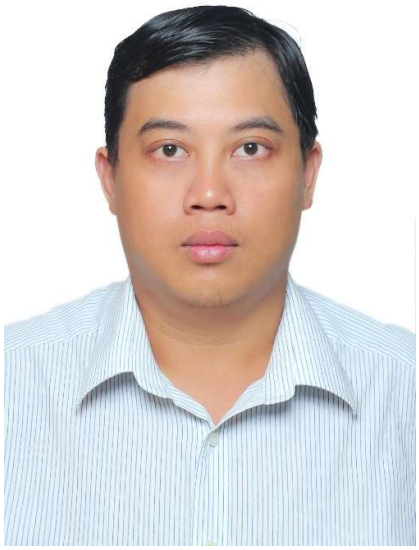
THÔNG TIN CHUNG HỌC VIÊN

- Họ và Tên: **Huỳnh Trần Trọng Nghĩa**
- MSHV: 250101047



- Lớp: CS2205.CH201
- Tự đánh giá (điểm tổng kết môn): 9.5/10
- Số buổi vắng: 0
- Link Github: https://github.com/nghiahuynh8620/CS2205.CH201
- Link Slide: https://github.com/nghiahuynh8620/CS2205.CH201/blob/main/Ngh%C4%A9a%2C%20Hu%E1%BB%B3nh%20Tr%E1%BA%A7n%20Tr%E1%BB%8Dng%20-%20CS2205.SEP2025.DeCuong.FinalReport.Slide.pdf
- Link Youtube: https://youtu.be/8iREqJGTmyA

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI

ỨNG DỤNG PHÁT HIỆN HÀNH VI BẤT THƯỜNG NGƯỜI DÙNG TRONG MẠNG DOANH NGHIỆP BẰNG KHAI PHÁ DỮ LIỆU VÀ HỌC MÁY

TÊN ĐỀ TÀI TIẾNG ANH

ANOMALOUS USER BEHAVIOR DETECTION IN ENTERPRISE NETWORKS USING DATA MINING AND MACHINE LEARNING APPLICATION

TÓM TẮT

Trong môi trường mạng doanh nghiệp, các hành vi lệch chuẩn của người dùng - như đăng nhập bất thường theo thời gian/địa điểm, truy cập tài nguyên nhạy cảm ngoài quy trình, tải xuống khối lượng lớn, di chuyển ngang hoặc lạm dụng đặc quyền - thường là dấu hiệu sớm của tài khoản bị chiếm quyền và/hoặc là mối đe doạ nội bộ. Đề tài tập trung xây dựng ứng dụng phát hiện hành vi bất thường dựa trên khai phá dữ liệu nhật ký (log) và các thuật toán học máy.

Dữ liệu đầu vào gồm các nguồn log phổ biến trong doanh nghiệp: xác thực (AD/SSO), truy cập file, lưu lượng mạng/Proxy, hoạt động endpoint (EDR/Process) và (tuỳ điều kiện) audit log cloud/email/USB. Hệ thống thực hiện chuẩn hoá schema (user–host–resource–action–timestamp), làm sạch, đồng bộ thời gian và gom sự kiện theo cửa sổ thời gian để trích xuất đặc trưng theo ngữ cảnh (tần suất, độ hiếm, xu hướng, chuỗi hành động, quan hệ user–host–resource). Trên nền đặc trưng này, đề tài triển khai và so sánh hai hướng tiếp cận:

(i) Mô hình Deel Learning như Autoencoder/Transformer

(ii) Mô hình Supervised Learning (ví dụ Random Forest/XGBoost) khi có nhãn sự kiện.

Đầu ra của hệ thống là điểm bất thường và cảnh báo kèm giải thích (đặc trưng đóng góp

chính/chuỗi hành động nổi bật) và giao diện tra cứu theo user/timestamp. Đề tài đánh giá trên bộ dữ liệu công khai phù hợp và/hoặc dữ liệu mô phỏng theo kịch bản (tham chiếu MITRE ATT&CK), sử dụng các chỉ số Precision/Recall/F1, PR-AUC/ROC-AUC, tỉ lệ cảnh báo sai (FPR) và thời gian phát hiện. Sản phẩm cuối gồm pipeline xử lý dữ liệu, mô hình, báo cáo thực nghiệm và demo ứng dụng.

GIỚI THIỆU

Bối cảnh & vấn đề: Trong các hệ thống giám sát An Toàn Thông Tin, cách tiếp cận dựa trên luật (rule/signature) có ưu điểm dễ triển khai nhưng khó bao phủ các hành vi nội bộ tinh vi hoặc các chuỗi hành động "hợp lệ" về mặt cú pháp nhưng bất thường theo ngữ cảnh. UEBA (User and Entity Behavior Analytics) khắc phục hạn chế này bằng cách học "baseline" hành vi bình thường theo từng người dùng/nhóm (peer group), từ đó phát hiện độ lệch và ưu tiên cảnh báo.

Đầu vào: log sự kiện từ các nguồn trong mạng doanh nghiệp (authentication, file access, network/proxy, endpoint/process, cloud audit…).

Đầu ra:

(i) cảnh báo hành vi bất thường kèm điểm số;

(ii) giải thích/ngữ cảnh phục vụ điều tra (nguồn log, chuỗi hành động, các đặc trưng nổi bật);

(iii) báo cáo đánh giá mô hình.

Động lực & tính ứng dụng: hỗ trợ SOC/Blue Team phát hiện sớm Insider Threat, Account Takeover, Privilege Misuse; giảm thời gian điều tra bằng cách ưu tiên các cảnh báo có độ tin cậy cao và cung cấp giải thích có thể kiểm chứng.

Đóng góp dự kiến:

(1) Xây dựng pipeline khai phá log và bộ đặc trưng UEBA có thể tái sử dụng;

(2) Triển khai và so sánh nhóm mô hình ML/DL cho bài toán phát hiện bất thường

theo log;

(3) Phát triển demo ứng dụng tra cứu, trực quan hoá và xuất danh sách cảnh báo.

MỤC TIÊU

1) Xây dựng bộ dữ liệu và pipeline tiền xử lý log (chuẩn hoá, làm sạch, ghép nguồn, gom theo cửa sổ thời gian) cùng bộ đặc trưng UEBA.

2) Thiết kế, huấn luyện và so sánh các mô hình phát hiện bất thường (không giám sát/bán giám sát và có giám sát khi có nhãn), tối ưu hoá để giảm cảnh báo sai.

3) Xây dựng ứng dụng demo phục vụ tra cứu và cảnh báo; đánh giá định lượng trên bộ dữ liệu lựa chọn, kèm báo cáo và hướng dẫn tái lập.

NỘI DUNG VÀ PHƯƠNG PHÁP

Khảo sát & lựa chọn dữ liệu:

Ưu tiên bộ dữ liệu log/insider threat công khai phù hợp; trường hợp thiếu nhãn, xây dựng log mô phỏng từ kịch bản (tham chiếu MITRE ATT&CK) để kiểm thử.

Tiền xử lý & chuẩn hoá:

Parse log, chuẩn hoá schema (user, host, action, resource, timestamp), xử lý trùng lặp/thiếu dữ liệu, đồng bộ múi giờ.

Gom sự kiện theo time window (ví dụ 5 phút/1 giờ/1 ngày) và tạo chuỗi hành động theo user/host.

Khai phá dữ liệu & thiết kế đặc trưng (feature engineering):

Thống kê: tần suất đăng nhập, số file truy cập, dung lượng tải xuống, số host liên quan, entropy/độ hiếm hành vi.

Chuỗi & ngữ cảnh: n-gram hành động, khoảng thời gian giữa sự kiện, đặc trưng theo vai trò và peer group.

Đồ thị (tuỳ chọn): user–host–resource graph để phát hiện quan hệ/đường đi bất thường.

Mô hình hoá:

Baseline Supervised Learning: Random Forest/XGBoost; xử lý mất cân bằng bằng Sampling/Weight.

Deep Learning: Autoencoder/Transformer tự giám sát cho chuỗi sự kiện hoặc embedding log.

Đánh giá & giải thích:

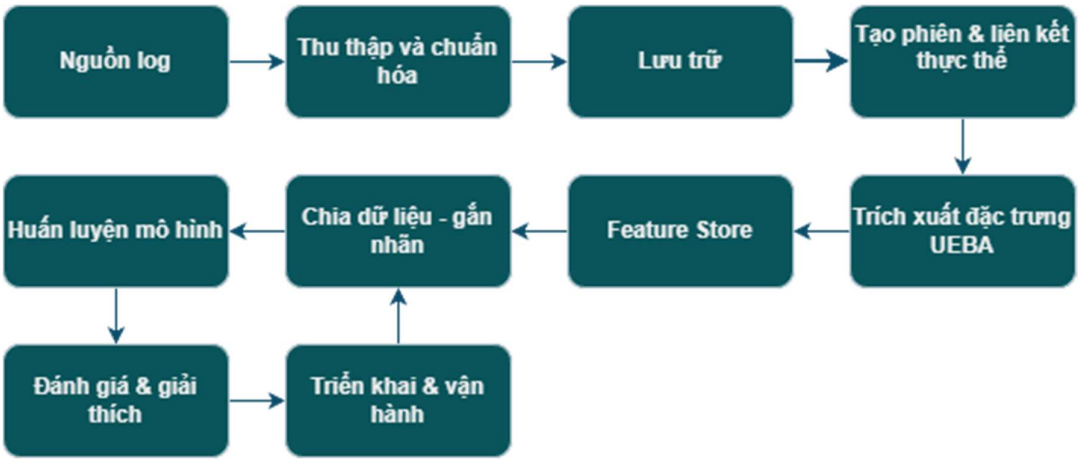
Chia Train/Test/Val theo thời gian để tránh rò rỉ dữ liệu (data leakage); đánh giá bằng Precision/Recall/F1, PR-AUC/ROC-AUC, FPR và thời gian suy luận.

Giải thích: feature importance/SHAP (tree-based), phân rã reconstruction error (Autoencoder).

Triển khai demo:

Pipeline batch (offline) kết hợp API/GUI (Streamlit/FastAPI); hiển thị timeline, top alerts và drill-down theo user.

Sơ đồ pipeline:



KẾT QUẢ MONG ĐỢI

Bộ dữ liệu đã chuẩn hoá (hoặc bộ dữ liệu công khai được chuẩn bị sẵn) và mã nguồn pipeline ETL/feature.

Tối thiểu 2–3 mô hình Baseline + 1 mô hình DL, kèm so sánh định lượng và phân tích lỗi.

Ứng dụng demo: tra cứu hành vi theo user/timestamp, hiển thị điểm bất thường, xuất danh sách cảnh báo.

Báo cáo đề tài (kèm phụ lục cấu hình, hướng dẫn chạy) và slide/video demo

TÀI LIỆU THAM KHẢO

[1]. David Alvarez Muniz, Luis Perez Miguel, et al.: Design and Generation of a Dataset for Training Insider Threat Prevention and Detection Models: The SPEDIA Dataset. Computers & Security 2025: 104743.

[2]. W. Feng, et al.: Multi-Granularity User Anomalous Behavior Detection. Applied Sciences 2024, 15(1):128.

[3]. Crispin Almodovar, Fariza Sabrina, Sarvnaz Karimi, et al.: LogFiT: Log Anomaly Detection Using Fine-Tuned Language Models. IEEE Transactions on Network and Service Management 2024, 21(2):1715-1723.

[4]. S. Song, et al.: Confront Insider Threat: Precise Anomaly Detection in Behavior Logs Based on LLM Fine-Tuning. COLING 2025.

[5]. (Scientific Reports): System log anomaly detection based on contrastive learning and retrieval augmented. Scientific Reports 2025.

[6]. Max Landauer, Florian Skopik, et al.: Cloud-based User Entity Behavior Analytics Log Data Set (CLUE-LDS). Zenodo 2022.