전염병 밀접접촉자 모니터링 알고리즘 성능 개선 방법 연구

Epidemic Close Contact Monitoring Algorithm Research on how to improve performance

유래이
컴퓨터학과
단국대학교
경기도 용인시
yulei0805@dankook.ac.kr

전성환
데이터지식서비스공학과
단국대학교
경기도 용인시
72160261@dankook.ac.kr

나연묵*
컴퓨터학과
단국대학교
경기도 용인시
ymnah@dankook.ac.kr

요 약

코로나 19 전염병 발생 이후 감염자의 수는 엄청났고, 대부분의 국가는 전염병을 피할 수 없었다. 전염병 예방 및 통제 기간 동안 모니터링을 이용해 사용자는 본인의 동선이 확진자와 겹쳤는지 등 많은 번거로움이 있었다. 중국에서는 동형암호화 기술을 기반으로 한 제품이 확진자와 밀접접촉자를 모니터링한다. 하지만, 동형 암호화 기술 알고리즘은 공개 키가 길어 한정적 자원에서 계산하기 많은 어려움이 있다. 본 논문에서는 동형 암호화 기술을 대신할 알고리즘을 소개하여 성능 개선 방법에 연구한다.

키워드 : 동형 암호화 기술, 해시집합, ECC, 코사인 유사도 계산

1. 서론

코로나 19 는 거의 종식되었지만, 전염병 발생 이후 대부분의 국가는 많은 피해를 입었다. 중국에서는 동형 암호화 알고리즘을 사용한 알리페이를 통해 모니터링을 할 수 있었다. 동형 암호화 알고리즘은 광범위하게 밀접접촉자를 선별할 수 있지만 계산 효율이 낮고 공개키 키 관리가 복잡한 단점이 있다. 본 논문에서는 먼저 해시함수를 사용하여 밀접접촉자 의심자를 선별한 다음 타원 곡선 암호와 코사인 간의 유사성을 계산해 밀접 접촉자인지 여부를 결정하는 방법을 제안한다.[3]

2. 관련연구

* 교신저자

2.1 동형 암호화 기술

동형 암호화 기술은 암호문을 먼저 해독할 필요 없이 암호문 상태에서 덧셈 또는 곱셈 작업을 수행할 수 있는 특수 암호화 기술이다. 원본 데이터를 노출하지 않고도 다양한 계산을 수행할 수 있어 데이터 개인정보 보호 효과를 높인다.[5]

2.2 해시집합

해시집합의 계산속도는 빨라야 하며 해시값에서 원시 데이터를 추출할 수 없는 단방향 함수이다. 이러한 특징은 데이터 암호화 및 요약 등의 응용에 사용할 수 있다. 장점으로는 빠른 첨삭 작업과 대규모 데이터 중복 제거, 연관성 등과 같은 데이터 처리 시나리오의 고효율에 적합하다.

2.3 타원 곡선 암호 ECC

ECC(Elliptic Curve Cryptography, ECC)는 타원곡선 이론에 기반한 공개 키 암호 알고리즘이다. 기존의 RSA 및 Diffie-Hellman 알고리즘과 비교하여 ECC 는 동일한 수준의 보안을 보장하고 필요한 공개 키 암호화 길이가 더 짧아 계산량이 적다. 이런 이유로 모바일 장치 및 사물 인터넷 장치와 같은 자원이 제한된환경에서사용하기에더적합하다.[1]

2.4 코사인 유사도 계산 프로토콜

코사인 유사도 계산 프로토콜은 벡터에 대한 정보를 유출하지 않고 두 벡터의 코사인 유사도를 계산해 개인정보 보호 계산 프로토콜이다.[2]

3. 밀접접촉자 모니터링 성능 개선 방법

3.1 시스템 구조

전체 설계는 위치 정보를 획득하고 업로드 서버가 해시집합을 사용하여 작은 범위를 선별한다. 그 다음 타원 곡선 암호 및 코사인 유사도 계산 프로토콜을 사용하여 계산한다. 밀착 여부를 확인하고 서버에 업로드하여 후속 처리를 진행한다.

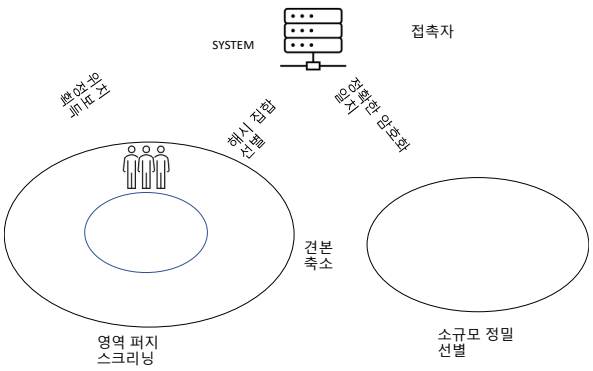

그림 1. 시스템 구조

3.2 알고리즘

3.2.1 해시집합

한 확진자가 목표 좌표를 (x,y)로 설정했다. 위치 퍼지 처리 함수를 통해 정확한 점을 퍼지점으로 변경하고 각 퍼지점을 찾아 해시를 집합으로 만든다. 서버 측 $Has\,h_1$ 과 클라이언트 $Has\,h_2$ 에서 각각 해당 집합을 찾고 교차 집합을 찾아, 공집합이라면 안전, 교집합일 경우 밀착이 의심되는 것으로 판단돼 다음 단계로 넘어간다.

$$Has\,h_1\{(x_n,y_n)\}\cap Has\,h_2\{(x_n,y_n)\}=\varnothing \text{ 안전}$$
$$Has\,h_1\{(x_n,y_n)\}\cap Has\,h_2\{(x_n,y_n)\}\neq\varnothing \text{ 밀착의심}$$

그림 2 에서 감염자 데이터 암호화 정보를 생성할 때 이 프로토콜이 사용하는 해시 집합의 성능이 동형 암호화 알고리즘의 약 250 배임을 확인할 수 있다.

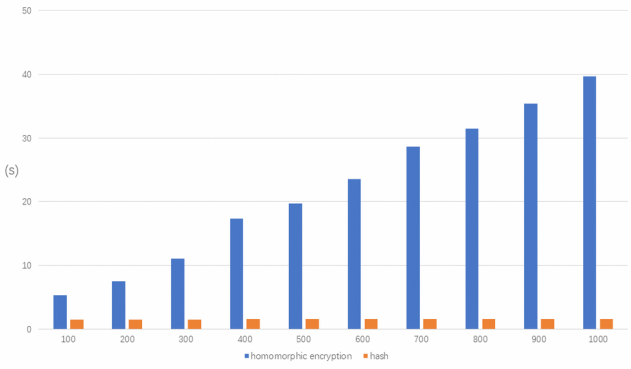

그림 2. 감염자 데이터 암호화

3.2.2 소범위 정밀 알고리즘

ECC 인 타원 곡선 암호화 알고리즘을 사용하여 해시

집합에서 계산된 결과를 암호화하여 서버에 업로드한 다음 서버 측에서 위치 정보를 해독한다.[4][6]

•암호화 방법: 난수 r 을 선택하여 메시지 M 을 암호문 C, C=(rp, m+rQ)로 생성.

•암호 해독: M+rQ-k(rP)=M+r(KP)=M

내 위치 정보와 해독된 위치 정보를 코사인 유사도로 계산하여 내가 밀착자인지 여부를 판단한다.

그림 3 에서 사용자의 위치를 2 차 암호화하여 업로드한 서버 단말기를 알 수 있으며, 이 두 알고리즘의 성능은 데이터가 많을수록 동형 암호화 알고리즘이 더 많은 시간을 소비한다는 것을 알 수 있다.

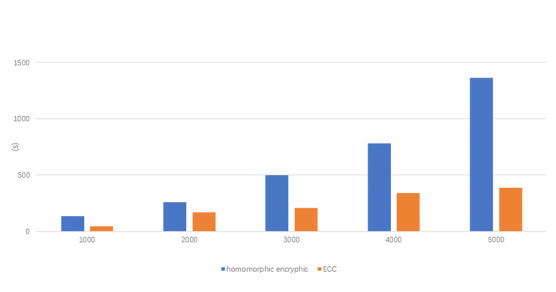

그림 3. 사용자 위치 2 차 암호화

•코사인 유사도 계산 공식:

$$\cos(\theta) = \frac{a \cdot b}{\|a\| \times \|b\|}$$

$$¿ \frac{(x_1, y_1) \cdot (x_2, y_2)}{\sqrt{x_1^2 + y_1^2} \times \sqrt{x_2^2 + y_2^2}}$$

$$¿ \frac{x_1 x_2 + y_1 y_2}{\sqrt{x_1^2 + y_1^2} \times \sqrt{x_2^2 + y_2^2}}$$

그림 4 에서 감염자와 피검자의 위치를 판별할 때 코사인 유사도 알고리즘에 필요한 시간 소모량을 거의 알 수 없는 반면, 동형 암호화 알고리즘은 정보의 양이 증가함에 따라 소모되는 시간 또한 증가함을 알 수 있다.

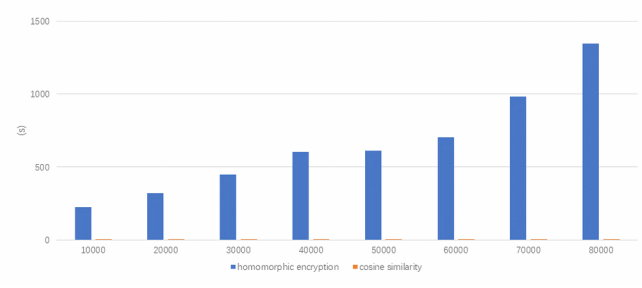

그림 4. 감염자와 검사자의 위치 판별

5. 결론

본 논문에서는 동형 암호화 알고리즘과 다른 알고리즘을 적용한 새로운 밀착접촉자 선별 방법을 제안한다. 성능을 개선시키기 위해 다양한 알고리즘을 결합했다. 또한, 이 알고리즘은 밀착접촉자를 선별하는 데만 사용할 수 있는 것이 아니다. 예를 들어 택시 서비스 소프트웨어도 사용할 수 있으며 계획의 실현은 사회 응용 분야 및 전염병 확산 연구 분야에서 좋은 발전 전망을 가지고 있다. 향후 연구로는 직접 시스템을 구축하여 새로운 알고리즘을 적용했을 때 개선된 성능을 측정한다.

Acknowledgement

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 2023 년도 SW 전문인재양성사업의 결과로 수행되었음"(2022-0-01110)

참 고 문 헌

[1] M. -Q. Hong, P. -Y. Wang and W. -B. Zhao, "Homomorphic Encryption Scheme Based on Elliptic Curve Cryptography for Privacy Protection of Cloud Computing," *2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS)*, New York, NY, USA, 2016, pp. 152-157.

[2] 우균, 송지원, 임은주 and 하성윤. "음원의 주파수 변화율과 코사인 유사도 알고리즘을 이용한 음악 검색 시스템 개발." 한국정보처리학회 학술대회논문집, vol. 21, no. 2, pp. 1027-1030, 2014.

[3] 김병규, 류범종 and 심형섭. "COVID-19 확진자 이동경로 정보 공유 체계 구축 연구" 한국컴퓨터정보학회논문지 25, no.12 (2020) : 155-163.

[4] 정윤수, 김용태 and 이상호. "ECC 기반의 클러스터간 노드들의 안전한 인증 프로토콜." , vol. 13, no. 2, pp. 167-176, 2008.

[5] S. J. Mohammed and D. B. Taha, "Performance Evaluation of RSA, ElGamal, and Paillier Partial Homomorphic Encryption Algorithms," *2022 International Conference on Computer Science and Software Engineering (CSASE)*, Duhok, Iraq, 2022, pp. 89-94.

[6] X. Fang and Y. Wu, "Investigation into the elliptic curve cryptography," *2017 3rd International Conference on Information Management (ICIM)*, Chengdu, China, 2017, pp. 412-415.