HỆ THỐNG PHÁT HIỆN XÂM NHẬP TỰ THÍCH NGHI ĐỂ PHÁT HIỆN CÁC KIỂU TẤN CÔNG MỚI

ADAPTIVE INTRUSION DETECTION SYSTEM FOR DETECTING NEW ATTACK PATTERNS

GVHD: PGS.TS Lê Đình Duy
Trương Gia Thạch - 240202012

Tóm tắt

- Lớp: CS2205.CH123
- Link Github của nhóm: https://github.com/TruongGiaThach/Final_Project-CS2205.CH183-240202012
- Link YouTube video: https://youtu.be/ud42ZaGuBDc
- Trương Gia Thạch

Giới thiệu

- Hệ thống phát hiện xâm nhập (Intrusion Detection System - IDS.
- IDS truyền thống gặp phải hạn chế lớn khi đối mặt với **zero-day attacks** và các phương thức tấn công mới do phụ thuộc vào **mẫu dữ liệu tĩnh** và **thiếu khả năng tự thích nghi**.
- Nghiên cứu này đề xuất một hệ thống **IDS tự thích nghi (Adaptive IDS) kết hợp Improved Random Forest (IRF) và Adaptive Learning Mechanism,** thử nghiệm trên bộ dữ liệu thực tế CICIDS2017 và UNSW-NB15.

Mục tiêu

- Phát triển hệ thống IDS tự thích nghi (Adaptive IDS), có khả năng học liên tục từ dữ liệu mới mà không cần huấn luyện lại toàn bộ.
- Tích hợp thuật toán Improved RF và Adaptive Learning Mechanism để tối ưu hóa khả năng phát hiện tấn công.
- Đánh giá hiệu suất mô hình dựa trên:
  - Tỷ lệ phát hiện tấn công mới (Detection Rate - DR).
  - Độ chính xác (Accuracy), Precision, Recall, F1-score.
  - Thời gian thích nghi của IDS khi có kiểu tấn công mới.

Nội dung và Phương pháp

Nội dung 1: Khảo sát các phương pháp phát hiện xâm nhập truyền thống và đề xuất mô hình IDS tự thích nghi

- Khảo sát các hệ thống phát hiện xâm nhập (IDS) hiện nay.
- Phân tích hạn chế của IDS truyền thống
- Xây dựng đề xuất mô hình Adaptive IDS, kết hợp Improved Random Forest (IRF) và Adaptive Learning Mechanism.
- Phương pháp thực hiện:
  - Nghiên cứu tài liệu
  - Thống kê và phân tích hiệu suất

Nội dung và Phương pháp

Nội dung 2: Phát triển mô hình IDS tự thích nghi dựa trên Improved RF và Adaptive Learning Mechanism

- Phát triển hệ thống IDS có khả năng cập nhật và học hỏi từ dữ liệu tấn công mới mà không cần huấn luyện lại toàn bộ mô hình.
- Tối ưu hóa hiệu suất bằng cách sử dụng Improved RF (IRF) và cơ chế Adaptive Learning.
- Phương pháp thực hiện:
  - Cải tiến thuật toán Random Forest
  - Thiết lập cơ chế Adaptive Learning

Nội dung và Phương pháp

Nội dung 3: Kiểm thử mô hình và đánh giá hiệu suất IDS tự thích nghi

- Đánh giá độ chính xác, tốc độ phản hồi, khả năng thích nghi của IDS so với IDS truyền thống.
- So sánh hiệu suất của IDS tự thích nghi với IDS không có cơ chế học liên tục.
- Phương pháp thực hiện:
  - Chạy thử nghiệm trên tập dữ liệu thực tế
  - Đánh giá hiệu suất dựa trên các tiêu chí chính: Accuracy, Detection Rate, False Positive Rate, Response Time

Kết quả dự kiến

- Cải thiện đáng kể hiệu suất IDS truyền thống:
  - Hệ thống IDS có thể thích nghi với các kiểu tấn công mới mà không cần đào tạo lại.
  - Độ chính xác trên 90%, giảm False Positive Rate xuống dưới 5%.
  - Tăng tốc độ phát hiện nhanh hơn ít nhất 30% so với IDS truyền thống.
- Báo cáo đánh giá đầy đủ về:
  - Khả năng học thích nghi của IDS khi dữ liệu tấn công thay đổi.
  - Hiệu suất trên dữ liệu thực tế và so sánh với IDS truyền thống.

Tài liệu tham khảo

[1] CICIDS2017 Dataset. [Online]. Available: https://www.unb.ca/cic/datasets/ids-2017.html
[2] UNSW-NB15 Dataset. [Online]. Available: https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/
[3] L. Breiman, “Random Forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001. doi: 10.1023/A:1010933404324. [Online]. Available: https://doi.org/10.1023/A:1010933404324
[4] J. Zhang and D.-M. Zhao, “Network Malicious Data Intrusion Detection Combining Distributed Network and Improved RF Algorithm under Spark Framework,” Journal of Network Intelligence, vol. 9, no. 3, pp. 1820-1835, 2024. doi: 10.32604/jni.2024.1835. [Online]. Available: https://doi.org/10.32604/jni.2024.1835
[5] M. Zaharia et al., “Spark: Cluster Computing with Working Sets,” in USENIX Conference on Hot Topics in Cloud Computing, 2010. [Online]. Available: https://www.usenix.org/conference/hotcloud-10/spark-cluster-computing-working-sets
[6] T. Kajiura and T. Nakamura, “Practical Performance of a Distributed Processing Framework for Machine-Learning-based NIDS,” IEEE Transactions on Dependable and Secure Computing, 2024. doi: 10.1109/TDSC.2024.10633597. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10633597/

Tài liệu tham khảo

[7] M. A. Shyaa, N. F. Ibrahim, Z. B. Zainol, R. Abdullah, and F. M. J. Ibrahim, “Reinforcement Learning-Based Voting for Feature Drift-Aware Intrusion Detection: An Incremental Learning Framework,” IEEE, 2025. doi: 10.1109/ICCEAI.2025.10896652. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10896652/
[8] A. A. Darem, F. A. Ghaleb, A. A. Al-Hashmi et al., “An Adaptive Behavioral-Based Incremental Batch Learning Malware Variants Detection Model Using Concept Drift Detection and Sequential Deep Learning,” IEEE Access, vol. 9, pp. 4321-4338, 2021. doi: 10.1109/ACCESS.2021.9467300. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9467300/
[9] Z. Jin, J. Zhou, B. Li, X. Wu, and C. Duan, “FL-IIDS: A Novel Federated Learning-Based Incremental Intrusion Detection System,” Future Generation Computer Systems, vol. 148, pp. 184-202, 2024. doi: 10.1016/j.future.2024.01.003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167739X23003503
[10] E. Mahdavi, A. Fanian, and A. Mirzaei, “ITL-IDS: Incremental Transfer Learning for Intrusion Detection Systems,” Applied Soft Computing, vol. 114, p. 107778, 2022. doi: 10.1016/j.asoc.2021.107778. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705122007778
[11] L. Cui, Z. Wu, P. Gao, and J. Chen, “An Incremental Learning Method Based on Dynamic Ensemble RVM for Intrusion Detection,” IEEE Transactions on Network and Service Management, vol. 18, no. 3, pp. 3215-3227, 2021. doi: 10.1109/TNSM.2021.9506882. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9506882/