딥러닝 기반의 SQL injection 공격 탐지 동향

강예준*, 김원웅*, 김현지*, 임세진*, 서화정*†

*한성대학교 (대학원생)

*† 한성대학교 (교수)

Deep Learning-based SQL Injection Attack Detection Trend

Yea-jun Kang*, Won-woong Kim*, Hyun-ji Kim*, Se-jin Lim*, Hwa-jeong Seo*†

*Hansung University(Graduate student)

*† Hansung University(Professor)

요 약

웹 어플리케이션 보안 프로젝트인 OWASP에서 발표한 문서에 따르면 2021년 10대 웹 애플리케이션 취약점 중 SQL injection 공격을 3위로 선정하였다. SQL injection 공격이란 데이터베이스에 대해 접근 시 사용되는 명령어인 SQL 쿼리문을 조작하여 악의적인 SQL 쿼리문을 실행시키는 것이다. SQL injection 공격에 경우 데이터베이스에 큰 피해를 입힐 수 있음으로, 이에 대한 적절한 대응 방안이 요구된다. 최근에는 딥러닝을 활용하여 SQL injection 공격을 탐지하는 연구가 많이 진행되고 있다. 본 논문에서는 딥러닝 기반의 SQL injection 공격 탐지 동향에 대해 살펴본다. SQL 쿼리문을 데이터프레임으로 바꾸거나 디코딩하는 등의 전처리 과정을 거쳐 모델에 입력하여 탐지하는 연구가 주로 진행되었다. 탐지 모델 CNN, KNN, SVM, LSTM 등 다양한 모델에 대해 성능을 측정하고 비교하였다. SQL 쿼리문의 경우 사용자의 편의성과 직결되는 문제이므로 성능 뿐만 아니라 속도도 중요한 요소이다. 따라서 모델의 속도를 개선시키기 위한 연구도 필요할 것으로 생각된다.

I. 서론

웹 보안 문제 중 하나인 SQL injection 공격은 임의의 SQL 쿼리문을 주입하여 데이터베이스가 비정상적으로 동작하게 하는 공격 기법이다. 따라서 데이터베이스와 연동되고 있는 웹 어플리케이션의 경우, 개인정보가 유출되거나 데이터가 조작되는 등의 위험이 있다. 웹 애플리케이션 보안 프로젝트인 OWASP(Open Web Application Security Project)에서 발표한 문서에 따르면, 2021년 10대 웹 어플리케이션 취약점 중 SQL injection을 3위로 선정하였다[1]. 이를 보아 SQL injection 공격은 치명적인 공격을 가할 수 있는 네트워크 공격 수단 중 하나이며, 이러한 위협에 대응하기 위해 SQL injection 탐지 기술이 요구된다.

II. 관련 연구

2.1 SQL injection

SQL injection이란 SQL 쿼리문을 조작하여 데이터베이스를 비정상적으로 동작하게 하는 공격 기법이다[2]. 즉, 공격자가 보안상의 취약점을 악용하여 악의적인 SQL 문을 주입하는 행위이다. 공격에 성공하게 될 경우, 공격자가 데이터베이스에 접근하게 되어 민감한 데이터나 개인 정보가 유출될 위험이 있으며, 조직의 데이터 전체를 장악하여 조직에 큰 피해를 입힐 수 있다.

2.2 인공신경망

인공신경망이란 인간의 뇌 속에 있는 뉴런의 연결 구조로부터 영감을 받아 만든 네트워크 구조이다[3]. 인간의 뇌에는 수많은 뉴런이 존재

하고 각 뉴런은 다른 뉴런들로부터 신호를 받기도 하고 신호를 전달하기도 한다. 이를 컴퓨터로 구현한 것이 인공신경망이며, 딥러닝은 이러한 인공신경망을 통해 학습을 수행한다. 인공신경망에는 입력층, 은닉층 그리고 출력층이 존재한다. 입력층을 통해 학습하고자하는 데이터를 입력받고, 신경망 외부에서 접근 불가능한 은닉층을 거쳐 출력층을 통해 결과가 출력된다. 머신러닝과 다르게 데이터로부터 특징을 스스로 추출하여 학습한다. 딥러닝은 현재 컴퓨터 비전, 음성 인식, 자연어 처리, 신호 처리 등 다양한 분야에서 활용되고 있다.

2.3 합성곱 신경망(Convolutional Neural Network : CNN)

컴퓨터 비전 분야에서 주로 사용되는 합성곱 신경망은 심층 신경망과 다르게 데이터의 형상을 유지한 채로 데이터를 학습한다[4]. 심층 신경망의 경우 데이터를 1차원 형태로만 입력받을 수 있는데, 이미지와 같이 3차원 데이터를 입력받기 위해서는 평탄화 작업을 반드시 수행해야한다. 하지만 이미지의 경우 인접한 픽셀들끼리는 RGB 채널에 대해 서로 연관성이 높으며, 멀리 떨어져있는 픽셀들끼리는 연관성이 떨어진다. 따라서 이미지를 평탄화시킬 경우 이미지에 담긴 공간적 정보가 유실되어 학습하는데에 한계가 존재한다. 이러한 문제점을 해결한 합성곱 신경망은 형상을 유지한 채로 데이터를 학습하기 때문에 컴퓨터 비전 분야에서 많이 사용되는 신경망이다.

III. 딥러닝 기반 SQL injection 공격 탐지 동향

본 논문에서는 딥러닝 기반의 SQL injection을 탐지하는 기법의 동향을 살펴본다. 대다수의 연구가 SQL 쿼리문에 대해 전처리 과정을 거친 후에 모델을 학습시켜 SQL injection을 탐지하였다.

[5]에서는 모든 유형의 SQL Injection 공격을 다루는 데이터셋을 직접 생성하여, 해당 데이터셋을 통해 모델을 학습시켰다. 데이터셋에 SQL 삽입 공격에 사용할 수 있는 모든 유형의 페이로드 및 쿼리를 포함시켜 범용적인 데이터셋을 생성하였으며, SQL injection을 탐지하는 Deep learning의 성능을 분석하기 위해 ROC 곡선, 정확도, 정밀도, 재현율과 같은 다양한 성능 지표를 사용하여 성능을 측정하였다. 데이터셋은 모델에 입력되기 전 전처리 과정을 거친다. 테이터는 SQL injection인 경우에 1, 일반 SQL 쿼리문인 경우에는 0으로 레이블을 지정해주고 데이터셋을 알고리즘이 더 잘 이해할 수 있도록 데이터 프레임으로 컴파일한다. 해당 논문에서는 의사결정트리(Decision Tree:DT), 나이브 베이즈 분류(Gaussian Naive Bayes Classifier:GNB), 서포트 벡터 머신(Support Vector Machine:SVM), 최근접 이웃 (Knearest Neighbors and DT:KNN) 그리고 합성곱신경망(Convolutional Neural Networks) 총 5가지 분류기에 대해 성능을 평가하고 비교하였다. 측정결과 GNB에서 가장 높은 정확도가 측정되었고 SVM에서 가장 높은 정밀도가 측정되었다. 하지만 CNN이 정확도, 정밀도 그리고 재현율에 대해 각각 94.84%, 85.67%, 96.56%로 일관성 있고 가장 좋은 성능을 보여주었다. 해당 논문에서는 향후 성능을 개선시키기 위해 다른 딥러닝 접근 방식을 고려중이며 비지도 학습에 대해서도 연구하고 SQL injection 공격 데이터를 추가하여 데이터셋을 확장시켜 모델을 개선시킬 예정이다.

[6]에서는 데이터셋 30919개를 SQL injection인 경우에 1로, 일반 SQL 쿼리문인 경우에는 0으로 레이블을 지정하였다. 데이터셋을 토큰화하여 분석해본 결과 "select", "*", "from"과 같은 단어는 SQL injection 공격과 일반 SQL 쿼리문 모두에서 자주 등장하는 단어이다. 또한 특정 단어는 SQL injection 공격과 일반 SQL 쿼리문에서 사용되는 빈도수가 매우 낮았다. 즉, 분류하기 위해 요구되는 데이터 특성으로써의 역할을 수행하지 못하는 단어가 존재하였다. 이러한 문제점을 해결하기 위해 TF-IDF 알고리즘을 통해 단어에 가중치를 부여하고 단어별로 중요도를 측정하여 사용되는 빈도수가

너무 높거나 낮은 단어는 제외하였다. 그 후 토큰화된 데이터를 벡터화하여 벡터행렬을 형성하고, 벡터행렬을 희소행렬로 변환하였다. 위와 같은 과정을 수행함으로써 신경망을 훈련시킬 때, 계산을 용이하게 하고 성능을 개선시켰다. 딥러닝 모델 SQLNN은 옵티마이저로써 Adam을 사용하였고, 손실함수는 Cross-Entropy를 사용하였다. 또한 모델에 Dropout 기법을 적용함으로써 과적합을 방지하였다. 장단기 메모리(Lonh Short-Term Memory:LSTM), KNN, DT와 SQLNN을 비교하여 성능을 비교하였다. 측정 결과 KNN 및 DT 알고리즘은 과적합 문제가 발생하였으며 정확도는 각각 82.96%, 92.33%로 측정되었다. 또한 LSTM의 경우 시간이 오래 걸리고 긴 시퀀스를 다루는데 한계가 있어 정확도가 62.32%로 상당히 낮은 수치를 보였다. 마지막으로 해당 모델에서 제안한 SQLNN에서는 정확도, 정밀도, 재현율 그리고 F1-Score에 대해 각각 96.16%, 97.28%, 92.23% 그리고 94.68%로 다른 알고리즘에 비해 높은 성능을 보였다.

[7]에서는 단어 임베딩, CNN, MLP(Multi-Layer Perceptron)을 사용하여 SQL injection 탐지하였다. 먼저 HTTP 요청 데이터는 urlencode, querystring, JSON, PHP serialize, base64등 다양한 방식으로 인코딩 될 수 있다. 따라서 모델에 입력되기 전에 디코딩을 수행하고, 특징을 강화하기 위해 데이터에 대해 일반화를 수행한다. 일반화 규칙은 숫자를 모두 "0"으로, URL은 "http://u"로 변환한다. 그 후 단어별로 나누면 디코딩이 완료된다. 그 후 디코딩된 데이터를 통해 CNN과 MLP 분류기를 학습시키고 해당 분류기를 통해 SQL injection을 탐지한다. CNN은 합성공 신경망 3개와 pooling layer 3개로 구성되어 있고 마지막 계층은 완전 연결 계층으로 구성되어 있다. 합성곱 신경망과 완전 연결 계층은 ReLu 활성화함수를 사용하였고 최종적으로 Softmax를 활성화 함수로써 사용하였다. MLP는 2개의 은닉층으로만 설계되었으며 활성화 함수는 CNN과 마찬가지로 ReLu와 Softmax를 활성화 함수로써 사용하였다. 성능 측정 결과 CNN의 경우 정확도, 정밀도, 재현율 그리고 F1-Score에 대해 98.25%, 97.46%, 99.07%, 98.26%로 측정되었다. 또한 MLP의 경우 정확도, 정밀도, 재현율 그리고 F1-Score에에 대해 98.57%, 97.95%, 99.22%, 98.58%로 측정되었다. CNN과 MLP 모델 모두 높은 성능을 나타냄을 확인할 수 있었지만, MLP는 데이터 4000개에 대해 탐지 속도로 약 12초를 사용하였고 CNN은 26초를 사용하였다. 이는 모델에서 사용하는 파라미터의 개수가 CNN보다 MLP가 더 적기 때문에 탐지 속도가 높은 것이다.

IV. 결론

본 논문에서는 딥러닝을 통해 SQL injection을 탐지해내는 연구들을 살펴보았다. 대다수의 연구들은 SQL 쿼리문을 그대로 사용하여 탐지하기 보다, 전처리 과정을 거친 후에 탐지하였다. 매우 다양한 모델에서 탐지를 시도하였으며, 대부분의 연구가 90% 이상의 성능을 보였다. 하지만 SQL injection 공격의 경우 모델의 성능 뿐만 아니라 속도도 중요하기 때문에, 더 빠른 속도로 SQL injection을 탐지할 수 있는 방법에 대한 연구도 진행되어야 할 것으로 보인다.

V.Acknowledgement

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임(No.2018-0-00264, IoT 융합형 블록체인 플랫폼 보안 원천 기술 연구, 50%) 그리고 이 성과는 2022년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2020R1F1A1048478, 50%)

[1] OWASP, https://owasp.org/www-project-top-ten/ , 2020.

[2] Halfond, William G., Jeremy Viegas, and

Alessandro Orso. "A classification of SQL-injection attacks and countermeasures." Proceedings of the IEEE international symposium on secure software engineering. Vol. 1. IEEE, 2006.

[3] Wang, Sun-Chong. "Artificial neural network." Interdisciplinary computing in java programming. Springer, Boston, MA, 2003. 81-100.

[4] O'Shea, Keiron, and Ryan Nash. "An introduction to convolutional neural networks." arXiv preprint arXiv:1511.08458 (2015).

[5] Falor, Ayush, et al. "A Deep Learning Approach for Detection of SQL Injection Attacks Using Convolutional Neural Networks." Proceedings of Data Analytics and Management. Springer, Singapore, 2022. 293-304.

[6] Zhang, Wei, et al. "Deep Neural Network-Based SQL Injection Detection Method." Security and Communication Networks 2022 (2022).

[7] Chen, Ding, et al. "Sql injection attack detection and prevention techniques using deep learning." Journal of Physics: Conference Series. Vol. 1757. No. 1. IOP Publishing, 2021.