

# TRIỂN KHAI CÔNG NGHỆ Đám Mây KẾT HỢP CÁC PHƯƠNG PHÁP ĐỂ PHÂN TÍCH PHẦN MỀM MÃ ĐỘC TRONG ỨNG DỤNG ANDROID

Hoàng Trường Thịnh - 230201031

<sup>1</sup> Trường Đại học Công nghệ Thông Tin

<sup>2</sup> University of Information Technology  
HCMC, Vietnam

<sup>3</sup> Information Security

## What ?

Chúng tôi thiết kế và triển khai nghiên cứu ứng dụng hệ thống đám mây (cloud system) để phân tích phần mềm mã độc trong ứng dụng Android:

- Phân tích tĩnh: Kiểm tra cấu trúc APK, mã nguồn và các thông tin khác không cần thực thi.
- Phân tích động: Chạy các ứng dụng trong môi trường ảo (Sanbox) để quan sát hành vi thực tế, xác định hoạt động đáng ngờ.

## Why ?

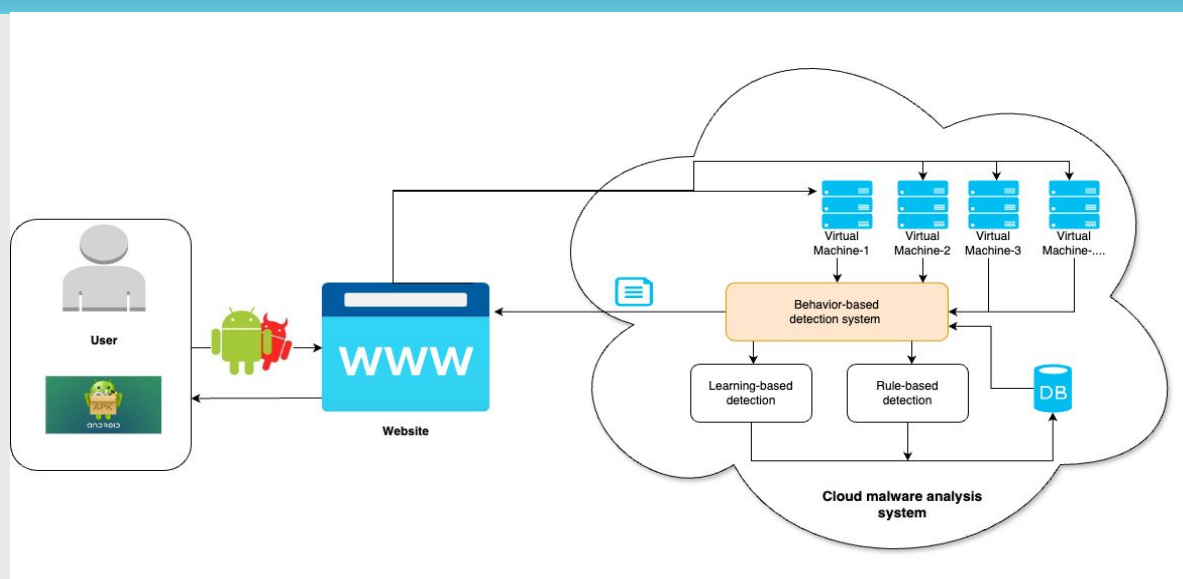
- Gia tăng mã độc trên Android: Với sự phát triển nhanh chóng của hệ điều hành Android, mã độc cũng tăng theo, gây ra các vấn đề bảo mật nghiêm trọng.
- Giới hạn của phương pháp truyền thống: Phương pháp phân tích mã độc truyền thống đòi hỏi tài nguyên tính toán lớn và tốn thời gian, không phù hợp với tốc độ phát triển của các mối đe dọa
- Ưu điểm của công nghệ đám mây: Cung cấp khả năng mở rộng và sức mạnh tính toán vượt trội, giúp nâng cao độ chính xác và tốc độ phân tích mã độc.

## Overview

Trong thời đại công nghệ phát triển nhanh chóng, hệ điều hành Android phổ biến mang lại nhiều tiện ích nhưng tạo ra nhiều lỗ hổng bảo mật nghiêm trọng cho phần mềm độc hại xâm nhập.

Phương pháp phân tích truyền thống dù hiệu quả nhưng bị thách thức bởi sự tinh vi và số lượng mã độc ngày càng tăng.

Công nghệ đám mây trở nên chiếm ưu thế với những bất lợi của các phương pháp truyền thống, giải quyết được các nhu cầu sử dụng dữ liệu và tài nguyên phân bổ sử dụng để hoạt động phân tích mã độc.



## Description

### 1. Nội dung

- Thiết kế và triển khai một hệ thống dựa trên nền tảng đám mây để phát hiện và phân tích mã độc trong các ứng dụng Android.
- Kết hợp các công nghệ dịch vụ đám mây để tối ưu hoá việc xử lý và lưu trữ dữ liệu lớn.
- Phát triển các mô-đun phân tích tĩnh và động để phân tích các tệp APK.
- Tối ưu hoá hệ thống phân tích để có thể đạt được độ chính xác cao hơn.
- Đảm bảo hệ thống có khả năng cập nhật và mở rộng với các thuật toán mới nhất.
- Áp dụng trí tuệ nhân tạo và các thuật toán máy học để tăng cường khả năng phát triển mã độc.

### 2. Phương pháp

- Có thể đánh giá và lựa chọn nền tảng đám mây như AWS, Google cloud, Microsoft Azure đánh giá bằng hiệu suất, bảo mật và chi phí.
- Xây dựng kiến trúc bao gồm các thành phần chính như bộ thu thập dữ liệu, mô-đun phân tích tĩnh, mô-đun phân tích động và hệ thống quản lý kết quả

- Xây dựng mô-đun phân tích tĩnh để kiểm tra các thuộc tính của tệp APK như cấu trúc tệp, mã nguồn và siêu dữ liệu.
- Xây dựng mô-đun phân tích động để chạy các tệp APK trong môi trường sandbox và giám sát hành vi của chúng.
- Tối ưu hoá lưu trữ và truy xuất trên nền tảng đám mây để giảm thời gian xử lý.
- Phát triển các mô hình máy học dựa trên các tập dữ liệu chứa mã độc và không chứa mã độc để cải thiện khả năng phân loại (URL dataset: <https://www.unb.ca/cic/datasets/andmal2017.html>)
- Xem xét các giải pháp mã nguồn mở và thương mại để chọn ra các công cụ tối ưu.
- Thiết lập hệ thống với khả năng cập nhật và tích hợp thêm các công cụ và thuật toán khi cần thiết.
- Đánh giá so sánh các kết quả trên các thuật toán và công cụ khác nhau để có được kết quả chính xác nhất.

### 3. Kết quả dự kiến

- Hệ thống sẽ cung cấp giao diện người dùng thân thiện cho phép người dùng có thể tải lên các tệp APK cần phân tích. Sau đó, hệ thống tiến hành phân tích và thông báo kết quả phát hiện mã độc.
- Khả năng cảnh báo thời gian thực, người dùng sẽ nhận được thông báo ngay lập tức qua email hoặc giao diện hệ thống khi phát hiện mã độc. Báo cáo chi tiết sẽ được tự động gửi tới người dùng.
- Hệ thống có thể được cập nhật và mở rộng một cách dễ dàng mà không cần thay đổi cấu trúc chính. Các thuật toán và công cụ mới có thể thêm vào để cải thiện hiệu quả phát hiện mã độc.
- Hệ thống có khả năng xử lý và phân tích ít nhất 100 tệp APK mỗi giờ, với độ chính xác phát hiện mã độc đạt 95%.