ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

BÁO CÁO GIỮA KỲ

MÁY HỌC TRONG BẢO MẬT MẠNG VÀ HỆ THỐNG

TÊN ĐỀ TÀI: MAGIC: DETECTING ADVANCED PERSISTENT THREATS VIA MASKED GRAPH REPRESENTATION LEARNING

GIẢNG VIÊN HƯỚNG DẪN
TS. LÊ KIM HÙNG

HỌC VIÊN THỰC HIỆN

230202002 - TÔ THỊ MỸ ÂU
220202022 - NGUYỄN HỒNG SƠN
230202006 - NGÔ THÁI HƯNG

TP. HỒ CHÍ MINH, NĂM 2024

Mục lục

1 Tóm tắt bài báo

MAGIC là một phương pháp học có giám sát để phát hiện các cuộc tấn công APTs bằng cách sử dụng phương pháp học masked graph representation. Nó cải thiện một số phương pháp trước đó bằng cách trích xuất efficient deep feature và structure abstraction on provenance graphs.

MAGIC có khả năng phát hiện multi-granularity, handle concept drift with a model adaption mechanism và có khả năng mở rộng. Nó sử dụng phương pháp phát hiện ngoại lệ để xác định hành vi bất thường.

MAGIC được đánh giá dựa trên 3 dataset, cho kết quả phát hiện với chỉ số precision và recall cao, chi phí thấp và nhanh hơn nhiều so với các phương pháp hiện tại

Các kết quả của bài báo là một công trình được phát hành dưới dạng open-source, với tiềm năng phát triển nhiều trong lương lai.

2 Advance Persistent Threats

Advance Persistent Threats là một quá trình tấn công mạng tinh vi trong thời gian dài, kẻ tấn công không bị phát hiện trong một thời gian dài và do đó, tiếp cận, đánh cắp, phá hoại dữ liệu quan trọng trong hệ thống. Nó sử dụng kết hợp nhiều công cụ và kỹ thuật phức tạp để xâm nhập và duy trì quyền truy cập tới hệ thống. Kẻ tấn công có mục tiêu dài hạn cụ thể (như gián điệp) và liên tục theo dõi và duy trì tương tác với hệ thống. Kẻ tấn công thường là tổ chức, có kỹ thuật và tổ chức tốt, được tài trợ thường xuyên, đôi khi do nhà nước tài trợ. Tóm lại, Advanced Persistent Threats (APTs) là những cuộc tấn công mạng kéo dài, tinh vi, có chủ đích rõ ràng và do những tổ chức lành nghề được tài trợ thực hiện [1]. Hầu hết các cuộc tấn công APT đều liên quan đến lỗ hổng zero-day và rất khó phát hiện.

Các nỗ lực nhằm phát hiện APT chủ yếu dựa vào: (1) xây dựng rules-base dựa trên các các mẫu APT phổ biến và so khớp với audit logs, (2) sử dụng thống kê các thành phần trong hệ thống:system entities, tương tác mạng... để phát hiện

bất thường. (3) sử dụng các kỹ thuật học sâu để mô hình hóa tấn công APT hoặc hành vi hệ thống, sau đó phát hiện APT bằng cách phân loại hoặc phát hiện bất thường.

Tài liệu tham khảo

[1] Adel Alshamrani et al. “A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities”. In: *IEEE Communications Surveys & Tutorials* 21.2 (2019), pp. 1851–1877.