TĂNG CƯỜNG PHÁT HIỆN MÃ ĐỘC ANDROID VỚI BẢO ĐẢM QUYỀN RIÊNG TƯ THÔNG QUA FEDERATED LEARNING

Lê Vũ Tuấn Anh - 230201061

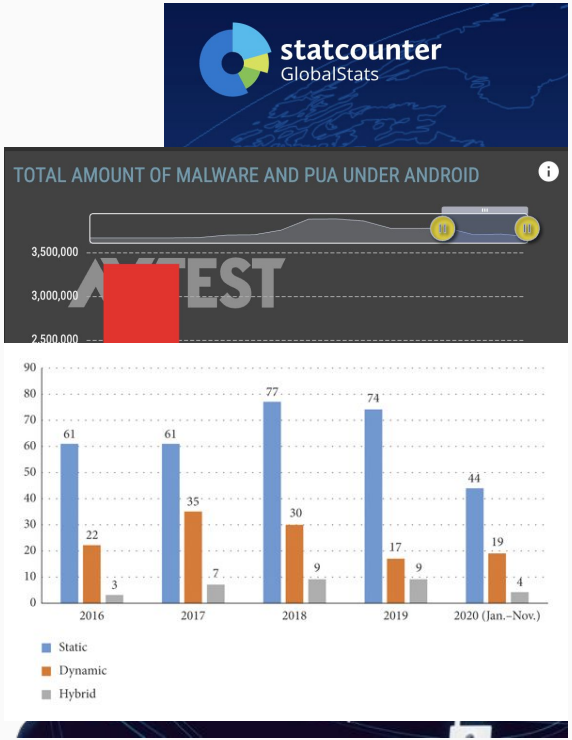
Tóm tắt

- Lớp: CS2205.CH181
- Link Github: https://github.com/anhlvt-uit/CS2205.APR2023
- Link YouTube video: https://youtu.be/eHBN9GvN2Jg
- Họ và Tên: Lê Vũ Tuấn Anh
- MSHV: 230201061

Giới thiệu

- Android là hệ điều hành di động phổ biến nhất thế giới.
- Mã độc Android đánh cắp thông tin cá nhân và gây hậu quả nghiêm trọng. Số lượng nhiều và biến đổi liên tục.
- Nhiều nghiên cứu phát hiện mã độc đã được thực hiện nhưng vẫn còn có những hạn chế.

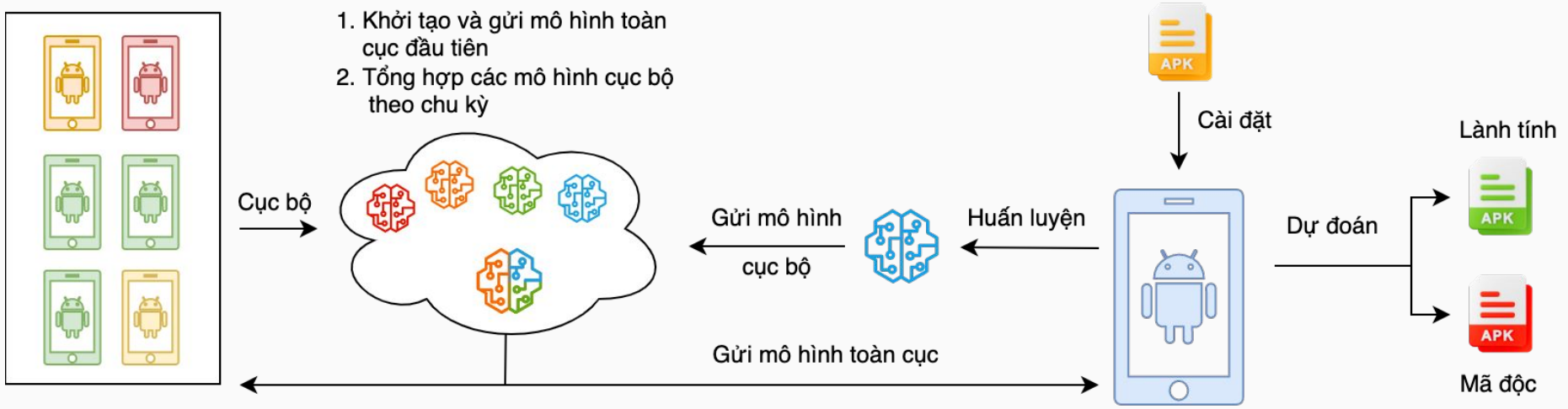
=> Đặt ra bài toán tăng cường khả năng phát hiện mã độc, tận dụng ưu điểm huấn luyện phi tập trung nhưng phải bảo vệ thông tin riêng tư của người dùng: danh sách apps, tài khoản, thói quen…



Khảo sát phương pháp trích xuất đặc trưng [3]

Giới thiệu

- Federated Learning: huấn luyện mô hình phi tập trung, đảm bảo quyền riêng tư dữ liệu cục bộ.
- Semi-supervised Learning: huấn luyện với số lượng lớn mẫu không gắn nhãn.



Mục tiêu

- Thu thập Dataset gồm các file APK mã độc và lành tính Android với số lượng khoảng **25,000 mỗi loại**, để trích xuất đặc trưng cho việc huấn luyện mô hình.
- Thiết kế tối ưu thuật toán huấn luyện cục bộ và tổng hợp tham số mô hình toàn cục cho hệ thống Federated Learning, đảm bảo độ chính xác tăng dần qua nhiều chu kỳ cho đến khi tiến đến mức **trên 93%**, thời gian huấn luyện **dưới 10 giây**.
- Triển khai kiểm thử hệ thống trên **1 máy chủ** vật lý và **40 máy ảo** Android với nhiều kịch bản để thu thập số liệu thống kê, đánh giá hiệu suất và tính khả thi.

Nội dung và Phương pháp

- Thu thập Dataset:
  - Tải xuống các Dataset mã độc Android: CIC, VirusTotal…
  - Tải thêm các mẫu lành tính từ nhóm ứng dụng được tải nhiều trên Google Play Store.
  - Hash MD5 các mẫu để loại bỏ các mẫu trùng lặp.

Nội dung và Phương pháp

- Thiết kế thuật toán huấn luyện cục bộ:
  - Tìm hiểu các thuật toán máy học phân loại nhị phân có thể huấn luyện dạng Semi-supervised Learning: SVM, LR…
  - Dự đoán rằng kết quả các mô hình trên có thể tổng hợp tạo ra dự đoán chính xác hơn thông qua trọng số.
  - Tìm hiểu Safe Semi-supervised Learning.
- Thiết kế thuật toán tổng hợp mô hình toàn cục:
  - Tìm hiểu các thuật toán: FedAvg, FedDyn.
  - Phân tích kết quả các nghiên cứu về mối liên hệ giữa số lượng thiết bị tham gia, thời gian chu kỳ lên chất lượng mô hình.

Nội dung và Phương pháp

- Triển khai kiểm thử và thống kê kết quả:
  - Tìm hiểu các thư viện hỗ trợ trích xuất đặc trưng.
  - Tìm hiểu các thư viện máy học và framework Federated Learning do TensorFlow cung cấp.
  - Tìm hiểu triển khai hệ thống máy ảo Android trên Docker.
  - Lập bảng thống kê các giá trị độ đo F1, accuracy… theo các kịch bản thử nghiệm khác nhau.
  - Phân tích kết quả thu được.

Kết quả dự kiến

- Tổng hợp được đủ số lượng mẫu mã độc và lành tính. Các mẫu này đều có thể trích xuất đặc trưng mà không gặp lỗi.
- Xây dựng thuật toán huấn luyện cục bộ có thể chạy được trên thiết bị Android mà không chiếm dụng nhiều tài nguyên, thuật toán tổng hợp mô hình toàn cục ổn định và cải thiện liên tục. Hệ thống đạt được các mức đo lường đã đề ra.
- Triển khai được hệ thống thông qua các công cụ và nền tảng dự kiến. Từ kết quả thử nghiệm, chứng minh được hiệu quả của mô hình huấn luyện đề xuất và ghi nhận những vấn đề mới cần giải quyết trong tương lai.

Tài liệu tham khảo

- [1]. Mobile Operating System Market Share Worldwide. In: StatCounter Global Stats. https://gs.statcounter.com/os-market-share/mobile/worldwide. Accessed 30 May 2024
- [2]. AV-TEST - The Independent IT-Security Institute Atlas - Malware & pua. In: AV. https://portal.av-atlas.org/malware/statistics. Accessed 30 May 2024
- [3]. Wu Q, Zhu X, Liu B (2021) A survey of Android Malware Static Detection Technology based on machine learning. Mobile Information Systems 2021:1–18
- [4]. Zaabi AA, Mouheb D (2020) Android malware detection using static features and machine learning. 2020 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)
- [5]. Ma J, Yu G, Xiong W, Zhu X (2023) Safe semi-supervised learning for pattern classification. Engineering Applications of Artificial Intelligence 121:106021
- [6]. Gálvez, R., Moonsamy, V., & Díaz, C. (2020). Less is More: A privacy-respecting Android malware classifier using federated learning. Proceedings on Privacy Enhancing Technologies, 2021, 96 - 116
- [7]. Fang W, He J, Li W, et al (2023) Comprehensive Android malware detection based on Federated Learning Architecture. IEEE Transactions on Information Forensics and Security 18:3977–3990
- [8]. Drainakis G, Pantazopoulos P, Katsaros KV, et al (2023) From centralized to Federated Learning: Exploring Performance and end-to-end resource consumption. Computer Networks 225:109657