프라이버시 보존 데이터 학습을 위한 고효율 동형 암호 기법

심혜연 [1], 전유란 [1], 이일구 [2]
[1] 성신여자대학교 미래융합기술공학과 박사과정
[2] 성신여자대학교 융합보안공학과 교수

220237062@sungshin.ac.kr, 220247016@sungshin.ac.kr, iglee@sungshin.ac.kr

High-Efficiency Homomorphic Encryption Techniques for Privacy-Preserving Data Learning

Hye Yeon Shim[1], Yu-Ran Jeon[1], Il-Gu Lee[1,2]
[1]Dept. of Future Convergence Technology Engineering, Sungshin Women's University
[2]Dept. of Convergence Security Engineering, Sungshin Women's University

요 약

최근 인공지능 기술의 발전과 함께 기계학습과 빅데이터를 융합한 서비스가 증가하게 되었고, 무분별한 데이터 수집과 학습으로 인한 개인정보 유출 위험도가 커졌다. 따라서 프라이버시를 보호하면서 기계학습을 수행할 수 있는 기술이 중요해졌다. 동형암호 기술은 정보 주체자의 개인정보 기밀성을 유지하면서 기계학습을 할 수 있는 방법 중 하나이다. 그러나 평문 크기에 비례하여 암호문 크기와 연산 결과의 노이즈가 커지는 동형암호의 특징으로 인해 기계학습 모델의 예측 정확도가 감소하고 학습 시간이 오래 소요되는 문제가 발생한다. 본 논문에서는 부분 동형암호화된 데이터셋으로 로지스틱 회귀 모델을 학습할 수 있는 기법을 제안한다. 실험 결과에 따르면 제안하는 기법이 종래 기법보다 예측 정확도를 59.4% 향상시킬 수 있었고, 학습 소요 시간을 63.6% 개선할 수 있었다.

1. 서론

4 차 산업혁명 시대가 도래함에 따라 빅데이터와 기계학습을 결합한 서비스가 기하급수적으로 증가하고 있다. 특히 의료, 금융 등과 같이 사용자와 밀접하게 연계된 분야에서 사용자의 정보를 이용한 학습을 통해 이전보다 질 좋은 서비스 제공이 가능하게 됐다[1],[2]. 그러나 기계학습 과정에서 개인정보가 서버에 노출될 수 있는 문제로 인해 정보 주체자의 프라이버시 유출 우려가 커지고 있다[3]. 이러한 문제를 해결하기 위해 차등 프라이버시 기술을 활용하여 데이터를 보호하며 학습하는 기술에 관한 연구가 진행되고 있다. 그러나 차등 프라이버시는 유틸리티와 프라이버시의 트레이드오프 문제가 있다[4]. 그리고 데이터를 보호하기 위한 또 다른 방법으로는 AES(Advanced Encryption Standard), RSA(Rivest Shamir Adleman)와 같은 암호 알고리즘이 있지만, 기계학습에 사용할 수 없는 한계가 있다.

이러한 유틸리티와 프라이버시의 트레이드오프 문제를 해결하기 위해 프라이버시 보존 데이터 학습이 가능한 동형암호 기술이 연구되고 있다. 동형암호는 암호화된 상태로 덧셈, 곱셈과 같은 사칙연산이 가능하고, 반복 연산을 요구하는 기계학습을 수행할 수 있다[5]. 그러나 동형암호로 정보를 암호화하면 암호문의 크기가 원본 정보에 비해 지나치게 커지는 문제가 있다[6]. 암호문 연산은 많은 자원을 소모하고, 상대적으로 긴 시간이 소요되기 때문에 실제 산업 환경에는 동형암호 학습을 적용하기 어렵다. 또한 반복되는 암호문 연산의 횟수가 늘어날수록 노이즈가 증가하기 때문에 기계학습의 epoch 가 클수록 암호문에 포함되는 노이즈가 비약적으로 커진다[7]. 반복된 연산으로 증가한 노이즈는 모델 학습 시 가중치와 편향의 업데이트에 영향을 주기 때문에 모델 정확도 감소를 야기한다.

본 논문에서는 프라이버시 보존 데이터 학습을 위해 동형암호를 활용할 때 기계학습 시 발생하는 속도

와 정확도 문제를 해결하기 위해서 학습에 사용되는 데이터셋의 일부 feature 를 동형암호로 암호화한 데이터셋으로 로지스틱 회귀 학습을 수행하는 기법을 제안한다. 본 논문의 기여점은 다음과 같다

- 프라이버시 보존 데이터 학습을 위한 부분 암호화 기반 고효율 동형 암호 기법을 제안한다.
- 종래의 동형암호 기술과 부분 암호 기반의 동형암호 기술의 성능을 평가하는 프레임워크를 제안한다.

본 논문의 구성은 다음과 같다. 2 장에서는 관련 연구를 비교 및 분석하고, 3 장에서는 부분 암호화된 데이터셋에 대한 로지스틱 회귀 학습 기법을 제안한다. 4 장에서는 실험 환경과 설정에 관하여 설명하고, 실험 결과를 분석한다. 마지막으로 5 장에서는 결론을 맺는다.

2. 관련 연구

본 장에서는 동형암호화된 데이터셋의 고효율 기계학습 기법 관련한 종래 대표 연구의 기여점과 한계점을 비교 및 분석한다.

Sgaglione 등[8]은 서명기반 데이터 검사를 통해 동형암호로 암호화된 정보에 포함된 공격을 성공적으로 검출했다. 그러나 시간 복잡도를 최적화하지 않았기 때문에 실제 서비스에 적합하지 않은 문제가 있다.

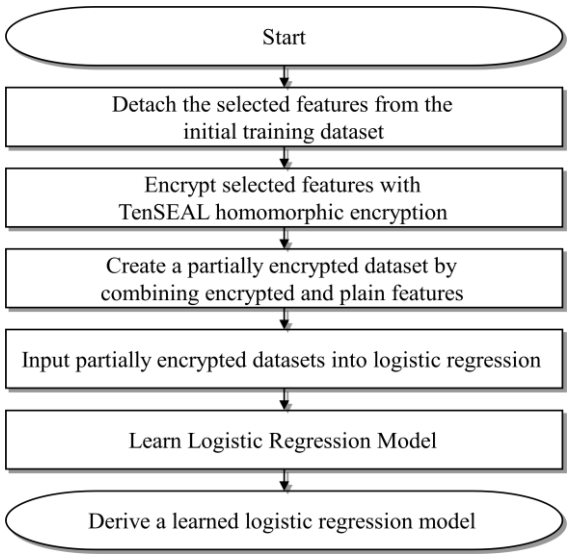
Spathoulas 등[9]은 침입탐지시스템에서 수집된 알림을 활용한 동형연산 기반 클러스터링 모델을 제안했다. 제안하는 모델은 서버가 처리할 수 없는 연산을 클라이언트가 대신 연산하는 협력적 연산 방식을 적용하여 동형암호 연산의 처리 효율을 개선했다. 그러나, 서버와 클라이언트 사이의 통신량이 증가하는 한계점이 있다.

Alabdulatif 등[10]은 스마트 시티 기술을 위한 완전 동형암호화 기반 기계학습 기법을 제안했다. 제안 모델은 클라우드 컴퓨팅 자원을 활용하여 동형암호화 연산을 병렬적으로 수행함으로써 모델 학습 시간을 줄였다. 그러나 병렬연산으로 인해 CPU 및 메모리 오버헤드가 증가하는 한계가 있다.

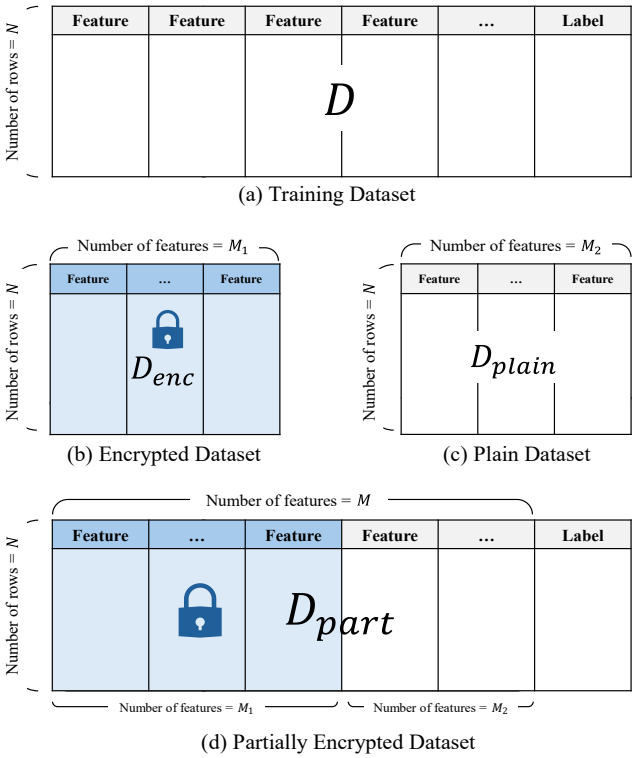
3. 제안하는 아이디어

본 장에서는 일부 feature 를 동형암호화한 학습 데이터셋을 이용하여 예측 모델을 생성하는 기법을 제안한다. 제안하는 기법에서 동형암호는 마이크로소프트의 SEAL 기반의 기계학습 가능한 완전 동형암호화 알고리즘인 TenSEAL[12]을 사용했고, 가장 단순한 모델인 로지스틱 회귀 기반의 기계학습 모델을 이용했다.

제안 기법의 전체 동작 흐름은 그림 1 과 같다. 우선 그림 2(a)와 같이 초기 학습 데이터셋 D의 feature 중 프라이버시가 중요하거나 기밀성 유지가 필요한 feature 와 그렇지 않은 feature 를 구분한다. 구분된 feature 의 집합 중 기밀성 유지가 필요한 집합은 동형암호인 TenSEAL 을 통해 암호화하여 그림 2(b)와 같은 암호화된 데이터셋 D_{enc}를 구성한다. 반대로 분리된 feature 집합 중 기밀 유지가 필요하지 않은 feature 의 집합을 그림 2(c)와 같이 일반 데이터셋 D_{plain}으로 구분한다. 이후 D_{enc}와 D_{plain}을 결합하여 그림 2(d)와 같이 부분 암호화된 데이터셋 D_{part}를 생성한다. 그다



(그림 1) 제안하는 기법의 동작 흐름도



(그림 2) 제안하는 기법의 학습 데이터셋 구조

Algorithm 1 Learning a Logistic Regression Model with a Partially Encrypted Dataset

Input dataset D_{part}, epoch E
Output weight W, bias B

row_number N = length(D_{part})
Initialize weight W, bias B
W_{enc}, W_{plain} = DataSeparation(W)
W_{enc}, B = TenSEAL_Encryption(W_{enc}, B)
sigmoid(x) = 0.5 + 0.197 * x - 0.0004 * x^3
while e $\geq$ E do
  Initialize $delta_{W_{enc}}$, $delta_{W_{plain}}$, $delta_B$
  while row >= N do
    y_{true} = D_{part}[row, label]
    y_{out} = sigmoid((D_{part}[row, enc].dot(W_{enc})) +($D_{part}[row, plain].dot(W_{plain})$) + B)
    $delta_{W_{enc}}$ += D_{part}[row, enc] * (y_{out} - y_{true})
    $delta_{W_{plain}}$ += D_{part}[row, plain] * (y_{out} - y_{true})
    $delta_B$ += y_{out} - y_{true}
    row ++
  end while
  W_{enc} -= $delta_{W_{enc}}$ / N + 0.05 * W_{enc}
  W_{plain} -= $delta_{W_{plain}}$ / N + 0.05 * W_{plain}
  B -= $delta_B$ / N
  e ++
end while
W_{enc}, B = TenSEAL_decryption(W_{enc}, B)
W = concatenate(W_{enc}, W_{plain})

음에 생성된 D_{part}를 이용하여 로지스틱 회귀 모델을 학습하고 라벨을 예측할 수 있는 모델을 생성한다.

부분 암호화된 데이터셋 D_{part}를 이용한 로지스틱 회귀 모델 기반의 학습 알고리즘은 종래의 로지스틱 회귀 알고리즘을 수정하여 생성하였으며, 수정된 로지스틱 회귀 모델 학습은 Algorithm 1 과 같은 구조로 동작한다. 먼저 초기 weight W와 bias B를 무작위 값으로 설정한다. 그리고 W를 D_{enc}에 대한 weight W_{enc}와 D_{plain}에 대한 weight W_{plain}으로 분리한다. 그다음 암호화된 데이터에 대한 연산을 수행하기 위해 TenSEAL 동형암호를 통해 W_{enc}와 B 각각을 암호화한다. 이후 W_{enc}, W_{plain}, B를 업데이트 해주기 $delta_{W_{enc}}$, $delta_{W_{plain}}$, $delta_B$를 각각 선언하고 초기화한다. 다음으로 D_{part}의 모든 데이터에 대해서 라벨 예측을 수행하여 y_{out}을 도출한다. 이때 동형암호 연산을 가능하게 하기 위해서 비선형성을 갖는 sigmoid 함수를 [12]에서 제안한 다항식으로 표현된 sigmoid 함수로 대체한다. 이후 y_{out}을 실제 label 값인 y_{true}와 비교하고, 이를 토대로 $delta_{W_{enc}}$, $delta_{W_{plain}}$, $delta_B$에 대한 업데이트를 D_{part}의 행 개수만큼 진행한다. 한 번의 epoch에서 업데이트된 $delta_{W_{enc}}$, $delta_{W_{plain}}$, $delta_B$는 W_{enc}, W_{plain}, B 각각의 업데이트에 사용된다. 모든 epoch 를 마치고 암호화된 W와 B를 복호화하면 최종적으로 학습된 로지스틱 회귀 모델이 도출된다.

(표 1) 전체 암호화된 데이터셋과 부분 암호화된 데이터셋에 대한 성능 평가

	Prediction accuracy (%)	Time(s)		Memory usage(MB)	
		Dataset encryption	Training	Dataset encryption	Training
Conventional	61.83	15.03	706.52	1,985.32	1.63
Proposed	98.56	11.67	257.13	1,714.50	0.05

4. 성능 평가 및 분석

성능 비교 분석을 위해 동일한 환경에서 부분 암호화 기반의 제안하는 동형 암호 모델과 종래의 동형 암호 모델을 구현했다. 실험에 사용한 데이터셋은 CIC-IDS2017 이며, 균등한 라벨 분포를 갖는 2,000 개의 행을 추출했다. 실험의 성능 지표는 메모리 사용량, 모델 예측 정확도, 소요된 시간이며, 100 회 반복을 통해 성능을 측정했다. 동형암호로 암호화한 feature는 5 개이고, 로지스틱 회귀 학습의 epoch 는 5 이다.

표 1 은 전체 데이터셋을 동형암호로 암호화했을 때, 제안하는 기법과 종래 기법의 성능을 비교한 결과이다. 실험 결과에 따르면 전체 암호화한 데이터셋을 사용한 종래 기법보다 부분 암호화한 데이터셋을 사용하는 제안 기법이 59.4% 높은 예측 정확도를 보였다. 또한, 학습에 사용할 데이터셋을 생성하는데 걸리는 시간은 제안하는 기법이 종래 기법보다 22.36% 단축시켰다. 특히, 로지스틱 회귀 모델 학습에 소요된 시간의 경우 제안 기법은 평균 257 초이고 종래 기법은 평균 707 초로, 제안하는 기법이 종래보다 63.61% 학습 시간을 줄일 수 있었다. 메모리 사용량의 경우에는 학습 데이터를 암호화할 때와 로지스틱 회귀 모델을 학습할 때 모두에서 제안하는 기법의 메모리 사용량이 적었다. 특히, 데이터셋을 암호화하는 과정에서 종래 기법보다 제안하는 기법이 13.64% 메모리 사용량을 줄일 수 있었다. 전체적으로 제안하는 기법이 종래 기법보다 항상 높은 성능을 보이는 점을 알 수 있다.

5. 결론

본 논문에서는 부분 암호화된 데이터셋을 통해 로지스틱 회귀 모델을 학습하는 기법을 제안했다. 실험 결과에 따르면 전체 데이터셋을 동형암호로 암호화한 경우보다 제안하는 기법이 예측 정확도를 약 59.4% 개선했고, 학습에 소요되는 시간을 약 63.61% 단축했다. 또한 데이터 암호화 시간 및 메모리 사용량을 줄일 수 있음을 보였다.

제안하는 기법은 로지스틱 회귀 모델을 통해 구현되었으며, 단순한 상황을 가정하여 실험을 진행했기

때문에 제안하는 기법의 효과를 명확하게 보여주지 못했다는 한계가 있다. 후속 연구에서는 제안하는 부분 암호화된 데이터셋을 이용한 기계학습 방식을 DNN, CNN 과 같은 딥러닝에 적용하여, 복잡한 모델에서의 효과를 증명할 계획이다.

Acknowledgement

본 논문은 2024 년도 산업통상자원부 및 한국산업기술진흥원의 산업혁신인재성장지원사업 (RS-2024-00415520)과 과학기술정보통신부 및 정보통신기획평가원의 ICT 혁신인재 4.0 사업의 연구결과로 수행되었음 (No. IITP-2022-RS-2022-00156310)

참고문헌

[1] Mohammad Shehab, Laith Abualigah, Qusai Shambour, Muhannad A. Abu-Hashem, Mohd Khaled Yousef Shambour, Ahmed Izzat Alsalibi, and Amir H. Gandomi, "Machine learning in medical applications: A review of state-of-the-art methods," Computers in Biology and Medicine, 145, 2022.

[2] Satish Kumar, Dipasha Sharma, Sandeep Rao, Weng Marc Lim, and Sachin Kumar Mangla, "Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research," Annals of Operations Research, pp. 1-44, 2022.

[3] Bo Liu, Ming Ding, Sina Shaham, Wenny Rahayu, Farhad Farokhi, and Zihuai Lin, "When machine learning meets privacy: A survey and outlook," ACM Computing Surveys (CSUR), 54, 2, pp. 1-36, 2021.

[4] Maria Rigaki, and Sebastian Garcia, "A survey of privacy attacks in machine learning," ACM Computing Surveys, 56, 4, pp. 1-34, 2023.

[5] Robert Podschwadt, Parsa Ghazvinian, Mohammad GhasemiGol, and Daniel Takabi, "Memory Efficient Privacy-Preserving Machine Learning Based on Homomorphic Encryption," International Conference on Applied Cryptography and Network Security. Abu Dhabi, United Arab Emirates, 2024, pp. 313-339.

[6] Martin Albrecht, Melissa Chase, Hao Chen, Jintai Ding, Shafi Goldwasser, Sergey Gorbunov, Shai Halevi, Jeffrey Hoffstein, Kim Laine, Kristin Lauter, Satya Lokam, Daniele Micciancio, Dustin Moody, Travis Morrison, Amit Sahai, and Vinod Vaikuntanathan, "Homomorphic encryption standard," Protecting privacy through homomorphic encryption, pp. 31-62, 2021.

[7] Luis Bernardo Pulido-Gaytan, Andrei Tchernykh, Jorge M. Cortés-Mendoza, Mikhail Babenko, and Gleb Radchenko, "A survey on privacy-preserving machine learning with fully homomorphic encryption," Latin American High Performance Computing Conference, Cuenca, Ecuador, 2020, pp. 115-129.

[8] Luigi Sgaglione, Luigi Coppolino, Salvatore D'Antonio, Giovanni Mazzeo, Luigi Romano, Domenico Cotroneo, and Andrea Scognamiglio, "Privacy preserving Intrusion Detection via Homomorphic Encryption," 2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), Napoli, Italy, 2019, pp. 321-326.

[9] Georgios Spathoulas, Georgios Theodoridis, and Georgios-Paraskevas Damiris, "Using homomorphic encryption for privacy-preserving clustering of intrusion detection alerts," International Journal of Information Security, 20, 3, pp. 347-370, 2021.

[10] Abdulatif Alabdulatif, Ibrahim Khalil, Heshan Kumarage, Albert Y. Zomaya, and Xun Yi, "Privacy-preserving anomaly detection in the cloud for quality assured decision-making in smart cities," Journal of Parallel and Distributed Computing, 127, pp. 209-223, 2019.

[11] Ayoub Benaissa, Bilal Retiat, Bogdan Cebere, and Alaa Eddine Belfedhal, "Tenseal: A library for encrypted tensor operations using homomorphic encryption," arXiv preprint arXiv:2104.03152, 2021.

[12] Hao Chen, Ran Gilad-Bachrach, Kyoohyung Han, Zhicong Huang, Amir Jalali, Kim Laine, and Kristin Lauter, "Logistic regression over encrypted data from fully homomorphic encryption," BMC medical genomics 11, pp. 3-12, 2018.