ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
KHOA MẠNG MÁY TÍNH VÀ TRUYỀN THÔNG

ẨN THÔNG TIN TRÊN DỮ LIỆU SỐ VÀ ỨNG DỤNG

CÁC KỸ THUẬT ẨN GIẤU THÔNG TIN TRONG ỨNG DỤNG VÀ DỮ LIỆU

TÔ TRỌNG NGHĨA
NGUYỄN HỒNG SƠN
PHẠM TRẦN TIẾN ĐẠT
TẠ VIỆT HOÀNG

GIẢNG VIÊN HƯỚNG DẪN
TS. NGUYỄN NGỌC TỰ

TP. HỒ CHÍ MINH, NĂM 2023

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
KHOA MẠNG MÁY TÍNH VÀ TRUYỀN THÔNG

NGUYỄN HỒNG SƠN - 220202022
TÔ TRỌNG NGHĨA - 220202019
PHẠM TRẦN TIẾN ĐẠT - 220202017
TẠ VIỆT HOÀNG - 220202018

ẨN THÔNG TIN TRÊN DỮ LIỆU SỐ VÀ ỨNG DỤNG

CÁC KỸ THUẬT ẨN GIẤU THÔNG TIN TRONG ỨNG DỤNG VÀ DỮ LIỆU

GIẢNG VIÊN HƯỚNG DẪN
TS. NGUYỄN NGỌC TỰ

TP. HỒ CHÍ MINH, NĂM 2023

Mục lục

Danh sách hình vẽ

Danh sách bảng

Danh sách từ viết tắt

RL	**R**einforcement **L**earning
MDP	**M**arkov **D**ecision **P**rocesses

Tóm tắt đề tài

Sự phát triển của công nghệ mạng và truyền thông trong kỷ nguyên hiện đại đã làm tăng tốc độ truyền tải lên gấp hàng ngàn lần. Dữ liệu truyền tài trên mạng máy tính luân chuyển liên tục, đòi hỏi sự an toàn cho chúng.

An toàn thông tin trong lĩnh vực này chia thành mã hóa thông tin và ẩn giấu thông tin. Mã hóa thông tin chuyển đổi những dữ liệu bí mật thành loại dữ liệu khác mà kẻ tấn công không thể đọc được nó. Tuy vậy, dữ liệu mã hóa trở thành ốc đảo giữa sa mạc, gây sự chú ý rất lớn từ kẻ tấn công.

Do đó, một kỹ thuật khác nhằm bảo vệ dữ liệu là che giấu chính sự tồn tại của bí mật đó trước kẻ tấn công. Các dữ liệu được nhúng và gần như tàng hình trước kẻ xấu, và như vậy ít gây chú ý hơn. Trong lĩnh vực này, chia thành hai nội dung với những mục đích khác nhau. Trong khi kỹ thuật giấu tin ẩn vào dữ liệu được thực hiện nhằm mục đích bảo vệ sự bí mật của "dữ liệu được giấu" thì kỹ thuật watermark lại có mục đích bảo vệ chính dữ liệu đó. Với khả năng rút trích dữ liệu được giấu từ phiên bản số, ta dễ dàng chứng minh được tác quyền với nó [**subhedar2014current**].

Chương 1

Giới thiệu

1.1 Ẩn thông tin và các đặc điểm của nó

Ẩn giấu thông tin là phương pháp đang được sử dụng để bảo mật và bí mật cho việc trao đổi dữ liệu. Thay vì dựa vào việc mã hóa thông điệp để bảo vệ nó khỏi sự xâm nhập, ẩn giấu thông tin đặt mục tiêu vào việc nhúng các thông tin nhạy cảm - từ các tệp tin, tin nhắn, hình ảnh, âm thanh cho đến video - vào bên trong các tệp tin gốc khác, có thể thuộc cùng một loại hoặc khác loại. Quá trình này được thực hiện một cách khéo léo để đảm bảo rằng dữ liệu ẩn được bảo vệ chặt chẽ và không thể dễ dàng nhận biết bởi bất kỳ ai ngoài các bên liên quan, chẳng hạn như người gửi và người nhận.

Trong lĩnh vực bảo mật thông tin, mật mã thường được sử dụng để mã hóa và giải mã thông điệp, tạo ra một tầng bảo vệ vững chắc. Ẩn giấu thông tin tập trung vào việc chèn thông tin bí mật mà không gây ra bất kỳ thay đổi nào trong dữ liệu gốc. Điều này đặc biệt hữu ích khi cần duy trì tính nguyên vẹn của dữ liệu gốc mà vẫn muốn lưu trữ thông tin bí mật.

Tính không thể nhận thấy là một đặc trưng quan trọng của ẩn giấu thông tin, được gọi là tính trong suốt hoặc hiệu suất chống phát hiện. Khả năng này đảm bảo rằng dữ liệu ẩn không dễ dàng bị phát hiện bởi những phương pháp kiểm tra thông thường. Để nâng cao tính trong suốt của kỹ thuật này, có thể thực hiện các cải tiến trong phương pháp ẩn giấu thông tin hoặc tăng cường mối quan hệ giữa thông tin bí mật và tệp chứa thông tin.

1.2 Ẩn thông tin dựa trên bit ít quan trọng nhất (LSB Steganography)

Least Significant Bit (LSB) Steganography là một kỹ thuật trong lĩnh vực ẩn giấu thông tin, được sử dụng để nhúng thông tin bí mật vào trong một tập tin đa phương tiện, chẳng hạn như hình ảnh, âm thanh hoặc video. Kỹ thuật này tận dụng tính chất của các bit ít quan trọng nhất (Least Significant Bits) trong các dữ liệu số như điểm ảnh, mẫu âm thanh hoặc khung hình video. LSB Steganography cho phép nhúng thông tin ẩn vào những bit ít quan trọng này mà không gây ra sự thay đổi đáng kể cho dữ liệu gốc.

Nguyên tắc hoạt động của LSB Steganography rất đơn giản: trong một tập tin đa phương tiện, các dữ liệu như điểm ảnh thường được biểu diễn bằng các chuỗi bit. Các bit ít quan trọng nhất thường có giá trị thấp hơn và có xu hướng thay đổi ít ảnh hưởng đến hình ảnh hoặc âm thanh tổng thể. Điều này tạo ra một cơ hội tốt để thay thế những bit này bằng các bit của thông tin ẩn, giữ nguyên tính nguyên vẹn của dữ liệu gốc mà vẫn chèn thông tin bí mật vào.

Một ví dụ cụ thể có thể là nhúng một chuỗi văn bản thông điệp vào một hình ảnh bằng cách thay thế các bit ít quan trọng nhất của các điểm ảnh trong hình ảnh đó. Khi xem hình ảnh, sự thay đổi này thường không dễ dàng bị nhận thấy bởi mắt người. Tuy nhiên, người nhận có thể sử dụng một thuật toán tương ứng để trích xuất thông điệp ẩn ra khỏi hình ảnh.

Mặc dù LSB Steganography đơn giản và dễ triển khai, nó vẫn có nhược điểm. Việc nhúng thông tin quá nhiều có thể làm thay đổi tới mức có thể nhận thấy được trong dữ liệu đầu ra. Ngoài ra, nếu người tấn công biết rằng một hình ảnh hoặc tập tin âm thanh đã được sử dụng để nhúng thông tin, họ có thể dễ dàng tìm ra thông điệp ẩn.

Chương 2

Cơ sở lý thuyết và các nghiên cứu liên quan

Chương 3

Phương pháp và mô hình đề xuất

Chương 4

Thực nghiệm và đánh giá

Chương 5

Tổng kết

5.1 Kết quả đạt được

Trong nghiên cứu này, chúng tôi đã khám phá hiệu suất của DRL trong xây dựng một công cụ kiểm thử xâm nhập tự động thông qua Metasploit Framework để thực hiện quét và khai thác kèm tích lũy kinh nghiệm.

Chúng tôi cũng đã xây dựng các môi trường lỗ hổng và thực hiện các kịch bản thử nghiệm khác nhau để đánh giá toàn diện hiệu suất của công cụ, bao gồm đào tạo và thử nghiệm dựa trên hiệu quả khai thác đạt được. Trong quá trình thử nghiệm, công cụ của chúng tôi đã có thể khai thác thành công tất cả các lỗ hổng dịch vụ mà chúng tôi đã tạo trong môi trường khai thác lý tưởng.

Với kết quả xuất sắc này, công cụ này đã chứng minh được khả năng tích lũy kết quả học từ các môi trường trước để thành công khai thác lỗ hổng cho lần khai thác tiếp theo trong môi trường khác ngay lần đầu tiên.

Tuy nhiên, môi trường chúng tôi thử nghiệm là môi trường lý tưởng nhất, khả năng khai thác của công cụ có thể giảm đi nếu môi trường chứa tường lửa hoặc số cổng bị thay đổi.

5.2 Hướng phát triển

Dựa vào kết quả đạt được, chúng tôi nhận thấy công cụ có tiềm năng phát triển trong tương lai, tuy nhiên cần phải có nhiều chỉnh sửa hơn nữa để có thể đạt được hiệu quả học tập tốt nhất. Chúng ta có thể chuẩn bị thêm nhiều môi

trường chứa lỗ hổng hơn để huấn luyện và tạo ra sự đa dạng về kinh nghiệm khai thác cho công cụ, cung cấp môi trường hoạt động cho các tác nhân học song song độc lập với nhau, giải quyết nhược điểm đụng độ khai thác.

Bên cạnh đó, chúng ta cũng có thể mở rộng chức năng công cụ thêm ở các bước khác trong quy trình kiểm thử thâm nhập, hoặc sử dụng thêm hậu khai thác để tiếp tục tấn công vào các máy chủ mục tiêu khác trong mạng cục bộ của nó.

Ngoài ra, chúng ta có thể tạo thêm mô-đun phụ trợ cho công cụ để có thể sử dụng các mô-đun khai thác ở các mức xếp hạng khác nhau trên Metasploit, từ đó mở rộng khả năng khai thác của công cụ. Đây là một công việc phức tạp và đòi hỏi thời gian rất lớn, bởi các mô-đun khai thác (Mức xếp hạng normal) khác nhau sẽ có những thông tin cấu hình riêng cho phù hợp với các lỗ hổng dịch vụ khác nhau.