6. 차세대보안

6 차세대보안

□ 기획과제안

(단위 : 억원)

순번	과제명	총 수행 기간	'21년 (총) 출연금	공모 방식	연구 단계 (TRL)	주관 기관	비고
【 차세대보안 】							
1	저사양 디바이스 대상 고효율 PQC 안전성 및 성능 검증 기술 개발 ▸ **(특징)** 혁신성장동력R&D	4년	15 (75)	지정 공모	응용 (3~5)	제한 없음	
2	국가공공 정보시스템 안전성 및 활용성 제고를 위한 차세대 암호체계 개발 ▸ **(특징)** 혁신성장동력R&D, 다부처협력(국정원) ▸ **(단계)** 원천기술확보(2년) → 응용연구(2년)	4년 (2+2)	19.02 (94.02)	지정 공모	응용 (3~6)	제한 없음	
3	간결한 비대화형 연산 증명 시스템(SNARK)을 위한 암호 원천 기술 개발 ▸ **(특징)** 혁신성장동력R&D, 고위험도전(전문연구실) ▸ **(단계)** 타당성연구(3년) → 기술확보(3년)	6년 (3+3)	4 (29)	지정 공모	응용 (2~5)	대학	
4	범죄증거 확보를 위한 암호분석 기술 고도화 ▸ **(특징)** 다부처협력(대검찰청,경찰청), 사회문제해결	3년	9 (33)	지정 공모	응용 (4~6)	제한 없음	
5	HW지원 프라이버시 보장 암호데이터 고속처리 기술개발 ▸ **(특징)** 혁신성장동력R&D, 고위험도전(선도형) ▸ **(단계)** 원천기술확보(2년) → 응용연구(2년)	4년 (2+2)	19 (94)	지정 공모	응용 (3~6)	제한 없음	
6	동형암호기술 활용 데이터 프라이버시 보존 국가통계 분석시스템 개발 구축 ▸ **(특징)** 디지털뉴딜R&D, 다부처협력(통계청)	3년	12 (44)	지정 공모	개발 (4~7)	제한 없음 (통계청 참여)	
7	영상 등 멀티미디어 데이터의 온전한 AI 학습/활용이 가능한 복원불가형 개인식별정보 비식별 핵심기술 개발 ▸ **(특징)** 디지털뉴딜R&D, 다부처협력(보호위) ▸ **(단계)** 원천기술확보(2년) → 응용연구(2년)	4년 (2+2)	12 (57)	지정 공모	응용 (3~6)	제한 없음 (산업체 참여 필수)	
8	대용량 정형 데이터 대상 비식별처리 자동화 및 안전성 검증 기술개발 ▸ **(특징)** 디지털뉴딜R&D, 다부처협력(보호위), 고위험도전(선도형) ▸ **(단계)** 원천기술확보(2년) → 응용연구(2년)	4년 (2+2)	12 (57)	지정 공모	응용 (3~6)	제한 없음 (산업체 참여 필수)	
9	5G Massive 디바이스 공격접점 은닉을 통한 차세대 사이버공격 기만기술 개발 ▸ **(특징)** 디지털뉴딜R&D, 고위험도전(선도형)	4년	9 (45)	지정 공모	응용 (4~6)	제한 없음 (산업체 참여 필수)	

순번	과제명	총 수행 기간	'21년 (총) 출연금	공모 방식	연구 단계 (TRL)	주관 기관	비고
【 차세대보안 】							
10	위협헌팅 모델 기반 지능형 사이버 공격/방어 분석 프레임워크 기술 개발 ▸ **(특징)** 혁신성장동력R&D	4년	12 (57)	지정 공모	응용 (4~6)	제한 없음	
11	AI·빅데이터 기반 사이버 보안 오케스트레이션 및 자동 대응 기술 ▸ **(특징)** 디지털뉴딜R&D ▸ **(단계)** 원천기술확보(2년) → 응용연구(2년)	4년 (2+2)	19 (94)	지정 공모	응용 (4~6)	제한 없음 (산업체 참여 필수)	
12	언택트 시대의 기업망 보호를 위한 제로트러스트 기반 접근제어 및 이상징후 분석기술 개발 ▸ **(특징)** 디지털뉴딜R&D, 고위험도전(선도형) ▸ **(단계)** 원천기술확보(2년) → 응용연구(2년)	4년 (2+2)	15 (75)	지정 공모	응용 (4~6)	제한 없음 (산업체 참여 필수)	
13	상시적 보안품질 보장을 위한 6G 자율보안 내재화 기반기술 연구 ▸ **(특징)** 혁신성장동력R&D, 고위험도전(선도형)	4년	15 (75)	지정 공모	기초 (2~3)	제한 없음	
14	비대면 환경의 보안 편의성 개선을 위한 Usable Security 기술 개발 ▸ **(특징)** 디지털뉴딜R&D, 사회문제해결, 경쟁형 ▸ **(단계)** 원천기술확보(2년) → 응용/상용화(2년)	4년 (2+2)	15 (65)	지정 공모	개발 (3~7)	제한 없음	
15	임베디드 시스템 악성코드 탐지·복원을 위한 RISC-V 기반 보안 CPU 아키텍처 핵심기술 개발 ▸ **(특징)** 혁신성장동력R&D, 고위험도전(선도형) ▸ **(단계)** 원천기술확보(2년) → 응용연구(2년)	4년 (2+2)	15 (75)	지정 공모	응용 (4~6)	제한 없음	
16	고신뢰 온-디바이스 딥러닝 가속기 설계를 위한 물리채널 기반 취약점 검증 및 대응기술 개발 ▸ **(특징)** 혁신성장동력R&D, 고위험도전(선도형) ▸ **(단계)** 원천기술확보(2년) → 응용연구(2년)	4년 (2+2)	9 (45)	지정 공모	응용 (3~6)	제한 없음	
17	개인정보보호를 위한 신뢰계산 기반 데이터보호박스 개발 ▸ **(특징)** 혁신성장동력R&D, 고위험도전(전문연구실) ▸ **(단계)** 타당성연구(3년) → 기술확보(3년)	6년 (3+3)	4 (29)	지정 공모	응용 (3~6)	대학	
18	실환경 기반 마스크 착용자 얼굴인식 및 재인식(Re-ID) 기술 ▸ **(특징)** 디지털뉴딜R&D, 고위험도전(선도형)	3년	12 (42)	지정 공모	개발 (4~7)	제한 없음 (산업체 참여 필수)	
19	다중 스펙트럼 영상 통합 분석 기술 및 영상 감시 장치 개발 ▸ **(특징)** 디지털뉴딜R&D	3년	12 (42)	지정 공모	개발 (5~7)	제한 없음 (산업체 참여 필수)	

순번	과제명	총 수행 기간	'21년 (총) 출연금	공모 방식	연구 단계 (TRL)	주관 기관	비고
【 차세대보안 】							
20	무인점포 환경 대응형 2D/3D 영상 통합 분석기반 지능형 영상보안시스템 기술 개발 ▸ **(특징)** 디지털뉴딜R&D	3년	12 (44)	지정 공모	개발 (4~7)	제한 없음 (산업체 참여 필수)	
21	신뢰 가능한 엣지AI 시스템 검증 플랫폼 핵심기술 및 시험기술 개발 ▸ **(특징)** 혁신성장동력R&D, 고위험도전(전문연구실) ▸ **(단계)** 타당성연구(3년) → 기술확보(3년)	6년 (3+3)	4 (29)	지정 공모	응용 (2~5)	대학	
22	(Grant 과제) AI 보안 ▸ **(품목)** AI 보안	6년	0.75 (5.75)	자유 (품목)	응용 (2~5)	대학	
23	AI·빅데이터 기반 개인정보 노출·불법 유통탐지 기술 개발 ▸ **(특징)** 사회문제해결	3년	12 (42)	지정 공모	응용 (4~6)	제한 없음	
24	온라인 중고거래 등 신종 비대면 사이버 사기 탐지·추적 및 피해 예방 플랫폼 기술 ▸ **(특징)** 디지털뉴딜R&D, 사회문제해결	3년	7.5 (27.5)	지정 공모	응용 (4~6)	제한 없음	
25	시스템·SW 취약점 AI기반 진단 및 대응 기술 개발 ▸ **(특징)** 디지털뉴딜R&D	4년	12 (57)	지정 공모	응용 (3~6)	제한 없음	
26	자동차 내부 네트워크의 보안 취약점 분석 기술 개발 ▸ **(특징)** 혁신성장동력R&D	4년	9 (45)	지정 공모	응용 (4~6)	산업체, 대학	
27	AI 시스템 내의 정보흐름 추적 및 제어를 통한 프라이버시 위험 분석 및 대응 기술 개발 ▸ **(특징)** 디지털뉴딜R&D	4년	3 (15)	지정 공모	응용 (2~4)	대학	

관리번호	2021-차세대보안-01	(지정공모형)
과제명	저사양 디바이스 대상 고효율 PQC 안전성 및 성능 검증 기술 개발	

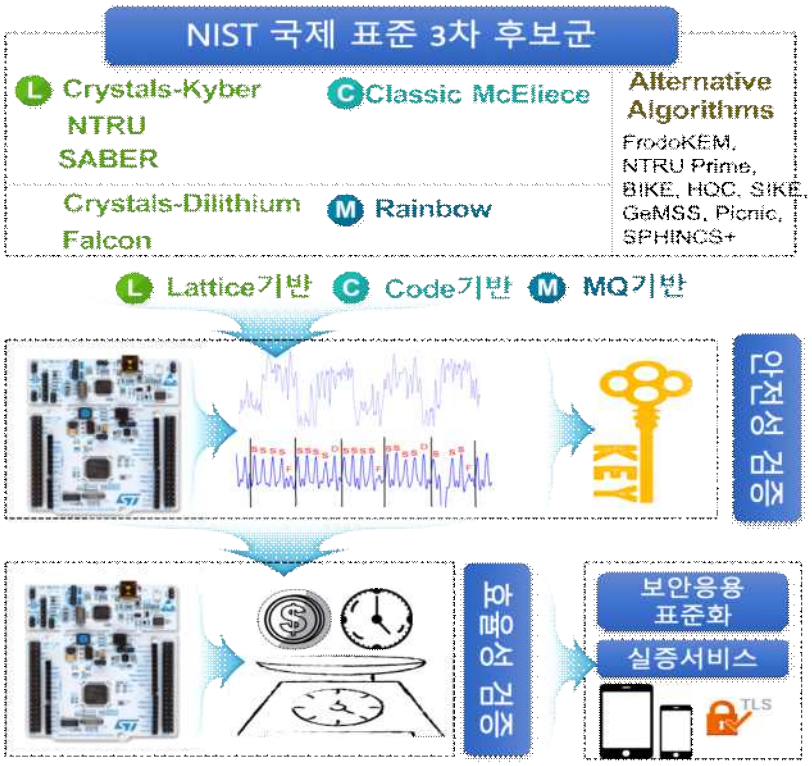
1. 개념

○ PQC* 암호 후보군에 대한 **저사양 디바이스**(Cortex M4급 MCU**) **환경**에서의 **부채널 등 구현 안전성 및 효율성 검증**을 통해 **고효율 PQC 설계 원천기술 확보 및 보안응용 기술 적용을 통한 실증**

* PQC: Post-Quantum Cryptography, 양자 컴퓨터에서도 안전한 암호

** MCU: Micro Controller Unit

※ '20년 현재 NIST(미국표준기술연구소) PQC 3차 후보군 선정, '22~'24년 표준화 예정



< 기술 개념도 >

2 필요성

○ **(정부 지원 필요성)** 대규모 양자 컴퓨터에 의한 현재 암호체계 붕괴가 예상되며, 이로 인해 전체 ICT 인프라가 무력화되므로 안전하고 효율적인 PQC의 선제적 발굴을 통해 **국가 ICT 인프라의 PQC로의 적시 전환 준비** 필요

- 양자 컴퓨터 등 첨단기술 리스크에 대비 가능하도록 중장기 대책 마련과 함께 현재의 암호체계 고도화를 위한 안전한 암호기술, PQC 등의 도입 필요성 권고(4차 산업혁명 대정부 권고안 부록)

○ **(기술성)** PQC 국제 표준 후보 등장에 따라 제안된 기술의 활용을 위한 **효율적인 암호기술 구현 기술, 실증 및 안전성·성능 검증 기술 개발** 필요

- '17년 이 후로 NIST 주도의 PQC 선정 및 표준화 진행 중임
- '20년 현재 NIST PQC 국제 표준의 3차 후보 선정 작업이 진행 중이며 15개 후보를 대상으로 NIST를 중심으로 성능 검증, 부채널 등 취약성을 비교·검증하고 있음
- 국제 표준화 작업이 진행됨에 따라 국외적으로 PQC의 SW/HW 고속 구현, TLS 등 보안 프로토콜 개발, PQC 안전성 검증을 위한 관련 연구가 진행되고 있음
- 특히, PQC 안전성 및 효율성 분석을 위해서는 제안된 NIST 후보 알고리즘에 대한 **성능 비교, 부채널 등 안전성 분석을 위한 벤치마크 플랫폼 개발을 추진**하고 있으나 **국내에서는 관련 연구가 미진한 실정**

※ OQS(Open Quantum Safe), pqm4 등 국외 프로젝트에서는 최적 구현 소스코드, 저사양 디바이스 환경에서의 구현 성능 비교, 분석을 위한 연구가 진행되고 있음

※ 미국 조지메이슨 대학 CERG 그룹에서는 다양한 마이크로프로세서를 대상으로 양자내성암호의 성능 벤치마크를 위한 XXBX(eXtended eXteranl Benchmarking eXtension) 플랫폼을 개발 함

- 양자 컴퓨터 시대에 국가 ICT 인프라의 신속한 암호체계 전환을 위해 필요한 양자내성암호의 안전하고 효율적인 구현 기술, 안전성·성능 검증 및 활성화 기반 마련을 위한 기술 개발 필요

○ **(경제성)** 안전하고 효율적인 PQC 구현 설계 기술을 선제적으로 개발하여 ICT 인프라에 필요한 **양자 저항 제품을 선점하여 글로벌 시장 조기 진입 필요**

- 양자컴퓨터에 취약한 현대 암호기술이 적용된 전자상거래 등 금융 분야, 국가/민간 네트워크 보안 분야 등 중요 인프라에 필요한 기술의 선제적 확보로 제품화에 기여
- 전 세계가 양자컴퓨팅 시대를 대비하여 중요 인프라의 암호체계 고도화를 준비하고 있어 양자 저항 기술/제품에 대한 수요가 증가 예상

3. 연구목표

○ **최종목표**

- **저사양 디바이스 환경**에서 PQC 글로벌 표준 후보군들에 대한 **안전성/성능 상호 비교·분석(벤치마크)을 통한 고안전성/고효율 PQC 기술 확보**와 이를 적용한 **보안응용 실증**

* NIST에서 진행 중인 PQC 공모와 연계하여 국제 표준 후보들에 대한 경쟁력 확보 (글로벌 R&D 협력 기술)

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	PQC 구현설계 취약성 검증 신규 기술 종류	종	4 이상	2종 (MQ기반, Lattice기반)	2종 (MQ기반, Lattice기반 한국/국민대, 인하대)

2	취약성 누출 부채널정보 종류(형태, 위치, 원인 등)	종	최소 7 (총 후보군의 1/2)	2 (MQ기반, Lattice기반)	2 (MQ기반, Lattice기반 한국/국민대, 인하대)
3	구현설계 안전성 검증 플랫폼 비용 절감	기능	계측장비 없는 부채널 정보 수집 가능	고가계측장비 기반 분석 (SCARF)	고가계측장비 기반 분석 (미국/램버스)
4	안전설계 적용 전 대비 메모리 증가율	%	100% 이하 (PQC, Cortex M4)	130% (대칭키기준)	130% (대칭키기준, 미국/램버스)
5	안전설계 적용 전 대비 시간 증가율	%	50% 이하 (PQC, Cortex M4)	50% (대칭키기준)	23% (RSA지수연산, Intel Core i5 프랑스/LIRMM)
6	PQC 적용 보안응용 프로토콜 종류	종	3종 (TLS/DTLS/Cert)	-	1종 (CECPQ2, 미국/구글)

○ 연차별 개발목표

구분	연도별 연구목표
2021년	- PQC 후보 대상 최적화 구현 및 부채널기반 취약성 검증 기반 기술 확보
2022년	- PQC 후보 대상 구현 안전성 보장 설계 구현 및 효율성 검증
2023년	- PQC 구현 안전성/효율성 통합 검증 시스템 및 고효율 안전설계 PQC 후보기반 보안응용 정합 기술 개발
2024년	- PQC기반 보안 적용 서비스 발굴 및 실 환경 적용을 통한 실증

4. 연구내용

○ 개발 기술 내용

① 저사양 디바이스기반 PQC 구현 안전성 및 성능 검증 기술
- 안전성/성능 검증용 부채널정보/사용량 수집장치 내장 Cortex M4기반 HW플랫폼
- PQC 연산단계 메모리 접근 정보기반 저비용 부채널 정보 수집 기술 개발
- 안전성 강화 PQC 구현설계 대한 효율성 검증 메커니즘 개발
- PQC 종류별 구현안전성 분석 및 성능 비교(벤치마크) 시스템 개발

② 저사양 디바이스대상 PQC 암호 최적화 기술
- [Lattice기반] PQC 종류별 Cortex M4기반 최적화 설계 기술 개발
- [Code기반] PQC 종류별 Cortex M4기반 최적화 설계 기술 개발
- [MQ기반 및 기타] PQC 종류별 Cortex M4기반 최적화 설계 기술 개발

③ 부채널 등 구현 취약성 분석 및 대응 기술
- PQC 연산단계 메모리접근 기반 부채널 정보 추출 및 분석 기술 개발
- PQC 연산단계 계측기반 부채널 정보 추출 및 분석 기술 개발
- PQC 암호 부채널분석 등 구현취약성 대응 기술 개발

④ PQC 기반 보안 프로토콜 표준 정합성 검증 기술
- PQC 최적화 기법 적용 TLS/DTLS 클라이언트/서버 라이브러리 개발 및 정합성 검증
- PQC 최적화 기법 적용 PKI인증서 라이브러리 개발 및 정합성 검증
- 부채널분석 대응 PQC 적용 TLS/DTLS 및 PKI인증서 라이브러리 개발 및 검증

⑤ 5G/6G 등 응용서비스에 양자내성암호 적용을 통한 실증
- 저사양/이기종 디바이스 동작 가능 PQC기반 5G/6G 서비스 적용 시나리오 개발
- 스마트시티, IoT 등에 양자내성암호 응용기술 활용사례 발굴 및 실증 테스트

○ **기존 (보유) 기술**

① 국산 보급형 부채널 분석 및 검증 시스템
- 국내외 표준 암호에 대한 부채널 정보 수집 및 분석 기술
- SW 및 HW기반 부채널 분석 대응기술

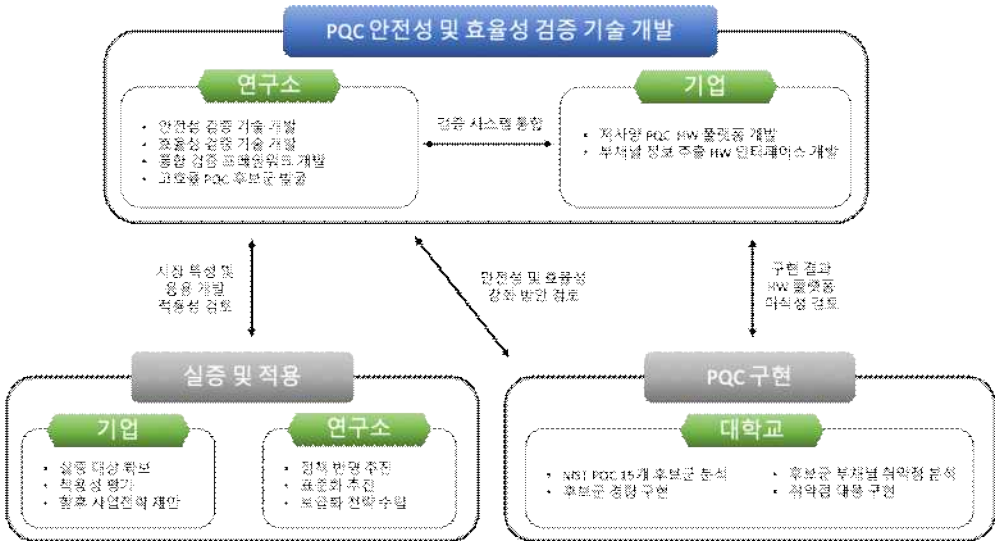
② 양자 컴퓨팅 환경에서의 PQC 안전성 분석 기술
- 분석 대상 암호 알고리즘의 양자분석에 소요되는 양자자원량 분석 기술

5. 지원기간/예산/추진체계

○ **기간** : 4년 이내

○ **정부출연금** : ‘21년 15억원 이내 (총 정부출연금 75억원 이내)

○ **주관기관** : 제한없음



< 추진 체계 >

기술분류	대분류(차세대보안) - 중분류(시스템 및 암호보안) - 소분류(암호기술)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (3) ~ (5)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-02	(지정공모형)
과제명	국가공공 정보시스템 안전성 및 활용성 제고를 위한 차세대 암호체계 개발	

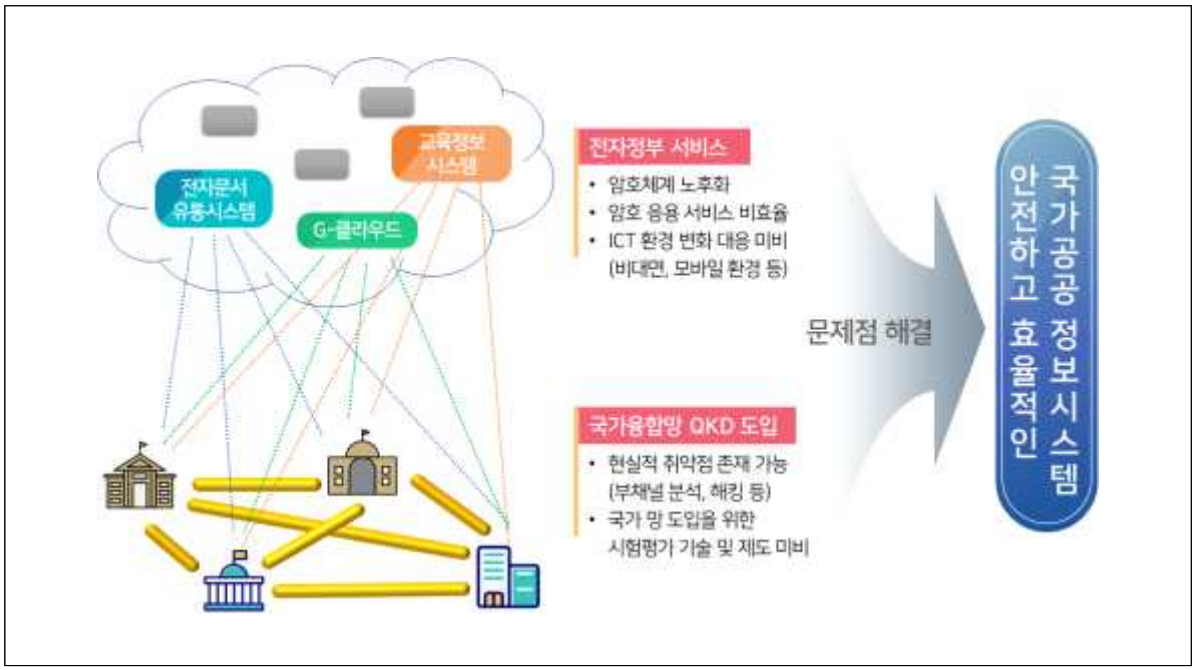
1. 개념

○ 국가공공 정보시스템의 잠재적 안전성 취약점을 제거하고 차세대 ICT 환경에서의 활용성을 제고하기 위한 국가공공용 암호체계의 고도화

○ 국가공공 정보시스템에 적합한 새로운 암호인프라(암호알고리즘, 양자내성암호, 양자키분배(QKD) 등)가 적용되어, 활용성이 강화되고 양자컴퓨터 등 미래위협에 대비할 수 있는 차세대 암호체계

- 차세대 국가공공 정보시스템의 종합적 안전성 보장을 위한 암호 요소기술(암호알고리즘, 암호 프로토콜 등), 암호 기반기술(암호키 생성/저장, 인증 체계 등), 암호 운영 방식 등을 포함한 암호체계 개발
- 양자컴퓨터 등 암호 해독기술의 발전과 신규 ICT 환경에 대응하기 위한 PQC 암호 정합 기술, QKD 안전성 평가 기준 수립 및 적용 기술 개발

< 국가공공 정보시스템 고도화 개념도>



2 필요성

○ **(정부 지원 필요성)** 행정 전자서명 및 국가융합망 등 국가공공 정보시스템은 국가 운영의 근간이 되는 기반 인프라로서 사회 전반에 미치는 영향 등을 고려할 때, 국가주도의 연구개발 사업으로 지원하는 것이 타당함

- 암호 분석기술의 발전과 ICT 환경변화에 대응하여 국가공공 정보시스템의 안전성과 효율성을 고도화하기 위해 암호체계의 개선 및 신기술 도입을 위한 적용 체계 확보 필요
- 급변하는 ICT 환경에서의 암호 활용성과 잠재적 위협에 대응한 안전성 제고를 위해 국가주도의 신속하고 체계적인 차세대 암호기술 연구개발 필요

* 美 NIST는 모바일 환경에 적합한 경량 인증암호(2019~)와 양자컴퓨터 위협 대응을 위한

포스트 퀀텀 암호(2016~)의 도입을 위한 공모사업 추진 中

* 중국은 'National Key R&D Programmes(총9.7억 위안, 37개 주제)' 사업 중 하나의 주제로 선정하여 신규 암호기술 개발에 투자하고 있으며, 중국암호학회 주관으로 블록암호와 공개키암호에 대한 공모사업(2018~2020) 추진 완료

- 정부주도 각종 QKD 기술 실용화 정책 지원을 위한 안전성 보장연구 시급
  * 세계 각국이 다양한 연구/제작으로 QKD 장치개발 완료(예정) 예상에 따라, 실용성 기반의 안전한 QKD 구현기술 및 사용환경 확보 연구 필요
  * QKD기술의 보안성 확보가 정부의 QKD 지원사업(디지털 뉴딜의 DNA 생태계 강화용 양자암호통신 인프라구축, 6G 등)의 주요목표 달성을 위한 핵심요소
  * 또한, NIA 주도의 양자암호통신 인프라 중 공공망 대상으로 시범망 구축 및 보안성, 안전성 검증(NSR, ETRI, TTA 등)을 추진 중이나, 암호적 안전성 분석기술 미흡

○ **(기술성)** 암호 분석기술 및 컴퓨팅 환경의 발전으로 인한 암호 안전성 저하와 ICT 환경 변화에 능동적 대처가 필요하며, 학문적 성숙도 대비 미숙한 QKD 장비의 實 사용을 고려한 안전성 측정기술 확보 필요

- 행정전자서명 등 국가공공 정보시스템의 기반 암호체계는 10여 년 전의 암호기술을 기반으로 구축되어 있어 암호 분석기술의 발전과 업무환경의 변화에 대응하기 위한 암호체계 개발 필요
  * 현재 표준에서 제외된 알고리즘(전자서명, 해시함수, 난수발생기 등)을 여전히 포함하고 있어 잠재적인 위협이 존재
  * 현 암호체계는 PC 환경을 기반으로 설계되어 모바일 환경 등에서 활용이 부적합
- 모바일, VPN 등의 환경에서의 활용성 강화를 위해서는 경량인증, 대체인증 등의 새로운 암호 응용기술 확보가 필요
  * 전자서명법 통과 이후(2020. 5.) SMS인증, 휴대폰인증, OTP, IC태깅, 생체인증, DID 등의 대체인증기술이 대두되고 있음
- QKD 요소기술에 대한 양자이론, 시스템 관점의 다양한 위협분석을 통해, 보다 안전한 QKD 장치 제작에 필요한 제반기술 마련이 중요
  * 국내는 국제적인 흐름보다 앞서 QKD 프로토콜(TTA, 2018년)과 QKD 보안 요구사항(TTA, 2019년)의 표준화를 완료하였으나, QKD 시험 요구사항 및 관련 요소기술 개발 연구가 필요

○ **(경제성)** 최신 암호기술을 탑재한 모듈의 보급과 함께 신규 양자암호기술의 국가공공용 암호인프라 도입 기반 확충을 통해 국내 암호 시장 활성화에 기여

- 국가공공 분야에 도입, 운영되고 있는 암호기술의 교체를 위해서는 막대한 시간과 비용이 소요되므로 새로운 암호체계 구축 시 철저한 안전성 검증 필요
  * 2005년에 발견된 SHA-1 취약점 대응, RSA 암호 안전성 수준 제고 등을 위한 공인인증서 고도화사업은 2012년 완료되었으며 교체에 약 100억 원 소요
  * QKD 시장은 5년간 매출 기준 19.9%의 CAGR(Compound Annul Growth Rate, 연평균 복합성장률)을 기록할 것으로 예상(2019, 도시바 시장 경제 분석)

3. 연구목표

○ **최종목표 : 국가공공 정보시스템의 안전성 및 활용성 제고를 위한 암호체계 고도화 및 양자키분배 기술 안전성 시험, 평가기술 개발**

- 국가공공 정보시스템 고도화를 위한 암호 알고리즘, 최적 구현기술, 암호모듈 개발
- 사회현안(Post COVID-19, 대체인증) 해결을 위한 차세대 암호인증기술 개발
- 국가융합망 등의 적용을 위한 QKD 시험/평가기술 개발
- QKD 구성요소별 양자광학적 성능 및 양자성 분석, 위협대응기술 개발을 통한 안전성 측정 및 고도화기술 확보

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	국가공공 정보시스템용 신규 암호체계 안전성	비트	128[주1] 이상	-	암호 기능 및 안전성 별 다수의 파라미터 보유 (미국, NIST)
2	암호모듈 지원 운영체제 종류	종	5 이상	-	8종[주2] (Cryptolib)
3	암호모듈 지원 범용 CPU 하드웨어 가속기 종류	종	5 이상	-	8종[주3] (Crypto++, Openssl)
4	KCMVP 보호함수 안전성 평가보고서	건	2	-	CryptREC (EU, NESSIE)
5	신규 암호체계/암호기술 실증[주4]	건	2 이상	-	-
6	시험가능한 QKD 요소별(조건별) 공격 분석/대응기술 개수	건	4 이상	-	3 (러시아, MISIS)
7	QKD 공격기술별 대응기술 시험방법론 개수	건	10 이상	-	-
8	QKD 안전성 분석도구용 시나리오	건	10 이상	-	-

주1) 기존 컴퓨팅 환경 기준이며, 양자 컴퓨팅 환경에서도 이와 동등한 수준의 안전성을 가져야 함
주2) Windows, Unix, Linux, Android, macOS, iOS, DOS, Embedded
주3) AES-NI, SSSE3/SSSE4.1, AVX/AVX2, RDRAND, VIA PadLick, AltiVec, ARMv7 NEON, ARMv8-A
주4) 신규 개발기술의 국가공공 정보시스템 적용 적합성 검증을 위한 환경(PC, 모바일 등), 서비스(전자문서유통 등) 실증

○ **연차별 개발목표**

구분	연도별 연구목표
2021년	• 기존 국가공공 암호체계 및 주요 전자정부 서비스의 안전성 및 취약점 분석 • QKD 광학모듈단위 공격기술 분석 및 안전성 확인용 토이모델/환경 구축
2022년	• 양자내성 암호 도입을 위한 구현 및 정합기술 개발 • 모바일 행정시스템 등을 위한 경량 인증기술 개발 • QKD 시스템 공격기술 분석 및 QKD 송신부 안전성 보장방법론 개발
2023년	• 국가공공 정보시스템용 신규 암호체계 개발 및 간편전이 체계/기술 개발 • 암호모듈 검증제도 보호함수 재개정을 위한 안전성 분석 • G 클라우드 안전성 강화를 위한 암호기술 개발 • QKD 시스템 공격기술 분석 및 QKD 송신부 안전성 보장방법론 개발
2024년	• 신규 암호체계의 최적 구현기술 및 국가공공 정보시스템용 암호모듈 개발 • 비대면 서비스 정보보호를 위한 암호 응용기술 개발 • 신규 암호체계/암호기술 국가공공 정보시스템 서비스 실증

	• 안전성 보장방법론 타당성 검증 및 QKD 시험도구 개발(토이모델 대상)

4. 연구내용

ㅇ 개발 기술 내용

① 국가공공 정보시스템 암호체계 고도화기술
- 기존 국가공공 암호체계 및 주요 전자정부 서비스의 안전성 및 취약점 분석기술
- 국가공공 정보시스템용 신규 암호체계 개발기술
- 국가공공 암호체계 전환을 위한 간편전이 체계 및 간편전이 기술
- 양자내성암호 도입을 위한 구현 및 정합기술
- 신규 암호체계의 최적 구현기술 및 국가공공 정보시스템용 암호모듈
- 암호모듈 검증제도 보호함수 목록 재개정을 위한 안전성 분석기술
- 신규 암호체계/암호기술 효율성 검증을 위한 서비스 실증기술

② Post COVID-19 시대 및 사용자 편의성 제고를 위한 차세대 암호 인증기술
- 모바일 행정시스템 등을 위한 경량 인증기술
- G 클라우드 안전성 강화를 위한 암호기술
- 비대면 서비스 정보보호를 위한 암호 응용기술

③ QKD에 대한 부채널 공격 등 안전성 위협 식별, 대응기술 및 시험기술
- QKD 장치 구성요소별 안전성 위협 분석기술
- QKD 장치용 양자성 구현기술의 안전성/신뢰성 분석기술 및 안전성 시험기술
- QKD 장치에 대한 부채널 공격 대응기술 및 시험방법론 개발기술

④ QKD 암호안전성 측정기준 및 안전성 시험도구 개발기술
- QKD 안전성 요구사항(TTA 표준) 대응되는 안전성 시험요구사항 도출
- QKD-암호장치 연동에 사용 가능한 암호키 연동기술 및 안전성 분석기술
- QKD 안전성 분석기술 적용을 위한 안전성 시험도구 개발기술

ㅇ 기존 (보유)기술

① 암호 알고리즘 : 비밀키 암호, 전자서명, 난수생성기
- 행정 전자서명에는 ARIA(블록암호), SHA-256(해시함수), KCDSA(서명) 등이 사용됨
- 최근에는 국내에서 개발된 블록암호 LEA, 해쉬함수 LSH 등이 사용되고 있으며, 국제적으로는 양자 안전성 확보를 위하여 PQC 알고리즘 표준화가 진행 중임
- 행정 전자서명에서 사용하는 FIPS 186-2, ANSI X9.62 난수생성기 이외에 NIST SP 800-90A에 수록된 CTR-, HASH-, HMAC-, EC- DRBG가 널리 사용됨

② 암호 모듈 개발기술
- S/W 암호모듈은 OpenSSL, Crypto++, RSA BSAFE 등이 사용되고 있음
- Bouncy Castle, RSA BSAFE, wolfCrypt 등은 CMVP(FIPS 140-2) 인증을 받음

③ QKD용 소자, 시스템 구현기술 및 시험망 운영기술
- KIST, ETRI, SKT 등: QKD용 소자, QKD장치, 운영/제어 소프트웨어 기술 개발

④ QKD 이론 및 프로토콜별 안전성 분석기술 연구
- KAIST, KIAS: QKD의 정보이론적 안전성 분석기술 연구 진행 중
- NSR: 유선 QKD 시스템 대상 물리적 안전성 분석기술

⑤ 실용화 기반 QKD 규격 표준화 기술(일부 진행 중)

- 국내 출연연 주도 QKD보안 표준화(TTA, QKD 프로토콜 표준(2018), QKD 보안 요구사항 표준(2019)), 통신업체 주도 네트워크 표준화(국제표준 인용 포함) 동시 진행

5. 지원기간/예산/추진체계

ㅇ **기간** : 4년 이내 (1단계 2년 → 2단계 2년)

ㅇ **정부출연금** : '21년 19.02억원 이내 (총 정부출연금 94.02억 원 이내)

ㅇ **주관기관** : 제한없음

기술분류	대분류(차세대보안) - 중분류(시스템 및 암호보안) - 소분류(암호기술)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (3) ~ (6)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-03	(지정공모형)
과제명	간결한 비대화형 연산 증명 시스템(SNARK)를 위한 암호 원천 기술 개발	

1. 개념

○ 처리장치나 저장공간 등 유효 연산자원이 비대칭적인 환경에서 제한된 자원을 보유한 사용자의 입력에 대한 서버의 연산 수행 및 연산결과의 무결성을 보장하기 위한 암호학적 기법들 중 “간결한 비대화형 방식”*

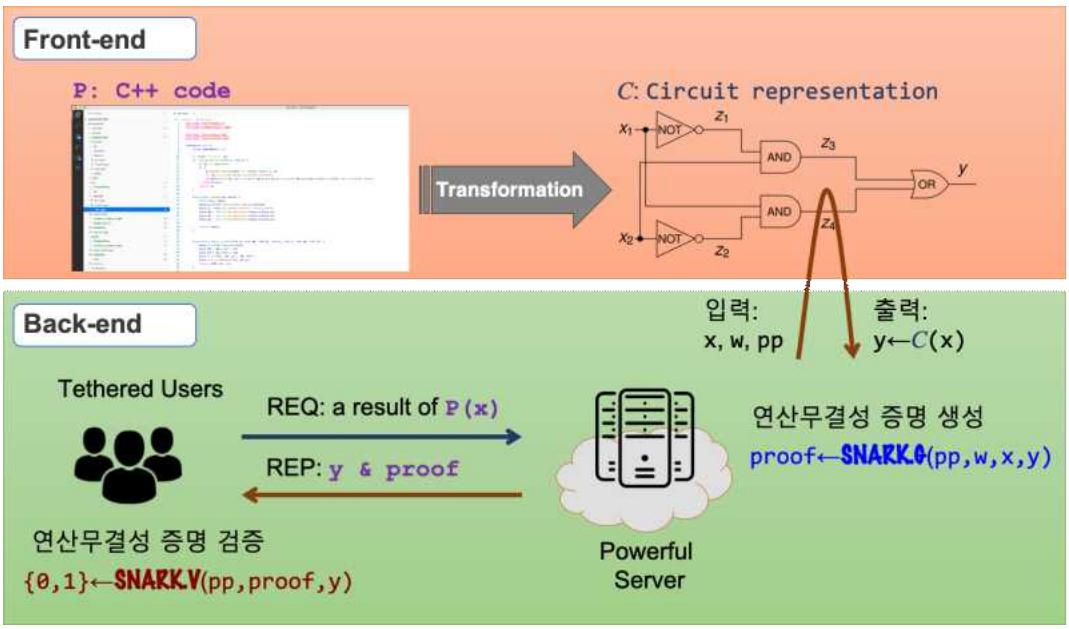
- 연산을 요청한 사용자는 연산결과의 무결성 요구뿐만 아니라, 제약된 연산자원으로 무결성의 검증을 위해 필요한 증명의 크기가 간결하고§ 비대화형※ 방식으로 요청을 전송한 후 결과 수신만 요구
- 연산을 수행하는 서버는 증거 생성과정에 증거값 (Witness)이 이용된 경우 서버의 비밀성 보장 요구†

* SNARK로 불리며 Succinct Non-interactive ARgument of Knowledge의 축약어

§ 간결한 (Succinct): 증거의 크기가 연산의 크기 (진술이나 증거값의 크기)와 독립적이거나 준선형 (Sub-linear)의 크기를 요구하는 특성

※ 비대화형 (Non-interactive): 요청을 수신하는 라운드와 결과를 전송하는 라운드 이외에 연산과정에서 별도의 통신을 요구하지 않는 특성

† 증거값의 비밀을 보장하는 SNARK은 zk-SNARK으로 불림



< SNARK 시스템 개념도>

2 필요성

○ **(정부 지원 필요성)** 빅데이터, 인공지능, 클라우드, 분산인증, 분산금융 등 여러 4차 산업혁명분야 개인정보호 활용 서비스의 신뢰성 향상을 위한 핵심 원천 기술인 SNARK 기술 필요

- 국제 경쟁력 확보 및 기술 선점 필요
- 2018년 MIT technology review에서는 영지식 증명(zero knowledge proof) 및 SNARK에 기반한 “Perfect online privacy“를 10대 혁신 기술(10 Breakthrough

technologies) 중 하나로 선정한 바 있음

- SNARK는 클라우드 연산결과, 빅데이터 분석 결과, 딥러닝 학습 결과 등 다양한 산업군에서 협업 결과에 대한 무결성 검증*을 위한 핵심 기술임

  * 서버-클라이언트, 분석서버-데이터소유주, 딥러닝서버-데이터제공자 간의 협업에서 서버의 결과에 대한 무결성 검증

- SNARK 관련 기술들은 분산인증, 분산금융 관련 산업들의 주요 문제점(확장성, 비밀성 보장) 해결을 위한 원천 기술로써 관련 산업분야의 시장 안착에 기여

○ **(기술성)** SNARK는 다양한 응용에 활용 가능한 암호 원천기술로써 범용 SNARK 기술의 고도화뿐만 아니라 실제 응용환경별 특화된 SNARK관련 기술개발 필요

- 일반적인 산술회로를 위한 범용 SNARK의 성능 개선뿐만 아니라, 클라우드 특화 SNARK, 딥러닝 특화 SNARK, 양자 내성 SNARK 등 여러 산업에서 주로 활용하는 형태의 데이터 및 연산 알고리즘에 특화된 SNARK 관련 기술 개발 필요*

  * '특정연산에 특화된 SNARK' 및 'SNARK 특화된 암호 시스템' 개발이 필요
  (예를 들어, 구간증명(range proof)에 특화된 SNARK 관련 기술로 Bulletproofs, SNARK에 특화된 해시함수로 MiMC과 Poseidon 등이 있음)

○ **(경제성)** SNARK는 이미 여러 산업분야에서 핵심 전략기술로 부상하였거나 가까운 미래에 필수 기술이 될 것으로 기대

- 빅데이터 분석 기술 및 딥러닝 등의 기계학습 기술이 다양한 산업군에 활용이 되면서 개인정보 등의 민감 데이터를 활용한 분석 및 학습이 빈번해지고, 이를 위해 다자간 협업으로 결과를 도출하는 경우가 다수 발생. 다양한 산업군에서 다자간 협업으로 도출된 결과를 신뢰하기 위해 분석 및 학습결과에 대한 무결성 검증 기술을 필요

- 딥러닝, 클라우드 이외에도 SNARK 관련 기술들은 분산인증(DID), 분산금융(Defi) 등 신뢰성이 중요한 역할을 하는 다양한 서비스들의 확장성 및 익명성 솔루션*으로 활용 가능하여 다양한 산업군으로의 활용 범위가 넓음

  * 익명성 제공을 위해서는 SNARK에 영지식(zero-knowledge) 성질이 추가된 zk-SNARK가 필요

3. 연구목표

○ **최종목표 : 간결한 비대화형 연산 증명 시스템(SNARK)을 위한 암호 원천기술 및 응용 기술 개발**

- 일반적인 산술회로 검증을 위한 범용 SNARK 암호 원천 기술
- 응용환경별 특화된 효율적인 SNARK 암호 원천 기술
- 응용환경별 SNARK 특화된 암호 기술 개발
- SNARK 최적화 구현 및 오픈 소스 라이브러리 개발
- SNARK 응용 기술 개발

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	산술회로 증명을 위한 범용 SNARK 증명크기[1]	공간 복잡도[2]	$\leq O(\sqrt{\log(N)})$	$O(\log(N))$[3]	$O(\log(N))$[4] (미국/영국, Stanford Univ. UCL, Blockstream)
2	페어링 연산 증명에 특화된 SNARK 검증시간	초 (sec)	≤ 2	-	15[5] (이스라엘/미국, Technion/Tel Aviv/MIT)
3	클라우드 연산 증명에 특화된 SNARK 검증시간	초 (sec)	≤ 500	-	2,346[6] (스페인/미국, IMDEA/CUNY)
4	딥러닝 연산 증명에 특화된 SNARK 검증시간	초 (sec)	≤ 1	-	2.2[7] (미국/New York University)
5	오픈 소스 라이브러리 지원 SNARK 수	종	3	-	-
6	오픈 소스 라이브러리 지원 프로그래밍 언어	종	2[8]	-	-

1) Trusted Third Party 없는 파라미터 생성(trustless setup)가능한 SNARK
2) 증명대상이 되는 산술회로의 곱셈 게이트 개수 N에 대한 증명크기의 공간 복잡도 (Big-O 기호로 표기)
3) Bulletproofs, https://eprint.iacr.org/2017/1066, IEEE S&P 2018
4) Bulletproofs+, https://eprint.iacr.org/2020/735
5) libsnark, https://eprint.iacr.org/2013/879, USENIX Security 2014
6) TPC-H #2 at vSQL, https://eprint.iacr.org/2017/1145, IEEE S&P 2017
7) SafetyNets: Verifiable Execution of Deep Neural Networks on an Untrusted Cloud, NIPS '17
8) C, C++, Java, Javascript, Rust, Go, OCaml, Haskell, Python 중 최소 2가지 이상

○ **연차별 개발목표**

구분	연도별 연구목표
2021년	○ 연산 무결성 검증을 위한 연산 증명 기법 조사 및 분석 연구 ○ SNARK 기초 기술 연구 (산술회로 변환 기술 등)
2022년	○ 일반적인 산술회로 검증을 위한 범용 SNARK 기술 연구 ○ SNARK 기초 기술 연구 (파라미터 갱신 등)
2023년	○ 페어링 연산 증명에 특화된 효율적인 SNARK 암호 원천 기술 ○ 딥러닝 연산 증명에 특화된 효율적인 SNARK 암호 원천 기술
2024년	○ 클라우드 연산 증명에 특화된 효율적인 SNARK 암호 원천 기술 ○ SNARK 확장 기술 연구 (양자 내성 안전성 등)
2025년	○ SNARK 최적화 구현 및 오픈 소스 라이브러리 개발 ○ SNARK 응용 기술 개발

4. 연구내용

ㅇ 개발 기술 내용

① SNARK 기초 기술 및 산술회로 증명을 위한 범용 SNARK 원천 기술
- 실행 프로그램을 산술회로(arithmetic circuit)로 효율적으로 변환하는 기술
- 신뢰성 향상을 위한 파라미터 갱신 (Updatable Common Reference String) 기술
- 일반적인 산술회로 연산을 검증하기 위한 SNARK 기술

② 페어링 연산 증명에 특화된 효율적인 SNARK 암호 원천 기술
- 페어링 연산은 다양한 암호 프로토콜 설계에 활용되는 수학 연산자임
- 페어링 연산 증명을 위해 산술회로로 변환시 비효율이 발생하기 때문에, 산술회로 변환하지 않고 효율적으로 증명 수행하는 SNARK 암호 원천 기술 개발

③ 딥러닝 연산 증명에 특화된 효율적인 SNARK 암호 원천 기술
- 딥러닝에서 많이 사용되는 연산*의 효율적인 검증에 특화된 SNARK 기술
  * 딥러닝에서는 ReLU, max pooling 등 실수기반 비선형 연산이 많이 활용
- 딥러닝에서 처리하는 데이터의 형태*에 적합한 SNARK 기술
  * 일반적인 SNARK가 유한체 원소 형태의 데이터 처리에 적합한 것에 반해, 딥러닝 데이터는 대부분이 실수이며 이러한 실제 데이터를 유한체 원소로 변환하는 데 비효율성이 발생

④ 클라우드 연산 증명에 특화된 효율적인 SNARK 암호 원천 기술
- 클라우드에서 데이터 처리시 많이 사용되는 연산*의 효율적인 증명에 특화된 SNARK 기술
  * 클라우드에서 많이 사용하는 연산의 예로는 검색, 산술연산, 통계분석 등이 있음
- 클라우드에서 처리하는 데이터의 형태*에 적합한 SNARK 암호 원천 기술
  * 클라우드에서 다루는 데이터는 주로, 정수, 실수, 문자열 등이며 이러한 실제 데이터를 유한체 원소로 변환하는 데 비효율성이 발생

⑤ 응용환경별 SNARK 특화된 암호 기술 개발
- 딥러닝에서 많이 활용되는 SNARK 특화된 암호 기술 개발
- 클라우드에서 많이 활용되는 SNARK 특화된 암호 기술 개발

⑥ SNARK 최적화 구현 연구 및 오픈 소스 라이브러리 개발
- 그룹 기반 SNARK의 최적화를 위한 SNARK 전용 타원곡선 고속연산 기술 개발
- SNARK CRS 생성, 증명, 검증 알고리즘 최적화 구현 기술 연구
- 다양한 프로그래밍 언어*로 SNARK 오픈 소스 라이브러리 개발
  * C, C++, Java, Javascript, Rust, Go, OCaml, Haskell, Python 중 최소 2가지 이상의 프로그래밍 언어로 오픈 소스 라이브러리 개발

⑦ SNARK 응용 기술 개발
- 개발한 SNARK 관련 기술들을 딥러닝/클라우드 등의 응용 서비스 적용 사례 실증

ㅇ 기존 (보유)기술

① 다양한 SNARK 기법 개발
- 대부분 서구권 연구진들에 의해 개발이 되었으며 Groth16, Sonic, Plonk, Marlin,

STARK, Virgo, Bulletproofs 등이 있음
- 아시아권에는 국내 연구진에 의해 개발된 Bulletproofs+가 유일
- 현재까지 제안된 기법들은 각각 성능과 안전성에서 장단점이 있으며, 절대적인 우위를 가지는 SNARK 기법은 없음

② SNARK 특화된 암호 기술 개발
- 유럽 연구진들에 의해 개발된 MiMC, Poseidon 해시함수 기술이 있음
- 전자서명, 동형암호 등 다양한 SNARK 특화된 암호 기술 개발이 필요

③ SNARK 오픈소스 라이브러리 개발
- 해외 연구진들에 의해 작성된 libsnark, Zokrates, libstark 등이 있음
- 국내 연구진에 의해 개발된 Bulletproofs+ (Rust) 라이브러리
- 특정 영지식 증명 기법에 특화 되어 있거나 (libsnark, libstark, Bulletproofs+), 플랫폼 제한적인 한계가 있음 (Zokrates)

5. 지원기간/예산/추진체계	

ㅇ **기간** : 5년 이내

ㅇ **정부출연금** : '21년 10억원 이내 (총 정부출연금 50억원 이내)

ㅇ **주관기관** : 대학, 연구기관

5. 지원기간/예산/추진체계	

ㅇ **기간** : 6년 이내 (1단계 3년 → 2단계 3년)

ㅇ **'21년 정부출연금** : 4억원 이내

ㅇ **총 정부출연금** : 29억원 이내(1단계 14억원 → 2단계 15억원)

ㅇ **주관기관** : 대학

기술분류	대분류(차세대 보안) - 중분류( ) - 소분류( )	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (2) ~ (5)
과제특징	정책지정( ), 혁신도약형(O), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-04	(지정공모형)
과제명	범죄증거 확보를 위한 암호분석 기술 고도화	

1. 개념

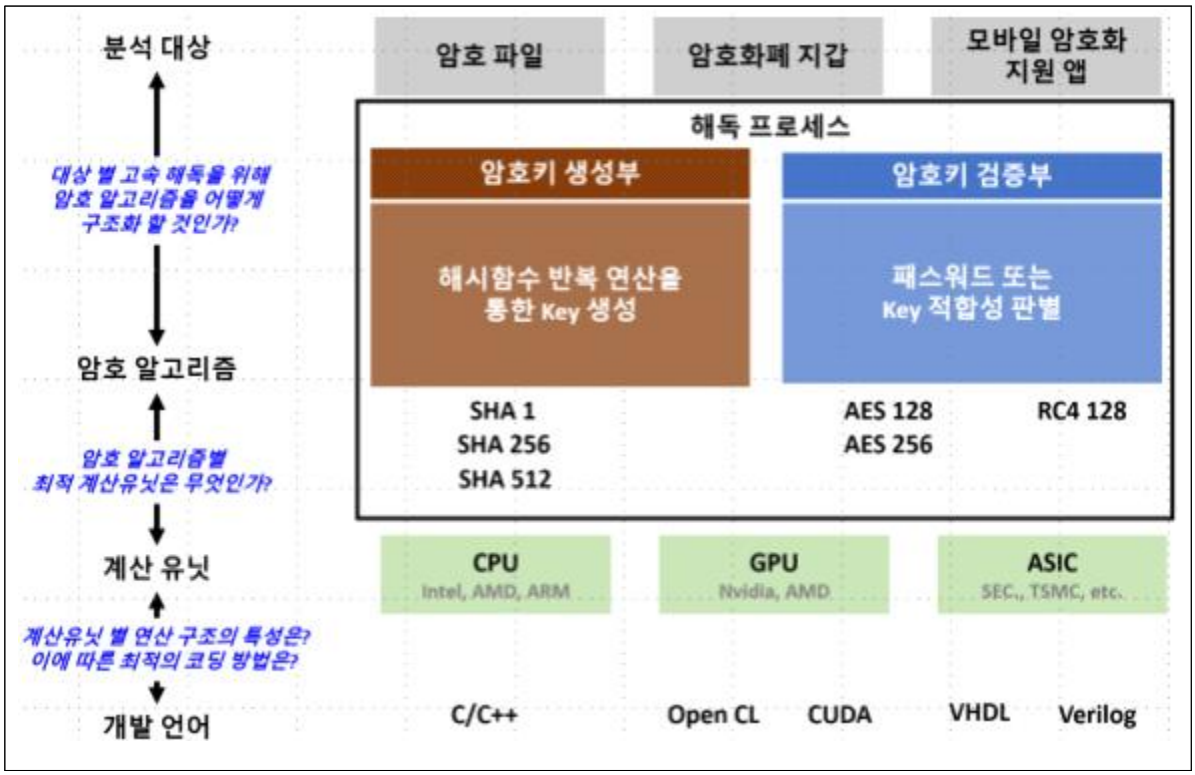
○ 암호기술이 일반에 확대·보급됨에 따라 정당한 법집행이 불가능한 경우 발생 중

- 1)비정형 파일, 암호화 앱, 암호지갑 등에 적용된 암호기술 등이 범죄증거를 은닉하는 데에 활용되어 실체적 진실발견과 범죄수익 환수에 막대한 지장을 초래
  * MS Offce, PDF, 압축파일 등에 고비도 암호기술이 적용되고 있으며, 모바일 기기내에 각종 앱 등에 암호 기술 적용되어 범죄단서 확보에 어려움이 가중되고 있음
  * 암호화폐 등이 범죄수익을 은닉하는 데에 활용되고 있으나, 암호 지갑 등에 보관되고 있어 피압수자의 협조없이 범죄수익 환수에 어려움 가중

○ 이에 법집행기관의 영장집행시 피압수자 협조여부와 상관없이 범죄단서를 확보하기 위한 기법 연구가 절실히 필요

- 암호해독 최적화기법과 CPU · 보조연산장치(GPU 등)를 활용한 암호연산 고속화 기법 뿐만 아니라 전용 ASIC기반의 고속화 기법 등의 연구개발 필요

< 개념도 >



2 필요성

○ **(정부 지원 필요성)**

- 암호해독 기술은 국가수요를 중심으로 요구되는 특성이 있으며, 최근 국가의 법집행기능을 유지하기 위한 필수 요소기술로서 대두됨
- 국가 법집행기능의 무력화 방지를 위한 형사사법체계 유지에 소요되는 기술로써 국가주도의 연구개발 전략 수립이 필요

- 디지털포렌식 기술은 법 통제하에서 디지털증거에 대한 정당한 접근권한 기술 확보가 필요한 분야임
  * 1997년, OECD는 “암호정책의 배경과 쟁점에 관한 보고서”에서 ‘암호사용의 권장이 공공의 안녕, 법집행 및 국가의 안보를 저해하지 않도록 하여야 한다’고 천명

○ (기술성)

- 암호해독은 암호이론, 계산이론, 병렬처리 기술 등을 포괄하는 고난도 기술이며, 민간수요 기술과 차별화되고 고도의 전문인력을 통한 지속적 연구개발이 필요
- 본 연구에서는 민간에서 활용되는 제품과 서비스의 암호해독 기법을 연구하여 공통 성능고도화 요소를 추출하고, 연산장치(CPU, GPU, ASIC 등)의 특성분석에 따른 고속화의 장단점을 분석하여야 하는 종합적 기술요소를 내포하고 있음
  * 암호해독은 암호해독 기법 연구, 암호연산 고속화기법 연구개발, 병렬컴퓨팅 인프라 구축 ·운영 등으로 구분할 수 있음

○ (경제성)

- 법집행 기능의 실패는 국가의 범죄 대응역량 약화를 초래하여 사회안전 유지의 어려움을 가중시켜 공공안녕을 저해함으로써 막대한 사회적 손실 발생 가능
- 암호해독은 막대한 계산자원을 바탕으로 구성되므로 분석기법의 최적화 및 고속 구현기술의 확보는 막대한 구축 및 운영비용 대폭 절감 가능
  * (1,000억원 규모 시스템을 구축·운영하는 경우) 해독성능을 기존 대비 50% 고속화에 성공한 경우, 시스템 구축 비용을 500억원 절감 가능하며, 설치 공간 및 전기사용료 등을 고려하는 경우 막대한 운용 비용 절감 가능

3. 연구목표

○ **최종목표 : 암호해독 매커니즘 고속화 기법 설계 및 개발**

- CPU, GPU기반의 암호알고리즘 고속화 기법 개발 및 구현
- CPU, GPU기반의 암호해독기법 고속화 설계 및 구현, 비용대비 성능기준 제시
- ASIC기반의 암호해독기법 고속화 설계 및 비용대비 성능기준 제시

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	ASIC기준 (칩당소모전력 10w기준)	GHash/sec	300[2)]	비공개	252.3/chip (중국, MicroBT)
2	GPU기준 (RTX2080Ti기준)	MSOffice2013+/sec	40,000	비공개	23,458 (미국, Passware)
3	GPU기준 (RTX2080Ti기준)	PDF(RC4 128)/sec	60백만	비공개	40백만 (미국, Passware)
4	GPU기준 (RTX2070기준)	PBKDF2-SHA512 129,977 iteration/sec	40,000	비공개	4,420 (러시아, Elcomsoft)

○ 연차별 개발목표

구분	연도별 연구목표
2021년	ASIC탑재 암호기법 선정 및 기준 도출(기준정의서)
	분석대상 정의 및 해독기법 연구(공통 구현요소 추출)
	CPU, GPU 특성에 적합한 암호알고리즘 구현
2022년	HDL기반으로 암호알고리즘 설계 및 구현(RTL 소스)
	CPU, GPU 특성에 최적화 암호기법 구현
	CPU, GPU 병행활용 방안 설계 및 최적화 암호기법 구현
2023년	FPGA기반 PCIe 보드 설계 및 RTL단계에서의 소스 검증

4. 연구내용

○ 개발 기술 내용

① 비정형파일(문서, 압축 등) 키생성알고리즘, 암호기법 내 암호알고리즘 특성 분석
- 비정형 파일대상(15종 이상) 공통 암호연산요소 추출 및 고속화 설계 구현
- 암호화앱 대상(7종 이상) 공통 암호연산요소 추출 및 고속화 설계 구현
- 암호지갑 대상(5종 이상) 공통 암호연산요소 추출 및 고속화 설계 구현

② GPU기반 암호연산 최적화 설계 및 구현(CUDA기반, OpenCL기반)
- 암호연산별, 분석대상별 해독메커니즘의 특성분석(동작주파수, 메모리 의존성 등)
  * OpenCL, CUDA구현에 따른 장단점 분석 및 개별 구현
- 암호키 생성부내 반복연산 최적화 방안 연구 및 구현

③ 해독메커니즘에 따른 CPU, GPU 연산 분배、병합 최적화 방안 연구 및 구현
- 개별 계산유닛별 최적화 방안 및 병합 구성 활용방안 연구

④ ASIC 기반 고속화 구현
- ASIC에 적용 가능한 암호알고리즘 선정 및 탑재기준 도출
  * 암호연산처리를 위한 통신량, 목표계산성능, 칩당 소모전력 등 결정요인에 따른 요구사양 정의 및 탑재 암호기법의 범위 선정
- HDL(Hardware Description Language)기반으로 암호알고리즘 설계 및 구현
- FPGA기반 PCIe 보드 설계 및 RTL(Register Transfer Level)단계에서의 소스 검증
  * FPGA검증용 칩 NETList, TestBench 도출
  * PCIe 보드 회로도, Gerber 데이터, IPC-D-365 표준 NETList 도출
  * 추정 성능목표치 및 기능 검증(소모전력, 계산성능, 통신량 등)

○ 기존 (보유)기술

①(비공개) 비정형파일 암호해독 기법 및 해독 프로그램
- CPU · GPU, ASIC기반 암호해독 메커니즘 구현 프로그램

②(비공개) 비정형파일 암호해독시스템
- CPU · GPU, ASIC기반 암호해독 시스템

5. 지원기간/예산/추진체계		
ㅇ **기간** : 3년 이내 ㅇ **정부출연금** : '21년 9억원 이내 (총 정부출연금 33억원 이내) ㅇ **주관기관** : 제한없음		
기술분류	대분류(차세대보안) - 중분류(시스템 및 암호보안) - 소분류(암호기술)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (4) ~ (6)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

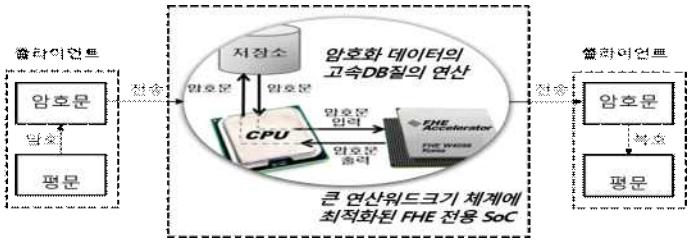
1) 비정형파일, 비정형데이터 : 고정된 필드에 저장되어 있지 않은 데이터를 의미함. 이미지파일, 문서파일, 압축파일 등 (cf. 정형파일, 정형데이터 : 고정된 필드에 저장된 데이터를 의미함. 관계형DB, Spreadsheet 등)
2) 본 사업의 경우 FPGA를 통한 기능검증 및 성능측정으로 추정 성능치 도출

관리번호	2021-차세대보안-05	(지정공모형)

과제명	HW지원 프라이버시 보장 암호데이터 고속처리 기술개발

1. 개념

○ 데이터 프라이버시가 보장된 상태 (암호화 상태)를 유지하면서 데이터 처리와 활용에 필요한 암호데이터 연산 (사칙연산 및 질의연산 등) 고속처리 HW 및 SW 기술



** FHE : Fully Homomorphic Encryption*

2 필요성

○ **(정부지원 필요성)** 데이터 3법통과 이후, 개인정보의 무차별적 활용에 따른 개인 프라이버시 침해에 대한 국민적 우려가 커지고 있어, 국가적 차원의 해결책 필요

- **데이터 유출이 원천적으로 불가능한 실용적 핵심기술** 부재
- **민간 기업은 축적된 기술이 없어서 국가차원의 공공주도 기술개발** 투자가 필요
- **세계적으로 초기단계인 핵심기술 확보 및 확산** 등을 위해 정부의 적극 지원이 필수적임

  * 미국 DARPA에서는 실용적인 동형암호 고속연산을 위해 4년 내외 330억 규모 연구비를 투입하여 “Data Protection in Virtual Environments (DPRIVE)” 프로젝트를 착수('20.3)

○ **(기술성)** 데이터 프라이버시 보호를 위한 현재의 비식별화나 차등 프라이버시 등의 해결책은 **재식별화 시의 위험성과 데이터의 질적 저하 등의 가능성**이 있고, 데이터 유출이 원천적으로 불가능한 완전동형암호가 대안이 될 수 있으나 **실용화를 위해서는 획기적인 성능 개선** 필요

- 실용적인 완전동형암호 기술을 실현하기 위해 현재 64비트 체계에 최적화된 CPU·GPU를 이용한 방식에서 **큰 연산워드크기에 최적화된 FHE 전용 SoC**를 이용한 구현기술이 필요

  * MS는 AI보호 위해 동형암호 선택, IBM은 향후 5년 동안 연구해야 할 기술로 꼽음(LG CNS, 2019.07)

  * Multi-CPU와 GPU를 사용한 처리능력과 전력소모 등의 한계 극복을 위해 비용 효율성, 저전력성, 고속성(연산횟수 낮춤) 측면에서 큰 연산워드크기 FHE 전용 SoC가 효과적임

- 의료, 금융, 공공·개인 데이터를 기반으로 하는 인공지능 모델의 고속 연산은 현재 CPU·GPU에서 전용 인공지능반도체로 진화하고 있으므로 데이터 프라이버시를 보장하면서 **암호화된 데이터 상에서 인공지능 모델 연산을 보장하는 전용 인공지능반도체용 암호기술** 요구 증대

  * 국내에서도 기존 동형암호의 연산속도가 느린 문제점을 응용연산 종류에 따라 최적화되도록 기술이 급속도로 발전하고 있으나, 여전히 **범용산술연산 기반으로 구현 및 검증하는 수준**임

  * 해외는 **완전동형암호의 성능을 개선하는 HW 가속기 개발에 착수**하는 등 동형암호 고속화에 많은 투자를 진행하고 있으며, 특히 완전동형암호 HW 가속기의 주요한 응용으로 **CNN 등의 머신러닝을 고려**하고 있음

- 민감 정보를 포함하는 빅데이터 급증으로 인해 집합데이터를 암호화된 상태로 고속 처리하기 위한 **HW지원 고속 DB암호데이터 처리 기술**이 요구됨

  * 전 세계데이터 양이 '18년에 33ZB에서 '25년에 175ZB까지 증가 전망 (출처:IDC, 2018)

○ **(경제성)** 민감 정보를 포함하는 데이터의 프라이버시를 보호하면서 실용적인 고속연산이 가능한 동형암호 HW 가속 기술을 개발하여 고속 성장(연간 8.06% 성장률)이 예상되는 세계 동형암호 시장을 선점

- **(가치창출)** 실용적인 고속연산이 가능한 동형암호 HW 가속기술을 개발하여 2025년까지 USD 201백만 달러 시장규모로 고속 성장이 예상되는 **세계 동형암호 시장 점유율 확보**

  * 출처: Global HE Market 2020, Forecast to 2025 (GlobalInfoResearch, '20.08)

- **(피해저감)** 데이터 유출이 원천적으로 불가능한 실용적인 완전동형암호 기술개발을 통해 수조 달러 수준에서 증가하는 **사이버 범죄 및 침해와 관련된 비용과 피해 사전예방**

  * 출처: 2020 글로벌 기업 데이터 유출 현황 (IBM, 2020.07)

3. 연구목표

○ **최종목표 :** 데이터의 전주기 프라이버시를 보장하고 동시에 **암호화된 데이터 상의 연산**을 보장하며 양자내성 암호 기술에 부합하는 특성을 지닌 고속 암호연산 처리를 위한 **큰 연산 워드크기 체계에 최적화된 FHE 전용 SoC**와 **고속 암호DB질의 연산 기술**을 개발

- (예상결과물)
  - 큰 연산워드크기 체계에 최적화된 FHE 전용 SoC 및 SDK(Software Development Kit)
  - 암호 데이터 상의 고속 DB질의연산 기술

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고 수준	세계최고수준 (보유국, 기업/기관명)
1	사칙연산 최대 비트	비트	4,096 (전용ALU기준)	N/A	64 (미국,Intel/AMD CPU기준)
2	암호데이터 연산오버헤드	배	10 이내 (평문연산대비)	N/A	10 이내 (평문연산대비)주1)
3	FHE 암호문 Modulus 가변	Range	2^{15}~2^{500}	N/A	2^{15}~2^{500} 주2)
4	FHE Ring Size 가변	Range	512~16384	N/A	512~16384주3)
5	7-layer CNN 추론모델 실행시간 (CIFAR-10이미지당)	ms	≤ 25 ms	N/A	≤ 25 ms주4)
6	7-layer CNN 학습모델 실행시간(CIFAR-10데이터 over 10 epoch)	시간	10	N/A	10주5)
7	암호DB질의연산자 제공 개수	종	4종 (=, <, >, AND, OR) 이상 제공	N/A	1종 (=)주6) (미국, Microsoft / Google)
8	암호화 상태 DB검색 효율성	시간	1ms(×결과 데이터수) 이하	N/A	1ms(×전체 집합데이터 수)이하 (미국, 스탠포드대학)

* 주1)~주5) 미국 DARPA에서 수행중인 '24년 목표 수준임

* 주6) 고정암호화 방식 사용

○ 연차별 개발목표

구분	연도별 연구목표
2021년	큰 연산워드크기 기반 FHE 고속 연산 알고리즘과 암호DB질의 단위연산 설계·검증
2022년	큰 연산워드크기 기반 FHE HW IP와 암호DB질의 복합연산 기술 설계·검증
2023년	최적화된 FHE 가속기 SoC 및 라이브러리와 고속 암호DB질의연산 라이브러리 개발
2024년	암호데이터 연산 고속처리 HW/SW SDK 개발 및 실증

4. 연구내용

○ 개발 기술 내용

① 높은 연산처리 부하를 낮출 HW에 최적화된 FHE 알고리즘 개발

- 큰 연산워드크기에 대해 이진컴퓨팅에 적합한 고속 연산 알고리즘 개발
- 큰 연산워드크기에서 높은 연산 부하를 갖는 곱셈연산에 대한 고속처리 기술 개발

② 가변적인 구조를 제공하는 완전동형암호 하드웨어 가속기 개발

- 가변형 데이터 사이즈 연산에 최적화된 고속 데이터 메모리 관리 기술 개발
- 암호 데이터에 대한 지속적 연산을 보장하는 부트스트래핑 회로 개발
- 완전동형암호 연산용 ALU(Arithmetic logic unit) 회로를 포함하는 고속 FHE 하드웨어 가속기 아키텍쳐 설계 및 시스템반도체 개발

③ 암호화된 상태에서의 DB질의연산 처리 기술 개발

- 동치비교, 대소비교, AND, OR 등을 지원하는 암호DB검색 기술 개발
- 키워드 검색, 범위 검색, 와일드카드 검색 등 다양한 유형 지원하는 암호DB검색 기술 개발
- 다종 암호데이터에 대한 결합검색 기술 개발

④ 하드웨어 가속기용 SDK 및 소프트웨어 설계 환경 개발

- FHE 가속기 상의 소프트웨어 개발을 지원하는 SDK 개발
- FHE 가속기 전용 프리미티브 및 라이브러리 개발
- HW지원 최적화된 고속 암호DB질의연산 라이브러리 개발

⑤ 고속 FHE SoC를 이용하는 실증 응용 개발

- 고속 FHE SoC 및 SDK를 활용한 머신러닝 라이브러리 개발
- 고속 FHE SoC 기반 머신러닝 라이브러리를 이용한 실증
- 공공데이터 기반의 암호DB질의연산 기술 실증

○ 기존 (보유)기술

① 완전동형암호 기계학습 알고리즘 개발 및 라이브러리 구현 (IITP, ‘20년~‘23년)
② 암호화된 데이터베이스 열람·저장을 위한 원천 기술

5. 지원기간/예산/추진체계		
ㅇ **기간** : 4년 이내 (1단계 2년 → 2단계 2년) ㅇ **정부출연금** : '21년 19억원 이내 (총 정부출연금 94억원 이내) ㅇ **주관기관** : 제한없음		
기술분류	대분류(차세대 보안) - 중분류(정보보호) - 소분류(암호기술)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (3) ~ (6)
과제특징	정책지정( ), 혁신도약형(O), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-06	(지정공모형)

과제명	동형암호기술 활용 데이터 프라이버시 보존 국가통계 분석시스템 개발 구축

1. 개념

o 동형암호 (Homomorphic Encryption)

- 동형암호란 데이터를 암호화한 상태에서 복호화 과정 없이 데이터 간의 덧셈, 곱셈 등 연산을 지원하는 4세대 암호기술
  * 즉, 평문 M_1, M_2에 대한 암호문 C_1, C_2가 주어져있을 때, 이 두 암호문으로부터 M_1+M_2에 대한 암호문 C_{add}, 혹은 $M_1 \times M_2$에 대한 암호문 C_{mult} 등을 얻을 수 있음

o 데이터 프라이버시 보존 국가통계 분석시스템

- 공공기관의 대용량 행정통계 데이터를 동형암호 기술을 활용하여 암호화된 상태에서 평균, 상관관계, 주성분분석 등 각종 통계분석을 안전하게 처리하는 분석시스템

2 필요성

ㅇ (정부 지원 필요성)

- 국가기관의 행정통계 데이터를 활용하는 과정에서 발생할 수 있는 의도적·비의도적 개인정보 침해 위협의 선제적 예방 필요
- 기존의 단편적 데이터 프라이버시 보호기술에서 탈피, 민감한 공공데이터를 안전하게 공유·활용하는 서비스 모델 구축을 위해서는 국가 차원의 연구개발 지원이 긴요
- 특히, 국가 행정통계 데이터는 그 유용성에도 불구, 개인정보보호 문제로 공개·활용이 어려운 상황으로 통계자료의 개인정보보호와 데이터 분석 활용이라는 문제를 근본적으로 해결할 수 있는 동형암호 기반의 행정통계 분석시스템 구축 운용 절실
  * 국가통계는 통계청장 승인(통계법)을 받아 426개 통계작성기관(국가, 지자체, 공공기관 등)이 통계자료를 각각 작성·관리하고, 그 외에도 세금, 의료, 보건, 교통 등 행정행위 과정에서 개인정보가 포함된 행정통계 데이터가 다량 수집·관리
- 따라서 국내 원천기술을 확보한 근사동형암호 기반의 프라이버시 보존 행정통계 분석시스템 구축 및 적용을 통해 공공분야는 물론 민간분야의 민감한 데이터를 안전하게 공유하고 활성화하는 데 기여

ㅇ (기술성)

- 국가행정 통계데이터 제공·활용시 K-익명화 등 기존의 가명처리 기술은 데이터의 질을 저하하거나 재식별화 우려 등으로 데이터 이용 활성화에 한계
- 근사동형암호 기술은 우리나라에서 확보한 원천기술로서 암호화된 상태에서 데이터 결합/분석/처리를 가능하게 하는 혁신적인 기술이며 최근 성능도 급속히 향상되어 기존 데이터 보호기술의 한계를 극복하는 4세대 암호기술임

ㅇ (경제성)

- 국세청, 통계청 등 국가기관이 보유하고 있는 민감한 행정통계 데이터의 보호와

데이터 융합·분석 간의 상충 문제를 해결함으로써 행정통계 데이터의 활용도를 크게 개선하여 국민의 프라이버시 보장 및 공공기관 서비스의 신뢰도 향상

- 특히, 통계 데이터의 이용자는 물리적으로 독립된 각 기관별 데이터센터를 방문해야 하는 불편함이 있었으나 클라우드 기반의 온라인 서비스로 사용자의 편의성 제공

※ 통계데이터센터 이용활성화를 위하여 비대면 서비스 지원방안 마련 요구('20년 국회)

- 디지털 뉴딜 세부사업인 데이터댐 구축에 유용한 고품질의 행정통계 데이터를 공유·활용할 수 있어 행정서비스의 질 제고와 국세통계센터, 서울시 데이터센터 등 여타 데이터센터에 同 시스템 기술이전을 통해 데이터시장 활성화에 기여

3. 연구목표

○ 최종목표 :

- 암호화된 상태로 데이터 프라이버시를 보존하면서 대용량 공공데이터에 대한 가공, 연계·분석 및 통계적 처리가 가능한 통계분석시스템 개발 구축 및 적용

○ 과제 목표

- **(정성적 목표)** 데이터보호 핵심 신기술을 통한 통계분석 기술 개발
  . 민감한 데이터를 안전하게 활용할 수 있는 행정통계 데이터 분석시스템 구축·운영
  . 암호화된 행정통계 데이터 융합·결합을 통해 새로운 데이터 가치를 창출할 수 있는 통계분석 라이브러리 개발

○ 정량적 개발목표

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	동형암호 기반 통계분석 기술	기능	일변량통계분석[1)] 이변량통계분석[2)] 다변량통계분석[3)]	NA[4)]	NA[4)]
2	프라이버시 보존 통계분석 시스템 실증	개	동형암호기반 통계분석시스템	NA	NA

1) 일변량분석: 합계, 평균, 중앙값, 최빈수, 범위, 분산, 표준편차 등
2) 이변량분석: 교차분석, T-TEST, 분산분석, 상관분석, 회귀분석 등
3) 다변량분석: 주성분분석, 요인분석, 판별분석, 군집분석 등
4) 세계 최초의 동형암호 기반의 통계분석기술 개발사업으로 비교대상 없음

○ 연차별 개발목표

구분	연도별 연구목표
2021년	- 클라우드 기반의 동형암호기술을 적용한 통계분석 모델 및 기술 개발 - 데이터 처리 및 동형데이터 분석라이브러리 구현과 실증(일변량분석)
2022년	- 데이터 처리 및 동형데이터 분석라이브러리 구현 및 실증(이변량분석) - 근사동형암호를 적용한 통계분석 콘텐츠 발굴
2023년	- 데이터 처리 및 동형데이터 분석라이브러리 구현과 실증(다변량분석) - 데이터 개방 연계·유통·분석·활용 확대를 위한 서비스 체계 구축
2024년	- 동형 대용량 데이터 분석시스템 구현과 실증(서비스) · 자료처리(Editing), 결과제공(시각화 등) 등

4. 연구내용

ㅇ 개발 기술 내용

① 클라우드 기반의 동형암호 통계분석 표준모델 개발

- 암호화 및 복호화 키, 데이터의 효과적인 보관(Archive) 기술 개선/개발
- 동형 데이터 결합 및 가공 라이브러리 구현·서비스를 위한 S/W 구축
- 안정적인 국가 표준의 클라우드 및 망분리 체계의 인프라 구축과 VM 할당 기반을 통한 응용 APP개발

② 대용량 데이터 처리 가능한 통계분석시스템 구축 및 공공기관 서비스 적용

1) 동형암호 기반의 통계분석시스템 구축

- 동형암호 기반의 통계분석 라이브러리를 구축하여 행정통계 자료의 개방·연계·분석 등 활용 확대 연구
- 실용화 측면의 기능목표를 정의하고 동형암호 통계자료의 연산처리 알고리즘 개발
- 통계자료의 연산·자료 처리 영향도 검증

2) 데이터 프라이버시가 보장된 통계분석 콘텐츠 개발 및 서비스 적용

- (자료관리) 근사동형암호를 적용한 인구·가구, 사업체, 기타 통계자료 콘텐츠 발굴
  . 자료간 연계분석이 가능하도록 데이터를 가공할 수 있는 콘텐츠 개발
  . 암호화된 상태에서 평균, 표준편차 등 통계분석이 가능한 함수 개발
  . 연차별 통계분석 함수 확대 개발 (일변량, 이변량, 다변량 분석 등)

- (자료활용) 데이터 개방을 통한 연계·분석·이용 활성화를 위한 서비스 표준모델 개발
  . 근사동형암호를 활용한 통계데이터 이용 절차 표준화
  . 신청승인에 따른 통계자료의 분석·집계 플랫폼 구축
  . 클라우드 컴퓨팅 기반의 근사동형암호 통계자료에 대한 원격접근서비스
  . 분석결과 반출자료의 비식별화(K-익명성, BSCA, 차등정보보호 등) 처리
  . 분석·처리 결과의 시각화(통계테이블, 그래프 등) 기능 구현

③ 기타 기술 실증

- 동형 암호화된 상태에서의 DB연산(검색, 정렬, 중복 제거 등) 처리기술 실증
- 국세, 보건, 복지, 금융, 교통 등 타분야 공공서비스로 확대가능한 표준모델 실증

ㅇ 기존 (보유)기술

① 동형암호 암·복호화 기술 및 기본연산 기술

② 근사계산이 가능한 동형암호 기술 및 재부팅 기술

③ G-클라우드 기술

- 중앙 행정 기관의 스마트 전자정부 서비스를 위해 행정기관의 공동 활용형 정보자원을 필요한 만큼 신속하게 제공하는 기술 및 서비스

5. 지원기간/예산/추진체계		
ㅇ **기간** : 3년 이내 ㅇ **정부출연금** : ‘21년 12억원(총 정부출연금 44억원) ㅇ **주관기관** : 제한없음 (부처 수요기관인 통계청 참여 및 검증)		
기술분류	대분류(차세대보안) - 중분류(공통기반보안) - 소분류(데이터보안)	
연구유형	기초연구 ( ), 응용연구 ( ), 개발연구 (O)	**TRL** (4) ~ (7)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-07	(지정공모형)
과제명	영상 등 멀티미디어 데이터의 온전한 AI 학습/활용이 가능한 복원불가형 개인식별정보 비식별 핵심기술 개발	

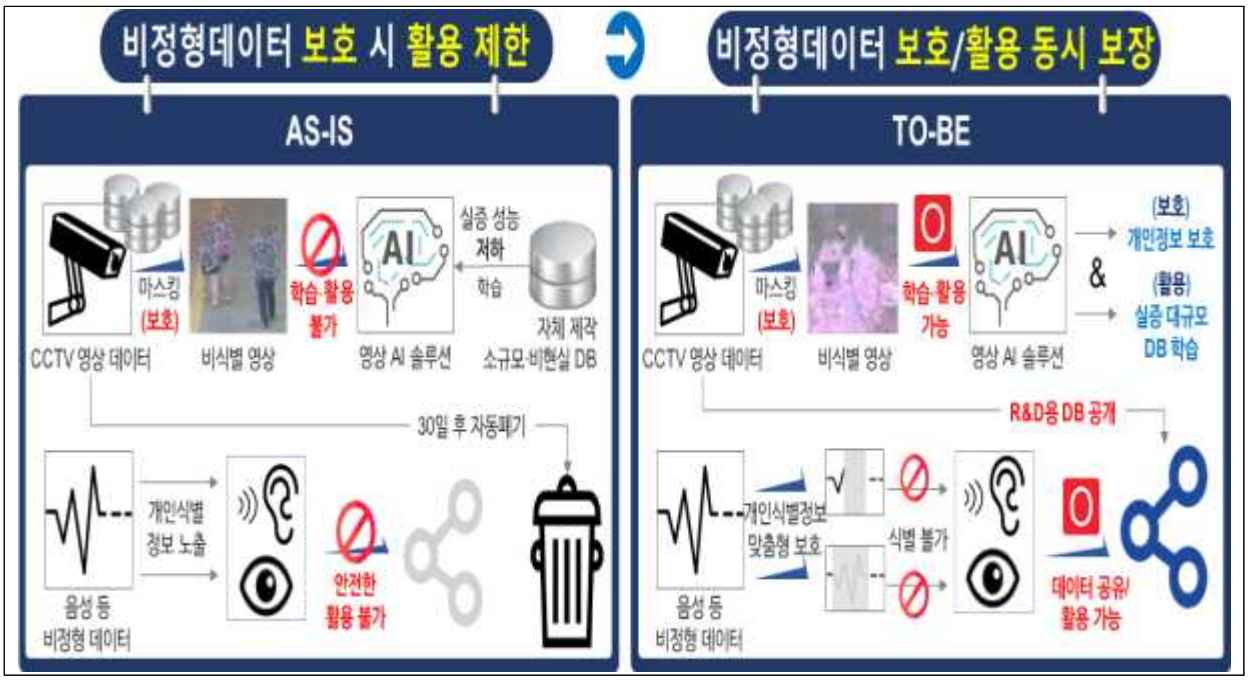
1. 개념

○ CCTV통합관제센터 등을 통해서 수집되는 **공공 영상 내 개인식별정보를 자동 비식별·공개**하여 범용적 **AI R&D에 아무런 제약없이 온전히 활용(학습, 시험 등)** 할 수 있도록 하고, **비정형 멀티미디어 데이터(음성 등)를 맞춤형 비식별화**하는 기술

- 사람의 눈과 시스템으로는 영상 내 개인식별정보를 인식할 수 없도록 복원불가형 비식별 처리하되 AI 학습과 시험에는 성능 저하를 최소화하여 활용할 수 있어야 함
  * 영상 내 개인식별정보(사람 등)는 비식별(보호) 대상이자 동시에 AI 학습/시험(활용)의 대상이므로, 기존 단순 비식별 처리만으로는 AI 연구에 활용이 불가능함
  * AI 학습/시험 분류 : 개인 검출/식별, 이상행동/상황 인식 등

- 기술적/활용적 요구사항
  * **(개인의 식별·유추 가능한 일부/전체 특징정보 노출 방지)** 한번 비식별화된 영상은 특정 개인을 식별·유추할 수 있는 영상특징(feature) 정보(일부/전체) 포함 불가
  * **(원영상 복원 불가)** 특정 솔루션이나 방식으로 원영상을 절대 복원할 수 없어야 함
  * **(범용적 AI에 공통 활용)** 비식별화된 오픈 DB는 모든(범용, 자체 개발한 신경망 포함) 영상보안솔루션(AI, 영상 비전처리 등)에 제약없이 공통 활용될 수 있어야 함 → 모든 신경망 모델 구조, 태스크에 대해서 동일한 비식별 처리 방식 필요
  * **(非개인식별정보 유지)** 영상 프레임 내 개인식별정보가 아닌 영역은 원영상 정보 유지

< 개념도 >



2 필요성

○ **(정부 지원 필요성)** 데이터 3법에 따라 CCTV통합관제센터의 공공 영상데이터의 안전한(개인식별정보 보호) 활용(AI 학습, 시험)을 위한 R&D 지원 필요

- 공공 안전을 위한 AI 영상보안기술은 대부분 지자체 CCTV통합관제센터 영상을 대상으로 하며, 정부차원에서 공공 영상정보를 안전하게 처리, 지원할 수 있는 공공 보안솔루션 제공 필요
- 의료 영상(CT, MR, 내시경, 초음파) 등 의료정보 대상 가명처리를 통한 데이터 활용 수요가 높아짐에 다양한 비정형 멀티미디어 데이터 대상 비식별 처리 기술 필요

○ **(기술성)** 국내에서 영상 기반 AI R&D를 수행하는 많은 산·학·연에서 공통으로 어려움을 겪고 있는 AI 학습용 데이터 확보, 활용에 관련된 이슈로, 글로벌 시장에서의 영상보안기술력 제고와 시장 확보에 반드시 필요

- 기존 영상 비식별 기술은 AI 학습에 필요한 핵심 개인식별정보를 제거하는데만 초점을 맞추고 있어 국내 AI 학습과 시험에 활용은 불가능한 상황임
- AI 학습가능한 개인영상보호, 사생활 보호와 활용이라는 모순적 요구사항을 동시에 보장하는, 새로운 기술 분야의 First Mover 선도 가능

○ **(경제성)** 공공 CCTV 영상/의료영상 등 비정형데이터를 AI 학습용으로 전면 공개하여 각 R&D 수행 시 불필요한 DB 구축 비용 및 시간 절감 가능

- 개인 프라이버시 정보가 삭제된 CCTV 영상과 의료영상을 AI R&D에 활용하기 위해서 온전한 DB 형태로 전면 공개가 가능하며 실환경과 유사한 연구환경 제공
  * 전국 지자체 CCTV통합관제센터 등 약 100만대의 공공용 CCTV 영상 대부분이 AI R&D에 활용되지 못하고 30일 후에 자동 폐기되고 있음
- 특정 AI 솔루션이나 산업체가 아니라 국내 산·학·연에서 수행하는 모든 R&D와 솔루션/제품에 공통 활용하고 비식별화된 DB 공개 가능
  * 국내 지능형 CCTV, VMS, VSaaS 서비스, 스마트시티 안전서비스, 의료기관(병원) 등

3. 연구목표

○ **최종목표 :** 영상 등 멀티미디어 데이터의 **안전한 활용**을 위해 **개인식별정보 보호**와 **온전한 AI 학습을 동시에 보장**하는 **복원불가형 개인식별정보 비식별 핵심기술** 개발

- AI 학습·활용이 가능한 영상 개인식별정보(사람)의 복원불가형 비식별 기술 개발
- 음성, 이미지 등 비정형 멀티미디어 데이터 맞춤형 비식별 기술 개발

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	원영상 학습 대비 비식별 영상 AI 학습 성능(mAP*)	%	95	99[1)]	대한민국,딥핑소스
2	영상 비식별 처리 속도	fps	21	31.3[2)]	중국, SOLOv2
3	음성 개인식별정보 선택적 비식별 성능(CER**)	%	7	7[3)]	미국, 구글
4	수요처 실증 서비스	개	2	-	-

* mAP : mean Average Precision, **CER: Character Error Rate

1) Resnet에 특화된 얼굴인식 학습 성능, 범용 신경망에서의 태스크(개인/행위/상황인식 등) 성능 지표는 없음

2) 영상 내 사람 세그멘테이션 성능 지표, 실시간 영상 비식별(후처리) 속도 성능 수치는 미포함

3) voice-to-text 성능 지표, 실시간 음성 비식별(후처리) 성능 수치는 미포함

○ 연차별 개발목표

구분	연도별 연구목표
2021년	AI 학습·활용이 가능한 영상 개인정보 비식별 프로토타입 개발
2022년	음성, 이미지 등 비정형 멀티미디어 데이터 맞춤형 비식별 기술 개발
2023년	AI 학습·활용이 가능한 영상 개인정보 비식별 AI 기술 개발 및 고도화
2024년	공공 CCTV/의료데이터 실증 및 스마트시티 빅데이터 플랫폼 연동

4. 연구내용

○ 개발 기술 내용

① AI 학습·활용 가능하고 복원불가한 영상 개인식별정보(사람) 비식별 기술 개발
- AI 학습·활용이 가능하되 프라이버시 없는 아이덴티티로 비식별 처리 기술 개발
  * 기존 AI 솔루션 수정이나 비식별 데이터의 변환없이 활용될 수 있어야 함
  * 비식별화된 데이터를 통해서 사람이나 시스템이 원영상의 (일부/전체) 식별 불가
  * 특정 API, 라이브러리, 플랫폼에 독립적이고 DB 공개가 가능해야 함
- 연속된 영상 프레임에서의 비식별화된 객체의 temporal consistency 유지 기술
  * 동영상의 연속된 프레임에서의 객체 특성을 분석하는 AI 기술(추적, 행위/상황분석 등)을 위해 비식별화된 객체의 아이덴티티 일관성 유지 필요
- 영상 내 개인식별정보 영역의 고정밀 세그멘테이션 AI 기술 개발
- 실증 학습용 DB 구축을 위한 자동 GT(Ground Truth) 분석 및 검증 기술 개발

② 공개용 비식별 영상에 대한 AI 학습 성능, 활용도 시험 및 검증 기술 개발
- 원영상 대비 비식별처리 영상에 기반한 AI 학습 성능 시험 및 검증 기술 개발
- 원영상으로 학습된 AI 솔루션을 적용해서 비식별화된 영상 내에서 인식 시험 및 성능 검증 기술 개발
- 원영상 복원 가능성 시험을 통한 영상 비식별화 보안성 검증 기술 개발

③ 실시간 음성 데이터 내 개인식별정보 선택적 비식별 기술 개발
- 음성 내 개인식별정보 영역 실시간 자동 검출 기술 개발
  * 음성 내 주민등록번호, 통장번호, 주소, 전화번호 등
- 실시간 검출된 개인식별음성의 선택적 비식별 및 인가 복원 기술 개발
- 공공장소에서 발화되는 음성의 온-에어 식별 난이도 증가 기술 개발

④ 이미지(의료) 데이터 맞춤형 자동 비식별 기술 개발
- 의료이미지(CT, MR, 내시경, 초음파) 內 개인정보(얼굴, 신체 등)가 복원가능한 영역의 부분·선택적 비식별 기술 개발
- 의료이미지 內 개인정보가 복원가능한 영상 영역 자동 검출 기술 개발

⑤ 공공 CCTV/의료데이터 실증 및 스마트시티 빅데이터 플랫폼 연동 기술 개발
- 지자체 주도형 공공데이터 안전 리빙랩 구축 및 의료기관 실증 플랫폼 연동
- 공공 CCTV 데이터를 처리하는 지자체 CCTV통합관제센터, 의료정보를 연구용으로 활용하는 CDW(Clinical Data Warehouse) 조기 탑재를 위한 협업체계 구축
- 법·제도적 분석 및 개선방안, 국내·외 표준 및 관련 기술 개발

○ 기존 (보유)기술

① 원영상 복원이 가능한 실시간 프라이버시 비식별 기술
- 부분 영상 프라이버시 마스킹/복원(코덱 독립형) 기술, 전체 영상 암·복호 기술 등
② 영상 내 개인식별정보 영역(바운딩 박스) 검출 및 단방향 비식별 기술
③ 개인특징정보를 포함한 타겟 신경망/태스크 종속형 영상 비식별/학습 기술
④ 음성 등 비정형 데이터 전체 정보 비식별 기술

5. 지원기간/예산/추진체계

○ **기간** : 4년 이내 (1단계 2년 → 2단계 2년)

○ **정부출연금** : '21년 12억원 이내 (총 정부출연금 57억원 이내)

○ **주관기관** : 제한없음 (산업체 참여 필수)

기술분류	대분류(차세대보안) - 중분류(데이터 및 응용서비스 보안) - 소분류(데이터보안)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (3) ~ (6)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

과제번호	2021-차세대보안-08	(지정공모형)
과제명	**대용량 정형 데이터 대상 비식별처리 자동화 및 안전성 검증 기술개발**	

1. 개념

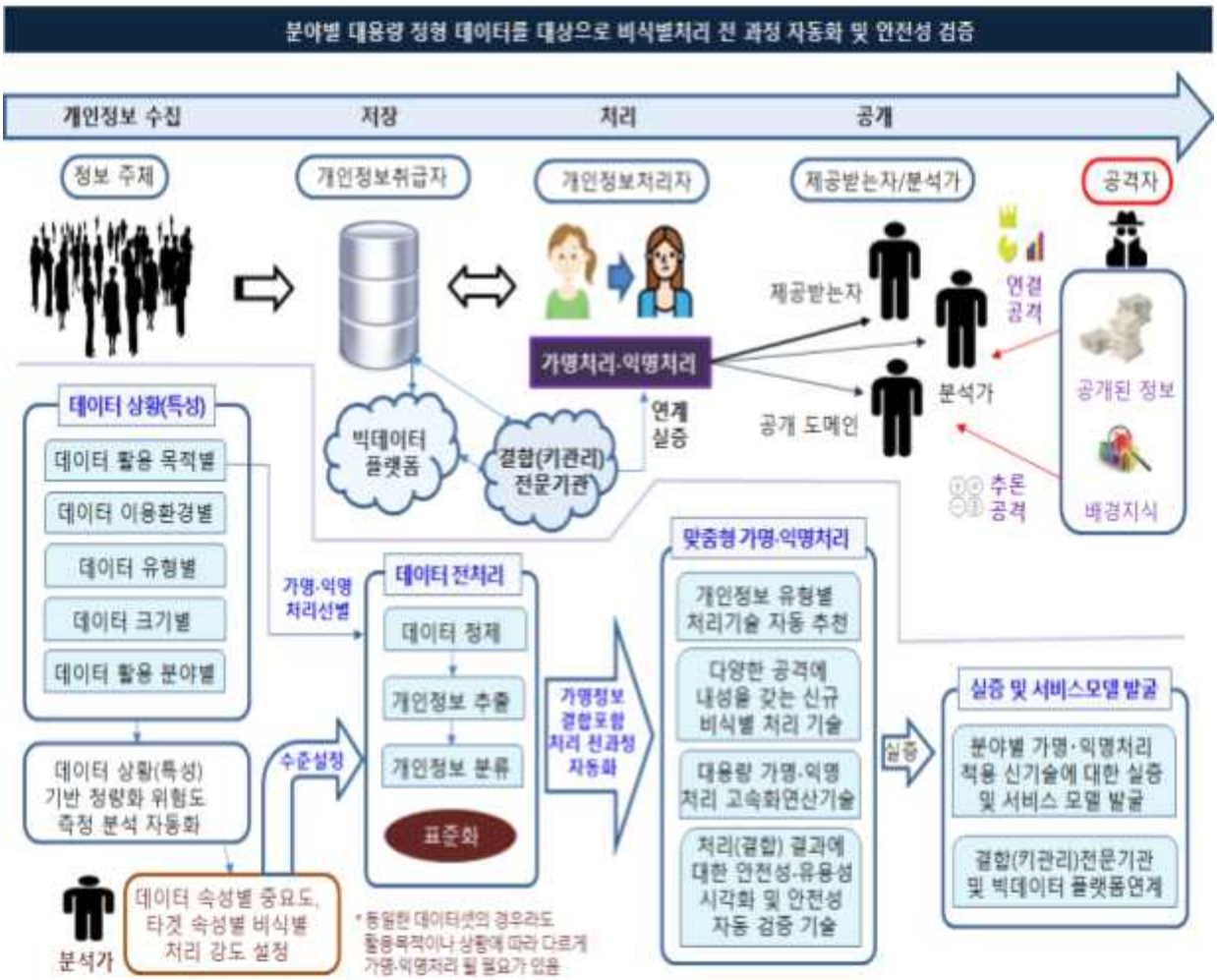
○ 산업 분야(금융, 통신, 의료, 유통 등)별 대용량 정형 데이터를 대상으로 개인정보를 자동으로 분류한 후 데이터 특성 및 환경을 고려한 위험도*에 따라 맞춤형으로 비식별 고속 연산 처리**하고 처리†된 결과의 안전성과 유용성을 정량적으로 시각화하고 검증‡하는 등 처리 전 과정을 자동화하는 기술

* 개인정보보호법 개정에 따른 데이터 활용목적에 따라 가명과 익명처리를 자동으로 선별하고 데이터 특성과 환경을 고려한 정량적 위험도를 자동으로 산출

** 가명처리와 익명처리를 포괄하며 국제표준과 기존 알려진 기술 이외 신기술 개발 및 대용량 처리에 적합한 고속화 연산처리 기술을 포함

† 처리시 개인정보보호법 개정에 따른 서로 다른 개인정보처리자간 가명정보 결합 처리 지원 포함

‡ 비식별 처리 기술의 안전성과 유용성 측정을 위한 정량지표 개발, 이를 비식별 처리전 위험도와 비교·시각화하고 처리 결과의 안전성을 자동으로 판단·검증하도록 함



< 대용량 정형 데이터 대상 비식별처리 전 과정 자동화 및 안전성 검증 개념도 >

2. 필요성

○ **(정부 지원 필요성)** 2020년 8월 데이터 3법 개정 이후로 비식별 조치(가명, 익명처리)를 통한 데이터 활용 수요가 높아짐에 따라 다양하고 규모가 큰 데이터 대상 비식별 처리 기술 개발 필요

- 기존 정형화된 데이터에서의 가명·익명처리 뿐만 아니라 전산업 분야와 이종산업 간 데이터 가명·익명처리 및 결합이 가능하도록 핵심 원천기술 확보가 요구됨
- 특히 개인정보 비식별처리, 이종산업 간 가명정보 결합 시 가명정보 활용 목적, 환경 등 전방위에 걸친 데이터 분석, 프라이버시 침해 위험도 측정·안전성 평가 및 검증 필요
- 민간 기업의 데이터 활용 수요를 담보하기 위해서는 개인정보의 안전성에 대한 대국민 우려를 불식시키기 위한 비식별 처리 분야 신기술 확보가 관건임

○ **(기술성)** 개정된 데이터 3법에 부합한 가명처리와 익명처리의 안전성을 제공하여 국민으로부터는 신뢰를 기업으로부터는 활용에 대한 수요를 담보하는 핵심 원천기술 확보 가능

- 가명처리에 대한 법적 명문화는 지난 2018년 5월 EU GDPR(유럽 개인정보보호법) 개정으로 이미 독일 뮌헨공대 의료정보연구실을 비롯한 EU내 스타트업 기업과 데이터 관련 기업들로부터 연구가 진행 중으로 핵심 원천기술 확보가 시급
- 미국의 경우 Google을 비롯한 글로벌 IT기업을 중심으로 비식별 처리 기술인 차등정보보호기술을 적용하여 마케팅 등에 활용하고 있으며, 의료분야에서는 HIPAA 개인정보보호 규정에 따라 HITRUST 등 기업들이 독자적인 솔루션을 이미 보급하고 있는 상황임

○ **(경제성)** 비식별 처리된 대용량 데이터의 안전성 보장을 통한 신기술, 개인정보 생명주기에 따른 전 과정 자동화 처리, 데이터 위험도 및 안전성과 유용성에 대한 정량화된 시각화 도구 제공 등 차별화된 신기술 제품화로 글로벌 시장에 진입

- 비식별 처리 서비스 전 과정의 자동화된 데이터의 안전성과 유용성 보장을 통해, 대국민의 안전성 우려로 시장 적용이 미흡했던 다양한 분야*에 적용·확산 가능
  * 의료, 통신, 유통 및 공공 등 개인정보 활용을 통해 공익이나 과학적 연구 등 가명처리나 혹은 마케팅 등 익명처리를 통해 다양한 이익을 창출하는 서비스 분야
- 전 세계 데이터 시장 규모는 '27년 1천억 달러에 달할 것으로 예상되며, 데이터 활용에 필수 요소 기술인 비식별 처리와 맞물리면서 향후 개인정보의 안전성 및 유용성 보장 기술의 폭발적인 수요 증가 예상
- 결합전문기관이나 기업이 가명정보 결합 등 데이터 활용에 따른 재식별 등 프라이버시 침해 위험요소를 체계적으로 진단 및 관리하여 관리 실수, 내부자에 의한 고의 유출 등 정보유출방지, 정보보호 솔루션, 컨설팅 관련 예산을 감소시키는 경제적 효과 예상

3. 연구목표

ㅇ 최종목표 : 분야별 대용량 정형 데이터 대상 개인정보 자동 추출·분류, 데이터 특성과 환경을 고려한 위험도에 따른 맞춤형 비식별 고속 연산 처리 및 안전성 검증까지의 처리 전 과정에 대한 자동화 기술 개발

ㅇ 정량적 개발목표

핵심 기술/제품 성능지표		단위	달성 목표	국내 최고수준	세계최고수준 (보유국, 기업/기관명)
1	테라급 데이터셋에 대한 전처리 성능[주3]	시간(분)	250분 이내	300분[주5]	284분[주7] (미국, Apache재단)
2	테라급 데이터셋에 대한 비식별 성능[주2]	시간(분)	250분 이내	300분[주5]	284분[주7] (미국, Apache재단)
3	테라급 두 데이터셋들 간의 결합 성능[주1]	시간(분)	250분 이내	500분[주4]	279분[주6] (미국, Apache재단)

주1) 결합 데이터셋은 정렬되지 않은 파일 크기 1TB 이상 두 개 데이터셋(CSV 형식)을 기준으로 하며 결합 결과 데이터셋(CSV 형식)이 총 파일 크기 1.8TB 이상(결합율 90% 가정), 열 개수 500개 이상, 행 개수 2억개 이상 기준.

결합 데이터셋은 일반적으로 결합된 큰 데이터를 저장하는 데 많은 시간이 소모되고 이를 내보내는 경우가 많으므로 하나의 결합 파일로 저장되는 시점까지의 시간을 모두 포함.

주2) 비식별 연산은 데이터셋 중 이름에 해당하는 식별정보 컬럼에 10byte salt 값을 덧붙여 사용한 SHA-256 알고리즘으로 해시하는 연산을 사용.

주1, 2, 3) 실제 고객의 대용량 데이터셋 처리 요구 사항(RFP 요건)을 바탕으로 산정. 결합 데이터셋에 대한 비식별 처리가 필요하므로 이에 기반 함

주4, 5) 실제 운영 환경 측정값 기준, 추정 계산

주6) 결합 성능은 실제 사용 시나리오를 고려하여 결합 대상 CSV 파일들의 적재, 결합 및 하나의 CSV 파일 형식으로 저장까지를 포함함

파일의 적재와 저장 속도는 하드웨어 의존적이므로 SSD SATA 구성(1.6 Gbps-2.4 Gbps)을 가정하여 Read는 300MB/sec, Write는 200MB/sec로 계산. 실제 결합에 걸린 시간은 Big Data에 대한 표준인 BigBench 성능 측정 결과를 기준으로 추정. 전체 BigBench 성능 쿼리 중 Join이 포함된 14개 SQL 유형에 대한 최고값을 보여준 환경의 Spark 프레임웍 값(1689초)의 산술 평균을 사용(120+6667+10000초)

[참고문헌 1] 'Characterizing BigBench queries, Hive, and Spark in multi-cloud environments', Barcelona Supercomputing Center, 2020. 7

주7) 기준치는 8 CPU core, 60GB of RAM로 구성된 16개 Worker 노드로 구성. 상기 [참고문헌 1]은 30개 쿼리에 대한 총 결과치를 측정한 값이므로 최고 속도로 수행된 환경인 오픈소스 병렬 처리 프레임워크 Spark의 값(3시간 13분 10초)을 산술평균하여 기준치를 산정하고 2TB의 Read, Write를 주6)에서와 같은 방법으로 합산(386+6667+10000초 = 284분)하여 계산

ㅇ 연차별 개발목표

구분	연도별 연구목표
2021년	데이터 특성과 상황을 고려한 위험도 측정 자동화 및 프레임워크 설계
2022년	맞춤형 가명·익명처리를 위한 데이터 전처리 자동화 및 신기술 개발
2023년	가명정보 결합을 포함한 가명·익명처리 전 과정 자동화 및 고속화 연산 처리 기술 개발
2024년	가명·익명처리 결과의 시각화 및 안전성 자동검증 도구 개발, 결과 실증 및 서비스 모델 발굴

4. 연구내용

ㅇ 개발 기술 내용

① 데이터 특성, 상황 및 처리 환경에 따른 프라이버시 침해 위험도 측정 자동화 기술

- 가명·익명정보 활용목적, 활용방법 및 이용환경에 대한 위험도 분석 지표 개발
- 가명·익명정보 데이터셋에 대한 위험도 분석 지표 개발
- 상기 분석 지표들을 바탕으로 한 프라이버시 침해 위험도 측정 자동화 도구 개발
  * 자동화 도구 개발 시 위험도 측정 결과에 따른 가명·익명처리 수준 정의 포함
  * 처리 수준에 따라 가명·익명처리 기술 선정 및 결합 등 처리가 가능하도록 위험도 분류

② 맞춤형 가명·익명처리를 위한 데이터 전처리 및 비식별처리 자동화 기술 개발

- 데이터 특성과 상황(데이터 활용 방법, 데이터 이용환경, 데이터 자체 등)에 대한 정의 및 분류 표준화
- 맞춤형 가명·익명처리를 위한 데이터 정제 기술
- 맞춤형 가명·익명처리 기술 적용을 위한 사전 개인정보 추출·분류 표준화 및 자동화 기술
  * 데이터 유형별, 크기별, 활용 분야별 특성 반영 필수
- 전처리 단계에서 분류된 개인정보 유형별 맞춤형 가명·익명처리 기술 추천 알고리즘 개발
- 개인정보보호법에 따른 데이터 활용 목적별 분류 및 가명·익명처리 선별 자동화 기술 개발
- 공격자로부터의 연결·추론 등 다양한 공격에 내성을 갖는 신규 비식별 처리 기술(예 : 다형성(Polymorphic) 암호기반 가명처리 등, 프라이버시 보호 모델 포함) 개발
  * 신기술은 최소 3종 이상 개발하되 개인정보보호위원회에서 발간한 가명정보 처리 가이드라인에서 소개하고 있는 비식별 처리 기술들을 모두 수용(단, 타과제와 중복된 동형암호 및 동형비밀분산 기법 제외)
- 가명정보 결합을 포함한 가명·익명처리 전 과정에 대한 자동화 처리 기술 개발
  * 단, 위험도(활용목적 포함)에 따라 분석가가 데이터 속성별 중요도, 타겟 속성별 비식별 처리 강도 설정 후 자동화 처리(왜냐하면 동일한 데이터셋의 경우라도 활용 목적이나 데이터 상황에 따라 다르게 비식별 처리될 필요가 있음)

③ 가명·익명처리된 개인정보의 안전성 및 유용성에 대한 안전성 검증 자동화 기술 및 정량지표 기반 시각화 도구 개발

  * 시각화 도구 개발 시 데이터 특성과 상황을 고려하여 측정한 위험도와의 비교 기능 탑재
  * 안전성 검증 시 평가기준은 국제표준 ISO/IEC 20889에서 제시한 특정가능성(single out), 연결가능성, 추론가능성 등 참조

④ 테라(Tera)급 이상 대용량 정형 데이터의 가명·익명·가명결합 처리 고속화 기술 개발

- 분산 병렬 처리 및 인메모리 컴퓨팅에 기반한 고속 연산 처리 엔진 개발
  * 병렬 처리 기반 가명처리 기법 적용(개인정보보호위원회 가명 처리 가이드라인 기준)
  * 고속 연산 처리시 컨테이너 기술(예: Docker, K8S 등) 수용
- 테라급 대용량 정형 데이터를 가명·익명 처리하기 위한 맵리듀스 알고리즘 구현 플랫폼 개발
  * 한정된 CPU 자원 외에 일부 가명·익명처리 알고리즘을 가속화할 수 있는 GPU 하드웨어 지원 알고리즘 기능 탑재(지원 GPU가 있는 경우와 없는 경우 동일한 알고리즘을 스위칭

하여 사용할 수 있도록 알고리즘 개발)

- 유한한 GPU 자원을 최대한 활용하고 CPU 자원과 함께 혼용하여 사용할 수 있는 하드웨어 자원 인식형 분산 리소스 스케줄러 엔진 개발

  * 테라급 데이터를 처리하기 위해 최대 50노드의 분산 처리가 가능할 것

⑤ 대용량 가명·익명처리, 안전성 검증 자동화 기술개발 결과 실증 및 서비스 모델 발굴

- 대용량 정형 데이터 대상 분야별 맞춤형 가명·익명처리, 안전성 검증 관련 기술 개발 결과에 대한 실증 및 서비스 모델 발굴

  * 실증 시 빅데이터 플랫폼과 연계하되 다양한 분야(의료, 통신, 유통, 금융 등)를 고려하여 최소 2곳 이상 지정할 것. 단, 최소 2곳 중 1곳은 가명정보 활용 계획, 처리, 데이터 결합과 반출 시 적정성 평가를 담당하는 해당 분야 결합(키관리)전문기관과 연계하여 실증

  * 상기 ①~④항에서 기 개발한 기술을 모두 적용하고 적용 결과에 따른 신규 서비스 모델을 도출할 것

o 기존 (보유)기술

① Google 등 글로벌 IT기업을 중심으로 자사 플랫폼 환경에서 차등 프라이버시를 이용한 비식별 자동화 처리 기술 개발 및 서비스 진행 중

② EU를 중심으로 국제표준 ISO/IEC 20889기반 비식별 처리기술을 표준화하고 이를 적용한 기술 개발 중

③ 비식별 처리를 위해 EU의 경우 데이터 상황, 캐나다, 호주의 경우 데이터 주변 맥락을 고려한 개인정보의 위험도 산출을 자국 가이드라인으로 제시

- 영국 UKAN, 미국 HITRUST, 캐나다 Privacy Analytics사를 중심으로 자국 혹은 자사의 데이터 위험도를 고려한 비식별 처리 방법론을 개발, 이를 적용한 비식별 처리와 컨설팅을 병행하여 서비스가 진행 중임

5. 지원기간/예산/추진체계

o 기간 : 4년 이내 (1단계 2년 → 2단계 2년)

o ’21년 정부출연금 : 12억원 이내

o 총 정부출연금 : 57억원 이내

o 주관기관 : 제한없음 (산업체 참여 필수)

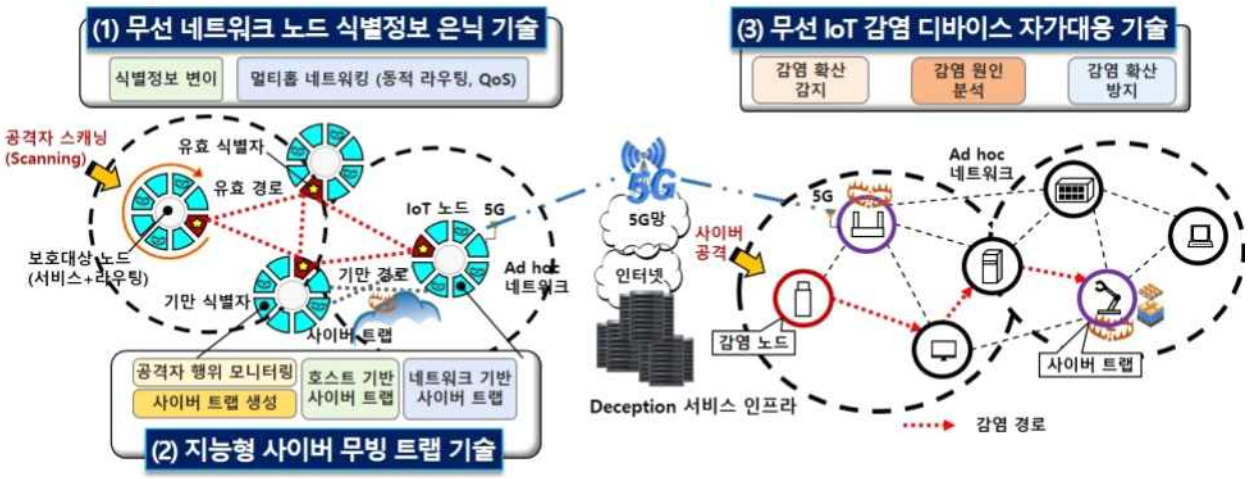
기술분류	대분류(차세대보안) - 중분류(데이터보안) - 소분류(프라이버시 보호)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	(3)~(6)
과제특징	정책지정( ), 혁신도약형(O), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리 연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-09	(지정공모형)
과제명	5G Massive 디바이스 공격접점 은닉을 통한 차세대 사이버공격 기만기술 개발	

1. 개념

○ 스마트시티 등 5G+ 서비스를 위협하는 고도의 사이버 공격에 선제 대응하기 위하여 상황에 따라 디바이스 식별정보 동적변이 및 지능형 사이버 무빙 트랩 기반의 악성코드 확산 자가대응 기술을 통한 IoT 디바이스 환경의 능동적 사이버 공격 기만기술 개발

- 디바이스 식별정보의 동적변이를 통해 공격에 소요되는 시간과 노력의 복잡도를 높일 뿐 아니라, 사이버 무빙 트랩이 악성코드 감염확산을 탐지하여, 사이버 공격의 사전예방과 사후대응이 융합된 능동형 사이버 공격 방어체계 구축



< 개념도 >

2 필요성

○ **(정부 지원 필요성)** 5G+ 융합서비스의 안전한 대국민 서비스 제공을 위해 5G 코어망 보안, 에지 클라우드 보안과 함께 5G 환경에 특화된 IoT 디바이스 보안에 대한 정부차원의 선제적 대응이 필요함

- 5G 네트워크를 통한 대규모의 IoT 디바이스들의 인터넷 연결은 사이버 공격에 활용될 수 있는 공격접점 (Attack Surface)의 확대를 의미함
  * APT (Advanced Persistent Threat)의 여러 단계에서 사용되는 접점 (Surface)들을 Exploration Surface와 Attack Surface로 구분하는 연구도 있지만, 본 문서에서는 이 접점들을 공격 접점이라는 용어로 표현함
- 공격자가 공격접점을 악용하지 못하도록 공격접점 자체를 사전에 은닉하는 사전예방 기술과 사이버 무빙트랩에 기반한 사후대응 기술이 융합된 능동형 사이버 공격 방어체계 구축에 대한 정부차원의 기술개발 추진이 필요함

○ **(기술성)** 사이버 공격 사전예방 기술은 선진국에서도 초기 연구단계로, 핵심 원천기술을 확보하여 사이버 공격 사전예방 분야의 기술수준 향상과 기술 경쟁력

확보를 통한 기술 선점이 필요함

- 대규모 IoT 디바이스들이 연결되는 5G 환경에서 공격자로부터 공격접점의 은닉을 위하여 활용 가능한 사이버 공격 사전예방 기술로 사이버 자가변이 (Moving Target Defense) 기술과 사이버 기만 (Deception) 기술이 존재함
- 5G를 활용한 대규모 IoT 디바이스 환경에서 사이버 공격 사전예방 기술 연구가 전무한 상태에서, 해당 분야의 핵심 원천기술 확보가 시급함

○ **(경제성)** IoT 디바이스에 대한 사이버 공격을 원천적으로 차단하여 국민생명을 보호하고, 대규모 공격으로 인한 사회적 비용을 최소화

- 5G IoT 디바이스 후보인 스마트 CCTV, 스마트 미터기 뿐만 아니라 자율주행차와 드론, 로봇 등의 무인 자율이동체 디바이스에 대한 사이버 공격으로 인한 심각한 사회/경제적 손실 발생 가능함

  * 軍은 드론부대를 창설(2018.10)하는 등 드론 부대나 로봇 병사와 같은 무인 자율이동체를 미래 국방전략으로 활용

  * 3조6천억 규모의 부산(23년 완료), 세종(21년 완료) 스마트시티 시범 도시 선정 구축, Massive IoT 디바이스(교통, 의료, 안전, 에너지, 로봇 서비스 등)로 구성
- 5G로 연결된 대규모 IoT 디바이스들 간의 초고속 악성코드 확산으로 인하여 더욱 강력하고 치명적인 DDoS 공격 발생 가능함

  * '16년 다인(Dyn)사가 미라이 악성코드에 감염되어 대규모 DDoS 공격을 받아 트위터(Twitter), 넷플릭스(Netflix), 뉴욕타임즈(NYT) 등 총 76개의 사이트가 일제히 마비되거나 서비스가 지연

  * 16년 wannacry 공격으로 샌프란시스코 경전철, 17년 영국 국영 철도망 티켓 발권 서비스로 교통서비스 마비 피해

  * 17년 NotPetya 공격으로 Fedex, Maersk 물류 운송, Merck 제약회사 등 모든 작업 중단, 총 12조원 이상의 복구 피해 발생

3. 연구목표

○ **최종목표 : 5G+ 서비스를 위협하는 고도의 사이버 공격에 선제 대응하기 위하여 상황에 따라 디바이스 식별정보 동적변이 및 지능형 사이버 무빙 트랩 기반의 악성코드 확산 자가대응 기술을 통한 IoT 디바이스 환경에서의 능동적 사이버 공격 기만 기술 개발**

- 기존 5G 코어망 보안 및 5G 에지 클라우드 보안에 추가로, 대규모 5G IoT 디바이스 환경에 적합한 보안 기술 개발
- 디바이스 식별정보 동적변이 기술: IoT 디바이스의 식별정보를 지속적으로 변경하여 공격자에게 식별정보 노출을 사전 차단하는 기술
- 지능형 사어버 무빙 트랩 기술: 정당한 디바이스를 위장하는 거짓 IoT 디바이스를 생성/운영하여 공격 복잡도를 높이고, 사이버 공격을 감시하는 기술
- IoT 악성코드 감염확산 자가대응 기술: 사이버 무빙 트랩을 이용하여 악성행위를 탐지하고, 악성코드의 확산에 자가 대응하는 기술

○ 정량적 개발목표

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	IoT 사이버 트랩 기반 사이버 공격 탐지율	%	> 90	NA	(> 90) (TrapX, 미국)
2	IoT 사이버 트랩 기반 사이버 공격 오탐율	%	<= 5	NA	(< 10) (TrapX, 미국)
3	공격 목표물 탐색 복잡도(탐색 시간 증가율)	배	10배 (기술 적용 후/전)	NA	*

* UNC Charlotte, College of William and Mary (미국)에서 서버급 장비의 주소변이 성능 수준으로는 30배이지만, IoT 디바이스의 주소변이 성능은 보고된바 없어서, 해당 부분을 표기하지 않음

○ 연차별 개발목표

구분	연도별 연구목표
2021년	테스트베드 생성 및 시스템 설계 요소기술(IoT 디바이스 식별ID 변이, 사이버 무빙트랩) 알고리즘 설계 국내외 IPR 확보
2022년	IoT 디바이스 식별ID 변이 기술 개발 사이버 무빙트랩 및 유도 기술 개발 웜 악성코드 감염 탐지 기술 설계 국내외 IPR 확보
2023년	웜 악성코드 감염 탐지 기술 개발 통합 시스템 프로토타입 개발 (테스트베드)
2024년	웜 악성코드 감염 탐지 실증 서비스 차세대 사이버공격 기반 기술 통합 시스템 검증

4. 연구내용

○ 개발 기술 내용

① 안전한 5G+ 서비스를 위한 무선 네트워크 환경의 주요 노드 식별정보 동적변이 기술

- IoT 디바이스 식별정보 동적변이 기술
- 멀티홉 무선 동적변이 네트워크 환경에서 끊김없는 고속 서비스를 위한 초저지연 연결 관리 기술
- 다중 디바이스 식별정보 동적변이를 위한 이기종 무선 네트워크 프로토콜 지원 미들웨어

② 사이버 공격 기만을 위한 지능형 사이버 무빙 트랩 기술

- 5G+ 서비스 맞춤형 기만 IoT 디바이스 관리 기술
- 공격자 기만을 위한 지능형 사이버 무빙 트랩 생성 및 유도 기술
- 사이버 트랩을 통한 위협(공격자+악성코드) 행위 모니터링 관리 기술

③ IoT 디바이스의 악성코드 감염확산 자가 대응기술 개발

- 공격대상 디바이스의 악성코드 감염 및 확산 감지 기술
- 공격대상 디바이스의 악성코드 (미라이봇넷, 워너크라이 등) 감염 원인분석 / 격리기술

④ 초신뢰 5G 서비스 제공을 위한 유무선 기만 네트워크 통합 기술
- 5G 기반 IoT 환경에서 발생 가능한 위협 데이터 수집 및 정형화
- 서비스 시나리오 기반 멀티 홉 무선 네트워크 테스트베드 생성
- Massive 디바이스 환경의 능동적 사이버 공격 예방 체계 구축

ㅇ **기존 (보유)기술**

① 엔터프라이즈급 ICT 인프라를 대상으로 지능형 사이버 위협에 사전에 대응하기 위하여 상황에 따라 공격자로부터 보호대상이 스스로 변이함으로써 능동적 사전 보안을 실현하는 사이버 자가변이 기술 개발 중

5. 지원기간/예산/추진체계

ㅇ **기간** : 4년 이내

ㅇ **정부출연금** : '21년 9억원 이내 (총 정부출연금 45억원 이내)

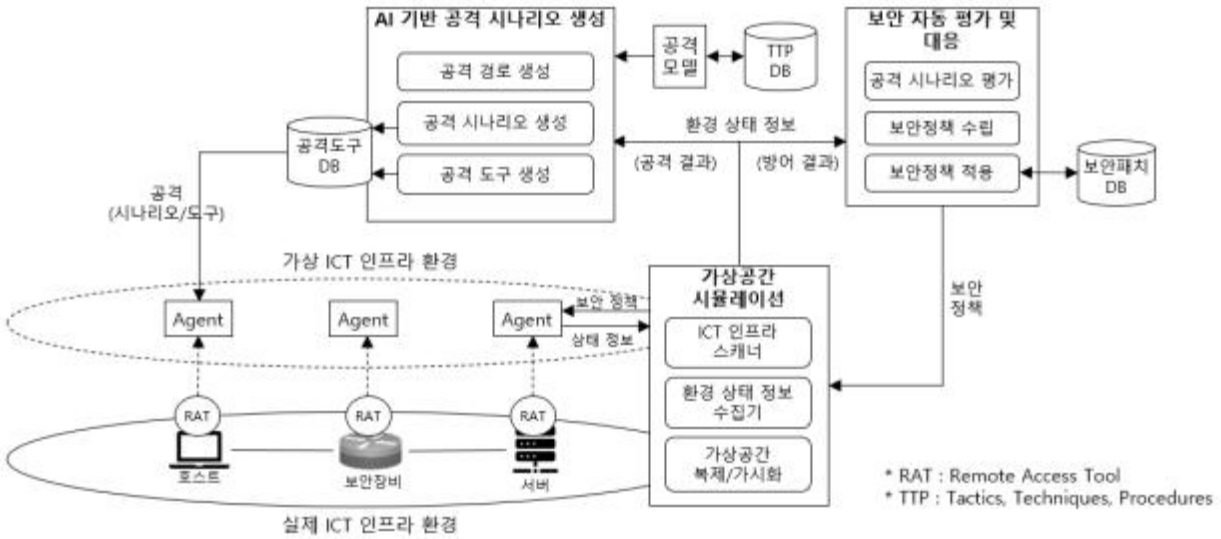
ㅇ **주관기관** : 제한없음 (산업체 참여필수)

기술분류	대분류(차세대보안) - 중분류(네트워크보안) - 소분류(무선네트워크보안)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (4) ~ (6)
과제특징	정책지정( ), 혁신도약형(O), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-10	(지정공모형)
과제명	위협헌팅 모델 기반 지능형 사이버 공격/방어 분석 프레임워크 기술 개발	

1. 개념

○ **위협헌팅 모델**에 기반한 사이버 **공격 시나리오를 자동으로 생성**하고 가상 ICT인프라 공간에서 시뮬레이션을 통해 보안장비의 사이버 공격 대응력 검증, 보안체계의 적합성 판단 및 궁극적으로 보안취약점을 사전에 확인하여 **사이버 공격을 선제 대응하는 Offensive 솔루션**



< 개념도 >

2 필요성

○ **(정부 지원 필요성)** 제안 주제는 R&D 전략 중 AI 보편화에 대한 활용 및 대응 기술로써 차세대보안 기술로드맵 중 **"AI기반 사이버공격 및 방어 시뮬레이션 기술"**과 관련성이 높음

- AI 강화 사이버 공격(AI 해커)은 전통적인 보안기술로는 대응이 불가하여 AI 기술을 활용한 해킹 예방 시뮬레이션 기술이 필요

○ **(기술성)** 주요 시설의 보안정책 수립/적용을 위해 복제된 ICT인프라를 대상으로 AI 기술을 활용한 위협헌팅 모델(예, MITRE ATT@CK 등)사이버 보안위협 공격, 방어 시뮬레이션 프레임워크 개발을 위한 연구주제임

- BAS 기술은 비대면 환경과 노동집약적 보안관제운영의 무인화를 위해 **기존 침투테스트 및 레드팀 평가 업무를 자동화**하기 위한 핵심 요소기술로 판단됨
- 5G+ 핵심서비스 중 하나인 스마트 팩토리 보안과 보안성 평가를 위한 모의해킹을 실제 ICT 인프라에 직접적으로 수행할 수 없는 **국가 주요기반 시설의 취약한 공격 경로 분석을 위한 시뮬레이션 도구**로 활용
- Real-world 공격을 자동 시뮬레이션하여 보안제품의 Security Control 효과성을 테스트하고 검증하여 제품 경쟁력 강화하고 기업의 보안수준 향상하는 측정도구

○ **(경제성)** AI 보편화에 따른 인공지능 활용 기술로 가트너는 BAS 기술이 2017년 가트너 하이프 사이클에 첫 등장하여 향후 **연평균 37.8% 고성장**할 것으로 예측되는 신기술 분야

* 2018년 93.9백만불 → 2027년 16.8억불(CAGR 37.8% 고성장, 마켓앤마켓)

- BAS(Breach and Attack Simulation) : Real 공격(TTP 시나리오) 시뮬레이션 - 방어기술의 효과성을 평가하는 과정을 자동화하여 침투테스트 및 모의 해킹(일시성, 고비용 인력) 이슈를 극복하기 위한 위험관리 기술
- AI기반 모의해킹을 통한 주요 국가기반시설(전력, 철도 등) 보안성 강화
- Rapid7, AttackIQ 등 글로벌 기업들이 BAS 기술 확보를 통해 시장 지배력을 높이고 있으나, 국내에는 관련 기술 개발이 전무하여 국산기술개발 지원이 시급

3. 연구목표

○ **최종목표 : 주요 시설에 대한 보안정책 수립/적용을 위해 가상 ICT 인프라를 대상으로 AI 기술을 활용한 위협헌팅 모델 기반 사이버 보안위협 공격·방어 시뮬레이션 프레임워크 개발**

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	AI 생성 공격 시나리오 실행 성공율	%	70%	-	-
2	위협헌팅 모델 기반 공격 테크닉 지원 개수*	%	≧ 80%	-	60 % (Automic Red Team/미국)**
3	공격대상 OS 개수	개	≧ 4개	-	3개 (Caldera/미국)

* Mitre Att@ck의 techiques 지원 비율 : (본 과제 분석 Tech. 수)/(Mitre Attack Tech.수)

** Mitre Att@ck Tech. 184개 존재(Automic Red Team에서 제공하는 Tech. 111개(60%) 수준임)

○ **연차별 개발목표**

구분	연도별 연구목표
2021년	지능형 사이버 공격/방어 분석 테스트베드 생성 위협헌팅 모델 기반 사이버 공격 시뮬레이션 프레임워크 및 요소기술 설계 국내외 IPR 확보
2022년	사이버 공격 시뮬레이션 요소기술(수집기술, 가상화 기술 등) 개발 인공지능 기반 공격 시나리오 자동 생성 기술 설계
2023년	위협헌팅 모델 기반 사이버 공격 시뮬레이션 시제품 개발 인공지능 기반 공격 시나리오 자동 생성 기술 개발
2024년	통합형 인공지능 기반 사이버공격/방어 플랫폼 시제품 개발 주요 기반시설(에너지, 교통 등) 대상 사이버공격/방어 플랫폼 실증

4. 연구내용	

ㅇ **개발 기술 내용**

① **(가상공간 시뮬레이션)** 보호대상 ICT 인프라 가상공간 생성 및 가시화기술
- 보호대상 ICT 인프라 Discovery 및 계층적 구조(Topology) 표현 기술 개발
- 지능형 OS/서비스 핑거프린트(Fingerprint) 분석 기술 개발
- 대상 도메인 가상환경 기반 가상공간(VM, 도커 등) 생성/가시화 기술 개발

② **(공격 시나리오 자동 생성)** 위협 헌팅 모델 기반 사이버공격 시나리오 자동생성 기술
- AI 기반 사이버 공격 경로 생성 기술 개발
- 대상 도메인 맞춤형 공격 시나리오 자동 생성 기술 개발
- 위협 헌팅 모델(사이버 킬체인, MITRE ATT@CK 등) 기반 공격도구 관리 플랫폼 개발
- 위협 헌팅 모델 기반 취약점 유형별 공격 도구 자동실행 기술 개발

③ **(AI기반 공격 시뮬레이션 플랫폼)** Security Automation Response 통합형 인공지능 기반 사이버공격/방어 플랫폼 기술
- 위협헌팅 모델 기반 사이버 공격 시뮬레이션 시제품 개발
- AI 기반 모의해킹·방어 시뮬레이션 통합 프레임워크 개발
- 주요 기반시설(에너지, 교통, 제조 등) 대상 사이버공격/방어 플랫폼 실증

③ **(사이버공격 대응전략 수립)** 사이버공격 시뮬레이션을 통한 보안성 평가 및 방어전략 수립 기술
- 사이버 공격 시뮬레이션을 통한 감내(보안) 수준 평가기술 개발
- 보안 위험(RISK)기반 사이버 방어전략 생성 기술 개발

ㅇ **기존 (보유)기술**

① 공격그래프(Attack Graph) 기반 보안성 평가 기술
- ICT 자산정보 및 취약점(CVE) 정보에 기반한 통계적 사이버공격 경로 생성 기술
- 생성된 공격경로에 대한 실제 실행여부 확인 불가능하며, 모든 가능한 공격경로 조합에 대한 단순한 우선순위를 제시하는 수준

5. 지원기간/예산/추진체계	

ㅇ **기간** : 4년 이내

ㅇ **정부출연금** : ‘21년 12억원 이내 (총 정부출연금 57억원 이내)

ㅇ **주관기관** : 제한없음

기술분류	대분류(차세대보안) - 중분류(네트워크·응용서비스 보안) - 소분류(위협분석 및 관제)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (4) ~ (6)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-11	(지정공모형)
과제명	AI·빅데이터 기반 사이버 보안 오케스트레이션 및 자동 대응 기술	

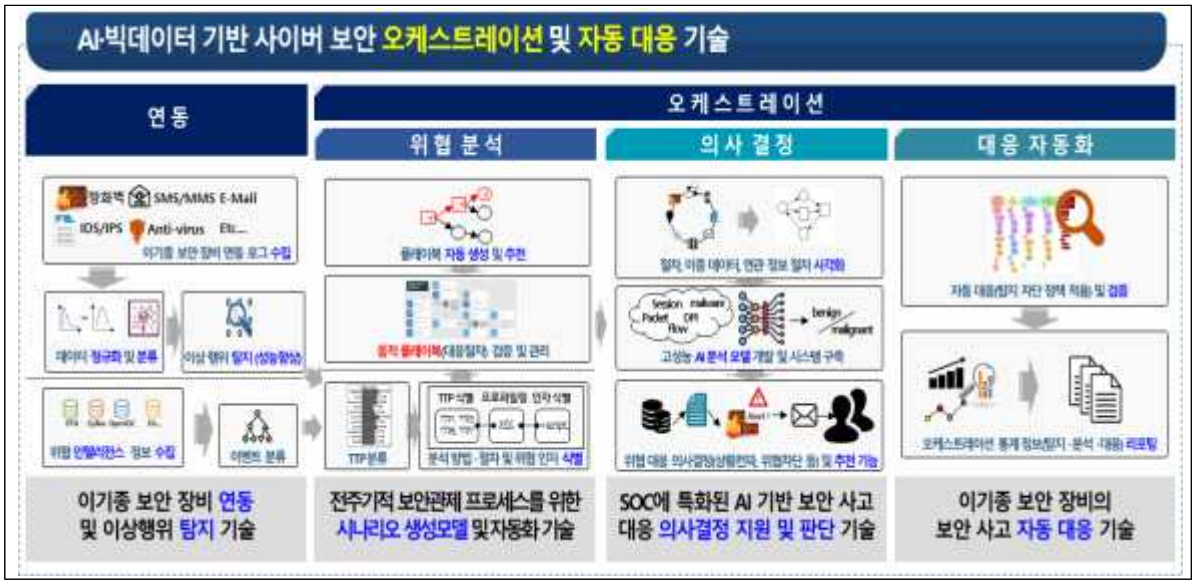
1. 개념

○ 사이버공격의 선제적 대응을 위해 **보안관제센터(SOC)** 全업무 영역에서 **공격·이상 징후의 실시간 탐지, 분석, 대응**을 아우르는 **프로세스의 자동화(오케스트레이션) 및 탐지성능 향상** 기술

- (문제점·현황) ①국내 보안장비의 상이한 이벤트 로그·제어 정책 등으로 인해 시스템·프로세스 간 연동 어려움, ②보안관제 담당자의 능력 수준에 따라 상이한 분석 정확도 및 대응 절차, ③탐지영역 이외의 분석·대응 업무의 자동화 기능 부재

※ IBM의 조사에 따르면 사이버보안 대응이 필요한 조직중 절반 이상이 역량을 갖추지 못했으며, 보안 전문가의 고용·유지 및 증가하는 공격 빈도, 복잡한 SOC 운영 환경을 사고대응의 어려움으로 선정함(2019기업 사이버공격 대응 실태, IBM)

< 개념도 >



2. 필요성

○ **(정부 지원 필요성) 비대면·디지털전환 환경**의 보편화 및 **신규 정보통신시설(스마트시티, 자율주행차 등)에 대한 보안관제 확대** 등에 따라 **국가 사이버 위협 대응력 강화**를 위한 사이버 보안관제 자동화(SOAR) 원천기술의 확보 필요(K-사이버방역, 과기정통부)

- SOAR* 기술은 기존의 보안 솔루션(SIEM, EDR 등)의 예방·탐지 기능을 포함한 보안 운영 자동화 솔루션으로 국내 보안업체의 규모에서는 기술력, 자금력, 실증환경 등을 모두 확보하기에 어려운 분야임(기관 주도의 기술 개발 필요)

* SOAR(Security Orchestration, Automation and Response) : IT 보안 분야의 숙련된 기술자의 부족으로 신속한 침해 사고 발생 시 즉각적으로 대응할 수 있는 자동화 기술이 대안으로 제시됨(Gartner, 2017)

○ **(기술성)** 전 세계적으로 AI 기술을 이용한 보안관제 자동화 제품이 출시되기 시작했으나 아직 연구 초기단계로, **다양한 국내외 보안장비의 연동, 복잡한 공격에 대한 동적 대응, 탐지 성능향상** 등의 기술 개발 및 선점 필요

- 현재 SIEM*은 로그수집, 탐지 등의 기능을 갖추었으나 대응 기능이 없으며, 최근 출시되는 해외 SOAR 솔루션의 경우 API를 통합한 분석·대응 정도만을 지원 하는 실정

  * SIEM(Security Information and Event Management) : 이기종 보안 장비의 이벤트 수집, 탐지 솔루션

  ※ 국산 보안 제품은 API 미지원 및 표준로그 미사용으로 해외 SOAR을 도입 하더라도 실효성이 없음

- 현재 개발되고 있는 보안관제 기술은 전문가의 시나리오 룰(Rule), 고정된 업무 절차를 따르기 때문에 지능화된 변형 공격 등에 유연함이 부족하므로 AI·빅데이터를 활용해 보다 효과적인 대응방안 및 우선순위 제시 등 신속한 의사결정 지원 요구

○ **(경제성)** 글로벌 SOAR 시장은 2025년까지 연평균 16.3% 상승 및 2.3B$ 시장으로 고성장 예측(2019 가트너 시장보고서)

- 향후 AI 기반 보안관제 기술은 정보보안 핵심 요소 기술로서, 해외 보안관제 기술과의 격차 해소를 위해 반드시 확보해야 할 원천기술임

  * SOAR 기술력 미확보 시 국내 보안관제 기술 및 제품은 외산에 종속될 가능성 높음

3. 연구목표

○ **최종목표 : 지능화된 사이버공격의 대응력 강화를 위한 보안 관제**(탐지기술 연동 - 위협분석 - 의사결정 - 대응) **자동화 오케스트레이션 기술 개발**

- (연동·탐지)이기종 보안 장비 연동 및 이상행위 탐지 기술

  ※ 국내 이기종 보안 장비의 로그 수집 및 대응 정책 적용을 위한 공통 포멧 개발/정책적 활성화 유도

- (위협분석·자동화)빅데이터 기반 전주기적 사이버 보안관제 업무 프로세스를 위한 시나리오 생성모델 및 자동화 기술
- (의사결정)보안관제센터(SOC : Security Operations Center)에 특화된 AI 기반 보안사고 대응 의사결정 지원 및 판단 기술
- (대응)이기종 보안 장비의 보안사고 자동 대응 기술

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	AI 기반 이상행위 탐지 정확도	AUC	0.95	-	0.93[1] (중국/CAS)
2	침해사고 대응 절차 자동화 수행 시간[2]	분	15	-	20[3] (미국/Fortinet)
3	보안 위협 대응 시나리오 (플레이북)	개	3,600	-	3,000[4] (미국/Fortinet)

1) Liu, Fucheng, et al. "Log2vec: A Heterogeneous Graph Embedding Based Approach for Detecting Cyber Threats within Enterprise." Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security. 2019.

2) 목표 기술이 침해사고 대응 절차의 자동화 수행 시간을 측정. 보안운영센터(SOC)팀(메뉴얼)의 침해사고 대응 시간은 평균 4.5~15시간이며, 미국 Fortinet 제품(자동화)의 경우 20분 소요됨.

※ 침해사고 대응절차 : IOC 식별을 위한 아티팩트 추가 수집, SIEM 이벤트 심사, 장치 격리, 사건 생성 및 분석, IOC 조치, 사고 대응, 요약보고서 작성 등

3) Fortinet, FortiSOAR Empowers Security Operations to Accelerate Incident Response, 2020.

4) Fortinet, FortiSOAR Datasheet, 2020

4. 연구내용

○ **개발 기술 내용**

① **(연동·탐지)** 이기종 보안 장비 연동 및 이상행위 탐지 기술

- 국내 이기종 보안 장비 로그 데이터 및 대응 정책 적용 연동 포맷 및 API 개발
- 빅데이터 기반의 이기종 보안 장비 데이터 정규화·분류 및 이상행위 탐지(연관 분석) 기술
- AI·빅데이터 기반 이기종 보안 장비 이상행위 탐지 성능 향상(정탐/미탐/오탐 분류) 기술
- 국내외 공개 보안 위협 인텔리전스 정보 수집

② **(위협분석/자동화)** 빅데이터 기반 전주기적 사이버 보안관제 업무 프로세스(탐지→분석→대응)를 위한 시나리오 생성모델 및 자동화 기술

- "MITRE ATT&CK" 기반 위협정보의 TTP(Tactics Techniques and Procedures) 분류 기술
- TTP 정보 기반 이상행위 탐지 자산의 분석 방법·절차 및 위협 인자 식별 기술
- 공격 유형 및 탐지 규칙 별 보안 이벤트 분석·대응을 위한 플레이북 모델 정의 및 자동 분석 실행 기술
- 보안관제 빅데이터(탐지 및 대응이력 등) 기반의 동적 플레이북(대응 절차) 검증·관리 기술

③ **(의사결정)** SOC에 특화된 AI 기반 보안사고 대응 의사결정 지원 및 판단 기술

- 플레이북을 통한 자동화 분석 절차 관리 및 시각화 기술
- 이종 보안 장비 데이터 및 TTP 정보의 연관정보 가시화 기능
- 신속한 의사결정 지원을 위한 고성능 AI 분석 모델 개발 및 시스템 구축
  - ※ 네트워크 특성이 보존된 보안관제 전용 데이터 전처리기술 개발
  - ※ 다종/다량의 보안이벤트를 포함한 보안관제 전용 대용량 학습데이터 구축
  - ※ 공격유형별 고유 특성을 반영한 고정밀 AI 분석·탐지모델 개발(APT 공격 및 이벤트 탐지 등)
- AI(강화학습)기반 공격 대응 최적 의사결정(대응 방안 및 우선순위 등) 자동 추천 기술

④ **(대응)** 이기종 보안 장비의 보안사고 자동 대응 기술

- 이기종 보안 장비의 보안사고 자동 대응(탐지·차단 정책 적용) 및 검증 기술
- 사이버 보안관제 업무 프로세스 전반에 대한 정량적 측정 기술
  - ※ 침해대응 신속도, 침해대응 정확도, 관제요원 1인당 분석대량 정보량 등
- SOC 운영에 필요한 제반 통계·보고서 자동생성 기술

○ **기존 (보유)기술**

① Security Analytics 기반의 이기종 보안솔루션 위협 분석 및 대응 기술

- 글로벌 사이버 위협정보 수집·분석 및 공격 대응방안 공유
- 이기종 보안 장비의 이벤트의 표준 형식으로 수집·저장 및 Security Analytics 기반 분석
- 위협정보를 통한 단말의 위협 대응을 위한 EDR(Endpoint Detection and Response) 시스템 개발

② SIEM을 위한 인공신경망 기반 네트워크 침해 위협 분석 및 탐지 기술

- 네트워크 보안 이벤트와 로그를 딥러닝 기술을 이용하여 학습한 모델을 기반으로 정탐 로그 분류

5. 지원기간/예산/추진체계

○ **기간** : 4년 이내 (1단계 2년 → 2단계 2년)

○ **정부출연금** : '21년 19억원 이내 (총 정부출연금 94억원 이내)

○ **주관기관** : 제한없음 (산업체 참여필수)

기술분류	대분류(차세대보안) - 중분류(시스템 및 암호보안) - 소분류(위협 분석 및 관제)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (4) ~ (6)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-12	(지정공모형)
과제명	언택트 시대의 기업망 보호를 위한 제로트러스트 기반 접근제어 및 이상징후 분석기술 개발	

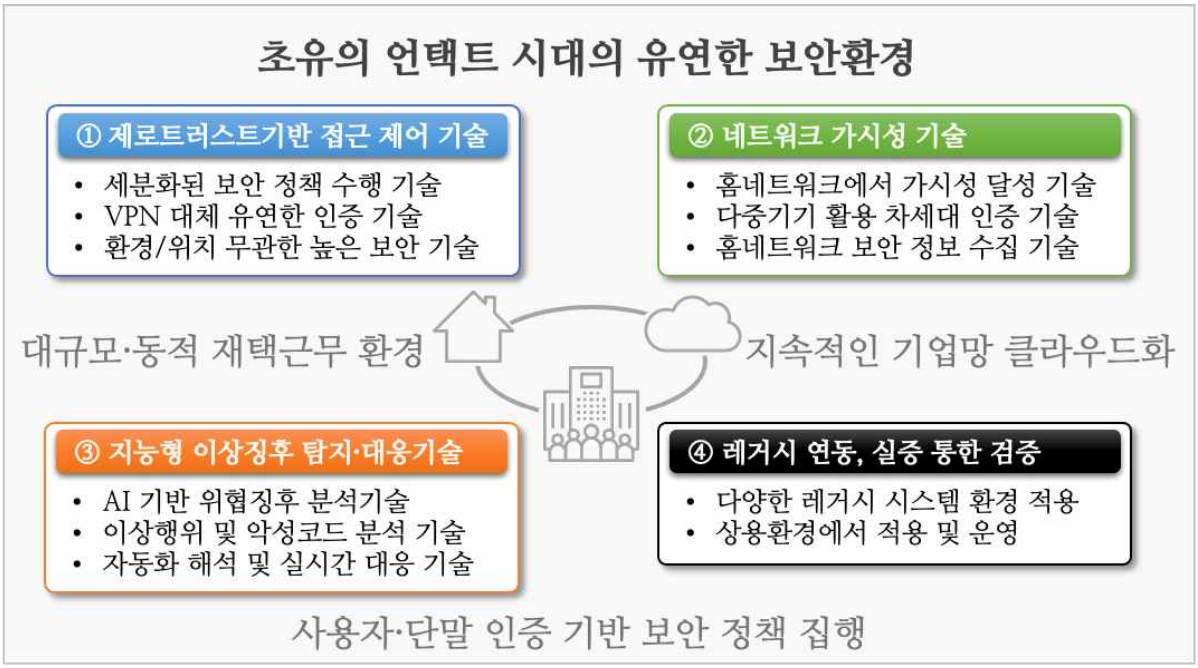
1. 개념

○ **기업망/홈네트워크/클라우드가 복잡하게 혼재된 새로운 기업망 구조를 위한 보안 모델링*, 언택트 시대에 적합한 제로트러스트 기반 접근제어** 및 이상징후 분석기술 개발**

- (문제점·현황) ①현재의 망분리 정책·기술은 지능형APT 공격 대응에 부적합하고, 업무 생산성/창의성에 큰 걸림돌이 됨, ②제로트러스트 기반의 기업망 기술이 대안으로 떠오르고 있지만, 이 또한 초유의 언택트 시대의 복합적인 기업망 환경을 고려하지는 않은 솔루션임, ③현재 운영중인 다양한 기업망 보안 솔루션을 확장·통합하는 선도적 기업망 기술 부재

* 서로 다른 기업망/홈네트워크/클라우드의 업무용기기/개인기기/기업망서버/클라우드 서버가 복잡하게 연결되어 기업의 디지털 서비스를 이룰 때 이들의 다양한 보안 정책을 중앙에서 관리·적용 가능케 하는 보안 모델링 기술

** 대규모 재택근무 상황이 불규칙적으로 발생할 시에도 높은 보안 항상성을 유지할 수 있는 유연한 기업망 망분리 완화 보안 기술



< 제로트러스트 기반 접근제어 및 이상징후 분석기술 >

2 필요성

○ **(정부 지원 필요성) 초유의 언택트 시대, 정부 지원 통한 기업망 보안 증진 필요**

- 기업의 탈망분리는 전세계적으로 판데믹 이전에 이미 시작되고 있었음 (예: 구글의 BeyondCorp, 넷플릭스의 LISA). 하지만 우리나라는 아직도 망분리에 대해서는 낮은 차원의 논의(예: 물리적 vs. 논리적 망분리)에 머물러 있는 상황으로 장기적인 연구 개발 계획이 필요한 시점임.
- 2020년 초유의 판데믹 발생 이후 정부기관 및 금융기관 등을 타겟으로 하는 사이버 범죄는 전세계적으로 급증*. 전문가들은 앞으로 점점 더 자주 판데믹 상황

이 벌어질 것이라 예상하는데, 이에 대비하여 정부 지원을 통해 공기업/사기업**의 광범위한 보안 증진 계획 필요.

* 대미 중국의 사이버공격 판데믹 기간 급증 (2020년 7월 더디플로맷)

** 미국 코로나 백신 기업 상대로 중국발 사이버공격 발생 (2020년 5월 뉴욕타임즈)

○ **(기술성) 현재 망분리의 보안 한계를 뛰어넘고 언택트 시대의 새로운 보안 환경에 미리 대응**

- 현재의 망분리 기술은 진화하는 APT공격에 대해 보안적으로도 취약*하고, 업무 생산성 및 창의성을 크게 훼손**시킴

  * 현재의 VPN 기반의 망분리는 임직원의 VPN 비밀번호 하나만 탈취되면 공격자들이 손쉽게 기업망 내부로 침투해서 다양한 기업망 내부 자원 접근 가능

  ** 금융권 망분리 정책으로 핀테크 개발자들이 심각한 생산성 저하로 고통받는 중 ("일할 땐 '인터넷 먹통' PC만 써라? 핀테크 '망분리하다 망할 판'" 한국경제 2020년 7월)

- 언택트 시대에 예측불가능한 기업망 환경변화*로 인해 필연적으로 공격표면 급증. 특히 대규모 재택근무 환경은 기업망 기기들이 홈네트워크의 저신뢰성 개인 PC 및 IoT 기기들과의 접점을 증폭시킴.

  * 방역지침으로 인해 기업의 일부 혹은 전체 인원이 예측불가능한 방식으로 재택근무 ("KT 광화문 본사서도 확진자 발생...재택근무 13일까지 연장" 2020년 9월 동아일보)

○ **(경제성) 제로트러스트 기반 망분리 완화 기술의 직·간접적인 큰 경제효과**

- 제로트러스트 글로벌 시장은 한동안 지속적으로 상장할 것으로 예상

  * 2025년까지 미화 380억 달러의 시장이 형성될 것으로 예측 (ADROIT 2020년 10월)

- 재택근무는 지속적으로 증가하여 주요 업무방식으로 정착할 것으로 예상

  * 페이스북등의 글로벌 기업들 2021년 이후로 재택근무 연장 (2020년 8월 조선일보)

  * 한국 100대 기업 대부분 재택근무 허용중이며 향후 지속 예정 (2020년 9월 중앙일보)

- 점진적 탈망분리로 인해 기업의 생산성이 급증하고, 신규 사업자의 시장진입 장벽을 낮추어 기업 생태계 정상화에 일조할 수 있음

3. 연구목표

○ **최종목표: 기업망/홈네트워크/클라우드가 복잡하게 혼재된 새로운 기업망 구조를 위한 보안 모델링, 언택트 시대에 적합한 제로트러스트 기반 접근제어* 및 이상징후 분석기술 개발**

- 유연한 고신뢰성 제로트러스트 기반 접근제어 기술
- 다중 홈네트워크 혼재 환경에서 네트워크 가시성 기술
- 대규모 재택근무 환경에서의 지능형 이상징후 탐지 및 대응기술
- 레거시 시스템과의 연동방안 마련 및 실증을 통한 기술 검증

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	보안정책을 위한 컨텍스트 개수	개	사용자/단말의 위험도, 상태, 인증레벨 등 4개 이상의 컨텍스트 구축	미달성 (현재의 SDP, CASB 기업망	SASE[주1)] 기술 통해 달성 노력 중 (미국,Cato

				접속기술로는 역부족)	Networks[주2], Palo Alto Networks[주3] 등)
2	혼재 상황 기업망 가시성 최대화 기술	%	혼재 상황 기업망에서 가능한 모든 APT공격 경로 99% 이상의 가시성 달성	N/A	가능한 모든 APT공격 경로 92% 가시성 (미국, CMU대/RedJack)[주4]
3	AI 기반 위협징후 탐지 시간	초	이상 공격징후 탐지시간 6초 이내* * 인가된 사용자의 침해공격 발생부터 탐지까지의 시간	N/A	9초 (미국, Darktrace)[주5]
4	개발기술 실증	개	망구성이 상이한 2개 이상 사이트에서의 실증	N/A	-

주1) Gartner는 2019년 리포트에서 SASE (Secure Access Service Edge) 기술이 향후 5년 안에 40%의 기업망에 직·간접적으로 사용될 것이라 예측. SASE는 기존의 유망한 여러 기술(SWG, CASB, FWaaS, SD-WAN 등)을 통합하는 네트워크 보안 기술이 될 것으로 예상.

주2) Cato Networks는 Gartner가 주창한 SASE의 모델에 가장 근접한 솔루션 현재 제공한다 평가됨

주3) Palo Alto Networks는 기존의 차세대 Firewall 기술 기반으로 SASE 기술 개발 중

주4) CMU대, RedJack 공동 연구 논문. T. Yu, et al., “PSI: Precise Security Instrumentation for Enterprise Networks” NDSS 2017.

주5) “Autonomous Response : Threat Report” Darktrace 2019.

○ 연차별 개발목표

구분	연도별 연구목표
2021년	o 제로트러스트 기반 기업망 보안 기술 배경 조사 o 기업망/홈네트워크/클라우드 혼재망 환경에서의 보안위협 연구 및 모델링
2022년	o 혼재망 환경에서 망분리 완화 구조 연구 및 시스템 구축 기술 연구 개발 o 혼재망 환경에서의 네트워크 가시화 기술 및 다중기기 인증 연구 개발
2023년	o 신뢰된 사용자 환경에서의 이상행위 탐지 및 악성코드 분석 기술 개발 o 상용환경에서의 개발기술 적용 및 운영을 통한 실효성 확보
2024년	o 비정상 접속징후에 대한 자동화된 해석 및 실시간 대응 기술 개발 o 다양한 레거시 시스템 환경에서의 유연한 적용 모델 개발

4. 연구내용

○ 개발 기술 내용

① 유연한 고신뢰성 제로트러스트 기반 접근제어 기술
- 기업망/홈네트워크/클라우드 혼재망 환경에서 세분화된 보안 정책 수행 기술
- 혼재망 환경에서 VPN기반의 단일 인증 시스템을 대체하는 유연한 인증 기술
- 다양한 클라이언트 환경, 서비스 위치에 무관한 높은 보안 항상성 유지 기술

② 다중 홈네트워크 혼재 환경에서 네트워크 가시성 기술
- 저신뢰성 홈네트워크 환경에서 기업망 보안을 위한 높은 가시성 달성 기술
- 재택근무시 업무용/개인용 다중기기 활용 인증 복잡도 저감 기술
- 기업망 보안시스템에서 안전하게 홈네트워크 보안 정보 수집 시스템 기술

③ 대규모 재택근무 환경에서의 지능형 이상징후 탐지 및 대응기술
- 지능형 침입대응을 위한 AI 기반 위협징후 분석기술
- 신뢰된 사용자 환경에서의 이상행위 및 악성코드 분석 기술
- 비정상 접속징후에 대한 자동화된 해석 및 실시간 대응 기술
④ 레거시 시스템과의 연동방안 마련 및 실증을 통한 기술 검증
- 상용환경에서의 개발기술 적용 및 운영을 통한 실효성 확보
- 다양한 레거시 시스템 환경에서의 유연한 적용 모델 개발

ㅇ **기존 (보유)기술**

① 레거시 VPN 대체 기술
- 지능형 VPN 기술
- Software-defined perimeter (소프트웨어 정의 경계) 기술
② 소프트웨어 정의 네트워크 (SDN) 기반 네트워크 보안 기술
- SDN 기반 네트워크 보안 함수 개발 및 DDoS 공격 대응 기술 (KAIST)
- 저신뢰성 기기들 사이의 신뢰 구축을 위한 시큐어 와이파이 기술 (삼성)
③ AI기반 호스트/네트워크 위협징후 분석기술
- 네트워크 상의 모든 시스템 행동패턴을 스스로 학습하여 이상행위 탐지(Darktrace)
- 다양한 머신러닝 모델을 통해 신규 악성코드 및 변종탐지(Cylance)

5. 지원기간/예산/추진체계

ㅇ **기간** : 4년 이내 (1단계 2년 → 2단계 2년)

ㅇ **정부출연금** : ‘21년 15억원 이내 (총 정부출연금 75억원 이내)

ㅇ **주관기관** : 제한없음 (산업체 참여필수)

기술분류	대분류(차세대보안) - 중분류(네트워크보안) - 소분류(유선네트워크보안)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (4) ~ (6)
과제특징	정책지정( ), 혁신도약형(O), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-13	(지정공모형)

과제명	상시적 보안품질 보장을 위한 6G 자율보안 내재화 기반기술 연구

1. 개념

○ 6G에서는 초성능, 초지능화된 통신인프라 기반의 대규모·지능화 공격이 일상화되고 新융합 서비스 대상의 보안위협 급증이 예상되므로, 이에 대한 보안분석을 기반으로 **6G 인프라 전반에 보안이 내재화될 수 있도록 하는 기술 정의 및 기반기술 확보**

- 안전한 6G 新융합 서비스 제공을 위한 AI 기반 6G 자율보안 내재화 기술
- 3차원 공간 이동통신의 가용성 보장을 위한 6G Flying 기지국 보안기술
- 6G 암호체계 내재화 기반의 최상급 기밀성 제공을 위한 6G 양자 보안기술

<상시적 보안품질 보장을 위한 6G 보안 내재화 기술 분야>

2 필요성

○ **(정부 지원 필요성)** 6G 미래 환경변화(AI보편화, 미션 크리티컬 서비스 확대, 양자컴퓨팅 시대 도래)에 따라 신뢰할 수 있는 인프라 제공을 위해 **6G 보안내재 아키텍처 수립 필요**

- 6G 新융합 서비스(원격수술, 완전자율주행차 등) 및 관련 산업이 안전하게 발전할 수 있도록 **6G 보안위협 대응 필요**
- **6G 이동통신 기술**의 초기 선점을 위한 선도형 기술개발 전략을 **'20년부터 추진 중**

* 과기정통부, "6G 시대를 선도하기 위한 「미래 이동통신 R&D 추진전략」(안)", 2020.8.6.

○ **(기술성)** 5G 보안기술 개발과 더불어, 최근 전 세계적으로 6G 기술개발이 시작된 점을 고려하여, **Security by Design 개념이 6G 구조에 조기 정착**하기 위한 **6G 보안 내재화 기술 필요**

- 신규 암호체계 및 QKD 요소기술에 대한 6G 구조의 적용 가능성 및 안전성 연구를 위하여 **6G는 설계단계부터 필수적 보안품질을 보장하는 내재화(Embedded Security) 아키텍처 개발** 필요

* 5G는 초기 설계단계에서 보안이 충분히 고려되지 않아, 보안 기능의 지속적 추가(Add-On Security)에 따른 성능 저하 발생

○ **(경제성)** 6G 분야 핵심 원천기술 확보 및 표준화 선도로 글로벌 시장 선점을 위해 각 국가는 경쟁적으로 **정부 주도의 선제적 R&D 투자 중**

- 정부는 2021년부터 5년간 6G 핵심기술개발 6대 중점분야(초성능, 초대역, 초정밀, 초공간, 초지능, 초신뢰)의 기술 확보에 2천억원 투자
- 핀란드 오울루 대학 주도 **'18년부터 6G 플래그십*** 설립
  * 오울루·알토대학, 핀란드 기술연구센터, 기업체(노키아, 인터디지털 등) 간 6G 플래그쉽 협업체를 구성, 내재화된 보안기술 기반의 6G 연구개발 착수(8년간 약 3,000억원 규모)

3. 연구목표

○ **최종목표 : 사이버위협 걱정없는 6G 新융합 서비스의 초신뢰 인프라 제공을 위한 상시적 보안 품질 보장을 바탕으로 6G 자율보안 내재화 아키텍처 및 프레임워크 개발**

- 6G 망에 대한 보안위협 대응을 위한 보안 내재화 프레임워크
- 6G 전역망 위협 대응을 위한 AI 기반 6G 자율보안 내재화 기술
- 6G 초공간 가용성 보장을 위한 Flying 기지국 보안기술
- 6G 망에서의 최상급 기밀성 제공을 위한 양자 보안기술

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	6G 보안 내재화 적용 시나리오 수	건	4	-	-
2	6G Flying 기지국 위협 탐지 범위[1)]	탐지 특성 수	디지털/전파/물리적	-	-
3	양자보안 6G 적용 시나리오 수	건	4	-	-

1) Flying 기지국 위협(사이버-물리 탈취)는 악성코드/백도어 등을 이용한 디지털 특성 기반의 제어권 탈취와 재밍/스푸핑 신호를 이용한 전파적 특성 기반의 오동작 유발, 그리고 그물/포획드론/독수리 등을 이용한 물리적 포획 등이 있음

○ **연차별 개발목표**

구분	연도별 연구목표
2021년	6G 보안위협의 위험도 분석 및 보안 내재화 세부 기능 정의
2022년	6G 보안 내재화 아키텍처 및 프레임워크 요소기술 분석 6G 보안 위협지표 및 보안 내재화 국제표준 아이템 발굴
2023년	AI 기반 6G 자율 보안관제 모델 및 프레임워크 요소기술 분석 6G 초공간 가용성 보장을 위한 Flying 기지국 보안 요소기술 분석 6G 자율 보안관제 모델 및 보안 내재화 요구사항 국제표준 개발
2024년	6G 암호체계 내재화 기반의 양자 보안 요소기술 및 적용 시나리오 개발 6G 자율보안 내재화 가이드라인 국제표준 개발

4. 연구내용

○ **개발 기술 내용**

① 6G 망에 대한 보안위협 대응을 위한 보안 내재화 프레임워크

- 6G 기반 보안위협(광대역 DDoS, 백도어 공격 등) 위험도 분석

- 6G 망의 보안위협 모델링 분석 및 보안 내재화 세부 기능 정의
- 6G 보안 내재화 아키텍처 및 프레임워크 요소기술 분석
- 6G 보안 위협 지표 및 보안 내재화 요구사항 국제표준 개발

② 6G 전역망 위협 대응을 위한 AI 기반 6G 자율보안 내재화 기술
- 6G 전역망 보안지능 내재화의 요소기술 분석 및 세부 기능 정의
- 6G 초지능·초신뢰 환경의 자율 보안관제 모델 및 프레임워크 분석
- 6G 위협대응을 위한 AI 기반 자율보안 구성 및 제어 요소기술 분석
- 6G 자율 보안관제 모델 및 보안 내재화 가이드라인 국제표준 개발

③ 6G 초공간 가용성 보장을 위한 Flying 기지국 보안기술
- 6G Flying 기지국 보안 요구사항 분석 및 보안 운용 시나리오 개발
- 6G Flying 기지국 환경에서의 경량/저전력/무손실 보안 라우팅 구조 연구
- 6G Flying 기지국 보안위협(사이버-물리 탈취공격 등) 대응 방안 연구

④ 6G 망에서의 최상급 기밀성 제공을 위한 양자 보안기술
- 6G 암호체계 내재화 기반의 양자 보안(PQC, QKD) 요소기술 분석
- 6G CIA(기밀성, 무결성, 가용성) 보장을 위한 양자 보안 기능/성분 분석
- 양자 보안 요소기술의 6G 망의 적용 시나리오 개발

ㅇ **기존 (보유)기술**

① 5G+ 기반 6G 정보보안 기술 연구 ('20.04~'27.12)
- 지능형 무선 접속망 및 핵심망용 정보보안과 암호기술 분석
- 초저지연 데이터 서비스 보안을 위한 정보보안 및 암호기술 분석
- 무인비행체 A2X 서비스 및 저궤도 위성통신용 정보보안과 암호기술 분석

5. 지원기간/예산/추진체계

ㅇ **기간** : 4년 이내

ㅇ **정부출연금** : '21년 15억원 이내 (총 정부출연금 75억원 이내)

ㅇ **주관기관** : 제한없음

기술분류	대분류(차세대보안) - 중분류(네트워크·응용서비스 보안) - 소분류(무선네트워크 보안)	
연구유형	기초연구 (O), 응용연구 ( ), 개발연구 ( )	**TRL** (2) ~ (3)
과제특징	정책지정( ), 혁신도약형(O), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-14	(지정공모형)
과제명	비대면 환경의 보안 편의성 개선을 위한 Usable Security 기술 개발	

1. 개념

○ 비대면 및 IoT 서비스 환경에서 빈번히 사용되는 보안 서비스의 편의성 향상을 위해서, IT 전문가 뿐만 아니라 일반인, 학생 등 일반 대중도 쉽게 보안 기능을 이용하고 관리할 수 있는 사용자 친화적 보안(Usable Security) 기술 개발

- 세계적인 보안 연구의 흐름은 시스템의 보안성 유지 및 프라이버시를 보장하면서 사용자 편의성을 개선하는 방향으로 진행되고 있으며 보안 솔루션의 요구사항으로 사용자의 이용 편의성이 중요한 항목으로 자리 잡음
- 종래의 사용자 인증 및 보안 관리 기술을 비대면 환경에 동일하게 적용할 경우, 사용자 인터페이스의 제약 및 이용 장치의 다양성으로 인하여 많은 한계점을 노출하고 있어, 이러한 한계 극복을 위한 비대면 환경의 사용자 친화적인 보안 연구가 필요함

< 개념도 >

2 필요성

○ **(정부 지원 필요성)** 비대면 환경의 정보기술 기반 서비스가 확대됨에 따라 일반 사용자들이 익숙하지 않은 서비스 이용 및 관리로 인한 혼란과 해킹 피해를 최소화할 수 있도록 편리하게 보안 기능을 이용할 수 있는 기술 개발 필요성 증가

- 노인층, 어린이 등 정보 약자가 ICT 서비스를 사용해야만 하는 상황 확대
- 현재 보안 기능은 이용 복잡성으로 인해 일반 사용자에게 많은 불편을 초래
- 또한, 보안 기능 향상을 위한 무분별한 사용자 (생체 및 행위) 데이터 수집으로 인한 프라이버시 침해 가능성 증가
- 비대면 환경의 보안 기능 이용 불편으로 인한 해킹, 프라이버시 침해 등 부작용*을 해소할 수 있는 연구가 부족하여, 정부 지원을 통한 대응 기술의 개발 필요

* 사용자의 보안 관리 미숙으로 인해, IP카메라, 스마트TV, 공유기 등 해킹 피해 증가

○ **(기술성)** 국내 보안기술 개발은 사용성보다는 보안성을 높이는 방향으로 추진되어, 사용자 친화적인 보안 연구 및 솔루션 설계 기술 확보는 해외에 비해 미흡함

- 대부분 보안 솔루션의 UX는 개발자와 연구자 중심으로 개발되어, 일반 사용자가 이용 및 접근하기 어려운 현실
- 특히, 모바일 및 IoT 기기의 경우에는 제한된 UI로 인하여 종래의 PC 및 서버 기반의 보안 솔루션을 직접적으로 적용하기 어려움
- 최근 AI 기반의 보안 솔루션은 단순히 정확도(accuracy) 향상 측면에서 접근하고

있으며, 오류에 대해서 사용자가 어떻게 대처할 수 있는지에 대한 해법을 제시하지 못하고 있음. 실제 적용 가능한 보안 솔루션 개발을 위해서는 “human in the loop”(휴먼 요소 고려) 개념으로 사용자를 대상으로 설계되어야 함

- 해외에서는 USENIX, SOUPS 등 주요 학술대회에서 지속적으로 연구 결과를 발표하고 있으며, CMU, NYIT 등 학계와 구글, 마이크로소프트 등 글로벌 기업은 사용자 친화적인 사용자인증*, 보안 설정, 보안 경고, 피싱방지** 등 연구를 진행 중

  * 패스워스 생성 규칙 및 패스워드 변경 정책 등이 실제 보안성을 향상시키는데 도움이 되지 않는다는 연구 결과에 따라서 NIST의 패스워드 가이드라인이 업데이트 됨

  ** 대부분 웹브라우저에서 제공되는 보안 경고 기능(URL 바의 Extended Validation Indicator)이 일반인에게 효과가 미흡하다는 연구 결과가 발표되어 브라우져에 반영됨

○ **(경제성)** 시장이 급격히 확대되고 있는 비대면 서비스 및 IoT 산업과의 연관성과 파급효과가 큰 기술로 고속 성장 예상

- 비대면 및 IoT 서비스에서 요구되는 사용자 친화형 보안 핵심기술 제공을 통해 글로벌 경쟁력 강화 및 보안 신시장 창출에 기여 가능
- 비대면 솔루션의 다양한 보안 기능에 대한 사용 편의성(usability) 평가 및 개선을 위한 객관적인 지침 제시
- 비대면 교육, 진료, 회의 등의 사용자 인증 및 보안성 강화에 활용하여 비대면 서비스에 대한 사이버 공격 피해로 인한 경제적, 사회적 비용 최소화

3. 연구목표

○ **최종목표 : 비대면 환경에서의 사용자 친화적인 보안 핵심 기술 및 비대면 보안 솔루션의 사용 편의성(usability) 평가 기술 개발**

- 비대면 환경에서의 사용자 친화적 보안 요구사항 분석
- 비대면 및 IoT 서비스를 위한 사용자 친화형 인증 기술 개발
- IT 기반 지식이 없는 일반 사용자 중심의 보안 관리 기술 개발
- 비대면 보안 솔루션의 사용 편의성에 대한 평가 기술 개발

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	사용자 행위 인식 알고리즘 성능*	EER	5% 이내	10%	8.6% (미국, 뉴욕공과대학)
2	사용자 행위 기반 인증 소요 시간**	초	90초 이내	-	-
3	비대면 환경에서의 보안 요구 사항	건	6건 이상	-	-
4	제안 솔루션에 대한 사용성 테스트***	건	종래의 솔루션 대비 사용성 테스트를 통하여 개선 정도를 증명 (관련 SCI 논문 1건 이상)	-	-

* 성능 평가는 실제 서비스의 사용자 행위 데이터를 수집하여 결과 분석

** 스마트폰에서 불법적인 행위를 수행하는데 걸리는 평균 시간이 90초를 초과하는 연

구 결과에 기반하여 인증 소요 시간 목표를 설정(출처: HMOG, IEEE Trans. on IFS, 2015)

*** 사용성 분석은 usable security 분야에서 요구하는 국제 수준에 맞춰서 진행하며, 관련 SCI 논문(CHI, SOUPS, S&P 등) 게재로 테스트 분석 결과 검증

○ **연차별 개발목표**

구분	연도별 연구목표
2021년	비대면 환경에서의 인증 및 보안 기술에 대한 사용자 요구사항 분석
2022년	비대면 환경에서의 인증, 접근 제어, 보안 설정 기술 개발
2023년	사용자의 신원 정보 노출을 최소화하는 인증, 접근제어, 보안설정 기술 개발
2024년	제안 기술의 고도화 및 보안 솔루션에 대한 사용성 평가 제시

4. 연구내용

○ **개발 기술 내용**

① 비대면 환경에서의 사용자 친화적 보안 요구사항 분석

- 비대면 및 IoT 서비스*에 대한 사용자 보안 위협 및 요구사항 분석
  * 서비스 예) 온라인 수업, 재택근무, 원격의료, IP카메라/공유기/스마트가전/기타IoT 등
- 사용성 개선이 필요한 비대면 응용 도메인 대상 선정 및 선정 도메인에서의 보안(인증 및 보안 관리) 편의성 강화 방안 도출

② 비대면 및 IoT 서비스를 위한 사용자 친화형 인증 기술 개발

- 사용자의 위치 및 행위 등의 컨텍스트(context)를 이용한 추가적인 인증 요소 확보 및 실제 적용할 수 있는 (“human in the loop” 고려) 사용자 인증 기술 개발
  * 비대면 환경의 실제 서비스 데이터를 활용하여 컨텍스트 정보 분석 필요
  * 비대면 환경에서의 데이터 기반 사용자 친화적인 다양한 인증 알고리즘 연구 및 개발
- 사용자 신원 정보의 노출을 최소화하는 효율적인 프라이버시 보존 (privacy-preserving) 생체 및 행위 인증 기술 개발

③ IT 기반 지식이 없는 일반 사용자 중심의 보안 관리 기술 개발

- IT 기반 지식이 없는 사용자도 손쉽게 설정할 수 있는 보안 정책 설정 관리 기술
- IT 기반 지식이 없는 사용자도 효과적으로 대응할 수 있는 보안 경고 가시화 기술
  * 가시화 장치는 사용자의 스마트폰 또는 별도의 개인장치로 구성하며, 가시화된 위협(비허용 접근, 정보 변조, 민감정보 노출, 서비스 중지 등)에 대한 구체적 대응 방안 제시

④ 비대면 보안 솔루션의 사용 편의성에 대한 평가 기술 개발

- 비대면 및 IoT 보안 솔루션의 사용성을 평가하기 위한 정량적인 평가 기준 및 방안 개발

○ **기존 (보유)기술**

① 모바일 환경하에서 모바일 인증과 보안 강화를 위해 직관적이며 사용하기 편하고 안전한 인간-컴퓨터 상호작용(HCI) 기반 Usable Security 원천기술

- HCI 기반 인증 프로토콜 및 피싱 공격 대응 Usable 캡차 기술

- 동작 인식 센서를 이용하여 태블릿PC, 스마트TV 등 다양한 스마트기기에 적용 가능한 비터치 방식의 3D 기반 사용자 인증 기술

② 인터넷 환경에서 안전한 비대면 원격회의를 위한 화상회의서비스 사용자 간편인증 및 자료 보안용 클라우드 기반 보안 플랫폼

- 다양한 화상회의 시스템 사용자의 편의성이 보장되는 간편인증 기술 제공 및 공유자료 보안 기술

5. 지원기간/예산/추진체계

ㅇ **기간** : 4년 이내 (1단계 2년(경쟁) → 2단계 2년(단독))

ㅇ **정부출연금** : '21년 15억원(1차년 7.5억원×2개) 이내 (총 정부출연금 65억원* 이내)

* 2차년도 10억×2개, 3~4차년도 30억×1개

ㅇ **주관기관** : 제한없음

기술분류	대분류(차세대보안) - 중분류(시스템보안 및 암호보안) - 소분류(인증/인가)	
연구유형	기초연구 ( ), 응용연구 ( ), 개발연구 (O)	**TRL** (3) ~ (7)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형(O), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-15	(지정공모형)
과제명	임베디드 시스템 악성코드 탐지·복원을 위한 RISC-V 기반 보안 CPU 아키텍처 핵심기술 개발	

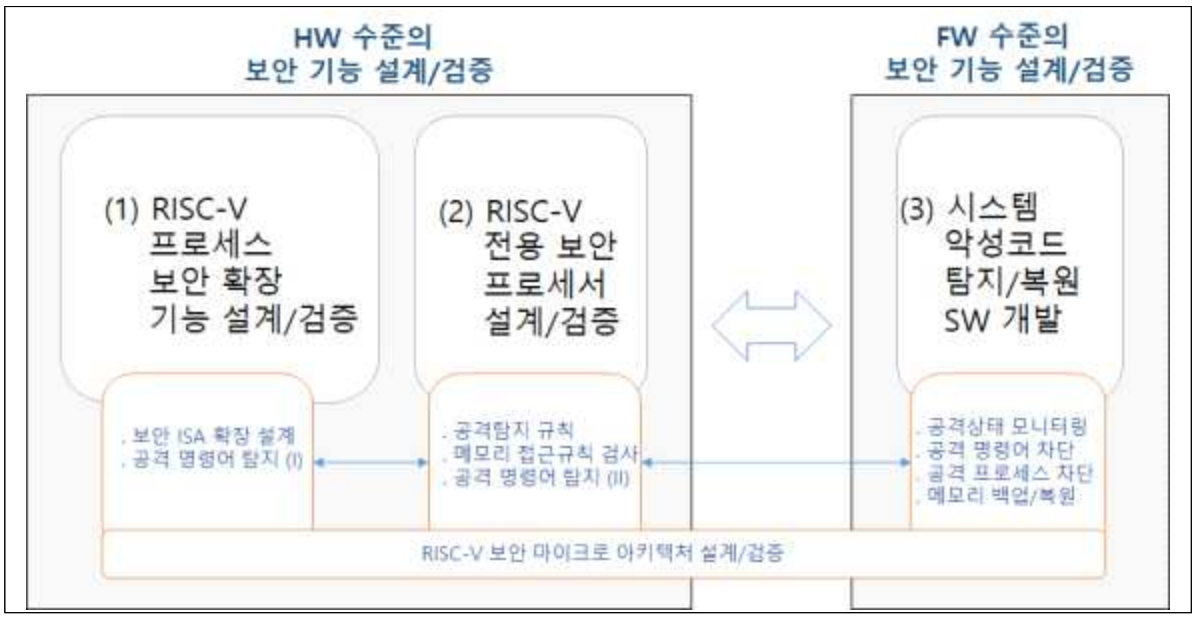
1. 개념

○ 임베디드 시스템 악성코드* 탐지 및 감염된 시스템의 복원을 위한 RISC-V 기반 보안 CPU 아키텍처 설계/검증 핵심 기술 개발

※ 임베디드 시스템 악성코드: 스마트 의료/가전/팩토리, 지능형 자동차, 전력/원자력 국가 ICS 제어망 등의 다양한 장치에 대한 시스템 잠금 및 파괴 목적의 악성코드 (시스템 랜섬웨어 등)

- 다양한 임베디드 장치에 대한 시스템 잠금 및 파괴 목적의 악성코드를 하드웨어 수준에서 탐지하기 위한 RISC-V ISA(Instruction Set Architecture) 및 마이크로아키텍처 확장 기술 설계
- 시스템 악성코드 공격에 대한 보안 ISA 확장 기반 실시간 탐지 및 복원 기술 개발
- RISC-V 보안 ISA 확장 핵심기술의 사실 표준화 및 실증 프레임워크 구축

< RISC-V 기반 보안 CPU 아키텍처 개념도 >



2 필요성

○ **(정부 지원 필요성)**

- (국가 사이버 주권확보) 국가 핵심 보안기술의 외산기술 종속성은 국가 사이버 주권확보에 치명적인 약점이 될 수 있음
- (민간기업의 보안분야 투자한계 극복) 국가차원의 新보안정책 실현을 위해 공공주도 핵심기술 개발 필요
- (국제협력 과학정책의 기술기반 제공) 제5차 한미 ICT 정책포럼(2020.09.10) 연구주제의 핵심기술로써 정부지원 과제로 추진 필요

○ **(기술성)** RISC-V는 IoT 및 경량 임베디드 시스템 분야의 CPU 시장을 주도할 것으로 전망되지만 경쟁력 확보를 위해 하드웨어 수준 보안 기능의 보완이 필수적임

- 다양한 임베디드 시스템에 대한 공격 피해가 증가하고 있으며, 악용된 취약점의 버퍼 오버플로우와 UAF(use after free) 취약점 등의 소프트웨어 버그가 대부분을 차지
- 소프트웨어 수준의 보안 기술은 성능 및 메모리 오버헤드로 인하여 임베디드 시스템에 적용하기 어려우며 하드웨어 수준의 보안 기술 필수적임
- RISC-V는 저전력과 저비용을 장점으로 경량 임베디드 시스템과 IoT 장치용 CPU 시장을 주도할 것으로 전망되지만 보안 기능은 걸음마 수준으로 경쟁력 확보를 위해 보완 필요

- 국내 PC 기반 백신 SW 수준으로는 임베디드 장치에 대한 실시간 대응 불가하며, 국외는 DARPA(SSITH), Google(OpenTitan), MS(Azure Sphere), SiFive(SiFiveShield) 등에서 임베디드 시스템 악성코드 대응을 위해 하드웨어 레벨의 보안 기술 연구시작

○ **(경제성)** 보안이 강조되고 있는 경량 임베디드 시스템 및 IoT 서비스 분야에서 RISC-V의 점유율 증가가 예상되면서 보안 기능이 탑재된 RISC-V CPU의 수요 증가 예상됨
- RISC-V CPU 코어의 매출은 2019년 이후 연간 146%씩 성장하여 2025년 624억 달러로 예상되며, 이 중 산업용 IoT 분야는 167억 달러를 차지하여 가장 많은 부분을 차지할 것으로 전망
  * RISC-V Market Analysis: The New Kid on the Block (Semico Research Corp., 2019)
- IoT 사이버 공격에 따른 국내 피해액이 2015년 13조 4,000억원에서 2020년 17조 7,000억원, 2030년 26조 7,000억원에 달할 것으로 예측
  * IoT 사이버 공격에 따른 피해액 (KT경제경영연구소, 2018)
- 경량 임베디드 시스템 및 IoT 서비스의 해킹 피해를 최소화하기 위해 하드웨어 수준의 보안 기능이 탑재된 RISC-V CPU의 수요가 증가할 것으로 전망

3. 연구목표

○ **최종목표 : 임베디드 시스템 악성코드 탐지 및 복원을 위해 RISC-V 기반 보안 CPU 아키텍처 설계/적용을 통한 하드웨어 레벨 보안 핵심 기술 개발**
- RISC-V 기반 보안 CPU 아키텍처 설계/검증 기술 개발
- 보안 전용 코프로세서 기반 악성코드 탐지 기술 개발
- 시스템 악성코드 공격에 대한 보안 ISA 확장 기반 실시간 시스템 복원 기술
- (표준화 및 실증) 보안 ISA 확장의 사실 표준화 및 실증 프레임워크 구축

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	탐지 가능한 공격/취약점 유형[주1]	EA	3	N/A	2[주2] (영국, ARM)
2	취약점 탐지율[주3]	기술 수준	≥1	N/A	1 (미국, 구글)
3	성능 오버헤드	%	≤15	N/A	17[주4] (미국, 오라클)
4	메모리 오버헤드	%	≤12.5	N/A	15.6[주5] (미국, 도버 마이크로시스템즈)
5	실시간 시스템 복원 응답 시간	ms	50	N/A	N/A

주1) stack overflow, heap overflow, use after free, privilege escalation, control hijacking 등의 공격/취약점 중 탐지 가능한 항목 개수

주2) ARMv8-A 는 BTI, PAC, PXN 등의 기능을 통해 control hijacking, privilege escalation 공격을 탐지
(* BTI: Branch Target Indicator, PAC: Pointer Authentication Code, PXN: Privilege eXecute Never)

주3) Google Address Sanitizer 대비 메모리 취약점 탐지율

주4) SPARC 프로세서는 ADI 기술 사용시 최대 성능(처리 시간) 오버헤드 17%
(* ADI: Application Data Integrity)

주5) Dover Microsystems 의 Coregurad 솔루션은 각 32비트 word에 대해 5비트의 컬러링(무결성보장) 정보와 접근 권한 정보를 추가로 사용함

○ **연차별 개발목표** * 기술분야별 특징에 따라 생략 가능 */

구분	연도별 연구목표
2021년	RISC-V 기반 ISA/마이크로아키텍처 설계
2022년	시뮬레이터 수준의 보안 아키텍처 및 악성코드 탐지 기술 구현
2023년	FPGA 수준의 보안 아키텍처 및 악성코드 탐지/복원 기술 구현
2024년	실증 프레임워크 구축 및 기능/성능 안정화, 사실 표준화

4. 연구내용

ㅇ **개발 기술 내용** /* 목표 달성을 위한 개발 내용을 구체적으로 기술 */

① RISC-V 기반 보안 CPU 아키텍처 설계/검증 기술 개발
- 악성코드/취약점 탐지 지원 명령어 및 레지스터 설계
- 오버헤드를 고려한 보안 마이크로아키텍처 설계
- 컴파일러/바이너리 계측 기반 보안 명령어 생성 및 주입 기능

② 보안 전용 코프로세서 기반 악성코드 탐지 기술 개발
- 공격/취약점 탐지 기반 정보(메타데이터) 추출을 위한 소프트웨어 분석 기술
- 코프로세서 기반 메타데이터 관리 및 메인 프로세서 모니터링 기술
- 메인프로세서-코프로세서 정보 공유 기술

③ 시스템 악성코드 공격에 대한 보안 ISA 확장 기반 실시간 시스템 복원 기술
- 보안 ISA 확장 기반 시스템 정보 추출 및 공격/취약점 진단 기능
- 진단 결과에 따른 실시간 시스템 복원 및 대응 기능
- 공격 명령어 및 악성 프로세스 차단 기능

④ (표준화 및 실증) 보안 ISA 확장의 사실 표준화 및 실증 프레임워크 구축
- RISC-V 파운데이션을 통한 사실 표준화 추진
- FPGA 기반의 실증 프레임워크 구축 및 기능 검증

ㅇ **기존 (보유)기술** */* 상용 기술 및 기존 정부 및 민간에서 개발중인 기술 포함 */*

① RISC-V 기반 보안 아키텍처 설계/검증 기술
- 제어흐름 무결성 기반 제어흐름 하이제킹(Control Hijacking) 공격 탐지 기술
- 메모리 실행 권한 모니터링 기반 권한 상승(Privilege Escalation) 공격 탐지 기술

② 임베디드 시스템 공격코드 검증을 위한 보안성 분석/검증 기술
- 하드웨어/펌웨어 역공학 기반의 펌웨어 분석 기술
- 펌웨어 정적/동적 분석 기반의 취약점 탐지 및 PoC(Proof-of-Concept) 기술

5. 지원기간/예산/추진체계

ㅇ **기간** : 4년 이내 (1단계 2년 → 2단계 2년)

ㅇ **정부출연금** : ‘21년 15억원 이내 (총 정부출연금 75억원 이내)

ㅇ **주관기관** : 제한없음

기술분류	대분류(차세대보안) - 중분류(시스템보안 및 암호보안) - 소분류(악성코드)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (4) ~ (6)
과제특징	정책지정( ), 혁신도약형(O), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-16	(지정공모형)
과제명	고신뢰 온-디바이스 딥러닝 가속기 설계를 위한 물리채널 기반 취약점 검증 및 대응기술 개발	

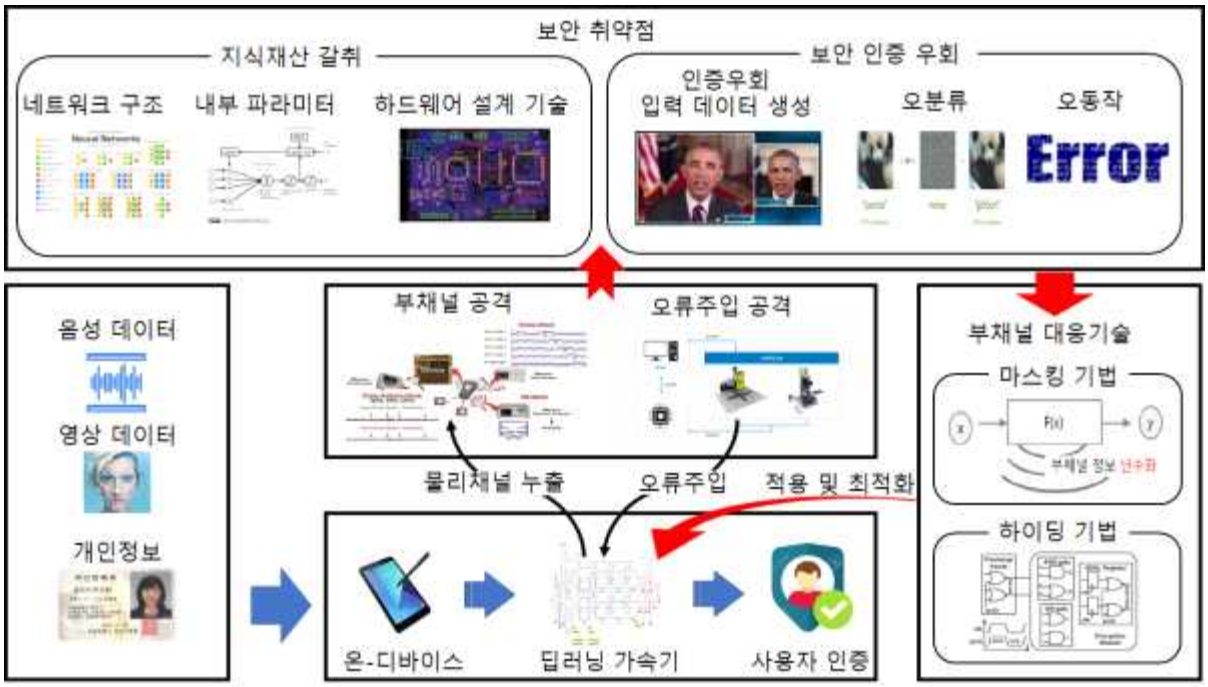
1. 개념

○ 모바일 기기 상에서 딥러닝(deep learning)을 활용하는 다양한 어플리케이션을 구동하기 위해 기기 자체에 탑재되는 온-디바이스 딥러닝 가속기*의 안정성을 보장할 수 있는 전력/전자파** 및 오류주입*** 기반 물리채널 취약점 검증 및 대응기술

* 인간의 뇌에서 일어나는 의사 결정 과정을 모방하여 만든 인공 신경망 네트워크를 에지 디바이스 상에 저장하여 사용하는 프로세서로서 그 활용범위와 시장규모가 급속도로 증대되고 있음

** 딥러닝 가속기의 소비 전력/전자파를 분석하거나 오류주입을 통해 model stealing (딥러닝 가속기 내부구조 및 파라미터 정보를 알아내는 기술) 하는 기술로서 이를 통해 사용자 인증을 우회하기 위한 입력데이터를 생성하는 등의 악의적인 행위가 가능

*** 딥러닝 가속기에 레이저 및 전자파 주입을 통해 가속기의 오분류·오동작을 야기하는 기술



< 기술 개념도 >

2 필요성

○ **(정부 지원 필요성)** 물리채널 기반 보안 취약점을 사전 차단한 고신뢰 온-디바이스 딥러닝 가속기 설계 원천·핵심 기술의 선제적 확보를 위한 정부의 적극적인 지원 필요

- 최근 딥러닝 기술이 컴퓨터 비전, 오디오, 생체의료 및 헬스케어 등 다양한 어플리케이션에 적용됨에 따라, IoT/스마트 센서, 웨어러블 디바이스 등의 모바일 에지(mobile edge) 디바이스와 같은 전력 소모가 제한된 기기에서의 딥러닝 기술 구현을 위한 저전력·저면적 인공 신경망 가속기 구현의 필요성이 증대됨. 이에 따라 국내에서도 정부의 지원 하에 다양한 딥러닝 가속기 설계 과제가 진행되고 있음
- 한편, 온-디바이스 딥러닝 가속기는 모바일 기기에서 음성, 이미지, 영상, 위치 정보 등과 같은 민감한 개인정보를 다룰 뿐만 아니라 사용자 인증 등의 과정에도

다양하게 활용되고 있음. 즉, 소비 전력/전자파 분석을 통한 딥러닝 가속기의 내부구조, 파라미터 정보를 찾아내 인증을 우회할 수 있는 입력 데이터를 생성하거나, 오류 주입을 통해 가속기가 오분류·오동작하도록 할 경우 심각한 위협을 초래할 수 있음

- 또한 딥러닝 기반 서비스는 입력에 대한 추론을 위하여 기학습된 신경망의 구조 및 파라미터 등을 사용하며, 신경망 내부정보들은 선별 및 학습에 막대한 비용이 소요됨. 즉, 온-디바이스 딥러닝 가속기의 내부구조 및 파라미터 정보는 이를 이용하는 산업의 핵심 기밀요소로 보호되어야만 하는 요소임
- 물리채널 기반 취약점에 대한 안전성이 보장된 고신뢰 온-디바이스 딥러닝 가속기는 산업 전반에서 활용도가 높고 보안 측면에서 굉장히 중요한 데이터를 다루고 있음. 이에, 해당 취약점 검증 및 대응기술을 선도적으로 연구하고 물리적으로 안전한 온-디바이스 딥러닝 가속기 및 이에 대한 안전성을 평가할 수 있는 검증도구를 개발하여 해당 분야를 우리나라가 선도할 수 있도록 정부의 적극적인 지원이 요구됨

ㅇ **(기술성)** 물리채널 기반 취약점에 대한 안전성을 보장할 수 있는 고신뢰 온-디바이스 딥러닝 가속기 설계 기술 요구 증대

- 온-디바이스 딥러닝 가속기의 현황 및 산업 전망을 고려할 때 현재 입출력 및 신경망의 구조/파라미터에 대한 물리채널 기반 취약점 검증 및 대응기술 관련 연구는 초기 단계임
- 컴퓨팅 파워가 제약적인 환경에서 사용되는 온-디바이스 딥러닝 가속기는 저면적/저전력/고성능을 유지해야 하여, 이러한 환경에서 물리채널 기반 취약점을 방지 가능한 설계기술은 거의 전무한 실정임. 그러나 이러한 설계기술은 딥러닝 가속기 기술 시장을 선도할 핵심 원천 기술로서 관련 연구에 대한 투자는 필수불가결함

ㅇ **(경제성)** 고부가 가치 창출이 예정될 시장의 우위 기술 선점을 통한 국내 관련 기업들의 세계 경쟁력 극대화 가능

- 모바일, IoT 환경 등에서 딥러닝 기술 기반 서비스가 증가함에 따라 온-디바이스 지능형 프로세서 탑재로 인한 시장의 규모가 급증하고 있음. 시장 분석 전문기관인 MarketsandMarkets*에 따르면 딥러닝 가속기의 시장 규모는 현재 약 76억 달러 내외로 추산되며, 연평균 40.1%의 속도로 성장하여 2026년에 약 578억 달러까지 증가할 것으로 예측됨

* https://www.marketsandmarkets.com/Market-Reports/artificial-intelligence-chipset-market-237558655.html

- 급증하는 온-디바이스 딥러닝 가속기 시장을 선점하기 위하여 전세계적으로 많은 IT 업체들이 저전력 가속기 구현 기술개발 및 연구에 박차를 가하고 있지만, 물리채널 취약점에 대한 안전성이 보장된 저전력 딥러닝 가속기 구현기술에 대한 연구 및 개발은 매우 미진한 실정임. 따라서 우리나라가 해당 분야에 대한 기술력 우위를 확보하여 시장을 선도하기 위해서는 기존의 접근법과는 차별화된 연구가 필요한 상황임

3. 연구목표

○ 최종목표 : 온-디바이스 딥러닝 가속기의 안전성을 검증하고 보장하기 위한 소비전력/전자파 및 오류주입 기반 물리채널 취약점 검증 도구 및 대응기술 개발

○ 정량적 개발목표

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	소비전력 기반 model stealing 취약점 검증 도구 검증 항목 수[주1]	개수	6	-	4[주2] (네덜란드, Radboud University)
2	오류주입 기반 오분류·오동작, model stealing 취약점 검증 도구 검증 항목 수[주3]	개수	5	-	3[주4] (싱가포르, 난양공대)
3	오류주입 기반 model stealing 취약점 분석 기술의 weight error 비율[주5]	%	$< 10^{-11}$	-	10^{-11} (싱가포르, 난양공대)
4	물리채널 기반 딥러닝 가속기 취약점 신규 대응기술 개발 항목 수	개수	2	-	1[주6] (미국, Intel)
5	물리채널 취약점 대응기술이 적용된 딥러닝 가속기 기준전력 대비 성능	TOPS/W	1.0[주7]	-	-

주1) 신규 개발된 소비전력 기반 model stealing 취약점 검증도구가 지원하는 취약점 검증항목 수

주2) 딥러닝 가속기의 소비전력 기반 model stealing 취약점을 평가할 수 있는 검증도구가 존재하지 않아 현재까지 보고된 취약점 개수 및 최고 기술 수준 보유국, 기업/기관을 기준으로 정의

주3) 신규 개발된 오류주입 기반 오분류·오동작 취약점 검증도구가 지원하는 취약점 검증항목 수

주4) 딥러닝 가속기의 오류주입 기반 오분류·오동작 취약점을 평가할 수 있는 검증도구가 존재하지 않아 현재까지 보고된 취약점 개수 및 최고 기술 수준 보유국, 기업/기관을 기준으로 정의

주5) 오류주입 기반 분석을 활용해 찾아낸 내부 파라미터 정보의 error 비율
참고자료 : Jakub Breier, Dirmanto Jap, Xiaolu Hou, Shivam Bhasin and Yang Liu, “SNIFF: Reverse Engineering of Neural Networks with Fault Attacks”

주6) 최근 딥러닝 가속기에 대한 물리채널 취약점 연구가 활발히 진행 중에 있으나 이에 대한 대응기술은 초기 연구단계로서 현재 BNN에 대한 대응기술 개발이 유일함
참고자료 : Anuj Dubey, Rosario Cammarota, Aydin Aysu, “MaskedNet: The First Hardware Inference Engine Aiming Power Side-Channel Protection”

주7) 암호 분석에 대한 하드웨어 대응기술인 Threshold Implementation의 성능 부하 정도(약 3~5배)와 Google Mobile TPU의 소모전력 대비 성능인 4.8 TOPS/W을 참고하여 목표 설정

○ 연차별 개발목표

구분	연도별 연구목표
2021년	- 딥러닝 코어 컴포넌트에 대한 오분류·오동작 취약점 검증 기술 개발 - 딥러닝 코어 컴포넌트에 대한 model stealing 취약점 검증 기술 개발 - 물리채널 기반 취약점 실험·검증을 위한 기본 딥러닝 네트워크 가속기 설계
2022년	- 대표적 음성/영상 딥러닝 네트워크에 대한 오분류·오동작 취약점 검증 기술 개발

	- 대표적 딥러닝 네트워크에 대한 대한 model stealing 취약점 검증 기술 개발 - 하드웨어 구성에 따른 물리채널 누출 정보 분석 기술 개발
2023년	- 딥러닝 코어 컴포넌트 및 네트워크 가속기 취약점 검증도구 시작품 개발 - 딥러닝 코어 컴포넌트에 대한 오분류·오동작 취약점 대응기술 개발 - 딥러닝 코어 컴포넌트에 대한 model stealing 취약점 대응기술 개발 - 물리채널 기반 취약점 대응기술이 적용된 딥러닝 코어 컴포넌트 설계
2024년	- 딥러닝 코어 컴포넌트 및 네트워크 가속기 취약점 검증도구 개발 및 최적화 - 대표적 음성/영상 딥러닝 네트워크에 대한 오분류·오동작 취약점 대응기술 개발 - 대표적 딥러닝 네트워크에 대한 model stealing 취약점 대응기술 개발 - 물리채널 기반 취약점 대응기술이 적용된 SW/HW 통합 딥러닝 가속기 최적화 설계

4. 연구내용

ㅇ 개발 기술 내용

① 딥러닝 코어 컴포넌트에 대한 물리채널 기반 신규 취약점 검증 기술 개발

- 딥러닝 코어 컴포넌트에서 발생하는 물리채널 종류 및 특징 분석
- 딥러닝 코어 컴포넌트 물리채널 기반 분석을 위한 맞춤형 취약점 검증 환경 구축
- 딥러닝 코어 컴포넌트에 대한 물리채널 기반 신규 취약점 검증 기술 개발
  - 오류주입 기반 오분류·오동작 취약점 검증 기술 개발
  - 오류주입 기반 model stealing 취약점 검증 기술 개발
  - 전력/전자파 기반 model stealing 취약점 검증 기술 개발

② MLP, CNN 등 대표적 음성/영상 딥러닝 네트워크 가속기에 대한 물리채널 기반 신규 취약점 검증 기술 개발

- 딥러닝 네트워크에 대한 오류주입 기반 오분류·오동작 취약점 검증 기술 개발
  - 레이저 및 전자파 주입을 통한 오분류·오동작 취약점 검증 기술 개발
- 딥러닝 네트워크에 대한 오류주입 기반 model stealing 취약점 검증 기술 개발
- 딥러닝 네트워크에 대한 전력/전자파 기반 model stealing 취약점 검증 기술 개발
  - 대표적인 부채널 공격인 Timing attack, Simple power attack, Differential power attack, Single trace attack, Template attack 등을 응용한 딥러닝 네트워크 가속기의 내부 구조 파악 기술, 입력 데이터 및 내부변수 복원 기술 개발
- 네트워크 프루닝 (pruning), 연산 Bit-width, bias, 데이터 플로우의 물리채널 누출 정보 분석 기술 개발
- 가속기 내의 코어 (core) 개수, 코어 내 MAC (multiplier accumulator) 개수, Global/Local 버퍼 사이즈, 한 번에 처리되는 이미지 숫자 (Batch Processing) 기반 가속기 하드웨어 구성요소 물리채널 누출 정보 분석 기술 개발

③ 딥러닝 코어 컴포넌트 및 네트워크 가속기에 대한 물리채널 취약점 안전성 검증 도구 시작품 및 프로토타입 개발·최적화

④ 딥러닝 코어 컴포넌트에 대한 물리채널 기반 취약점 신규 대응기술 개발 및 하드웨어 설계

- 코어 컴포넌트를 대상으로 Constant-time, 부채널 원자성(atomicity), 마스킹/하이딩 기법을 적용하기 위한 전력/전자파 기반 model stealing 취약점 신규 대

응기술 개발

- 코어 컴포넌트에 대한 오류주입 기반 오분류·오동작 취약점 신규 대응기술 개발
- 코어 컴포넌트에 대한 오류주입 기반 model stealing 취약점 신규 대응기술 개발
- 물리채널 기반 취약점 대응기술이 적용된 딥러닝 코어 컴포넌트 설계

⑤ 대표적 음성/영상 딥러닝 네트워크에 대한 물리채널 기반 취약점 신규 대응기술 개발

- 대응기술이 적용된 코어 컴포넌트를 활용, 대표적 음성/영상 딥러닝 네트워크에 대한 하드웨어 친화적인 (hardware friendly) 신규 대응기술 개발
  - 오분류·오동작 및 model stealing 취약점에 대한 딥러닝 네트워크 대응기술 개발

⑥ 물리채널 기반 취약점 대응기술이 적용된 SW/HW 통합 딥러닝 가속기 최적화 설계 및 안전성 검증

- 물리채널 기반 취약점 대응기술이 적용된 SW/HW 통합 딥러닝 가속기 최적화 설계
  - 신규 물리채널 기반 취약점 대응기술의 저전력·저면적 하드웨어 최적화 기술 개발
  - 하드웨어 구조 변화에 따른 물리채널 기반 취약점 분석 및 대응기술 경향성 분석
- 물리채널 취약점 안전성 검증도구를 활용하여 SW/HW 통합 딥러닝 가속기 취약성 검증

ㅇ 기존 (보유)기술

① 암호 알고리즘에 대한 물리채널 기반 취약성 검증 기술 및 평가 도구

- 대칭키/공개키 암호에 대한 전력/전자파 기반 취약성 검증 기술 및 평가 도구
- 대칭키/공개키 암호에 대한 오류주입 기반 취약성 검증 기술 및 평가 도구

② 펌웨어 기반 딥러닝 네트워크에 대한 물리채널 기반 취약점 분석 기술 개발 중

- 펌웨어 기반 딥러닝 네트워크에 대한 전력/전자파 기반 내부변수 복원 기술 개발 중
- 내부변수, 활성 함수에 오류주입 공격을 수행하여 오분류를 유도하는 공격 개발 중

③ 딥러닝 가속기 하드웨어 최적화 설계 기술

- TPU, NPU 등 딥러닝 가속기 하드웨어 최적화 설계 기술

5. 지원기간/예산/추진체계

ㅇ **기간** : 4년 이내 (1단계 2년 → 2단계 2년)

ㅇ **정부출연금** : ‘21년 9억원 이내 (총 정부출연금 45억원 이내)

ㅇ **주관기관** : 제한없음

기술분류	대분류(시스템·디바이스보안) - 중분류(보안취약성) - 소분류(HW취약점분석)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (3) ~ (6)
과제특징	정책지정( ), 혁신도약형(O), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-17	(지정공모형)
과제명	개인정보보호를 위한 신뢰계산 기반 데이터보호박스 개발	

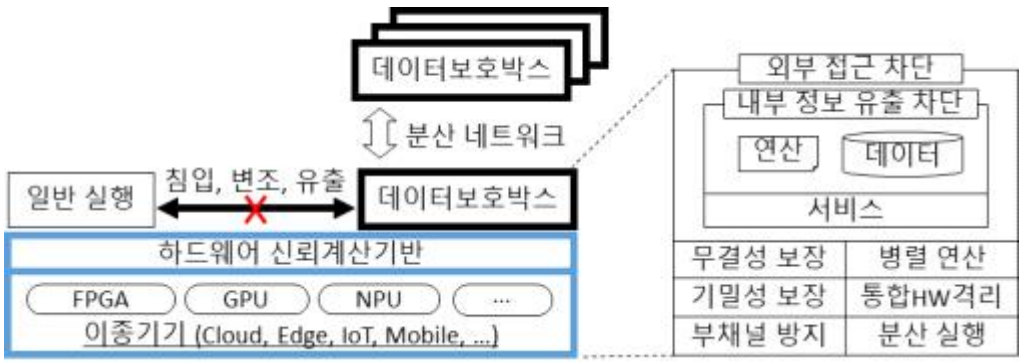
1. 개념

○ IoT・엣지・클라우드와 같이 이종의 기기* 간 경계가 모호해지는 차세대 초연결 컴퓨팅 환경에서 향상된 하드웨어기반의 신뢰계산 기술을 활용하여 안전하고 효율적인 분산 데이터처리에 특화된 데이터보호박스** 개발

- 소프트웨어-하드웨어 공동디자인을 통해, 이종의 기기에서 AI・빅데이터 등 폭넓은 연산능력을 요구하는 응용들에게 최적화된 성능/보안을 보장할 수 있는 데이터 처리기술을 개발하고, 이를 기반으로 다가오는 초연결 시대에 신뢰할 수 있는 효율적인 원격 분산 실행 환경 구현

* 상이한 하드웨어 아키텍처를 탑재한 기기

** 외부의 침투 및 감시가 불가능한 격리된 실행환경에서 데이터를 처리하는 실행환경



< 하드웨어 신뢰계산기반의 데이터보호박스 >

2 필요성

○ **(정부 지원 필요성)** 최근 개정된 데이터 3법등을 통해 4차 산업시대에는 데이터의 적극적인 유통 및 활용에 대한 필요성과 유출에 대한 우려가 동시에 커지고 있기 때문에, 데이터보호박스는 데이터의 보안우려를 해소함으로써 국가 공공기관 및 민간의 데이터 활용을 극대화하고 미래 데이터 주권을 보장하기 위해 필요

- 정부의 차세대 정보보호 정책*에서도 원격에서 데이터를 안전하게 처리하기 위한 보안기술의 필요성을 제기하여 이를 위한 계획수립 추진

* 과기정통부 차세대 정보보호 R&D 기술로드맵 - 데이터 활용 보안기술, 안전한 비대면 서비스를 위한 사용자 보호 및 보안 기술

○ **(기술성)** 하드웨어기반의 신뢰계산 기술을 통해 성능 저하를 최소화 하면서도 원격계산 데이터의 안정성을 동형암호에 비견할만한 수준으로 올림으로써, 신뢰할 수 있는 차세대 초연결 컴퓨팅 환경 구축을 위한 원천기술 확보 가능

- Intel, AMD, ARM 아키텍처에서 일반적인 격리실행 환경을 제공 중이며, 스마트폰, 클라우드 등 다양한 환경에서 이를 적극 활용하고 있으므로, 데이터 주권 확보를 위해 보다 신속하고 안전한 데이터 처리에 특화된 데이터보호박스의 하드웨어/소프트웨어 원천기술 확보 시급

- 최근 급격히 발전하고 있는 RISC-V, CPU/FPGA, NPU, SPU 등의 범용 하드웨어 아키텍처 및 IP의 활용성 극대화 및 기술선도를 뒷받침하기 위한 신뢰할 수 있는 데이터보호박스 기술 마련 필요

○ **(경제성)** 데이터가 중심이 되는 AI·빅데이터·IoT·클라우드 등의 4차산업 주축 기술의 데이터 훼손 및 유출 우려에 대비한 보안 차별화를 통해, 인프라 구축 및 신규 서비스 개발의 글로벌 경쟁력 확보

- 데이터보호박스 기반의 신뢰할 수 있는 데이터 처리를 통해 강화되고 있는 각국의 데이터 보호 요구*에 대한 장벽을 제거함으로써, 글로벌 시장 진출의 교두보 마련

  * EU의 GDPR (General Data Protection Regulation), 캘리포니아의 CCPA (California Consumer Privacy Act) 등

- 2025년 세계 데이터 시장의 규모는 약 900억 달러에 이르며, 그 양은 약 1조 1천억 GB에 이를 것으로 예상되므로, 안전한 데이터 확보/유통/처리를 위한 데이터 보호박스에 대한 수요가 폭발적으로 증가할 것으로 예상

3. 연구목표

○ **최종목표 : 차세대 초연결 컴퓨팅 환경에서 하드웨어에 신뢰를 기반으로 한 계산 기술을 활용해 안전하고 효율적인 원격 분산 데이터 처리를 보장하는 데이터보호박스 개발**

- (1단계:1~4년차) 하드웨어 신뢰계산기반의 데이터보호박스 개발
- (2단계:5~8년차) 동종/이종기기 간 분산 데이터보호박스 개발 및 PoC 검증

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	데이터보호박스가 구현된 아키텍처의 종류	개	6[주1)]	1	3 (영국, ARM)
2	Oblivious 실행의 성능 향상	배	100	1[주2)]	N/A
3	데이터보호박스 내 격리 실행으로 인한 성능 저하	%	< 10%	N/A	> 40%[주3)]
4	기기 간 데이터보호박스의 Seamless한 이전 (Migration) 시간	초	< 5	N/A	25[주4)]

주1) 가장 폭넓게 사용되고 연구되는 7종의 아키텍처 지원: Intel/AMD x86, ARM Cortex-A/R/M, RISC-V

주2) Intel x86에서 구현된 최신 Oblivious 실행 구현 기준
(참조논문: Ahmad, Adil, et al. "OBFUSCURO: A Commodity Obfuscation Engine on Intel SGX." NDSS. 2019.)

주3) Intel SGX를 활용하여 신뢰계산 환경에서 데이터베이스 연산을 수행할 경우
(참조논문: Weichbrodt, Nico, Pierre-Louis Aublin, and Rudiger Kapitza. "sgx-perf: A performance analysis tool for Intel SGX enclaves." Proceedings of the 19th International Middleware Conference. 2018.)

주4) Intel x86에서 SGX Enclave의 실시간 실행 이전 기준
(참조논문: Jinyu, Gu, et al. "Secure Live Migration of SGX Enclaves on Untrusted Cloud" DSN. 2017)

○ 연차별 개발목표

구분		연도별 연구목표
단계	년차	
1	1	- 각 이종기기의 아키텍처를 고려한 하드웨어 신뢰계산기반 설계 - 다양한 계산 유닛들을 포괄하는 통합 격리 실행 기술 개발
	2	- 내부정보의 유출 방지를 위한 양방향 샌드박스 개발 - 데이터보호박스 내 효율적 병렬처리 지원 기술 개발
	3	- 고속의 Oblivious 실행을 실현하는 기술 개발 - 예측실행 기반의 부채널 공격에 대한 방어 기술 개발
	4	- 하드웨어 신뢰계산기반에 대한 보안성 검증 - 데이터보호박스의 보안성 검증 - 프로그램을 데이터보호박스에서 실행하기 위한 개발도구 구축
2	5	- 데이터보호박스의 실시간 보안 상태 측정 (Measurement) 기술 개발 - 원격 데이터보호박스 증명 기술 개발
	6	- 동종 기기 간 분산된 데이터보호박스 체계 개발 - 동종 기기 간 Seamless한 데이터보호박스 이전 개발
	7	- 이종 기기 간 분산된 데이터보호박스 체계 개발 - 이종 기기 간 Seamless한 데이터보호박스 이전 개발
	8	- 분산 데이터보호박스 활용을 위한 개발도구 구축 - 분산 데이터보호박스 활용 Use Case 적용 검증

4. 연구내용

○ 개발 기술 내용

① 향상된 하드웨어 신뢰계산기반 개발
- 각 이종기기의 아키텍처를 고려한 하드웨어 신뢰계산기반 구축
- 프로그램의 실행상태를 실시간으로 수집 및 요약하는 기술
- 하드웨어 신뢰계산기반의 보안성에 대한 검증 기술

② 향상된 하드웨어 신뢰계산기반을 이용한 데이터보호박스 개발
- 외부에서 임의로 접근하려는 시도와, 은닉 채널을 통해 내부에서 임의로 정보를 유출하려는 시도를 차단하는 양방향 격리 실행 기술
- 예측실행을 이용한 부채널 공격에 대한 방어 기술
- 공유 하드웨어 자원 (캐시, 버스, TLB 등)의 할당/접근 패턴정보, 수행시간 정보를 이용한 부채널 공격에 대한 방어를 위한 Oblivious 실행 기술
- 데이터보호박스의 보안성에 대한 검증 기술

③ 데이터보호박스의 성능 최적화 기술 개발
- 병렬연산을 통해, 대용량 데이터의 고속 처리를 지원하는 격리 실행 기술
- 시스템에 존재하는 FPGA, GPU, NPU 등의 다양한 이종 계산 기기들을 지원하는 통합 격리 실행 기술
- 데이터보호박스 내 프로그램 개발을 위한 개발도구 구축

④ 이종/동종 기기 간 분산 데이터보호박스 체계 개발
- 분산된 데이터보호박스 간 안전한 통신 채널 수립 및 유지 기술
- 분산된 데이터보호박스 간에 코드/메모리의 무결성을 실시간으로 측정 및 증명하는 기술

- 동종/이종 기기 간에 Seamless하게 데이터보호박스를 이전하는 기술
- 다양한 이종 시스템에 적용 가능한 포터블 데이터보호박스 플랫폼 및 개발도구 구축
- 분산 데이터보호박스의 Use Case 발굴 및 효용성 검증(예: 개인정보를 보호하는 코로나 확진자 동선 검증 서비스, 개인정보를 보호하는 네트워크 방화벽 서비스 등)

ㅇ 기존 (보유)기술

① 데이터보호박스에 활용될 수 있는 부분 하드웨어/소프트웨어 요소들이 국내외 다양한 대학과 기업에서 연구 개발되고 있음
- 기밀성 보장을 위한 하드웨어 신뢰계산기반 기술
- 캐시/TLB 등의 부채널 공격을 방어하기 위한 하드웨어 확장
- 프로세서의 컨트롤 및 메모리 접근 정보 추출 기술
- x86 기반의 양방향 샌드박스 기술
- 병렬 데이터 처리를 가속하는 하드웨어 기술

② 원격 실행을 위한 기술이 현장에서 활용되고 있음
- 프로그램을 분산된 실행 환경에서 나누어 실행하는 기술
- 암호화 기법을 기초로 한 보안 채널 수립 기술
- 프로그램 코드의 무결성을 원격으로 검증하는 기술
- 실행중인 프로그램을 상태를 기기 간에 Seamless하게 실시간 이전하는 기술

5. 지원기간/예산/추진체계

ㅇ **기간** : 6년 이내 (1단계 3년 → 2단계 3년)

ㅇ **총 정부출연금** : '21년 4억원 (총 정부출연금 29억원 이내)

ㅇ **주관기관** : 대학

기술분류	대분류(차세대보안) - 중분류(시스템 및 암호보안) - 소분류(시스템 보안)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (3) ~ (6)
과제특징	정책지정( ), 혁신도약형(O), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-18	(지정공모형)
과제명	실환경 기반 마스크 착용자 얼굴인식 및 재인식(Re-ID) 기술	

1. 개념

○ CCTV 및 실내 출입통제시스템 환경을 포함하는 실환경 상황에서 마스크 착용 등으로 얼굴이 가려진 사람의 얼굴인식 및 해당 사람의 추적 및 재인식을 통하여 얼굴이 가려진 사람들에 대한 현 보안 시스템의 모니터링 기능 및 성능을 강화하고자 함

< 개념도 >

2 필요성

○ **(정부 지원 필요성)**

- 최근 발생하고 있는 코로나 사태 등으로 인해 전 세계 국가들은 공공 및 다중이용시설에서 마스크 착용을 의무화 하는 조치를 시행하고 있으며, 이로 인해 마스크 등으로 가려진 얼굴에 대한 확인, 대응 능력이 감소하여 CCTV 등 사회 안전 시스템의 효용성이 감소하고 있어 이에 대한 기술적 해결책 마련이 필요
- CCTV 환경에서의 얼굴 인식 및 재인식 기술은 상용화에 미치지 못하는 성능적 제약, 개인 정보 활용에 대한 제약, 대규모 연구 개발비용 부담으로 인하여 민간 개발 동력이 약하며, 따라서 정부 주도의 연구 개발이 필요함. 중국에서도 정부 주도로 연구 개발을 통한 CCTV 기반 인식 시스템 개발 및 도입이 이루어지고 있음

○ **(기술성)**

- 일반적인 얼굴인식 기술의 경우, 중국·미국을 중심으로 기술적 우위성을 확보하고 있는 상황에서 기술적 장벽을 극복하고 기술 경쟁력을 강화하기 위한 기반

기술의 확보가 필요

- CCTV 환경에서의 얼굴 인식 및 재인식 기술 개발을 위해서는 영상 처리 알고리즘과 카메라 시스템의 통합적인 개발이 필요하며 산학연 연계를 통한 체계적인 기술 개발이 필요

○ (경제성)

- 얼굴인식 솔루션 시장은 중국의 메그비(Megvii), 센스타임(SenseTime), 이투커지(Yitu) 등이 주도하고 있으며, 미국 IBM, MS 등은 경찰의 과잉 대응으로 인한 흑인 사망사건 이후 얼굴인식 기술 규제 마련 시까지 솔루션에 대한 사업화를 보류하고 있는 상황으로, 핵심 기술 확보를 통한 인공지능 기반의 영상보안시장 주도권 확보를 위해 개발이 필요
- CCTV 보급률은 지속적으로 증가하고 있는 반면 지능형 감시 기능의 도입은 정체되어 있는 현 상황에서, 상용화 가능한 효과적인 지능형 감시 기능의 개발을 통한 경제적 효과는 매우 클 것으로 보임

3. 연구목표

○ **최종목표 : CCTV 및 실내 출입통제시스템 환경을 포함하는 실환경 상황에서 마스크 착용 등으로 얼굴이 가려진 사람의 얼굴인식 및 해당 사람의 재인식을 지원하는 지능형 영상보안시스템 핵심 기술 개발**

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	가려진 얼굴 검출 기술 정확도(mAP)	%	92	-	88.3* (China, Beihang Univ.)
2	가려진 얼굴 인식 기술 정확도	%	70	-	56.34** (China, Tencent)
3	가려진 얼굴의 사람 재인식 기술 정확도	%	95	81.77****	94.5*** (China, Sun Yat-en Univ.)

* Face attention network: An effective face detector for the occluded faces, CoRR 2017

** Occlusion robust face recognition based on mask learning with pairwise differential siamese network, ICCV 2019

*** Spatial-temporal person re-identification, AAAI 2019 (가려지지 않은 얼굴의 사람 재인식 성능)

**** 사람과 자동차 재인식이 가능한 다중 손실함수 기반 심층 신경망 학습, 멀티미디어학회논문지 2020 (가려지지 않은 얼굴의 사람 재인식 성능)

○ **연차별 개발목표**

구분	연도별 연구목표
2021년	CCTV 환경에서 마스크 등으로 가려진 얼굴 검출 및 인식 기술 개발
2022년	마스크 등으로 가려진 얼굴 인식 기술 고도화 및 재인식 기술 개발
2023년	CCTV 환경 마스크 얼굴 검출, 인식, 재인식 기술 안정화 및 실증 테스트

4. 연구내용

ㅇ **개발 기술 내용**

① CCTV 환경에서 마스크 등 가려진 얼굴 검출 및 인식 기술 개발
- Occlusion type별 학습용 가려진 얼굴 DB 구축
- 원거리, 소형 객체를 중심으로 마스크 등으로 가려진 얼굴에 대한 정밀 검출 기술 개발
- 다양한 Occlusion Type에 강인한 가려진 얼굴 인식을 위한 신경망 모델링 기술 개발
- 가려진 얼굴의 정밀 복원을 위한 영상 해상도 향상 기술 개발
- CCTV 환경 및 walk-through형 출입통제시스템에 적용 가능한 실환경 기반 가려진 얼굴인식 기술 개발

② CCTV 환경에서 가려진 얼굴의 사람 추적 및 재인식 기술 개발
- 가려진 얼굴 사람의 추적 및 재인식을 위한 특징 분석 및 신경망 모델링 기술 개발
- 얼굴, 신체로부터 획득된 특징 별 융합을 통한 추적 및 재인식 성능 향상 기술 개발
- CCTV 및 출입통제시스템에 적용 가능한 가려진 얼굴의 사람 추적 및 재인식 기술 개발
- CCTV 및 출입통제시스템에 적용 가능한 추적 및 재인식 시스템 개발

③ 재인식 시스템 실증 테스트 및 데이터셋 공개
- CCTV 및 출입통제시스템에 적용 가능한 추적 및 재인식 시스템 통합 및 안정화
- 지자체, 출입통제시스템과 연계한 기술 유효성 검증 및 실증 테스트
- Occlusion type별 가려진 얼굴에 대한 학습 데이터셋(GT 포함) 공개

ㅇ **기존 (보유)기술**

① 얼굴 검출 및 인식 기술
- 근거리 또는 가려지지 않은 얼굴 검출 및 인식 기술
- CCTV 환경 또는 가려진 얼굴 검출 및 인식 기술은 개발 초기 단계이거나 상용화 성능에 미치지 못함

② 객체 추적 및 재인식 기술
- CCTV 환경에서의 객체 추적 및 재인식 기술은 성능 개선 중이나 상용화 수준에 미치지 못함
- 가려진 얼굴 객체의 추적 및 재인식 기술에 대한 연구는 시도되지 않음

5. 지원기간/예산/추진체계

ㅇ **기간** : 3년 이내

ㅇ **정부출연금** : ‘21년 12억원 이내 (총 정부출연금 42억원 이내) ㅇ **주관기관** : 제한없음(산업체 참여 필수)		
기술분류	대분류(차세대보안) - 중분류(물리보안) - 소분류(CCTV 감시/관제)	
연구유형	기초연구 ( ), 응용연구 ( ), 개발연구 (O)	**TRL** (4) ~ (7)
과제특징	정책지정( ), 혁신도약형(O), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-19	(지정공모형)
과제명	다중 스펙트럼 영상 통합 분석 기술 및 영상 감시 장치 개발	

1. 개념

○ 주요 보안시설 및 산업단지의 보안 감시와 안전 관리를 위하여 상황인식 정확도가 높고, 오감지율이 낮은 다중 스펙트럼 영상분석 시스템 기술을 개발함

- 정확한 상황인식을 위하여 가시광(VIS), 중·장파장 IR을 포함하는 다중 스펙트럼 영상 데이터셋 구축과 AI 영상분석 기술
  * VIS: 400~700nm, 중파장: 3~7㎛, 장파장: 8~14㎛
- 유해 화학물질과 가스를 조기 감지를 위해서는 중파장, 장파장 IR 영상 수집, 데이터셋 구축, 영상처리 및 AI분석 기술

< 다중 스펙트럼 영상 감시 및 분석 시스템 개요도 >

2 필요성

○ **(정부 지원 필요성)** 시민의 생활에 영향을 주는 주요 보안시설과 석유 및 가스 저장시설에 대한 무단 침입(월담)과 유해물질 및 가스 누출 사고 발생하고 있어 주요 보안시설 보안 감시와 안전 관리를 위한 핵심 기술 개발이 필요함

- 상황인식 정확도가 높고, 오감지률 낮은 지능형 CCTV 시스템 구현을 위해서는 다중 스펙트럼 영상분석 기술이 필요함
- 중화학공업 산업단지는 시설의 노후화 등으로 각종 유해 가스 누출 및 공해 물질 배출 사고가 발생하고 있어 산업단지 종사자와 시민 안전을 위한 관제 시스템이 필요함
  * 환경부, 대기환경보존법 시행규칙에 따라 불완전 연소 배출에 대한 기준 제시

○ **(기술성)** 종래의 가시광 영상(CCTV) 위주의 영상분석 기술이 딥러닝 기반으로 급속하게 발전하고 있으나, 다중 스펙트럼(중, 장파장 열화상) 영상에 대한 데이터셋 구축 및 분석 기술은 국내에서는 시작 단계임

- 글로벌 기술 경쟁력을 강화하기 위해서는 다중 스펙트럼 영상 수집, 데이터셋 구축, 인공지능 기반의 분석 기술 개발이 필요함

- 다양한 분야에 적용할 수 있는 OGI(Optical Gas Imaging) 기술은 세계적으로 미국, 프랑스 등의 제한한 국가 및 기업에서 점유하고 있어, 연구 개발을 집중함으로써 관련 분야를 선도할 수 있는 기회를 확보할 수 있음

* FLIR(미국), ULIS(프랑스)社는 센서 및 카메라 위주로 개발 공급하고 있음

○ **(경제성)** 주요 보안시설, 해안·해상 감시, 산업단지의 효율적인 보안 감시와 인체 유해 가스 누출을 신속하게 감지하기 위한 솔루션이 고가(수천만원 ~ 수억원/1세트) 이면서, 수입에 의존하고 있어 대체 기술 개발이 필요하고, 연구 개발을 통하여 글로벌 경쟁력을 확보할 필요가 있음

- FLIR,(미국), ULIS(프랑스)가 주도하는 외산 제품을 대부분 수입하여 사용하는 상황으로 실제 산업용으로 적용하는데 한계가 있음
- 특정 국가에서 주도하는 CCTV 카메라 및 분석 솔루션에 대응하여 특수 분야의 고부가가치 분야를 선도하는 것이 필요함
- 다중 스펙트럼 영상의 취득 및 분석 기술은 자동차, 로봇 분야에도 적용 가능성이 높아 기술 개발이 필요함

○ **(사회성)** 국내의 중화학산업단지의 노후화가 진행되고 있는 산업시설에 대한 유해 물질 누출 대응 및 개인정보의 유출 문제 없이 보안 서비스의 제공이 가능해 안전 사회 구현에 기여할 수 있음

3. 연구목표

○ **최종목표 : 주요 보안시설 및 산업단지의 보안 감시와 안전 관리를 위하여 가시광 및 중·장파장 IR 영상을 취득하고, AI 학습 데이터셋 구축하여 다중 스펙트럼 영상에 대하여 통합·연계 분석을 지원하는 다중 스펙트럼 영상분석 및 영상 보안감시 시스템 핵심 기술 개발**

- 다중 스펙트럼 영상 데이터셋 구축
- 다중 스펙트럼 영상 처리 및 분석 기계학습(딥러닝) 기술 개발
- 다중 스펙트럼 영상 통합 관제 솔루션 개발 및 상호 연동 실증
- 다중 스펙트럼 영상 수집을 위한 감지 장치 개발

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	다중 스펙트럼 영상 데이터셋	건	15,000건	-	200 (미국, FLIR)
2	다중 스펙트럼 영상 분석을 통한 이상행동 정확도/오감지율	%	95%/5%	-	- (미국, FLIR)
3	다중 스펙트럼 영상 분석을 통한 유해물질(가스) 감지 정확도	%	95%	90%	
4	IR 열화상 영상 해상도	화소	640x480	320×240	640x480 (미국 , FLIR)

○ 연차별 개발목표

구분	연도별 연구목표
2021년	VIS(가시광), 중·장파장 IR 영상 데이터셋 구축 IR영상 신호처리 및 분석 인공지능 모델 개발 VIS 및 IR 영상 정합 및 관제 솔루션 개발 다중 스펙트럼 영상 수집 장치 설계 및 개발
2022년	저조도 환경 VIS, 중·장파장 IR 영상 데이터셋 구축 VIS, 중·장파장 IR 영상 처리 및 분석 AI 모델 개발 IR 영상 통합 관제 솔루션 개발 다중 스펙트럼 통합 영상 수집 장치 설계 및 개발
2023년	시험 및 실증 다중 스펙트럼 데이터셋 구축 VIS, 중·장파장 IR 영상 분석 복합 인공지능 모델 개발 및 시험 VIS, IR 영상 통합 관제 솔루션 고도화 다중 스펙트럼 카메라 시작품 개발 및 실증

4. 연구내용

○ 개발 기술 내용

① 다중 스펙트럼(VIS 및 중·장파장) 영상 데이터셋 구축
- 저조도 및 다양한 환경의 이상행동 영상 수집
- 유해 물질 및 가스 누출 VIS 및 중·장파장 IR 영상 수집
- AI 학습 및 시험용 VIS 및 중·장파장 IR 영상 데이터셋 구축

② 다중 스펙트럼 영상분석 기술
- VIS 및 중·장파장 IR 영상 개선(enhancement) 기술
- VIS 및 중·장파장 IR 영상 분석 머신러닝(딥러닝) 기술
- 다중 채널 영상 분석 기술

③ 다중 스펙트럼 영상 통합 관제 솔루션 기술 및 구현
- VIS 및 IR 영상 통합 관제 기술
- VIS 영상 관제 시스템 호환성 고려한 상호 연동 기술

④ 다중 스펙트럼(VIS 및 중·장파장) 카메라 설계 및 구현
- VIS 및 중·장파장 IR 센서 제어 기술
- 중·장파장 IR 카메라 필터 및 렌즈 설계 기술
- 다중 스펙트럼 영상 처리, 합성, 전송 모듈 설계 및 제작 기술

○ 기존 (보유)기술

① VIS 및 IR 열화상 영상 분석 솔루션 및 데이터셋
- ADAS 및 보행자 데이터셋(가시광 및 장파장 IR 열화상 카메라 영상 데이터셋)
- IR 파장대별 OGI(Optical Gas Imaging) 카메라 및 가시화 기술
- 단일 종류의 유해가스(메탄 가스) 감지를 위한 중파장 IR 영상처리 및 딥러닝 영상 분석 기술
- IR 파장대별 영상 처리 기술 기반 영상 분석 솔루션

② VIS 및 IR 열화상 영상 감시 시스템 기술
- 냉각 방식 국방용 보안 감시 영상 감시 기술
- 체온 측정 및 사람 감지용 장파장 열화상 감지 기술

5. 지원기간/예산/추진체계		
ㅇ **기간** : 3년 이내 ㅇ **정부출연금** : ‘21년 12억원 이내 (총 정부출연금 42억원 이내) ㅇ **주관기관** : 제한없음(산업체 참여 필수)		
기술분류	대분류(차세대보안) - 중분류(물리보안) - 소분류(CCTV 감시/관제)	
연구유형	기초연구 ( ), 응용연구 ( ), 개발연구 (O)	**TRL** (5) ~ (7)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-20	(지정공모형)
과제명	무인점포 환경 대응형 2D/3D 영상 통합 분석기반 지능형 영상보안시스템 기술 개발	

1. 개념

○ 2D와 3D Depth 영상을 연계·분석하여 가려짐·혼잡·겹침 상황에서 객체 행위 중심의 이상행위를 인지하는 무인점포환경 대응형 영상보안시스템 기술 개발

< 무인점포환경 대응형 영상보안시스템 기술 개념도 >

2 필요성

○ **(정부 지원 필요성)** 코로나 사태 등으로 인해 비대면 무인점포 등의 확산이 가속화 될 것으로 예상되며, 무인점포 환경에서의 사고·범죄 예방, 안전 서비스 강화 및 범죄 등에 효율적 대응 등 사회적 요구에 대한 해결책 확보 필요

* 중소벤처기업부는 비대면 소비 확대 등 유통환경 변화에 대응하여 동네슈퍼의 스마트화를 지원하는 '스마트슈퍼 구축사업'을 '20년 시범사업을 시작으로, '21년부터 본격적으로 확대할 예정

○ **(기술성)** 개인정보의 유출 등의 이슈를 최소화 하면서 무인점포 보안에 활용 가능한 3D 정보기반 영상보안시스템 기술 확보를 통해 기술 경쟁력 강화 및 국외 기술 대비 우위 선점이 가능

- 기존 2D 영상만을 적용하는 영상보안시스템에 3D Depth 영상 기반 영상 보안 기술을 추가하여 2D 영상 기반 기술적 한계를 극복할 필요가 있음

* 3D Depth 영상은 2D 영상과는 다르게 개인의 사생활을 침해하지 않고 사람이나 사물을 입체적으로 표현함으로써 개인영상정보 보호가 가능함. 또한, 고도의 정밀한 거리 데이터를 통해 복수의 사람들이 상호 교차하거나 겹치는 환경에서도 이동경로 추적이 가능하여 신뢰성이 높음

○ **(경제성)** 2D/3D 연계 기반 지능형 보안시스템 솔루션 확보를 통해 중국, 미국에 비해 상대적으로 낮은 기존 CCTV 영상보안시스템의 산업 경쟁력을 강화하고 신 시장 창출 및 선점이 가능

3. 연구목표

○ **최종목표 : 가려짐, 혼잡, 겹침 등의 상황이 빈번하게 발생하는 무인점포환경에서 2D/3D 영상을 통합 신경망 기반으로 분석하여 이상행위를 인지하는 지능형 영상보안시스템 핵심기술 개발**

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	2D/3D 연계 검출 대상 이상행위 종류	종	5*	-	-
2	2D/3D 연계 이상행위 인지 정확도	%	90**	-	-
3	Multiple Object Tracking 정확도	%	80	-	61.8*** (미국, MS)
4	표준화	개	국내표준기고서(6) /국내표준제정(1)	TTA PG427 (표준화 기관)	-

* 활용 및 수요처의 의견을 반영하여 이상행위 5종이상 인지를 목표 성능 설정

** 2D/3D 영상을 연계하여 이상행위를 인지하는 기술에 대한 개발 사례가 없어, KISA CCTV 성능 시험 평가 인증 기준으로 목표 성능 설정

*** MOT20 데이터셋 기준의 Multiple Object Tracking 정확도(MOTA)

○ **연차별 개발목표**

구분	연도별 연구목표
2021년	무인점포 환경 기반 2D 영상보안시스템의 이상행위 인지 기술 고도화 개발 및 3D 분석 기반 객체 인식/추적 및 이상행위 인지 기술 개발
2022년	2D/3D 영상 기반 이상행위 인지를 위한 통합 신경망 기반 영상보안시스템 개발
2023년	무인점포 기반 2D/3D 영상보안시스템의 기술 활용성 검증 및 테스트베드 실증

4. 연구내용

○ **개발 기술 내용**

① **(무인점포환경에서 이상행위 인지)** 2D/3D Depth 영상기반 이상행위 인지 기술 개발

- 무인점포환경 기반 이상행위 정의
- 무인점포환경에 특화된 2D 영상보안시스템의 이상행위 인지기술 고도화 개발
- 혼잡·겹침 상황에 강인한 객체 및 이동 동선 감지 기술 개발
- 2D 영상/3D depth 정보기반 이상행위 인지 통합 신경망 기술 개발

② **(무인점포대응 통합형 영상보안시스템)** 무인점포 대응형 2D/3D 영상보안시스템

기술 개발

- 2D/3D 영상 취득 및 관리 기술 개발
- 무인점포 환경에 최적화된 2D/3D 통합형 영상보안시스템 개발
- 2D/3D 통합형 영상보안시스템에서 메타데이터 기반 검색 및 이상행위 인지 연계 지능형 보안서비스 개발
- 효율적인 하드웨어 자원 사용 기술 개발
- 무인점포 기반 2D/3D 영상보안시스템의 기술 활용성 검증 및 테스트베드 실증
- 서비스 모델 정의 및 서비스 모델 구축

③ **(이상행위 데이터셋 구축)** 2D/3D 영상기반 이상행위 데이터셋 구축 및 공개

- 무인점포환경에 특화된 2D/3D 영상기반 이상행위 학습 및 테스트 데이터셋 구축 및 공개
  * 레이블링된 학습 데이터셋 구축 규모 : UCF-Crime 데이터셋 기준 이상행위 종류별 200개 이상의 영상 데이터셋 구축 및 학습데이터 활용에 용이한 레이블링 구조 개발
  * UCF-Crime 데이터 셋: 폭행, 방화, 싸움, 공공 기물 파손 등 13종의 행위로 구성되며, 총 1900개의 영상 데이터셋 구축

④ **(영상보안시스템 연동 표준화)** 영상보안시스템 연동을 위한 인터페이스 규격 정의 및 표준화

- 2D/3D 영상 및 데이터 연동을 위한 인터페이스 규격 정의 및 표준화

○ 예상 수요처 : 편의점, 카페 등의 무인점포 및 비대면 서비스 매장

○ 기존 (보유)기술

① CCTV 영상 기반 객체 탐지 및 이상행위 인지 기술

- 사람, 차량 등 객체 검출, 인식, 추적 기술
- 기물파손, 쓰러짐, 폭행, 방화 등 이상행위 인식 기술

② 영상보안시스템 관련 인터페이스 및 메타데이터 규격 표준 기술

5. 지원기간/예산/추진체계

○ 기간 : 3년 이내

○ 정부출연금 : '21년 12억원 이내 (총 정부출연금 44억원 이내)

○ 주관기관 : 제한없음(산업체 참여 필수)

기술분류	대분류(차세대보안) - 중분류(물리보안) - 소분류(CCTV 감시/관제)	
연구유형	기초연구 ( ), 응용연구 ( ), 개발연구 (O)	**TRL** (4) ~ (7)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-21	(지정공모형)
과제명	신뢰 가능한 엣지AI 시스템 검증 플랫폼 핵심기술 및 시험기술 개발	

1. 개념

ㅇ **(목표)** 엣지AI 시스템*의 취약성을 평가하고 방어하여 신뢰 가능한 엣지 AI 시스템을 구축할 수 있는 핵심기술 및 이를 실증할 수 있는 검증 기술 개발

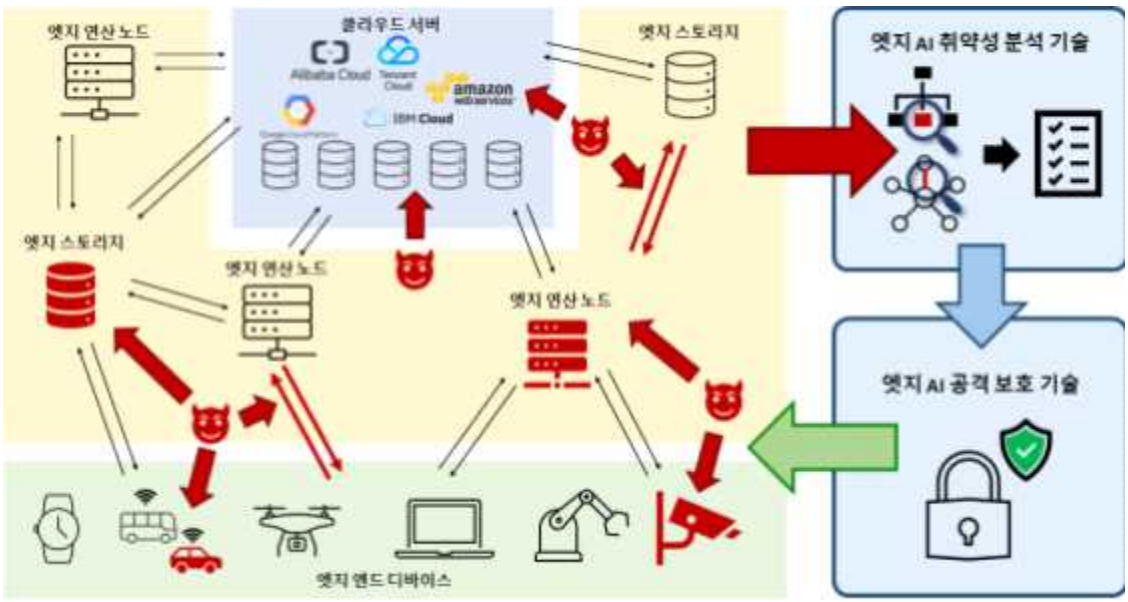
- 엣지AI의 특성에 의해 발생되는 취약점을 자동 분석하고 이러한 취약점 공격으로부터 엣지AI 시스템을 보호하여 신뢰성 있는 AI 시스템을 구현하는 기술 개발
- 개발된 취약점 분석 기술 및 공격 보호 기술을 검증하고 엣지 AI 시스템의 신뢰도를 평가할 수 있는 통합 지표와 검증 플랫폼 기술 개발

ㅇ **(개념)** *엣지AI 시스템이란 중앙화된 클라우드 서버, 로컬 엣지 컴퓨팅 노드, 그리고 엔드디바이스가 협업하는 가운데, 특히 클라우드의 개입을 최소화하고 엣지 노드와 엔드디바이스를 활용하여, 모델 학습 및 운용 전반에 걸친 비용, 응답성, 확장성, 정보보호 등에서 최적화를 이루기 위한 AI 증강 시스템을 지칭한다.

ㅇ **(문제점)**

- 엣지AI 환경은 다양한 요구조건 최적화를 위해 AI 모델 학습과 운용 과정에서 데이터 보관 및 전송과 연산 작업이 다양한 채널과 구성 요소에 분산되어 이루어짐. 5G/B5G/6G 고도 통신 기술 및 스마트 서라운딩 기술과 함께 지속 발전. 학습 데이터 오염, DoS, SW/HW 부채널 공격, API 접근, 하드웨어 근접 등 기존 클라우드 기반 AI 시스템과 비교하여 더욱 다양하고 취약한 공격 표면이 존재함.
- 다양한 형태의 서비스 제공 업체 및 이종 장비들로 이루어진 체계를 이룰 수 있으므로 구성요소별 취약성 분석과 구성 요소 간 상호 신뢰성을 확보해야 함
- 개별적인 취약점 분석 대응으로 이루어지던 AI 보안 기술을 엣지AI에 적합하게 개발하고, 통합적으로 적용하여 수많은 노드가 참여하는 엣지AI 시스템의 신뢰성을 종합적으로 평가하며 새로운 위협에 신속히 대응할 수 있는 체계가 필수적임.

<개념도>



2 필요성

○ **(정부 지원 필요성)** 엣지AI는 초연결 시대의 핵심 전략 기술로 국내외 연구가 활발히 진행중이므로 이를 선도할 수 있는 핵심기술 및 시험기술 개발에 대한 국가적 선제 지원이 필요

- 엣지AI 시스템을 위한 국내외 연구 및 개발이 활발
  - 삼성의 온디바이스 AI 전략, 퀄컴의 모바일 AI 전략
  - 애플의 CoreML, 페이스북의 Cafe2Go 등 엣지AI S/W, H/W 기술
  - 인텔의 IoT개발자를 위한 Intel Edge AI 나노학위과정 개설 (2020.4)
- 신뢰 가능한 AI 시스템의 연구개발 필요성 증대
  - 미국 DARPA의 GARD 프로젝트 (2019~)
  - 미국 NSTC 네트워킹/정보기술 R&D 및 ML/AI 커미티 보고서에서 미션크리티컬 응용을 위한 신뢰 가능 AI 시스템의 필요성 제고 (2020.03)
  - EU의 신뢰 가능한 AI의 윤리 강령 제정 (2018.12)
  - 미국 MITRE 의 Adversarial ML Thread Matrix는 머신러닝 시스템 대상 공격 방법을 정리. 알려진 29가지의 공격 방법 중 16가지는 기존 공격 방법과 달리 머신러닝 시스템에 특화된 공격 (2020.10)

○ **(기술성)**

- 엣지AI 시스템에는 센싱, 전처리, AI 모델, 학습, 출력 가공, 물리/사이버 세계로의 피드백 등 다양한 공격 표면이 존재
  - 요소 별, 요소 간 결합에 대한 취약성 분석 및 신뢰성 보장 필요
- 엣지AI 시스템에는 엣지디바이스의 소형화, 임베디드화, 분산화 등 기존 클라우드 기반의 AI 시스템과 차별화된 기술적 접근 필요
  - NVIDIA의 Jetson, Google의 Edge TPU 등 AI 전용 하드웨어 및 TensorFlow Lite, TensorRT와 같은 엣지AI용 S/W 프레임워크 등 기존 시스템과 차별화되는 환경에 대한 취약성 분석 및 신뢰성 보장 기술이 필요

○ **(경제성)**

- 엣지AI 시스템의 신뢰성 검증을 통해 사후 공격 대응과 복구에 필요한 막대한 경제적 손실을 줄이고 국제적 시험 인증이 가능한 검증 플랫폼 환경 구축 기반 마련
- 엣지AI 시스템의 신뢰성 증가를 통해 의료, 금융, 교통, 안전 등 다양한 산업군으로 확장 가능성 증대 (스마트팩토리, 커넥티드카, 5G/6G, IoT 등 도메인 확대)

3. 연구목표

○ **최종목표 : 엣지AI 시스템의 취약성을 평가하고 방어하여 신뢰 가능한 엣지 AI 시스템을 구축할 수 있는 핵심 기술 및 이를 실증할 수 있는 검증 기술 개발**

○ 정량적 개발목표

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	엣지AI 위협 분석을 위한 통합적 정량화 지표 개발 주1)	개	1	-	-
2	자동화된 엣지AI 취약점 분석 기술 개수	개	≥ 5	-	-주2) 미국, NYU주3) 미국, Iowa State Univ.주4)
3	방어가능 공격모델 종류	개	≥ 5	-	-주2) 미국, Facebook주5) 미국, Google주6)
4	동시 평가 가능 노드 수주7)	개	≥ 10	-	-

주1) 엣지AI 시스템에 대한 위험도를 판단하기 위해 세부 평가 항목들을 개발하고 종합하여 정량적인 위험도 수준을 나타내는 스코어로 도출하기 위한 통합 지표 개발

주2) 공격 모델별 개별 검증 및 방어기술은 다수 존재하나 통합 플랫폼 개발은 미진하며 특히 엣지AI에 적합한 취약점 분석 기술은 아직 기초적인 수준

주3) 참고자료: “SafetyNets: Verifiable Execution of Deep Neural Networks on an Untrusted Cloud,” NIPS 2017

주4) 참고자료: “Stealing Hyperparameters in Machine Learning,” SP 2018

주5) 참고자료: “Feature Denoising for Improving Adversarial Robustness,” CVPR 2019

주6) 참고자료: “Adversarial Logit Pairing,” arXiv: 1803.06373

주7) 클라우드 서버, 엣지 컴퓨팅 노드, 엔드 디바이스를 포함하여 엣지 AI시스템에 참여하는 노드의 수

○ 연차별 개발목표

구분		연도별 연구목표
1단계	2021년	엣지AI 시스템 위협 모델링 및 위험도 평가지표 도출 기술 개발 (다양한 엣지노드, 엔드디바이스 반영, SW 및 HW 선택 및 실험 연구 환경 구축)
	2022년	엣지AI 구성 요소별 취약점 자동 분석 기술 설계 및 구현 (엣지노드, 엔드디바이스의 SW 및 HW 취약점 분석)
	2023년	엣지AI 구성 요소별 취약점 분석 및 대응 실험 (1단계 위험도 평가지표 도출)
2단계	2024년	엣지AI 구성 요소간 취약점 분석 기술 설계 및 구현 (엣지노드, 엔드디바이스, 클라우드서버간 통신 및 협업 취약점 분석)
	2025년	엣지AI 구성 요소간 취약점 분석 및 대응 실험 (2단계 위험도 평가지표 도출)
	2026년	엣지AI 취약점 공격 대응 기술 개발 및 검증 플랫폼 구축 (전 단계 취약점 분석 기술 및 실험 환경 기반, 종합된 위험도 평가지표 도출)

4. 연구내용

ㅇ 개발 기술 내용

① 엣지AI 취약성 분석 기술
- 엣지AI 시스템 위협 모델링 기술 (공격 표면 도출 및 위험도 평가 지표 도출)
- 학습 데이터 획득(센싱), 전처리, AI 모델, 학습, 추론, 출력 가공, 물리/사이버 세계로의 피드백 등 각 구성 요소 및 구성 요소간 취약성 분석 기술
- AI H/W, S/W 프레임워크 및 응용 S/W의 취약성 분석 기술
- 엣지AI 학습모델의 오류유발 샘플 자동탐지를 위한 환경 구축과 퍼징 기술
- 온라인 학습 및 연합(federated) 학습 시 엣지 환경에서 이루어지는 학습 과정의 취약성 분석 기술
- 엣지AI 시스템의 취약성을 자동으로 분석/시험하기 위한 S/W 프레임워크 기술

② 취약점 공격 대응 기술
- 데이터 오염, 신뢰되지 않은 구성 요소, 통신 오류/변조/지연/감청, 쿼리 기반 공격, 디바이스 내 공격 등 엣지AI 운용 특성에 따른 취약점을 보완하기 위한 H/W 및 S/W 기반 기술
- 모델 탈취, 성능 저하, 추론 결과 변조, 오작동 유도 등 다양한 공격 목표에 대응 가능한 기술군 개발
- 취약점 공격에 강건한 AI 모델 다양화(ensemble 등) 및 실용 가동 상태 유지를 위한 AI 모델 다중화 기술 등 AI 알고리즘 기반 취약점 대응 기술
- 공격자의 시스템 접근 가능 정도(학습 과정 개입, 물리적 디바이스 접근, 외부 입력 조작)에 따른 방어 기술 체계화
- 엣지AI 시스템 장애 및 공격 여부를 실시간으로 판단 가능한 신뢰성 감독 기술

③ 신뢰가능 엣지AI 시스템의 시험 평가를 위한 검증 플랫폼 구축
- 실용 도메인(예: 스마트팩토리, 커넥티드카, 5G, IoT 등)을 대상으로 하여 엣지AI 시스템의 취약성 분석, 신뢰성 시험 평가 및 지표 도출이 가능한 검증 플랫폼 구축

ㅇ 기존 (보유)기술

① 클라우드, 엣지 컴퓨팅 노드, 엔드 디바이스간 협업을 통한 AI 모델 배포 및 연산 기술

② 학습모델의 취약성 분석 자동화 기법 연구 중
- 컬럼비아대학, CMU 등: 다양한 입력 변이를 통해 자동화된 취약성 검증을 수행하지만 생성 가능한 입력 범위가 제한적

③ 학습모델 취약성 공격과 그 대응 기법 연구 중
- MIT 등: 정상 샘플의 분포에서 벗어난 입력과 오류샘플을 재학습하는 등 일부 공격 방어를 시도함

④ 머신러닝 모델 탈취 공격에 대한 대응 기법 연구 중
- UCSD: 모델 소유권 증명을 위한 워터마킹 기법 등

⑤ AI 시스템에서의 개인 정보 보호를 위한 기술 연구 중
- Google: 학습 과정에서의 보호를 위한 차등정보보보호 기술 등
- Intel: 추론 과정에서의 보호를 위한 동형 암호화 기술 등
⑥ AI 시스템에 특화된 하드웨어 기반 신뢰실행환경 기술 개발 중
- Microsoft: GPU기반 신뢰실행환경 기술
- KAIST: 기존 CPU 기반 신뢰 실행환경의 GPU 확장 기술

5. 지원기간/예산/추진체계

ㅇ **기간** : 6년 이내 (1단계 3년 → 2단계 3년)

ㅇ **정부출연금** : '21년 4억원 이내 (총 정부출연금 29억원 이내)

ㅇ **주관기관** : 대학

기술분류	대분류(차세대보안) - 중분류(공통기반 보안) - 소분류(인공지능 보안)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (2) ~ (5)
과제특징	정책지정( ), 혁신도약형(O), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-23	(지정공모형)
과제명	AI·빅데이터 기반 개인정보 노출·불법 유통탐지 기술 개발	

1. 개념

○ 다변화되는 개인정보 유형·패턴 및 음성화되는 불법유통 행태 등을 자동으로 학습하여 대응 가능한 AI·빅데이터 기반 개인정보 노출·불법유통 탐지 기술개발

2 필요성

○ **(정부 지원 필요성)** 최근 데이터 3법 통과와 제4차 산업혁명으로 온·오프라인이 융합된 초연결 지능 정보 사회의 도래로 디지털 경제 가속화

- 정부는 D.N.A.(Data-Network-AI) 생태계, 비대면 산업 육성 등에 집중투자 하는 '디지털 뉴딜'을 추진
- 데이터 결합, 분석, 활용 증가로 개인정보 유·노출 등 침해사고의 위험성 증가

○ **(기술성)** 데이터 활용 확산으로 인하여 개인정보 유·노출이 증가하고 있으나, 개인정보 노출·불법유통 탐지 사각지대*가 지속 발생하여 국가 차원의 대응이 필요

* 비정형(이미지 등) 개인정보·불법유통 게시물 노출 및 음성화되는 불법유통 행태 등

- AI, 빅데이터 등 신기술을 적극 활용하여 새로운 유형의 개인정보 노출·불법유통 행태에 대한 선제적 대응 필요

○ **(경제성)** 개인정보 유·노출의 2차 피해(보이스 피싱 등) 발생에 따라 국가 차원의 대응이 필요한 상황

- 개인정보 노출·불법유통을 선제적으로 대응하여, 각종 범죄로부터 국민의 피해(경제적·물질적·정신적) 방지 기반 마련
- 신기술(AI 등) 대체 시 비용의 최소화 등 개인정보 노출·불법유통 대응 효율화

3. 연구목표

○ **최종목표 : 데이터 활용 증가에 따른 개인정보 유·노출 피해 예방을 위하여 AI·빅데이터 기반 개인정보 노출·불법유통 탐지 기술개발**

- **(지능형 크롤러)** 탐지이력, 유노출 동향 등 빅데이터 분석·학습을 통해 키워드 추천, 탐지패턴 자동화 등 지능형 크롤러 개발
- **(스코어링)** 개인정보 노출·불법유통 게시물에 대한 빅데이터 기반 스코어링(Scoring) 모델 개발
- **(예측 모델링)** 지능형 크롤러 및 스코어링 기술 등을 기반으로 웹사이트 내 개인정보 노출·불법유통 사전 예측 모델링 개발
- **(음성화 대응)** 음성화된 사이트 대상 개인정보 노출 및 불법유통 대응 프레임워크(수집/색인/검색/분류/분석) 기술 개발

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	게시물 분류 정확도[주1)]	%	80%	-	-

주1) (개념) 마신러닝 기반 분류 모델을 통해 웹사이트 내 개인정보 노출·불법유통 게시물을 분류한 확률
(측정 방법) 전체 탐지 게시물 분류 값(TP, TN, FN, FP)* 중 실제 개인정보 노출·불법유통 게시물을 분류한 경우(TP, TN)
* TP(True Positive), TN(True Negative), FN(False Negative), FP(False Positive)

$$\text{게시물 분류 정확도} = \frac{TP+TN}{TP+FN+FP+TN}$$

○ **연차별 개발목표**

구분	연도별 연구목표
2021년	데이터 기반 자동 학습 기능 및 지능형 크롤러 개발
2022년	스코어링 모델 및 예측 모델링 개발
2023년	음성화된 사이트 내 개인정보 노출·불법유통 대응 프레임워크 기술 개발

4. 연구내용

○ **개발 기술 내용**

① 데이터 기반(개인정보 노출·불법유통 동향, 탐지이력 등) 자동 학습 기능 및 지능형 크롤러 개발
- 자동 학습 기능을 적용한 AI 지능형 크롤러 개발
- AI 기반 분류 모델 등을 통한 개인정보 노출·불법유통 게시물 자동 분류 기술 개발
- 빅데이터 기반(탐지이력·유형 등)의 정형·비정형데이터 관리·분석 기술 개발

② 개인정보 노출·불법유통 게시물에 대한 빅데이터 기반 스코어링(Scoring) 모델 개발
- 스코어링을 위한 자연어 처리(형태소 분석) 기능 개발
- 웹 페이지 내 단어 빈도 및 연관성 분석 등 기술 개발
- 비정형·정형 데이터의 메타데이터 추출·분석 기술 개발
- 새로운 유형의 탐지 키워드, 패턴 등 추천 기술 개발

③ 개인정보 노출·불법유통 사전 예측 모델링 개발

- 개인정보 영향도, 침해발생 정도, 노출 정도에 따른 위험도 측정 모델링 개발
- 웹사이트의 개인정보 노출·불법유통 위험도 자동 측정 기술 개발
- 예측 모델링 기반으로 개인정보 노출·불법유통 사전 예측 기술 개발

④ 음성화된 사이트 대상 개인정보 노출 및 불법유통 대응 프레임워크(수집/색인/검색/분류/분석) 기술 개발
- 음성화된 사이트 대상 콘텐츠 수집용 웹로봇 개발
- 실시간 변동 등 음성화된 사이트 특성에 대응 가능한 웹 아카이빙(캐싱) 기술 개발
- 음성화 콘텐츠에 대한 데이터 저장/색인/검색/분류 체계 개발
- 불법유통 게시자 추적을 위한 음성화 사이트 및 서피스웹(Surface Web) 간 연관성 분석 기술 개발
- 음성화 사이트 내 개인정보 유출 현황(유출 사업자, 시점, 건수 등) 파악을 위한 데이터 분석 기술 개발

ㅇ 기존 (보유)기술

① 국내외 포털을 중심으로 서피스웹 상의 다양한 정보를 자동으로 검색하고 색인할 수 있는 크롤링 기술 보유
- 정형(텍스트)·비정형(이미지 등) 데이터에 대한 저장·색인·검색·분석 기술

② 국내외 IT기업 중심으로 빅데이터 수집·정제·전처리·분석 기술 보유
- 정형·비정형 데이터에 대한 수집 및 저장
- 연관성 분석, 희귀분석, 의사결정나무 등 다양한 분석 기술

5. 지원기간/예산/추진체계

ㅇ **기간** : 3년 이내

ㅇ **정부출연금** : '21년 12억원 이내 (총 정부출연금 42억원 이내)

ㅇ **주관기관** : 제한 없음

기술분류	대분류(차세대보안) - 중분류(데이터보안) - 소분류(프라이버시 보호)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (4) ~ (6)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( ),	

관리번호	2021-차세대보안-24	(지정공모형)
과제명	온라인 중고거래 등 신종 비대면 사이버 사기 탐지·추적 및 피해 예방 플랫폼 기술	

1. 개념

○ 온라인 중고거래, 메신저 피싱 등 비대면 사이버 사기를 조기 탐지 및 예방하기 위하여 감성 분석(Sentiment Analysis)* 기반 사기의도 분석·탐지 및 비대면 거래 피해 예방 플랫폼 기술 개발

* 텍스트에 나타난 사람들의 태도, 의견, 성향과 같은 주관적인 데이터를 분석하는 자연어 처리 기술

< 개념도 >

2 필요성

○ **(정부 지원 필요성)** 코로나 사태로 인한 비대면 거래가 활성화됨에 따라 다양한 경로를 통해 거래가 이용되고 있어, 포스트 코로나 시대를 대비하기 위한 안전한 환경 구축의 기술 개발 지원이 필요

- 기존 거래 환경은 사이버 환경과 직접 대면 거래 등이 함께 사용되면서, 신뢰성 측면을 고려한 환경 유도가 용이했으나, 비대면 거래가 활성화되면서 안전성 확보가 어려워짐
- 다양한 경로로 사이버환경을 이용한 거래가 발생하고 있으나, 기업에서 자체적으로 제공하는 서비스는 모든 환경을 고려하지 못함

* 온라인 게시글, 메신저/메일을 이용한 직접적인 접근 등 사용자가 위험한 환경에 노출될 수 있으나, 기업이 이에 대한 대응 어려움

○ **(기술성)** 포스트 코로나를 대비하기 위해서는 비대면이 주요한 이슈인 사이버 환경에서 발생하는 안전한 거래 환경 조성 및 사용자 주의 환기 등의 기술 개발이 시급

- 국내 헬로마켓은 '20년 10월부터 모든 거래를 100% 비대면으로 전환한다고 하였으며, 번개장터와 S2W랩이 AI 기반으로 사기 거래 유도 패턴을 인식해 차단하는

기술 협약을 체결하는 등 비대면 거래를 위한 구조 변화가 진행 중

* 기존 안전거래 시스템을 통한 비대면 거래로 안전성을 확보하고자 하나, 가짜 안전거래 사이트를 꾸며 이용하는 등 사기거래는 여전히 발생되고 있음(노컷뉴스, '20.07)

- 사회문제 해결형 과제, 기술로드맵 상 사이버범죄 예방 및 추적기술 개발 등을 통해 국가 주요 기술개발 추진전략을 세우고 추진하고 있음(ICT R&D 기술로드맵 2023, '18.12)

○ **(경제성)** 사이버상의 비대면 거래 증가로 사용자들을 위한 안전한 거래 환경 구성은 피해 예방 및 금전적 손실 방지 등을 위한 필수적인 기술로서 부상하고 있음

- 사이버사기로 인한 피해는 사용자들에게 직접적으로 영향을 발생시키고 있으며, 추적이 어려운 특성으로 피해자 검거 등에 어려움이 높아 점차 증가 추세

* '19년 발생한 사이버 사기는 136,074건으로 전년(112,000건)보다 21.49% 증가, 사이버 사기로 검거된 인원은 31,331명으로 전년(28,757명)보다 8.95% 증가(경찰청, '20.02)

- 코로나 사태로 인해 비대면 거래의 중요성이 증가하고 있는 시점으로 포스트 코로나 시대를 대비하는 기술로서 경제성이 높고, 국내외 온라인 거래 시장을 선도할 수 있는 주요 원천기술 개발 및 선점이 필요함

3. 연구목표

○ **최종목표 : 신종 비대면 사이버 사기 탐지 및 안전한 거래 환경 구성을 위하여 Agent 기반의 사용자 안전 거래 환경 구축 및 감성 분석 기반 비대면 사기 거래 분석/탐지 기술 개발**

- 텍스트 감성 분석 기반 사이버 사기 의심 거래 탐지 기술 개발
- Agent 기반 사용자 비대면 거래 탐지/모니터링 기술
- 사이버 사기 거래 피해 예방을 위한 연관분석·모니터링 플랫폼 개발

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	감성 분석 기반 사기 거래 탐지율*	%	90%	-	-
2	초당 텍스트 감성 분석 처리 성능	구문 수/초	16,000	-	16,000 (미국, Talkwalker)

* 대부분 고객 만족도 조사 등의 감성 분석을 진행하고 있으며, 정확도를 명시하고 있지 않음. 또한, 사기거래 분야에 대한 분석의 사례가 없어, 기존 AI 기반 탐지율(MIT의 AI^2 탐지율 85% 등) 등을 고려한 도전적 목표 설정

○ **연차별 개발목표**

구분	연도별 연구목표
2021년	텍스트 감성 분석을 통한 사기거래 탐지 기술
2022년	Agent 기반 사용자 비대면 거래 탐지/모니터링 기술
2023년	사이버 사기 거래 피해 예방을 위한 연관분석·모니터링 플랫폼

4. 연구내용

ㅇ 개발 기술 내용

① 텍스트 감성 분석 기반 사이버 사기 의심 거래 탐지 기술 개발

- 사이버 거래 정보 분석을 통한 중고거래 관련 주요 키워드 추출 기술
- 온라인 중고거래 사이트, SNS 등 웹 검색을 통한 키워드 기반 중고거래 게시글 자동 수집 기술
- 주관성 탐지(Subjectivity Detection)*를 위한 사기거래 연관 텍스트 추출 기술
  * 감성을 포함한 단어와 이를 포함하지 않는, 즉 주관성이 없는 부분을 제외하는 단계로 작성자 이름, 성별 같은 개인정보를 제외하는 과정도 포함
- AI를 이용한 텍스트 극성 분석(Polarity Detection)* 학습 및 사기거래 탐지 기술
  * 주어진(추출된) 데이터가 긍정 혹은 부정인지를 판단하는 단계로, 단어가 나타나는 빈도, 단어의 속성에 따른 가중치 등으로 전체 텍스트의 긍정/부정 여부를 판단

② Agent 기반 사용자 비대면 거래 탐지/모니터링 기술

- Agent 기반 중고거래, 메신저/메일 어플리케이션 내 사기거래 의심 텍스트 자동 추출 기술
  * 어플리케이션 내 데이터 접근, 수신 데이터 접근에 대한 보안사항 고려 필요
- 추출 텍스트 대상 피싱URL, 사기거래 계좌, 연락처 등의 실시간 모니터링을 통한 위험 알림 기술
- 사기의심 거래 대상 자동 이력 정보 조회 및 알림 기술

③ 사이버 사기 거래 피해 예방을 위한 연관분석·모니터링 플랫폼 개발

- 대량의 사기 거래 정보(게시글, 계좌, 연락처 등) 대상 유형 및 위험도 분류 기술
- 사기거래 정보* 연계분석을 통한 범죄그룹 식별 및 사기거래 사전 탐지·모니터링 기술
  * 사기 거래 게시글, 거래물품, 거래방식, 계좌 등 수집된 사기거래 정보
- 사기 거래 및 범죄 그룹 관련 정보와 사기 의심 계좌(더치트), 연락처(사이버캅) 등 사기 거래 이력 연계를 통한 경고 서비스 및 정보 공유 인터페이스

ㅇ 기존 (보유)기술

① 다양한 중고거래 및 온라인 상거래 업체에서 안전거래 서비스 사이트 제공

- 네이버 페이 안전거래, 딜앱, 헬로마켓, 번개장터 등

② 경찰청 사이버캅, 더치트 등을 통한 사기거래 관련 계좌 및 연락처 이력 조회

5. 지원기간/예산/추진체계

ㅇ 기간 : 3년 이내

ㅇ 정부출연금 : '21년 7.5억원 이내 (총 정부출연금 27.5억원 이내)

ㅇ 주관기관 : 제한없음

기술분류	대분류(차세대보안) - 중분류(데이터 및 응용서비스보안) - 소분류(전자화폐·핀테크 보안)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (4) ~ (6)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

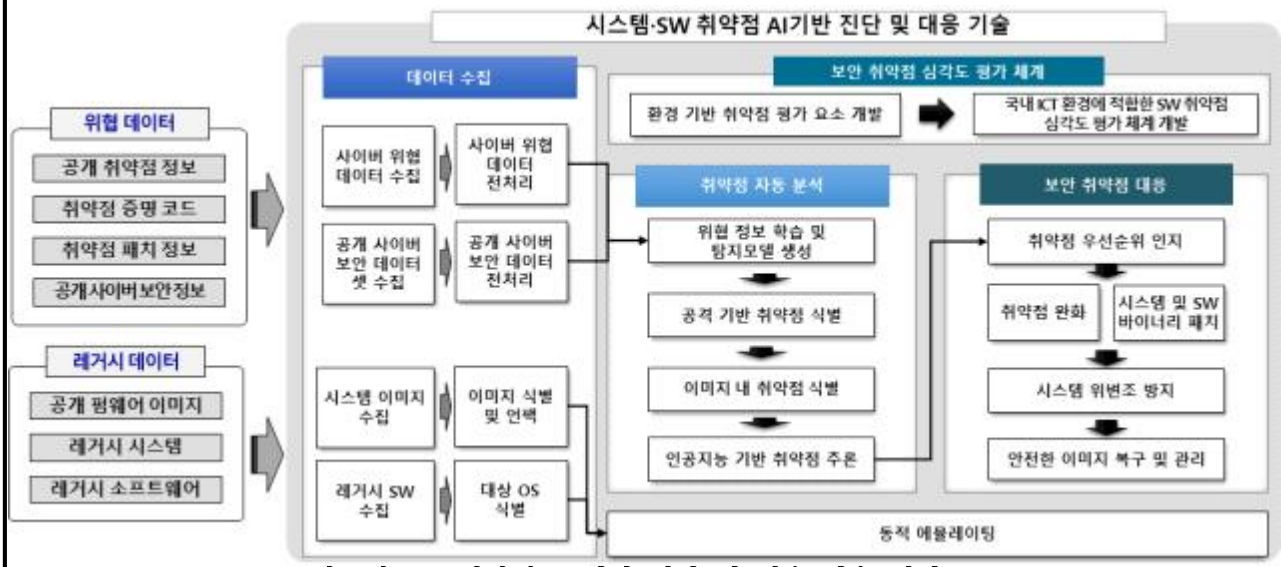
관리번호	2021-차세대보안-25	(지정공모형)
과제명	시스템·SW 취약점 AI기반 진단 및 대응 기술 개발	

1. 개념

○ 사이버공격의 원천적 차단이 힘들어지면서 위협에 대한 사전 예방부터 사후 대응까지 사이버 복원력*에 대한 중요성 부각

* 사이버 복원력은 사이버공격에 의한 시스템 및 서비스의 피해를 최소화하고, 장애/사고가 발생하기 이전 상태 혹은 그에 준하는 상태로 신속하게 돌아가는 역량을 의미함

○ 보안이 취약한 사물인터넷 기기 및 보안 업데이트 기술지원이 종료된 시스템의 커널, 라이브러리, 파일 시스템 및 서비스 취약점을 종합적으로 진단하고 악용/위협을 자동 방어하는 능동형 사이버방역 기술 개발 요구



< 시스템·SW 취약점 AI기반 진단 및 대응 기술 개념도 >

2 필요성

○ **(정부 지원 필요성)** 美 국방성(DARPA)는 보안 기술에 AI를 접목하기 위한 AI NEXT 프로그램에 연 20억달러 이상을 투자하면서 보안 기술력 확보 경쟁에서 우위를 점하고 있음(DARPA, AI Next Campaign)

- 국내 AI 기반 사이버 보안 제품 제작이 가능한 개발 선도 기업이 전무한 상황에서 사이버 복원력 확보 및 사이버방역 체계 구축의 일환으로 시스템/SW 취약점 대상 AI 기반 진단 및 대응 원천기술 확보를 위해 연구 경험 및 환경을 갖춘 기관의 주도가 필요함

* EU에서는 Foresight 프로젝트를 통해 항공·전력망·군 등에 대한 사이버 공격을 예방, 탐지, 대응 및 완화하기 위한 AI 기반 해킹시뮬레이션 플랫폼 개발을 추진

* 미 정보고등연구기획국에서는 자동화된 비정형 사이버공격 탐지 프로젝트 및 제로데이 취약점 탐색을 위한 컴퓨터-인간 협업 프로젝트(CHESS : Computer-Human Exploring Software Security) 등을 통해 능동형 대응 체계 구축 추진

○ **(기술성)** '16년 DEFCON에서 개최한 CGC대회 이후 소프트웨어의 취약점을 자동으로 탐지하고, 공격하기 위한 기술이 개발 중이지만, AI 기술 접목 및 시스템 레벨에서 취약점 탐지 및 자동 대응을 위한 기술 개발 미흡

- 현재 공개되어 있는 CRS(Cyber Reasoning System)는 특정 종류의 취약점에 대

한 자동 탐지 및 대응을 위주로 개발되어 있고 취약점 점검 대상 또한 SW 수준으로 시스템 및 펌웨어 레벨에서 대응은 불가능함

- 또한 취약점에 대응하기 위한 소스코드를 대상으로 하는 연구 및 제품은 많이 존재하지만, 바이너리 및 시스템 대상의 취약점 대응 기술은 연구 초기 단계임

○ **(경제성)** 인공지능 기반의 SVM*시장은 '20년(10억달러) 대비 '26년(36억달러)까지 약 3.6배 성장 예측('19년 Markets and Markets, Artificial Intelligence in Cybersecurity Market)

* SVM(Security & Vulnerability Management) 시장은 취약점에 대한 식별, 모니터링, 평가 및 패치를 포함한 전반적인 취약점 관리 시장을 의미함

- '17년 5월에 기술지원이 종료된 Windows XP 취약점을 악용하여 전 세계 150개국 30만대 PC를 랜섬웨어(워너크라이) 감염 등 대규모 피해 발생

3. 연구목표

○ **최종목표 : 보안 업데이트 기술지원이 종료된 시스템 및 IoT 기기의 커널, 파일시스템 및 서비스 취약점을 종합적으로 진단하고 악용 위협을 자동 방어하는 사이버방역 기술 개발**

- 시스템/펌웨어 원본 이미지 수집 및 분석 기술 개발
- AI 기반 보안 취약점 분석 자동화 플랫폼 개발
- 보안 취약점 심각도 평가·관리 체계 개발
- AI 기반 보안 취약점 대응 기술 개발

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	동적 에뮬레이팅 성공률	%	60%	-	41.18%[주1] (CMU, 미국)
2	학습에 활용된 공개 취약점 비율	%	99%	-	94%[주2] (Virginia Tech, 미국)
3	AI 기반 취약점 탐지율	%	95%	-	91%[주3] (Concordia Univ, Canada)
4	탐색된 취약점 대비 취약점 대응률	%	99%	-	97%[주4] (Shellphish, 미국)

주1) D.Chen et al, "Towards Automated Dynamic Analysis for Linux-based Embedded Firmware", 크롤러에 의해 수집된 시스템 및 펌웨어 원본 이미지를 대상으로 동적 에뮬레이팅에 성공한 이미지의 수를 평가, 해당 논문에서 23,035개 펌웨어 이미지 대상으로 9.486개 펌웨어에 대해 에뮬레이팅을 성공

주2) Jay Jacobs et al. "Improving vulnerability remediation through better exploit prediction", NVD를 기준으로 공개 취약점 DB에 등록된 취약점 수 대비 학습에 활용된 취약점 비율

주3) Paria Shirani et al. "BINARM: Scalable and Efficient Detection of Vulnerabilities in Firmware Images of Intelligent Electronic Devices", 펌웨어 대상 기존 취약점 대비 학습된 탐지 모델에 의해 탐지된 취약점 비율

주4) Cyber Grand Challenge 대회에서 취약 바이너리 대비 취약점 패치 및 완화 기술 적용 비율

4. 연구내용

○ **개발 기술 내용**

① 시스템/펌웨어 원본 이미지 수집 및 분석 기술 개발
- 제조사 배포 펌웨어 이미지, 지원 종료 OS 및 패키지 수집 기능
- 이미지 파일 학습 기반 압축 포맷 자동 식별 및 언팩(Unpack) 기능
- 가상화 플랫폼 기반 멀티 포맷 이미지 동적 에뮬레이팅 기능

② AI 기반 보안 취약점 분석 자동화 플랫폼 개발
- 시스템 이미지/SW 대상 동적·정적 분석 기반 특징 추출 기능
- 공개된 취약점, 개념증명코드(POC), 익스플로잇 수집 및 특징 추출 기능
- 취약점 데이터 기반 학습 및 취약점 탐지 모델 생성
- 시스템/SW 대상 AI 기반 보안 취약점 탐지 및 추론 기능

③ 보안 취약점 심각도 평가·관리 체계 개발
- 시스템/SW 구성 요소 식별·관리 및 SW부품표(SBOM) 자동 생성 기능
- 환경 기반 정성적 취약성 평가요소 선정 및 위협데이터 연계 취약점 영향도 분류
- 국내 ICT 환경에 적합한 SW 취약점 심각도 평가 체계 개발

④ AI 기반 보안 취약점 대응 기술 개발
- 보안 취약점 심각도 평가 체계에 따른 대응 우선순위 자동인지 기능
- 취약점 패치 학습 기반 시스템/SW 바이너리 패치 및 취약점 완화 기능
- 시스템/SW 패치 및 취약점 완화 모듈 대상 적합성 검증 기능
- 보호 기능이 탑재된 운영체제 이미지 파일 Repack 및 안전 배포 기능

ㅇ 기존 (보유)기술

① 자기학습형 사이버면역 기술
- 레거시 SW를 대상으로 바이너리 내 존재하는 취약점에 대해 퍼징 기술을 이용하여 탐색하고, 원인을 분석하여 패치를 수행하기 위한 기술

② IoT SW 보안취약점 점검 기술
- IoT SW를 대상으로 오픈소스에 존재하는 취약점을 학습하고 유사도 분석을 통해 SW 내 존재하는 취약점을 탐색하기 위한 기술

5. 지원기간/예산/추진체계

ㅇ **기간** : 4년 이내

ㅇ **정부출연금** : ‘21년 12억원 이내 (총 정부출연금 57억원 이내)

ㅇ **주관기관** : 제한없음

기술분류	대분류(차세대보안) - 중분류(시스템 및 암호보안) - 소분류(위협 분석 및 관제)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (3) ~ (6)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-기술분야-26	(지정공모형)

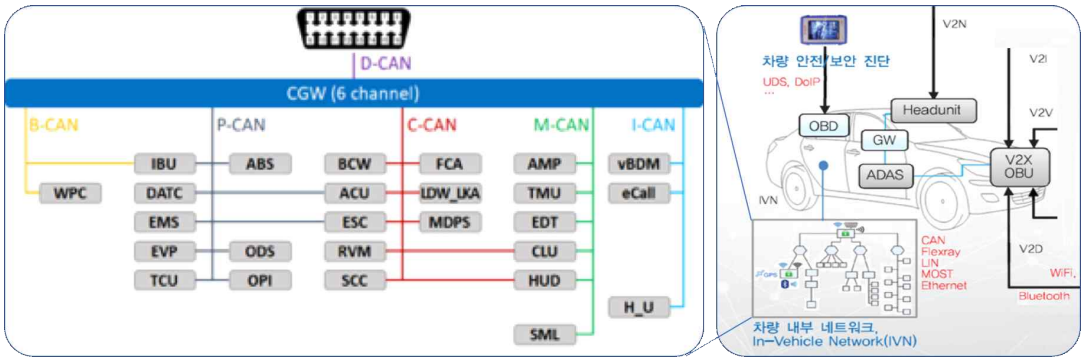
과제명	자동차 내부 네트워크의 보안 취약점 분석 기술 개발

1. 개념

○ 자동차 내부 네트워크 대상 사이버공격에 대한 우려가 현실화됨에 따라 자동차 내·외부 접점으로부터 내부 네트워크로 유입되는 악성 패킷을 발견하고 대응함으로써 자동차 내부 네트워크를 사이버공격으로부터 안전하게 보호할 수 있는 기술이 필요함

○ 자동차 내부 네트워크에 대한 공격 표면을 정의하고 이에 따라 공격에 활용 가능한 신규 취약점을 발굴하며 이를 실증할 수 있는 동적 분석 테스트베드의 개발이 요구됨

○ 취약점 발굴을 위한 동적분석과 발굴 취약점에 대한 공격 실증을 토대로 자동차 내부 네트워크 연관 자원에 대한 접근제어 기술 및 침입탐지 기술의 개발이 시급함



< 개념도 >

2 필요성

(정부 지원 필요성)

○ 정부는 일관된 정책을 통하여 미래자동차 산업에 지속적으로 투자하고 있음

- ('18.02) 「정부 R&D 투자 혁신방안」에서 자율주행차를 최우선 적용대상으로 선정하여 R&D PIE 모델 및 패키지형 연구개발전략 수립
- ('19.04) 「혁신성장을 위한 5G+ 전략」에서 10대 핵심산업으로 5G-V2X와 정보보안, 5대 핵심서비스으로 자율주행차 선정
- ('19.10) 「미래자동차 산업 발전 전략 2030 국가 로드맵」에서 전동화 기반, 자율주행 기능, 인프라, 제도 등의 집중 개선을 통하여 '24년까지 완전자율주행차를 상용화하고 '27년까지 핵심부품, 시스템 및 인프라 기술에 집중 투자하여 자율주행차 기술강국으로의 도약 추진
- ('20.10) 「지역균형 뉴딜 추진방안」에서 중앙정부, 지자체 및 민간이 공동으로 매칭하는 디지털 뉴딜 프로젝트를 통하여 10대 대표과제로 친환경 미래모빌리티 및 자율주행차 산업 생태계 조성 계획

○ 자동차 내부 네트워크에 대한 다양한 공격 가능성이 제기되거나 전문가 그룹에 의하여 그 가능성이 검증되고 있음에도 불구하고 보안 문제 대응을 위한 변화에 신속하지 못한 상황임

- ('15.08) Charlie Miller와 Chris Valasek은 "Remote Exploitation of anUnaltered Passenger Vehicle"을 통하여 모바일 네트워크를 경유하여 자동차 내부 네트워크로 악성 패킷을 주입하여 원격으로 자동차를 임의제어가 가능함을 입증함
- ('18.08) Sen Nie, Ling Liu , Wenkai Zhang, Yuefeng Du 등은 "Over-the-Air: How we Remotely Compromised the Gateway, BCM, and Autopilot ECUs of Tesla Cars"에서 WiFi에 연동된 자동차에 대하여 원격으로 익스플로잇이 가능함을 입증함

(기술성)

○ 자동차 내부 네트워크 자체의 보안문제를 포함하여 외부 접점으로부터 기인한 내부 네트워크로의 악성 패킷 유입에 의한 보안위협을 조기 발견하기 위해서는 외부 접점으로부터의 정보와 연계한 다양한 각도로의 메시지 분석이 요구됨

○ 신규 보안 취약점을 발굴하거나 기존 취약점에 대한 기술적 검증을 통하여 잠재적 보안 위협이 사고로 이어지는 것을 방지할 필요가 있음

○ 자동차와 관련한 사이버보안 강화 요구가 잠재적 수요에서 현실적 의무사항으로 변화함에 따라 이에 대한 긴급 대응이 필요함

- ('20.06) 유엔유럽경제위원회 UNECE의 WP.29의 채택에 따라 자동차 사이버 보안 규정이 2021년부터 발효될 예정으로 상용 자동차에 대한 기술적 대응이 요구됨
  * UNECE WP.29는 자동차 제조사가 사이버보안관리시스템을 구축하여 자동차의 보안 수준을 지속적으로 관리하는 것을 의무화함
  * UNECE WP.29는 ('22.07) 모든 새로운 차량 유형에 대하여, ('24.07) 모든 생산 차량에 대하여 사이버보안을 의무 적용하는 것을 규정하고 있음
- ISO21434의 표준화 완료에 따라 상용 자동차에 대한 사이버보안 대응기술의 적용이 가시화됨

(경제성)

○ ('20.03) McKinsey 보고서에 의하면 글로벌 자동차 보안 시장은 '20년 49억 달러에서 '30년 97억 달러로 급성장할 것으로 전망됨

○ ('16.09) IHS 오토모티브의 자동차 사이버 보안 관련 보고에 의하면 전 세계적으로 약 1억 1,200만대의 자동차가 네트워크로 연결되어 있으며 이를 기반으로 하는 자동차 사이버 보안 시장은 2023년에 7억 5,900만 달러로 성장할 것으로 예상함

○ 향후 자동차 내부 네트워크에 대한 보안기술 적용은 네트워크 자체의 진화와 함께 지속적으로 이루어질 것으로 예상되며 이를 위하여 자동차 부품사와 제조사의 지속적 투자가 이어질 것으로 전망됨

3. 연구목표

○ **최종목표 : 자동차 내부 네트워크 취약점 분석 및 실증 테스트베드 개발**

- 자동차 내·외부 통신에 대한 취약점 분석 및 공격 가능성 검증
- 자동차 내·외부 자원에 대한 접근제어 무력화 가능성 검증
- 자동차 내부 네트워크에 대한 침입탐지 기술 개발
- 자동차 보안취약점 검증 및 진단을 위한 테스트베드 개발

○ **정량적 개발목표**

핵심 기술/제품 성능지표		단위	달성목표	국내최고수준	세계최고수준 (보유국, 기업/기관명)
1	취약점 공격 PoC 수	개	10		
2	이상 패킷 탐지율	%	85		
3	주행 데이터 수집율	%	75		
4	이종 데이터셋 종류	종	5		
5	테스트베드 지원 ECU 수	종	20		
6	이종 IVN 지원	종	3		

*1) 중요 기술/제품 개발목표를 계량화하여 제시하되, 연구내용과 기술분야별 특성에 따라 변경 가능

* 정량화된 목표 제시가 곤란한 경우 정성적 목표 제시 가능

○ **연차별 개발목표**

구분	연도별 연구목표
2021년	자동차 내부 네트워크 대상 취약점 공격벡터 도출 및 개념증명기술 개발
2022년	자동차 내부 네트워크 공격 실증을 위한 실차 기반 테스트베드 개발
2023년	실차 기반 자동차 내부 네트워크 공격 시나리오 및 침입탐지기술 개발
2024년	자동차 내부 네트워크 가상화 기술 및 실차 기반 침입탐지 적용기술 개발

4. 연구내용

○ **개발 기술 내용**

① 자동차 내·외부 통신에 대한 취약점 분석 및 공격 가능성 검증

- CAN, MOST, FlexRay, LIN, Ethernet 기반의 자동차 내부 평면적 또는 계층적 네트워크에 대한 이종 패킷 수집 기술 개발 (Domain, Zone 기반)
- 초음파, 카메라, GPS, V2X 인터페이스 등 외부정보 오류주입에 따른 내부 네트워크 패킷 영향성 평가 및 검증

② 자동차 내·외부 자원에 대한 접근제어 무력화 가능성 검증

- 자동차 센서, 입출력 인터페이스 및 진단기 등으로부터의 침입 가능성 검증
- 자동차 헤드유닛 (AVN, IVI)의 운영체제, 시스템 소프트웨어, 어플리케이션 등의 임의 설치 및 변조 가능성 분석 (키관리, 인증서관리, 서명검증 취약점 분석)
- 자동차의 ECU, 모뎀, 원격 서버, 클라이언트 소프트웨어 등의 업데이트 메커니즘 및 통신 프로토콜 취약점 분석

③ 자동차 내부 네트워크에 대한 침입탐지 기술 개발
- 이종 메시지셋에 대한 상호운용성 보장을 위한 자동차 내부 네트워크 메시지 가상화 기술 개발
- 자동차 내부 네트워크 패킷에 특화된 Signature 및 Anomaly 기반 침입탐지기술 개발
- 자동차 내부 네트워크 및 자동차 행위에 대한 복합정보 연동기술 개발
- 빅데이터 및 머신러닝 기반 자동차 이상 행위 및 유해 메시지 분석기술 개발

④ 자동차 보안취약점 검증 및 진단을 위한 테스트베드 개발
- 현장 실차 주행 데이터셋 기반 주행 재현을 위한 동적분석 테스트베드 플랫폼 개발
- 자동차 내부 네트워크, 내장 센서 데이터 및 운전자 행위 정보 라벨링, 동기화 및 메타파일 생성기술 개발
- 실차 주행을 통한 내부 네트워크 패킷 수집 및 통계적 분석을 통한 형상화 기술 개발
- 실차 주행 현장 데이터셋에 대한 시나리오 기반 오류 주입 및 이에 대한 영향 평가기술 개발

○ 기존 (보유)기술

① 자동차 내부 네트워크 패킷 정적 분석기술
- 자동차 내부 네트워크로부터 수집한 패킷을 대상으로 통계적 분석기법을 적용하여 패킷으을 분류하고 검증하는 기술

② 자동차 내외부 취약점분석 기술
- 자동차 내외부 접점으로부터의 이상 신호 주입과 이에 따라 변화하는 내부 네트워크 패킷을 수집하는 기술

③ 자동차 테스트베드 구현 기술
- 실차용 ECU 등을 이용하여 실차와 동일하게 테스트베드를 구성하여 실차용 패킷을 주입하여 재현할 수 있는 기술

5. 지원기간/예산/추진체계

○ 기간 : 4년 이내

○ 정부출연금 : ‘21년 9억원 이내 (총 정부출연금 45억원 이내)

○ 주관기관 : 대학 또는 산업체

기술분류	대분류(차세대보안)-중분류(융합보안) -소분류(자동차보안)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (4) ~ (6)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	

관리번호	2021-차세대보안-27	(지정공모형)
과제명	AI 시스템 내의 정보흐름 추적 및 제어를 통한 프라이버시 위험 분석 및 대응 기술 개발	

1. 개념

○ **(개념)** 개인정보를 사용하는 AI 및 기계학습 시스템의 학습에서 활용까지의 전주기에서 발생하는 정보흐름 및 사용을 추적하여 프라이버시 위험을 분석하고 이에 대응하는 기술을 개발함

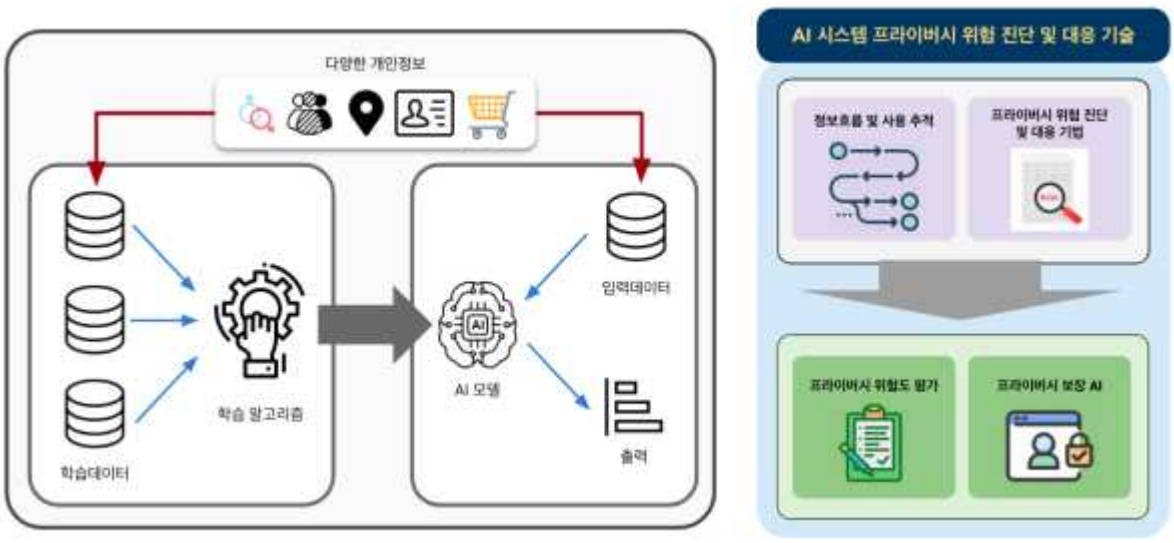
○ **(문제점)** 4차 산업혁명으로 많은 기업들이 기계학습·AI 기반 시스템을 활용하여 제품 및 서비스 상용화에 초점을 맞추고 있지만, 이러한 모델에서 발생할 수 있는 개인정보의 유출 및 오·남용 등 프라이버시 위험에 대한 인식과 대응기술은 부족한 상황임

- AI 시스템의 학습에서부터 활용까지 학습데이터에 사용되는 개인정보가 유출되거나 개인정보의 오·남용이 발생하여 AI 모델 자체의 신뢰성과 안전성 하락을 초래할 수 있음
- 서비스로서의 AI(AIaaS)를 개발하고자 하는 벤처 및 일반 기업들은 개인정보 보호가 아닌 성능을 우선시하며, 정보보호 관련 전문성의 부족 등으로 프라이버시 위협에 대한 대응 조치가 힘든 상황임

○ **(해결방안)** AI 시스템에서 발생할 수 있는 개인정보의 유출 및 오·남용 방지를 위한 프라이버시 위험 분석 및 대응 기술개발

- AI의 학습에서 활용까지 발생할 수 있는 개인정보의 흐름 및 사용을 추적 및 분석함으로서 프라이버시 위험도 측정 및 분석
- 개인정보의 유출 및 오·남용을 방지하고 프라이버시 위험도를 최소화하기 위한 AI 시스템 내의 각종 개인정보의 흐름 및 사용 제어 기술개발
- 프라이버시가 보장되는 AI 시스템 개발을 위해 AI 시스템 개발자가 이용 가능한 자동화 프라이버시 진단 도구 생성

< 개념도 >



2 필요성

○ **(정부 지원 필요성)** AI 진흥을 위한 학습데이터 구축 및 차세대 AI 개발 등이 활발하나, 개인정보를 활용하는 AI로 인해 발생 가능한 프라이버시 침해에 대한 이해 및 그 대응이 미흡

- AI의 학습 및 활용까지의 과정에는 다양한 개인정보가 내재되어 있음
- AI 활용과정에서 발생하는 프라이버시 침해는 AI의 신뢰도를 하락시켜 AI 진흥에 큰 걸림돌이 될 수 있음
- 산업계에서는 더 나은 성능의 AI 개발에 주력하고 있으므로, AI의 활용과정에서의 개인정보 보호는 정부 주도의 기술개발 지원이 필요함
- AI 시스템의 정보흐름에 관한 기술 연구는 프라이버시에 직접적으로 연관되어 있을 뿐만 아니라 AI의 안전성 신뢰성과 밀접한 관련이 있으므로 차세대 보안기술 기술 중 하나로 추진되어야 함

○ **(기술성)** AI를 활용하기 전 학습 데이터에 대한 프라이버시 보호기술 개발은 많으나, AI 학습 및 활용 전주기 상의 프라이버시 노출, 침해 위험에 대한 대응기술은 미흡

- 허술하게 설계된 AI 시스템을 통해 AI 학습데이터 뿐만 아니라 학습되어 운용되는 AI 시스템으로부터 다양한 개인정보 유출 및 오·남용이 발생 가능
- EU에서는 GDPR 등의 제도적인 장치를 통해 AI 시스템에서 발생하는 무분별한 개인정보의 사용을 제한하지만, 실질적으로 이를 기술적인 뒷받침은 미비한 상태
- 해외에서는 미국의 DARPA의 Accountable Information Use 프로그램 등 관련 대응이 활발
- AI 시스템에서 발생하는 정보흐름 추적 및 제어 기술 개발을 통하여 프라이버시를 보장하는 AI 개발에 세계적인 선두 그룹으로 도약할 수 있는 발판 마련

○ **(경제성)** 개인정보 등 광범위한 데이터를 활용하는 AI 시스템의 프라이버시 위협 대응기술에 대한 수요가 증가하고 있으며, 개인정보 유출 및 오·남용 방지를 통해 막대한 거시적 경제적 효과가 발생할 것으로 예상됨

- 가트너는 2021년 10대 전략기술에 "AI 엔지니어링" 및 "개인정보보호 강화 컴퓨팅" 기술을 포함하며, 2025년경 대기업의 50%는 개인정보보호가 강화 컴퓨팅을 도입하여 상용 데이터를 보호할 것으로 전망. 이를 위해 데이터 처리활동 및 분석 과정에서 발생할 수 있는 프라이버시 위험 분석이 필요

  ※ 가트너,「Gartner Top Strategic Technology Trends for 2021」
- 가트너는 또한 AI의 실용화를 위해서는 "AI 엔지니어링"을 통해 라이프사이클을 관리하는 것이 중요하게 될 것으로 예상. 이를 위해서는 AI 시스템의 정보흐름 추적을 통한 위험 분석 및 대응이 필요.

3. 연구목표

○ **최종목표 : AI 시스템에서 발생할 수 있는 개인정보의 유출 및 오·남용 방지를 위한 프라이버시 위험 분석 및 대응 기술 개발**

- AI 시스템의 학습 및 활용 전주기에서 발생하는 정보흐름 및 사용에 대한 추적 기술 연구 및 개발
- 허술한 AI 시스템에서의 개인정보 유출 및 오·남용 가능성 분석 기술

- AI 모델 내의 정보흐름 및 사용에 기반한 프라이버시 위험 분석 기술
- 비식별화 등의 프라이버시 보호를 위한 AI 모델 내의 정보흐름 제어 기술
- 어플리케이션 특화 AI 모델에 적용 가능한 프라이버시 위험 자동화 분석 도구

○ 정량적 개발목표

핵심 기술/제품 성능지표		단위	달성목표	국내 최고수준	세계최고수준 (보유국, 기업/기관명)
1	개발된 프라이버시 위험에 대한 정량화된 측도 개수	개	4(a)	N/A	2(b) (Singapore, NUS)
2	개발된 프라이버시 침해 대응 기술 개수	개	3(a)	N/A	N/A
3	개발된 프라이버시 침해 대응 기술 적용 시 평균 위험도 감소 정도	%	50	N/A	N/A
4	진단 도구에서 지원하는 AI 프레임워크 종류 (c)	개	2	N/A	1(b) (Singapore, NUS)
5	진단 도구에서 진단 가능한 프라이버시 침해 종류	개	3	N/A	2(b) (Singapore, NUS)

(a) 한국 정보과학회가 정의한 최우수보안학회에 발표된 논문들을 참고하여 목표치 설정
(b) ML Privacy Meter, https://github.com/privacytrustlab/ml_privacy_meter
(c) AI 프레임워크 종류: PyTorch, TensorFlow 등의 기계학습 API에서 지원하는 모델 바이너리 혹은 소스 파일

○ 연차별 개발목표

구분	연도별 연구목표
2021년	AI 학습 및 활용 전주기에서 발생하는 정보흐름 및 사용 추적 기술 개발
2022년	정보흐름 및 사용 정도에 기반한 프라이버시 위험 분석 기술 개발
2023년	AI 모델에서의 프라이버시 침해 방지 대응 기술 개발
2024년	선택 어플리케이션 특화 AI에서의 프라이버시 위험 자동화 분석 도구 개발

4. 연구내용

○ 개발 기술 내용

① 선택 어플리케이션에 특화된 개인정보 수집, 전처리 및 학습 기술
- 마케팅, 위치 정보, 메디컬 등 특정 종류의 개인정보를 사용하는 AI 어플리케이션에 대한 지정
- 선택 어플리케이션에 특화된 개인정보 데이터 수집 및 전처리 기술
- 선택 어플리케이션에 특화된 AI 시스템의 학습

② AI 시스템의 학습에서 활용까지의 전주기 정보흐름 및 사용 추적 기술
- 기존 AI 시스템에서의 정보흐름 및 사용 추적 기술 현황 분류
- AI의 학습데이터로서 활용되는 개인정보에 대한 정보흐름 분석 기술
- AI 시스템이 입력으로 활용하는 개인정보의 정보흐름 및 사용 추적 분석 기술

③ 정보흐름 및 사용에 기반한 AI 시스템에서의 프라이버시 위험 진단 기술
- 기존 AI 시스템에서의 프라이버시 공격기술의 분류
- 정보흐름에 따른 AI 시스템에서의 입력 및 학습데이터 유추 공격기술의 연구
- 유추 공격에 기반한 범용 AI 시스템에서의 프라이버시 위험 진단 기술
- AI 시스템의 입력으로 활용되는 개인정보에 대한 오·남용 분석 및 진단 기술
- AI 시스템에서의 프라이버시 침해 위험도 측도(measure)의 개발

④ AI 시스템에서의 개인정보의 유출 및 오·남용 방지를 위한 대응 기술
- 입력 및 학습데이터에 적용가능한 기존 비식별화 기법들의 분류
- 기존 비식별화 기법들의 효과성 진단 기술 연구
- AI 시스템에 특화된 개인정보 유추 및 추출 방지를 위한 입력 및 학습데이터 비식별화 기술
- 개인정보 오·남용 방지를 위한 AI 시스템에서의 정보흐름 및 사용 제한 기술
- 인공지능 모델의 성능을 떨어트리지 않는 프라이버시 위험 대응 기술

⑤ 선택 어플리케이션 특화 AI 시스템에 적용 가능한 자동화된 프라이버시 위험 분석 및 대응 기술 플랫폼 개발
- 도메인 특화 인공지능 모델 질의를 위한 인터페이스 설계 및 개발
- 도메인 특화 개인정보 노출 및 오·남용 구현을 위한 선택적 필터 설계
- 프라이버시 위험 분석 결과에 대한 추천을 사용자에게 알리는 인터페이스 설계
- 도메인 특화 AI 시스템에 대한 프라이버시 분석 플랫폼 검증

○ 기존 (보유)기술

① Membership inference attack, Model inversion attack 등 허술하게 설계된 AI 모델에 대해 각각 학습 및 입력 데이터에 사용된 개인정보를 유추 및 추출하는 공격 기법

② ML Privacy Meter (NUS)
- 기계학습 모델에서 사용하는 학습데이터 및 입력데이터 등의 개인정보에 대해 프라이버시 위험도를 계측하는 도구
  · https://github.com/privacytrustlab/ml_privacy_meter

③ 설명가능한 인공지능 기법인 특성기여도 분석법은 각 입력 특성이 출력에 미치는 영향도를 정량화함으로서 입/출력의 정보흐름을 나타냄.
- iNNvestigate (Technische Universit¨at Berlin, Fraunhofer Heinrich Hertz Institute)
  · 다양한 특성기여도 분석법들에 대해 구현해 놓은 라이브러리
  · https://github.com/albermax/innvestigate

5. 지원기간/예산/추진체계

○ 기간 : 4년 이내

○ 정부출연금 : '21년 3억원 이내 (총 정부출연금 15억원 이내)

○ 주관기관 : 대학

기술분류	대분류(차세대보안) - 중분류(공통기반 보안) - 소분류(인공지능 보안)	
연구유형	기초연구 ( ), 응용연구 (O), 개발연구 ( )	**TRL** (2) ~ (4)
과제특징	정책지정( ), 혁신도약형( ), 경쟁형( ), 표준화연계( ), SW자산뱅크등록( ), 공개SW( ), 기술료비징수( ), 일자리연계( ), 규제샌드박스( ), 사업화연계( ), 소재부품장비( ), 연구데이터공개( )	