NGHIÊN CỨU VÀ XÂY DỰNG CÔNG CỤ KIỂM THỬ BẢO MẬT TỰ ĐỘNG SỬ DỤNG PHƯƠNG PHÁP HỌC TĂNG CƯỜNG

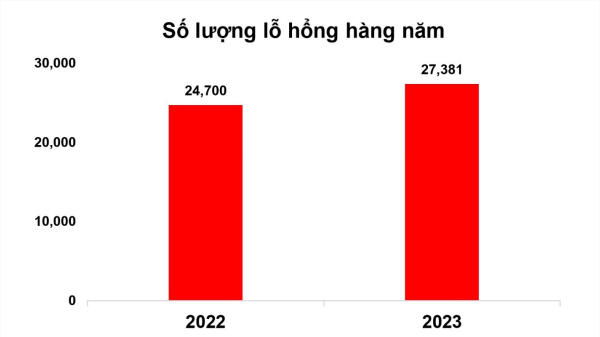
Lê Thành Đạt - 240101041

Tóm tắt

- Lớp: CS2205.CH190
- Link Github của nhóm: https://github.com/ltd-uit/CS2205.CH190
- Link YouTube video: https://youtu.be/4fKUBxvE3Mg
- Họ và tên: Lê Thành Đạt
  - 

Giới thiệu

- Sau đại dịch COVID-19, các quốc gia buộc phải áp dụng công nghệ thông tin vào gần như mọi lĩnh vực.
- Kiểm thử xâm nhập (Penetration Testing) trở thành một công cụ không thể thiếu để đảm bảo an ninh và bảo mật hệ thống. Tuy nhiên, phương pháp kiểm thử thâm nhập **truyền thống** đã **không còn phù hợp** với các hệ thống hiện đại.
- Bài toán: Tiết kiệm tối đa chi phí về thời gian và tài nguyên, đáp ứng nhu cầu về nguồn nhân lực an toàn thông tin đang thiếu hụt hiện nay ?
- **Nghiên cứu của nhóm đề xuất ứng dụng tự động hoá vào quá trình kiểm thử xâm nhập với sự hỗ trợ của AI**



Hình 1: Thống kê số lượng lỗ hổng bảo mật hàng năm (theo Viettel Threat Intelligence)

Mục tiêu

Nghiên cứu xây dựng công cụ kiểm thử bảo mật hệ thống có khả năng:

- **Thực hiện tấn công**: Rà soát các lỗ hổng đang có, xâm nhập sâu vào mạng nội bộ.
- **Tích hợp học tăng cường** vào công cụ đã thiết kế → Tự động hóa quy trình khai thác
- **Mở rộng khả năng khai thác các lỗ hổng phức tạp**: Yêu cầu nhiều bước khai thác hơn so với các công cụ đã có.
- **Tích lũy kinh nghiệm** thông qua mỗi lần thực thi.

So sánh kết quả với các công trình nghiên cứu khác.

Nội dung và Phương pháp

Tìm hiểu chức năng, cách thức hoạt động của một số công cụ hỗ trợ kiểm thử bảo mật phổ biến.

- Nghiên cứu, thực hiện khảo sát các công cụ theo từng mục đích sử dụng
- Tiến hành thử nghiệm với các công cụ trên.

Nội dung và Phương pháp

Tìm hiểu và xây dựng công cụ kiểm thử bảo mật toàn diện sử dụng kết hợp các công cụ khác một cách có chiến lược.

- Xây dựng chiến lược khai thác theo mục đích.
- So sánh, lựa chọn công cụ phù hợp.
- Kết hợp sử dụng các công cụ theo chiến lược đã xây dựng

Nội dung và Phương pháp

Tìm hiểu về học tăng cường sâu (Deep Reinforcement Learning) và tìm cách tích hợp vào công cụ đã xây dựng.

- Tham khảo cơ sở lý thuyết, thực nghiệm của các nghiên cứu ứng dụng(**DRL**)
- Xây dựng framework **DRL** giúp tự động hóa các bước kiểm thử bảo mật (sử dụng thư viện Keras, Tensorflow, ...)
- Tìm hiểu các tham số cần cho bài toán, tinh chỉnh và đào tạo mô hình có hiệu quả tốt nhất.

Nội dung và Phương pháp

Thực nghiệm và đánh giá kết quả.

- Xây dựng mục tiêu cần kiểm thử là máy server chứa các lỗ hổng
- Dùng kịch bản thử nghiệm khác nhau để đánh giá công cụ (hiệu năng, độ chính xác, thuật toán **DRL**, …)
- So sánh với các công cụ đã biết như DeepExploit, Shennina

Kết quả dự kiến

- Đề xuất chiến lược thu thập thông tin và triển khai module tối ưu.
- Xây dựng module khai thác có khả năng đánh giá bảo mật mục tiêu toàn diện và thực hiện các hình thức tấn công phức tạp.
- Triển khai mô hình học tăng cường và tích hợp vào công cụ đã xây dựng.
- Có được kết quả thực nghiệm và đưa ra được báo cáo tổng quan về quá trình thực hiện đề tài.

Tài liệu tham khảo

[1]. Shah, Mujahid, et al. "Penetration testing active reconnaissance phase–optimized port scanning with nmap tool." *2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*. IEEE, 2019.

[2]. Li, Zegang, Qian Zhang, and Guangwen Yang. "EPPTA: Efficient partially observable reinforcement learning agent for penetration testing applications." *Engineering Reports* 7.1 (2025): e12818.

[3]. Ghanem, M. C., Chen, T. M., & Nepomuceno, E. G. (2023). Hierarchical reinforcement learning for efficient and effective automated penetration testing of large networks. *Journal of Intelligent Information Systems*, *60*(2), 281-303.

[4]. Ovidiu Valea, Ciprian Oprisa: Towards Pentesting Automation Using the Metasploit Framework. ICCP 2020: 171-178.

[5]. Fabio Massimo Zennaro, László Erdodi: Modelling penetration testing with reinforcement learning using capture-the-flag challenges: Trade-offs between model-free learning and a priori knowledge. IET Inf. Secur. 17(3): 441-457 (2023).

[6]. Tyler Cody: A Layered Reference Model for Penetration Testing with Reinforcement Learning and Attack Graphs. CoRR abs/2206.06934 (2022)

[7]. Li, Qianyu, et al. "An intelligent penetration testing method using human feedback." *IEEE Transactions on Industrial Informatics* (2024).