**ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**KHOA MẠNG MÁY TÍNH VÀ TRUYỂN THÔNG**

A blue logo with a planet and a flower

Description automatically generated

**BÁO CÁO THỰC TẬP DOANH NGHIỆP**

**TRUNG TÂM AN NINH MẠNG**

**Người hướng dẫn: anh Hà Vũ Minh Ngọc**

**Giảng viên hướng dẫn: Ths. Nghi Hoàng Khoa**

**Sinh viên thực hiện: Huỳnh Minh Tân Tiến - 21521520**

**Lớp: NT215.P11.ATCL**

**TP. HỒ CHÍ MINH, 12-2024**

|  |  |
| --- | --- |
| TRƯỜNG ĐẠI HỌC  CÔNG NGHỆ THÔNG TIN **TRUNG TÂM AN NINH MẠNG** | **CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**  **Độc lập – Tự do – Hạnh phúc** |

**PHIẾU XÁC NHẬN THỰC TẬP**

**Trung tâm An ninh Mạng xác nhận:**

Anh/chị: Huỳnh Minh Tân Tiến Sinh ngày: 13/03/2003

Sinh viên năm thứ: 4 Khoa: Mạng máy tính & Truyền thông

Trường: Đại học Công Nghệ thông tin, ĐHQG-HCM

Đã thực tập tại: Phòng E8.1, Trường ĐH Công nghệ Thông tin, Khu phố 6, phường Linh Trung, Thành phố Thủ Đức, Tp. HCM

Thời gian thực tập: 3 tháng

Vị trí thực tập: Thực tập sinh

Nhiệm vụ được giao và hướng dẫn: Nghiên cứu Phân loại họ mã độc PowerShell dựa trên kết hợp ngữ nghĩa đa phương thức và học sâu.

Cán bộ phụ trách và hướng dẫn: Hà Vũ Minh Ngọc

Chức vụ: Chuyên viên

**Đánh giá quá trình thực tập của sinh viên:**

Các kết quả sinh viên đã thực hiện được:

|  |  |  |
| --- | --- | --- |
| **Các tiêu chí đánh giá** | **Điểm đánh giá**  **(Theo thang điểm 10)** | **Ghi chú** |
| 1. Năng lực chuyên môn |  |  |
| 1. Chất lượng công việc |  |  |
| 1. Tinh thần trách nhiệm |  |  |
| 1. Tính chủ động sáng tạo |  |  |
| 1. Tính kỷ luật |  |  |
| **Tổng điểm:** |  |  |

Đánh giá khác:

|  |  |  |
| --- | --- | --- |
|  |  | …, ngày…tháng …năm … |
| **TRƯỞNG ĐƠN VỊ** | **NHÂN SỰ** | **CÁN BỘ QUẢN LÝ TRỰC TIẾP** |

|  |  |
| --- | --- |
| TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN  **KHOA MẠNG MÁY TÍNH**  **VÀ TRUYỀN THÔNG** | **CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM**  **Độc lập – Tự do – Hạnh phúc** |

**NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN**

Họ và tên sinh viên: Huỳnh Minh Tân Tiến

MSSV: 21521520

Công ty thực tập: Trung tâm an ninh mạng

Thời gian thực tập: 3 tháng

Vị trí thực tập: Thực tập sinh

Nhiệm vụ được giao: Nghiên cứu Phân loại họ mã độc PowerShell dựa trên kết hợp ngữ nghĩa đa phương thức và học sâu.

**Đánh giá quá trình thực tập của sinh viên:**

Các kết quả sinh viên đã thực hiện được:

Điểm: Bằng chữ:

|  |  |  |
| --- | --- | --- |
|  |  | ……., ngày…..tháng …..năm …… |
|  |  | **Giáo viên hướng dẫn** |

[Chương 1. Giới thiệu về doanh nghiệp thực tập 9](#_Toc186382723)

[1.1. Thông tin doanh nghiệp 9](#_Toc186382724)

[1.2. Lĩnh vực hoạt động 9](#_Toc186382725)

[1.3. Thông tin thêm 9](#_Toc186382726)

[Chương 2. VỊ TRÍ VÀ YÊU CẦU 10](#_Toc186382727)

[2.1. Vị trí thực tập 10](#_Toc186382728)

[2.2. Yêu cầu với vị trí thực tập 10](#_Toc186382729)

[2.3. Thời gian thực tập 10](#_Toc186382730)

[2.4. Mô tả công việc thực tập 10](#_Toc186382731)

[Chương 3. NỘI DUNG THỰC TẬP 11](#_Toc186382732)

[3.1. Kanban Work Management 11](#_Toc186382733)

[3.2. Git Workflow 12](#_Toc186382734)

[3.3. Dự án: nghiên cứu phân loại họ mã độc powershell dựa trên kết hợp ngữ nghĩa đa phương thức và học sâu 14](#_Toc186382735)

[3.3.1. Tìm hiểu lý thuyết, đề tài và case study 15](#_Toc186382736)

[3.3.2. Khảo sát các nghiên cứu liên quan 16](#_Toc186382737)

[3.3.3. Tái hiện và phát triển 20](#_Toc186382738)

[Chương 4. KẾT QUẢ VÀ ĐÁNH GIÁ 28](#_Toc186382739)

[4.1. Kết quả đạt được 28](#_Toc186382740)

[4.2. Đánh giá và nhận xét 28](#_Toc186382741)

DANH MỤC HÌNH VẼ

[Hình 1: Website doanh nghiệp 9](#_Toc186382657)

[Hình 2: Minh họa về kanban board 11](#_Toc186382658)

[Hình 3: Mô tả về Git workflow 13](#_Toc186382659)

[Hình 4: Chuỗi lây nhiễm Log4J 16](#_Toc186382660)

[Hình 5: Kiến trúc 5 lớp của data provenance 17](#_Toc186382661)

[Hình 6: Kiến trúc của PowerDetector 18](#_Toc186382662)

[Hình 7: Kiến trúc chuyển Payload Thành Ảnh và Phát Hiện Bằng CNN 20](#_Toc186382663)

[Hình 8: Kiến trúc mô hình kết hợp 25](#_Toc186382664)

[Hình 9: Ma trận nhầm lẫn 27](#_Toc186382665)

DANH MỤC BẢNG

[Bảng 1: Lệnh Git 14](#_Toc186382678)

[Bảng 2: Độ hiệu quả của các kỹ thuật giảm log 17](#_Toc186382679)

[Bảng 3: Tập data dữ liệu PowerShell script 26](#_Toc186382680)

[Bảng 4: Chỉ số hiệu suất của mô hình 26](#_Toc186382681)

LỜI CẢM ƠN

Lời cảm ơn đầu tiên, trân trọng gửi đến thầy Nghi Hoàng Khoa, đã giới thiệu và tạo điều kiện hỗ trợ em trong suốt kỳ thực tập, cùng với đó là lời cảm ơn đến Trung tâm an ninh mạng đã cho em cơ hội thực tập và cuối cùng là lời cảm ơn đến anh Hà Vũ Minh Ngọc đã hỗ trợ, dẫn dắt em trong suốt thời gian thực tập. Nhờ sự hỗ trợ và giúp đỡ từ mọi người mà em đã hoàn thành suôn sẻ được kỳ thực tập tại trung tâm an ninh mạng.

Em xin chân thành cảm cảm ơn!

Huỳnh Minh Tân Tiến

MỞ ĐẦU

Thực tập là một phần không thể thiếu trong quá trình học tập, đó là giai đoạn để sinh viên sử dụng những kiến thức đã học được từ nhà trường để áp dụng và giải quyết các vấn đề thực tiễn của doanh nghiệp. Thông qua kỳ thực tập, sinh viên sẽ được trao dồi thêm những kiến thức thực tế, đồng thời tự đánh giá bản thân, tìm hiểu đượcnhững điểm mạnh, những việc làm phù hợp với bản thân hay biết bản thân còn đang khiếm khuyết ở những điểm nào để cải thiện và bổ sung cho phù hợp với nhu cầu hiện tại của doanh nghiệp.

Trong khoảng thời gian thực tập tại trung tâm an ninh mạng, em đã học tập được nhiều thứ về kiến thứ, kỹ năng và tư duy tốt cùng với đó là có cơ hội thực hiện một dự án nghiên cứu. Qua tất cả những điều trên đã xây dựng cho em một nền tảng mạnh mẽ để cho bản thân có đủ khả năng đáp ứng một nhu cầu ngày càng cao của các nhà tuyển dụng hiện tại.

Với môi trường làm việc gắn chặt đến chuyên ngành hiện tại em đang theo học và trong 3 tháng thực tập, em xin tổng hợp và trình bày lại toàn bộ quá trình thông qua báo cáo với các nội dung:

Chương 1: Giới thiệu về doanh nghiệp thực tập

Chương 2: Vị trí và yêu cầu

Chương 3: Nội dung thực tập

Chương 4: Kết quả và đánh giá

# NỘI DUNG

## Dự án: nghiên cứu phân loại họ mã độc powershell dựa trên kết hợp ngữ nghĩa đa phương thức và học sâu

Yêu cầu: Nghiên cứu phân loại họ mã độc PowerShell bằng phương pháp kết hợp ngữ nghĩa đa phương thức và học sâu. Cần tìm hiểu khái niệm về mã độc PowerShell, phân loại mã độc, ngữ nghĩa đa phương thức, học sâu, và ứng dụng thông qua các case study tiêu biểu. Khảo sát các nghiên cứu liên quan, đặc biệt những phương pháp sử dụng học sâu và tích hợp ngữ nghĩa trong phát hiện mã độc. Tiếp theo, tái hiện nghiên cứu têu biểu để hiểu rõ quy trình, sau đó cải tiến, phát triển dựa trên việc áp dụng ngữ nghĩa đa phương thức nhằm nâng cao độ chính xác và hiệu quả trong phân loại mã độc.

### Tìm hiểu lý thuyết, đề tài và case study

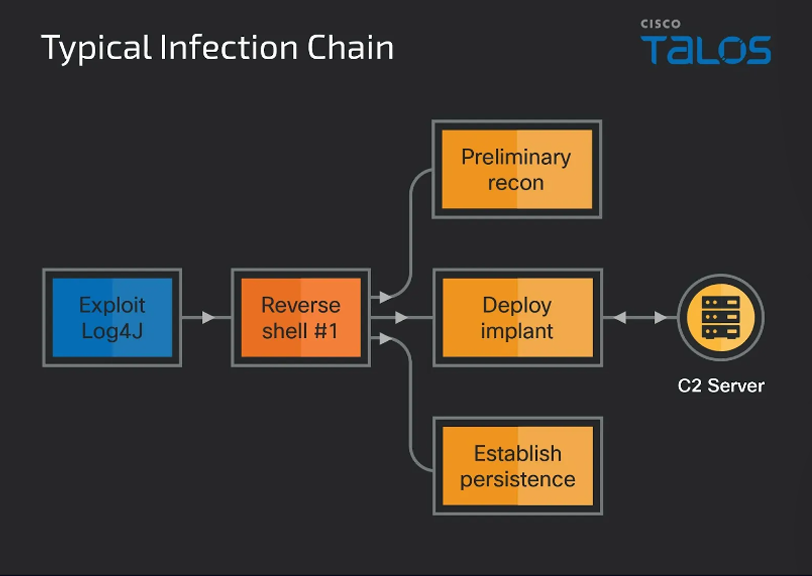
**Fileless Malware**   
 Fileless Malware là loại phần mềm độc hại không được lưu trữ thành dạng file trên máy tính mà chỉ tồn tại trong vùng nhớ máy tính. Để tránh việc phải lưu trữ, mã độc tồn tại bằng cách tận dụng các công cụ và ứng dụng có sẵn, mã độc có thể xâm nhập, tổn hại và thao túng hệ thống cùng với khả năng ẩn nấp khỏi các công cụ truy vết virus.

Thông thường Fileless malware là giai đoạn đầu tiên của các cuộc tấn công của các cuộc tấn công. Mặc dù không có lưu dưới dạng file, nhưng khi thực hiện tấn công, malware có thể sử dụng các tập lệnh hoặc tải xuống các tập lệnh.

**Casestudy: Operation Blacksmith**

Kỹ thuật được sử dụng trong Operation Blacksmith:

* Khai thác lỗ hổng Log4Shell: Nhóm Lazarus đã lợi dụng lỗ hổng thực thi mã từ xa trong Log4j để xâm nhập vào hệ thống mục tiêu.
* Triển khai phần mềm độc hại mới: Sau khi xâm nhập, nhóm đã triển khai ba phần mềm độc hại mới được viết bằng ngôn ngữ DLang:
* NineRAT: Một trojan truy cập từ xa sử dụng API của Telegram để giao tiếp với máy chủ điều khiển (C2), giúp truyền tệp và tránh bị phát hiện.
* DLRAT: Một trojan truy cập từ xa khác cho phép triển khai thêm phần mềm độc hại và thực hiện các lệnh từ máy chủ C2.
* BottomLoader: Một trình tải xuống phần mềm độc hại, có khả năng truy xuất và thực thi các tải trọng giai đoạn tiếp theo từ máy chủ từ xa.
* Sử dụng công cụ proxy HazyLoad: Nhóm đã sử dụng HazyLoad để thiết lập quyền truy cập liên tục vào hệ thống bị xâm nhập, cho phép thực hiện các lệnh trinh sát và tạo tài khoản người dùng đặc quyền.



Hình : Chuỗi lây nhiễm Log4J

### Khảo sát các nghiên cứu liên quan

#### Nghiên cứu: “SoK: History is a Vast Early Warning System: Auditing the Provenance of System Intrusions” [1]

Tổng quan về nghiên cứu:

* Mục đích của nghiên cứu: Tài liệu trình bày tổng quan về các phương pháp dựa trên "data provenance" (nguồn gốc dữ liệu) nhằm cải thiện khả năng phát hiện và điều tra các xâm nhập hệ thống.
* Cách tiếp cận với đề tài: Tác giả thực hiện khảo sát và phân loại các nghiên cứu về auditing dựa trên "data provenance" thành năm lớp (Capture, Reduction, Infrastructure, Detection, Investigation) trong chuỗi phân tích log. Đồng thời, họ đánh giá hiệu quả của 8 kỹ thuật giảm dung lượng log bằng bộ dữ liệu DARPA Transparent Computing.

Phương pháp và kỹ thuật thực hiện:

* Phân loại các kỹ thuật auditing dựa trên các tầng trong quy trình thu thập và phân tích log.
* So sánh các kỹ thuật giảm dung lượng log về hiệu quả lưu trữ, khả năng truy vấn và tính toàn vẹn của dữ liệu.
* Thực hiện đánh giá độc lập các kỹ thuật giảm log trên các bộ dữ liệu công khai.
* Các kỹ thuật giảm log như LogGC, DPR, và CPR để tối ưu hóa kích thước và bảo toàn thông tin trong log.
* Phân tích đồ thị dựa trên quan hệ nhân quả để trích xuất thông tin có liên quan.

A diagram of a process

Description automatically generated

Hình : Kiến trúc 5 lớp của data provenance

Kết quả của nghiên cứu:

Đề cập đến các thông tin liên quan kỹ thuật giảm log có thể giảm kích thước log tới 185 lần trong hệ thống audit khi kết hợp với nén dữ liệu, nhưng một số kỹ thuật như DPR có thể làm mất thông tin quan trọng liên quan đến các cuộc tấn công.

A table with numbers and text

Description automatically generated

Bảng : Độ hiệu quả của các kỹ thuật giảm log

#### Nghiên cứu: “PowerDetector: Malicious PowerShell Script Family Classification Based on Multi-Modal Semantic Fusion and Deep Learning” [2]

Tổng quan về nghiên cứu:

* Mục đích của nghiên cứu: Giới thiệu PowerDetector, một hệ thống phân loại tập lệnh PowerShell độc hại dựa trên việc hợp nhất ngữ nghĩa đa chiều và học sâu, nhằm cải thiện độ chính xác trong nhận dạng các tấn công fileless và APT.
* Cách tiếp cận với đề tài: Nghiên cứu tập trung vào trích xuất các đặc trưng từ bốn chiều khác nhau của tập lệnh PowerShell, sau đó hợp nhất các vector đặc trưng thông qua phương pháp học sâu kết hợp Transformer và CNN-BiLSTM để thực hiện phân loại đa lớp.

A diagram of a computer

Description automatically generated with medium confidence

Hình : Kiến trúc của PowerDetector

Phương pháp và kỹ thuật thực hiện:

* Thiết kế bốn phương pháp trích xuất đặc trưng bao gồm: ký tự (Char2Vec), token (Token2Vec), cây cú pháp trừu tượng (AST2Vec), và đồ thị tri thức (Rela2Vec).
* Xây dựng một mô hình kết hợp Transformer và CNN-BiLSTM, sử dụng cơ chế attention để tăng cường khả năng phát hiện.
* Sử dụng tập dữ liệu PowerShell từ các nguồn thực tế và công cụ tấn công để đào tạo và kiểm thử.
* Multi-modal embedding: hợp nhất các vector đặc trưng từ nhiều chiều.
* Học sâu với BiLSTM và Transformer để xử lý mối quan hệ ngữ nghĩa dài và phức tạp.

Kết quả nghiên cứu:

* PowerDetector đạt hiệu suất cao, vượt trội hơn các phương pháp hiện tại.
* Hệ thống phát hiện hiệu quả các tấn công obfuscated và nhận diện được các tập lệnh PowerShell thuộc năm nhóm chức năng: Bypass, Downloader, Injector, Payload, và TaskExecution.

#### Nghiên cứu: “Machine learning based fileless malware traffic classification using image visualization” [3]

Tổng quan về nghiên cứu:

* Mục đích của nghiên cứu: Đề xuất một phương pháp mới sử dụng kỹ thuật chuyển đổi lưu lượng mạng thành hình ảnh và mạng nơ-ron tích chập (CNN) để phát hiện lưu lượng mạng liên quan đến fileless malware (mã độc không file).
* Cách tiếp cận với đề tài: Tài liệu tập trung vào chuyển đổi dữ liệu lưu lượng mạng thành định dạng hình ảnh, áp dụng các kỹ thuật học sâu để phát hiện mẫu mã độc trong dữ liệu mạng thực tế

Phương pháp và kỹ thuật thực hiện:

* Sử dụng Cobalt Strike beacon payloads để thu thập và chuyển đổi lưu lượng mạng thành hình ảnh grayscale.
* Thiết kế mô hình CNN với 3 lớp tích chập, normalization và pooling để phân loại hình ảnh lưu lượng mạng thành dữ liệu hợp lệ hoặc mã độc.
* Chuyển đổi dữ liệu payload thành hình ảnh 64x64 pixel sử dụng thư viện NumPy và OpenCV.
* Tăng cường dữ liệu bằng cách xáo trộn các trường header trong HTTP request và xử lý mất cân bằng dữ liệu với các kỹ thuật downsampling và upsampling.

A diagram of a computer system

Description automatically generated

Hình : Kiến trúc chuyển Payload Thành Ảnh và Phát Hiện Bằng CNN

Kết quả nghiên cứu:

* Hệ thống đạt được độ chính cao trong việc phát hiện và phân loại fileless malware.
* Hệ thống hiệu quả trong việc phát hiện mã độc liên quan đến Cobalt Strike, ngay cả khi lưu lượng mạng được ngụy trang dưới các phương thức hợp pháp.

### Tái hiện và phát triển

“PowerDetector: Malicious PowerShell Script Family Classification Based on Multi-Modal Semantic Fusion and Deep Learning” với nhiều ưu điểm nổi bật:

Phương pháp trích xuất đặc trưng: Một hệ thống kết hợp nhiều phương pháp trích xuất đặc trưng từ mã PowerShell, giúp mô hình hiểu sâu hơn về hành vi và cấu trúc mã độc. Các phương pháp này bao gồm:

* Char2Vec: Trích xuất đặc trưng ở cấp độ ký tự, tập trung vào cách các ký tự và chuỗi ký tự xuất hiện trong mã. Phương pháp này rất hiệu quả với các tập lệnh đã bị làm rối (obfuscated).
* Token2Vec: Phân tích và chuyển đổi các token (từ khóa, biến, hàm, ký hiệu) trong mã thành các vector ngữ nghĩa. Phương pháp này giúp nắm bắt mối quan hệ cú pháp và ngữ nghĩa giữa các phần tử trong mã.
* AST2Vec (Abstract Syntax Tree): Phân tích cây cú pháp trừu tượng của tập lệnh để hiểu cách tổ chức cấu trúc của mã và xác định các mô hình cú pháp phức tạp.
* Rela2Vec (Relation Graph): Tạo biểu đồ quan hệ giữa các phần tử của tập lệnh (ví dụ: đường dẫn thực thi, tham số, payload). Phương pháp này giúp mô hình phân tích các mối quan hệ ở cấp cao hơn, từ đó phát hiện các hành vi bất thường

Kiến trúc học sâu tích hợp (Fusion Deep Learning):sử dụng một kiến trúc học sâu tiên tiến để tối ưu hóa khả năng phân loại:

* Transformer: Được sử dụng để nắm bắt mối quan hệ ngữ nghĩa dài hạn và phức tạp trong mã độc.
* CNN-BiLSTM (Convolutional Neural Network - Bidirectional Long Short-Term Memory):
* CNN được áp dụng để xử lý và trích xuất các đặc trưng không gian từ các đặc trưng đầu vào (như từ Char2Vec hoặc Token2Vec).
* BiLSTM giúp mô hình nắm bắt thông tin tuần tự, từ đó hiểu được ngữ cảnh trước và sau trong mã.
* Sự kết hợp giữa CNN và BiLSTM tạo ra một kiến trúc mạnh mẽ, vừa xử lý tốt thông tin không gian vừa nắm bắt thông tin tuần tự.
* Attention Mechanism: Cơ chế attention cho phép hệ thống tập trung vào các phần quan trọng nhất của mã, giúp cải thiện hiệu suất phân loại.

Phân loại mã độc dựa trên nhóm chức năng: Không chỉ phát hiện mã độc mà còn phân loại chúng thành các nhóm chức năng cụ thể, việc phân loại này giúp hệ thống cung cấp cái nhìn chi tiết về mục đích và hành vi của mã độc, hỗ trợ các nhà phân tích bảo mật trong việc đưa ra biện pháp phòng ngừa phù hợp.

* Bypass: Các tập lệnh né tránh cơ chế phát hiện bảo mật.
* Downloader: Các tập lệnh tải xuống mã độc từ một nguồn bên ngoài.
* Injector: Tập lệnh chèn mã độc vào các tiến trình hợp pháp.
* Payload: Chứa các thành phần mã độc thực thi hành vi nguy hiểm.
* TaskExecution: Các tập lệnh thực hiện các tác vụ như lập lịch trình hoặc chạy lệnh hệ thống.

Ứng dụng, tiềm năng thực hiện phát triển:

* Ứng dụng thực tiễn: Hệ thống có thể tích hợp vào các công cụ bảo mật hiện đại để tự động hóa việc phát hiện mã độc trong môi trường thực tế.
* Tiềm năng phát triển: Mở rộng tập dữ liệu để cải thiện khả năng phát hiện các nhóm khó nhận diện, như Bypass cùng với đó là phát triển các kỹ thuật trích xuất đặc trưng mới, đặc biệt tập trung vào mã độc tinh vi và tích hợp thêm các phương pháp như phân tích đồ thị quan hệ sâu để nâng cao hiệu quả phân loại.

#### Dataset

Dữ liệu được thu thập và báo cáo bởi đơn vị Unit42 thuộc Palo Alto Network liên quan đến Encoded Command PowerShell Attacks, báo cáo "Pulling Back the Curtains on EncodedCommand PowerShell Attacks" của Unit 42 phân tích việc lạm dụng tham số -EncodedCommand trong PowerShell để che giấu mã độc. Với 4100 mẫu Powershell Command sử dụng kỹ thuật encode base64 Chứa các tập lệnh thuộc năm nhóm chức năng mã độc: Bypass, Downloader, Injector, Payload, và TaskExecution. Các mẫu mã độc có mức độ phức tạp khác nhau, bao gồm cả các tập lệnh đã được obfuscated (làm rối) để khó bị phát hiện được thu thập từ hai nguồn chính:

Tấn công thực tế (real-world attacks): Tập hợp các mẫu mã độc từ các chiến dịch tấn công fileless malware trong thực tế.

Công cụ tấn công mã nguồn mở: Bao gồm các mẫu được tạo từ các công cụ như Metasploit và Empire, vốn phổ biến trong các cuộc tấn công PowerShell.

#### Data processing

Quá trình xử lý dữ liệu tập trung vào việc chuẩn bị các tập lệnh PowerShell sao cho phù hợp với mô hình học sâu. Điều này bao gồm các bước xử lý tiền dữ liệu và chuẩn hóa đầu vào:

* Làm sạch dữ liệu (Data Cleaning):
* Loại bỏ các phần tử không liên quan hoặc gây nhiễu trong tập lệnh PowerShell như các ký tự dư thừa, dòng trống không cần thiết.
* Xử lý mã hóa đặc biệt hoặc ký tự escape được sử dụng trong các tập lệnh để làm phức tạp quá trình phân tích.
* Tách các phần tử quan trọng:
* Các thành phần chính như lệnh PowerShell, đường dẫn thực thi, tham số (arguments), và các hàm được tách riêng để xử lý sâu hơn.
* Đường dẫn đầy đủ (fully qualified paths) và các biến được ánh xạ thành dạng tiêu chuẩn để giảm tính phức tạp.
* Chuẩn hóa và mã hóa (Normalization and Encoding):
* Các từ khóa, hàm, và biến được mã hóa thành dạng token thống nhất để đảm bảo tính nhất quán.
* Chuẩn hóa các tên biến và hàm để giảm thiểu sự khác biệt giữa các tập lệnh tương tự.
* Xử lý dữ liệu bị làm rối (Obfuscated Data):
* Các tập lệnh PowerShell bị làm rối được chuyển về dạng "clear-text" bằng cách giải mã các chuỗi base64, phân tích cú pháp hàm Invoke-Expression, hoặc các đoạn mã được bao bọc trong các hàm như Compress-String.
* Xây dựng các thuật toán để phát hiện và giải mã tự động các tập lệnh sử dụng Concatenation, String Splitting, hoặc các kỹ thuật tương tự để ẩn mã độc.

#### Data Extraction

Char2Vec (Character-Level Feature Embedding): Phương pháp này trích xuất đặc trưng từ các ký tự trong mã PowerShell. Mỗi ký tự trong mã được ánh xạ vào một vector có ý nghĩa, giúp mô hình học các mẫu tần suất xuất hiện của các chuỗi ký tự trong mã độc. Tận dụng công cụ là Các kỹ thuật embedding trong Word2Vec được sử dụng để tạo ra các vector ký tự.

Token2Vec (Token-Level Feature Embedding): Phương pháp này phân tách mã PowerShell thành các token, bao gồm từ khóa, hàm, biến, và tham số. Mỗi token được chuyển thành một vector ngữ nghĩa, giúp mô hình hiểu được các quan hệ cú pháp và ngữ nghĩa giữa các phần tử trong mã. Tận dụng công cụ là Mô hình Word2Vec có thể được sử dụng để tạo các vector đại diện cho các token, từ đó giúp mô hình phân loại các hành vi của mã độc dựa trên cấu trúc ngữ nghĩa của nó.

AST2Vec (Abstract Syntax Tree Embedding): Trong phương pháp này, mã PowerShell được chuyển thành cây cú pháp trừu tượng (AST), và các đặc trưng cấu trúc của AST được ánh xạ thành các vector. AST giúp mô hình hiểu được cách mã được tổ chức và các quan hệ giữa các phần tử trong mã. Tận dụng công cụ là các phương pháp như Tree-LSTM hoặc Recursive Neural Networks (RNN) được áp dụng để ánh xạ các cấu trúc AST thành vector biểu diễn.

Rela2Vec (Relation Graph Embedding): Phương pháp này xây dựng một biểu đồ quan hệ (relation graph) giữa các phần tử trong mã, chẳng hạn như các đường dẫn thực thi, tham số, và payload. Mỗi quan hệ giữa các phần tử được ánh xạ thành các vector, giúp mô hình hiểu được cách thức các phần tử trong mã tương tác với nhau. Tận dụng công cụ là các phương pháp như Graph Neural Networks (GNN), Node2Vec, được sử dụng để học các mối quan hệ trong biểu đồ và tạo các vector quan hệ.

Sau khi trích xuất được từng loại vector cụ thể, sẽ tiến hành kết hợp kết quả lại thành một vector đa chiều (multi-model vector). Mục đích tạo ra một biểu diễn thống nhất của mã PowerShell, giúp mô hình học được sâu và có thể hiểu và phân tích dữ liệu một cách toàn diện hơn

#### Xây dựng mô hình

Sử dụng CNN-BiLSTM và Transformer:

* CNN (Convolutional Neural Networks) được áp dụng để xử lý các đặc trưng không gian của vector kết hợp, giúp mô hình học các mẫu không gian trong dữ liệu. CNN có khả năng trích xuất các đặc trưng cục bộ, giúp nhận diện các hành vi đặc trưng trong mã độc.
* BiLSTM (Bidirectional Long Short-Term Memory) được sử dụng để nắm bắt thông tin tuần tự trong dữ liệu. Phương pháp này cho phép mô hình xem xét cả ngữ cảnh trước và sau của mỗi phần tử trong tập lệnh, giúp phân tích mối quan hệ thời gian và tuần tự của các lệnh trong mã.
* Transformer: Được sử dụng để nắm bắt các mối quan hệ ngữ nghĩa phức tạp và dài hạn trong mã PowerShell, đặc biệt hữu ích trong các tình huống mà các lệnh có mối quan hệ lâu dài hoặc không liên tục. Transformer giúp mô hình hiểu được các hành vi mã độc phức tạp và có sự tương tác giữa các phần tử mã ở mức độ cao.

Mô hình học sâu kết hợp:

* Các vector đặc trưng kết hợp (multi-modal vectors) được đưa vào mô hình học sâu, bắt đầu từ các lớp tích chập (Convolutional Layers) để học các đặc trưng cục bộ, tiếp theo là các lớp LSTM để học các mối quan hệ tuần tự, và cuối cùng là các lớp Transformer để cải thiện khả năng học ngữ nghĩa phức tạp.

A diagram of a multi-mode light

Description automatically generated

Hình : Kiến trúc mô hình kết hợp

#### Huấn luyện và đánh giá mô hình

Quá trình huấn luyện và đánh giá mô hình được thực hiện trên bộ dữ liệu gồm 3700 mẫu mã độc PowerShell đã được xử lý, được phân chia theo tỷ lệ 80:20 cho tập huấn luyện và tập kiểm tra.

A white rectangular table with black text

Description automatically generated with medium confidence

Bảng : Tập data dữ liệu PowerShell script

Kết quả về các chỉ số của mô hình khi tái hiện: Đánh giá mô hình được thực hiện thông qua các chỉ số chính bao gồm độ chính xác (accuracy), F1-score, độ nhạy (recall) và độ đặc hiệu (specificity). Kết quả thử nghiệm cho thấy mô hình đạt độ chính xác trung bình 92,7%.

|  |  |
| --- | --- |
| **Metric** | **Value** |
| Loss | 0.24614659 |
| Accuracy | 0.927984002 |
| Precision | 0.92014193 |
| Recall | 0.927983979 |
| F1 Score | 0.924033537 |

Bảng : Chỉ số hiệu suất của mô hình

A diagram of a diagram

Description automatically generated with medium confidence

Hình : Ma trận nhầm lẫn

# KẾT QUẢ VÀ ĐÁNH GIÁ

## Kết quả đạt được

* Về mặt quy trình làm việc
* Đã nắm vững và áp dụng thành thạo phương pháp Kanban Work Management trong quản lý công việc, giúp tổ chức và theo dõi tiến độ các nhiệm vụ một cách hiệu quả.
* Thành thạo quy trình Git Workflow, có khả năng làm việc với các nhánh (branch), quản lý phiên bản và hợp nhất mã nguồn một cách chuyên nghiệp.
* Phát triển kỹ năng làm việc nhóm thông qua việc sử dụng các công cụ quản lý dự án và kiểm soát phiên bản.
* Về mặt nghiên cứu và phát triển
* Đã hoàn thành việc nghiên cứu và tái hiện thành công mô hình PowerDetector, một hệ thống phân loại mã độc PowerShell tiên tiến với độ chính xác trung bình đạt 92.7%.
* Nắm được các kỹ thuật trích xuất đặc trưng đa phương thức và dựng và huấn luyện mô hình học sâu kết hợp CNN-BiLSTM và Transformer, cho phép phân loại chính xác các họ mã độc PowerShell thành các nhóm chức năng khác nhau.
* Kiến thức và kỹ năng đạt được
* Kiến thức chuyên môn:
* Hiểu sâu về các loại mã độc PowerShell và cách thức hoạt động của chúng
* Nắm vững các kỹ thuật học máy và học sâu trong lĩnh vực an ninh mạng
* Làm quen với quy trình nghiên cứu khoa học và phát triển hệ thống thực tế
* Hiểu rõ về cách thức xử lý và phân tích dữ liệu lớn trong lĩnh vực an ninh mạng
* Kỹ năng mềm
* Phát triển khả năng nghiên cứu và tổng hợp tài liệu khoa học
* Nâng cao kỹ năng giải quyết vấn đề và tư duy phân tích
* Cải thiện kỹ năng làm việc độc lập và quản lý thời gian
* Tăng cường khả năng thuyết trình và viết báo cáo kỹ thuật

## Đánh giá và nhận xét

* Điểm mạnh
* Có khả năng tiếp thu nhanh và áp dụng hiệu quả các công nghệ mới
* Thể hiện tính chủ động và sáng tạo trong quá trình nghiên cứu
* Đạt được các mục tiêu đề ra trong thời gian thực tập
* Có khả năng làm việc độc lập và theo nhóm tốt
* Điểm cần cải thiện
* Cần tăng cường thêm kiến thức về các kỹ thuật liên quan đến chuyên môn
* Cần phát triển thêm khả năng phân tích và xử lý các trường hợp khác nhau
* Cần mở rộng hiểu biết về các công nghệ và framework mới trong lĩnh vực an ninh mạng

TÀI LIỆU THAM KHẢO

|  |  |
| --- | --- |
| [1] | Y. C. J. L. J. M. Muhammad Adil Inam, "SoK: History is a vast early warning system: Auditing the provenance of system intrusions.," 2023. |
| [2] | G. P. D. Z. Y. G. C. L. Xiuzhang Yang, "PowerDetector: Malicious PowerShell Script Family Classification Based on Multi-Modal Semantic Fusion and Deep Learning," 2023. |
| [3] | A. N. S. K. M. W. K. R. Y. F. Fikirte Ayalke Demmese, "Machine learning based fileless malware traffic classification using image visualization," 2023. |