연합학습을 통한 보안성 높은 무선통신 변조기술 탐지 모델에 관한 연구

서중하, *우태희, 박찬호, 강준혁
한국과학기술원, *충남대학교
junghaa.seo@kaist.ac.kr, seeles@o.cnu.ac.kr, kmapark@kaist.ac.kr, jhkang@kaist.edu

A Study on the Secured Radio Modulation Classification via Federated Learning

Seo Junghaa, *Woo Taehee, Park Chanho, Kang Joonhyuk
Korea Advanced Institute of Science and Technology, *Chungnam National University

요 약

본 논문은 각각의 수신기에서 수집된 데이터로 개별 학습된 모델의 가중치 파라미터를 중앙 서버로 보낸 후 글로벌 수신기 모델을 학습하는 과정의 반복을 통해 신호의 변조기술 탐지 성능을 향상시키는 연합학습 모델을 제안한다. 그리고 실험을 통해 개별 수신기가 악의적 사용자에 의해 오염되었을 때 글로벌 수신기 모델의 보안성을 확보할 수 있음을 확인하였다.

Ⅰ. 서 론

심층 신경망(DNN), 순환신경망(RNN) 등 다양한 기계학습 방법을 통해 인공지능은 여러 분야에서 인간의 인지능력을 뛰어넘는 성과를 보여주고 있다. 무선통신 분야에서도 인공지능을 통해 비선형적 채널 환경을 예측하고[1], 신호를 탐지하며[2], 기존 통신이론 기반의 분리된 소스코딩, 채널코딩을 하나의 시스템으로 연결하려는 End-to-end 오토인코더[3] 등이 심도있게 연구되고 있다. 기존의 기계학습 방법은 대용량의 데이터를 고성능 학습용 서버를 통해 처리/학습한다. 신호 변조기술 탐지 모델을 예로 들면, 수신기에서 수집된 신호 데이터는 학습용 서버로 수집, 학습된 탐지 모델을 다시 수신기로 보내게 된다. 이렇게 데이터가 이동하는 동안 네트워크 자원이 소모되며, 민감한 데이터의 경우 악의적 사용자가 이를 유출시키거나 학습 데이터를 오염시킬 수도 있다. 또한 서버에서 학습을 담당하기 때문에 학습 방법, 데이터의 양에 따라 많은 시간이 소요된다는 단점도 존재한다. 이를 해결하기 위해 연합학습(Federated Learning)[4] 개념이 등장하였다, 이 개념은 서버에 비해 성능이 다소 부족하지만 여러 대의 분산된 컴퓨팅 자원을 활용하여 각각 로컬 모델을 학습한 후 학습된 모델의 정보만을 중앙 서버에서 통합하여 글로벌 모델을 만드는 방법이다. 본 논문에서는 각각의 수신기에서 수집된 데이터로 개별 학습된 모델의 가중치 파라미터를 중앙 서버로 보낸 후 글로벌 수신기 모델을 학습하는 과정의 반복을 통해 신호의 변조기술 탐지 성능을 향상시키는 연합학습 모델을 제안한다, 그리고 실험을 통해 개별 수신기가 악의적 사용자에 의해 오염되었을 때 글로벌 수신기 모델의 보안성이 높음을 확인하였다.

Ⅱ. 본론

본 장에서는 연합학습 소개 및 연합학습을 기반으로 무선통신 변조기술을 탐지하는 수신기 모델을 제안하며, 글로벌 수신기 모델의 보안성 측정 실험을 통해 결과를 분석한다.

2-1 연합학습

일반적인 기계학습이 대용량의 학습용 데이터를 고성능 서버를 통해 학습하는 것과 달리 연합학습은 그림 1 과 같이 서버가 아닌 분산된 개별 컴퓨팅 자원에서 데이터를 처리하고 학습한다. 그리고 학습된 모델을 서버로 보내면(aggregation) 서버는 개별 컴퓨팅 자원에서 받은 모델을 글로벌 모델로 종합한 후 다시 개별 컴퓨팅 자원으로 보낸다(broadcast). 개별 컴퓨팅 자원에서는 서버와 연결되어 있지 않더라도 이전에 서버로부터 받은 글로벌 모델로 구동할 수 있으며, 학습 데이터가 아닌 모델만을 서버로 보내기 때문에 대용량의 네트워크를 유지할 필요가 없다. 또한 민감한 학습 데이터의 경우 내부에서만 처리되기 때문에 보안적 측면에서도 안정되어 있다. 서버에서는 회차별 개별 컴퓨팅 자원을 선정하고 이들로부터 받은 모델을 글로벌 모델로 종합을 하는 알고리즘을 통해 상대적으로 적은 데이터 처리만 하면 된다. 더욱 효율적인 네트워크 전송을 위해 가중치 파라미터를 압축하거나, 모델의 가중치 변화값을 전송, 손실함수의 경사도값을 전송하는 방법 등도 널리 연구되고 있다. 다음 장에서는 가중치 변화값 전송을 통해 효율적으로 네트워크 사용량을 줄인 연합학습 기반 무선통신 변조기술 탐지 모델을 제안한다.

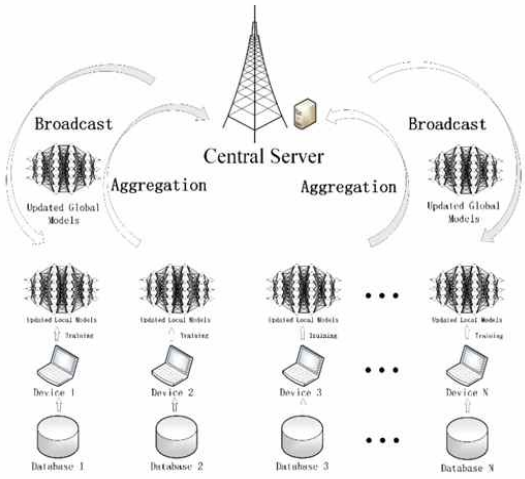

그림 1 연합학습 과정

2-2 연합학습 기반 무선통신 변조기술 탐지 모델

서버로부터 받은 $k-1$번째 글로벌 모델의 가중치 w_{k-1}를 기반으로 i번 수신기에서는 수신 신호 $r(t)$와 변조기술 y를 통해 수신기 모델 $f(w_k^i;)$ 를 학습하고, 가중치 변화값

$$\Delta w_k^i = w_k^i - w_{k-1},$$
$$\arg\min_{w_k^i} L(f(w_k^i; r(t)), y) \quad (1)$$

을 계산하여 서버로 보낸다. 이때, f는 DNN 기반 수신기 모델, L은 손실함수로, catrgorical cross-entropy를 사용한다. 서버는 개별 수신기로부터 받은 k번째 가중치 변화값들을 종합(Federated averaging)하여 글로벌 가중치 변화값

$$\Delta w_k = \frac{1}{N}\sum_{i=1}^{n} \Delta w_k^i \quad (2)$$

을 개별 수신기로 보내면 i번 개별 수신기는 다시 수신기 모델 $f(w_{k+1}^i;)$을 학습하고, 가중치 변화값

$$\Delta w_{k+1}^i = w_{k+1}^i - w_k^i + \Delta w_k \quad (3)$$

을 서버로 보내는 반복적인 과정을 진행한다.

실험에는 24개의 변조기술에 상응하는 무선신호[5]를 VGG[5] 구조로 학습되는 5개의 개별 수신기 모델을 사용하고 {1, 5}의 개별 수신기 반복학습 횟수(local epoch), {32, 64, 128}의 배치 사이즈로 구분하였다. 그림 2 는 학습 회차에 따라 글로벌 수신기 모델의 분류정확도를 측정한 것으로, 개별 수신기의 반복학습 횟수가 많을수록, 배치 사이즈가 작을수록 분류정확도가 높지만 학습시간도 길어진다.

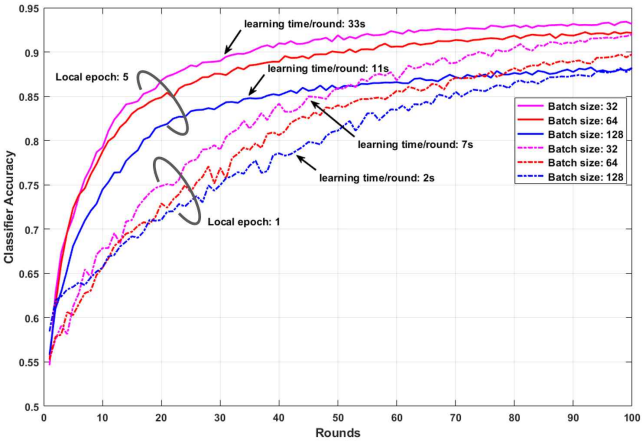

그림 2 연합 학습 회차별 글로벌 수신기 분류정확도 측정

2-3 위협 모델

연합학습은 서버가 개별 수신기의 학습 데이터를 볼 수 없고, 수신기를 통제하지 않기 때문에 악의적 사용자에 의해 오염되었을 때 학습 데이터의 무결성을 보장할 수 없다. 이 경우 오염된 데이터로 학습된 개별 수신기 모델이 글로벌 모델의 보안성에 영향을 미치는지를 확인하기 위해 개별 수신기의 학습데이터를 오염된 것(poisoning attack)으로 가정하였다.

실험에는 5개의 개별 수신기 모델이 1회의 반복학습 횟수와 32의 배치 사이즈로 학습되며, {11, 22, 33, 44}번째 회차에 하나의 수신기 모델의 학습 데이터를 무작위로 섞어 학습시킨 후 글로벌 수신기 모델의 분류정확도를 측정하였다. 그림 3 과 같이 연합학습 회차가 낮을 때는 개별 수신기의 오염이 글로벌 모델에 영향을 크게 미치지만, 연합학습 회차가 늘어날수록 개별 수신기의 오염도가 낮아지고, 이에 따라 글로벌 모델에 미치는 영향도 낮아진다.

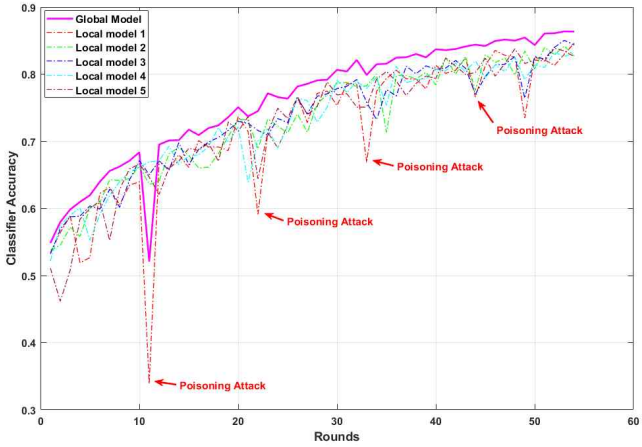

그림 3 개별 수신기의 학습데이터 오염시 분류정확도 측정

Ⅲ. 결론

본 논문은 기계학습의 방법 중 하나인 연합학습에 대해 소개하고 각각의 수신기에서 수집된 무선통신 신호로 개별 학습된 모델의 가중치 변화값을 서버로 보낸 후 글로벌 수신기 모델을 학습하는 과정의 반복을 통해 신호의 변조기술 탐지 성능을 향상시키는 연합학습 모델을 제안하고, 개별 수신기가 악의적 사용자에 의해 오염되었을 때 글로벌 수신기 모델의 보안성이 높음을 실험을 통해 확인하였다. 본 논문에서의 악의적 사용자에 의해 학습 신호가 오염된 것(포이즈닝 공격)을 가정하였으며, 추가적으로 악의적 사용자가 수신기 모델을 탈취하여 적대적 예제(adversarial example)나 백도어 공격을 하는 경우 글로벌 수신기 모델의 보안성에 대해서도 연구 중이다.

ACKNOWLEDGMENT

이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2018-0-00831, 이종 무선 네트워크를 위한 물리 계층 보안 기술 연구).

참 고 문 헌

[1] W. Lee, M. Kim, and D.-H. Cho, “Deep cooperative sensing: Cooperative spectrum sensing based on convolutional neural networks,” IEEE Transactions on Vehicular Technology, vol. 68, no. 3, pp. 3005 - 3009, 2019.

[2] H. Ye, G. Y. Li, and B.-H. Juang, “Power of deep learning for channel estimation and signal detection in ofdm systems,” IEEE Wireless Communications Letters, vol. 7, no. 1, pp. 114 - 117, 2017.

[3] T. O’Shea and J. Hoydis, “An introduction to deep learning for the physical layer,” IEEE Transactions on Cognitive communications and Networking, vol. 3, no. 4, pp. 563 - 575, 2017.

[4] Konečný, Jakub, et al. “Federated learning: Strategies for improving communication efficiency,” arXiv preprint arXiv:1610.05492, 2016.

[5] T. J. O’Shea, T. Roy, and T. C. Clancy, “Over-the-air deep learning based radio signal classification,” IEEE Journal of Selected Topics in Signal Processing, vol. 12, no. 1, pp. 168-179, 2018.