동형암호를 이용한 인공지능 학습 데이터 보안 연구 동향

오지현, 김명현, 이준영, 박영호

경북대학교

chldlstnr071@knu.ac.kr, kimmyeong123@knu.ac.kr, harry250@knu.ac.kr, parkyh@knu.ac.kr

Research on Security of Artificial Intelligence Learning Data based on Homomorphic Encryption

Oh Ji Hyeon, Kim Myeong Hyun, Lee Joon Young, Park Young Ho

Kyungpook National Univ.

요 약

동형암호는 암호화된 데이터에 대한 연산 및 처리를 가능하게 하는 암호시스템으로 데이터의 기밀성 및 처리 효율성을 보장하기 위해 스마트 그리드, IoT 및 의료 진단 시스템 등 다양한 환경에서 동형암호를 적용한 연구가 이루어지고 있다. 본 논문에서는 다양한 환경에서 동형암호 기반의 인공지능 학습 데이터 보안 시스템을 제안한 기존 연구들의 동향 및 문제점을 분석하고 향후 연구 방향을 제시한다.

Ⅰ. 서 론

최근 정보통신 및 인터넷의 발전으로 스마트 그리드, IoT(Internet of Things)[1] 및 의료 진단 시스템 등 다양한 환경에서 서비스를 이용하는 사용자의 수가 증가함에 따라 데이터도 방대하게 생성되고 있다. 이러한 빅데이터를 효율적으로 처리 및 분석하기 위해 인공지능을 적용하는 연구가 지속해서 이루어지고 있다. 그러나 인공지능은 모델 구축 단계에서 사용자의 원본 데이터를 분석하기 때문에 개인정보가 침해된다는 문제점이 존재한다. 이러한 문제를 보완하기 위해 사용자들은 자신의 데이터를 암호화할 수 있지만 인공지능은 일반적인 암호시스템으로 암호화된 데이터를 분석할 수 없어 다량의 데이터를 복호화해야 한다. 이는 효율적인 데이터 처리에 대한 어려움이 있으며 기존의 개인정보 유출 문제를 해결하지 못한다. 이러한 문제들을 해결하기 위해 암호화된 데이터의 연산 및 처리 가능한 동형암호를 기반으로 연구가 활발히 진행되고 있다.

본 논문에서는 2019년에 제안된 Yao 등의 Paillier 암호를 이용한 스마트 그리드 에너지 도난감지 시스템[2]과 2020년에 제안된 Zhang 등의 Paillier 암호를 이용한 산업 IoT의 개인정보 보호 시스템[3] 및 Zhang 등의 OU(Okamoto-Uchiyama) 암호를 이용한 클라우드 기반 의료 진단 시스템[4]의 보안 연구 동향을 분석한다. 또한 이들 방식의 문제점을 분석하고 향후 연구 방향을 제시한다.

Ⅱ. 동형암호

동형암호는 암호화된 데이터에 대해 연산을 가능하게 하는 암호시스템으로 평문의 연산 후 암호화한 결과와 암호화된 상태에서의 연산 결과가 동일하다는 특성을 가진다. 또한 암호화된 형식으로 연산을 수행한 최종적인 결과는 해당 암호문의 복호 가능한 키 소유자만이 볼 수 있어 데이터의 기밀성을 보장할 수 있다. 동형암호는 아래 그림 1과 같이 1978년 RSA 암호를 시작으로 BGN, Gen 및 BGV 암호까지 활발히 연구되고 있으며 연산 가능 범위에 따라 다음과 같이 3가지로 분류된다.

- 부분동형암호: 횟수 제한 없이 한가지 유형의 연산만 가능하며 ElGamal, Paillier 및 OU 암호 등이 존재한다.
- 준동형암호: 제한된 횟수로 일부 연산이 가능하며 SYY 및 BGN 암호 등이 존재한다.
- 완전동형암호: 횟수 제한 없이 모든 연산이 가능하며 Gen 및 BGV 암호 등이 존재한다.

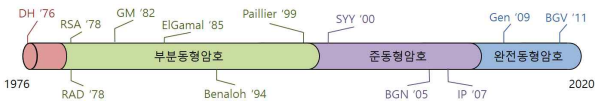


그림 1. 동형암호 연대표.

본 논문에서는 대표적인 부분동형암호인 Paillier 및 OU 암호를 분석한다.

2.1. Paillier 암호

Paillier 암호는 암호화된 데이터에 대해 곱셈 연산이 가능한 동형암호이며, 다음과 같은 3단계의 알고리즘으로 구성된다.

- $KeyGen(1^{\lambda}) \rightarrow (pk, sk)$: 보안 파라미터 λ로부터 공개키 $pk=(n,g)$와 개인키 $sk=(p,q)$를 생성한다.
- $Enc(pk,m) \rightarrow c = g^m r^n \bmod n^2$: 평문 m과 공개키 pk로 암호문 c를 계산한다.
- $Dec(sk,c) \rightarrow m = \frac{L(c^{\lambda(p-1,q-1)} \bmod n^2)}{L(g^{\lambda(p-1,q-1)} \bmod n^2)} \bmod n$: 암호문 c와 개인키 sk로 평문 m을 계산한다.

2.2. OU(Okamoto-Uchiyama) 암호

OU 암호는 암호화된 데이터에 대해 곱셈 연산이 가능한 동형암호로, 다음과 같은 3단계의 알고리즘으로 구성된다.

- $KeyGen(1^{\lambda}) \rightarrow (pk, sk)$: 보안 파라미터 λ로부터 공개키 $pk=(n,g,h,l)$와 개인키 $sk=(p,q)$를 생성한다.
- $Enc(pk,m) \rightarrow c = g^m h^r \bmod n$: 평문 m과 공개키 pk로 암호문 c를 계산한다.
- $Dec(sk,c) \rightarrow m = \frac{L(c^{p-1} \bmod p^2)}{L(g^{p-1} \bmod p^2)} \bmod p$: 암호문 c와 개인키 sk로 평문 m을 계산한다.

Paillier 및 OU 암호는 각각 $E(m_1)\cdot E(m_2)\,(\mathrm{mod}\,n^2)=E(m_1+m_2\,(\mathrm{mod}\,n))$ 및 $E(m_1)\cdot E(m_2)\,(\mathrm{mod}\,n)=E(m_1+m_2\,(\mathrm{mod}\,p))$와 같이 암호화된 데이터의 곱셈 결과와 평문을 더하여 암호화한 결과가 같다는 특성을 가진다.

Ⅲ. 동형암호를 이용한 인공지능 학습 데이터 보안 시스템

본 논문에서는 제안된 Yao 등, Zhang 등 및 Zhang 등의 동형암호를 이용한 보안 시스템을 통해 최근 보안 연구 동향을 분석한다.

3.1. Yao 등의 스마트 그리드에서의 에너지 도난감지 시스템

스마트 미터가 제공하는 사용자의 실시간 에너지 소비 데이터가 보안 위협에 노출될 경우 사생활 침해 및 에너지 도난 문제가 발생할 수 있다. Yao 등은 이러한 문제들의 해결 및 에너지 소비 데이터의 비정상적인 동작 감지를 위해 Paillier 암호 및 CNN(Convolutional Neural Network)을 이용한 에너지 도난감지 시스템을 그림 2와 같이 제안하였다. 각 에너지 사용자는 Paillier 암호로 자신의 에너지 소비 데이터를 암호화하며 최종적으로 암호화된 데이터를 집계한 결과가 관제 센터로 보내진다. 관제 센터는 집계 결과만 얻기 때문에 사용자 개별의 데이터 노출을 막음으로써 개인정보 보호 및 데이터 기밀성을 보장할 수 있다. 그러나 Yao 등의 시스템은 각 객체의 인증 없이 에너지 소비 데이터를 교환하기 때문에 악의적인 에너지 사용자가 거짓 에너지 소비를 신고하여 스마트 그리드에 경제적 손실을 초래할 수 있다는 문제점이 존재한다. 따라서 안전한 에너지 서비스를 위한 객체 간의 상호인증 과정이 필요하다[5].

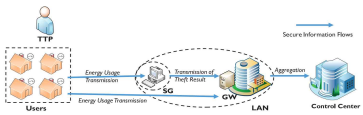
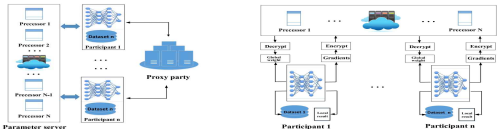


그림 2. Yao 등의 Paillier 암호 기반 에너지 도난감지 시스템.

3.2. Zhang 등이 제안한 산업 IoT 환경의 개인정보 보호 딥러닝 시스템

산업 IoT는 상호연결된 분산 기기에서 생성되는 데이터를 효율적으로 처리하기 위해 분산된 데이터를 병렬로 처리하는 연합 딥러닝을 많이 사용한다. 그러나 연합 딥러닝은 학습 권한을 가진 로컬 참가자들의 데이터로 학습되기 때문에 개인정보 유출 문제가 존재한다. Zhang 등은 이러한 문제를 해결하기 위해 Paillier 암호를 이용한 두 가지 개인정보 보호 비동기 딥러닝 시스템을 그림 3과 같이 제안하였다. 서버가 Paillier 암호 공개키로 암호화된 글로벌 파라미터를 전달하면 그림 2-(a)는 프록시가 복호화한 후 로컬 참가자들에게 전달하고 그림 2-(b)는 각 로컬 참가자가 복호화한다. 로컬 참가자들은 글로벌 파라미터로 학습된 로컬 파라미터들을 Paillier 암호 공개키로 암호화하여 서버로 보내고 서버는 해당 로컬 파라미터들을 집계하여 최종적인 암호화된 글로벌 파라미터로 업데이트한다. 로컬 참가자 이외의 객체는 글로벌 파라미터 이외의 데이터를 얻을 수 없으므로 참가자들의 개인정보를 보호할 수 있다. 그러나 Zhang 등의 시스템은 서버와 로컬 참가자 간의 인증 없이 학습을 진행하기 때문에 악의적인 로컬 참가자가 잘못된 학습 데이터를 주입하여 딥러닝 시스템을 손상시킬 수 있다는 문제점이 존재한다. 따라서 서버와 로컬 참가자 간의 상호인증 후 안전한 학습 과정이 이루어져야 한다.



(a) (b)

그림 3. Zhang 등의 Paillier 암호 기반 개인정보 보호 딥러닝 시스템.

3.3. Zhang 등이 제안한 클라우드 환경의 의료 진단 시스템

클라우드 서버에 의료 데이터를 저장하는 기존의 시스템은 클라우드를 완전히 신뢰할 수 없기 때문에 환자의 개인정보 유출 문제가 존재한다. Zhang 등은 이러한 문제점의 해결 및 효율적인 진단 결정을 위해 OU 암호 및 다중분류 SVM(Support Vector Machine)을 이용한 의료 진단 시스템을 그림 4와 같이 제안하였다. 먼저 데이터 사용자는 환자 데이터의 SVM 결정 함수를 자신의 OU 암호 공개키로 암호화한 후 클라우드 서버에 전송한다. 클라우드 서버는 데이터 소유자와 협력하여 데이터 사용자의 OU 암호 공개키로 암호화된 진단 결과를 데이터 사용자에게 전달한다. 암호화된 진단 결과는 데이터 사용자만이 복호화하여 결과를 확인할 수 있으므로 환자 데이터의 기밀성을 보장할 수 있다. 그러나 Zhang 등의 시스템은 각 객체 간의 인증 없이 환자를 진단 결정하기 때문에 악의적인 데이터 소유자가 잘못된 진단 결과를 데이터 사용자에게 전달하여 환자의 생명을 위협할 수 있다는 문제점이 존재한다. 따라서 데이터 교환 전 각 객체 사이의 인증 과정이 필요하다[6].

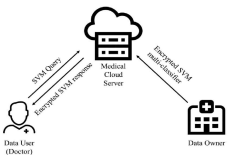


그림 4. Zhang 등의 OU 암호 기반 의료 진단 시스템.

Ⅳ. 결론

최근 스마트 그리드, IoT 및 의료 진단 시스템 등 다양한 환경에서 동형암호를 이용한 연구가 활발히 이루어지고 있다. 본 논문에서는 Yao 등, Zhang 등 및 Zhang 등의 동형암호를 이용한 보안 시스템을 분석하였다. 이들이 제안한 보안 시스템들은 데이터에 대한 안전성 및 처리 효율성만 다루고 있으며 데이터 주체 간의 인증은 다루지 않고 있다. 이러한 문제점을 보완하여 향후 모바일 IoT 환경에서 Paillier 및 연합 학습을 이용한 데이터 객체 사이의 상호인증을 구현하는 보안 연구를 진행할 것이다.

참 고 문 헌

[1] Lee, J., Yu, S., Park, K., Park, Y., and Park, Y. “Secure Three-Factor Authentication Protocol for Multi-Gateway IoT Environments,” Sensors, vol. 19, no. 10, pp. 1-25, 2019.

[2] Yao, D., Wen, M., Liang, X., Fu, Z., Zhang, K., and Yang, B. “Energy Theft Detection With Energy Privacy Preservation in the Smart Grid,” IEEE Internet of Things Journal, vol. 6, no. 5, pp. 7659-7669, Oct. 2019.

[3] Zhang, X., Chen, X., Liu J. K., and Xiang, Y. “DeepPAR and DeepDPA: Privacy Preserving and Asynchronous Deep Learning for Industrial IoT,” IEEE Transactions on Industrial Informatics, vol. 16, no. 3, pp. 2081-2090, Mar. 2020.

[4] Zhang, M., Song, W., and Zhang, J. “A Secure Clinical Diagnosis With Privacy-Preserving Multiclass Support Vector Machine in Clouds,” IEEE Systems Journal, pp.1-12, 2020.

[5] Yu, S., Park, K., Lee, J., Park, Y., Park, Y., Lee, S., and Chung, B. “Privacy-Preserving Lightweight Authentication Protocol for Demand Response Management in Smart Grid Environment,” Applied Sciences, vol. 10, no. 5, pp. 1-26, 2020.

[6] Son, S., Lee, J., Kim, M., Yu, S., Das A. K., and Park, Y. “Design of Secure Authentication Protocol for Cloud-Assisted Telecare Medical Information System Using Blockchain,” IEEE Access, vol. 8, pp. 192177-192191, 2020.