동형암호화 기반
데이터 활용 플랫폼 구축 사례

2023. 3. 31

이홍일. (주)아이넷

목차

1 데이터 공유와 개인정보 보호

1. 데이터 결합/공유/활용 동향
2. 데이터 공유와 개인정보
3. 암호화를 통한 개인정보 공유 절차

1.1 데이터 결합/공유/활용 동향

- 산업계의 데이터 협력 배경 - 고객의 데이터를 확보하고 융합 연계하여 활용

주요 동향

'22.1월 API 방식의 마이데이터 시스템 구축

초거대 AI 업체의 빅데이터 기반 고도화

데이터 중립성 관련 입법 논의 진행

데이터 결합/공유 현황

전 금융권의 금융 데이터가 개방/공유

산업 여러 분야 데이터를 결합/융합
새로운 비즈니스 모델 창출

플랫폼 보유 데이터의 차별없는 제공
(데이터의 개방과 공유)

1.2 데이터 공유와 개인정보 보호

- 데이터의 결합 및 공유 ➔ 개인정보를 비식별화 후 자유로운 활용이 가능

	개념	활용가능 범위	기술적 보호조치
개인정보	개인에 관한 정보 개인을 식별할 수 있는 정보	사전적이고 구체적인 동의를 받은 범위 내 활용 가능	전달 및 저장시 암호화
가명정보	추가정보의 사용 없이는 특정 개인을 식별할 수 없도록 조치한 정보	통계작성 연구 공익적 기록보존	비식별화(가명화)
익명정보	더 이상 개인을 식별할 수 없게 조치한 정보	개인정보가 아니므로 제한없이 이용 가능	비식별화(익명화)

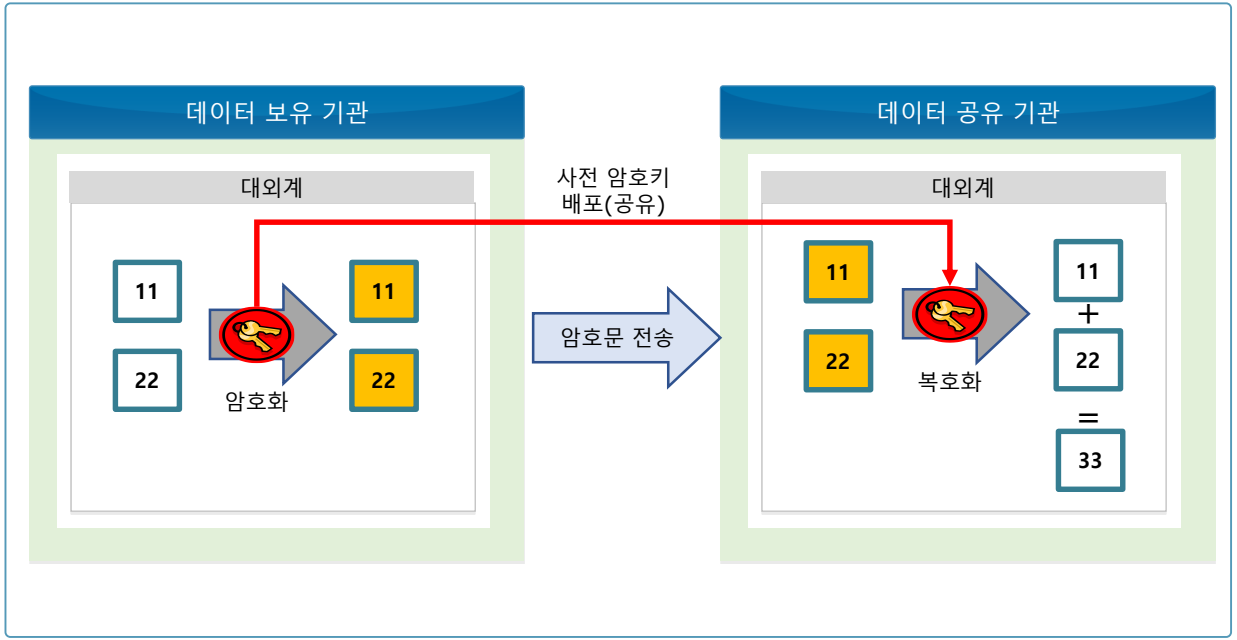
1.2 데이터 공유와 개인정보 보호

- 개인정보의 비식별화 후 데이터 활용의 제약사항

	기술적 보호조치	제약사항	기술적 문제점
개인정보	전달 및 저장시 암호화	이용 범위에 대한 사전 동의	- 공유 및 활용시 암/복호화 절차 필요 - 공유하는 제3자에게 암호화키와 평문 노출
가명정보	비식별화(가명화)	활용 범위의 제약 - 통계작성 - 연구 - 공익적 기록보존	
익명정보	비식별화(익명화)	데이터의 손상	- 데이터의 활용도 낮음

1.3 암호화를 통한 개인정보 공유 절차

- 금융권 대외기관 고객 데이터 공유 절차



데이터 공유 범위

- 사전 이용범위에 대한 동의
- 데이터 공유 기관에 평문형태의 개인정보 제공
- 데이터 전달 채널에 대한 보안성/기밀성 확보를 위해 암호화 전송

데이터 공유 절차

- 공유기관에 사전 암호키 전달
- 공유기관 전송 데이터 암호화
- 전송받은 암호화 데이터를 공유기관에서 복호화
- 복호화된 평문데이터를 통한 연산 및 활용

문제점

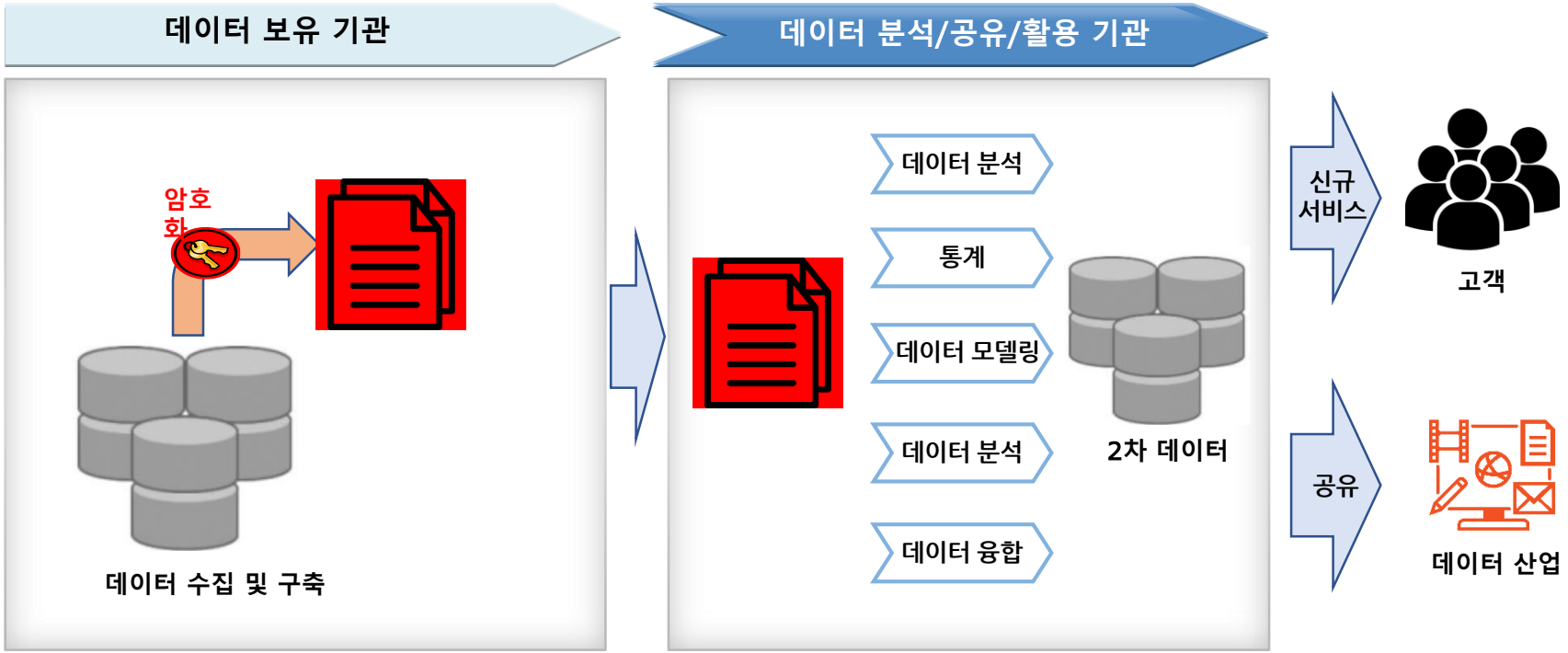
- 공유기관에 암호키 및 평문 노출
- 공유기관의 데이터 활용 목적은 연산 결과임
- 연산 결과 이외에 원천 평문데이터는 공유 불필요

2 동형암호화 기술

1. 새로운 암호화 패러다임의 필요성
2. 동형암호화
3. 동형암호화 장/단점

2.1 새로운 암호화 패러다임의 필요성

- 데이터의 활용만 가능 ➔ 데이터의 내용은 노출 불가능



2.1 새로운 암호화 패러다임의 필요성

기존 암호화 기반 공유 플랫폼의 문제점

데이터의 공유 및 활용시 반드시 복호화 필요

복호화에 따른 암호화 키 공유 절차 필요
암호키 유출방지를 위한 조치 및 비용 발생

데이터를 공유하는 제3자에게 암호화 키와 평문이 노출

데이터 공유시 개인정보에 대한 비식별화 처리
➔ 익명처리등으로 인한 데이터 손실

새로운 암호화 패러다임의 필요성

암호문 형태로 기존의 연산 등의 처리 절차를 그대로 수행

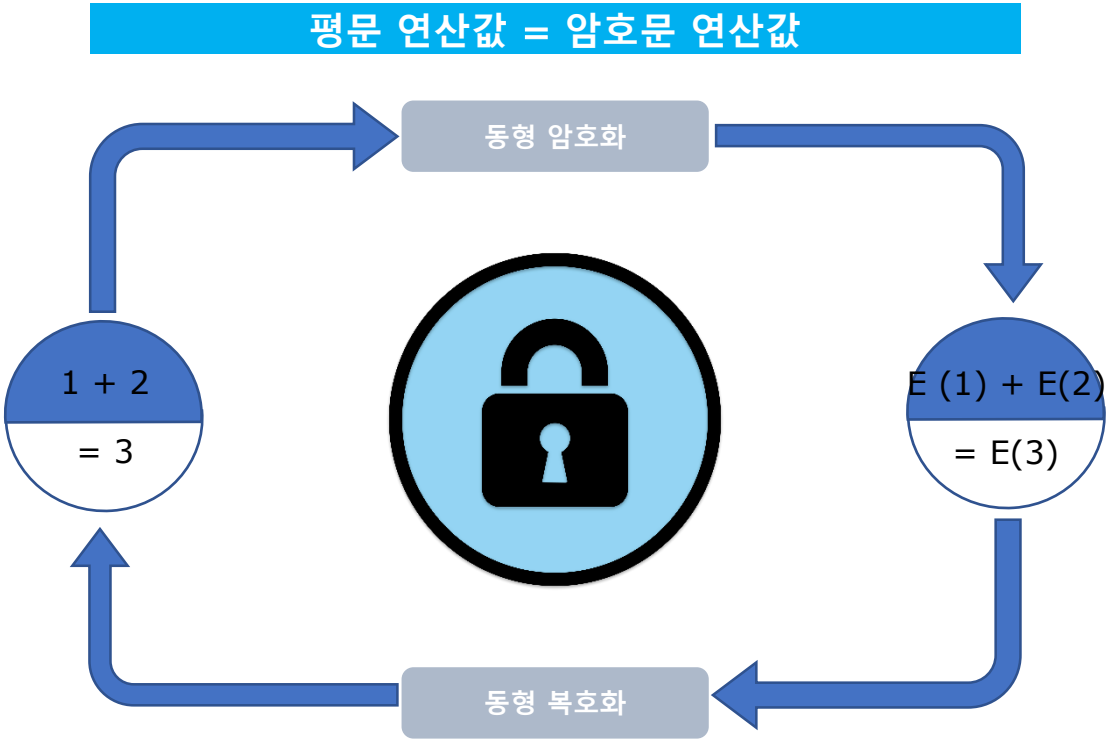
안전성과 협업성을 동시에 제공하는 완전동형암호의 도입

데이터를 공유하고 활용하는 제3자에게 암호화 키와 평문정보의 비노출

비식별화(익명화) 처리가 불필요하므로 데이터 손실없이 개인정보 처리가 가능

2.2 동형 암호화

- 동형 암호화 ➔ 암호문 형태로 연산



암호문에 대한 연산 지원

- 데이터가 암호화된 상태로 연산할 수 있도록 지원하는 암호기술
- 암호화된 상태로 데이터 분석 가능

개인정보 유출 원천 차단

- 개인정보의 소유자(Client)단에서 암호화
- 고객으로부터 제공받은 암호화된 개인정보를 복호화 하지 않고 분석 등의 업무처리가 가능

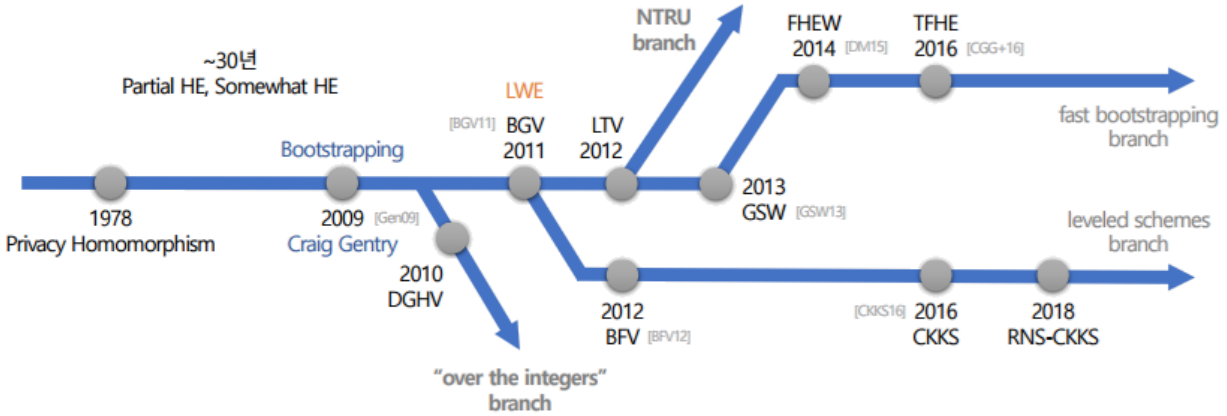
2.2 동형 암호화

- 동형암호화 키

	소유자	기능	사용 시점
공개키	데이터 소유자 등	데이터 암호화 시 사용	- 데이터의 소유자가 공유하고자 하는 데이터를 공개키로 암호화
연산키 (공개키)	데이터 공유 및 활용자	암호문간 혹은 암호문과 평문 연산시 사용	- 데이터 공유 및 활용자가 기존 보유 데이터와 암호문 혹은 암호문간에 연산 수행시 사용
비밀키	데이터 소유자	암호문 혹은 연산결과의 복호화 시 사용	- 데이터의 소유자가 연산 결과값을 전달받아 복호화 수행

2.2 동형 암호화

- 동형 암호화 Scheme의 발전사



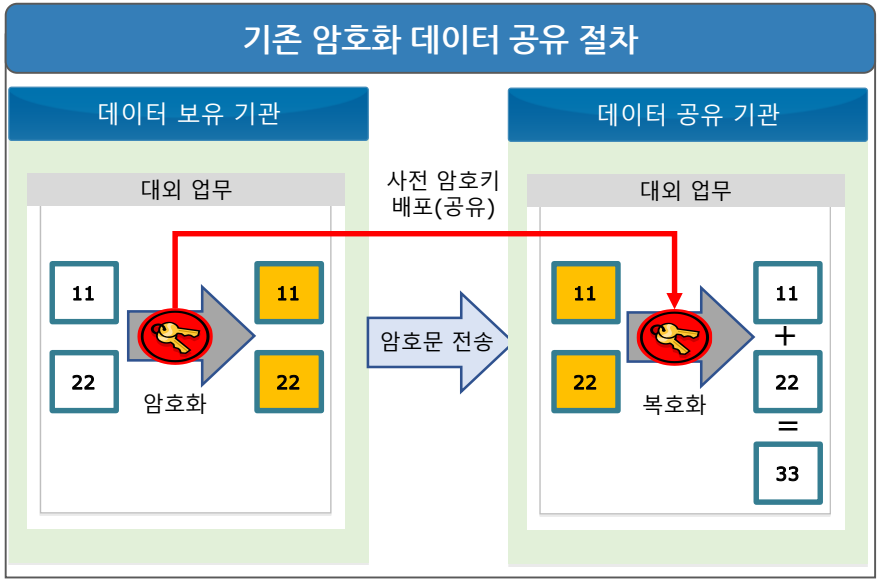
구분	연도	특징	연산 단위	Scheme
1세대	2009	최초의 완전동형암호	비트	
2세대	2011	실제 데이터에 적용 가능	정수	BGV, BFV
3세대	2013	작은 데이터 처리에 효과적	비트	TFHE
4세대	2016	실수까지 연산을 지원하는 동형암호	실수(정수)	CKKS

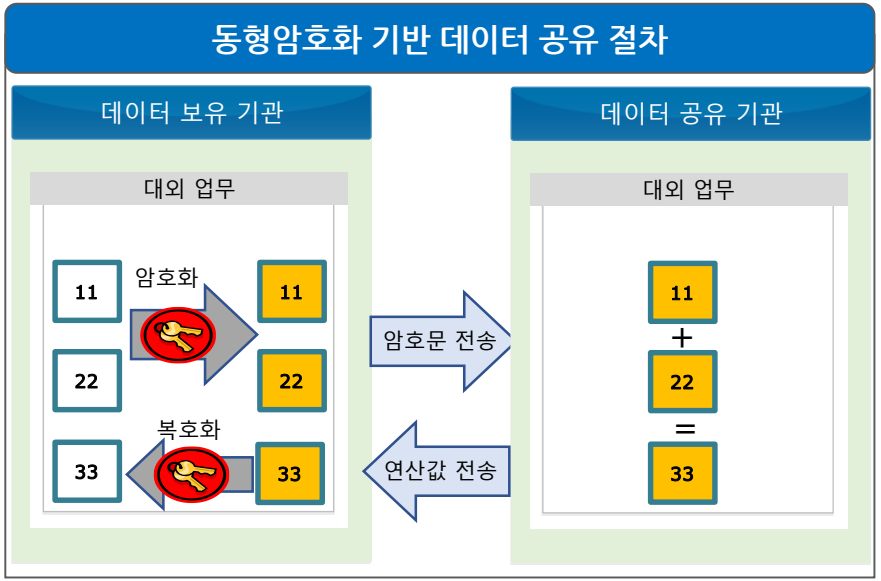
2.2 동형 암호화

- 동형암호화 Scheme의 분류
  - 암호에서 다루는 평문의 데이터 타입과 연산 종류에 따라 Boolean Circuit, Exact Arithmetic, Approximate Arithmetic 세 가지로 분류되며, 응용에서 필요한 조건을 고려하여, 적절한 동형암호 스킴을 선택하여 사용해야됨

	Boolean Circuit	Exact Arithmetic	Approximate Arithmetic
Message Data Type	Boolean	Integer	Real / Complex
Scheme	TFHE, FHEW	BGV, BFV	CKKS
Operation	Boolean Gates, MUX	ADD, MUL	ADD, MUL, Rescale
Library	TFHE, Concrete	HElib, SEAL, Palisade	HEAAN, SEAL, Palisade, Lattigo

2.3 기존 암호화 vs 동형 암호화





데이터 공유 범위 및 보안성 취약

- 데이터 보유기관과 데이터 공유기관간 사전 암호키 공유과정 필요
- 복호화 된 모든 평문에 대한 접근성 부여

데이터 공유의 가용성 및 보안성 확보

- 복호화에 사용하는 비밀키는 보유기관만이 소유
- 데이터 공유기관은 암호문에만 접근 가능하며, 암호문 형태로 연산
- 연산 결과에 대한 복호화는 보유기관에서 수행 후 전달

2.3 동형 암호화의 장/단점

✓ 동형암호화 장점

- 복호화 후 평문 상태의 연산을 수행하는 기존 방식과 달리 **암호문 형태로 연산이 가능**
- 익명화 등의 비식별화 처리 불필요 ➔ **데이터의 손실이 없음**
- 평문 그대로 암호화하여 암호문 형태로 연산 처리 ➔ 데이터의 정확성 보존

✓ 동형암호화 단점

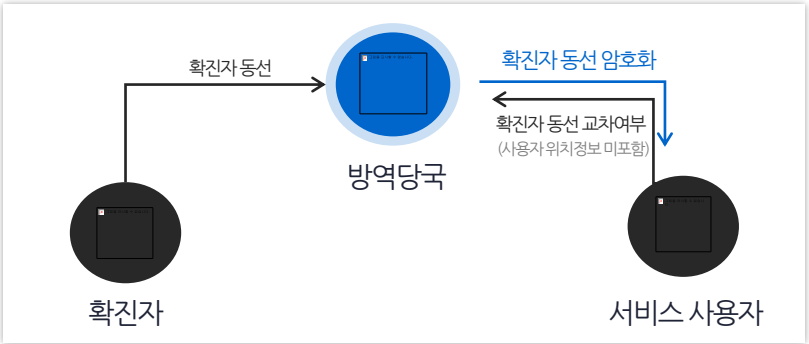
- **암호화 키 사이즈가 매우 큼**
- **암호문의 사이즈가 매우 큼** ➔ 평문 대비 수십 ~ 수백배 (저장 및 활용의 제약)
- 노이즈 감소를 위한 재부팅(Bootstrap) 과정이 필요
- **성능 저하**
  - ➔ 암/복호화 : 수십 ms
  - ➔ 간단한 통계연산 : 수십 ms ~ 0.x s
  - ➔ 간단한 기계학습 훈련 과정 : 수십 분 ~ hour

3 동형암호화 활용 사례

1. 동형 암호화 국내 적용 사례 – 위치정보
2. H은행 동형암호화 데이터 활용 플랫폼 적용 사례

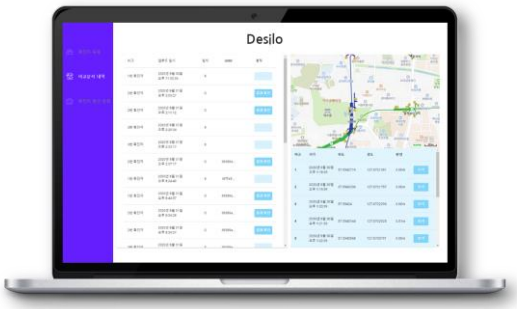
3.1 동형 암호화 국내 적용 사례 - 위치정보

카톨릭대학교병원과 함께 개인정보 및 사생활을 보호하는 동시에 뛰어난 방역 효과를 발휘할 수 있는 동형암호화 기반 코로나 19 방역 솔루션 적용 사례 (Desilo)



〈서비스 개념도〉

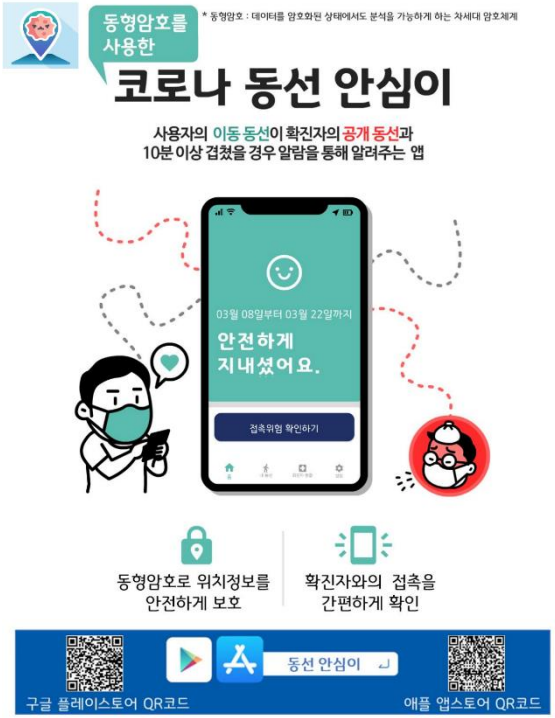
〈구축된 모바일 앱〉

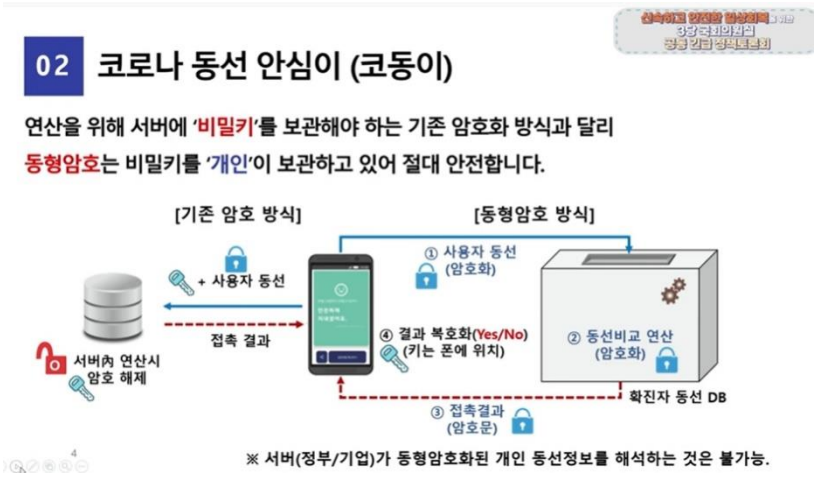


〈구축된 관리자 서버〉

3.1 동형 암호화 국내 적용 사례 - 위치정보

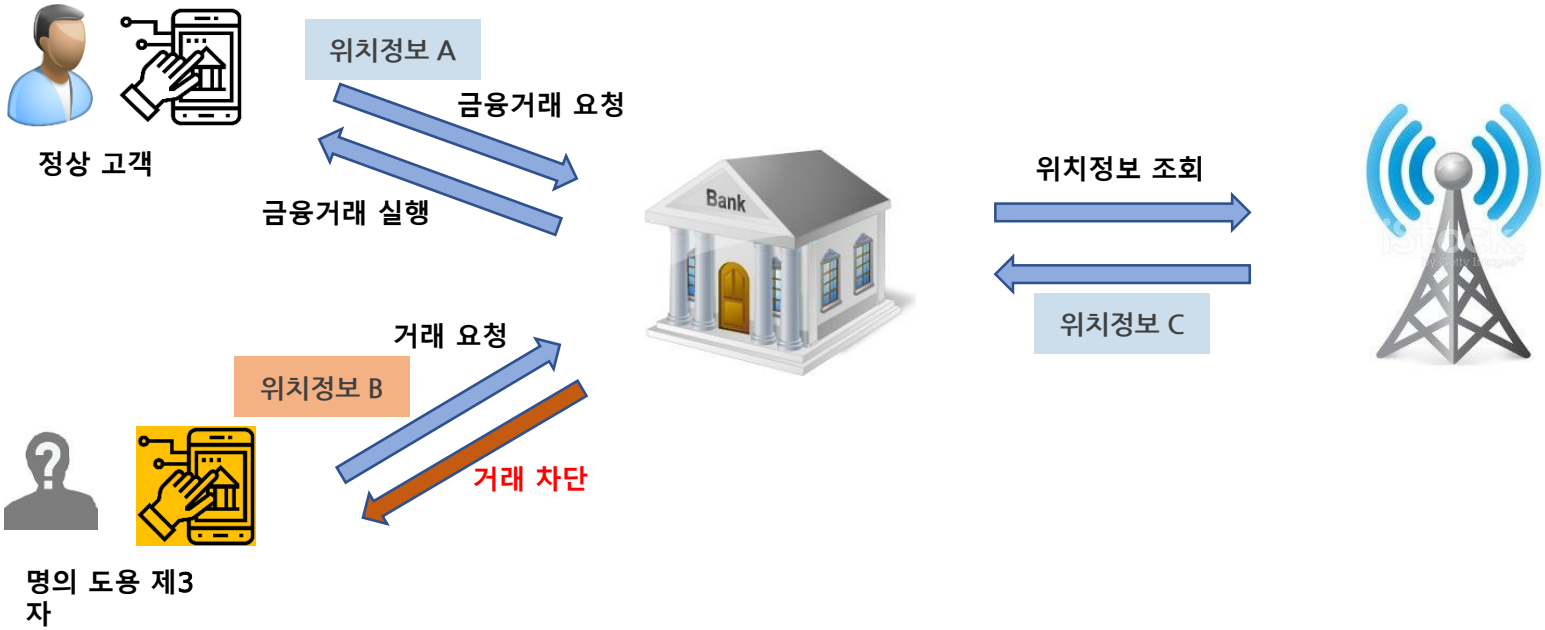
코동이는 사용자의 최근 1주 동선과 공개된 확진자의 동선을 비교하여 접촉 여부를 확인. 모든 사용자 위치 정보는 동형 암호화된 상태로 비교, 연산되므로 안전하게 프라이버시를 보호 (크립토랩)





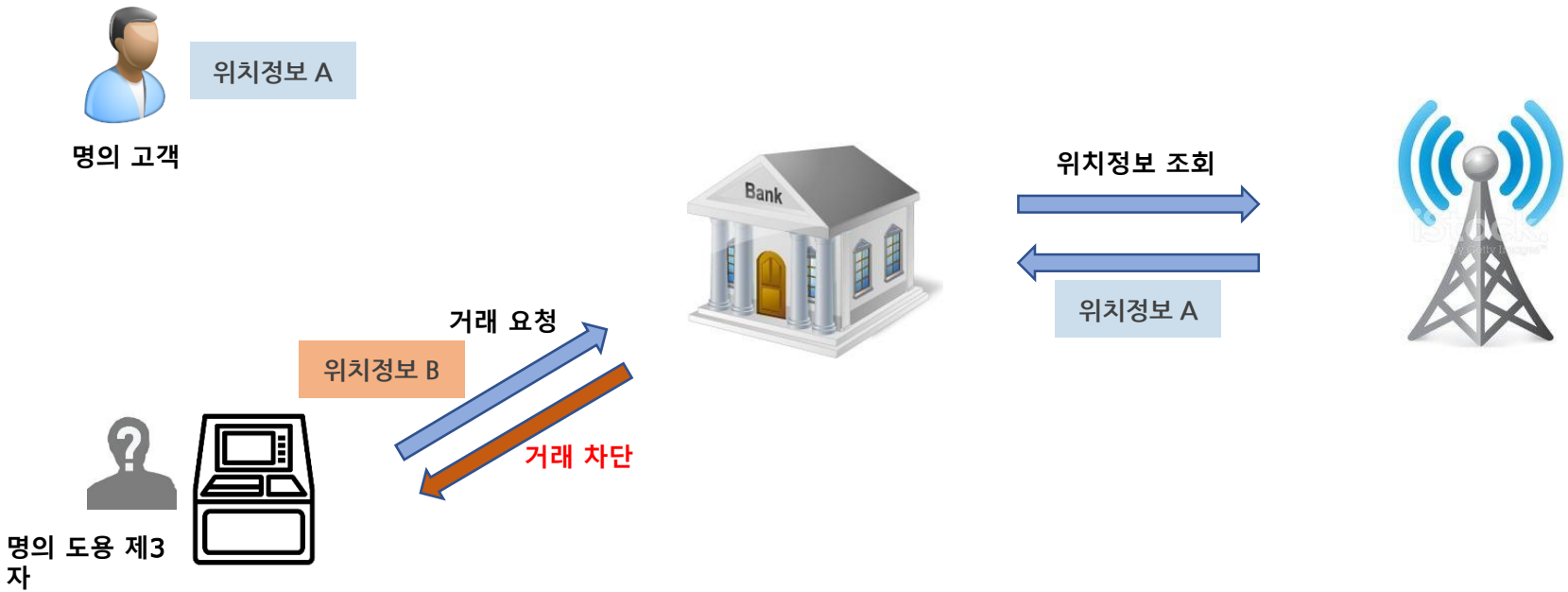
3.2 H은행 동형암호화 데이터 활용 플랫폼 적용 사례

- H은행 위치정보 활용 이상거래 탐지 - 모바일 명의 도용 거래 차단
  - ✓ 모바일 금융앱과 통신사 두채널을 통한 위치정보 조회
  - ✓ 노출 된 개인정보를 도용하여 거래를 하는 제3자에 의한 거래를 탐지 차단



3.2 H은행 동형암호화 데이터 활용 플랫폼 적용 사례

- H은행 위치정보 활용 이상거래 탐지 - 제3자 ATM 인출 거래 차단
  - ✓ ATM의 위치정보와 거래 고객의 휴대폰 번호로 조회한 통신사의 위치정보 조회
  - ✓ 거래 고객이 실제로 ATM 앞에 위치하는지 여부를 판단하여 제3자의 거래시 차단



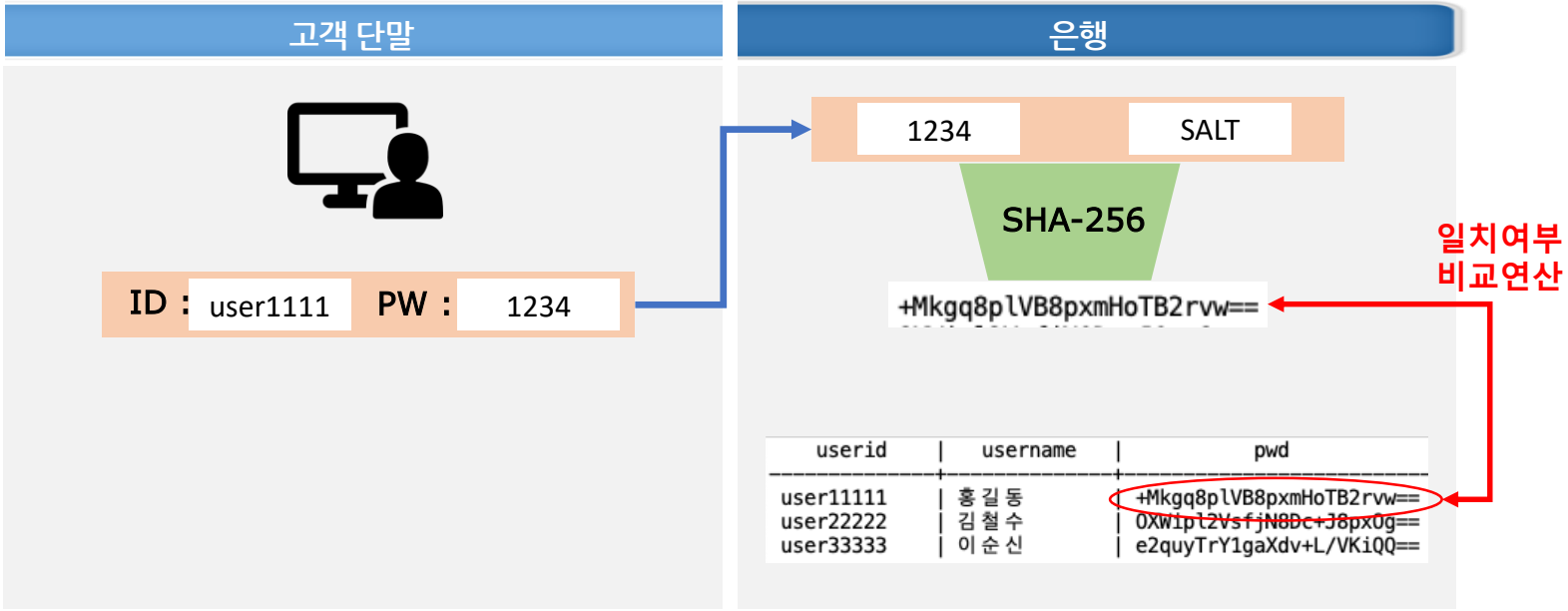
3.2 H은행 동형암호화 데이터 활용 플랫폼 적용 사례

- H은행 동형암호화 위치정보 활용 플랫폼 구축 요건

요건	상세 요건	기술적 조치
금융 거래 고객 식별 및 인증	▪ 모바일 앱을 이용하여 **거래를 실행하는 고객의 식별** ▪ **모바일 앱의 위치정보**와 **통신사의 실시간 위치정보**를 비교 연산하여 **실제 은행 고객의 거래여부 확인**	▪ 이체 거래시 실시간 모바일상의 위치정보 조회 ▪ 거래 실행시 은행에 등록된 고객 전화번호로 해당 통신사에 실시간 위치정보 조회 ▪ 두개의 위치정보 일치여부를 비교 연산
이상거래 탐지	▪ ATM 거래를 실행하는 **고객의 실제 위치 확인** ▪ ATM의 위치정보와 고객의 **통신사 실시간 위치정보**를 비교 연산하여 **실제 은행 고객의 ATM 거래여부 확인**	▪ 피싱 등의 이상거래를 탐지하기 위해 ATM상의 거래 고객의 위치정보를 통신사에 조회 ▪ 사전 등록된 ATM의 위치정보와 실시간 통신사 고객 위치정보를 비교 연산
고객 위치정보 수집 및 조회 불가	▪ **모든 위치정보는 생성되는 위치, 즉 모바일과 통신사쪽에서 암호화** 후 은행 내 처리 ▪ 암호화된 위치정보는 **암호화된 형태로 비교 연산을 수행**하여야 함 ▪ 은행은 해당 암호문에 대한 **복호화 수단이 없어야 함**	▪ 모든 고객 위치정보는 생성되는 위치에서 동형 암호화 적용 ▪ 암호화된 위치정보는 은행내에서 암호문 형태로 비교 연산을 수행 ▪ 연산 결과에 대한 복호화는 암호화를 수행한 위치(모바일, 통신사)쪽에서 수행

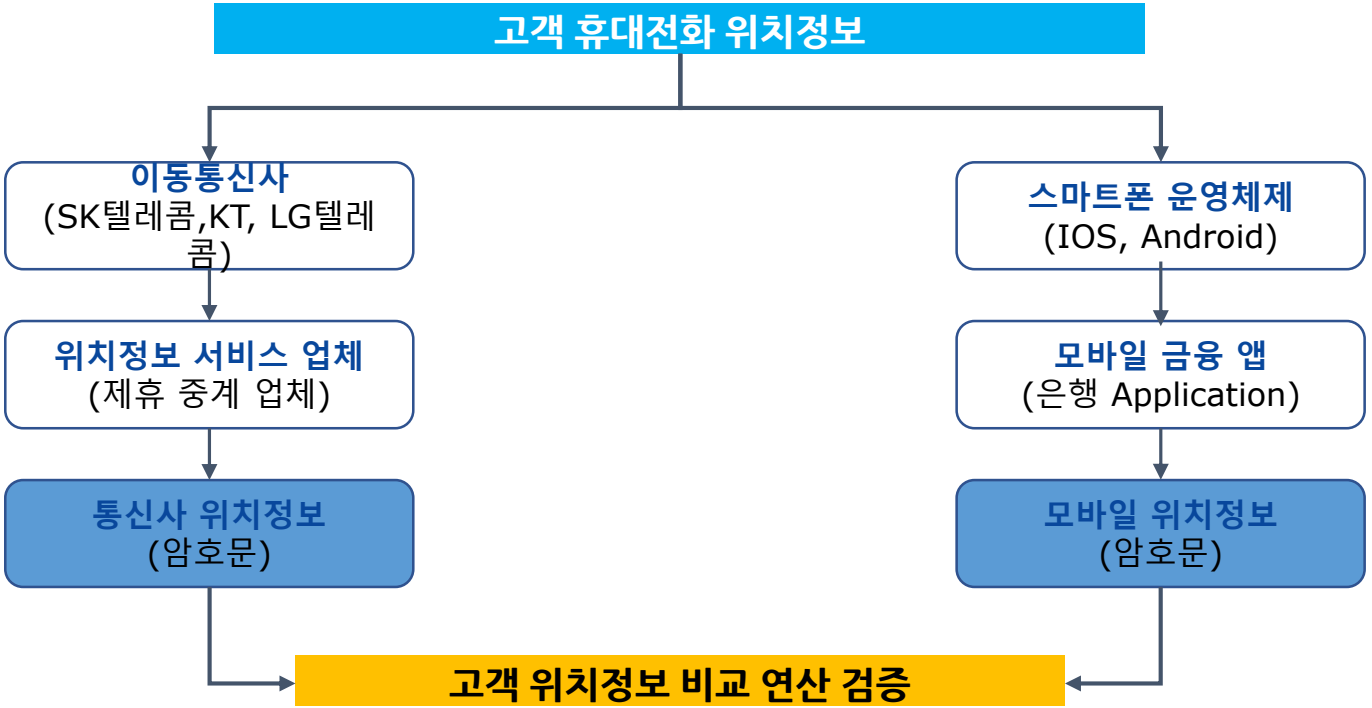
3.2 H은행 동형암호화 데이터 활용 플랫폼 적용 사례

- 기존 암호문 연산 사례 ➔ 단방향 알고리즘
  - ✓ 고객 비밀번호 저장시 단방향 암호화 알고리즘 적용
  - ✓ 은행 내 비밀번호 저장시 단방향 암호화된 Hash/HMAC 값을 저장 ➔ 복호화 불가능
  - ✓ 비밀번호 인증시 저장된 암호화 값과 고객 입력값을 암호화 한 값을 비교 연산하여 일치 여부 확인



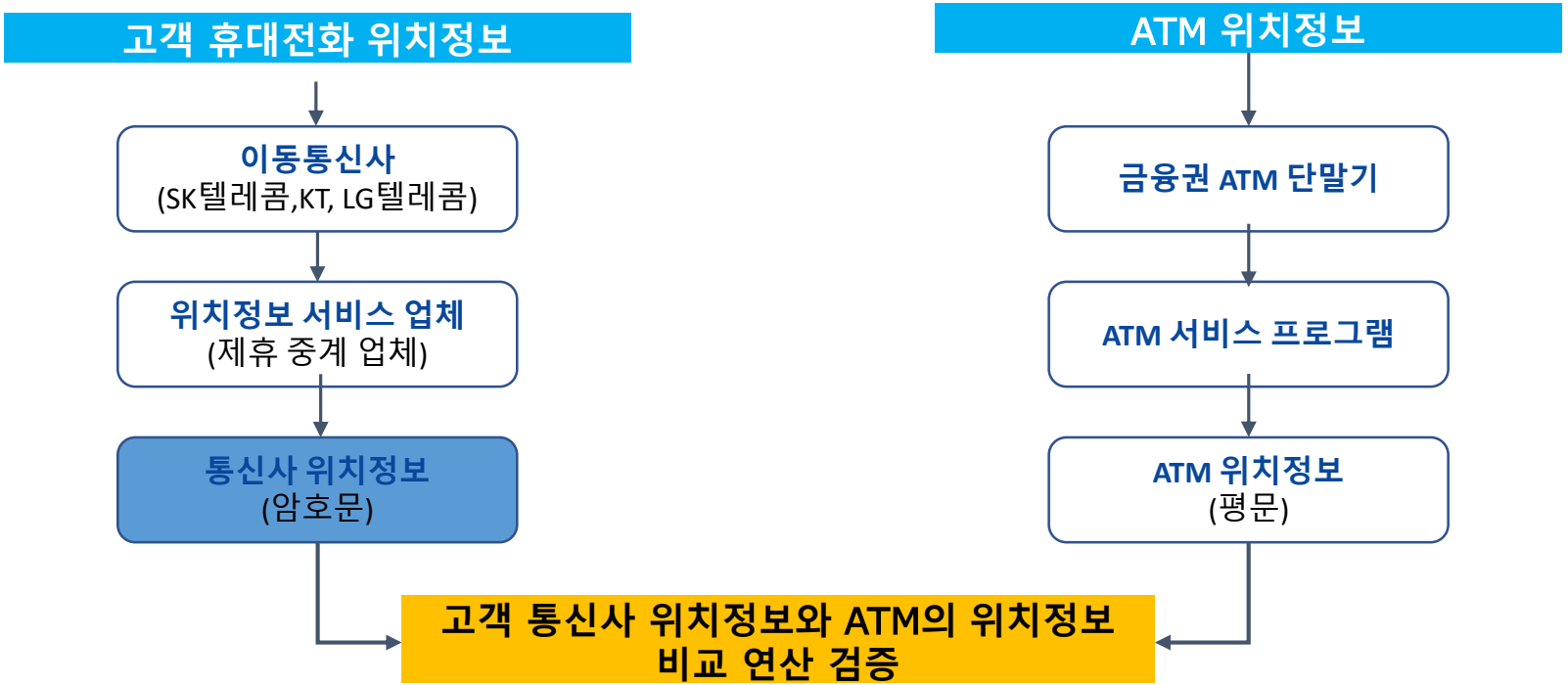
3.2 H은행 동형암호화 데이터 활용 플랫폼 적용 사례

- 데이터의 결합 및 공유 ➔ 개인정보를 비식별화 후 자유로운 활용이 가능



3.2 H은행 동형암호화 데이터 활용 플랫폼 적용 사례

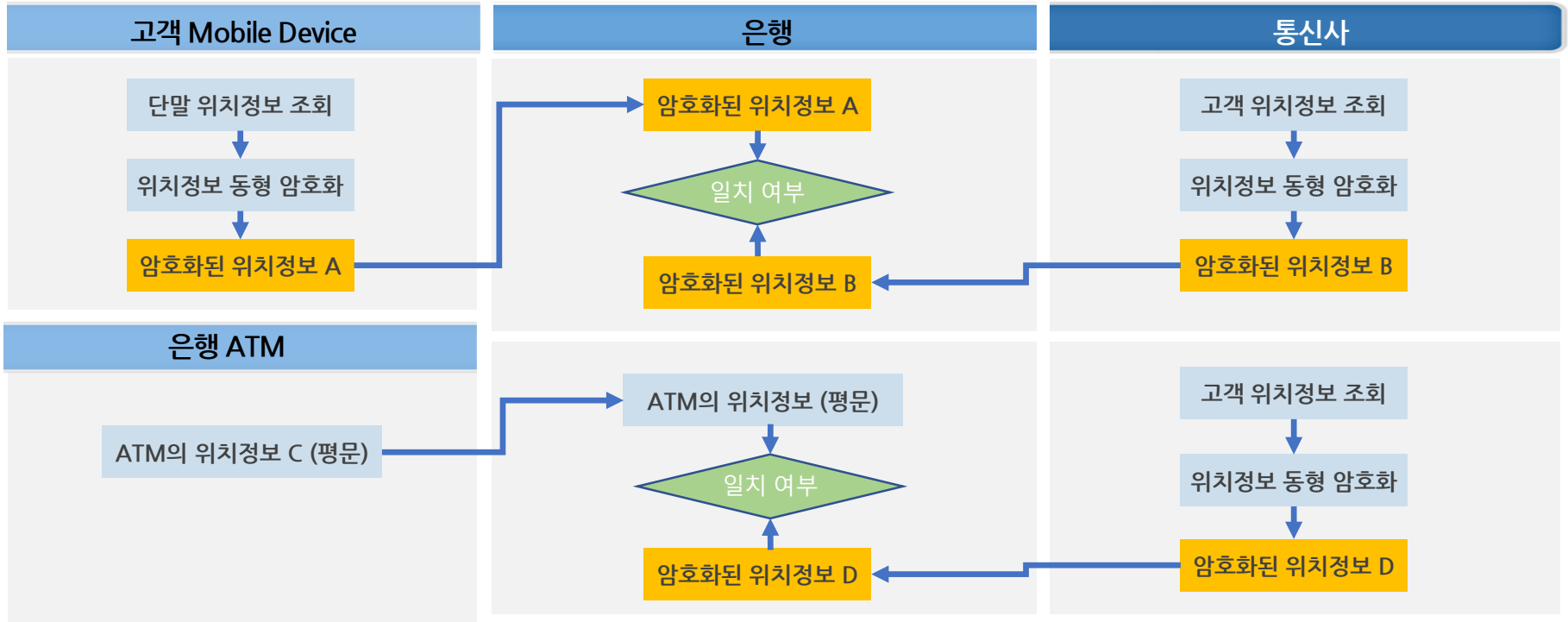
- 데이터의 결합 및 공유 ➔ 개인정보를 비식별화 후 자유로운 활용이 가능



3.2 H은행 동형암호화 데이터 활용 플랫폼 적용 사례

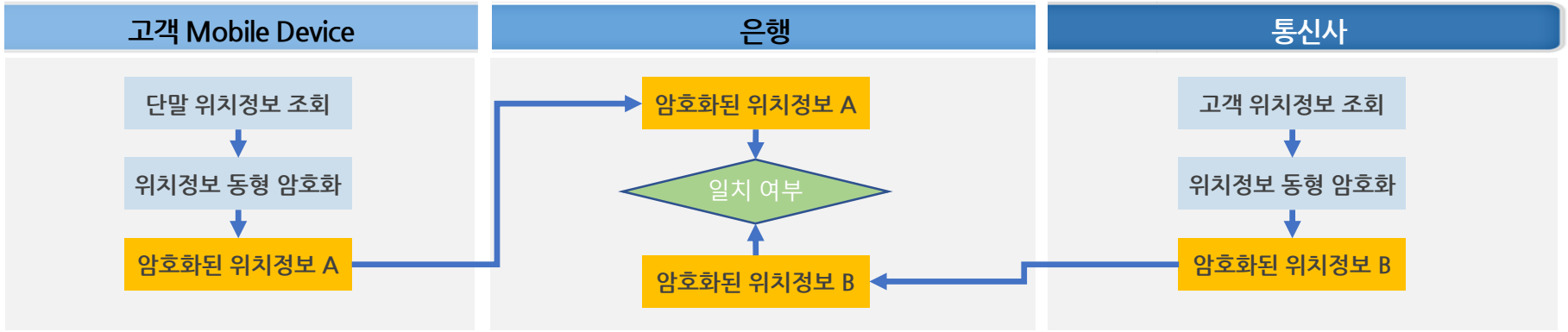
- 기대효과 - 고객 위치정보의 익명화

✓ 은행 내 모든 고객의 위치정보는 동형암호화를 적용하여 암호문으로만 저장/활용 ➔ 복호화 불가능 (비밀키 고객 소유)



3.2 H은행 동형암호화 데이터 활용 플랫폼 적용 사례

- 위치정보 동형암호화 활용 플랫폼 구축 리뷰



위치정보 보호 (비노출)

- 공개키, 비밀키 소유
- **위치정보 생성 시점에 공개키로 암호화**
- 복호화에 필요한 비밀키 전송 없음
- 은행에서 연산된 결과값 복호화 후 전송

위치정보의 익명화

- 고객의 정확한 위치정보의 활용
- 고객의 위치정보는 암호문 형태로 비교 연산 수행
- **은행은 고객의 정확한 위치정보를 알 수 없으나 두가지 위치정보의 비교 연산이 가능**
- 고객의 위치정보 추적, 노출에 대한 리스크 없음

위치정보 보호 (비노출)

- 공개키, 비밀키 소유
- **위치정보 생성 시점에 공개키로 암호화**
- 복호화에 필요한 비밀키 전송 없음

감사합니다