**ĐẠI HỌC QUỐC GIA TP.HCM**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**MÔN HỌC: AN TOÀN MẠNG KHÔNG DÂY VÀ DI ĐỘNG**

**BÁO CÁO ĐỒ ÁN CUỐI KÌ**

***Mitigating adversarial evasion attacks of ransomware using ensemble learning***

**Giảng viên hướng dẫn: Phan Thế Duy  
 NT230.O21.ATCL**

**Sinh viên thực hiện :**

**Nhóm G1**

**21522620 – Hồ Ngọc Thiện**

**21522483 - Chu Nguyễn Hoàng Phương**

**TP.HCM, Ngày 28 tháng 5 năm 2024**

MỤC LỤC

[I. Tổng quan bài báo 2](#_Toc167825574)

[II. Sơ lược về phương pháp phân tích tổng hợp (ensemble analyze) 3](#_Toc167825575)

[1. Phân tích tĩnh 3](#_Toc167825576)

[2. Phân tích động 4](#_Toc167825577)

[III. Demo 5](#_Toc167825578)

[1. Phân tích tĩnh 5](#_Toc167825579)

[2. Phân tích động 5](#_Toc167825580)

[3. Vận dụng mô hình tổ hợp 5](#_Toc167825581)

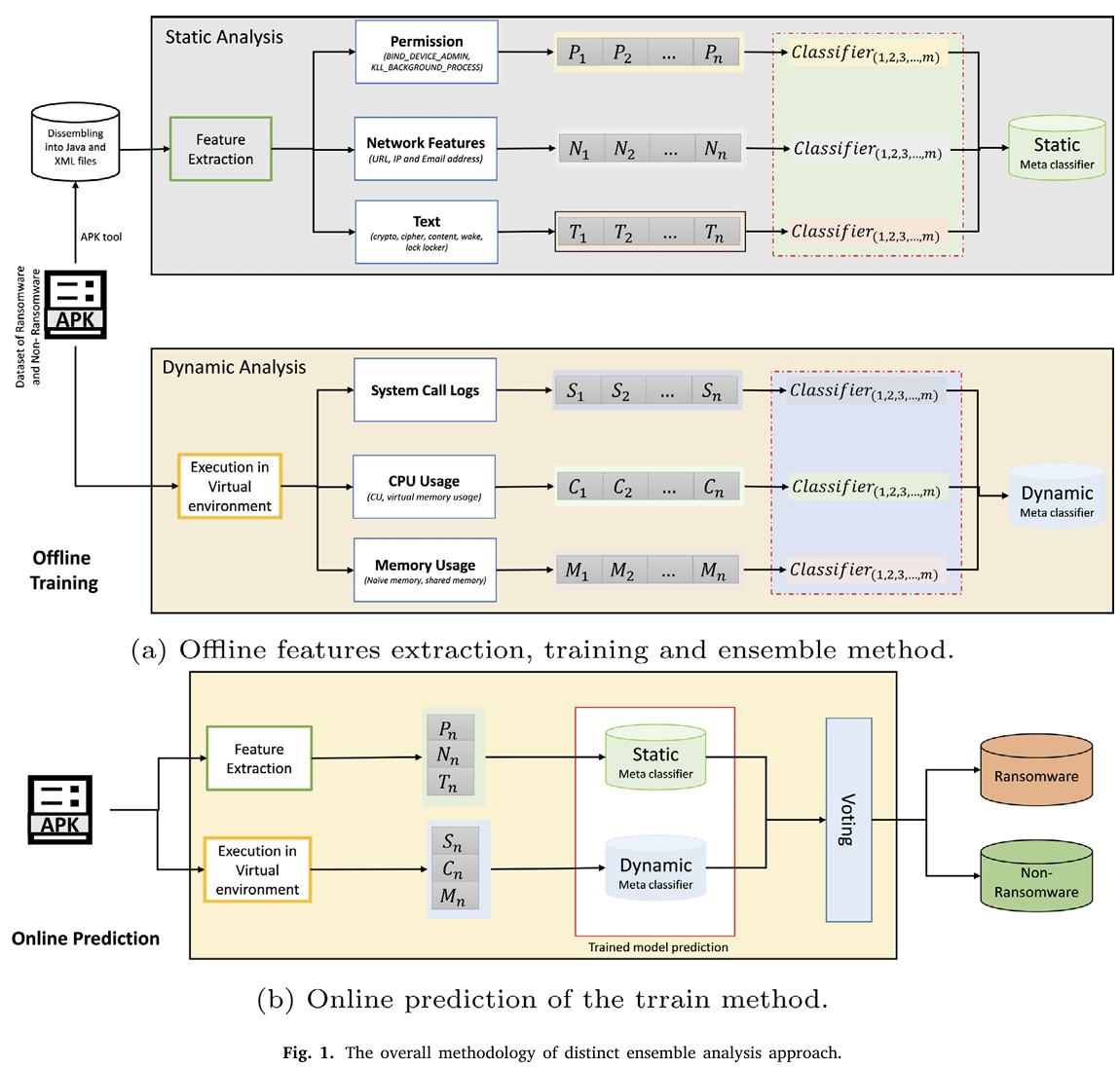
**BÁO CÁO CHI TIẾT**

# I. Tổng quan bài báo

Ransomware tiếp tục gây ra mối đe dọa đáng kể đối với an ninh mạng bằng cách tống tiền người dùng bằng cách khóa thiết bị và dữ liệu cá nhân của họ. Những kẻ tấn công buộc phải trả tiền chuộc để khôi phục quyền truy cập vào các tệp cá nhân. Do sự tương đồng về cấu trúc, việc phát hiện ransomware và các ứng dụng lành tính trở nên dễ bị tấn công trốn tránh (evasion attack). Học tập kết hợp (ensemble learning) có thể cung cấp các biện pháp đối phó, trong khi những kẻ tấn công có thể sử dụng cùng một kỹ thuật để cải thiện hiệu quả của các cuộc tấn công tương tự của chúng. Điều này thúc đẩy việc điều tra xem liệu những phương pháp kết hợp riêng biệt có thể đạt được hiệu suất tốt hơn khi kết hợp với phương pháp dựa trên bỏ phiếu hay không. Nghiên cứu được đề cập trong bài báo này đề xuất một cách tiếp cận lai kiểm tra quyền, văn bản và các tính năng dựa trên mạng cả tĩnh và động bằng cách giám sát việc sử dụng bộ nhớ, nhật ký cuộc gọi hệ thống và sử dụng CPU. Các máy phân tích tập hợp trên các tính năng tĩnh và động được trích xuất từ các ứng dụng phần mềm độc hại Android (ransomware và không ransomware) sau đó được đào tạo theo mô hình được thiết kế. Kết quả thử nghiệm của bài báo cho thấy kỹ thuật phát hiện và phân loại tập thể được đề xuất có thể phân loại hành vi ransomware tĩnh và động không xác định để giảm thiểu các cuộc tấn công trốn tránh đối thủ (adversarial evasion attack).

Nghiên cứu này tập trung vào việc giảm thiểu các cuộc tấn công trốn tránh trong phát hiện ransomware Android, chẳng hạn như xáo trộn mã (obfuscation) được sử dụng để tránh việc bị phát hiện là phần mềm độc hại hoặc ransomware. Hầu hết các kỹ thuật ransomware hiện tại không sửa đổi vectơ tính năng đầu vào để phân tích (khi một khía cạnh được sử dụng để xáo trộn, nó sẽ thay đổi toàn bộ vectơ tính năng, khiến trình phân loại được đào tạo phân loại sai ransomware. Do đó, nghiên cứu này đề xuất một cơ chế phân tích dựa trên tổng hợp để phát hiện ransomware Android và giảm thiểu các cuộc tấn công trốn tránh. Các đặc điểm Android được sử dụng để phân tích ở đây không dễ sửa đổi và sử dụng cho các nỗ lực trốn tránh. Dựa trên các đặc tính này, tác giả đề xuất một kỹ thuật kết hợp hiệu quả của cả tính năng tĩnh (ví dụ: quyền, văn bản, tính năng dựa trên mạng, v.v.) và các tính năng động (như nhật ký cuộc gọi hệ thống, CPU và sử dụng bộ nhớ) để phát hiện phần mềm tống tiền Android bằng cách sử dụng mô hình máy học tổng hợp.

# II. Sơ lược về phương pháp phân tích tổng hợp (ensemble analyze)

Mô hình của phương pháp phân tích tổng hợp sẽ bao gồm 2 phần: offline training và online prediction. Quá trình offline training bao gồm các công đoạn như trích xuất đặc trưng, chọn lựa đặc trưng, huấn luyện và kiểm thử hiệu suất của mô hình. Còn trong quá trình online prediction, bộ phân loại đã được huấn luyện trước đó sẽ được sử dụng để phân loại các ứng dụng độc hại và lành tính. Việc trích xuất đặc trưng dựa trên phân tích tĩnh và động, ở quá trình offline training, phân tích tĩnh sẽ được sử dụng để trích xuất quyền hạn, mạng và các văn bản chứa trong tập tin APK. Với phân tích động, ứng dụng sẽ được chạy trong môi trường ảo được kiểm soát, sau đó các đặc trưng sẽ được trích xuất. Các thông tin của hệ thống như system call, tài nguyên CPU, memory được ghi lại, sau đó trích xuất và tiến hành phân tích.

## 1. Phân tích tĩnh

Dựa theo yêu cầu bài báo, tập dữ liệu được phân tích sẽ bao gồm các file APK, trong đó 50

############################################################################

## 2. Phân tích động

Phân tích động bao gồm việc chạy ứng dụng Android trên môi trường ảo để kiểm tra các hành vi của ứng dụng theo thời gian thực. Phân tích động được sử dụng để phát hiện các hành vi độc hại của ứng dụng mà ta không thể phát hiện chúng dựa vào phân tích tĩnh. Như có đề cập ở trên, mục đích của việc phân tích động là để trích xuất ra được các đặc trưng “động” (Lượng CPU sử dụng, thống kê system call, lượng memory sử dụng,…) , những thông tin này rất hữu ích cho việc phát hiện ransomware. Do đó, các dấu vết thực thi có chứa dữ liệu này phải được thu thập bằng cách chạy ứng dụng trong môi trường được kiểm soát. Những dấu vết này được ghi lại thủ công bằng cách chạy từng ứng dụng một trong 10 phút trong trình giả lập Android. Tuy nhiên, một số dấu vết ngắn hơn vì trình giả lập có những điểm yếu nhỏ. Tuy nhiên, thời gian thực hiện dài hơn mang lại cho chúng ta kết quả có ý nghĩa hơn. Trình giả lập Android Genymotion được chọn để phân tích động các ứng dụng phần mềm độc hại Android. Công cụ Genymotion được sử dụng vì nó là phần mềm mã nguồn mở và hỗ trợ Android Studio. Lý do sử dụng phần mềm giả lập Android thay vì thiết bị gốc là môi trường mô phỏng cung cấp nhiều dung lượng hơn để chạy một số lượng lớn các chương trình phần mềm độc hại trong một thời gian hợp lý. Phân tích động trên thiết bị miễn nhiễm với các kỹ thuật bỏ qua trình giả lập. Tuy nhiên, trong trường hợp ransomware Android, thiết bị vật lý không thể được đặt lại về trạng thái sạch. Ngược lại, trình giả lập có thể được khởi tạo lại sau khi phân tích từng ứng dụng. Tuy nhiên, việc chống mô phỏng có thể có một số tác động đến việc trích xuất tính năng động được thực hiện trên Genymotion. Do đó, trong mô hình đề xuất của chúng tôi, khi sử dụng các tính năng động (sử dụng bộ nhớ và CPU), chúng tôi giả định rằng ngay cả khi các ứng dụng được đóng gói lại, làm xáo trộn hoặc được trang bị các kỹ thuật để tránh lỗi phát hiện, chúng vẫn sẽ hiển thị các dấu vết hành vi tương tự trong quá trình thực thi. Một trình phân tích tổng hợp được đào tạo về các đặc điểm này có thể phân biệt đầy đủ các ứng dụng ransomware Android với các ứng dụng phần mềm độc hại khác. Thiết bị ảo được khởi tạo lại mỗi lần trước khi một ứng dụng độc hại mới được thực thi để tránh sự can thiệp từ các ứng dụng đã thực thi trước đó, chẳng hạn như thay đổi cài đặt, thực thi các quy trình nền, thay đổi liên quan đến cấu hình hệ điều hành, v.v. Android Debug Bridge (ADB) được sử dụng để giám sát việc sử dụng bộ nhớ và CPU của các ứng dụng. ADB là một công cụ dòng lệnh cho phép PC giao tiếp với phiên bản giả lập hoặc thiết bị Android. Strace (một công cụ theo dõi cuộc gọi hệ thống) được sử dụng để thu thập các cuộc gọi hệ thống từ các ứng dụng. Để trích xuất tính năng động, các bước sau được thực hiện cho từng ứng dụng như được đề cập trong đoạn mã giả dưới đây.

A screenshot of a computer

Description automatically generated

# III. Demo

## 1. Phân tích tĩnh

## 2. Phân tích động

## 3. Vận dụng mô hình tổ hợp