THÔNG TIN CHUNG CỦA NHÓM

- Link YouTube video của báo cáo (tối đa 5 phút):
  https://www.youtube.com/watch?v=wCjZTluYy1c
- Link slides (dạng .pdf đặt trên Github của nhóm):
  https://github.com/VanDo-27-2-2024/CS2205.CH201

- Họ và Tên: Đỗ Đình VạVạn
- MSSV: 250202027

- Lớp: CS2205.CH201
- Tự đánh giá (điểm tổng kết môn): 8.5/10
- Số buổi vắng: 1
- Số câu hỏi QT cá nhân: 4
- Số câu hỏi QT của cả nhóm:
- Link Github:
  https://github.com/VanDo-27-2-2024/CS2205.CH201

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

XÂY DỰNG GATEWAY BẢO MẬT THÔNG MINH TẠI BIÊN PHÁT HIỆN MÃ ĐỘC IOT SỬ DỤNG MẠNG NƠ-RON SÂU

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

BUILDING AN INTELLIGENT EDGE SECURITY GATEWAY FOR IOT MALWARE DETECTION USING DEEP NEURAL NETWORKS

TÓM TẮT *(Tối đa 400 từ)*

Sự bùng nổ của các thiết bị Internet vạn vật (IoT) đã mở ra kỷ nguyên kết nối mới, nhưng đồng thời cũng tạo ra bề mặt tấn công rộng lớn cho tội phạm mạng. Các thiết bị IoT với khả năng bảo mật yếu kém thường xuyên bị khai thác để tạo thành các mạng máy tính ma (Botnet) khổng lồ như Mirai, thực hiện các cuộc tấn công từ chối dịch vụ (DDoS). Các giải pháp tường lửa truyền thống thường dựa trên chữ ký (Signature-based) hoặc đòi hỏi tài nguyên lớn, không phù hợp với đặc thù của mạng IoT.

Nghiên cứu này đề xuất giải pháp **AI Security Gateway** hoạt động ngay tại biên mạng (Edge Computing), sử dụng thiết bị phần cứng giá rẻ Raspberry Pi. Hệ thống kết hợp công cụ giám sát mạng **Zeek** để trích xuất siêu dữ liệu (Metadata) và mô hình **Mạng nơ-ron sâu (Deep Neural Network - DNN)** để phân tích hành vi lưu lượng. Bằng cách sử dụng bộ dữ liệu chuẩn **IoT-23** và loại bỏ các đặc trưng định danh (IP/Port), mô hình tập trung học các quy luật thống kê của dòng chảy dữ liệu (Flow-based features). Kết quả thực nghiệm cho thấy hệ thống đạt độ chính xác **99.99%** trong việc phát hiện các cuộc tấn công DDoS và Port Scan, với thời gian xử lý thời gian thực (<10ms) và đảm bảo quyền riêng tư người dùng do không giải mã nội dung gói tin.

GIỚI THIỆU

Trong thập kỷ qua, số lượng thiết bị IoT đã vượt qua dân số toàn cầu, len lỏi vào mọi ngõ ngách từ nhà thông minh đến công nghiệp (IIoT). Tuy nhiên, đặc điểm chung của các thiết bị này là tài nguyên tính toán hạn chế, firmware ít được cập nhật và thường sử dụng mật khẩu mặc định. Đây là điều kiện lý tưởng để các mã độc (Malware) lây nhiễm và biến chúng thành các "zombie" trong mạng Botnet. Các cuộc tấn công quy mô lớn như Mirai (2016) đã chứng minh sức tàn phá của Botnet IoT khi có thể đánh sập các dịch vụ DNS lớn trên thế giới.

Các phương pháp phát hiện xâm nhập hiện nay chủ yếu chia làm hai hướng: dựa trên chữ ký (Signature-based) và dựa trên bất thường (Anomaly-based). Phương pháp dựa trên chữ ký tuy chính xác nhưng bất lực trước các biến thể mã độc mới (Zero-day). Phương pháp dựa trên bất thường sử dụng Machine Learning truyền thống lại thường gặp vấn đề về tỷ lệ báo động giả cao. Hơn nữa, xu hướng đưa dữ liệu lên đám mây (Cloud) để phân tích gây ra độ trễ lớn và lo ngại về quyền riêng tư dữ liệu.

Để giải quyết vấn đề này, xu hướng Edge AI (Trí tuệ nhân tạo tại biên) đang lên ngôi. Việc đưa khả năng xử lý thông minh xuống ngay Gateway - cửa ngõ kết nối của mạng IoT - giúp phát hiện và ngăn chặn tấn công tức thì. Đề tài này tập trung nghiên cứu xây dựng một Gateway bảo mật sử dụng Deep Learning, cụ thể là kiến trúc DNN hình phễu, để phân loại lưu lượng mạng dựa trên hành vi. Giải pháp tận dụng Zeek Network Monitor làm tầng trích xuất đặc trưng mạnh mẽ, cho phép mô hình AI "hiểu" được ngữ cảnh mạng mà không cần can thiệp sâu vào gói tin, tạo ra một lớp bảo vệ chủ động, nhẹ và hiệu quả cho người dùng cuối.

MỤC TIÊU

Xây dựng hệ thống giám sát và trích xuất dữ liệu tại biên: Thiết lập Gateway trên nền tảng Raspberry Pi tích hợp công cụ Zeek để thu thập, phân tích và trích xuất các đặc trưng thống kê của luồng mạng (Flow statistics) theo thời gian thực.

Phát triển mô hình Deep Learning nhận diện tấn công: Thiết kế, huấn luyện và tối ưu hóa mô hình Mạng nơ-ron sâu (DNN) trên bộ dữ liệu IoT-23 để phân loại chính xác lưu lượng bình thường và lưu lượng mã độc (DDoS, Scanning) dựa trên hành vi, không phụ thuộc vào địa chỉ IP.
Triển khai và đánh giá thực nghiệm: Tích hợp mô hình vào Gateway, kiểm thử khả năng hoạt động trong môi trường thực tế với các kịch bản tấn công giả lập, đánh giá hiệu năng dựa trên độ chính xác, độ trễ và khả năng chịu tải.

NỘI DUNG VÀ PHƯƠNG PHÁP

1. Thu thập và Tiền xử lý dữ liệu:
    - Sử dụng bộ dữ liệu IoT-23), chứa traffic thực của các botnet nổi tiếng (Mirai, Torii...).
    - Sử dụng Zeek để chuyển đổi file PCAP thành `conn.log`.
    - Tiền xử lý: Loại bỏ IP nguồn/đích, Port nguồn/đích. Mã hóa (Label Encoding) các trường định tính (`proto`, `conn_state`, `history`) và Chuẩn hóa (Standard Scaling) các trường định lượng (`duration`, `orig_pkts`, `orig_ip_bytes`...).
2. Kiến trúc Mô hình Neural:
    - Sử dụng kiến trúc Feed-forward Deep Neural Network (DNN).
    - Cấu trúc hình phễu: Input (9 đặc trưng) $\rightarrow$ Dense (512) $\rightarrow$ Dense (256) $\rightarrow$ Dense (128) $\rightarrow$ Output (2 - Softmax).
    - Kỹ thuật tối ưu: Sử dụng hàm kích hoạt ReLU, áp dụng Dropout (0.5) và L2 Regularization để chống Overfitting.
3. Phương pháp Triển khai hệ thống:
    - Mô hình mạng: Sensor => AI Gateway (Raspberry Pi) => Target/Internet.
    - Phần mềm: Module Python (ai_guard.py) thực hiện đọc log thời gian thực từ Zeek (tail -f), nạp mô hình .h5 đã huấn luyện để dự đoán và đưa ra cảnh báo trên giao diện dòng lệnh.

4. Phương pháp Đánh giá:
    - Sử dụng các chỉ số: Accuracy, Precision, Recall, F1-Score và Loss.
    - Công cụ tấn công kiểm thử: `hping3` (Flood), `nmap` (Scan), `curl` (Normal).

KẾT QUẢ MONG ĐỢI

Sản phẩm phần mềm: Mã nguồn hoàn chỉnh của hệ thống AI Gateway, bao gồm các script tiền xử lý, file mô hình đã huấn luyện (`model.h5`, `scaler.pkl`) và chương trình giám sát thời gian thực.

Sản phẩm phần cứng: Một thiết bị Gateway (trên nền tảng Raspberry Pi hoặc máy ảo Ubuntu) hoạt động ổn định, có khả năng bắt và phân tích gói tin đi qua nó.

Chỉ số hiệu năng:

- Độ chính xác nhận diện (Accuracy) trên tập kiểm thử đạt > 95% (Thực tế đạt 99.99%).
- Thời gian xử lý trung bình mỗi dòng log < 50ms.
- Phát hiện chính xác các hành vi DDoS SYN Flood, UDP Flood và Port Scanning.

TÀI LIỆU THAM KHẢO

[1]. Sebastian Garcia, Agustin Parmisano, Maria Jose Erquiaga: IoT-23: A labeled dataset with malicious and benign IoT network traffic. Zenodo 2020.

[2]. Vern Paxson: Bro: a system for detecting network intruders in real-time. Computer Networks 31(23-24): 2435-2463 (1999).

[3]. Arnan Sae-Tang, M. L. Dennis Wong, Steven R. Weller: NID-Net: A Deep

Neural Network for Network Intrusion Detection. ISCIT 2021: 271-276

[4]. Chuanlong Yin, Yuefei Zhu, Jinlong Fei, Xinzheng He: A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks. IEEE Access 5: 21954-21961 (2017)