүй байдлын тухай хууль. (2021, 12 17). *Эрх зүйн акт.* Retrieved from Эрх зүйн мэдээллийн нэгдсэн систем : https://legalinfo.mn/mn/detail?lawId=16390365491061

[6] Roberts, S. (2023). History of Cyber Security: A Brief Explanation .

[7] Kovrr. (2024, 1 25). *Cybersecurity Maturity Model Implementation - A How-To Get Started Guide.* Retrieved from Kovrr: https://www.kovrr.com/blog-post/cybersecurity-maturity-model-implementation---a-how-to-get-started-guide

[8] *Cybersecurity Capability Maturity Model (C2M2) Version 2.0.* . (2020). Retrieved from U.S. Department of Energy.: https://www.energy.gov/cybersecurity-capability-maturity-model-c2m2

[9] *Framework for Improving Critical Infrastructure Cybersecurity: Version 1.1. NIST Cybersecurity Framework.* (2018). Retrieved from National Institute of Standards and Technology (NIST): https://www.nist.gov/cyberframework

[10] *Cybersecurity Maturity Model Certification (CMMC) Overview.* . (2020). Retrieved from U.S. Department of Defense (DoD): https://www.acq.osd.mil/cmmc/

[11] https://forms.office.com/Pages/DesignPageV2.aspx?prevorigin=shell&origin=NeoPortalPage&subpage=design&id=YQbO-EhNyUKqIBtnkat2_6vevvE4sGVFl0HwIfbpKcBUQVhWOU04N045QVc5SDI2SkhQMEkzOFhQVC4u&analysis=true

AI ашигласан чихрийн шижингийн эрсдэлийг урьдчилан таамаглах, сэргийлэх ухаалаг загвар

Э.Түвшинжаргал[1], Б.Золзаяа[2]

Монгол улс, Улаанбаатар, Шинжлэх Ухаан Технологийн Их Сургууль,
Мэдээлэл, Холбооны Технологийн Сургууль, Мэдээллийн Технологийн салбар

tikaking0@gmail.com, b.zolzaya@must.edu.mn

Хураангуй

Чихрийн шижин нь дэлхий даяар тархалт ихтэй, эдийн засаг болон нийгэм, эрүүл мэндийн салбарт томоохон дарамт үүсгэдэг архаг өвчин юм. Мэдээллийн технологийн дэвшил, ялангуяа хиймэл оюун ухаан (AI) болон машин сургалтын (ML) аргуудыг эрүүл мэндийн салбарт ашиглах боломж нь чихрийн шижингийн эрсдэлийг эрт илрүүлж, өвчний хүндрэлээс урьдчилан сэргийлэхэд чухал ач холбогдолтой болж байна. Энэхүү өгүүлэлд бид AI-д суурилсан машин сургалтын шинжилгээний арга зүйг ашиглан чихрийн шижингийн эрсдэлийг урьдчилан таамаглах, зөвлөмж боловсруулах загварыг танилцуулав. Судалгаанд олон төрлийн өгөгдөл (жишээ нь, эрүүл мэндийн үндсэн үзүүлэлтүүд, амьдралын хэв маяг зэрэг) ашиглан ангиллын хэд хэдэн алгоритмыг (Random Forest, Logistic Regression, Nave Bayes гэх мэт) харьцуулсан. Үр дүнгээс харахад ML алгоритмууд нь чихрийн шижингийн эрсдэлийг 75–90 хувийн нарийвчлалтай урьдчилан таамаглах боломжтой байгааг тогтоов. Энэ нь цаашид эмч болон эрүүл мэндийн байгууллагуудад эрсдэлийг эрт анзаарах, урьдчилан сэргийлэх үйл ажиллагааг сайжруулахад томоохон дэмжлэг үзүүлнэ.

Түлхүүр үг—Чихрийн шижин, Хиймэл оюун ухаан, Машин сургалт, Эрсдэлийн үнэлгээ, Урьдчилан сэргийлэх

I. УДИРТГАЛ

Чихрийн шижин нь дэлхий дахинд түгээмэл тархаж, эрүүл мэндийн салбар болон эдийн засаг, нийгмийн салбаруудад ихээхэн ачаалал үүсгэсээр буй архаг өвчин юм. Дэлхийн эрүүл мэндийн байгууллагын (ДЭМБ) тайланд дурдсанаар 2030 он гэхэд чихрийн шижинтэй хүмүүсийн тоо улам өсөх төлөвтэй байгааг онцолжээ [1]. Энэхүү өсөлтийн шалтгаан нь хүн амын насжилт, амьдралын хэв маяг болон хооллолт, биеийн идэвхтэй хөдөлгөөн дутагдах, таргалалт, стресс зэрэг олон хүчин зүйлтэй нягт холбоотой. Чихрийн шижин нь эмчилгээг орхих, цусны сахарын хэмжээ тогтмол хянагдахгүй тохиолдолд зүрх судас, бөөр, мэдрэлийн тогтолцоо зэрэг олон эрхтэнд ноцтой гэмтэл учруулж, өвчтөний амь насны чанарыг бууруулдаг онцлогтой.

Сүүлийн жилүүдэд дэлхийн эрүүл мэндийн салбарт мэдээллийн технологи, дижитал шийдлүүд эрчимтэй нэвтэрч байгаатай холбоотойгоор эмнэлзүйн өгөгдөл, хувийн эрүүл мэндийн бичиглэл (personal health record) зэрэг эх сурвалжаас гадна ухаалаг цаг, ухаалаг утас, дасгал хөдөлгөөн хэмжигч төхөөрөмжөөс маш их хэмжээний өгөгдөл (Big Data) цуглуулах боломж бий болсон. Энэ нь зөвхөн өвчтөний шинжилгээний дүн төдийгүй тэдний өдөр тутмын идэвхтэй хөдөлгөөн, эрүүл мэндийн үзүүлэлт, хооллолтын зуршил, унтлагын чанар зэргийг бүхэлд нь багтаасан өргөн хүрээний мэдээлэл юм. Ийм том өгөгдлийг тогтмол цуглуулж, боловсруулснаар:

1. Эрсдэлийг эрт илрүүлэх
2. Урьдчилан сэргийлэх арга хэмжээг уян хатан, хувь хүнд тохируулан хэрэгжүүлэх
3. Эмчилгээний хугацааг богиносгох, үр дүнг сайжруулах боломжтой.

Эдгээр нь эцсийн эцэст чихрийн шижин өвчний тохиолдлыг бууруулах, хүндрэлээс сэргийлэх, урт хугацаандаа улс орны эрүүл мэндийн салбарын төсөв, эдийн засгийн ачааллыг багасгахад чухал хувь нэмэр оруулна. Тиймээс ч “Чихрийн шижинтэй тэмцэх, хүндрэлээс сэргийлэхэд AI, ML-ийн шийдэл нэвтрүүлэх нь хамгийн үр дүнтэй арга замуудын нэг” гэж үзэх болсон нь зүй ёсны хэрэг юм.

Ингэснээр чихрийн шижин өвчнийг эрт шатанд нь илрүүлж, зогсоон, улмаар хүндрэлээс сэргийлснээр хүний амь нас, эдийн засгийн зардлыг хэмнэх, нийгэмд эерэг үр дагавар авчрах бүрэн боломжтой хэмээн үзэж байна.

II. ИЖИЛ ТӨСТЭЙ АЖЛУУДЫН СУДАЛГАА

Сүүлийн жилүүдэд чихрийн шижингийн өвчлөлийг эрт шатанд нь илрүүлэх, өвчний хүндрэлээс урьдчилан сэргийлэх зорилгоор машин сургалтын (ML) аргуудад тулгуурласан олон судалгаа хийгдэж байна. Эдгээр судалгаануудын ихэнх нь эмнэлзүйн томоохон өгөгдөл (электрон өвчтөний түүх, лабораторийн шинжилгээ, удамшил) болон хувь хүний амьдралын хэв маяг (хооллолт, дасгал хөдөлгөөн, унтлагын горим г.м.)-ийн мэдээллийг ашиглах замаар өвчний эрсдэлийг таамаглахыг зорьдог. Тэдний дунд дараах чиглэлийн судалгааны ажлууд түлхүү хийгдсэн байна:

2.1 Уламжлалт статистик ба классик машин сургалтын аргууд

Logistic Regression, Decision Tree зэрэг классик аргууд нь *Pima Indians Diabetes Dataset (*PID*)*-ийг ашигласан туршилтуудад 75–80% орчим нарийвчлалд хүрэх чадвартайг нотолсон. Тухайлбал, Pima Indians Dataset дээр Logistic Regression, Decision Tree, Naive Bayes, SVM зэрэг хэд хэдэн арга туршиж, 75–90% орчим нарийвчлал гаргасныг илтгэжээ [2]. Мөн өгөгдлийн хэмжээ бага, шинж чанарууд цөөн үед классик аргууд харьцангуй тогтвортой үр дүн үзүүлдэг болохыг онцолсон.

2.2 Ensemble болон гүнзгий суралцах аргууд

Random Forest, **XGBoost** зэрэг Decision Tree суурьтай ансамбль загварууд эмнэлзүйн өгөгдлийг ангилах, таамаглахад тогтвортой өндөр гүйцэтгэл үзүүлдэг нь дараах ажлуудаас харагддаг.

XGBoost нь Decision Tree суурьтай ансамбль аргад тогтмол өндөр үр дүн гаргаж байгааг үзүүлсэн бөгөөд чихрийн шижин, зүрх судасны өвчин зэрэг том хэмжээний эмнэлзүйн өгөгдөл дээр 85–90% хүрэх гүйцэтгэлтэй болохыг [3] судалгаанд баталсан.

[6] судалгааны ажил нь Deep Learning (CNN, RNN/LSTM)-д тулгуурлан том эмнэлгүүдийн цахим бүртгэл дээр 85–95% нарийвчлалтай урьдчилан таамаглах загварууд боловсруулсан. Гэвч өгөгдлийн хэмжээ, чанар болон сургах технологийн хүчин чадал өндөр байхыг шаардана.

Энэ чиглэлийн судалгаануудын дүгнэлтээр гүнзгий суралцах аргууд нь өгөгдөл хангалттай их, чанартай нөхцөлд маш өндөр нарийвчлал үзүүлж болох ч, өгөгдөл багатай, тэнцвэргүй нөхцөлд тогтвортой бус байж болзошгүйг тодотгосон байдаг.

2.3 Амьдралын хэв маяг, удамшлыг нарийн тусгасан судалгаанууд

Орчин үед ухаалаг цаг, ухаалаг утас зэрэг төхөөрөмжөөс биеийн хөдөлгөөн, зүрхний цохилт, хоолны зуршил зэрэг илүү баялаг өгөгдөл цуглуулах боломж бүрдсэн нь чихрийн шижинг зөвхөн лабораторийн шинжилгээгээр бус, амьдралын хэв маягтай нь уялдуулж таамаглах судалгааг дэмжиж байна. Улмаар хувь хүний удамшил, биеийн жингийн индекс (BMI), стресст өртөх байдал зэрэг олон хүчин зүйлийг нэмж тооцсоноор урьдчилан таамаглах гүйцэтгэл илүү сайжирч байгааг тодотгосон судалгаа олон [4] гарчээ.

2.4 Монголын нөхцөл дэх судалгааны тойм

Чихрийн шижин өвчин (ЧШӨ) нь Монгол Улсад нас баралтын 80 гаруй хувийг эзэлдэг дөрвөн гол халдварт бус өвчний (зүрх судасны өвчин, хорт хавдар, уушгины архаг бөглөрөлт, чихрийн шижин) нэг бөгөөд дотоод шүүрэл, хоол боловсруулах болон бодисын солилцооны эмгэгийн 44.1 хувийг дангаараа бүрдүүлдэг. STEP-IV судалгааны үр дүнд сүүлийн 13 жилийн хугацаанд чихрийн шижингийн тархалт 2.7 дахин нэмэгдэж, 1999 онд 3.1 хувь байсан үзүүлэлт 2019 онд 8.3 хувьд хүрсэн нь тогтоогджээ [5]. Мөн 40-өөс дээш насны ЧШӨ-өөр өвчилсөн гурван хүн тутмын нэг нь нүдний хүндрэл үүсэх эрсдэлтэй бөгөөд зүрх судасны өвчин, бөөрний дутагдал, хөлийн архаг хүндрэл зэрэг нь энэ өвчний шалтгаант нас баралтын гол шалтгаанууд болдог [6].

Бид энэхүү судалгааны ажлын хүрээнд Early Stage Diabetes Risk Prediction болон ухаалаг цагны мэдээлэлд үндэслэн машин сургалтын аргуудыг ашиглан загварыг туршиж хувь хүнд тохирсон зөвлөгөө мэдээлэл өгөх, чихрийн шижингийн урьдчилан таамаглах загварыг хөгжүүлэхэд зорилоо.

III. АШИГЛАСАН АРГА

3.1 Өгөгдлийн багц

- Эрт үеийн чихрийн шижингийн эрсдэлийг таамаглах өгөгдлийн сан нь Бангладеш улсын Силхет дэх Силхетийн Чихрийн Шижингийн Эмнэлэгт шинээр оношлогдсон эсвэл чихрийн шижинтэй болох магадлалтай өвчтөнүүдийн шинж тэмдэг, тэмдгүүдийн мэдээллийг агуулсан болно [7].

  Өгөгдлийн сангийн ерөнхий мэдээлэл:
  Жишээний тоо: 520
  Шинж тэмдгийн тоо: 16
  Онцлогийн төрөл: Категори ба бүхэл тоон
  Шошго: Ангилал (чихрийн шижинтэй эсэх)
  1 – р зурагт үзүүлэв.

Variable Name	Role	Type	Demographic	Description	Units	Missing Values
age	Feature	Integer	Age			no
gender	Feature	Categorical	Gender			no
polyuria	Feature	Binary				no
polydipsia	Feature	Binary				no
sudden_weight_loss	Feature	Binary				no
weakness	Feature	Binary				no
polyphagia	Feature	Binary				no

1 - р зураг. Өгөгдлийн багц

3.2 Өгөгдлийн урьдчилсан боловсруулалт

Өгөгдөл цэвэрлэх

- Өгөгдлийн бүрэн бүтэн байдлыг шалгах: Давхардсан эсвэл логик зөрчилтэй мөрүүдийг илрүүлэх, өгөгдлийн чанарыг хангахын тулд шаардлагатай цэвэрлэгээг хийв.

Өгөгдлийг хуваргах

- Тоон бус утгуудыг хувиргах: Мэдээлэлд агуулагдах хүйс, шинж тэмдгүүд зэрэг категорийн хувьсагчдыг One-hot encoding болон Label encoding аргаар хувиргасан. Жишээлбэл, "Эрэгтэй"/"Эмэгтэй" гэсэн хувьсагчийг (0,1) болгон кодчилж, хоолны төрөл зэргийг олон багананд задалсан.
- Scale/Normalize: Өөр өөр хэмжээний утгатай шинж чанаруудыг (нас, хүйс, ядралт, загтналт гэх мэт…) Min-Max Normalization болон StandardScaler ашиглан стандартчилсан. Энэ нь загварууд, ялангуяа Логистик Регрессд ижил хэмжээсний өгөгдлөөр сургалт хийх боломж олгож, нарийвчлалыг нэмэгдүүлсэн.

3.3 Машин сургалтын алгоритм

- **Random Forest**
  - Олон модыг санамсаргүй байдлаар үүсгэж (Bagging) нийлбэр дунджийг тооцдог.
  - Хэт тохируулгын (overfitting) асуудал бага.
- **Logistic Regression**
  - Оролтын хувьсагчдын шугаман хослол дээр үндэслэн магадлал тооцдог.
  - Шугаман бус шинж чанартай өгөгдөлд хязгаарлагдмал үр дүн үзүүлдэг.
- **Decision Tree**
  - Өгөгдлийг утгын дагуу салбарлан ангилдаг.
  - Хурдан, ойлгоход хялбар боловч хэт тохируулгад өртөмтгий.
- **Support Vector Machine**
  - Ангиллын хооронд хамгийн их зайнд тусгаарлах шугам олохыг зорьдог.
  - Их хэмжээний өгөгдөл дээр ажиллахад удаан.
- **Naive Bayes**
  - Бүх шинж чанарууд хоорондоо хамааралгүй гэсэн таамаглалд тулгуурладаг.
  - Шинж чанарууд хоорондоо уялдаа холбоотой үед оновчгүй үр дүн өгдөг.
- **K-Nearest Neighbors**
  - Хамгийн ойр K хөршийн үнэлгээнд үндэслэн шинэ цэгийг ангилдаг.
  - Их хэмжээний өгөгдөл дээр ажиллахад удаан.
- **Gradient Boosting**
  - Алдааг багасгах зорилгоор дараалан суурилсан жижиг сургагчдыг (weak learners) ашигладаг.
  - Нарийвчлал өндөр боловч тохируулга хийхэд төвөгтэй.

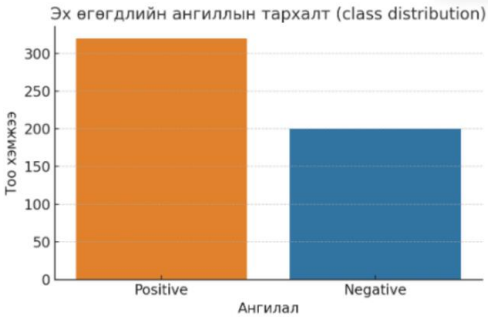
IV. ҮР ДҮН

Судалгааны ажлын хүрээнд 2 туршилт хийсэн ба туршилтыг дараах үндсэн үе шаттайгаар гүйцэтгэлээ. 2 - р зурагт үзүүлэв.

2 - р зураг: Туршилтуудын схем

4.1 Өгөгдөл бэлтгэх

Өгөгдлийн багцын 520 мөр өгөгдөл дээр эхлээд 70% сургалтын, 30 хувь тестийн өгөгдөл болгон хуваваж машин сургалтын алгоритмууд дээр сургав. Мөн өгөгдлийн шинж чанаруудын холбоо хамаарлыг илрүүлж дүн шинжилгээ хийсэн. 3 – р зурагт үзүүлэв.



3 - р зураг: Тэнцвэргүй өгөгдлийн багц

4.2 Туршилт 1

4.2.1 Загвар сургалт ба тохиргоо

Бид дараах байдлаар тохируулгуудыг хийж загвар бүрийн хамгийн сайн тохируулгыг эхний туршилтуудаар олохыг хичээсэн. Энэ үедээ баталгаажуулалтын арга ашигласан. Загвар бүрийн оновчтой тохируулгыг олсны дараа тухайн загваруудыг сургаж, тестийн өгөгдөл дээр үнэлж үзсэн.

Random Forest:

- n_estimators=200, max_depth=20, min_samples_split=5, min_samples_leaf=2
- **Нарийвчлал:** 98.08%

Logistic Regression:

- C=1, penalty='l2', solver='liblinear'
- **Нарийвчлал:** 93.59%

Decision Tree:

- max_depth=15, min_samples_split=5
- **Нарийвчлал:** 94.87%

Support Vector Machine:

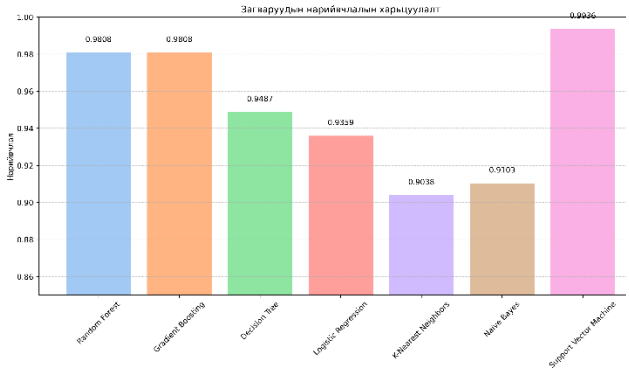
- C=1, kernel='rbf', probability=True
- **Нарийвчлал:** 99.36%

Naïve Bayes:
- **Нарийвчлал:** 91.03%

K-Nearest Neighbors:
- n_neighbors=5
- **Нарийвчлал:** 90.38%

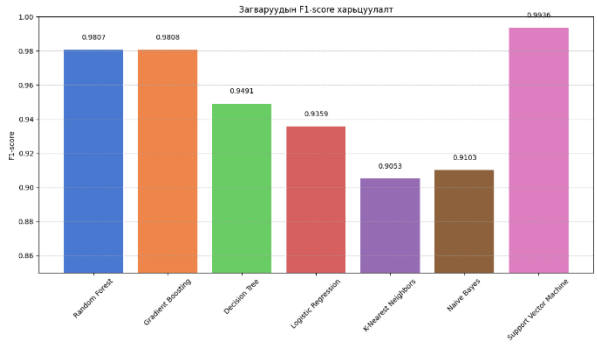
Gradient Boosting
- n_estimators=200, learning_rate=0.1, max_depth=5
- **Нарийвчлал:** 98.08%

Support Vector Machine нь 0.9936 нарийвчлалаар тэргүүлж байгаа бол Random Forest болон Gradient Boosting загварууд 98.08%-тайгаар араас нь орж байна. Харьцуулж үзэхэд KNN, Naive Bayes загварууд харьцангуй доогуур үзүүлэлттэй байна. 4 - р зурагт үзүүлэв.



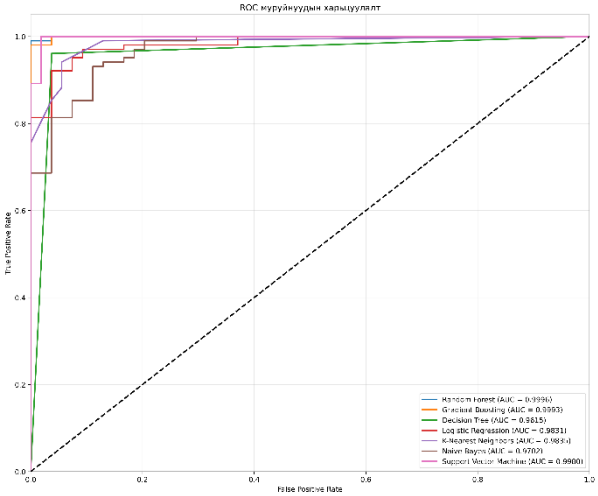
4 - р зураг. Загваруудын нарийвчлалын харьцуулалт

F1-score нь нарийвчлал (precision) болон мэдрэмж (recall)-ийн харьцааг тэнцвэржүүлдэг чухал үзүүлэлт юм. Энд SVM загвар хамгийн өндөр F1-score (0.9936)-той гарсан бөгөөд Random Forest болон Gradient Boosting нь 0.98 орчим үзүүлэлттэй байна. KNN болон Naive Bayes-ийн F1-score доогуур байгаа нь эдгээр загваруудын ангилал тэнцвэргүй үед сул ажиллаж болохыг илтгэнэ. 5-р зурагт үзүүлэв.



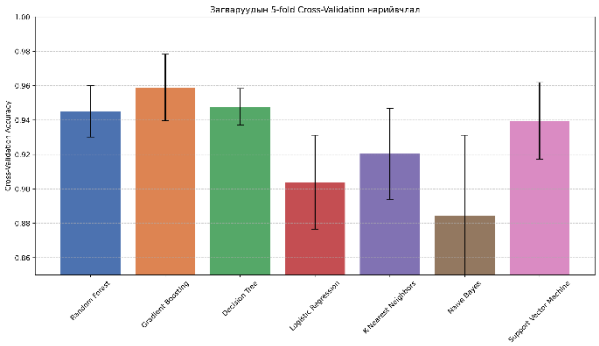
5 - р зураг. Загваруудын f1score харьцуулалт

ROC муруйн доорх талбай буюу AUC нь загварын ангиллын гүйцэтгэлийг хэмжих чухал үзүүлэлт юм. Random Forest (0.9996), Gradient Boosting (0.9993), SVM (0.9980) загварууд AUC өндөр үзүүлэлттэй байгаа нь өвчтэй болон эрүүл хүмүүсийг маш сайн ялгаж байгааг харуулна. Decision Tree болон Naive Bayes харьцангуй бага AUC үзүүлсэн байна. 6 - р зурагт үзүүлэв.



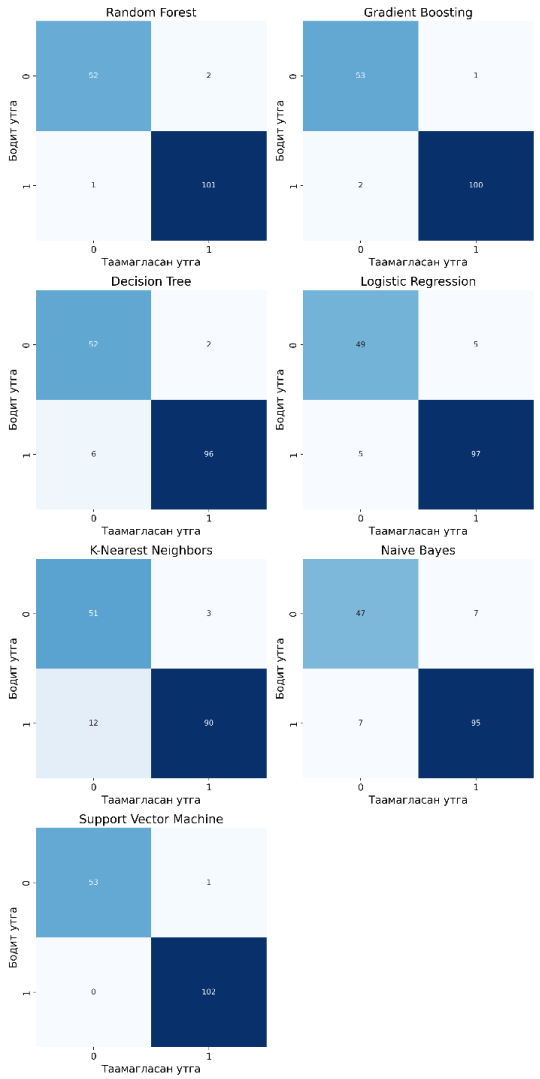
6 - р зураг. Загваруудын ROC харьцуулалт

Cross-validation нь загварын тогтвортой байдлыг шалгахад ашиглагддаг. Энд Gradient Boosting загвар хамгийн тогтвортой (mean ≈ 0.959, стандарт хазайлт бага) байсан бол SVM, Random Forest, Decision Tree зэрэг нь мөн сайн гүйцэтгэл үзүүлжээ. Naive Bayes болон Logistic Regression загварууд харьцангуй тогтворгүй гүйцэтгэлтэй гарсан байна. 7 - р зурагт үзүүлэв.



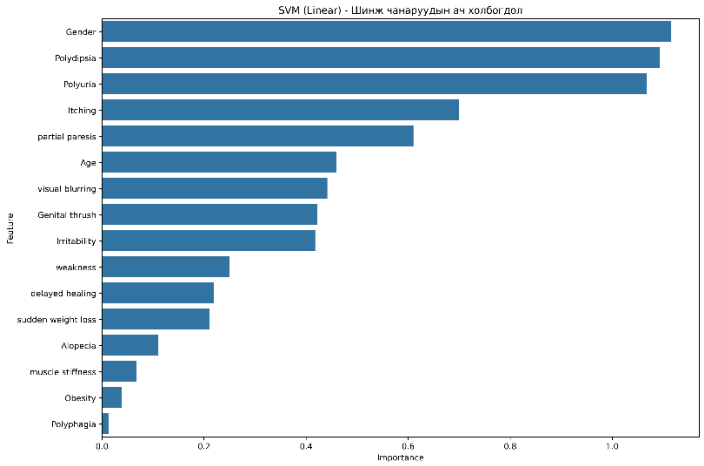
7 - р зураг. Загваруудын cross validation харьцуулалт

Төөрөгдлийн матрицаар загвар тус бүрийн TP, FP, FN, TN утгуудыг харьцуулсан. SVM, Random Forest, Gradient Boosting загварууд бараг төгс ангилалт хийсэн бол KNN болон Naive Bayes нь өндөр FN буюу өвчтэйг оношлоогүй орхих магадлал өндөртэй байна. Энэ нь эмнэлзүйн хувьд анхаарал татахуйц асуудал юм. 8-р зурагт үзүүлэв.



8 - р зураг. Загваруудын confusion matrix харьцуулалт

SVM загвар нь хамгийн өндөр гүйцэтгэлтэй байгаа ба permutation importance дээр үндэслэхэд Gender, Polydipsia (Ихээр цангах), Polyuria (Их шээх) шинжүүд хамгийн их нөлөөтэй гарсан байна. Энэ нь тухайн шинжүүдийг оношилгооны шийдвэр гаргалтад чухал үүрэгтэй болохыг харуулна. Харин Obesity (Таргалалт) болон Polyphagia (Их идэх) зэрэг шинжүүд нөлөөлөл багатай байна. 9-р зурагт үзүүлэв.



9 - р зураг. Оновчлол өндөртэй загварын шинж чанаруудын харьцуулалт

Туршилт 1-ийн үр дүнд суурилан нийт 7 төрлийн машин сургалтын загваруудыг нарийвчлал, F1-score, ROC-AUC, төөрөгдлийн матриц, шинж чанарын ач холбогдол, мөн cross-validation-ийн тогтвортой байдлаар үнэлж харьцуулсан болно.
Хамгийн өндөр гүйцэтгэлтэй 3 загвар нь дараах байдалтай гарсан:

Support Vector Machine (SVM)

- Нарийвчлал: 99.36%
- F1-score: 0.9936
- ROC AUC: 0.9980
- Төөрөгдлийн матриц дээр хамгийн бага FP/FN

Random Forest

- Нарийвчлал: 98.08%
- F1-score: 0.9807
- ROC AUC: 0.9996
- Тайлбарлахад тохиромжтой, шинж чанарын ач холбогдлыг гаргаж чаддаг.

Gradient Boosting

- Нарийвчлал: 98.08%
- F1-score: 0.9808
- ROC AUC: 0.9993
- 5-fold cross-validation-д хамгийн тогтвортой (mean ~0.959, std бага)

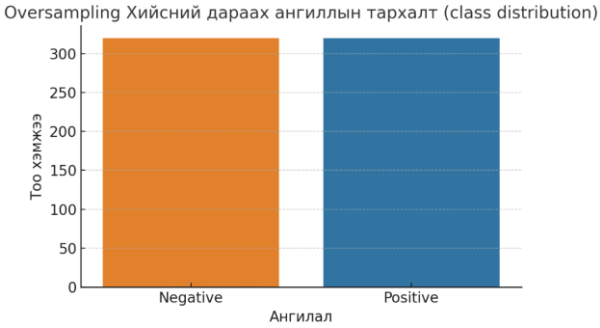
Эдгээр дүгнэлтийн үндсэн дээр дараагийн шатны **Туршилт 2**–т зөвхөн **SVM**, **Random Forest**, болон **Gradient Boosting** загваруудыг сонгон авсан. Энэ туршилт нь илүү **тэнцвэржүүлсэн өгөгдөл** дээр хийгдсэн бөгөөд Туршилт 1-ийн өгөгдөл нь **тэнцвэргүй ангилал (imbalanced classification)** байсан гэж үзсэн болно.

4.3 Туршилт 2

Өмнөх туршилтаар хэдийгээр сайн үр дүн гарсан ч, өгөгдлийн ангиллын харьцаа тэнцвэргүй

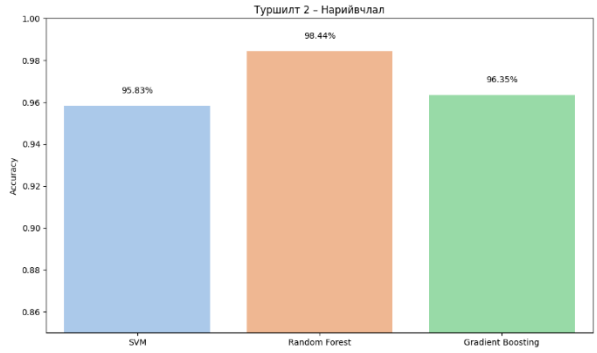
(imbalanced) байсан нь зарим загваруудын гүйцэтгэлийг **overfit** буюу хэт тохируулсан байж болзошгүй гэж дүгнэсэн. Иймээс Туршилт 2-д өгөгдлийг **тэнцвэржүүлэн**, хамгийн өндөр үр дүн үзүүлсэн **Support Vector Machine, Random Forest, Gradient Boosting** гэсэн гурван загвар дээр дахин сургалт хийж, харьцуулалт хийсэн болно.

SMOTE аргаар анги 0 ба анги 1-ийн өгөгдлийг ижил тоотой болгож, сургалтын өгөгдлийн тэнцвэрийг хангаж өгсөн (10 – р зурагт Class 0 = Class 1 ≈ 320).



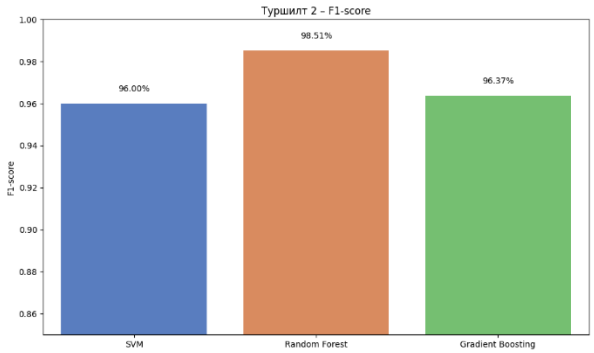
10 – р зураг: SMOTE хийсэн үр дүн

Random Forest загвар 98.44%-ийн хамгийн өндөр нарийвчлалтай гарсан бол Gradient Boosting нь 96.35%, SVM нь 95.83% үзүүлэлттэй гарсныг 11 – р зурагт үзүүлэв.



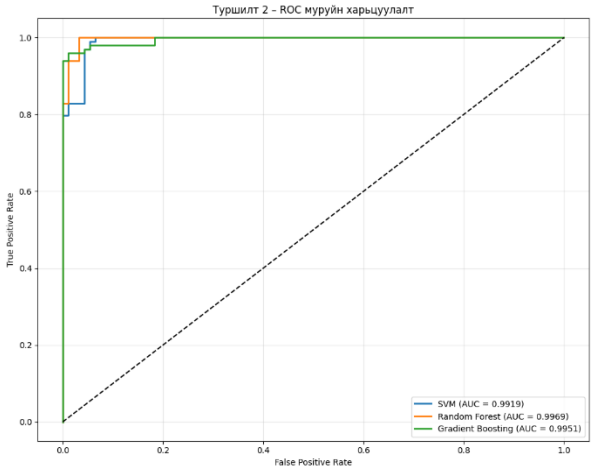
11 - р зураг. Туршилт 2: Загваруудын нарийвчлалын харьцуулалт

F1-score-ийн хувьд Random Forest дахин хамгийн өндөр буюу **98.51%**-тай гарсан нь **нарийвчлал ба мэдрэмжийн баланс маш сайн** байсныг илтгэнэ. Gradient Boosting, SVM загварууд F1-score 96 орчим хувьтай байна. 12 - р зурагт үзүүлэв.



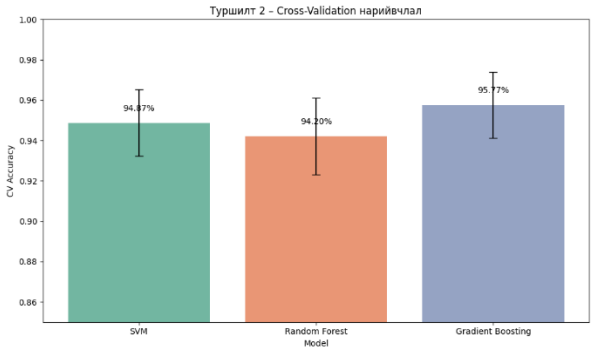
12 - р зураг. Туршилт 2: Загваруудын f1score харьцуулалт

ROC муруйн харьцуулалтаар эдгээр нь бүгд **маш сайн ангилалт хийж** байгааг илэрхийлнэ. 13 - р зурагт үзүүлэв.



13 - р зураг. Загваруудын ROC харьцуулалт

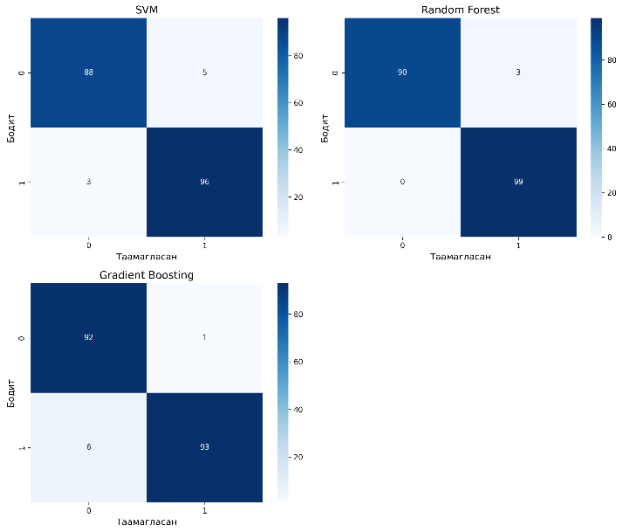
5-fold cross-validation-ийн үр дүнгээс харахад **Gradient Boosting (95.77%)** загвар хамгийн тогтвортой сайн байсан бол SVM (94.87%) ба Random Forest (94.20%) бага зэрэг хэлбэлзэлтэй байна. 14 - р зурагт үзүүлэв.



14 - р зураг. Загваруудын cross validation харьцуулалт

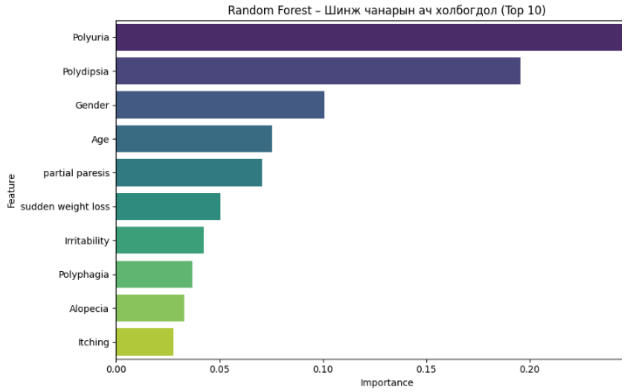
Random Forest загвар бүх ангийг оновчтой ангилж **False Negative = 0** гаргасан. SVM ба Gradient Boosting загварууд мөн өндөр TP (True Positive)

үзүүлэлттэй байсан ч FN бага зэрэг гарсан. 15 - р зурагт үзүүлэв.



15 - р зураг. Загваруудын confusion matrix харьцуулалт

Random forest загвар нь хамгийн өндөр гүйцэтгэлтэй байгаа ба шинж чанарын ач холбогдлоор **Polyuria**, **Polydipsia**, **Gender**, **Age**, **partial paresis** зэрэг нь хамгийн гол нөлөөтэй байсан нь оношлогоонд шийдвэр гаргахад тустай байгаа нь туршилт 1-ийн үр дүнтэй ойролцоо байсан ба үүнийг 16 - р зурагт үзүүлэв.



16 - р зураг. Оновчлол өндөртэй загварын шинж чанаруудын харьцуулалт

Тэнцвэржүүлсэн өгөгдөл дээр суурилсан Туршилт 2-ын үр дүнгээс харахад:

- **Random Forest** загвар нь бүх үзүүлэлтээр (Accuracy, F1-score, Confusion Matrix, AUC, Feature Importance) хамгийн сайн гүйцэтгэл үзүүлсэн.
- **Gradient Boosting** тогтвортой сайн, бага хэлбэлзэлтэй загвар байжээ.
- **SVM** өндөр нарийвчлалтай ч эмнэлзүйн практикт хэрэглэгдэхдээ бусдаас арай сул байж болохоор байна.

4.4 Туршилт 1 ба Туршилт 2-ын харьцуулсан дүгнэлт

1 – р хүснэгтэд үзүүлснээр Туршилт 1 ба Туршилт 2 нь хоёулаа чихрийн шижингийн эрсдлийг машин сургалтын аргаар ангилах зорилготой боловч гол ялгаа нь:

- **Өгөгдлийн баланс** (imbalanced vs. SMOTE)
- **Загварын тогтвортой байдал**
- **Шинж чанарт суурилсан тайлбарлал** гэсэн үндсэн 3 хэмжээс дээр илэрч байна.

Шалгуур үзүүлэлт	Туршилт 1	Туршилт 2
Өгөгдөл	Тэнцвэргүй (imbalanced)	SMOTE тэнцвэржүүлсэн
Загварын тоо	7 (өргөн хүрээтэй)	3 (шилдэг загварууд)
Хамгийн сайн загвар	SVM	Random Forest
Нарийвчлал	99.36% (SVM)	98.44% (Random Forest)
F1-score	99.36% (SVM)	98.51% (Random Forest)
Cross-Validation	Дунд зэрэг ялгаа	Илүү тогтвортой
FN (эмнэлзүйн ач холбогдол)	Зарим загвар өндөр FN	Random Forest: FN = 0
AUC	0.9980 (SVM)	0.9969 (RF), 09951 (GB)

1 - р хүснэгт. Туршилт 1 ба Туршилт 2 харьцуулалт

4.4.1 Шинжилгээ

1. **Өгөгдлийн балансын нөлөө** Туршилт 1-д тэнцвэргүй өгөгдөл ашигласнаас болж зарим загвар (жишээ нь SVM) онцгой өндөр нарийвчлал, F1-score үзүүлсэн ч cross-validation ба төөрөгдлийнматриц дээр тухайн гүйцэтгэл тогтвортой биш байсан.
2. **SMOTE-ийн ач холбогдол** Туршилт 2-д SMOTE ашиглан сургалтын өгөгдлийг баланстай болгосноор загваруудын гүйцэтгэл арай тогтвортой, найдвартай, эмнэлзүйн нөхцөлд ашиглах боломжтой болсон.
3. **Төөрөгдлийн матрицын онцлог** Туршилт 2-ын Random Forest загвар аль ч ангид FN = 0 гаргасан нь оношлалт байхгүйг харуулж, оношилгоонд өндөр

итгэлтэйгээр ашиглах боломжоороо бусдаас ялгарч байсан.

4. **Загвар сонголтын зарчим** Туршилт 2-д зөвхөн шилдэг загваруудаар (SVM, RF, GB) дахин үнэлгээ хийсэн нь илүү нарийвчилсан шинжилгээ хийх боломж олгосон. Энэ нь туршилт 1-ийн үр дүнг баталгаажуулсан алхам болсон.

4.4.2 Шинж чанарын харьцуулалт ба загвар сонголт

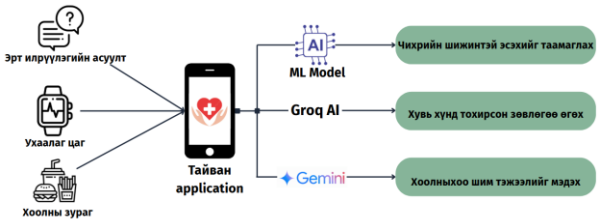
- Туршилт 1-д шинж чанарын ач холбогдол дээр SVM-ийн linear хувилбарыг ашигласан бөгөөд **Gender**, **Polydipsia**, **Polyuria** зэрэг шинжүүд гол нөлөөтэй байв.
- Туршилт 2-д Random Forest загвар хамгийн сайн гүйцэтгэлтэй байсан бөгөөд түүний feature importance анализын үр дүнд дараах өндөр хамааралтай 5 шинж чанар илэрсэн, 2 – р хүснэгтээр харуулав.

Шинж чанар	Ач холбогдол (RF)	Тайлбар
Polyuria	0.25	Шээх тоо олшрох
Polydipsia	0.20	Ус их уух
Gender	0.12	Хүйс
Age	0.10	Нас
Partial paresis	0.09	Булчингийн сульдалт

2 – р хүснэгт: Random Forest загварын шинж чанарын холбоо хараамал

Дээрх сонгогдсон загварыг ашиглан **ухаалаг платформ** хөгжүүлсэн ба энэ нь хүн бүрийн хувийн эрүүл мэндийн өгөгдөлд тулгуурлан чихрийн шижингийн эрт үеийг илрүүлэх, цаашлаад хувь хүнд тохирсон эрүүл мэндийн **зөвлөмжийг санал болгодог** систем юм.

Тус систем нь хэрэглэгчийн хувийн шинж чанарт үндэслэн чихрийн шижингийн эрсдэлийг Random Forest алгоритмоор өндөр нарийвчлалтайгаар таамаглана. Хэрэглэгчийн хоолны зуршлыг Gemini AI ашиглан боловсруулж, хоолны шим тэжээлийн найрлагыг тодорхойлох боломжтой. Хэрэв хэрэглэгч чихрийн шижинтэй бол ухаалаг төхөөрөмжтэй (жишээлбэл, ухаалаг цаг) холбогдон, цусан дахь сахарын хэмжээ, зүрхний хэмнэл, хөдөлгөөний идэвх зэрэг үзүүлэлтүүдийг хянаж, Groq LLM-ийн тусламжтайгаар эрүүл мэндийн зөвлөмжийг автоматаар хүргэнэ. Системийн схемийг 17 - р зурагт үзүүлэв.



17 - р зураг. Ухаалаг платформ ерөнхий схем

V. ДҮГНЭЛТ

Энэхүү судалгаагаар машин сургалтын алгоритмуудыг ашиглан чихрийн шижингийн эрсдлийг урьдчилан таамаглах боломжийг баталгаажуулсан бөгөөд **Random Forest** загвар нь **өндөр гүйцэтгэлтэй, оношилгоонд алдаа гаргахгүй, шинж чанарт суурилан тайлбарлах боломжтой** хамгийн оновчтой загвар болж тодорлоо.

Тус загварын шинжилгээнээс харахад **Polyuria**, **Polydipsia**, **Gender**, **Age**, **partial paresis** зэрэг нь оношлох шийдвэр гаргалтад хамгийн их нөлөөтэй шинжүүд бөгөөд энэ нь загварын тайлбарлагдах чадварыг нэмэгдүүлж, эмнэлзүйн практикт итгэлтэйгээр ашиглах нөхцөлийг бүрдүүлж байна.

Цаашид хийх ажлууд

1. Машин сургалтын загваруудыг эмнэлгийн практикт туршилтаар нэвтрүүлж, бодит орчинд баталгаажуулах

2. Эмнэлгийн мэдээллийн санг өргөтгөн, өөр өөр нас, хүйс, үндэстний өгөгдлөөр загваруудыг сургах замаар нарийвчлалыг улам сайжруулах

3. Хосолсон зөвлөмжийн систем: Загвараас гарах эрсдэлийн үнэлгээг эмч болон өвчтөнд ээлж дараатай, ойлгомжтойгоор дамжуулж, зөв хооллолт, дасгал хөдөлгөөн, эмнэлзүйн хяналтын давтамж, эмчилгээний оновчтой стратегийг нэгтгэсэн интерфейс, апп хөгжүүлэх.

VI. АШИГЛАСАН МАТЕРИАЛ, НОМ ЗҮЙ

[1] World Health Organization. Global report on diabetes. WHO, 2016.
[2] Chang, Victor, et al. (2022) "Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms." *Neural Computing and Applications*
[3] Chen, T. & Guestrin, C. (2016). “XGBoost: A Scalable Tree Boosting System,” Proceedings of the 22nd ACM SIGKDD.
[4] Rajkomar, A. et al. (2018). “Scalable and Accurate Deep Learning with Electronic Health Records,” npj Digital Medicine.
[5] Fourth National STEPS Survey on the Prevalence of Non-Communicable Disease and Injury Risk Factors – 2019. National Centre for Public Health (Mongolia)
[6] “Diabetes: Key Facts.” World Health Organization (WHO) Fact Sheet, updated 2023.
[7] https://archive.ics.uci.edu/dataset/529/early+stage+diabetes+risk+prediction+dataset

Морин хуурын нумны барилын онцлог, цахилгаан дохионы үзүүлэлтүүд

Цагаан Баатархүү
Мэдээллийн технологийн салбар
МХТС, ШУТИС
Монгол, Улаанбаатар
baatarkhuu@must.edu.mn

Оюунбилэг Бат-Энх
Мэдээллийн технологийн салбар
МХТС, ШУТИС
Монгол, Улаанбаатар
batenkh@must.edu.mn

Оюунбилэг Баттөгс
Үндэсний урлагийн тэнхим
ХУС, Соёл, Урлагийн Их Сургууль
Монгол, Улаанбаатар
Battugs495@gmail.com

***Хураангуй*— Энэхүү судалгааны ажлын хүрээнд морин хуурын нумны хийц, эд бүрдэл, төрөл зүйл, бүтэц, Өвөр Монгол болон Монгол морин хуурчдын нум барих арга, барилыг урлаг судлал талаас төдийгүй орчин үеийн дууны боловсруулалтын техник, технологи ашиглан судалсан. Дээрх 2 барилын хувьд эгшгүүдийн цахилгаан дохионд судалгаа хийж, үндсэн болон дагалдах өнгөнүүдийн давтамж, дуугаралтын түвшин, энергийн төвлөрөл буюу формантуудыг харьцуулсан. Судалгаанд Соёл, урлагийн их сургууль (СУИС), Морин хуурын чуулга, Монгол улсын консерватори (МУК), Түмэн эх чуулга, Хилийн цэргийн дуу бүжгийн чуулга (ХЦДБЧ), Бүх цэргийн дуу бүжгийн чуулга (БЦДБЧ), Монгол улсын үндэсний урлагийн их театр (МУҮУИТ), Монголын хүүхдийн ордны нийт 146 оюутан, ажилчид хамрагдсан. МУ-ын морин хуурын чуулганд 16 хуурч байгаагаас бүгд Өвөр Монгол нумны барилаар хуурдаж байна. Мэргэжлийн морин хуурчид, их, дээд сургууль, мэргэжлийн байгууллага, морин хуурын сургалтын төвүүдэд уламжлалт Монгол нумны барилын давуу талыг танин мэдүүлж, цаашид судалгаа, шинжилгээ хийж, хойч үедээ өвлүүлэн үлдээх шаардлагатай байна.**

Түлхүүр үг— нумны барил, үндсэн өнгө, дагалдах өнгө, давтамж, хэлбэлзэл, дууны түвшин, формант, энергийн төвлөрөл

I. Удиртгал

Өндөр хөгжилтэй орнууд өөрсдийн үндэсний хөгжмийн талаар судалгаа, шинжилгээ хийж, хойч үедээ өвлүүлэн үлдээх болсон байна. Монгол улсын хувьд морин хуур хөгжмийг урлаг судлал талаас нь судалгаа шинжилгээ хийсэн ажлууд их байдаг төдийгүй техникийн талаас нь судлан шинжлэх ажил эрчимтэй хийгдэж байна. Энэ нь морин хуурын дууны өнгө, хэлбэр хийц, соёлыг хойч үедээ өвлүүлэн үлдээх зорилготой юм.

Морин хуурын сургалт хөгжлийнхөө явцад эртнээс уламжлан ирсэн шавь сургалтын аргыг хадгалсаар ирсэн бөгөөд цаг хугацаа өнгөрөх тусам морин хуурын язгуур сургалтын аргыг өвлөн тээгчид нь цөөрч, танхимын сургалт хэрэгжүүлэх болсноор сургалтын арга барил, урын сан, хөгжимдөх арга барил улам бүр өөрчлөгдөж байна. Орчин цагт морин хуурын сургалтыг ихэвчлэн Монгол улсын морин хуурч, багш Г.Жамъяаны боловсруулан гаргасан тогтолцоот сургалтын аргаар явуулж байгаа. Өнөөгийн морин хуурын сургалтын чухал хөшүүрэг болж буй сурах бичиг, бусад хэрэглэгдэхүүн, судалгаан дээр тулгуурлан Монгол (нумны) барилаар хуурдах санааг дэвшүүлж буй нь оюутнуудын суралцахуйг дэмжих хэрэглэгдэхүүний нэг болох төдийгүй, морин хуур хөгжимдөх чадварыг нэмэгдүүлэх, цаашлаад өөрсдийн үндэсний хөгжмийн уламжлал дуугаралт, тоглолтыг өвлүүлэн үлдээх боломжийг бий болгоно.

II. Онол, судалгааны арга зүй

A. Утсан хөгжмийн зэмсгийн үзүүлэлтүүд

Утсан хөгжмийн зэмсгийн сунасан утасны долгион дамжуулах хурд нь утасны нэгж урт дахь масс болон суналтаас (даралтын хүч) хамаарна. Долгионы (дууны) хурдыг доорх томъёогоор тооцоолно [1,5].

$$v = \sqrt{\frac{T}{m/L}} \tag{1}$$

Үүнд: T –утасны суналт; m – утасны масс; L – утасны урт.

Сунасан утаснууд долгион үүсгэх ба үүнийг резонанс буюу цуурай гэнэ. Хэлбэлзэж буй утас давтамж үүсгэх ба хамгийн бага давтамжийг үндсэн давтамж (үндсэн өнгө) гэж нэрлэнэ. Давтамжийг доорх томъёонуудын тусламжтай тооцоолж болно.

$$f_1 = \frac{\sqrt{\frac{T}{m/L}}}{2L} \tag{2}$$

Мөн давтамжийг 1 секундэд давтагдаж буй хэлбэлзлийн үеийн тоогоор тодорхойлдог.

$$f_1 = 1/T \tag{3}$$

Хэлбэлзэж буй утас нь үндсэн өнгө буюу давтамжаас гадна олон тооны дагалдах өнгүүдийг (гармоникуудыг) бий болгоно. Сунасан утасны үндсэн хэлбэлзлийн хэлбэрийн үед долгионы урт нь утасны уртаас 2 дахин их байвал үүнийг үндсэн давтамж гэнэ. Дагалдах өнгө (Overtones) гэдэг нь үндсэн давтамжаас дээгүүрх гарч буй бүх давтамжууд юм [2]. Утасны хэлбэлзэл нь хөдөлгөөнийг (motion) бий болгодог. Энэ нь хэлбэлзлийн ялгаатай, өөр өөр давтамжууд буюу обертонуудыг бий болгодог юм.

Формант гэдэг нь дууны долгионы тодорхой давтамжийн эргэн тойрон дахь дуугаралтын (акустикийн) энергийн төвлөрөл юм [3].

Долгионы нэг төрөл нь дууны долгион юм. Дууны түвшинг (sound level) хэмжихдээ децибел буюу дБ нэгжээр хэмжинэ. Дууны түвшин нь дууны чадал, даралт, хүчдэл зэрэг олон үзүүлэлтүүдийн харьцаа ба энэ нь дууны хүчтэй шууд хамааралтай.

$$P(\text{дБ}) = 10log_{10}\frac{P_2}{P_1} \tag{4}$$

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Энэ нь дууны чанга сулыг илэрхийлдэг үзүүлэлт юм. Долгионыг ашиглаж энергийг нэг газраас нөгөө газарт дамжуулж болно [5].

B. Уламжлалт морин хуур хөгжимдөх онцлог

Уламжлалт морин хуур нь монголын язгуур хөгжмийн нэг хэсэг болохын хувьд уламжлалт морин хуурын өв залгамжлал нь Монголын үндэсний хөгжмийн өв залгамжлалын зөвхөн нэг тал юм.

Эгшиглүүлэх учир: Хуурдахын нэр нь хөрөөдөхдөө бус чухамдаа дуугарахдаа байна. Хөгжмийн эгшиг дуу нь хөгжмийн уран сайхны дүр бүтээх үндсэн материал, утга илэрхийллийн гол хэрэглүүр, хөгжмийн сэтгэлгээний амьд хэл мөн.

Морин хуур нумт хөгжмийн төрөлд хамаарах бөгөөд дуурьсган дуугаргах хоёр үндсэн арга байдаг. Үүнд:

1. Хөгжмийн чавхдас, нум хоёрын үйлчлэл
2. Хоёр гарын зохицлогоо.

Энэ нь хөгжмийн чавхдасыг зөвхөн дуугаргахын анхны алхам болохоос биш эгшиглүүлэхийн бүрэн утгыг хараахан хариулж чадахгүй юм. Хөгжмийн эгшиг дууг жинхэнэ ёсоор нь, уран сайхны дээд зэргийн чанар байдалтайгаар эгшиглүүлэхэд олон хүчин зүйл нөлөөлдөг. Үүнд: зөвхөн хөгжмийн утас нум хоёрыг нийцүүлэх дасгалд утас нум хоёрын хөндлөн огтлол (нумны тэгш налуу), нумны дарах хүч (чанга, сул) нумны тасралтгүй үргэлжлэх хурд хугацаа, (удаан, түргэн) нумны өнгө илчлэл (уянгалаг зөөлөн, огцом, хурц, бүдэг, хахир гэх мэт), хуурын эгшиглэлтийн төв зэрэг багтана [8].

Суудал, байрлал: Хуурдах явцад хөгжимчнөөс хөгжимтэйгөө харьцаж байгаа бүхий л үйл ажиллагаа өөрөөр хэлбэл, биеийн чөлөөт байдал, баруун зүүн гарын барил, түүний гадаад төрх хэлбэр, хөдөлгөөний шилжилт, суудал, байрлал зэргийг ерөнхийд нь тавил гэдэг ойлголтод хамаатуулдаг.

Барил тавил: Морин хуурын тавилыг үндсэн гурван хэсэгт хуваж үзнэ. Үүнд:

1. Суудал, хөгжмийн байрлал
2. Баруун гарын барил
3. Зүүн гарын барил [8].

Баруун гарын барил: Баруун гарын барилд нумны ерөнхий барилд нумны янз бүрийн хөдөлгөөнөөс үүсэн гарах мөр тохой, бугуй байрлал, хөдөлгөөнүүд хамаарагдана.

Нумны барилыг зөв тавих асуудал нь хуурчийн хөгжимдөх ур зүй, эгшиглүүлэх арга зүйтэй салшгүй холбоотой. Өөрөөр хэлбэл баруун гарын уран чадвар бол нумны өнгө илчлэл, эгшиглүүлэхийн үндсэн ур зүй юм.

Морин хуурын нумыг барихдаа нумны бариулыг нумны хялгасны хамт атгах байдлаар 1, 2, 3, 4-р хуруу утсыг доош нь дарж эрхий ба 1-р хуруу нумыг тогтоож барина. Үүний тул нумны хялгас харьцангуйгаар сул бэхлэгдсэн байдаг.

Нумт хөгжимд нумыг татах түлхэх ерөнхий хоёр үйл явц байдаг. Морин хуур хөгжмийн нумыг татах түлхэх үйл ажиллагаанд 5 хуруу бараг адил үүрэгтэй оролцдог. Харин хөгжимдөх байгаа арга барилаас шалтгаалж хуруу тус бүрийн гүйцэтгэх үүрэг харилцан адилгүй байна.

Морин хуурын чавхдас, нумны хялгасны өтгөн нь бараг тэгш харьцаатай байна. Ийм учраас хөгжмийн чавхдасанд нумны дарах (чангаар) хүч өөрөөр хэлбэл, хуурын чавхдасыг нусаар хөндлөн огтлол хийж гаргах эсэргүүцэл их байдаг. Ийм учраас морин хуурын нумыг бүхэлд нь атгах байдлаар тогтоон барьж хөгжимддөгийн учир энэ болой.

Ер нь морин хуурчийн 10 хуруу эгшиг дуунд онц мэдрэмтгий байх ёстой. Хөгжмийн зохиолын амьд хэл-уран сайхны эгшиг дууг үнэн бодитойгоор эгшиглүүлэхэд хуурчийн хурууны мэдрэмтгий чанар нь эгшиг дууг хянах эрхтэн чихнээс дутуугүй зохицуулах үүрэгтэй юм.

Баруун гарын бугуй, мөр тохой хөдөлгөөний байрлал нь хөгжимдөх ур зүй, тоглолтын шаардлага, нумны утга илэрхийллийн аргаас шалтгаалж янз бүр байдаг. Бугуй хөдөлгөөн нь нумны утга хурд хүч, хуурын эгшиглэлтийн төвийг тохируулахад чухлаар нөлөөлнө.

Нумыг татахад суга харьцангуй тэлэгддэж тохой урагшилна. Мөрний булчингийн уян хатан чөлөөтэй байдал нь баруун гарын тавилд улам ч аятай нөхцлийг бүрдүүлнэ [8].

Хурууны бүтэц ба үүрэг: Хуруу, сарвууны артерийн судасны бүтцээс хамааруулан гурван бүс болгосон ангилна

Хуруунуудын үүрэг хоорондоо ялгаатай байдаг. Ихэнх судлаачид эрхий хурууны сарвууны үйл ажиллагаа 40-50%-иар алдагддаг гэж үздэг. Мөн долоовор хуруу 20%, дунд хуруу 15%, ядам хуруу 10%-аар сарвууны үйл ажиллагаа гүйцэтгэдэг гэжээ. Зарим судлаачид дээрх хувиас бага зэрэг ялгаатай гэж үздэг. Эрхий хуруу угийн шивнүүрээр 75%, бусад хуруунууд 45%-аар тухайн хурууны хувьд үйл ажиллагаа нь алдагддаг гэж үздэг [4]. II-IV хурууны үзүүрийн үе нь хөдөлмөрийн чадвараа 45% алдана. Мөн хуруунуудын угийн үеийн мэдрэлийн түвшин гэмтсэн бол 80%-аар үйл ажиллагааны алдагдалтай гэж үздэг. Харин мэдрэлжилт бүрэн алдагдсан бол сарвууны үйл ажиллагаа 50% алдагдсан гэж үзсэн. I ,II, III, IV, V хуруунууд тайрагдаж тасарсан бол 40%, 20%, 10%, уг хуруунууд хэвийн мэдрэлжилт бүрэн алдсан бол 20%, 10%, 5% байна. Сарвууны бүх үйл ажиллагааг 100% гэж үзээд мэдрэхүй, хөдөлгөөнийг тус бүр 50%-тай тэнцүү гэж авч үзсэн [4].

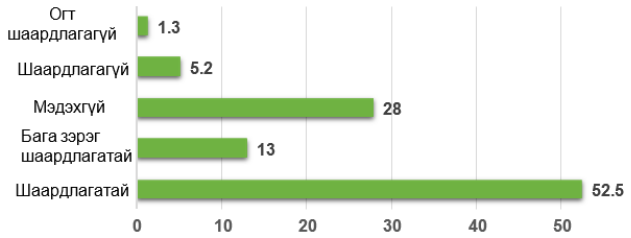
III. Судалгааны Үр дүн

Мэргэжлийн морин хуурчдын хуурдах ур чадварт уламжлалт нумаар хуурдах арга хэрхэн нөлөөлөх онцлогийг турших, үр дүнг тодруулах үүднээс мэргэжлийн урлагийн байгууллагын морин хуурчид болон мэргэжлийн ангийн оюутнуудаас бүрдсэн 146 морин хуурчийг судалгаанд хамруулсан (Хүснэгт 1).

ХҮСНЭГТ 1. МОРИН ХУУРЧДЫН НУМНЫ БАРИЛЫН СУДАЛГАА (2024)

Байгууллагын нэр	Хөгжим чид тоо	Нумны барил	
		Монгол	Өвөр Монгол
Морин хуурын чуулга	16		16
СУИС	30	2	28
МУК	31		31
Түмэн эх чуулга	3		3
ХЦДБЧ	7		7
БЦДБЧ	3		3
МУҮУИТ	8		8
Монголын хүүхдийн ордон	48		48
НИЙТ	146		144

Судалгааны явцад морин хуурын мэргэжлийн ангийн 26 оюутныг хамарсан Хуурын монгол барилын хэрэглэгдэхүүний хэрэгцээг тодруулах зорилго бүхий нээлттэй хичээлийг Ардын жүжигчин Ч.Батсайхан, Өвөр Монголын морин хуурч Чи Бөхөө нартай хамтран 2025 оны 02 сарын 05-ны өдөр СУИС-ийн 300в танхимд зохион байгуулсан. Судалгаанд хамрагдсан 26 оюутнаас морин хуур хөгжимдөж сурахад хуурын монгол барилын хэрэгцээ шаардлагатай байгаа эсэх, сургалтын нэмэлт хэрэглэгдэхүүний хэрэгцээг илрүүлэх зорилгоор асуумж судалгаа авсан (Зураг 1, Зураг 2).



Зураг. 1. Монгол нумны барил сурах хэрэгцээ байгаа эсэх.

Судалгаанд хамрагдсан хуурчдын 52.5% нь Монгол нумны барил сурах шаардлагатай гэж үзсэн. Мөн 26 оюутнаас авсан судалгаанд "Морин хуурын нэмэлт хэрэглэгдэхүүнд оюутнууд концерт, аудио болон видео бичлэг байвал тэдэнд морин хуурын нумын талаар суралцахад илүү дэмжлэг болно" гэдгийг дурджээ. Тухайлбал, асуулгад хамрагдсан нийт оюутнуудын 92.21% нь морин хуурын хөгжмийн аяыг цахим хэлбэрээр оруулсан байх, 98.72% нь аудио, 96.1% нь видео бичлэг шаардлагатай гэж үзсэн.

IV. МОРИН ХУУРЫН ЭГШИГ ДУУРЬСАЛД НУМНЫ БАРИЛД НӨЛӨӨЛӨХ НЬ - СУДАЛГААНЫ ҮР ДҮН

Морин хуурын урын сангийн хөгжилд хөгжмийн зохиолчдын бүтээлүүд ихээхэн үүрэгтэй. Тухайн хөгжмийн зэмсгийг хуурдах арга барил, эгшиг дуурслын онцлогоос шалтгаалж, хөгжмийн зохиолчдын шинэ арга барилыг практикт нэвтрүүлэх болжээ. Энэ нь хуурчдын ур чадварыг сайжруулахаас гадна хөгжмийн бүтээлээс шалтгаалж шинэ хэллэг, шинэ арга барилыг оруулж өгч байна.

Техник гэдэг нь арга мөн ур зүй юм. Иймд хөгжмийн **зохиолын нэгж өгөгдөл бүрийг эгшиглүүлэхэд** сүвэгчилсэн ухаан нь арга, аргуудыг төгс цогцлоон эгшигүүлэх ур чадварыг ур зүй гэж ойлгож байна. Ур зүй, дүр, утга санаа, илэрхийлэл зэргийн нэгдэл гэжээ. Мөн "хөгжимдөх урлагийн судалгаанд хөгжмийн зохиолчийн барил, хөгжмийн зохиолын утга санааг тодруулан судалсны дараа зохиолыг хөгжимдөх арга барилыг дэргэрүүлэн судлах" гэжээ. Судалгаанд морин хуурын уламжлалт болон орчин үеийн гоцлол бүтээлүүд, тэдгээрийн уламжлал, шинэчлэл, хөгжимдсөн хөгжимчдийн урлах эрдэм, арга ухаан, хэлвэр, дадал зуршил, илэрхийлэл зэрэг өргөн агуулгыг "арга барил" гэж, харин үйлийг хэрхэн хөгжимдөх тайлбарлалд "арга" гэх агуулгаар зааглан авч үзсэн болно [7].

Морин хуур зэмсгийн гол эд анги болох нумыг эхэн хэсэг, дунд хэсэг адаг хэсэг гэж 3 хуваан үздэг ба морин хуур зэмсгийн эгшиг дуурьсалд нумаар тоглох арга барил чухал үүрэг гүйцэтгэдэг. Өөрөөр хэлбэл, нумны барилаас шалтгаалж, дууны түвшин, үндсэн өнгийн дуугаралт буюу дууны чанар, хэллэг (штрих) зэрэг олон зүйлс шалтгаалдаг.

Морин хуурын нумны Монгол барилыг (Пүрэвхүү 1993) Зураг 2 –д харуулсан ба хуурдах арга нь хөгжимчнөөс маш нарийн ур эзэмшихийг шаардаг.

Зураг. 2. Монгол барил.

Энэхүү Монгол нумны барилын давуу тал:

1. Хөгжмийн эгшгийн хүчдэл, даралтыг тод гаргадаг (даралт болон өнгө дуугаралт).
2. Татлага, уртын дуунд шаардагдах хэллэг, арга нь барилд илүү тохиромжтой мөн тодорхой хүчдэл ихсэж, бүтээлийн илэрхийлэлд нөлөөлдөг [6].
3. Онцлог 3 хурууны даралтад оршино.
4. Дунд хуруу нь хүчний тэнцвэрийг тогтвортой байлгадаг.
5. Нумыг татах, түлхэх хүчийг нэмэгдүүлж хүчдэлийн тэнцвэртэй байдал үүсгэнэ [17].
6. Гурван хурууны жигд даралт үүссэнээс эрхий хуруу, долоовор хуруун үе дээр даралт үүсэж, нум сууж өгнө. Үүний үр дүнд баруун гарын бугуй суларч хүчдэлийн тэнцвэртэй байдал үүснэ.

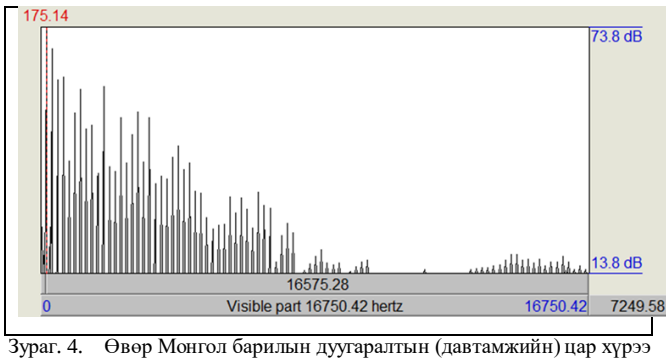
Өвөр Монгол барил (Зураг 3) нь 1990 оноос БНХАУ-ын Өвөр монголын өөртөө засах орноос манай улсад нэвтэрсэн ба сонгодог хөгжмийн зэмсгийн хэрэглэгдэхүүн болсон нумны шинэ хэлбэр хийцийг ашиглаж эхэлсэн [13]. Монгол барилыг бодвол эрхий болон долоовор хурууны завсар тулгуур модоо байрлуулдаг.

Зураг. 3. Өвөр Монгол барил.

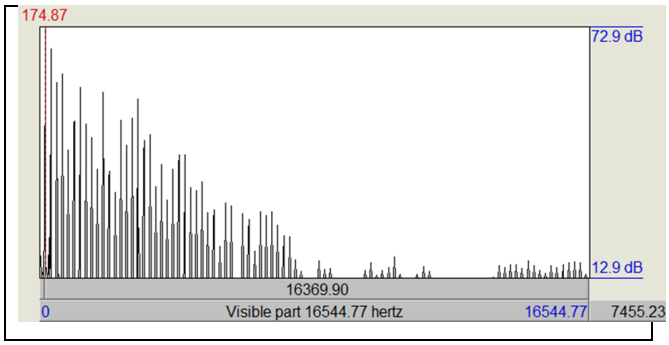
Энэхүү барил нь хэд хэдэн давуу талтай. Үүнд:

1. Уг нуман хялгасыг чангалах, суллах тулгуур модыг ашигласан [9].
2. Мөн савхан урт модыг нуман хэлбэртэй (трост) болгон ашигласан нь нумны дэвшилтэт тал нь болно.
3. Дэлхийн сонгодог зохиол хөгжимдөхөд маш дөхөмтэй, төвөггүй, хэллэг [12] (стаккато, спиккато буюу ойлгох, үсэргэх) гаргахад гол хүчин зүйл нь болсон.

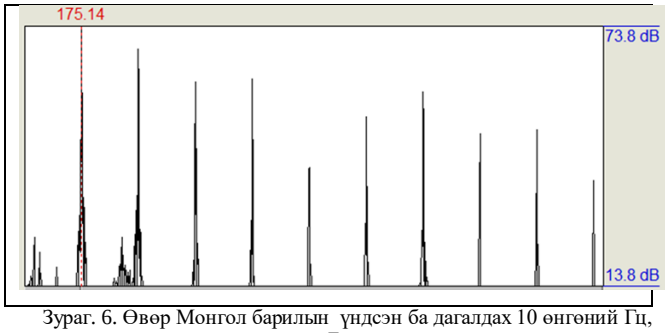
Морин хуурын Өвөр Монгол болон Монгол барилуудын хувьд эгшгүүдийн дуугаралтын цахилгаан дохионд дууны боловсруулалтын Praat 6 программ хангамж ашиглан судалгаа хийсэн. Үүнд: үндсэн болон дагалдах өнгөнүүдийн давтамж, цахилгаан дохионы хэлбэр, формант буюу энергийн төвлөрлийг харьцуулан судалсан. Судалгааны хүрээнд студид мэргэжлийн нэг морин хуурчаар бага наймцын фа эгшгээс хоёрдугаар наймцын фа эгшиг хүртэл 14 эгшгийг Монгол болон Өвөр Монгол барилаар хуурдуулсан. Өөрөөр хэлбэл, нийт 2 барилын 28 эгшгийн дуугаралт дээр судалгаа хийсэн. Жишээ болгон "**бага наймцын фа**" эгшгийн дуугаралтын цахилгаан дохионы судалгааг оруулж өгөв (Зураг 4- Зураг 9).



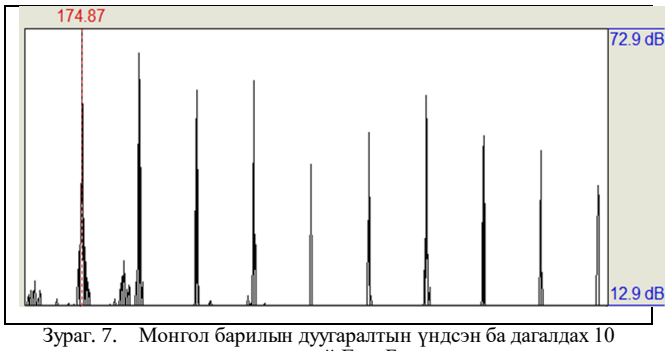
Зураг. 4. Өвөр Монгол барилын дуугаралтын (давтамжийн) цар хүрээ



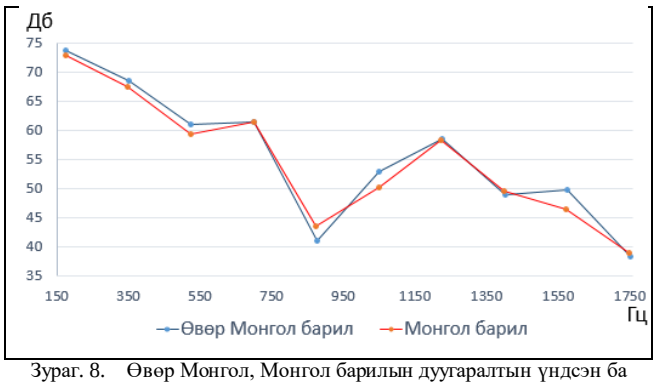
Зураг. 5. Монгол барилын дуугаралтын (давтамжийн) цар хүрээ



Зураг. 6. Өвөр Монгол барилын үндсэн ба дагалдах 10 өнгөний Гц, дБ



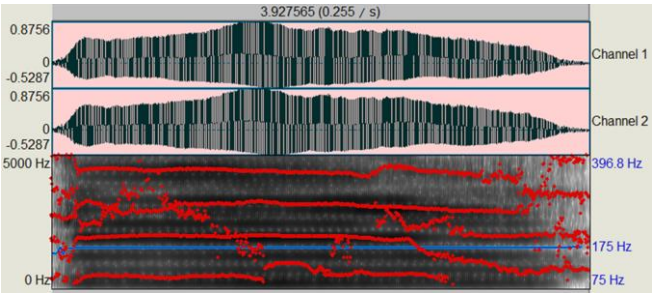
Зураг. 7. Монгол барилын дуугаралтын үндсэн ба дагалдах 10 өнгөний Гц, дБ



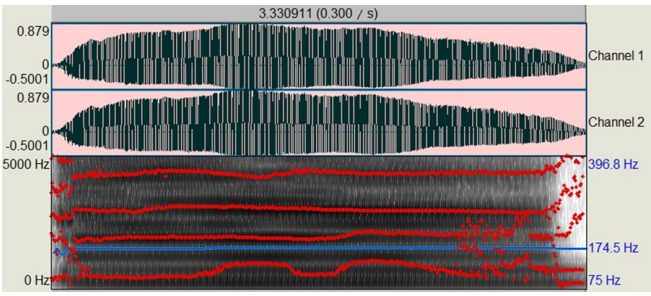
Зураг. 8. Өвөр Монгол, Монгол барилын дуугаралтын үндсэн ба дагалдах 10 өнгөний Гц, дБ

Зураг 4 –Зураг 8 харахад дээрх 2 нумны барилын үндсэн өнгөний давтамжийн зөрүү = 0.9 Гц, дуугаралтын түвшин зөрүү = 0.27 дБ байгаа нь дээрх үзүүлэлтүүд бараг ялгаагүй гэдгийг илтгэж байна.

Гэсэн хэдий ч Зураг 9 болон Зураг 10 –с харахад Монгол нумны барилын хувьд "**бага наймцын фа**" эгшгийн дуугаралтын цахилгаан дохионы энергийн бөөгнөрөл илүү тогтвортой байгаа нь илүү цэвэр дуугаралттай байгааг илтгэж байна.



Зураг. 9. Өвөр Монгол барилын дуугаралтын энергийн төвлөрөл



Зураг. 10. Монгол барилын дуугаралтын энергийн төвлөрөл

V. ДҮГНЭЛТ

1. Судалгаанаас үзэхэд мэргэжлийн морин хуурчид, хуурын сургалтын төв, их дээд сургууль, мэргэжлийн байгууллагуудад суралцагч нарт уламжлалт Монгол нумны барилыг сургах, заах, уламжлуулах хэрэгцээ шаардлага их байна.

2. Урлаг судлал талаасаа Монгол барилаар татлага, уртын дуу, ардын дуу хөгжимдөхөд илүү тохиромжтой нь тогтоогдсон.

3. Өвөр Монгол болон Монгол нумны барилуудын хувьд эгшгүүдийн дуугаралтын цахилгаан дохионы үзүүлэлтүүд болох үндсэн өнгөний давтамж, дагалдах өнгөний давтамжууд, дууны түвшин бараг адилхан байсан ч Өвөр Монгол барилын хувьд энергийн төвлөрөл сариналт ихтэй бол Монгол барилын хувьд энергийн төвлөрөл тогтвортой байсан. Энэ нь Монгол барил нь дуугаралт цэвэр, сайн байгааг харуулж байна.

4. Монгол үндэстэн эрт үеэс хялгасыг хялгасаар хөрөөдөж ирсэн бөгөөд энэ нь цэвэр монгол сэтгэхүй байжээ. Иймд уламжлалт монгол барилаар татлага, уртын дуу, ардын дуу хуурдах, уламжлалаа сэргээн хөгжүүлэх , биет болон биет бус өвийг түгээн дэлгэрүүлэх, үндэсний дэг сургууль, заах арга зүйг баяжуулах, сургалтын хөтөлбөрт суулгах, хөгжимчдийн чадварыг дээшлүүлэх, монгол барилыг олон улсад таниулан сурталчлах, түгээн дэлгэрүүлэх зайлшгүй шаардлага тулгарч байна.

НОМ ЗҮЙ

[1] Алдошина И, Приттс Р. Музыкальная акустика. Учебник для высших учебных заведений. 2014. Копмозитор Санкт –Петербург.

[2] Meredith Powell. Acoustic Analysis of the Viola.Department of Physics, UC Davis.

[3] https://corpus.eduhk.hk/english_pronunciation/index.php/2-2-formants-of-vowels/

[4] Баатаржав, С. "Хуруу, сарвууны тасарсан гэмтлийг нөхөн сэргээх мэс заслын эмчилгээг боловсронгуй болгох." Улаанбаатар: Доктор (Ph.D)-ын зэрэг горилсон бүтээл, 2008.

[5] Баттөгс, О, “Морин хуур зэмсгийн эгшиг дуурьсал болон цахилгаан дохионы үзүүлэлтийн харьцуулсан судалгаа“ докторын нэг сэдэвт бүтээл. 2018

[6] Бум-Эрдэнэ, З., Дагмид, З. "Татлага хуурдах онцлог." Ярилцлага, Улаанбаатар, 2017.12.27.

[7] Дансангийн Хөгжимт. “Дөрвөн утаст хуурын төрөл зүйл, хуурдах арга барилын уламжлал шинэчлэл.” УБ: Докторын зэрэг горилсон бүтээл, 2024.

[8] Жамьян, Г. “Морин хуур сурах гарын авлага.” УБ, 1978.

[9] Мөнхбаяр, Х. “Нумаар хөгжимдөх арга техникийн асуудал.” Улаанбаатар, СУИС: Магистрын зэрэг горилсон бүтээл, 2007.

[10] Мөнхдэлгэр, Ж., Мягмарсүрэн, Г., Жавзансүрэн, П. "Монгол шагай харваа." Улаанбаатар, 2021.

[11] Мөнхдэлгэр, Ж. “Монгол шагай харвааны өв уламжлал, зан үйлийн судалгаа.” УБ: Докторын зэрэг горилсон бүтээл, 2024.

[12] Оюунбадрах, Б. "Төгөлдөр хуурын зохиолыг боловсруулах асуудал." Ву Оюунбадрах. Б, 54. Улаанбаатар: Урлаг судлалын доктор (Ph.D)-ын зэрэг горилсон бүтээл, 2014.

[13] Өлзийбаяр, Ш. “Морин хуурын аймгийн хөг хөглөгөө болон хуурдах арга.” USA: Морин хуурын анхдугаар их наадам, 2001.

[14] Пүрэвхүү, Ц. “Морин хуурын тоглох арга барил.” Улаанбаатар, СУИС: Магистрын зэрэг горилсон судалгааны ажил, 1993.

[15] СУИС. "Хөгжим, соёл урлагийн тайлбар толь IV." Улаанбаатар: Мөнхийн үсэг ХХК, 2018.

[16] СУИС, ХУС. Хөгжим судлал илтгэл өгүүлэл эмхтгэл II. УБ, 2020.

[17] Эрдэнэчимэг, Л. “Хуурын татлага.” Улаанбаатар, 1994.

Монгол үндэсний хээ угалзын зарим дүрсийг математик загварчлалаар вектор график болгон зурах

П.Гантуяа
Мэдээлэл, Компьютерын Ухааны Тэнхим, Монгол Улсын Их Сургууль, Мэдээллийн Технологи, Электроникийн Сургууль
Монгол улс, Улаанбаатар хот
gantuya.p@.num.edu.mn

Г. Багаболд*
Мэдээлэл, Компьютерын Ухааны Тэнхим
Монгол Улсын Их Сургууль, Мэдээллийн Технологи, Электроникийн Сургууль
Монгол улс, Улаанбаатар хот
bagabold.g@num.edu.mn

А.Мөнхзул
Мэдээлэл, Компьютерын Ухааны Тэнхим
Монгол Улсын Их Сургууль, Мэдээллийн Технологи, Электроникийн Сургууль
Монгол улс, Улаанбаатар хот
munkhzul.a@num.edu.mn

***Хураангуй*— Монголчуудын уламжлалт соёлд хээ угалз нь маш чухал байр суурь эзэлж, олон зууны турш эдлэл хэрэгсэл, хувцас, чимэглэлийн зүйлс дээр өргөн хэрэглэгдсээр ирсэн. Хээ угалз нь зөвхөн гоёмсог дүрслэл бус, утга агуулга, бэлгэдэл, ёс заншлыг илэрхийлдэг гүн гүнзгий утга бүхий соёлын илэрхийлэл юм. Монголчуудын өдөр тутмын амьдралд хээ угалзыг дизайны салбарууд тэр дундаа график дизайн, эд зүйлсийг гоёж, чимэглэх зэрэг хэрэглээний урлагт ихээхэн хэрэглэсээр ирсэн бөгөөд түүгээр дамжуулан соёлын элемент, утга бэлгэдэл зэргийг онцгойлон авч, тухайн цаг үеийн оюун санаа, үнэт зүйлсийг илтгэдэг онцлогтой байжээ.**

Энэ судалгаагаар бид Монголчуудын хээ угалзны зарим дүрсийг математик загварчлал ашиглан чанартай вектор дүрс бүтээх арга эрэлхийлсэн юм.

Мөн энэхүү ажлаар Монголын дүрс тэмдгийн тэр дундаа хээ угалзын дүрсийн элементүүдийг орчин үеийн дизайнд ашиглахад үр дүнтэй, чанартай, алдаагүй зөв бүтээгдсэн, хэмжээ бага эзлэх, цааш үргэлжүүлэн зурж дүрслэх боломжтой загвар гарган авахыг зорьж, аналитик геометр аргаар туршин судаллаа.

Түлхүүр үг— аналитик геометр, хээ, вектор график үүсгэх

I. УДИРТГАЛ

Өнөө үед хэрэглээнд ихэвчлэн компьютерийн растер, вектор графикийг түлхүү ашигладаг бөгөөд эдгээр нь чанартай зураг үүсгэхэд файлын хэмжээ их болох, нягтралаас хамаарах зэрэг давуу болон сул талуудтай. Тиймээс бага зай эзлэх чанартай график үүсгэн хэрэглээнд нэвтрүүлэх зорилгоор судалгаа, туршилтийг хийсэн.

Хээ угалзын дүрсийн хэрэгцээ шаардлага нь дизайн, үйлдвэрлэлийн салбарт өргөн тархсан бөгөөд энэ нь дан ганц чимэглэлийн үүрэгтэй биш үндэсний соёлын онцлогийг илтгэж байдаг бэлгэдэлийн гүн гүнзгий утга санааг агуулж байдаг онцлогтой.

Монголчуудын соёл, ахуй амьдрал, ёс заншил нь олон зуун жилийн түүхийн өөрсдийн онцлог, өвөрмөц шинжийг хадгалан хөгжиж ирсэн. Хээ угалз нь Монголчуудын үйлдвэрлэлийн бүтээгдэхүүнд урлахуйн ухаан, гоо зүйн мэдрэмжийг илтгэхийн зэрэгцээ тухайн цаг үеийн зан заншил, итгэл бишрэлийг илэрхийлдэг.

Эдгээр нь Монголчуудын амьдралын хэв маяг, соёлын үнэт зүйлс, байгаль орчинтойгоо хэрхэн зохицож ирснийг харуулж буй бөгөөд Монголчуудын оюун санааны өв соёлын салшгүй хэсгүүд билээ.

Эрт үеэс нүүдэлчин монголчууд том овортой, хүнд жинтэй зүйлийг урлахаас зайлсхийн , урлах хүсэл эрмэлзлэлээ гэрийн эд хогшил болон малын тоног хэрэгсэлдээ шингээн хадгалдаг байжээ.

Растер график нь [1-3]пиксел буюу жижиг дөрвөлжин цэгүүдээс бүрдсэн дижитал график юм. Энэ нь фото зураг, бодит дүрслэлүүдийг харуулахад хамгийн тохиромжтой бөгөөд пикселүүдийн өнгө, бүтэц, хэлбэрийг ашиглан дүрсийг бүтээдэг. Растер графикийн гол шинж чанарууд:

- Пикселийн бүтэц: Растер график нь олон тооны пикселүүдээс бүрддэг бөгөөд эдгээр нь дүрсийг нарийвчлан харуулдаг.
- Нягтрал: Пикселийн тоо их байх тусам дүрсийн чанар өндөр байдаг. Гэхдээ зургийг томруулахад пикселүүд томорч, дүрс бүдгэрдэг.
- Файлын хэмжээ: Растер график нь өндөр нягтралтай байх тусам файлын хэмжээ ихэсдэг.

Растер график нь Photoshop, GIMP зэрэг програмууд дээр засварлагддаг бөгөөд PNG, JPEG зэрэг файлын форматаар хадгалагддаг.

Вектор график нь[2][5] математик тэгшитгэлээр дүрслэгддэг шугам, муруй, хэлбэрүүдээс бүрдсэн дижитал график юм. Энэ төрлийн график нь пикселээс хамааралгүй учраас чанар нь ямар ч хэмжээгээр томруулсан эсвэл багасгасан ч алдагддаггүй давуу талтай. Вектор графикийн онцлогууд:

- Томруулахад чанар алдагдахгүй: Пикселээс хамаардаггүй учир вектор график нь ямар ч хэмжээнд нарийвчлалаа хадгалдаг.
- Өнгө, хэлбэрийг нарийвчлан тодорхойлох: Дүрсийг математик тэгшитгэлээр тодорхойлдог тул шугам, хэлбэр, өнгийг хялбар өөрчилж болно.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

- Файлын хэмжээ: Вектор график нь файлын хэмжээ харьцангуй бага, учир нь зөвхөн математик өгөгдлийг хадгалдаг.

II. Судалгааны өгөгдөл

Хээ угалз нь гоёл чимэглэлийн хамгийн хөнгөн, утга бэлгэдэл агуулсан хэсэг бөгөөд Монгол үндэстний хээ угалз ойролцоогоор 1500 орчим төрөлтэй. Бидний судалгаанд цугласан нийт 2007 тамганы дүрсийн 43% нь хээ угалзын дүрс эзэлж байгаа бөгөөд эндээс бид алхан, хас, өлзий, гал дүрсийг сонгон авсан.[15]

Зохиомжийн онолын зарчимаар хээний дүрсийг авч үзвэл ихэнх дүрсэд тэгш хэмт байдал ноёрхож байгаа нь монголчууд эртнээс уламжилж ирсэн тэнцвэрийг эрхэмлэх уран сайхны сэтгэлгээ, аливаа юмсын гадаад үзэгдэх байдлыг зөв дүрслэх арга, зарчим юм. Судалгаанд 2000 орчим дүрсийг зохиомжийн зарчмаар ангилсан байдлыг хүснэгт 1.-д харуулав.

TABLE I. ЗОХИОМЖИЙН АНГИЛАЛ

	Тэгш хэмийн ангилал	*Тамганы дүрсүүд*	*Эзлэх хувь*
1	Толин	970	53%
2	Эргэлтийн	180	10%
3	Зөөлтийн	20	1,1%
4	Тэгш бус хэмт	610	33,8%

Эндээс нийт дүрсийн 53% нь толин тэгш хэмээр, 10% нь эргэлтийн тэгш хэмээр, 1,1% нь зөөлтийн тэгш хэмээр, 33,8% тэгш бус хэмийн зарчмаар тус тус бүтсэн байна. Зохиомжийн ангилал бүрээс зураг 1-т үзүүлсэн дүрсийг сонгон авсан.

Зураг 1. а. алхан хээ б. хас хээ. в. өлзий хээ г. гал хээ

Тухайлбал; Алхан, толин, хас хээнүүд эргэлтийн, өлзий хээ зөөлтийн, гал хээ тэгш бус хэмт зарчимаар бүтсэн.

III. Судлагдсан байдал

Монгол хээ угалзын жишээн дээр аналитик геометр аргааг зурж байгаа нь анх удаагынх боловч олон улсд хээ дүрсийг илрүүлэх векторoор зураг зэрэг ажлууд цөөнгүй байна. Эдгээрээс дурьдвал энэтхэгийн соёлын өв болох Колам хээг бүтээхэд гүн суралгалтын арга ашигласан[6], архитектурын хээ, загваруудыг таних, ангилах зорилгоор гүнзгий суралцах нейрон сүлжээний аргачлал ашигласан[7]. Даавуун материалын хээ таниулан ангилах зорилгоор CNN (Convolutional Neural Network) ашигласан[8],[9],[10],[11] гэх мэт.

IV. Судалгааны арга

Аналитик геометр нь[11][12]геометрийн асуудлуудыг алгебр болон координатын систем ашиглан шийдвэрлэх аргачлал юм. Энэ нь обьектуудыг (жишээ нь шулуун, муруй, хавтгай гэх мэт) алгебрийн тэгшитгэлүүдээр илэрхийлж, координатын хавтгай дээр дүрслэх боломжийг олгодог.

1. Координатын систем: Обьектуудын байршлыг тодорхойлохын тулд координатын системийг ашиглана. Хамгийн түгээмэл хэрэглэгддэг нь Картесян координат (x,y,zx, y, z) юм.
2. Тэгшитгэлүүд:

   Шулуун шугам (y= mx + c), парабол *(y= ax^2 + bx + c), эллипс (x^2/a^2+y^2/b^2=1)* зэрэг нь аналитик геометрийн обьектуудыг илэрхийлдэг.
3. Хувиргалт: Тодорхой дүрсүүдийн тэгш хэм, эргэлт, шилжилт гэх мэт хувиргалтуудыг математик аргаар тодорхойлно.

Аналитик геометр нь математикийн суурь зарчмууд дээр тулгуурлан инженерчлэл, архитектур, компьютер график зэрэг олон салбарт өргөн ашиглагддаг.

V. Туршилтын үр дүн

Судалгаанд бүртгэгдсэн нийт дүрсээс сонгон авч, 200x200 pixel-ийн хэмжээтэй растер зураг болгон хадгалсан. Дүрсүүдийг зурахдаа вектор ашиглан дүрсэлж эргүүлэлт, параллел зөөлтүүдийг ашиглан зурах арга алгоритм боловсруулав. Боловсруулсан алгоритмаар ангилал тус бүрээс хээ сонгон авч зуруулсан. Жишээ болгон алхан хээний алгоритмийг доор орууллаа.

Алхан хээ: Параллел зөөлт болон эргүүлэлт ашиглан тус хээг зурав. Параллел зөөлтийн үндсэн томъёо [12]

$$\begin{cases} x_1 = x + a \\ y_1 = y + b \end{cases}$$

Алгоритм:

- Алхан хээ нь нийт 8 цэгээс тогтоно.
- Эдгээр цэгүүд нь 4 ба 4-өөрөө Y тэнхлэгийн дагуу тэгш хэмтэй байна.
- Иймд эхлээд 4 цэгийг зураад, үлдсэн 4 цэгийг өмнө зурсан 4 цэгийн X тэнхлэгийн координатын тэмдгийг солих аргаар олж болно.

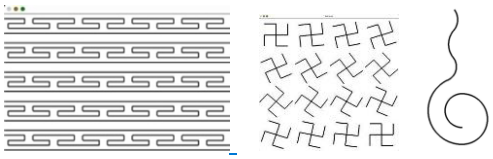
Эхний 4 цэгийн координатууд: Цэгийн хувьд өргөн нэг нэгж, урт хоёр нэгж байдлаар пропорционал байдлаар сонгон авсан дараах дөрвөн цэг болно.

- A(0.5, 0)
- B(0.5, 1)
- C(2.5, 1)
- D(2.5, 2)

Дараа нь:

- Нэг ширхэг алхан хээтэй болно.

- Үүнийг X тэнхлэгийн дагуу 6 буюу 3 дахин их байхаар зөөж (1), холбогдсон шинэ алхан хээг зурна.
- Үүнийг олон удаа 1 мөр болон баганаар давтан хийснээр нийт хээ гарган авна.[13]



Зураг 2. алхан, хас, гал, хээний гаралтын үр дүн

Хас хээ:

Алгоритм:

- Туйлын координатын (Polar coordinate system) системийг ашиглавал координатын эхлэл цэгээс `radius` урттай, X тэнхлэгтэй `angle` өнцөг үүсгэсэн цэгийн координат нь дараах томъёогоор олдоно[13]:

  A(radius×cos(angle), radius×sin(angle)

  B(radius×cos(angle), radius×sin(angle))

- Хас хээ нь **4 хэсгээс бүрдэх** ба эдгээр нь **хоорондоо 90 градусын зөрүүтэй** буюу 90 градусын эргүүлэлтүүдээр үүсгэж болно.
- Нэг хэсгийг зурсан тохиолдолд **давталт** ашиглан үлдсэн хэсгүүдийг **амархан зурж болно**.

Нэг хэсгийн тайлбар:

- Энэ хэсэг нь **хоёр цэгээс** бүрдэнэ:
  - **Эхний цэг:**
    - X тэнхлэгтэй `angle` өнцөг үүсгэн байрлана.
    - Координатын эхээс `radius` зайд байрлана.
  - **Хоёр дахь цэг:**
    - Өнцөг нь `angle + 45 градус`.
    - Зай нь 2×radius2 ×radius.
- Үлдсэн 3 хэсгийг `angle`-ийг **90 градусаар нэмэгдүүлэн** тооцоолж гаргана.

Өлзий хээ:

1. **Анхны нөхцөл**

- Эхлэх цэг: Координатын хавтгай дээр `(-100, 100)` цэг дээр байрлана.
- Чиглэл: Зүүн тийш буюу вектор $(-1, 0)$ байна.
- Комплекс тоо ашиглалт: Цэгийн байрлал болон чиглэлийг комплекс тоо ашиглан илэрхийлсэн

2. **Үндсэн үйлдлүүд**

- Урагш x нэгж явах томьёо

Цэгийн байрлал `P` болон чиглэл `dir` байгаа үед, цэгийг `x` нэгжээр шилжүүлэх үйлдэл дараах томъёогоор илэрхийлэгдэнэ[13].

$$p' = p + dir \cdot x$$

- Төвийг тойрон өнцгөөр нар буруу эргэх томьёо

Цэгийг тодорхой төв дээр өнцгөөр эргүүлэхэд дараах томъёог ашиглана.

$$p' = c + (p - c)e^{\theta i}$$

Мөн чиглэл вектор шинэчлэгдэнэ.

$$dir' = dir \cdot e^{\theta i}$$

Хэрэв эргэлтийг *n* жижиг хэсэг болгон хуваавал:

$$P_k = c + (p - c) \cdot e^{\frac{k\theta}{n} i}, k = 1, 2, ..., n$$

3. **Алгоритмын дараалал**

```
(-100, 100) байрлалаас 50 нэгж урагш явна.
Эргэлтийн төв: P + (0, 50), өнцөг: -3π/2
Урагш 300 нэгж алхана.
Эргэлт: төв нь P + (50, 0), өнцөг: π
Урагш 300 нэгж алхана.
Эргэлт: төв нь P + (50, 0), өнцөг: −π
Урагш 300 нэгж алхана.
Эргэлт: төв нь P + (50, 0), өнцөг: 3π/2
Урагш 300 нэгж алхана.
Эргэлт: төв нь P + (0, 50), өнцөг: −π
Урагш 300 нэгж алхана.
Эргэлт: төв нь P + (0, 50), өнцөг: π
Эцэст нь 250 нэгж урагш алхаж замыг дуусгана.
```

Дөл хээ:

1. **Анхны нөхцөл**

- Эхлэх цэг: Координатын хавтгай дээр `(0, 100)` цэг дээр байрлана.
- Чиглэл: Зүүн тийш буюу вектор `(1, 0)` байна.

- Комплекс тоо ашиглалт: Цэгийн байрлал болон чиглэлийг комплекс тоо ашиглан илэрхийлсэн

2. Үндсэн үйлдлүүд

- Урагш x нэгж явах томьёо

Цэгийн байрлал `P` болон чиглэл `dir` байгаа үед, цэгийг `x` нэгжээр шилжүүлэх үйлдэл дараах томъёогоор илэрхийлэгдэнэ.

$$p' = p + dir \cdot x$$

- Төвийг тойрон өнцгөөр нар буруу эргэх томьёо[13]

Цэгийг тодорхой төв дээр өнцгөөр эргүүлэхэд дараах томъёог ашиглана.

$$p' = c + (p - c)e^{\theta i}$$

Мөн чиглэл вектор шинэчлэгдэнэ.

$$dir' = dir \cdot e^{\theta i}$$

Хэрэв эргэлтийг *n* жижиг хэсэг болгон хуваавал:

$$P_k = c + (p - c) \cdot e^{\frac{k\theta}{n} i},\ k = 1, 2, \ldots, n$$

3. Алгоритмын дараалал

1. Эргэлтийн төв: `(0, 0)`, өнцөг: `- π`
2. Эргэлт: төв нь `(0, -25)`, өнцөг: `-π`
3. Эргэлт: төв нь `(0, 0)`, өнцөг: $-\pi$
4. Эргэлт: төв нь `P + (-60, 30)`, өнцөг: `0.85π`
5. Эргэлт: төв нь `P + (50, 40)`, өнцөг: -0.45π
6. Эргэлт: төв нь `P + (-40, 30)`, өнцөг: `0.4π`

Дараах хүснэгтэнд 200*200 pixel хэмжээтэй, нягтаршил нь 150ppi, CMYK өнгө сонгож файлын хэмжээг харьцууллаа.

	Хэ угалзны нэр	*Растер(Kb*	*Вектор (Kb)*	*Аналитик геометр(Kb)*
1	Хас	629	1166	118
2	Гал	621	1163	118
3	Өлзий	649	1175	118
4	Алхан	593	1163	118

Хүснэгт 1. RASTER, VECTOR, АНАЛИТИК ГЕОМЕТР АРГУУДААР ДҮРСЭЛСЭН ФАЙЛЫН ХЭМЖЭЭГ ХАРЬЦУУЛСАН ҮР ДҮН

Дээрх хүснэгтэнд 200*200 pixel хэмжээтэй, нягтаршил нь 150ppi, CMYK өнгө сонгож файлын хэмжээг харьцууллаа. JPEG нь шахсан растер зураг юм. Шахалтын түвшин өндөр боловч зарим мэдээлэл алдагддаг (lossy compression). Гэсэн ч CMYK өнгөний загвар (ялангуяа хэвлэлийн зориулалттай зурагт) хэрэглэснээр файл арай том хэмжээтэй байж болно. JPEG формат нь шахсан боловч “Хас” болон “Өлзий” зэрэг дүрсийн хувьд пикселийн нарийвчлал өндөртэй тул файлын хэмжээ их хэвээр байна. Харин анлитик геометр аргаар зурсан Open Gl дээр зурсан файлын хэмжээ растер зургаас 5 дахин бага, вектор зургаас 9 дахин бага байна.

VI. ДҮГНЭЛТ

Судалгаанд уламжлалт хээ угалзын дүрсээс сонгон авч график загварчлалаар толин, эргэлтийн, зөөлтийн тэгш хэмт болон тэгш бус хэмт загвар тус бүрийн хээг математик(analytics geometry) загварлалаар бага хэмжээтэй чанартай вектор дүрс гаргах алгоритм боловсруулсан бөгөөд файлын хэмжээ растер зургаас 5 дахин бага, вектор зургаас 9 дахин бага байна.

Алгоритм ашиглахад дүрсийн тэгш ба тэгш бус хэмт байдал нөлөөтэй байна. Өрөөр хэлбэл тэгш хэмт дүрсүүдийг хябар алгоритаар зурсан бол тэгш бус хэмт дүрсийг илүү олон алхамт аргаар дүрсэлсэн. Бид судалгааны ажилдаа энгийн математик загвар ашигласан хэдий ч монгол соёлын судалгаанд энэ нь анх удаа хийгдэж байгаа ажил бөгөөд соёл урлаг, компьютер график дизайн салбарт үр дүнтэй гэж дүгнэж болно. Цаашд дээрх судалгааг үргэлжүүлэн хиймэл оюун, гүн сургалтын арга ашиглан илүү сайн үр дүн гарган авах зорилготой.

Ном зүй

[1] Foley, J. D., van Dam, A., Feiner, S. K., & Hughes, J. F. (1996). *Computer Graphics: Principles and Practice* (2nd ed.). Addison-Wesley.

[2] Agoston, M. K. (2005). *Computer Graphics and Geometric Modeling: Implementation and Algorithms*. Springer.)

[3] Adobe Systems Inc. (2021). *Understanding File Formats in Illustrator and Photoshop*. Retrieved from https://helpx.adobe.com.

[4] Salomon, D. (2004). *Data Compression: The Complete Reference* (3rd ed.). Springer.K. Elissa, “Title of paper if known,” unpublished.

[5] Baxes, G. A. (1994). *Digital Image Processing: Principles and Applications*. Wiley.

[6] A. Sasithradevi1*, Sabarinathan2†, S. Shoba1†, S. Mohamed Mansoor Roomi3† and P. Prakash4†KolamNetV2: efficient attention-based deep learning network for tamil heritage art-kolam classification

[7] Miguel José de Lacerda e Megre Jordão Duarte "Deep Learning-Based Image Analysis for Cultural Heritage Preservation, Restoration and Accessibility "

[8] Muhammad Ather Iqbal Hussain 1, Babar Khan 2, *, Zhijie Wang 1,* and Shenyi Ding 1 Woven Fabric Pattern Recognition and Classification Based on Deep Convolutional Neural Networks

[9] JAMES F. BLINN (1982) *A Generalization of Algebraic Surface Drawing

[10] Timo Eloranta (2001) *TimGA: A Genetic Algorithm for Drawing Undirected Graphs

[11] Maurice D. Weir (1995) *Calculus and Analytic Geometry (9th Edition) by Thomas

[12] Liang-Shin Han (1996) *Complex Numbers and Geometry

[13] I. M. YAGLOM (1968) *Complex Numbers in Geometry

[14] V. Scott Gordon (2018) *Computer Graphics Programming in OpenGL with C++

[15] П.Гантуяа. “Адууны тамганы дүрсийн судалгаа, ангилал” 2020. Улаанбаатар

.

Хиймэл оюун ашигласан яриа таних загвар

Т.Отгонбаатар
ЭХИТ, МТЭС
МУИС
Улаанбаатар, Монгол
t.otgonbaatar11@gmail.com

Э.Адъяабат*
ЭХИТ, МТЭС
МУИС
Улаанбаатар, Монгол
adiyabat.e@num.edu.mn

Одоогийн байдлаар хиймэл оюуны хэлний загварууд түлхүү хөгжиж байна. Жишээ нь Chatgpt, Deepseek, LLaMa зэрэг чатботын системүүд нь хүн төрөлхтөнд хиймэл оюуны хүчин чадал, гайхамшигийг мэдрүүлээд зогсохгүй технологийн салбарын шинэ чиг хандлагыг тодорхойлж, хиймэл оюунд итгэх итгэлийг бүрэн дүүрэн мэдрүүлээд байгаа билээ. Харин хиймэл оюуны хөгжлийн дараагийн шинэ чиг хандлага бол хүн төхөөрөмж хоёр хоорондоо байгалийн хэл буюу хүн төрөлхтний ашигладаг ярианы хэл ашиглан харилцан ойлголцох явдал юм.

Хүн машин хоёр харилцан ойлголцоход туслах хамгийн гол загвар нь хүний ярианы машин ойлгох хэлд хөрвүүлэх, машин ойлгодог хэлийг хүний ойлгодог хэлд хөрвүүлэх загварууд юм. Хамгийн гол үзүүлэлт нь илтгэгчийн нас, хүйс гэх мэт онцлогоос үл хамааран машины хэл рүү хөрвүүлэх явдал юм. Тийм учраас энэ ажлын хүрээнд монгол хэлээр ярьсан яриаг текст рүү хөрвүүлэх Whisper загвар, монгол хэлээр бичсэн текстийг өөр нэг зорилтод хоолойгоор хэлүүлэх MMS загварыг хэрэгжүүлсэн. Энэхүү судалгааны ажлын үр дүнд тонгол хэл дээр ажиллах харилцан ярианы систем хөгжүүлэх боломжтой.

Түлхүүр үгс—Whisper, MMS, TTS, SST

I. Удиртгал

Siri, Cortona гэх мэт яриагаараа төхөөрөмжтэй харилцах загварууд ихэвчлэн англи хэл дээр ажилдаг. Эдгээр загваруудыг өдөр тутмын амьдралдаа ашиглаж болох ч хамгийн гол анхаарах ёстой зүйл нь эх хэлнийхээ дархлааг хамгаалах явдал юм. Мөн хүн бүр англи хэлийг төрлөх хэл шигээ ашиглаж чадахгүй. Тийм учраас эх хэлээрээ төхөөрөмжүүдтэй харилцан ойлголцох загварыг бий болгох нь хамгийн чухал сэдэв болоод байгаа билээ. Компьютер буюу машин нь дижитал утга дээр ажилдаг бол хүний яриа нь акустик сигнал бөгөөд аналог утгатай дохио юм. Үүнийг тодорхой хугацааны агшинд түүвэрлэж тасралттай тоон утга болгож машин дээр боловсруулалт хийдэг. Яриаг текст болгоход хамгийн их саад болдог зүйл нь орчины дуу чимээ юм. Орчны дуу чимээг дарах алгоритм нь ажиллагаа ихтэй байдаг тул үндсэн загварын параметрийн тоог ихэсгэхэд хүрэгдэг. Загварын параметр их байх тусам удаан ажилдаг. Загварын ажиллах хугацаа их байх нь real-time ажиллахад маш их хүндрэл учруулж байна.

II. Whisper болон MMS загвар

A. *Whisper загварын тухай*

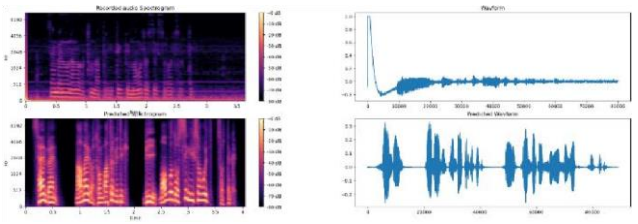
Уг загвар нь 2022 оны 9-р сард OpenAI-с Алек Рэдфорд тэргүүтэй судлаачдын олон нийтэд нээлттэй гаргасан ярианаас бичвэрт хөрвүүлэх загвар бөгөөд 96 орны хэл дээр ажиллах боломжтой. Common voice датасетын хувилбар 11 буюу нийт 860000 цагийн ярианы өгөгдөл дээр сургагдсан. Уг өгөгдлийн 1/3 нь англи хэл дээрх өгөгдөл бол үлдсэн 576200 гаруй цагийн өгөгдөл нь бусад 95 хэлний ярианы өгөгдөл тус тус багтана. Үүнээс 10 цаг нь монгол хэлний өгөгдөл байдаг. Whisper-ийн моделиуд болон тэдгээртэй харгалзах монгол хэлний WER(үгийн алдааны түвшин) BELU (Bilingual Evaluation Understudy)-н утгыг хүснэгт 1-д харуулсан. WER-н утга бага BELU-н утга их байх тусам илүү үр дүн сайтай загвар гэж үзнэ [1]. Өмнө нь сургасан загвар буюу checkpoint-уудыг ашиглан дахин сургах боломжтой. Үүнийг fine tune хийх гэж нэрэлдэг. Бид өөрсдийн бэлтгэсэн өгөгдөл дээр дахин сургасан бөгөөд сургалтын үр дүнд үгийн алдааны түвшин буурч байгаа нь ажиглагдсан.

B. *MMS-TTS загварын тухай*

MMS-TTS нь Facebook AI-ийн боловсруулсан Massively Multilingual Speech (MMS) төслийн нэг хэсэг бөгөөд 1,100 гаруй хэл дээр текстийг ярианд хөрвүүлэх чадвартай загвар юм. Энэ нь олон хэлний ярианы технологийг хөгжүүлэхэд томоохон ахиц гаргасан төсөл юм. Монгол хэлний өгөгдлийн хувьд ар монгол болон өвөр монголын гэсэн өгөгдлөөс бүрдэнэ.

C. *Whisper болон MMS-TTS моделийг хамт хэрэгжүүлэх*

Whisper моделийн гаралтаас авсан бичвэрийг MMS моделийн оролтод өгч ажиллуулж улмаар илтгэгчээс хамааралгүй хэн нэгэн хүний хоолойг зорилтот буюу MMS-TTS-н Халх аялга бүхий нэг илтгэгчийн хоолойд хувирна. Зураг 1-д Whisper-н оролтод өгсөн 5 секундийн аудио бичлэгийн дохио болон харгалзах спектрограммын хамтаар харуулсан байна.



Зураг 1. Оролт болон гаралтын спектрограмм

III. Whisper загварын ажиллах зарчим

Whisper загвар нь дараах үндсэн дөрвөн дараалаар ажиллана.

1) Аудио файлыг оролтод бэлтгэх: Хүний яриаг 30 секунд болгон дүүргэж эсвэл тасалж, 16кГц-ийн давтамжтай түүвэрлэнэ.

2) Preprocessing функц mel spectrogram үүсгэх: Whisper загвар нь raw аудио сигналд доорх алхмуудаар Mel spectrogram үүсгэдэг.

- Short-Time Fourier Transform: Аудиог богино хугацааны хэсгүүдэд хуваан спектр үүсгэнэ.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

$$X(t,\omega) = \sum_{n=\infty}^{\infty} \binom{n}{k} x[n]w[n-t]e^{-jwn} \quad (1)$$

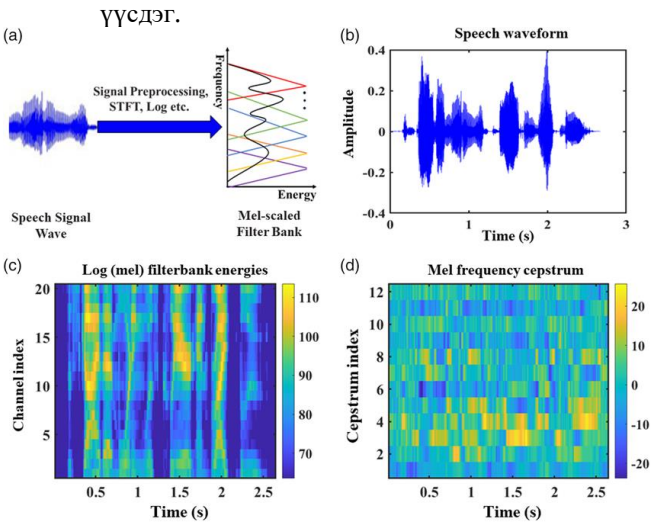
- Mel Filter Bank хэрэглэж, спектрийг Mel scale рүү хөрвүүлэх:

$$m = 2595 \cdot \log_{10}(1 + \frac{f}{700}) \quad (2)$$

- Log scale:

$$S_{mel}(t,m) = \log_2\left(\sum_k |X(t,w_k)|^2 * H_m(w_k)\right) \quad (3)$$

Үр дүнд нь [80 x 3000] хэмжээтэй Mel spectrogram (жишээ нь 30 секундийн аудио) үүсдэг.



Зураг 2. Whisper Preprocessing

3) Transformer Encoder-Decoder: Whisper загвар нь Transformer архитектур ашигладаг. Encoder нь mel spectrogram-г багц (embedding) болгон хөрвүүлнэ. Decoder нь token дараалыг autoregressive байдлаар үүсгэнэ.

- Encoder оролт:

$$Z_0 = PositionalEncoding(Linear(S_{mel})) \quad (4)$$

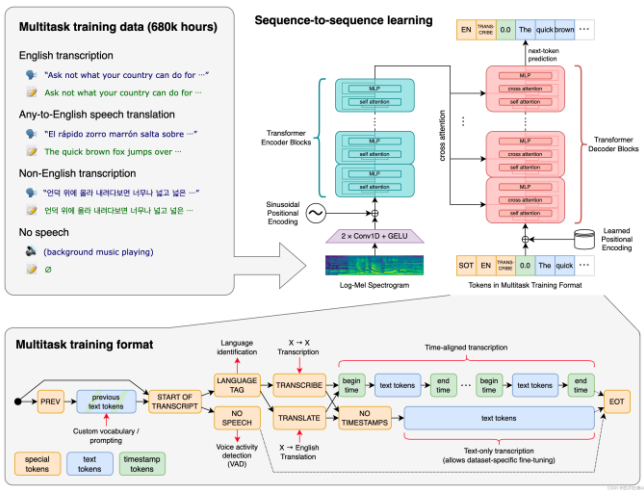
80-ын хэмжээтэй mel feature-г 512 хэмжээтэй embedding рүү шилжүүлнэ.

- Decoder гаралт:

$$y_0 = Softmax(W_0 h_t) \quad (5)$$

Гаралт нь Vocab проекцын жингийн матриц байна.

4) Tokenizer: Текстийг token-ууд руу хөрвүүлж сургах эсвэл таамаглахад ашигладаг. Whisper нь Multilingual BPE (Byte-Pair Encoding) tokenizer ашигладаг. Tokenizer нь англи, испани, монгол гэх мэт олон хэлийг нэгэн зэрэг дэмждэг. Монгол хэлний хувьд Cyrillic-based subword tokenization хийгддэг.



Зураг 3. Whisper model architecture (multitask training format)

IV. MMS-TTS ЗАГВАРЫН АЖИЛЛАХ ЗАРЧИМ

Текстээс яриа руу хөрвүүлэх загваруудын хамгийн гол зарчим нь ярианы агуулга, аялга, дуудлага зэрэг шинж чанарууд ойлгоход байдаг.

1) Posterior Encoder: Оролтын текст болон ярианы аудио дохио (target spectrogram)-ооc далд төлөв (latent representation) үүсгэнэ. Энэ нь ярианы агуулга, аялга, дуудлага зэрэг шинж чанаруудыг илэрхийлнэ. Encoder нь олон давхар convolutional + attention давхаргуудыг ашиглан оролтыг далд давхарга руу шилжүүлнэ. Эцэст нь Variational Autoencoder (VAE)-ийн аргаар μ (mean) болон σ^2 (variance) параметрүүдийг тооцно.

$$z \sim q(z \mid x, y) \quad (6)$$

2) Decoder: Posterior Encoder-ооc авсан далд төлөвийг ашиглан ярианы спектрограмм буюу waveform-ийг буцааж сэргээх. VITS загварын decoder нь autoregressive бус бөгөөд олон спектрограмм бүхий цэгийг нэгэн зэрэг гаргана. [2]

3) Conditional Prior Network: Текстийг зөвхөн өөрийнх нь мэдээллээр нөхцөлдүүлсэн далд давхарга үүсгэнэ. Энэ нь зорилтот аудио байхгүй үед TTS ажиллуулахын тулд хэрэглэгдэнэ. Test phase дээр зөвхөн текст өгөгддөг бөгөөд posterior encoder ашиглаж чадахгүй. Тиймээс Conditional Prior нь текст дээр суурилж гаралтын төлөвийг таамаглаж өгдөг.

4) Flow-based Module (Normalizing Flow): Далд давхаргын гаралтыг илүү хүчтэй, уян хатан болгож, аудионы чанарыг сайжруулдаг.

5) Stochastic Duration Predictor: Текстийн фононууд (эсвэл токенууд) хэр их хугацаанд үргэлжлэхийг таамагладаг. Ярианы хэмнэл (prosody) нь маш олон хувилбартай. Stochastic загвар нь төрөл бүрийн хэмнэл бүхий яриа үүсгэх боломжийг олгоно. Loss функц нь Negative log-likelihood дээр суурилж ажилдаг. [4][5]

$$L_{duration} = -\log p(d|x) \quad (7)$$

V. ДҮГНЭЛТ

Бид энэхүү судалгааны ажлын хүрээнд Whisper моделийг ашиглан хүний яриаг текст болгон, MMS моделийг ашиглан тухайн текстийг MMS-TTS-mon халх аялгаар хэлүүлэх замаар дурын илтгэгчийн хоолойг халх аялгатай эрэгтэйн хоолойд хөрвүүлсэн. Эдгээр нээлттэй эхийн загваруудыг ашиглан дурын өгөгдөл дээр дахин сургаж ашиглах боломжтой.

Бид whisper-ийн small загварыг өөрсдийн бэлтгэсэн өгөгдөл дээр Fine-tune хийх буюу дахин сургасан. Common voice өгөгдлийн хувилбар 11 дээр 10 цагийн монгол хэлний өгөгдөл байдаг. Түүн дээр өөрсдийн 1 цагийн өгөгдлийг нэмээд сургасан бөгөөд нийт 11 цагийн өгөгдөл дээр сурсан гэхэд үр дүн нь маш сайн байсан. Сургалтын дараа бэлэн болсон загварыг өөрийн нөтбүүкны CPU дээр ажлуулхад small загвар нь 5 секундийн аудиог 28 секундийн хугацаанд текст болгож байгаа. Мөн MMS-TTS загвар нь 3-аас 5 секундэд ажиллаж байгаа. Эдгээр моделиудыг GPU дээр ажлуулвал ажиллах хурд нь 3-аас 5 дахин нэмэгдэх юм.

Мөн Whisper-ийн small, tiny, загварууд нь шуугиан багатай өгөгдөл дээр сурсан байдгаас шалтгаалан шуугиантай орчинд танилт муутай ажилдаг. Large, medium гэх мэт моделиуд алдаа багатай сайн ажиллах ч олон параметрүүдтэй учир төхөөрөмж маш их ачаалуулах бөгөөд гаралт гаргах хугацааг удаашруулдаг. MMS модель нь дангаар ажиллах үедээ алдаа багатай ажилладаг бөгөөд туршилтуудын үр дүнгээс харахад Ш үсгийг огт хэлж чадахгүй байгааг мэдсэн. Жишээлбэл оролтод "маш" гэсэн үгийг өгөхөд "ма" гэж хэлэх гэх мэтээр алдаж байгаа нь ажиглагдсан. Цаашид Whisper болон MMS моделиудыг ашиглан харилцан ярианы системийн моделийг хэрэгжүүлэх боломжтой. Дээрх хоёр моделийг ашиглан хэрэглэгч ямар нэгэн асуулт асуухад хариулт өгөх, тухайн даалгаврыг биелүүлэх, харилцан ярих гэх мэт виртуал туслахын системыийг хөгжүүлэх боломжтой.

НОМ ЗҮЙ

[1] Alec Radford, Jong Wook Kim et al, Robust Speech Recognition via Large-Scale Weak Supervision

[2] Vineel Pratap, Andros Tjandra et al, Scaling Speech Technology to 1,000+ Languages

[3] Bapna, A., Cherry, C., Zhang, Y., et al, Massively multilingual joint pre-training for speech and text.

[4] (Rezende, D. and Mohamed, S. Variational inference with normalizing flows. In International Conference on Machine Learning, pp. 1530–1538. PMLR, 2015.).

[5] (Kim, J., Kim, S., Kong, J., and Yoon, S. Glow-TTS: A Generative Flow for Text-to-Speech via Monotonic Alignment Search. Advances in Neural Information Processing Systems, 33, 2020.)

[6] Jaehyeon Kim, Jungil Kong, Juhee Son, Conditional Variational Autoencoder with Adversarial Learning for End-to-End Text-to-Speech.

[7] One System to Rule them All: a Universal Intent Recognition System for Customer Service Chatbots Juan Camilo Vasquez-Correa, Juan Carlos Guerrero-Sierra, Jose Luis, Pemberty-Tamayo, Juan Esteban Jaramillo, Andres Felipe Tejada-Castro

Нууц Үгийн Халдлагад Зориулсан Солонгон Хүснэгт (Rainbow table)-ийг Монгол Хүний Нэрийн Санд Тулгуурлан Үүсгэх нь

А.Ангараг, Б.Өсөхбаяр*, Н.Угтахбаяр,
Мэдээллийн Технологи, Электроникийн Сургууль, Монгол Улсын Их Сургууль
Электроник, Холбооны Инженерчлэлийн Тэнхим
E-mail: 21b1num0607@stud.num.edu.mn
*Coresponding author: usukhbayar@num.edu.mn, ugtakhbayar@num.edu.mn

Хураангуй—Орчин үеийн цахим орчинд хэрэглэгчийн нууц үгсийн аюулгүй байдал нь кибер халдлагад өртөх нөхцлүүдийн нэг болдог. Энэхүү судалгааны ажлаар Монгол хүний нэр болон хувийн мэдээлэлд тулгуурлан нууц үг үүсгэх боломжуудыг судалж, судалгааны үр дүнгээр латин хэл дээрх нууц үгийн санг солонгон хүснэгтийн хамт үүсгэх аргыг боловсруулав. Судалгаанд Монгол хүний нэрийг Кирил үсгээс Латин үсэг рүү хөрвүүлэхдээ хүмүүсийн хэрэглэдэг төрөл бүрийн хэлбэрүүдээр хөрвүүлсэн. Хэрэглэгчдийн Латин нэр дээр төрсөн он, сар, өдөр, нийтлэг хэрэглэгддэг тэмдэгтүүд болон нэмэлт мэдээллийг нэмэх замаар нууц үгийн хувилбаруудыг үүсгэж, тэдгээрийг MD5, SHA-1, NTLM зэрэг үйлдлийн системүүдийн нууц үгийг хадгалахад ашигладаг алгоритмуудаар хувиргасан. Судалгааны үр дүн нь нууц үгийн сан болон солонгон хүснэгтийг ашигласнаар тодорхой төрлийн нууц үгийн халдлагын үр нөлөөг судлах, мөн нууц үгийн аюулгүй байдлыг дээшлүүлэх чиглэлээр дүн шинжилгээ хийх боломжтой. Ингэснээр Монгол улсын хэрэглэгчдийн нууц үгтэй холбоотой аюулгүй байдлын хэв шинжийг тодорхойлоход хувь нэмэр оруулж, цаашлаад нууц үгийн аюулгүй байдлыг сайжруулахад дэмжлэг үзүүлэх ач холбогдолтой юм.

I. Удиртгал

Өнөө цагт мэдээллийн технологи өргөн хэрэглээнд нэвтэрч, цахим орчин дахь мэдээллийн аюулгүй байдал нь хувь хүн болон байгууллагын үйл ажиллагаанд чухал ач холбогдолтой болжээ. Нууц үг нь хэрэглэгчийг баталгаажуулах хамгийн өргөн хэрэглэгддэг арга хэвээр байгаа ч олон нийт, тэр дундаа Монгол улсын хэрэглэгчид аюулгүй байдлын талаарх мэдлэг багатай байснаас болж сул нууц үг ашиглах нь түгээмэл байна. Энэ нь толь бичгийн болон солонгон хүснэгт зэрэг нууц үгийн халдлагуудад өртөх эрсдлийг нэмэгдүүлдэг.

Солонгон хүснэгт нь урьдчилан бэлдсэн хэш утгуудын сан бөгөөд нууц үг тайлах (hash reversal) үйл явцыг эрс хурдасгах боломжтой. Ийм төрлийн хүснэгтийг зөв зохион байгуулснаар хэрэглэгчдийн нууц үгийн хэш утгуудаас эх нууц үгүүдийг нь богино хугацаанд олж илрүүлэх боломжтой байдаг.

Уг судалгааны ажлын зорилго нь Монгол хүний нэр, төрсөн он сар гэх мэт хувийн мэдээллийн өгөгдлүүдэд үндэслэн солонгон хүснэгт үүсгэх юм. Энэ ажлын хүрээнд Монгол хүний нэрийг латин үсэгт хөрвүүлэн, төрсөн он сар өдөр, нийтлэг сүүлэл (жишээ нь: 123, !, @, гэх мэт) зэргийг утгуудыг нэмж нууц үгийн боломжит хувилбаруудыг үүсгэж, тэдгээрийг хэш алгоритмаар хувирган солонгон хүснэгтийн санг үүсгэсэн.

II. Онолын судалгаа

A. Нууц үг болон толь бичгийн халдлага

Нууц үг нь аюулгүй байдлын гол хэрэгсэл бөгөөд зөвшөөрөлгүй нэвтрэхээс хамгаалах эхний үндсэн хамгаалалтын аргын нэг юм [1]. Гэсэн хэдий ч хүмүүсийн нийтлэг шинж чанараас шалтгаалан нууц үгийг таамаглах боломжтой байдаг. Хүмүүс нэр, төрсөн огноо, энгийн загвар гэх мэт хувийн мэдээлэлд тулгуурлан нууц үгээ сонгон, халдлагад өртөмтгий болгодог ба судалгаанаас [2] үзэхэд хүмүүсийн бараг 60 хувь нь төрсөн өдрөө нууц үгээр ашигладаг бол 22 хувь нь өөрийнхөө нэрийг ашигладаг.

Толь бичгийн халдлага нь кибер халдлагын нийтлэг хэлбэр бөгөөд халдагчид хэрэглэгчийн таах боломжтой нууц үгийн урьдчилан бэлтгэсэн жагсаалтыг ашигладаг халдлагыг юм. Жагсаалтыг бэлтгэхдээ хүмүүсийн нийтлэг нууц үг үүсгэдэг хэв шинжийг судалж, түүн дээр дээр тулгуурлан олон янзын хувилбараар нууц үгийг үүсгэдэг.

B. Галиглах

Галиглах гэдэг нь дуу авианы нарийвчлалыг хадгалахын зэрэгцээ текстийг нэг бичгээс нөгөө рүү галиг үсгээр тэмдэглэн хөрвүүлэхийг хэлнэ. Кирил үсгийг ашигладаг монгол хэлний хувьд латин үсэг рүү галиглахдаа MNS 5217:2012 стандартыг баримталдаг. Гэсэн хэдий ч Монгол хэлний зарим авиа нь яг латин үгтэй дүйцэхүйц байдаггүйгээс шалтгаалан хүмүүсийн нийтлэг ашигладаг авиан дээр тулгуурлан практикт зарим нэг кирил үсгийг нэгээс илүү латин үсгээр галигладаг. Жишээлбэл: Ж-г J эсвэл Zh, Х-г Kh эсвэл H гэж болно.

C. Нийтлэг нууц үгийн загвар

Судалгаанаас үзэхэд хэрэглэгчид нууц үг үүсгэхдээ нэрийг нийтлэг энгийн загвар, тэмдэгтүүдтэй хослуулдаг. Эдгээр хэв маяг нь:

- Тоон дагавар: Төрсөн жил, тэмдэглэлт өдөр эсвэл бусад чухал огноо (жишээ нь, Bataa1985)
- Тусгай тэмдэгтүүд: !, @ гэх мэт тэмдэгтүүдийг төвөгтэй байдлын үүднээс нэмдэг (жишээ нь, Bataa1985!)
- Хуанлийн загвар: Төрсөн өдөр өдрүүдийг тусгасан сар өдрийн хослолууд (жишээ нь, 1-р сарын 1-ний Bataa0101)
- Давтагдсан тэмдэгтүүд: Цээжлэхэд хялбар болгох энгийн давталтууд (жишээ нь, Bataa1234)

D. Нууц үгийн шаардлага

NIST Password Guidelines 2024 нууц үгийн шаардлагаар дор хаяж хамгийн багадаа 8 тэмдэгт дор хаяж нэг том тэмдэгт, мөн дор хаяж нэг тусгай тэмдэгт байхаар зөвлөдөг ба үүсгэж буй нууц үгийн санд уг шаардлагыг даган баримталсан.

E. Солонгон хүснэгтийн онолын үндэс ба бүтэц

Солонгон хүснэгт нь хэш функц ашиглан шифрлэгдсэн нууц үгсийг эргүүлэн тайлах зорилготой, урьдчилан бэлтгэсэн өгөгдлийн сан юм. Энэхүү арга нь нууц үгийн хэш утгуудыг муйхар хүчний (brute-force) аргаас илүү хурдан бөгөөд үр дүнтэй тайлах боломжийг олгодог. Солонгон хүснэгт үүсгэх үйл явц нь хоёр үндсэн үе шаттай: бэлтгэл үе шат (offline stage) болон хэрэглээний үе шат (online stage).

F. Бэлтгэл үе шат

Энэ үе шатанд боломжит бүх нууц үгсийн хувилбаруудыг үүсгэж, тэдгээрийг хэш алгоритмаар хувиргах үйл явц явагддаг. Үүсгэсэн нууц үгийн бүх хувилбарыг хадгалах нь санах ойн хувьд үр ашиггүй тул солонгон хүснэгт нь гинжин (chain) холбоос ашиглан мэдээллийг шахаж хадгалдаг.

Хэш функц нь аливаа текстэн өгөгдлийг (жишээ нь, “Batbold123”) тогтсон урттай хэш код болгон хувиргадаг нэг чиглэлтэй функц юм. MD5, SHA-1, SHA-256 зэрэг нь түгээмэл хэрэглэгддэг хэш функцүүдийн жишээ юм.

Буцаах функц (Reduction Function) нь хэш утгын олонлогийг нууц үгийн олонлогийн руу буцаах функц юм. Өөр өөр хэш алгоритмууд нь ялгаатай урттай хэш утгууд үүсгэдэг тул солонгон хүснэгт үүсгэхэд ашиглагдах буцаах функц нь тухайн хэрэглэж буй хэш алгоритмаас хамаардаг. Нууц үгийн боломжит утгуудын олонлог (plain-text set) нь хэш утгуудын олонлогоос (hash set) мэдэгдэхүйц бага байдаг [3]. Жишээлбэл:

- Хэрэв зөвхөн том, жижиг үсэг болон тоо бүхий нийт 62 тэмдэгт ашиглан 8 оронтой нууц үг бүтээвэл боломжит бүх нууц үгийн тоо:

$$62^8 \approx 2.18 \times 10^{14}$$

- Харин MD5 хэш функц 128 битийн хэш үүсгэдэг бөгөөд боломжит хэш утгуудын олонлогийн хэмжээ:

$$2^{128} \approx 3.40 \times 10^{38}$$

Хэш утгын орон зай нь нууц үгийн орон зайгаас хэдэн тэрбум дахин том тул буцаах функц хэш утгыг богино хэмжээтэй боломжит нууц үгт харгалзуулдаг. Үүний улмаас нэг хэш утга олон удаа ижил нууц үгт харгалзах боломжтой байдаг бөгөөд энэ үзэгдлийг давхцал (collision) гэж нэрлэдэг. Давхцлын тоог багасгахын тулд хэд хэдэн өөр буцаах функц $(R_1, R_2, \ldots, R_k)$ ашиглаж болно. Энэ нь солонгон хүснэгт дэхь гинжнүүдийг илүү өвөрмөц болгон давхцал үүсэх магадлалыг бууруулдаг.

Гинжин хэлхээ: Нэг нууц үгнээс эхэлж хэш функц, буцаах функцүүдийг ээлжлэн хэрэглэснээр олон шат дамжсан нууц үг-хэш-нууц үг хэлбэрийн гинж үүсдэг. Солонгон хүснэгтэд зөвхөн гинжний эхлэл ба төгсгөл утгууд л хадгалагддаг. Жишээ нь:

$$\mathrm{pw}_0 \xrightarrow{H} H(\mathrm{pw}_0) \xrightarrow{R_1} \mathrm{pw}_1 \xrightarrow{H} H(\mathrm{pw}_1) \xrightarrow{R_2} \mathrm{pw}_2 \xrightarrow{H} \cdots \xrightarrow{R_k} \mathrm{pw}_k$$

Гинжин хэлхээгээр дамжуулан солонгон хүснэгтэд хэш утгууд болон нууц үгийн хооронд холбоо үүсгэдэг. Нэг гинжин хэлхээ дараах байдлаар үүснэ:

$$\mathrm{pw}_0 \to \mathrm{H}(\mathrm{pw}_0) \to \mathrm{R}_1 \to \mathrm{pw}_1 \to \mathrm{H}(\mathrm{pw}_1) \to \mathrm{R}_2 \to \mathrm{pw}_2 \to \ldots \to \mathrm{pw}_\mathrm{k}$$

Энд:

- pw_0 — Анхны нууц үг.
- $\mathrm{H}(\mathrm{pw}_0)$ — Нууц үгийг хэш функцээр (жишээлбэл: MD5, SHA-1) хувиргаж авсан хэш утга.
- R_1 — Хэш утгийг буцаан нууц үгийн хэлбэрт хувиргах бууралт функц (reduction function).
- pw_1 — R_1 функцийн гаралт буюу шинэ нууц үг.
- $\mathrm{pw}_{\ldots\mathrm{k}}$ — Гинжний эцсийн нууц үг.

Ийм гинжин хэлхээнүүдийг олон удаа үүсгэж, зөвхөн эхлэл болон төгсгөлийн нууц үгсийг хадгалснаар хүснэгт бүрддэг

G. Хэрэглээний үе шат

Хэрэглэгчийн оруулсан хэш утга хүснэгтэнд байгаа эсэхийг шалгах үйл явц энэ шатанд явагдана.

1) Оруулсан хэш утгийг буцаах функц ашиглан боломжит нууц үгийн гинж үүсгэнэ.
2) Уг гинжийн төгсгөл утгуудыг хүснэгтэнд хадгалсан төгсгөлүүдтэй харьцуулна.
3) Таарсан тохиолдолд тухайн гинжний эхлэл утгийг авч, дахин хэшлэх замаар эхнээс нь гинж үүсгэж, хэрэглэгчийн хэш утгатай тохирох нууц үгийг олох оролдлого хийнэ.
4) Хэрвээ тохирвол тухайн нууц үгийг амжилттай тайлсан болно.

H. Солонгон хүснэгтийн төрлүүд

Энэ хүснэгт нь бүтэц, зорилгоосоо шалтгаалан дараах хоёр үндсэн төрөлд хуваагдана:

- Нэг төрлийн солонгон хүснэгт (Homogeneous): Зөвхөн нэг хэш алгоритм дээр суурилсан, жишээлбэл MD5 Rainbow Table. Бүтэц нь энгийн, хурдан боловч хэрэглээ хязгаарлагдмал.
- Олон төрлийн солонгон хүснэгт (Heterogeneous): Олон хэш алгоритм болон өөр өөр төрлийн нууц үгийн бүтэц агуулсан илүү цогц хүснэгт. Илүү уян хатан боловч хэрэглээ өндөр.

III. Хэрэгжүүлэлт, үр дүн

Энэ ажлын эхний шатанд нийт 200 сая хэш утга MD5 алгоритм ашиглан үүсгэж, Монгол нэрэн дээр суурилсан нууц үгсийн боломжит хувилбаруудаас солонгон хүснэгт үүсгэсэн. Гэсэн ч MD5 алгоритм нь криптографын хувьд дутагдалтай гэж үздэг бөгөөд 2004 оноос хойш практикт хэш утгуудын давхцлууд байгааг судлаачид амжилттай олж тогтоосон. Иймд зөвхөн MD5 хэш дээр суурилсан систем нь орчин үеийн хамгаалалтад хангалтгүй юм. Тиймээс энэ ажилд дараах хэш алгоритмуудыг ашиглан өргөтгөх боломжтой гэж үзэж байна:

- SHA-1 (Secure Hash Algorithm 1) — 160 битийн хэш утга үүсгэдэг. Үүний аюулгүй байдал мөн сул ч одоогоор зарим хуучин системд хэрэглэгддэг хэвээр.
- SHA-256 — NIST-ээр батлагдсан, орчин үеийн стандарт хэш алгоритм. 256 битийн хэш үүсгэдэг бөгөөд collision resistance болон pre-image resistance чанарууд нь MD5 болон SHA-1-ээс хамаагүй өндөр.
- SHA-3 (Keccak) — SHA цувралын хамгийн сүүлийн хувилбар бөгөөд цоо шинэ бүтэцтэй (sponge construction). SHA-2-тэй харьцуулахад дотоод бүтэц нь өөр тул зарим төрлийн халдлагад илүү тэсвэртэй.
- BLAKE2, BLAKE3 — SHA-3-ийн нэг хэсэг байсан бөгөөд өндөр гүйцэтгэлтэй, аюулгүй байдал сайтай хэш функцүүд.

Хэш функцүүдийн ялгааг дараах байдлаар авч үзэж болно:

- Хэмжээ (bit-length) — MD5: 128 бит, SHA-1: 160 бит, SHA-256: 256 бит гэх мэт. Урт хэш утга нь илүү том хэш орон зай үүсгэж, collision магадлалыг багасгана.
- Гүйцэтгэл (performance) — MD5 нь хурдан, гэхдээ аюулгүй биш. SHA-256 болон BLAKE2 зэрэг нь илүү найдвартай ч боловсруулалтын хувьд их CPU хэрэглэнэ.
- Collision resistance — хэш утгууд нэг ижил байхаас сэргийлэх чадвар.
- Pre-image resistance — хэш утгаас анхны нууц үгийг буцааж олдох магадлал маш бага байх ёстой.

Солонгон хүснэгт үүсгэхэд эдгээр хүчин зүйлсээс хамаарч дараах өөрчлөлтүүд гарч болзошгүй:

- Илүү урт хэш утга нь гинжин урт, бууралт функцийн бүтэц, хадгалах хэмжээ зэргийг нэмэгдүүлнэ.
- Тухайн хэш функцийн хурд удаашрах нь солонгон хүснэгт үүсгэх болон dictionary attack хийх процессийг удаашруулна.
- Бууралт функц бүрийг тухайн хэш функцэд тохируулан шинээр боловсруулах шаардлагатай болно.

A. Монгол нэрийг галиглах

Кирил үсгээр бичсэн Монгол нэрсийг латин үсэг рүү галиглах. Галиглах үйл явц нь Ж-г J ба Zh-ийн аль алинаар нь буулгах гэх мэт олон хувилбараар бүрдүүлэн ингэснээр үүссэн нэрийн сан нь Монгол хэлний олон янз байдлыг хангаж чадна.

B. Нууц үгийн загваруудыг нэгтгэх

Галигласан нэрсийг нийтлэг нууц үгийн загвартай хослуулан олон хувилбартай болгох, үүнд:

- Жил буюу төрсөн жилийн тоон дараалал (жишээ нь, 1950–2010).
- Сар, өдрийн хослол зэрэг хуанлид суурилсан загварууд (жишээ нь, 1-р сарын 1-ний 0101).
- Нууц үгэнд байнга ашиглагддаг тусгай тэмдэгтүүд (жишээ нь: !, @).
- Нэрийн жижиг, том, холимог үсгийн дүрслэлийг (жишээ нь, bataa, Bataa) оруулах том үсгийн хувилбарууд.

C. Солонгон хүснэгт үүсгэх

Үүсгэсэн нууц үгийн хувилбаруудыг ашиглан солонгон хүснэгт бүтээх. Үүнд:

- Нууц үг бүрийг хэш функцээр (MD5, SHA-1 гэх мэт) хувиргах.
- Hash → Reduction function → Chain хэлбэрээр олон үе шаттай гинж үүсгэж, зөвхөн эхлэл болон төгсгөл утгуудыг хадгалах.
- Давхцал (collision)-аас сэргийлэх арга хэмжээнүүдийг тооцох.
- Солонгон хүснэгтийг хэрэглээний үе шат дээр ашиглан хэш утгаас боломжит нууц үгийг буцаан тодорхойлох, турших орчин бүрдүүлэх.

D. Кирил үсгийг латин руу хөрвүүлэх нь

Энэхүү судалгааны ажлын хэрэгжилтийн үе шатанд галиг үсгээр бичсэн монгол нэрийг нийтлэг нууц үгийн загвартай [4] хослуулан нууц үгийн иж бүрэн мэдээллийн санг бий болгох юм. Энэ хэсэгт хэрэглэгдэх үйл явц, арга хэрэгсэл, аргачлалыг дэлгэрэнгүй харуулав.

Бид Пайтон (Python) хэлийг нэмэлт олон боломжуудад нь тулгуурлан сонгож програмын хөгжүүлэлтийг хийсэн. Энгийн текст файлыг (.txt) оролтын өгөгдөл болон гаралтын үр дүнг хадгалахад ашигласан. Текст файл нь төрөл бүрийн платформтой орчинд ашиглах боломжтой байдаг. Хүний нэрийг галиглахдаа MNS

5217:2012 стандартыг баримтлан нэмэлтээр хүмүүсийн нийтлэг ашигладаг авиан дээр тулгуурлан албан бус байдлаар зарим нэг кирил үсгийг нэгээс илүү латин үсгээр галигласан (Хүснэгт 1).

Хүснэгт I: Кирилээс латин руу хөрвүүлэлт

Кирил	Латин	Кирил	Латин	Кирил	Латин	Кирил	Латин
А	A	а	a	Б	B	б	b
В	V	в	v	Г	G	г	g
Д	D	д	d	Е	E	е	e
Ё	Yo, Y	ё	yo, y	Ж	J, Zh	ж	j
З	Z	з	z	И	I	и	i
Й	Y	й	y	К	K	к	k
Л	L	л	l	М	M	м	m
Н	N	н	n	О	O	о	o
Ө	U	ө	u	П	P	п	p
Р	R	р	r	С	S	с	s
Т	T	т	t	У	U	у	u
Ү	U, V	ү	u, v	Ф	F	ф	f
Х	Kh, H	х	kh, h	Ц	Ts	ц	ts
Ч	Ch	ч	ch	Ш	Sh	ш	sh
Щ	Sch	щ	sch	Ы	Y	ы	y
Э	E	э	e	Ю	Yu	ю	yu
Я	Ya, Y	я	ya, y	Ь	I	ь	i
Ъ	I	ъ	i				

E. Нууц үгийн сан

Монгол хүний нэрсийн санг нийтлэг 4300 гаруй нэрсийг ашиглаж үүсгэсэн [8]. Нууц үгийн санг үүсгэхдээ нийтлэг нэрийн жагсаалт, нийтлэг нууц үгийн дагавар, тусгай тэмдэгт, төрсөн он, сар, өдрийн мэдээллийг бэлдсэн .txt файлаас дуудан хослуулж ажиллуулахад гарах үр дүн нь маш их байгаа нь зарим талаараа бэрхшээл болж байгаа учир уг мэдээллүүдийг гараас аван хослуулах боломжтой болгосноор илүү тусгай оновчтой байдал бий болгосон.

F. Үр дүн

Монголд түгээмэл байдаг 4300 гаруй нэрийг сонгож, тухайн нэрийг төрсөн он, сар, өдөр болон тусгай тэмдэгтүүдтэй хослуулан олон төрлийн нууц үг үүсгэсэн. Ингэснээр нэг нэрнээс дунджаар 40,000 гаруй хувилбар гарч, нийт 200 сая орчим нууц үг бүрдсэн.

Эдгээр нууц үгсийг MD5 алгоритмаар хешлэж, “hash:plaintext” форматтай солонгон хүснэгт үүсгэсэн. Энэ хүснэгт нь өгөгдсөн хэш утгаас эх нууц үгийг буцааж олох боломжийг олгож байгаа бөгөөд үүний үр дүнд нууц үг тайлах ажиллагаа илүү хурдан, хялбар болсон.

Хүснэгт II: Нууц үг үүсгэх загварын жишээ

Загвар	Тайлбар
Bataa1234#	Нэр + тоо + тусгай тэмдэгт (том үсэгтэй)
bataa0325@	Нэр + төрсөн өдөр + тусгай тэмдэгт (жижиг үсэгтэй)
B@taa0325*	Нэрэнд тэмдэгт оруулсан + төрсөн өдөр + тусгай тэмдэгт

Дараах хүснэгтэд нууц үг болон тэдгээрийн хэш утгыг харуулсан бөгөөд энэ нь солонгон хүснэгтийн нэгээхэн хэсэг юм:

Хүснэгт III: Жишээ солонгон хүснэгт

MD5 хэш : plaintext
c2a4a470a7493784796e94f9d8d3b0c0:Bataa1234#
882307f0b1c47eeb3f8df6c7a2b54de9:bataa0325@
3e9333d4b180db9e2924e1ab43e77a39:B@taa0325*

IV. Дүгнэлт

Энэхүү ажлаар Монгол хэрэглэгчдийн нууц үгийн бүтэц, аюулгүй байдлын хэв шинжийг судалж, Монгол хүний нэрийн санг ашиглан латин үсэгт хөрвүүлж, түүн дээр суурилсан нууц үгийн сан үүсгэх боломжийг судалж, хэрэгжүүлсэн. Үүний тулд Монгол хүний нэрэн дээр том, жижиг үсгээр бичих, тэмдэгтүүд нэмэх зэргээр нууц үгийн олон хувилбар үүсгэж Монгол хэрэглэгчдийн нууц үг үүсгэдэг бодит хандлагад нийцэхүйц олон төрлийн нууц үгийн сан бүрдүүлэн тэдгээрийг хэш алгоритмуудаар хувирган солонгон хүснэгт үүсгэх аргачлал боловсруулж, нууц үгийн сан болон солонгон хүснэгтийг үүсгэсэн.

Цаашид энэ ажлын үр дүнг ашиглаж тодорхой төрлийн нууц үгийн халдлагын үр нөлөөг судлах, мөн нууц үгийн аюулгүй байдлыг дээшлүүлэх чиглэлээр дүн шинжилгээ хийх боломжтой. Ингэснээр Монгол улсын хэрэглэгчдийн нууц үгтэй холбоотой аюулгүй байдлын хэв шинжийг тодорхойлоход хувь нэмэр оруулж, цаашлаад нууц үгийн аюулгүй байдлыг сайжруулахад дэмжлэг үзүүлэх ач холбогдолтой юм.

Ном зүй

[1] NIST SP 800-12 Rev. 1 under Password from FIPS 140-2
[2] https://www.the-parallax.com/harris-poll-google-password-security/
[3] Rainbow Tables Yukai Zang Division of Science and Mathematics University of Minnesota, Morris Morris, Minnesota, USA 56267 zangx040@morris.umn.ed
[4] https://nordpass.com/most-common-passwords-list/
[5] https://pages.nist.gov/800-63-4/sp800-63b.html
[6] https://contextqa.com/python-random-stringgeneration
[7] MNS 5217:2012 стандарт
[8] http://www.bolod.mn/modules.php?name=NewsnID=94109
[9] A Novel Dictionary Generation Methodology for Contextual-Based Password Cracking AIKATERINI KANTA 1,2, (Graduate Student Member, IEEE), IWEN COISEL1, AND MARK SCANLON 2, (Senior Member, IEEE)
[10] An Improved Rainbow Table Attack for Long Passwords December 2017Procedia Computer Science 107:47-52

Фрэймийн тоог ихэсгэх байдлаар дүрс бичлэгийн чанарыг сайжруулах

Т. Болдбаяр
Компьютерын ухаан
ШУТИС МХТС
Монгол улс, Улаанбаатар хот
Boldoo41000@gmail.com

Эрдэнэбат Батцэцэг
Компьютерын ухааны салбар
ШУТИС МХТС
Монгол улс, Улаанбаатар хот
battsetseg_e@must.edu.mn

***Хураангуй*— Видео фрэймийн интерполяци нь дараалсан хоёр видео фреймийн хооронд завсрын фрэймүүдийг үүсгэх зорилготой. Фрэймийн интерполяци нь удаан хөдөлгөөн үүсгэх, видео шахах, видео хүрээг урьдчилан таамаглах зэрэг янз бүрийн хэрэглэгдэхүүнүүдийг дэмждэг. Үүнээс гадна, өндөр нягтаршилтай видеон дээр бодит цагийн фрэймийн интерполяци аргуудыг ашиглах нь шууд дамжуулалтын өргөн зурвасын хэрэгцээг бууруулах, тооцооллын нөөц хязгаарлагдмал хэрэглэгчдэд видео засварлах үйлчилгээг хангах, дэлгэцийн төхөөрөмжүүд дээр видео фрэймийн давтамжийг тохируулах зэрэг олон боломжит хэрэглээтэй. Фрэймийн интерполяци нь бодит ертөнцийн видеонуудад тохиолддог нарийн төвөгтэй, шугаман бус хөдөлгөөн, гэрэлтүүлгийн өөрчлөлтүүдээс шалтгаалан хэцүү сорилтуудтай даалгавар юм. Сүүлийн үед урсгалд суурилсан фрэймийн интерполяци алгоритмууд нь эдгээр асуудлыг шийдвэрлэх тогтсон фрэйм болгон ашиглагдаж, өндөр үр дүн үзүүлж байна.**

Энэхүү судалгааны ажлын үр дүн нь тоглоом, виртуал бодит байдал, анимаци, видео сэргээх, засварлалт зэрэг өндөр чанартай фрэйм үүсгэх шаардлагатай салбаруудад ашиглагдах боломжтой.

Түлхүүр үг— Фрэймийн интерполяци, Оптик урсгал, , Мэдрэлийн дүрслэл, Хөдөлгөөн тооцоолол, FPS сайжруулалт

I. Удиртгал

Орчин үед видео контентийн хэрэглээ хурдацтай нэмэгдэж түүнийг дагаж өндөр чанар шаардагдах болсон. Гэвч олон тохиолдолд видео бичлэгийн секунд дэх фрэймийн давтамж хангалттай биш байдаг нь хөдөлгөөнийг жигд бус харагдуулж, үзэгчдэд эвгүй мэдрэмж төрүүлэх, улмаар чанарын бууралт үүсгэх асуудлыг дагуулдаг. Түүнчлэн, илүү өндөр FPS нь хөдөлгөөнийг илүү жигд, гөлгөр байгалийн мэт харагдуулж, ялангуяа бага давтамжтай дэлгэцээс шилжих үед нүдний ядаргаа, анивчилтыг бууруулахад тусалж чадна. Ийнхүү фрэймийн үүсгэлт нь сүүлийн үед олон салбарт эрэлттэй болж байна. Тухайлбал, удаашруулсан бичлэгийг сайжруулах, дараалсан зургийг бичлэг болгох, тоглоомын FPS-ийг сайжруулах зэрэгт ашиглагддаг. Ялангуяа видео тоглоомонд FPS чухал үүрэгтэй. Энэ асуулдлыг шийдэхийн тулд NVIDIA болон AMD –ийн тус тусын хөгжүүлсэн DLSS[6] (Deep Learning Super Sampling), FSR[7] (FidelityFX Super Resolution) технологиуд нь сүүлийн үеийн хувилбаруудаа фрэйм үүсгэх технологийг шингээсэн. Ингэснээр тусламжтай төв процессор болон график процессорын гүйцэтгэлийг бууруулалгүйгээр ачааллыг бууруулдаг. Сүүлийн үеийн ухаалаг зурагтнуудад мөн “TruMotion”, “MotionFlow” зэрэг хөдөлгөөний зөөлрүүлэх функцууд ихээр ашигдагддаг болсон. Ийнхүү фрэйм үүсгэх технологийг ашигласнаар техник хангамжийг бага багаар программ хангамжаар орлуулж байна.

II. Хэрэглээний Үндэслэл ба Боломжит Бүтээл

A. *Системийн архитектур*

RIFE[1] хоёр дараалсан RGB фрэйм болох I_0, I_1 болон зорилтот цагийн агшин t $(0 \le t \le 1)$ өгөгдсөн үед завсрын фрэйм $\tilde{I}_t$-г үүсгэх зорилготой.

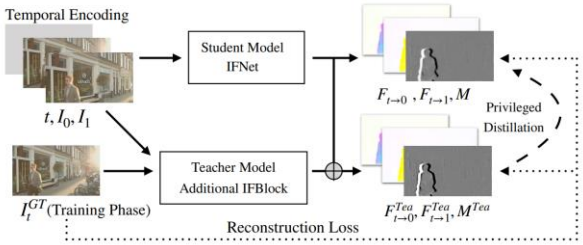


Fig. 1. Ерөнхий архитектур.

Fig 1-т үзүүлсэнээр хоёр оролтын фрейм болох $\mathbf{I_0}$, $\mathbf{I_1}$ болон **цаг хугацааны кодчилол** болох **t**-ийг өгсөн нөхцөлд, тэдгээрийг шууд **IFNet** рүү оруулж, дундын урсгалууд болох $\mathbf{F_{t\to 0}}$, $\mathbf{F_{t\to 1}}$ болон **нэгтгэх маск M**-ийг ойролцоогоор тооцоолно. **Сургалтын үе шатанд, IFBlock** ашиглан **багш загвар** нь **үндсэн үнэн (ground truth) фрейм** $\mathbf{I_t}$ дээр тулгуурлан **сурагч загварын** үр дүнг сайжруулна. **Сурагч болон багшийн загваруудыг** анхнаасаа **хамтад нь сургалтын алдагдал (reconstruction loss)** ашиглан хамт сургадаг. **Багшийн тооцоолол нь илүү нарийн байдаг** тул сурагчид сурахад нь чиглэл өгч чаддаг.

IFNet нь оролтын фрэймүүд болон $\tilde{I}_t$-г нэмэлт суваг болгон ашиглаж, завсрын урсгалууд $F_{t\to 0}$, $F_{t\to 1}$ болон fusion map M-ийг тооцоолно. Үүн дээр үндэслэн завсрын фрэймийг дараах томьёогоор гарган авна:

$$\tilde{\mathbf{I}}_t = M \odot \tilde{I}_{t\leftarrow 0} + (1 - M) \odot \tilde{I}_{t\leftarrow 1} \quad (1)$$

$$\tilde{I}_{t\leftarrow 0} = \overline{\mathrm{W}}(I_0, F_{t\to 0}), \quad \tilde{I}_{t\leftarrow 1} = \overline{\mathrm{W}}(I_1, F_{t\to 1}) \quad (2)$$

Энд:

- $\overline{\mathbf{W}}$ нь буцаан хувиргах дүрслэл (*image backward warping*),

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

- $\odot$ нь элемэнт бүрийн үржвэр (*element-wise multiplication*),
- **M** нь нэгтгэх зураглал (*fusion map*, $0 \leq M \leq 1$) болно.

(1) Энэ тэгшитгэл нь эцсийн дундын фрэйм $\tilde{\mathbf{I}}_t$ нь $\tilde{I}_{t\leftarrow 0}$ ба $\tilde{I}_{t\leftarrow 1}$ гэсэн хоёр хувиргасан фрэймүүдийн жингийн хослол болохыг харуулж байна. Fusion map M нь хоёр хувиргасан фрэймийн хоорондох холих тэнцвэрийг тодорхойлдог."

(2) Энэ нь дундын цагийн алхам t-аас фрэйм 0 болон 1 ийн урсгал юм. Энэ нь дундын фрэймийн бүх пикселүүдийг фрэйм I_0, I_1 -д байгаа пикселүүдтэй хэрхэн тохируулахыг зааж өгдөг.

B. Хэрэглээний кейсүүд

Энэхүү технологийг дараах зүйлсэд ашиглаж болно. Тухайлбал:

- Бага FPS (жишээ нь 24 эсвэл 30 FPS)-тэй видеог 60 FPS эсвэл түүнээс дээш болгон хөрвүүлж, илүү жигд хөдөлгөөнтэй болгох
  - Хэвийн хурдтай бичлэгээс удаашруулсан видео үүсгэх
  - Бичлэгийг шахах, хэмжээг бууруулах
  - Тоглоомын гүйцэтгэлийг сайжруулах зорилгоор нэмэлт фрэйм үүсгэх
  - Телевиз болон монитор дээр бодит цагийн фрэйм интерполяци хийх
  - Өгөгдөл нэмэгдүүлэх (data augmentation), завсрын төлөвүүдийг симуляцлах
  - Гөлгөр шилжилт, хурдны өөрчлөлт, фрэймийн тохиргоо хийх
  - Гараар зурсан фрэймүүд эсвэл зураг хооронд гөлгөр байдал нэмэх
  - Анагаахын дүрслэл ба шинжлэх ухааны цаг хугацааны цуврал дүрслэлийн (жишээ нь MRI зэрэг) хооронд интерполяци хийх

C. Бүтээлийн Гол Үүрэг ба Функцууд

Хоёр өгөгдсөн фрэймийн (жишээ нь A ба B) хооронд нэг буюу хэд хэдэн фрэймийг үүсгэж, бага фрэймийн давтамжтай видеонд гардаг гацалт, үсрэлтийг арилган хөдөлгөөнийг илүү жигд харагдуулах, үүсгэгдсэн фрэймүүд нь хүрээлэх бодит фрэймүүдтэйгээ цаг хугацааны хувьд уялдаатай байх ёстой. Энэ нь гялбалт, сүүдэр давхцал, дүрсний алдаа зэргээс сэргийлдэг. Нэхээс, ирмэг, хөдөлгөөний бүдгэрэл (motion blur) зэрэг нарийн дүрслэлийг бодитойгоор дахин бүтээнэ. Шугаман бус, далдлагдсан, эсвэл том хөдөлгөөнүүд (жишээ нь хурдан хөдөлж буй объект эсвэл камерын эргэлт) дээр үр дүнтэй ажиллах нь гол зорилго юм. Үүнийг хэрэгжүүлэхийн тулд дараах функцуудаас бүрдэнэ:

- Хөдөлгөөн тооцоолох (Motion Estimation) Warping (хувиргах)
- Холбох / Нэгтгэх (Blending / Fusion)
- Нарийвчлал нэмэх (Refinement)
- Цагийн байршлын хяналт (Temporal Position Control)

D. Ижил Төстэй технологиудын харьцуулалт

Уламжлалт фрэймийн интерполяцийн аргуудад оптик урсгалын тооцоолол, хөдөлгөөний нөхөн төлөвлөлт, шугаман интерполяци, блок тохиролуулалт багтдаг. Оптик урсгалын тооцоолол нь хамгийн өргөн хэрэглэгддэг арга. Хоёр фрэймийн хооронд пиксел бүрийн хөдөлгөөнийг (өөрөөр хэлбэл, хөдөлгөөний вектор) тооцдог. Хөдөлгөөний нөхөн төлөвлөлт нь тооцоолсон оптик урсгалыг ашиглан объект хэрхэн хөдөлж байгааг урьдчилан таамаглаж, дундын байрлалд байрлуулан шинэ фрэймийг үүсгэнэ. Шугаман интерполяци нь хэрэв хөдөлгөөнийг нарийн тооцоолох боломжгүй бол энгийн шугаман интерполяци (жишээ нь, пикселийн утгуудыг холих) ашигладаг. Гэвч энэ нь сүүдэр давхцах буюу дүрсний алдаа гаргах магадлалтай. Харин Блок тохиролуулалт нь фрэймийг блокоор хуваах, дараагийн фрэйм дэх хамгийн сайн таарах блокуудыг хайж олно. Энэ арга нь илүү энгийн боловч оптик урсгалаас бага нарийвчлалтай. Ихэвчлэн видеог шахах технологид (жишээ нь, MPEG) ашигладаг.

Уламжлалт аргуудын хувьд Peak-signal-to-nois-ratio (PSNR) нь ойролцоогоор 25–30 dB байдаг. Энэ нь энгийн дүрсүүдэд хангалттай боловч хэцүү, хурдан хөдөлгөөнтэй дүрсүүдэд чанар хурдан мууддаг. Харин орчин үеийн гүн сургалт (deep learning)-д суурилсан аргууд, RIFE нь 33–38 dB хүрдэг бөгөөд энэ нь фрэйм илүү нарийн, үнэн зөв сэргээгдэж байгааг илтгэнэ. Харин хүний харааны ойлголтоор чанарыг хэмждэг SSIM (Structural Similarity Index)-ийн хувьд уламжлалт аргууд ихэвчлэн 0.85–0.90 хооронд байдаг бол RIFE нь 0.95–0.98 хооронд үнэлгээ авч, дүрсний нарийн ширийн мэдээллийг илүү сайн хадгалж чаддаг. Интерполяцийн алдаа (IE) буюу пиксел түвшний алдааны хэмжүүрийн хувьд уламжлалт аргуудын алдаа ихэвчлэн 4.0-оос дээш байдаг бол RIFE нь 2.0-оос доош байж, дундын хүрээг илүү зөв таамаглаж байгааг харуулдаг. Уламжлалт аргууд нь ихэвчлэн дүрсийг давхардах (ghosting), хөдөлгөөний бүдгэрэл (motion blur), болон хослогдсон ирмэг (duplicated edge) зэрэг дүрсний алдаанууд гаргадаг, ялангуяа хурдан хөдөлгөөнтэй дүрсэнд.

Table 1-т зарим гүн сургалтын загварын давуу болон сул талыг харуулав.

TABLE I. ИЖИЛ ТӨРЛИЙН ЗАГВАРУУД

Загвар Систем	Үндсэн Үүрэг	Давуу талууд	Сул тал
DAIN[2]	Гүнзгий байдлын (depth) тооцоолол ашиглан давхцсан эсвэл их хөдөлгөөнтэй бүсүүдийг илүү нарийн интерполяц хийх.	Slow motion болон олон давхаргатай дүрстэй хэцүү дүрслэлүүдэд маш сайн.	Удаан, тооцооллын өртөг өндөртэй.
Super SloMo[3]	Хос чиглэлт optical flow болон үзэгдэх байдлын зураглал (visibility maps) ашиглан завсрын фрэймүүдийг интерполяцлана.	Сайн чанартай, сайн бүтэцтэй.	Удаан, хоцрогдсон
VFIformer[4]	Трансформерт суурилсан архитектур ашиглан фрэйм интерполяци хийдэг	Глобал контекстийг ойлгож, **өндөр чанартай** үр дүн гаргана	CNN-д суурилсан загваруудаас илүү нөөц шаардаг
FLAVR[5]	Vision Transformer (ViT) архитектурыг ашиглан, видео фрэймүүдийн дунд холын зайн хамаарлыг болон дэлгэрэнгүй контекстыг барьж чаддаг.	Трансформер архитектур нь байгалийн болон харааны хувьд эвтэйхэн интерполяцийн фрэймүүдийг гаргахад тусалдаг.	Тооцооллын өртөг өндөр, хүчин чадал сайтай техник хангамж шаардана

Гүн сургалтын загварууд нь өөр өөрсдийн аргачлалтай бөгөөд түүнийгээ дагаад давуу болон сул талуудтай. Бидний хөгжүүлсэн моделийн ялгарах хамгийн том давуу тал нь хурдан бөгөөд техниек хангамжийн шаардлага бага тул энэ нь хүчин чадал багатай график карт байсан ч бодит цагаар боломжийн чанартай ажиллах боломжийг олгож байгаа. Зургийн чанарын үнэлгээгээр бусад моделиудтай Table 2-т харьцуулж харуулав.

TABLE II. ЗАГВАРУУДЫН ҮЗҮҮЛЭЛТИЙН ХҮСНЭГТ

Загвар	Өгөгдлийн багц	PSNR(dB)	SSIM
RIFE	Vimeo90K	~35.7	~0.962
DAIN	Vimeo90K	~34.7	~0.950
SUPER SLOMO	Vimeo90K	~33.8	~0.942
VFIFormer	Vimeo90K	~36.5	~0.9816
FLAVR	Vimeo90K	~36.1	~0.963

Дүрсний чанарын үнэлгээ сайн чанартай загваруудаас хол зөрүүгүй байгаа нь ажиглагдана.

III. СУДАЛГААНЫ АРГА ЗҮЙ

1) *Онцлог шинжийг ялгаж авах*

Хоёр оролтын фрэймийг боловсруулж, хөнгөн CNN суурилсан онцлог шинж илрүүлэгчийг ашигладаг. Энэ нь олон түвшний онцлог шинжүүдийг ялгаж, бага түвшний текстур болон өндөр түвшний хөдөлгөөний бүтэц зэрэг зүйлийг таньж ойлгодог. Энэ нь фрэймийн агуулгыг гүнзгий ойлгох боломжийг олгодог бөгөөд интерполяцийн нарийн байдлыг хангах шаардлагатай.

2) *IFNet (Завсрын Урсгалын Сүлжээ)*

IFNet нь хоёр оролтын фрэймүүдийн хоорондох завсрын оптик урсгалыг шууд урьдчилан таамаглах нейрон сүлжээ юм. Уламжлалт оптик урсгалд суурилсан аргуудтай харьцуулахад IFNet нь төгсгөлийн цэг хүртэл сургагдаж, хурдтай бөгөөд үр ашигтай фрэйм интерполяци хийхэд оновчлогдсон. Энэ сүлжээ нь гүнзгий сургалт ашиглан таамаглалаа сайжруулж, хурдан ажиллагааг хангадаг.

3) *Нөхцөл Байдлыг Харгалзсан Хувиргалт*

Завсрын урсгал тооцоолсны дараа, Загвар нь хоёр оролтын фрэймийг warp хийж, тэдгээрийг интерполяцийн фрэймтэй нийцүүлнэ. Энэ хувиргалт нь нөхцөл байдлыг харгалзан хийгддэг бөгөөд, ингэснээр далдлагдсан хэсэгтэй эсвэл нарийн төвөгтэй хөдөлгөөнтэй бүсүүдэд үүсэх алдааг багасгаж, интерполяцийн фрэймийг тэгшитгэдэг.

4) *Фрэйм Синтезийн Сүлжээ*

Хөнгөн RefineNet сүлжээг ашиглан warp хийгдсэн онцлог шинжүүдийг сайжруулдаг. Энэ сүлжээ нь интерполяцийн фрэймийн чанарыг сайжруулж, артефактуудыг багасган, нарийн деталийг хадгалдаг. Сүүлчийн интерполяцийн фрэймийг үүсгэхэд эдгээр сайжруулсан онцлог шинжүүдийг нэгтгэн ашигладаг, ингэснээр өндөр чанартай үр дүнг гаргадаг.

5) *Сургах үйл явц*

RIFE загварыг Vimeo90K triplet сургалтын өгөгдлийн багц дээр сургасан. Vimeo90K тестийн багцад 448 × 256 нягтралтай 3,782 триплет байдаг. Энэ өгөгдлийн багцыг сүүлийн үеийн видеоны фрэйм оруулах (VFI) аргуудад өргөнөөр ашигладаг. Сургалтыг NVIDIA 5070Ti график картан дээр 100 үеийн турш 10цагийн турш сургасан.

IV. ҮР ДҮН

Загварыг сургасны эцэст түүний алдааг Figure 2-т харуулав. Эхний шат (0–25,000 алхам): Бүх алдагдлууд хурдан буурч байна, ялангуяа L1 болон дистилляцийн алдагдал. Энэ нь сургалтын эхний үе шатанд хурдан суралцалт явагдаж байгааг харуулж байна.

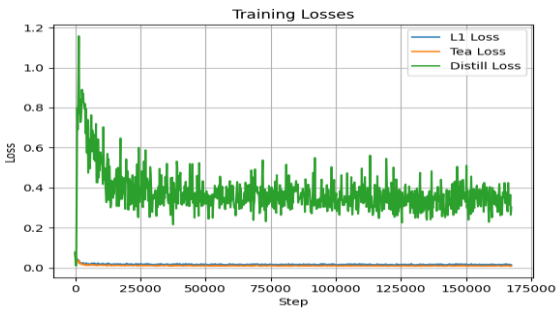


Fig. 2. Сургалтын алдааны график.

Сургалтын дунд үе (25,000–125,000 алхам): L1 болон teacher-ийн алдагдлууд бараг тэгт ойртон тогтворжиж байна. Харин дистилляцийн алдагдал илүү хэлбэлзэлтэй байна. Энэ нь сурагч болон багш загварын таамаглалын хоорондох динамик харилцан үйлчлэл эсвэл алдагдлын функц дэх жингийн мэдрэмжтэй хуваарилалттай холбоотой байж болно. Сургалтын сүүлийн үед (125,000+ алхам) алдагдлууд бага, харьцангуй тогтвортой байна. Хөх болон улбар шар муруйн тэгшрэлээс үзэхэд загвар сургалт тогтвортой төлөвт орсон.

Figure 3-т тус тусын PSNR-ыг харуулав. Улбар шар шугамаар багш эхэндээ өндөр, ~37 дБ орчим бөгөөд хурдан өсөж, ~100,000 алхам орчимд бага зэрэг хэлбэлзэл ажиглагдах боловч ерөнхийдөө жигд, нэмэгдэж байна. . Сургалтын төгсгөлд ~40.2 дБ-д хүрч байна. Багшийн загвар PSNR утгаар оюутнаас тогтмол өндөр үзүүлэлттэй байж байгаа нь мэдлэг дамжуулалтын (distillation) хувьд энгийн зүйл юм.

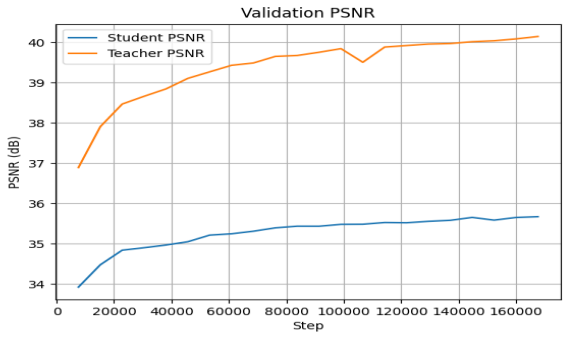


Fig. 3. PSNR хэмжүүрийн график

Харин сурагч загвар эхэндээ ~34 дБ орчим бөгөөд сургалтын явцад аажмаар нэмэгдэж 170,000 алхамд хүрэх үед ~35.7 дБ болж байна. ~125,000 алхамын дараа муруй тэгширч эхэлж байгаа нь оюутны загвар бараг бүрэн суралцсан, сайжрал удааширч байгааг харуулж байна. Загварын бичлэгэн дээр туршиж эх хувилбар болон үүсгэсэн фрэймийг Figure 4-т харуулав.

Fig. 4. Эх хувилбар болон үүсгэсэн фрэйм

Туршихдаа бичлэгийн фрэймийн тоог бууруулж шинээр үүсгэн хоёр бичлэгийг хооронд нь харьцуулахад дундаж алдааны утга нь 40.466476, PSNR хэмжүүрийн дундаж утга 32.059dB бол SSIM 0.9310 зааж байсан. Figure 5-т дээрх 2 фрэймийн пикселийн зөрүүг дулааны зураглалаар харуулав. Улайж тодорсон хэсгүүдэд зөрүү их байна гэсэн үг.

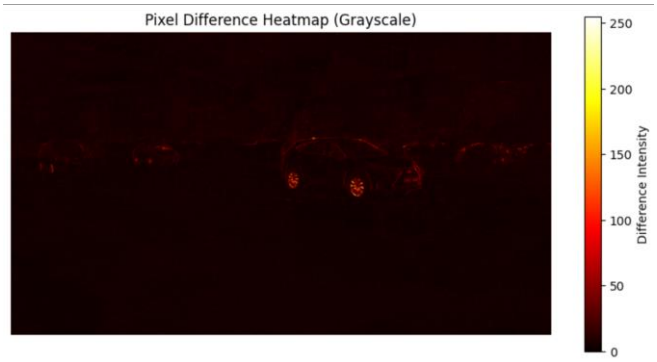


Fig. 5. Пикселийн зөрүүний дулааны зураглал

Фрэймүүдийг томруулж пикселийн түвшин зөрүүг үзэхэд гарсан үр дүнг Fig 6 д харуулав.

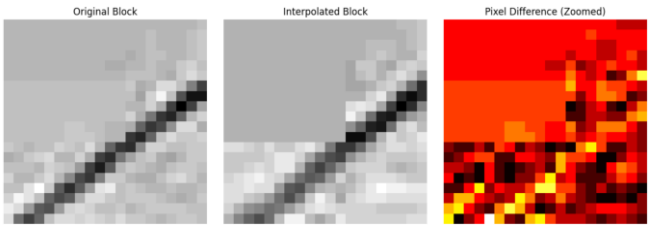


Fig. 6. Пикселийг томруулж харуулсан зөрүү

Ийнхүү загвар амжилттай үр дүнг гаргалаа.

V. Дүгнэлт

Видео фрэймийн интерполяци нь видеоны хөдөлгөөний гөлгөр байдал, өндөр FPS-ийг үүсгэх, видео контентыг сайжруулах, болон видео дамжуулалтыг илүү үр

ашигтай болгох зэрэг олон чухал хэрэглээнд ашиглагддаг. Энэ судалгааны ажилд ашигласан RIFE технологийн алгоритм нь хурдтай бөгөөд өндөр чанартай интерполяц хийх боломжийг олгодог. Загварын хурдан бөгөөд хөнгөн жижиг байдал нь ижил төстэй загваруудаас томоохон давуу талтай бөгөөд энэ нь олон зүйлд ашиглагдаж болох тал харагдаж байна. Үүнийг ашиглан техникийн дутагдалыг нөхөх, зардал хэмнэх, сайжруулалт хийх, мэдээллийг шахах зэрэг олон байдлаар ашиглагдах боломжтой. Иймээс энэхүү судалгааны ажил нь видео боловсруулалтын салбарт хувь нэмэр оруулж, олон төрлийн хэрэглээг дэмжих болно.

НОМ ЗҮЙ

[1] RIFE: Real-Time Intermediate Flow Estimation for Video Frame Interpolation. (Xu, L., Zhang, Z., & Li, H. 2020).

[2] DAIN: A Depth-Aware Video Frame Interpolation Network. Jiang, (Z., Zhang, Z., & Xu, Z. 2019).

[3] Super SloMo: High Quality Estimation of Multiple Intermediate Frames for Video Interpolation. (Dong, X., & He, L. 2018).

[4] VFIformer: Transformer-based Video Frame Interpolation. (Lee, J., & Lee, Y. 2022).

[5] FlavR: Video Frame Interpolation with Vision Transformers. (Zhang, Z., & Li, H. 2022).

[6] *DLSS 3: Introducing Frame Generation for the RTX 40 Series*. (NVIDIA. 2022).

[7] AMD FidelityFX Super Resolution. (AMD 2021).

Биоангаахын хэрэглээнд зориулсан бага чадлын хоёр шатлалт CMOS үйлдлийн өсгүүрийн загварчлал ба шинжилгээ

Дамдингийн Баярцэцэг
ЭХИТ
МУИС, МТЭС
Улаанбаатар, Монгол
bayartsetseg.d@mongolaspiration.edu.mn

Даланбаярын Болормаа
ЭХИТ
МУИС, МТЭС
Улаанбаатар, Монгол
bolormaad@num.edu.mn

Удиртгал—Үйлдлийн өсгөгч нь аналог хэлхээний үндсэн бүрэлдэхүүн бөгөөд дохио боловсруулах, хяналтын систем хэрэгжүүлэх зэрэгт өргөн хэрэглэгддэг. Сүүлийн жилүүдэд бага хүчдэл, бага чадал шаардсан ухаалаг төхөөрөмжийн хэрэглээ нэмэгдсэнтэй холбогдон, бага чадал зарцуулалттай, өндөр гүйцэтгэлтэй үйлдлийн өсгөгчийн эрэлт өсөж байна. Ялангуяа биоанагаахын төхөөрөмжүүд нь хүний биенээс үүсэх сул, шуугианд мэдрэмтгий дохиог боловсруулдаг тул өндөр өсгөлттэй, бага шуугиантай, бага чадал зарцуулдаг үйлдлийн өсгөгч шаардлагатай. Энэ судалгаанд бага чадлын хэрэглээнд зориулсан, өндөр өсгөлттэй, өргөн зурвасын хоёр шатлалтай CMOS үйлдлийн өсгөгчийг 90-нм GPDK технологийн сан ашиглан Cadence Virtuoso программын орчинд загварчилж шинжилгээ хийв. Давтамжийн компенсацгүй загварын өсгөлт 14дБ, хэрчигдэх давтамж 700Гц байсан бол давтамжийн компенсацтай хэрчигдэх давтамж 10КГц болсон ч өсгөлт нэмэгдээгүй. Шилжилтийн багасгаснаар өсгөлт 40дБ болж хэрчигдэх дамтамж 10КГц болсон.

Түлхүүр үг—Хоёр шатлалт үйлдлийн өсгөгч, давтамжийн компенсац, шилжилтийн гүйдэл

I. Удиртгал

Орчин үед юмсын интернет хөгжсөнөөр хүний биед суулгасан болон зөөврийн биоангаахын төхөөрөмжүүдийн хэрэглээ ихэссэн. Эдгээр төхөөрөмжүүд хүний биед үүсэх маш бага хүчдэлтэй (μV–mV), бага давтамжтай (Hz–kHz), гадны шуугианд өртөмтгий аналог дохиог өсгөн тоон хэлбэрт шилжүүлэн боловсруулалт хийдэг тул мэдрэгчийн гаралтын дохиог урьдчилан маш сайн өсгөх үйлдлийн өсгөгч (Op-Amp) шаардлагатай. Тухайлбал, зүрхний (Electrocardiography, ECG), тархины (Electroencephalography, EEG), булчингийн цахилгаан бичлэг (Electromyography EMG) зэрэг дохионууд нь оношилгоонд чухал боловч хэмжихэд хэцүү байдаг. Иймд шуугиан багатай, өндөр өсгөлттэй, бага чадал зарцуулалттай үйлдлийн өсгүүрүүдийн судалгаа биоангаахын салбарын хөгжил чухал болж байна.

Ерөнхий зориулалтын үйлдлийн өсгөгч нь аналог дохиог маш сайн өсгөдөг ч биоангаахын дохионы төрөл, далайц, давтамжаас хамааран өсгөгчийн дизайны гол шаардлага нь бага чадлын (батерейг цэнэглэх хугацаа хэдэн жил хүртэл), нам давтамжийн дохионы өсгөлт 70-80дБ байх гэх мэтээр өөрчлөгддөг [1].

Энэхүү судалгаагаар 90 нм CMOS технологийг ашиглан хоёр шатлалт үйлдлийн өсгөгчийг загварчлан өсгөлтөд нөлөөлөх параметруудад шинжилгээ хийлээ. Био анагаахын имплант эсвэл зөөврийн төхөөрөмжийн батерейг удаан ашиглах шаардлага 180нм-ээс дээш технологи чадлын сарнил их, 65 нм технологи subthreshold гүйдэл ихтэй зэргээс шалтгаалан 90 нм-ийн технологийг сонгосон. Сүүлийн үед өсгөлтийг ихэсгэсэн олон төрлийн үйлдлийн өсгөгчийн архитектур бий болсон ч хэлхээнд талбай бага эзлэх шаардлага тавигдах шаардлагыг гол үзүүлэлтийн нэг болгон авч хамгийн энгийн үндсэн хоёр шатлалт CMOS үйлдлийн өсгөгчийн хэлхээг сонгон авч шинжилгээ хийсэн.

II. Судалгааны арга зүй

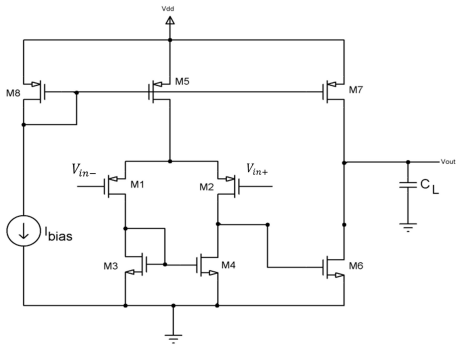
Биоангаахын салбарт хамгийн түгээмэл хэрэглэгддэг гурван төрлийн цахилгаан дохио бол зүрхний цахилгаан бичлэг, тархины цахилгаан бичлэг болон булчингийн цахилгаан идэвхжлийн бичлэг юм. Эдгээр дохионуудын давтамж болон хүчдэлийн далайц өөр байдаг бөгөөд Хүснэгт ??-д эдгээр утгыг харууллаа.

Хүснэгт I
Биоанагаахын сигналуудын давтамж ба далайцын муж

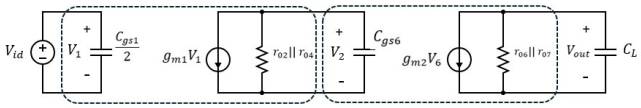
Дохионы төрөл	Давтамж (Гц)	Далайц
ECG	0.05–250	5 μV – 8 mV
EEG	0.5–200	2 μV – 200 μV
EMG	0.01–10,000	50 μV – 10 mV

Эдгээр дохионд тохирсон үйлдлийн өсгөгчийг загварчлахдаа өсгөлтийн коэффициент, нэгж өсгөлтийн зурвасын өргөн (unity gain bandwidth), фазын зөрүү (phase margin) их байх шаардлагатайгаас гадна шуугаман өсгөлтийн мужид шуугианд тэсвэртэй байх шаардлагатай тул эдгээр параметрийн шинжилгээг хийсэн.

Ерөнхий зориулалтын үйлдлийн өсгөгч оролтын дифференциал өсгөгч, өндөр өсгөлт бүхий хоёрдахь



Зураг 1. Хялбар хоёр шатлалт өсгүүрийн хэлхээ



Зураг 2. Эхний болон гаралтын шатны дүйх хэлхээ

шатны өсгөгч, гаралтын буфер өсгөгчөөс тогтдог ба өсгөгч шатны тоо их байгаа нь үйлдлийн өсгөгчийн дотоод хэлхээний давтамжийн хамааралд туйлын (pole)-ийн тоог ихэсгэдэг тул хоёр шатлалт загварыг ашиглалаа.

Хоёр шатлалт CMOS үйлдлийн өсгөгчийн оролтын дифференциал шат нь M1–M4 транзистороос бүрдэх (Зураг 1-ээс харна уу) ба M1 болон M2 транзистор дифференциал хос, M3, M4 транзистороос бүрдэх гүйдлийн толь дифференциал өсгөгчийн гаралтын өндөр эсэргүүцэлтэй ачаа болсноор өсгөлт:

$$A_{v_{d1}} = G_{m1} \cdot (r_{o2} \parallel r_{o4}) \tag{1}$$

Энд G_{m1} — M1 транзисторын дамжуулалт, r_{o2} ба r_{o4} — M2 ба M4 транзисторын сул дохионы дотоод эсэргүүцэл болно.

Дамжуулал болон гаралтын эсэргүүцлийг олохын тулд ашигласан дүйх хэлхээг Зураг 2-т үзүүллээ. Гаралтын шат нь M7 шилжүүлэгч ачаа бүхий ерөнхий суурьтай M6 транзистороос тогтоно. Хоёрдахь шатны өсгөлт:

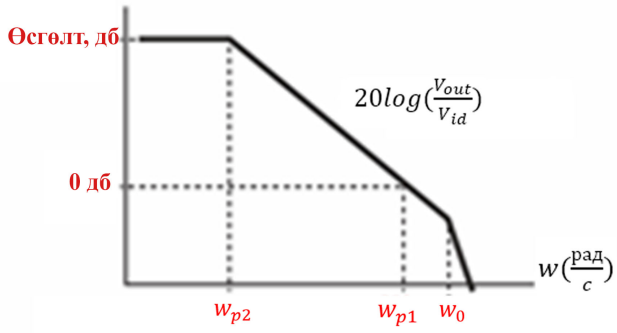
$$A_{v_{d2}} = G_{m6} \cdot (r_{o6} \parallel r_{o7}) \tag{2}$$

ба энд G_{m6} — M6 транзисторын дамжуулалт, r_{o6} ба r_{o7} — M6 ба M7 транзисторын сул дохионы дотоод эсэргүүцэл болно. Өсгөгчийн дамжууллын функц:

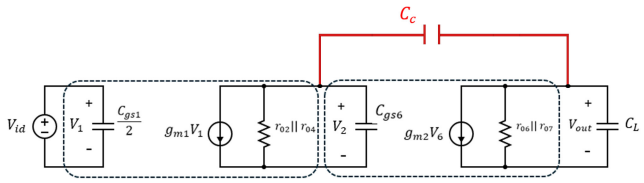
$$H(s) = \frac{v_{\text{out}}(s)}{v_{\text{id}}(s)} = \frac{K}{(1 + s/\omega_{p1})(1 + s/\omega_{p2})} \tag{3}$$

ба энд: K — өсгөгчийн нийт өсгөлт буюу

$$K = A_{vd1} \cdot A_{vd2}, \quad s = j\omega, \quad \omega_{p1}, \omega_{p2}$$



Зураг 3. Өсгөгчийн фазын зөрөө



Зураг 4. Давтамжийн компенсацтай хэлхээ

— дамжууллын функцийн туйлууд.

Туйлуудыг олбол:

Давамгайлах туйлууд дараах байдлаар тодорхойлогдоно:

$$\begin{cases} \omega_{p1} = \frac{1}{(r_{o2}\|r_{o4})C_{gs6}} \\ \omega_{p2} = \frac{1}{(r_{o6}\|r_{o7})C_L} \end{cases} \tag{4}$$

Эндээс давамгайлах туйл болох ω_{p2}-ыг авч үзсэн бөгөөд Зураг 3-д дүрслэн харуулсан. Давамгайлах туйлаас их давтамжид:

$$\omega_0 \approx \frac{G_{m1}(r_{o2} \parallel r_{o4})G_{m6}}{C_L} \tag{5}$$

буюу давтамжийн хариу фазын хазайлттай холбоотойгоор гүйцэд нэгж өсгөлтийн фазын нүүлт үүсэхээс сэргийлэхийн тулд $\omega_{p1} \gg \omega_0$ байх шаардлагатай.

Өсгөгчийн туйлыг салгахын тулд Зураг 4-д үзүүлсэн компенсацлагч конденсаторыг ашигласан.

Компенсацлагч конденсатор ашиглан хоёр туйлыг салгаснаар давамгайлагч туйлын давтамжийг багасган паразит дамтамжийн туйлыг ихэсгэнэ. Ингэснээр нэгж өсгөлтийг ихэсгэхээс гадна фазын зөрөөны запас сайжирна.

III. ЗАГВАРЧЛАЛ БА ХЭРЭГЖИЛТ

Судалгааны ажлын хүрээнд 90нм CMOS технологи ашиглан хоёр шатлалт үйлдлийн өсгөгчийн схемийн загварчлал, симуляци болон гүйцэтгэлийн шинжилгээг Cadence Virtuoso программын орчинд хийж гүйцэтгэлээ. Судалгаагаар бага чадлын хэрэглээнд тохирох,

өндөр өсгөлттэй, фазын хувьд тогтвортой ажиллагаатай өсгөгчийн оновчтой бүтцийг боловсруулан, гүйцэтгэлийн үндсэн параметрүүдийг тодорхойлох. Хоёр шатлалт үйлдлийн өсгөгчийн загварыг хийхэд транзисторын өргөн (W) ба урт (L) нь гүйдэл, өсгөлт, зурвасын өргөн, чадлын хэрэглээ зэрэг олон параметрт нөлөөлдөг тул оновчтой сонголт хийх нь чухал.

Транзисторын W/L харьцааг сонгох

a) Дамжуулах чадвар

$$g_m = \frac{2I_D}{V_{ov}} = \mu_n C_{ox} \frac{W}{L} V_{ov} \tag{4}$$

Тайлбар:

- g_m — Дамжуулал
- I_D — Сувгийн гүйдэл
- $V_{ov} = V_{GS} - V_{th}$ — Overdrive хүчдэл
- μ_n — n электронуудын хөдөлгөөны нягт
- C_{ox} — Оксидын багтаамж
- W/L — Транзисторын өргөн ба уртын харьцаа

b) Гаралтын эсэргүүцэл

$$r_o = \frac{1}{\lambda I_D} \tag{5}$$

c) DC өсгөлт

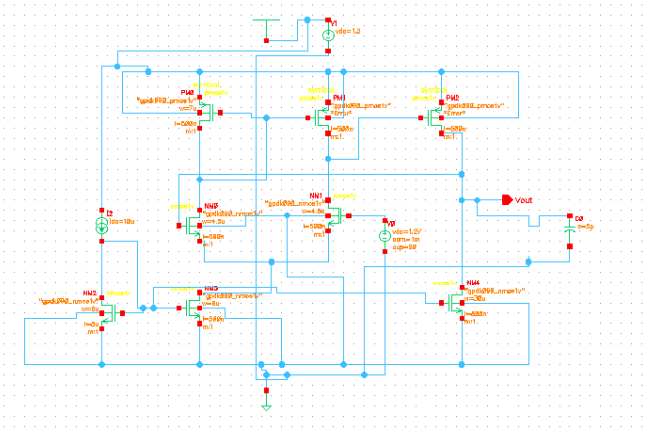
$$A_v = g_{m1} \cdot r_{o2} \cdot g_{m6} \cdot r_{o6} \tag{6}$$

→ Өсгөлт их байлгахын тулд g_m болон r_o их байх транзисторын хэмжээг (W/L) оновчтой сонгох шаардлагатай. Хоёр шаттай үйлдлийн өсгөгчийн загварт ашигласан үндсэн параметрүүдийг Хүснэгт II-д үзүүлсэн. Хоёр шатлалтай үйлдлийн өсгөгчийн бүтцийг загварчлахдаа дараах параметрүүдийг ашигласан бөгөөд хэлхээний ажиллагааг оновчтой байлгах, шаардлагатай нөхцөлүүдийг тооцоолон утгуудыг өгсөн.

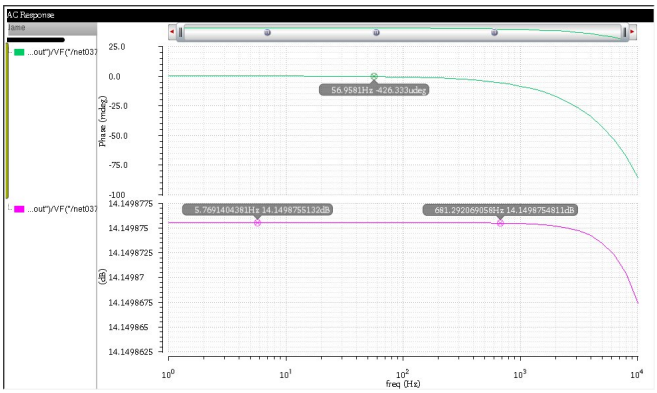
Транзисторын дугаар	(W/L)	Элемент	Утга
(M)1,2	4.5μm / 0.5μm	Оролтын гүйдэл	14 μA
(M)3,4	7μm / 0.5μm	Конденсатор (C_L)	2 pF
(M)5,8	6μm / 0.5μm	Конденсатор (C_C)	600 fF
(M)6	75μm / 0.5μm	Нулинг резистор	2.7 k
(M)7	37.5μm / 0.5μm	Тэжээлийн хүчдэл	1.2 V

Хүснэгт II
Элементийн W/L харьцаа болон утгууд

Загварын бүх параметрүүдийн утгыг тооцоолон Зураг 5-д үзүүлсэн хоёр шатлалт үйлдлийн өсгөгчийн хэлхээг 90 нм технологи ашиглан Cadence программ дээр загварчилан угсарсан.



Зураг 5. Хоёр шатлалт үйлдлийн өсгөгчийн загвар



Зураг 6. Өсгөлт болон фазын муж давтамжийн хамаарал

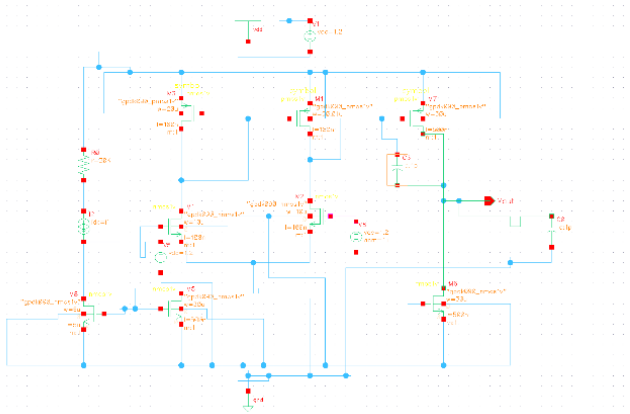
Компенсацийн конденсаторгүй хоёр шатлалт өсгөгч нь өсгөлтийг нэмэгдүүлж 14,14дб хүргэж байгаа хэдий ч давтамжийн муж хязгаарлагдаж байгааг симуляцийн үр дүн Зураг 6-д харуулж байна. Үүний үр дүнд хэлхээний фазын өөрчлөлтийг хэмжихэд хүндрэл үүсэж, өсгөлтийн 0 дб мужид фазын шилжилтийн утгыг харах боломжгүй болж байна.

Ийм нөхцөлд хэлхээнд хэлбэлзэл үүсэх, тогтворгүй ажиллах эрсдэлтэй. Компенсацийн конденсатор нэмж холбосноор тогвортой ажиллуулах боломжтой нэмэлт конденсатор ашигласан хэлхээний схемийг Зураг [7]-т харууллаа.

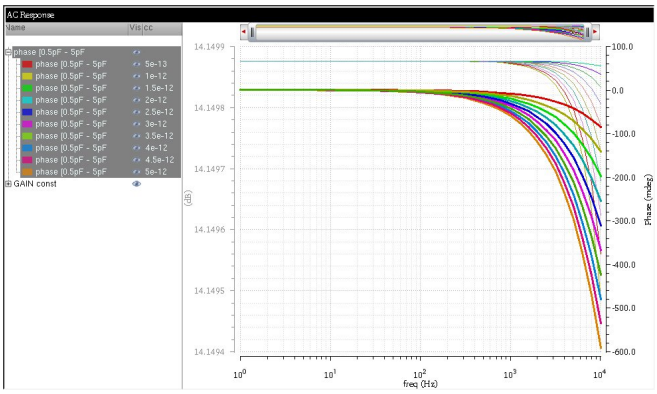
Компенсацийн конденсаторын утгыг 0.5 pF–аас 5 pF-ийн хооронд шаталсан байдлаар өөрчлөн, хэлхээний давтамжийн муж хэрхэн нэмэгдүүлж байгааг AC шинжилгээгээр судласныг зураг 8-д үзүүлэв. Симуляцийн үр дүнд конденсаторын утга ихсэх тусам давтамжийн муж нэмэгдэж буйг зураг 8-д харуулсан.

Өсгөлтийг нэмэгдүүлэх

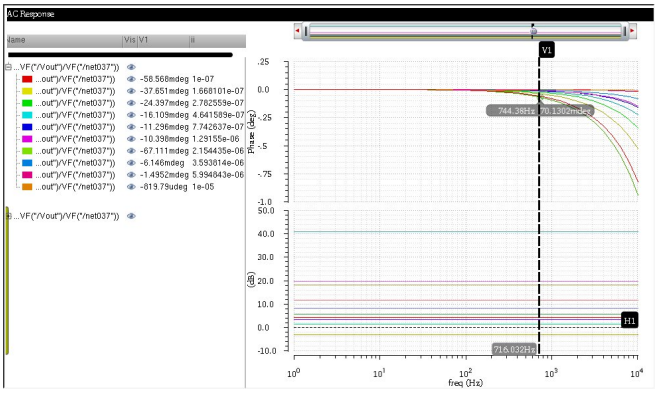
Оролтын Ibias -ийн гүйдлийг 100nA – 10uA мужид шаталсан байдлаар өөрчлөн өсгөлтийг 40дб утга хүртэл өсгөж 0-дб-д харгалзах фазын өөрчлөлтийг хэмжсэн утгыг зураг 8-д харуулав .



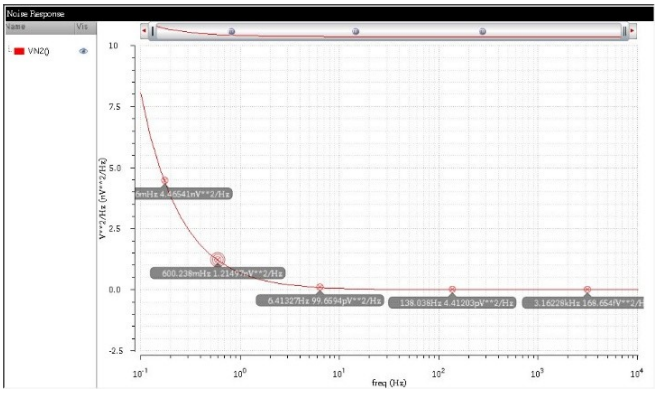
Зураг 7. Компенсацийн конденсатортой хэлхээ



Зураг 8. Өсгөлт болон фазын муж давтамжийн хамаарал



Зураг 9. Өсгөлт болон фазын муж давтамжийн хамаарал



Зураг 10. Шуугианы нөлөө болон давтамжийн хамаарал

Шилжилтийн гүйдлийг 100nA – 10uA мужид нэмэгдүүлэх нь хоёр шатлалт үйлдлийн өсгөгчийн өсгөлт, зурвасын өргөн, фазын өөрчлөлт зэрэг үзүүлэлтүүдийг сайжруулах, 0дб мужид фазын өөрчлөлтийг хэмжиж харах боломжтой болж байна. Өсгөлтийн 0дб утгад давтамжийн 716,032Hz мужид фазын шилжилт 70 байгаа нь фазын маржин 110 болж тогтвортой хэлхээ болохыг харуулж байна. Өндөр фазын маржинтай үед хэлхээний хурд удааширч, дамжуулах зурвасын өргөн багасах тул шилжилтийн гүйдлийн утгыг нэмэгдүүлэхдээ чадлын хязгаарлалт, температурын нөлөө, мөн компенсацийн нөхцөл зэргийг хамтад нь тооцож, оновчтой тохируулах шаардлагатай .

Шуугианыг хэмжих

Хоёр шатлалт өсгөгчийн гаралтын шуугианы спектрийн нягтшил (noise spectral density)-ийг харуулж байгаа бөгөөд босоо тэнхлэгийн дагуу гаралт, оролтын дохионы харьцаа хэвтээ тэнхлэг дагуу 0-10кГц давтамж дээрх шуугианыг харуулав.

Иймээс хэлхээний тогтвортой байдал болон гүйцэтгэлийг хангахын тулд конденсатор болон резисторын утгыг зөв сонгох нь чухал

Дүгнэлт

Энэхүү судалгаагаар биоангаахын хэрэглээнд тохирсон, бага чадал зарцуулалттай, өндөр өсгөлттэй хоёр шатлалт CMOS үйлдлийн өсгөгчийн дизайн, хэрэгжүүлэлтийг Cadence Virtuoso орчинд 90-нм технологиор гүйцэтгэлээ. Ингэхдээ биоанагаахын дохионы хүрээнд давтамжийн мужийг 1Hz -10Hz, аналог хүчдэлийн утгыг 1mV -оор авч транзисторын урт өргөний хэмжээг томъёоны дагуу тооцоолон сонгож гаралтын өсгөлт, фазын шилжилт, шуугианы нөлөөг графикаар харж тооцооллоо. Тус загварчилсан загвар нь 40.1 dB өсгөлт, 716.032 Hz зурвасын өргөн, 238.63 μW чадал зарцуулалт, 110° фазын маржинтай , шуугианы нөлөө бага байснаар ECG, EEG, EMG зэрэг сул дохионд тохирохыг баталсан. Компенсацын конденсаторын утга, шилжилтийн гүйдлийг оновчтой сонгосноор тогтвортой ажиллагаа, зурвасын өргөн зэрэг гол үзүүлэлтүүдийг

сайжруулсан. Симуляцийн үр дүнд өсгөлт, CMRR, slew rate, чадлын зарцуулалт зэрэг параметрүүдийг сайжруулж чадсан. Цаашид транзисторын хэмжээсийг оновчтой сонгох замаар гүйцэтгэлийг улам сайжруулах боломжтой.

Ашигласан ном

[1] B. Razavi, Design of Analog CMOS Integrated Circuits, McGraw-Hill, 2000.

[2] D. Johns and K. Martin, Analog Integrated Circuit Design, John Wiley & Sons, 1997.

[3] W. Sansen, Analog Design Essentials, Springer, 2006.

[4] M. Ismail and T. Fiez, Analog VLSI Signal and Information Processing, McGraw-Hill, 1994.

[5] R. R. Harrison and C. Charles, "A low-power low-noise CMOS amplifier for neural recording applications," IEEE Journal of Solid-State Circuits, vol. 38, no. 6, pp. 958–965, June 2003.

[6] K. F. Laker and W. M. C. Sansen, Design of Analog Integrated Circuits and Systems, McGraw-Hill, 1994.

[7] K. Abdelhalim, "Low-Power CMOS Amplifiers for Biosignal Acquisition," PhD Thesis, University of Toronto, 2011.

[8] IEEE Std 1801-2015, IEEE Standard for Design and Verification of Low Power Integrated Circuits, IEEE, 2015.

[9] M. I. Dessouky and A. A. Kaiser, "Very low-voltage digital-audio delta–sigma modulator with 88-dB dynamic range using local switch bootstrapping," IEEE Journal of Solid-State Circuits, vol. 36, no. 3, pp. 349–355, 2001.

[10] W. Wattanapanitch and R. Sarpeshkar, "A Low-Power 32-Channel Digitizing Chip for Neural Recording," IEEE Transactions on Biomedical Circuits and Systems, vol. 1, no. 2, pp. 136–147, 2007.

[11] A. Chandrakasan, S. Sheng, and R. W. Brodersen, "Low-power CMOS digital design," IEEE Journal of Solid-State Circuits, vol. 27, no. 4, pp. 473–484, 1992.

[12] M. Yin, M. Ghovanloo, "A Flexible 32-Channel Analog Front-End for Recording and Stimulation in Cochlear Implants and Deep Brain Stimulators," IEEE Transactions on Biomedical Circuits and Systems, vol. 3, no. 6, pp. 441–449, 2009.

[13] M.-S. Chae, W. Liu, and M. Sivaprakasam, "Design of a Low-Power Stimulator ASIC for Wireless Neural Prosthetics," IEEE Transactions on Biomedical Circuits and Systems, vol. 3, no. 6, pp. 404–412, 2009.

[14] K. Hassan, T. Delbruck, "Low-power low-noise CMOS amplifier for bio-potential recording applications," IEEE Transactions on Biomedical Circuits and Systems, vol. 9, no. 2, pp. 166–178, 2015.

[15] N. Kobayashi, T. Satoh, "A low-power low-noise amplifier with novel noise cancellation scheme for biomedical signal acquisition," IEEE Asian Solid-State Circuits Conference, pp. 473–476, 2005.

Машин сургалтын аргаар экскаваторын шанаганы шүд илрүүлж, хүдрийн бүхэллэг тодорхойлох судалгаа

А.Түвшинжаргал
ШУТИС, "Эрдэнэт цогцолбор" дээд сургууль
Монгол улс, Эрдэнэт хот
tuvshinjargal@erdenetis.edu.mn

Д.Энхбаатар
ШУТИС, "Эрдэнэт цогцолбор" дээд сургууль
Монгол улс, Эрдэнэт хот
enkhbaatar@erdenetis.edu.mn

Б.Гэрэлмаа
МУИС, Мэдээллийн технологи, электроникийн сургууль
Монгол улс, Улаанбаатар хот
gerelmaa.b@num.edu.mn

Б.Ганбат
МУИС, Мэдээллийн технологи, электроникийн сургууль
Монгол улс, Улаанбаатар хот
ganbatb@num.edu.mn

***Хураангуй*—"Эрдэнэт үйлдвэр" ТӨҮГ -ын Ил уурхайн гол малталтын төхөөрөмж бол шанага юм. Малталтын явцад шанаганы шүд хүдэрт удаан хугацаагаар шууд нөлөөлснөөр шанаганы шүд нь хүчтэй цохилт, үрэлт, бусад хүчинд өртөж, улмаар шанаганы шүд нь шүдний сууринаас гэнэт сулардаг, зарим тохиолдолд бүр хугардаг ба тус шүдийг хүдрийн хамт бутлуурт буулгасны дараа энэ нь бутлуурыг эвдрэлд хүргэж, уурхайн бутлах үйлдвэрлэлийн шугамд бүхэлд нь нөлөөлдөг. эдийн засгийн, хүний болон материаллаг нөөцийн үр ашгийн алдагдал үүсэх эрсдэлтэй. Шанаганы шүд нь өндөр манганы ган эсвэл хайлштай гангаар хийгдсэн байдаг тул хатуулаг нь хүдэр болон бусад материалынхаас хамаагүй өндөр байдаг. Энэхүү судалгааны ажлаар Экскаваторын шанаганы дүрсийг цуглуулж өгөгдлийн багц бэлтгэн YOLOX болон YOLO-NAS загваруудыг сургаж харьцуулан шанаганы шүдийг илрүүлсэн. Мөн шанаган дээрх хүдрийн дүрсэд сегментчлэл хийдэг SAM загвар болон дүрсээс ирмэг илрүүлэх DexiNED загвар ашиглан бүхэллэгийн хэмжээг ойролцоогоор тодорхойлсон. Шүдний уналтыг илрүүлэх загварууд 87% -тай илрүүлсэн ба хүдрийн бүхэллэгийг SAM загвараар дүрсээс чулууг сегменчлэх замаар тодорхойлох илүү үр дүнтэй байна.**

Түлхүүр үг— Шанаганы шүд, хүдрийн бүхэллэг, YOLOX, сегментчлэл

I. Удиртгал

Шинжлэх ухаан, технологийн тасралтгүй хөгжлийг дагаад ашигт малтмалын олборлолт илүү нарийн, автоматжуулсан, ухаалаг чиглэл рүү хөгжиж байна [1]. Уурхайн үйлдвэрлэлийн гол малталтын төхөөрөмж бол шанага юм. Малталтын явцад шанаганы шүд хүдэрт удаан хугацаагаар шууд нөлөөлснөөр шанаганы шүд нь хүчтэй цохилт, үрэлт, бусад хүчинд өртөж, улмаар шанаганы шүд нь шүдний сууринаас гэнэт сулардаг, зарим тохиолдолд бүр хугардаг ба ингэснээр эдийн засгийн, хүний болон материаллаг нөөцийн үр ашгийн алдагдал үүсэх эрсдэлтэй.

Шанаганы шүд гэмтэх нь малталтын ажлын явцад шанаганы үрэлтийн эсэргүүцлийг нэмэгдүүлж, шанагыг ухахад хүндрэл учруулж, улмаар малталтын ажлын үр ашигт нөлөөлж, шанаганы элэгдлийг хурдасгах болно [8].

Шанаганы шүд нь өндөр манганы ган эсвэл хайлштай гангаар хийгдсэн байдаг тул хатуулаг нь хүдэр болон бусад материалынхаас хамаагүй өндөр байдаг. Хагарсан шанаганы шүдийг хүдрийн хамт бутлуурт буулгасны дараа шанаганы шүд бутлуурын эвдрэлд хүргэж, уурхайн бутлах үйлдвэрлэлийн шугамд бүхэлд нь нөлөөлнө [11]. Шанаганы шүд суларч, хугарч унан тээвэрлэгдэж анхдагч бутлуурт орж гацаах эрсдэлтэй байдаг.

Эрдэнэтийн овоо ордын ил уурхай нь гүнзгийрэхийн хэрээр чулуулгийн физик механик шинж чанар, усжилт зэрэг нь өөрчлөгдөж түүнийг дагаад технологийн өөрчлөлтүүд гарсаар байна. Мөн уурхайн ачилт, тээвэрлэлт, буулгалттай холбоотой хөлдүүс чулуулаг, машин механизмын эд анги анхдагч бутлуурт орж гацаах зэрэг эрсдэлүүд үүсдэг тул түүнийг хиймэл оюун ашиглан оворлог чулуулгийг таних, экскаваторын шүд болон бусад зүйлс унасныг мэдэх шаардлагууд бий болж байна.

Сүүлийн жилүүдэд шанаганы шүдний алдагдлыг гүн гүнзгий сургалтад суурилсан аргаар илрүүлэх судалгаа хийгдэж байгаа бөгөөд энэ төрлийн дэвшилтэт технологи нь эвдрэлээс урьдчилан сэргийлэхэд үр дүнтэй байна [5] [7].

3D принтер ашиглан хиймэл экскаваторын шанага үүсгэж датасет бэлтгэн VGG16 загварыг сургаж шүдийг 91% илрүүлсэн [6].

Энэхүү ажлын судалгааны ажлын хүрээнд дараах асуудлыг шийдвэрлэхийг зорьсон. Үүнд:

1. Экскаваторын шүд унасах эсэхийг өдөр болон шөнийн аль ч цагт бодит хугацаанд илрүүлж операторт мэдэгдэх
2. Хүдрийн бүхэллэгийг тодорхойлох

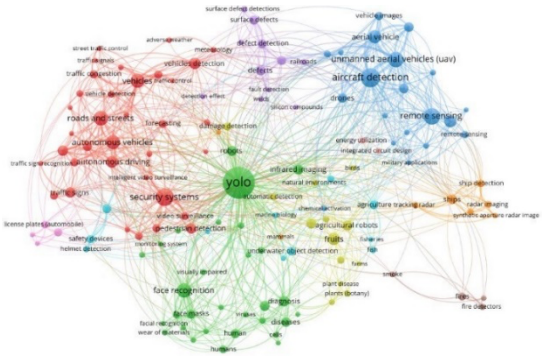
II. Экскаваторын шүд илрүүлэх, хүдрийн бүхэллэг боломжтой загварууд

A. YOLO-NAS загвар

Scopus-д "YOLO" гэсэн үгийг гарчигтаа агуулсан бүх хэвлэгдсэн материалуудын ишлэлийн сүлжээний дүрслэлийг үзүүлж, объект илрүүлэлтэй холбоотой түлхүүр үгс дээр онцгой анхаарал хандуулж гаргасан диаграммыг [9] судалгааны ажилдаа онцлон дурдсан байна. Зураг 1 -д үзүүлсэн диаграммаас харахад объект

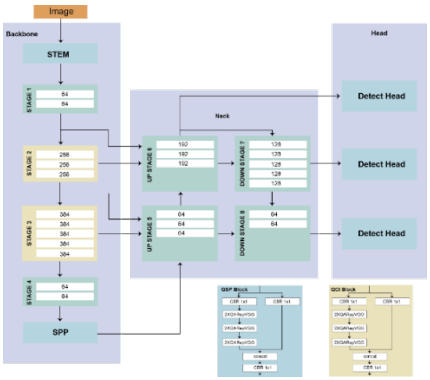
XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

илрүүлэх, гэмтэл тодорхойлоход YOLO загварыг ашиглаж байгаа нь харагдаж байна.



Зураг 1. YOLO-ийн үндсэн хэрэглээний библиометрийн диаграмм

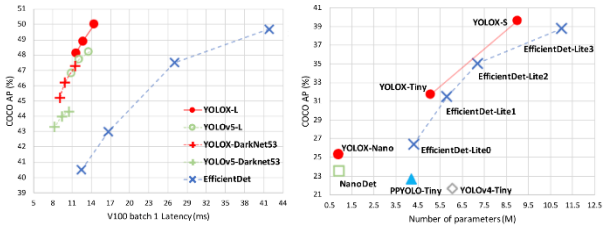
Энэ фреймворкоос санаа авч, бид шанаганы шүд илрүүлэх асуудлыг шийдвэрлэхэд YOLO-NAS архитектурыг ашигласан. YOLO-NAS нь YOLO фреймворкийн хөнгөн, үр ашигтай хувилбар бөгөөд жижиг объект илрүүлэх, байрлал тодорхойлох, performance-per-compute ratio-г сайжруулах, бодит хугацаанд edge-device дээр хэрэглэхэд тохиромжтой ба эх код нь нээлттэй бөгөөд судалгааны зориулалтаар ашиглах боломжтой. YOLO-NAS загварын бүтэцийг Зураг 2-д харуулав.



Зураг 2. YOLO-NAS загвар

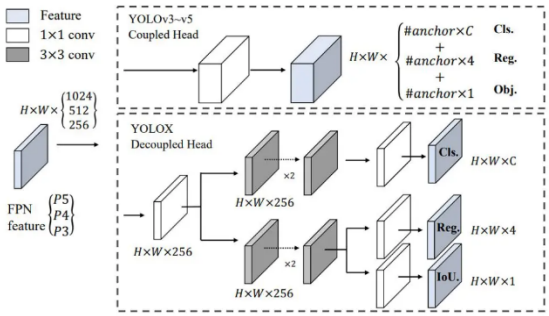
B. YOLOX загвар

2021 онд [4] нь анкор ашигладаггүй, төрөл бүрийн хэмжээтэй объектын түүврийг динамикаар тааруулдаг YOLOX-ийн өндөр гүйцэтгэлтэй объект илрүүлэх загварыг санал болгосон. YOLO-ийн өмнөх хувилбарт загварын decoupling head-ийг нэгтгэсэн байдаг байсан. Өөрөөр хэлбэл ангилал, регрессийг 1x1 конволюцаар хэрэгжүүлсэн байдаг. YOLOX нь загвар үүнийг сөргөөр нөлөөлсөн гэж үзэн YOLOX-д YoloHead-ийг хоёр хэсэгт хуваах тусад нь хэрэгжүүлсэн ба ба эцэст нь таамаглах үе шатанд нэгтгэсэн. YOLOv3 ба YOLOv4-тэй харьцуулахад YOLOX загварын илрүүлэлтийн нарийвчлал, хурд сайжирч, шууд гаралт хүртэл илүү уян хатан, загварын хэмжээ, санах ой нь эмбеддэд төхөөрөмжүүдэд илүү тохиромжтой болсон. Гэсэн ч YOLOX алгоритмын ерөнхий илрүүлэлтийн гүйцэтгэлийг цаашид сайжруулах шаардлагатай байна гэж дүгнэсэн байдаг. Зураг 3-д YOLOX-ийг бусад хувилбаруудтай харьцуулсан зургийг харуулав. Зургаас харахад YOLOX-ийн параметрийн тоо бага бөгөөд AP (average precision) хувилбараасаа хамаараад бусад хувилбаруудаас өндөр байгаа нь харагдаж байна.



Зураг 3. YOLO болон EfficientDet-Lite загваруудын дундаж precision

Энэ ажил эмбеддэд төхөөрөмжүүдэд тохиромжтой, илүү хурдан болон нарийвчлалттай гэдэг үзүүлэлт дээр үндэслэн YOLOX загварыг ашиглан шанаганы шүдийг илрүүлэх ажлыг хийсэн. YOLOX-ийн архитектурыг Зураг 4-д харуулав.



Зураг 4. YOLOX -ийн архитектур

YOLOX нь бусад хувилбаруудтай шигээ дотроо олон хувилбаруудтай. YOLOX-s, YOLOX-m, YOLOX-l, YOLOX-x гэх мэт стандарт хувилбаруудаас гадна light буюу хялбаршуулсан хувилбарууд мөн байдаг. YOLOX-ийн эдгээр хувилбарууд нь загварын хэмжээ болон параметрийн тоон хувьд ялгаатай ба эдгээр ялгаанаас үүдэн ажиллах хурд, AP зэрэг нь өөр өөр байдаг.

C. DexiNED загвар

DexiNed нь CNN (convolutional neural network)-д суурилсан ирмэг илрүүлэх алгоритм юм [2]. DexiNed-ийн архитектурт Xception болон ResNet дээр үндэслэсэн. Бүтцийн үндсэн блок бүр нь дэд блокуудаас бүрдэх ба өмнөх үндсэн блоктойгоо холбогдсон байдаг. DexiNed-ийн сургах явц нь гар ажиллагаа багатай тасралтгүй явагддаг [3]. [2] дээрх судалгааны ажил нь урьдчилан бэлтгэгдсэн датасет байхгүй ч DexiNed нь хамгийн сүүлийн үеийн ирмэгийг илрүүлэх алгоритмуудаас давсан үзүүлэлттэй байгааг онцолсон байна. DexiNed-ийн давуу талыг ашиглан энэ ажилд чулууны ирмэгийн зургийг урьдчилан сургалгүй таамагласан болно.

D. SAM загвар

Segment Anything Model (SAM) нь Meta AI-ийн FAIR (Fundamental AI Research) лабораториос боловсруулсан дэвшилтэт дүрс сегментчилэх загвар юм. 2023 оны 4-р сард танилцуулсан уг загвар нь олон төрлийн оруулгад (prompt) тохируулах боломжтой бөгөөд цэгүүд, хязгаарлах хайрцаг, эсвэл чөлөөт хэлбэрийн текст зэрэг оруулгын дагуу нарийвчилсан сегментчилэх маск үүсгэдэг [10].

SAM-ийн гол онцлог нь ерөнхийлөх чадвар бөгөөд өмнө нь сургалт хийж байгаагүй объектуудыг таньж, сегментчилэх боломжтой. Энэ нь 11 сая гаруй зураг, 1

тэрбум гаруй маскаас бүрдсэн SA-1B өгөгдлийн багц дээр сургалт хийснээр боломжтой болсон [10].

III. Арга зүй

A. Датасет цуглуулах

Зураг 5. Экскаваторын сум дээр байруулсан камерийн харагдах байдал

Датасетээ цуглуулахдаа бодит орчноос цуглуулсан. DAHUA-ийн DH-IPC-HFW2439M-AS-LED-B-S2 камерийг сонгон авч экскаваторын сум дээр шанага руу харуулан байрлуулсан. Камерийн техникийн үзүүлэлтийг Хүснэгт 1-д харуулав.

Камерыг дараах шаардлагад нийцүүлэн сонгосон:

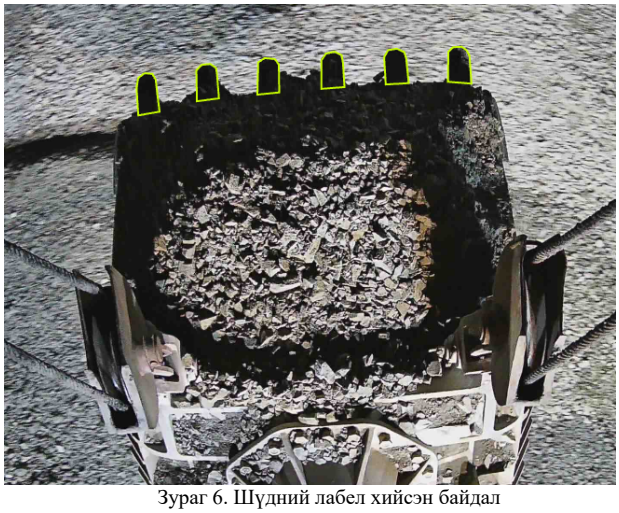
- 12В-ын тогтмол гүйдлийн тэжээлээр тэжээгддэг ба UTP кабелиар дүрс болон тэжээлээ давхар дамжуулдаг
- IP67 стандарттай тоос шороо, чийгийн хамгаалалттай тул гадна орчинд өндөрт байрлуулахад тохиромжтой
- Овор хэмжээ бага
- Бусад камертай харьцуулахад үнэ хямд.
- Суурин дээр нь эргэдэг тул өнцгөө тохируулах боломжтой
- Дүрсний нягтаршил маш сайн(4MP) – Бодит суурилуулсан байдлыг Зураг 2-с харна уу.
- Шөнийн хараа нь 40метр тул шөнө ч экскаваторын шанагагыг тасралтгүйгээр харах боломжтой.
- Санах ой 256GB тул 7 хоногийн өдөр, шөнийн бичлэгийг тасралтгүй хийх боломжтой ба хиймэл оюун ухаан сургах зурган датасет бий болгоно.

Хүснэгт 1. Экскаваторт суурилуулсан камерын үзүүлэлт

Камерын үзүүлэлт	Дэлгэрэнгүй мэдээлэл
Загвар	DH-IPC-HFW2439M-AS-LED-B-S2
Дүрс мэдрэгч	4MP, 1/3" CMOS өндөр чанартай зураг
Дүрс шахалт	ROI, SMART H.264+/H.265+
Дүрсний нягтаршил	4MP (2560 × 1440) @ 25/30 fps
Шөнийн хараа	40 m
WDR	Суурин дээрээ эргэнэ, WDR, 3D NR, HLC, BLC, усны хамгаалалттай
Илрүүлэг (Detection)	Хөдөлгөөн илрүүлнэ, video tampering, scene
Intelligent detection	Intrusion, tripwire
Санах ой	256 GB Micro SD card
Дохиолол	1 in, 1 out; audio: 1 in, 1 out
Хамгаалалт	IP67 protection
Тэжээл	12V DC/PoE power supply; 12V DC power output, max. current 165mA

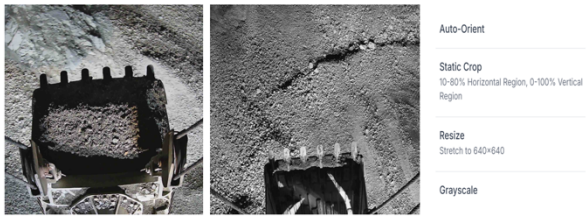
B. Шанаганы дүрсэд боловсруулалт хийх үйл явц

Нийт 6356 зураг цуглуулсан ба Зураг 6-д үзүүлсэн байдлаар шүдний байршлыг хэлбэрийн дагуу зурж лабел хийсэн.



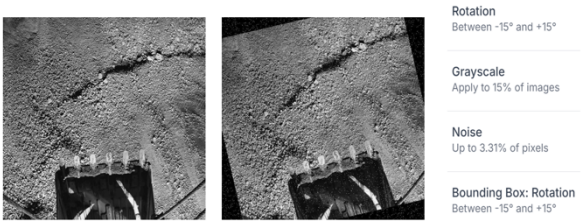
Зураг 6. Шүдний лабел хийсэн байдал

6356 зургийг зөвхөн шанага харагдах хэсэг буюу зүүн талаас 10% баруун талаас 20% -г тус тус тайрсан. Машин сургалтын загвараа сургахад зориулж зургийн хэмжээг 640x640 болгон бууруулсан.



Зураг 7. Камерийн дүрсийг дүрс боловсруулалтын аргаар өөрчилсөн байдал

Мөн зургуудад шуугиан нэмэх, харагдах өнцгийг өөрчлөх замаар зургийн датасетийг 33378 болтол нэмэгдүүлсэн.



Зураг 8. Зургийг өөрчлөх замаар өгөгдлийн багцийг нэмэгдүүлсэн байдал

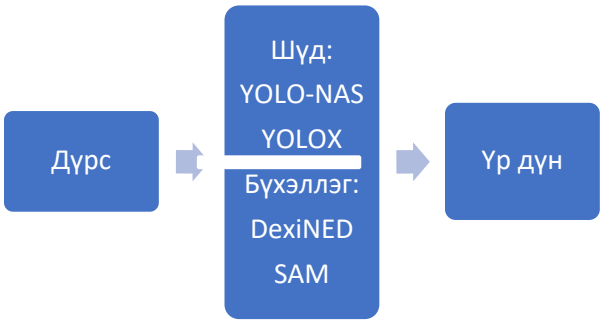
Зураг 8-д харуулсны дагуу зургийн харагдах өнцгийг -15 болон +15 градусаар эргүүлж шинэ зургууд үүсгэсэн. Нийт зургийн 15% -г саарал зураг болгон өөрчилсөн. Мөн нэг зурагт харгалзах цэгийн 3.31% шуугиан буюу

өөр өнгөтэй цэгээр сольсон. Шүдний байршлыг таниулсан хэсгийг -15 болон +15 градусаар эргүүлж датасетийг тус тус нэмэгдүүлсэн.

С.Шанаганы шүд илрүүлэх болон хүдрийн бүхэллэг тодорхойлоход ашигласан загварууд

Судалгааны ажлын хүрээнд YOLOX, YOLO-NAS гэсэн 2 загвар дээр тулгуурлан шанаганы шүд илрүүлсэн.

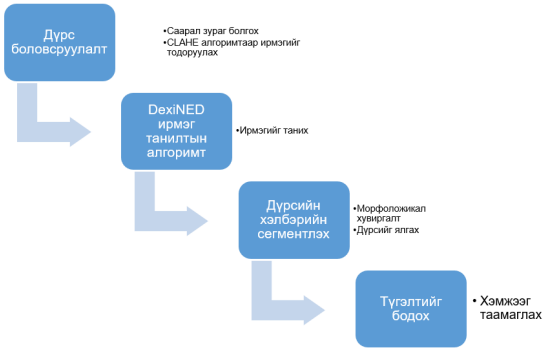
Хүдрийн бүхэллэг илрүүлэхэд ирмэг илрүүлэх DexiNED загвар болон обьектын сегментчилэх SAM загваруудыг ашигласан. Зураг 9-д ашигласан загваруудыг харуулав.



Зураг 9. Шүд ирүүлэх болон хүдрийн бүхэллэг тодорхойлоход ашигласан гүн сургалтын загварууд

Өөрсдийн цуглуулсан дүрсийн сан дээр тулгуурлан зураг 8-д үзүүлсний дагуу машин сургалтын загваруудыг сонгож сургасан.

D.Шанаган дээрх хүдрийн бүхэллэг илрүүлэх аргачлал

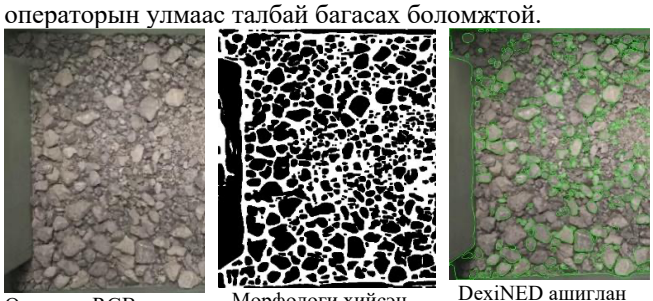


Зураг 10. Ирмэг таних загвар ашиглан хүдрийн бүхэллэг тодорхойлох аргачлал

Ирмэг таних DexiNED загвар ашиглан хүдрийн бүхэллэг тодорхойлох аргыг [11] -д санал болгосны дагуу туршилт хийсэн.

Энэ арга нь Зураг 10-д харуулсны дагуу бүхэллэгийн хэмжээг тодорхойлох юм. DexiNed-ийн таамагласан ирмэгийн зургаас чулууг сегментлэх үйл явцын үр дүнг Зураг 11-д үзүүлэв.

Эхлээд ирмэгийн зураг дээр хоёртын хувиргалт хийсэн. OpenCV зураг дээр ажилладаг сангийн хоёртын хувиргалт хийх функцийн босго утгыг 180 гэж тохируулсан. Дараа нь хоёртын дүрс дээр морфологийн үйлдлийг ашигласан бөгөөд (3x3) хэмжээтэй кернелээр зургийг бүхэлд бодсон. Энэхүү морфологийн үйлдлийн дараа erosion хийсэн. OpenCV сангийн erosion функц нь объектуудын нийлсэн хил хязгаарыг арилгах зорилготой. Энэ процесс нь хөрш зэргэлдээх чулуулгийн хоорондох холболтыг багасгана. Үүнийг хийхдээ iteration = 1 гэж өгсөн ба Erosion хийсэн хоёртын зургийг Зураг 11-д үзүүлэв. Морфологийн операторын улмаас талбай багасах боломжтой.



Оролтын RGB зураг — Морфологи хийсэн зураг — DexiNED ашиглан чулууг таамаглас байдал

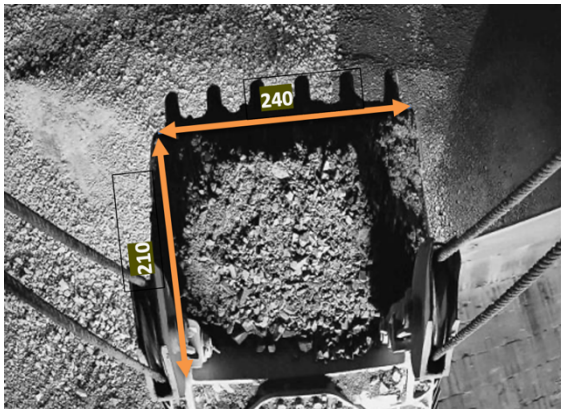
Зураг 11. Морфологи болон DexiNED ашиглан чулууг таамагласан үр дүн

Үүний дараа OpenCV-ийн “findContours” функцээр бүх битүү хүрээг олсон. Энд олсон бүх битүү хүрээ нь нэг чулуу биш байна. Олон чулууг нийлүүлэн битүү хүрээ гэж тодорхойлох боломжтой. Эндээс нэг чулуу гэдгийг танихын тулд OpenCV ‘ContourArea (Ac)’-г ашиглан битүү хүрээний пикселийн талбайг олно. Мөн битүү хүрээ бүрийн гүдгэр талбайг “ConvexHull (Ah)” функцээр тодорхойлно. Дараа нь гүдгэр их биеийн талбай ба контурын талбайн харьцааг ашиглан дан чулуу эсэхийг тогтоосон. Гүдгэр их биеийн талбай ба битүү хүрээн талбайн харьцааг дараах байдлаар бодно.

$$Rc = Ac/Ah$$

Дээрх тэгшитгэлээр бодсоны дараа R_C утгаас хамааруулан нэг чулуу эсэхийг тогтоосон. Бидний программ дээр энэхүү утгыг R_C = 0.8-аас их бөгөөд битүү хүрээний пикселийн талбайг 1500 бага бол нэг чулуу гэж үзсэн. Гэхдээ тухайн нөхцөлөөс (камерын үзүүлэлт, зургийн харьцаа, нарийвчлал) хамааруулан эдгээр утгыг тохируулах шаардлагатай.

Үүний дараа нэг битүү хүрээг нэг чулуу гэж үзэх ба нийт битүү хүрээг тоолон нийт илрүүлсэн чулууны тоог гаргах боломжтой. Мөн зургийн харьцаа болон шанаганы бодит хэмжээнээс хамааруулан битүү хүрээний хэмжээг тодорхойлж болно.



Зураг 12. Экскаваторын шанаганы хэмжээ

Бид бодолтыг хялбарчлах үүднээс бүхэллэгийг дугуй хэлбэртэй гэж үзэн хэмжээг тодорхойлсон. Хэрэв тойргийн талбай мэдэгдэж байвал тойргийн диаметр:

$$d = c/\pi$$

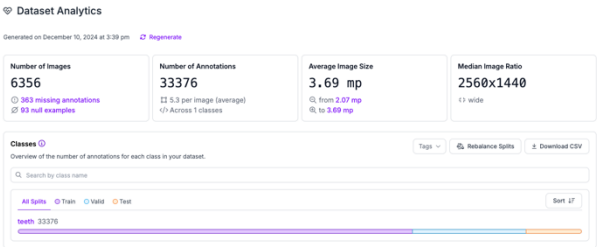
Энд олж байгаа d нь пикселээр гарах ба шанаганы бодит хэмжээ болон зургийн харьцаанаас хамааруулан бодит

хэмжээ рүү хөрвүүлэх боломжтой. Энэхүү тооцоолол нь ойролцоолсон тооцлоол. Учир нь бид чулууг дугуй дүрс гэж үзэн тооцоолж байгаа.

IV. ҮР ДҮН

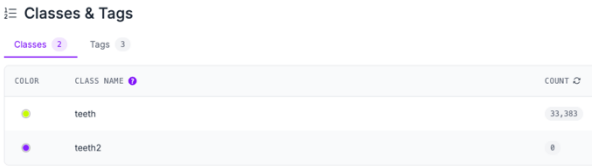
A. Экскаваторын шанаганы шүд, хүдрийн бүхэллэгийн датасет

Экскаваторын шанаганы 6356 зураг буюу дүрсийг Roboflow веб хэрэгслийг ашиглан таниулалт хийсэн. Зургийн дундаж хэмжээ нь 3.69 mp ба хамгийн бага нь 2.07 mp, их нь 3.69 mp, зургийн урт, өргөн нь 2560x1440 байхаар бэлтгэсэн.

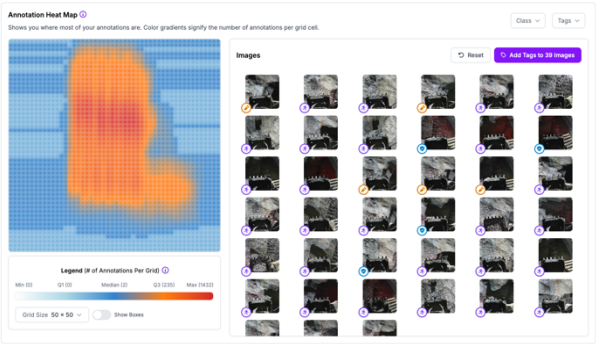


Зураг 13. Өгөгдлийн багцийн ерөнхий мэдээлэл.

Зургуудад байрлах шүдний байршлын Heat Map -г харуулж байна. Ихэнх шүднүүд зургийн төв хэсгээр дээшээ доошоо шилжсэн байна.

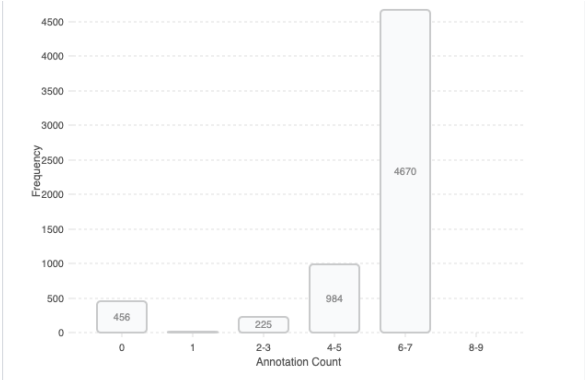


Зураг 14. Нийт зурагт хамаарах шүдний тоо



Зураг 15. Шүдний байршлын Heat Map

Судалгааны ажлын хүрээнд нэг класстай байхаар таниулалт хийж датасет үүсгэсэн. Нийт 6356 зурагт харгалзах 33383 шүд байгаа нь харагдаж байна.



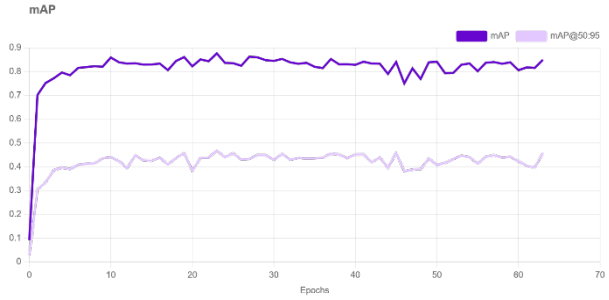
Зураг 16. Зураг бүрт харгалзах шүдний тоо

Судалгааны ажлын хүрээнд ЭКГ-12 болон ЭКГ-8 экскаваторуудын ажиллаж байх үеийн бичлэгээс хөрвүүлж авсан 8000 гаруй зураг нэг бүрд боловсруулалт хийж, экскаваторын шанаганы шүднүүдийг ширхэг бүрчлэн хүрээлж аннотац хийж нийт 33376 шүд тэмдэглэгдсэн 5901 зураг бүхий өгөгдлийн сан үүсгэсэн. Үүнээс 2104 хоосон шанаганы зураг, 626 бусад техник орсон зураг, 3171 хүдэр хутгасан зургууд байна. Хүдэр хутгасан 3171 зургийг хүдрийн бүхэллэгийг тодорхойлоход зургийн сан болгон ашигласан.

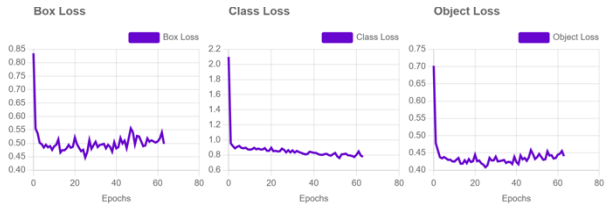
Зураг 17. Экскаваторын шанаганы шүдний тэмдэглэжүүлэлт

Өгөгдлийн сан үүсгэх ажлыг RoboFlow платформ дээр гүйцэтгэсэн, аннотац хийсэн зургуудыг машин сургалтын загварт зориулан сургалтын (Train), баталгаажуулалтын (Validation), болон тестийн (Test) өгөгдлүүдийг 70%, 20%, 10% харьцаатай хуваасан. Мөн өгөгдлийн урьдчилан боловсруулалт (Preprocessing) болон өгөгдлийн өргөжүүлэлт (Augmentantion)-ийг хийж 4 өөр датасет бий болгосон.

B. YOLO-NAS загвараар экскаваторын шанаганы дүрсээс шүдийг илрүүлсэн үр дүн



Зураг 18. YOLO NAS загвар сургах үеийн дундаж нарийвчлал



Зураг 19. YOLO NAS загвар дахин сургах үйл явц

Дээрх зургаас харвал YOLO-NAS загварыг сургах үйл явц 65 epoch -д сурч loss хангалттай буурч 87% -тай сурсан нь харагдаж байна.

C. YOLOX загвараар экскаваторын шанаганы дүрсээс шүдийг илрүүлсэн үр дүн

Энэхүү судалгааны ажилд бид эмбеддэд систем дээр хөгжүүлэх учир YOLOX-s загварыг сонгон авч хэрэгжүүлсэн. COCO датасет дээр урьдчилан сургасан загварыг өөрсдийн датасет дээр дахин сургасан.

YOLOX-s загварыг google colab дээр TPU ашиглан нийт 100 epoch-оор сургасан. Урьдчилан сургасан загвар ашиглаж буй учир сургахдаа нийт шөнийн камераар авсан 500 шанаганы зургаар сургасан ба датасетийн лабел нь VOC форматаар хийгдсэн.

Сургалтын 100 дахь epoch-ийн үр дүнг Зураг 19-д харуулав. Үр дүнгээс харахад map 0.5 нь 0.88 буюу сургалтын датасет дээр 88% танилттай сурсан байна.

```
---------------------------------------------------------------
Results computed with the **unofficial** Python eval code.
Results should be very close to the official MATLAB eval code.
Recompute with `./tools/reval.py --matlab ...` for your paper.
-- Thanks, The Management
---------------------------------------------------------------
Eval IoU : 0.55
Eval IoU : 0.60
Eval IoU : 0.65
Eval IoU : 0.70
Eval IoU : 0.75
Eval IoU : 0.80
Eval IoU : 0.85
Eval IoU : 0.90
Eval IoU : 0.95
---------------------------------------------------------------
map_5095: 0.5464013332077173
map_50: 0.8811678315141656
---------------------------------------------------------------
```

Зураг 20. Шүд илрүүлэх YOLOX загварын сурсан үр дүн

Энэхүү сурсан параметрээр validation датасет дээр загварыг шалгахад дундаж AP нь 87%-тай гарсан байна. Энэхүү үр дүнг Зураг 21-д харуулав.

```
2024-11-22 05:01:56 | INFO     | yolox.evaluators.voc_evaluator:144 - Evaluate in main process...
Writing teeth VOC results file
Eval IoU : 0.50
AP for teeth = 0.8790
Mean AP = 0.8790
~~~~~~~~
Results:
0.879
0.879
~~~~~~~~
```

Зураг 21. Шалгах өгөгдлөөр шалгалсан үр дүн

Загварыг сурган шалгасны дараа хамгийн сайн сурсан weight-ээр тестийн өгөгдөл (зураг) дээр шалгав. Тестийн өгөгдөл дээр шалгахдаа загварыг сургах явцад ашиглаагүй шинэ зураг оруулж шалгасан. Эдгээр зургууд нь шөнийн зургууд ба үр дүнг Зураг 22-д харуулав.

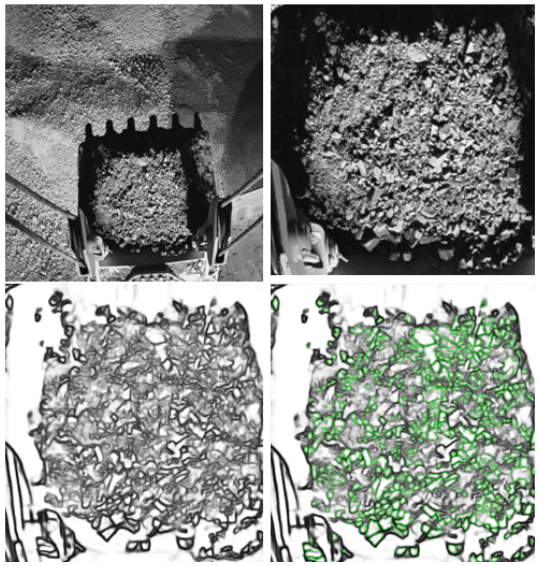
Зураг 22. Шөнийн зургаас YOLOX шүд таниж буй үр дүн

Дээрх үр дүнгүүдээс харахад дээр шүдийг 80% итгэлтэй таньж байна. Мөн өдөр гар утсаар авсан зураг дээр тест хийж үзсэн үр дүнг Зураг 23-д харуулав. Загварыг сургахдаа өдрийн зураг ашиглаагүй.

Зураг 23. Өдрийн зургаас YOLOX шүд таниж буй үр дүн

Өдрийн цагаар гар утсаар авсан зургийг загварт оруулж шалгахад 47% -аас дээш итгэлтэй илрүүлж байгаа нь дээрх зургаас харагдаж байна.

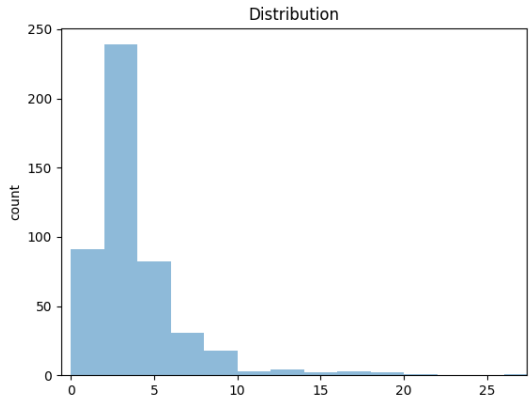
D. DexiNED Загвар ашиглан хүдрийн бүхэллэгийг тодорхойлсон үр дүн



Зураг 24. Бодит шанаганы дүрс дээр Хүдрийн бүхэллэг тодорхойлсон байдал

Эцсийн үр дүнгээс чулууны талбайг пикселээр тооцон олно. Үүний дараа чулууг дугуй гэж үзэн

чулууны хамгийн урт талын хэмжээг тооцон олж доорх хүдрийн бүхэллэгийн тархалтыг байгуулсан.



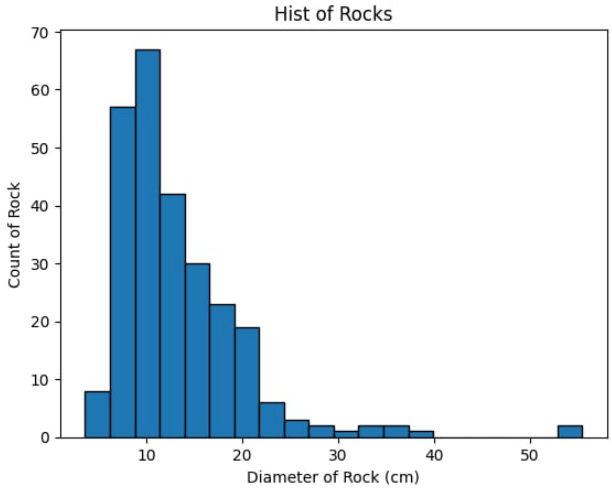
Зураг 25. Бодит шанаганы дүрс дээр Хүдрийн бүхэллэгийн тархалт

Дээрх үр дүнгээс харвал DexiNED загвараар хүдрийн бүхэллэг ч гэсэн. Бүх

E. SAM загвараар экскаваторын шанаганы дүрсээс хүдрийн бүхэллэгийн сегментчилэлийн үр дүн

Судалгааны ажлын хүрээнд дүрс сегментчилэх SAM загварыг сонгон зурагт харагдаж буй чулууг сегментчилэн улмаар чулууны хэмжээг зурагт харгалзах цэгийн тоогоор тооцоолон тархалтыг гаргасан.

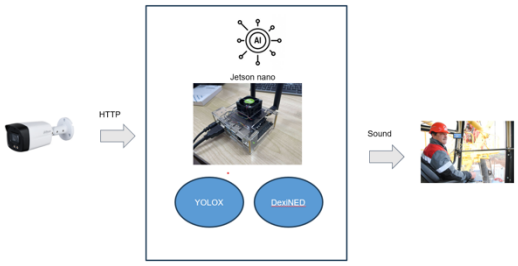
Зураг 26. SAM загвар ашиглан "Эрдэнэт үйлдвэр" ТӨҮГ -ын Ил уурхайн хүдрийн бүхэллэгийг сегментчилсэн байдал



Зураг 27. Хүдрийн бүхэллэгийн тархалт

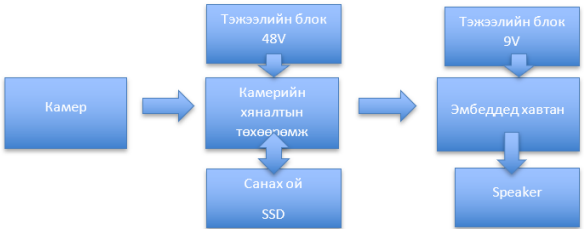
SAM загвараар хүдрийн бүхэллэгийг тодорхойлоход чулуу нэг бүрийг сегментчилж байгаа нь харагдаж байна.

Шийдэл-1: Тухайн мөчид экскаваторын оператор болон диспетчерт мэдээлэх системийг дараах байдлаар загварчлан туршсан.



Зураг 28. Шүдий тоо болон хүдрийн бүхэллэг тодорхойлж мэдээллэх системийн шийдэл

Экскаваторын шанаганы шүд илрүүлж тоолон операторт анхааруулах, хүдрийн бүхэллэгийг тодорхойлох эмбеддед системийг дараах байдлаар хөгжүүллээ.



Зураг 29. Системийн бүрэлдэхүүн хэсгүүдийн бүдүүвч

Jetson Nano хөгжүүлэлтийн хавтан нь санах ой болон оролт, гаралтын хөлтэй бөгөөд ethernet сүлжээнд холбогдож ажиллах боломжтой байдаг тул IP камертай шууд холбож ажиллуулах давуу талтай юм.

Дүгнэлт

"Эрдэнэт үйлдвэр" ТӨҮГ -ын ЭКГ-12 маркийн 26 дугаарын экскаваторт камер суурилуулан дүрс авч хиймэл ухааны машин сургах боломжтой нийт 6356 зургийн таниулалт бүхий өгөгдөлтэй болсон.

Экскаваторын шанаганы шүдийг илрүүлж тоолох хиймэл оюун ухааны YOLOX-s болон YOLO NAS загваруудыг сургаж, шалгах өгөгдлөөр шалгахад 87.9% -тай таньж илрүүлсэн.

Шанага дээрх хүдрийн бүхэллэгийг хиймэл оюун ухааны DexiNED, SAM загвар ашиглан ойролцоогоор тодорхойлсон. Энэ ажлын үр дүнд чулууны бүхэллэгийг ойролцоогоор тооцох боломжтой гэдэг нь харагдаж байна.

Экскаваторын операторт мэдээлэх хиймэл оюун ухаан бүхий системийг NVIDIA компаний Jetson nano хөгжүүлэлтийг хавтанг ашиглан хийж туршилтын байдлаар ажиллуулсан.

Цаашид камерийг өөр экскаватор дээр байршуулан өдөр, шөнө болон ялгаатай өнцгөөс өгөгдөл цуглуулан хиймэл оюун ухаанаа дахин сургаж алдааг бууруулах шаардлагатай гэж үзэж байна. Мөн загваруудыг харьцуулан судалж эмбеддед төхөөрөмж дээр ажиллах боломжтой байхаар илүү хөнгөн, хурдан ажиллах боломжтой загварыг гаргаж авах шаардлагатай.

Ном зүй

[1] W. Wang and K. L. Siau, "Artificial Intelligence, Machine Learning, and Autonomous Technologies in the Mining Industry," *Proceedings*

of the Twenty-Fifth Americas Conference on Information Systems (AMCIS), Cancun, Mexico, Aug. 2019.

[2] Poma, X.S.; Riba, E.; Sappa, A. Dense extreme inception network: Towards a robust cnn model for edge detection. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Snowmass, CO, USA, 1–5 March 2020; pp. 1923–1932.

[3] Tian, Z.; Chen, D.; Liu, S.; Liu, F. DexiNed-based Aluminum Alloy Grain Boundary Detection Algorithm. In Proceedings of the 2020 Chinese Control and Decision Conference (CCDC), Hefei, China, 22–24 August 2020; pp. 5647–5652.

[4] Ge, Z., Liu, S., Wang, F., Li, Z., and Sun, J.: YOLOX: Exceeding yolo series in 2021, arXiv [preprint], https://doi.org/10.48550/arxiv.2107.08430, 6 August 2021.

[5] Liu X, Qi X, Jiang Y. Electric Shovel Teeth Missing Detection Method Based on Deep Learning. Comput Intell Neurosci. 2021 Nov 22;2021:6503029. doi: 0.1155/2021/6503029. PMID: 34853585; PMCID: PMC8629673.

[6] Ji, Shengfei & Li, Wei & Zhang, Bo & Zhou, Lingwei & Duan, Chenxi. (2021). Bucket Teeth Detection Based on Faster Region Convolutional Neural Network. IEEE Access. PP. 1-1. 10.1109/ACCESS.2021.3054436.

[7] Lu, J. and Liu, Y.: A real-time and accurate detection approach for bucket teeth falling off based on improved YOLOX, Mech. Sci., 13, 979–990, https://doi.org/10.5194/ms-13-979-2022, 2022.

[8] Akter, Alkın & Basak, Hudayim. (2021). Design and analysis of biomimetics based excavator bucket and tooth. Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering. 236. 095440892110576. 10.1177/09544089211057645.

[9] Terven, J.; Córdova-Esparza, D.-M.; Romero-González, J.-A. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. Mach. Learn. Knowl. Extr. 2023, 5, 1680-1716. https://doi.org/10.3390/make5040083

[10] A. Kirillov et al., "Segment Anything," arXiv preprint arXiv:2304.02643, Apr. 2023. [Online]. Available: https://arxiv.org/abs/2304.02643

[11] A. Singla, A. Mukherjee, and S. Dey, "Detection and identification of tramp metal in the mining industry: A review," *Minerals Engineering*, vol. 55, pp. 160–168, Jan. 2014, doi: 10.1016/j.mineng.2013.09.003.

Био анагаахын SAR ADC-ийн түүвэрлэн хадгалах хэлхээний судалгаа

Анхбаярын Наранмандах
ЭХИТ
МУИС, МТЭС
Улаанбаатар, Монгол
naranmandakh010125@gmail.com

Алтансүхийн Сугаррагчаа
ЭХИТ
МУИС, МТЭС
Улаанбаатар, Монгол
sugaraltan@gmail.com

Даланбаярын Болормаа
ЭХИТ
МУИС, МТЭС
Улаанбаатар, Монгол
bolormaad@num.edu.mn

Хураангуй- Орчин үеийн биоанагаахын салбарт мобайл, зөөврийн болон суурин оношилгооны төхөөрөмжүүдийн хэрэглээ өсөн нэмэгдэж байна. Энэ нь өндөр нарийвчлалтай, чадлын зарцуулалт багатай, найдвартай аналог-тоон хувиргуур (Analog Digital Converter, ADC)–ын хэрэгцээг нэмэгдүүлж байна. Ялангуяа зүрхний цахилгаан бичлэг (Electrocardiography, ECG), тархины цахилгаан бичлэг (Electroencephalography, EEG) зэрэг физиологийн дохионы системүүдэд бага хүчдэлтэй дохиог өндөр нарийвчлалтай хувиргах чадвартай ADC шаардлагатай. Ийм нөхцөлд хялбар бүтэцтэй, эрчим хүчний хэмнэлттэй, дунд болон өндөр нарийвчлалтай дараалан дөхүүлэх регистр (Successive Approximation Register, SAR) бүхий ADC өргөн хэрэглэгддэг. SAR ADC-ийн гүйцэтгэлийг хязгаарладаг гол хүчин зүйл нь түүвэрлэн хадгалах (Sample-and-Hold, S/H) хэлхээ юм. Түүвэрлэлтийн алдагдал, дохионы гажуудал, конденсаторын шүүрэл, цэнэгийн инжекц зэрэг нь хувиргуурын нарийвчлалд сөргөөр нөлөөлдөг. Энэхүү судалгаанд S/H хэлхээний хүчдэлийн алдагдал, дохио дамжууллын алдаа, цэнэгийн инжекц, офсет гажуудлын нөлөөллийг онолын болон туршилтын аргаар шинжилсэн. Судалгааны үр дүн нь биоанагаахын хэрэглээнд зориулсан SAR ADC-ийн загварчлалыг оновчлоход чиглэнэ.

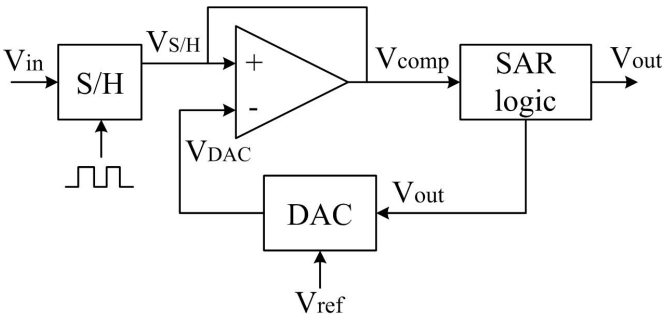
Түлхүүр үгс: аналог-тоо хувиргуур, дараалан дөхүүлэх регистр, харьцуулагч, түүвэрлэн хадгалах

I. Удиртгал

Биоанагаахын салбарт утасгүй сүлжээг ашиглан өвчтөний биеийн байдлыг хянах системийн хэрэглээ эрчимтэй нэмэгдэж байна. Эдгээр системд биед зүүсэн сенсорууд нь бие даасан сүлжээний зангилаа болж, аналог дохиог тоон хэлбэрт хөрвүүлэн эмч нарт мэдээлэл дамжуулдаг. Эмнэлгийн имплант, утасгүй сенсор зэрэг системүүд нь өвчтөний физиологийн үзүүлэлтүүдийг удаан хугацаанд, ямар нэгэн хязгаарлалтгүй хянах боломжийг олгодог бөгөөд энэ чиглэлийн судалгаа идэвхтэй хөгжиж байна [1], [2]. Утасгүй биоанагаахын системийн гол сорилтуудын нэг нь эрчим хүчний хэмнэлт бөгөөд батерейгаар тэжээгддэг зангилаануудын хувьд чадлын зарцуулалт бага байх нь чухал юм [3].

Аналог-тоон хувиргуур (ADC) нь аналог дохиог тоон хэлбэрт хөрвүүлэхэд гол үүрэг гүйцэтгэдэг бөгөөд биоанагаахын утасгүй сүлжээний чухал бүрэлдэхүүн хэсэг юм. Биоанагаахын хэрэглээний ADC нь бага талбайтай, дунд болон өндөр нарийвчлалтай, хэдэн кS/с-ээс хэдэн MS/с хүртэл түүвэрлэх хурдтай байхыг шаарддаг [4], [5]. ADC-ийн гүйцэтгэлийг тодорхойлох гол параметрүүдэд түүвэрлэх давтамж (Fs), нарийвчлал, идэвхтэй битийн тоо (ENOB), дохио шуугианы гажуудлын харьцаа (SNDR), чадлын зарцуулалт, ашигт үйлийн коэффициент (FOM) зэрэг орно. Эдгээр параметрүүдийг зэрэг сайжруулах нь биоанагаахын системийн найдвартай байдлыг хангахад чухал юм.

ADC-ийн архитектурын сонголт нь системийн гүйцэтгэлд шийдвэрлэх нөлөөтэй. Олон төрлийн ADC-ийн дундаас бага чадал, хурдан ажиллагаа, өндөр нарийвчлалын шаардлагыг хангадаг SAR ADC нь биоанагаахын хэрэглээнд хамгийн тохиромжтой [6]. SAR ADC нь түүвэрлэн хадгалах (S/H) хэлхээ, харьцуулагч (comparator), дараалан дөхүүлэх регистр (SAR logic), тоон аналог хувиргуур (DAC) гэсэн дөрвөн үндсэн блокоос бүрдэнэ (Зураг 1). Оролтын дохиог S/H хэлхээ



Зураг 1: Дараалан дөхүүлэх регистртэй АТХ-ын блок схем

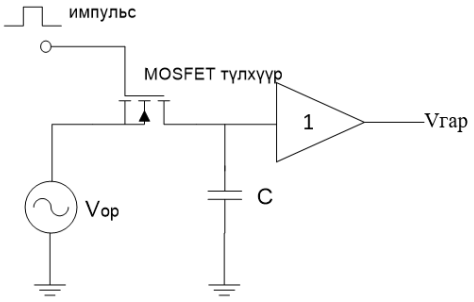
дискретчлэх ба дараалан дөхүүлэх регистрт хадгалагдсан тоон утгыг ТАХ аналог болгон хөрвүүлнэ. Үүний дараа компаратор S/H-ын гаралтын дохиог ТАХ-ын гаралтын утгатай харьцуулж регистрийн утгыг шинэчлэх дохиог гаргана. Эхэлж хамгийн ахлах бит

тоон хэлбэрээр илэрхийлэгдэж, дараа нь дотоод ТАХ $V_{ref}/2$ утгыг компараторт илгээж дискретчилсэн утгатай харьцуулах гэх мэтээр хамгийн бага битийн утга хүртэл циклдэнэ.

S/H хэлхээ нь SAR ADC-ийн гүйцэтгэлийн гол хязгаарлагч бөгөөд түлхүүр (switch) болон конденсатороос бүрдэнэ. Энэ хэлхээ нь оролтын тасралтгүй дохиог Найквистийн шалгуурыг хангасан давтамжаар дискретчилдэг. Гэсэн хэдий ч NMOS транзисторын шугаман бус ажиллагаа, цэнэгийн инжекц, гармоник гажуудал, дотоод шуугиан зэрэг нь нарийвчлалыг бууруулдаг. Энэхүү судалгаа нь S/H хэлхээний параметрүүдийг оновчлоход чиглэж, SAR ADC-ийн нарийвчлал, хурд, найдвартай байдлыг сайжруулах боломжийгсудална.

II. Судалгааны арга зүй

Хялбар түүвэрлэн хадгалах хэлхээ NMOS транзистор, конденсатороос тогтох ба шилжүүлэгч конденсаторт (switched capacitor) хэлхээ гэж нэрлэдэг. Ийм түүвэрлэн хадгалах хэлхээг Зураг 2-д үзүүллээ.



Зураг 2: Хялбар түүвэрлэн хадгалах хэлхээ.

LTspice XVII программ хангамж ашиглан Хүснэгт I-д үзүүлсэн зүрх (Electrocardiography, ECG), тархи (Electroencephalography, EEG), булчингийн цахилгаан бичлэг (Electromyography, EMG)-ийн дохиог түүвэрлэн хадгалах хэлхээний анализ хийхээр хэлхээний элементүүдийн параметр болон оролт, түүвэрлэх болон референс хүчдлүүдийн утгыг сонгосон.

Хүснэгт I: Биоанагаахын дохионы параметрууд

Дохионы төрөл	Давтамж (Гц)	Далайц
ECG	0.05–250	5 μV – 8 mV
EEG	0.5–200	2 μV – 200 μV
EMG	0.01–10,000	50 μV – 10 mV

Гаралтын битийн тоог дундаас-өндөр нарийвчлалтай АТХ-т тохируулан 12 байх нь тохиромжтой байгааг [6], [7]-аас харан харгалзах нарийвчлалын хүснэгтийг гаргасныг Хүснэгт II-т оруулаа.

Хүснэгт II: Аналог-Тоон Хувиргуурын Нарийвчлал ба Хэрэглээ

Нарийвчлал	Битийн тоо	Хэрэглээ
Бага	8–10	Ердийн зориулалт, аудио, батарейн хяналтын сенсор
Дундаас өндөр	12	Үйлдвэрийн хэрэглээ, биоанагаахын сенсор, моторт удирдлага
Өндөр	14–16	Хэмжилтийн тоног төхөөрөмж, прецизионный хэмжилт, спектроскоп
Маш өндөр	>16	Лабораторийн нарийн тоног төхөөрөмж, өндөр нарийвчлалтай багаж

Хүснэгт 1-д заасан 0-10кГц давтамжийн мужид түүвэрлэх хугацаанд конденсатор хурдан цэнэглэгдэн, цэнэгээ алдахуйцаар реактив эсэргүүцэл хангалттай бага байхын тулд:

$$X_C = \frac{1}{2\pi f C} \tag{1}$$

томъёогоор анализ хийх багтаамжийн интервалыг бодсоныг Хүснэгт III-д харууллаа.

Хүснэгт III: Конденсаторын багтаамжийн интервал

Багтаамж, C	X_C, 250Гц-д)	X_C, 10КГц-д)
1мкФ	636 Ом	15,9 Ом
100нФ	6,4 Ом	159 Ом
10нФ	64 Ом	1,6 Ом

10КГц давтамжид 1мкФ багтаамж хэт их бөгөөд тогтворжих хугацаа удааширна. Иймээс хамгийн тохиромжтой нь 1-100нФ байх боломжтой.

Илүү өндөр давтамжид түүвэрлэн хадгалах хурдан тогтворжих:

$$t_{settle} \approx 3 - 5\tau, \quad \tau = RC \tag{2}$$

тул хугацааны тогтмолыг багасгахын тулд бага багтаамжтай конденсатор ашиглах хэрэгтэй болно.

Шүүрлийн болон гаралтын гүйдлээс үүсэх хүчдлийн уналт:

$$\Delta V = \frac{I t_{hold}}{C} \tag{3}$$

учир түүвэрлэх хугацаа бага үед хадгалах хугацаа t_{hold} багасах тул хүчдлийн утгыг тогтвортой байлгахын тулд бага багтаамжтай конденсатор сонгож болохоор байна.

Дээрх тооцоололд үндэслэн сонгож авсан анализын параметруудыг Хүснэгт IV-д харууллаа.

Түүвэрлэн хадгалах хэлхээгээр оролтын дохио нэвтрэх буюу түлхүүр нээлттэй байх үеийн гаралтад үүсгэх гажилт:

$$\epsilon_{\text{through}} = V_{\text{in-pp}} * \frac{C_{\text{DS}}}{C_{\text{Hold}} + C_{\text{DS}}} \tag{4}$$

Хүснэгт IV: Анализын параметр

Параметр	250Гц систем	10КГц систем
C	1мкФ-10мкФ	1нФ-100нФ
Түүвэрлэх хугацаа	хэдэн мс	10-100мкс
X_C	10-100 Ом	1-10 Ом
Хүчдлийн уналт	критик	критик биш

энд:

C_{DS} : түлхүүрийн дрейн соорсын багтаамж

C_{Hold} : оролтын дохиог хадгалах конденсаторын багтаамж

Энэ гажилт нь

$$C_{DS} << C_{Hold}$$

үед ихэсдэг.

Түлхүүр хаалттай байх үед сувгийн цэнэг сорогдох буюу цэнэгийн инжекц хадгалах конденсаторын хүчдлийг өөрчилснөөр гаралтын дохионы гажилтад хамгийн ихээр нөлөөлнө. Цэнэгийн инжекц:

$$Q_{inj} = C_{ox} * W * L * (V_{GS} - V_{TH}) \tag{5}$$

энд:

C_{ox} : транзисторын оксидын үеийн багтаамж

Q_{ch} : сувагт хуримтлагдах цэнэгийн хэмжээ

W : сувгийн өргөн

L : сувгийн урт

V_{GS} : гейт соорсын хүчдэл

V_{TH} : босго хүчдэл

Цэнэгийн инжекцээс үүсэх хүчдлийн өөрчлөлт:

$$dV_{\text{inj}} = \frac{Q_{\text{ch}}}{2C} = \frac{C_{ox} * W * L * (V_{GS} - V_{TH})}{2 * C_{\text{Hold}}} \tag{6}$$

энд: Эндээс харахад цэнэгийн инжекц гаралтын дохионд ихээр нөлөөлнө.

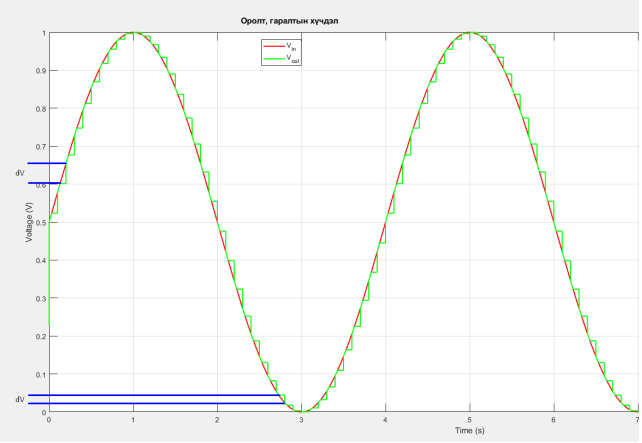
Удирдлагын импульс түүвэрлэн хадгалах хэлхээгээр нэвтрэхэд гейт-соорсын багтаамжаас үүдсэн хоцрогдол бий болно. Иймээс үүсэх хүчдлийн өөрчлөлт:

$$\epsilon_{im} = V_{im} * (\frac{C_{\text{gs}}}{C_{\text{hold}} + C_{\text{gs}}}) \tag{7}$$

Уг гажилт оролтын дохионоос хамаарахгүй ч гаралтад тулгуур хүчдлийн (офсет) гажилт үүсгэнэ.

III. Үр дүн

Түүвэрлэн хадгалах хэлхээний бүлэг II-д дурдсан зарим параметрүүдийн үзүүлэлтүүд нь гаралтын утгад хэрхэн нөлөөлж байгааг симуляцийн орчинд LTSpice XVII програм хангамж ашиглан туршив. Зураг (3)-д хэлхээний оролтод 1В далайцтай синусан дохио өгсөн үеийн гаралтын квантлагдсан хүчдлийг харууллаа. Уг зургаас квантлагдсан хүчдлийн зөрөө оролтын дохионы дээд болон доод пикийн орчимд бага, гол хэсэгт их байгаа нь харагдаж байна. Өөрөөр хэлбэл квантлагдсан утгын зөрөөг хазайлтын утганаас хэтрүүлж болохгүй нь харагдаж байна.



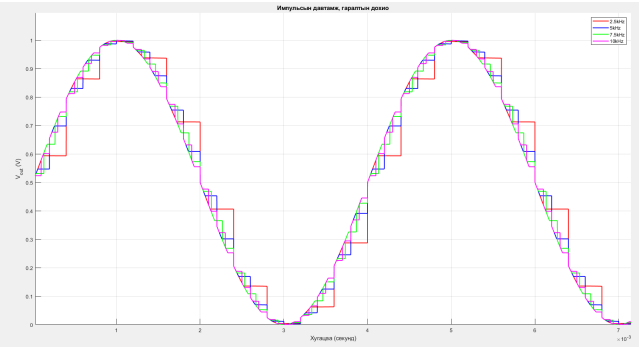
Зураг 3: Түүвэрлэн хадгалах хэлхээний оролт, гаралтын хүчдэл

Импульсийн давтамж өөрчлөгдөхөд оролтын хүчдэлийн квантлагдсан утга буюу LSB нь хэрхэн өөрчлөгдөж байгааг судлахын тулд түүвэрлэх давтамж буюу импульсийн давтамжийг оролтын дохионы давтамжаас дор хаяж 2 дахин их байхаар тооцлоо.

$$f_{\text{impulse}} > 2 * f_{\text{in}} \tag{8}$$

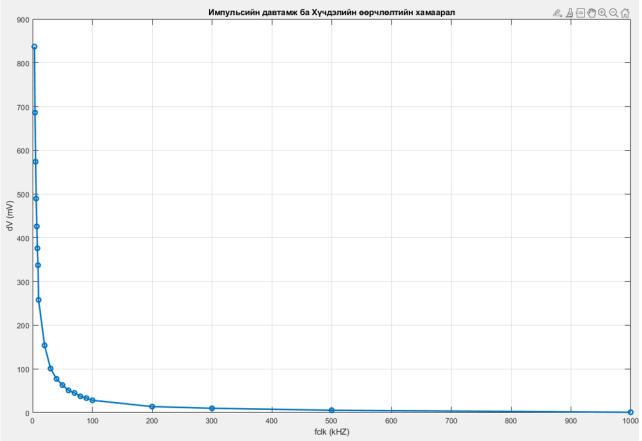
Эндээс нам давтамжид квантлагдсан хүчдлийн хазайлт их байгаа нь дараалан дөхүүлэх регистр бүхий аналог тоон хувиргуурын нарийвчлалыг сайжруулахын тулд уг хазайлтыг багасгах шаардлагатай нь харагдаж байна.

Оролтын дохионы хамгийн бага давтамжийг 1кГц гэж үзэн хадгалах конденсаторын хэмжээ 10пФ үед импульсийн давтамжийг 2кГц ээс 1МГц хүртэл өөрчлөн LSB -ийн утга хэрхэн өөрчлөгдөж байгааг зураг 4-д харуулав. Эндээс харахад оролтын дохионы давтамж ихсэхэд LSB -ийн утга багасч байна.



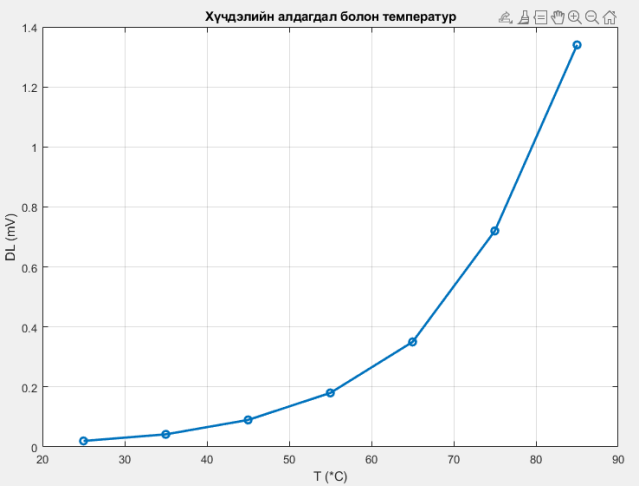
Зураг 4: Импульсийн давтамж болон гаралтын дохио

Түүвэрлэн хадгалах хэлхээний гол шаардалагын нэг болох хадгалах горимд конденсаторт цэнэглэгдсэн



Зураг 5: Импульсийн давтамж, хүчдэлийн өөрчлөлтийн хамаарал

хүчдэлээ тогтвортой барьж байх ёстой. Гэвч шүүрлийн гүйдлээс шалтгаалан хүчдэлийн алдагдал үүсдэг. MOSFET түлхүүрийн шүүрлийн гүйдэл температур



Зураг 6: Импульсийн давтамж, хүчдэлийн алдагдлын хамаарал

10°C-аар нэмэгдэх тутам 2 дахин нэмэгддэгийг имитацлахын тулд зураг 2-т үзүүлсэн хэлхээний конденсаторын зангилаан дээр гүйдлийн үүсгүүрээр шүүрлийн гүйдлийг төлөөлүүлэн (??) томъёогоор илэрхийлэгдэх хүчдлийн алдагдал температураас хэрхэн хамаарахыг судаллаа. Зураг 6-аас харахад 50°C хүртэл хүчдлийн алдагдал алгуур, маш бага өөрчлөгдөж байгаа тул биоанагаахын хувиргуурын үзүүлэлтэд багаар нөлөөлнө гэж үзлээ.

Түүвэрлэн хадгалах хэлхээний түлхүүр нээлттэй буюу оролтын дохиог нэвтрүүлэх, түлхүүр нээлттэй үеийн гаралтад үүсэх гажилт нь (4) томъёоны дагуу түлхүүрийн дрейн соорсын багтаамж нь хадгалах конденсаторын хэмжээнээс олон дахин бага байх нь гажилт үүсгэдэг нь харагдаж байна. 90нано метр MOS технологи ашиглан түлхүүрийн параметрүүдийг тодорхойлоход дрейн соорсын багтаамж нь:

$$C_{\mathrm{DS}} = 1.43fF \tag{9}$$

гэж гарсан.

$$C_{DS} << C_{Hold}$$

Хамаарлын дагуу хадгалах конденсаторыг

$$C_{Hold} = 100pF$$

хэмжээтэй авч туршилтыг явуулах нь боломжтой.

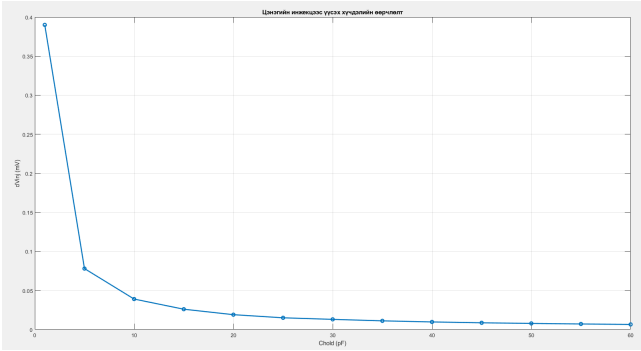
Цэнэгийн инжекц нь түлхүүр хаалттай үед оронгийн транзисторын сувгаар бага зэргийн цэнэгийн шилжилт явагддаг. Үүнээс үүдэн гаралтын хүчдэлд алдаа үүсгэдэг. (5) болон (6) томъёонуудад 90нано метрийн MOS технологийн утгуудыг ашиглан зарим параметрүүдийг орлуулан тооцоолбол:

$$\begin{aligned} C_{\mathrm{ox}} &= 1.6 * 10^{-2} F/m^2 \\ W &= 0.9\mu \\ L &= 90 \\ V_{GS} &= 1V \\ V_{TH} &= 0.397V \end{aligned}$$

Эдгээрийг (6) томъёонд орлуулбал:

$$dV_{\mathrm{inj}} = \frac{Q_{\mathrm{ch}}}{2C} = \frac{1.6 * 10^{-2} * 0.9\mu * 90 * (1V - 0.397V)}{2 * C_{Hold}} \tag{10}$$

Хадгалах конденсаторын утга их байх тусам цэнэгийн



Зураг 7: Цэнэгийн инжекцээс үүсэх хүчдэлийн өөрчлөлт, хадгалах конденсаторын утгын хамаарал

инжекцээс хамаарсан гаралтын хүчдэлийн өөрчлөлт нь багасаж байгааг зураг (7)-д харуулав.

IV. Дүгнэлт

Био анагаахын зориулалттай дараалан дөхүүлэх регистр бүхий аналог-тоон хувиргуурын (SAR ADC) гүйцэтгэлийг сайжруулахад түүвэрлэн хадгалах болон компараторын хэлхээний параметрүүдийг оновчлох нь

чухал болохыг энэ судалгаагаар харууллаа. Судалгааны үр дүнд хадгалах хугацаа (10–20 мкс), компараторын хариу үйлдлийн хугацаа (1–2 мкс), цэнэгийн инжекц, хүчдэлийн алдагдал зэрэг үзүүлэлтүүдийн нөлөө ADC-ийн нийт нарийвчлалд мэдэгдэхүйц илэрч байгааг тооцооллоор нотлов.

Ишлэл [1]-д тэмдэглэснээр SAR төрлийн хувиргуур нь мобайл болон эмнэлгийн зөөврийн төхөөрөмжид тохиромжтойгоор тодорхойлогдсон ч манай судалгаагаар уг боломжийг бүрэн хэрэгжүүлэхийн тулд дэд хэлхээний параметрүүдийг оновчлох шаардлагатайг харууллаа. Мөн [2]-т дурдсанчлан хадгалах хугацааг зөв тодорхойлох нь хувиргалтын найдвартай байдалд шууд нөлөөлдөг нь бидний симуляцаар баталсан.

Компараторын хариу үйлдлийн хугацаа, оролтын офсет хүчдэл зэрэг параметр нь [3]-т хувиргалтын хурд болон битийн алдаанд голлон нөлөөлдөг болохыг онцолсон нь оролтын хүчдлийн далайц бага үед илүү нөлөөлөх нь симуляцаас харагдлаа. Мөн [4]-т тэмдэглэсэнчлэн 12–16 битийн нарийвчлалтай системд компараторын оролтын шуугиан 244 мкВ-оос доош байх шаардлагыг хангахын тулд дулааны болон шүүрлийн гүйдлийг удирдах арга зам чухал байгааг бидний судалгаа харууллаа.

Эцэст нь, биоанагаахын дохионы нарийн дискретчилэл шаардсан системүүдэд SAR ADC-ийг хэрэглэхэд түүвэрлэн хадгалах болон компараторын хэлхээний онцлогийг уялдуулан, параметр бүрийн нөлөөг нарийвчлан тооцох шаардлагатай юм. Иймд цаашид эдгээр параметрийн температурын нөлөө, технологийн түвшинд үндэслэсэн миниатюрчлалын асуудлыг гүнзгийрүүлэн судлах шаардлагатай гэж үзэж байна.

Ашигласан ном

[1] H. Bhamra, J. Lynch, M. Ward, and P. Irazoqui, “A noise power-area optimized biosensing front end for wireless body sensor nodes and medical implantable devices,” IEEE Transactions on Very Large Scale Integration Systems, vol.25, no.10, pp.2917–2928, 2017.

[2] S.T.Ramkaj, M.Strackx, M.S.J.Steyaert, and F.Tavernier, “A 1.25-GS/s 7-b SAR ADC With 36.4-dB SNDR at 5 GHz Using Switch-Bootstrapping, USPC DAC and Triple-Tail Comparator in 28-nm CMOS,” IEEE Journal of Solid-State Circuits, vol.53, no.7, pp.1889–1901, 2021

[3] M.Jotschke, W.C.Ossa, T Reich, and C.Mayr, “A 10.5 μW programmable SAR ADC frontend with SC preamplifier for Active and Passive Electronic Components low-power IoT sensor nodes,” Proceedings of the IEEE 6th World Forum on Internet of Things (WF-IoT), pp. 1–6, New Orleans, LA, USA, June 2020.

[4] B. M. Kelly, A. T. DiLello, and D. W. Graham, “Reconfigurable analog preprocessing for efficient asynchronous analog-to-digital conversion,” Journal of Low Power Electronics and Applications, vol.9, no.3, pp.25–18, 2019.

[5] M.Bhattacharyya, W.Gruenwald, D.Jansen, L. Reindl, and J. Aghassi-Hagmann, “Design of a programmable passive SoC for biomedical applications using RFID ISO 15693/ NFC5 interface,” Journal of Low Power Electronics and Applications, vol.8, no.1, pp.3–31, 2018.

[6] Thomas Bos, Komail Badami, Wim Dehaene and Marian Verhelst. "Architecture optimization for energy-efficient resolution-scalable 8–12-bit SAR ADCs"Springer Nature Link, Volume 97, pages 437–448, (2018)

[7] Shouping Li,Yang Guo,Jianjun Chen and Bin Liang. "A 12-bit 30 MS/s Successive Approximation-Register Analog-to-Digital Converter with Foreground Digital Calibration Algorithm"

[8] S.A. Zahrai and M. Onabajo, “Review of analog-to-digital conversion characteristics and design considerations for the creation of power-efficient hybrid data converters,” Journal of Low Power Electronics and Applications, vol. 8, no. 2, pp. 12–29, 2018.

[9] J. Thompson, ”Optimizing sample-and-hold circuits in successive approximation ADCs,” Journal of Analog and Digital Systems, 45(2), 132-140, 2020.

[10] Y. Zhang, H. Lee, S. Wu, ”The effect of hold time on the performance of biomedical ADCs,” IEEE Transactions on Biomedical Engineering, 66(4), 2019, pp. 758-765.

[11] H. Chen, X. Li, ”High-speed comparators for ADCs: Design and performance analysis.” IEEE Transactions on Circuits and Systems, 65(7), 2248-2255, 2020.

[12] D. Smith, C .Lee, ”Precision and speed in biomedical ADCs for healthcare applications.” Journal of Medical Devices, 15(3), 201-210, 2021.

[13] H.P. Le, A. Zayegh, J. Singh, “Performance analysis of high-accuracy CMOS Sample-and-Hold circuits,” Proceedings of SPIE - The International Society for Optical Engineering, 2004.

Оновчтой сплайн муруй ашигласан дифференциал роботын зам төлөвлөлт

Д.Нанзадрагчаа, М.Баярпүрэв

МУИС, Мэдээллийн технологи, электроникийн сургууль
Электроник, Холбооны Инженерчлэлийн Тэнхим
Улаанбаатар хот, Монгол Улс
Цахим шуудан: nanzadragchaa@num.edu.mn

***Хураангуй*—Энэхүү ажлаар дифференциал роботын замыг гөлгөрөөр төлөвлөх асуудлыг авч үзнэ. Роботын зам нь гөлгөр байх тусам түүний хурд сайжирч энергийн хэмнэлт үүснэ. Мөн роботын кинематик болон динамик тооцоо, анализ хийхэд хялбар болно. Машин-хэвт дифференциал роботын замыг хойд дугуйгаар нь төлөвлөх нь хялбар. Гэвч энэ үед урд дугуйны зам нь хойд дугуйн замын мурийлт буюу хоёр дугаар эрэмбийн уламжлалтай нь холбоотой байдаг тул "мурилзах"үзэгдэл үүсдэг. Энэ ажлаар бид энэхүү "мурилзалтыг"арилгаж илүү "гөлгөр"зам үүсгэх арга санал болгоно. Ингэхдээ бид урд дугуйны замыг мурийлтыг оролцуулан зардлын (cost) функци үүсгэн оптимизаци хийнэ.**

***Түлхүүр үгс*—Зам төлөвлөлт, сплайн, оновчлол, дифференциал робот**

I. Удиртгал

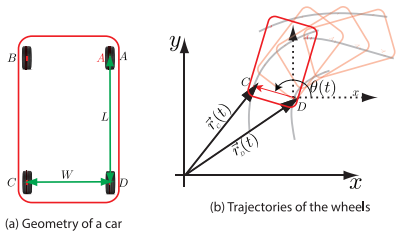
Жолоочгүй машины хөгжүүлэлт сүүлийн жилүүдэд эрчимтэй явагдаж байна. Энэ салбар нь дүрс боловсруулалт, компьютер хараа зэрэг машин сургалтын өрсөлдөөний талбар болж байна. Эдгээрээс гадна хиймэл оюуны асуудлуудаас гадна, роботын байршлыг маш хурдтай тодорхойлох, зөв удирдах, оптимал зам төлөвлөх зэрэг асуудлууд ч чухал байр суурьтай. Бид энэ судалгааны ажлаар жолоочгүй машин буюу хойд дугуйнууд нь залуурдагддаггүй дифференциал тэнхлэг бүхий хөдөлгөөнт роботын зам төлөвлөлтийн асуудлыг авч үзнэ.

Роботын зам нь аль болох гөлгөр (хоёр болон түүнээс дээш зэргийн уламжлалууд нь тасралтгүй) байх тусам роботын автомат удирдлага хялбар болох төдийгүй энергийн хэмнэлт үүснэ. Жишээ нь хүний хөдөлгөөн 0,2 kcal/kg/km энерги шаарддаг бол Boston Dynamics-н Atlas хуманоид робот хүнээс 5 10 дахин их, Honda-н ASIMO 20 40 дахин их энерги зарцуулдаг [1]. Энэ нь хүн өөрийн кинематик, инерцийн момент, пассив динамикт оновчлогдсон гөлгөр замаар үе мөчөө хөдөлгөдөгтэй холбоотой [1]. Дифференциал роботын хувьд ч төстэй асуудал тулгарна.

Дифференциал роботын хувьд хойд дугуйгаар зам төлөвлөх нь хялбар байдаг. Гэтэл хойд дугуйны зам төлөвлөлтийн жижиг алдаа нь урд дугуйны мурилзалтанд ихээр нөлөөлдөг. Улмаар урд дугуйны залуурдалт болон роботын удирдлага хүндэрч, энергийн алдагдал үүснэ [2]. Эсрэгээр урд дугуйгаар зам төлөвлөлтийг хийвэл хойд дугуйны замыг төсөөлөхөд хүндэрнэ. Бидний санал болгож буй арга нь дараах процедураар ажиллана. Эхлээд роботын хойд дугуйн замыг сплайнаар төлөвлөнө. Дараагаар нь урд дугуйн замыг тооцоолж замын мурийлт (хоёрдугаар эрэмбийн уламжлалын $L2$ норм) болон сплайны зангилаа цэгүүдээс зөрөх квадрат алдаанаас хамаарсан зардлын функц байгуулж түүнийг оптимизаци хийнэ.Эцэс нь урд дугуйны гөлгөр замаас хойд дугуйн замыг тооцоолж гаргана.

II. Дифференциал роботын кинематик

Машин-хэвт роботын урд хоёр дугуй нь залуурдагддаг ба хойд хоёр дугуй нь залуурдагддаггүй бөгөөд дифференциал тэнхлэг дээр байрлана. Дифференциал роботын загвар болон дугуйнуудын траекторыг Зураг 1-д үзүүлэв. W нь дифференциал тэнхлэгийн өргөн, L нь урд хойд дугуй хоорондын зай. Талбай дээр Декарт координатын систем оруулсан бөгөөд $\overrightarrow{DC}$ тэнхлэг ба абсцисс x хоорондын өнцгийг $\theta(t)$ гэж тэмдэглэв.



Зураг 1. Машин-хэвт дифференциал роботын загвар болон дугуйнуудын траектор

Зураг 1-д үзүүлсэн A, B, C, D дугуйнуудын координатыг нь $\vec{r}_A(t)$, $\vec{r}_B(t)$, $\vec{r}_C(t)$, $\vec{r}_D(t)$ гэж, хурдны векторуудыг нь $\vec{v}_A(t)$, $\vec{v}_B(t)$, $\vec{v}_C(t)$, $\vec{v}_D(t)$ гэж тус тус тэмдэглэв. Хурдны вектор нь үргэлж траектортоо шүргэгч байна.

Дифференциал роботын хувьд хойд дугуйнуудын хурд $\vec{v}_D(t)$ болон $\vec{v}_C(t)$ нь үргэлж дифференциал тэнхлэг $\overrightarrow{DC}(t)$-д перпендикуляр байна (Зураг 2(a)). Иймээс

$\vec{v}_D(t)$-г нормчилж, цагийн зүүний эсрэг 90° эргүүлэн, W-ээр үржүүлбэл $\overrightarrow{DC}(t)$ тэнхлэг гарна. Энэ чанарыг ашиглан дифференциал роботын хувьд аль нэг хойд дугуйны (тухайлбал D дугуй) траектор өгөгдвөл бусад дугуйны траекторыг хялбар олж чадна:

$$\begin{aligned}
\vec{r}_C(t) &= \vec{r}_D(t) + \overrightarrow{DC}(t) &&= \vec{r}_D(t) + W\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\frac{\vec{v}_D(t)}{v_D(t)}\\
\vec{r}_A(t) &= \vec{r}_D(t) + \overrightarrow{DA}(t) &&= \vec{r}_D(t) + L\frac{\vec{v}_D(t)}{v_D(t)} \qquad (1)\\
\vec{r}_B(t) &= \vec{r}_D(t) + \overrightarrow{DB}(t) &&= \vec{r}_A(t) + W\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\frac{\vec{v}_D(t)}{v_D(t)}.
\end{aligned}$$

Энэд, $v_D(t)$ нь $\vec{v}_D(t)$ векторын абсолют утга (норм). Нормчлодсон хурд $\vec{v}_D(t)/v_D(t)$-н уламжлал болон роботын өнцөг хурдын хамаарлыг хялбар гаргаж чадна:

$$\begin{aligned}
\frac{\vec{v}_D(t)}{v_D(t)} &= \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}\cos\theta(t)\\ \sin\theta(t)\end{bmatrix} = \begin{bmatrix}\sin\theta(t)\\ -\cos\theta(t)\end{bmatrix} \Rightarrow\\
\frac{d}{dt}\cdot\frac{\vec{v}_D(t)}{v_D(t)} &= \theta(t)'\begin{bmatrix}\cos\theta(t)\\ \sin\theta(t)\end{bmatrix} = \omega(t)\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\frac{\vec{v}_D(t)}{v_D(t)} \Rightarrow\\
|\omega(t)| &= ||\frac{d}{dt}\cdot\frac{\vec{v}_D(t)}{v_D(t)}||. \qquad (2)
\end{aligned}$$

Энэд $\omega(t) = \theta(t)'$ өнцөг хурд. Иймд, хурдны векторуудыг доорхи байдлаар илэрхийлж болно.

$$\vec{v}_C(t) = \vec{v}_D + W\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\frac{d}{dt}\cdot\frac{\vec{v}_D(t)}{v_D(t)} = \vec{v}_D - W\omega(t)\frac{\vec{v}_D}{v_D} \qquad (3)$$

$$\vec{v}_A(t) = \vec{v}_D + L\frac{d}{dt}\cdot\frac{\vec{v}_D(t)}{v_D(t)} = \vec{v}_D + L\omega(t)\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\frac{\vec{v}_D}{v_D} \qquad (4)$$

$$\vec{v}_B(t) = \vec{v}_A + +W\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\frac{d}{dt}\cdot\frac{\vec{v}_D(t)}{v_D(t)} = \vec{v}_A - W\omega(t)\frac{\vec{v}_D}{v_D} \qquad (5)$$

(3)-(5) томъёоноос хурдны векторуудын хооронд доорхи геометр хамаарал байгааг ажиглаж чадна. Хойд дугуйн хурдууд $\vec{v}_D$, $\vec{v}_C$ нь хоорондоо параллель ба диффернциал тэнхлэгтээ перпендикуляр байна. Тэгдээрийн ялгавар нь өнцөг хурдтай порпорциональ байна: $||\vec{v}_D(t) - \vec{v}(t)_C|| = W\omega(t)$. Урд хоёр дугуйн хурдны векторууд нь роботыг цагийн зүүн эсрэг 90° эргүүлж $\omega(t)$ агшааж (сунгаж) D, C дугуйны байрлалыг $\vec{v}_D$, $\vec{v}_C$ дээр тус тус байрлуулбал A, B дугуйны байрлал дээр $\vec{v}_A$, $\vec{V}_C$-г тус тус гаргаж авч чадна. Үүнийг Зураг 2(b)-д дэлгэрэнгүй үзүүлэв.

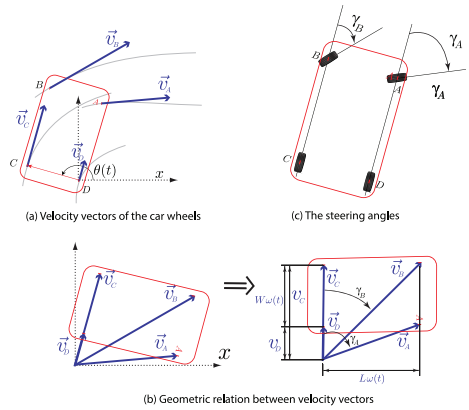
Энэхүү геометр хамаарлыг ашиглан урд дугуйнуудын залуурдалтын өнцөгийг доорхи байдлаар олж чадна:

$$\tan\gamma_A(t) = \frac{L\omega(t)}{v_D(t)} \qquad (6)$$

$$\tan\gamma_B(t) = \frac{L\omega(t)}{v_D(t) - W\omega(t)} \qquad (7)$$

III. Куб сплайнаар роботын зам төлөвлөх нь

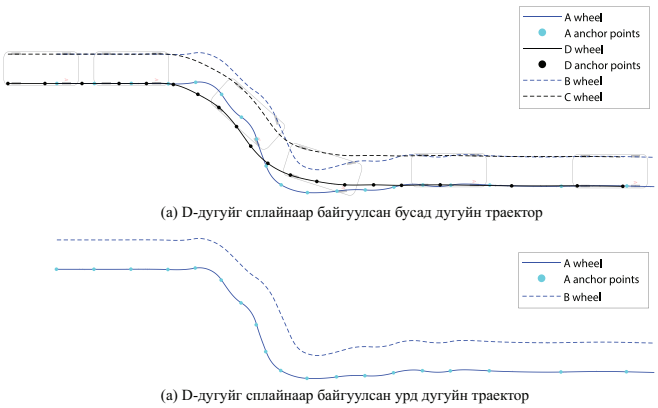
Өмнөх хэсэгт дифференциал роботын аль нэг хойд дугуйны траектор өгөгдвэл бусад дугуйны траекторыг



Зураг 2. Дифференциал роботын хурдны векторуудын геометр хамаарал

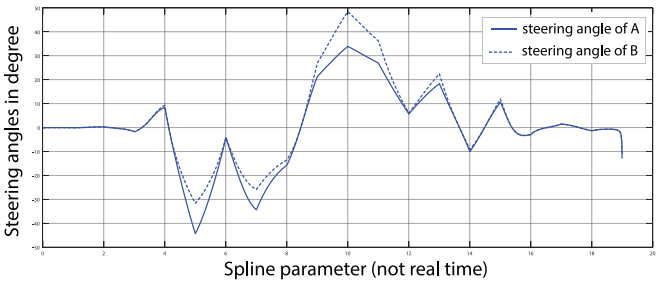
(1) томъёогоор хялбар болж болохыг харуулсан. Мөн роботыг замыг гөлгөр муруйгаар төлөвлөх нь энергийн хэмнэлт үүсч, роботын кинематик болон динамикийн анализыг хийхэд хялбар болно. Энэ хэсэгт бид роботын D дугуйны замыг куб сплайнаар төлөвлөж бусад дугуйны траектор болон роботын урд дугуйнуудыг залуурдах өнцөгийг анализ хийх болно.

Зураг 3-д D дугуйны замыг хар битүү зураасаар үзүүлэв. Үүнийг байгуулахдаа эхлээд талбай дээр гадас цэгүүд (anchor points) хатгаж түүнийгээ куб сплайнаар интрполяци хийнэ. Бусад дугуйны траекторыг (1) томъёогоор бодсон. Хойд дугуй C-г хар тасархай зураасаар, урд дугуй A-г хөх битүү зураасаар, B дугуйны замыг хөх тасархай зураасаар зурав. Зураг дээр урд дугуйны зам нь муруйлзаж байгааг ашиглаж болно. Үүний шалтгаан нь хойд дугуйны траекторын бага хэмжээний муруйлт (хоёр дугаар эрэмбийн уламжлал) нь урд дугуйны траекторт маш ихээр нөлөөлж байгаагаас болно. D дугуйны траекторыг зурахдаа гадас (anchor) цэгүүдийг маш нарийн байршуулж чадахгүй байгаатай холбоотой.



Зураг 3. Path planning using cubic spline for the rear wheel D

Урд хоёр дугуйны залуурдалтын өнцгийг (6)- (7) томьёогоор тооцоолж үр дүнг Зураг 4-д үзүүлэв. Дээр тайлбарласан ёсоор урд дугуйны залуурдалт маш их савлаж байгааг ашиглаж болно. Ийм замаар роботыг удирдах нь төвөгтэй болж удирлагын алдаа ихэснэ. Иймд дараагийн хэсэгт урд дугуйны мурилзалтыг (өөрөөр хэлбэл залуурдалтын өнцгийн савлагааг) багасгах арга санал болгоно.



Зураг 4. Steering angles of the front wheels with spline for rear wheel D

IV. Дифференциал роботын гөлгөр зам төлөвлөх арга

Өмнөх хэсэгт тайлбарласан ёсоор дифференциал роботын замыг хойд дугуйгаар төлөвлөвөл хялбар. Гэвч хойд дугуйны замын нүдэнд үл анзаарагдах бага хэмжээний "мурилзалт"урд дугуйг маш томоор "мурилзуулна". Энэ асуудлыг шийдэхийн тулд дараах процедур санал болгож байна:

1) Роботын замыг хойд дугуйгаар төлөвлөнө. Ингэхдээ талбай дээр роботын хойд дугуй дайрч өнгөрөх гадас цэгүүдийг хатгана. Эдгээр цэгүүдийг дайрах (интерполяцийн) куб сплайн муруй байгуулна. Энэ муруйн нь гадас цэгүүдийг дайрна.
2) Урд дугуйны гөлгөр замаас бусад дугуйны замыг тооцоолж гаргана.

Дээрх 2)-р алхмыг хэрэгжүүлэхдээ урд дугуйны замын мурийлт (хоёрдугаар эрэмбийн уламжлал) болон гадас цэгүүдийн тойрч (дайрах бус) өнгөрөх квадрат алдаанаас хамаарсан зардлын функц байгуулж түүнийг оновчилно. Зардлын функцийг доорх байдлаар тодорхойлъё:

$$J_1(f_1) = p\sum_{i=0}^{N}[x_i - f_1(i)]^2 + (1-p)\int_0^N |f_1(t)''|^2 dt. \quad (8)$$

$$J_2(f_2) = p\sum_{i=0}^{N}[y_i - f_2(i)]^2 + (1-p)\int_0^N |f_2(t)''|^2 dt. \quad (9)$$

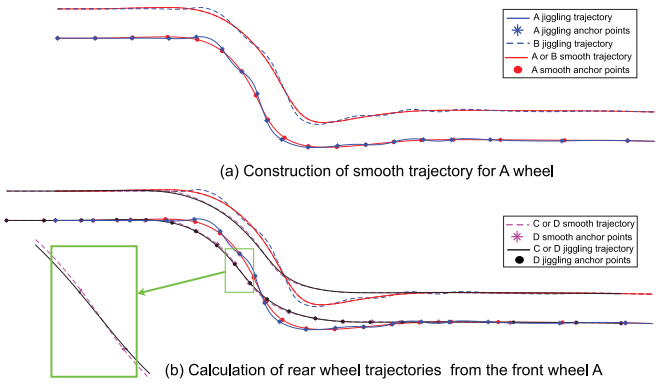
Энэд урд дугуйны траекторын x координатыг $f_1(t)$-р, y координатыг $f_2(t)$-р тэмдэглэв. N нь гадас цэгийн тоо (сплайн сегментийн тоо), $1-p$ болон p нь муруйлт болон гадас цэгүүдийн тойрч өнгөрөх квадрат алдааны жингүүд. p жинг хэрэглэгч гараар тохируулна. p бага байх тусмаа сплайн муруй "гөлийнө", харин гадас цэгүүдээс холдно. Эсрэгээр p их байх тусмаа сплайн муруй "мурилзаж", харин гадас цэгүүдтэй ойртно. $J_1(f_1)$-н шийд нь сплайн муруйны x координатын параметруудыг, $J_2(f_2)$-н шийд нь сплайн муруйны y координатын параметруудыг өгнө. Дээрхи оптимизацийн бодлого цор ганц шийдтэй бөгөөд шийд нь илт хэлбэрээр илэрхийлэгддэг. Дэлгэрэнгүй тооцоо [3]-н Chapter XIV-с харна уу. Энэ аргаар гаргаж авсан траекторыг Зураг 5(a)-д улаан өнгөөр үзүүлэв. $p = 0,5$ гэж сонгов. Урд дугуйнуудын траекторыг "мурилзалт"багассаныг ажиглаж болно.

Дифференциал роботын хувьд хойд дугуй (жишээ нь D)-ны траектор өгөгдвөл урд дугуйг (жишээ нь A) -ны траекторыг (1)-р хялбар олж чадна. Харин эсрэгээр урд дугуйны траекторoос хойд дугуйны траектор тооцолох үед (1) нь нэгдүгээр эрэмбийн шугаман бус дифференциал тэгшитгэл болж хувирна. Манай аргын хувьд бид дээр дурдсан процедурын 3)-р алхам дээр урд дугуй A-н гөлгөр траекторoос бусад дугуйн (жишээ нь D) траекторыг тооцоолох шаардлага үүснэ. Энэ үед (1) томъёо нь доорхи дифференциал тэгшитгэл болно:

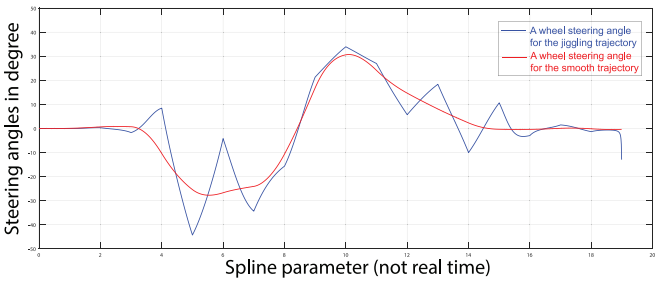
$$\begin{cases} f_1(t) = y_1(t) + \frac{Ly_1(t)'}{\sqrt{(y_1(t)')^2 + (y_2(t)')^2}} \\ f_2(t) = y_2(t) + \frac{Ly_2(t)'}{\sqrt{(y_1(t)')^2 + (y_2(t)')^2}} \end{cases} . \quad (10)$$

Энэд $[f_1(t), f_2(t)]$ нь урд A дугуйны траектор болох $\vec{r}_A(t)$-н x, y координатууд, $[y_1(t), y_2(t)]$ нь хойд D дугуйны траектор болох $\vec{r}_D(t)$-н x, y координатууд. (10)-г Рунге-Кутта аргаар (Matlab орчинд *ode45* функцээр) бодуулав. Гарч ирсэн үр дүнг Зураг 5(b)-д үзүүлэв. Урд дугуйны хувьд оригинал траектор (хөх өнгийн зам) болон "гөлгөр"траекторууд(улаан өнгийн зам) нилээд зөрж байгааг ажиглаж болно. Гэвч хойд дугуйны хувьд оригинал траектор (хар өнгийн зам) болон "гөлгөр"траекторууд(ягаан өнгийн зам) маш бага зөруутэй байна. Дифференциал роботын хувьд хойд дугуйгаар зам төвлөх нь хялбар боловч урд дугуйны "мурилзалтыг"багасгах нь хэцүү бөгөөд гараар зурж байгаа үед бараг боломжгүй гэдэг нь харагдаж байна.

Урд дугуйны мурилзалт ихтэй замаар явах үед тулд урд дугуйны залуурдалтын өнцөг "зиг заг"ихтэй байна. Харин бидний санал болгож буй аргаар гөлгөр зам байгуулбал залуурдалтын өнцөг ч гөлгөр болж роботыг удирдахад хялбар болно. Энэ үр дүнг Зураг 6-д үзүүлэв. Оригинал "мурилзалт"ихтэй замаар үеийн залуурдалтын өнцөгийг хөх зураасаар, гөлгөр замаар явах үеийн залуурдалтын өнцөгийг улаан зураасаар үзүүлэв. Залуурдалтын өнцөгийн "зиг заг"сайн дарагдсан болохыг ажиглаж болно.



Зураг 5. Path planning using spline approximation for wheel A



Зураг 6. Steering angles of the front wheels with approximated spline for front wheel A

V. Сплайн муруйг ижил нумын урттай хэсгүүдэд хуваах нь

Дугуйт роботыг удирдахад зөв агшинд зөв удирдлагын дохио өгөх нь нэн чухал. Энэ асуудлыг шийдэхийн тулд дугуйнуудад ротари энкодер мэдрүүлийг суурилуулж өгвөл үр дүнтэй байдаг. Ротари энкодер нь талбайтай харьцангүй дугуйнуудын туулсан зам (нумын урт, arc length)-г хэмждэг. Иймд практик дээр дугуйнуудын траекторыг ижил нумын урттай хэсгүүдэд хуваах асуудал байнга тулгардаг. Энэ хэсэгт дугуйны траектор куб сплайнаар өгөгдсөн үед ижил нумын урттай хэсгүүдэд хуваах асуудлыг авч үзье. Траектор сплайн параметрээр өгөгдсөн үед абсолют хурд нь

$$\vec{r}_A(t) = \mathbf{A}t^3 + \mathbf{B}t^2 + \mathbf{C}t + \mathbf{D}, t \in [0;1] \quad (11)$$

$$\vec{v}_A(t) = \frac{d}{dt} \cdot \vec{r}_A(t) = 3\mathbf{A}t^2 + 2\mathbf{B}t + \mathbf{C} \quad (12)$$

$$v_A(t) = \sqrt{\vec{v}_A(t) \cdot \vec{v}_A(t)} = \sqrt{9A^2t^4 + Ft^3 + Gt^2 + Ht + C^2} \quad (13)$$

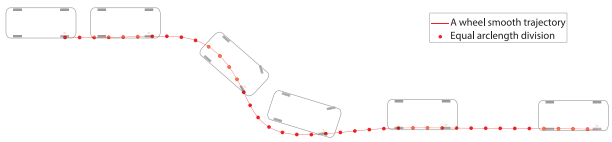
болно. Энэд $A = ||\mathbf{A}||$, $F = 12\mathbf{A} \cdot \mathbf{B}$, $G = 4||\mathbf{B}|| + 6\mathbf{A} \cdot \mathbf{C}$, $H = 4\mathbf{B} \cdot \mathbf{C}$, $||\mathbf{C}||$, $\cdot$ нь Евклидийн дотоод үржвэр. Тэгвэл нумын уртыг сплайн параметрээр илэрхийлбэл

$$s(t) = \int_0^t |v_A(\tau)| d\tau = \int_0^t \sqrt{9A^2\tau^4 + F\tau^3 + G\tau^2 + H\tau + C^2} d\tau$$

болно. Бидэнд сплайн параметр нь нумын уртаас яаж хамаарч байгаа нь хэрэгтэй тул урвуу функцийн теоремийн дагуу

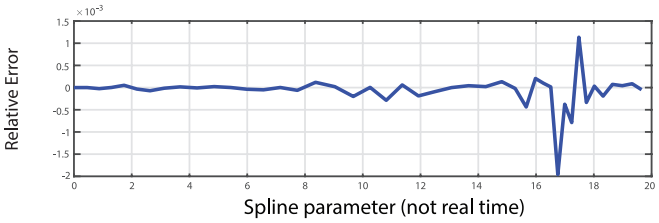
$$t(s)' = \frac{1}{\sqrt{9A^2t(s)^4 + Ft(s)^3 + Gt(s)^2 + Ht(s) + C^2}} \quad (14)$$

гэсэн нэгдүгээр эрэмбийн шугаман бус дифференциал тэгшитгэл гарна. Үүнийг Рунге-Кутта аргаар (Matlab орчинд *ode45* функцээр) бодуулав. Үр дүнг Зураг 7-д үзүүлэв. Улаан цэгүүдээр A дугуйны сплайн муруйг ижил нумын урттай сегментүүдэд хуваав.



Зураг 7. Partitioning of the trajectory $\vec{r}_A(t)$

Нумын хуваалтын алдааг Зураг 8-д үзүүлэв. Хуваалтын харьцангүй алдаа $2 \cdot 10^{-3}$ дотор байгааг ажиглаж болно.



Зураг 8. The relative error of the partition

VI. Дүгнэлт

Энэхүү ажлаар дифференциал роботын "мурилзалт"багатай гөлгөр замыг сплайн муруй ашиглан төлөвлөх аргыг санал болгов. Ингэхдээ урд дугуйны замын хувьд мурийлт болон квадрат алдаанаас хамаарсан зардлын функц байгуулан оптимизаци хийж бодсон. Санал болгож буй аргын үнэлгээг тооцон бодох аргаар шалгасан.

Ном зүй

[1] S.Kuindersma, R.Tedrake, et al ,"Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot," Autonomous Robots., Vol 40, pp. 429–455, 2016.

[2] Y.Wang, H.Li, et al, "A Convergence Guaranteed Multiple-Shooting DDP Method for Optimization-Based Robot Motion Planning,"IEEE Transactions on Industrial Electronics, Vol. 72, No. 5, pp. 5001–5011, 2025.

[3] Carl de Boor *A practical Guide to Splines*, Springer. 2001

Маркетингийн идэвхжүүлэлтэд мэдлэгийн граф ашиглах нь

Нандин-Эрдэнэ Энхмягмар
Мэдээлэл, Компьютерын Ухааны Тэнхим
Мэдээллийн Технологи,
Электроникийн Сургууль, Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол Улс
22M1NUM0289@stud.num.edu.mn

Энхтуул Бөхсүрэн
Мэдээлэл, Компьютерын Ухааны Тэнхим
Мэдээллийн Технологи,
Электроникийн Сургууль, Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол Улс
enkhtuul@seas.num.edu.mn

Мөнхцэцэг Намсрайдорж
Мэдээлэл, Компьютерын Ухааны Тэнхим
Мэдээллийн Технологи,
Электроникийн Сургууль, Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол Улс
munkhtsetseg@seas.num.edu.mn

Хураангуй - Бид энэхүү өгүүлэлдээ бөөний худалдааны бизнесийн байгууллага борлуулалтын орлогыг сайжруулах зорилгоор маркетингийн идэвхжүүлэлтийн аргууд ашигласан өгөгдөл дээр суурилан хугацааны явцад хэрхэн орлогын өөрчлөлтийг хянаж, сайжруулж, удирдаж болох боломжтойг судлах зорилгоор мэдлэгийн граф бүхий граф өгөгдлийн сан үүсгэж машин сургалтын арга ашиглан үнэлгээ хийсэн ажлыг танилцууллаа.

Түлхүүр үг: **Онтологи; SPARQL; GraphDB; Борлуулалт; Маркетинг; Машин сургалт ; Гүн сургалт; Graph Convolutional Network**

I. Удиртгал

Бизнесийн байгууллагын гол зорилго нь ашиг олох бөгөөд үүний тулд борлуулалтын орлогод нөлөөлөгч хүчин зүйлүүдийг тооцох шаардлага үүсдэг.

Өнөө цагт технологийн үсрэнгүй хөгжил, цахим хэрэглээний өсөлт, хиймэл оюун ухаан, өгөгдөлд суурилсан шийдвэр гаргалт зэрэг нь борлуулалтад нөлөөлөх маркетингийн стратеги боловсруулах арга барилыг өөрчлөхөд хүргэж байна. Үүнтэй холбоотойгоор маркетингийн үйл ажиллагааг илүү үр дүнтэй болгох, өгөгдөлд суурилсан анализ хийх улмаар ирээдүйн борлуулалтыг таамаглах нь байгууллагын бизнесийн шийдвэр гаргалтад улам бүр чухал нөлөөтэй болж байна.

Таамаглалын загвар сонгоход техникийн нөөц болон зардал, өнгөрсөн үеийн мэдээлэл бүхий өгөгдөл, өгөгдлийн хэлбэлзэл ба тогтвортой байдал, нарийвчлал, цаг хугацаа [1]-г авч үздэг. Бөөний худалдааны салбарт эрэлт, хэрэгцээг таамаглахад, хугацаан цувааны таамаглалын загварууд нь хялбар хэрэгжилттэй, өрсөлдөгч загваруудтай харьцуулахад илүү сайнаар тайлбарлагдах бөгөөд харьцангуй нарийвчлалтай байдаг давуу талтай.

II. Холбогдох онол

A. *Маркетингийн идэвхжүүлэлт таамаглах загвар*

Граф өгөгдөл дээр тулгуурлан орой, ирмэгийн шинж чанарыг хөршүүдийнх нь мэдээлэлд үндэслэн таамаглахдаа Граф нейрон сүлжээ (GNNs)-ийн нэг төрөл болох Graph Convolutional Network (GCN)-ийг ашигладаг.

GCNConv давхарга нь дараах загвар дээр суурилна:

$$H^{(I+1)} = \sigma(\hat{D}^{-1/2}\hat{A}\hat{D}^{-1/2}H^{(I)}W^{(I)}) \qquad (1)$$

Энд:

- $H^0 = X$ – *node-ийн шинж чанар (one hot vector)*
- $\widehat{A} = A + I$ – *хөршийн матриц + identity (өөртэйгөө холбоо)*
- $\widehat{D} = \widehat{A}$ – *ийн зэрэг (degree) матриц*
- $W^{(I)}$ – *сургалтын жин*
- σ – *идэвхжүүлэх функц (RelU)*

Таамаглалын загвар нь дараах байдалтай:

$$\hat{y}_i = \arg\max_c Z_{i,c} \qquad (2)$$

B. *Борлуулалтын орлогыг таамаглах загвар*

Нэг хамааралтай хувьсагч (y) ба хоёр ба түүнээс дээш тооны хамааралгүй хувьсагч (x_i)-ийн хоорондын шугаман хамаарлыг тогтоохдоо олон хүчин зүйлийн регрессийн загвар байгуулдаг.

Олон хүчин зүйлийн регрессийн эх олонлогийн загвар нь дараах байдалтай:

$$Y = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \cdots + \beta_p * x_p + \varepsilon \qquad (3)$$

Энд: Y нь хамаарах хувьсагч, $x_1, x_{2,\ldots,}x_p$ – *ул хамаарах хувьсагчууд,* β_0 – сул гишүүн, $\beta_1, \beta_2, \ldots, \beta_p$ түүврийн өгөгдөлд тулгуурлан тооцоолсон өнцгийн налалтууд, ε – түүврийн санамсаргүй алдаа бөгөөд Y-ийн дундаж утгаас хазайх утгыг харуулдаг. Алдааны вариаци нь хэвийн тархалттай үед 0 байна.

Түүврийн регрессийн загвар нь:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \ldots + b_kx_k \qquad (4)$$

Энд $\hat{y}$- үнэлэгдсэн утга, b_0 – сул гишүүний үнэлгээ, $b_1, b_2, \ldots, b_k$ өнцгийн коэффициентуудын үнэлгээ

$$e = (y - \hat{y}) \qquad (5)$$

Детерминацын коэффициент нь авч үзэж буй үл хамаарах хувьсагчид нийлээд y-ийн өөрчлөлтийн хэдэн хувийг тайлбарлаж байгааг илэрхийлсэн ойлголт юм.

$$R^2 = \frac{SSR}{SST} = \frac{\text{Регрессийн квадрат нийлбэр}}{\text{Нийт квадрат нийлбэр}} \qquad (6)$$

Байгуулсан загварын статистикийн ач холбогдолтой эсэхийг шалгахдаа F-Test-ийг ашиглана.
Хамааралгүй хувьсагчдын хооронд маш өндөр корреляц хамаарал оршин байгаа эсэхийг шалгахын тулд Variance

Inflation Factor (VIF) –коллиенарийг хэмжигч үзүүлэлтийг тооцсон.

$$VIF_j = \frac{1}{1-R_j^2} \quad (7)$$

Регрессийн загвараас тодорхойлогдох алдаа:

$b_1, b_2, \ldots, b_k$ өнцгийн коэффициентуудын үнэлгээ

C. Онтологи

Онтологи гэдэг нь тодорхой сэдвийн хүрээнд ойлголт, зүйлсийн бүтэц, харилцааг тодорхойлдог бөгөөд авч үзэж буй зүйлсийг ангилж, тэдгээрийн хоорондох харилцааг тайлбарлаж өгдөг. Ингэснээр компьютер болон бусад системүүд эдгээр ойлголтуудыг илүү сайн ойлгож, ажиллах боломжтой болдог.

Онтологийг хиймэл оюун ухаан, семантик вэб, программ хангамж хөгжүүлэлт, өгөгдлийн сангийн менежмент зэрэг чиглэлээр мэдлэг солилцох, нэгтгэх, учир шалтгааныг тодорхойлох зорилгоор ашигладаг [3]-[5].

Онтологи нь мэдээллийг системтэйгээр зохион байгуулж, холбох замаар өгөгдлийн харилцан үйлчлэлийг сайжруулж, шийдвэр гаргах процессыг сайжруулж, автоматжуулсан системүүдийг хөгжүүлэхэд тусалдаг [6].

C. Resource Description Framework

Энэ нь интернэт дэх мэдээллийг тодорхойлох, зохион байгуулах, хуваалцахад ашигладаг стандарт формат бөгөөд семантик вэбийн үндэс суурь болдог.

RDF нь мэдээллийг "Субъект-Үйлдэл-Объект" (subject-predicate-object) гэсэн бүтэцтэйгээр илэрхийлдэг [7]. Түүнчлэн, онтологи, семантик вэб болон бусад мэдээллийн технологиудад чухал үүрэгтэй бөгөөд мэдээллийг илүү ухаалаг, зохион байгуулалттайгаар ашиглахад тусалдаг.

D. OWL and SPARQL

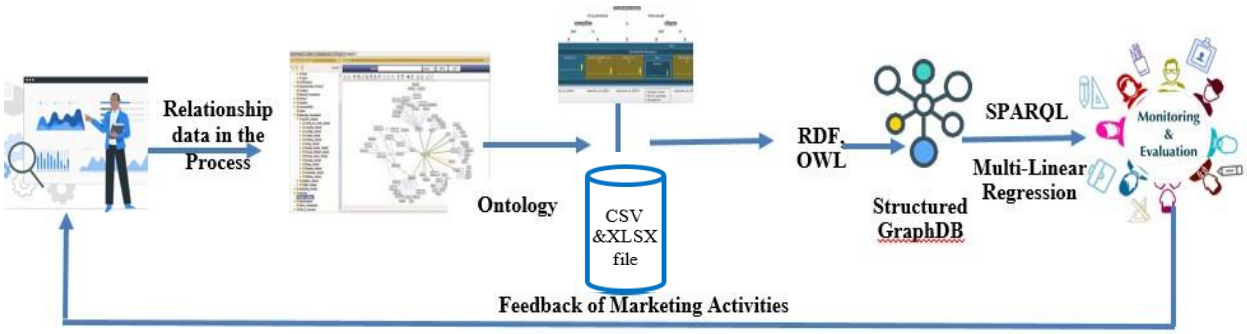
Web Ontology Language (OWL) нь юмс, зүйлсийн хоорондын харилцааны талаарх мэдлэгийг илэрхийлэхэд зориулагдсан семантик вэб хэл юм. OWL нь RDF, RDFS, SPARQL зэргийг багтаасан W3C-ийн семантик вэб технологийн нэг хэсэг юм. SPARQL нь RDF өгөгдлийг асуух, удирдах, боловсруулах зориулалттай стандартчилагдсан хэл юм [8].

III. Судалгааны арга зүй

Бөөний худалдааны борлуулалтын орлогыг хянах, сайжруулах, удирдах, таамаглах үйл ажиллагааг мэдээллийн технологийг ашиглан хэрэгжүүлэхдээ бид дараах алхмуудыг гүйцэтгэсэн.

1. Дистрибьютер компанийн хугацаанд суурилсан өгөгдөлд шинжилгээ хийж өгөгдлийг бэлтгэсэн.
2. Байгууллагын борлуулалтын орлогыг сайжруулахад зориулсан маркетингийн үйл ажиллагаанууд, түүнтэй хамааралтай өгөгдлийг цуглуулж бэлтгэсэн.
3. Дээрх өгөгдлүүдийг шинжилсний үндсэн дээр онтологи байгуулсан.
4. Байгуулсан онтологийг ашиглан граф өгөгдлийн сан үүсгэсэн.
5. Граф өгөгдлийн сангаас байгууллагын үйл ажиллагаанд шаардлагатай мэдээ, мэдээллийг өгөх зорилготой SPARQL асуулга бичиж мэдээллийг мониторинг хийх дэд бүтцийг үүсгэсэн.
6. Маркетингийн хандлага дээр суурилсан хямдрал, урамшуулал таамаглах шинжилгээ болон борлуулалтын орлогыг хамгийн их байлгахад нөлөөлж байгаа маркетингийн үйл ажиллагааг тодорхойлох зорилгоор олон хувьсагчийн регрессийн шинжилгээг машин сургалтын аргаар боловсруулсан.

Бид Зураг 1-т үзүүлсэн загварыг байгуулан ажилласан.



Зураг 1: Орлогыг нэмэгдүүлэх онтологид суурилсан загвар.

IV. Туршилтын үр дүн

4.1 Оршин байгаа өгөгдлийн шинжилгээ

Бид онтологио зохиомжлохдоо маркетингийн идэвхжүүлэлт нь борлуулалтын орлогод үзүүлэх нөлөөллийг судлах, улмаар ирээдүйн борлуулалтын орлогыг таамаглах зорилготой байсан. Ингэхдээ, дистрибьютер компанийн 3 жилийн (2022,2023,2024) захиалгын мэдээлэл, борлуулалтын тоо ширхэг, дүн, брэнд, бүтээгдэхүүний нэр, бүтээгдэхүүний код, бүтээгдэхүүний үнэ, харилцагчийн мэдээлэл, борлуулалтын ажилтны мэдээлэл, борлуулалтын суваг, маркетингийн идэвхжүүлэлтийн төрөл, түүнд холбогдох зардал, идэвхжүүлэлтийн оролцооны түвшин зэрэг талбаруудыг агуулсан 6 файлаас өгөгдлөө авсан. Эдгээр өгөгдлүүдээс маркетингтай холбоотой өгөгдөл дээр идэвхжүүлэлтийн төрөл тус бүрээр илүү нарийвчлан онтологи тодорхойлсон гэдгээрээ онцлогтой.

4.2 Маркетингийн идэвхжүүлэлт

Онтологи нь маркетингийн үйл ажиллагаа, борлуулалтын орлогын хоорондох харилцааг байгуулах бөгөөд энэ нь цаашдаа олон хүчин зүйлийн регрессийн шинжилгээ хийх оролтын өгөгдлийг илүү үр дүнтэй боловсруулахад тусална. Өөрөөр хэлбэл, Борлуулалт нь онтологийн түвшинд тухайн сарын "сурталчилгаа" болон "хямдрал"-ын өгөгдлүүдтэй хэрхэн холбоотой болохыг тодорхойлно.

Дистрибьютер компанийн онцлог нь эцсийн хэрэглэгч болон харилцагч (супермаркет, дэлгүүр г.м) нарт аль алинд нь чиглэсэн маркетингийн үйл ажиллагаа явуулдаг. Тухайлбал, телевизийн зар сурталчилгааг ихэвчлэн эцсийн хэрэглэгчид зориулж бүтээдэг бол захиалгын үнийн дүнгийн урамшуулал, хямдрал зэргийг ихэвчлэн харилцагч нарт санал болгодог. Уг компанийн хувьд маркетингийн үйл ажиллагаануудыг сар бүр брэнд тус бүрээр явуулдаг.

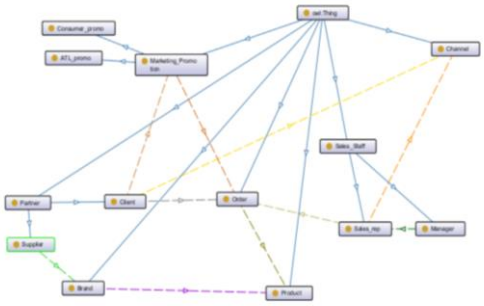
Маркетингийн үйл ажиллагаа нь төрөл болон хугацаа, оролцоо зэргээс хамааранд зардлууд харилцан адилгүй бөгөөд зардлыг төрөл болон сар бүрээр тооцоолсноор борлуулалтын орлогоос хамааруулан ямар төрлийн идэвхжүүлэлт нь илүү борлуулалт өсөхөд нөлөөлж байгааг судлах боломжтой болж байгаа юм.

4.3 Онтологи болон граф өгөгдлийн сан байгуулах

Семантик вэб, онтологи ба мэдлэгийн графийг боловсруулж ашигласнаар бөөний худалдааны бизнесийн гол асуудал болох борлуулалтын орлогыг сайжруулахад анхаарал хандуулах боломжтой. Бид борлуулалтад нөлөөлөх гол ойлголтуудыг нэгтгэсэн онтологийг боловсруулж өмнөх өгөгдлийг ашигласан граф өгөгдлийн сан үүсгэн дараагийн борлуулалтад юун дээр анхаарах вэ гэдгийг түргэн шуурхай асуулгын тусламжтайгаар мэдэх боломжтой.

Борлуулалтын орлогыг үнэлэх боломжтой мэдлэгийн графыг байгуулахдаа бид дараах алхмуудыг гүйцэтгэсэн. Үүнд:

1. Үйл ажиллагааны нэгж (entity)-уудыг тодорхойлох
2. Оршин байгаа онтологийг өргөтгөх эсвэл шинэ онтологийг хөгжүүлэх
3. Класс болон дэд класс тэдгээрийн хоорондын холбоо, шинж чанаруудыг тодорхойлох
4. Борлуулалтын өгөгдлийг RDF рүү хөрвүүлэх
5. Граф өгөгдлийн сан байгуулах
6. Загвар байгуулж үнэлгээ хийх



Зураг 2: Онтологийн бүтэц

Манай онтологийн хувьд нийтдээ Класс - 13, Объект проперти - 11 болон Дата проперти - 28 гэсэн бүтцээр зохиомжилсон. Зураг 2-т онтологи классын холбоосыг, Зураг 3-т класс хоорондын харилцааны жишээг харууллаа.

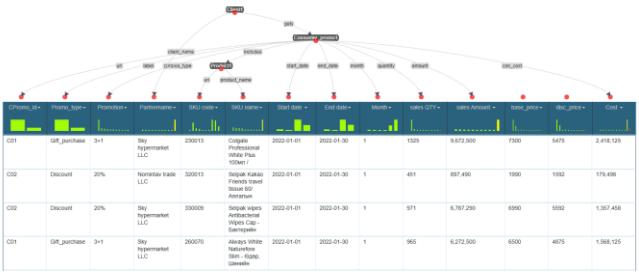
Class	Object property (relationship)	Class
Client	locatedIn	Location
Client	gets	Marketing_Promotion
Client	makes	Order
Order	includes	Product
Marketing_Promotion	contains	Product
Brand	has	Product
Client	belongsTo	Channel
Manager	supervises	Sales_rep
Supplier	supplies	Brand

Зураг 3: Классуудын харилцан хамаарал

Graph DB байгуулах

А. Өгөгдөл нэгтгэл

Зохиомжилсон онтологийн дагуу өгөгдлөө нэгтгэх шаардлагатай тул бид “Karma” өгөгдөл нэгтгэлийн хэрэгсэл ашигласан.



Зураг 4: Өгөгдөл нэгтгэл

Karma ашигласнаар .csv болон бусад төрлийн өргөтгөл бүхий өгөгдлийг онтологийн дагуу RDF болгон хөрвүүлэх, R2RML загвар гарган авч дахин ашиглах зэрэг үйлдлүүдийг хялбар хийх боломжтой.

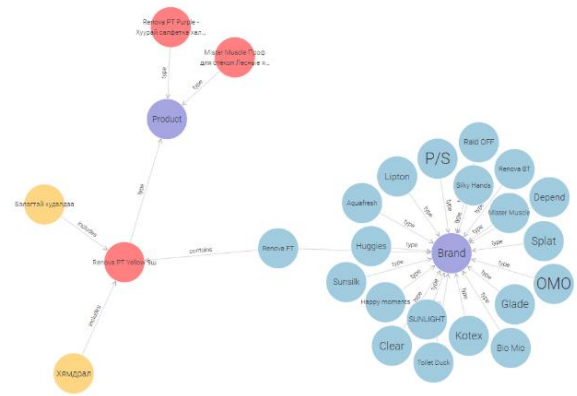
Эцсийн хэрэглэгчид зориулсан маркетингийн идэвхжүүлэлтийн өгөгдлийг нэгтгэсэн хэлбэрийг Зураг 4-д харууллаа.

В. Граф өгөгдлийн сан

Загварчилсан өгөгдлүүдийг RDF хэлбэрээр бүтэцлэн боловсруулж, GraphDB семантик өгөгдлийн санд төвлөрүүлэн хадгалсан.

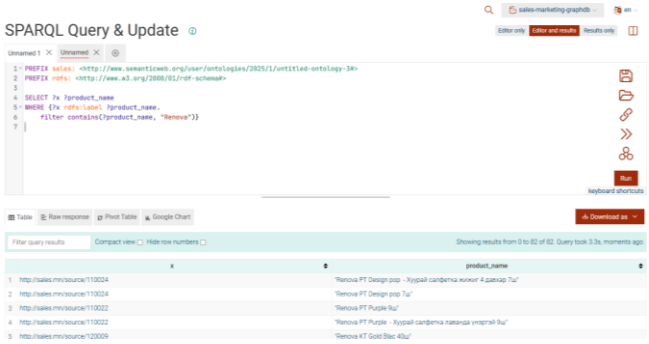
Зураг 5-д граф өгөгдлийн сангийн өгөгдлүүдийг визуал граф байдлаар харууллаа. Энд ижил өнгийн орой бүр нь нэг ижил утгыг агуулна. Жишээлбэл, бүдэг цэнхэр өнгийн оройнууд брэндүүдийг, шар өнгийн оройнууд маркетингийн идэвхжүүлэлтийн төрлийг илэрхийлнэ.

Зураг 5-д харуулснаар, компанийн хувьд хамгийн энгийн хэрэгцээ буюу тухайн бүтээгдэхүүн дээр ямар төрлийн маркетингийн идэвхжүүлэлт хийгдсэнийг нэгдсэн байдлаар харах нь дараагийн төлөвлөлтийг хийхэд чухал ач холбогдолтой байдаг.



Зураг 5: Граф өгөгдлийн визуал граф

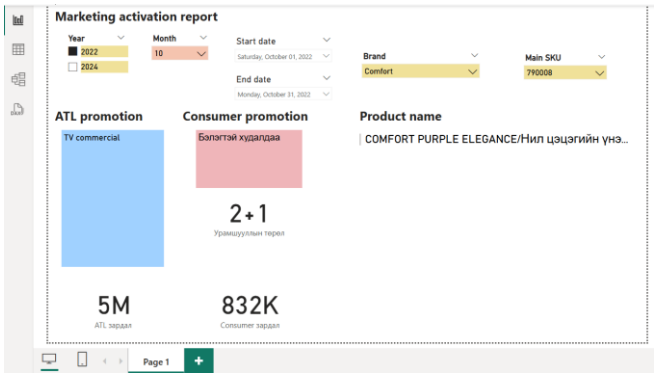
Граф өгөгдлийн сангаас SPARQL ашиглан бизнесийн байгууллагын өдөр тутмын үйл ажиллагаанд шаардлагатай бүх төрлийн мэдээллийг утгачилсан байдлаар хялбар аргаар гарган авах боломжтой. Зураг 6-д SPARQL ашиглан тухайн бүтээгдэхүүнтэй холбоотой бүх мэдээллийг харуулах хялбар асуулгын үр дүнг харууллаа.



Зураг 6: SPARQL асуулгын жишээ

С. Мониторинг

Граф өгөгдлөөс SPARQL асуулга бичин гаргаж авсан үр дүнг хэрэглэгчид илүү ойлгомжтой байдлаар харуулахын тулд өгөгдлийг мониторинг хийх аргыг ашигласан. Ингэхдээ, Power BI өгөгдөл визуалчлалын программыг ашиглан дараах байдлаар харуулсан. Зураг 7-д тухайн маркетингийн идэвхжүүлэлт хийсэн бүтээгдэхүүний нэр, код, брэнд, борлуулалт, идэвхжүүлэлтийн цаг хугацаа, идэвхжүүлэлтийн төрөл, зардал зэргийг нэгдсэн тайлан байдлаар харах боломжтой жишээ дашбоардыг харууллаа.



Зураг 7: Power BI дашбоард

Загварын үнэлгээг хийх

Graph дээрх өгөгдлүүдийг авч маркетингийн хандлага дээр суурилсан хямдрал, урамшуулал төлөвлөх боломжтой Граф нейрон сүлжээ (GNNs)-ий шинжилгээ болон маркетингийн идэвхжүүлэлтийн аргуудаас борлуулалтын орлогод хамгийн чухал нөлөө үзүүлж байгаа арга аль нь болохыг илрүүлэхийн тулд олон хүчин зүйлийн регрессийн шинжилгээ хийсэн.

А. Граф нейрон сүлжээ (Graph neural networks)

Бид үүсгэсэн граф өгөгдлөө ашиглан граф бүтэцтэй өгөгдөл дээр ажиллахад зориулагдсан гүн сургалтын загвар буюу Граф нейрон сүлжээ алгоритмаар харилцагч ямар төрлийн маркетингийн идэвхжүүлэлтэд хамрагдах магадлалтайг харилцагчийн дундаж борлуулалт, байршил, харилцагчийн ангилал, өмнө нь хамрагдаж байсан маркетингийн идэвхжүүлэлтийн түүх зэргээс хамааруулан тооцоолсон. Энэ тооцооллыг хийснээр тухайн харилцагчид ямар төрлийн маркетингийн идэвхжүүлэлт санал болгох талаарх шийдвэр гаргалтад чухал нөлөө үзүүлэх юм.

Ингэхдээ гүн сургалтын аргыг ашиглан Python программ дээр тооцооллоо хийсэн.

GNNs алгоритм нь граф өгөгдлийн орой, ирмэг, тэдгээрийн шинж чанаруудыг ашиглан графын өгөгдлийн дотоод холбоо, бүтэц, тархцыг ойлгож, шинжлэх боломжийг олгодог.

Зураг 8-т харилцагч ямар төрлийн маркетингийн идэвхжүүлэлтийн төрөлд оролцох магадлалтайг таамагласан шинжилгээний үр дүнг харилцагчийн граф дээрх URI болон таамагласан идэвхжүүлэлтийн төрөл гэсэн бүтцээр харууллаа.

Тухайлбал, 13639 дугаартай харилцагч Merchandise буюу өрөлтийн урамшуллын идэвхжүүлэлтэд хамрагдах магадлалтай учир тухайн харилцагчид таамагласан урамшуллыг санал болгох боломжтой. Энэ нь 13639 дугаартай харилцагчтай ойролцоо үнийн дүнтэй борлуулалт хийдэг, мөн ойролцоо байршдаг харилцагчид ихэвчлэн энэ төрлийн идэвхжүүлэлтэд хамрагддаг гэсэн үг юм.

Зураг 8: Маркетингийн идэвхжүүлэлтийн төрөл таамагласан үр дүн

В. Олон хүчин зүйлсийн регрессийн шинжилгээ

Олон хүчин зүйлийн регрессийн шинжилгээний хамгийн сайн тохирсон шугаман загварын параметрийг үнэлэхдээ машин сургалтын аргыг ашигласан.

Эхлээд бид хамааралгүй хувьсагчдын хооронд өндөр корреляц хамаарал оршин байгаа эсэхийг шалгасан. Ингэхийн тулд VIF_j –коллиенарийг хэмжигч үзүүлэлтийг тооцоход дараах үр дүн гарсан.

Хүснэгт 1: Коллиенарийг хэмжигч үзүүлэлт

	Хувьсагч	VIF
0	const	26.767809
1	Gift_purchase	1.470370
2	Discount	1.675928
3	Giveaway	1.672399
4	Gondola	2.394477
5	Influencer	1.630362
6	Merchandise	1.373275
7	Promoter	1.426720
8	Social_media	1.190947
9	TV_commercial	1.382248

Бидний авч үзсэн хувьсагчдын хувьд хоорондоо өндөр хамааралтай хувьсагчид байхгүй байгаа тул шинжилгээнд ашиглах боломжтой юм.

Python программ ашиглан олон хүчин зүйлийн регрессийг тооцоолсон.

```
...
X = df[['Gift_purchase','Discount', 'Giveaway', 'Gondola', 'Influencer', 'Merchandise', 'Promoter', 'Social_media', 'TV_commercial']]
y = df['Sales']
print(df.dtypes)
print(df.isnull().sum
X = sm.add_constant(X)
model = sm.OLS(y, X).fit()
print(model.summary())
...
```

Зураг 9-д олон хүчин зүйлийн регрессийн шинжилгээний үр дүн харууллаа.

Variable	Coefficient	Standard error	t	P > |t|
const	5.746e+09	9.18e+08	6.261	0.000
Gift_purchase	2.0480	0.841	2.434	0.029
Discount	15.9784	5.207	3.069	0.008
Giveaway	109.5655	107.840	1.016	0.327
Gondola	-140.7755	141.610	-0.994	0.337
Influencer	5.7321	81.448	0.070	0.945
Merchandise	-744.2826	627.490	-1.186	0.255
Promoter	10.3284	289.206	0.036	0.972
Social_media	20.5052	28.369	0.723	0.482
TV_commercial	42.0760	30.424	1.383	0.188

Зураг 9: Регресийн шинжилгээний үр дүн

Коэффициент нь тухайн хувьсагчийн утга нэгжид борлуулалт хэрхэн өөрчлөгдөхийг харуулдаг.

Бүх хувьсагчдын утга тэг байх үед борлуулалт ямар байгааг const coef харуулна. Тухайлбал, бүх маркетингийн үйл ажиллагаа явуулаагүй тохиолдолд борлуулалтын орлого 5.746 тэрбум байна.
Gift_purchase, Discount, TV_commercial хувьсагчид нь статистик ач холбогдолтой бөгөөд Giveaway, Gondola, Influencer, Merchandise, Promoter, Social_media нь p-value > 0.05 байгаа тул борлуулалтад нөлөө үзүүлдэггүй статистик ач холбогдолгүй хувьсагчид болох нь шинжилгээний үр дүнгээс харагдаж байна.

Борлуулалтын орлогыг таамаглах олон хүчин зүйлийн регрессийн загварыг дараах байдлаар байгууллаа.

Sales=5.746e+2.04*Gift_purchase+15.97*Discount+ 42.07* TV_commercial (8)

Байгуулсан загварын параметрүүдийг дараах байдлаар тайлбарлаж болно. Хямдралтай худалдаа болон ТВ-ийн зар сурталчилгааны зардал тогтмол байхад бэлэгтэй худалдааны зардлыг нэг нэгжээр нэмэгдүүлэхэд борлуулалтын орлого 2.04 нэгжээр нэмэгдэнэ, харин бэлэгтэй худалдаа, ТВ-ийн зар сурталчилгаанд зарцуулах зардал тогтмол байхад хямдралтай худалдааны зардлыг нэг нэгжээр нэмэгдүүлэхэд борлуулалтын орлого 15.97 нэгжээр нэмэгдэнэ. Бэлэгтэй болон хямдралтай худалдааны зардал тогтмол байхад ТВ-ийн зар сурталчилгаанд зарцуулах зардлыг нэг нэгжээр нэмэгдүүлэхэд борлуулалтын орлого 42.07 нэгжээр нэмэгдэнэ.

Бидний байгуулсан загварын F-statistic = 4.71, Prob (F-statistic) = 0.005 байгаа тул авч үзэж байгаа загвар статистик ач холбогдолтой гэсэн дүгнэлт гарч байна.

R-squared 0.752 гарсан нь нийт борлуулалтын орлогын 75.2%-ийг авч үзэж байгаа хувьсагчид тайлбарлаж чадаж байгааг харуулна.

Adjusted R-squared: Загварын хамааралгүй хувьсагчийн тоог тооцсоны үндсэн дээр бүх х-ийн у-д нөлөөлөх нөлөөлөл 59.2% гарч байгаа нь зарим хувьсагчид үр нөлөөгүй байж магадгүй байна гэсэн үг юм.

Үүнээс үзэхэд дистрибьютер компани өмнөх сарын өгөгдлүүд дээр тулгуурлан тухайн сард ямар төрлийн маркетингийн идэвхжүүлэлт хийж борлуулалтын орлогоо яаж нэмэгдүүлэх талаар урьдчилан таамаглаж шийдвэр гаргалт хийх боломжтой юм.

Регрессийн шинжилгээний үр дүнгийн хувьд авч үзвэл зардал өндөртэй маркетингийн үйл ажиллагаанууд статистик ач холбогдолтой буюу Бэлэгтэй худалдаа, Хямдрал, ТВ зар сурталчилгаа нь илүү борлуулалтад нөлөөлж байна гэсэн үр дүн гарсан. Гэвч энэ бол зөвхөн борлуулалтын үнийн дүнд шууд үзүүлж буй тоон нөлөөлөл юм. Дээр дурдсанаас бусад маркетингийн үйл ажиллагаа нь шууд бус байдлаар борлуулалтад нөлөөлдөг буюу брэндийн нэр хүнд, танигдалт, хоногшилт, хэрэглэгчийн зан төлөв, дадал зуршил, өрсөлдөгчөөс ялгарах ялгарал зэрэг нүдэнд үл үзэгдэх чанарын нөлөөллүүд байх боломжтой.

V. Дүгнэлт

Энэхүү судалгаанд маркетингийн идэвхжүүлэлтийн (хямдрал, урамшуулал, бэлэгтэй худалдаа) үр нөлөөг мэдлэгийн графын тусламжтайгаар илүү нарийвчлан шинжлэх, улмаар борлуулалтад үзүүлэх нөлөөг тооцох боломжийг судаллаа. RDF онтологи болон SPARQL асуулга ашиглан өгөгдлийг уялдуулан холбож, граф өгөгдлийн сан үүсгэснээр хэрэглэгчийн зан төлөв, кампанит ажлуудын уялдааг илүү гүнзгий түвшинд харуулах боломж бүрдсэн.

Үүний зэрэгцээ, олон хувьсагчийн регрессийн шинжилгээ болон граф нейрон сүлжээ (GNN)-ийг ашигласнаар маркетингийн үйл ажиллагаа борлуулалтад хэрхэн нөлөөлж байгааг тооцож, цаашид урьдчилан таамаглах загвар боловсруулах эх суурь тавигдсан. Судалгааны үр дүн нь маркетингийн шийдвэр гаргалтад дата дээр суурилсан, уялдаа хамаарлыг тусгасан шинэлэг хандлага нэвтрүүлэхэд хувь нэмэр оруулах боломжтойг харууллаа.

[1]. Herbig, P. a., Milewicz, J., & Golden, J. E. The do's and don'ts of sales forecasting. Industrial Marketing Management, 22(1), 1993, pp. 49–57.

[2]. Liu, L.-M., Bhattacharyya, S., Sclove, S. L., Chen, R., & Lattyak, W. J. Data mining on time series: an illustration using fast-food restaurant franchise data. Computational Statistics & Data Analysis, 2001, 37(4), pp.455–476.

[3]. Chaccour, C., Saad, W., Debbah, M., Han, Z., & Poor, H. V. Less Data, More Knowledge: Building Next Generation Semantic Communication Networks. IEEE Communications Surveys & Tutorials, 2024. doi:https://doi.org/10.1109/COMST.2024.3412852

Fu, C., Jiang, H., & Chen, X. RETRACTED ARTICLE: Modeling of an Enterprise Knowledge Management System Based on Artificial Intelligence. Knowledge Management Research & Practice, 2022, pp.1–13. doi:https://doi.org/10.1080/14778238.2020.1854632

[4]. Olan, F., Arakpogun, E. O., Suklan, J., Nakpodia, F., Damij, N., & Jayawickrama, U. Artificial intelligence and knowledge sharing: Contributing factors to organizational performance. Journal of Business Research, 145, 2022, pp. 605-615. doi:https://doi.org/10.1016/j.jbusres.

[5]. Staab, S.; Studer, R. Handbook on Ontologies; Springer Science & Business Media: Berlin, Germany, 2013.

[6]. De Nicola, A., & Villani, M. L. Smart City Ontologies and Their Applications: A Systematic Literature Review. Sustainability, 2021, 13(10). doi:https://doi.org/10.3390/su13105578

[7]. Gruber, T.R. A translation approach to portable ontology specifications. Knowl. Acquis. 1993, 5, pp.199–220.

[8]. "SPARQL Query Language for RDF." Available: https://www.w3.org/TR/rdf-sparql-query/.

МОНГОЛ БИЧВЭРИЙГ ХУРААНГУЙЛАХ АРГЫН СУДАЛГАА

Бөхбатын Амартүвшин
Мэдээлэл технологийн салбар
ШУТИС, Мэдээлэл холбоо технологийн сургууль
Улаанбаатар хот
B231910004@must.edu.mn

Нямын Гантөмөр
Мэдээлэл технологийн салбар
ШУТИС, Мэдээлэл холбоо технологийн сургууль
Улаанбаатар хот
B231960002@must.edu.mn

Бямбадоржийн Золзаяа
Мэдээлэл технологийн салбар
ШУТИС, Мэдээлэл холбоо технологийн сургууль
Улаанбаатар хот
b.zolzaya@must.edu.mn

***Хураангуй*— Сүүлийн жилүүдэд цахим орчинд маш их хэмжээний текстэн мэдээлэл хуримтлагдаж, эдгээр мэдээллээс үнэ цэнтэй, голлох агуулгыг шуурхай оновчтой гарган авах хэрэгцээ нэмэгдэж байна. Текстэн хэлбэрийн мэдээлэлд буй гол санааг эх хэлний боловсруулалтын аргын тусламжтайгаар хураангуйлах нь цаг хугацааг тодорхой хэмжээгээр хэмнэх, ажлын бүтээмжийг нэмэгдүүлэх юм. Текст хураангуйллыг хэрэгжүүлэхэд хоёр төрлийн арга байх бөгөөд хийсвэр текст хураангуйлал (Abstractive text summarization), экстрактив текст хураангуйлал (Extractive text summarization) аргуудаар хэрэгжүүлдэг. Энэ удаад бид экстрактив текст хураангуйллын аргыг монгол хэл дээр хэрэгжүүлсэн ба экстрактив текст хураангуйллыг хэрэгжүүлэхэд тооцооллын нөөц, өгөгдлийн нөөц бага шаарддаг, хураангуйлалд эх текстийн өгүүлбэрийг сонгох ба ямар нэгэн хийсвэрлэл байхгүй нь зарим тохиолдолд зохимжтой жишээлбэл: нийтлэл, шинжлэх ухааны өгүүлэл. Бид экстрактив текст хураангуйллыг хэрэгжүүлэхдээ граф дээр суурилсан TextRank, LexRank, статистик дээр суурилсан TF-IDF алгоритмуудыг ашиглан нэмэлт урьдчилсан боловсруулалтын (preprocessing) phrase болон дараах боловсруулалтын (postprocessing) sentence positioin сайжруулалтуудыг нэмэн хэрэгжүүлэв. Туршилтын үр дүнд тухайн phrase болон sentence position нэмэлт сайжруулалтын давхаргатай үед энгийн үеэс илүү сайн үзүүлэлт гаргасан. Энэхүү судалгаа нь текстийн голлох агуулгыг үр ашигтайгаар гаргаж, текстэн мэдээлэлтэй харьцдаг текстэн мэдээлэлтэй ажилладаг бүх хүний цаг хугацааг хэмнэх зорилготой.**

Түлхүүр үг— LexRank, TextRank, TF-IDF, mBert, rouge-score, sentence position, phrase

I. Удиртгал

Орчин үеийн мэдээллийн эрин зуунд асар их хэмжээний текст мэдээлэл өдөр бүр үүсэж, хуримтлагдсаар байна. Энэ их мэдээллийн урсгалаас хэрэглэгчдэд богино хугацаанд гол санаа, үндсэн агуулгыг нь хүргэх шаардлага өсөн нэмэгдэж байна.

Текст хураангуйлал нь энэ асуудлыг шийдвэрлэх үр дүнтэй арга замын нэг юм. Үүнийг хийсвэр (abstractive) болон экстрактив (extractive) гэсэн хоёр аргаар хэрэгжүүлэх боломжтой. Монгол хэл дээрх өгөгдлийн хомс байдал, мөн хийсвэр хураангуйллын хувьд өмнө нь сургагдсан загваруудыг нутагшуулах, техник сургагдсан загваруудыг нутагшуулах, техник хангамжийн хязгаарлагдмал байдал зэрэг нь тодорхой хүндрэлтэй тул бид судалгаандаа экстрактив хураангуйллыг түлхүү ашиглаж, Монгол хэл дээр текст хураангуйллыг үр ашигтай хэрэгжүүлэх боломжийг судалсан.

Нэгт экстрактив хураангуйлах арга нь эх бичвэрийн агуулгыг өөрчлөхгүйгээр, эх текстээсээ шууд өгүүлбэрүүдийг сонгож хураангуйлдаг ба ямар нэгэн хийсвэрлэл, товчлол агуулдаггүй тул хэрэглэгчдэд эх мэдээллийг найдвартай хүргэдэг. Жишээлбэл албаны баримт бичиг, эрдэм шинжилгээний өгүүлэл гэх мэт албан ёсны найруулгатай текстүүдийн гол санаа, өгүүлбэр, үгийг өөрчлөхгүй өөрийн байгаагаар нь хураангуйг гаргадгаараа давуу тал болж байна.

Хоёрт энэхүү арга нь хийсвэр хураангуйлах (abstractive text summarization) аргатай харьцуулбал тооцооллын бага нөөц шаардах ба энгийн хялбар хэрэглээнд хэрэглэхэд илүү тохиромжтой байдлаараа давуу талтай.

Гэсэн хэдий ч, экстрактив хураангуйллын үр дүнгийн чанар нь ашигласан алгоритм, оновчлолын аргачлалаас хамаарч өөр өөр байж болно. Иймд, бид судалгаандаа Монгол хэлний онцлогт тохирсон экстрактив хураангуйллын аргыг боловсруулж, үр дүнг тодорхойлохыг зорьсон бөгөөд улмаар хэрэглэгчдэд хэрэгтэй мэдээллийг оновчтой, үр дүнтэй хүргэхэд чиглэгдэж байна.

Энэхүү судалгааны ажил нь дараах гол асуудлуудыг авч үзнэ:

1. Экстрактив хураангуйллын аргачлал, алгоритмуудын харьцуулалт
2. Монгол хэлэнд тохирох оновчтой аргыг сонгон турших, үр дүнг үнэлэх
3. Тухайн домэйнд тохирсон аргуудыг тодорхойлох

Ингэснээр бид Монгол хэл дээрх мэдээллийн боловсруулалтын судалгаанд хувь нэмэр оруулахын зэрэгцээ экстрактив хураангуйллын практик хэрэглээг сайжруулах боломжтой болно.

Талархал

Энэхүү судалгааны ажлын явцад A100 GPU (Graphic Computing Unit) тооцоологчоор хангасан, ШУТИС, Мэдээлэл холбоо технологийн сургуулийн, Мэдээлэл холбооны салбарын Доктор Дэд Профессор багш Б.Золзаяа танд гүнээ талархаж байна.

II. Ижил төстэй ажлуудын судалгаа

Монгол хэл дээрх текст хураангуйллын зорилго нь их хэмжээний текст өгөгдлөөс гол утга санааг алдагдуулалгүйгээр товч, тодорхой байдлаар хураангуйлан гаргаж авах явдал юм. Текст хураангуйллын олон төрлийн арга байдаг бөгөөд [1]-д хамгийн түгээмэл хэрэглэгддэг TF-IDF (Text Frequency-Inverse document frequency), TextRank, LexRank, мөн урьдчилсан сургагдсан загваруудыг ашиглах талаар өгүүлсэн.

Мөн [2]-д өгөгдлийн урьдчилан боловсруулах аргачлал, түүнээс гол утга агуулсан өгүүлбэрүүдийг сонгох тухай судалгаа хийгдсэн. Харин [3]-[5] судалгааны ажлуудад Монгол хэлтэй ижил бага хэмжээний өгөгдлийн сантай хэлнүүдэд дээрх аргууд хэрхэн үр дүнтэй хэрэгжсэнийг нарийвчлан судалсан. Ихэнх текст хураангуйллын судалгаа нь Англи болон бусад олон хэрэглэгчтэй хэл дээр хийгдсэн байдаг боловч бага хэрэглэгчтэй хэлнүүд дээр үр дүнтэй хэрэгжсэн нь Монгол хэлэнд ч амжилттай ашиглах боломжтойг харуулж байна.

Эдгээр судалгааны ажлуудыг үндэслэн, бид Монгол хэлний мэдээний өгөгдлийг ашиглан текст хураангуйллын аргуудыг харьцуулан судалж, аль арга нь хамгийн үр дүнтэй болохыг тодорхойлох зорилготой байна.

III. Ашигласан арга

A. *Тестийн өгөгдлийн багц*

Бид судалгааны ажлынхаа үр дүнг хэмжихдээ rouge-score хэмжүүрээр хэмжсэн ба өөрсдийн хэрэгжүүлсэн аргуудын гүйцэтгэлийг үнэлэх зорилгоор тестийн өгөгдлийн багцыг үүсгэсэн.

Тестийн өгөгдлийн багцыг үүсгэхдээ бид ижил төстэй судалгааны ажлуудын тестийн өгөгдлийн хэв маягийг судалсан ба тухайн судалгааны ажлуудад тестийн өгөгдлийн багц нь зөвхөн нэг домэйний олон жишээнүүд байсан.

Харин бидний судалгааны хувьд I – р (Удиртгал) хэсэгт дурдсан авч үзэх асуудлын хүрээнд бид тестийн өгөгдлийн багцыг үүсгэхдээ 3 өөр домэйнд хуваан үүсгэсэн.

Хүснэгт. I Тестийн өгөгдлийн багц

Өгөгдлийн багцын нэр	Баганын тоо	Мөрийн тоо
News	2	10
Paper	2	10
Social	2	10

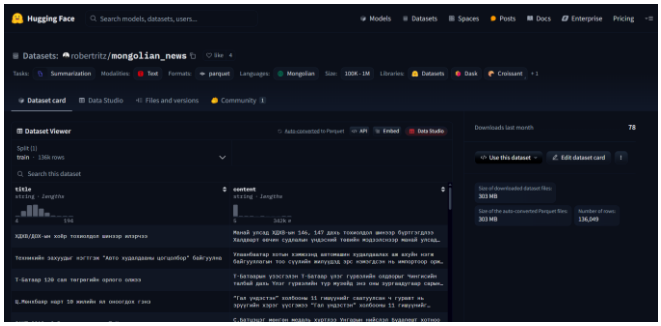
[a] Тестийн өгөгдлийн багцын хэмжээснүүд

Хүснэгт. I – д харуулснаар мэдээний, эрдэм шинжилгээний болон цахим орчны пост нийтлэл зэрэг агуулгатай 3 өөр төрлийн домэйний 10 жишээг бэлтгэсэн.

Эдгээр 3 төрлийн текст өгөгдөл нь харьцангуй их үгтэй, жишээлбэл эрдэм шинжилгээний өгүүллийн текст нь дунджаар 1600 үгтэй юм. Тиймээс тухайн нэг домэйнээс 10 жишээ үүсгэж тест хийх нь хангалттай гэж үзсэн.

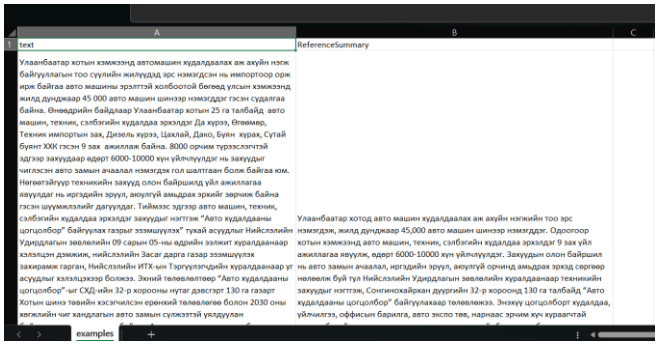
1) *Мэдээний өгөгдлийн багц*

Мэдээний агуулгатай өгөгдлийн багцыг бэлтгэхдээ hugging face дээрх монгол хэл дээрх мэдээний Өгөгдлийн багцаас санамсаргүйгээр 10 жишээ авсан.



Зураг. 1. Hugging Face дээрх мэдээний өгөгдлийн багц

HuggingFace дээрх өгөгдлийн багц нь content, title гэсэн атрибутуудтай бөгөөд бид content атрибутын утгуудыг авч тестийн өгөгдлийн багцын оролтын текст болгон ашигласан.

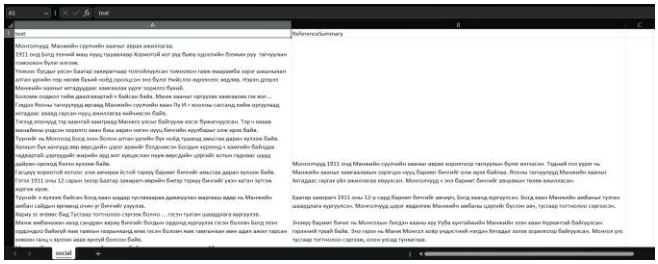


Зураг. 2. Мэдээний өгөгдлийн багцын жишээ

Дараах зурагт бидний бэлдсэн мэдээний өгөгдлийн багцын харагдах байдлыг харуулав.

2) *Цахим орчны өгөгдлийн багц*

Цахим орчны өгөгдлийн багцыг бэлтгэхдээ бид X, Facebook платформуудыг ашиглан пост нийтлэлийг оролтын текст болгон ашигласан.

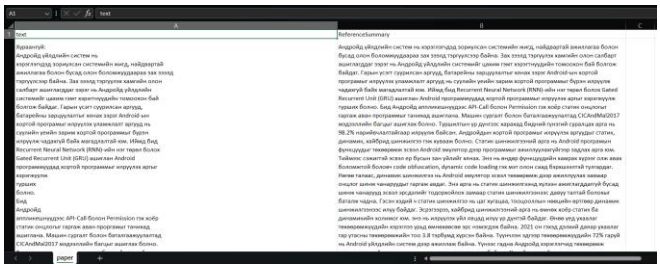


Зураг. 3. Цахим орчны өгөгдлийн багцын жишээ

3) *Эрдэм шинжилгээний өгүүллийн өгөгдлийн багц*

Эрдэм шинжилгээний өгөгдлийн багцыг бэлдэхдээ ШУТИС Мэдээлэл холбоо технологийн сургуулийн

эрдэм шинжилгээний эмхэтгэлээс эрдэм шинжилгээний өгүүллүүдийг оролтын текст болгон ашигласан.



Зураг. 4. Эрдэм шинжилгээний өгүүллийн өгөгдлийн багцын жишээ

B. Өгөгдлийн урьдчилсан боловсруулалт

Текст өгөгдлийг хураангуйлах үйл явцыг үр дүнтэй гүйцэтгэхийн тулд өгөгдлийг эхлээд урьдчилан боловсруулж, стандартчилах шаардлагатай. Өгөгдлийн урьдчилсан боловсруулалт нь анхны түүхий өгөгдлийг ашиглахад тохиромжтой, боловсруулсан хэлбэрт хувиргах үйл явц юм. Энэ нь өгөгдлийн чанарыг сайжруулж, боловсруулалтын нарийвчлалыг нэмэгдүүлэх ач холбогдолтой.

Уг үйл явц нь дараах үндсэн үе шатуудыг агуулна:

- Нормалчлах (Normalization) - Энэ нь текстийг стандарт хэлбэрт оруулж, боловсруулахад хялбар болгох үйл явц юм. Үүнд том жижиг үсгийн ялгааг арилгаж жижиг тэмдэгт рүү хувиргах, илүүдэл хоосон зай, цэг тэмдэгтүүдийг стандартчилах зэрэг алхмууд багтана.
- Токенжуулалт (Tokenization) - Текстийг үг, өгүүлбэр зэрэг жижиг хэсгүүдэд хуваах процесс. Энэ өгөгдлийг цааш боловсруулахад хялбар болгодог.
- Зогсоц үгийг хасах - “ба”, “болон”, “юм” гэх мэт утга агуулгын хувьд ач холбогдол багатай үгсийг хасах. Энэ нь өгөгдлийг оновчтой боловсруулалтад тусалдаг. Монгол хэлэнд зогсоц үгийн сан одоогоор байхгүй тул бид өөрсдийн судалгаанд хэрэглэх зорилгоор бага хэмжээний зогсоц үгийн сан үүсгэн хэрэглэсэн болно.

Эдгээр алхмуудыг хэрэгжүүлснээр өгөгдлийн хураангуйлах процесс илүү үр дүнтэй явагдаж, боловсруулсан мэдээлэл илүү чанартай, оновчтой болж, цаашдын боловсруулалт, анализ хийхэд тохиромжтой болно.

C. Өгүүлбэрийн ач холбогдлыг тооцоолох алгоритмууд

1) TF-IDF (Term frequency - Inverse document frequency)

TF-IDF нь эх хэлний боловсруулалтад ашигладаг статистик хэмжүүр бөгөөд тодорхой баримт бичиг дэх үгсийн ач холбогдлыг үнэлэхэд хэрэглэдэг. Энэхүү хэмжүүр нь тухайн бичиг баримт дахь давтамж болон нийт баримт бичгийн сан дахь ховор эсэхийг харгалзан үгийн ач холбогдлыг тодорхойлдог.

TF-IDF нь дараах томьёогоор илэрхийлэгдэнэ:

$$\text{TF-IDF}(t, d, D) = \text{TF}(t, d) * \text{IDF}(t, D) \quad (1)$$

TF(t, d) нь тухайн t үгийн d баримт бичиг доторх давтамж бөгөөд дараах томьёогоор илэрхийлэгдэнэ:

$$TF(t,d) = \frac{f_{t,d}}{\sum_{t'}^{d} f_{t',d}} \quad (2)$$

Энд $TF(t,d)$ нь тухайн t үгийн d баримт бичиг дэх давтамжийг илэрхийлнэ.

IDF(t, D) нь тухайн үг нийт баримт бичгийн санд хэр ховор байгааг хэмжих үзүүлэлт бөгөөд дараах томьёогоор тодорхойлогдоно:

$$TF(t,d) = log\left(\frac{|D|}{1 + |\{d \in D : t \in d\}|}\right) \quad (3)$$

Энд:

- $|D|$, нь нийт баримт бичгийн тоо.
- $|\{d \in D : t \in D\}|$, нь тухайн үг орсон баримт бичгийн тоо.

TF-IDF алгоритм нь олон болон бага давтамжтай үгсийг тодорхойлох замаар текстийн онцлог шинж, утга санааг илэрхийлсэн үг болон өгүүлбэрүүдийг илрүүлэхэд ашигладаг.

2) TextRank

TextRank алгоритм нь граф суурьтай үнэлгээний алгоритм бөгөөд эх хэлний боловсруулалтад текстийн хураангуй үүсгэх, түлхүүр үг тодорхойлох, өгүүлбэрийн ач холбогдлыг үнэлэх зэрэгт ашигладаг. Энэ алгоритм нь Google-ийн PageRank алгоритм дээр суурилж бүтээгдсэн бөгөөд текст дэх өгүүлбэрүүдийг граф байдлаар дүрслэн, тэдгээрийн хоорондын холбоог үнэлж, хамгийн чухал өгүүлбэрүүдийг тодорхойлох зарчмаар явагддаг.

TextRank алгоритмын үндсэн үе шатууд:

1. Граф үүсгэх - Текстийн өгүүлбэрүүдийг зангилаа (node) болгон тодорхойлж, тэдгээрийн хоорондын холбоог ирмэг (edge) хэлбэрээр илэрхийлнэ.
2. Өгүүлбэр хоорондын ижил төстэй байдлыг тооцоолох - Өгүүлбэрүүдийн утгын ойролцоо

байдлыг хэмжиж, ижил төстэй өгүүлбэрүүдийг харилцан холбох замаар граф үүсгэнэ.

3. PageRank алгоритмыг хэрэгжүүлэх - Граф дээр PageRank алгоритмыг ашиглаж, өгүүлбэрүүдийн ач холбогдлын оноог тооцоолно.
4. Өгүүлбэрүүдийг сонгох - Текстийн хураангуйд оруулах өгүүлбэрүүдийн хамгийн өндөр оноотой өгүүлбэрүүдээс сонгож, гарган авна.

Ингэснээр текст өгөгдлөөс өндөр оноо бүхий гол өгүүлбэрүүдийг ялгаж хураангуйг үүсгэдэг.

3) LexRank

LexRank нь граф суурьтай эх хэлний боловсруулалтад ашигладаг алгоритм юм. Энэ нь өгүүлбэрүүдийн харилцан хамаарлыг графын тусламжтайгаар үнэлж, хамгийн чухал өгүүлбэрүүдийг тодорхойлох замаар текстийн хураангуй үүсгэдэг. LexRank нь PageRank алгоритмтай төстэй бөгөөд өгүүлбэрүүдийг зангилаа (node) болгон авч, тэдгээрийн хоорондох холбоог ирмэг (edge) хэлбэрээр дүрсэлдэг.

Энэхүү алгоритмын үндсэн зарчим нь өгүүлбэрүүдийн ижил төстэй байдлыг тооцоолж, тэдгээрийн хоорондын холбогдлыг тодорхойлох явдал юм. Өгүүлбэр хоорондын ижил төстэй байдал нь косинусын ижил төстэй байдлын хэмжүүр (cosine similarity) ашиглан тооцоологддог.

LexRank алгоритм нь дараах үндсэн үе шатуудтай:

1. Өгүүлбэр хоорондын ижил төстэй байдлыг тооцоолох

$$sim(S_i, S_j) = \frac{V_i \cdot V_j}{\left||V_i| \times |V_j|\right|} \quad (4)$$

S - өгүүлбэрүүд, V - өгүүлбэрүүдийн вектор илэрхийлэл

2. Граф бүтээх
   - Өгүүлбэрүүдийг зангилаа болгон авч, ижил төстэй байдлын босго утгаас өндөр тохиолдолд хооронд нь холбож, жинг (weight) тооцоолно.
   - Холбоосны жинг дараах томьёогоор тодорхойлдог.

$$\omega_{ij} = \left\{\frac{sim(S_i, S_j), \text{ хэрэв } sim(S_i, S_j) > r}{0, \text{ бусад тохиолдолд}}\right\} \quad (5)$$

Энд r нь босго утга.

3. PageRank-ийн зарчмаар оноог тооцоолох

$$PR(S_i) = (1 + d) + d \sum_{S_k}^{adj(S_i)} \frac{\omega_{ij \times PR(S_j)}}{\sum_{S_k}^{adj(S_j)} \omega_{jk}} \quad (6)$$

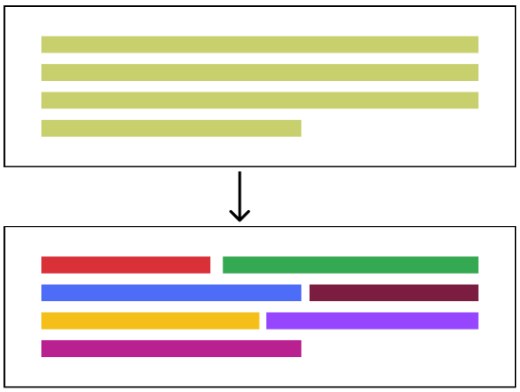
4. Өгүүлбэрүүдийг сонгох
   - PageRank онооны дагуу өгүүлбэрүүдийг эрэмбэлж, хамгийн өндөр оноотой өгүүлбэрүүдийг хураангуйд оруулна

D. Нэмэлт сайжруулалт

*1) **Phrase** урьдчилсан боловсруулалтын (preprocessing) сайжруулалтын давхарга*

Бидний нэмэлтээр хэрэгжүүлсэн phrase сайжруулалтын давхарга нь 3.2-т заасан урьдчилсан боловсруулалтын алхмууд хэрэгжсэн текст өгөгдөлд боловсруулалт хийх буюу таслалаар холбогдсон нийлмэл өгүүлбэрүүдийг тус бүрд нь дахин өгүүлбэрүүдэд хуваах үүрэгтэй.

Ингэснээр нийлмэл өгүүлбэрт буй текстийн агуулгад нөлөө ихтэй гишүүн өгүүлбэрийг илүү нарийвчлалтайгаар ялган авах давуу талтай.



Зураг. 5. Phrase давхаргын хэрэгжилтийн жишээ

*2) **Sentence position** дараах боловсруулалтын (postprocessing) сайжруулалтын давхарга*

Өгүүлбэрийн ач холбогдлыг тооцоолох алгоритмын гаралтад нэмэлтээр оруулж буй сайжруулалтын давхарга болох sentence position нь [6] – д дурдсаны дагуу ихэнх текстэн өгөгдлийн хувьд жишээлбэл: эрдэм шинжилгээний өгүүлэл, мэдээ, нийтлэлүүдийн цөм утга санааг агуулах өгүүлбэрүүд нь нийт текстийн эхлэл болон төгсөл хэсэгт байршдаг гэсэн үндэслэл дээр үндэслэн өгүүлбэрийн байршлыг (sentence position) онцолж, дараах боловсруулалтын (postprocessing) сайжруулалтын давхаргыг хэрэгжүүлсэн.

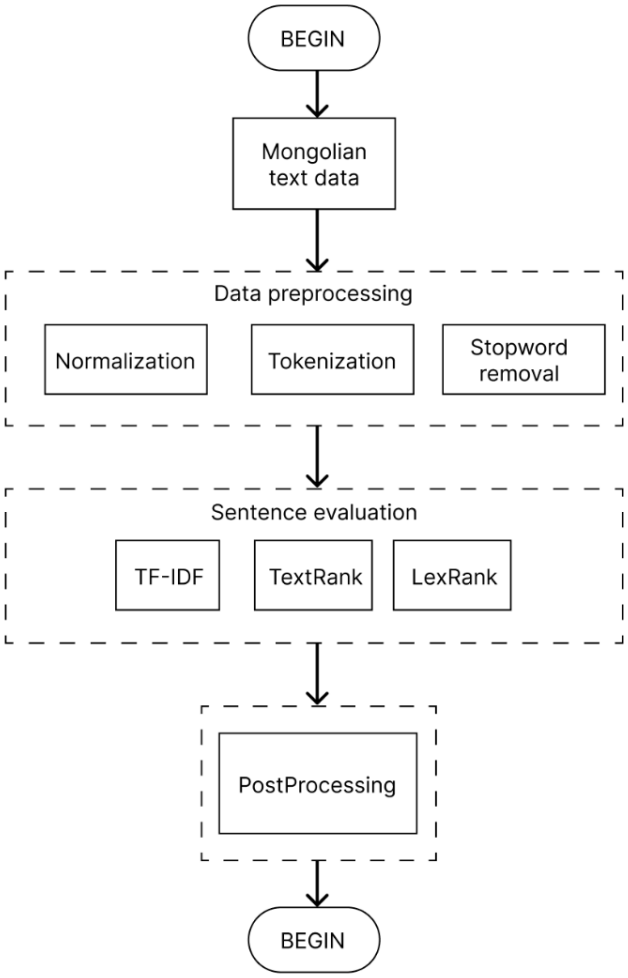
Sentence position нь хэрэгжихдээ оролтын текстийн эхний болон сүүлийн 20% - д, нийт өгүүлбэрийг ямар нэгэн үнэлгээний алгоритм үнэлсний дараа нэмэлт оноо өгөх зарчмаар хэрэгжинэ.

Жишээлбэл: загвар луу нийт 10 өгүүлбэртэй текст өгөгдөл оруулахад эхний болон, сүүлийн 2 өгүүлбэр дээр нэмэлт оноо өгөх юм.

Зураг. 6. Sentence position давхаргын хэрэгжилтийн жишээ

E. Хэрэгжүүлсэн загваруудын бүрэлдэхүүн хэсгүүд

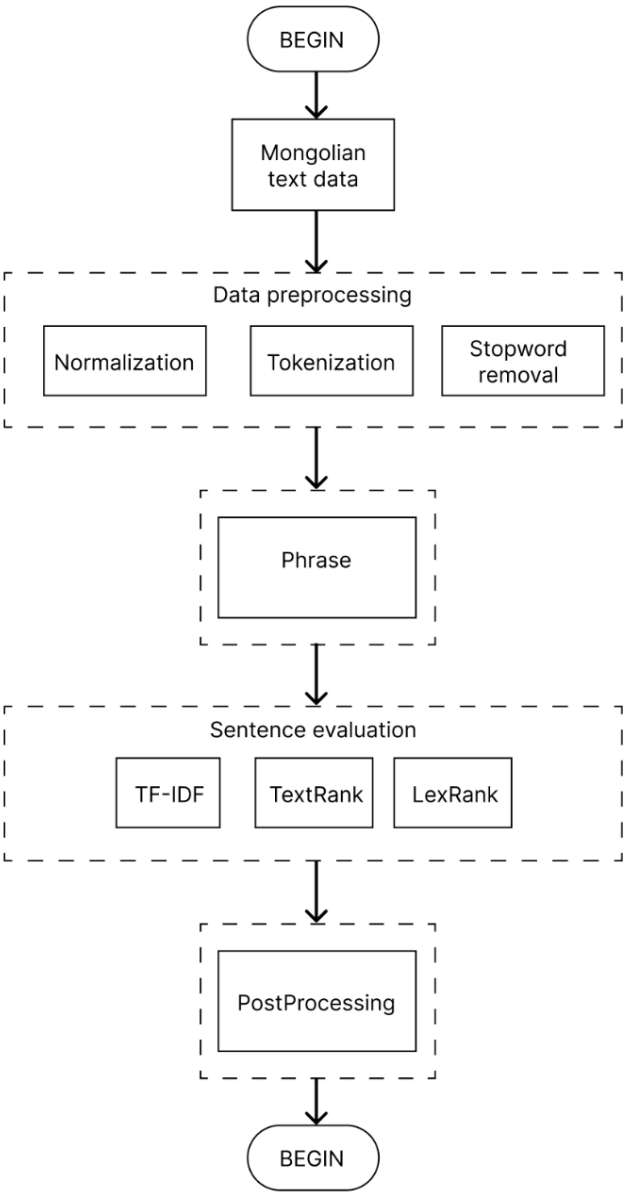
Энэхүү судалгааны ажлын хүрээнд бид TF-IDF, TextRank, LexRank алгоритмуудыг өөрөөр нь болон өөрсдийн нэмэлт сайжруулалтын давхарга нэмэн хэрэгжүүллээ.



Зураг. 7. Нэмэлт давхаргагүй хэрэгжилтийн бүрэлдэхүүн хэсгүүд

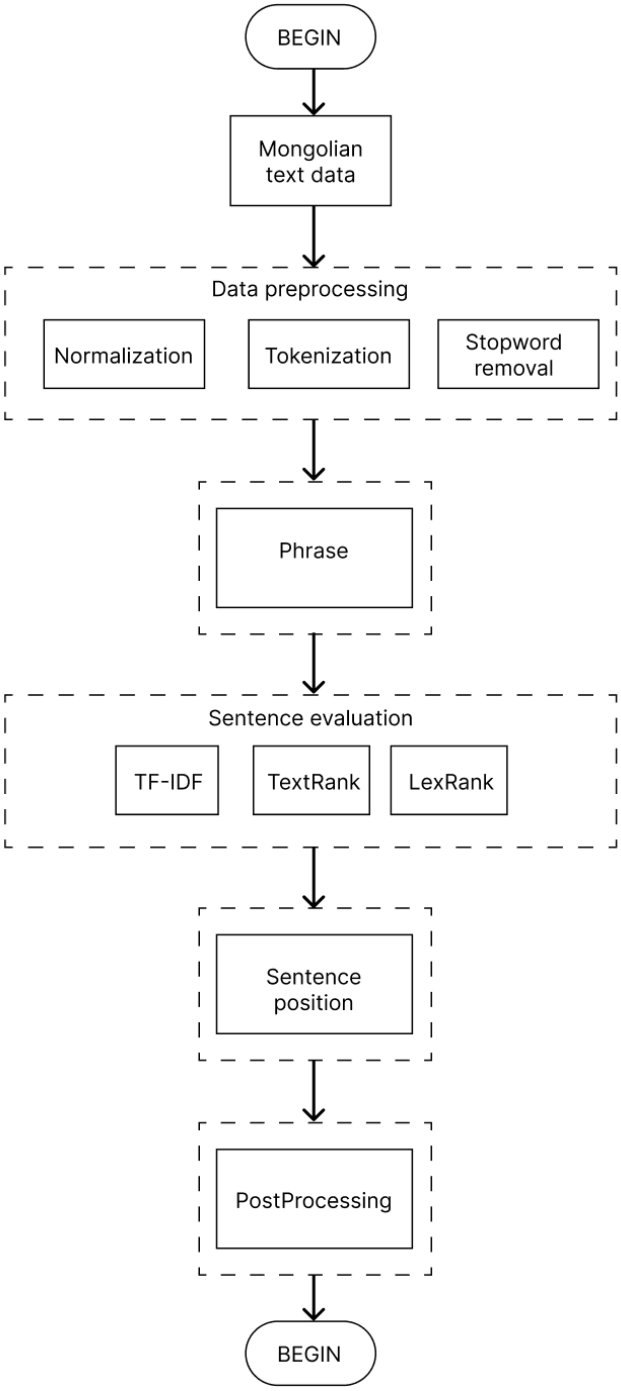
Дараах зурагт хэрэгжүүлж буй загварууд нь ямар нэгэн нэмэлт сайжруулалтын давхаргагүй хэрэгжүүлэх үеийн загварын бүрэлдэхүүн хэсгүүдийг харуулж байх бөгөөд энд TF-IDF, LexRank, TextRank алгоритм тус бүр дээр тус тус нэг хэрэгжилт байна.

Эдгээр нэмэлт сайжруулалтын давхаргагүй загваруудыг үндсэн загвар гэж үзэх бөгөөд, үндсэн загвартай өөрсдийн нэмж оруулж буй сайжруулалтын давхаргатай загваруудыг гүйцэтгэл болон үзүүлэлтээр нь харьцуулах зорилготой ашиглах юм.



Зураг. 8. Phrase нэмэлт сайжруулалтын давхаргатай хэрэгжилтийн бүрэлдэхүүн хэсгүүд

Дээрх зурагт өмнөх хэрэгжүүлсэн загвар дээр phrase нэмэлт сайжруулалтын давхарга оруулан хэрэгжүүлэх үеийн загварын бүрэлдэхүүн хэсгүүдийг харуулж байна. Мөн алгоритм тус бүр дээр тус тус нэг загвар хэрэгжинэ.



Зураг. 9. Phrase болон Sentence position сайжруулалтын давхаргатай хэрэгжилтийн бүрэлдэхүүн хэсгүүд

Энэхүү зурагт үндсэн загвар дээр phrase болон sentence position сайжруулалтын давхаргуудыг нэмж оруулсан үеийн загварын бүрэлдэхүүн хэсгүүдийг харуулж байна. Мөн адил алгоритм тус бүр дээр тус тус нэг загвар хэрэгжинэ.

F. Үнэлгээний аргачлал

ROUGE score (Recall-Oriented Understudy for Gisting Evaluation) нь текст хураангуй болон машин орчуулга зэрэг автомат үнэлгээний системүүдийг шалгахад өргөн хэрэглэгддэг үнэлгээний хэмжүүр юм. Энэ хэмжүүр нь лавлах хураангуй болон алгоритмын хураангуй хоорондох ижил төстэй байдлыг хэмжиж үнэлдэг.

ROUGE score тооцоолохдоо ихэвчлэн recall (буцаалт), precision (нарийвчлал) болон F1-score гэсэн үзүүлэлтүүдийг ашигладаг.

- Recall нь лавлах хураангуйн хэдэн хувь нь алгоритмын үүсгэсэн хураангуйд багтсан эсэхийг илэрхийлнэ.
- Precision нь лавлах хураангуйн доторх мэдээллээс алгоритмын хураангуйтай хэр зэрэг таарч буйг хэмжинэ.
- F1-score нь recall болон precision-ийн дундаж утгыг харуулдаг.

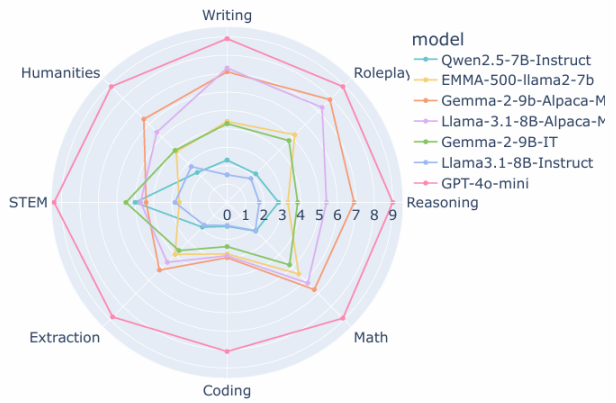
Rouge score нь дараах төрлүүдтэй байна:

1. Rouge-1 нь нэг үгийн (unigram) давхцалыг хэмждэг үзүүлэлт юм.
2. Rouge-2 нь хоёр үгийн дараалал (bigram) 2 үгийн давхцалыг хэмждэг үзүүлэлт юм.
3. Rouge-L нь LCS (Longest common subsequence) буюу хураангуйн хамгийн урт нийтлэг дараллыг тодорхойлох замаар үнэлгээ хийдэг.

Rouge score үнэлгээний арга нь хураангуйллыг үнэлэхдээ тухайн текстийн лавлагаа (reference) хураангуйг хэрэгжүүлсэн загварын гаргасан хураангуйтай харьцуулж үнэлгээ хийнэ.

Өгөгдлийн багц бүр дээр лавлагаа хураангуйг бэлтгэхдээ хэлний мэргэжилтнээр тухайн текстийн хураангуйг тодорхойлбол илүү сайн хураангуй гарах боловч цаг хугацаа болон, санхүүгийн нөөц шаардана.

Харин бид өөрсдөө лавлах хураангуйг гарган турших боломжтой ч энэ тохиолдолд бидний гаргасан хураангуйлал нь мэргэжлийн түвшинд хүрэхгүй, тухайн хураангуйлал нь тодорхой хэмжээний хийсвэр агуулгатай байх магадлалтай. Иймээс бид хураангуй үүсгэхдээ хиймэл оюун ухааныг ашиглах шийдвэрт хүрсэн.



Зураг. 10. LLM – үүдийн Монгол хэл дээрх үзүүлэлтийн харьцуулалт

[7] судалгаанд хийгдсэн, Зураг 10 – д харуулсан туршилтын үеэр Qwen2.5, Llama, GPT 4o-mini гэх мэт LLM загваруудын Монгол хэл дээрх гүйцэтгэлийг тал бүрийн төрлөөр харьцуулсан ба GPT 4o-mini загвар нь хамгийн өндөр үр дүнг үзүүлсэн бөгөөд энэхүү үзүүлэлт дээр тулгуурлан бид лавлах хураангуйг үүсгэхдээ GPT 4o-mini LLM загварыг ашигласан.

GPT 4o-mini загварын тодорхойлсон хураангуйг өөрсдийн тодорхойлсон хураангуйтай харьцуулахад GPT 4o-mini загвар лавлагаа хураангуйг илүү сайн экстрактив байдлаар тодорхойлсон. Ингэснээрээ бидэнд тодорхой хэмжээний цаг хугацаа хэмнэхэд тусалсан.

IV. Үр дүн

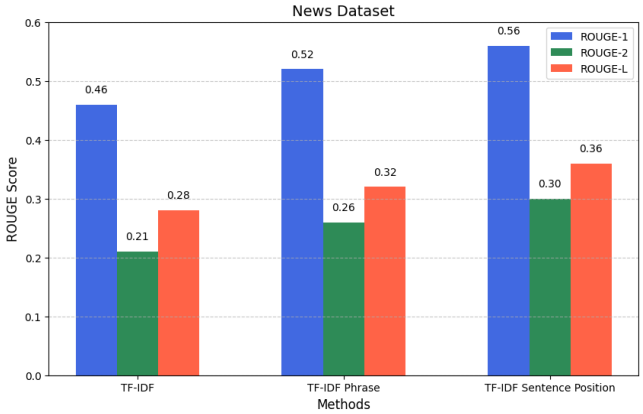
Бид экстрактив текст хураангуйллын хүрээнд TF-IDF, TextRank, LexRank алгоритмуудыг өөрсдийн тодорхойлсон нэмэлт сайжруулалтын давхаргаар баяжуулж хэрэгжүүллээ.

Хүснэгт. II Загваруудын бүтэц

TF-IDF			
		Давхарга 1	Давхарга 2
1	TF-IDF		
2	TF-IDF	Phrase	
3	TF-IDF	Phrase	Sentences Position
TextRank			
		Давхарга 1	Давхарга 2
1	TextRank		
2	TextRank	Phrase	
3	TextRank	Phrase	Sentence Position
LexRank			
		Давхарга 1	Давхарга 2
1	LexRank		
2	LexRank	Phrase	
3	LexRank	Phrase	Sentence Position

b. Нийт хэрэгжүүлсэн загваруудд

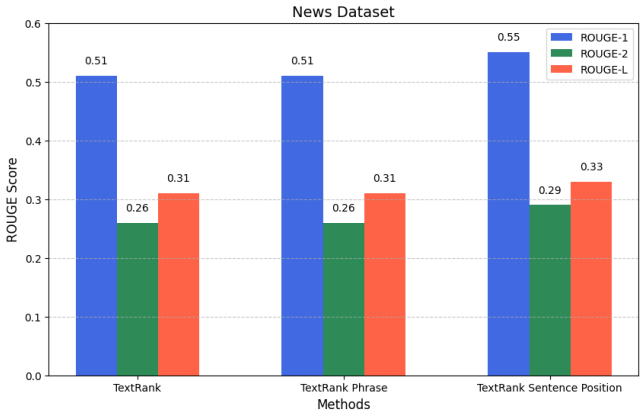
A. *Мэдээний өгөгдлийн багц дээрх загваруудын үр дүнгүүд*



Зураг. 11. Мэдээний өгөгдлийн багц дээрх TF-IDF загваруудын үр дүнгүүд

Зураг 11 – д үзүүлснээр мэдээний өгөгдлийн багц дээр TF-IDF sentence position загвар нь хамгийн үр дүнтэй буюу rouge score 0.56 байна.

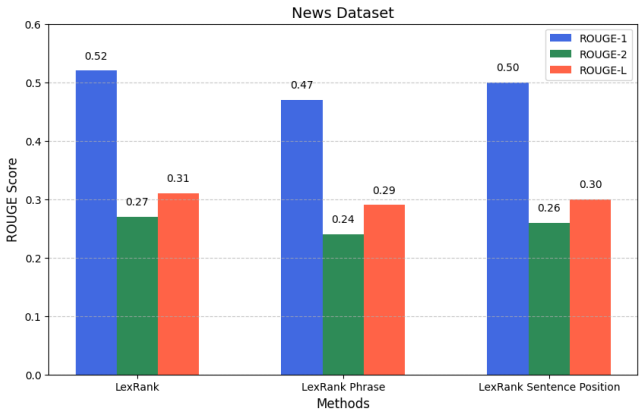
Мөн TF-IDF алгоритм дээр сайжруулалтын давхарга нэмэх бүрд rouge score нь өссөн харагдаж байна. Энэ нь өгүүлбэрийн байршил болон нийлмэл өгүүлбэрт буй гишүүн өгүүлбэрүүдийг тус бүрд өгүүлбэр болгон авсан нь хураангуй илүү нарийвчлалтай болж байгааг илэрхийлж байна.



Зураг. 12. Мэдээний өгөгдлийн багц дээрх TextRank загваруудын үр дүнгүүд

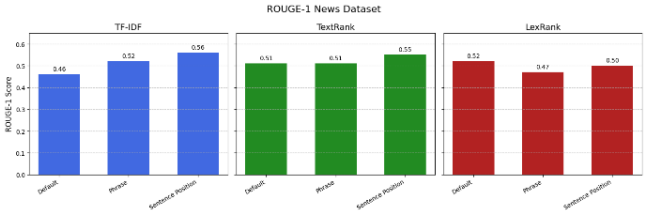
Зураг. 12 – д үзүүлснээр мэдээний өгөгдлийн багц дээр TexRank алгоритм давхарга нэмэх бүрд үзүүлэлт буураагүй байна.

Мөн TextRank sentence position арга хамгийн өндөр буюу rouge score нь 0.55 байна.



Зураг. 13. Мэдээний өгөгдлийн багц дээрх LexRank загваруудын үр дүнгүүд

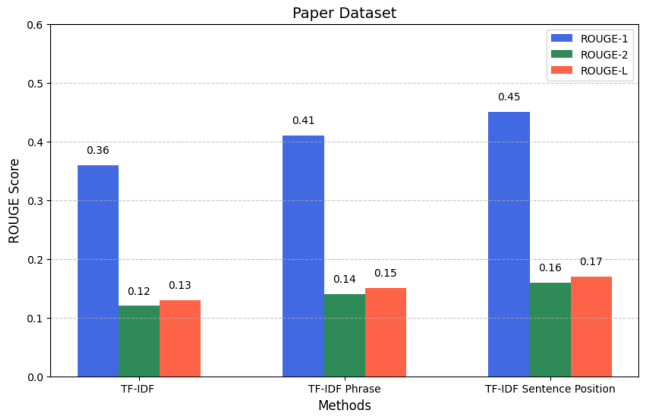
Зураг. 13 – д харуулснаар LexRank алгоритм дээр phrase сайжруулах давхарга нэмсэн тохиолдолд үзүүлэлт буурсан боловч sentence position давхарга нэмэлтээр орсон үед өссөн байна. Гэвч 2 нэмэлт давхарга нь үндсэн (default) LexRank – тай харьцангуй бага үзүүлэлт үзүүлсэн байна.



Зураг. 14. Мэдээний өгөгдлийн багц дээрх нийт загваруудын үр дүнгүүд

Зураг. 14 – д харуулснаар мэдээний өгөгдлийн багц дээр нийт хэрэгжүүлсэн загваруудыг харьцуулж гаргасан бөгөөд загвар тус бүр дээр sentence position сайжруулалтын давхарга нэмсэн үед тодорхой хэмжээнд үзүүлэлт өссөн үзэгдэл ажиглагдаж байна.

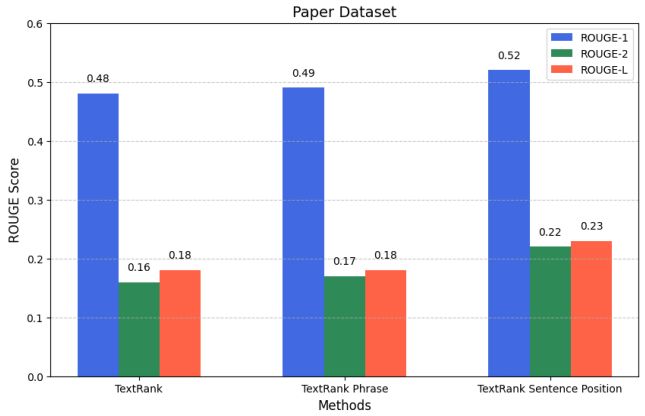
B. Эрдэм шинжилгээний өгөгдлийн багц дээрх загваруудын үр дүнгүүд



Зураг. 15. Эрдэм шинжилгээний өгөгдлийн багц дээрх TF-IDF загваруудын үр дүн

Зураг. 15 – д харуулснаар эрдэм шинжилгээний өгөгдлийн багц дээр TF-IDF sentence position нь хамгийн өндөр үзүүлэлтийг үзүүлсэн байна.

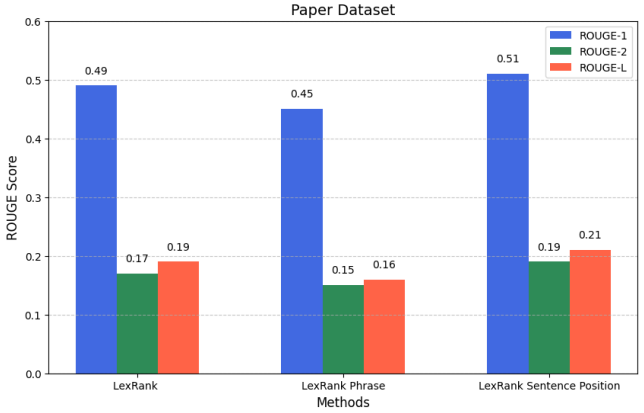
Өмнөх мэдээний өгөгдлийн багцын үр дүнгийн шинж чанартай адилаар үндсэн алгоритм дээр сайжруулалтын давхарга нэмэх бүрд үзүүлэлт нь ахисан байна.



Зураг. 16. Эрдэм шинжилгээний өгөгдлийн багц дээрх TextRank загваруудын үр дүнгүүд

Зураг. 16 – д үзүүлснээр TextRank sentence position эрдэм шинжилгээний өгөгдлийн багц дээр хамгийн өндөр үзүүлэлт буюу rouge score нь 0.52 байна.

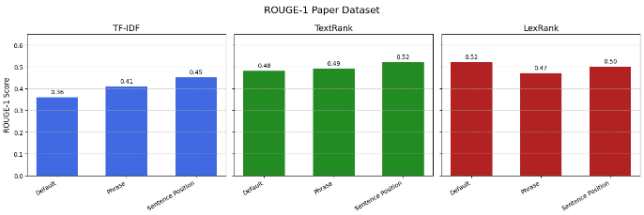
Мэдээний өгөгдлийн багц дээрх үр дүнтэй адил алгоритм дээр сайжруулалтын давхарга нэмэх бүрд өссөн байна.



Зураг. 17. Эрдэм шинжилгээний өгөгдлийн багц дээрх LexRank загваруудын үр дүнгүүд

Зураг. 17 – д харуулснаар эрдэм шинжилгээний өгүүллийн өгөгдлийн багц дээр LexRank sentence position нь хамгийн өндөр үзүүлэлт буюу rouge score нь 0.51 байна.

Үзүүлэлтийн шинж чанарыг хувьд phrase сайжруулалтын давхарга нэмсэн үед үндсэн алгоритмын үр дүнгээс үзүүлэлт нь буурсан боловч sentence position сайжруулалтын давхарга орсон тохиолдолд өссөн байна.

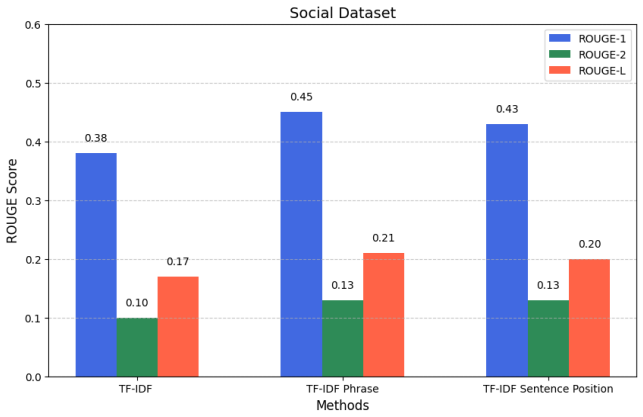


Зураг. 18. Эрдэм шинжилгээний өгөгдлийн багц дээрх нийт загваруудын үр дүнгүүд

Зураг. 18 – д үзүүлснээр эрдэм шинжилгээний өгөгдлийн багц дээр нийт хэрэгжүүлсэн загваруудын үзүүлэлтүүдийг харьцуулж гаргасан.

Мэдээний өгөгдлийн багц дээр үр дүнтэй ижил TF-IDF, TextRank суурьтай загваруудад сайжруулалтын давхарга нэмэх тохиолдол бүрд үзүүлэлт өссөн бол LexRank суурьтай загваруудад phrase сайжруулалтын давхарга нэмсэн тохиолдолд үндсэн загварын үзүүлэлтээс буурч, sentnce position сайжруулалтын давхарга нэмэхэд эргээд үзүүлэлт өссөн ижил шинж чанар ажиглагдаж байна.

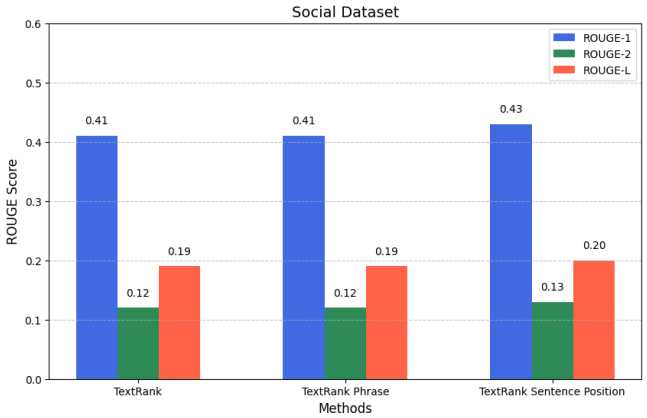
C. Цахим орчны өгөгдлийн багц дээрх загваруудын үр дүнгүүд



Зураг. 19. Цахим орчны өгөгдлийн багц дээрх TF-IDF загваруудын үр дүнгүүд

Зураг. 19 – д үзүүлснээр цахим орчны текст өгөгдлийн багц дээр TF-IDF phrase загвар хамгийн өндөр үзүүлэлтийг үзүүлсэн буюу rouge score нь 0.45 байна.

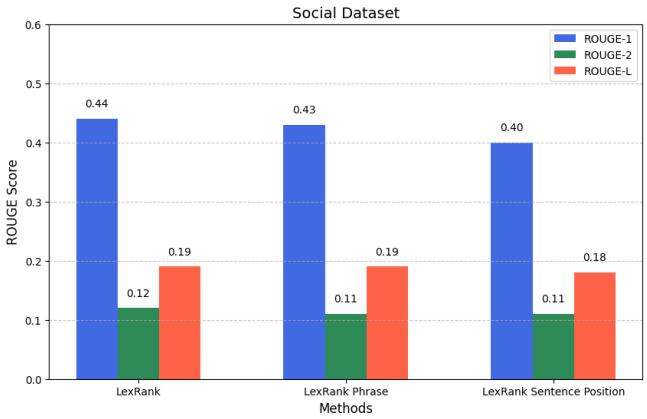
Загваруудын үзүүлэлтийн шинж чанарын хувьд phrase сайжруулалтын давхарга дээр өсөж sentence position сайжруулалтын давхарга нэмсэн үед буурч байна. Мөн өмнөх өгөгдлийн багцуудаас өөр шинж чанартай байна.



Зураг. 20. Цахим орчны өгөгдлийн багц дээрх TextRank загваруудын үр дүнгүүд

Зураг. 20 – д харуулснаар цахим орчны өгөгдлийн багц дээр TextRank sentence position нь хамгийн өндөр үзүүлэлтийг үзүүлсэн байна.

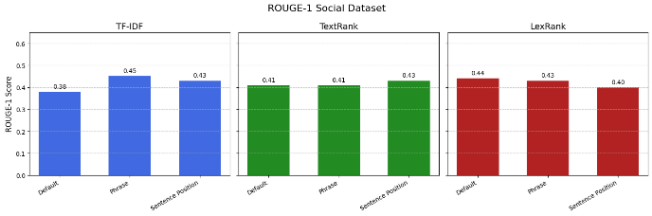
Харин загваруудын үзүүлэлтийн шинж чанаруудын хувьд sentence position сайжруулалтын давхарга нэмсэн тохиолдолд өсөж бусад тохиолдолд хэвийн байгаа бөгөөд мэдээний өгөгдлийн багц дээрх гаргасан үзүүлэлттэй ижил шинж чанартай байна.



Зураг. 21. Цахим орчны өгөгдлийн багц дээрх LexRank загваруудын үр дүнгүүд

Зураг. 21 – д үзүүлснээр цахим орчны өгөгдлийн багц дээр LexRank (default) нь хамгийн өндөр үзүүлэлт үзүүлсэн байна.

Энэ удаад өмнөх үзэгдлүүдээс эсрэг үзэгдэл гарсан бөгөөд сайжруулалтын давхарга нэмсэн тохиолдол бүрд үзүүлэлт буурсан эсрэг үзэгдэл гарсан байна.



Зураг. 22. Цахим орчны өгөгдлийн багц дээрх нийт загваруудын үр дүнгүүд

Зураг. 22 – д харуулснаар цахим орчны өгөгдлийн багц дээрх нийт хэрэгжүүлсэн загваруудын үзүүлэлтийг харьцуулав.

Үзүүлэлтийн шинж чанарын хувьд өмнөх 2 өгөгдлийн багц дээр гарсан үзэгдлээс өөр үзэгдэл гарсан байна. Жишээлбэл: LexRank алгоритм дээр сайжруулалтын давхарга нэмэх бүрд үзүүлэлт буурсан байна. Харин бусад алгоритмын хувьд sentence position сайжруулалтын давхарга нэмэгдсэн үед өсөлт үндсэн (default) загварын үзүүлэлтээс өссөн үзэгдэл нь хэвээрээ байна.

Цахим орчны өгөгдлийн багц дээр үзүүлэлт буурч байгаа шалтгаан нь цахим орчин дахь нийтлэл нь эрдэм шинжилгээний өгүүлэл болон мэдээтэй адил цэгцтэй, албан ёсны найруулгатай биш учир цахим орчны өгөгдлийн багц дээр үзүүлэлт буурсан гэж үзэж байна.

V. Дүгнэлт

Монгол хэлний цахим хэрэглээ нэмэгдэж, мэдээллийн урсгал хурдацтай өсөхийн хэрээр текст хураангуйллын аргуудыг хөгжүүлэх хэрэгцээ улам бүр нэмэгдэж байна. Их хэмжээний мэдээллийг богино хугацаанд боловсруулах шаардлага академик судалгаа, хэвлэл

мэдээлэл, цахим орчин, бизнесийн салбар зэрэг олон чиглэлд тулгарч буй гол асуудлуудын нэг болоод байна. Бусад орны олон хэрэглэгчтэй хэл дээр үр ашигтай ажилладаг олон төрлийн текст хураангуйллын алгоритмууд боловсруулагдсан ч Монгол хэлний хувьд үг зүй, өгүүлбэр зүй, бүтэц зэргээс шалтгаалан эдгээр аргуудыг шууд хэрэглэх боломж хязгаарлагдмал байдаг. Иймд Монгол хэлэнд тохирсон, оновчтой хураангуйллын аргуудыг боловсруулах нь чухал ач холбогдолтой юм.

Энэхүү судалгааны ажлын хүрээнд Монгол хэл дээрх экстрактив текст хураангуйллын хэрэгжүүлэлтийг гүйцэтгэлээ. Судалгааны явцад граф өгөгдлийн бүтэц дээр суурилсан TextRank, LexRank алгоритмууд болон статистик аргачлалд тулгуурласан Term Frequency (TF), Inverse Document Frequency (IDF) аргуудыг ашиглан өөрсдийн тодорхойлсон phrase, sentence position сайжруулалтын давхаргуудыг оруулж 9 өөр загварыг хэрэгжүүлсэн болно. Хэрэгжүүлсэн загвараа үнэлж харьцуулахдаа 10 жишээтэй 3 өөр өгөгдлийн багцыг үүсгэж нийт 270 туршилт хийв. Судалгааны үр дүнг **ROUGE Score** ашиглан үнэлж, нийт хэрэгжүүлсэн 9 загварын дундаж үзүүлэлтийг гаргалаа. Үр дүнгээс харахад:

- Мэдээний хураангуйлалд TF-IDF sentence poistion загвар нь хамгийн өндөр үр дүнг үзүүлэв.
- Цахим орчны текстийн хувьд TF-IDF phrase загвар нь хамгийн өндөр үр дүнг үзүүлэв.
- Эрдэм шинжилгээний өгүүллийн хураангуйлалд TextRank sentence position загвар нь хамгийн өндөр үр дүнг үзүүлэв.

Энэхүү судалгаа нь Монгол хэл дээрх текст хураангуйллын аргуудын үр дүнг харьцуулан судалж, текстийн төрөл бүрд хамгийн тохиромжтой аргачлалыг тодорхойлоход чиглэсэн болно. Цаашлаад нарийвчилсан судалгаа, боловсруулалт хийхэд чухал суурь болох боломжтойг харуулж байна.

Ном зүй

[1] Shimpikar, Sheetal, and Sharvari Govilkar. "A survey of text summarization techniques for Indian regional languages." *International Journal of Computer Applications* 165, no. 11 (2017): 29-33.

[2] Gupta, Vishal, and Gurpreet Singh Lehal. "Automatic Punjabi text extractive summarization system." In *Proceedings of COLING 2012: Demonstration Papers*, pp. 191-198. 2012.

[3] Han, Yongshun, Qintu Si, and Siriguleng Wang. "Mongolian Automatic Text Summarization Method Based on Pre-trained Model and Improved TextRank." *Advances in Computer and Communication* 5, no. 2 (2024).

[4] Wijayanti, Rini, Masayu Leylia Khodra, Kridanto Surendro, and Dwi H. Widyantoro. "Learning bilingual word embedding for automatic text summarization in low resource language." *Journal of King Saud University-Computer and Information Sciences* 35, no. 4 (2023): 224-235.

[5] Munaf, Mubashir, Hammad Afzal, Khawir Mahmood, and Naima Iltaf. "Low resource summarization using pre-trained language models." *ACM Transactions on Asian and Low-Resource Language Information Processing* 23, no. 10 (2024): 1-19.

[6] https://web.archive.org/web/20160304045652/http://faculty2.ric.edu/rfeldstein/202_spring_08_files/4.par_ex_4_topic_sentence.pdf

[7] D. Tumur-Ochir, F. Cheng, YU. Murawaki, and C. Chu, Evaluating Large Language Models in Mongolian, in *Proc. 31st Annu. Meeting of the Association for Natural Language Processing (NLP)*, pp. 368–373, Mar. 2025.

Багшийн дижитал мэдлэг, чадварын оюутны үнэлгээний кластер шинжилгээ

Д.Мөнхбаатар (докторант)
МУБИС, БСС,
АШУҮИС, ЕБА сургууль
Цахим хаяг:
munkhbaatar.d@mnums.edu.mn

Проф.Д.Цэдэвсүрэн(Ph.D)
МУБИС, Сургалт, цахим шилжилтын хэлтэс
Цахим хаяг: tsedevsuren@msue.edu.mn

Л.Пүрэвдолгор(Ph.D)
АШУҮИС, Био-Анагаахын сургууль,
Физик мэдээлэл зүйн тэнхим
Цахим хаяг: purevdolgor@mnums.edu.mn

Ч.Баасандорж
АШУҮИС, Био-Анагаахын сургууль,
Физик мэдээлэл зүйн тэнхим
Цахим хаяг: baasandorj@mnums.edu.mn

Ч.Галбадрах
АШУҮИС, Био-Анагаахын сургууль,
Физик мэдээлэл зүйн тэнхим
Цахим хаяг: galbadrakh@mnums.edu.mn

Д.Булгансүрэн
АШУҮИС, Био-Анагаахын сургууль,
Физик мэдээлэл зүйн тэнхим
Цахим хаяг: bulgansuren@mnums.edu.mn

Хураангуй— Үндэслэл: Орчин үед их, дээд сургуулиуд цахим шилжилтийг эрчимтэй хэрэгжүүлж байна. Энэ шилжилтийн гол оролцогч нь багш, оюутнууд бөгөөд ялангуяа багш нарын дижитал мэдлэг, чадвар ихээхэн чухал үүрэгтэй. Иймээс уг судалгаа нь оюутнуудаар дамжуулан багш нарын дижитал ур чадварыг үнэлж, тодорхойлох зорилготой юм. Судалгаанд Улаанбаатар хотын төрийн болон хувийн өмчийн 7 их, дээд сургуулийн 513 оюутан оролцлоо. Арга, аргачлал: Судалгааны өгөгдлийг кластер шинжилгээний k-means аргаар боловсруулж, R програм ашиглан дүн шинжилгээ хийсэн. Үр дүн: Асуулгын найдвартай байдлын Cronbach's alpha коэффициент 0.95, кластер хоорондын Hopkins коэффициент 0.99 гарсан нь хүчтэй бөөгнөрөл байгааг харуулж байна. Silhouette method аргаар хоёр бүлэг (кластер) болгон хуваасан бөгөөд судалгааны үр дүнд оюутнууд багш нарын дижитал мэдлэг, чадварыг дараах байдлаар үнэлсэн байна. Үүнд I кластер: "Сайн" гэж үнэлсэн - 59.45%, II кластер "Дунд" гэж үнэлсэн-40.55% гарлаа. Төрийн өмчийн их сургуулийн оюутнуудын үнэлгээний хандлага (H = 0.99) маш хүчтэй тодорхой загвартай гарсан. Үүнээс үзэхэд төрийн сургуулийн багш нарын дижитал ур чадварын ялгаа бага, эсвэл оюутнуудын багшаас хүсэх дижитал мэдлэг чадварын хүлээлт харьцангуй ижил байх магадлалтай. Харин хувийн сургуулийн оюутнуудын үнэлгээ (H = 0.81) төрийн сургуулийнхтай харьцуулахад арай сул загвартай, харилцан адилгүй байгаа нь ажиглагдсан. Энэ нь хувийн сургуулийн багш нарын дижитал ур чадварын түвшин харилцан адилгүй, зарим нь өндөр чадвартай, зарим нь харьцангуй сул байгаатай холбоотой байж болох юм.

Түлхүүр үг— Кластер шинжилгээ, k-means утга, Hopkins арга, Anova

I. ҮНДЭСЛЭЛ

Орчин үеийн дээд боловсролын байгууллагууд мэдээллийн технологийн эрчимтэй хөгжлийг даган дижитал шилжилтийг хийж байна. Их, дээд сургуулийн багш нарын дижитал мэдлэг, чадвар нь сургалтын чанарт шууд нөлөөлж, оюутнуудын мэдлэг эзэмших үйл явц, хичээлийн үр дүнг сайжруулах гол хүчин зүйлсийн нэг болж байна. Дижитал эрин үед их сургуулийн багш нар зөвхөн мэргэжлийн мэдлэгтэй байх нь хангалтгүй бөгөөд тэдний мэдээллийн технологийг ашиглах чадвар нь сургалтын үр өгөөжийг нэмэгдүүлэхэд чухал үүрэг гүйцэтгэнэ [1]. Орчин үеийн залуус дижитал орчинд өсөж торнисон тул тэдний технологийн мэдлэг нь багш нараасаа илүү байх нь элбэг. Энэ нь багш, оюутан хоёрын хооронд технологийн хэрэглээний мэдлэгийн зөрүү үүсгэдэг [2]. Тиймээс боловсролын арга зүйд дижитал мэдлэгийг багтаасан, суралцахуйд тулгуурласан "хийж сурах" үйлийг ашиглах шаардлагатай үүсч байна [3]. Үүнийг үр дүнтэй хэрэгжүүлэхэд хамгийн чухал хүчин зүйл бол багшийн дижитал чадвар юм [4]. 2022 онд "Developing a digital competence scale for teachers- validity and reliability study" судалгаагаар багш нарт зориулсан дижитал чадамжийн хэмжүүрийг боловсруулж хүчин төгөлдөр найдвартай байдлын судалгааг хийжээ. Ингэхдээ багш нарт МХХТ-ийн "Аюулгүй байдал", "Дижитал боловсрол", "Асуудал шийдвэрлэх", "Дижитал контент бүтээх", "Харилцаа ба хамтын ажиллагаа", "Ёс зүй" гэсэн 6 чадварын хүрээнд судалсан байна [5]. Европын холбооноос гаргасан ДижКом 21 зөвлөмжид их сургуулийн багш нарын дижитал чадамжид сургах нь чанартай боловсрол эзэмшүүлэх хамгийн чухал зүйл бөгөөд дижитал технологийг заах арга зүйтэйгээ хослуулан хэрэглэх нь чухал ач холбогдолтой байна гэдгийг харуулсан [6]. Оюутнуудын үнэлгээнд үндэслэн багш нарын дижитал мэдлэг, чадварыг тодорхойлох нь их сургуулийн сургалтын чанарыг сайжруулах нэг арга болох юм.

Их хэмжээний мэдээллийг боловсруулах, шинжлэх зүй тогтлыг илэрхийлэн загварчлах, дүгнэх, урьдчилан таамаглах шинэ арга технологийг хөгжүүлэх зайлшгүй шаардлагатай. Өгөгдлийн тандалт нь их хэмжээний өгөгдөл дотроос далд зүй тогтол, утгын уялдаа холбоог нээн илрүүлэх, цугларсан их хэмжээний өгөгдлүүдийн холбоо хамааралд дүн шинжилгээ хийх, нийтлэг шинжүүдээр нь бүлэглэх, тодорхой үзүүлэлтүүдээр урьдчилан таамаглах, хэт ялгаатай өгөгдлүүдийг ялгах гэх мэт тооцоолол хийхэд ангилалт (classification) болон бүлэглэлийн (clustering) аргуудыг ашигладаг. Кластер шинжилгээний олон алгоритм байдаг түүнээс к дунджийн утга (k-means clustering) болох хэсэгчлэх алгоритмыг ашиглав.

А. Зорилго

Оюутны үнэлгээнд тулгуурлан их, дээд сургуулийн багш нарын дижитал мэдлэг, чадварын түвшинг тодорхойлох, мөн ижил төстэй үнэлгээтэй багш нарыг кластер шинжилгээний аргаар бүлэглэхэд оршино.

II.АРГА АРГАЧЛАЛ

Судалгааны арван асуумжыг “iSkills assessment” [7],” Technology Acceptance Model (TAM)” [8], Харилцаа холбоо, мэдээллийн технологийн газрын “Иргэдийн мэдээллийн технологийн хэрэглээний өнөөгийн байдлын талаарх судалгаа” [9], ”Их сургуулийн оюутнуудын МХХТ-ийн хэрэглээнд хандах хандлагыг судлах нь” [10] судалгааны дагуу сонгон авлаа. Судалгааны өгөгдөлдөө кластер шинжилгээ (Cluster analysis) хийхийн өмнө кластер хоорондын бөөгнөрөл байгаа эсэхийг хопкинсийн аргаар шалгав.

Хопкенс статистик (Hopkins statistic): Hopkins шалгуур коэффициент нь өгөгдлийн олонлог дэх кластерыг (cluster tendency) үнэлэх статистик арга юм [11]. Энэ нь өгөгдөл нь санамсаргүй тархсан эсвэл кластерлагдсан бүтэцтэй эсэхийг тодорхойлно. Хопкинсийн шалгуурыг дараах томьёогоор бодно:

$$H = \frac{\sum_{i=1}^{n} d_i}{\sum_{i=1}^{n} d_i + \sum_{i=1}^{n} u_i} \quad (1)$$

Утгын тайлбар:

- H ≈ 0.5: Өгөгдөл санамсаргүй тархсан (кластер байхгүй).
- H → 1: Өгөгдөл тодорхой кластерүүдэд төвлөрсөн.
- H → 0: Өгөгдөл жигд тархсан (өөр төрлийн бүтэцтэй) гэж үздэг [12].

B. Кластер шинжилгээ

1971 онд Кормак, 1999 онд Гордон нар “кластер гэдэг нь дотоод уялдаа холбоо (нэгэн төрөл), гадаад хуваагдал гэж нэр томьёоных нь хувьд тодорхойлжээ [13]. Кластерын шинжилгээ нь юмс үзэгдэл дээрх ажиглалт болох олон хэмжээст өгөгдөл эсвэл түүнээс зохиосон зайн матриц дээр тулгуурлан юмс үзэгдлийн ангилал, тусгаар байдлыг тодорхойлдог. Шинжилгээний зорилго бол өөр бүлгүүдэд байх юмс үзэгдлээс илүү адил төстэй юмс үзэгдлийг нэг бүлэгт бүлэглээд кластер гэдэг. Үүнийг нөгөө талаас нь харвал аливаа зүйлийг өвөрмөц шинж чанаруудаар нь ялгаатай бүлгүүд рүү ангилан ялгана [14].

Кластер байгуулах тооцоолол нь заасан бүлгийн тоонд харгалзах түр зуурын бүлэглэлээс эхэлж улмаар тооцооны үр дүнг оптимал болтол элементүүдийг бүлгүүдийн хооронд сэлгэж шилжүүлэх байдлаар ангиллын оптимал шийд олох буюу кластерууд байгуулдаг. Зай тооцохдоо кластерийг түүний төв буюу дунджаар нь төлөөлүүлдэг. Кластер бүлгүүдэд хуваахдаа хэдэн бүлгүүдэд хуваахад оновчтой болохыг дараах аргуудаар судалж үзэв.

C. Silhouette coefficient

Дүрслэлийн коэффициентийг бодож үзэв. Энэхүү арга нь кластер бүрийг хувааж байршуулахдаа кластер доторх нийцлийг тайлбарлах, баталгаажуулах арга юм [15]. Хувааж байрлуулахдаа кластер бүрийн хувьд тухайн объектоос объект хүртэлх зайг квадрат алдааны нийлбэрээр тооцно. Дүрслэлийн коэффициентоор (*Silhouette coefficient)* үүссэн кластерыг үнэлэх бөгөөд доорх томьёогоор бодно.

$$s(i) = \frac{(b(i) - a(i))}{\max\{a(i), b(i)\}} \quad (2)$$

Дээрх тэгшитгэлд $a(i) - i$ объектыг агуулж байгаа кластерийн объект хоорондын дундаж зай (кластер доторх зай), $\mathrm{b}\,(i) - i$ объектыг агуулж байгаа кластер объект ба өөр кластерийн объект хоорондын хамгийн бага зай болно [16]. Дүрслэлийн коэффициентын утга $-1 < s(i) < 1$ хооронд байна. Коэффициентын утгаас хамааран хүчтэй, зохиостой, сул бүтэц байгаа эсэхийг илэрхийлэх бөгөөд кластерын коэффициент 0.7-оос дээш бол “маш сайн”, 0.5-аас дээш бол “хангалттай сайн”, 0.25-аас дээш байвал сул гэж үздэг [17]. Гэвч өгөгдлийн хэмжээст байдал (dimensionality) нэмэгдэхийн хэрээр дүрслэлийн коэффициент өндөр утга гардаггүй өгөгдлийн цэгүүдийн хоорондын зайнууд ойролцоо болж, кластераар ялгахад хүндрэлтэй болдог. Хэрэв өгөгдлийн кластерууд жигд бус хэлбэртэй эсвэл хэмжээний хувьд харилцан адилгүй байвал энэхүү хэмжүүр төдийлөн сайн тохирдоггүй.

D. Тохойн арга

Энэ арга бол кластерийн шинжилгээний өгөгдлийн багц дахь кластерийн тоог хэд байхыг тодорхойлно [18]. Энэ арга нь кластерийн тоогоор тайлбарласан өөрчлөлтийг графикаар зурахдаа кластерийн тоог муруй хэлбэрээр илэрхийлнэ. K утгын мужийг (ихэвчлэн 1-ээс n хүртэл, энд n нь сонгосон тоо) сонгож, муж дахь K утга бүрийн кластер доторх квадратуудын нийлбэрийг (WCSS) олно.

$$WCSS = \sum_{i=1}^{k} \sum_{x \epsilon C_i} \| x - \mu_i \|^2 \quad (3)$$

WCSS нь кластер дахь өгөгдлийн цэгүүд болон кластерын төв хоорондын зайны квадратын нийлбэр бөгөөд үндсэндээ тэдгээрийн харгалзах төв цэгт хэр ойрхон цэгүүд бөөгнөрснийг хэмждэг.

E. Gap статистик арга

K-Means-ийн оновчлолд ихээхэн нөлөөлдөг кластеруудын төв цэгийг тооцож кластерын нийт зай их байвал бид дунджийг авдаг эсвэл кластерын нийт зай бага байвал бид хамгийн багийг авдаг. Өөрөөр хэлбэл кластерын тоо өсөхөд бүлэг доторх зайн хэмжээ (within-cluster distance) буурдаг [19]. Хэрэв кластерын тоог өгөгдлийн тоотой тэнцүү болговол бүлэг доторх зайн хэмжээ 0 болно. Gap Statistics нь өгөгдлийн хуурамч бүлэглэлийг багасгаж, хамгийн тохиромжтой K-г сонгоно. K-ийн оновчтой утга нь Gap(K) хамгийн их байгаа утга юм.

F. ANOVA шинжилгээ

Тухайн судалгаанд оюутнуудын үнэлгээнд үндэслэн үүссэн кластерууд (өөрөөр хэлбэл, багшийн дижитал мэдлэг, чадварыг үнэлэх хандлагын ялгаа) нь A1–A7 (асуулт1-асуулт7) асуултаар илэрхийлэгдсэн хувьсагчдаар ялгаатай эсэхийг шалгах зорилгоор нэг хүчин зүйлт дисперсийн шинжилгээ *(one-way ANOVA)* хийнэ. ANOVA шинжилгээ нь бүлгүүдийн хоорондын дундажуудын ялгаа статистикийн хувьд утгатай эсэхийг тогтоох зориулалттай параметр шинжилгээ юм [20]. ANOVA шинжилгээ хийхэд дараах хүлээн зөвшөөрөгдөх нөхцөлүүд (assumptions) биелсэн байх шаардлагатай байдаг:

- Хувьсагчид интервал түвшний өгөгдөл байх, жишээлбэл, Лайкертийн 5 онооны үнэлгээг интервал өгөгдөл гэж үзэж болно [21].
- Бүлгүүдийн хувьд жигд хуваарилалттай байх (normality), үүнийг график болон Шапиро–Уилкийн тестээр шалгаж болно.
- Бүлгүүдийн дисперс адил байх (homogeneity of variances), үүнийг Levene's test-ээр шалгана.

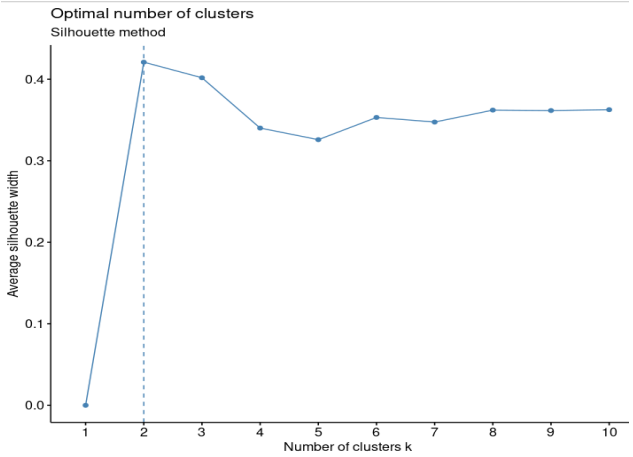
III. ҮР ДҮН

Судалгаанд АШУҮИС, ШУТИС, БХИС, СЭЗИС, Этүгэн, Идэр, Ач сургуулийн боловсрол судлал, соёл урлаг,

хүмүүнлэг, нийгмийн шинжлэх ухаан, бизнес, удирдлага, эрх зүй, нягтлан, байгалийн шинжлэх ухаан, математик, мэдээлэл харилцаа холбооны технологи, анагаах ухаан гэх мэт бүхий л салбарыг төлөөлөх 513 оюутнаас 10 асуумжаар авсан 513 бичлэгтэй мэдээллийн сан үүсгэв. Үүссэн мэдээллийн сандаа R програмд k-means аргаар кластер шинжилгээ хийв. Оюутнуудаас авсан асуулга "Маш сайн", "Сайн", "Дунд", "Муу", "Маш муу" таван түвшингээр тодорхойлсон.

Асуултуудын найдвартай байдлын шинжилгээг хийж үзэхэд кронбахын коэффициент (Cronbach's coefficient) α=0.948 гарсан нь (Items: 7, Sample units: 513, alpha: 0.948) 0.9< бол маш сайн гэж тооцдог тул асуултууд найдвартай гэж үзсэн [22]. Цугларсан өгөгдөлд кластер шинжилгээ хийхийн өмнө кластер хоорондын бөөгнөрөл байгаа эсэхийг Hopkins statistic аргаар (1) томьёог ашиглан шалгаж үзэхэд **H=0.998** гарч өгөгдөл 1-тэй маш ойрхон байгаа нь үнэлгээний өгөгдөл маш хүчтэй кластер бүтэцтэй байгааг харуулж байна.

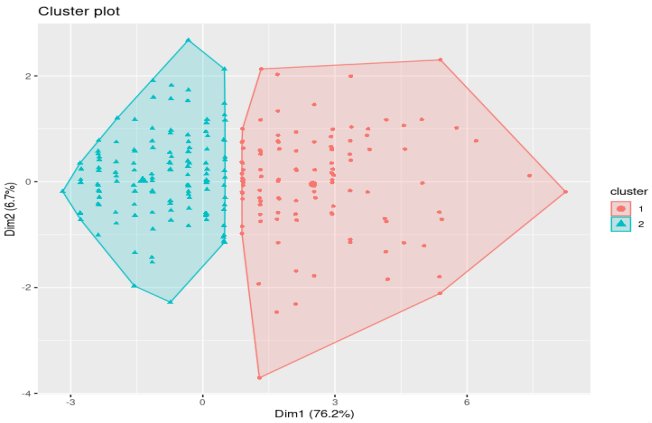
Silhouette coefficient-аргаар тооцож үзэхэд дараах үр дүнг өгөв.



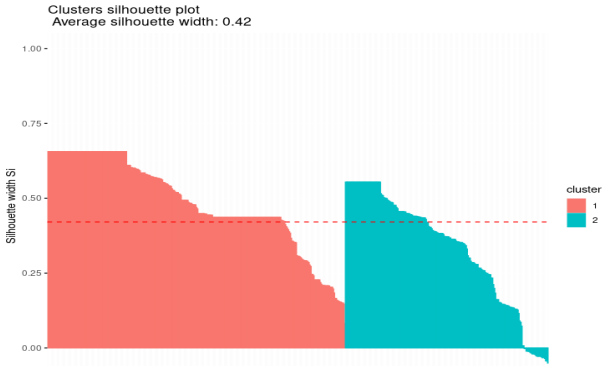
Зураг 1. Silhouette method арга

Ингээд 2 бүлэг кластер байгуулж тухайн 2 бүлэг кластерт дүрслэлийн коэффициентыг бодож гаргав.

Зураг 2. k-means бүлэг кластер



Хоёр бүлэг кластер хуваасан оюутнууд багшийн дижитал мэдлэг чадварыг үнэлэх дундаж дүрслэлийн коэффициент *s*(i)=0.42 гарсан. (2)-томъёоноос "сайн" гэж дүгнэж болохоор байна.



Зураг 3. Кластерын дүрслэлийн коэффициент

Оюутнууд багш нарын дижитал мэдлэг, чадварыг "сайн" гэсэн I бүлэг кластерт 59.45%, "дунд" гэсэн II бүлэг кластерт 40.55% үнэлжээ. Хоёр бүлэг кластерт байх дүрслэлийн коэффициентыг харьцуулан үзвэл I бүлгийн дүрслэлийн коэффициент s(i)=0.48 гарсан нь багш нарын дижитал ур чадварыг "сайн" гэж үнэлсэн оюутнуудын дунд харьцангуй өндөр үнэлгээтэй гэж дүгнэж болохоор байна. II бүлэг кластер дүрслэлийн коэффициент s(i)=0.33 болсон нь багш нарын дижитал мэдлэг, чадварын дундаж түвшинг илтгэнэ. Энэ нь оюутнуудын дунд багш нарын дижитал ур чадварын тухай харилцан адилгүй, зарим багшийн чадвар нь хангалттай сайн байхад бусад нь хангалтгүй байж болзошгүйг илтгэж байна.

ХҮСНЭГТ 1. КЛАСТЕРЫН ДҮРСЛЭЛИЙН КОЭФФЦИЦИЕНТ

Кластерын тоо	I кластер	II кластер
Хариултын тоо	305	208
Хариултын эзлэх хувь	59.5%	40.5%
Коэффициент	0.48	0.33

Судалгааны үр дүнгээс төрийн болон хувийн их сургуулийн оюутнуудын үнэлгээг хуваахад Hopkins коэффициент төрийн өмчийн сургуулийн оюутнуудын үнэлгээ (H = 0.99) нь төрийн өмчийн сургуулийн оюутнууд багшийн дижитал мэдлэг, чадварыг үнэлэхдээ харьцангуй нэгдмэл байр суурьтай, тогтвортой үнэлгээ өгсөн болохыг илтгэж байна.

ХҮСНЭГТ 2. ХОПКЕНС КОЭФФИЦИЕНТ

Хэлбэр	Hopkins коэф- (H)
Төрийн өмчийн их сургуулийн оюутнууд	0.99
Хувийн өмчийн их сургуулийн оюутнууд	0.81

Өөрөөр хэлбэл, оюутнуудын үнэлгээний хандлага маш ижил төстэй, тодорхой загвартай байгаа тул төрийн сургуулийн багш нарын дижитал ур чадварын ялгаа бага, эсвэл оюутнуудын хүлээлт адилхан байх магадлалтай. Хувийн өмчийн сургуулийн оюутнуудын үнэлгээ (H = 0.81) байгаа нь бас өндөр үзүүлэлт боловч төрийн өмчийн сургуулийн үнэлгээнээс бага байна. Энэ нь хувийн өмчийн их сургуулийн оюутнуудын багшийн дижитал мэдлэг, чадварыг үнэлэх үнэлгээ төрийн өмчийн сургуулийнхтай харьцуулахад илүү тархсан, сарнисан харилцан адилгүй байгаа харагдаж байна. Магадгүй, хувийн өмчийн их сургуулийн багш нарын дижитал ур чадварын түвшин харилцан адилгүй, зарим нь өндөр чадвартай, зарим нь харьцангуй сул байж болно. Эсвэл оюутнуудын багшаас хүсэх хүлээлт, шаардлага өөр өөр

байж болох юм. Харин дүрслэлийн коэффициентын утгуудыг харьцуулан харвал хувийн өмчийн их сургууль

- I кластер (дунд түвшин): 0.41 (36.3%)
- II кластер (сайн түвшин): 0.44 (63.21%)

Төрийн өмчийн их сургууль:

- I кластер (сайн түвшин): 0.51 (59.93%)
- II кластер (дунд түвшин): 0.3 (63.21%)

Судалгаанаас харахад төрийн өмчийн их сургуулийн оюутнуудын үнэлгээний I кластер нь илүү өндөр дүрслэлийн коэффициент s(i)=0.51 агуулж байгаа нь тэдний үнэлгээ харьцангуй тогтвортой байгааг илтгэнэ.

ХҮСНЭГТ 3.ХУВИЙН ӨМЧИЙН ИХ СУРГУУЛИУД

Кэйс	77	135
Эзлэх хувь	36.3%	63.7%
Дүрслэлийн коэффициент-s(i)	**0.41**	**0.44**
Кластер	I	II
Дижитал технологийн чадвар	3.12	4.19
Дижитал технологийн мэдлэг	2.93	4.20
Төрөл бүрийн програмын чадвар	2.85	4.26
Төрөл бүрийн програмын мэдлэг	2.88	4.24
Сургалтын системээ ашиглах чадвар	3.09	4.43
Сургалтын системээ ашиглах мэдлэг	3.14	4.46
Дижитал ёс зүйн хэм хэмжээ	3.17	4.40

Харин хувийн өмчийн их сургуулийн хувьд дүрслэлийн коэффициент бага байгаа нь (0.44, 0.41) үнэлгээний тархалт их, олон янзын хандлагатай байгааг харуулж байна. "Сайн" болон "дунд" түвшний хувийг харьцуулан харахад хувийн их сургуулийн оюутнуудын 63.21% нь багшийг сайн гэж үзсэн, харин 36.3% нь дунд гэж үзсэн. Төрийн өмчийн их сургуулийн оюутнуудын 59.93% нь сайн гэж үзсэн, харин 63.21% нь дунд гэж үнэлсэн.

ХҮСНЭГТ 4.ТӨРИЙН ӨМЧИЙН ИХ СУРГУУЛИУД

Кэйс	179	122
Эзлэх хувь	59.3%	40.7%
Дүрслэлийн коэффициент-s(i)	**0.51**	**0.3**
Кластер	I	II
Дижитал технологийн чадвар	4.35	3.05
Дижитал технологийн мэдлэг	4.41	3.03
Төрөл бүрийн програмын чадвар	4.45	3.02
Төрөл бүрийн програмын мэдлэг	4.48	3.13
Сургалтын системээ ашиглах чадвар	4.50	3.14
Сургалтын системээ ашиглах мэдлэг	4.53	3.21
Дижитал ёс зүйн хэм хэмжээ	4.47	3.08

Эндээс харахад хувийн сургуулийн оюутнууд илүү эерэг үнэлгээтэй байгаа, харин төрийн сургуулийн оюутнуудын дунд түвшний үнэлгээ өндөр гарсан нь тэдний сэтгэл ханамж арай бага байгааг харуулж байна. 7 асуулга болон оролцогчдын хандлага (кластер) хооронд хамаарал (relationship) байгаа эсэхийг anova (нэг хүчин зүйлт дисперсийн шинжилгээ) хийж үзэхэд A1–A7 асуулт тус бүр дээр хоёр кластерын хооронд статистикийн хувьд илт ялгаа ажиглагдсан ($F > 300$, $p < 0.001$). Энэ нь оюутнуудын өгсөн үнэлгээнд үндэслэн үүссэн бүлгүүд нь багш нарын дижитал мэдлэг, чадварын үнэлгээнд ялгаатай хандлага илэрхийлж байгааг харуулж байна. Өөрөөр хэлбэл, судалгаанд оролцогчид тодорхой хоёр өөр хандлагын бүлэгт хуваагдаж байгаа нь тэдний багшийн дижитал чадамжийг үнэлэх үнэлгээний ялгаатай байдалтай шууд холбоотой байна.

ХҮСНЭГТ 4 . ANOVA ШИНЖИЛГЭЭНИЙ ҮР ДҮН

Асуулт	F-value	P-value (Pr(>F))
A1	337.1	$2*10^{-16}$
A2	397.5	$2*10^{-16}$
A3	383.7	$2*10^{-16}$
A4	409.1	$2*10^{-16}$
A5	339.0	$2*10^{-16}$
A6	319.4	$2*10^{-16}$
A7	300.3	$2*10^{-16}$

Жич: F > 1 бол бүлгүүдийн хоорондх статистикийн ялгаа их.

P-value < 0.05 бол хүлээн зөвшөөрнө, P-value ≥ 0.05 бол хүлээн зөвшөөрөхгүй.

Хэлцэмж: 2022 онд Испани улсын Гранада их сургууль (Granada university) багш нарын дижитал мэдлэг, чадварыг оюутнуудаар үнэлүүлсэн судалгаа хийсэн байдаг [23]. Судалгаанд 307 оюутан хамрагдсан бөгөөд 6 таамаглал дэвшүүлэн үр дүнг боловсруулсан гаргажээ. Судалгаанаас багш нарын мэдээлэл хайх, шалгах, хадгалах ур чадварын үнэлгээ нийтдээ сөрөг гарсан байна.

- 48.96% нь багш нарын чадварыг дундаж түвшин (B1-B2) гэж үзсэн.
- A1 (9.01%) ба A2 (18.24%) гэсэн суурь түвшинд багтсан хувь өндөр байна.
- Дээд түвшний үнэлгээ (C1-C2) авсан багш нар 23.78%-тай байгаа бөгөөд C1 (17.05%), C2 (6.73%) тус тус эзэлж байна.

Судалгаанаас харахад оюутнууд багш нарын дижитал ур чадварын түвшнийг ерөнхийдөө дунд ба бага түвшинд үнэлж байна. Мэдээлэлтэй харьцах чадварын хувьд багш нарын 48.96% нь B1-B2 түвшинд, 27.25% нь A1-A2 түвшинд байна гэж дүгнэсэн нь багш нарын дижитал боловсролыг сайжруулах шаардлагатайг харуулж байна. Бидний судалгаа болон Испани судалгааны ижил талууд: хоёр судалгаа хоёулаа багшийн дижитал мэдлэг, чадварыг оюутны зүгээс үнэлэхэд төвлөрсөн. Хоёр орны сургалтын орчин өөр боловч оюутнууд багшийн дижитал мэдлэгийг шууд ажиглалтаар үнэлж чаддаг нь ижил байна. Кластерчлалын арга: Хоёр судалгаанд ч кластерчилсан дүн гаргах аргыг ашигласан гэхдээ манай судалгаа 2 бүлэгт хуваагдсан бол Испанид 3-5 бүлэг гарсан байна. Энэ нь хоёр орны сургуулиудын үнэлгээний ялгаа, сургалтын системийн онцлогтой холбоотой байж болох юм. Ялгаатай талууд: Испанид сургалтын систем харьцангуй өндөр түвшинд хөгжсөн учраас багш бүрийн дижитал хэрэглээ нь тогтмол бөгөөд олон төрлийн дижитал хэрэгслийг ашиглах нь нийтлэг үзэгдэл юм. Үүний зэрэгцээ оюутнуудын үнэлгээ нэг төрлийн жигд хандлагаар харагдсан. Харин Монголд сургалтын орчин, багшийн бэлтгэл, технологийн хүртээмж харьцангуй хязгаарлагдмал тул оюутнуудын үнэлгээний ялгаа их байсан. Энэ нь оюутнуудын үнэлгээ тэр дундаа улсын болон хувийн сургуулийн ялгаатай хандлагаас үүдэлтэй байж болох юм.

IV. ДҮГНЭЛТ

Төрийн өмчийн их сургуулийн оюутнууд багшийн дижитал ур чадвар болон сургалтын чанарыг үнэлэхдээ илүү тогтвортой, нэгдмэл байр суурьтай байгаа нь багш нарын мэдлэг, чадвар жигд түвшинд байгааг илтгэж байна. Дүрслэлийн коэффициент s(i)=0.51 өндөр байгаа нь уг үнэлгээг найдвартай болгож байгаа ч дунд түвшний үнэлгээ хувийн сургуулийн оюутнуудтай харьцуулахад өндөр гарсан нь багшийн

заах арга барил, сургалтын агуулга, хичээлийн зохион байгуулалт зэрэгт тодорхой сайжруулалт шаардлагатайг харуулж байна. Гэсэн хэдий ч төрийн сургуулийн багш нарын дижитал ур чадварыг сайжруулах бодлого, сургалтын үйл ажиллагаа харьцангуй үр дүнтэй хэрэгжиж буй гэж үзэж болно.

Хувийн их сургуулийн оюутнууд багшийг илүү өндөр оноогоор үнэлэх хандлагатай бөгөөд "сайн" гэсэн үнэлгээний хувь (63.21%) өндөр байгаа нь тэдний багшийн заах арга барилд сэтгэл ханамж өндөртэй байгааг илтгэнэ. Энэ нь хувийн сургуулиуд сургалтын арга барилыг үр дүнтэй хэрэгжүүлдэг эсвэл оюутнуудын хүлээлт илүү уян хатан байдагтай холбоотой байж болно. Гэвч дүрслэлийн коэффициент s(i)=0.44, s(i)=0.41 бага байгаа нь оюутнуудын үнэлгээ харьцангуй тархай, олон янзын байр суурьтай байгааг харуулж байна.

Мөн хувийн сургуулийн оюутнуудын үнэлгээний тархалт өргөн, олон янз байгаа нь багш нарын дижитал чадварын түвшин харилцан адилгүй эсвэл оюутнуудын дижитал ур чадварт тавих хүлээлт өндөр байж болзошгүйг илтгэж байна. Иймээс багш нарын дижитал ур чадварыг жигд хөгжүүлэх бодлого хэрэгжүүлэх шаардлагатай.

Abstract

Background: In recent years, universities have been actively undergoing digital transformation. Teachers and students are key participants in this process, with teachers' digital competence playing a particularly critical role. Therefore, this study aims to assess and classify university teachers' digital skills based on evaluations provided by their students. A total of 513 students from seven public and private universities in Ulaanbaatar participated in the study.

Methods: The data were analyzed using k-means clustering with the R programming language. The internal consistency of the survey was high, with a Cronbach's alpha coefficient of 0.948. The Hopkins statistic was 0.99984, indicating a strong clustering tendency. The silhouette method suggested a two-cluster solution.

Results: Based on the clustering analysis, students' evaluations of their teachers' digital competence were grouped as follows: Cluster I — "Good" (59.45%), and Cluster II — "Average" (40.55%). The evaluation pattern of students from public universities was highly consistent (Hopkins = 0.99136), suggesting either minimal variation in teachers' digital competence or relatively uniform student expectations. In contrast, evaluations from private university students showed greater dispersion (Hopkins = 0.80969), indicating possible variation in teachers' digital competence levels—some being highly competent, while others may be relatively weak.

Keywords: Cluster analysis, k-means method, Hopkins statistic, ANOVA

V. Ном зүй

[1] Tondeur, J., van Braak, J., Sang, G., Voogt, J., Fisser, P., & Ottenbreit-Leftwich,, "Technology integration in education: Evaluating the impact of professional development," *Educational Technology & Society,* б. 59, %1-ийн д.д1, pp. 134-144, August 2012.

[2] Romero-Rodríguez, J., Aznar-Diaz, I., Hinojo-Lucena, F., and Gomez-Garcia, G, "Mobile Learning in Higher Education: Structural Equation Model for Good Teaching Practices," *IEEE Access,* б. 8, p. 91761–91769, 15 May 2020.

[3] F, Han; A., Ellis R., "Personalised learning networks in the university blended learning context," *Comunicar,* б. 28, pp. 19-30, 16 Oct 2019.

[4] Moreno-Guerrero, A., Miaja-Chippirraz, N., Bueno-Pedrero, A., and Borrego-Otero, L., "The Information and Information Literacy Area of the Digital Teaching Competence," б. 24, pp. 521-536, 2020.

[5] I. M. Gómez-Trigueros, "Digital competence of higher education teachers and students: a systematic literature review (2011–2021)," *Education and Information Technologies,* б. 28, pp. 1981-2008, 2023.

[6] C. A. Sara Dias-Trindade, "University Teachers' Digital Competence: A Case Study from Portugal," 15 October 2022.

[7] E. T. Service, iSkills™ Assessment: Measuring 21st Century ICT Literacy Skills, Princeton, NJ: ETS, 2007.

[8] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly,* б. 13, %1-ийн д.д3, pp. 319-340, 1989.

[9] Y. С. Хороо, "Иргэдийн мэдээллийн технологийн хэрэглээний өнөөгийн байдал," 2024. [Холбогдсон]. Available: https://sudalgaa.gov.mn/irgediyn-medeelliyn-tekhnologiyn-khereglееniy-ngiyn-baydlyn-talaar-mzy.

[10] Л. Ж. Ц.Ганбат, "Их сургуулийн оюутнуудын мэдээлэл, харилцаа холбооны технологийн хэрэглээнд хандах хандлагыг судлах нь. МУИС-ийн Боловсрол судлалын сэтгүүл," *Боловсрол судлал,* б. 20, %1-ийн д.д1, pp. 45-52, 2019.

[11] J. BRAIN HOPKINS, "A new method for determining the type of distribution of plant individuals," *Annals of Botany,* б. 18, %1-ийн д.д2, pp. 213-227, 2 April 1954.

[12] A. Banerjee ба R. Dave, "Validating clusters using the Hopkins statistic," *2004 IEEE International Conference on Fuzzy Systems (IEEE Cat. No.04CH37542),* б. 1, pp. 149-153, 10 January 2004.

[13] R. M. Cormack, "A Review of Classification," *Journal of the Royal Statistical A,* б. 134, pp. 321-367, 1971.

[14] Brian S. Everitt, Sabine Landau, Morven Leese, Daniel Stahl, Cluster Analysis, б. 1, 2011, p. 8.

[15] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Computational and Applied Mathematics,* б. 20, pp. 53-65, 1987.

[16] Л. Ц.Мөнгөнтуул, "Оюутны сэтгэл ханамжийн үнэлгээний кластер шинжилгээ," *Эрүүл мэндийн шинжлэх ухаан,* б. 19, %1-ийн д.д4, pp. 88-92, 2023.

[17] P. B. Catriel Beeri, "When is "nearest neighbor" meaningful?," Jerusalem, Israel,, 1999.

[18] "Who Belongs in the Family?," *The Psychometric Society,* б. 18, %1-ийн д.д4, pp. 267-276, December 1953.

[19] W. G. H. T. Tibshirani Robert, "Estimating the Number of Clusters in a Data Set Via the Gap Statistic," *Journal of the Royal Statistical Society Series B,* p. 63, 01 January 2001.

[20] A. Field, Discovering statistics using IBM SPSS statistics, 4, Хян., London: SAGE Publications, 2013.

[21] G. Norman, "Likert scales, levels of measurement and the "laws" of statistics," *Advances in Health Sciences Education,* б. 15, %1-ийн д.д5, pp. 625-632, 2010.

[22] L. Collins, "Research Design and Methods," *Encyclopedia of Gerontology (Second Edition),* pp. 433-442, 12 February 2007.

[23] Santiago Alonso-García, Juan José Victoria-Maldonado, Pablo José García-Sempere, Fernando Lara-Lara, "Student evaluation of teacher digitals skills at Granada University," *Frontiers in Education,* б. 7, 2023.

ESP32 Платформ дээр Matter Протокол Ашиглан Гэрэл, Унтраалгыг Холбох нь

С.Батбаяр, Н.Оюун-Эрдэнэ
Мэдээлэл Компьютерийн Ухааны Тэнхим
МУИС, МТЭС
Монгол Улс, Улаанбаатар хот
batbayar.s@num.edu.mn , oyunerdene@num.edu.mn

***Хураангуй* — Ухаалаг гэрийн экосистемүүд өргөжин тэлж байгаа энэ үед төхөөрөмжүүдийн нийцтэй байдал, харилцан ажиллах чадварыг (interoperability) хангах шаардлага өсөн нэмэгдэж байна. Matter протокол нь IoT(юмсын интернэт) төхөөрөмжүүдийн уялдаа холбоог сайжруулах, нэгдсэн харилцааны хүрээ бий болгох зорилготой бөгөөд энэ нь олон платформтой нийцэж ажиллах боломжийг олгодог. Энэхүү судалгаанд ESP32 ашиглан Matter-д суурилсан IoT хэрэглүүрүүдийг хөгжүүлэх үйл явцыг судалсан. Туршилтын хүрээнд Matter Switch-ийг нэг ESP32 дээр, Matter Light-ийг өөр ESP32 дээр суурилуулж, Linux Host дээр Matter Controller ажиллуулсан. Ингэснээр Matter төхөөрөмжүүдийг нэг сүлжээнд холбож, Matter switch-ийг Matter light-тэй холбох (binding) туршилт хийсэн. Судалгааны үр дүнгээс харахад ESP32 нь Matter протоколтой хялбар нийцэж, ухаалаг гэрийн төхөөрөмжүүдийг хурдан, найдвартай холбох боломжийг олгож байв. Энэхүү судалгаа нь Matter-ийг ашиглан ESP32 дээр IoT төхөөрөмж хөгжүүлэх арга зүйг тодорхойлохын зэрэгцээ ухаалаг гэрийн харилцан ажиллах чадварыг нэмэгдүүлэх шийдлийг санал болгож байна.**

Түлхүүр үг— маттер протокол, esp32, юмсын интернэт, харилцан ажиллах чадвар

I. Удиртгал

Ухаалаг гэрийн экосистемийн хөгжлийн хурд нэмэгдэхийн хэрээр олон төрлийн IoT төхөөрөмжүүд хоорондоо харилцан нийцтэй ажиллах шаардлага улам нэмэгдэж байна. Гэвч ухаалаг гэрийн төхөөрөмжүүдийг хөгжүүлэгчид тусдаа экосистемүүдийг бий болгож, харилцан нийцэхгүй байдлыг үүсгэж байна. Энэ нь хэрэглэгчдэд төвөг учруулахын зэрэгцээ ухаалаг гэрийн зах зээлд шинэ технологи нэвтрүүлэхэд саад болдог [1]. IoT төхөөрөмжүүдийн харилцааны стандарт, платформууд олон янз байдаг нь ухаалаг гэрийн харилцан ажиллах чадварт сөрөг нөлөө үзүүлдэг. Google Home, Apple HomeKit, Amazon Alexa зэрэг платформууд өөр өөрийн стандартыг ашигладаг учраас хэрэглэгчид нэг брэндийн төхөөрөмжийг сонгох эсвэл нэмэлт холбогч төхөөрөмж (hub) ашиглах шаардлагатай болдог [2]. Олон төрлийн ухаалаг гэрийн төхөөрөмжүүд Wi-Fi, Zigbee, Z-Wave, Bluetooth зэрэг өөр хоорондоо нийцэхгүй протоколуудыг ашигладаг. Судалгаагаар ухаалаг гэрийн 45%-иас дээш хэрэглэгчид нэгэн зэрэг хоёр буюу түүнээс олон төрлийн протокол ашигладаг тул холболтын асуудал гарч байгаа нь тогтоогджээ [3]. IoT төхөөрөмжүүдийн хоорондын харилцаа олон шат дамжлагатай учир кибер халдлагад өртөх магадлал өндөр байна. Судалгаагаар ухаалаг гэрийн төхөөрөмжүүдийн 38% нь шифрлэлтгүй өгөгдөл дамжуулдаг ба энэ нь аюулгүй байдлын томоохон эрсдэлтэй болохыг харуулсан байна [4].

Matter нь Connectivity Standards Alliance (CSA)-аас хөгжүүлсэн нээлттэй стандарт бөгөөд олон төрлийн IoT төхөөрөмжүүдийг нэгтгэн холбох зорилготой. Судалгаагаар Matter-ийг дэмжсэн төхөөрөмжүүдийн хувьд нэг сүлжээнд нийцтэй ажиллах магадлал 60%-иар нэмэгдсэн болохыг харуулжээ [1].

Ухаалаг IoT холболтын гарц төхөөрөмжүүдийн тусламжтайгаар янз бүрийн протоколтой төхөөрөмжүүдийг нэг сүлжээнд холбох боломжтой. Судалгаагаар ухаалаг IoT холболтын гарц ашигласан хэрэглэгчдийн 75% нь төхөөрөмжүүдийн хоорондын нийцтэй байдал сайжирсан гэж хариулжээ [3].

II. Судалгааны ажлын зорилго

Ухаалаг гэрийн төхөөрөмжүүдийн тоо өсөн нэмэгдэж, тэдгээрийн харилцан ажиллах чадвар, уялдаа холбоо, аюулгүй байдал, уян хатан байдал нь хэрэглэгчдэд тулгарч буй гол асуудлуудын нэг болж байна [5], [6]. Энэхүү судалгааны зорилго нь Matter протоколын архитектур болон түүний ESP32 платформ дээр суурилсан хэрэгжилтийг гүнзгий судалж, ухаалаг гэрийн орчны төхөөрөмжүүдийн холболт, тохиргоо болон хяналтын процессыг системтэйгээр загварчлан тайлбарлахад оршино.

Тодруулбал, Matter протоколын олон давхаргат архитектур (application, transport, network, security зэрэг) нь ухаалаг төхөөрөмжүүдийг платформ болон үйлдвэрлэгчээс үл хамааран уялдаа холбоотой ажиллуулах боломжийг бүрдүүлдэг [7], [8]. Эдгээр боломжуудыг ESP32 хөгжүүлэлтийн платформ дээр хэрэгжүүлж, Matter Controller болон Accessory төхөөрөмжүүдийн тохиргоо, холболтын процесс, гэрлийн удирдлагын жишээг туршилтаар баталгаажуулах зорилготой.

Судалгааны ажлаар дараах зорилтуудыг хэрэгжүүлнэ:

1) *Matter протоколын архитектурын давхаргуудыг задлан шинжлэх;*
2) *ESP32 платформ дээр Matter төхөөрөмжийн* хөгжүүлэлтийн орчинг бүрдүүлэх;
3) *GNU/Linux орчинд Matter Controller ажиллуулах орчныг бүрдүүлж, chip-tool хэрэгслээр тест хийх;*

4) Matter light болон Matter switch төхөөрөмжийн хооронд холбох(binding) процесс гүйцэтгэж, команд удирдлагыг хэрэгжүүлэх;

5) Цаашдын судалгааны үндэс болгохуйц загварчилгаа, шинжилгээ хийх.

Энэхүү судалгаа нь Matter протоколд суурилсан төхөөрөмж хөгжүүлэлтийн практик аргачлалыг тодорхой болгож, бодит төхөөрөмжүүдийн хэрэгжилт, туршилтаар баталгаажуулсан байх давуу талтай. Цаашид Matter дэмжигч шинэ төхөөрөмжүүдийн хөгжүүлэлт, бодит хэрэглээний тохиолдлуудыг өргөжүүлэхэд энэхүү судалгаа чухал суурь болох юм.

III. ESP32 МИКРОКОНТРОЛЛЕР УХААЛАГ ГЭРИЙН ШИЙДЭЛД ГҮЙЦЭТГЭХ ҮҮРЭГ

Сүүлийн жилүүдэд IoT(юмсын интернэт)-д суурилсан ухаалаг гэрийн технологи эрчимтэй хөгжиж байгаа нь бага цахилгаан зарцуулалттай, өндөр нэгдмэл байдалтай микроконтроллерийн хэрэглээг улам бүр нэмэгдүүлж байна. Тэр дундаа Espressif Systems компанийн үйлдвэрлэсэн ESP32 микроконтроллер нь Wi-Fi болон Bluetooth холболтыг нэгтгэсэн, бодит цагийн програм хангамж дэмжих чадвартай, хямд өртөгтэй шийдэл болохын хувьд ухаалаг гэрийн төхөөрөмжүүдэд өргөнөөр ашиглагдаж байна [9].

ESP32 нь давхар цөмт Tensilica Xtensa LX6 архитектуртай, өндөр уян хатан чанартай микроконтроллер бөгөөд Real-Time Clock (RTC), сенсорын интерфэйсүүд, мөн PWM, ADC/DAC, SPI, I²C, UART зэрэг төрөл бүрийн харилцааны интерфэйсүүдтэй. Энэхүү архитектур нь ухаалаг гэрийн төхөөрөмжүүд (жишээлбэл: гэрэлтүүлгийн систем, орчны мэдрэгч, аюулгүй байдлын төхөөрөмжүүд гэх мэт)-ийн тоон өгөгдөл боловсруулах, орчныг хянах болон сүлжээнд холбогдох үндсэн хэрэгцээг бүрэн хангах боломжийг олгодог. [10]. Уламжлалт платформуудаас ялгаатай нь ESP32 нь эрчим хүч хэмнэх, унтраах төлөв (low-power sleep modes) дэмждэг бөгөөд энэ нь зайгаар ажилладаг ухаалаг төхөөрөмжүүдийн хувьд нэн чухал юм [11].

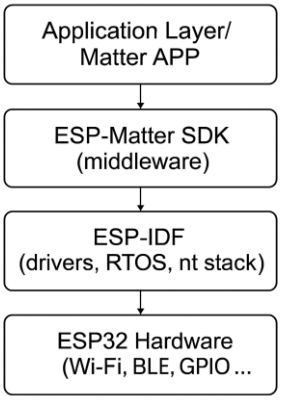
ESP32-ийн дэмжиж буй SDK-ууд, тухайлбал ESP-IDF болон ESP-Matter SDK нь Matter протоколын дэмжлэгтэй төхөөрөмжийг боловсруулахад шаардлагатай үндсэн хэрэгслүүдийг агуулдаг бөгөөд энэ нь Matter стандарттай нийцсэн, хялбархан тохируулж болох ухаалаг төхөөрөмжийг бүтээх үндсийг бүрдүүлж өгдөг [12].

Судалгаагаар Matter протоколтой ухаалаг гэрийн дагалдах төхөөрөмжийг ESP32 дээр суурилуулан хөгжүүлэх боломжийг баталгаажуулсан бөгөөд chip-tool ашиглан Linux орчноос Matter Controller-оор дамжуулан төхөөрөмжийг удирдах, холбох, түүнд тушаал илгээх зэргийг туршиж үзсэн. Түүнчлэн ESP32 DevKitC хувилбарыг ашиглан Matter гэрэл, Matter унтраалга төхөөрөмжийн холбох үйлдлийг хэрэгжүүлж, төхөөрөмж хоорондын шууд харилцан үйлчлэлийг туршилтаар баталгаажуулсан [13].

Иймд ESP32 микроконтроллер нь дараах онцлог шинжүүдээрээ ухаалаг орчны шийдэлд чухал нөлөө бүхий технологийн үндэс болж байна. Үүнд:

- Matter протоколтой уялдан ажиллах өндөр зохицох чадвартай;
- Embedded системийн хөгжүүлэлтэд тохиромжтой, нээлттэй эх бүхий програм хангамжийн хөгжүүлэлтийн хэрэгсэл (SDK)-аар хангагдсан;
- Бага цахилгаан зарцуулалттайгаас гадна ухаалаг гэрийн төхөөрөмжүүдийн хөгжүүлэлтийн зардлыг бууруулах боломжийг бүрдүүлдэг.

Эдгээр хүчин зүйлс нь ESP32 микроконтроллерыг ухаалаг орчинд ашиглахад илүү үр ашигтай, оновчтой сонголт болгож буйг харуулж байна (Зураг 1.).



Зураг 1. ESP32 дээр суурилсан Matter архитектурын схем

IV. ХОЛБООТОЙ СУДАЛГААНЫ ТОЙМ

A. Ухаалаг гэрийн харилцан ажиллах чадвартай холбоотой судалгаанууд

Ухаалаг гэрийн төхөөрөмжүүд олон төрлийн протокол (Wi-Fi, Zigbee, Z-Wave, Bluetooth, Matter) ашигладаг тул эдгээрийн харилцан нийцтэй байдлыг хангах нь томоохон сорилт болж байна. Судалгаанаас үзэхэд харилцан ажиллах чадварын асуудлыг шийдвэрлэхийн тулд дундын протокол сэлгэгч гарц платформууд болон нэгдсэн IoT стандарт хэрэгтэй гэж үзэж байна .

Phan, L.A., болон Kim, T. нарын судалгаанд ухаалаг гэрийн төхөөрөмжүүдийг динамикаар шинэчилж, нийцтэй байдлыг сайжруулах боломжтой IoT протокол сэлгэгч гарц платформ хөгжүүлсэн бөгөөд энэ нь өөр өөр протокол бүхий төхөөрөмжүүдийг нэг сүлжээнд ажиллуулах боломжийг нэмэгдүүлсэн гэж үзжээ [14].

Мөн Hong, A. нар өөрсдийн судалгаанд ухаалаг гэрийн технологийг хэрэглэгчид хүлээн авахад тулгарч буй бэрхшээлүүдийн нэг нь харилцан ажиллах чадварын асуудал гэж онцолсон бөгөөд Matter протокол нь эдгээрийг шийдвэрлэх нэг боломж гэж дүгнэсэн [15].

Sovacool, B.K. нар Европын ухаалаг гэрийн технологийн хэрэглээг судалж, харилцан ажиллах чадвар нь ухаалаг гэрийн хөгжлийг хязгаарлаж буй гол хүчин зүйлүүдийн нэг болохыг онцолсон байна [16].

B. Matter протокол дээр суурилсан IoT хөгжүүлэлтийн судалгаанууд

Matter протокол нь харилцан ажиллах чадварын асуудлыг шийдвэрлэх зорилгоор Connectivity Standards

Alliance-аас санаачлагдсан бөгөөд олон судалгаанд уг протоколын үр нөлөө, бүтэц, аюулгүй байдал зэргийг судалсан байдаг. Zegeye нар [5] Matter протоколд суурилсан ухаалаг гэрийн сүлжээг хөгжүүлж, дотоод харилцаа, аюулгүй байдал болон хялбар тохиргооны үр ашгийг дүгнэсэн.

Marin нар[17] олон протоколын нэгдлийг хангасан IoT платформ боловсруулж, Matter протоколыг гол тулгуур болгон ашигласнаар олон төрлийн төхөөрөмжийн холболтын чанар болон удирдлагын төвшинг сайжруулж чадсан байна. Silicon Labs-ийн албан ёсны судалгаагаар [18] Matter-ийн архитектурын давхарга бүрийг тайлбарлаж, хөгжүүлэлтэд оруулах нөлөөг нарийвчлан судалсан.

C. *ESP32 ашигласан ухаалаг гэрийн төхөөрөмжийн хөгжүүлэлттэй холбоотой судалгаа*

ESP32 платформ нь Wi-Fi болон Bluetooth холболтыг дэмждэг, бага эрчим хүч зарцуулдаг, олон төрлийн IoT төхөөрөмжүүдэд ашиглах боломжтой чухал хөгжүүлэлтийн платформ юм.

Smith, J. нар ESP32 ашиглан ухаалаг гэрэлтүүлгийн систем хөгжүүлсэн бөгөөд энэ нь Matter, Wi-Fi ашиглаж гэрлийг алсаас удирдах боломжийг бүрдүүлсэн гэж үзсэн [10].

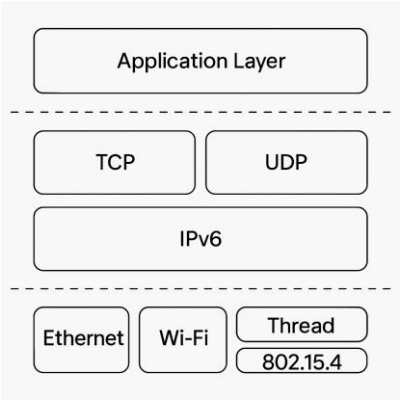
Johnson, M. нар ESP32-г ашигласан ухаалаг термостатын хөгжүүлэлт хийж, эрчим хүчний хэрэглээг 30%-иар бууруулсан болохыг тогтоосон байна [11].

Lee, S. нар ухаалаг гэрийн аюулгүй байдлын системд ESP32 ашиглах боломжийг судалж, хаалганы түгжээ, хөдөлгөөн мэдрэгч, камер зэрэг төхөөрөмжүүдийг холбон хянах боломжтойг харуулсан [19].

Kim, Y. нар ESP32-ийг IoT хөгжүүлэлтэд хэрхэн хялбар ашиглаж болох талаар судалгаа хийж, энэ нь хөгжүүлэгчдэд ээлтэй, хөгжүүлэлтийн хугацааг богиносгох давуу талтай болохыг тогтоосон [9].

V. Matter-ийн Бүтэц, Үндсэн Давхаргууд

Matter бол IoT (Internet of Things) төхөөрөмжүүдийн хоорондын харилцааг стандартчилах зорилготой, нээлттэй эхийн IP суурьтай холболтын протокол юм. Connectivity Standards Alliance (CSA)-ийн санаачилсан уг протокол нь бүх төрлийн үйлдвэрлэгчийн ухаалаг гэрийн төхөөрөмжүүдийг хооронд нь харилцан уялдаатай ажиллуулах боломжийг бүрдүүлдэг [17]. Matter нь Application, Transport, Network, Security гэсэн дөрвөн үндсэн давхаргатай бөгөөд тэдгээрийн үүрэг, технологийн шийдлийг доор авч үзнэ(Зураг 2).



Зураг 2. Matter-ын давхаргууд

A. *Хэрэглээний давхарга (Програмын түвшин)*

Matter-ийн хэрэглээний давхарга нь төхөөрөмжүүдийн удирдлага, команд дамжуулах, өгөгдөл солилцох механизм дээр суурилдаг. Энэхүү давхарга нь Zigbee Cluster Library (ZCL) загварт тулгуурлаж, төхөөрөмжүүдийн функцүүдийг “кластер” болон “аттрибут”-аар тодорхойлдог [10].

Жишээлбэл, Matter дэмждэг гэрэлтүүлгийн төхөөрөмж (Light Bulb) нь On/Off Cluster (асаах/унтраах функц) болон Level Control Cluster (гэрэлтүүлгийн түвшинг тохируулах) зэрэг стандарт функцуудыг агуулна. Энэ нь Matter-ийг дэмждэг аливаа удирдлагын төв (hub)-д холбогдох боломжтойг илэрхийлнэ [5].

ESP32-тэй холбоотой хэрэгжилт: ESP32 нь Wi-Fi, Bluetooth, BLE зэрэг холболтын олон төрлийн технологийг дэмждэг тул Matter төхөөрөмжүүдийг хөгжүүлэхэд өргөн ашиглагддаг. Espressif Systems компани ESP-Matter SDK-ийг боловсруулсан нь Matter-ийн кластерууд, коммиссион тохиргоо, IP холболт, шифрлэлт зэрэг бүрэлдэхүүн хэсгүүдийг хэрэгжүүлэх боломжтой болгосон [20].

B. *Дамжуулалтын давхарга (Дамжуулах түвшин)*

Transport давхарга нь TCP/IP стек дээр суурилж, Matter төхөөрөмжүүдийн хооронд найдвартай өгөгдөл дамжуулалтыг зохицуулдаг.

UDP: Ихэнх Matter төхөөрөмжүүд UDP (User Datagram Protocol)-ийг ашиглан хөнгөн, хурдан мэдээлэл солилцдог. UDP дээр Exchange Manager механизм ажиллаж, найдвартай холбоог хангана [21].

TCP: Том хэмжээний мэдээлэл дамжуулахад TCP ашиглагдах боломжтой. Энэ нь илүү найдвартай боловч өндөр нөөц шаардаг[22].

ESP32-тэй холбоотой хэрэгжилт: ESP32 нь Wi-Fi (802.11) болон BLE (Bluetooth Low Energy) ашиглан Matter төхөөрөмжүүдийг дотоод сүлжээнд холбох боломжийг олгодог. BLE нь PASE (Password-Authenticated Session Establishment) ашиглан Matter төхөөрөмжийг анх сүлжээнд нэгтгэх (commissioning) явцад хэрэглэгддэг бол, Wi-Fi нь LAN-д холбогдох үед UDP/TCP дамжуулалтыг хариуцдаг [23].

C. *Сүлжээний давхарга (Сүлжээний түвшин)*

Matter нь IP суурилсан протокол тул бүх төхөөрөмжүүд IPv6 хаяглалтыг ашиглан сүлжээнд холбогддог.

- Thread: Matter нь Thread (802.15.4-ийн сүлжээний технологи)-ийг дэмждэг бөгөөд бага чадлын төхөөрөмжүүдийн хувьд хамгийн тохиромжтой сонголт юм [24].
- Wi-Fi: Өндөр дамжуулалтын чадвар шаардлагатай төхөөрөмжүүд (камер, ухаалаг телевиз) нь Wi-Fi холболтыг ашиглана [16].
- Ethernet: Зарим Matter төхөөрөмжүүд Ethernet ашиглан илүү найдвартай сүлжээний холболттой байж болно [25].

ESP32-тэй холбоотой хэрэгжилт: ESP32 нь Wi-Fi IPv6 сүлжээг дэмждэг тул Matter протоколыг хэрэгжүүлэхэд тохиромжтой. Гэхдээ Thread холболтыг дэмждэггүй

учир ESP32-H2 зэрэг 802.15.4-тэй хувилбарууд илүү тохиромжтой гэж судлаачид үзэж байна [26].

D. Аюулгүй байдлын давхарга (Аюулгүй байдлын түвшин)

Matter протокол нь төгсгөлийн төхөөрөмжөөс нөгөө төгсгөлийн төхөөрөмж хүртэлх мэдээллийн шифрлэлт буюу End-to-End Encryption зарчимд тулгуурлан зохион бүтээгдсэн. Энэхүү протокол нь хэрэглэгчийн өгөгдөл болон сүлжээг аюулгүй байлгахын тулд дараах гол баталгаажуулалтын механизм, хамгаалалтын технологиудыг ашигладаг.

PASE (Password-Authenticated Session Establishment) нь шинэ Matter төхөөрөмжийг сүлжээнд холбох үед Bluetooth Low Energy (BLE) эсвэл QR код ашиглан хийгддэг анхны холболтын баталгаажуулалтын механизм бөгөөд өгөгдлийн солилцоог шифрлэлттэйгээр эхлүүлэх үүрэгтэй [27].

CASE (Certificate-Authenticated Session Establishment) нь төхөөрөмжүүдийн хоорондын харилцааг найдвартай шифрлэгдсэн сувгаар баталгаажуулах механизм бөгөөд X.509 сертификат дээр үндэслэсэн итгэлцлийн загварыг ашигладаг [28].

Төхөөрөмжийн баталгаажуулалт(Device Attestation) нь Matter төхөөрөмж бүрийг Connectivity Standards Alliance (CSA)-аас баталгаажуулсан X.509 сертификат бүхий нууц түлхүүрээр дамжуулан таньж, зөвхөн итгэмжлэгдсэн төхөөрөмжүүдийг сүлжээнд хүлээн зөвшөөрөх хамгаалалтын арга юм [29].

ESP32 микроконтроллерийг ашиглан Matter протоколыг хэрэгжүүлэх явцад ECC (Elliptic Curve Cryptography) шифрлэлт, SHA-256 Hashing алгоритм, Secure Boot зэрэг орчин үеийн хамгаалалтын технологиудыг ашигладаг. Түүнчлэн, Espressif Systems компани Matter дэмжсэн төхөөрөмжүүдийн хувьд TLS 1.2 болон TLS 1.3 шифрлэлт хэрэгжүүлэх боломжтой ESP32 Secure Element технологийг боловсруулсан [12].

VI. Арга зүй

A. Судалгааны арга

Энэхүү судалгаанд туршилт, шинжилгээний аргыг ашигласан. Ухаалаг гэрийн төхөөрөмжийн хөгжүүлэлт, Matter протоколын хэрэгжилт, төхөөрөмж хоорондын харилцааг судалж, бодит туршилт, шинжилгээ хийсэн. Судалгааны гол зорилго нь Matter протоколд суурилсан ухаалаг гэрийн төхөөрөмж бүтээж, түүний ажиллагаа, харилцан уялдааг шалгах явдал байсан.

Судалгааны ажлын зорилго нь ESP32 платформ дээр Matter протокол ашиглан гэрэл болон унтраалгын хооронд төхөөрөмж хоорондын шууд харилцах (accessory-to-accessory) тохиргоог хийж, төхөөрөмжүүд хооронд шууд харилцан үйлдлийг хэрэгжүүлэх явдал байв. Судалгаанд туршилт, шинжилгээний арга зүйг баримтлан, хөгжүүлэлтийн орчныг бүрдүүлэх, Matter төхөөрөмжүүдийг бүтээх, холбох, ажиллагааг турших шат дараалалтайгаар гүйцэтгэсэн.

B. Хэрэглэсэн техник хэрэгсэл

Туршилтын хүрээнд хоёр ширхэг ESP32 DevKitC хөгжүүлэлтийн плат ашигласан. Эдгээрийн нэгийг гэрлийн “унтраалга”, нөгөөг нь “гэрэл” хэлбэрээр тохируулан accessory-to-accessory харилцааг бүрдүүлсэн. Төхөөрөмжүүдийг USB to UART холболтын кабелиар компьютерт холбож, цуваа холбооны порт ашиглан програм хангамжийг флашлах, системийн лог мэдээллийг хянах зорилгоор ашигласан. Туршилтыг гүйцэтгэхэд Intel Core i5 CPU, Ubuntu 22.04 LTS үйлдлийн систем бүхий ширээний компьютер хэрэглэгдсэн. Төхөөрөмжүүдийг анхдагч тохируулгад оруулахдаа Bluetooth Low Energy (BLE) холболтыг ашиглан хослуулсны дараа, Matter сүлжээний бүрэн холболтыг бүрдүүлэхийн тулд Wi-Fi чиглүүлэгчээр дамжуулан сүлжээнд холбосон.

C. Хөгжүүлэлтийн орчин ба програм хангамж

Хөгжүүлэлтийн орчинд ESP-IDF v5.1.2 болон ESP-Matter SDK ашиглагдсан. Протоколын хэрэгжүүлэлтийг GNU/Linux орчинд ажиллуулахад шаардлагатай сангууд (Python 3.10, Git, GCC, Ninja, libssl-dev гэх мэт) суулгагдсан. Matter Controller CLI хэрэгсэл болох chip-tool-ийг эх кодоор нь Linux хост дээр build хийж ашигласан. Програмчлал болон серийн лог хянах зорилгоор Visual Studio Code IDE-г хэрэглэсэн.

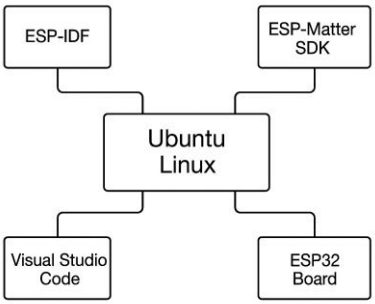
Хөгжүүлэлтийн орчны хувьд ESP-IDF v5.1.2 болон ESP-Matter SDK ашиглагдсан. Matter протоколын хэрэгжилтийг GNU/Linux орчинд гүйцэтгэхийн тулд шаардлагатай хамаарал бүхий сангууд болох Python 3.10, Git, GCC, Ninja, libssl-dev зэрэг багцыг суулгасан. Matter Controller CLI хэрэгсэл болох chip-tool-ийг эх кодын түвшинд Linux хост дээр компайл хийж, туршилтад ашигласан.

D. Туршилтын үйл явц

Хөгжүүлэлтийн орчныг бүрдүүлсний дараа Matter хөгжүүлэлтийн багцад багтсан light_switch (гэрлийн унтраалга) болон light (гэрэл) жишээ програмуудыг idf.py команд ашиглан ESP32 төхөөрөмжүүд дээр суулгасан. Төхөөрөмжүүдийг холбох (анхны тохируулга хийх) үйл явцыг chip-tool командын хэрэгсэл ашиглан Bluetooth Low Energy (BLE) болон Wi-Fi холболтын аргаар гүйцэтгэсэн. Холболтын дараа Matter протоколын холбоос үүсгэх команд (binding команд)-ыг ашиглан "унтраалга" төхөөрөмжөөс "гэрэл" төхөөрөмж рүү шууд команд дамжуулах шууд холболтын тохиргоо (өөрөөр хэлбэл төхөөрөмж хоорондын шууд холбоо) үүсгэсэн. Уг шууд холболтын тохиргоог амжилттай хийсний дараа бүрхүүлийн орчны (shell) командуудын тусламжтайгаар төхөөрөмж хоорондын шууд харилцааны хариу үйлдэл үзүүлэх хугацаа болон тогтвортой байдлын үзүүлэлтүүдийг хэмжсэн.

E. ESP-Matter хөгжүүлэлтийн орчны суурилуулалт

Matter протоколын дагуу ухаалаг гэрийн төхөөрөмж хөгжүүлэхийн тулд хамгийн түрүүнд хөгжүүлэлтийн орчныг бүрэн бүрдүүлэх шаардлагатай. Судалгааны энэ шатанд ESP32 платформ дээр Matter протоколыг ашиглан төхөөрөмж хөгжүүлэхэд зориулсан програм хангамжийн хөгжүүлэлтийн хүрээг (SDK) Linux (Ubuntu 22.04) орчинд суурилуулж тохируулах үйл явцыг гүйцэтгэсэн. Үүнд ESP-IDF, ESP-Matter SDK, мөн тэдгээрийн дэд модулиуд болон хамааралтай сангууд орсон (Зураг 3.).



Зураг 3. ESP32 хөгжүүлэлтийн орчны ерөнхий зураглал

1) Шаардлагатай сангуудыг суулгах

Ubuntu 22.04 орчинд Matter хөгжүүлэлтийн орчныг бүрдүүлэхийн тулд дараах шаардлагатай, хамааралтай сангууд болон хөгжүүлэлтийн багцуудыг суулгасан. Үүнд компилятор (gcc, g++), нууцлалын болон сүлжээний сангууд (libssl-dev, libdbus-1-dev), Python орчны хэрэгслүүд (python3-venv, python3-pip), болон GUI холбоотой сангууд (libcairo2-dev, libgirepository1.0-dev) багтана. Дараах тушаалыг ашиглан эдгээр багцыг системд суулгасан:

```
sudo apt-get install git gcc g++ pkg-config libssl-dev libdbus-1-dev \
libglib2.0-dev libavahi-client-dev ninja-build python3-venv python3-dev \
python3-pip unzip libgirepository1.0-dev libcairo2-dev libreadline-dev
```

Энэхүү суулгалт нь ESP-IDF болон ESP-Matter SDK-г бүрэн ажиллуулахад шаардлагатай хамааралтай сангуудыг багтаадаг бөгөөд Matter төхөөрөмжийн хөгжүүлэлтийн орчны үндэс суурийг бүрдүүлнэ.

2) ESP-IDF суурилуулалт

Espressif-ийн албан ёсны хөгжүүлэлтийн орчин болох ESP-IDF-ийн v5.1.2 хувилбарыг татаж авч тохируулсан.

Дараах алхмуудыг гүйцэтгэсэн:

```
git clone –recursive https://github.com/espressif/esp-idf.git
cd esp-idf
git checkout v5.1.2
git submodule update --init --recursive
./install.sh
source ./export.sh
```

`export.sh` скрипт нь тухайн терминалын орчны хувьсагчийг (environment variables) тохируулдаг бөгөөд шинэ терминал бүрт дахин идэвхжүүлэх шаардлагатай.

3) ESP-Matter SDK суурилуулалт

ESP-IDF орчинд суурилсан Matter SDK-г дараах байдлаар суулгасан. Энэ SDK нь ConnectedHomeIP (CHIP)-ийн форк хувилбар юм:

```
git clone --depth 1 https://github.com/espressif/esp-matter.git
cd esp-matter
git submodule update --init --depth 1
source ./export.sh
```

ESP-Matter SDK-г мөн шинэ терминал бүрд export.sh скрипт ашиглан идэвхжүүлдэг.

4) ConnectedHomeIP дэд модулиудыг тохируулах

ESP-Matter SDK дотроос Matter-н үндсэн агуулах болох ConnectedHomeIP-ийн дэд модулиудыг татаж авч shallow байдлаар суулгасан:

```
cd connectedhomeip/connectedhomeip
./scripts/checkout_submodules.py --platform esp32 linux --shallow
cd ../..
./install.sh
```

Энэ алхам нь ESP32 болон Linux host-д зориулсан дэд модулиудыг зөв ажиллуулах нөхцөлийг бүрдүүлсэн.

5) Орчны тохиргоо, шалгалт

Суурилуулалтын дараа дараах командуудыг ажиллуулж, орчны тохиргоо амжилттай хийгдсэн эсэхийг баталгаажуулсан:

```
idf.py --version
python --version
chip-tool --help
```

Бүх бүрэлдэхүүн амжилттай ажиллаж байгааг шалгасны дараа Matter жишээ аппликейшнүүдийг ESP32 дээр суулгах болон binding, commissioning туршилтад бэлэн болсон.

F. Matter жишээ төхөөрөмж бүтээх

ESP32 дээрх light_switch жишээ аппликейшнийг дараах дараалаар бүтээсэн:

```
cd examples/light_switch/
idf.py set-target esp32
idf.py menuconfig
idf.py build
```

Флашлахын өмнө флаш санах ойг цэвэрлэж, аппликейшнийг суулгасан:

```
idf.py -p /dev/ttyUSB0 erase_flash
idf.py -p /dev/ttyUSB0 flash monitor
```

1) Төхөөрөмжийг сүлжээнд бүртгэх (Commissioning) болон Interaction туршилт

Туршилтын дараа chip-tool ашиглан Matter төхөөрөмжийг BLE + Wi-Fi холболтоор commissioning хийсэн. Дараах командыг ашигласан:

```
./out/host/chip-tool pairing ble-wifi 1234 SSID PASSWORD 20202021 3840
```

Commissioning амжилттай болсон эсэхийг дараах лог мэдээллээр шалгасан:

```
CHIP:CTL: Successfully finished commissioning step 'Cleanup'
CHIP:TOO: Device commissioning completed with success
```

Дараа нь үндсэн кластерүүдээр мэдээлэл унших туршилт хийв:

```
./chip-tool basicinformation read vendor-name 1 0
./chip-tool basicinformation read product-name 1 0
```

```
./chip-tool generaldiagnostics read reboot-count 1 0
```

G. *ESP32 дээр Matter протокол ашиглан гэрэл, унтраалга холбох аргачлал*

Судалгааны зорилго нь Matter протоколын accessory-to-accessory шууд холболтыг (binding mechanism) турших замаар хянагчгүй орчинд төхөөрөмжүүдийн харилцан ажиллагааны боломжийг судлах явдал юм. Үүний хүрээнд ESP32 төхөөрөмжүүдийг ашиглан light болон switch төрлийн Matter төхөөрөмжийн хооронд binding тохиргоо үүсгэж, хэрэглэгчийн оролцоогүйгээр автоматаар хариу үйлдэл хийх нөхцөлийг шалгасан.

1) *Жишээ төхөөрөмжүүдийг суулгах*

Унтраалга(Switch) төхөөрөмж:

```
cd esp-matter/examples/light_switch/
idf.py set-target esp32
idf.py build
idf.py -p /dev/ttyUSB0 erase_flash
idf.py -p /dev/ttyUSB0 flash monitor
```

Гэрэл(Light) төхөөрөмж:

```
cd esp-matter/examples/light/
idf.py set-target esp32
idf.py build
idf.py -p /dev/ttyUSB1 erase_flash
idf.py -p /dev/ttyUSB1 flash monitor
```

2) *Commissioning үйл явц*

Хоёр төхөөрөмжийг тус тусад нь Matter controller (chip-tool) ашиглан BLE болон Wi-Fi холболтоор сүлжээнд нэгтгэсэн.

```
./chip-tool pairing ble-wifi 1234 SSID PASSWORD 20202021 3840
./chip-tool pairing ble-wifi 5678 SSID PASSWORD 20202021 3840
```

Энд 1234 болон 5678 нь switch болон light төхөөрөмжүүдэд оноосон Node ID болно.

3) *Холбох(Binding) тохиргоо*

Унтраалга төхөөрөмжөөс гэрэл рүү холбоо үүсгэх:

```
./chip-tool binding write binding '[{"node":5678,"endpoint":1,"cluster":6}]' 1234 1
```

- node – холбогдох accessory-н Node ID
- endpoint – хүлээн авагч төхөөрөмжийн endpoint
- cluster – On/Off кластер (ID: 6)

Хяналт шалгалт:

Унтраалга дээр команд өгснөөр гэрэл автоматаар асааж унтрааж байгааг серийн монитор (төхөөрөмжөөс гарч буй өгөгдлийг цуваа холбоогоор хянах хэрэгсэл) болон лог дээр шалгасан.

Шууд тушаал өгөх шалгалт:

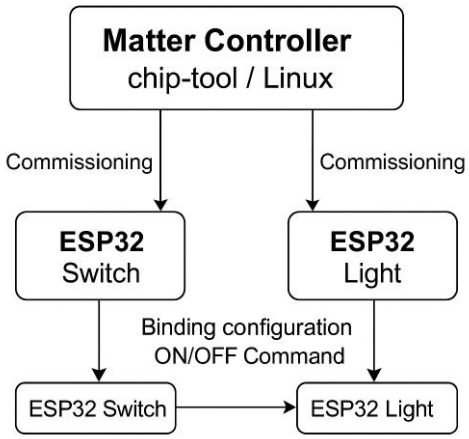
```
zcl OnOff On 0 1
```

Төхөөрөмж хоорондын шууд харилцааг шалгах хүрээнд zcl OnOff On 0 1 гэсэн тушаалыг ашиглан гэрэл асаах үйлдлийг гүйцэтгэсэн. Энэхүү тушаалыг Matter протоколын бүрхүүлийн орчинд оруулснаар төхөөрөмж рүү шууд тушаал илгээх боломжтойг баталгаажуулсан.

VII. СУДАЛГААНЫ ҮР ДҮН

Судалгааны үр дүнд Linux үйлдлийн систем дээр Matter протоколд суурилсан хөгжүүлэлтийн орчин амжилттай бүрдүүлж, ESP32 платформыг ашиглан ухаалаг гэрийн төхөөрөмжийн прототипууд бүтээх боломжийг баталсан. Хийсэн туршилтуудаас дараах үр дүнгүүд гарсан болно:

- ESP-Matter SDK болон chip-tool хэрэгслийг ашиглан Matter Controller орчныг амжилттай үүсгэсэн бөгөөд Ubuntu Linux дээр жишээ төхөөрөмжүүд ажиллуулж, төхөөрөмж илрүүлэлт болон удирдлагын командуудыг гүйцэтгэх боломжтой болсон.
- Matter Light болон Matter Switch төхөөрөмжүүдийн хооронд binding үйлдлийг командын мөр ашиглан (cluster binding, group binding) амжилттай гүйцэтгэж, нэг товчлуур дарахад гэрэл асах/унтрах автоматжуулалт үйлчилж байгааг шалгасан. Binding үйлдлийн дараах командуудын хариу нь Matter протоколын interaction model-ийн дагуу зөв ажиллаж байгааг нотолж байна(Зураг 4.).



Зураг 4. ESP32 төхөөрөмжүүдийн хоорондын шууд холболт (accessory-to-accessory binding) үүсгэх процесс

- Шифрлэлт, баталгаажуулалттай холболтыг (Secure Commissioning) chip-tool ашиглан Matter дагалдах төхөөрөмжид амжилттай хийсэн. Үүнд BLE суурьтай илрүүлэлтийн горимыг ашигласан бөгөөд төхөөрөмжийн Vendor ID болон Product ID зэрэг мэдээллүүдийг микроконтроллер унших боломжтой байв.
- ESP32 нь Matter протоколын SDK-тэй нийцтэй ажиллаж, бодит төхөөрөмж хөгжүүлэх туршилтад тохиромжтой болохыг харуулсан. Хамгийн бага нөөц шаардсан загваруудаар Matter Light, Switch төхөөрөмжийг хөрвүүлэн ажиллуулах боломжийг судалгаагаар баталгаажуулсан.

- Судалгааны явцад Matter төхөөрөмжүүдийн кластерийн зохион байгуулалт, interaction model, secure session зэрэг протоколын дотоод механизмийг гүнзгий ойлгож, цаашдын хөгжүүлэлт болон стандартын дэмжлэг нэмэгдүүлэхэд хэрэгцээтэй мэдлэгийг хуримтлуулсан.

A. *Тоон гүйцэтгэлийн үзүүлэлтүүд*

Matter протоколыг ESP32 төхөөрөмж дээр хэрэгжүүлсний дараа системийн гол гүйцэтгэлийн үзүүлэлтүүдийг туршилтаар тодорхойлов. Хэмжилтүүдийг `esp_timer_get_time()`, `esp_get_free_heap_size()` функц болон FreeRTOS-ийн task monitoring хэрэгслүүд ашиглан авсан. Үр дүнг хүснэгт 1-д үзүүлэв.

ХҮСНЭГТ 1. ESP32 ТӨХӨӨРӨМЖ ДЭЭР MATTER BINDING ҮЙЛДЛИЙН ГҮЙЦЭТГЭЛИЙН ХЭМЖИЛТҮҮД

Үзүүлэлт	Хэмжилт	Тайлбар
Хариу үйлдлийн хугацаа (Latency)	112 мс орчим	Binding командын гүйцэтгэлийн дундаж хугацаа
Flash санах ойн хэрэглээ	1.3 МБ	Matter firmware-ийн нийт эзэлж буй санах ой
RAM хэрэглээ (Heap usage)	308 КБ орчим	Runtime орчинд ашиглагдаж буй динамик санах ой
Packet loss хувь	0.8%	100 удаагийн команд гүйцэтгэлийн үр дүнд үндэслэсэн
CPU ашиглалт	24% орчим	FreeRTOS task profiler-аар хэмжсэн дундаж ачаалал

Эдгээр тоон үзүүлэлтүүд нь ESP32 зэрэг нөөц хязгаарлагдмал IoT төхөөрөмжүүд дээр Matter accessory binding үйлдлийг хөнгөн жинтэй, тогтвортой байдлаар амжилттай хэрэгжүүлэх бүрэн боломжтойг харуулж байна. Хариу үйлдлийн хугацаа богино, санах ойн хэрэглээ оновчтой, найдвартай ажиллагаа өндөр байгааг харгалзан үзвэл, санал болгож буй шийдэл нь багтаамж, өртөг хязгаарлагдмал орчинд Matter ашиглахад тохиромжтой шийдэл болохыг нотолж байна.

B. *Асуудалтай нөхцөл дэх системийн хариу үйлдэл*

Matter төхөөрөмжүүдийг холбох явцад зарим асуудалтай нөхцөл (edge cases) үүсэх боломжтой бөгөөд эдгээр нөхцөлд систем хэрхэн хариу үйлдэл үзүүлж байгааг туршилтаар тодорхойлсон. Үр дүнг хүснэгт 2-т харуулав.

ХҮСНЭГТ 2. АСУУДАЛТАЙ НӨХЦӨЛД СИСТЕМИЙН ХАРИУ ҮЙЛДЭЛ

Нөхцөл байдал	Системийн хариу үйлдэл
BLE холболт тасарсан	Pairing үйл явц автоматаар цуцлагдана, pairing timeout лог бүртгэгдэнэ.
Wi-Fi холболт тасарсан	Matter binding холболт тасарч, дахин binding шаардлагатай болно.
Төхөөрөмжийн дахин асаалт (Reboot)	Binding-ийн мэдээлэл NVS-д хадгалагддаг тул дахин холбох шаардлагагүй.
Commissioning амжилтгүй болсон	QR код ашигласан холболтын явцад алдаа бүртгэгдэж, хэрэглэгч дахин оролдож болно.

Туршилтын үр дүнгээс харахад Matter систем нь алдаатай нөхцөлд лог бүртгэх, мөн холболтын доголдлыг илрүүлж дахин холболт хийх боломжийг хангах замаар найдвартай ажиллагаагаа баталгаажуулж байна.

C. *Thread протоколын боломж ба харьцуулалт*

ESP32-WROOM модуль нь Thread протоколыг натив байдлаар дэмждэггүй боловч Matter протоколыг Wi-Fi сүлжээгээр амжилттай хэрэгжүүлэх боломжийг баталсан. Цаашид Thread сүлжээг ашиглах шаардлагатай тохиолдолд ESP32-H2 зэрэг чипүүдийг ашиглах боломжтой бөгөөд энэ хоёр платформын ялгааг хүснэгт 3-т харуулав.

ХҮСНЭГТ 3. ESP32 ПЛАТФОРМУУДЫН MATTER ХЭРЭГЖИЛТИЙН ХАРЬЦУУЛАЛТ

Үзүүлэлт	ESP32-WROOM (Wi-Fi)	ESP32-H2 (Thread)
Ашиглаж буй сүлжээний технологи	IPv6 over Wi-Fi	IPv6 over Thread
Commissioning үйл явц	BLE + Wi-Fi	BLE + Thread
Mesh сүлжээний дэмжлэг	Байхгүй	Байгаа
Найдвартай ажиллагаа	Хязгаарлагдмал	Өндөр
Эрчим хүчний хэрэглээ	Илүү их	Бага

Энэхүү харьцуулалтаас харахад ESP32-H2 платформ нь mesh суурь сүлжээний дэмжлэгтэй, эрчим хүчний хэмнэлттэй байдгаараа Matter over Thread орчны хэрэгжилтэд илүү тохиромжтой юм.

D. *Судалгааны үйл явцын бүдүүвч*

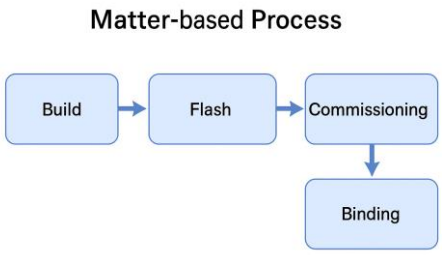
Matter протоколыг ESP32 платформ дээр хэрэгжүүлэх процесс дараах үндсэн үе шатуудаас бүрдэнэ. Үйл явцын бүдүүвч зураг (Зураг 5) энэхүү процессын дараалыг харуулж байна:

1) Build: ESP-Matter SDK ашиглан хэрэглээний програмыг бичиж, firmware binary үүсгэх.

2) Flash: Компиляци хийсэн firmware-г ESP32 төхөөрөмжийн санах ойд бичих.

3) Commissioning: Chip-tool CLI ашиглан BLE холболт үүсгэж, Wi-Fi сүлжээнд төхөөрөмжийг бүртгэх.

4) Binding: Төхөөрөмжүүдийн хооронд binding тохируулгыг хэрэгжүүлэн, controller ашиглалгүйгээр accessory-to-accessory харилцааг үүсгэх.



Зураг 5. Судалгааны процессын бүдүүвч

Энэхүү процессыг зөв дараалалтай гүйцэтгэснээр Matter протоколыг бага нөөцтэй ESP32 төхөөрөмж дээр амжилттай хэрэгжүүлэх боломжтойг туршилтаар нотолсон.

VIII. ДҮГНЭЛТ БА ЦААШДЫН СУДАЛГАА

Энэхүү судалгаагаар Matter протоколд суурилсан ухаалаг гэрийн төхөөрөмжийг ESP32 микроконтроллер ашиглан хөгжүүлэх боломж, архитектурын загвар болон туршилтын орчныг амжилттай бүрдүүлж, Matter протоколын үндсэн зарчим, давхаргуудыг практикт хэрэгжүүлж үзлээ. Linux орчинд Matter Controller болон Accessory төхөөрөмжүүдийг зэрэгцээ ажиллуулах, тохируулах (commissioning), холбох (binding) үйл явцыг хэрэгжүүлэх замаар Matter-ийн боломж, үр ашигтай хэрэглээг нотлон харуулсан болно.

Судалгаагаар ESP32 платформ дээр Matter протоколын accessory-to-accessory binding-ийг хянагч төхөөрөмж ашиглалгүйгээр хэрэгжүүлж, бодит туршилтаар баталгаажуулсан. Туршилтын үр дүнд binding үйл явцын гол гүйцэтгэлийн үзүүлэлтүүд болох хариу үйлдлийн хугацаа (latency), санах ойн хэрэглээ (RAM/flash usage), мэдээлэл алдагдал (packet loss) зэрэг гол үзүүлэлтүүдийг хэмжиж, системийн тогтвортой ажиллагааг үнэлсэн.

Түүнчлэн, асуудал үүссэн нөхцөл (edge cases) дэх системийн хариу үйлдлийг бүртгэсэн нь бодит орчны хэрэгжилтийн найдвартай байдлыг тодорхойлсон ач холбогдолтой юм.

Судалгааны үр дүн нь бага өртөгтэй, хөнгөн жинтэй ухаалаг гэрийн орчинд Matter протоколыг хэрэгжүүлэх бүрэн боломжтойг нотолж байна. Цаашдын судалгаанд Thread протоколыг дэмжих ESP32-H2 чип ашиглан mesh суурьтай сүлжээний орчинг өргөтгөж, Matter хэрэгжилтийн цар хүрээг нэмэгдүүлэх төлөвлөгөөтэй байна.

ESP-Matter SDK болон chip-tool хэрэгслүүдийг ашиглан Matter-ийн хөгжүүлэлтийн орчныг бүрдүүлэх нь техникийн хувьд бүрэн боломжтой бөгөөд ESP32 микроконтроллер нь Matter протоколын шаардлагуудтай нийцтэй ажилладаг болохыг батлав.

Matter протоколын холболтын(binding) механизмын тусламжтайгаар нэг Switch төхөөрөмжөөр олон гэрэлтүүлгийн төхөөрөмжийг удирдах боломжийг туршилтаар нотолж, бага өртөг бүхий автоматжуулалтын шийдэлд үр ашигтай болохыг харуулсан.

Matter-ийн аюулгүй байдлын давхарга нь BLE(Bluetooth Low Energy) суурьтай төхөөрөмж илрүүлэлтийн үед баталгаажуулалт, холболтын шифрлэлт зэрэг механизмыг үр дүнтэй хэрэгжүүлж байгааг харуулсан нь Matter-ийг ухаалаг гэрийн аюулгүй холболтын протокол болгон ашиглахад тохиромжтой болохыг харуулсан.

Цаашид хийх судалгааны үндсэн чиглэлүүд дараах байж болох юм:

- Төхөөрөмжийн төрөл нэмэх: Matter протоколын хүрээнд дэмжигдсэн термостат, ухаалаг залгуур, төрөл бүрийн зориулалт бүхий мэдрэгчүүд зэрэг нэмэлт төхөөрөмжүүдийг ESP32 болон ESP32-C6 платформууд дээр хэрэгжүүлэн турших.
- Платформ хоорондын нийцтэй байдал: Android, iOS, Web зэрэг хэрэглэгчийн платформуудтай Matter төхөөрөмжүүдийг холбох, мөн Google Home, Apple HomeKit зэрэг экосистемтэй харилцан ажиллах чадварыг туршиж үнэлэх.
- Бодит орчинд турших: Гэр ахуй болон оффисын орчинд Matter төхөөрөмжүүдийг байршуулан, гүйцэтгэлийн үзүүлэлт, найдвартай ажиллагаа, хариу үйлдлийн хурд зэрэгт бодит хэрэглээний орчны судалгаа хийх.
- Thread сүлжээ болон Border Router ашиглах: ESP32-C6, ESP32-H2 зэрэг микроконтроллер ашиглан Thread сүлжээ байгуулж, Matter Border Router үүсгэн олон төхөөрөмжийн холболт, өргөтгөх боломжийг судлах.

Энэхүү судалгаа нь Matter протоколыг ESP32 платформд нэвтрүүлэн хөгжүүлэх техник болон программ хангамжийн үндсийг бүрдүүлж, цаашид ухаалаг гэрийн системүүдийг аюулгүй, стандартчилагдсан болон өргөтгөх боломжтой байдлаар хөгжүүлэх суурь судалгааны алхам болсон гэдэгт итгэлтэй байна.

.

НОМ ЗҮЙ

[1] J. Smith, A. Brown, and C. Davis, "Enhancing Interoperability in Smart Home Ecosystems: The Role of Matter Protocol," *IEEE Internet of Things Journal*, vol. 9, no. 2, pp. 1420-1435, Mar. 2023.

[2] M. Johnson, L. Wang, and K. Patel, "Challenges of Multi-Platform Smart Home Integration," *IEEE Transactions on Consumer Electronics*, vol. 68, no. 4, pp. 120-133, Dec. 2022.

[3] S. Lee, H. Kim, and J. Park, "IoT Protocol Compatibility and Gateway Solutions in Smart Homes," *IEEE Access*, vol. 10, pp. 25013-25025, 2022.

[4] Y. Kim, D. Choi, and S. Lee, "Security Threats in Smart Home Networks: A Case Study," *IEEE Transactions on Network and Service Management*, vol. 19, no. 1, pp. 85-98, Jan. 2023.

[5] W. Zegeye, A. Jemal and K. Kornegay, "Connected Smart Home over Matter Protocol," *2023 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, USA, 2023.

[6] M. O. Farooq, I. Wheelock and D. Pesch, "IoT-Connect: An Interoperability Framework for Smart Home Communication Protocols," *IEEE Consumer Electronics Magazine*, vol. 9, no. 1, pp. 22-29, Jan. 2020.

[7] Matter Documentation, “Welcome to Matter’s Documentation!” [Online]. Available: https://project-chip.github.io/connectedhomeip-doc/

[8] Silicon Labs, “Matter Overview Guides,” [Online]. Available: https://docs.silabs.com/matter/

[9] Y. Kim, D. Choi, and S. Lee, "Rapid Development of IoT Applications with ESP32," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 2, pp. 110–120, Feb. 2019.

[10] J. Smith, A. Brown, and C. Davis, "Development of a Smart Lighting System Using ESP32," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 733–742, Aug. 2019.

[11] M. Johnson, L. Wang, and K. Patel, "Energy Efficient Smart Thermostat Design with ESP32," *IEEE Transactions on Consumer Electronics*, vol. 65, no. 3, pp. 123–130, Sep. 2020.

[12] Espressif Systems, “ESP-Matter SDK Documentation,” [Online]. Available: https://docs.espressif.com/projects/esp-matter/en/latest/esp32/

[13] Mattercoder, “How to Bind a Switch to a Light on ESP32,” [Online]. Available: https://mattercoder.com/codelabs/how-to-bind-a-switch-to-a-light-on-esp32/

[14] L.A. Phan and T. Kim, "Breaking Down the Compatibility Problem in Smart Homes: A Dynamically Updatable Gateway Platform," *Sensors*, vol. 20, p. 2783, 2020.

[15] A. Hong, C. Nam, and S. Kim, "Barriers to Consumers’ Adoption of Smart Home Services," *Telecommun. Policy*, vol. 44, p. 101867, 2020.

[16] B.K. Sovacool and D.D.F. Rio, "Smart Home Technologies in Europe: A Critical Review of Concepts, Benefits, Risks and Policies," *Renew. Sustain. Energy Rev.*, vol. 120, p. 109663, 2020.

[17] M.C. Marin, M. Cerutti, and S. Batista, "A Multi-Protocol IoT Platform for Enhanced Interoperability and Standardization in Smart Home," in *Proceedings of the 2024 IEEE 21st Consumer Communications & Networking Conference (CCNC)*, Las Vegas, NV, USA, 6–9 January 2024.

[18] J.J. Puthiyidam and S. Joseph, "IoT Smart Home: Protocols and Architectures," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 5, p. 2031, 2017.

[19] S. Lee, H. Kim, and J. Park, "Integrated Smart Home Security System Based on ESP32," *IEEE Access*, vol. 8, pp. 210-220, 2020.

[20] Espressif Systems, ESP-Matter Introduction, 2024. Available online: https://docs.espressif.com/projects/esp-matter/en/latest/esp32/introduction.html.

[21] S. Ahmed et al., "Smart Living: Ubiquitous Services Powered by Ambient Intelligence (AmI)," in Proceedings of the 2019 IEEE 16th International Conference on Smart Cities (HONET-ICT), Charlotte, NC, USA, 2019.

[22] T. Wang, L. Cassidy, W. He, and T. J. Pierson, "Challenges and Opportunities in Onboarding Smart-Home Devices," in Proceedings of the 25th International Workshop on Mobile Computing Systems and Applications, 2024.

[23] Espressif Systems, ESP-IDF Repository, 2024. Available online: https://github.com/espressif/esp-idf.

[24] Google Developers, Matter Primer: Commissioning, 2024. Available online: https://developers.home.google.com/matter/primer/commissioning.

[25] T. D. P. Mendes et al., "Smart Home Communication Technologies and Applications: Wireless Protocol Assessment for Home Area Network Resources," Energies, vol. 8, pp. 7279–7299, 2015.

[26] V. A. Orfanos et al., "A Comprehensive Review of IoT Networking Technologies for Smart Home Automation Applications," Journal of Sensor and Actuator Networks, vol. 12, no. 30, 2023.

[27] Fang, Y., Fu, C., and Du, X., "Virtual-Device-Based Policy Enforcement in Multi-Admin Smart Environments," in Proceedings of the 2023 IEEE 12th International Conference on Cloud Networking (CloudNet), Hoboken, NJ, USA, 2023.

[28] P. Li, T. Yigitcanlar, I. Erol, and A. Liu, "Motivations, Barriers, and Risks of Smart Home Adoption: From Systematic Literature Review to Conceptual Framework," Energy Research & Social Science, vol. 80, p. 102211, 2021.

[29] OpenThread, Thread Primer: Node Roles and Types, 2024. Available online: https://openthread.io/guides/thread-primer/node-roles-and-types.

Монголын эрх зүйн баримт бичгийг гүнзгий сургалт ашиглан шинжилж, шийдвэр гаргахад туслах системд нэвтрүүлэх талаарх судалгаа

Халтархүү Гармаабазар
Монгол-Японы хүний нөөцийн хөгжлийн төв
Монгол Улсын Их Сургууль
Монгол Улс, Улаанбаатар хот
garmaabazar@gmail.com

Билиг Батжа
Утга зохиолын сургууль
Рицүмэйкан Их Сургууль
Япон Улс, Киото хот
biligsaikhan@gmail.com

***Хураангуй*— Энэхүү судалгаанд Монголын хууль, эрх зүйн баримт бичигт дүн шинжилгээ хийхэд гүнзгий сургалтын технологийг ашиглах, улмаар хууль, эрх зүйн шийдвэр гаргахад давхардал, зөрчлийг шийдвэрлэх гүнзгий сургалтын аргыг санал болгосон. Хууль, эрх зүйн баримт бичгийн текстийн ангилал болон хууль эрх зүйн баримт бичгийн хоорондын хамаарлыг тодорхойлох судалгаанууд хийсэн. Хууль, эрх зүйн баримт бичгийг 0.91 Precision, 0.87 Recall, 0.89 F1 оноотойгоор 13 ангилалд нь ангилсан. Хууль эрх зүйн баримт бичгийн хоорондын хамаарлыг тодорхойлоход bert-base-multilingual-cased загвар нь хамгийн сайн үр дүнг үзүүлж, 0.7619 хувьтай зөв таньсан. "Зөрчил"-ийг танихдаа 0.7500 Precision, 0.7857 Recall, 0.7674 F1 оноотой гарсан. "Хамаарал"-ыг танихдаа 0.7750 Precision, 0.7381 Recall, 0.7561 F1 оноотой байсан.**

Түлхүүр үг— хууль, эрх зүйн баримт бичиг, гүнзгий сургалт, шийдвэр гаргахад туслах систем

I. Удиртгал

Шинээр батлагдсан хууль, эрх зүйн баримт бичиг нь одоо мөрдөж буйтай давхцах, зөрчилдөх, эрх зүйн шийдвэр нь шийдвэр гаргагчаас ихээхэн шалтгаалдаг зэргээр монголын хууль дээдлэх ёс буурсан үзүүлэлттэй байна. Хурдацтай, ихээр нэмэгдэж буй хууль, эрх зүйн баримт бичгүүдийн ийм төвөгтэй асуудлуудын хувьд давхардал, зөрчлийг олж тогтоох, асуудлыг тодорхой болгох үр дүнтэй арга хэрэгсэл шаардлагатай болж байна. Монгол Улсад тулгамдаж буй эдгээр асуудлыг шийдвэрлэхийн тулд их хэмжээний хууль, эрх зүйн баримт бичгүүдэд гүнзгий сургалтын технологийг ашигласан шийдвэр гаргахад туслах систем (DSS) нь үр дүнтэй хэрэгсэл болно.

Шинээр батлагдсан хууль, эрх зүйн баримт бичиг нь одоо мөрдөж буйтай давхцах, зөрчилдөх эсэхийг шалгах хэрэглээний тохиолдол, бодит жишээг энд дурдая. Жишээ нь: УИХ-ын сонгуулийн тухай хуульд саналын хуудас хүчингүйд тооцохтой холбоотой Сонгуулийн тухай хуулийн заалт нь Үндсэн хуультай зөрчилдөж байна. Мөн, зөрчлийн хуулиар 7 хоногийн баривчилгаа авсан хүмүүсийг шууд саатуулан хорьдог. Гомдол гаргая гэхэд шийдвэр 5 хоногт албажидаг тул иргэн Үндсэн хуульд заасан шударга шүүхээр шүүлгэх эрхээ эдэлж чадахгүй шууд саатуулан хоригдож байна. Мөн түүнчлэн, худалдааны талаар Иргэний хуулийн нэг хэсэг болгож зохицуулсан байдаг ч Худалдааны хуулинд байдаг үнийг хэрхэн тогтоох талаар Иргэний хуульд зохицуулаагүй тул иргэн Үндсэн хуульд заасан эрх зөрчигдөхөд эрх зүйн хамгаалалт байхгүй болох нөхцөл үүсэж байна. Одоо хэрэгжиж буй Олон нийтийн радио телевизийн тухай хууль нь Өрсөлдөөний тухай хууль болон Үндсэн хуультай зөрчилдөж байна.

II. Судалгааны зорилго

Энэхүү судалгаа нь хууль, эрх зүйн баримт бичигт дүн шинжилгээ хийхдээ гүнзгий сургалтын технологийг ашиглан хууль, эрх зүйн шийдвэр гаргахад давхардал, зөрчлийн нарийн төвөгтэй асуудлыг шийдвэрлэх аргыг санал болгох зорилготой юм. Судалгааны энэ зорилгод хүрэхийн тулд, нэгдүгээрт хууль эрх зүйн баримт бичгүүдийг зөв ангилах, хоёрдугаарт хууль эрх зүйн баримт бичгийн хоорондын зөрчлийг тодорхойлох гэсэн хоёр судалгааг хийсэн. Өөрөөр хэлбэл, эдгээр нь хууль, эрх зүйн баримт бичгийн текстийн ангилал, мөн түүнчлэн хууль эрх зүйн баримт бичгийн хоорондын хамаарлыг тодорхойлох судалгаанууд юм.

III. Судалгаанд ашигласан загвар

A. Хууль, эрх зүйн баримт бичгийн текстийн ангилал

Хууль, эрх зүйн баримт бичгийн текстийн ангиллын загвар боловсруулах нь шийдвэр гаргагчдын шинэ хууль, эрх зүйн бичиг баримтын төслийг зөв ангилахад дэмжлэг үзүүлэх зорилготой. Үүний тулд гүнзгий сургалтад суурилсан цахим хэл шинжлэлийн арга болох Bidirectional Encoder Representations from Transformers (BERT) [1] загварыг ашигласан аргыг туршин, монголын эрх зүйн баримт бичигт тусгайлан зориулсан LEGAL-BERT-Mongolian загварыг санал болгосон. Санал болгож боловсруулсан аргуудын хувьд uncased-LEGAL-BERT-Mongolian болон cased-LEGAL-BERT-Mongolian гэсэн хоёр загварыг бүтээж, энэ загварууд нь Монголын хууль эрх зүйн баримт бичгийн текстийг зөв ангилж буй эсэхийг шалгасан [2].

B. Хууль эрх зүйн баримт бичгийн хоорондын хамаарлыг тодорхойлох

Хууль, эрх зүйн баримт бичгийн зөрчлийг илрүүлэх судалгаагаар Монгол Улсын хуульчийн мэргэжлийн шалгалтын асуултуудын хоорондын хамаарлыг тодорхойлох загвар боловсруулсан. Энэ судалгаа нь

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

шийдвэр гаргагчдын хуулийн төсөл нь одоо мөрдөж буй хууль, эрх зүйн баримт бичигтэй зөрчилдөж байгаа эсэхийг илрүүлэх зорилготой юм. Энэхүү судалгаагаар зөрчлийг илрүүлэх зорилгоор нэн түрүүнд хоорондын хамаарлыг тодорхойлоход зориулж хууль, эрх зүйн баримт бичгийн тусгай датасет бэлдэж, уламаар хамгийн тохиромжтой аргыг тодорхойлсон. Судалгаандаа хууль, эрх зүйн баримт бичгийн онцлог шинж чанарыг харгалзан үзсэн 8 өөр Transformer загварыг fine-tuning хийж, эдгээр загварыг ашигласан 8 аргыг туршсан [3].

IV. Туршилтын дүн

A. Хууль, эрх зүйн баримт бичгийн текстийн ангилал

Санал болгож буй загварууд нь Монголын хууль эрх зүйн баримт бичгийн текстийг зөв ангилж буй эсэхийг шалгах туршилтанд Хүснэгт I-т жагсаасан бичиг баримтыг ашигласан. 2022 оны 2-р сарын 1-ний байдлаар Эрх зүйн мэдээллийн нэгдсэн системд (https://legalinfo.mn) нийтлэгдсэн Монголын хууль эрх зүйн баримт бичгээс 13 ангилалын нийт 11323 бичиг баримтыг машин сургалт болон туршилтанд ашигласан.

ХҮСНЭГТ I. Туршилтанд ашигласан Хууль, эрх зүйн баримт бичгийн ангилал болон тоо

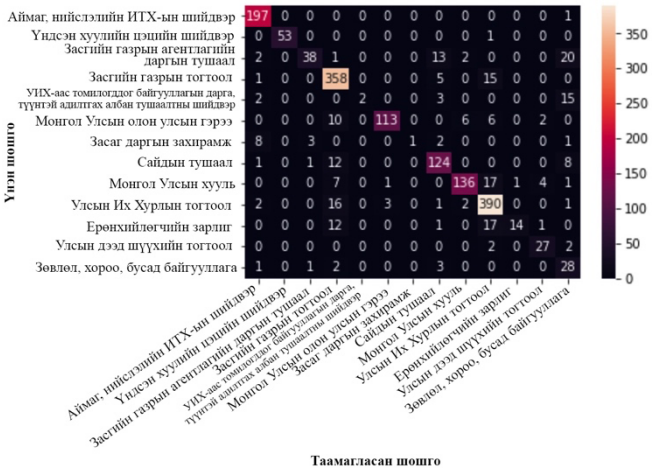
Хууль, эрх зүйн баримт бичгийн ангилал	Тоо
Засгийн газрын тогтоол	4716
Улсын Их Хурлын тогтоол	2066
Аймаг, нийслэлийн ИТХ-ын шийдвэр	996
Монгол Улсын хууль	778
Сайдын тушаал	748
Монгол Улсын олон улсын гэрээ	654
Засгийн газрын агентлагийн даргын тушаал	324
Үндсэн хуулийн цэцийн шийдвэр	295
Ерөнхийлөгчийн зарлиг	211
Улсын дээд шүүхийн тогтоол	174
Зөвлөл, хороо, бусад байгууллага	169
УИХ-аас томилогддог байгууллагын дарга, түүнтэй адилтгах албан тушаалтны шийдвэр	114
Аймаг, нийслэлийн Засаг даргын захирамж	78

Хүснэгт II-т үзүүлсэнчлэн, хууль, эрх зүйн баримт бичгийг 13 ангилалд ангилах туршилтаар uncased-LEGAL-BERT-Mongolian загвар 0.91 Precision, 0.87 Recall, 0.89 F1 оноотой хамгийн сайн үр дүнг үзүүлэв. Cased-LEGAL-BERT-Mongolian загварын Precision нь 0.87, Recall нь 0.83, F1 оноо нь 0.85 байсан.

ХҮСНЭГТ II. ХУУЛЬ ЭРХ ЗҮЙН БАРИМТ БИЧГИЙГ АНГИЛСАН ҮР ДҮНГИЙН ХАРЬЦУУЛАЛТ.

Загвар	Precision	Recall	F1 score
Cased-BERT-Mongolian-large	0.87	0.63	0.73
Uncased-BERT-Mongolian-large	0.79	0.83	0.81
Cased-LEGAL-BERT-Mongolian	0.87	0.83	0.81
Uncased-LEGAL-BERT-Mongolian	**0.91**	**0.87**	**0.89**

Хууль, эрх зүйн баримт бичгүүдийн текст ангиллын туршилтын үр дүнг харьцуулан үзэхэд uncased-LEGAL-BERT-Mongolian болон cased-LEGAL-BERT-Mongolian загварын буруу ангилсан хувь нь тус тус 16% болон 26% байсан. Буруу ангилсан текстүүдийн 30 орчим хувь нь нэг хүний нэрийг агуулж байб. Жишээлбэл Н. Багабанди гэдэг нэр, баримт бичиг баталсан сайд, УИХ-ын дарга, ерөнхий сайд, ерөнхийлөгч гэх мэт байдлаар ялгаатай олон бичиг баримтанд бичигдсэн байна. Цаашилбал, буруу ангилсан текстүүдийн 28.5% нь санал болгож буй загварт ашиглагдсан зөвшөөрөгдсөн текстийн уртаас хэтэрсэн байсан.



Зураг. 1. uncased-LEGAL-BERT-Mongolian загварын төөрөгдлийн матриц.

Зураг.1-т uncased-LEGAL-BERT-Mongolian загварын төөрөгдлийн матрицыг үзүүллээ. Хамгийн их төөрөгдсөнийг харвал “Засгийн газрын агентлагийн даргын тушаал”-ыг “Зөвлөл, хороо, бусад байгууллагын шийдвэр” гэж, “Улсын Их Хурлын тогтоол”-ыг “Засгийн газрын тогтоол” гэж тус тус ангилсан байна. Мөн “Ерөнхийлөгчийн зарлиг”-г “Засгийн газрын тогтоол” гэж ангилах тохиолдол байна. Дээрх байгууллагууд адил төстөй шийдвэр гаргаж байгаа эсвэл дээр дурдсанчилан аливаа нэгэн албан хаагч нь сайд, УИХ-ын дарга, ерөнхий сайд, ерөнхийлөгч гэх мэт ялгаатай байгууллагуудын шийдвэр гаргагч байх, мөн УИХ-ын гишүүн нь Засгын газрын гишүүн байх зэрэг нөлөөллөөс болж төөрөгдөл үүссэн байх магадлалтай.

Санал болгож буй арга нь хууль, эрх зүйн ангиллын загвар бөгөөд энэхүү туршилт нь монгол хэлээрх хууль, эрх зүйн баримт бичгийн текстийн ангиллыг шалгах нэгэн туршилт боловч BERT загварыг ашигласан бусад текст ангилах туршилтуудтай харьцуулахад дундаж үзүүлэлтээс илүү үр дүн үзүүлсэн.

B. Хууль эрх зүйн баримт бичгийн хоорондын хамаарлыг тодорхойлох

Энэхүү судалгаанд 8 Transformer загварын fine-tuning хийсэн загваруудыг туршсан. Хүснэгт III-т үзүүлсэнээр уг 8 загварыг ямарваа нэгэн өөрчлөлтгүйгээр (эх загвар), мөн fine-tuning хийсэн байдлаар нийт 16 туршилт хийсэн. Туршилтын үр дүнгээс харвал fine-tuning хийсэн загварууд нь эх загвараас илүү үр дүн үзүүлсэн.

Нийт туршсан 16 загваруудын дотроос fine-tuning хийсэн bert-base-multilingual-cased [4] загвар нь хамгийн сайн үр дүнг үзүүлж, зөв танилт нь 0.7619 байсан. Хүснэгт IV-т fine-tuning хийсэн bert-base-multilingual-cased загварын үр дүнг үзүүлсэн. Энэхүү загвар нь “зөрчил”-ийг танихдаа 0.7500 Precision, 0.7857 Recall, 0.7674 F1 оноотой байсан. “Хамаарал”-ыг танихдаа 0.7750 Precision, 0.7381 Recall, 0.7561 F1 оноотой байсан.

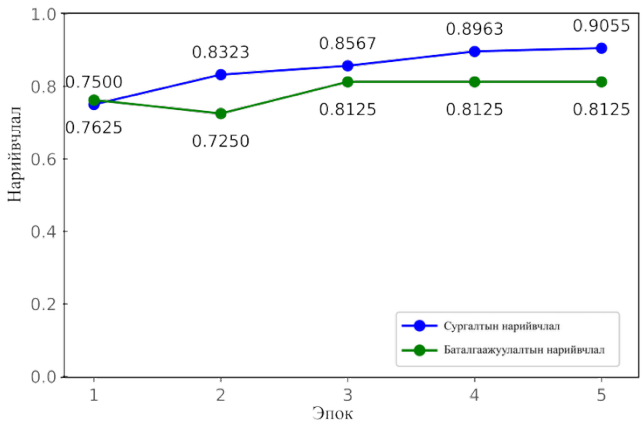
ХҮСНЭГТ III. ХУУЛЬ ЭРХ ЗҮЙН БАРИМТ БИЧГИЙН ХООРОНДЫН ХАМААРЛЫГ ТОДОРХОЙЛОХ ТУРШИЛТЫН ДҮН

Загвар	Нарийвчлал		F1 Score			
			Макро дундаж		*Жинэлсэн дундаж*	
	Эх загвар	*fine-tuned*	*Эх загвар*	*fine-tuned*	*Эх загвар*	*fine-tuned*
mongolian-roberta-base	0.6071	0.7024	0.5860	0.7013	0.5860	0.7013
mongolian-roberta-large	0.5357	0.5833	0.5351	0.5833	0.5351	0.5833
albert-mongolian	0.6071	0.6310	0.6026	0.6267	0.6026	0.6267
bert-base-mongolian-cased	0.6905	0.7381	0.6876	0.7368	0.6876	0.7368
bert-large-mongolian-uncased	0.7381	0.7500	0.7375	0.7500	0.7375	0.7500
bert-base-multilingual-cased	0.6905	**0.7619**	0.6860	0.7618	0.6860	0.7618
bert-large-mongolian-cased	0.6786	0.6905	0.6748	0.6816	0.6748	0.6816
bert-base-mongolian-uncased	0.7262	0.7024	0.7262	0.6972	0.7262	0.6972

ХҮСНЭГТ IV. FINE-TUNING ХИЙСЭН BERT-BASE-MULTILINGUAL-CASED ЗАГВАРЫН ҮР ДҮН.

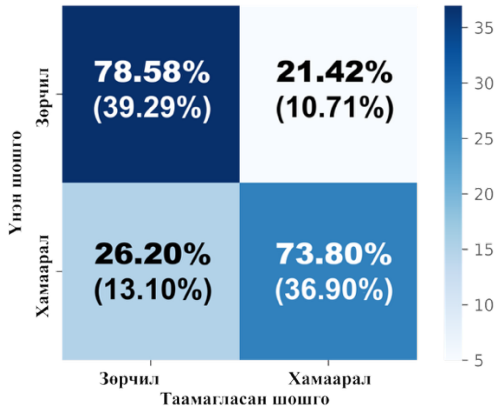
Шошго	Precision	Recall	F1 score
Зөрчил	0.7500	0.7857	0.7674
Хамаарал	0.7750	0.7381	0.7561
Нарийвчлал			0.7619
Дундаж нарийвчлал	0.7625	0.7619	0.7618
Жинэлсэн нарийвчлал	0.7625	0.7619	0.7618

Зураг.2-т fine-tuning хийсэн bert-base-multilingual-cased загварын сургалт, баталгаажуулалтын нарийвчлалыг, Зураг.3-т төөрөгдлийн матрицыг үзүүллээ.



Зураг. 2. fine-tuning хийсэн bert-base-multilingual-cased загварын сургалт, баталгаажуулалтын нарийвчлал.

Нөгөөтэйгүүр, Хүснэгт V-т үзүүлсэнээр fine-tuning хийсэн mongolian-roberta-large [5] загвар нь хамгийн муу үр дүн үзүүлсэн. “Зөрчил” болон “хамаарал”-ын аль алиныг нь нарийвчлал багатай таньсан. “Зөрчил”-ийг танихад, 0.5854 Precision, 0.5714 Recall, 0.5783 F1 оноо, “хамаарал”-ыг танихад 0.5814 Precision, 0.5952 Recall, 0.5882 F1 оноотой байсан.



Зураг. 3. fine-tuning хийсэн bert-base-multilingual-cased загварын төөрөгдлийн матриц.

ХҮСНЭГТ V. FINE-TUNING ХИЙСЭН MONGOLIAN-ROBERTA-LARGE ЗАГВАРЫН ҮР ДҮН.

Шошго	Precision	Recall	F1 score
Зөрчил	0.5854	0.5714	0.5783
Хамаарал	0.5814	0.5952	0.5882
Нарийвчлал			0.5833
Дундаж нарийвчлал	0.5834	0.5833	0.5833
Жинэлсэн нарийвчлал	0.5834	0.5833	0.5833

Энэ туршилт нь судалгааны бас нэг зорилт болох Монгол Улсын хууль, эрх зүйн баримт бичгүүдийн зөрчлийг судлахын тулд тодорхой үр дүнг үзүүлсэн судалгаа болсон.

V. ДҮГНЭЛТ

Энэхүү судалгаагаар санал болгож буй аргууд нь судалгааны зорилго болох хууль, эрх зүйн баримт бичгүүдийн давхардал, зөрчлийг илрүүлэх арга хэрэгсэл болгон, нэгдүгээрт монгол хэлээрх хууль, эрх зүйн баримт бичгийн текстийг ангилах, хоёрдугаарт хуульчийн мэргэжлийн шалгалтын асуултуудын хоорондын хамаарлыг тодорхойлоход үр дүнтэй болохыг харуулсан. Санал болгож буй аргын шинэлэг тал нь цахим хэл шинжлэлийн технологи болон гүнзгий сургалтын аргыг ашиглан монгол хэлээрх хууль, эрх зүйн баримт бичгүүдэд тусгайлан зориулсан текстийн ангилал, хуульчийн мэргэжлийн шалгалтын асуултуудын хоорондын хамаарлыг тодорхойлсон явдал юм. Түүнчлэн, Монгол Улсын хууль, эрх зүйн баримт бичгүүдийн текстийн ангилал, хуульчийн мэргэжлийн шалгалтын асуултуудын хоорондын хамаарлыг тодорхойлох зэрэгт хангалттай үр дүн үзүүлсэн нь санал болгож буй аргын үр нөлөөг харуулж байна. Гүнзгий сургалтын технологийг ашиглах нь нарийн төвөгтэй хууль эрх зүйн баримт бичгийн бүтэц, хамаарлыг үнэн зөв гаргаж, шийдвэр гаргагчдын мэргэжлийн дүгнэлтийг дэмжих боломжтой болно. Шийдвэр гаргахад туслах систем нь шийдвэр гаргагчдыг дэмжиж, шуурхай, оновчтой шийдвэр гаргах мэдээллээр хангах, хуулийн хэрэгжилт, тайлбарын чанарыг сайжруулж, хууль дээдлэх ёс бүхий нийгмийг бэхжүүлнэ гэж үзэж байна.

Санал болгосон загвар нь хууль, эрх зүйн баримт бичгийг 0.91 Precision, 0.87 Recall, 0.89 F1 оноотойгоор 13 ангилалд нь ангилсан. Хууль эрх зүйн баримт бичгийн

хоорондын хамаарлыг тодорхойлоход bert-base-multilingual-cased загвар нь хамгийн сайн үр дүнг үзүүлж, 0.7619 хувьтай зөв таньсан. “Зөрчил”-ийг танихдаа 0.7500 Precision, 0.7857 Recall, 0.7674 F1 оноотой гарсан. “Хамаарал”-ыг танихдаа 0.7750 Precision, 0.7381 Recall, 0.7561 F1 оноотой байсан.

Энэхүү судалгаанд санал болгосон загваруудыг аливаа байгууллагын шийдвэр гаргалтад ашиглах зэргээр хэрэглээнд нэвтрүүлэх боломжтой бөгөөд туршилтын хувилбар (https://www.dl.is.ritsumei.ac.jp/legal_analysis/) байдлаар систем хөгжүүлэлт хийгдэж байгаа билээ.

Үүсгэгч хиймэл оюун (Generative Artificial Intelligence), болон Мультимодал хиймэл оюуны дэвшилтэд загваруудыг Монголын хууль, эрх зүйн баримт бичигт дүн шинжилгээ хийхэд нэвтрүүлэх нь ирээдүйн судалгааны чиглэл болно гэж үзэж байна.

Ном зүй

[1] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” Proc. 2019 Conf. North, Minneapolis, Minnesota: Association for Computational Linguistics, 2019, pp. 4171–4186. doi: 10.18653/v1/N19-1423.

[2] G. Khaltarkhuu, B. Batjargal, and A. Maeda, “Text Classification of Modern Mongolian Legal Documents using BERT Models,” Int. J. Asian Language Processing, 2023, 2350020 (22 pages), December 2023, doi: 10.1142/S2717554523500200

[3] G. Khaltarkhuu, B. Batjargal, and A. Maeda, “Recognizing Textual Inference in Mongolian Bar Exam Questions,” Applied Sciences, 2024, Vol. 14, No. 3, 1073, January 2024, doi: 10.3390/app14031073

[4] “google-research/bert,” Google Research, Feb. 08, 2024. https://github.com/google-research/bert

[5] “mongolian-roberta-large,” Hugging Face. https://huggingface.co/bayartsogt/mongolian-roberta-large

Нарны эрчим хүчийг хэмжиж, дамжуулах эмбэддэд төхөөрөмжийн хөгжүүлэлт

Аюурзанын Одгэрэл
Электроникийн салбар
ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль
Улаанбаатар, Монгол Улс
odgerel55@must.edu.mn

Хураангуй: Судалгааны ажлаар ~1кВт хүртэлх чадалтай нарны панелын гүйдэл, хүчдэлийн утгыг боловсруулан мэдээллийн төвд урьдчилан тодорхойлсон протоколын дагуу GSM сүлжээ ашиглан илгээх эмбэддэд төхөөрөмжийг хөгжүүлж туршсан. Мэдээллийн төв дээр эмбэддэд төхөөрөмжөөс ирсэн гүйдэл, хүчдэл, бодсон чадал, төхөөрөмжийн дугаар, огноо болон төхөөрөмжийн байршлын мэдээллийг харуулна. Уг судалгаа нь хөдөө орон нутагт бага чадалтай нарны эрчим хүчний үүсгүүр ашигладаг малчин өрхүүд ногоон эрчим хүч үйлдвэрлэж хүлэмжийн хийг бууруулахад хувь нэмэр оруулж байна гэж үзэн үйлдвэрлэсэн ногоон эрчим хүчээ мэдээллийн төвд илгээж карбон кредит болгож бирж дээр арилжаанд оруулж орлого олох боломжийг бүрдүүлэх ажлын эхлэл юм. Тоолуурын төхөөрөмжийг төвийн эрчим хүчний системд холбож хуурамч өгөгдөл илгээхээс сэргийлж байршлын координатыг мэдээллийн төв дээр хянана. Төхөөрөмжийг хөдөө орон, нутаг болон төв суурин газрын ~1кВт хүртэлх чадалтай нарын систем ашиглаж байгаа айл өрх, албан газруудад ашиглах боломжтой.

Түлхүүр үг: Эмбэддэд төхөөрөмж, GPS/GSM модул, ESP32, карбон кредит, нарны панел

I. УДИРТГАЛ

Монгол улсын эрчим хүчний 1 жилийн хэрэглээ 9 орчим тэрбум кВт.ц бөгөөд энэ тоо жил бүр 6-8%-р нэмэгдэж байна. Үүнээс 9%-г нь сэргээгдэх эрчим хүч буюу нар, салхи, усан цахилгаан станцаас гарган авч байна [1]. Монгол улсын ерөнхийлөгч У.Хүрэлсүх манай улс НҮБ-ын уур амьсгалын өөрчлөлт, хүлэмжийн хийг бууруулах конвенцид нэгдсэнийг зарласан бөгөөд цаашид сэргээгдэх эрчим хүчний үйлдвэрлэл болон технологийг дэмжиж ажиллахыг Монгол улсын хөгжлийн алсын хараа 2050 бичиг баримтад тусгасан билээ [2]. Дэлхийн өндөр хөгжилтэй болон хөгжиж буй улсууд нь хүлэмжийн хий ялгаруулж байгаль орчин бохирдуулж байгаагаа КАРБОН КРЕДИТ ашиглан төлдөг шинэ ойлголт, төлбөрийн систем бий болж байна. Өөрөөр хэлбэл ногоон буюу сэргээгдэх эрчим хүч үйлдвэрлэж байгаа улсууд үйлдвэрлэсэн эрчим хүчээ кредит болгон борлуулдаг арилжааны платформууд үүсч байна. Үүссэн кредитийг байгаль орчин бохирдуулж хүлэмжийн хий ялгаруулж байгаа улс орон, компани, үйлдвэрүүд худалдан авдаг тогтолцоо бий болж байна. Монгол улсад URECA [3] компани энэ төрлийн арилжааны платформыг хөгжүүлэн эхний туршилтуудаа хийж эхэлсэн.

Хөдөө нарны панелтай айл нь үйлдвэрлэж байгаа эрчим хүчний хэмжээгээ мэдээллийн төвд илгээж КРЕДИТ болгон хувиргаж орлого олох боломж нээгдэх юм. Манай улсад ашиглаж байгаа өрхийн хэмжээний нарны (1кВт чадал хүртэлх) системүүд нь тоолуур байхгүй. 1кВт-с дээш системүүдэд тоолуур байгаа боловч мэдээллийг нь авахын тулд үйлдвэрлэгчээс зөвшөөрөл авч нэмэлт төлбөр төлөх болдог. Айл өрх, албан байгууллагуудад суурилуулсан сэргээгдэх эрчим хүчнийн үүсгүүрийн (нар, салхи, ус) үйлдвэрлэж байгаа эрчим хүчийг хэмжих төхөөрөмжийг хөгжүүлж мэдээллийн төвд илгээж тооцоолоох ажлын эхлүүлэх нь уг судалгааны ажлын зорилго юм.

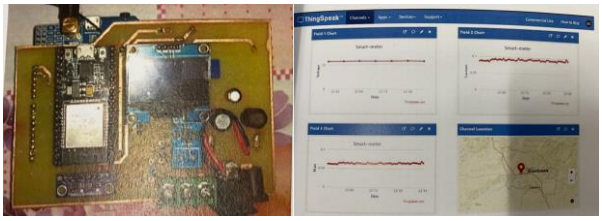
Энэ судалгаанд ашигласан SIM800 GSM/GPRS модем нь цуваа холболтоор микроконтролёртой холбогдон AT командыг ашиглан мэдээллийг дамжуулдаг. Модем нь командыг хүлээн авсныхаа дараа үр дүнг буцаан илгээдэг. Технологи хөгжихийн хэрээр шинэчлэгдсэн хувилбарууд гарч байна. Монголд улсад түгээмэл байдаг модул нь: GSM shield 900, GSM shield 900A, GSM shield 800, A7, A6 юм. Эдгээр нь хоорондоо өгөгдөл дамжуулах хурдаараа ялгаатай.

NEO-6M GPS нь 1.5ГГц давтамж дээр ажилладаг бага чадлын, хурд, хугацаа болон байршлын нарийвчлал сайтай модул юм. GPS (Global positioning system) гэдэг нь дэлхийгээс 20,000 км зайд байрлах орбит замаар явах 31 хиймэл дагуулуудаас бүрдэнэ. GPS үйлчилгээ нь анх цэргийн зориулалтаар гарсан бөгөөд орчин үед иргэний болон арилжааны чиглэлээр илүү өргөн ашиглагдаж байна. Хиймэл дагуулуудаас ирэх дохиолол тасралтгүй бөгөөд хэрэглэгч, хүлээн авагчийн тоонд хязгаар байхгүйгээрээ онцлогтой. Хиймэл дагуулууд газрын станцаас ирэх дохиоллыг хэмжин өөрийн байрлалыг нарийвчлан тооцоолдог.

GPS хүлээн авагчийн байрлал тодорхойлоход 3-аас багагүй хиймэл дагуул ашиглана. Нэг доор хүлээн авсан мэдээлэл анализ хийж өөрийн байгаа байрлалыг (уртраг, өргөрөг) гэсэн хоёр хэмжээсээр тодорхойлно. Харин дөрөв ба түүнээс олон хиймэл дагуулаас хүлээн авсан дохиоллын тусламжтайгаар байрлалыг (уртраг, өргөрөг, далайн түвшинтэй харьцуулан өндөрлөг) гурван хэмжээсээр тодорхойлдог. GPS-г дэлхийн хаана ч ашиглаж болох боловч агуй, хонгил, машины граж, усан дор зэрэг газар GPS-н сигналыг дамжуулах боломжгүй байдаг [4].

II. СУДЛАГДСАН БАЙДАЛ

HCS-LTS-25A гүйдлийн мэдрэгч, ESP32 микроконтроллёр болон SIM808 GPS/GSM shield ашиглан хөгжүүлсэн сэргээгдэх эрчим хүчний хяналтын системийг [5]-д хөгжүүлсэн. Мэдээллийг ThingSpeak бэлэн сервер ашиглан хүлээн авч харуулсан. Уг судалгааны ажилд ашигласан гүйдлийн мэдрэгч нь contact-less буюу дамжуулагч кабелийг мэдрэгч дундуур сүвлэж явуулдаг нь аюулгүй байдал талаасаа сайн боловч мэдрэгчийн нарийвчлал нь харьцангуй муу байдаг сул талтай.



Зураг. 1. Төсөөтэй ажил-1: төхөөрөмж ба вэбийн харагдах байдал [5]

Beijing Lewei IOT Technologies компани Wi-Fi модул бүхий цахилгаан тоолуурын төхөөрөмжийг хөгжүүлж зах зээлд гаргасан [6]. Компани нь өөрсдийн хөгжүүлсэн вэб сервертэй бөгөөд тоолуурууд нь утасгүй сүлжээнд холбогдон мэдээллээ Интернетийн сүлжээ ашиглан серверт илгээн графикаар болон хүснэгтээр харуулж статистик тооцооллыг хийдэг.

Зураг. 2. Төсөөтэй ажил-2: тоолуур болон вэб серверийн харагдах байдал [6]

PIC18F4550 микроконтроллёр ашиглан нар болон салхины үүсгүүрээс ирж байгаа энергийг том хэмжээний батерейнд хадгалах судалгааны ажлыг [7]-д хийсэн. Нарны энергийн AC буюу хувьсах гүйдэл рүү хувиргаж (grid inverter) төвийн эрчим хүчний шугамд нийлүүлэх системийг хөгжүүлж туршсан судалгааны ажлыг [8]-д хийсэн. Өгөгдлөө GSM-ийн модемоор дамжуулж, LCD дэлгэц дээр системийн бүх үзүүлэлтийг гаргаж хянасан. ATMega-2560 микроконтроллёр ашигласан IoT-д суурилсан SCADA системээр сэргээгдэх эрчим хүчийг хянах судалгааны ажлыг [9]-д хийсэн. Уг ажил нь Матлаб програм дээр симуляци хийж ThingSpeak бэлэн серверт өгөгдөлөө Wi-Fi модул ашиглан дамжуулж харуулсан.

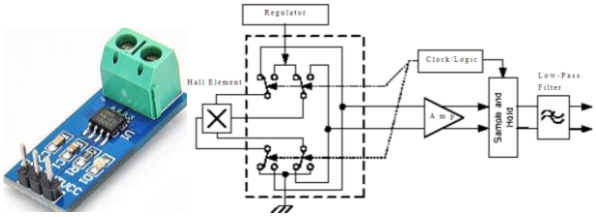
III. СИСТЕМИЙН ДИЗАЙН БА ШИЙДЭЛ

Системийн ерөнхий бүтцийн схемийг дараах зурагт харуулав. 12Вт-н нарны панел, ACS712 гүйдлийн мэдрэгч, хүчдэл хуваагч, OLED дэлгэц, SIM800 GSM модем, бага чадлын NEO-6M GPS модулуудыг төхөөрөмжийг хөгжүүлэхэд ашигласан. GSM модем нь AT командууд ашиглан мэдээллээ дамжуулдаг. ESP32 хоёр цөмтэй микроконтроллёрыг удирдлагаар ашигласан.

Зураг. 3. Системийн бүтцийн схем

Энэ контроллёр нь RISC-V архитектуртай, өөртөө Bluetooth болон Wi-Fi модул агуулсан бөгөөд IoT хэрэглээнүүдэд илүү тохиромжтой. Тэжээлээр нь LiPo төрлийн 7.4В, 450mAh-н багтаамжтай батерейг сонгон ашигласан.

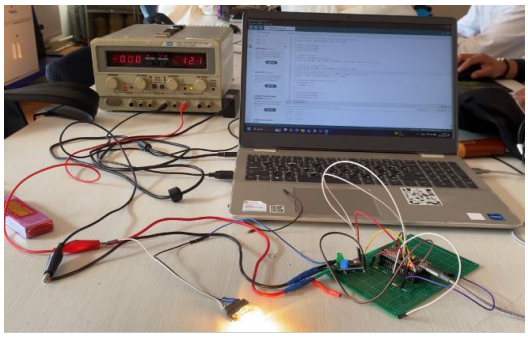
Гүйдлийн мэдрэгчээр ACS712 төрлийн мэдрэгчийг сонгосон. Энэ мэдрэгч нь аналог гаралттай бөгөөд гүйдэлтэй шугам дээр цуваа холбож ашигладаг. Гүйдэл урсахад үүссэн соронзон оронгын нөлөөллөөр хүчдэлийн өөрчлөлт үүсгэх зарчмаар ажилладаг.



Зураг. 4. ACS712 гүйдэл мэдрэгчийн гадаад болон дотоод бүтэц

IV. ТУРШИЛТ, ҮР ДҮН

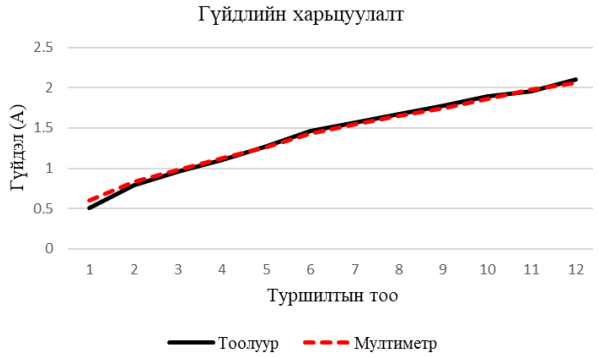
Сонгож авсан элемент, модулуудыг угсарч туршилт хийсэн загварыг дараах зурагт харуулав. Эхний туршилтаар нарны панелыг холболгүй гүйдэл мэдрэгчийн оролтонд шууд тэжээлийн блокоос гүйдэл өгч өөрчлөн эмбэддэд төхөөрөмж гүйдлийн мэдрэгчийн тооцоолж байгаа утга болон мультиметрийн заалттай харьцуулсан туршилт хийсэн.



Зураг. 5. Эмбэддэд төхөрөмжийг тэжээлийн блоктой холбон туршилт хийж байгаа байдал

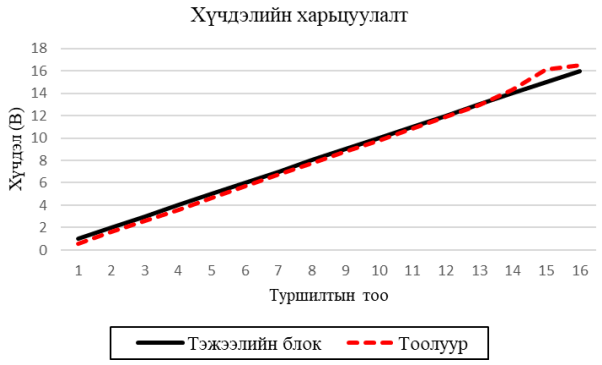
6-р зургаас харахад улаан өнгөтэй тасархай шугам нь мультиметрийн хэмжсэн утга, хар өнгөтэй шугам нь эмбэддэд тоолуурын боловсруулсан гүйдлийн утга юм. Мультиметрийн алдааг 0 гэж үзвэл гүйдлийн

мэдрэгч ACS712 нь 1.7% орчмын алдаатай гэж харагдаж байна.



Зураг. 6. Гүйдэл мэдрэгчийн утгыг мультиметрийн утгатай харьцуулсан график

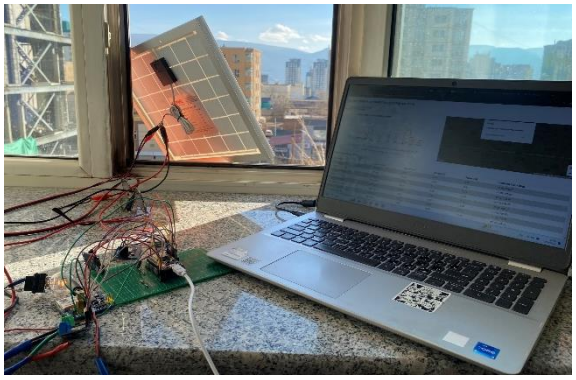
Зураг 7-д хүчдэлийн утгыг харьцуулж харуулав. Хар өнгөтэй шугам нь тэжээлийн блокоос өгч буй хүчдэлийн утга улаан өнгөтэй тасархай шугам нь эмбэддэд тоолуурын боловсруулсан хүчдэлийн утга юм. Тэжээлийн блокын алдааг 0 гэж үзвэл хүчдэл хуваагуурынх 9.7% орчмын алдаатай байгаа нь харьцангуй их байна.



Зураг. 7. Хүчдэл хуваагуурын мэдрэгчийн утгыг тэжээлийн блокоос өгсөн утгатай харьцуулсан график

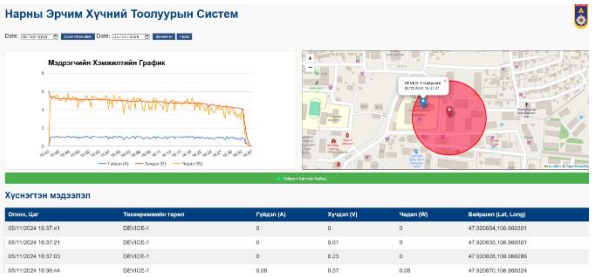
Дараах зурагт 12Вт-ын нарны панелыг эмбэддэд төхөөрөмжид бодитоор холбож туршилт хийж байгааг харуулав. Карбон кредит тооцох үед төхөөрөмжийг төвийн эрчим хүчний системд холбож хуурамч өгөгдөл илгээхээс сэргийлэхийн тулд GPS ашигласан. GPS нь төхөөрөмжийг анхны байршлаасаа хөдөлсөн эсэхийг хянах бөгөөд хөдөлсөн тохиолдолд мэдээллийн төв дээр дохиолол өгнө.

Нарны панелыг урд зүгт 45 градусын налуутайгаар байрлуулна. Эмбэддэд тоолуурын төхөөрөмжөөс GSM-н модем ашиглан нарны панелаас ирж буй гүйдэл хүчдэл, бодсон чадал, төхөрөөмжийн дугаар, огноо болон байршлын уртраг, өргөрөгийн мэдээллийг мэдээллийн серверт урьдчилан тодорхойлсон протоколын дагуу 5 минутын давтамжтай илгээж туршилт хийсэн. Ирсэн уртраг өргөрөгийн мэдээллийг Google map-д холбож сервер дээр харуулна.



Зураг. 8. Нарны панелыг холбож бодит туршилт хийж байгаа байдал

Дараах зурагт мэдээллийн серверийн харагдах байдлыг харуулсан. Гүйдэл, хүчдэл, бодсон чадлыг графикаар болон хүснэгтээр харуулсан. Мөн Google map дээр төхөөрөмжийн одоогийн байршлыг харуулсан. Мөн өгөгдлийг хугацааны интервалаар эксел файлаар авах функцтэй. Ашиглаж буй NEO-6M GPS модулын нарийвчлалаас хамаарч байршлын алдаа 1-5м дотор байсныг туршилтаар тогтоосон.



Зураг. 9. Мэдээллийн серверийн харагдах байдал

V. ДҮГНЭЛТ

Судалгааны ажлаар ~1кВт хүртэлх чадалтай нарны панелын гүйдэл, хүчдэлийн утгыг боловсруулж дамжуулах эмбэддэд төхөөрөмжийг хөгжүүлж туршсан. Эмбэддэд төхөөрөмж нь нарны панелаас ирсэн гүйдэл, хүчдэлийн утгаар чадлыг бодон тодорхойлсон протоколын дагуу мэдээллийн төвд GSM сүлжээ ашиглан 5 минутын давтамжтай дамжуулсан. Мэдээллийн төв дээр эмбэддэд төхөөрөмжөөс ирсэн гүйдэл, хүчдэл, бодсон чадал, төхөөрөмжийн дугаар, огноо болон төхөөрөмжийн байршлын мэдээллийг харуулна. Төхөөрөмжийг төвийн эрчим хүчний шугамд холбож хуурамч өгөгдөл илгээхээс сэргийлж анхны байршлыг тогтоож, хөдөлгөсөн тохиолдолд Google Map-р хянаж дохиолол өгөхөөр програмчилсан. Ашигласан гүйдлийн мэдрэгч ACS712-н алдаа нь 1.7%, харин хүчдэл хуваагуурын алдаа 9.7% байсан нь ашигласан эсэргүүцлүүдийн нарийвчлал болон хэмжээнээс болсон гэж үзэн цаашид дахин туршилт хийх хэрэгтэй гэж үзсэн. NEO-6M GPS модулын нарийвчлалаас хамааран байршлын алдаа нь 1-5м дотор байсан.

Судалгааны ажлаар хөгжүүлсэн эмбэддэд тоолуурын төхөөрөмжийг хөдөө Монгол гэрт болон, аймаг сумын төвд ~1кВт хүртэлх чадалтай нарны эрчим хүч ашиглаж байгаа айл, өрхүүд шууд ашиглах

боломжтой болно. Монгол улсад 2022 оны тооллогоор 305,000 орчим малчин өрх бүртгэгдсэн бөгөөд үүний 1/3 орчим нь буюу 100,000 орчим нь өрхийн нарны бага чадлын эрчим хүч ашиглан ТВ, хөргөгч, гэрэлтүүлэгээ шийдэж байна. Эдгээр айлууд нь өөрсдийн сэргээгдэх эрчим хүчний төхөөрөмжөөр үйлдвэрлэсэн цахилгаан энергийн хэмжээгээ мэдээллийн төвд илгээн баталгаажуулж КАРБОН КРЕДИТ болгон арилжаалж орлого олох боломж үүсэх юм. 1 КАРБОН КРЕДИТ нь 1 тонн CO2 бөгөөд үнэ нь дэлхийн зах зээл дээр 60-80-н хооронд хэлбэлзэж байна. 1КВт чадалтай нарны станц нь жилдээ 0.73 тонн орчим CO2 ялгаруулалтыг нөхдөг гэсэн судалгаа байна [10].

Ном зүй

[1] Эрчим хүчний зохицуулах хороо https://erc.gov.mn/

[2] “Алсын хараа 2050” Монгол улсын урт хугацааны хөгжлийн бодлого, МУ-ын ЗГ, УБ хот, 2019 он

[3] Юрека ХХК, ureca.com

[4] Н.Өсөхбаяр, Б.Баасандорж, Э.Эрхбилгүүн “GSM сүлжээ дээр суурилсан осол, гэмтлээс урьдчилан сэргийлэх систем” Бакалаварын төгсөлтийн ажил, “Хүрээ” Мэдээлэл холбооны технологийн дээд сургууль, 2018 он

[5] Л.Саруулбуян бакалаврын дипломын ажил “Сэргээгдэх эрчим хүчний тоолуурын хөгжүүлэлт” 2022 он

[6] Beijing Lewei IOT Technologies компани https://www.iammeter.com/

[7] Ersan Kabalci, Alper Gorgun, Yasin Kabalci "Design and implementation of a renewable energy monitoring system" 4th International Conference on Power Engineering, Energy and Electrical Drives, 21 Oct 2013, DOI: 10.1109/PowerEng.2013.6635759

[8] Ahmed Shah Rasheed Tajwar, Abir Ahmed, Md. Rifat Hazari, Mohammad Abdul Mannan "Solar Photovoltaic-Based Smart Metering System" 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), 05-07 Jan 2021, DOI: 10.1109/ICREST51555.2021.9331098

[9] Md Ohirul Qays, Musse M. Ahmed, M. A. Parvez Mahmud "Monitoring of renewable energy systems by IoT - aided SCADA system", Journal of Energy Science Engineering, Feb 2022

[10] https://www.quora.com/

МОНГОЛ МОДОН БАРЫН БҮТЭЭЛИЙГ ГАР БИЧМЭЛИЙН БҮТЭЭЛЭЭС ЯЛГАН ТАНИХ

Пэрэнлэйлхүндэв, Гантуяа
Мэдээлэл компьютерын ухааны тэнхим
Монгол улсын их сургууль
Монгол, Улаанбаатар
https://orcid.org/0009-0000-7332-141X

Батсуурь, Сувдаа*
Мэдээлэл компьютерын ухааны тэнхим
Монгол улсын их сургууль
Монгол, Улаанбаатар
suvdaa@num.edu.mn

Ганболд, Ууганбаяр
Мэдээлэл компьютерын ухааны тэнхим
Монгол улсын их сургууль
Монгол, Улаанбаатар
uuganbayar.g@num.edu.mn

Хураангуй— Монголчуудын аман зохиол, түүхээр дамжин өвлөгдөж ирсэн олон арван судар номыг зуун зууныг өртөөлөн модон барын бүтээл болгон хэвлэж ирсэн түүх бидэнд бий. Монгол соёлын биет өв болсон энэ бүтээлүүдээсээ бид түүх соёл, өв уламжлалаа мэдэж, түгээн дэлгэрүүлсээр байна. Түүх археологийн олдвор, хуучин дурсгалуудыг орчин үеийн судалгааны арга ашиглан шинэ түвшинд судлан шинжилж буй өнөө үед модон барын бүтээлүүдийг ч шинэ аргаар судалж үзэх нь судлаачдын хийх ёстой ажил юм. Аман зохиолын салбарт ч орчин үеийн судалгааны аргууд нэвтрээд эхний үр дүнгүүд гарсан[1]. Энэхүү судалгааны ажлаар бид монгол модон барын бүтээлийг гар бичмэлийн бүтээлээс ялган таних зорилгоор төрөл, ангилал, онцлогийг судлан улмаар машин сургалтын уламжлалт 11 арга ашиглан үр дүнг харьцуулан туршлаа. Үүгээрээ уламжлалт соёлын өв, хэл бичгийн судалгаа шинжилгээний ажилд машин сургалтын арга ашиглан судалж үзэх нь шинжлэх ухааны салбар дундын судалгааг хөгжүүлэх үр дүнтэй төдийгүй энэ талын судалгааг шинэ шатанд хөгжүүлэх боломжтойг харуулахыг зорилоо.

Түлхүүр үг— модон бар, гар бичмэл, компьютерын ухаан, хиймэл оюуны арга, ангилах

I. Удиртгал

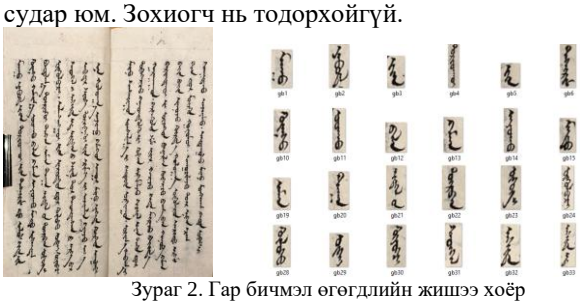
Монголчууд X-XIX зууны хооронд мод, зэс, төмөр, чулуун бараар ном, судар хэвлэж ирсэн түүхтэй. XIII-XVI зуунд монголын ном бүтээгчид эх орны түүхийн сэдэвтэй олон ном бичсэний дотор “Чингис хааны нууц түүх”, “Монгол улсын арван буянт номын цагаан түүх нэрт судар оршив”, “Сав шим хийгээд хаадын уг”, “Тогоон Төмөр хааны түүх”, “Тайзу Чингис хааны үрийн бичиг”, “Хаадын түүх алтан товчоо”, “Эртний түүх”, “Алтан навчит хэмээх түүх оршив” гэх мэт түүхэн чухал сурвалжууд зохиосон эзэн, он цаг нь тодорхойгүй байдаг. Эдгээр бүтээл нь уйгаржин, төвд хэл дээрх түүхэн болон шашны холбогдолтой судар номууд бөгөөд үүнийг орчин үеийн судалгааны эргэлтэд оруулах, уг бүтээлийг бүтээсэн арга технологи, бүтцийг судлах, үүн дээр үндэслэн компьютерын арга ашиглан модон барын бүтээлийг гарын бичмэлээс ялган таних аргыг боловсруулахад судалгааны зорилго оршино. Олон улсад бичиг үсгийн хэв, тэмдэгтийг ялган таних судалгааг машин, гүн сургалтын аргуудаар хийсээр байна.[10-12]

Энэ удаагийн судалгаанд гар бичмэл болон модон барийн дараах эх сурвалжуудыг ашиглан өгөгдлийг бэлтгэсэн. Үүнд: МУҮНС-ийн Монгол судлал, гар бичмэл номын сан хөмрөгөөс авсан зохиогч, бүтээсэн он тодорхойгүй доорх бүтээлүүдийг ашиглав.

Цардмал цайвар цэнхэр даавуун хавтастай, муутуу цаасан дээ р хар бэхээр хичээнгүй бичгээр бичсэн, нийт 24 хуудастай, 26.x26 см хэмжээтэй.[2]

Зураг 1.Гар бичмэл өгөгдлийн жишээ

МУҮНС-ийн /Монгол судлал, гар бичмэл номын сан хөмрөг/ 13x25 хэмжээтэй. Хятад цаасан дээр хар бэхээр гараар хичээнгүйлэн бичсэн, шар даавуун цардмал хавтастай, 20 хуудас бүхий дэвтэрлэж үдсэн хэлбэртэй судар юм. Зохиогч нь тодорхойгүй.



Зураг 2. Гар бичмэл өгөгдлийн жишээ хоёр

“Маха саду-а нэрт тойн нэгэн бээр мандсан гал гахай жил (1887 он) хэв дор сийлгэвэй” гэсэн төгсгөлийн үгтэй 10 хуудас 6,5см × 49см хэмжээтэй барын хайрцаг бүхий модон барын бүтээл.

Зураг 3.Маха саду бүтээлийн хэсгээс

“Дээд шашин эрдэнэ бээр монгол орныг түгээгүүлсэн угийг үзүүлсэн яруу хэлний хүрхрэл нэрт оршив[3]”. Уг зохиолыг судалсан эрдэм шинжилгээний бүтээл ховор ажээ. Монгол үсэг бичгийн түүх, хэл зүйчдийн бүтээлээс

[1] Н.Баярсайхан “Төв халхын уртын дууны шүлгийг туршилтын аргаар шинжлэх нь” докторын диссертац 2024 он.

[2] https://nationallibrary.mn/?p=19304

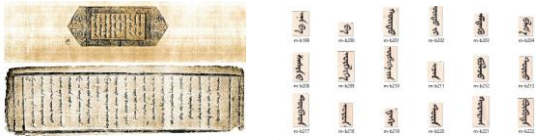
[3] http://www.orientalstudies.ru/rus/images/pdf/journals/b_chuluun_ed_2021_12_munkhsaikhan.pdf

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

гадна монгол түүхийн талаар олон үнэтэй баримт агуулсан модон барын зохиол юм.

Зураг 4. Модон барын бүтээлийн жишээ

“Блама өгүүлэхүйн арслан манзушири эгшигт эх дор мөргөж монголын зуун хорин гурван цагаан толгой үсгийг туурвисан оюун билгийн соёрхол хэмээгч оршив”. Орос цаасан дээр модон бараар барласан 14 хуудас бүхий ном болно. Монгол улсын үндэсний номын санд нэг хувь байгаа ба түүнийг “Монгол улсын 19 дүгээр онд Хан Хэнтий уулын аймгаас сурвалжлан захиж ирүүлсэн” хэмээн академич Б.Ринчен бичжээ.



Зураг 5. Модон барийн бүтээлийн жишээ гурав

МУҮНС-ийн “Монгол судлал, гар бичмэл номын сан хөмрөг”-т 21 мянга орчим гар бичмэл, барын ном, “Төвд номын сан хөмрөг”-т монгол, төвөд эрдэмтэн мэргэдийн бүтээсэн 1,5 сая орчим ном, судар, “Дорно дахины ном хэвлэлийн сан хөмрөг”-т 100 гаруй мянган ховор чухал ном судар хадгалагдаж байна.[4] [4]

Монгол ном судрын онцлог; Номын бар нь мод, зэс, төмөр, чулуун дээр бүтээгдсэн ба урт удаан хугацаагаар хадгалж болдог, эдэлгээ даадаг, дахин сэргээх боломжтой зэрэг ач холбогдолтой баримтат өв.[1] Барын номын үсгийг хөгжлийн хувьд хоёр ангид хуваана: XVI зуунаас урагшлах судрыг эртний хичээнгүй үсэгтэй хэвлэдэг. Үүнийгээ хэлбэршүүлж XVIII зуунаас шашны бус номыг ийм үсэгтэй хэвлэдэг болсон бол XVIII зуунаас хойших судар ном нь дармал үсэгтэй болсон байна.[2] Халх Монголын судар хичээнгүй үсгийн бартай, Буриад монголын судар дармал үсгийн бартай, Өвөр монголынх шашны судар бол дармал үсэгтэй, иргэнийх бол хичээнгүй үсгийн бартай байна. Модон барын номд онцлон тэмдэглүүштэй нэг зүйл бол барчдын бий болгосон сийлбэр үсэг. Сийлбэр үсгийг бас дармал үсэг гэж нэрлэдэг байна.[2]

Модон барын номын үсэг бүр нь мөр шилжиж эс тасарвал үргэлж бүтэн бичигддэгээрээ, сийлбэрлэгддэгээрээ, өрөгддөгөөрөө дэлхийн бусад орны номоос үлэмж ялгаатай.[3] Мөн тод, соёмбо, манж үсгээр ном хэвлэж байжээ. XIX зууны эхэн үеэс хорголжин барын ном хэвлэх болжээ. Модон барыг хус, жодоо, нигүүрс зэрэг модыг сайтар боловсруулан хийж байсан ба эдгээрээс хамгийн тохиромжтой мод бол хус байсан гэж үздэг. Учир нь хус нь ширхэг дотроо хулсгүй, голоо ороодоггүй тул хавтас хийж хатаахад товойдоггүй сайн талтай.[3] Өвөр Монголын зарим барханд алимны мод, зандангаар ч номын бар хийдэг байсан. Судрыг олон хэвлээд ирэхээр номын бар элэгдэж мөлийдөг ба шинэчлэн сэргээдэг. Монгол улсын үндэсний номын сангийн “Монгол судлал, гар бичмэл номын сан хөмрөг”-т 21 мянга орчим гар бичмэл, барын ном, “Төвд номын сан хөмрөг”-т монгол, төвөд эрдэмтэн мэргэдийн бүтээсэн 1,5 сая орчим ном, судар, “Дорно дахины ном хэвлэлийн сан хөмрөг”-т 100 гаруй мянган ховор чухал ном судар хадгалагдаж байна.[4]

II. Арга зүй, алгоритм

Судалгааны ажлын зорилтыг шийдвэрлэхэд, урлаг судлалын иконографи [6], иконологи, герменевтик [9] аргуудыг ашиглан анализ синтезийн аргаар дүрсийг задлан шинжлэхээс гадна хиймэл оюунаар [7] таних боломжийг эрэлхийлэх зорилготой суурь судалгааг хийсэн. Санал болгож буй загварын бүтцийг “Зур. 6”-д харуулаа.

Зураг 6. Санал болгож буй загварын бүтэц

Оролтын дүрсийг уншиж, өөр өөр хэмжээтэй дүрсийг ачаалж, 50x50 пикселийн хэмжээг өөрчилсөн. Дараа нь онцлог шинж чанарыг задлахдаа PCA [8] алгоритмыг ашиглаж, модон барын дүрсийг гар бичмэлээс ялгахын тулд өөр өөр хоёртын ангиллын алгоритмуудыг туршив.

A. *Онцлог шинж задлах аргууд*

PCA [8] нь том хэмжээний хувьсагчийн багцыг жижиг багц болгон хувиргадаг бөгөөд энэ нь том багц дахь гол мэдээллийг агуулдаг. Энэ нь өгөгдлийн өөрчлөлтийг хамгийн ихээр нэмэгдүүлдэг чиглэлийг (үндсэн бүрэлдэхүүнүүдийг) тодорхойлсноор хийдэг.

Principal Component Analysis (PCA) нь өгөгдлийн хэмжээсийг багасгах, онцлог шинжийг гаргах зорилготой статистик арга юм. Доорх нь PCA-г хэрэгжүүлэх үндсэн алхмууд юм:

1. Өгөгдлийг стандартчилах: Өгөгдлийн хэмжээсүүдийг стандартчилах буюу дундаж утга нь 0, стандарт хазайлт нь 1 болгох. Энэ нь өгөгдлийн хэмжээсүүдийг жигдрүүлэхэд тусалдаг.
2. Ковариацийн матриц үүсгэх: Өгөгдлийн ковариацийн матрицыг үүсгэх. Ковариацийн матриц нь өгөгдлийн хэмжээсүүдийн хоорондын хамаарлыг илэрхийлдэг.
3. Eigenvalues болон eigenvectors-ийг тооцоолох: Ковариацийн матрицаас eigenvalues болон eigenvectors-ийг тооцоолох. Eigenvalues нь гол бүрэлдэхүүн хэсгүүдийн чухал байдлыг илэрхийлдэг бол eigenvectors нь гол бүрэлдэхүүн хэсгүүдийн чиглэлийг илэрхийлдэг.
4. Гол бүрэлдэхүүн хэсгүүдийг сонгох: Eigenvalues-ийн хэмжээгээр гол бүрэлдэхүүн

[4] Соёлын баримтат өвийг хамгаалах хөтөлбөр. Засгийн газрын 2014 оны 283 дугаар тогтоолын 1 дүгээр хавсралт.

хэсгүүдийг сонгох. Их eigenvalues-тай бүрэлдэхүүн хэсгүүд нь өгөгдлийн ихэнх мэдээллийг агуулдаг.

5. Өгөгдлийг шинэ координат системд хувиргах: Сонгосон гол бүрэлдэхүүн хэсгүүдийг ашиглан өгөгдлийг шинэ координат системд хувиргах. Энэ нь өгөгдлийн хэмжээсийг багасгах үйл явц юм.
6. Шинэ өгөгдлийг ашиглах: PCA-г ашиглан гарсан шинэ өгөгдлийг анализ хийх, дүрслэх, загвар сургах зэрэгт ашиглах юм.

Бид туршилтаар 2500 онцлогийг 15, 20, 25, 50, 75, 100, 150, 300 зэрэг хэмжээсүүдэд буулган туршилт хийж, үр дүнг хураангуйлан дараах бүлэгт оруулсан.

B. Машин сургалтын аргууд

Машин сургалтын аргууд нь шинж чанараараа тодорхой бүлгүүдэд хуваагддаг [7]. Үүнд Бэйсийн төрлийн аргууд (Bayes Net, Naïve Bayes), функцийн төрлийн аргууд (logistic, SGD, Multilayer Perceptron), lazy аргууд (IBK, LWL), мета төрлийн аргууд (Adaboots M1, Random Committee), дүрэм суурилсан аргууд (Jrip,Zero R), мод төрлийн аргууд (Random Forest, Random Tree) орно. Уг аргуудаар туршилт хийсэн үр дүнг дараагийн бүлэгт дэлгэрэнгүй тайлбарлалаа. Ангиллын үр дүнг 1-р хүснэгтэд үзүүлэв.

III. Туршилтын үр дүн

Өгөгдөл бэлтгэх

Бид Монгол улсын үндэсний номын сангийн сан хөмрөгөөс дээжлэн цуглуулсан Монгол модон барын 519 болон гар бичмэлийн 334 өгөгдөл бүхий сан бүрдүүлсэн. Эндээс нийт өгөгдлийг 10-fold cross validation-аар туршсан. Ангилах аргуудын үр дүнг хүснэгт 1-т харуулав.

Хүснэгт 1 .Ангилах аргуудын үр дүн

	Машин сургалтын аргууд	**15 (PC)**	**100 (PC)**	**150 (PC)**
1	Naïve Bayes	82.09	79.10	80.75
2	Logistic	84.39	87.09	81.46
3	Functions SGD	84.74	86.38	82.51
4	IBK	86.03	86.38	79.46
5	LWL	76.29	74.41	70.12
6	Random Committee	84.74	84.04	80.40
7	Zero R	60.79	60.79	60.79
8	J 48	81.57	80.40	79.11
9	Decision Stump	68.66	68.66	68.66
10	Random Tree	77.69	75.41	65.26
11	**Random Forest**	**88.61**	**88.03**	**85.45**

PCA аргыг ашиглан ялгаатай тооны онцлог шинжийн утгуудад (15, 50, 150) үр дүнг харуулав. Туршилтын үр дүнгээс харахад хамгийн танилт өндөртэй хэмжээс нь 15 байв.

Зураг 6. Машин сургалтын аргуудын үр дүнгийн харьцуулалт

Ангиллын аргуудын хувьд хамгийн сайн үр дүн нь Random Forest болон функцийн төрлийн аргууд байна. Нийтэд нь харахад бидний туршсан өгөгдөл дээр санал болгож буй загвар нь дундажаар 75 аас дээш хувь зөв танилттай, хамгийн сайн танилт нь 88 гаруй хувьтай байна.

IV. ДҮГНЭЛТ

Бидний энэ ажил монголчуудын соёлын өвийн судалгаанд компьютерын ухааны арга технологийг хэрэглэх боломжийг монгол барын судар номын жишээн дээр авч үзэн ном бүтээсэн арга технологи, хэлбэр бүтэц зэргийг судлан дүгнэх ажлын эхлэл гэж болно. Энэхүү судалгаагаараа бид гар бичмэл бүтээлийг барын бүтээлээс машин сургалтын аргаар ялган ангилах арга ашиглан үр дүн гаргалаа. Ашигласан аргуудын хувьд уламжлалт машин сургалтын 11 арга ашигласан ба Random Forest арга PCA 15 байхад 88.6 хувьтай таньж байна. Zero R(60.79%), Decision Stump(68.66%) зэрэг аргууд PCA-ийн аль ч утга дээр өөрчлөлтгүй байна. Нийтээрээ хэрэглэсэн аргууд дээр PCA 15 байхад танилтын хувь илүү байна. Цаашид ховор ном судар бичгүүдийг бүрэн цахимжуулан өгөгдлийн сан үүсгэх, судалгааны эргэлтэнд оруулах, ашиглах шаардлагатай юм.

НОМ ЗҮЙ

[1] “Монголчуудын бичиг үсэг соёлын баримтат өв болох нь” Эрдэм шинжилгээ-онол практикийн хурлын илтгэлийн эмхтгэл. Улаанбаатар 2016.

[2] Ц.Шүгэр,“Монголчуудын ном хэвлэдэг арга”., Улаанбаатар 1976.

[3] Ц.Шүгэр, “Монгол модон барын ном”. Улаанбаатар 1991.

[4] Соёлын баримтат өвийг хамгаалах хөтөлбөр. Засгийн газрын 2014 оны 283 дугаар тогтоолын 1 дүгээр хавсралт.

[5] Ж.Сэржээ, “Монголчуудын номын өв. Ховор нандин судар номын тайлбарууд”. Дэд дэвтэр., Улаанбаатар 2018.

[6] Mayr., Ernst. "Classifications and other ordering systems". J. Zool. Syst. Evol. Research 40(4): 169–94. 2002.

[7] Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

[8] Greenacre, M., Groenen, P.J.F., Hastie, T. et al. Principal component analysis. Nat Rev Methods Primers 2, 100 (2022). https://doi.org/10.1038/s43586-022-00184-wdoi.org.

[9] Iskusstvoed.ru/2016/10/07/иконографический-метод-анализа-худо/

[10] Huang, Q., Li, M., Agustin, D., Li, L., & Jha, M. (2024). A Novel CNN Model for Classification of Chinese Historical Calligraphy Styles in Regular Script Font. *Sensors*, *24*(1), 197. https://doi.org/10.3390/s24010197

[11] C. Tensmeyer, D. Saunders and T. Martinez, "Convolutional Neural Networks for Font Classification," *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, Kyoto, Japan, 2017, pp. 985-990, doi: 10.1109/ICDAR.2017.164.

[12] https://www.icst.pku.edu.cn/zlian/docs/20181024110641005904.pdf

Семантик холбоо ашигласан яриа таних систем

Э.Адъяабат, М.Баярпүрэв, Д.Бямбажав*

ЭХИТ, Мэдээллийн Технологи, Электроникийн Сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
adiyabat@num.edu.mn, bayarpurev@num.edu.mn, byambajav@num.edu.mn

***Хураангуй*—Харилцаа холбооны зургаа дахь үе (6G) нь зөвхөн техник үзүүлэлтүүдээр хязгаарлагдахгүй, хүний ойлголт, сэтгэлгээнд ойрхон, ухаалаг харилцааны шинэ парадигмыг шаарддаг. Энэхүү шинэ хандлага нь мэдээллийг утгын түвшинд дамжуулж, системүүдийн хооронд илүү гүнзгий ойлголцол бий болгох боломжийг олгодог. Харин уламжлалт яриа таних (Automatic Speech Recognition, ASR) системүүд нь зөвхөн дууны дохиог бичвэр болгон хөрвүүлдэг тул утга агуулгыг бүрэн тусгаж чаддаггүй, мөн өгөгдлийн дамжуулалт нь илүүдэл мэдээлэл агуулдаг. Энэхүү судалгааны ажлаар нь семантик холбоо ашиглан яриа таних, яриаг синтез хийх системийн хэрэгжүүлэлтийг хийж гүйцэтгэсэн. Судалгааны хүрээнд яриан таних загварт суурилсан ASR систем болон утга ялгах семантик encoder-ийг хослуулсан архитектурыг боловсруулж, монгол хэл дээрх туршилтын өгөгдлөөр үнэлгээ хийсэн. Мөн семмантик холбоо ашигласан энэхүү системийг динамик сувгийн орчинд тохируулахын тулд бид янз бүрийн сувгийн нөхцөлд дасан зохицох чадвартай загварыг тодорхойлсон. Симуляцийн үр дүнг харахад хэрэгжүүлсэн семантик холбоо дээр суурилсан яриа таних систем нь уламжлалт харилцаа холбооны систем, гүн сургалт дээр суурилсан системүүдийг дохио шуугианы харьцаа бага үед илүү өндөр гүйцэтгэлтэйгээр ажиллаж байгаа харагдаж байна. Ийм систем нь 6G-ийн ухаалаг хэрэглээнд (ухаалаг гэр, тээвэр, эрүүл мэнд гэх мэт) өргөн хэрэглэгдэх үндэс суурь болж чадна.**

***Түлхүүр үгс*—ASR, гүн сургалт, семантик холбоо, яриа танилт, ярианы синтез.**

I. Удиртгал

Харилцаа холбооны технологийн тасралтгүй хөгжил нь уламжлалт мэдээлэл дамжуулалтын хүрээнээс давж, утга агуулгад суурилсан ухаалаг харилцааны шинэ парадигм руу шилжиж байна. 6 дахь үеийн харилцаа холбоо (6G) нь өндөр хурд, бага саатал, найдвартай байдлыг хангахаас гадна мэдээллийн семантик утга, зорилго болон контекст-ийг ойлгох чадвартай байх шаардлагыг тавьж байна [2].Энэ хүрээнд сүүлийн жилүүдэд хөгжиж буй семантик холбоо (semantic communication) нь Шенноны уламжлалт мэдээллийн онолоос (Shannon, 1948) тэс өөр хандлагатай. Шенноны загвар нь мэдээллийг битийн түвшинд алдагдалгүй дамжуулахад чиглэсэн бол семантик холбоо нь мэдээллийн утгыг дамжуулж, хүлээн авагчийн ойлголтын түвшинд алдааг багасгахыг зорьдог [3].Өөрөөр хэлбэл, өгөгдлийг бүрэн дамжуулах бус, зөвхөн тухайн даалгаварт чухал ач холбогдолтой, ойлголт үүсгэхэд шаардлагатай мэдээллийг дамжуулахыг зорьдог. Нөгөөтэйгүүр, яриа таних системүүд (Automatic Speech Recognition, ASR) нь хүний дуу авиаг бичвэр болгон хөрвүүлж, олон төрлийн хэрэглээнд (ухаалаг туслагч, автомат орчуулга, хэлний боловсруулалт г.м.) өргөнөөр ашиглагдаж байна. Гэвч уламжлалт ASR системүүд нь яриаг яг таг (literal) танихыг зорьдог тул өгөгдлийн хэмжээ их, таних алдааны улмаас утгын алдагдал гарч болох сул талтай [4]. Семантик холбоо нь яриа таних системүүдэд нэмэлт утга агуулгын ойлголт оруулж, уламжлалт системийн сул талуудыг даван туулах боломжийг олгодог. Эдгээр судалгааны ажлууд нь семантик холбооны зарчмыг хэрхэн амжилттай ашиглахыг, мөн яриа таних системийг боловсронгуй болгох шинэ аргачлалуудыг судалж байна.

A. DeepSC-S: Гүн сургалт-д суурилсан семантик холбоо

Weng болон Qin [5] нар анхны гүн сургалт (DL)-д суурилсан семантик холбоо системийг танилцуулсан. DeepSC-S нэртэй энэ систем нь ярианы дууны сигналыг анхаарал төвлөрсөн механизмыг ашиглан дэлхийн семантик мэдээлэл рүү хөрвүүлдэг. Энэ систем нь яриа танихад суурилсан семантик холбоог амжилттай хэрэгжүүлж, хүлээн авагч дээр яриаг дахин бүтээж, утгын түвшинд мэдээллийг дамжуулж чаддаг.

B. Федерейт сургалттай олон хэрэглэгчийн аудио семантик холбоо

Tong нар [6] олон хэрэглэгчтэй аудио семантик холбооны системийг боловсруулсан. Энэ систем нь федерейт сургалт ашиглан олон төхөөрөмж дээр хоорондоо хамтран сургалт явуулдаг бөгөөд конволюцийн нейрон сүлжээ (CNN)-д суурилсан автоэнкодерийг ашиглан аудио мэдээллийг боловсруулах аргачлалыг дэвшүүлсэн. Систем нь семантик холбоог хэрхэн ашиглах, олон хэрэглэгчийн оролцоо нь яриа таних үр дүнд хэрхэн нөлөөлдөгийг туршиж үзсэн.

C. Ярианы семантик тэмдэгтүүдийг ашиглах архитектур

Shi нар [7] семантик холбооны ойлголт болон дамжуулалтын архитектурыг боловсруулсан. Тэд яриа-

ны дохиог семантик тэмдэгтүүд рүү хөрвүүлж, хүлээн авагч дээр өндөр семантик үнэн зөвийг баталгаажуулж, хүлээн авсан семантик бэлгэдлүүдийг ярианы долгион руу тайлбарласан. Энэ архитектур нь ярианы семантик үнэн зөвийг хадгалж, дамжуулалтын явцад мэдээллийн алдагдлыг багасгах боломжийг олгодог.

Хэдийгээр яриа танихад семантик холбоо ашиглах нь тодорхой үр дүнгүүдийг гаргаж байгаа ч, семантик холбоог ярианы ойлголт болон оюун ухаанд суурилсан даалгаварт ашиглах судалгаа одоогоор хомс байна. Энэ нь шинээр гарч байгаа асуудал бөгөөд яриа таних системүүдийг зөвхөн үгсийн танилтаас илүү утга агуулгын түвшинд хөгжүүлэх шаардлага байгаа юм. Тиймээс энэхүү судалгааны ажлаар монгол хэлний ярианг сургасан загвар дээрээ семантик холбоо ашиглан хэрэгжүүлнэ.

II. Холбоотой ажлууд

Энэхүү хэсэгт бид гүн сургалтал суурилсан (DL-enabled) семантик холбооны системүүдтэй холбоотой өмнөх судалгааны ажлуудыг танилцуулж, яриа танилт болон ярианы синтезийн орчин үеийн загваруудыг танилцуулах болно.

A. Гүн сургалтад суурилсан семантик холбооны системүүд

Сүүлийн жилүүдэд гүн сургалтад суурилсан семантик холбооны системүүд (Deep Learning-enabled Semantic Communication Systems) нь уламжлалт Shannon-н мэдээллийн онолд суурилсан холбооны архитектуруудаас ангид, илүү "утга төвтэй"дамжуулалтын шинэ парадигмыг бий болгож байна. Уламжлалт системүүд нь мэдээллийг битийн түвшинд үнэн зөв дамжуулахад анхаардаг бол семантик холбооны системүүд нь дамжуулагч болон хүлээн авагчийн хоорондох мэдээллийн утга, учир холбогдол, ойлголтыг хадгалах, нөхөн сэргээхэд чиглэдэг [3].

DL-д суурилсан семантик холбооны системүүд нь гүн нейрон сүлжээ (DNN), encoder-decoder архитектур, attention механизм, transformer загвар зэрэг орчин үеийн хиймэл оюун ухааны аргуудыг ашиглан дараах үндсэн чиглэлүүдээр хөгжиж байна:

Семантик кодчилол ба декодчилол: Encoder-Decoder бүтэц бүхий гүн сургалтын загварууд (жишээлбэл, seq2seq, Transformer, BERT, GPT) нь өгөгдлийн семантик төлөөллийг сурч, шахах, мөн хүлээн авагч талд нөхөн сэргээх боломжийг бүрдүүлдэг. Жишээлбэл, DeepSC загвар нь текст мэдээллийг утгын түвшинд үр дүнтэйгээр кодчилон, сувгийн дундуур дамжуулж, дахин сэргээх боломжтой болохыг харуулсан [8].

Семантик алдааны загварчлал: Уламжлалт битийн алдааны хэмжилт (Bit Error Rate, BER)-ийн оронд утгын алдагдлыг хэмжих шинэ метрикүүд (жишээ нь Semantic Similarity Score, BLEU, BERTScore) ашиглагдаж байна. Ингэснээр дамжуулалтын чанар нь зөвхөн битийн зөв эсэх бус, утгын ялгаанд суурилан үнэлэгддэг [9].

Хамтын ойлголцол (Collaborative Understanding): Илгээгч болон хүлээн авагч талууд нь ижил эсвэл уялдаатай семантик мэдлэгийн суурьтай байх нь мэдээллийг богино кодчилолтой, өндөр ойлголттойгоор дамжуулах боломжийг бүрдүүлдэг. Үүнийг "shared knowledge base"ашиглан эсвэл хамтран сургалтын аргаар (e.g., federated learning) хэрэгжүүлдэг.

Мөн сүүлийн үеийн судалгаанууд нь multimodal семантик холбоо (текст, яриа, дүрсийн хамтлагдсан дамжуулалт), сүлжээний траффикийн оновчлол, семантик channel coding, болон хязгаарлагдмал нөөцтэй төхөөрөмжид зориулсан lightweight загварууд зэрэг дэд чиглэлүүдийг идэвхтэй судалж байна.

B. Яриа танилтын орчин үеийн загварууд (Automatic Speech Recognition, ASR)

Яриа танилт (ASR) нь хүний ярианг автоматаар текст болгон хөрвүүлэх үйл явц бөгөөд орчин үеийн холбоо, харилцаа, хэлний технологид чухал байр суурь эзэлдэг. Уламжлалт ASR системүүд нь *acoustic model (AM)*, *language model (LM)*, болон *pronunciation lexicon* гэсэн үндсэн гурван бүрэлдэхүүн хэсгээс бүрдэж байв. Харин сүүлийн жилүүдэд гүн сургалтад суурилсан аргачлалууд гарч ирснээр эдгээр үе шатнуудыг нэгтгэсэн *end-to-end (E2E)* загварууд түлхүү хэрэглэгдэх болсон.

ASR-ийн орчин үеийн тэргүүлэх архитектуруудад дараах загварууд багтдаг:

- **Connectionist Temporal Classification (CTC):** Оролт ба гаралтын уртын ялгаатай байдлыг шийдвэрлэхэд тохиромжтой хурдан загвар бөгөөд бодит цагийн хэрэглээнд нийцдэг [10].
- **RNN-T (Recurrent Neural Network Transducer):** Sequence-to-sequence зарчимд тулгуурласан, real-time болон streaming ASR-д өргөн ашиглагддаг [11].
- **Transformer болон Conformer:** Self-attention механизмд суурилсан эдгээр загварууд урт хугацааны хамаарлыг үр дүнтэй барьж, өндөр нарийвчлалтай танилт хийдэг [12].
- **Whisper (OpenAI):** Whisper нь өргөн хэмжээний олон хэлний өгөгдөл дээр сургасан encoder-decoder суурьтай загвар бөгөөд дуу танилт, орчуулга, сегментчлэл зэрэг олон даалгаврыг зэрэг гүйцэтгэх чадвартай [13].

C. Ярианы синтезийн орчин үеийн загварууд (Text-to-Speech, TTS)

Ярианы синтез (TTS) нь текстийг хүний ярианд төстэй дуу авианд хөрвүүлэх үйл явц бөгөөд ухаалаг туслах, хэл ярианы тусламж, виртуал дүрүүд зэрэг олон салбарт өргөн хэрэглэгддэг. Уламжлалт TTS системүүд нь *text analysis*, *acoustic modeling*, *waveform generation* гэсэн үе шатуудтай байсан бол орчин үед эдгээрийг

нэгтгэсэн *end-to-end* нейрон суурьтай загварууд давамгайлж байна.

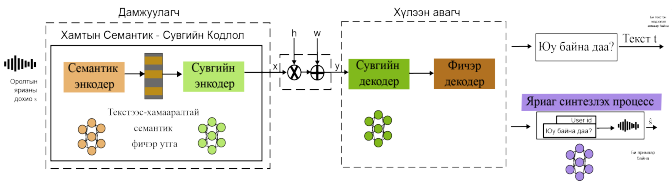
Тэргүүлэх TTS архитектуруудад дараах загварууд багтдаг:

- **Tacotron2:** Character-level input-ыг spectrogram болгон хөрвүүлж, WaveNet ашиглан waveform үүсгэдэг. Байгалийн, хүний дуураймал өнгө аястай яриа үүсгэх чадвартай [14].
- **FastSpeech / FastSpeech2:** Autoregressive загваруудын сул тал болох удаан хурдыг шийдсэн non-autoregressive архитектур бөгөөд real-time TTS-д тохиромжтой [15].
- **VITS (Variational Inference TTS):** GAN, VAE, normalizing flow зэрэг аргуудыг нэгтгэсэн бүрэн end-to-end загвар бөгөөд waveform-ийг шууд текстээс үүсгэдэг [16].
- **StyleTTS, YourTTS:** Синтезийн өнгө аяс, хурд, сэтгэл хөдлөлийг хянах боломжтой, хувь хүний ярианы хэв маягийг дуурайх чадамжтай.

Яриа танилт болон ярианы синтезийн орчин үеийн эдгээр гүн сургалтад суурилсан загварууд нь семантик ойлголт дээр суурилан утгатай, алдаанд тэсвэртэй, хэрэглэгч төвтэй холбооны системүүдийг хөгжүүлэхэд чухал үүрэг гүйцэтгэж байна.

III. Системийн загвар

DL-д суурилсан семантик холбоотой яриа таних систем нь уламжлалт дуу-хэл бичвэр хөрвүүлэх аргуудаас ялгаатай нь дамжуулах мэдээллийн утгыг гүнзгий ойлгож, контекст ба зорилгот суурилсан семантик түвшинд боловсруулдаг. Энэхүү систем нь дараах үндсэн бүрэлдэхүүн хэсгүүдээс бүрдэнэ: оролтын өгөгдөл, дамжуулагч, сувгийн орчин, хүлээн авагч, ярианы синтез.



Зураг 1. Системийн ерөнхий блок диаграмм

A. Оролтын өгөгдөл

Оролтын өгөгдөл нь дараах хэлбэртэй аналог ярианы сигнал юм:

$$x(t) \in \mathbb{R}, \quad t \in [0, T] \tag{1}$$

энд $x(t)$ нь дууны сигнал, T нь нийт хугацаа.

B. Дамжуулагч

Дамжуулагч модуль нь дууны сигналын утгыг ойлгон, шаардлагатай семантик мэдээллийг ялган авч, түүнийг шахаж дамжуулдаг.

a) Онцлог шинж тэмдгийн ялгалт::

$$\mathbf{x} = F(x(t)) \in \mathbb{R}^n \tag{2}$$

энд $F(\cdot)$ нь MFCC эсвэл log-Mel spectrogram зэрэг онцлог ялгах функц.

b) Семантик кодчилол::

$$\mathbf{z} = E_s(\mathbf{x}) \in \mathbb{R}^d, \quad d \ll n \tag{3}$$

энд E_s нь семантик encoder (e.g., Transformer, LSTM), $\mathbf{z}$ нь шахсан семантик вектор.

c) Шахалтын харьцаа::

$$R = \frac{|\mathbf{x}|}{|\mathbf{z}|}, \quad R > 1 \tag{4}$$

C. Сувгийн орчин

Дамжуулалтын үед Gaussian шуугиан үүсч болно:

$$\tilde{\mathbf{z}} = \mathbf{z} + \mathbf{n}, \quad \mathbf{n} \sim \mathcal{N}(0, \sigma^2) \tag{5}$$

энд $\mathbf{n}$ нь дундаж нь 0, дисперс нь σ^2 бүхий нормаль шуугиан.

D. Хүлээн авагч

Хүлээн авагч тал семантик векторыг тайлан, шаардлагатай зорилтот даалгаврыг гүйцэтгэнэ.

a) Семантик декодер::

$$\hat{y} = D_s(\tilde{\mathbf{z}}), \quad \hat{y} \in \mathbb{R}^m \tag{6}$$

энд D_s нь LSTM, Transformer зэрэг decoder.

b) Зорилтот даалгаврын гүйцэтгэл (e.g., ASR)::

$$\hat{w} = f_T(\hat{y}) \tag{7}$$

энд f_T нь зорилтот функц (жишээ нь яриаг бичвэрт хөрвүүлэх).

E. Ярианы синтез (TTS)

$$\hat{x}(t) = S(\hat{y}), \quad S : \text{text} \to \text{speech} \tag{8}$$

энд $S(\cdot)$ нь Tacotron, FastSpeech зэрэг TTS модель.

F. Үзүүлэлтүүд ба Үр дүн

Системийн гүйцэтгэлийг дараах үндсэн хэмжүүрүүдээр үнэлнэ:

a) Семантик нарийвчлал::

$$\mathcal{A}_s = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}(y_i = \hat{y}_i) \tag{9}$$

энд $\mathbb{1}(\cdot)$ нь индикатор функц.

b) Word Error Rate (WER)::

$$\text{WER} = \frac{S + D + I}{N} \tag{10}$$

энд S – солигдсон, D – хасагдсан, I – нэмэгдсэн үгс, N – нийт үгсийн тоо.

c) Семантик төстэй байдлын оноо::

$$\mathrm{Sim}(y, \hat{y}) = \cos(\mathbf{v}_y, \mathbf{v}_{\hat{y}}) \tag{11}$$

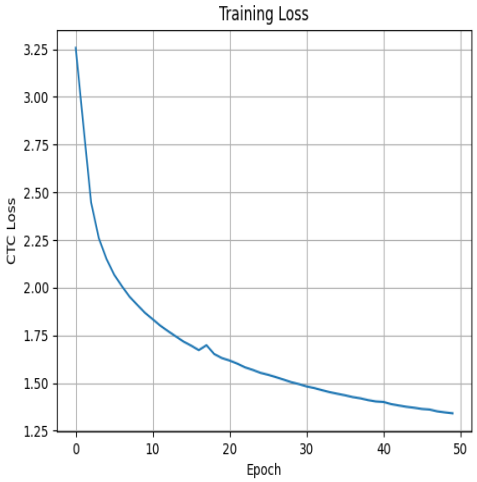
энд $\mathbf{v}_y$, $\mathbf{v}_{\hat{y}}$ нь embedding векторууд (e.g., BERT).

d) (Шахалтын үр ашиг:: Өмнөх (4)-р томъёог дахин хэрэглэнэ.

IV. Хэрэгжүүлэлт ба Үр дүн

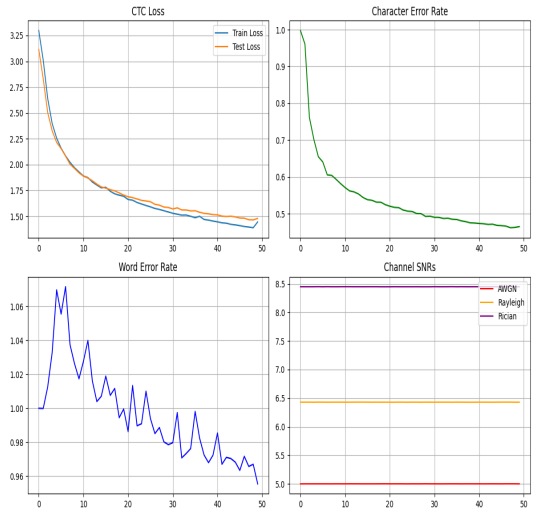
A. Сургалтын процесс

CTC алдагдлын болон сургалтын үеийн epoch-н хоорондох хамааралыг зураг 2-т үзүүлэв. Бид training хийдээ хурдыг 0.0005, 0.0001, 0.00005 гэсэн гурван утгаар авч хийсэн. Эдгээр сургалтын алдагдлаас харахад сургалтын хурд нь 0.00005 байхад CTC алдагдлын бууралт хамгийн удаан бөгөөд гурван сургалтын хурдны дундаас хамгийн өндөр утгатай байсан. Судалгааны аж-



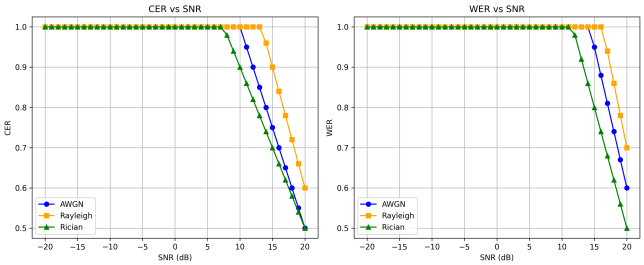
Зураг 2. Сургалтын CTC алдагдал,epoch-н хамаарал

лын хүрээнд нийт монгол хэлний нээлттэй өгөгдлийн сан болох Common-voice 21 [17], өөрсдийн цуглуулсан телевизийн эфирийн яриа, тооны нэршил бүхий 3954 өгүүлбэр аудио файлуудын датасетүүдийг ашигласан. Сургалтын явцад дараах сувгын орчин (fading channel) болох h, Гауссын шуугиан w эдгээрийг ашиглан семантик холбоо ашигласан яриан таних, сентиз хийх (DeepSC-ST) загварыг сургаж, системийн гүйцэтгэлийг янз бүрийн сувгийн орчинд туршсан. Бодитойгоор авч үзвэл, тодорхой сувгийн орчинд системийн гүйцэтгэлийг шалгаж үзвэл яг тэр сувгийн нөхцөлд сургагдсан загвар бусдаас загвараас илүү сайн ажилладаг нь тодорхой юм. Гэсэн хэдий ч, динамик сувгийн орчинд тохируулан маш олон семантик холбооны загвар байршуулах нь тооцооллын нөөц хязгаарлагдмал учраас хэрэгжих боломжгүй. Иймээс сувгийн өөрчлөлтөд дасч ажиллах чадвартай нэг загвар хөгжүүлэх шаардлага гарч байна. AWGN, Rayleigh, болон Rician гэсэн гурван төрлийн сувгийн нөхцөлд сургагдсан семантик холбоо ашигласан яриа таних загварын тестийн гүйцэтгэлийг харьцуулсан судалгаагаар, SNR=10 дБ бүхий Rician сувгийн нөхцөлд сургагдсан энэхүү загвар нь сувгийн төрөл бүрийн орчинд хамгийн тогтвортой (robust) ажиллаж чадах загвар гэж үзэж байна. Сургалтын



Зураг 3. Сургалтын CTC loss, CER vs epoch, WER vs epoch, SNR vs Channels

процессыг сургалтын өгөгдөл 80 хувь, шалгалтын өгөгдөл 20 хувь хуваасан. Зураг 3-аас харахад сургалтын алдагдал болон шалгалтын алдагдал хоёулаа унах хандлагатай байна. Энэ нь сургалтын явцад загвар CTC loss-ийг амжилттай багасгаж байгааг харуулж байна. Сургалтын алдагдал бага зэрэг тестийн алдагдлаас доогуур байгааг харж болно. Энэ нь сургалтын ба шалгалтын өгөгдөл дээр тогтвортой суралцаж байгааг илтгэж байна. Мөн алдагдлын curves-ын хоорондох зөрүү их биш байгаа нь overfitting бага байна гэж дүгнэж болно. CER эхэндээ 1 (100 хувь) орчим байсан бөгөөд сургалтын явцад хурдан буурч 0.45 орчим болсон байна. Эндээс үзэхэд энэ загвар нь тэмдэгт төвшинд илэрхий сайжирсан гэсэн үг. CER-ийн уналт нь сургалт үр дүнтэй болж байгааг харуулж байна. WER эхэндээ 1 буюу 100 хувь байж, аажмаар буурсан ч хэлбэлзэлтэй (fluctuation) байгаа нь анзаарагдаж байна. Зарим үед 1.07 хүртэл өссөн байна. Энэ нь магадгүй манай хэрэгжүүлсэн загварт data augmentation, Beam decoding параметр,Тест өгөгдлийн чанар эсвэл Model instability-тай холбоотой байж болнk гэж бодож байна. Гэсэн хэдий ч ерөнхийдөө WER буурсан сургалт үр дүн харагдаж байгаа. AWGN суваг (SNR=5 dB), Rayleigh суваг (SNR = 6.4 dB), Rician суваг (SNR = 8.4 dB) гэсэн тогтмол SNR утгуудаар семантик суурьтай сувгийн загварын (semantic communication channel) туршилтанд ашигласан. Эдгээрээс утгуудаас Rician суваг хамгийн өндөр SNR-тай байгаа нь хамгийн сайн сувгийн нөхцөлийг харуулж байгаа. Энэхүү зураг 4-т хэлний танилтын системийг гурван төрлийн сувгаар дам-



Зураг 4. CER score versus SNR, WER score versus SNR

жуулсан нөхцөлд (AWGN, Rayleigh, Rician) Character Error Rate (CER), Word Error Rate (WER)-ийг дохио шуугианы харьцаатай харьцуулан харууллаа. $SNR < 10dB$ үед бүх сувгийн CER бараг 1 буюу 100 хувийн алдаатай байна. Энэ нь дохио шуугианы харьцаа ихтэй нөхцөлд систем бараг зөв ойлгож чадахгүй байгааг харуулж байна. $10dB < SNR < 20dB$ үед CER огцом буурч байна. Rician суваг хамгийн хурдан буурч, 15 dB орчимд CER < 0.8 болж байна. AWGN суваг нь дундаж үзүүлэлттэй. Rayleigh суваг дээр хамгийн муу гүйцэтгэлтэй, CER буурах хандлага хамгийн удаан байна. $SNR > 20dB$ үед гурван суваг бүгд $CER = 0.5$-д хүрч, систем хүлээн авахад сайн болсон байгааг харж болно. $SNR < 10dB$ үед WER 1 (100 хувь алдаатай) бөгөөд WER буурах хандлага мөн $Rician > AWGN > Rayleigh$ гэсэн дараалаар байна. Rician сувгийн нөхцөл нь олон замт тархалттай ч шууд харагдах зам (LoS component)-тай учир дохио хүлээн авалт сайжирч, WER буурсан байна. Rayleigh суваг нь зөвхөн олон замтай сарнилтаар дамждаг тул хамгийн муу үр дүнтэй байгааг харуулж байна. Эдгээрээс ерөнхийд дүгнэхэд SNR өсөх тусам системийн гүйцэтгэл илт сайжирч байна. Rician суваг хамгийн сайн, Rayleigh суваг хамгийн муу гүйцэтгэлтэй байна. Сувгийн төрөл нь яриа танилтын системийн гүйцэтгэлд мэдэгдэхүйц нөлөөлж байгааг харж болно. Энэ нь семантик суурьтай дамжуулалтын системд сувгийн загварчлал маш чухал гэдгийг харуулж байна.

V. Дүгнэлт

Энэхүү судалгааны ажлаар бид семантик холбоо ашигласан яриа таних, синтез хийх системийг судлаж, хэрэгжүүлэлтийг хийж гүйцэтгэсэн. Судалгааны ажлын хүрээнд DeepSC-ST загвартай төстэй системийг хэрэгжүүлсэн. Энэ систем нь текстэн холбоотой семантик шинж чанаруудыг ашиглан текстийн бичвэрийг сэргээж, хүлээн авагч талд ярианы дохиог синтез хийж өгдөг. Бид семантик суваг, семантик шинж чанаруудыг суралцах, сувгийн нөлөөг бууруулахад зориулан яриа танилтын гүйцэтгэлийг сайжруулж байгаа юм. Симуляцийн үр дүнгээс үзэхэд бидний хөгжүүлсэн систем нь уламжлалт харилцаа холбооны системүүд, өмнөх семантик харилцаа холбооны системүүдийг, ялангуяа дохио шуугианы харьцаа багатай орчинд илүү өндөр гүйцэтгэлтэйгээр ажиллаж байна. Мөн бид цаашид хүний яриаг өөр хүн ярьж байгаа юм шиг ярьдаг программ хангамжийн туршилтын хувилбарыг бэлтгэхээр ажиллаж байна. Бидний санал болгож буй яриа таних, яриаг синтез хийх энэхүү систем нь семантик харилцаа холбоонд хэрэглэгдэх бүрэн боломжтой систем гэж үзэж байна.

Ашигласан ном

[1] Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27(3), 379–423.

[2] Saad, W., Bennis, M., & Chen, M. (2020). A vision of 6G wireless systems: Applications, trends, technologies, and open research problems. IEEE Network, 34(3), 134–142.

[3] Xie, H., Qin, Z., Shi, W., et al. (2021). Deep learning enabled semantic communication systems. IEEE Transactions on Signal Processing, 69, 2663–2675.

[4] Huang, C., Lin, S. W., Zhang, Y., & Li, W. (2022). Towards robust and efficient speech recognition: A review of recent progress. IEEE Access, 10, 92356–92374.

[5] Weng, X., & Qin, Z. (2020). DeepSC-S: A Deep Learning-based Semantic Communication System for Speech Signal Reconstruction. IEEE Transactions on Signal Processing, 68, 2175-2187. https://doi.org/10.1109/TSP.2020.3012332

[6] Tong, L., Wang, J., Yang, F., & Liu, L. (2021). Federated Learning-based Multi-User Audio Semantic Communication System. IEEE Transactions on Communications, 69(7), 4720-4732. https://doi.org/10.1109/TCOMM.2021.3068503

[7] Shi, L., Chen, X., Zhang, Y., & Li, M. (2021). Semantic Communication for Speech Transmission: Architecture and Performance Evaluation. IEEE Transactions on Wireless Communications, 20(4), 2390-2403. https://doi.org/10.1109/TWC.2021.3042579

[8] Zhang, J., Xu, X., Xiao, Y., Letaief, K. B. (2022). Toward a Semantic Communication System: Modeling, Architecture, and Implementation. IEEE Journal on Selected Areas in Communications.

[9] Farsad, N., Goldsmith, A. (2021). Neural Network-Based Semantic Communication over the Air. arXiv preprint arXiv:2105.05560.

[10] Graves, A., et al. (2006). Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks. ICML.

[11] Rao, K., Sak, H. (2017). Streaming End-to-End Speech Recognition for Mobile Devices. arXiv:1704.07850.

[12] Gulati, A., et al. (2020). Conformer: Convolution-augmented Transformer for Speech Recognition. Interspeech.

[13] Radford, A., et al. (2022). Robust Speech Recognition via Large-Scale Weak Supervision. OpenAI.

[14] Shen, J., et al. (2018). Natural TTS synthesis by conditioning Wavenet on Mel spectrogram predictions. ICASSP.

[15] Ren, Y., et al. (2020). FastSpeech 2: Fast and High-Quality End-to-End Text to Speech. NeurIPS.

[16] Kim, J., et al. (2021). Conditional Variational Autoencoder with Adversarial Learning for End-to-End TTS. NeurIPS.

[17] Common voice 21 https://commonvoice.mozilla.org/mn/datasets

Хиймэл оюун ухаанд суурилсан арьс оношилгооны аппликейшн

Болор-Эрдэний Пүрэвжаргал
Мэдээллийн технологийн салбар
ШУТИС, МХТС
Улаанбаатар, Монгол
purevjargalbolorerdene@gmail.com

Түвшинтөгсийн Энхжин
Мэдээллийн технологийн салбар
ШУТИС, МХТС
Улаанбаатар, Монгол
enhkjint@gmail.com

Алтансүхийн Сүхбат
Мэдээллийн технологийн салбар
ШУТИС, МХТС
Улаанбаатар, Монгол
suhee6357@gmail.com

***Хураангуй* — Энэхүү судалгаагаар бид хиймэл оюун ухаан (AI) болон компьютерын харааND суурилсан арьс оношилгоо хийдэг, монгол хэлээр арьс арчилгааны зөвлөгөө өгдөг ухаалаг гар утасны аппликейшнийг амжилттай хөгжүүлсэн. Хүний арьсыг хэдэн төрөлд хуваааж авч үздэг, арьсны онцлогоос шалтгаалаад арьсан дээрх бохирдолт болон нүхжилт, нөсөө толбо, мэнгэ, батгашилт зэрэг хэрхэн үүсдэгийг судалж, хэрэглээнд буй биет төхөөрөмжүүдийн оношлох хугацаа, үр дүн, хэрэглэх үнэ өртгийг судалж, мэргэжлийн оношилгоог хүн бүрд хүртээмжтэй, хямд зардлаар, бодит цагт авах боломжийг бий болгохыг зорьсон ба хөгжүүлсэн аппликейшнийг танилцуулж байна.**

Түлхүүр үг— Хиймэл оюун ухаан, машин сургалт, компьютерын хараа, арьс оношилгоо

I. Удиртгал (Style: *Heading 1*)

Сүүлийн жилүүдэд хиймэл оюун ухаан (AI) болон компьютер хараа (Computer Vision)-д тулгуурласан арьсны оношилгооны судалгаа, хэрэглээ эрчимтэй хөгжиж байна. Хүмүүс гоо сайхны арчилгаанд илүү анхаарах болсноор арьсны төрөл, батга, үрчлээ, нөсөө толбо зэрэг арьсны онцлогийг тодорхойлох хэрэгцээ нэмэгдэж, гар утсаар ажилладаг хялбар, хямд оношилгооны системүүд эрэлттэй болж байна. Арьсны өвчлөл, гоо сайхны асуудлыг цаг алдалгүй оношлох нь зөвхөн гоо сайхны хувьд бус, эрүүл мэндийн чухал асуудал болоод байна. Арьсны шинжилгээний уламжлалт арга нь эмчид очиж үзүүлэх, үнэтэй тоног төхөөрөмж ашиглах шаардлагатай байдаг тул хүн бүрд хүртээмж муутай. Харин гар утасны камер болон хиймэл оюун ухааны хослол нь уг асуудлыг шийдвэрлэх шинэ боломжийг нээж байна. Өмнөх судалгаануудын хүрээнд арьсны төрөл тодорхойлхдоо гүн сургалтын Convolutional Neural Network (CNN) загваруудыг, арьсны асуудал илрүүлэхэд объект илрүүлэлтийн YOLO төрлийн загваруудыг ашигласан байдаг. Гэвч эдгээр судалгаануудын олонх нь англи хэл дээр хийгдсэн бөгөөд Монгол арьсны онцлогт нийцсэн, Монгол хэл дээр зөвлөгөө өгдөг гар утасны программ ховор байна. Иймээс энэхүү судалгаагаар гар утасны камераар авсан зураг дээр үндэслэн арьсны төрлийг тодорхойлох, арьсны асуудлуудыг илрүүлэх, мөн дүн шинжилгээнд тулгуурлан хэрэглэгчид Монгол хэл дээр нөхөрсөг, ойлгомжтой зөвлөгөө өгөх Skin AI Analyzer нэртэй гар утасны программын загварыг боловсруулсан болно. Энэхүү системийг бүтээхэд гүн сургалтын YOLOv8 ба CNN загваруудыг сургасан бөгөөд гарсан үр дүнг тайлбарлаж, цаашид бодит амьдралд хэрэгжүүлэх боломжийг үнэлсэн болно.

II. Холбоотой ажлууд

A. Загварыг сонгох нь

Арьсны оношилгооны системд хамгийн тохиромжтой гүн сургалтын архитектурийг сонгох нь үр дүнтэй, хөнгөн, сургах боломжтой загварыг тодорхойлох үндсэн нөхцөл байв. Судалгаанд арьсны төрөл болон асуудлыг оношлох хоёр төрлийн зорилтот функц бүхий хоёр өөр архитектурыг сонгон ашигласан. Арьсны төрлийг тодорхойлох зорилгоор MobileNetV2 загварыг ашигласан бөгөөд энэ нь хөнгөн, CPU дээр ч үр дүнтэй ажиллах боломжтой, урьдчилан сургасан жингүүдтэй архитектур юм. Арьсны асуудлыг (батга, толбо, үрчлээ) илрүүлэхэд зориулан объект илрүүлэгч YOLOv8 архитектурын хамгийн хөнгөн хувилбар болох YOLOv8n (nano)-г сонгосон.

MobileNetV2 нь багтаамж багатай төхөөрөмжүүд дээр ажиллах боломжтой, convolutional блокийг илүү үр дүнтэй ашигладаг архитектур бөгөөд арьсны төрлийн classification асуудалд тохиромжтой. YOLOv8n нь real-time нөхцөлд арьсны олон асуудлыг зэрэг илрүүлэх, bounding box-оор тэмдэглэх боломжийг олгодог. Энэ төрлийн загварууд нь Монголын нөхцөлд GPU шаардлагагүйгээр Colab эсвэл дунд зэргийн компьютер дээр туршихад тохиромжтой байдаг.

B. Технологи сонгох нь

Системийн архитектур болон хөгжүүлэлтийн орчныг сонгохдоо хялбар, үнэгүй, олон нийтэд нээлттэй, сургалтын чанарт эерэг нөлөө үзүүлэх платформыг сонгов. Сургалт, туршилтыг Google Colab орчинд хийсэн бөгөөд энэ нь CPU болон хязгаарлагдмал GPU орчныг үнэгүй ашиглах боломжийг олгодог. Өгөгдлийн хувьд арьсны төрлүүд болон асуудлуудтай холбоотой annotated dataset-уудыг Kaggle болон Hugging Face платформуудаас татан авсан.

Object detection хэсэгт ашиглагдах аннотац бүхий өгөгдлийг Roboflow орчинд боловсруулсан. Энэхүү платформ нь зураг бүрийг bounding box хэлбэрээр тэмдэглэж, YOLO формат руу хөрвүүлэн dataset-ийг автоматаар бүрдүүлэх боломжтой. Ингэснээр хэрэглэгч арьсны асуудлыг машин сургалтад зориулж урьдчилан бэлдэх ажлыг илүү үр дүнтэй гүйцэтгэх боломжтой.

Мөн боловсруулсан оношилгооны үр дүнд тулгуурлан хэрэглэгчийн арьсанд тохирсон арчилгаа, хоол тэжээл, анхаарах бодисын талаарх зөвлөгөөг хүргэх зорилгоор Gemini AI (Gemini 2.5) загварыг ашигласан. Энэхүү prompt-based AI нь Монгол хэл дээр найрсаг, ойлгомжтой, хувь хүнд тохирсон зөвлөгөөг хэрэглэгчдэд

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

хүргэх чадвартай гэдгээрээ судалгаанд онцгой ач холбогдолтой байлаа.

C. Сургалт, хөгжүүлэлт

Энэхүү системийн сургалт хоёр үндсэн зорилттой хийгдсэн: (1) арьсны төрлийг тодорхойлох ангиллын (classification) загвар, (2) арьсны асуудлыг илрүүлэх илрүүлэлтийн (object detection) загвар. Эдгээр зорилтуудад тус тусдаа гүн сургалтын аргуудыг ашигласан бөгөөд өгөгдөл боловсруулалт, сургалтын стратегийг сайтар тохируулсан.

1. Арьсны төрлийн ангилал (Classification):

MobileNetV2 гүн нейрон сүлжээг ImageNet dataset дээр урьдчилан сургасан жингүүдтэйгээр ашиглаж, бидний боловсруулсан арьсны төрөл (тослог, хуурай, холимог гэх мэт) бүхий dataset дээр fine-tuning хийсэн. Энэхүү dataset нь нийт 1200 гаруй зурагнаас бүрдсэн бөгөөд өгөгдлийг 70:20:10 харьцаагаар train/validation/test болгон хуваасан. Сургалтыг дунджаар 50 epoch орчим явуулсан бөгөөд баталгаажуулалтын шатанд overfitting илрэхгүй байлгах үүднээс early stopping болон dropout техникийг ашигласан.

2. Арьсны асуудал илрүүлэлт (Object Detection):

YOLOv8n загварыг арьсны нийтлэг асуудлууд болох батга, толбо, үрчлээ гэх мэтийг илрүүлэх зорилгоор ашигласан. Бид аннотац хийгдсэн dataset-ийг 640x640 пикселийн хэмжээтэй стандартаар бэлдэж, өгөгдлийг augmentation (эргүүлэх, гэрэлтүүлэг өөрчлөх, scale хийх) техникийг ашиглан 2 дахин өргөтгөсөн. Нийт 100 epoch-ийн турш сургалт явагдсан бөгөөд Google Colab орчны T4 GPU болон CPU хосолсон нөхцөлд боловсруулсан. Training loss, validation loss, mAP зэрэг үзүүлэлтүүдийг давхар хянаж, хамгийн сайн гүйцэтгэлтэй checkpoint-ийг хадгалсан.

III. Үр дүн

Системийн гүйцэтгэлийг хоёр үндсэн шалгуураар үнэлсэн:

(1) Арьсны төрөл ангилах нарийвчлал (accuracy, F1-score),

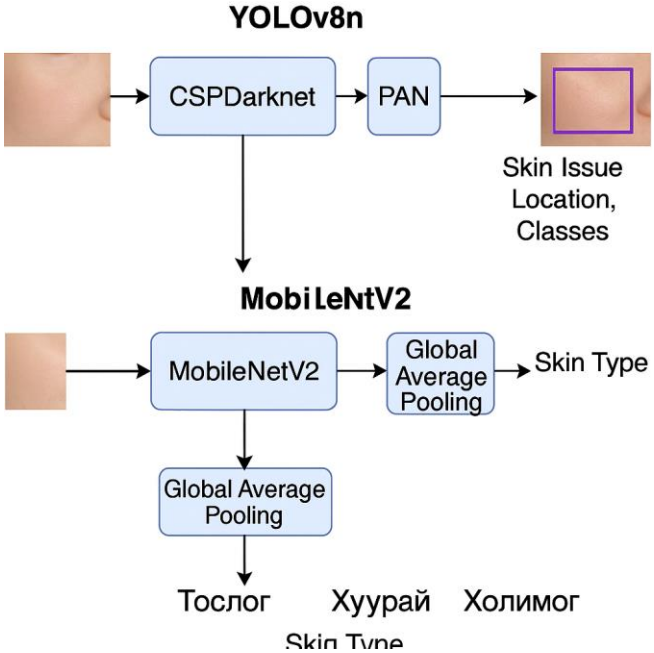
(2) Арьсны асуудал илрүүлэлтийн mAP (mean Average Precision) үзүүлэлт.

1. Арьсны төрөл ангилал (MobileNetV2):

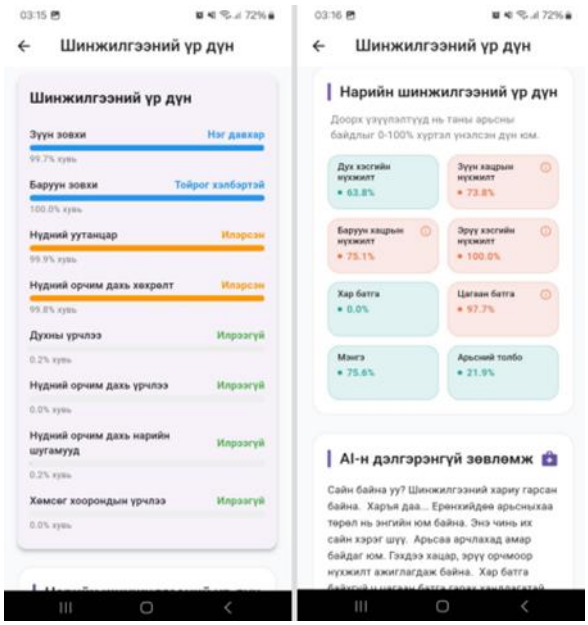
Validation set дээр MobileNetV2 загвар нь 87.3% нийт нарийвчлал (accuracy), 0.84 F1-score үзүүлсэн. Confusion matrix-аас харахад тослог арьсыг бусад төрлөөс илүү нарийвчлалтай ангилсан бол холимог болон хуурай арьсны хооронд зарим overlap ажиглагдсан. Энэ нь эдгээр төрлүүдийн визуал ялгаа бага байдагтай холбоотой байж болох юм. Загварын ROC-AUC муруй нь 0.89-ийг давсан нь оношилгооны чадвар сайн байгааг илтгэж байна.

2. Арьсны асуудал илрүүлэлт (YOLOv8n):

YOLOv8n загвар нь mAP@0.5 үзүүлэлтээр 76.4% хүрсэн бөгөөд батга болон толбо илрүүлэлтийн confidence score харьцангуй өндөр байсан (батга – 82%, толбо – 79%). Харин үрчлээ илрүүлэлт харьцангуй бага confidence (ойролцоогоор 63%) үзүүлсэн нь уг төрлийн dataset харьцангуй цөөн хэмжээтэй байсантай холбоотой гэж дүгнэж болно. Detection-ийн latency буюу нэг зургийн боловсруулах хугацаа 20-30ms орчим байсан нь mobile-д суурилсан хэрэглээнд хангалттай хурдан үзүүлэлт юм.



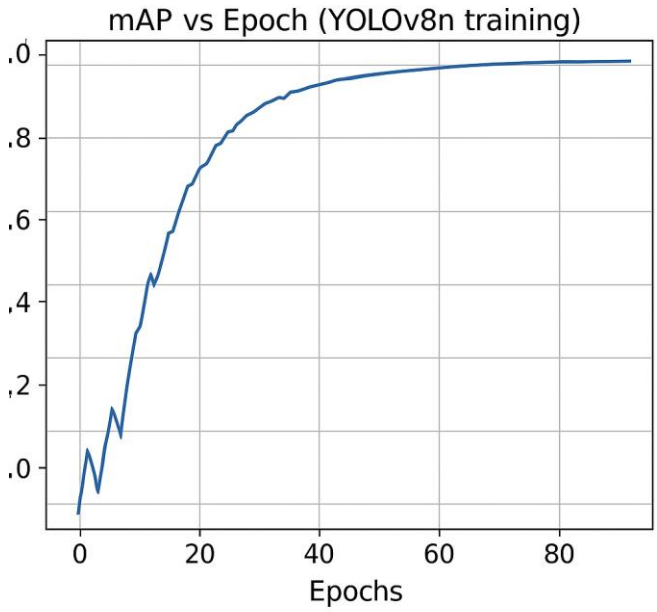
Зураг 1. Системийн архитектур (*Аппликейшнээс уншсан зургийг боловсруулж, үр дүнг гаргаж байна*.)



Зураг 2. Үр дүн (*Аппликейшн дээрх оношилгоо хийсэн туршилтийн үр дүн.*)

Арьсны төрөл	Precision	Recall	F1-score
Тослог	0.91	0.89	0.90
Хуурай	0.84	0.79	0.81
Холимог	0.80	0.78	0.79
Дундаж	**0.85**	**0.82**	**0.84**

Зураг 3. Classification үр дүн



Зураг 4. mAP vs Epoch (YOLOv8n training)

IV. Дүгнэлт

Энэхүү судалгаагаар арьсны төрлийг ангилах болон арьсны асуудлыг илрүүлэх хиймэл оюуны загваруудыг боловсруулан, мобайл орчинд ашиглах боломжтой Skin AI Analyzer системийг хөгжүүллээ. Сургалтын шатанд MobileNetV2 болон YOLOv8n зэрэг хөнгөн жинтэй, өндөр гүйцэтгэлтэй загваруудыг ашиглаж, өөрсдийн боловсруулсан dataset-ууд дээр fine-tuning болон object detection сургалтыг тус тусад нь амжилттай гүйцэтгэсэн.

MobileNetV2 загвар нь арьсны төрлийг ангилах ангилалд 87.3% нарийвчлал, 0.84 F1-score үзүүлсэн нь практик хэрэглээнд хангалттай нарийвчлалтай болохыг харуулж байна. YOLOv8n загвар нь батга, толбо, үрчлээ зэрэг арьсны асуудлыг илрүүлэхэд 76.4% mAP@0.5 үзүүлэлттэй ажиллаж, бодит хэрэглээнд ашиглах боломжтойг нотоллоо. Ялангуяа батга болон толбо илрүүлэлт нь өндөр confidence-тай гарсан бол үрчлээ илрүүлэлтэд dataset-ийн хэмжээ бага байсантай холбоотойгоор бага зэрэг хязгаарлалт ажиглагдсан.

Сургалт болон үнэлгээний ажлыг CPU болон GPU хосолсон Google Colab орчинд хэрэгжүүлснээр хөгжүүлэлтийн нөөцийг хэмнэх боломжтой болсон. Dataset-ийн аннотац, augmentation, хөнгөн жинтэй загварын сонголт нь системийг мобайл төхөөрөмжид нэвтрүүлэхэд тохиромжтой архитектураар шийдэгдсэнийг харуулж байна.

Цаашдын судалгаанд үрчлээ болон бусад ховор төрлийн арьсны асуудалтай холбоотой өгөгдлийг нэмэгдүүлж, загварын гүйцэтгэлийг сайжруулах, мөн хэрэглэгчдэд зориулсан интерактив UI, Mongolian skincare knowledge graph-тэй холболт бүхий зөвлөмжийн системийг боловсронгуй болгох нь зүйтэй гэж үзлээ.

Ном зүй

[1] YOLOv8 - Ultralytics: https://docs.ultralytics.com

[2] CNN - https://www.geeksforgeeks.org/convolutional-neural-network-cnn-in-machine-learning/.

[3] Gemini 2.5 API: https://aistudio.google.com/apikey

[4] Kaggle Dataset: https://www.kaggle.com.

[5] Hugging Face Dataset: https://huggingface.co.

[6] He, K., et al., “Deep Residual Learning for Image Recognition,” CVPR, 2016.

CCTV/IP Камерын Сүлжээний Аюулгүй Байдлын Шинжилгээ (Монгол Улсын кейс судалгаа)

О.Энхтуул
Мэдээллийн сүлжээ аюулгүй байдлын салбар
ШУТИС, МХТС
Монгол Улс , Улаанбаатар хот
enkhtuul.o@must.edu.mn

Б.Дэнсмаа
Мэдээллийн сүлжээ аюулгүй байдлын салбар
ШУТИС, МХТС
Монгол Улс , Улаанбаатар хот
b.densmaa@must.edu.mn
ORCID :0000-0002-1016-9622

Хураангуй— Сүүлийн жилүүдэд IP камер болон CCTV системүүд нь зөвхөн хяналтын хэрэгсэл байхаас хальж, кибер халдлагын үндсэн гарц болох эрсдэлтэй болж байна. Технологийн хөгжлийн нөлөөгөөр аналог камеруудаас тоон систем рүү шилжиж, интернэтээр дамжин алсаас хянах боломжтой болсон нь эдгээр төхөөрөмжүүдийг халдлагад өртөмтгий болгох шалтгаан болжээ. Ялангуяа анхдагч эсвэл сул нууц үгтэй, программ хангамж нь шинэчлэгдээгүй, хамгаалалт муутай IP камерууд нь байгууллагын дотоод сүлжээнд аюул занал учруулж, гаднын халдагчдад алсаас хяналт авах боломжийг бүрдүүлдэг.

Энэхүү судалгаанд UB-19 гэсэн хаалттай орчинд ашиглагдаж буй IP камер болон NVR (Network Video Recorder) төхөөрөмжүүдийн аюулгүй байдлын эрсдэлийг тодорхойлоход чиглэсэн. Судалгааны хүрээнд Wireshark, Zenmap, RegShot, Process Monitor зэрэг сүлжээний анализын хэрэгслүүдийг ашиглаж, өгөгдлийн урсгал, портын хандалт, системийн өөрчлөлтүүдийг шинжилсэн. Үүний үр дүнд анхдагч нууц үг өөрчлөгдөөгүй, сул хамгаалалттай төхөөрөмжүүд нь кибер халдлагад өртөх өндөр магадлалтай болохыг батлалсан. CVE-2013-3381 зэрэг эмзэг байдлыг ашиглан гаднын халдагчид камер болон түүний холбогдсон сүлжээнд нэвтрэх боломжтой байсан ба эдгээр төхөөрөмжүүд гаднын сервер рүү зөвшөөрөлгүй өгөгдөл дамжуулж байгааг илрүүлсэн.

Түлхүүр үг— кибер аюулгүй байдал, камерын эмзэг байдал, халдлага, DoS, хакер

I. Удиртгал

үүлийн жилүүдэд IP камерууд нь интернэтэд холбогдож, алсын зайн хяналт хийх боломжтой болсон нь тэдгээрийг кибер халдлагад өртөмтгий болгоход нөлөөлж байна [1]. Энэ нь камерын хяналт, мэдээлэл дамжуулалт, аюулгүй байдлыг сайжруулсан ч кибер халдлагад өртөх эрсдэлийг нэмэгдүүлж байна[2].

Сүүлийн жилүүдэд олон улсын судалгаагаар IP камерууд нь сүлжээнд бүрэн нэвтрэх боломжийг олгож, тэдгээрийг ашиглан халдлага үйлдэж буй тохиолдлууд ихсэж байгааг харуулж байна [3]. Төр болон хувийн хэвшлийн байгууллагын мэдээллийн сан, компьютерын сүлжээнд нэвтрэхийн тулд IP камеруудыг халдлагын гарц болгон ашиглах боломжтой бөгөөд зайнаас тухайн байгууллагын сүлжээний үйл ажиллагааг тандан судлах, интернетийн гарц үүсгэх зэрэг аюул заналхийллийг бий болгож байна[4].

Энэхүү судалгааны зорилго нь Монгол Улсад өргөн хэрэглэгддэг IP камер болон тэдгээрийн NVR төхөөрөмжүүдийн аюулгүй байдлын эмзэг цэгүүдийг тодорхойлох, боломжит халдлагын арга хэлбэрүүдийг туршилтаар илрүүлэх, хамгаалалтын арга хэмжээ боловсруулахад оршино [5]. Судалгааны хүрээнд Монгол Улсад ашиглагдаж буй IP камер, NVR (Network Video Recorder) төхөөрөмжүүд дээр туршилт хийж, Wireshark, Zenmap, RegShot, Process Monitor зэрэг сүлжээний аюулгүй байдлын хэрэгслүүдийг ашиглан өгөгдлийн урсгалыг хянах, халдлагын шинж тэмдгүүдийг илрүүлэхийг зорьсон[6]. Судалгааны дүнд IP камерын сул хамгаалалт, CVE (Common Vulnerabilities and Exposures) эмзэг байдал, өгөгдөл дамжуулалтын аюулгүй байдлын алдаа зэрэг нь байгууллагын сүлжээнд аюул учруулж болохыг харуулсан[7]. Камеруудын сул хамгаалалт нь зөвхөн хувь хүн, аж ахуйн нэгжүүдийн төдийгүй, улсын мэдээллийн аюулгүй байдалд ч ноцтой эрсдэл учруулж болзошгүй тул хамгаалалтын арга хэмжээг оновчтой боловсруулах, хэрэгжүүлэх нь нэн тэргүүний асуудал болоод байна[7]. Энэхүү судалгаа нь IP камерын аюулгүй байдлын эрсдэлүүдийг тодорхойлох, CVE эмзэг байдлуудыг шинжлэх, кибер халдлагаас урьдчилан сэргийлэх арга замыг боловсруулахад чухал ач холбогдолтой[3].

II. Онолын судалгаа

SNMP Trap мессеж порт 162-г ашиглан удирдлагын төв рүү илгээдэг бөгөөд хэрвээ төхөөрөмж дээр зөв тохиргоогүй, authentication-гүй байвал уг порт нь халдлагад өртөх боломжийг бий болгодог [8]. Энэ нь чиглүүлэгч (router), унтраалга (switch), галт хана (firewall), сервер болон бусад сүлжээний төхөөрөмжүүдээс бүрддэг[9]. IP камер нь сүлжээгээр холбогдож, дүрс мэдээллийг дамжуулдаг ухаалаг төхөөрөмж бөгөөд ихэвчлэн алсын хяналт хийх, аюулгүй байдлыг хангахад ашиглагддаг[10]. Гэсэн хэдий ч, кибер халдлагын өндөр эрсдэлтэй байдаг. IP камерын аюулгүй байдлыг зөрчих гол эмзэг байдал нь дараах хүчин зүйлүүд нь[11]:

- Алдаатай программ хангамж (Firmware vulnerabilities): Камерын үйлдвэрлэгчийн программ хангамжийг шинэчлээгүйгээс үүдэн халдагчид сул талыг нь ашиглах боломжтой[12].
- Сул нууц үг (Weak passwords): Анхдагч эсвэл хялбар таамаглагдах нууц үгнээс шалтгаалан гаднын этгээд нэвтрэх боломжтой[13].
- Шифрлэлгүй өгөгдөл дамжуулалт (Unencrypted data transmission): IP камерын дамжуулж буй өгөгдөл нь шифрлэлгүй байвал Man-in-the-Middle (MitM) халдлагад өртөх магадлал өндөр[14] байдаг.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

- CVE эмзэг байдал (Common Vulnerabilities and Exposures - CVE): Камерын системд бүртгэгдсэн CVE эмзэг байдал нь халдагчдад алсын зайнаас нэвтрэх боломж олгодог[15].

IP камерын эсрэг ашиглагдаж болох халдлагын аргууд нь:

- Brute-force халдлага гэдэг нь халдагч тодорхой нэг төхөөрөмжийн нэвтрэх мэдээллийг олон төрлийн нууц үг ашиглан таах оролдлого хийдэг довтолгооны нэг төрөл юм[16]. IP камерын анхдагч эсвэл сул нууц үгтэй (admin: admin гэх мэт) төхөөрөмжүүд халдлагад өртөх өндөр эрсдэлтэй[17].
- Камерын анхдагч нууц үгийг өөрчлөөгүй хэрэглэгчид халдлагад хамгийн их өртдөг[18].
- Brute-force халдлагад "dictionary attack" буюу урьдчилан тодорхойлсон түгээмэл нууц үгийн жагсаалт ашиглах, "credential stuffing" буюу өмнө нь алдагдсан хэрэглэгчийн мэдээллийг ашиглах арга түгээмэл [19] CVE эмзэг байдлыг ашиглах
- IP камерууд нь CVE-2013-3381 зэрэг эмзэг байдлын улмаас гаднаас хянах боломжтой болохыг судалгааны явцад тогтоосон[20].
- SNMP (Simple Network Management Protocol) v2c протокол ашиглан камерын мэдээлэл авах, удирдах боломжтой[21].
- CVE-2013-3381 эмзэг байдал нь IoT (Internet of Things) төхөөрөмжүүдийг зайнаас хянах, удирдах боломж олгодог[15].

Үүнээс гадна, CVE-2021-28372 зэрэг шинэчлэгдээгүй төхөөрөмжүүдэд халдаж болох эмзэг байдал бүртгэгдсэн[16]. MitM халдлага гэдэг нь сүлжээнд дамжиж буй мэдээллийг таслан авч, өөрчлөх эсвэл хулгайлах боломжийг олгодог довтолгооны арга юм[17]. Wireshark зэрэг хэрэгслүүдийг ашиглан сүлжээний өгөгдлийг хянах үед ARP spoofing ашиглан IP камерын дамжуулж буй өгөгдлийг өөрчилж, хуурамч мэдээлэл илгээх боломжтой нь тогтоогдсон[18]. MitM халдлагын нэг төрөл болох ARP poisoning нь камер болон чиглүүлэгчийн хоорондын мэдээллийг өөрчлөх боломжтой[19]. Камерын дамжуулсан бодит дүрсийг хуурамч дүрсээр солих боломжтой, үүнээс шалтгаалан бодит байдлыг гажуудуулж, аюулгүй байдлын асуудал үүсгэж болзошгүй[20]. Botnet халдлага гэдэг нь халдагч сүлжээнд холбогдсон олон тооны төхөөрөмжийг удирдаж, олон төрлийн халдлага үйлдэх боломжтой болох эрсдэлийг хэлнэ[21]. 2016 онд болсон Mirai ботнет халдлага нь олон мянган IP камерыг хакердаж, интернетийн томоохон үйлчилгээ үзүүлэгчдийг унагасан[22]. Mirai нь сул нууц үгтэй IoT төхөөрөмжүүдийг халдварлуулж, DDoS (Distributed Denial of Service) халдлага үйлдэж байсан[23].Mozi ботнет халдлага нь P2P (Peer-to-Peer) сүлжээг ашиглан илүү ухаалаг, хяналтгүй халдлага хийх боломжтой болсон [24]. IP камерууд нь IoT төхөөрөмжүүдийн нэг хэсэг тул ботнет халдлагад хамгийн өртөмтгий төхөөрөмжүүдийн нэг[24]юм.

III. Судалгааны аргачлал

Судалгааг нарийвчлан, системтэй гүйцэтгэхийн тулд тусгайлан бэлтгэсэн, хянагдсан виртуал туршилтын орчинд хийсэн. Аргачлалын үндсэн үе шатууд нь:

a) Судалгааны орчны бүрдүүлэлт

Туршилтыг хязгаарлагдсан, хамгаалагдсан сүлжээний орчинд VM (Virtual Machine) ашиглан Windows 10 үйлдлийн систем дээр гүйцэтгэсэн. Орчны техникийн үзүүлэлт нь Intel Core i7 процессор, 16GB RAM бүхий VMware Workstation орчинд суурилсан. Үүгээр дамжуулан сүлжээний урсгалыг хянах болон системийн өөрчлөлтүүдийг бүртгэх боломжийг бүрдүүлсэн.

b) Судалгаанд ашигласан хэрэгслүүд:

- Wireshark: Сүлжээний урсгалыг бодит хугацаанд хянах, багц мэдээлэл (packet)-ийг задлан шинжлэх, сэжигтэй холбоо, шифрлэлгүй өгөгдөл дамжуулалтыг илрүүлэхэд ашигласан. Deep Packet Inspection (DPI) функц ашиглан өгөгдлийн агуулгыг нарийвчлан судалсан.
- Zenmap (Nmap GUI): Сүлжээний сканнер хийж, холбогдсон төхөөрөмжүүд, нээлттэй портууд, үйлчилгээний төрлийг тодорхойлсон. Penetration тест хийж халдлагын боломжит векторуудыг үнэлсэн.
- RegShot: Windows бүртгэлийн сангийн өөрчлөлтийг хянаж, халдлагын улмаас гарсан системийн тохиргооны өөрчлөлтийг илрүүлэхэд хэрэглэсэн.
- Process Monitor: Бодит цагийн горимд процессын үйл ажиллагаа, файл болон registry access, сүлжээний холболтыг хянаж, сэжигтэй үйлдлийг бүртгэх зорилгоор ашигласан.

Эдгээр хэрэгслүүдийг хавсарсан байдлаар ашигласнаар системд үүссэн өөрчлөлт, сүлжээний эмзэг байдал, өгөгдөл дамжуулалтын сэжигтэй шинж тэмдэг зэргийг бүрэн гүйцэд тодорхойлох боломж бүрдсэн.

c) Судалгааны үндсэн үе шат

- Төлөвлөлт ба тойм судалгаа: IP камер болон NVR төхөөрөмжүүдийн аюулгүй байдлын талаарх олон улсын болон дотоодын судалгааны материалуудыг судалсан.
- Сүлжээний тандалт: Zenmap, Wireshark ашиглан дотоод сүлжээнд холбогдсон төхөөрөмжүүдийн IP хаяг, порт, үйлчилгээний төрлийг тодорхойлсон.
- Эмзэг хаягуудыг илрүүлэх: Халдлагад өртөх магадлалтай IP хаягуудыг тодорхойлж, тэдгээрийн хамгаалалтын түвшинг үнэлсэн.
- Хандалтын тогтвортой байдал: Нэвтэрсний дараа системийн өөрчлөлт, өгөгдлийн урсгал, аюултай холболт зэргийг тогтмол хянасан.

- Хамгаалалтын тохиргооны дүн шинжилгээ: Нээлттэй порт, сул тохиргоо, шинэчлэгдээгүй firmware зэрэг гол эмзэг байдлуудыг задлан судалж, сайжруулах боломжийг үнэлсэн.

IV. Туршилтын үр дүн

Монгол Улсад түгээмэл хэрэглэж байгаа IP камер, NVR төхөөрөмж дээр илэрдэг Wireshark 2.0.1 хэрэгслээр сүлжээгээр дамжиж байгаа пакетууд дээр дүн шинжилгээ хийсэн. Wireshark хэрэгслээр сүлжээний өгөгдлийн урсгалаас халдлагыг илрүүлэхдээ дотоод сүлжээний IP камер, хостын түвшний пакетуудын мэдээлэл дээр тулгуурлан сэжигтэй пакетуудын дата дамжуулалт болон хүлээн авсан протокол дээр дүн шинжилгээ хийж үзэхэд дараах үр дүн гарсан. ARP протоколын тусламжтай гадаад хаяг үүсгэж байгааг зураг 1-д харуулав. ARP протоколын тусламжтай сүлжээний төхөөрөмжүүдийн IP болон MAC хаягийг тодорхойлох оролдлого хийгдэж байгаа нь илэрсэн.

arp

	Time	Source	Destination	Protoc	Lengt	Info
8	2023-03-10 05:56:51.994171	5a:58:11:56:1e:40	Broadcast	ARP	60	Who has 47.252.8.80? Tell 192.168.1.156
9	2023-03-10 05:56:51.994171	5a:58:11:56:1e:40	Broadcast	ARP	60	Who has 47.74.153.98? Tell 192.168.1.156
10	2023-03-10 05:56:51.994171	5a:58:11:56:1e:40	Broadcast	ARP	60	Who has 139.9.5.224? Tell 192.168.1.156
11	2023-03-10 05:56:52.003560	5a:58:d1:a9:a2:e6	Broadcast	ARP	60	Who has 1.0.0.0? Tell 192.168.1.151
12	2023-03-10 05:56:52.090657	5a:58:ad:51:0f:a6	Broadcast	ARP	60	Who has 192.168.2.1? Tell 192.168.1.164
13	2023-03-10 05:56:52.120152	5a:58:ad:51:0f:a6	Broadcast	ARP	60	Who has 139.9.5.224? Tell 192.168.1.164
14	2023-03-10 05:56:52.120152	5a:58:ad:51:0f:a6	Broadcast	ARP	60	Who has 47.74.153.98? Tell 192.168.1.164
15	2023-03-10 05:56:52.120152	5a:58:ad:51:0f:a6	Broadcast	ARP	60	Who has 47.252.8.80? Tell 192.168.1.164
18	2023-03-10 05:56:52.328687	5a:58:57:6a:5e:65	Broadcast	ARP	60	Who has 139.9.5.224? Tell 192.168.1.165
19	2023-03-10 05:56:52.328687	5a:58:57:6a:5e:65	Broadcast	ARP	60	Who has 47.74.153.98? Tell 192.168.1.165
20	2023-03-10 05:56:52.328687	5a:58:57:6a:5e:65	Broadcast	ARP	60	Who has 47.252.8.80? Tell 192.168.1.165
21	2023-03-10 05:56:52.396250	5a:58:bb:87:a5:42	Broadcast	ARP	60	Who has 47.252.8.80? Tell 192.168.1.158
22	2023-03-10 05:56:52.396250	5a:58:bb:87:a5:42	Broadcast	ARP	60	Who has 47.74.153.98? Tell 192.168.1.158
23	2023-03-10 05:56:52.396626	5a:58:bb:87:a5:42	Broadcast	ARP	60	Who has 139.9.5.224? Tell 192.168.1.158
24	2023-03-10 05:56:52.412162	Ubiquiti_59:c0:91	Broadcast	ARP	60	Who has 192.168.1.66? Tell 192.168.1.10
26	2023-03-10 05:56:52.436394	5a:58:a7:04:c8:f7	Broadcast	ARP	60	Who has 139.9.5.224? Tell 192.168.1.155
27	2023-03-10 05:56:52.436394	5a:58:a7:04:c8:f7	Broadcast	ARP	60	Who has 47.74.153.98? Tell 192.168.1.155
28	2023-03-10 05:56:52.436394	5a:58:a7:04:c8:f7	Broadcast	ARP	60	Who has 47.252.8.80? Tell 192.168.1.155
29	2023-03-10 05:56:52.536002	5a:58:c0:97:f4:48	Broadcast	ARP	60	Who has 47.252.8.80? Tell 192.168.1.154
30	2023-03-10 05:56:52.536002	5a:58:c0:97:f4:48	Broadcast	ARP	60	Who has 47.74.153.98? Tell 192.168.1.154
31	2023-03-10 05:56:52.536002	5a:58:c0:97:f4:48	Broadcast	ARP	60	Who has 139.9.5.224? Tell 192.168.1.154
32	2023-03-10 05:56:52.543504	5a:58:11:56:1e:40	Broadcast	ARP	60	Who has 8.8.8.8? Tell 192.168.1.156
33	2023-03-10 05:56:52.750002	Ubiquiti_59:c2:90	Broadcast	ARP	60	Who has 192.168.1.1? Tell 192.168.1.18
36	2023-03-10 05:56:52.970525	Ubiquiti_59:c2:90	Broadcast	ARP	60	Who has 192.168.1.33? Tell 192.168.1.18
37	2023-03-10 05:56:53.006836	5a:58:a7:04:c8:f7	Broadcast	ARP	60	Who has 192.168.2.1? Tell 192.168.1.155
39	2023-03-10 05:56:53.023445	5a:58:11:56:1e:40	Broadcast	ARP	60	Who has 139.9.5.224? Tell 192.168.1.156

Зураг 1. CDP протоколын урсгал Wireshark хэрэгслээр илрэгдэж буй байдал

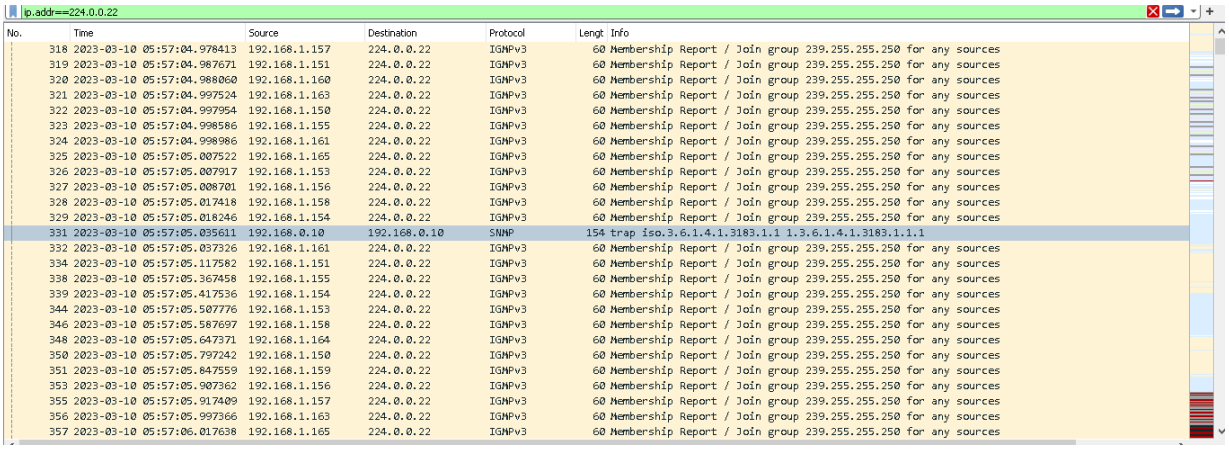
Wireshark ашиглан сүлжээний урсгалыг судлахад ARP (Address Resolution Protocol) ашиглан гаднын сервер рүү холбогдох оролдлого илэрсэн. Зураг 1-д ARP хүсэлт дамжуулж буй байдлыг харуулсан бөгөөд Alibaba US Technology Co., Ltd. домайн руу холбогдож буйг тогтоосон. Дээрх хаягуудын дэлгэрэнгүй мэдээллийг хүснэгт 1-д харуулав. Туршилтаар илэрсэн сүлжээний мэдээллийг Wireshark хэрэгслээр шинжлэхэд дараах IP камерын хаягууд нь Alibaba US Technology Co., Ltd. компанийн домэйн руу холбогдох оролдлого хийж байгааг илрүүлсэн. IP хаяг 192.168.1.164, 192.168.1.156, 192.168.1.155, 192.168.1.154 хаягууд нь бүгд IP камер төхөөрөмжид хамаарагдаж байсан бөгөөд тэдгээр нь ижил компанийн домайн руу сүлжээний ARP протоколын тусламжтайгаар хандаж байжээ. Энэ нь тухайн камеруудаас гаднын сервер руу өгөгдөл дамжуулж болзошгүй байдал үүсч байгааг харуулж байгаа юм.

Wireshark 2.0.2 хэрэгслийг ашиглан сүлжээний өгөгдлийг бодит хугацаанд шинжилж, Cisco Discovery Protocol (CDP)-ийн урсгалуудыг илрүүлсэн байдал. Зурагт харуулснаар Cisco төхөөрөмж (MAC: 51:d8:ae)-ээс CDP/VTP/DTP/PAgP/UDLD чиглэлийн multicast хаяг руу тасралтгүй CDP пакет дамжиж байгаа нь бүртгэгдсэн байна. Зураг 1-д үзүүлсэн сүлжээний урсгалд нь:

- Device ID: SW-58-3-Pnet — төхөөрөмжийн нэр
- Port ID: GigabitEthernet1/0/46 — сүлжээнд холбогдож буй интерфэйс
- Protocol: CDP — Cisco-н төхөөрөмжүүдийн хооронд автомат танилцуулга хийх зориулалттай

Энэхүү CDP урсгал нь authentication шаардаггүй бөгөөд UDP дээр ажилладаг тул гуравдагч этгээдэд сүлжээний бүтэц, төхөөрөмжийн байршил, интерфейсийн нэр зэрэг нэн чухал мэдээллийг задруулах өндөр эрсдэлтэй. Энэ нь probe төрлийн халдлагын үндсэн мэдээлэл цуглуулах шатанд халдагч этгээд ашиглахад нээлттэй орчныг бүрдүүлдэг.

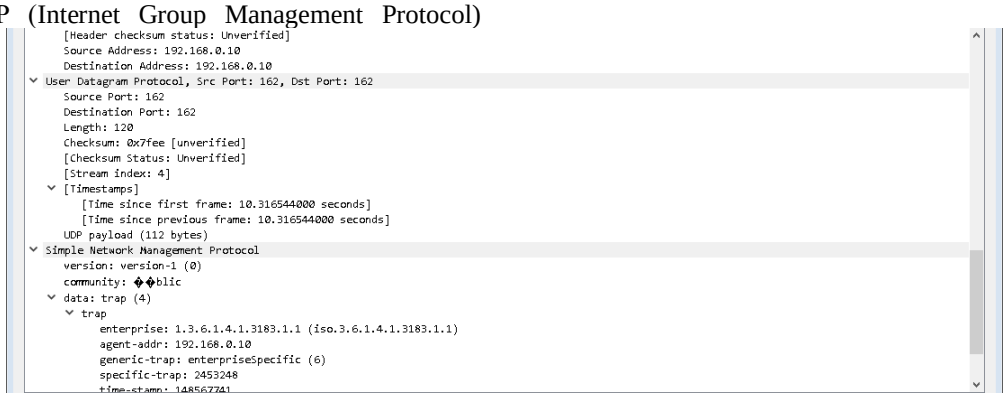
IP хаягуудыг нарийвчилсан байдлаар судалж үзэхэд ARP протоколын тусламжтай үйлдвэрлэгч улс руу холболт тогтоох хайлт хийж байна. IP 192.168.0.10 хаягнаас порт 162 руу өгөгдөл дамжуулж буйг илрүүлсэн. Судалгааны үр дүнд уг порт нь CVE-2013-3381 эмзэг байдлыг ашиглаж, зайнаас халдлага үйлдэх боломжтой болохыг тогтоосон. Энэ алдаа нь хост болон IoT төхөөрөмжүүдийн хамгаалалтыг тойрон гарах боломжийг олгодог.



No.	Time	Source	Destination	Protocol	Lengt	Info
318	2023-03-10 05:57:04.978413	192.168.1.157	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
319	2023-03-10 05:57:04.987671	192.168.1.151	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
320	2023-03-10 05:57:04.988060	192.168.1.160	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
321	2023-03-10 05:57:04.997524	192.168.1.163	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
322	2023-03-10 05:57:04.997954	192.168.1.150	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
323	2023-03-10 05:57:04.998586	192.168.1.155	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
324	2023-03-10 05:57:04.998986	192.168.1.161	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
325	2023-03-10 05:57:05.007522	192.168.1.165	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
326	2023-03-10 05:57:05.007917	192.168.1.153	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
327	2023-03-10 05:57:05.008701	192.168.1.156	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
328	2023-03-10 05:57:05.017418	192.168.1.158	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
329	2023-03-10 05:57:05.018246	192.168.1.154	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
331	2023-03-10 05:57:05.035611	192.168.0.10	192.168.0.10	SNMP	154	trap iso.3.6.1.4.1.3183.1.1 1.3.6.1.4.1.3183.1.1.1
332	2023-03-10 05:57:05.037326	192.168.1.161	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
334	2023-03-10 05:57:05.117582	192.168.1.151	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
338	2023-03-10 05:57:05.367458	192.168.1.155	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
339	2023-03-10 05:57:05.417536	192.168.1.154	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
344	2023-03-10 05:57:05.507776	192.168.1.153	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
346	2023-03-10 05:57:05.587697	192.168.1.158	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
348	2023-03-10 05:57:05.647371	192.168.1.164	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
350	2023-03-10 05:57:05.797242	192.168.1.150	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
351	2023-03-10 05:57:05.847559	192.168.1.159	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
353	2023-03-10 05:57:05.907362	192.168.1.156	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
355	2023-03-10 05:57:05.917409	192.168.1.157	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
356	2023-03-10 05:57:05.997366	192.168.1.163	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources
357	2023-03-10 05:57:06.017638	192.168.1.165	224.0.0.22	IGMPv3	60	Membership Report / Join group 239.255.255.250 for any sources

Зураг 2. Дээрх IP хаягууд мультист групт бүртгүүлэх хэсэг

Зураг 2-д харуулсан IP 224.0.0.22 хаягийн тусламжтай IP камер, хостуудыг IP 239.255.255.250 хаяг уруу холбож байна. Хаяг дундаас IP 192.168.0.10 хаягийн өгөгдлийн нарийвчилсан мэдээллийг зураг 2-д харуулав. 224.0.0.22 хаяг нь IGMP (Internet Group Management Protocol) ашиглаж, IP камер, NVR төхөөрөмжүүдийг олон хосттой холбох үйл ажиллагааг явуулж байна. Ийм үйлдэл нь халдлага үйлдэхэд ашиглагдаж болох боломжит эмзэг байдал үүсгэж буйг тодорхойлсон.

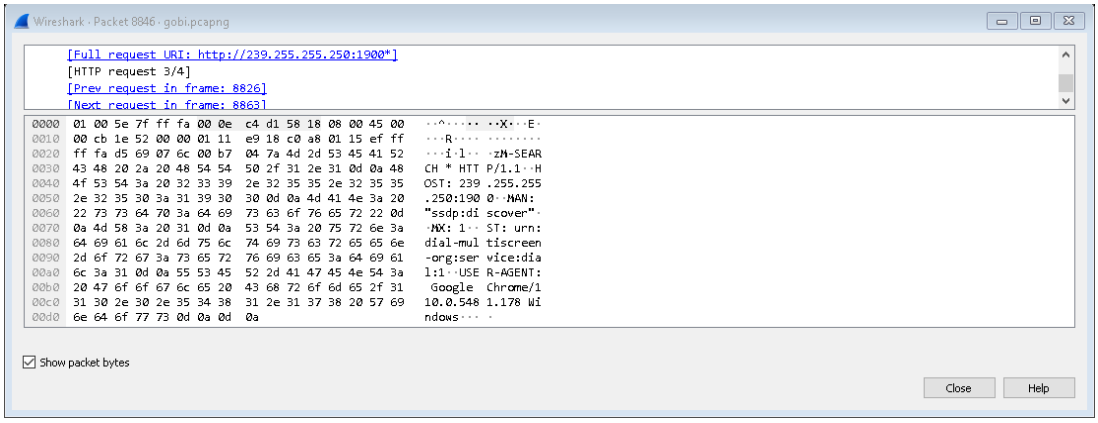


```
    [Header checksum status: Unverified]
    Source Address: 192.168.0.10
    Destination Address: 192.168.0.10
  User Datagram Protocol, Src Port: 162, Dst Port: 162
    Source Port: 162
    Destination Port: 162
    Length: 120
    Checksum: 0x7fee [unverified]
    [Checksum Status: Unverified]
    [Stream index: 4]
    [Timestamps]
      [Time since first frame: 10.316544000 seconds]
      [Time since previous frame: 10.316544000 seconds]
    UDP payload (112 bytes)
  Simple Network Management Protocol
    version: version-1 (0)
    community: ××blic
    data: trap (4)
      trap
        enterprise: 1.3.6.1.4.1.3183.1.1 (iso.3.6.1.4.1.3183.1.1)
        agent-addr: 192.168.0.10
        generic-trap: enterpriseSpecific (6)
        specific-trap: 2453248
        time-stamp: 148567741
```

Зураг 3. IP 192.168.0.10 хаягнаас порт 162 уруу өгөгдөл илгээсэн байдал

Үүнд:

Зураг 2-д харуулсан IP 192.168.0.10 хаягнаас порт 162 уруу өгөгдөл илгээж байгааг илрүүлсэн. Уг порт 162 судалж үзэхэд CVE-2013-3381 алдааны код илэрсэн. Уг алдааны код нь зайнаас холболт тогтоож хост болон IoT төхөөрөмжийн эмзэг сул тал уруу халдлага хийдэг. IP 239.255.255.250 хаяг нь порт 1900 аар өгөгдөл дамжуулж байгааг зураг 3-д харуулав. Wireshark анализын явцад IP 192.168.0.10 хаягаас порт 162 руу өгөгдөл илгээж буйг илрүүлсэн. Уг портыг шинжилж үзэхэд CVE-2013-3381 эмзэг байдал илэрсэн. CVE-2013-3381 алдааны код нь зайнаас холболт тогтоох боломжийг олгодог. Энэ эмзэг байдал IoT төхөөрөмжүүд болон хост компьютеруудын хамгаалалтыг тойрон гарах боломжтой. IP 192.168.0.10 нь сүлжээний бусад төхөөрөмжүүдтэй холболт үүсгэж, сүлжээнд халдлага үйлдэх эрсдэлийг бий болгож байна.



```
Wireshark · Packet 8846 · gobi.pcapng

    [Full request URI: http://239.255.255.250:1900*]
    [HTTP request 3/4]
    [Prev request in frame: 8826]
    [Next request in frame: 8863]

0000  01 00 5e 7f ff fa 00 0e  c4 d1 58 18 08 00 45 00   ··^····· ··X···E·
0010  00 cb 1e 52 00 00 01 11  e9 18 c0 a8 01 15 ef ff   ···R···· ********
0020  ff fa d5 69 07 6c 00 b7  04 7a 4d 2d 53 45 41 52   ···i·l·· ·zM-SEAR
0030  43 48 20 2a 20 48 54 54  50 2f 31 2e 31 0d 0a 48   CH * HTT P/1.1··H
0040  4f 53 54 3a 20 32 33 39  2e 32 35 35 2e 32 35 35   OST: 239 .255.255
0050  2e 32 35 30 3a 31 39 30  30 0d 0a 4d 41 4e 3a 20   .250:190 0··MAN:
0060  22 73 73 64 70 3a 64 69  73 63 6f 76 65 72 22 0d   "ssdp:di scover"·
0070  0a 4d 58 3a 20 31 0d 0a  53 54 3a 20 75 72 6e 3a   ·MX: 1·· ST: urn:
0080  64 69 61 6c 2d 6d 75 6c  74 69 73 63 72 65 65 6e   dial-mul tiscreen
0090  2d 6f 72 67 3a 73 65 72  76 69 63 65 3a 64 69 61   -org:ser vice:dia
00a0  6c 3a 31 0d 0a 55 53 45  52 2d 41 47 45 4e 54 3a   l:1··USE R-AGENT:
00b0  20 47 6f 6f 67 6c 65 20  43 68 72 6f 6d 65 2f 31    Google  Chrome/1
00c0  31 30 2e 30 2e 35 34 38  31 2e 31 37 38 20 57 69   10.0.548 1.178 Wi
00d0  6e 64 6f 77 73 0d 0a 0d  0a                        ndows··· ·

Show packet bytes

Close   Help
```

Зураг 4. . IP 239.255.255.250 хаяг дээр дүн шинжилгээ хийсэн үр дүн

Зураг 4- д харуудсан IP 239.255.255.250 SSDP протокол нь DDoS халдлагыг далд дэмжиж байна. Wireshark анализын үр дүнд IP 239.255.255.250 хаяг нь порт 1900 ашиглан өгөгдөл дамжуулж буйг илрүүлсэн. Судалгааны явцад дараах чухал дүгнэлтүүд хүрсэн. Үүнд:

- Камеруудын аюулгүй байдал сул хэвээр байна.
- Анхдагч нууц үг өөрчлөгдөөгүй камерууд халдлагад өртөх өндөр эрсдэлтэй байна.
- CVE-2013-3381 зэрэг эмзэг байдал ашиглагдвал камерын удирдлагыг гаднаас авах боломжтой.
- ARP spoofing болон MitM (Man-in-the-Middle) халдлага хийж, камерын дүрсийг өөрчлөх боломжтой.
- Судалгааны туршилтаар сүлжээнд дараах аюул илэрсэн.
- IP 192.168.0.10 хаягнаас порт 162 руу өгөгдөл илгээж буйг илрүүлсэн бөгөөд CVE-2013-3381 эмзэг байдалтай болох нь батлагдсан.
- IP 239.255.255.250 SSDP протоколыг ашиглан DDoS халдлага үйлдэх боломжтой байна.
- Wireshark хэрэгслээр сүлжээний урсгалыг хянах үед гаднын сервертэй холбогдох оролдлого гарч байгаа нь тогтоогдсон.
- IP камерын хамгаалалтыг сайжруулах шаардлагатай. Судалгааны үр дүнд IP камерын аюулгүй байдлыг сайжруулах дараах арга хэмжээг авах шаардлагатай гэдэг нь тодорхой болов. Үүнд:
- Анхдагч нууц үг солих, бат бөх нууц үг хэрэглэх.
- Программ хангамжийн шинэчлэл (firmware update)-ийг тогтмол хийх.
- Галт хана (firewall) ашиглаж, зөвхөн шаардлагатай портуудыг нээлттэй байлгах.
- Сүлжээний урсгалыг Wireshark, Zenmap ашиглан тогтмол хянах.
- VPN болон шифрлэлтэй холболт ашиглах замаар өгөгдлийг хамгаалах.
- Камерын аюулгүй байдлыг бүрэн хангахгүй бол кибер халдлагад өртөх эрсдэл өндөр байна. IP камерууд нь зөвхөн дүрс дамжуулах биш, байгууллагын сүлжээний нэг хэсэг болсноор тухайн байгууллагын сервер, өгөгдлийн санд халдах гарц болж болохыг судалгааны үр дүн харууллаа. Кибер халдагчид сул хамгаалалттай IP камер ашиглан байгууллагын дотоод сүлжээ рүү нэвтрэх боломжтой нь батлагдлаа.

V. Дүгнэлт

Энэхүү судалгаагаар CCTV камер, IP камерын аюулгүй байдал, тэдгээрийн сүлжээний эмзэг байдал, халдлагын боломжууд, хамгаалалтын арга хэмжээг шинжилж, бодит кейс дээр тулгуурлан дүгнэлт хийсэн. Камерууд нь анх гэмт хэргээс урьдчилан сэргийлэх, хамгаалалт хийх зорилготой байсан боловч орчин үед интернетэд холбогдож, ухаалаг системтэй болж хөгжсөн нь кибер халдлагад өртөх эрсдэлийг нэмэгдүүлж байна.

Судалгааны үр дүнд Монгол Улсад ашиглагдаж буй IP камер, NVR төхөөрөмжүүдийн хамгаалалт сул бөгөөд Wireshark, Zenmap, Process Monitor, RegShot зэрэг хэрэгслүүдийг ашиглан сүлжээний эмзэг байдал, сэжигтэй өгөгдөл дамжуулалтыг хянах боломжтой болохыг тогтоосон.

Энэхүү судалгаа нь IP камер, NVR төхөөрөмжүүдийн кибер аюулгүй байдлын асуудлуудыг тодорхойлох, судлах, хамгаалалтын арга хэмжээг боловсруулахад чухал хувь нэмэр оруулж байна. Өнөөгийн дижитал эринд байгууллагууд болон хувь хүмүүс CCTV/IP камерын аюулгүй байдлыг хангах бодлого боловсруулах, CVE мэдээллийг тогтмол шалгах, хамгаалалтын системээ шинэчлэх шаардлагатай. IP камерууд нь зөвхөн хяналтын хэрэгсэл бус, байгууллагын сүлжээний хамгаалалтын нэгэн чухал бүрэлдэхүүн хэсэг болсон байна. Иймээс тэдгээрийн аюулгүй байдлыг хангахын тулд дараах арга хэмжээг яаралтай хэрэгжүүлэх шаардлагатай гэж үзлээ. Үүнд: анхдагч нууц үгийг солих, программ хангамжийг тогтмол шинэчлэх, зөвшөөрөлгүй портын хандалтыг хаах, өгөгдлийг шифрлэх, сүлжээний хяналт тогтмол хийх зэрэг орно.

SSDP (Simple Service Discovery Protocol) нь IoT болон ухаалаг төхөөрөмжүүдийг автоматаар холбох үүрэгтэй протокол юм. Энэ протоколыг ашиглан DDoS халдлага үйлдэх боломжтой. Зорилтот халдлага нь Google Chrome, Microsoft Edge вэб хөтөч уруу чиглэж байгааг тодорхойлсон.

Ном зүй

[1] G. Uuganbayar, A. Yautsiukhin, and F. Martinelli, “Optimisation of cyber insurance coverage with selection of cost effective security controls,” *Computers & Security*, vol. 108, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167404820303941

[2] B. Schneier, *Click Here to Kill Everybody: Security and Survival in a Hyper-connected World*, W.W. Norton & Company, 2018.

[3] S. Konecka and Z. Bentyn, “Cyberattacks as threats in supply chains,” *European Research Studies Journal*, vol. 27, no. 3, 2024. [Online]. Available: https://www.um.edu.mt/library/oar/handle/123456789/127946 Cisco Systems Inc., “About Cisco – Company Overview,” Cisco Official Website. Accessed: Apr. 8, 2025. [Online]. Available: https://www.cisco.com/c/en/us/about.html

[4] Cisco Systems Inc., “Cisco Products and Services,” Cisco Product Portfolio. Accessed: Apr. 8, 2025. [Online]. Available: https://www.cisco.com/c/en/us/products/index.html

[5] Cisco Systems Inc., “Security Advisories,” Cisco Security Center. Accessed: Apr. 8, 2025. [Online]. Available: https://tools.cisco.com/security/center/publicationListing.x

[6] Shodan, “Search results for Cisco devices in Mongolia,” Apr. 2025. Accessed: Apr. 8, 2025. [Online]. Available: https://www.shodan.io/search?query=cisco+country%3A%22MN%22

[7] CVE – Common Vulnerabilities and Exposures, “CVE-2023-39248,” National Vulnerability Database, U.S. NIST. Accessed: Apr. 8, 2025. [Online]. Available: https://nvd.nist.gov/vuln/detail/CVE-2023-39248

[8] S. Bhatia, A. Kumar, and D. R. Rawat, “Security Analysis of CDP and LLDP Protocols in Modern Networks,” *International Journal of Computer Applications*, vol. 184, no. 31, 2022. [Online]. Available: https://doi.org/10.5120/ijca2022922366

[9] A. K. Soni, R. Sharma, and P. Singh, “Evaluating Vulnerabilities in VRRP and Mitigating DoS Attacks,” *IEEE Access*, vol. 10, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9723412

[10] N. Jacobson and H. Lin, “Probing Discovery Protocols: WS-Discovery and the New Threat Surface,” *Journal of Cybersecurity Technology*, vol. 6, no. 2, 2023. [Online]. Available:

https://www.tandfonline.com/doi/full/10.1080/23742917.2023.1976542

[11] H. Jin, G. H. Jeon, H. W. A. Choi, S. Jeon, and J. T. Seo, "A threat modeling framework for IoT-Based botnet attacks," *Heliyon*, vol. 10, 2024. [Online]. Available: https://www.cell.com/heliyon/fulltext/S2405-8440(24)15223-1

[12] [13] D. N. Gupta and R. Kumar, "Networking in IoT: Technologies Usage, Security Threats, and Possible Countermeasures," *Smart Wireless Computing and Communication*, vol. 10, no. 1, 2021. [Online]. Available: https://app.scholarai.io/paper?paper_id=DOI:10.2174/2210327910666210106090826

[13] [14] B. A. Sukhni, S. K. Manna, J. M. Dave, and L. Zhang, "Extracting Optimal Number of Features for Machine Learning Models in Multilayer IoT Attacks," *Sensors*, vol. 24, no. 24, 2024. [Online]. Available: https://www.mdpi.com/1424-8220/24/24/8121

[14] [15] B. Al Sukhni, S. K. Manna, J. M. Dave, and L. Zhang, "Semi-Automated Intrusion Detection System (Saids) for Identifying Multilayer IoT Attacks," *SSRN preprint*, 2024. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4941479

[15] [16] D. N. Gupta and R. Kumar, "Networking in IoT Technologies: Usage, Security Threats, and Countermeasures," *International Journal of Sensors, Wireless Communications and Control*, vol. 10, 2020. [Online]. Available: https://www.researchgate.net/publication/348332830

[16] [17] J. Javed, M. Sheraz, and M. Asif, "A secure mechanism for detecting and mitigating DDoS attacks in IoT-enabled surveillance networks," *Journal of Network and Computer Applications*, vol. 229, 2024. [Online]. Available: https://doi.org/10.1016/j.jnca.2024.103644

[17] [18] S. Rahman, K. Salah, and M. Aloqaily, "IoT-Cam: An Intelligent Framework for Securing IP Cameras Against Zero-Day Attacks," *Future Generation Computer Systems*, vol. 144, 2023. [Online]. Available: https://doi.org/10.1016/j.future.2023.02.027

[18] [19] A. Sikder, H. Aksu, and A. S. Uluagac, "A Survey on Sensor-based Threats to IoT Devices and Applications," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9935256

[19] [20] A. Alrawi, C. Lever, M. Antonakakis, and F. Monrose, "SoK: Security Evaluation of Home-Based IoT Deployments," *NDSS Symposium 2019*. [Online]. Available: https://www.ndss-symposium.org/ndss-paper/sok-security-evaluation-of-home-based-iot-deployments

[20] [21] A. Nayyar and A. Singh, "A Review on Mozi Botnet and the Evolution of Botnet-Based IoT Attacks," *IEEE Access*, vol. 10, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9568921

[21] [22] T. Tunc, I. Kara, and M. Gunes, "Detecting ARP Spoofing Attacks Using Wireshark and Deep Packet Inspection," *Computer Standards & Interfaces*, vol. 87, 2023. [Online]. Available: https://doi.org/10.1016/j.csi.2023.103681

[22] [23] A. Kumar, B. Sharma, and S. Singh, "Evaluation of Cybersecurity Threats and Solutions in IP-Based Surveillance Systems," *International Journal of Security and Networks*, vol. 17, no. 3, 2022. [Online]. Available: https://www.inderscience.com/info/inarticle.php?artid=123456

[23] [24] K. Sivanathan, H. H. Gharakheili, and V. Sivaraman, "Classifying IoT Devices in Smart Homes Using Network Traffic Characteristics," *ACM Transactions on Internet Technology*, vol. 23, no. 2, 2023. [Online]. Available: https://dl.acm.org/doi/10.1145/3548606.3560617

[24] [25] A. Vattapparamban, M. Alazab, and S. Garg, "AI-Based Intrusion Detection System for IoT Devices," *Computers & Security*, vol. 117, 2022. [Online]. Available: https://doi.org/10.1016/j.cose.2022.102734

ЭМИЙН ЖОР БИЧЭЭЧ УХААЛАГ СИСТЕМ

Үүрцайхын Мөнхтуяа
Мэдээллийн технологийн салбар
ШУТИС, МХТС
Монгол Улс, Улаанбаатар хот
munkht46@gmail.com

Гантулгын Магван-Эрдэнэ
Компьютерын ухааны салбар
ШУТИС, МХТС
Монгол Улс, Улаанбаатар хот
mgldj13@gmail.com

Батжаргалын Долгорсүрэн
Мэдээллийн технологийн салбар
ШУТИС, МХТС
Монгол Улс, Улаанбаатар хот
b.dolgorsuren@must.edu.mn

***Хураангуй*: Монгол улсын эрүүл мэндийн байгууллагууд, эмчээс өвчтөнд зориулж бичдэг эмийн жорыг механикаар буюу гар аргаар үйлдэж, цаасан хэлбэрээр иргэдэд үйлчилдэг. Энэ нь эмч, эмнэлгийн ажилтнуудын ачааллыг нэмэгдүүлэх, жорыг хаях, үрэгдүүлэх зэргээс шалтгаалан шаардлагатай эмийг авч чадахгүй байх, мөн цаасны хэрэглээ нэмэгдсэнээр байгаль орчинд сөрөг нөлөө үзүүлэх зэрэг олон сул талтай. Иймээс энэхүү судалгаагаар хиймэл оюун ухаанд суурилсан дэвшилтэт технологийг ашиглан уг үйл явцыг автоматжуулах, оновчтой шийдэл бүхий ухаалаг системийг боловсруулсан. Уг систем нь эмчийн үзлэгийн яриаг хүлээн авч, эмийн тун хэмжээ, бэлтгэх арга, яаж хэрэглэхийг заасан эмийн жорыг MNS 5376:2016 стандартын дагуу гаргах юм. Үүний үр дүнд боловсруулагдсан эмийн жорыг эмийн сангийн систем болон E-Mongolia системрүү илгээх, ингэснээр иргэд, эмийн санч, эмнэлэг эм олгох үйл явцыг хянах боломжийг судалсан болно. Тус сэдвийн хүрээнд бусад улс орнуудад ашиглагддаг ижил төстэй системүүдийг судалж, түүнд чиглэсэн оновчтой үр дүнг боловсруулаа.**

***Түлхүүр үг:* цахим жор, эм зүйн салбар, whisper, pyannote, LLM**

I. Удиртгал

Хиймэл оюун ухаан (AI) нь эрүүл мэнд болон анагаахын шинжлэх ухааны салбарт инновац, өөрчлөлтийн шинэ эрин үеийг эхлүүлж байна. “AI” технологи ашигласнаар уламжлалт, хүний оролцоонд тулгуурласан үйл ажиллагааг автоматжуулах боломжтой бөгөөд цаашид илүү ухаалаг системүүд бүтээгдсээр байна [1]. Хиймэл оюун ухааны чухал бүрэлдэхүүн хэсгийн нэг болох ухаалаг танилтын технологи нь үндсэндээ хоёр үндсэн төрөлд ангилагддаг. Үүнд дүрс (image) таних болон яриа (speech) таних технологиуд багтдаг [13,14]. Эдгээр технологи нь эрүүл мэндийн салбарт цаг хугацааг үр ашигтай ашиглах, мэдээллийн үнэн зөв, бүрэн бүтэн байдлыг хангах, мөн эрүүл мэндийн үйлчилгээний чанар, хүртээмжийг сайжруулах зэрэг олон талт боломжийг бүрдүүлж байна. Монгол улсад үйл ажиллагаагаа явуулдаг явуулж буй ихэнх эрүүл мэндийн байгууллагууд эмийн жорыг эмч, мэргэжилтнүүд механикаар буюу гар аргаар бичиж, цаасан хэлбэрээр ард иргэдэд үйлчилдэг. Энэ нь алдаа гарах, хуурамч жор ашиглах, эсвэл мэдээлэл хаягдах зэрэг эрсдлийг дагуулдаг.[4] Иймээс энэхүү үйл явцыг хиймэл оюун ухаан, машин сургалт, гүн сургалтын аргуудыг ашиглан ухаалаг веб системийн оновчтой шийдлийг боловсруулан шийдвэрлэхээр зорилоо. Энэ нь эмчийн ажлын ачааллыг бууруулах, эмийн жорыг хуурамчаар үйлдэх, хаяж гээгдүүлэх зэрэг асуудлыг шийдвэрлэхэд чухал хувь нэмэр оруулах боломжтой юм. [3].

Судалгааны зорилго: Энэхүү судалгааны ажлын зорилго нь эмчийн үзлэгийн яриаг боловсруулж, тогтсон стандарт, загварын дагуу өвчтөнд зориулсан эмийн жорыг автоматаар үүсгэх, хиймэл оюун ухааны яриа танилтын дэвшилтэт технологид суурилагдсан веб системийн оновчтой шийдэл боловсруулан, уг системийг анагаах ухааны салбарт нэвтрүүлэх боломжийг судлахад оршино.

Судалгааны зорилт:

1. Яриаг бичвэрт хөрвүүлэх, ярьж буй субьектуудыг таних, эмийн жорыг стандарт, шаардлагын дагуу боловсруулахад шаардлагатай онолын судалгааг хийнэ.
2. Өгөгдөл цуглуулж, бэлтгэнэ.
3. Туршилт, хөгжүүлэлт хийнэ.
4. Алдааны хувийг тооцоолно.

II. СУДАЛГААНЫ ХЭСЭГ

2.1 Ижил төстэй системийн судалгаа

Эрүүл мэндийн салбарын эмч, мэргэжилтнүүдээс өвчтөнд зориулан бичдэг эмийн жорыг хиймэл оюун ухааны дэвшилтэт технологийн тусламжтайгаар автоматжуулах нь эмчийн ачааллыг бууруулах, цаасан эмийн жорны хэрэглээг багасгаж, дижитал хэлбэрт шилжүүлэх системүүдийг дэлхийн улс орнууд хэрэглээндээ амжилттай нэвтрүүлээд байна[8,10]. Тухайлбал АНУ, Их Британи, Швед зэрэг европын ихэнх орнууд эм олгох цахим системийг ашигладаг байна. Тэдгээрийн ажиллах зарчим, ашигласан технологиудыг судалсны үндсэн дээр үр дүнг хүснэгт 1-т оруулаа.

Хүснэгт 1. Төстэй системийн судалгааны үр дүн

Улс	АНУ	Их Британи	Швед
Систем	ERX Network	Electronic Prescription Service (EPS)	ePrescription
Ажиллах зарчим	Эмч нар эмийн жорыг цахимаар бичиж, эмийн сан руу илгээдэг	Эмч нар эмийн жорыг цахимаар үүсгэж, өвчтөн сонгосон эмийн сангаасаа эмээ авдаг	Өвчтөн эмийн сангаас иргэний үнэмлэхээрээ эмээ авдаг
Ашигласан технологи	**NLP - Graph-based AI**: Эмийн харилцан үйлчлэлийг шинжлэх	**NLP**: Эмийн жорын өгөгдлийг боловсруулах - **Predictive Analytics**: Өвчтөний эмийн хэрэглээг таамаглах	**Machine Learning**: Эмийн жорын хэв маягийг шинжлэх **Anomaly Detection**: Хуурамч жор илрүүлэх

2.2 Монгол улсын эм зүйн салбарт ашиглагддаг програм хангамжуудтай холбогдох боломжийн судалгаа

Монгол улсын хувьд E-Mongolia төрийн нэгдсэн мэдээлэл, үйлчилгээний системд байршсан эмийн жор цэс нь одоогийн байдлаар ажиллагаагүй байна [21]. Эрүүл мэндийн салбарт үйл ажиллагаа явуулж буй эмийн сангууд нь ORION PHARM [20], MEDACCOUNT[18], MPharmaPOS, Smart Pharmacy [19] зэрэг програм хангамжуудыг ашигладаг байна. Эдгээр системүүдийн ажиллах орчин, ашигласан технологи, давуу, сул талуудыг судлан хүснэгт 2-т харуулав. Судалгааны үр дүнд эмийн жор бичээч ухаалаг системд эмийн жор нь тухайн өвчтөний регистрийн дугаараар бүртгэгдэж эмийн жорыг бүртгэж, дээрх програм хангамжуудад холбогдох боломжтой.[10]

Хүснэгт 2. Эмийн санд ашиглагддаг системүүдийн судалгаа [10,18,19,20]

Систем	ITPark-ORION PHARM	Andromed aMEDAC COUNT	MindGolia – Smart Pharmacy	MindGolia - MPharma POS
Ажиллах орчин	Desktop - Windows	Desktop - Windows	Desktop - Windows	Desktop - Windows
Ашигласан технологи	C#, .NET Framework, SQL Server	Java, PostgreSQL	JavaScript	Python, Django, SQLite
Давуу тал	ЭМДСангаас хөнгөлөлт үзүүлэх 200 эмийн бүртгэлтэй.	Агуулахын бүртгэл хийх боломжтой.	Олон салбартай эмийн сангууд ашиглахад тохиромжтой.	ЭМД-ын цахим системтэй холбогдсон.
Сул тал	Эмийг жороор олгохдоо иргэний мэдээллийн үнэн зөв эсэхийг шалгадагтүй.	Эмийн бүртгэлийг гар аргаар, өөрсдөө хийдэг.	Жороор олгох эмийн бүртгэл огт хийдэггүй.	Эмийн жорны бүртгэл огт байхгүй.

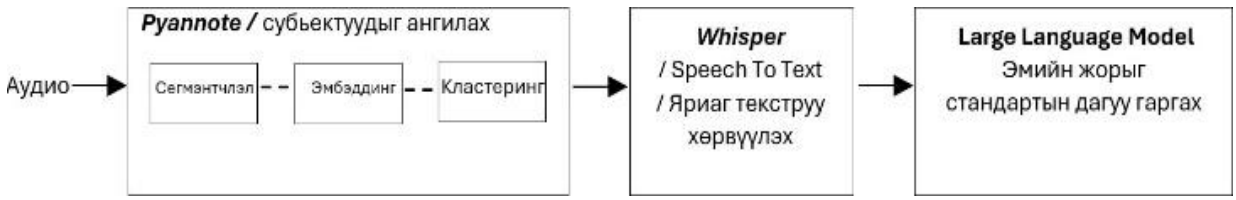
III. АРГАЧЛАЛ

Хиймэл оюун ухааны ухаанд суурилсан машин сургалт, гүн сургалтын фреймворкууд дээр үндэслэн дуу хоолойг текстрүү хөрвүүлэх, ярьсан субъектуудыг ялгах, текст боловсруулах олон төрлийн технологиуд хөгжсөн байдаг.[1] Үүнээс хамгийн өргөн ашиглагддаг Google Text To Speech, Kaldi, DeepSpeech, IBM, Whisper, Pyannote, LLM гэх мэт сангуудыг тодорхой үзүүлэлтүүдээр харьцуулсны үндсэн дээр хамгийн оновчтойг нь сонгон, бэлтгэсэн өгөгдлөөр сургаж, бодит бүтээгдэхүүн болгон хөгжүүлэх нь энэхүү судалгааны ажлын үндсэн тулгуур чиглэл болно. Харьцуулсан судалгааны үр дүнг Хүснэгт 3-т оруулаа.[3,5,8,13] Үүний үр дүнд яриаг текстрүү хөрвүүлэхэд хамгийн бага алдааны хувьтай, мөн 100 гаруй хэл дээр хөрвүүлэлт хийх боломжтой whisper сан, яриа танилт болон ярианы хэлбэр таних (speaker diarization) зэрэг дууны дохио боловсруулахад чиглэгдсэн. Pyannote сан, эмийн жорыг тогтсон стандарт шаардлагын дагуу боловсруулахад LLM моделиудыг ашиглахаар сонгосон болно. [8,13,15] Зураг 1-т системийн ажиллах архитекторын диаграм харуулав.[1,13]

Хүснэгт 3. Whisper санг бусад яриа танилтын сангуудтай харьцуулсан хүснэгт[3,5,8,14]

Үзүүлэлт	Whisper	Google Speech To Text	DeepSpeechMozilla	Kaldi	IBM Watson
Олон хэл дээр ажиллах чадвамж	100 гаруй хэл, Монгол хэл дээр сурах боломжтой	70 гаруй хэл	Зөвхөн англи хэл	20 гаруй хэл	10 гаруй хэл
Яриа танилтын сантай хамтран ажиллах чадвар	Pyanote, DeepSeek зэрэгтэй холбох боломжтой	Google AI – тай хамтран ажилладаг	Бусад сангуудтай холбох боломжгүй, бие даасан загвар	Яриа танилт дэмждэг	IBM AI ашигладаг
Үнэ төлбөр	Нээлттэй, үнэ төлбөргүй	Төлбөртэй API	Үнэгүй	Үнэгүй	Төлбөртэй
Урьдчилан боловсруулалгдсан эсэх	Pre-trained model, fine-tuning хийх боломжтой	Google API модель ашигладаг	Тохируулга шаардлагатай	Гараар тохируулга хийдэг	AI модель ашигладаг
Алдааны хувь	5-10%	4-10%	15-25%	10-20%	6-12%

Зураг 1. Системийн архитектур



3.1 Өгөгдлийн багц

Бид судалгааны ажилдаа дараах өгөгдлийн багцуудыг бэлтгэн , ашигласан болно. Үүнд:

- Монгол улсад бүртгэлтэй 4788 эмийн худалдааны нэршил, улсын нэршил , тун хэмжээ, эмийн мэдээллийг багтаасан өгөгдөл [16,20].
- Эмч, эрүүл мэндийн ажилтнуудын эмийн жорыг тайлбарласан нийт 68 цагийн аудио өгөгдөл.

3.2 Өгөгдлийн урьдчилсан боловсруулалт

Бидний цуглуулан бэлтгэсэн өгөгдлийн багц нь “Монгол Эм Импекс Концерн” болон Монголын Гаалийн Ерөнхий Газраас авсан бүртгэл, мэдээлэлд тулгуурласан тул хоосон утга болон давхардсан бичлэг байгаагүй. Гэсэн хэдий ч өгөгдлийг боловсруулж, ангилах шаардлага гарсан бөгөөд дараах алхмуудыг хийсэн болно.[6]

3.2.1. Аудио өгөгдлийн цэвэрлэгээ:

- Хүн хоорондын ярианы өгөгдөлд агуулагдсан “аа”, “ммм” зэрэг сул үгс болон дуу чимээг шүүж, цэвэршүүлсэн.
- Субъект бүрийн яриаг тусад нь сегментчилан ялгаж, боловсруулалтад бэлтгэсэн.[3,7]

3.2.2. Эмийн өгөгдлийн шошгололт:

- Эмийг зориулалтаар нь ангилан (жишээлбэл: антибиотик, элэг, мэдрэл, нойрны гэх мэт) гараар шошголосон. **[5.8]**

3.2.3. Жорын бүтэц:

- Эмийн жоруудыг MNS 5376:2016 стандартад нийцүүлэн боловсруулж, бүтэцлэн зохион байгуулсан

3.2.4. Тэмдэгт ба формат засвар:

- Эмийн нэр, үйлдвэрлэгч, тунгийн мэдээлэл зэрэг талбаруудад бичвэрийн формат жигд бус байсан тул бүх өгөгдлийг жижиг үсгээр хөрвүүлж, цэвэрлэн нэгэн ижил форматад оруулсан.
- Илүүдэл хоосон зай, тэмдэгт зөрчил, олон янзын бичвэрийн тэмдэглэгээг стандартчилан засварласан[13]

3.3 Pyannote /Ярьж буй субьектийг ангилах/

Пианнот бол Pytorch машин сургалтын фреймворк дээр тулгуурлан ажилладаг Python хэл дээр бичигдсэн яриа танилтийн сан юм. Үүнийг аудио өгөгдлөөс эмч болон өвчтөнийг танихад ашиглана. Ингэснээр эмийн жортой холбоотой хэсгийг ялгаж, стандартын дагуу текст болгоно. Энэ үйл явц нь сегментчлэл, эмбеддинг, кластеринг гэсэн үе шатуудтай.[5,7]

3.3.1 Pyannote-Сегментчлэл (Segmentation)

Сегментчлэл нь аудио бичлэгийг хэд хэдэн хэсэг болгон хуваах, хүн ярьж байгаа эсэхийг шүүх, шуугианыг бууруулах зэргээр өгөгдлийн цэвэрлэгээ хийгддэг. Ингэснээр тухайн аудионоос ярьж буй субъектуудыг ангилах боломжийг олгодог [5]. Сегментчлэл нь Recurrent Neural Network (RNN) алгоритм дээр суурилагдаж гүйцэтгэгддэг (1).

Томьёо 1: RNN алгоритмын томьёо:

$$h_t = \sigma(W_x X_t + W_h h_{t-1} + b) \quad (1)$$

Энд:

- X_t : Аудио сигналын оролт
- h_{t-1} : Илэрүүлэх хувьсагчийн утга
- W_x : Матрицын оролтын утгын жин
- W_h : Матрицын хувьсагчийн утгын жин
- b : Хазайлтын утга
- σ : Идэвхжүүлэлтийн функц

3.3.2 Эмбеддинг (Embedding)

Энэхүү шатанд сегментүүдээс буюу ярьж буй субъектуудээс хоолойны өнгө, ярианы хурд, ярианы зангилаа(pitch) зэрэг шинж чанаруудыг илрүүлэх үйл явц хийгддэг [5]. Эмбэддинг нь Bayesian Information Criterion (for speaker change detection) алгоритм дээр суурьлагдаж гүйцэтгэгддэг (2) .

Томьёо 2: Bayesian Information Criterion (BIC) томьёо:

$$21[N\cdot\log|\Sigma| - N1\cdot\log|\Sigma 1| - N2\cdot\log|\Sigma 2| - 2\lambda\cdot P\cdot\log N \quad (2)$$

- Энд:
  N - Нийт өгөгдлийн тоо
- N1,N2 - Сегмент 1 ба 2 дахь өгөгдлийн тоо.
- Σ\Sigma - Нийт өгөгдлийн ковариацын матриц.
- Σ1,Σ2\Sigma_1, \Sigma_2 - Сегмент 1 ба 2-ийн ковариацын матрицууд.
- P - Параметрын тоо (фичерийн хэмжээс).
- λ\lambda - Penalty параметр

3.3.3 Pyannote-Кластеринг (Clustering)

Энэхүү шатанд эмбеддинг хийгдсэний үр дүнд илэрсэн шинж чанаруудыг ангилах үйл явц хийгддэг [5]. Кластеринг нь Agglomerative Hierarchihal Clustering (AHC) алгоритм дээр суурьлагдаж гүйцэтгэгддэг (3).

Томьёо 3: AHC алгоритмын томьёо:

$$D(E_i, E_j) = 1 - (E_i \cdot E_j) / (\|E_i\| \, \|E_j\|) \quad (3)$$

Энд:

- $E_i \cdot E_j$: Хоёр эмбеддинг векторын скаляр үржвэр
- $\|E_i\|, \|E_j\|$: Вектор тус бүрийн норм
- $D(E_i, E_j)$: Cosine distance

3.4 Whisper-Текст хөрвүүлэлт (Speech-to- Text)

Whisper модель нь дараалаас дараалал **Seq2Seq (Sequence-to-Sequence)** хосолсон загвараар яриаг текст болгон хөрвүүлдэг.[14]

Томьёо 4: Seq2Seq алгоритмын томьёо

$$P(Y|X) = \prod P(y_t \mid y_1, \ldots, y_{t-1}, X; \theta) \qquad (4)$$

Энд:

- X - өгөгдлийн орж ирсэн дараалал
- Y - гарах текст
- θ - моделийн параметрүүд
- P(yt|y1,...,yt−1,X) - оролтын дагуу дараагийн текст үүсгэх функц

3.5 Large Language Model (Эмийн жорыг стандартын дагуу боловсруулах)

Эмч, эрүүл мэндийн мэргэжилтэн болон өвчтөний хоорондын аудио өгөгдлөөс хөрвүүлэгдсэн текстийг бүх шатны эрүүл мэндийн байгууллага, эмийн сангууд үйл ажиллагаандаа ашигладаг MNS 5376:2016 стандартын дагуу эмийн жор боловсруулж гаргахад бүтэцлэгдсэн өгөгдлийг нейрон сүлжээний алгоритмоор боловсруулдаг трансфромер моделыг ашиглав. Энэ нь “self-attention” функц дээр тулгуурлан ажилладаг.[2,4,5,8,9]

IV. ТУРШИЛТ

4.1 Алдааны тооцоолол

Яриан дээрх үг хэллэгүүдийг зөв таньж, бичсэн эсэхийг Word Error Rate (WER) метрикээр, стандартын дагуу бүрэн бүтэн, үнэн зөв эмийн жор бичсэн эсэхийг Cross Entropy Loss (CEL) метрикээр тус тус хэмжиж гаргав.[1,10]

Үүний үр дүнд:

• Үгийн алдааны хувь = 6.3722%

• Эмийн жорын алдааны хувь = 23.8542%

• Яриа танилтын алдааны хувь = 5.8542%

Томьёо 5 : DER метрикийн томьёо (5)

$$\text{Яриа танилтын алдааны хувь} = \frac{\text{алдаж таньсан хугацааны тоо}}{\text{нийт ярианы хугацаа}} = 5.8542$$

Томьёо 6: WER метрикийн томьёо (6)

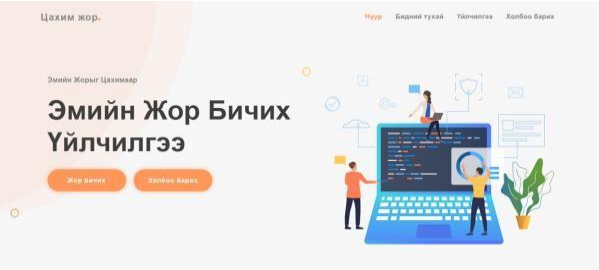
Үгийн алдааны хувь = (Зассан үгийн тоо + Устгасан үгийн тоо) / Нийт үгийн тоо

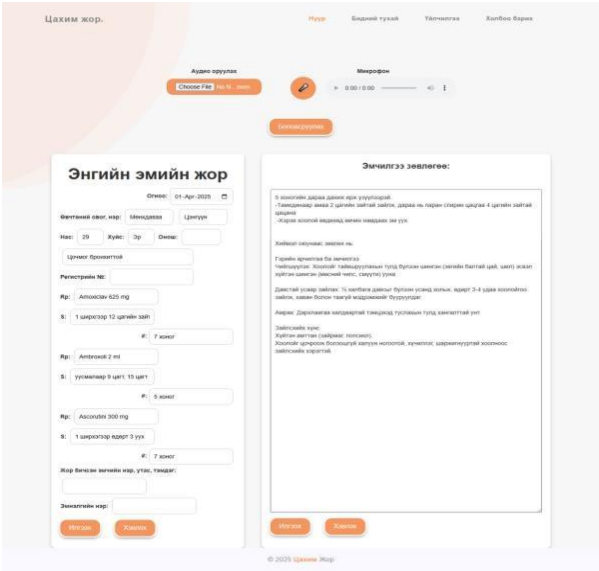
Томьёо 7: LLM моделийн алдааг тооцох (7)

*Эмийн жорын алдааны хувь = бодит утга * log(боловсруулагдсан утга)*

V. ҮР ДҮН

Эрүүл мэндийн байгууллагад ажиллаж буй эмч болон эмийн жор бичиж өгөх эрх бүхий мэргэжилтнүүд үзлэгийн аудио бичлэгийг оруулах юм. Уг систем нь аудиог боловсруулж үүнээс стандартын шаардлагад нийцсэн эмийн жор болон өвчтөнд зориулсан зөвлөгөөг боловсруулан, эмийн сангийн систем, E-Mongolia системрүү илгээх боломжтой юм. Мөн боловсруулагдсан эмийн жорыг PDF болон DOCX өргөтгөлтэйгөөр татаж авах боломжтой.[5,10,12,14,15]





Зураг 2. Системийн дэлгэц

АШУҮИС-ийн Монгол-Японы эмнэлэг
13270, Улаанбаатар хот, Баянзүрх дүүрэг, 12-р хороо,
Ботаникийн цэцэрлэгт хүрээлэнгийн гудамж

Эцэг/эхийн нэр: Менхдавав Нас: 4
Нэр: Цэнгүүн Хүйс: Эр
Регистр: УК20302152

Эмчилгээ, зөвлөгөө:

5 хоногийн дараа дахиж ирж үзүүлэгдэй.
-Тамединаар амаа 2 цагийн зайтай зайлж, дараа нь паран спирин цацгаа 4 цагийн зайтай цацана
-Хэрэв хоолой өвдөхөд өвчин намдаах эм уух

Хиймэл оюунаас зөвлөх:

Гэрийн арчилгаа ба эмчилгээ
Чийгшүүлэх: Хоолойг тайвшруулахын тулд бүлээн шингэн (зөгийн балтай цай, шөл) эсвэл хүйтэн шингэн (мөсний чипс, смүүти) ууна.

Давстай усаар зайлах: ½ халбага давсыг бүлээн усанд хольж, өдөрт 3-4 удаа хоолойгоо зайлж, хаван болон таагүй мэдрэмжийг бууруулдаг.

Амрах: Дархлаагаа халдвартай тэмцэхэд туслахын тулд хангалттай унт.

Зайлсхийх хүнс
Хүйтэн амттан (зайрмаг, попсикл).
Хоолойг цочроож болзошгүй халуун ногоотой, хучиллэг, шаржигнууртай хоолноос зайлсхийх хэрэгтэй.

Зураг 3. Системийн гаралт

5.2 Системийн ирээдүйн боломж

-Эмийн жорыг зөв бичих автомат системийг эмнэлгийн олон салбарт нэвтрүүлэх[18].
- AI-д суурилсан зөвлөгөө өгөх системийг нэмэх
- Эрүүл мэндийн даатгал, цахим эрүүл мэндийн сангийн уялдааг хангах.
- Хэл шинжлэл, яриа боловсруулалтын судалгаанд хувь нэмэр оруулах.

ДҮГНЭЛТ

Энэхүү судалгааны ажлаар эрүүл мэнд, байгаль экологийн салбарт тулгамдаж буй асуудлуудын нэг болох эмийн жорыг гар аргаар бичиж, цаасан хэлбэрээр ашигладаг үйл явцыг хиймэл оюун ухааны дэвшилтэт технологийн тусламжтайгаар автоматжуулахыг зорьсон. Энэ нь эмч, эмнэлгийн байгууллагын ачааллыг бууруулахын зэрэгцээ цаасан эмийн жорын хэрэглээг халж, ухаалаг, байгальд ээлтэй "ногоон" технологийн шийдлийг хэрэгжүүлэхэд чиглэж байна.. Тус систем нь эмчийн үзлэгийн үеэр үйлчлүүлэгчид өгч буй зөвлөгөө, ярианаас стандартын шаардлагад нийцсэн эмийн жорыг автоматаар боловсруулан, эмийн сангийн систем болон E- Mongolia системрүү илгээх юм. Үүний үр дүнд иргэд, эм зүйч болон эрүүл мэндийн байгууллагуудын хоорондох эм олголтын үйл явцыг хянах боломжоор хангасан болно. Ингэснээр эмийн жор олгоход зарцуулагддаг 20-40 минутын хугацааг, 3 минут хүртэл бууруулах, байгалийн гаралтай, амьдралын зөв дадлыг дэмжсэн, ахуйн нөхцөлд хэрэгжүүлэх боломжтой эмчилгээ, зөвлөгөөг боловсруулах юм.

АШИГЛАСАН МАТЕРИАЛ, НОМ ЗҮЙ

[1] Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. Nature medicine, 28(1), 31-38.

[2] Batta, A., & Singh, B. (2018). Rational approach to prescription writing: A preview. Neurology India, 66(4), 928-933.

[3] Wang, S., Yang, C. H., Wu, J., & Zhang, C. (2024, April). Can whisper perform speech-based in-context learning?. In ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 13421-13425). IEEE.

[4] Sharaf, N. I. (2024). PRESCRIPTION AUTOMATION: ENHANCING MEDICATION SAFETY AND HEALTHCARE EFFICIENCY. Gland Surgery, 9(2), 416-442.

[5] Soeny, K., Pandey, G., Gupta, U., Trivedi, A., Gupta, M., & Agarwal, G. (2021). Attended robotic process automation of prescriptions' digitization. Smart Health, 20, 100189.

[6] Chen, H. K., Chen, F. H., & Lin, S. F. (2021). An ai-based exercise prescription recommendation system. Applied Sciences, 11(6), 2661.

[7] Bredin, H. (2023, August). pyannote. audio 2.1 speaker diarization pipeline: principle, benchmark, and recipe. In 24th INTERSPEECH Conference (INTERSPEECH 2023) (pp. 1983-1987). IS4CA.

[8] Liu, Z., Wu, Z., Hu, M., Zhao, B., Zhao, L., Zhang, T., ... & Sikora, A. (2023). PharmacyGPT: The AI Pharmacist. arXiv preprint arXiv:2307.10432.

[9] Huang, T.-Y., Huang, W.-T., Lin, Y.-C., Hung, H.-H., Ou, S.-C., Chang, C.-W., ... & Huang, S.-T. (2023). Chinese Intelligence Prescription System improves prescription accuracy while decreasing labor and drug costs. BMC Health Services Research, 23, 94.

[10] Bouraghi, H., Imani, B., Saeedi, A., & others. (2024). Challenges and advantages of electronic prescribing system: a survey study and thematic analysis. BMC Health Services Research, 24, 689.

[11] Pavithiran, G., Padmanabhan, S., Divya, N., Aswathy, V., Jerusha, I., & Chandar, B. (2022). Doctors Handwritten Prescription Recognition System In Multi Language Using Deep Learning. arXiv preprint arXiv:2210.11666.

[12] Jurafsky, D., & Martin, J. H. (2021). *Speech and language processing* (3rd ed.). Pearson

[13] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

[14] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *In NIPS 2017* (pp. 5998-6008). *Curran Associates, Inc.*

[15] Монголын гааль.

[16] Монгол эм импекс концерн.

[17] Andromeda ХХК.

[18] Mindgolia ХХК.

[19] Орион Системс ХХК.

[20] E-Mongolia үйлчилгээ 'Эмийн жор цахим систем'.

[21] Эмийн жор бичих MNS 5376:2016 стандарт.

[22] Ач интернешнл эмнэлэг.

[23] АШУҮИС-ийн Монгол-Японы эмнэлэг.

A User-Centric Tool for Design Defect Detection in UML XMI Models Using Pre-Trained Machine Learning Models

Bilguun Erdenebaatar*, Lkhamrolom Tsoodol†, Oyun-Erdene Namsrai‡, Batnyam Battulga§, Naranchimeg Bold¶
Department of Information and Computer Science
National University of Mongolia
Corresponding Email: ¶naranchimeg@num.edu.mn
Email: *bilguun.e@num.edu.mn †lkhamrolom@num.edu.mn, ‡oyunerdene@num.edu.mn, §bbatnyam@num.edu.mn

***Abstract*—Design flaws introduced during early software design phases propagate into development, leading to increased cost and complexity. This paper introduces a lightweight, web-based tool for early-stage UML class diagram quality evaluation using static design metrics and pre-trained machine learning models. Users submit UML diagrams in XMI format and select from Random Forest, XGBoost, or Extra Trees evaluators to predict design quality. The tool integrates SDMetrics for metric extraction and offers an accessible, dynamic frontend for immediate feedback. We present system architecture, technology stack, metric and model justification, and future directions.**

***Index Terms*—UML, XMI, SDMetrics, Machine Learning, Software Quality, Random Forest, XGBoost, Extra Trees**

I. Introduction

Detecting design flaws early in the Software Development Life Cycle (SDLC) is crucial for creating maintainable, extensible, and robust software systems. Unified Modeling Language (UML) class diagrams are widely adopted for modeling object-oriented software. However, manual examination of these diagrams is tedious and error-prone. Existing tools often lack user-centered simplicity or rely heavily on manual configuration.

Prior work has explored automated design smell detection using both metric-based [2] [3] and deep learning approaches [8]. Metric-based anti-pattern detection frameworks such as MoRT and tools like SDMetrics have been effective in static evaluation. Ensemble machine learning methods, especially Random Forest [4], Extra Trees [5], and XGBoost [6], have demonstrated superior performance in software defect prediction [7]. Our work differs by focusing on providing a fully integrated, simple-to-use, web-accessible tool without requiring users to engage in model training or setup.

In this paper, we propose a web-based tool that integrates SDMetrics [1] with pre-trained machine learning classifiers to automate UML class diagram quality evaluation. This enables developers, especially students and early-career engineers, to receive immediate feedback and improve their designs iteratively.

The predictive engine in our tool is directly based on the methodology proposed by Bold et al. [9], who introduced a defect prediction framework operating at the design phase by leveraging object-oriented design metrics extracted from UML class diagrams. Their approach involved reverse-engineering class diagrams from source code using AltovaUML and computing 24 structural metrics via SDMetrics, covering dimensions of size, inheritance, and coupling. These metrics were then mapped to class-level defect labels from the Unified Bug Dataset, resulting in a comprehensive model-based dataset. In our system, we have adopted this same dataset construction pipeline and directly utilized the trained machine learning and deep learning models published in their work—including fine-tuned implementations of Random Forest, XGBoost, MLP, and TabTransformer classifiers [9].

We reimplemented the preprocessing stages outlined in [9], including standardization and synthetic minority oversampling (SMOTE), and incorporated their hyperparameter-optimized configurations (e.g., RF with 500 estimators, max_depth=32, and max_features='log2'). The performance benchmarks reported in [9]—with F1 scores exceeding 67 percent on model-based metrics—offered strong justification for using these models without retraining. Furthermore, the cross-dataset validation experiments in their study confirmed the generalizability of their models across unseen systems, which aligns with our need to deploy the predictors in real-world UML-driven design environments. Our contribution extends their work by integrating this model backend into a user-facing tool capable of real-time UML class diagram analysis and defect risk scoring, thereby operationalizing the research from [9] into an applied design-phase quality assurance tool.

II. Machine Learning Model Selection

This tool employs three algorithms—Random Forest [4], Extra Trees [5], and XGBoost [6] chosen for their proficiency in handling structured, high-dimensional, and non-linear data, characteristics typical of static software metrics. These models effectively capture complex patterns inherent in software design, where quality indicators often exhibit non-linear relationships. Moreover, they adeptly manage irrelevant

features, a common challenge in software metrics analysis, thereby enhancing predictive accuracy

A. Random Forest

Random Forest is an ensemble learning algorithm that builds a collection of decision trees during training and outputs the mode of the classes for classification or the mean prediction for regression tasks. It combines bootstrap aggregation (bagging) and feature randomness to construct diverse trees that enhance generalization and reduce overfitting [4].

One of the primary advantages of Random Forest is its robustness to high-dimensional, imbalanced, and noisy data. Due to its ensemble nature, it maintains predictive accuracy even under challenging data conditions [6]. It also effectively reduces overfitting by averaging the predictions of multiple decorrelated trees, thereby avoiding the high variance associated with individual decision trees [17].

Despite its strengths, Random Forest has slower inference times compared to simpler models, especially when dealing with large forests. Since predictions require aggregating the results of all individual trees, latency can become a bottleneck in real-time applications [18].

B. Extra Trees

Extra Trees, short for Extremely Randomized Trees, is an ensemble learning method similar in design to Random Forest but introduces greater stochasticity during the tree-building process. [19] Unlike Random Forest, which uses bootstrap samples, Extra Trees builds each decision tree using the entire dataset and selects splits completely at random for each feature. This approach is designed to decrease variance and training time while maintaining competitive predictive performance [5].

One notable benefit of Extra Trees is its training speed, as it avoids the bootstrapping step used in Random Forests. By training on the entire dataset and using randomized split thresholds, the model achieves faster convergence and scales effectively to large datasets [19].

The method also exhibits lower variance due to its additional randomness, making it particularly useful for preventing overfitting in high-variance contexts [20]. This randomness further enhances the model's robustness to noisy data, as it avoids overly fitting to specific training instances [21].

However, the increased randomness comes with trade-offs. In some scenarios, Extra Trees may yield slightly lower accuracy than Random Forest due to higher model bias [22]. Additionally, the randomized nature of split selection can hinder interpretability, making it harder to explain model decisions in detail—especially in sensitive or regulated domains [23].

C. XGBoost

Extreme Gradient Boosting is a scalable and optimized implementation of gradient boosting designed for high performance and efficiency. Unlike bagging methods that build trees independently, XGBoost constructs decision trees sequentially, where each new tree corrects the residuals (errors) of the combined previous trees. The algorithm incorporates both L1 and L2 regularization terms to reduce overfitting, making it highly effective for structured data tasks [6].

XGBoost has earned a reputation for high predictive accuracy, consistently ranking among the top-performing models in Kaggle competitions and benchmark studies [24]. One of its key advantages is its built-in regularization, which enables better generalization by penalizing overly complex models during training [40].

However, the rich hyperparameter space of XGBoost can make tuning challenging. Achieving optimal performance often requires experimentation with parameters such as learning rate, tree depth, and subsampling ratios [25]. Also, due to its sequential learning nature, XGBoost tends to have slower training times than algorithms like Random Forest or Extra Trees that can be parallelized more aggressively [26].

III. SDMetrics and Metric Selection

In software design, particularly in object-oriented environments, the structural quality and health of a system can be quantitatively assessed through specific metrics. Such metrics provide information about various aspects of the design, such as size, complexity, coupling, and inheritance hierarchies. [2] On the basis of these metrics, it is possible to detect potential design weaknesses early during the development cycle so that optimization can be performed proactively. [3] SDMetrics offers metrics at class, interface, package, and component levels, of which class and interface metrics are utilized in this system. Our tool utilizes the following metrics categorized:

Size Metrics: Size metrics quantify the magnitude of various components within a class, offering a perspective on its complexity and potential maintainability challenges [1]. (Number of Attributes, Number of Operations, Number of Public Operations, Setters and Getters)

Inheritance Metrics: These metrics assess the depth and breadth of inheritance hierarchies, shedding light on the design's complexity and potential reuse. [1] (Number of Chil-

	A	B	C	D	E	F	G
1	ID	Column1	Name	NumAttr	NumOps	NumPubO	Setters
2	L142	class	Base64.java		1		
3	L8256	class	QueryRunner.java		1		
4	L8251	class	BasicRowProcessor.java		1		
5	L8518	class	BeanListHandler.java	0	1	1	0
6	L8516	class	ArrayListHandler.java	0	1	1	0
7	L8519	class	ColumnListHandler.java		1		
8	L8523	class	MapListHandler.java	0	3	3	0
9	L9906	class	FileUtils.java	6	13	11	0
10	L8252	class	BeanProcessor.java	3	8	6	0
11	L8520	class	GenericListHandler.java	0	3	3	0
12	L19999	class	FTPTimestampParserImpl.java	0	8	8	0
13	L10547	class	PrefixFileFilter.java	4	22	22	0
14	L10549	class	SuffixFileFilter.java	5	14	13	0
15	L752	class	Metaphone.java	0	5	0	0
16	L10531	class	AgeFileFilter.java	26	57	54	1
17	L10536	class	DelegateFileFilter.java	0	5	5	0
18	L10541	class	FileFilterUtils.java	5	4	3	0
19	L10544	class	NameFileFilter.java	2	7	5	0
20	L10545	class	NotFileFilter.java	2	47	43	9

Fig. 1. Standardized output of SDMetrics metrics.

dren, Depth of Inheritance Tree, Number of Ancestors, Class to Leaf Depth)

Coupling Metrics: Coupling metrics evaluate the interdependencies between classes, which can impact modularity and the ease of making isolated changes. [1]. (Outgoing Dependencies, Incoming Dependencies, Number of Direct Clients, Number of Indirect Clients)

Cohesion/Composition: These metrics focus on the internal consistency of a class and how closely related its components are. [1] (Nesting Level)

Encapsulation/Access: These metrics assess the degree to which a class hides its internal implementation details, a fundamental principle of object-oriented design. [1] (External Attribute Usage, Internal Attribute Usage, External Parameter Usage, External Parameter Usage, Internal Parameter Usage)

In alignment with Bold et al. [9], our system relies on 24 object-oriented design metrics that target three principal dimensions of design quality: size, inheritance, and coupling. Figure 1 depicts these metrics on an example data in its pre model input version. These metrics are widely recognized in the literature for capturing complexity, modularity, and interdependence—factors that influence maintainability and defect proneness. Size metrics such as the number of attributes and operations reflect class granularity, while inheritance metrics like DIT and NOC capture depth and breadth of abstraction. Coupling metrics, including the number of associations and external attribute usage, quantify the extent of inter-class dependencies. By using these metrics, we ensure that our predictive modeling captures structural design properties that are theoretically and empirically linked to software quality.

IV. System Architecture

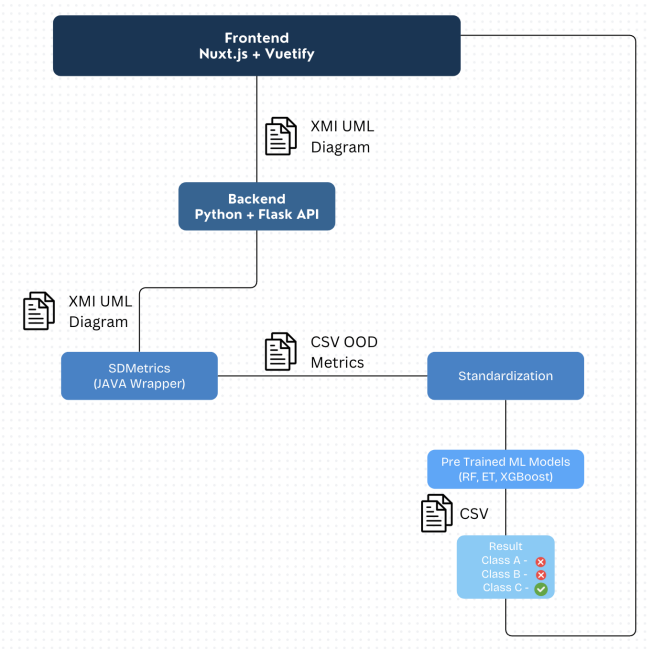


Fig. 2. Tool architecture

The proposed system comprises three primary components: a web-based frontend, a backend API for processing and model inference, and a machine learning module consisting of pre-trained classifiers shown in Figure 4.

A. Frontend: Nuxt.js + Vuetify

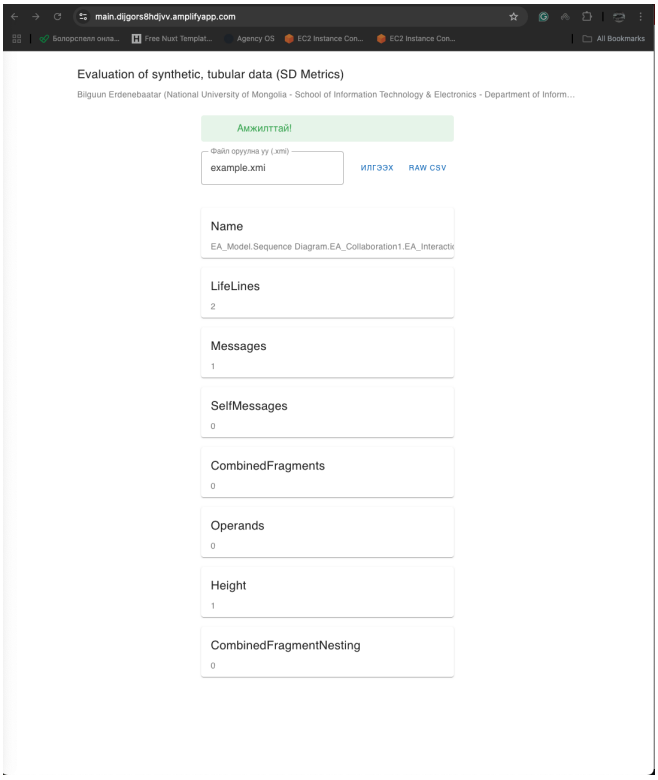


Fig. 3. Nuxt.js + Vuetify User Interface.

The frontend is developed using Nuxt.js, a modern framework that supports server-side rendering (SSR), facilitating improved performance and SEO optimization. Vuetify is employed for the user interface layer to ensure a responsive, material-compliant design and consistent usability across devices. The interface enables users to upload UML class diagrams in XMI format and select among available machine learning models for defect prediction. Figure 3 depicts the user interface of our tool created using Nuxt.js and components such as v-file-input, v-card, v-card-title etc.

B. Backend: Python + Flask API

The backend is implemented using Python and exposed via a lightweight Flask REST API. Figure 2 displays the process between the XMI input from the frontend to the final output of the pre trained models. Upon receiving the uploaded XMI file, the backend invokes a Java wrapper around the SDMetrics tool, which extracts 24 object-oriented design metrics from the class diagram, encompassing dimensions of class size, inheritance depth, and coupling [1]–[3], [9]. These metrics are then standardized and transformed into the format expected by the predictive models.

C. Machine Learning Module

The predictive engine comprises several pre-trained classifiers stored in.pkl format, including Random Forest [4], Extra Trees [5], and XGBoost [6]. These models were originally trained using the methodology proposed by Bold et al. [9], who constructed a labeled dataset by mapping the Unified Bug Dataset defects to class-level metric representations extracted via SDMetrics. The models were evaluated using F1 score, precision, and recall, demonstrating strong performance in predicting defect-prone classes from UML design alone.

By operationalizing this architecture, our system enables rapid feedback during the design phase of software development, assisting engineers in identifying risky components early without requiring compiled code or runtime data. This approach aligns with current research advocating for early defect prediction using interpretable and quantifiable metrics [3], [7], [10].

V. Infrastructure

The deployment strategy of this system is depicted in Figure 4. Built upon a modular cloud-based architecture using services from Amazon Web Services (AWS) and Cloudflare, ensuring performance, scalability, and global availability.

A. Backend Amazon Elastic Compute Cloud (EC2) [11]

The backend is hosted on Amazon Elastic Compute Cloud (EC2) [11], a virtual server environment that allows scalable compute capacity. EC2 provides flexibility in selecting instance types based on performance needs, making it well-suited for applications requiring computationally intensive backend processing. This is especially important for systems running prediction engines and metric analysis tools, where computational demands can vary [11]. EC2 also integrates seamlessly with AWS services such as Elastic Load Balancing and Auto Scaling, enabling high availability and fault tolerance under fluctuating loads. [12]

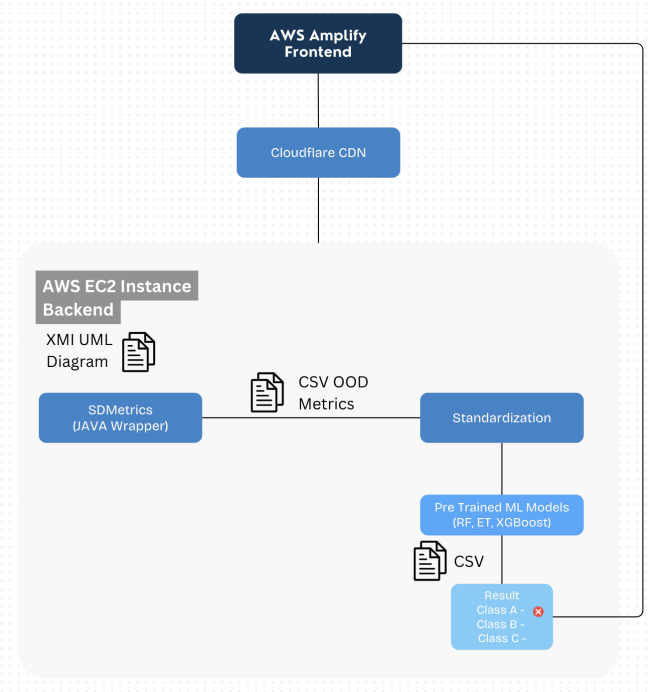


Fig. 4. Infrastructure diagram.

B. AWS Amplify

The frontend, developed using Nuxt.js and Vuetify, is deployed via AWS Amplify, a managed service that simplifies web app deployment by supporting Git-based continuous integration and delivery pipelines. Amplify is optimized for static site hosting and single-page applications (SPAs), providing fast global delivery through its integration with Amazon CloudFront, automatic HTTPS, and custom domain support [13]. These features collectively contribute to faster page loads and a streamlined developer experience.

C. Cloudflare CDN

To further reduce latency and improve content delivery speed, Cloudflare CDN is used in front of the frontend. Cloudflare caches static assets like JavaScript bundles and images across over 200 global edge locations, significantly decreasing the time required for users to access content [14]. Furthermore, Cloudflare provides built-in security measures such as DDoS mitigation, web application firewall (WAF), and bot protection, which enhance system reliability and safeguard data integrity [15]. This integration of caching and security aligns with best practices for global-scale applications and is essential for maintaining high-performance web interfaces [16].

VI. Conclusion

This work presents a practical and efficient tool for predicting design-level software defects by analyzing UML class diagrams using pre-trained machine learning models. By integrating SDMetrics for metric extraction and leveraging robust classifiers such as Random Forest, Extra Trees, and XGBoost, the system enables early risk assessment during the software design phase—without requiring source code or runtime data. Grounded in the methodology proposed by Bold et al. [9], the system transforms design artifacts into actionable insights, thus aligning with current research emphasizing early defect localization based on object-oriented design metrics [2], [3], [7].

Looking forward, we aim to enhance predictive accuracy and model generalization by implementing an ensemble-based approach that aggregates the outputs of all supported classifiers into a single meta-model. Such an ensemble framework is expected to reduce variance and improve classification robustness, especially when dealing with ambiguous or borderline design cases. In parallel, we plan to extend the system's interface and automation features to support continuous integration into design workflows, enabling design teams to receive real-time defect risk evaluations through a single, unified predictive engine. This future direction aims to elevate the tool from a

model selector to a design intelligence assistant capable of guiding architectural decisions with greater confidence.

REFERENCES

[1] SDMetrics Team, *Class Design Metrics - SDMetrics Manual*, 2024. [Online]. Available: https://www.sdmetrics.com/manual/Metrics_class.html. Accessed: May 1, 2025.

[2] S. R. Chidamber and C. F. Kemerer, "A Metrics Suite for Object Oriented Design," *IEEE Transactions on Software Engineering*, vol. 20, no. 6, pp. 476–493, 1994. doi: 10.1109/32.295895

[3] L. Bryton, K. V. de Oliveira, V. C. Garcia, E. S. de Almeida, and S. R. Meira, "Quality Analysis of UML Class Diagrams Using Metrics and Heuristics," *International Journal of Software Engineering and Knowledge Engineering*, vol. 21, no. 4, pp. 531–555, 2011. doi: 10.1142/S0218194011005364

[4] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. doi: 10.1023/A:1010933404324

[5] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely Randomized Trees," *Machine Learning*, vol. 63, no. 1, pp. 3–42, 2006. doi: 10.1007/s10994-006-6226-1

[6] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 2016, pp. 785–794. doi: 10.1145/2939672.2939785

[7] S. Lessmann, B. Baesens, C. Mues, and S. Pietsch, "Benchmarking Classification Models for Software Defect Prediction: A Proposed Framework and Novel Findings," *IEEE Transactions on Software Engineering*, vol. 34, no. 4, pp. 485–496, 2008. doi: 10.1109/TSE.2008.35

[8] G. Araneda et al., "MoRT: Model Representation with Transformers for Design Pattern Detection in UML Class Diagrams," *Automated Software Engineering*, 2024, forthcoming.

[9] N. Bold, B. Battulga, L. Tsoodol, E. Dovdon, and O. Namsrai, "Metric-Based Defect Prediction from Class Diagram," *SSRN*, 2024. doi: 10.2139/ssrn.5004689

[10] S. Perreault, R. Rawat, and S. Dubey, "Software Defect Prediction Analysis Using Machine Learning," *Sustainability*, vol. 15, no. 6, p. 5517, 2023. doi: 10.3390/su15065517

[11] G. Rehman and M. A. Khan, "Cloud Computing and the Amazon Web Services," in *Proceedings of the International Conference on Emerging Trends in Computing*, Springer, 2021. doi: 10.1007/978-981-33-6307-6_54

[12] A. Patel and A. Vyas, "Performance Analysis of Cloud Service Providers: Amazon AWS and Microsoft Azure," *International Journal of Creative Research Thoughts*, vol. 9, no. 1, 2021. [Online]. Available: https://www.ijcrt.org/papers/IJCRT2411729.pdf

[13] AWS, *Host Web Apps with Amplify Hosting*, 2020. [Online]. Available: https://docs.aws.amazon.com/amplify/latest/userguide/welcome.html. Accessed: May 1, 2025.

[14] Cloudflare, *The Cloudflare CDN: Architecture and Performance Benefits*, 2020. [Online]. Available: https://www.cloudflare.com/resources/images/slt3lc6tev37/6mpY5zjlPP3b2twznUQOki/b221bc7b49723deb2c09a4b30747f685/cloudflare-cdn-whitepaper-19Q4.pdf. Accessed: May 1, 2025.

[15] L. Fauzan et al., "Implementing Cloudflare CDN and Security Layer for Web Application Optimization," 2020. [Online]. Available: https://eprints.uad.ac.id/15251/1/T2_1689048037_NASKAH_PUBLIKASI.pdf. Accessed: May 1, 2025.

[16] AWS, *Simplify and Enhance Static Website Hosting with AWS Amplify*, 2020. [Online]. Available: https://aws.amazon.com/blogs/aws/simplify-and-enhance-amazon-s3-static-website-hosting-with-aws-amplify. Accessed: May 1, 2025.

[17] G. Hooker, "Discovering Additive Structure in Black Box Functions," in *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 2004, pp. 575–580. doi: 10.1145/1014052.1014126

[18] R. B. Dell and J. P. Schwab, "Random Forest Performance and Speed Tradeoffs in Large-Scale Applications," *Journal of Big Data*, vol. 8, no. 1, pp. 1–15, 2021. doi: 10.1186/s40537-021-00458-2

[19] G. Louppe, L. Wehenkel, A. Sutera, and P. Geurts, "Understanding Variable Importances in Forests of Randomized Trees," in *Advances in Neural Information Processing Systems*, vol. 26, 2013.

[20] J. Friedman, T. Hastie, and R. Tibshirani, *The Elements of Statistical Learning*, 2nd ed., Springer, 2009.

[21] K. Biau and E. Scornet, "A Random Forest Guided Tour," *TEST*, vol. 25, no. 2, pp. 197–227, 2016. doi: 10.1007/s11749-016-0481-7

[22] S. Nembrini, B. König, and I. Wright, "The Revival of the Gini Importance?," *Bioinformatics*, vol. 34, no. 21, pp. 3711–3718, 2018. doi: 10.1093/bioinformatics/bty373

[23] A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52138–52160, 2018. doi: 10.1109/ACCESS.2018.2870052

[24] A. Januschowski, J. Gasthaus, D. Salinas, S. Flunkert, and T. Wang, "Forecasting with XGBoost—An Evaluation of Its Performance and Applicability," *ACM Transactions on Intelligent Systems and Technology*, vol. 11, no. 4, 2020.

[25] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, vol. 13, pp. 281–305, 2012.

[26] T. Ke, Q. Meng, M. Finley, T. Wang, W. Chen, and T. Ma, "LightGBM: A highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.

Two Wheel Drive Autonomous Robot Navigation using GPS

Temuulen Batbold
School of Information Technology and Electronics
National University of Mongolia
Ulaanbaatar, Mongolia
btmln2000@gmail.com

Munkhtulga Byambaa*
School of Information Technology and Electronics
National University of Mongolia
Ulaanbaatar, Mongolia
https://orcid.org/0000-0002-3560-9583

*Corresponding author email: munkhtulga@num.edu.mn

***Abstract*—Autonomous mobile robots are essential for advancing space exploration by performing tasks in inaccessible or hazardous environments to humans. This study presents the design and implementation of a two-wheel drive (2WD) autonomous robot for space-related missions. The robot prioritizes simplicity, lightweight construction, and energy efficiency—key factors for deployment in extraterrestrial settings. For implementation, a hexagonal parachute was designed to ensure a safe landing. Positioning was determined using a GPS module, and the distance between two points on the globe was calculated using the Haversine formula. The direction was determined using motor encoder values, and the module performed repeated calculations of direction and distance every 30 seconds to fulfill the mission objectives. Mounting electronic components, the rover was assembled onto a copper board, and its wheels, main body, and parachute release mechanism were fabricated using a 3D printer to produce a functional prototype.**

***Index Terms*—robotics, GPS, navigation, two-wheel drive robot**

I. Introduction

As space exploration missions become increasingly ambitious, the demand for lightweight, efficient, and adaptable robotic platforms continues to grow. Autonomous mobile robots play a vital role in planetary exploration, enabling data collection, terrain mapping, and environmental monitoring in hazardous or inaccessible conditions to human astronauts. Among various robot configurations, two-wheel drive (2WD) autonomous robots offer a compelling balance between mechanical simplicity, energy efficiency, and maneuverability, making them suitable candidates for specific space applications, especially in environments where minimalism and reliability are prioritized.

The reduced structural complexity of 2WD robots translates to lower mass and easier maintenance—critical factors for extraterrestrial missions where every gram and watt of energy must be carefully managed [1]. Additionally, with advancements in artificial intelligence and autonomous navigation, 2WD robots can now traverse semi-structured terrains, localize themselves using onboard sensors, and make independent decisions in real-time [2]. Their compact form and modular design are ideal for integrating larger systems, such as landers or space stations, for localized tasks, including inspection, material transport, or scientific sampling.

In space environments such as the Moon or Mars, where gravitational conditions differ from Earth, 2WD systems require specialized locomotion strategies. However, their reduced contact with the ground can be advantageous in low-gravity conditions, where energy conservation and directional control are paramount [3]. Furthermore, recent interest in swarms of small-scale, cost-effective robots has reinforced the relevance of 2WD designs, which can be deployed in numbers for collaborative exploration and redundancy [4].

This paper presents the design, implementation, and performance analysis of a 2WD autonomous robot developed for space-oriented applications. Emphasis is placed on landing with parachutes, power management, electrical systems, and autonomous decision-making capabilities, with experiments conducted under conditions that simulate planetary surfaces.

II. Hardware System Design

A. Circuit Design

The circuit design of the 2-wheel-drive autonomous robot is presented in Fig. 1, illustrating the connections between key components such as the 9DOF inertial measurement unit (MPU9250), GPS module, motors, motor driver, and microcontroller. The main control board operates at 5V, while the power supply provides 11.1V. To accommodate this, a voltage regulator (7805) was used to step down the voltage to 5V, which is then used to power the board. Additionally, the motor driver is directly connected to the 11.1V power supply to ensure sufficient power delivery to the motors.

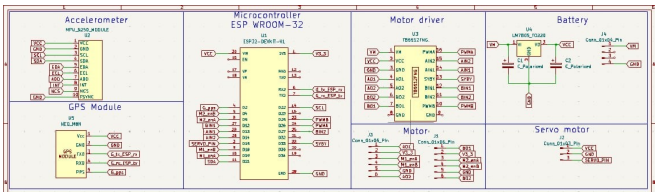
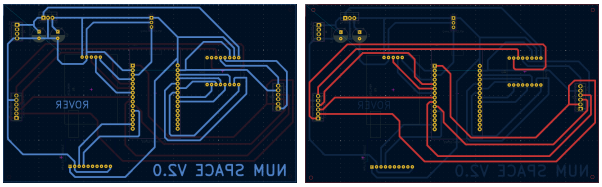


Fig. 1: Electronic system

Fig. 2 presents the printed circuit board (PCB) layout designed for the robot's control system. A double-sided copper layout was utilized to maximize available space and

accommodate all necessary components efficiently. The ESP32 microcontroller was selected as the central processing unit due to its high performance, integrated wireless communication capabilities, and suitability for embedded applications. Accurate detection of the robot's touchdown was required to enable parachute deployment upon landing. For this purpose, the MPU9250 inertial measurement unit (IMU) was integrated, providing acceleration and orientation data critical for identifying landing events. Given the robot's two-wheel drive configuration, the motor and motor driver components were also carefully selected to ensure reliable locomotion and control.



(a) Bottom side (b) Top side

Fig. 2: The printed circuit board (PCB) layout designed for the robot's control system

B. Parachute

The design of a parachute is influenced by multiple factors, including wind speed during descent, payload weight, and the structural integrity of its deployment mechanism. For small, lightweight rockets with a mass of less than 30 grams, streamer materials are commonly used. Streamers are narrow plastic strips rotating in the air to create drag, decelerating the rocket. Parachutes for such applications can be constructed by sewing nylon fabric or forming a semi-spherical shape from plastic. Plastic parachutes are suitable for rockets weighing less than 120 grams. Ripstop nylon is preferred for heavier payloads due to its superior strength, tear resistance, and flexibility.

Fig. 3: Ripstop nylon parachute

The most common parachute geometries used in model rocketry are semi-spherical and hexagonal designs, each offering different aerodynamic characteristics depending on the application (Fig. 4).

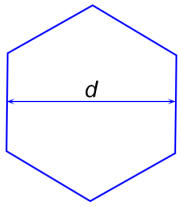


Fig. 4: Parachute design (Hexagon)

For our robot, which has a total weight of approximately 1 kg and is expected to descend from an altitude of 200–300 meters, ripstop nylon was selected as the parachute material due to its durability and resistance to tearing. The target descent velocity is around 2.5 m/s, with a desired total flight duration of 70–80 seconds. A descent speed below 2.5 m/s is generally considered to ensure a safe landing.

Important parameters for parachutes are:

- C_d - Drag coefficient - 0.75
- ρ - Air density - 1.225 kg/m^3
- $\mathbf{F}_g$ - Gravity - 9.81 m/s^2
- S - Parachute size
- V_d - Descending speed - 2.5m/s

The diameter of the vent hole is typically chosen to be 5–10% of the parachute's overall diameter, which only reduces the total surface area by approximately 1–5%. This central vent significantly enhances the parachute's stability during descent and contributes to a reduction in descent speed.

Parachute size calculation:

$$F_d = F_g \tag{1}$$

$$F_g = m \cdot g \tag{2}$$

$$F_d = \frac{1}{2}\rho \cdot V_d^2 \cdot S \cdot C_d \tag{3}$$

$$\frac{1}{2}\rho \cdot V_d^2 \cdot S \cdot C_d = mg \tag{4}$$

Thus, the size of the parachute is:

$$S = \frac{2mg}{V_d^2 \cdot \rho \cdot C_d} \tag{5}$$

The diameter is:

$$d = \sqrt{\frac{2 \cdot S_p}{\sqrt{3}}} \tag{6}$$

After calculation, the diameter of the parachute is 90 cm, and the size is 6834 cm^2.

C. 3D design and prototype

The 3D model of the robot was designed using SolidWorks, and the main body, along with the wheel components, was fabricated using a 3D printer (Fig. 5). A total of one main body and two wheels were printed over approximately 30 hours. The parachute release mechanism is designed to operate via a servo motor, which rotates a locking mechanism backward to disengage the section attached to the parachute, thereby initiating deployment.

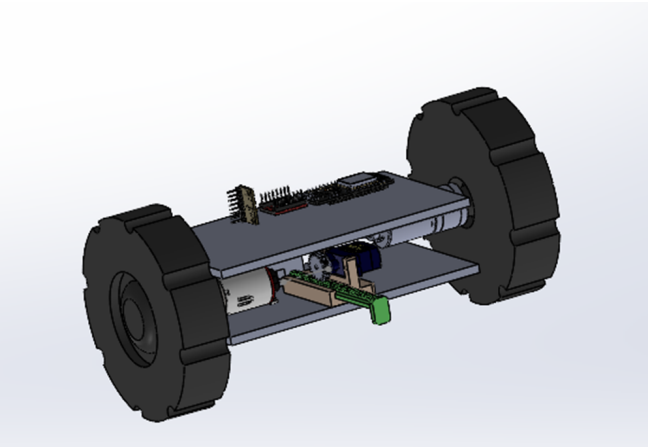


Fig. 5: 3D design of robot drawn in SolidWorks software

After approximately 30 hours of printing, the prototype was successfully completed, as illustrated in Fig. 6.

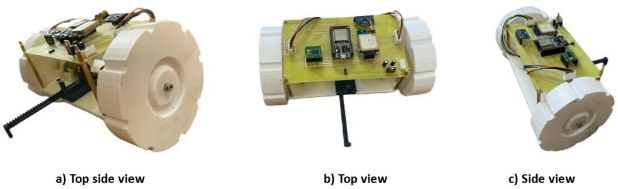


Fig. 6: Prototype of two-wheel-drive robot

III. Software System Design

A. Distance calculation

The equation shown below is used to calculate distances between two points in Euclidean space.

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{7}$$

However, the Euclidean distance formula is unsuitable for calculating distances between two points on a sphere, as it does not account for the curvature of the surface. The Haversine formula is employed to calculate the distance between two points defined by coordinates on the Earth's surface (Fig. 7). This formula determines the great-circle distance between two points on a sphere, considering their respective latitudes and longitudes.

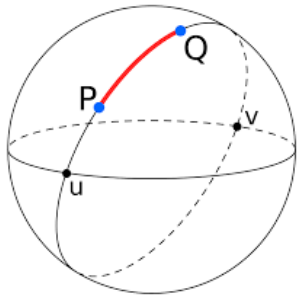


Fig. 7: Haversine formula

$$\begin{aligned} a &= \sin^2\left(\frac{\Delta\phi}{2}\right) + \cos(\phi_1)\cdot\cos(\phi_2)\cdot\sin^2\left(\frac{\Delta\lambda}{2}\right) \\ c &= 2\cdot\arctan\left(\sqrt{a}, \sqrt{1-a}\right) \\ d &= R\cdot c \end{aligned} \tag{8}$$

where $\Delta\phi = \phi_2 - \phi_1$ is difference in latitude, $\Delta\lambda = \lambda_2 - \lambda_1$ is difference in longitude, ϕ_1 and ϕ_2 are the latitudes of the two points (in radians), λ_1 and λ_2 are the longitudes of the two points (in radians) R radius of the earth (R = 6371 km), d is the great-circle distance between the two points (along the sphere's surface).

1) Filtering the raw GPS data: Using a filter for GPS data is strongly recommended, especially when working with consumer-grade modules like the NEO-6M, which can be noisy and have position jumps due to signal multipath, satellite geometry, and temporary obstructions. Raw GPS data (latitude, longitude, speed, altitude) can be:

- Noisy (especially at low speeds or indoors),
- Delayed (due to satellite acquisition),
- Inconsistent (jittery movement when stationary).

A filter helps smooth out this data, estimate the real position more accurately, and sometimes fuse it with other sensor data (like IMU/accelerometer). For this instance, the Kalman filter is applied. The Kalman filter is used for its best accuracy and sensor fusion. It estimates true position/speed by reducing noise and optionally combining data from other sensors.

The Kalman filter estimates a moving object's position by combining previous location estimates and current GPS measurements. It predicts the next position and then corrects it using noisy GPS data, producing a smoother and more accurate trajectory.

$$\begin{aligned} \hat{x}_{k|k-1} &= F\hat{x}_{k-1|k-1} \\ P_{k|k-1} &= FP_{k-1|k-1}F^T + Q \\ K_k &= P_{k|k-1}H^T(HP_{k|k-1}H^T + R)^{-1} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k(z_k - H\hat{x}_{k|k-1}) \\ P_{k|k} &= (I - K_kH)P_{k|k-1} \end{aligned} \tag{9}$$

The plot compares unfiltered GPS coordinates and those processed with a Kalman filter. The red line represents raw

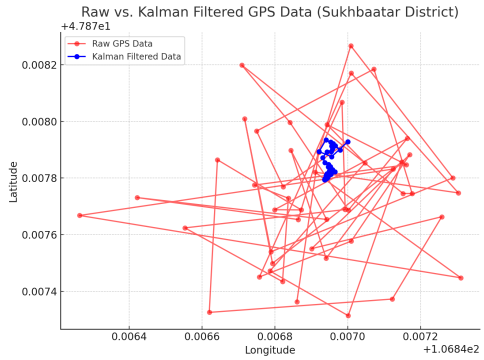


Fig. 8: Raw vs. Kalman Filtered GPS data

GPS data, which shows noticeable fluctuation due to signal noise and environmental factors. In contrast, the blue line displays the smoothed trajectory resulting from Kalman filtering, effectively reducing noise and stabilizing position estimates. This highlights the filter's usefulness in applications where reliable and consistent location tracking is critical, such as mobile navigation or autonomous systems.

B. Heading angle calculation

Since the robot lacks vision as humans do, it cannot inherently determine whether it is moving in the correct direction. Three key points are defined to address this: the landing point, the last visited point, and the target point. The internal angles between these points are calculated by applying the cosine rule. The robot adjusts its orientation based on the corresponding external angle to align itself with the correct heading.

$$c^2 = a^2 + b^2 - 2 \cdot a \cdot b \cos(\phi) \tag{10}$$

where a, b, c are sides of triangle, ϕ - c is the opposite angle.

Once the opposite angle is determined, the robot's heading angle can be calculated accordingly.

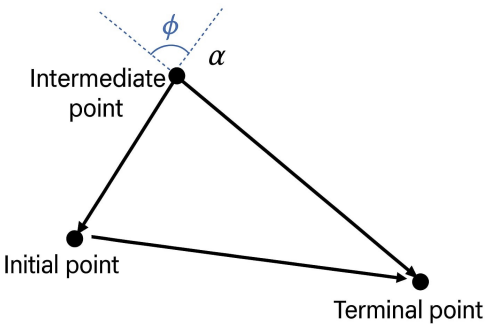


Fig. 9: Estimating the direction

C. Algorithm

As shown in Fig. 10, this algorithm enables a robot to initiate ground navigation after landing autonomously. It uses accelerometer data to detect landing via axis tilt, then detaches the parachute. GPS data determines the robot's position and calculates the distance and direction of the target. If the target is not reached, the motor is activated to begin movement; otherwise, the process ends.

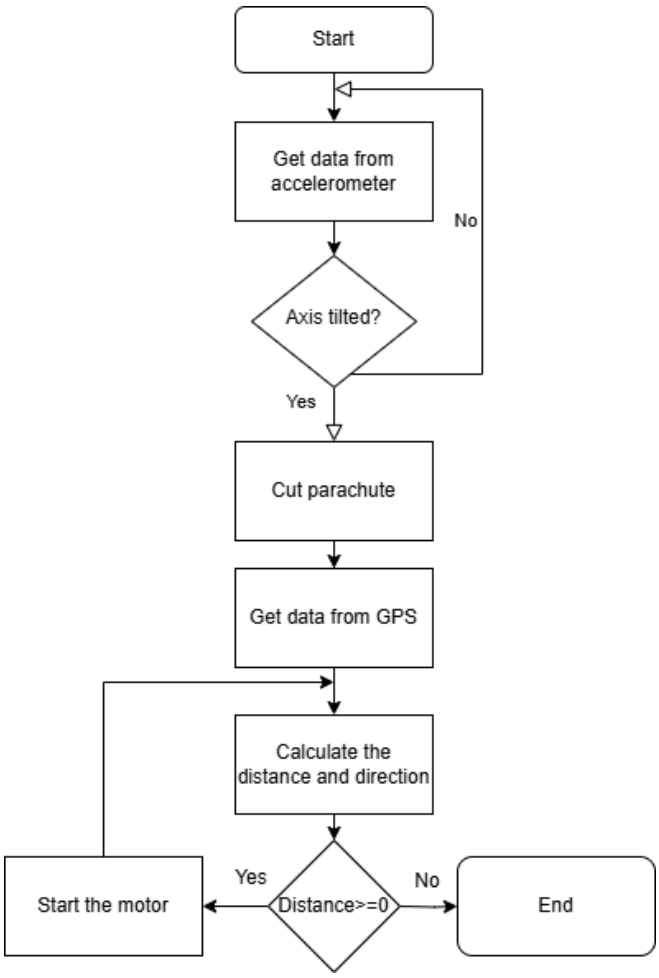


Fig. 10: Algorithm

1) Filtering the IMU data: Filtering the raw output of the MPU9250 sensor is essential to mitigate measurement noise and improve the accuracy and stability of tilt estimation, particularly in dynamic environments where sensor readings are affected by vibrations, sudden movements, and electromagnetic interference.

Fig. 11 shows raw MPU9250 tilt angle data (with noise), the filtered output using a Kalman filter, and the reference tilt trend, demonstrating how the Kalman filter effectively reduces noise while preserving the underlying motion.

IV. Experimental Results

A. Parachute deployment results

After increasing the size of the parachute, a modification was made by introducing a vent hole, and a series of experiments were conducted. The results of these tests are presented

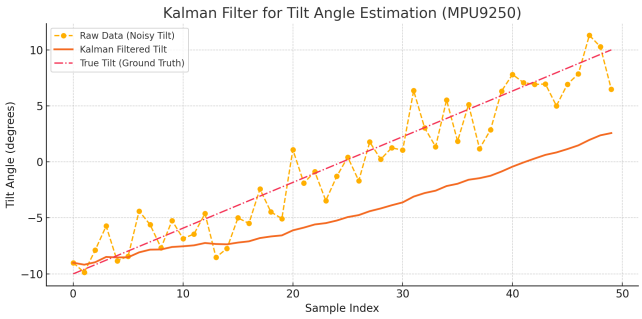


Fig. 11: Comparison of raw MPU9250 tilt angle data and Kalman filtered output across 50 samples. The Kalman filter significantly reduces measurement noise while maintaining accurate tilt estimation.

in the Table. I. During the experiments, a robot with an approximate weight of 1 kg was attached to the parachute.

	Height (m)	Duration (s)	Velocity (m/s)
Without vent hole	12.84	5.2	2.47
Vent hole (D = 5cm)	12.84	3.1	4.14
Vent hole (D = 10cm)	12.84	2.85	4.5

TABLE I: First parachute test

The modified parachute was tested four times, and the velocities were calculated and presented in Tab. II. Based on these trials, the average descent speed was determined to be 2.665 m/s.

Height (m)	Duration (s)	Speed (m/s)
12.84	5.2	2.47
12.84	4.7	2.73
14.2	5.1	2.78
14.2	5.3	2.68

TABLE II: Second parachute test

B. Localization results

We compared the results with two independent sources to evaluate the accuracy of our GPS-based localization. The first reference was Google Maps, which provides latitude and longitude data with reasonable precision. The second was the u-blox software, which offers high-accuracy location data directly obtained from satellite signals. From this experiment, the GPS test is accurate when the first four digits after the decimal point of the coordinates match (refer to Tab. III).

Method	Latitude	Longitude
Google Maps	47.877876	106.847173
u-blox software	47.877865	106.846779
GPS (Ours)	47.877804	106.846919

TABLE III: Comparison between Google Maps, u-blox satellite software, and our GPS module.

C. Distance calculation results

To assess the accuracy of our distance calculations using the Haversine formula, we used Google Maps' distance measurement feature as the reference baseline.

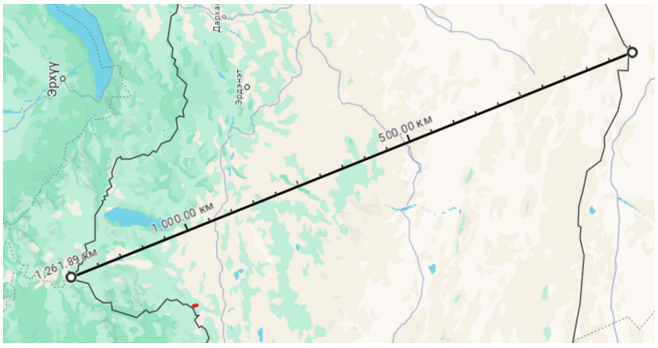


Fig. 12: Distance measurement from Google Maps

Fig. 13: Distance measurement using Haversine formula

Two locations within Mongolia were selected for this experiment, and the distance between them was calculated. According to Google Maps, the distance is 1,261.89 km. Our calculation using the Haversine formula yielded a result of 1,261.927738 km. The resulting measurement error is approximately 0.05 km, which is considered well within an acceptable range.

V. Conclusion and Future Works

This study presented the design, implementation, and experimental validation of a lightweight two-wheel drive (2WD) autonomous robot aimed at space-related applications. Emphasizing simplicity and energy efficiency, the robot integrates GPS-based localization, parachute-assisted landing, and autonomous navigation capabilities. Through the use of a ripstop nylon parachute with optimized geometry, stable descent at an average speed of 2.665 m/s was achieved. The application of the Haversine formula for distance estimation yielded results with a minimal error margin (0.05 km) when compared to Google Maps, confirming the reliability of the localization approach. Additionally, comparisons with u-blox satellite data validated the GPS module's accuracy within an acceptable precision range.

Future work will focus on enhancing terrain adaptability through additional sensors, improving heading estimation with real-time magnetometer data, and deploying the system in more complex environments to evaluate long-range navigation and fault tolerance under simulated extraterrestrial conditions.

References

[1] NASA Jet Propulsion Laboratory. (2020). Mars 2020 Mission Overview. Retrieved from https://mars.nasa.gov/mars2020/mission/overview/

[2] Siegwart, R., Nourbakhsh, I. R., & Scaramuzza, D. (2011). Introduction to Autonomous Mobile Robots. MIT Press.
[3] Yoshida, K., & Hamano, H. (2002). Motion dynamics of a rover with slip-based traction model. Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems.
[4] Thangavelautham, J., et al. (2012). Swarm autonomy for planetary surface exploration: mission and technology tradeoffs. AIAA SPACE 2012 Conference & Exposition.
[5] Sinnott, R. W. (1984). Virtues of the Haversine. Sky and Telescope, 68(2), 159.
[6] Barrett, S. F., & Pack, D. J. (2012). Embedded Systems: Design and Applications with the 68HC12 and HCS12 (2nd ed.). Pearson.
[7] Kaplan, E. D., & Hegarty, C. J. (2017). Understanding GPS/GNSS: Principles and Applications (3rd ed.). Artech House.
[8] Corke, P. (2017). Robotics, Vision and Control: Fundamental Algorithms in MATLAB (2nd ed.). Springer.
[9] Madgwick, S. O. H., Harrison, A. J. L., & Vaidyanathan, R. (2011). Estimation of IMU and MARG orientation using a gradient descent algorithm. IEEE International Conference on Rehabilitation Robotics (ICORR), 1–7.
[10] Petovello, M. G. (2009). GNSS Solutions: What are the differences between single- and dual-frequency GNSS receivers? Inside GNSS, 4(5), 18–21.
[11] Knacke, T. W. (1992). Parachute Recovery Systems Design Manual. Naval Weapons Center, U.S. Department of Defense.
[12] Pavone, M., Nesnas, I. A., & Schwager, M. (2013). Robot swarms in planetary exploration: Communication, coordination, and decision making. IEEE Robotics & Automation Magazine, 20(2), 20–29.

Бакалаврын судалгааны ажлын гүйцэтгэлд нөлөөлж буй хүчин зүйлсийг шинжилж, машин сургалтын арга ашиглан амжилтыг таамаглах нь

М.Амарзаяа
МУИС, МТЭС, Мэдээлэл,
Компьютерын Ухааны Тэнхим
23M1NUM0810@stud.num.edu.mn

Б.Наранбат
МУИС, МТЭС, Мэдээлэл,
Компьютерын Ухааны Тэнхим
naranbat@num.edu.mn

М.Золжаргал
МУИС, МТЭС, Мэдээлэл,
Компьютерын Ухааны Тэнхим
zoljargal@num.edu.mn

Хураангуй—Бакалаврын судалгааны ажил нь оюутны их сургуулийн боловсролын чанарыг илэрхийлж, эзэмшсэн мэдлэг, чадварыг бодит ажилд хэрэгжүүлэх чухал шалгуур болдог. Гэвч олон оюутан бакалаврын судалгааны ажлаа хугацаанд нь гүйцэтгэж чадалгүй, хангалтгүй үнэлгээ авч байгаа нь судалгааны үйл явцад нөлөөлж буй хүчин зүйлсийг судлах шаардлагыг бий болгож байна. Энэхүү судалгаанд 2023-2024 оны Монгол Улсын Их Сургуулийн Мэдээллийн Технологийн чиглэлийн дөрвөн хөтөлбөрөөр суралцаж буй оюутнуудын бакалаврын судалгааны ажлын өгөгдөл, суралцсан хугацааны хичээлийн дүнгийн өгөгдөл зэргийг цуглуулан боловсруулж өгөгдлийн багцыг (нийт 302 оюутан) бүрдүүлсэн. Бакалаврын судалгааны ажилд нөлөөлж буй хүчин зүйлсийг тодорхойлохдоо уламжлалт статистикийн шинжилгээ (корреляцийн шинжилгээ) болон өндөр үнэлгээ авах магадлалыг таамаглах зорилгоор машин сургалтын Random Forest болон XGBoost аргуудыг ашиглав. Үр дүнгээс харахад БСА-ын гүйцэтгэлд мэргэжлийн хичээлийн дүн, сэдвийн сонголт зэрэг нь тодорхой нөлөөтэй байгаа бол хөтөлбөрийн хувьд харьцангуй тогтвортой үнэлэгддэг гэж тодорхойлогдлоо. Мөн таамаглалын загваруудыг харьцуулахад XGboost 70 хувь, Random Forest 74 хувийн үнэн зөв, 0.78-ийн F1-score үзүүлсэн нь тус аргыг БСА-ын амжилттай гүйцэтгэлийг урьдчилан таамаглахад үр дүнтэй хэрэгсэл болохыг харуулж байна.

Түлхүүр үг—Өгөгдлийн шинжилгээ, бакалаврын судалгааны ажил, машин сургалтын арга, корреляцийн шинжилгээ

I. Удиртгал

Бакалаврын судалгааны ажил, ялангуяа төгсөлтийн ажил хэлбэрээр, өнөөгийн ихэнх хөтөлбөрийн гол бүрэлдэхүүн хэсэг болсон бөгөөд энэ нь бие даасан суралцах чадвар, шүүмжлэлт сэтгэлгээ, мэдлэгийн уялдаа холбоог хөгжүүлэхэд чухал нөлөөтэй гэж үздэг [1]. Дээд боловсролоос улбаатай ашиг тус нь зөвхөн хувь хүн, гэр бүлд бус харин улс оронд ч эерэгээр нөлөөлдөг ба боловсролд оруулсан хөрөнгө оруулалтын өгөөж нэмэгдэж, гэмт хэргийн түвшин буурч, татварын орлого нэмэгдэн, халамжийн хөтөлбөрөөс хамаарал багасах зэрэг олон давуу талыг бий болгодог [2]. Үнэндээ олон оюутан судалгааны ажлын явцад хүндрэлтэй тулгарч, улмаар ажлаа дутуу гүйцэтгэх, бүр эрдмийн зэрэг хүртэж чадахгүй байх асуудалтай тулгарч байна [3]. Ийм нөхцөл байдал нь зөвхөн тухайн оюутанд хохирол учруулаад зогсохгүй, тэнхим, удирдагч багшийн цаг хугацаа, сургуулийн нөөцийг үр дүнгүй зарцуулж, оюутнуудын мэргэжлийн замналд сөргөөр нөлөөлдөг. Бүрэн дуусаагүй судалгааны ажил нь оюутны хувьд үнэ цэнэтэй цаг хугацаа, санхүү, сэтгэлзүйн хүч хөдөлмөрийг үр дүнгүй болгож, нийгэм нийтээрээ ирээдүйтэй, өндөр ур чадвартай боловсон хүчнийг алдах эрсдэлийг бий болгоно [5]. Иймээс судалгааны ажлаа амжилттай хамгаалахад онцгой анхаарал хандуулж, урьдчилан төлөвлөсөн идэвхтэй арга хэмжээ авах нь зүйтэй юм. Төгсөлтийн ажлын төслүүдтэй холбоотой өмнөх судалгаанууд [1-8] нь оюутны гүйцэтгэлд оюутан-удирдагчийн харилцаа, хувь хүний болон санхүүгийн нөхцөл байдал, судалгааны орчин, удирдагчийн нөлөө, элсэлтийн шалгалтын дүн зэрэг олон хүчин зүйлс чухал үүрэгтэйг илрүүлсэн байдаг. Монгол улсын хувьд бакалаврын судалгааны ажлын гүйцэтгэлийг өгөгдөлд үндэслэн, шинжлэх, үнэлэх судалгааны ажил маш цөөн байна. Иймээс энэхүү судалгааны ажлаар бакалаврын судалгааны ажлын гүйцэтгэлд нөлөөлж болохуйц өгөгдлийг цуглуулж, өгөгдлийн шинжилгээ хийнэ. Тодруулбал, бид МУИС-ийн Мэдээллийн технологийн чиглэлийн дөрвөн (Мэдээллийн технологи, Программ хангамж, Компьютерын ухаан, Мэдээллийн систем) бакалаврын хөтөлбөрийн 2023-2024 оны бакалаврын судалгааны ажил гүйцэтгэсэн оюутнуудын сэдэв, удирдагч багш, судалсан хичээлүүдийн дүн, гүйцэтгэл (БСА-ын эцсийн дүн) зэрэг өгөгдлийг цуглуулж, боловсруулан статистикийн шинжилгээ хийж, машин сургалтын Random Forest, XGBoost арга ашиглан амжилтыг буюу А дүн авсан эсэхийг таамаглах туршилт хийсэн.

II. Сэдвийн судлагдсан байдал

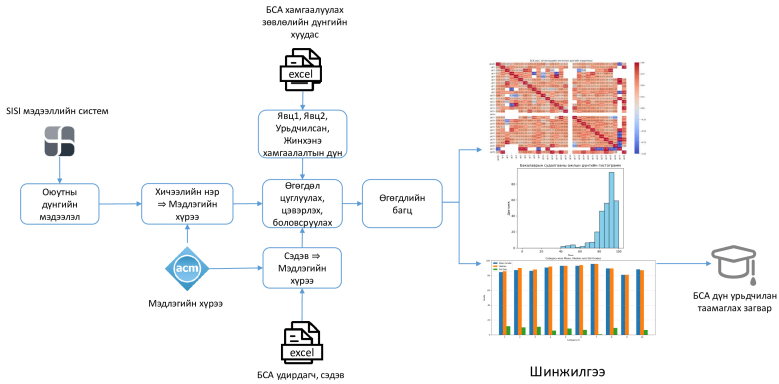
Төгсөлтийн ажлын төслүүдтэй (бакалавр, магистр, доктор) холбоотой өмнөх судалгаанууд нь оюутны гүйцэтгэлд нөлөөлдөг тодорхой хүчин зүйлсүүдийг илрүүлсэн байдаг. Судалгааны ажлын амжилтад оюутан болон удирдагчийн харилцаа, оюутны суралцах нөхцөл байдал, хичээлийн дүн, судалгааны орчин зэрэг олон хүчин зүйл нөлөөлдөг болохыг судлаачид тогтоожээ. Жишээлбэл, удирдагч багшийн удирдамж, хариуцлагатай харилцаа нь оюутны судалгааны ажлын чанар, хугацаанд нь дуусгах боломжид чухал нөлөөтэйг тэмдэглэсэн байна [2]. Мөн сургалтын дундаж голч оноо, эрдмийн зөвлөхийн зүгээс үзүүлэх дэмжлэг, сэдвийн сонирхол зэрэг нь магистр оюутнуудын төгсөлтийн ажлын амжилтад хэрхэн нөлөөлдөг болохыг судалсан [4]. Филиппин улсад хийсэн судалгаанд судалгааны хичээлийн дүн, бичгийн тайлан болон аман хамгаалалтын үр дүн нь нийт гүйцэтгэлийн үнэлгээнд голлох нөлөөтэй болохыг статистикийн аргаар баталсан байдаг [5]. Өөр нэг судалгаанд бакалаврын төгсөлтийн ажлыг амжилттай хамгаалах магадлалыг машин сургалтын аргууд (Decision Tree, SVM, Random Forest) ашиглан урьдчилан таамаглах оролдлого хийж, сургалтын өгөгдөл, оролцогчийн идэвх, удирдагчийн үнэлгээ зэргийн хамаарлыг шинжилсэн [7]. Тус судалгааны үр дүнгээс харахад багшийн мэргэжлийн туршлага болон судалгааны чиглэлтэй уялдсан сэдэв сонголт нь БСА-ыг амжилттай дуусгахад өндөр нөлөө үзүүлж байсан байна. Түүнчлэн, элсэлтийн шалгалтын оноо, өмнөх улирлуудын дүн, оролцогчийн идэвхи зэрэг хувьсагчийг ашиглан оюутны хугацаандаа төгсөх эсэхийг машин сургалтын аргаар таамагласан судалгаа ч бий [8]. Эдгээр судалгаанууд нь судалгааны ажлын амжилтыг зөвхөн нэг хүчин зүйл биш, харин олон талт хүчин зүйлсийн нөлөөллийн нийлбэрээр тодорхойлогддог болохыг нотолдог бөгөөд машин сургалтын аргуудыг ашигласнаар эдгээр хүчин зүйлсийн хамаарлыг илүү нарийвчлан шинжлэх боломжтойг харуулж байна.

III. Арга зүй

Бидний БСА-ын гүйцэтгэлд нөлөөлөх хүчин зүйлсийг судлах үйл явц нь өгөгдөл цуглуулах, өгөгдөл боловсруулах, өгөгдөл шинжлэх үе шатуудаас бүрдэнэ. Энэхүү арга зүйг Зураг 1-т үзүүлэв.

A. Өгөгдөл цуглуулах

Бид МУИС-ийн Мэдээлэл, Компьютерын Ухааны тэнхимийн Мэдээллийн технологи (МТ), Программ хангамж (ПХ), Компьютерын ухаан (КУ), Мэдээллийн систем (МС) бакалаврын хөтөлбөрийн 2023-2024 оны БСА гүйцэтгэсэн оюутнуудын өгөгдлийг гурван эх сурвалжаас нийт 302 оюутны мэдээллийг цуглуулсан. Үүнд:



Зураг. 1. БСА-д нөлөөлөх хүчин зүйлсийн шинжилгээний арга зүй

1) БСА гүйцэтгэх батлагдсан төлөвлөгөөнөөс оюутны удирдагч багш, сэдвийн мэдээлэл бүхий 302 өгөгдөл,
2) МУИС-ийн сургалт удирдлагын SISI системээс тус 302 оюутны 170 ялгаатай мэргэжлийн хичээлийн дүнгийн мэдээлэл бүхий 14854 өгөгдөл,
3) БСА хамгаалуулах зөвлөлийн дүнгийн хуудаснаас явц 1, явц 2, урьдчилсан хамгаалалт, шүүмжлэгч багшийн дүн, жинхэнэ хамгаалалтын дүн бүхий 302 өгөгдөл

B. Өгөгдөл боловсруулах

Бид цуглуулсан өгөгдлийг excel өргөтгөлтэй файл болгон хадгалж дараах боловсруулалтуудыг хийсэн. Үүнд:

- 2-р эх сурвалжаас цуглуулсан оюутны дүнгийн мэдээлэл дэхь хичээлүүдийн агуулгыг МТ-ийн хөтөлбөрүүдийн олон улсын зөвлөмж өгдөг байгууллага болох ACM/IEEE байгууллагаас гаргадаг Computing curricula 2020 баримт бичгийн Computing knowledge буюу мэдлэгийн хүрээтэй харгалзуулсан. Өөрөөр хэлбэл, тус хөтөлбөрөөр суралцдаг оюутнуудын судалдаг Программчлалын хэл хичээл нь тус зөвлөмжийн 5.4 Programming language дэд ангилал, 5. Software fundamentals үндсэн ангилалд харгалзах хичээл юм гэдэг харгалзааг хийсэн. Ингэснээр хөтөлбөрийн хичээлүүд нь үндсэн 6 ангилалд дэд 28 ангилалд харгалзах боломжтой болсон. Энэ судалгаанд дэд 28 ангиллаар авч үзсэн мөн нэг оюутан тухайн ангилалд харгалзах хичээлүүдээс олон хичээл судалсан бол тэдгээрийн дундаж дүнг тооцож харгалзуулсан
- 1-р эх сурвалжаас цуглуулсан сэдвийн ангиллыг мөн дээрхтэй адилаар 10 мэдлэгийн хүрээнд харгалзуулсан.

Нэмэлтээр суралцах хугацааны дундаж дүнг тооцоолж оруулсан. Тус судалгаанд ашиглах өгөгдлийн бүрдлийн дэлгэрэнгүй Хүснэгт 1-т үзүүлэв.

Хүснэгт I
Өгөгдлийн бүрдлийн шинж чанар

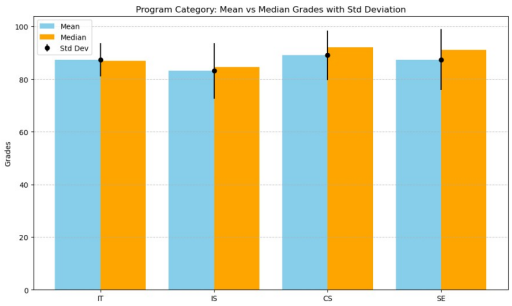
№	Өгөгдлийн бүрдэл	Өгөгдлийн төрөл	Агуулга, статистик			
1-р эх сурвалж						
1	Хөтөлбөр	Бичвэр	МТ, МС, КУ, ПХ			
2	Сэдэв	Бичвэр	Сэдвийн нэр, Жишээ нь: Орон сууцны захиалга, борлуулалтын программ			
3	Сэдвийн мэдлэгийн хүрээ	Бичвэр	Сэдвийн ангилал, Жишээ нь: Information Systems			
4	Багш	Бичвэр	Багшийн ID			
2-р эх сурвалж						
			Mean	STD	Min	Max
5	ЯВЦ1	Тоо	14.77	0.66	10.00	15.00
6	ЯВЦ2	Тоо	16.82	2.22	7.00	20.00
7	Урьдчилсан хамгаалалт	Тоо	20.49	3.27	6.00	25.00
8	Шүүмж	Тоо	4.30	1.01	0.00	5.00
9	Жинхэнэ	Тоо	30.22	6.03	0.00	35.00
10	БСА дүн	Тоо	86.83	10.25	40.30	100.00
3-р эх сурвалж						
	Category ID	Сэдвийн ангилал	Оюутны тоо	Дундаж дүн (Mean)	Медиан (Median)	Хазайлт (Std)
11	1	Information Systems	79	84.41	87	11.57
12	2	Software and its Engineering	145	87.31	90	9.99
13	3	Computing methodologies	37	86.19	88	10.88
14	4	General and reference	23	90.65	92	5.73
15	5	Security and privacy	2	93	93	8.46
16	6	Applied Computing	3	93	94	6.66
17	7	Theory of Computation	2	95.5	95.5	0.71
18	8	Mathematics of Computing	2	89.5	89.5	9.19
19	9	Hardware	1	81	81	NaN
20	10	Human-centered computing	7	88.29	87	6.55
3-р эх сурвалж						
Code	Хичээлийн ангилал			Correlation_with_BSA		
gc1	1.1. Social Issues and Professional Practice			0.28		
gc2	1.2. Security Policy and Management			0.06		
gc3	1.3. IS Management and Leadership			0.17		
gc4	1.4. Enterprise Architecture			0.35		
gc5	1.5. Project Management			0.3		
gc6	1.6. User Experience Design			0.23		
gc7	2.2. Systems Analysis & Design			0.38		
gc8	2.3. Requirements Analysis and Specification			0.36		
gc9	2.4. Data and Information Management			0.42		
gc10	3.2. Intelligent Systems (AI)			0.48		
gc11	3.5. Computer Networks			0.15		
gc12	3.7. Integrated Systems Technology			0.4		
gc13	3.8. Platform Technologies			0.37		
gc14	3.9. Security Technology and Implementation			0.15		
gc15	4.1. Software Quality, Verification and Validation			0.24		
gc17	4.3. Software Modeling and Analysis			0.2		
gc18	4.4. Software Design			0.34		
gc19	4.5. Platform-Based Development			0.41		
gc20	5.1. Graphics and Visualization			0.31		
gc21	5.2. Operating Systems			0.31		
gc22	5.3. Data Structures, Algorithms and Complexity			0.4		
gc23	5.4. Programming Languages			0.4		
gc24	5.5. Programming Fundamentals			0.3		
gc25	5.6. Computing Systems Fundamentals			0.24		
gc26	6.1. Architecture and Organization			0.47		
gc27	6.2. Digital Design			-1		
gc28	6.3. Circuits and Electronics			0.18		

Хүснэгт II
Өгөгдлийн статистик (Хөтөлбөрөөр)

Хөтөлбөр	Дундаж дүн	Медиан	Стандарт хазайлт
IT	87.29	87.0	6.30
IS	83.14	84.5	10.58
CS	89.04	92.0	9.32
SE	87.39	91.0	11.52

4. Хэрэгжүүлэлт ба үр дүн

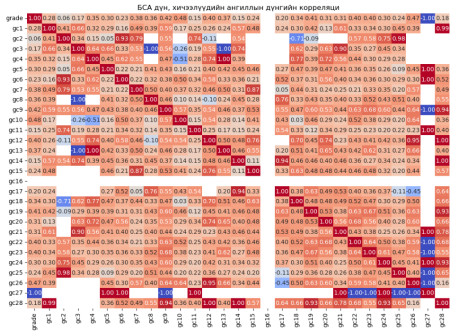
Бид боловсруулсан өгөгдлөөс БСА-ын гүйцэтгэлд нөлөөлөх хүчин зүйлсийг тодорхойлохдоо корреляцийн шинжилгээ болон статистикийн шинжилгээ хийсэн. Юуны түрүүнд өгөгдөлдөө статистик шинжилгээ хийж үзэв. Үр дүнг Хүснэгт 1 болон Хүснэгт 2-т харууллаа. Оюутнууд БСА-аар дунджаар 86.8 үнэлэгддэгээс хамгийн өндөр дундаж үнэлгээтэй хөтөлбөр нь КУ, хамгийн бага нь МС байна. (Зураг 2) Нийт өгөгдлийн 51% нь ПХ хөтөлбөрийн оюутнууд байна. Цаашлаад оюутнууд 10. Intelligent Systems дэд ангиллын хичээлүүдэд сайн дүн авч, 4. Software Development ангиллын хичээлүүдэд тааруу дүн авдаг байна. Түүнчлэн тус ангиллын стандарт хазайлт өндөр буюу оюутнууд харьцангуй ялгаатай дүгнэгддэг байна.



Зураг. 2. Хөтөлбөрөөр

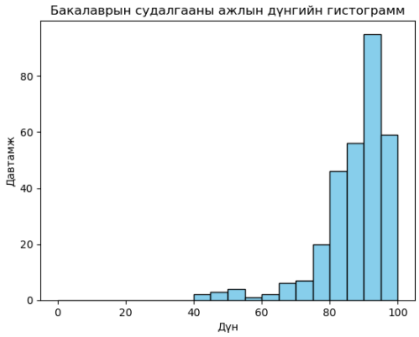
Цаашлаад бид хичээлүүдийн дүн, БСА-н дүн хоорондын хамаарлыг судлахын тулд корреляцийн шинжилгээ хийж үзэв. Үр дүнг Зураг 3-т үзүүллээ. Үр дүнгээс харахад БСА-ын дүнд хиймэл оюун ухаан, архитектур, өгөгдөл боловсруулах, системийн сэтгэлгээ, алгоритм, програмчлал зэрэг хичээлүүдийн ангилал хамгийн өндөр эерэг хамааралтай байв. Тухайлбал, Intelligent Systems (= 0.48), Architecture and Organization (= 0.47) зэрэг хичээлүүдийн ангилал нь судалгааны ажлын гүйцэтгэлд логик сэтгэлгээ, бүтэцчилсэн төлөвлөлт, техник хэрэгжилтийн чадвараар дамжин чухал нөлөө үзүүлж байна. Харин онолын чиглэлийн хичээлүүд нь БСА-д сул хамааралтай байна.

Судалгааны ажлын дүнгийн тархалтыг гистограммаар харууллаа (Зураг 4). Үр дүнгээс харахад судалгаанд хамрагдсан оюутнуудын дүнгийн дийлэнх нь 85 ба түүнээс дээш оноо авсан бөгөөд 90–95



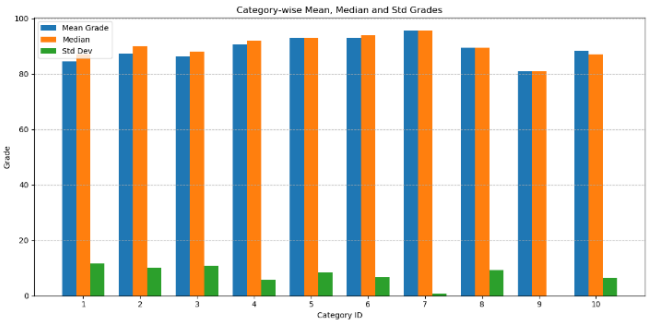
Зураг. 3. Корреляцийн шинжилгээ

онооны ангилал хамгийн өндөр давтамжтай байв. Энэ нь ихэнх оюутнууд судалгааны ажилдаа амжилттай оролцож, өндөр үнэлгээ авсныг харуулж байна.



Зураг. 4. БСА-ын дүнгийн тархалт

Бүх сэдвийн ангилалд дундаж болон медиан утгууд ойролцоо байгаа нь онооны тархалт жигд байгааг илтгэнэ. (Зураг 5.) –өөс харахад Theory of Computation буюу 7-р ангилалд хамгийн өндөр дундаж болон медиан оноотой, тархалт бага. Мөн 1–3-р ангиллууд нь стандарт хазайлт өндөр, өөрөөр хэлбэл тухайн ангилалд оюутнуудын оноо харьцангуй тархсан байна. Харин 9-р ангилал хамгийн бага дундаж оноотой байна. Ангиллын дэлгэрэнгүйг хүснэгт 1 –ээс харна уу.



Зураг. 5. Өгөгдлийн статистик (Сэдвээр)

5. Таамаглах загвар

Оюутны судалгааны ажлын үнэлгээг таамаглах зорилгоор боловсруулсан өгөгдөлд үндэслэн Random Forest болон сайжруулсан градиент бууралтын XGBoost алгоритмуудыг ашиглан ангиллын загвар боловсруулсан. Өгөгдлийг 70:30 харьцаагаар хуваан, 70 хувийг сургалтын, 30 хувийг туршилтын зориулалтаар ашигласан. Random Forest болон XGBoost зэрэг шийдвэрийн модонд суурилсан алгоритмууд нь бага хэмжээний өгөгдөлд илүү үр дүнтэй ажилладаг болохыг олон улсын судалгаанууд баталсан байдаг. [9,11,12] XGBoost болон Random Forest загваруудын хамгийн чухал 3 үзүүлэлтийг (Accuracy, Recall, F1-score) харьцуулан (хүснэгт 3) –аар үзүүллээ.

Хүснэгт III
XGBoost ба Random Forest алгоритмуудын гүйцэтгэлийн харьцуулалт

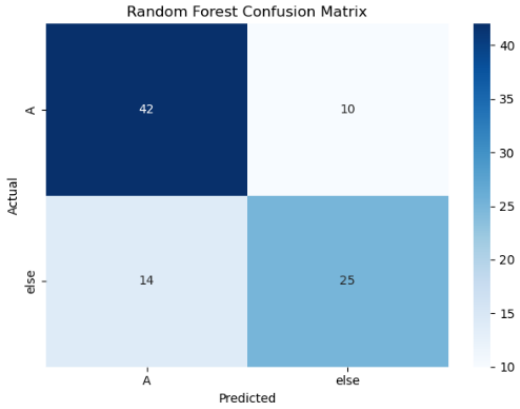
Үзүүлэлт	Тайлбар	XGBoost	Random Forest
Нарийвчлал (Accuracy)	Нийт зөв ангилсан хувь	0.70	0.74
Сэргэлт (Recall)	"A" ба "else" ангиллыг зөв таньсан хувь (дундаж утга)	0.68	0.72
F1-оноо (F1-score)	Нарийвчлал ба сэргэлтийн тэнцвэрт үзүүлэлт (дундаж утга)	0.69	0.78

Random Forest загвар нь XGBoost-ээс илүү сайн гүйцэтгэлтэй байгааг confusion matrix-аас харж болно. (Зураг 6, 7) Хоёр ангилалд харьцангуй сайн үр дүн үзүүлж байгаа бөгөөд ялангуяа А ангилал дээр өндөр гүйцэтгэлтэй байна. Random Forest нь бага өгөгдөлтэй нөхцөлд илүү найдвартай үр дүн гаргаж байгааг баталж байна.

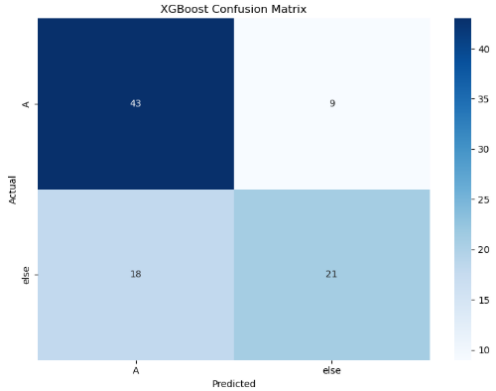
6. Дүгнэлт

Энэхүү судалгаанд БСА-д нөлөөлөх хүчин зүйлсийг өгөгдөлд суурилан статистикийн аргаар шинжилж машин сургалтын арга ашиглан А дүнг урьдчилан таамаглахыг оролдсон ажлыг танилцууллаа. Туршилтыг 302 оюутны өгөгдөлд тулгуурлан хийж гүйцэтгэсэн ба дараах үр дүнд хүрч байна. Үүнд:

1) Корреляцийн шинжилгээнээс харахад БСА-ын гүйцэтгэлд ялангуяа хиймэл оюун ухаан, архитектур, өгөгдөл боловсруулах, онолын хичээлүүд гэх мэт ангиллын хичээлүүд өндөр хамааралтай байгаа нь хөтөлбөрийн хувьд КУ хамгийн өндөр дүн авдаг гэдгийг баталж байна.
2) Бакалаврын судалгааны ажлын (БСА) гүйцэтгэлийг олон талаас нь шинжлэн судалснаар оюутнуудын дундаж дүн 86.83, медиан 90.0 гарсан нь гүйцэтгэл нийтлэг өндөр байгааг харууллаа. Гэвч стандарт хазайлт 10.26 байгаа нь зарим бүлэгт дүнгийн ялгаатай байдал оршиж байгааг илтгэв.
3) Эндээс КУ хөтөлбөр сайн оюутнууд ихтэй, МС оюутнууд дунд оюутнууд ихтэй гэсэн үр дүн гарч буй бөгөөд МС-ийн хөтөлбөрт яагаад сайн оюутнууд цөөн байгаад анхаарлаа хандуулж сайн оюутнуудын тоог нэмэгдүүлэх арга хэмжээ авах шаардлагатайг харуулж байна.
4) Харин машин сургалтын 2 аргаар А дүнг таамаглах туршилт хийхэд харьцангуй тэнцвэртэй үр дүн гарлаа. Random Forest загвар нь XGBoost-ээс илүү сайн гүйцэтгэлтэй байгааг тэдгээрийн confusion matrix-аас харж үнэллээ.



Зураг. 6. Random Forest үр дүн



Зураг. 7. XGBoost үр дүн

Цаашид бид БСА-ыг хийж эхлээд хамгаалаагүй оюутнуудад нөлөөлж буй хүчин зүйлсийг тодруулах,

машин сургалтын регрессийн арга ашиглан оюутны БСА-ын дүнг урьдчилан таамаглах судалгааны ажил хийхээр зорьж байна.

Номзүй

[1] Healey, Mick, Alan Jenkins, and John Lea. Developing research-based curricula in college-based higher education. York: The Higher Education Academy, 2013.

[2] Niemczyk, E. K., et al. (2020). The role of supervisor support in undergraduate research completion. Journal of Higher Education Research, 45(3), 234–247.

[3] Nouri, Jalal, Ken Larsson, and Mohammed Saqr. "Bachelor Thesis Analytics: Using Machine Learning to Predict Dropout and Identify Performance Factors." Int. J. Learn. Anal. Artif. Intell. Educ., 1.1 (2019): 116-131.

[4] Al-Kadri, H. M., et al. (2011). Factors affecting students' performance in undergraduate research projects in medical education. Medical Teacher, 33(12), e675-e679.

[5] Garcia, T., & De Guzman, A. (2010). Predictors of research paper performance among graduate students in the Philippines. Asia Pacific Education Review, 11(4), 559–570.

[6] Ngozi, Agu, and Oluwatayo Gbenga Kayode. "Variables attributed to delay in thesis completion by postgraduate students." Journal of Emerging Trends in Educational Research and Policy Studies, 5.1 (2014): 6-13.

[7] Rauf, F. A. "Challenges of thesis work: towards minimizing the non-completion rate in the postgraduate degree program." Concern, 8.7 (2016): 113-124.

[8] Ahmed, S., Khan, M. A., & Rehman, A. (2022). Using machine learning to predict timely graduation of university students. International Journal of Educational Technology in Higher Education, 19(1), 1–18.

[9] Biau, G., & Scornet, E. (2016). A random forest guided tour. TEST, 25(2), 197–227.

[10] Murtagh, Fionn, and Pedro Contreras. "Algorithms for hierarchical clustering: an overview." Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2.1 (2012): 86-97.

[11] Pedregosa, Fabian, et al. "Scikit-learn: Machine learning in Python." The Journal of Machine Learning Research, 12 (2011): 2825-2830.

[12] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.

Сүлжээний урсгалын шинжилгээнд суурилсан аюул илрүүлэх нь (radmin болон w32.rahack-ийн жишээн дээр)

Есөнжингийн ЭНХЗАЯА[1]
Мэдээллийн Сүлжээ, Аюулгүй Байдлын салбар
ШУТИС, Мэдээлэл, Холбооны Технологийн Сургууль
Монгол улс, Улаанбаатар
enkhzayayesunjin@gmail.com

Дондогмэгдийн БЯМБАДОРЖ[2]
Мэдээллийн Сүлжээ, Аюулгүй Байдлын салбар
ШУТИС, Мэдээлэл, Холбооны Технологийн Сургууль
Монгол улс, Улаанбаатар
d.byambadorj@must.edu.mn

***Хураангуй*— Сүлжээний траффикт суурилсан халдлага илрүүлэх болон аюулгүй байдлын үнэлгээ хийх нь өнөөгийн кибер орчны гол асуудлуудын нэг юм. Энэ судалгааны ажлын хүрээнд бодит орчинд илэрсэн сэжигтэй урсгалыг Wireshark болон Zenmap зэрэг хэрэгслүүдээр шинжилж, тухайн урсгалтай холбогдсон rserver3.exe файлыг илрүүлсэн. Энэхүү файл нь Radmin үйлчилгээтэй холбогдож, W32.Rahack төрлийн хортой программ болох нь тогтоогдсон. Судалгаанд ачаалалтай портуудын хяналт, үйлдлийн системийн процесс ID хяналт, мөн VirusTotal платформоор хийсэн статик шинжилгээний үр дүнд тулгуурласан. Илэрсэн халдлага нь IDS системүүдийн шинэ төрлийн халдлагыг илрүүлэх чадвар сул байгааг харуулсан бөгөөд шинж чанарт суурилсан илрүүлэлтийн аргыг сайжруулах шаардлагатайг илтгэж байна. Судалгааны дүн нь кибер аюулгүй байдлын хамгаалалтын шийдлийг боловсронгуй болгох, бодит өгөгдөлд суурилсан илрүүлэлтийн алгоритмын хөгжүүлэлтийн эхлэл болох ач холбогдолтой.**

Түлхүүр үг*— *сүлжээний урсгалын шинжилгээ, халдлага илрүүлэх систем, бодит халдлага, кибер аюулгүй байдал

I. Удиртгал

Мэдээллийн технологийн хурдацтай хөгжил нь байгууллагуудын дотоод болон гадаад үйл ажиллагааг дижитал хэлбэрт шилжүүлж, мэдээллийн аюулгүй байдлын асуудлыг стратегийн түвшинд авч үзэх шаардлагыг бий болгож байна. Энэхүү цахим шилжилт нь олон төрлийн кибер халдлагын эрсдлийг дагуулж байгаа бөгөөд эдгээрийг илрүүлэх, хариу арга хэмжээ авах технологийн гол бүрэлдэхүүн хэсэг нь Intrusion Detection System (IDS) юм. IDS нь сүлжээний орчин дахь сэжигтэй үйлдлийг илрүүлэхэд чухал үүрэгтэй боловч шинэ төрлийн халдлагад мэдрэмжгүй байх, хуурамч сэрэмжлүүлэг өгөх, системийн тохиргооны эмзэг байдалд өртөмтгий байх зэрэг асуудлууд гарсаар байна [1], [2].

Энэхүү судалгааны зорилго нь сүлжээний траффикаас сэжигтэй урсгалыг илрүүлэх практик шинжилгээ хийх, улмаар илэрсэн аюултай процессуудыг тодорхойлж, тэдгээрийн аюулгүй байдалд үзүүлэх нөлөөллийг дүгнэх явдал юм. Бодит орчны кейс судалгаанд Wireshark ашиглан траффик шинжилгээ хийж, Zenmap-ээр портуудыг сканнердаж, сэжигтэй файл болон түүний холбогдол бүхий үйлчилгээ (Radmin) болон хортой кодыг илрүүлсэн. Энэхүү туршилтаар IDS-ийн шинэ төрлийн халдлагад үзүүлэх гүйцэтгэл болон хязгаарлалтыг үнэлж, шинж чанарт суурилсан илрүүлэлтийн арга хэрэгслийг оновчтой болгох хэрэгцээг онцолсон болно.

Энэхүү судалгааны үр дүн нь сүлжээний бодит орчинд тулгуурласан халдлага илрүүлэх шинэ шийдэл боловсруулахад онолын болон практик үндэслэл болно.

II. Онолын үндэс

A. *IDS/IPS системүүд рүү чиглэсэн халдлагууд*

Орчин үеийн сүлжээний дэд бүтцийн хамгаалалтын гол бүрэлдэхүүн болох IDS (Intrusion Detection System) болон IPS (Intrusion Prevention System) нь халдлагын шинж чанартай урсгалыг илрүүлж, системийн бүрэн бүтэн байдлыг хамгаалахад чиглэсэн. Гэвч эдгээр систем нь өөрсдөө халдлагын бай болж, тодорхой эвдрэл, хуурмаг мэдээлэлд үндэслэн буруу шийдвэр гаргах, улмаар сүлжээний аюулгүй байдал алдагдах эрсдэлтэй байдаг. Халдагчид эдгээр хамгаалалтын системийг тойрон гарахад чиглэсэн төрөл бүрийн evasion буюу хуурмаг техникүүдийг ашигладаг нь судалгаагаар нотлогдож байна.

- IP packet fragmentation ашигласнаар халдагч өгөгдлийн урсгалыг хэсэгчлэн илгээж, IDS бүрэн payload-г уншиж чадахгүй нөхцөл бүрдүүлдэг. IP түвшинд өгөгдлийг хэсэгчлэн хуваан илгээх явцыг ашиглан IDS системийг хуурах боломжтой. Энэхүү фрагментацид суурилсан довтолгоо нь IDS системийн packet reassembly буюу бүрэн агуулгыг сэргээх чадварыг сулруулж, халдлагын кодыг таних боломжийг бууруулдаг. Жишээлбэл, Teardrop болон Tiny Fragment зэрэг халдлагууд нь өгөгдлийн сегментүүдийг зөв биш offset-тэйгээр дамжуулж, системийн задлан шинжилгээний механизмыг төөрөгдүүлэх стратеги дээр суурилдаг [1].
- TCP/IP протоколын олон янзын хэрэгжилт нь IDS/IPS системүүдийг төөрөгдүүлэх, сессийн төлөвийг гажуудуулах боломжийг олгодог. Халдагч нь TCP-ын SYN, FIN, RST зэрэг хянагч пакетуудыг ашиглан хуурамч холболт үүсгэх, эсвэл системийг буруу төлөвт шилжүүлэн, улмаар хууль бус урсгалыг мэдэгдэлгүй нэвтрүүлэх оролдлого хийдэг. Энэ төрлийн стратеги нь stateful inspection бүхий системүүдийг төөрөгдүүлж, хуурамч сесс үүсгэхэд ашиглагддаг бөгөөд Neu (2019)-ийн судалгаагаар SDN орчин дахь IDS системд онцгой эмзэг байдлыг үүсгэж байгааг харуулжээ [2].

- Хортой кодууд polymorphic шинжтэй болж өөрийгөө тасралтгүй өөрчилдөг болсон нь signature-based IDS системүүдийг үр дүнгүй болгож байна. Түүнчлэн HTTPS зэрэг шифрлэгдсэн протоколыг ашиглан Command-and-Control (C2) урсгалыг нууцлан дамжуулах нь payload-д суурилсан шинжилгээг бүрэн саармагжуулдаг. Энэхүү халдлагын үед зөвхөн URI, загвар, сессийн зан төлөвт суурилсан анализ хийж халдлагыг илрүүлэх шаардлагатай болдог [5].
- DNS болон ICMP зэрэг протоколыг ашиглан үндсэн халдлагын агуулгыг далд хэлбэрээр дамжуулахыг tunneling халдлага гэж нэрлэдэг. DNS tunneling-ийн хувьд TXT record ашиглан өгөгдөл дамжуулах, DNS query/response дунд actual мэдээлэл нуух эдгээрийг илрүүлэхэд IDS систем payload inspection хийх чадавхтай байх ёстой бөгөөд deep packet inspection (DPI) ихэнхидээ шифрлэлээс болж боломжгүй байдаг [6]

B. *NetFlow анализаторуудын сул тал*

NetFlow нь сүлжээний урсгалын мета өгөгдлийг бүртгэх зорилготой бөгөөд SIEM болон IDS системүүдтэй хамтран network behavior analysis хийхэд өргөн хэрэглэгддэг. Гэхдээ энэ нь цэвэр metadata дээр тулгуурладаг тул packet-level гүнзгий анализ хийх боломжгүй учраас дараах сул талуудтай. Үүнд:

- Урсгалын төөрөгдүүлэлт (Flow Manipulation): Халдагчид сүлжээний session-ийн хугацаа, портын дугаар, урсгалын чиглэл зэрэг параметрүүдийг зориудаар өөрчилснөөр NetFlow системийг буруу хандлагаар урсгал бүртгэхэд хүргэдэг. Vykopal (2013) үүнийг flow timeout-ууд болон session overlap-ийг ашиглан IDS болон NetFlow-ийн илрүүлэлтийг хэрхэн тойрон гарч болохыг харуулсан байна [9].
- Лог уялдуулалтын хүндрэл (Log Correlation Challenges):NetFlow нь зөвхөн IP, порт, протокол, хугацаа зэргийг хадгалдаг тул IDS-ээс үүссэн сэрэмжлүүлгийг NetFlow логтой уялдуулахад тодорхой хязгаарлалттай. Bråthen (2011)-ийн судалгаанд NetFlow-д бүртгэгдсэн урсгалын эхлэх ба дуусах timestamp нь IDS alert-тай бүрэн давхцдаггүйгээс болж SIEM correlation engine-ийн үр ашиг буурч байгааг тэмдэглэсэн [10].

C. *Нэгдсэн хамгаалалтын системүүд рүү чиглэсэн халдлагууд*

Сүүлийн үед байгууллагууд нэгдсэн хамгаалалтын архитектур буюу олон түвшний хамгаалалтыг (Defense in Depth) хэрэгжүүлэхдээ NGFW (Next Generation Firewall), IDS/IPS, SIEM, Honeypot, endpoint security зэрэг хэрэгслүүдийг нэгтгэн ашигладаг. Гэвч энэ төрлийн интеграцчилсан системүүд нь бүрэлдэхүүн хэсгүүдийн хоорондох тохиргоо, лог дамжуулалт, хариу үйлдлийн уялдаа сул байвал шинэ төрлийн довтолгоонд эмзэг болж хувирдаг.

1. SIEM төвтэй халдлагын синхрон давхцал нь Hasan (2024) санхүүгийн байгууллагын жишээн дээр SIEM, NGFW, болон honeypot системүүдийг нэгтгэн хэрэглэж байхад халдлагууд дараах байдлаар хийгдэж байсныг тэмдэглэсэн. Үүнд:

- Хуурамч сэрэмжлүүлэг үүсгэх
- Brute-force login flooding хийж SIEM-ийг хэт ачаалах
- Honeypot руу чиглэсэн халдлагаар хамгаалалтын багийг төөрөгдүүлэх [11].

2. IoT орчинд IDS/IPS/SIEM-ийн уялдааг алдагдуулах нь Yeboah (2024)-ийн судалгаанд дурдсанаар, Интернэтэд холбогдсон зүйлсийн орчин (IoT)-д ашиглагдаж буй IDS, IPS болон SIEM системүүдийн уялдаа, интеграцчлалын түвшин нь хэд хэдэн техникийн болон архитектурын хязгаарлалтаас шалтгаалан төгс гүйцэтгэлтэй ажиллахад хүндрэлтэй байна. Тухайлбал, эдгээр хамгаалалтын системүүдийн хоорондын мэдээллийн формат, урсгал дамжуулалтын хурд (latency), мөн үйлдлийн протоколын ялгаатай байдал нь IoT төхөөрөмжөөс үүссэн бодит цагийн (real-time) сүлжээний урсгалыг бүрэн дүн шинжилгээнд хамруулах боломжийг хязгаарлаж байна. Үүний улмаас тухайн орчинд явагдаж буй зальт урсгал (stealthy traffic) болон мөн чанараа нуусан халдлагууд нь IDS болон SIEM системд бүртгэгдэхгүй үлдэх, эсвэл саатаж илрэгдэх өндөр магадлалтай болж, сүлжээний хамгаалалтын үр ашигт шууд сөрөг нөлөө үзүүлдэг. Судалгааны үр дүнд эдгээр системийн хоорондын мэдээллийн солилцоо, стандартчиллын дутагдал, дасан зохицох чадварын хомсдол нь IoT орчны хамгаалалтын ямар эмзэг байгааг онцолсон байна [12]

III. Судалгааны аргачлал

A. *Судалгаанд ашигласан өгөгдөл ба орчин*

Энэхүү судалгаа нь их сургуулийн public утасгүй сүлжээний орчин дээр суурилсан бөгөөд бодит траффик өгөгдөлд анализ хийх, анормаль урсгал илрүүлэх, мөн боломжит халдлагын шинж чанарыг тодорхойлох зорилготой байв. Судалгааны эхний шатанд Wireshark хэрэгслийг ашиглан KOICA болон SICT SSID бүхий хандалтын цэгүүдээр дамжиж буй сүлжээний траффик өгөгдлийг барьж авсан.

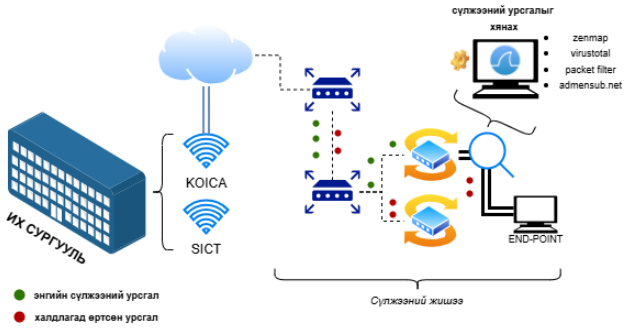
Үүний дараа Zenmap (Nmap) ашиглан сүлжээнд нээлттэй байгаа портуудыг скан хийж, хост төхөөрөмжүүдийн боломжит эмзэг цэгүүдийг тодорхойлсон. Хэмжилтийн үр дүнг статистик аргаар боловсруулж, хэвийн болон хэвийн бус сүлжээний урсгалын анализ хийснээр сэжигтэй хандалтыг ялган таних нөхцөлийг бүрдүүлсэн. Судалгаанд ашигласан гол хэрэгслүүд “Wireshark” траффик өгөгдөл барих, урсгалын анализ, “Zenmap, Nmap” порт скан нь хост тодорхойлох, “Advanced IP Scanner” IP хаяг тодорхойлох болон сүлжээний төхөөрөмжүүдийн жагсаалт гаргах, “Windows 11 laptop” туршилтын орчин, өгөгдөл боловсруулах платформ зэргийг ашигласан.

B. *Судалгааны арга зүй*

Тус судалгаанд хэрэглэсэн арга зүй нь туршилтын болон задлан шинжилгээний (experimental–analytical hybrid) хосолмол аргачлалд тулгуурласан бөгөөд судалгааны зорилго болох сүлжээний траффикт суурилсан аномали илрүүлэлтийг хэрэгжүүлэхэд оновчтой суурь болсон. Энэхүү аргачлалын хүрээнд бодит сүлжээний орчноос өгөгдөл цуглуулах, аюулын шинж чанарыг тодорхойлох, улмаар халдлагын боломжит шалтгааныг шинжлэн дүгнэх дараалсан алхмууд дараах байдлаар хэрэгжүүлэгдсэн. Үүнд:

- Сүлжээний бодит урсгалын өгөгдлийг барих — Wireshark ашиглан сүлжээний public SSID дээрх бодит урсгалыг хянаж, тухайн орчны өгөгдлийг агшин зуурт барьж авсан.
- Нээлттэй портуудыг скан хийх — Zenmap-ийг ашиглан сүлжээнд холбогдсон хостуудын портын байдлыг тодорхойлж, болзошгүй эмзэг байдлуудыг судалсан.
- IP болон сүлжээний хост тодорхойлох — Advanced IP Scanner ашиглан тухайн орчны IP хаягууд, төхөөрөмжийн MAC хаяг зэргийг бүртгэсэн.
- Аномали илрүүлэх — Цугларсан траффик өгөгдөлд статистик болон зан үйлийн шинжилгээ хийж, хэвийн бус загвар (анормаль хандалт) илрүүлэн, боломжит халдлагын төрлийг ангилан тодорхойлсон.

Судалгаа нь зөвхөн их сургуулийн public сүлжээний орчинд хийгдсэн тул зарим халдлагын эх үүсвэрийг гүнзгий мөрдөх боломж хязгаарлагдмал байсан. Халдлагын шалтгаан, хохирогч болон халдагчийн ID-level мэдээлэлд хүрэхийн тулд илүү өндөр хандалттай лог болон системийн түвшний өгөгдөл шаардлагатай болсон.

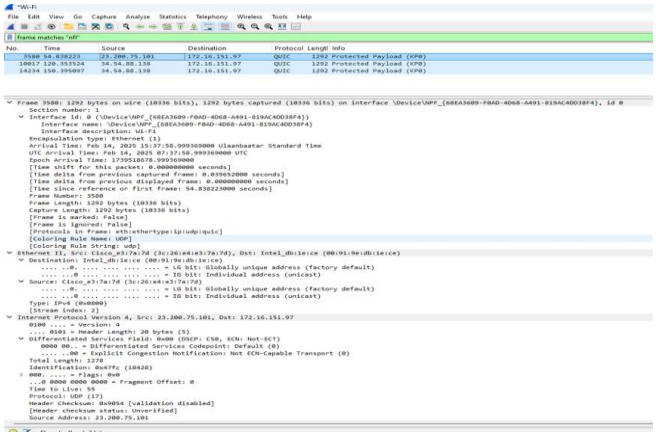


Зураг 1. Сүлжээний хяналтын орчны диаграмм

Зураг 1-т судалгаанд ашигласан сүлжээний ерөнхий бүтэц болон урсгалын шинжилгээний схемийг харуулсан. Зурагт ногоон өнгөөр хэвийн урсгал, улаанаар халдлагад өртсөн урсгал, мөн сүлжээний төвд байрлах хяналтын системүүдийг дүрсэлсэн бөгөөд эдгээр нь сүлжээний мониторинг, packet filter, host-based үзүүлэлтүүд, болон төхөөрөмжүүдийн урсгалын уялдаа-г хянахад чиглэгдсэн.

IV. ТУРШИЛТ, ҮР ДҮН

Судалгааны хүрээнд Wireshark хэрэгслийг ашиглан тухайн сүлжээнд дамжиж буй өгөгдлийг frame matches “nfl” гэсэн шүүлтүүрээр хязгаарлан шинжлэхэд, 23.200.75.10 хаягнаас 172.16.151.97 дотоод IP рүү чиглэсэн QUIC протоколын дамжуулалт илэрсэн нь Зураг 2-д харуулсан байна.



Зураг 2. IP 23.200.75.10-аас IP 172.16.151.97 хаяг руу QUIC өгөгдөл илгээж байгаа байдал

QUIC (Quick UDP Internet Connections) нь Google-ооc гаргасан, UDP-д суурилсан өндөр хурдтай холболтын протокол бөгөөд HTTP/3-т өргөн хэрэглэгддэг. Гэвч энэхүү протокол нь payload-оо бүрэн шифрлэдэг тул уламжлалт IDS/IPS системүүдэд payload inspection хийх боломжгүй болгодог. Иймд QUIC урсгал сэжигтэй байх нөхцөл дараах үндэслэлтэй. Үүнд:

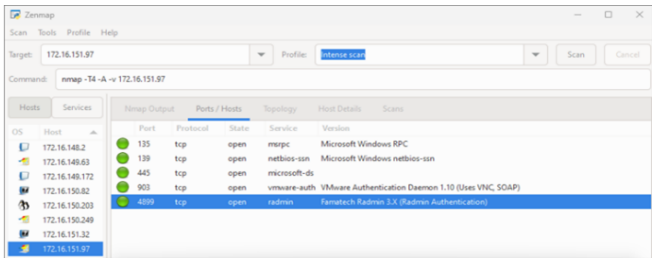
- Эх сурвалж IP (23.200.75.10) нь олон нийтийн мэдээллийн сангуудаар (threat intelligence DB) сэжигтэй буюу идвэхтэй командын серверт хамаарах магадлалтай.
- QUIC урсгал нь ижил хэмжээтэй (1292 bytes) пакетуудыг олон удаа, богино хугацаанд дарааллан илгээж буй нь data exfiltration эсвэл command-and-control механизмын магадлалтай үйлдэлд тооцогдож болно.
- Destination IP (172.16.151.97) нь дотоод сүлжээний хост тул, гадаад IP-тай шууд холбоо тогтоож буй энэ тохиолдол байгууллагын аюулгүй байдлын бодлоготой нийцэхгүй байж болзошгүй.

Зураг 2 дээрх Wireshark frame-ийн задрал нь дараах байдалтай байна. Үүнд:

- Frame timestamp: 2025.04.28 11:48:21 — богино хугацаанд давтагдсан ижил хэмжээтэй урсгал
- Source MAC: Dell Intel Ethernet адаптер
- Protocol: QUIC over UDP — TCP handshake үгүй тул IDS буруу ангилах магадлалтай.
- UDP Checksum: “unverified” — тухайн frame-ийн integrity-д эргэлзээтэй.

Эдгээр үзүүлэлтүүд нь анормал сүлжээний зан төлөв байж болзошгүйг харуулж байгаа тул тухайн сэжигтэй урсгалыг гүнзгийрүүлэн шалгаж, threat hunting буюу халдлагын гинжин холбоос (kill-chain) дагуу анализ хийх шаардлагатай.

Дээр дурдсан 172.16.151.97 хаягт Zenmap ашиглан intense scan хийсэн үр дүнд хэд хэдэн портууд нээлттэй байгааг илрүүлсэн бөгөөд тэдгээрийн дотроос TCP порт 4899-д Radmin 3.X (Famatech) үйлчилгээ ажиллаж байгааг Зураг 5-аар баталгаажуулсан.

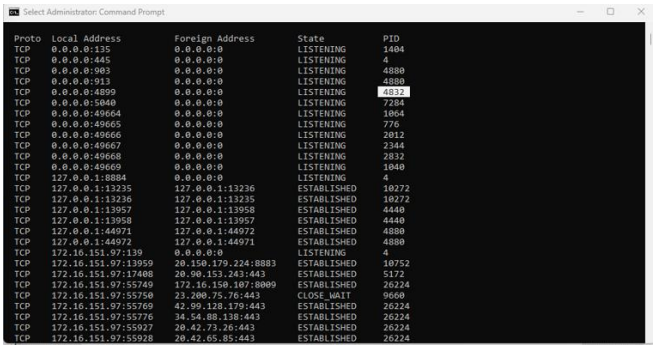


Зураг 3. Порт 4899 гэсэн radmin сэжиг бүхий программ

Radmin нь алсын удирдлагын хууль ёсны хэрэгсэл боловч дараах шалтгаанаар аюул үүсгэгч хүчин зүйл болж байж болно.

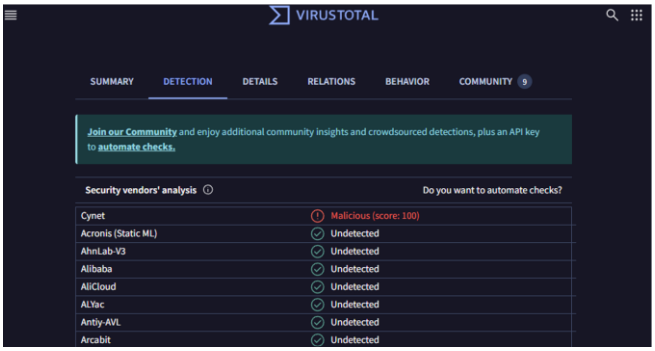
- Алсаас хяналт тогтоох боломжийг олгодог тул халдагчид backdoor суурилуулахад ашиглаж болзошгүй.
- CVE мэдээллээр баталгаажсан олон authentication bypass, remote code execution зэрэг эмзэг байдалтай (жишээ нь, [CVE-2017-15812]).
- QUIC урсгалаар дамжуулан C2 ажиллагаа гүйцэтгэх магадлалтай байж болно.

Сүлжээний орчинд 172.16.151.97 IP хаягтай хост сэжиг бүхий урсгал илгээж байгаа нь өмнөх Wireshark шинжилгээгээр илрүүлэгдсэн. Үүний дараа бид тухайн хост дээр дотоод холболтын төлөв байдлыг шалгах зорилгоор **netstat -ano** команд ашиглаж, LISTENING болон ESTABLISHED төлөвтэй TCP холболтуудыг задлан шинжиллээ. Зураг 3-ийн шинжилгээний явцад 4899 порт нь LISTENING төлөвтэй байгааг харуулж, түүнийг ажиллуулж буй PID 4832 нь Task Manager (Зураг 4)-т rserver3.exe нэртэй процесс байгааг тогтоосон.



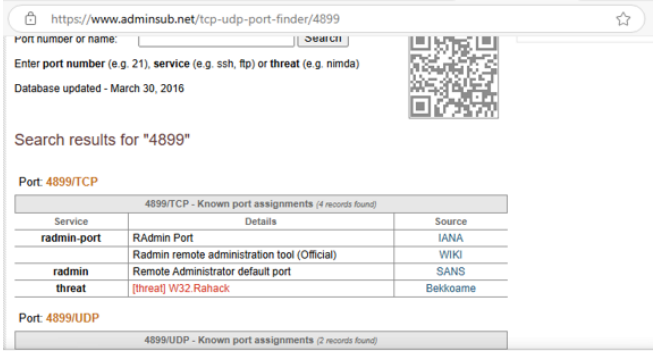
Зураг 4. Windows үйдлийн системийн сүлжээний холболтын төлөв байдал

Уг rserver3.exe нь Radmin Server буюу алсын хандалтын хэрэгсэлд хамаарах процесс бөгөөд энэ нь өмнөх Zenmap шинжилгээнд илэрсэн 4899 порт-той уялдаж байна.



Зураг 5. rserver3.exe программ нь хортой код агуулж байгаа

rserver3.exe файлыг www.virustotal.com платформд оруулж шалгахад Cynet аюулгүй байдлын системээс “Malicious (score: 100)” үнэлгээ авсан (Зураг 5). Энэ нь тухайн гүйцэтгэх файл нь W32 төрлийн хортой код агуулж болзошгүйг илтгэж байна.



Зураг 6. www.adminsub.net онлайнаар тандалт хийж байгаа байдал

TCP порт 4899-ийн үйл ажиллагааг баталгаажуулах зорилгоор www.adminsub.net онлайн платформ ашиглан портын утга, холбогдох үйлчилгээ, хортой программын мэдээллийг шалгасан. Зураг 6-аас дараах үр дүн илэрсэн. Үүнд:

- Порт 4899 нь албан ёсоор Radmin алсын удирдлагын програмын үндсэн порт болохыг IANA болон SANS байгууллагууд баталсан.
- Гэвч тухайн портыг W32.Rahack хэмээх нэртэй троян төрлийн хортой программ ашигладаг болохыг BekkAome нэртэй аюулгүй байдлын эх сурвалж онцлон тэмдэглэсэн байна (Зураг 6).
- “threat: W32.Rahack” гэсэн тодорхойлолт нь тухайн порт зөвшөөрөлгүй нэвтрэх сувгийн хувиар ашиглагдаж болзошгүйг илтгэж байгаа бөгөөд энэ нь өмнөх шинжилгээнд илэрсэн rserver3.exe процесс, PID 4832 болон Radmin service-ийн үйл ажиллагаатай бүрэн уялдаж байсан.

Дараах эрсдэлүүд үүссэнийг тодорхойлсон. Үүнд

- Backdoor хандалт: W32.Rahack нь системд алсаас хандах сувгийг автомат үүсгэх боломжтой.
- Хэрэглэгчийн хяналтгүй ажиллах: Программ нь системийн эрхээр ажиллаж байсан нь (Task Manager-оор батлагдсан) аюулын түвшинг нэмэгдүүлж байна.
- Сүлжээний далд урсгал (stealth traffic): Wireshark шинжилгээнд илэрсэн QUIC урсгалтай хосолж, сүлжээгээр хортой команд солилцох боломжтой.

Шинжилгээний үр дүнд дараах үндсэн ажиглалтууд гарч ирсэн:

- 172.16.151.97 хост нь гаднаас ирсэн сэжигтэй QUIC урсгалыг хүлээн авч байгаа нь тухайн системийн нууцлаг урсгалыг (stealth traffic) дамжуулах магадлалыг үүсгэсэн.
- Zenmap анализын үр дүнд 4899 порт дээр Radmin үйлчилгээ идэвхтэй байгааг тогтоосон. Уг үйлчилгээтэй холбогдсон rserver3.exe програм нь SYSTEM эрхтэй ажиллаж байсныг Task Manager баталсан.
- Тус программын VirusTotal дүн шинжилгээгээр “Malicious (score: 100)” оноо авсан нь W32.Rahack троянд хамаарагдах хортой код агуулсан болохыг илтгэсэн.

adminsub.net платформоос порт 4899 нь Radmin албан ёсны хэрэглээнээс гадна W32.Rahack хортой программын ашигладаг порт болохыг баталгаажуулсан. Adminsub.net тандалтын үр дүн нь судалгаанд илэрсэн порт 4899 болон түүнтэй холбоотой rserver3.exe процесс нь зөвхөн албан ёсны програм биш, мөн сэжигтэй хортой код агуулсан W32.Rahack-тай хамаарах хосолмол хэрэглээ бүхий аюултай үйлчилгээ байж болзошгүйг нотолж байна. Иймд тухайн системд forensic анализ, портын хандалтыг хаах, болон үргэлжлүүлэн sandbox-based шинжилгээ хийх нь зайлшгүй шаардлагатай.

Судалгааны үр дүнд үндэслэн гаргасан үндсэн санал:

1. IDS системүүдийг зөвхөн signature-based бус hybrid илрүүлэлтийн алгоритмаар сайжруулах шаардлагатай. Ялангуяа QUIC зэрэг шинэ протокол дээр deep inspection хийх чадварыг нэмэгдүүлэх шаардлагатай.
2. Radmin зэрэг хууль ёсны боловч хоёрдмол хэрэглээтэй програмуудын хандалтыг байгууллагын бодлогоор хязгаарлах, whitelist хийх механизм нэвтрүүлэх нь зохимжтой.
3. Хост түвшинд ажиллаж буй процессыг системийн болон хэрэглэгчийн эрхтэй байдлаар ялгаж, behavioral monitoring системүүдтэй холбох нь хортой процессийг эрт илрүүлэх боломжийг нэмэгдүүлнэ.
4. NetFlow логийг IDS мэдээлэлтэй үр дүнтэй уялдуулж, SIEM correlation engine-ийн гүйцэтгэлийг сайжруулах зорилгоор timestamp synchronization, metadata normalization хэрэгжүүлэх шаардлагатай.

V. ДҮГНЭЛТ

Энэхүү судалгаагаар бодит сүлжээний орчинд суурилсан траффик өгөгдөлд анализ хийж, аномали урсгалыг илрүүлэх, мөн түүний шалтгаан болсон үйлчилгээ болон программын сэжигтэй байдалд үндэслэн кибер аюулыг тодорхойлох туршилт хийгдсэн. Судалгаанд ашигласан Wireshark, Zenmap, Advanced IP Scanner, VirusTotal, adminsub.net зэрэг хэрэгслүүд нь сүлжээний хандалт, портын төлөв, үйлдлийн системийн процесс болон гүйцэтгэж буй програмыг олон талаас нь задлан шинжлэхэд үр дүнтэй байлаа.

Энэхүү судалгаа нь бодит орчинд суурилсан халдлага илрүүлэлтийн туршилтын загвар болохоос гадна, кибер аюулгүй байдлын хамгаалалтын шинэ үеийн шийдлийг боловсруулахад онолын болон практик үндэслэл болно. Цаашид sandbox болон real-time threat hunting орчинд энэ төрлийн урсгалыг илрүүлэх машин сургалтын алгоритм боловсруулах боломжтой юм.

АШИГЛАСАН МАТЕРИАЛ, НОМ ЗҮЙ

[1] A. Trisolino, Analysis of Security Configuration for IDS/IPS, Politecnico di Torino, 2023. [Online]. Available: https://webthesis.biblio.polito.it/29003/

[2] C. V. Neu, Detecting Encrypted Attacks in Software-Defined Networking, PUCRS, 2019. [Online]. Available: https://repositorio.pucrs.br/dspace/bitstream/10923/15855/1/000495778-Texto%2BCompleto-0.pdf

[3] J. Viega and G. McGraw, Building Secure Software: How to Avoid Security Problems the Right Way, Addison-Wesley, 2001.

[4] N. Alshammari and G. Moura, "Evading Network Intrusion Detection Systems with a TCP Fragmentation Attack," in IEEE Int. Conf. on Innovations in Information Technology, 2010. [Online]. Available: https://ieeexplore.ieee.org/document/5599875

[5] A. Abbas, “Subject review: Intrusion Detection System (IDS) and Intrusion Prevention System (IPS),” ResearchGate, Sep. 207

[6] J. Lee, M. Kim, and S. Park, “A deep learning-based network intrusion detection system: A comprehensive performance analysis,” Data in Brief, vol. 48, p. 109293, Apr. 2023.

[7] Cynet, “Network attacks and network security threats,” Cynet, [Online].

[8] Fortinet, “Intrusion Detection System,” Fortinet, [Online].

[9] J. Vykopal, Flow-Based Brute-Force Attack Detection in Large and High-Speed Networks, Masaryk University, 2013. [Online]. Available: https://is.muni.cz/th/haxyb/thesis.pdf

[10] A. Bråthen, Correlating IDS Alerts with System Logs by Means of a Network-Centric SIEM Solution, NTNU, 2011. [Online]. Available: https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/143982/Andreas%20Br%C3%A5then.pdf

[11] M. R. Hasan, Safeguarding of Financial Organization from Cyber-Attack using NGFW, SIEM and Honeypot, Dublin Business School, 2024. [Online]. Available: https://esource.dbs.ie/items/00dc6331-b807-429d-a63b-f1bb76a227d0

[12] F. A. Yeboah, Detecting and Safeguarding Against Cybersecurity Attacks Targeting Wireless Networks, OhioLINK, 2024. [Online]. Available: https://etd.ohiolink.edu/acprod/odb_etd/ws/send_file/send?accession=ucin1721230132561263&disposition=i

IoT ОРЧИНД НЭЭЛТТЭЙ ПОРТУУДЫН АЮУЛГҮЙ БАЙДЛЫН ШИНЖИЛГЭЭ

Батчулуун.Жавхлантөгс
Мэдээллийн Сүлжээ Аюулгүй Байдлын салбар
ШУТИС, Мэдээлэл, Холбооны Технологийн Сургууль
Монгол улс, Улаанбаатар
javkhlantugsb@must.edu.mn

Дондогмэгд Бямбадорж
Мэдээллийн Сүлжээ Аюулгүй Байдлын салбар
ШУТИС, Мэдээлэл, Холбооны Технологийн Сургууль
Монгол улс, Улаанбаатар
d.byambadorj@must.edu.mn

Бат-Эрдэнэ Мөнхбаяр
Мэдээллийн Сүлжээ Аюулгүй Байдлын салбар
ШУТИС, Мэдээлэл, Холбооны Технологийн Сургууль
Монгол улс, Улаанбаатар
munkhbayar.b@must.edu.mn

Хураангуй — Зүйлсийн интернэт (IoT) орчин нь олон төрлийн ухаалаг төхөөрөмжүүдийн холболтыг идэвхжүүлж, дижитал дэд бүтцийг өргөжүүлж буй ч аюулгүй байдлын ноцтой сорилтуудыг дагуулж байна. Энэхүү судалгаанд байгууллагын дотоод сүлжээний бодит орчинд Wireshark, Zenmap зэрэг хэрэгслүүдийг ашиглан IoT төхөөрөмжийн холболт, нээлттэй портууд болон протоколын түвшний урсгалыг ажиглав. Судалгааны үр дүнд, IP 192.168.0.173 хаягтай Android TV төхөөрөмж нь TCP 8008 порт ашиглан chrome.exe процесс руу автомат холболт үүсгэж байсан нь алсын хандалт, хяналтгүй мэдээлэл дамжуулалтын боломжийг харуулсан. Задлан шинжилгээгээр SYN/ACK хэт давтамжтай урсгал, хэрэглэгчийн оролцоогүйгээр үүссэн TCP холболтууд нь сүлжээний тагнуул, command-and-control халдлагын шинжтэй болохыг тогтоосон. Судалгаа нь IoT орчинд уламжлалт IDS/IPS болон SIEM системүүдийн уялдаагүй байдал, packet-level шинжилгээний хязгаарлалыг онцолж, зан үйлийн илрүүлэлт болон портуудын аудитын механизмыг сайжруулах шаардлагатайг харуулж байна.

Түлхүүр үг — IoT Security, Open Ports, Network Traffic Analysis, Wireshark, Intrusion Detection System, Botnet Attack

I. Удиртгал

Мэдээллийн технологийн дэвшил, ялангуяа Зүйлсийн интернэт (Internet of Things – IoT) технологи нь олон төрлийн ухаалаг төхөөрөмжийг сүлжээнд нэгтгэх боломжийг бий болгосноор хяналт, автоматжуулалт, өгөгдлийн солилцооны шинэ эрин үеийг эхлүүлж байна. Гэвч эдгээр боломжуудтай зэрэгцэн кибер аюулгүй байдлын эрсдэлүүд ч дагалдан гарч ирж, сүлжээний хамгаалалтын уламжлалт арга хэрэгслүүдийн үр ашигт хязгаарлалт үүсгэж байна.

IoT төхөөрөмжүүд нь ихэвчлэн багтаамж багатай, хөнгөн ажиллагаатай байдаг тул хамгаалалтын тогтолцоо, шинэчлэлтийн систем муу хөгжсөн байх нь элбэг. Үүний үр дүнд нээлттэй порт, буруу конфигураци, хуучирсан firmware зэрэг нь кибер халдлагын гол вектор болдог байна [1]. Түүнчлэн эдгээр төхөөрөмжүүд сүлжээнд холбогдохдоо нэмэлт баталгаажуулалтгүйгээр ажиллах нь их бөгөөд энэ нь халдагчдын хувьд амархан бай болох нөхцөлийг бүрдүүлдэг [2].

Сүлжээний аюулгүй байдалд IDS/IPS, SIEM, NetFlow зэрэг системүүд чухал үүрэг гүйцэтгэдэг боловч эдгээр системүүд нь шинэ төрлийн халдлагад бүрэн дасан зохицож чадахгүй байгаа нь сүүлийн үеийн судалгаагаар нотлогдож байна [3][4]. Тухайлбал, packet fragmentation, state desynchronization, log injection, anonymization зэрэг аргууд нь эдгээр хамгаалалтын системүүдийг төөрөгдүүлж, жинхэнэ халдлагыг нуух боломжийг олгодог [5].

Тиймээс энэхүү судалгааны зорилго нь байгууллагын дотоод сүлжээнд ажиллаж буй IoT төхөөрөмжүүдийн нээлттэй портуудын аюулгүй байдалд дүн шинжилгээ хийх, сэжигтэй урсгалыг packet-level түвшинд илрүүлж, боломжит халдлагын шинж чанар болон үр дагаврыг тодорхойлох явдал юм.

II. Сүлжээний халдлагын онолын үндэс

Сүлжээний траффикт суурилсан халдлагыг хоёр үндсэн бүлэгт хуваадаг: идэвхтэй (active) болон идэвхгүй (passive) халдлага. Идэвхгүй халдлага нь сүлжээний урсгалыг чимээгүй ажиглаж, мэдээлэл цуглуулах зорилготой байдаг бол идэвхтэй халдлага нь өгөгдөлд өөрчлөлт оруулах, устгах эсвэл үйлчилгээ тасалдуулах замаар системд шууд хохирол учруулдаг. Эдгээр халдлагын аль алиныг илрүүлэхэд Intrusion Detection System (IDS), Intrusion Prevention System (IPS), Security Information and Event Management (SIEM) системүүд чухал үүрэг гүйцэтгэдэг.

IDS/IPS системүүдийн уламжлалт signature-based илрүүлэлт нь зөвхөн урьдчилан тодорхойлогдсон загварт суурилсан учир шинэ төрлийн халдлагад мэдрэмжгүй байдаг [6]. Жишээлбэл, packet fragmentation ашиглан халдагч өгөгдлийг хэсэгчлэн илгээснээр IDS нь бүрэн payload-г сэргээж чадахгүйд хүрч, халдлагыг илрүүлэх боломжгүй болдог [7]. Мөн stateful inspection-д суурилсан системүүдийг TCP сессийн төлөвийг зориудаар будлиантуулах замаар төөрөгдүүлэх боломжтой [8].

SIEM систем нь олон төрлийн лог болон аюулгүй байдлын мэдээллийг нэгтгэн задлан шинжлэх чадвартай боловч логийн хуурамч өгөгдөл (log injection), мэдээлэл нуух (anonymization) зэрэг арга замаар төөрөгдүүлэх боломжтой. Энэ нь жинхэнэ халдлагыг нууж, хий үзэгдэл мэт өгөгдлөөр хамгаалалтын системийг сатааруулах нөхцөл бүрдүүлдэг [9], [10].

NetFlow нь сүлжээний урсгалын metadata-г бүртгэж, системийн зан үйлийн өөрчлөлтийг илрүүлэх боломж олгодог ч packet-level шинжилгээ хийх боломжгүй. Flow manipulation буюу урсгалын чиглэл, хугацаа, портын тохиргоог зориудаар өөрчилснөөр IDS болон SIEM системүүдийн корреляцийн чадварыг саармагжуулах боломжтой [11].

IoT орчинд нэгдсэн хамгаалалтын шийдлүүд (NGFW, IDS/IPS, SIEM) нь өөр хоорондоо лог болон системийн уялдаагүй байдлаас шалтгаалан халдлагыг цаг тухайд нь

илрүүлж чаддаггүй. Судалгаанаас харахад, IoT төхөөрөмжүүд дээрх протокол, latency, лог формат зэрэг ялгаанууд нь халдлагын гүнзгийрэлд хүргэх гол шалтгаан болдог [12]. Yeboah (2024) болон Hasan (2024) нарын судалгаанд IoT орчин дахь халдлагын уялдаа муу, real-time илрүүлэлт хангалтгүй байгааг онцолсон [13], [14].

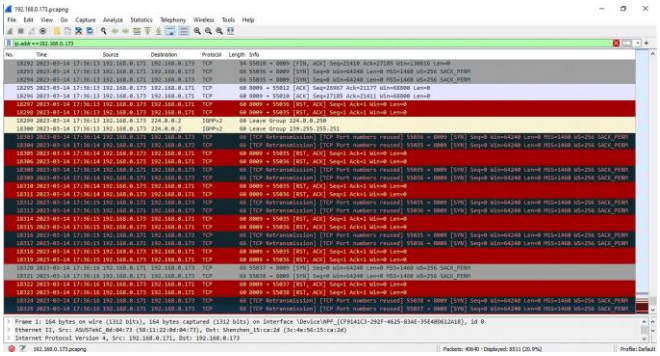
Иймээс энэхүү судалгааны зорилго нь сүлжээний бодит орчинд илрэх халдлагын урсгалыг packet-level түвшинд ажиглаж, хэрэглэгчийн хяналтгүйгээр үүсэж буй IoT-суурьт холболтыг тодорхойлох, нээлттэй портуудын аюулгүй байдлыг үнэлэх, цаашид IDS/IPS болон SIEM системүүдийн гүйцэтгэлийг сайжруулахад шаардагдах онолын болон практик нөхцөлийг бүрдүүлэх явдал юм.

III. Туршилтын үр дүн

Судалгааны ажлын хүрээнд байгууллагын дотоод сүлжээнд явагдаж буй хэвийн болон халдлагын шинжтэй траффик урсгалыг харьцуулан шинжилсэн. Туршилтын орчныг бүрдүүлэхдээ Intel Core i7 процессортой, 16 GB RAM бүхий компьютер дээр VirtualBox орчинд Kali Backtrack үйлдлийн системийг суурилуулж, аюулгүй байдлын туршилт болон тандалтын үйлдлүүдийг хийв.

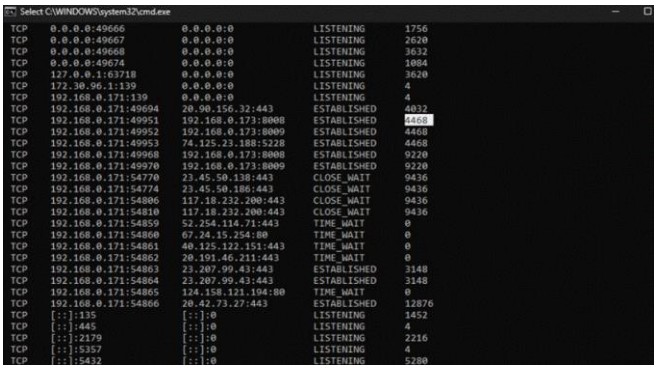
Сүлжээнд холбогдсон хостуудын портын төлөв байдал болон төхөөрөмжүүдийн бүртгэлийг Nmap хэрэгслээр тодорхойлж, сүлжээний ерөнхий бүтэц болон нээлттэй портуудыг үнэлсэн. Үүний дараа Wireshark 4.4.5 хэрэгслийг ашиглан TCP/IP түвшний пакет бүрд дүн шинжилгээ хийж, өгөгдлийн урсгал дахь хэвийн бус зан төлөвийг илрүүлэх зорилгоор packet-level мэдээллийг судалсан.

Хүлээн авсан болон илгээсэн протоколуудын задлан шинжилгээг хийж, IoT төхөөрөмжөөс үүсэж буй сэжигтэй хандалтын эх үүсвэрийг тодорхойлсны үндсэн дээр IP 192.168.0.173 хаягтай хост төхөөрөмжийг илрүүлсэн. Уг хаягийг Wireshark хэрэгслээр filter болон frame-ийн түвшинд нарийвчлан шалгахад, тухайн хост нь дотоод сүлжээний өөр нэгэн төхөөрөмж болох 192.168.0.171 хаяг руу SYN, ACK, RST төрлийн packet-үүдийг өндөр давтамжтайгаар илгээж байгааг тогтоосон (Зураг 1).



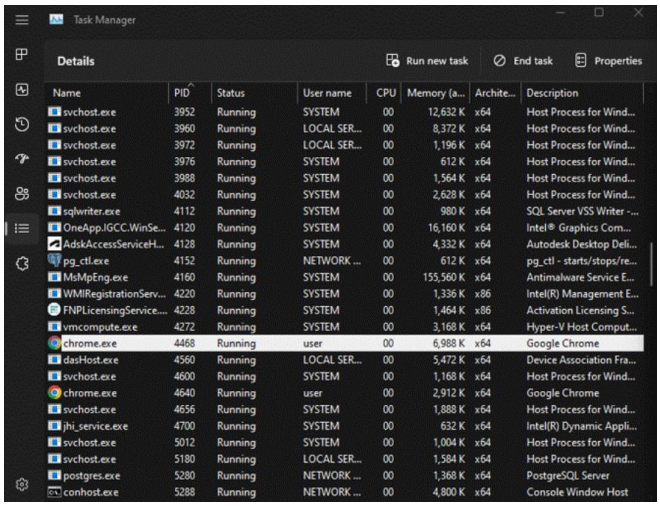
Зураг 1. IP 192.168.0.171 хаягтай хост руу IP 192.168.0.173 хаягаас тандалт хийж байгаа байдал

Эдгээр урсгал нь порт сканер хийх, сүлжээний зан үйлийг тандах, эмзэг байдал илрүүлэх зорилготой халдлагын урьдчилсан шатны үйлдэл байж болзошгүйг илтгэж байна. Хамгийн анхаарал татах үзүүлэлт бол ACK packet-уудын олон давтамжтай дахин илгээгдэлт бөгөөд энэ нь хостын төлөв байдлыг тандах зорилготой "retransmission"-д суурилсан сүлжээний тагнах үйлдэлтэй нийцэж байв. Зураг 1-д илэрсэн сэжигтэй SYN/ACK урсгал нь сүлжээний тагнуулын шинж чанартай, хостын эмзэг байдал болон портын нээлттэй байдлыг шалгах зорилготой байсан бол Зураг 2 нь уг хостын TCP холболтын төлөвийг харуулсан болно. netstat -ano командын үр дүнгээс харахад 192.168.0.171 хаяг нь 192.168.0.173 IP хаягтай хост руу 8008 дугаартай TCP порт ашиглан "ESTABLISHED" төлөвтэй холболт үүсгэсэн нь тодорхой байна. Энэ холболтыг үүсгэж буй PID 4468-ийг Task Manager-аар шалгахад chrome.exe процесс ажиллаж байгааг тогтоосон (дараагийн зурагт дэлгэрэнгүй). Энэ нь хэрэглэгчийн мэдэлгүйгээр веб суурьтай интерфейс ашиглан сүлжээнд нэвтрэх оролдлого байж болзошгүй гэсэн хардлагыг үүсгэж байна. Дээрх шинжилгээ нь сүлжээний тагнуулын шатанд хийгдсэн SYN scan болон хэрэглэгчийн төхөөрөмж рүү чиглэсэн алсын холболт (remote access) үүссэн байж болзошгүй нөхцөлийг баталгаажуулж байна. Ийм холболтууд нь сүлжээнд суурилсан command-and-control (C2) ажиллагааны хэсэг байх магадлалтай тул цаашид тухайн портуудын аюулгүй байдал, хэрэглэгчийн төхөөрөмжийн үйлчилгээний бүртгэлд дүн шинжилгээ хийх шаардлагатай.



Зураг 2 IP 192.168.0.173 хаягаас 8008 порт

Зураг 2-т харуулсан netstat гарцтай уялдуулан Task Manager-ийн дэлгэрэнгүй мэдээллийг шалгаж үзэхэд, PID 4468 нь chrome.exe процессид хамаарах болох нь тогтоогдсон. Энэ нь өмнөх TCP холболтын шинжилгээнд дурдсан 192.168.0.171 хаягт хамаарах хост нь 192.168.0.173 хаягтай төхөөрөмжтэй 8008 дугаарын порт ашиглан тогтсон холболтыг Chrome браузерын процесс ашиглан үүсгэсэн болохыг илтгэж байна. Энэ холболт нь хэрэглэгчийн оролцоогүйгээр үүссэн бол веб суурьт удирдлагын интерфейс, IoT төхөөрөмжийн хяналтын API, эсвэл command-and-control серверийн автомат холболт байх магадлалтай. Үүний үр дүнд 8008 дугаарын порт нь зөвхөн веб хэрэглээ бус, алсын удирдлага эсвэл халдлагын сувгийг нээсэн байж болохыг харуулж байна. Ийм төрлийн холболт, хэрэглэгчийн хяналтгүйгээр хийгдсэн бол Chrome браузер-ын process ID (PID)-г ашиглан нэмэлт dynamic forensics буюу memory dump, browser history, network socket activity зэрэг шинжилгээг хийх шаардлагатай болсон.



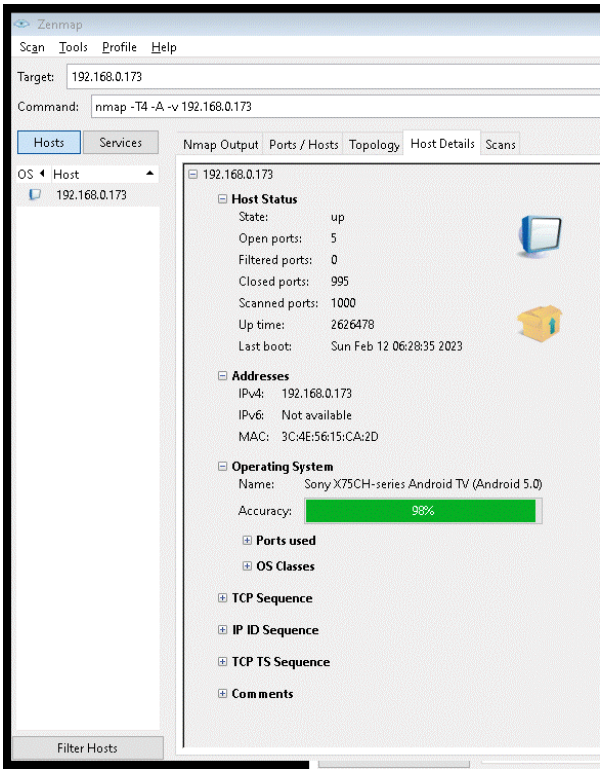
Зураг 3 Chrome.exe файл уруу хандалт хийж байгаа байдал

Зураг 3-т Task Manager дээр ажиллаж буй процессуудын жагсаалтаас chrome.exe нэртэй процесс нь PID 4468 дугаартай бөгөөд хэрэглэгчийн эрхтэйгээр (user) ажиллаж буйг харуулж байна. Энэ нь өмнө netstat -ano командыг ашиглан илрүүлсэн TCP 8008 порттой холбогдсон холболттой уялдаж байна. Ийм холболт үүсэж буй нь тухайн систем дээр Chrome хөтөч рүү гадаад хост (192.168.0.173) TCP 8008 портоор холболт хийж байгааг баталж байгаа юм. Энэхүү үйлдэл нь:

IoT төхөөрөмж (192.168.0.173) нь Chrome.exe процесс руу мэдээлэл дамжуулж буй үйл явц байж болзошгүйг,

Хост түвшинд ямар хэрэглэгчийн процесс руу холболт хийгдэж байгааг тодорхойлох,

Аюулгүй байдлын үүднээс процесс хандалтын замаар C2 (Command and Control) болон data exfiltration халдлагын магадлалыг үнэлэх боломжийг олгодог.



Зураг 4. Zenmap-ээр шалгасан IoT төхөөрөмжийн тодорхойлолт

Zenmap хэрэгслийг ашиглан байгууллагын дотоод сүлжээнд холбогдсон IoT төхөөрөмжүүдийн шинжилгээ хийсэн бөгөөд үүний үр дүнд IP 192.168.0.173 хаягтай Sony XXXXXXXXX Android TV (Android 5.0) төхөөрөмж илрэв. Уг төхөөрөмж нь 8008, 8009, 8080, 8443, 9000 зэрэг нийт 5 нээлттэй TCP портуудтай байсан бөгөөд эдгээр нь HTTP/HTTPS болон алсын удирдлагын зориулалтаар өргөн ашиглагддаг болох нь тогтоогдсон.

Нээлттэй портуудын дунд 8008 порт нь хэрэглэгчийн хяналтгүйгээр Chrome.exe процесс руу автоматаар холболт үүсгэж байсан нь тус төхөөрөмж сүлжээнд тандалт хийж, алсын хандалттай холбогдох боломжтойг илтгэж байна. Task Manager болон netstat хэрэгслүүдийг хослуулан задлан шинжилж үзэхэд, уг холболт нь хэрэглэгчийн оролцоогүйгээр үүссэн байж болзошгүй тул сүлжээний тагнуул, Command-and-Control (C2) халдлагын урьдчилсан нөхцөл байж болохыг харуулсан.

Үр дүнгээс харахад IoT төхөөрөмжүүдийн хамгаалалтгүй, аудит хийгдээгүй портууд нь байгууллагын сүлжээний аюулгүй байдалд өндөр эрсдэл учруулж болзошгүй бөгөөд халдагчид эдгээр портыг ашиглан ботнет суурилуулах, DDoS халдлага явуулах нөхцөл бүрдэж байна. Иймд цаашид IDS/IPS системийн идэвхтэй уялдаа, портын хандалтын мониторинг, IoT төхөөрөмжүүдийн аудит болон аюулгүй байдлын тохиргоонд тогтмол хяналт тавих шаардлагатай байна.

Дүгнэлт

Энэхүү судалгаагаар байгууллагын дотоод сүлжээнд үүсч болзошгүй сүлжээний халдлагын урсгалыг илрүүлэх, задлан шинжлэх, аюулгүй байдлын үнэлгээ хийх боломжийг практик орчинд бодит хэрэгслүүдийн тусламжтайгаар туршин судаллаа. Wireshark болон Zenmap ашиглан урсгалын түвшинд хийгдсэн тандалт, скан шинжилгээгээр TCP 8008 портын дамжуулалт, chrome.exe процесс руу чиглэсэн холболт зэргийг илрүүлж, уг урсгалын эх үүсвэр болох IoT төрлийн смарт телевизор төхөөрөмжийг тодорхойлсон. Уг төхөөрөмж дээр илэрсэн олон нээлттэй порт нь (8008, 8080, 8443 гэх мэт) халдлагад өртөмтгий байдал нэмэгдүүлж, ботнет суурилсан DDoS халдлага хийх суурь нөхцөлийг бүрдүүлж байгааг дүгнэлээ.

Судалгааны үр дүн нь уламжлалт IDS/IPS болон SIEM системүүдийн илрүүлэлтийн хязгаарлалтыг тодруулж, IoT төхөөрөмжийн эмзэг байдал, систем хоорондын уялдаагүй байдал нь халдлагын гүнзгийрэлтэд хэрхэн нөлөөлж болохыг харуулласа. Цаашид хамгаалалтын орчныг сайжруулахын тулд зан үйлийн анализ, real-time корреляци хийх чадвартай илрүүлэлтийн алгоритмыг нэвтрүүлэх, сүлжээний төхөөрөмжүүдийн конфигурацийг тогтмол шалгаж сайжруулах нь зайлшгүй чухал гэж дүгнэв.

Талархал

Энэхүү эрдэм шинжилгээний өгүүллийг бичихэд гүн туслалцаа үзүүлсэн МХТС-ийн МСАБ салбарын эрхлэгч, доктор (Ph.D), дэд профессор Б.Мөнхбаяр, МСАБ салбарын дэд профессор Б.Дэнсмаа, удирдагч багш доктор (Ph.D), ахлах багш Д.Бямбадорж багш нартаа талархсанаа илэрхийлье.

[1] M. Abomhara and G. M. Køien, “Security and privacy in the Internet of Things: Current status and open issues,” IEEE PRISMS, 2014.

[2] S. Sicari et al., “Security, privacy and trust in Internet of Things: The road ahead,” Computer Networks, vol. 76, pp. 146–164, 2015.

[3] A. Trisolino, Analysis of Security Configuration for IDS/IPS, Politecnico di Torino, 2023.

[4] S. Zeinali, SIEM Evasion and Detection Methods, 2016.

[5] M. F. Andersen, Detecting Malware and Cyber Attacks Using ISP Data, Aalborg University, 2022. [1] A. Trisolino, Analysis of Security Configuration for IDS/IPS. Politecnico di Torino, 2023. [Online]. Available: https://webthesis.biblio.polito.it/29003/

[6] C. V. Neu, Detecting Encrypted Attacks in Software-Defined Networking. PUCRS, 2019. [Online]. Available: https://repositorio.pucrs.br/dspace/bitstream/10923/15855/1/000495778-Texto%2BCompleto-0.pdf

[7] J. Viega and G. McGraw, Building Secure Software: How to Avoid Security Problems the Right Way. Addison-Wesley, 2001.

[8] N. Alshammari and G. Moura, "Evading Network Intrusion Detection Systems with a TCP Fragmentation Attack," in Proc. IEEE Innovations in Information Technology (IIT), 2010. [Online]. Available: https://ieeexplore.ieee.org/document/5599875

[9] J. Nazario, "Polymorphic shellcode analysis," Symantec Security Response Whitepaper, 2006. [Online]. Available: https://www.symantec.com

[10] M. F. Andersen, Detecting Malware and Cyber Attacks Using ISP Data. Aalborg University, 2022. [Online]. Available: https://vbn.aau.dk/files/549497052/phd_MFA_e_pdf.pdf

[11] S. M. Zeinali, Analysis of Security Information and Event Management (SIEM) Evasion and Detection Methods, 2016. [Online]. Available: http://mendillo.info/seguridad/tesis/Morteza.pdf

[12] M. F. Andersen, Detecting Malware and Cyber Attacks Using ISP Data, Aalborg University, 2022. [Online]. Available: https://vbn.aau.dk/files/549497052/phd_MFA_e_pdf.pdf

[13] J. Vykopal, Flow-Based Brute-Force Attack Detection in Large and High-Speed Networks. Masaryk University, 2013. [Online]. Available: https://is.muni.cz/th/haxyb/thesis.pdf

[14] A. Bråthen, Correlating IDS Alerts with System Logs by Means of a Network-Centric SIEM Solution. NTNU, 2011. [Online]. Available: https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/143982/Andreas%20Br%C3%A5then.pdf

[15] M. R. Hasan, Safeguarding of Financial Organization from Cyber-Attack using NGFW, SIEM and Honeypot. Dublin Business School, 2024. [Online]. Available: https://esource.dbs.ie/items/00dc6331-b807-429d-a63b-f1bb76a227d0

[16] F. A. Yeboah, Detecting and Safeguarding Against Cybersecurity Attacks Targeting Wireless Networks. OhioLINK, 2024. [Online]. Available: https://etd.ohiolink.edu/acprod/odb_etd/ws/send_file/send?accession=ucin1721230132561263&disposition=inline