GNN 기반 실시간 URL 컨텍스트 분석을 이용한 지능형 위협 탐지 및 위험도 정량화 연구

허현우*, 송종근*
*동서대학교 정보보안학과
e-mail: acehurwoo@naver.com, ssongjg@dongseo.ac.kr*

A Study on Intelligent Treat Detection and Risk Quantification using Graph Neural Network-Based Real-time URL Context Analysis

Hyeonwoo Heo and Song Jong Gun*
*Dept. of Cyber Security, Dongseo University

요 약

2024년 메신저 스미싱 급증으로 금융 피해액이 전년 대비 4배 증가하였다. 공격자는 식별 어려운 URL과 사회 공학 기법을 사용해 탐지를 어렵게 만든다. 기존 블랙리스트 방식은 실시간으로 생성되는 신종 URL 대응에 한계가 있다. 이에 GNN 기반 실시간 URL 컨텍스트 분석으로 알려지지 않은 위협을 탐지하고 위험도를 정량화한다. 위험도 시각화, 문자 격리, 신고 연동으로 피해를 원천 차단하는 지능형 시스템을 제안한다.

Ⅰ. 서 론

최근 비대면 소통의 확산으로 인해, 메신저가 주요 소통 수단으로 자리 잡았다. 메신저를 이용한 스미싱, 피싱 범죄가 해마다 급증하고 있으며, 통계에 따르면 2024년 금융 피해액은 전년 대비 약 4배 증가한 것으로 나타났다. 현재 사용되는 보안 솔루션은 신고된 악성 URL을 차단하는 블랙리스트 방식에 의존하여, 실시간으로 생성되는 신종 및 변종 URL을 이용한 지능형 공격에 대한 선제적 피해 예방에 한계가 있다. 본 논문에서는 이러한 문제를 해결하기 위해 그래프 신경망을 활용하여 URL의 정보를 다차원적인 컨텍스트를 실시간으로 분석하고, 위협까지 탐지 및 정량화하여 사용자에게 위험도와 판단 근거를 시각적으로 제공하는 지능형 스미싱 대응 시스템을 제안한다.

Ⅱ. 본론

2.1 그래프 신경망(Graph Neural Network)

그래프 신경망[1]은 그래프 형태의 데이터를 효과적으로 처리하기 위해 제안된 딥러닝 기반 모델이다. 그래프 신경망은 그래프의 각 노드를 벡터로 표현하고, 이웃 노드 간의 메시지 전달 및 집계 과정을 통해 노드 간의 복잡한 구조적 관계를 학습하여 데이터에 내재된 패턴을 파악한다. 이러한 특성 덕분에 그래프 신경망은 기존의 정형 데이터 분석 방법론으로는 파악하기 어려운 비선형적이고 복잡한 상호작용을 효과적으로 모델링할 수 있다. 본 연구에서는 URL과 관련된 다차원적 컨텍스트 피처들의 상호 관계를 그래프 형태로 모델링하는 데 활용하였다. 이를 통해 단일 피처만으로는 식별하기 어려운, 여러 컨텍스트가 복합적으로 작용하여 발생하는 지능형 스미싱 공격의 잠재적 패턴을 탐지하고자 한다.

2.2 URL 컨텍스트 분석 (URL Context Analysis)

URL 컨텍스트 분석[2]은 URL 문자열 자체의 통계적 특성뿐만 아니라, 해당 URL과 연관된 도메인 정보, IP 주소의 평판, SSL 인증서 유무, 웹페이지 키워드 등 다차원적인 데이터를 종합적으로 분석해 잠재적 위협을 판단하는 분석 방법론이다. 이는 URL 문자열만으로는 위협을 판단하기 어려운 리다이렉션이나 동적 생성 URL을 이용한 최신 공격 기법에 대응하기 위해 필수적이다. 본 연구에서는 이러한 URL 컨텍스트 분석을 통해 도메인 생성 정보, IP 평판 점수, 인증서 유효성 등을 그래프 신경망이 학습할 핵심 피처로 추출하고, 이를 그래프 데이터의 노드 속성으로 정의하는 데 활용하였다.

2.3 시스템 아키텍처

Fig. 1.은 본 논문에서 제안하는 지능형 스미싱 예방 시스템의 전체 구조를 나타낸다. 제안 시스템은 그래프 신경망 기반 URL 컨텍스트 분석 모듈, 위험도 정량화 모듈, 사용자 시각화 인터페이스로 구성된다.

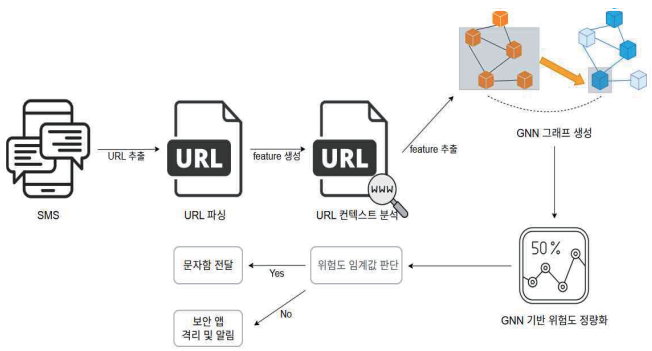

Fig. 1. System Structure

2.4 URL 컨텍스트 기반 그래프 생성

수신된 문자 메시지에서 URL을 추출하고, URL 컨텍스트 분석을 통해 도메인 생성 정보, IP 주소, 인증서 유효성, 관련 키워드로 구성된 다차원 메타데이터를 수집한다. 수집된 메타데이터를 기반으로 URL과 주변 요소들 간의 관계를 나타내는 그래프를 실시간으로 생성한다. 이 그래프 구조에서 URL, 도메인, IP 주소, 도메인 등록자는 그래프의 고유한 노드로 표현된다. 노드 간의 엣지는 이들 엔티티 간의 논리적, 물리적 연결성을 나타내며, 구체적으로 소유 관계, 호스팅 관계, 리다이렉션 상호 관계를 반영하여 정의된다.[3] 여러 의심 URL 노드가 동일한 신생 도메인 등록자 노드나 특정 악성 IP 주소 노드에 연결되는 패턴을 포착할 수 있다. 이러한 그래프 구조는 URL을 둘러싼 다차원적인 컨텍스트의 상호 관계를 종합적으로 반영하여, 단일 정보로는 파악하기 어려운 고도화된 스미싱 공격 패턴을 효과적으로 탐지할 수 있도록 설계되었다.[4]

2.5 GNN 노드 및 엣지 특성 표현 연산

본 연구에서 사용된 그래프 신경망 모델은 특정 URL 노드 v의 위험도를 이웃 노드 $u \in N(v)$의 정보와 엣지 관계를 종합적으로 반영하여 다음과 같이 계산한다.

$$h_v = \sigma\left(\sum_{u \in N(v)} \alpha_{vu} \cdot W h_u\right) \quad (1)$$

h_v는 대상 URL 노드 v의 위험도 특성, h_u는 이웃 노드의 입력 특성을 나타내는 벡터이다. W는 학습 가능한 가중치 행렬, α_{vu}는 노드 간의 관계 중요도를 나타내는 attention 계수이며, σ는 비선형 활성화 함수이다.[4] α_{vu}는 도메인 생성일과 현재 시점의 차이와 IP 평판 점수를 컨텍스트 정보를 함께 고려하여 계산되며, 이를 통해 URL의 잠재적 위험도를 정확히 계산할 수 있다.

2.6 사용자 알림 및 격리 대응

그래프 신경망 모델이 사전에 정의된 특정 임계값 이상으로 위험도를 측정한 URL이 포함된 문자는 즉시 애플리케이션 내 별도로 분리된 안전한 공간으로 격리된다. 이는 사용자가 악성 URL에 의도치 않게 노출되거나 클릭하는 것을 원천적으로 방지하는 조치이다. 위협 탐지와 동시에 사용자에게는 실시간 알림이 전송되며 알림에는 그래프 신경망 모델이 산출한 정량화된 위험도 수치와 판단의 핵심 근거가 된 컨텍스트 메타데이터가 시각적으로 포함된다. 사용자는 제공된 명확한 근거 정보를 통해 잠재적 위협을 직관적으로 인지하고, 해당 메시지를 검토 후 삭제하거나 즉시 신고하는 등 스스로 안전을 확보할 수 있는 결정을 내릴 수 있다. 이 기능은 고도화된 기술적 탐지 결과를 사용자의 실질적인 피해 예방 행동으로 유기적으로 연결하여, 수동적인 방어 체계를 능동적이고 실용적인 보안 체계로 구축하는 핵심 역할을 수행한다.

Ⅲ. 결론

본 연구는 실시간 URL 컨텍스트 정보를 그래프 신경망을 모델링하여, 알려지지 않은 신종 스미싱 위협 패턴의 위험도를 정량화하였다. 이를 통해 의심 문자를 격리하고 사용자에게 위험도와 근거를 시각적으로 제공하여, 피해를 선제적으로 예방하는 지능형 사용자 보호 시스템을 제안한다. 향후 연구로는 그래프 신경망 모델 최적화를 통해 탐지 성능을 고도화하고, 모바일 환경에 최적화된 경량화 모델 개발을 계획하고 탐지 범위를 URL뿐만 아니라 이미지, QR 코드와 같은 새로운 형태의 피싱 공격까지 확장하는 연구를 진행하고자 한다.

참 고 문 헌

[1] M.-H. Zhong, M. Lin, C. Zhang, and Z. Xu, "A survey on graph neural networks for intrusion detection systems: Methods, trends and challenges," Comput. Secur., vol. 141, p. 103821, 2024. DOI: 10.1016/j.cose.2024.103821.

[2] A. K. Jain and B. B. Gupta, "A survey of phishing detection using traditional machine learning and deep learning," Inf. Sci., vol. 643, p. 128532, 2023. DOI: 10.1016/j.ins.2023.128532.

[3] P. Prakash, M. Kumar, R. R. Kompella, and M. Gupta, "Phishnet: predictive blacklisting to detect phishing attacks," in Proc. IEEE INFOCOM, 2010, pp. 1–9. DOI: 10.1109/INFCOM.2010.5462194.

[4] Z. Zhang, P. Cui, and W. Zhu, "A comprehensive survey on graph neural networks," IEEE Trans. Neural Netw. Learn. Syst., vol. 34, no. 1, pp. 1–21, Jan. 2023. DOI: 10.1109/TNNLS.2021.3054358.