THÔNG TIN CHUNG CỦA BÁO CÁO

- Link YouTube video của báo cáo (tối đa 5 phút): **https://youtu.be/6ZriumP1MZs**
- Link slides (dạng .pdf đặt trên Github): https://github.com/17130225Hoangthinh/CS2205.CH181/blob/main/DEPLOYING_CLOUD_TECHNOLOGY_IN_COMBINATION_WITH_METHODS_FOR_ANALYZING_MALWARE_IN_ANDROID_APPLICATIONS.pdf

• Họ và Tên: Hoàng Trường Thịnh • MSSV: 230201031 	• Lớp: CS2205.CH181 • Tự đánh giá (điểm tổng kết môn): 8/10 • Số buổi vắng: 0 • Số câu hỏi QT cá nhân: • Link Github: https://github.com/17130225Hoangthinh/CS2205.CH181

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

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

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

DEPLOYING CLOUD TECHNOLOGY IN COMBINATION WITH METHODS FOR ANALYZING MALWARE IN ANDROID APPLICATIONS

TÓM TẮT *(Tối đa 400 từ)*

Ứng dụng sẽ tập trung vào việc phát triển những giải pháp phân tích dựa trên công nghệ đám mây để phát hiện, phân tích và đối phó với mã độc trong các ứng dụng Android. Với sự gia tăng các mối đe dọa an ninh mạng trong những năm gần đây, đặc biệt là trên các thiết bị di động, việc tìm ra một phương pháp hiệu quả để xác định và ngăn chặn mã độc trở nên cần thiết. Phương pháp tiếp cận này sử dụng lợi thế của việc xử lý và lưu trữ dữ liệu trên đám mây để quản lý và phân tích một lượng lớn ứng dụng Android mà không cần phụ thuộc vào nguồn lực phần cứng cục bộ. Qua đó, giải pháp này cho phép phân tích đồng thời nhiều mẫu mã độc, từ đó cải thiện độ chính xác và tốc độ của quá trình phân tích.

Quá trình phân tích mã độc trong đồ án được chia thành hai phần chính: phân tích tĩnh và phân tích động. Phân tích tĩnh bao gồm việc kiểm tra các tệp APK mà không thực thi chúng, nhằm phát hiện các dấu hiệu đáng ngờ và tiềm ẩn nguy hiểm thông qua việc phân tích cấu trúc tệp, mã nguồn và các thông tin khác. Mục tiêu là để lọc ra những ứng dụng có khả năng chứa mã độc ngay từ bước đầu tiên. Tiếp theo, phân tích động được thực hiện bằng cách chạy các ứng dụng trong một môi trường ảo an toàn (sandbox) để quan sát hành vi thực tế của chúng [1]. Quá trình này giúp xác định các hoạt động như kết nối mạng, thay đổi hệ thống tệp, hoặc cố gắng lây nhiễm các thiết bị khác. Từ đó, có thể phân tích sâu hơn về mục đích và tác động tiềm tàng của mã độc.

GIỚI THIỆU *(Tối đa 1 trang A4)*

Trong thời đại công nghệ phát triển nhanh chóng, sự phổ biến của hệ điều hành Android mang lại nhiều tiến bộ nhưng cũng tạo ra lỗ hổng của bảo mật nghiêm trọng qua phần mềm độc hại, gây ra truy cập dữ liệu trái phép, vi phạm quyền riêng tư và tổn thất tài chính. Phương pháp phân tích phần mềm độc hại truyền thống, dù hiệu quả ở mức độ nhất định, đang bị thách thức bởi sự tinh vi và số lượng phần mềm độc hại ngày càng tăng. Chúng đòi hỏi tài nguyên tính toán lớn và tốn thời gian, không phù hợp với tốc độ phát triển mối đe doạ.

Công nghệ đám mây (cloud) là giải pháp hiệu quả cho các hạn chế này. Với khả năng mở rộng và sức mạnh tính toán của cloud, việc phân tích phần mềm độc hạy trở nên năng động và hiệu quả hơn. Luận án này nghiên cứu việc sử dụng công nghệ cloud để phân tích phần mềm độc hại trên Android, nâng cao độ chính xác và tốc độ phân tích. Nghiên cứu này chứng minh các giải pháp đám mây giúp phát hiện và giảm thiểu phần mềm độc hại hiệu quả hơn phương pháp truyền thống. Tài nguyên đám mây cho phép sử lý dữ liệu lớn, thích ứng nhanh với kỹ thuật mới và bảo vệ thời gian thực cho người dùng.

Đầu vào của ứng dụng:

- Ứng dụng Android cần phân tích: (File APK nghi ngờ chứa phần mềm độc hại)
- Dữ liệu về phần mềm độc hại: Dataset, Signatures, behaviors và dữ liệu khác.
- Thông tin môi trường cloud: cấu hình và tài nguyên của nền tảng cloud sử dụng
- Các thuật toán và công cụ phân tích: Các thuật toán phân tích tĩnh và động, công cụ phát hiện phần mềm

Đầu ra của ứng dụng:

- Báo cáo phân tích nhanh chóng mã độc có tồn tại trong mẫu APK

MỤC TIÊU

(Viết trong vòng 3 mục tiêu, lưu ý về tính khả thi và có thể đánh giá được)

- Tạo ra một hệ thống sử dụng công nghệ đám mây để phát hiện và phân tích mã

độc trong các ứng dụng Android thông qua hai phương pháp phân tích chính: phân tích tĩnh và phân tích động.

- Tận dụng lợi thế của công nghệ đám mây để cải thiện độ chính xác và tốc độ của quá trình phân tích mã độc so với các phương pháp truyền thống.
- Sử dụng các thuật toán và công cụ phân tích hiện đại, bao gồm cả phân tích tĩnh và phân tích động, để phát hiện và phân tích mã độc trong các ứng dụng Android.

NỘI DUNG VÀ PHƯƠNG PHÁP

(Viết nội dung và phương pháp thực hiện để đạt được các mục tiêu đã nêu)

Nội dung thực hiện:

- 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[2].
- 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.

Phương pháp thực hiện:

- Có thể đánh giá và lựa chọn nền tảng đám mây như AWS, Google cloud, Microsoft Azure đánh giá bằng hiệu xuất, 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 ảo sandbox và giám sát hành vi của chúng[3].
- Tối ưu hoá lưu trữ và truy xuất trên nền tảng đám mây để giảm thời gian xử lý[4].
- 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 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.

KẾT QUẢ MONG ĐỢI

(Viết kết quả phù hợp với mục tiêu đặt ra, trên cơ sở nội dung nghiên cứu ở trên)

- 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 thông 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%.

TÀI LIỆU THAM KHẢO *(Định dạng DBLP)*

[1] Cui, S. J., Sun, G. X., Bin, S., & Zhou, X. C. (2016). An android malware detection system based on cloud computing. *chemical engineering transactions*, *51*, 691-696.

[2] Aslan, Ö., Ozkan-Okay, M., & Gupta, D. (2021). Intelligent behavior-based malware detection system on cloud computing environment. *IEEE Access*, *9*, 83252-83271.

[3] Kumar, R., Sethi, K., Prajapati, N., Rout, R. R., & Bera, P. (2020, July). Machine learning based malware detection in cloud environment using clustering approach. In *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-7). IEEE.

[4] Walls, J., & Choo, K. K. R. (2015). A review of free cloud-based anti-malware apps for android. *2015 IEEE Trustcom/BigDataSE/ISPA*, *1*, 1053-1058.

[5] Tian, D., Zhao, R., Ma, R., Jia, X., Shen, Q., Hu, C., & Liu, W. (2022). MDCD: A malware detection approach in cloud using deep learning. *Transactions on Emerging Telecommunications Technologies*, *33*(11), e4584.