Graph Transformer Network 기반 무선 네트워크 침입 탐지 시스템

홍석원[1], 김진성[1], 김민재[2], 최석환[3]
[1]연세대학교 전산학과 석사과정
[2]연세대학교 정보통신공학 학부생
[3]연세대학교 소프트웨어학부 교수
sw.hong@yonsei.ac.kr, acmiheee@yonsei.ac.kr, joy4mj@yonsei.ac.kr, sh.choi@yonsei.ac.kr

Graph Transformer Network based Wireless Network Intrusion Detection System

Seok-Won Hong[1], Jin-Seong Kim[1], Min-Jae Kim[2], Seok-Hwan Choi[3]
[1]Dept. of Computer Science, Yonsei University
[2]Major in Information & Communication Engineering, Yonsei University
[3]Division of Software, Yonsei University

요 약

수많은 무선 네트워크 서비스의 등장과 함께 무선 네트워크를 대상으로 한 공격이 증가하고 있다. 이러한 공격을 탐지하기 위해 최근 많은 연구가 진행되고 있다. 특히 네트워크의 복잡한 연결 구조와 패턴을 효율적으로 분석할 수 있는 그래프 기반 인공지능 모델이 적용된 네트워크 침입 탐지 시스템(Network Intrusion Detection System, NIDS)에 관한 다양한 연구가 진행되고 있다. 이러한 배경을 바탕으로 본 논문에서는 무선 네트워크를 대상으로 한 공격의 정확하고 신속한 탐지를 위한 Graph Transformer Network(GTN) 기반 네트워크 침입 탐지 시스템을 제안하고 AWID3 데이터셋을 이용한 실험을 통해 GTN 기반 NIDS의 우수성을 검증한다.

1. 서론

무선 통신 기술의 발전과 함께 수많은 무선 네트워크 서비스가 등장하면서 무선 네트워크를 대상으로 한 공격이 가파르게 증가하고 있다. 이와 더불어 무선 네트워크에 대한 공격을 빠르게 탐지하기 위한 인공지능 기반의 네트워크 침입 탐지 시스템에 관한 다양한 연구가 진행되고 있다. 특히, 네트워크의 구조와 패턴을 효율적으로 학습하고 분석할 수 있는 그래프 기반 인공지능 모델이 적용된 NIDS의 연구가 활발히 수행되고 있다[1]. 그래프 구조는 네트워크의 복잡한 연결 구조와 패턴을 표현하는 데 최적화되어 있어 네트워크 침입 탐지에 적합한 형태를 띠고 있다[2].

본 논문에서는 더 빠르고 효과적인 무선 네트워크 침입 탐지를 위해 Graph Transformer Network 기반 네트워크 침입 탐지 시스템을 제안한다. 이를 검증하기 위해 Recurrent Neural Network(RNN), Gated Recurrent Unit(GRU), Long Short-Term Memory(LSTM)와 같은 전통적인 딥러닝 모델이 적용된 기존 무선 침입 탐지 모델과 Graph Convolution Network(GCN), Graph Attention Network(GAT), Graph Isomorphism Network(GIN), Dynamic Graph CNN(DGCNN), Graph Diffusion Convolution(GDC)과 같은 최신 그래프 기반 딥러닝 모델이 적용된 그래프 기반 무선 침입 탐지 모델과의 성능을 AWID3 데이터셋을 사용한 실험을 통해 비교 분석한다.

2. 본론

본 장에서는 무선 네트워크의 효율적이고 빠른 탐지를 위한 GTN 기반 NIDS를 제안하고 실험을 통해 검증한다.

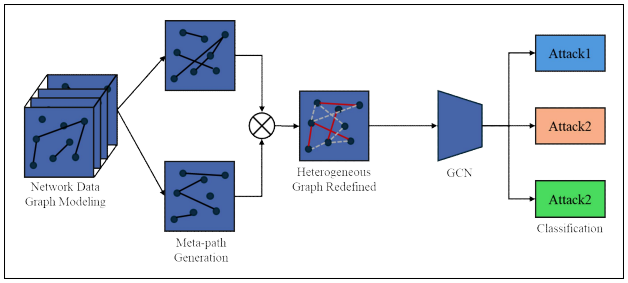

그림 1 : GTN 기반 네트워크 침입 탐지 시스템 개요도

표 1 : 실험 결과 및 성능 평가

Data Type	Model	Accuracy	Precision	Recall	F1-score	Training Time(sec)	Test Time(sec)
Non Graph	RNN	0.9349	0.9356	0.9445	0.9400	317.7	98.2
	GRU	**0.9401**	**0.9510**	**0.9629**	**0.9569**	**303**	**96.5**
	LSTM	0.9277	0.9334	0.9413	0.9373	329.4	120.3
	Average	0.9342	0.9400	0.9495	0.9447	316.7	105
Graph	GCN	0.9410	0.9682	0.9650	0.9665	267.5	82.6
	GAT	0.9363	0.9554	0.9508	0.9531	264.9	79.7
	GIN	0.9347	0.9372	0.9421	0.9396	259.1	81.4
	DGCNN	0.9448	0.9621	0.9597	0.9610	273.8	81.7
	GDC	0.9440	0.9618	0.9611	0.9614	262.2	80.6
	GTN	**0.9451**	**0.9699**	**0.9690**	**0.9694**	**258.8**	**79.3**
	Average	0.9410	0.9591	0.9580	0.9585	264.4	80.9

2.1 Graph Transformer Network based NIDS

GTN은 다양한 엣지와 노드 유형을 포함하는 이기종 그래프를 효율적으로 다루기 위해 설계된 모델이다. 일반적인 GNN 계열 모델들이 처리하기 어려운 다양한 유형의 노드와 엣지로 구성된 이기종 그래프에서 유의미한 정보를 추출하고 그래프 구조의 복잡한 관계를 모델링할 수 있는 모델이다.

GTN의 핵심 구성 요소는 메타-패스 생성, 이기종 그래프 재정의, GCN Layer 통합이다. 메타-패스 생성은 이기종 그래프에서 노드 간의 복잡한 관계를 표현하여 downstream task에 가장 적합한 그래프 구조를 도출한다. 이기종 그래프 재정의는 생성된 메타-패스를 통해 복잡한 이기종 그래프를 더 단순하고 처리하기 쉬운 형태로 변환한다. 이후 GCN Layer를 통합하여 그래프의 구조적 정보와 노드의 특성 정보를 결합하고 분류를 수행한다.

2.2 데이터 전처리

본 논문에서는 GTN 기반 NIDS의 실험 및 검증을 위해 AWID3 데이터셋을 사용하였다. AWID3 데이터셋은 University of The Aegean에서 IEEE 802.1X EAP 환경에서 발생하는 다양한 공격들을 캡처하여 무선 네트워크 패킷을 수집 및 가공하여 만든 데이터셋이다[3].

본 논문에서는 AWID3 데이터셋을 기존 무선 침입 탐지 모델 및 그래프 생성 알고리즘 기반 무선 침입 탐지 모델에 적용하기 위하여 데이터 전처리를 진행하였다. 먼저 각 열의 데이터 형식을 통일하기 위하여 모든 데이터를 숫자 형태로 변환하였다. 그 후 데이터의 범위에 따른 민감도를 낮추기 위해 데이터 범위를 [0,1]로 제한하였으며, 모델의 안정성을 높이기 위해 MinMaxScaler를 통해 정규화하였다. 마지막으로 데이터의 분포 및 특성을 유지하기 위해 NaN값을 해당 열의 평균값으로 대체하였다. 이후 그래프 기반 NIDS의 데이터 입력을 위해 데이터셋을 그래프 형태로 변환하는 과정을 거친다.

2.3 실험 및 성능 비교

본 논문에서는 제안하는 GTN 기반 NIDS의 성능을 평가하기 위해 Accuracy와 Precision, Recall, F1-score, 학습 및 테스트 시간을 측정하였다. 실험은 Python 3.10.9, Torch 2.0.1, i9-13900F, 32GB RAM, RTX 4070Ti 환경에서 진행하였으며, RNN, GRU, LSTM 3가지 모델을 통해 기존 무선 침입 탐지 모델을 구현하고, GCN, GAT, GIN, DGCNN, GDC, GTN 6가지 모델을 통해 그래프 기반 무선 침입 탐지 모델을 구현하였다. RNN, GRU, LSTM, GCN, GAT, GIN의 경우 Pytorch 라이브러리로 구현하였으며, GTN, DGCNN, GDC의 경우 Pytorch에서 라이브러리를 제공하지 않기 때문에 각 모델의 논문에서 제안하는 구조를 직접 구현하였다[4][5][6].

표 1은 Non Graph 방식의 기존 침입 탐지 모델과 Graph 방식의 침입 탐지 모델의 Accuracy, Precision, Recall, F1-score, Training Time, Test

Time을 측정한 결과를 나타낸다. 기존 침입 탐지 모델 중 GRU의 성능이 가장 우수하였으며, 그래프 기반 침입 탐지 모델에서는 GTN의 성능이 가장 우수하였다. 모든 침입 탐지 모델의 성능을 비교하였을 때, Non Graph 방식보다 Graph 방식의 침입 탐지 모델이 평균적으로 더 높은 탐지 성능을 제공함을 확인할 수 있다. Graph 방식의 침입 탐지 모델 중 제안하는 GTN 기반 네트워크 침입 탐지 모델의 성능이 가장 우수함을 확인할 수 있다. 또한 학습 시간 및 테스트 시간 측면에서도 제안하는 기법의 성능이 가장 좋은 것을 알 수 있다.

3. 결론

본 논문에서는 무선 네트워크 침입 탐지를 위한 GTN 기반 네트워크 침입 탐지 시스템을 제안하고 그래프 기반 딥러닝 기법을 활용한 실험을 통해 제안한 NIDS의 유효성을 검증하였다. AWID3 데이터셋을 활용한 실험 결과를 통해, GTN을 적용했을 때 침입 탐지 성능이 증가함을 보였다. 또한, GTN 기반 NIDS의 학습 및 테스트 시간이 다른 모델이 적용된 NIDS에 비해 평균 20% 감소함을 보였다.

사사문구

본 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단(RS-2023-00243075) 및 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업(IITP-2023-RS-2023-00259967)의 지원을 받아 수행된 연구임.

참고문헌

[1] T. Pourhabibi, et al., "Fraud detection: A systematic literature review of graph-based anomaly detection approaches," Decision Support Systems, Vol. 133, 113303, 2020.

[2] L. Akoglu, H. Tong, & D. Koutra, "Graph based anomaly detection and description: a survey," Data Mining and Knowledge Discovery, Vol. 29, pp. 626-688, 2015.

[3] CHATZOGLOU, Efstratios; KAMBOURAKIS, Georgios; KOLIAS, Constantinos. Empirical evaluation of attacks against IEEE 802.11 enterprise networks: The AWID3 dataset. IEEE Access, 2021, 9: 34188-34205.

[4] Yun, Seongjun, et al. "Graph transformer networks." Advances in neural information processing systems 32 (2019).

[5] WANG, Yue, et al. “Dynamic graph cnn for learning on point clouds.” ACM Transactions on Graphics (tog), 2019, 38.5: 1-12.

[6] Gasteiger, Johannes, Stefan Weißenberger, and Stephan Günnemann. "Diffusion improves graph learning." Advances in neural information processing systems 32 (2019).