

**ĐỀ CƯƠNG CHI TIẾT KHÓA LUẬN TỐT NGHIỆP**  
**HỌC KỲ 1 NĂM HỌC 2025 - 2026**

<b>Tên đề tài (viết chữ thường):</b> <ul style="list-style-type: none"><li>- Tên tiếng Việt: Phát hiện mã độc PowerShell bằng mạng nơ-ron và kỹ thuật giải rối mã</li><li>- Tên tiếng Anh: Detecting Malicious PowerShell Scripts using Neural Networks and Deobfuscation</li></ul>
<b>Cán bộ hướng dẫn</b> (ghi đầy đủ họ tên và học hàm, học vị): ThS. Nghi Hoàng Khoa
<b>Thời gian thực hiện:</b> Từ ngày 20/09/2025 đến ngày 20/12/2025
<b>Sinh viên thực hiện:</b> Hồ Vi Khánh – 22520633 – ATTT2022.1 - 0869008902
<b>Nội dung đề tài</b> (Mô tả chi tiết mục tiêu, phạm vi, đối tượng, phương pháp thực hiện, kết quả mong đợi của đề tài): Mã độc phi mã đang là mối đe dọa và là thách thức lớn trong lĩnh vực an ninh mạng. Do khả năng ẩn mình và trốn tránh tinh vi có thể vượt qua các phương pháp phát hiện mã độc truyền thống gây ra không ít khó khăn trong việc nhận biết, phát hiện và đề phòng bởi các cuộc tấn công phi mã. Nhằm mục đích khắc phục hạn chế của các phương pháp trước đây (chỉ dựa vào code hoặc chỉ dựa vào AST) bằng cách tiếp cận bằng cách kết hợp nhiều nguồn thông tin như: mã nguồn, cây cấu trúc trườ tượng (AST), đồ thị luồng điều khiển (CFG) và embedding từ các mô hình ngôn ngữ hiện đại. Phạm vi đề tài tập trung vào phân tích tĩnh mã nguồn, biểu diễn dưới dạng đồ thị và đánh giá các mã độc phi mã sử dụng PowerShell được thu thập trên các bộ dữ liệu công khai. Đối tượng nghiên cứu là các đoạn mã PowerShell scripts lành tính, mã độc và các mã độc đã được làm rối. Các thành phần biểu diễn mã độc như: AST nodes, function calls, CFG, embedding từ BERT/Electra/CodeBERT, GraphCodeBERT. Phương pháp thực hiện là phân tích và trích xuất đặc trưng của các mã nguồn thu thập được như: phân tích shellcode, URL, IP, entropy. Trích xuất hàm, node đặc trưng từ AST và biểu diễn các quan hệ thực thi của hàm từ CFG. Xử lý bằng mô hình ngôn ngữ (LMs) và Transformer. Sau đó kết hợp tất cả vector đặc trưng và phân loại chúng bằng mô hình dự đoán hiện đại. Đồng thời ứng dụng XAI giải thích các kết quả thu được để định lượng mức độ ảnh hưởng của từng đặc trưng đến quyết định mô hình. Kết quả mong đợi có thể ứng dụng máy học và các phương pháp hiện đại nhất để tạo ra một mô hình nhận diện được các mã độc phi mã hiệu quả. Phát hiện tốt các mã độc bị làm rối và giải thích để tăng độ minh bạch và tin

*cậy cho mô hình. Có thể tích hợp vào SOC hoặc hệ thống phát hiện malware hiện có, cung cấp vừa cảnh báo vừa insight cho nhà phân tích mã độc trên thực tế.*

**Kế hoạch thực hiện**(Mô tả kế hoạch làm việc, thời gian biểu và phân công công việc cho từng sinh viên tham gia):

21/09/2025 – 30/09/2025: Chuẩn bị, nghiên cứu và tìm hiểu các đề tài liên quan đến mã độc phi mã. Thiết lập môi trường.

01/10/2025 – 14/10/2025: Thu thập dữ liệu và tiền xử lý các dữ liệu đã thu thập được. Thống kê và phân tích sơ bộ.

15/10/2025 – 29/10/2025: Trích xuất đặc trưng xây dựng và trích xuất các nodes của AST. Tạo CFG từ AST. Trích xuất feature: entropy, shellcode, URL/IP.

30/10/2025 – 13/11/2025: Fine-tune các mô hình ngôn ngữ để tăng hiệu quả. Đồng thời tạo các embedding chuẩn bị đầu vào cho mô hình dự đoán.

14/11/2025 – 28/11/2025: Kết hợp tất cả features + embeddings để huấn luyện và đánh giá mô hình.

28/11/2025 – 06/12/2025: Tích hợp thêm XAI để giải thích kết quả.

07/12/2025 – 20/12/2025: Hoàn thiện báo cáo khóa luận, chuẩn bị slide, demo.

**Tài liệu tham khảo** (theo chuẩn IEEE):

[1] N. Basheer, B. Pranggono, S. Islam, S. Papastergiou, and H. Mouratidis, “Enhancing malware detection through machine learning using XAI with SHAP framework,” in IFIP International Conference on Artificial Intelligence Applications and Innovations. Springer, 2024.

[2] A. Y. M. Benselloua, S. A. Messadi, and A. E. Belfedhal, “Effective malicious PowerShell scripts detection using DistilBERT,” in Proc. IEEE Afro-Mediterranean Conference on Artificial Intelligence (AMCAI), 2023.

[3] Y. Fang, X. Zhou, and C. Huang, “Effective method for detecting malicious PowerShell scripts based on hybrid features,” Neurocomputing, 2021.

**Xác nhận của CBHD**

(ghi rõ họ tên)

**TP. HCM, ngày 21 tháng 09 năm 2025**

**Sinh viên**

(ghi rõ họ tên)

**Hồ Vĩ Khánh**