생성형 AI 기반 네트워크 데이터 생성 및 유효성 평가 방법

박노삼*, 이종훈
한국전자통신연구원

siru23@etri.re.kr, mine@etri.re.kr

Generative AI based network data synthesis and data validation method

Noh-Sam Park*, Jonghoon Lee
Electronics and Telecommunications Research Institute

요 약

본 논문은 네트워크 환경에서 데이터 부족 문제를 해결하기 위해, 생성형 AI 기반 합성 데이터 생성 및 유효성 평가 방법을 제안한다. 원본 네트워크 데이터를 기반으로 다양한 트래픽 패턴을 반영한 합성 데이터를 생성하고, 이를 통해 학습된 침입 탐지 모델의 성능을 분석하였다. 제안된 방법은 실제 데이터와 유사한 수준의 정확도를 확보하면서도 데이터 품질과 신뢰성을 향상시켜, 효율적인 보안 위협 예측이 가능한 지능형 네트워크 보안 시스템 구축에 기여할 수 있다

Ⅰ. 서 론

5G 네트워크는 초고속, 초저지연, 대용량 통신을 가능하게 하며, 산업, 국방, 스마트시티 등 다양한 분야에서 활용되고 있다. 그러나 네트워크 구조의 복잡성과 트래픽 다양성 증대로 인해 새로운 형태의 보안 위협이 지속적으로 발생하고 있다. 기존의 침입 탐지 시스템은 알려진 공격 패턴에 의존하기 때문에 새로운 위협을 효과적으로 탐지하기 어렵다. 특히, 학습용 네트워크 데이터는 수집과 라벨링 과정에서 높은 비용과 시간이 소요되어, AI 기반 탐지 모델의 성능 향상에 제약이 존재한다. 이에 본 논문은 생성형 AI 를 활용하여 5G 환경과 같은 유무선 네트워크에 최적화된 합성 데이터를 생성하고, 해당 데이터의 품질과 유효성을 평가함으로써 학습 데이터 부족 문제를 해결하고자 한다.

본 논문의 구성은 다음과 같다. 2 장에서 지능형 네트워크 침입 탐지를 위한 생성형 AI 기반 네트워크 데이터 합성 방법을 논한다. 3 장에서 합성된 네트워크 데이터에 대한 유효성을 검증하는 방법 및 결과를 분석한 후 결론을 도출한다.

Ⅱ. 생성형 AI 기반 네트워크 데이터 합성

지능형 5G 특화망 보안 위협 탐지를 위해서는, 지능화된 네트워크 침입 및 이상행위를 인공지능 기술을 통해 탐지하고 분석할 기술이 필요하고, 정상적인 행위와 각종 비정상/위협/공격 행위를 수집하여, 다양한 학습 알고리즘을 적용하는 고도화된 보안 분석 기술이 필요하다. 기존의 네트워크 침입 탐지 연구는 주로 지도학습 기반 접근을 사용하였으며, 정상 트래픽과 공격 트래픽의 차이를 모델링하여 위협을 식별하였다. 그러나 실환경에서는 공격 데이터가 불균형하거나 부족하여 모델의 일반화 성능이 저하되는 문제가 존재한다. 이러한 한계를 보완하기 위해 최근 오토인코더(autoencoder), GAN(Generative Adversarial Network), LLM(Large Language Model) 등을 활용한 데이터 합성 연구가 활발히 진행되고 있다.

생성형 모델 중 확산(diffusion) 모델은 이미지·신호 생성에서 GAN 의 한계를 극복하기 위해 등장한 생성형 모델 기술로서, 원본 데이터에 점차 노이즈를 추가하고 이를 역으로 제거하는 과정을 학습해 고품질의 데이터를 생성한다. 이는 이미지 생성, 음성 합성, 텍스트-이미지 변환, 시계열 예측 등 다양한 분야에 활용되며, 최근에는 네트워크 트래픽이나 센서 데이터 등 복잡한 구조를 지닌 데이터의 합성에도 응용되고 있다.확산 모델을 이용하여 네트워크 데이터를 합성하기 위한 연구로는 NetDiffusion[1], Diff-IDS[2] 등이 있다.

본 논문은 고품질의 합성 데이터를 확보하고, 이를 기반으로 침입 탐지 모델의 신뢰성과 성능을 향상하기 위해, CIC-IDS 2017 데이터셋[3]을 이용하여 생성형 모델을 기반으로 네트워크 플로우 데이터를 생성하고 데이터 유효성을 검증한다.

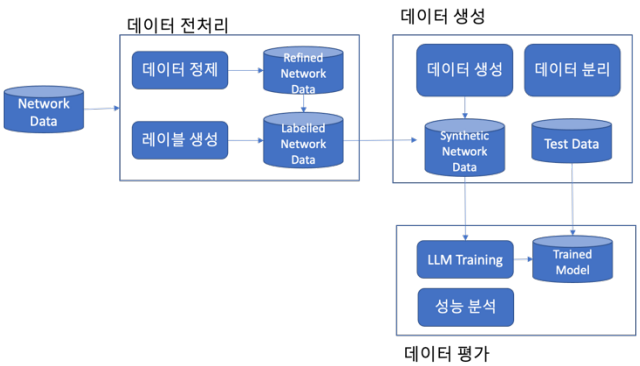

그림 1. 생성형 AI 기반 네트워크 데이터 합성 구조

생성형 AI 기반 네트워크 데이터 합성은 데이터 전처리, 데이터 생성, 데이터 평가 모듈로 구성된다. 데이터 전처리 모듈은 원본 네트워크 플로우 데이터에서 불필요한 항목을 제거하고 결측치 및 이상치를 처리하는 데이터 전처리 과정을 수행한다. 이후 전처리된 데이터를 입력으로 하여 데이터 생성 모듈에서는 확산 모델을 기반으로 합성 네트워크 데이터를 생성한다. 생성된 데이터는 정상 및 이상 트래픽의 다양한 패턴을 포함하며, 실제 네트워크 프로토콜 규칙을 준수하도록 설계된다.

합성된 데이터를 이용하여 침입 탐지 모델을 학습하고, 실제 데이터로 학습된 모델과의 성능을 비교함으로써 합성 데이터의 품질과 유효성을 평가한다. 평가 지표로는 정확도, 정밀도, 재현율 등을 사용하며, 생성 데이터가 실제 데이터와 유사한 수준의 성능을 보일 경우 데이터 품질이 우수한 것으로 평가한다. 이러한 과정을 반복하여 데이터 합성 과정을 최적화하고, 최종적으로 신뢰성 높은 네트워크 침입 탐지 시스템을 구현한다.

확산 모델 기반 데이터 생성은 noise 에 많이 영향을 받으며, noise 가 높을 수록 네트워크 데이터 합성에서는 원본 데이터와 패턴이 상이한 결과를 보여 주었다. 향후 높은 noise 에서도 원본과 유사한 데이터 패턴을 보이도록 학습을 해야 데이터 강건성을 확보할 수 있다.

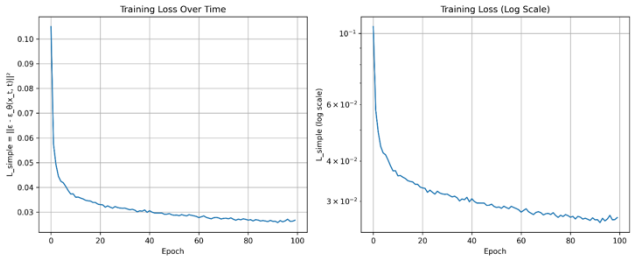

그림 2. 생성형 AI 기반 네트워크 합성 모델 학습 과정

III. 합성 데이터 유효성 검증 결과

합성된 데이터의 유효성을 검증하기 위해서, 네트워크 침해 위협 탐지에서 많이 사용되는 CIC-IDS 2017 데이터셋을 기반으로 네트워크 공격 탐지를 위한 Transformer 기반 이진 분류 모델을 학습하였다.

CIC-IDS 2017 데이터셋은 79 개의 피쳐를 가진 네트워크 트래픽으로서 정상(benign) 트래픽과 다양한 공격 유형을 포함하고 있다. 본 논문에서는 정상과 공격(attack)으로 분류하였으며. 트랜스포머 기반의 이진 분류 모델을 이용하여 성능을 분석하였다. 네트워크 플로우 파일을 입력으로 데이터 전처리 과정을 거친 후 클래스 불균형 문제를 완화하기 위해 공격 및 정상 데이터를 균형화한 뒤 학습, 검증, 테스트용으로 분할하였다. 모델은 총 35 epoch 동안 학습되었으며, 학습 배치 크기는 512, 검증 배치 크기는 1024 로 설정하였다. 학습 도중 체크포인트를 생성하여 최적의 가중치를 저장하였다.

성능 평가는 세 가지 시나리오로 수행되었다. 원본 데이터만 사용했을 경우를 기준선(baseline)으로 선정하고, 합성 데이터만을 사용하거나 원본과 합성을 같이 사용할 경우의 분류 성능을 비교 분석하였다.

1) 원본 데이터만 사용한 경우 성능

CIC-IDS 2017 데이터의 10%를 시험 데이터로 사용하여 243,951 건 데이터를 사용한 결과, 공격 탐지의 정밀도(precision)과 f1-score 는 각각 0.97, 0.9 을 기록하였다.

2) 합성 데이터만 사용한 경우 성능

합성 데이터만을 이용하여 공격 탐지를 수행한 결과, 정밀도 0.99, f1-score 0.99 로 원본 데이터보다 더 높은 성능을 보였다.

3) 합성 데이터를 원본 데이터에 추가한 경우 성능

원본 데이터와 동일한 개수의 합성 데이터를 사용하였으며, 정밀도 0.98, f1-score 0.99 로 나타났다.

실험 결과 합성 데이터의 품질이 우수하며, 원본 데이터의 다양성과 결합하여 학습했을 때도 안정적인 탐지 성능을 유지함으로써 합성 데이터의 유효성을 검증할 수 있었다.

IV. 결론

본 논문에서는 생성형 AI 를 활용하여 네트워크 플로우 데이터를 합성 생성하고, 이를 통해 침입 탐지 모델의 성능과 데이터 유효성을 향상시키는 방법을 제시하였다. 또한 합성된 데이터의 유효성을 검증 하기 위해 트랜스포머 기반 이진 분류 모델을 구축하고, 합성 데이터를 활용한 학습 효과를 분석하였다. 실험 결과, 합성 데이터만으로도 높은 탐지 성능을 확보하였으며, 원본 데이터와의 결합 시에도 안정적인 성능을 유지하였다. 이는 제안된 데이터 생성 및 학습 방식이 실제 네트워크 보안 환경에서도 적용 가능함을 시사한다.

제안된 접근은 데이터 부족 문제를 해결함과 동시에 예측 정확도를 높여, 향후 지능형 네트워크 보안 기술 발전에 기여할 것으로 기대된다

ACKNOWLEDGMENT

이 논문은 2025 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구 결과임[No. RS-2024-00397469, 특화망·기업망 통합보안을 위한 5G 특화망 보안 기술개발].

참 고 문 헌

[1] Jiang, X., Liu, S., Gember-Jacobson, A., Bhagoji, A.N., Schmitt, P., Bronzino, F., Feamster, N.: NetDiffusion: Network data augmentation through protocol-constrained traffic generation. Proc. ACM Meas. Anal. Comput. Syst. 8(1), 1- 32, 2024.

[2] Y. Yang, X. Tang, Z. Liu, J. Cheng, H. Fang, and C. Zhang, “Diff-IDS: A Network Intrusion Detection Model Based on Diffusion Model for Imbalanced Data Samples,” Comput. Mater. Contin., vol. 82, no. 3, pp. 4389- 4408, 2025. https://doi.org/10.32604/cmc.2025.060357

[3] Iman Sharafaldin, Arash Habibi Lashkari, and Ali A. Ghorbani, “Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization”, 4th International Conference on Information Systems Security and Privacy (ICISSP), Portugal, January 2018.