HIDS에서 위험 허용도에 따른 윈도우 크기와 임계치의 성능 영향 분석

박상준[1], 김미수[2]
[1]전남대학교 컴퓨터정보통신공학과 학부생
[2]전남대학교 인공지능융합학과 교수
191416@jnu.ac.kr , misoo.kim@jnu.ac.kr

Performance Impact Analysis of Window Size and Threshold Based on Risk Tolerance in HIDS

SangJoon Park[1] , Mi-Soo Kim[2]
[1]Dept. of Computer Engineering, Chonnam University
[2]Dept. of Artificial Intelligence Convergence, Chonnam University

요 약

본 연구는 호스트기반 침입탐지시스템(HIDS)에서 윈도우 크기와 임계치 설정이 성능에 미치는 영향을 분석하였다. 그 결과, 윈도우 크기와 임계치에 따라 성능 변동이 크게 나타났으며, 고정된 최적값 대신 각 시스템의 환경과 위험 허용도에 맞춰 윈도우 크기와 임계치를 조정하는 방안을 제안한다.

1. 서론

마이크로서비스의 확산으로 보안 표면적이 증가하고 있다[1]. 이를 해결하기 위해, 서비스 운영 호스트 내 상태를 모니터링하여 이상 징후를 탐지 하는 호스트 기반 침입 탐지 시스템(Host based intrusion detection system, HIDS)에 대한 연구가 활발히 진행되고 있다[2].

HIDS는 1) 실시간으로 시스템 콜을 수집하고, 2) 수집된 데이터를 바탕으로 이상 여부를 탐지한 후, 3) 이상 여부의 시계열 정보를 기반으로 최종적으로 침입 여부를 관리자에게 알린다. 이 과정에서 다양한 탐지 방식이 활용될 수 있으며, 최종 알람을 위해 슬라이딩 윈도우 방식이 사용된다[2] (그림 1).

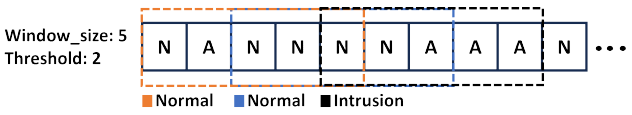


그림 1. 슬라이딩 윈도우 방식 (A:이상, N:정상)

이 분야에 적용된 슬라이딩 윈도우 방식은 먼저 각 구간에서 시스템 콜이나 n-gram 패턴의 이상 여부를 예측하고, 그 예측 결과를 슬라이딩 윈도우로 묶어 종합적인 판단을 내리는 방식이다. 일정 크기의 윈도우를 시간 축을 따라 이동시키며, 설정된 임계치를 초과하면 이를 침입으로 간주한다. 예를 들어, 그림 1에서 윈도우 크기를 5로 설정하고, 임계치를 2로 설정한 경우 첫 번째, 두 번째 윈도우에서는 이상 여부가 1개만 있어 정상 상태로 간주되지만, 세 번째 윈도우에서는 A가 3개로 침입 상태로 간주된다.

즉, 윈도우 크기와 임계치는 최종 침입 탐지 성능에 큰 영향을 미칠 수 있다. 기존 연구들은 실험 데이터 셋 전체를 기준으로 최적 성능을 보이는 값으로 설정하나, 각 탐지 환경에서 요구되는 보안 수준이나 위험 허용도가 다르기 때문에 환경에 맞춰 최적 성능을 발휘할 수 있는 값도 달라질 수 있다. 본 연구에서는 탐지 환경, 즉 위험 허용도를 기준으로 윈도우 크기와 임계치가 성능에 미치는 영향을 분석하고, 이를 바탕으로 HIDS에서 최종 침입 탐지를 위한 윈도우 크기와 임계치를 위험 허용도를 기준으로 설정할 것을 제안한다.

2. 배경 지식: 위험 허용도

보안 시스템의 성능 평가에 있어, Dempsey et al. (2011)[3]과 Antunes와 Vieira (2015)[4]는 세 가지 위험 허용도 기반 시나리오를 제시하였다. 이 연구들은 조직의 시스템 중요도와 위험 허용 범위에 따라 우선시 해야하는 성능 지표가 달라질 수 있음을 강조하였다 (표 1).

<표 1> 위험 허용도 기반 시나리오

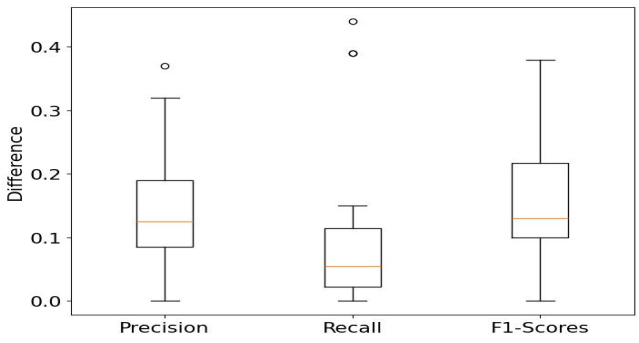
위험 허용	목표	주요 지표	예시
낮음	모든 공격 탐지	Recall	항공 관제시스템, 금융 거래시스템
중간	정밀도와 탐지율의 균형	F1-Score + P/R 균형	기업 네트워크 보안시스템
높음	정확한 경보	Precsion	공개 개시판, 일반 웹사이트

3. 분석 방법

마이크로서비스 환경에서 다양한 CVE 공격을 포함한 데이터셋[1]을 사용한다. 정상 데이터를 학습하고, 이상 패턴을 탐지하는 대표적인 비지도 학습 기반 오토인코더 모델을 사용한다[5]. 윈도우 크기와 임계치 설정은 관련 연구에서 사용된 값을 기반으로 진행한다[2]. 윈도우 크기는 {10, 20, 30, 40, 50}, 임계치는 {5%, 10%, 15%, 20%}로 설정하여 침입 탐지 성능을 평가한다. 평가 결과를 바탕으로 위험 허용 시나리오 별로 윈도우 크기와 임계치에 따라 성능이 어떻게 변화되는지 확인한다.

4. 분석 결과

그림 2는 오토인코더 모델의 윈도우 크기와 임계치에 따른 성능 변화폭을 보여준다. 분석 결과. Precision은 최대 37%까지 변화하였고, Recall은 최대 48%, F1-Score는 최대 38%까지 변화하는 등, 윈도우 크기와 임계치에 따라 성능이 크게 달라질 수 있음을 확인하였다. 즉 위험 허용도 시나리오에 따라 적절한 윈도우 크기와 임계치를 설정하지 않으면, 시나리오에 맞는 성능을 보장할 수 없는 문제가 있음을 확인하였다.



(그림 2) 윈도우 크기와 임계치에 따른 성능 변화폭

표 2는 위험 허용도에 따른 최적의 값과 그 성능을 보여준다. 각 위험 허용 시나리오에 적합한 윈도우 크기와 임계치는 해당 시나리오에서 주요 지표가 가장 높은 성능을 발휘할 수 있도록 설정하였다. 낮은 위험의 경우 작은 윈도우 크기와 낮은 임계치에서 최고의 성능을 보였으며, 중간 위험의 경우에는 비교적 큰 윈도우 크기와 높은 임계치가, 높은 위험의 경우 가장 큰 윈도우 크기와 높은 임계치에서 최고의 성능을 보였다. 이는 각 위험 허용 시나리오에 따라 더 적합한 값들이 존재하며, 각 시나리오의 목표에 맞춰 설정을 조정해야 한다는 것을 보여준다.

<표 2> 위험 허용도에 따른 최적의 윈도우 크기, 임계치, 및 주요 성능 지표

기준	Window size	임계치	성능 (주요 지표)
낮은 위험	10	5%	1.00 (Recall)
중간 위험	40	20%	0.49 (F1-score)
높은 위험	50	20%	0.48 (Precision)

5. 결론

본 연구는 HIDS에서 윈도우 크기와 임계치 설정이 탐지 성능에 미치는 영향을 분석하였으며, 그 결과 이러한 설정에 따라 성능이 크게 달라짐을 확인하였다. 따라서 고정된 최적값을 사용하는 대신, 각 시스템의 위험 허용도와 운영 환경에 맞춰 윈도우 크기와 임계치를 조정해야 할 것을 제안한다. 이를 통해 상황에 맞는 탐지 설정을 적용함으로써 실질적인 보안 성능을 극대화할 수 있다.

사사

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 소프트웨어중심대학사업(2021-0-01409), 인공지능융합혁신인재양성사업(IITP-2023-RS-2023-00256629)과 정보통신기획평가원의 지원(No.RS-2024-00438686, 비정상 오픈소스 식별 및 DevSecOps 자동 적용을 통한 소프트웨어 신뢰성 향상 기술 개발)을 받아 수행된 연구임.

참고문헌

[1] Flora, José. Evaluating intrusion detection for microservice applications. Journal of Systems and Software, 216, 112142, 2024.

[2] Araujo, Iury; Vieira, Marco. Evaluation of Machine Learning for Intrusion Detection. In: 12th Latin-American Symposium on Dependable and Secure Computing, 2023, pp. 126-135.

[3] Dempsey, Kelley L., et al. Information security continuous monitoring (ISCM). 2011.

[4] Antunes, Nuno; Vieira, Marco. On the metrics for benchmarking vulnerability detection tools. In: 45th IEEE/IFIP International Conference on Dependable Systems and Networks, 2015, pp. 505-516.

[5] Cheng, Zhen, et al. Improved autoencoder for unsupervised anomaly detection. International Journal of Intelligent Systems, 36(12), 7103-7125, 2021.