BỘ THÔNG TIN VÀ TRUYỀN THÔNG
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

TÓM TẮT LUẬN ÁN

NGHIÊN CỨU CÁC KỸ THUẬT PHÁT HIỆN DGA BOTNET

NCS: VŨ XUÂN HẠNH

TẬP THỂ HƯỚNG DẪN: PGS. TS. HOÀNG XUÂN DẬU
TS. NGÔ QUỐC DŨNG

HÀ NỘI - 2022

.

Công trình được hoàn thành tại:

Học viện Công nghệ Bưu chính Viễn thông

Người hướng dẫn khoa học: 1. TS. Hoàng Xuân Dậu

2. TS. Ngô Quốc Dũng

Phản biện 1: .....

.....

Phản biện 2: .....

.....

Phản biện 3: .....

.....

Luận án được bảo vệ trước Hội đồng chấm luận án cấp Học viện

họp tại:……

……...

Vào hồi giờ ngày tháng năm

Có thể tìm hiểu luận án tại thư viện:……

(*ghi tên các thư viện nộp luận án*)

PHẦN MỞ ĐẦU

1. GIỚI THIỆU

Bot là một dạng phần mềm độc hại cho phép các nhóm kẻ tấn công, hay tin tặc kiểm soát từ xa các máy tính hoặc các hệ thống tính toán (gọi chung là máy tính) có kết nối Internet. Khi một máy tính bị lây nhiễm bot, nó được gọi là *máy tính ma*, hay *zombie*. Tập hợp các máy bot do một nhóm tin tặc kiểm soát (*botmaster*) được gọi là *botnet* - hay mạng của các bot. Botmaster thường điều khiển các bot trong botnet do mình kiểm soát thông qua hệ thống các máy chủ chỉ huy và kiểm soát (Command and Control, hoặc C&C, hoặc CnC). Khác với các phần mềm độc hại thông thường, các bot trong một botnet có khả năng tương tác với nhau và kết nối đến máy chủ CnC của botnet để nhận lệnh và mã cập nhật từ botmaster. Hơn nữa, các bot cũng được trang bị các kỹ thuật ẩn mình tiên tiến, như đóng gói, xáo trộn mã, mã hóa, nâng cấp, cập nhật mã nhị phân... giúp cho chúng có khả năng tồn tại lâu dài trên hệ thống nạn nhân. Quy mô của các botnet có thể rất khác nhau, từ hàng hàng chục ngàn đến hàng trăm ngàn bot phân tán ở mọi vị trí địa lý trên mạng Internet. Đặc biệt, một số botnet như Conficker theo ước tính có hơn 10.5 triệu bot.

Trong những năm gần đây, các botnet được xem là một trong những mối đe dọa an ninh chủ yếu đối với các hệ thống thông tin, các thiết bị có kết nối và người dùng Internet. Điều này là do các botnet có liên hệ trực tiếp đến nhiều dạng tấn công và lạm dụng trên mạng Internet, như các cuộc tấn công từ chối dịch vụ (DDoS) trên qui mô lớn và rất lớn, gửi thư rác, truyền tải và phát tán các loại mã độc, sinh click và like ảo và đánh cắp các thông tin nhạy cảm. Hơn nữa, các dạng tấn công nguy hiểm do botnet hỗ trợ thực hiện còn bao gồm giả mạo địa chỉ URL, giả mạo hệ thống tên miền (DNS), tấn công chèn mã độc trên các ứng dụng web và thu thập các thông tin nhạy cảm từ người dùng. Các tổ chức tài chính và các cơ quan chính phủ thường là các mục tiêu chính của các dạng tấn công do botnet hỗ trợ thực hiện. Một vấn đề khác khiến cho các mối đe dọa từ botnet các trở lên nghiêm trọng, khó bị phát hiện và loại bỏ là do trong quá trình phát triển của mình các botnet liên tục tiến hóa trên mạng Internet về cả qui mô và mức độ tinh vi của các kỹ thuật điều khiển.

Do tính chất nguy hiểm của botnet và các dạng mã độc mà botnet hỗ trợ truyền tải và phát tán, nhiều giải pháp đã được nghiên cứu, phát triển và triển khai trên thực tế cho giám sát, phát hiện và loại bỏ botnet. Có thể chia các giải pháp giám sát, phát hiện botnet thành 2 nhóm: (1) các giải pháp dựa trên honeynet và (2) các giải pháp dựa trên hệ thống phát hiện xâm nhập (IDS). Các giải pháp thuộc nhóm (1) xây dựng các honeynet - là các mạng bẫy để thu thập các thông tin về các botnet đang hoạt động và sau đó sử dụng các thông tin thu thập được để phân tích các đặc tính và hành vi của botnet. Nhìn chung, các giải pháp dựa trên honeynet có ưu điểm là dễ xây dựng và không yêu cầu lớn về tài nguyên tính toán. Tuy vậy, các giải pháp này thường bị hạn chế về khả năng mở rộng và khả năng tương tác với mã độc botnet. Các giải pháp thuộc nhóm (2) sử dụng các kỹ thuật giám sát, phát hiện của IDS để giám sát, phát hiện botnet. Dựa trên kỹ thuật phát hiện, các giải pháp dựa trên IDS lại có thể được chia thành (i) phát hiện dựa trên dấu hiệu, chữ ký và (2) phát hiện dựa trên bất thường. Trong các hướng phát hiện dựa trên bất thường, hướng phát hiện botnet dựa trên giám sát lưu lượng mạng, giám sát các truy vấn hệ thống DNS sử dụng học máy được quan tâm nghiên cứu, phát triển và cho nhiều kết quả khả quan.

Luận án này tập trung nghiên cứu các phương pháp, kỹ thuật phát hiện các dấu hiệu hoạt động của các botnet sử dụng dữ liệu truy vấn hệ thống DNS dựa trên học máy. Trước hết, luận án sẽ thực hiện khảo sát về botnet, kiến trúc và hoạt động của botnet, và khảo sát, hệ thống hóa các giải pháp giám sát, phát hiện botnet. Sau đó, luận án phát triển và thử nghiệm một số mô hình phát hiện DGA botnet dựa trên các kỹ thuật học máy sử dụng dữ liệu truy vấn hệ thống DNS.

2. TÍNH CẤP THIẾT CỦA LUẬN ÁN

Như đã đề cập trong mục Giới thiệu, các botnet đã thực sự trở thành một trong các mối đe dọa lớn nhất đối với mạng Internet toàn cầu do chúng đã và đang phát triển rất mạnh về cả quy mô, mức độ phân tán, kỹ thuật điều khiển và trực tiếp thực hiện, hoặc có liên quan chặt chẽ đến nhiều hoạt động độc hại, như tấn công DDoS, phát tán thư rác, quảng bá, phát tán các loại phần mềm độc hại, phần mềm gián điệp, quảng cáo, giả mạo địa chỉ URL, giả mạo hệ thống DNS, tấn công chèn mã độc trên các ứng dụng web và đánh cắp các thông tin nhạy cảm trên các hệ thống máy chủ cũng trên hệ thống máy người dùng cuối. Một số họ mã độc tống tiền (ransomware) được phát hiện gần đây có khả năng tự quảng bá, truyền thông qua mạng botnet và thậm chí các cuộc tấn công có chủ đích (APT) cũng đã bắt đầu sử dụng các botnet để triển khai thực hiện. Trong vài năm qua, một xu hướng mới của mạng botnet như một dịch vụ (Botnet as a Service - BaaS) đã hình thành, làm giảm chi phí của tội phạm mạng khi thực hiện các cuộc tấn công liên tục với qui mô rất lớn và mặt khác, giúp chúng kiểm soát botnet dễ dàng hơn. Cùng với xu hướng này, ngày càng có nhiều mạng botnet với quy mô ngày càng tăng với mức độ phân tán rất cao, tạo ra mối đe dọa nghiêm trọng đối với hệ sinh thái Internet.

Do mối đe dọa của các botnet đối với mạng Internet toàn cầu, các hệ thống, dịch vụ và người dùng Internet ngày càng lớn, việc nghiên cứu, phát triển và ứng dụng các giải pháp giám sát, phát hiện và loại trừ botnet là rất cấp thiết. Tuy vậy, do các bot trong botnet thường có tính phân tán, khả năng giấu mình và tính tự động (autonomy) rất cao, nên việc giám sát, phát hiện và loại trừ botnet gặp rất nhiều thách thức. Giải pháp tổng thể để khắc chế mối đe dọa từ botnet cần sự phối hợp hành động từ nhiều bên có liên quan, bao gồm các cơ quan chính quyền, các nhà cung cấp dịch vụ Internet (ISP), các tổ chức, doanh nghiệp và cả người dùng Internet. Chẳng hạn, cần có khung pháp lý về an toàn thông tin mạng từ các cơ quan chính quyền; cần có các hệ thống giám sát, phát hiện hoạt động của mã độc, các bot, botnet trên các cổng dịch vụ của các ISP, các cơ quan, tổ chức, doanh nghiệp; và ý thức cảnh giác của người dùng Internet. Trong đó, các giải pháp, kỹ thuật giám sát, phát hiện hoạt động và loại trừ các bot, botnet đóng vai trò trọng yếu và đây cũng là hướng nghiên cứu của đề tài luận án này - tập trung nghiên cứu phát hiện botnet sử dụng kỹ thuật phát hiện xâm nhập dựa trên bất thường.

Luận án sử dụng kỹ thuật phát hiện xâm nhập dựa trên bất thường cho phát hiện botnet do kỹ thuật này có ưu điểm nổi bật là có khả năng phát hiện các dạng bot, botnet mới mà không đòi hỏi phải có trước các thông tin về chúng như kỹ thuật phát hiện dựa trên dấu hiệu, chữ ký. Hơn nữa, phát hiện dựa trên bất thường cho phép tự động hóa quá trình xây dựng mô hình phát hiện botnet từ tập dữ liệu huấn luyện, nhờ đó giảm thiểu việc sử dụng nhân lực chuyên gia cho xây dựng thủ công các tập luật phát hiện. Nhược điểm chính của phát hiện botnet dựa trên bất thường là tỷ lệ cảnh báo sai (gồm tỷ lệ dương tính giả và và tỷ lệ âm tính giả) còn tương đối cao so với kỹ thuật phát hiện dựa trên dấu hiệu, chữ ký.

Trong nhóm các kỹ thuật phát hiện botnet dựa trên bất thường, các hướng (1) phát hiện botnet dựa trên giám sát lưu lượng mạng và (2) phát hiện dựa trên giám sát và phân tích truy vấn DNS thu hút được sự quan tâm lớn của cộng đồng nghiên cứu và các hãng bảo mật. Nổi bật trong hướng (1) là các hệ thống giám sát, phát hiện botnet đã được phát triển và triển khai, như BotHunter, BotSniffer, BotTrack, BotMiner, BotFinder và BotProbe. Các hệ thống trên đã được triển khai và đã giám sát, thu thập được một lượng lớn dữ liệu lưu lượng mạng có liên quan đến hoạt động của các bot và botnet phục vụ cho phân tích. Nhằm hỗ trợ cho các nhóm nghiên cứu, Garcia và cộng sự đã xây dựng bộ dữ liệu thu thập lưu lượng mạng botnet với nhiều kịch bản khác nhau với tên là CTU-13. Nhược điểm chính của các hệ thống dạng này là yêu cầu rất cao về năng lực bắt, xử lý và lưu trữ một lượng rất lớn các gói tin lưu thông qua các cổng mạng. Điều này có thể làm giảm khả năng triển khai và vận hành hiệu quả các giải pháp dạng này trên thực tế, đặc biệt là trên các cổng mạng có lưu lượng lớn.

Hướng (2) phát hiện botnet dựa trên giám sát và phân tích các truy vấn DNS được đông đảo cộng đồng nghiên cứu quan tâm trong những năm gần đây, đặc biệt với sự phát triển vượt trội của các họ DGA botnet. DGA botnet gồm các họ botnet sử dụng các thuật toán để tự động sinh và đăng ký tên miền cho các máy chủ CnC của chúng. Đây là kỹ thuật mà các botnet sử dụng để thay thế cho các tên miền và địa chỉ IP cố định cho các máy chủ CnC của chúng nhằm lẩn tránh các kỹ thuật rà quét và chặn lọc. Trong quá trình hoạt động của botnet, botmaster tự động định kỳ sinh các tên miền sử dụng kỹ thuật DGA cho các máy chủ CnC của botnet và đăng ký với hệ thống DNS động. Trong khi đó, các bot trong botnet được lập trình để tự động kết nối máy chủ máy chủ CnC của botnet để tải các lệnh và mã cập nhật. Để thực hiện kết nối, các bot định kỳ tự sinh tên miền của máy chủ CnC sử dụng cùng kỹ thuật DGA và gửi tên miền này lên hệ thống DNS cục bộ để tìm địa chỉ IP của máy chủ CnC. Nếu bot nhận được địa chỉ IP từ hệ thống DNS, nó tạo kết nối đến máy chủ CnC để tải các lệnh và mã cập nhật. Nếu tên miền truy vấn không tồn tại, bot lại sinh một tên miền mới và thực hiện lại quá trình truy vấn hệ thống DNS ở chu kỳ kế tiếp. Mỗi họ DGA botnet sử dụng các thuật toán DGA sinh tên miền khác nhau và số lượng, tần suất sinh tên miền mới cũng khác nhau. Một số họ botnet sử dụng thuật toán DGA sinh tên miền dựa trên thời gian, hoặc dựa trên việc tổ hợp ngẫu nhiên các ký tự (character-based DGA), hoặc dựa trên việc tổ hợp các từ lấy trong từ điển (word-based DGA), hoặc dựa trên sự kết hợp giữa tổ hợp ngẫu nhiên các ký tự và tổ hợp các từ lấy trong từ điển (mixed DGA). Về số lượng tên miền sinh, một số botnet chỉ sinh vài chục tên miền trong cả vòng đời hoạt động, trong khi đó cũng có những botnet sinh hàng chục, thậm chí hàng trăm ngàn tên miền trong vòng đời hoạt động của chúng.

Như vậy, do hoạt động của các DGA botnet gắn liền với việc truy vấn hệ thống DNS, nên có thể giám sát và phân tích các truy vấn các máy chủ DNS có thể tìm được các bằng chứng về sự tồn tại các bot và hoạt động của botnet. Có nhiều giải pháp, kỹ thuật được sử dụng cho giám sát, phân tích lưu lượng truy vấn DNS và nhận dạng, phân loại các tên miền được sử dụng bởi botnet và các tên miền hợp lệ. Trong thời gian gần đây, các phương pháp học máy được sử dụng rộng rãi trong nhận dạng, phân loại các tên miền được sử dụng bởi botnet và các tên miền hợp lệ nhờ đạt độ chính xác cao và khả năng tự động hóa xây dựng mô hình phát hiện từ tập dữ liệu huấn luyện. Ưu điểm của các đề xuất đã nêu là độ chính xác tương đối cao khi thử nghiệm với từng tập dữ liệu cụ thể và khả năng tự động hóa việc xây dựng mô hình phát hiện. Tuy vậy, tỷ lệ cảnh báo sai của các đề xuất này còn khá cao, đến hơn 10% với, ảnh hưởng đến khả năng triển khai thực tế. Lý do cho vấn đề này là tập đặc trưng, hoặc phương pháp phân loại sử dụng trong các đề xuất đã có chưa thực sự phù hợp để nhận dạng sự khác biệt giữa các tên miền DGA và các tên miền hợp lệ. Ngoài ra, do một số họ DGA botnet liên tục sử dụng các thuật toán sinh tên miền mới, như các họ word-based và mixed DGA cho phép sinh các tên miền DGA rất giống với các tên miền hợp lệ và do vậy một số đề xuất đã có không có khả năng phát hiện các họ DGA botnet này.

Đề tài "Nghiên cứu các kỹ thuật phát hiện DGA botnet" được thực hiện trong phạm vi luận án tiến sĩ chuyên ngành hệ thống thông tin nhằm góp phần giải quyết một số vấn đề còn tồn tại trong các kỹ thuật, giải pháp phát hiện các dạng DGA botnet, bao gồm: (1) lựa chọn, trích xuất tập đặc trưng mới phù hợp hơn để phân biệt tốt hơn các tên miền DGA và tên miền hợp lệ, nhằm tăng độ chính xác phát hiện, giảm tỷ lệ cảnh báo sai và (2) phát triển mô hình kết hợp có khả năng phát hiện đồng thời nhiều họ DGA botnet.

3. MỤC TIÊU CỦA LUẬN ÁN

Mục tiêu của luận án là nghiên cứu, đề xuất một số mô hình phát hiện botnet dựa trên các kỹ thuật học máy. Cụ thể, luận án tập trung vào các mục tiêu sau: (1) Nghiên cứu, đánh giá các phương pháp, kỹ thuật, giải pháp, công cụphát hiện botnet hiện có; (2) Nghiên cứu, đề xuất các mô hình phát hiện botnet dựa trên học máy có giám sát và học kết hợp sử dụng các tập đặc trưng phân loại tên miền mới nhằm nâng cao độ chính xác, giảm cảnh báo sai,

đồng thời cho phép phát hiện nhiều dạng DGA botnet; (3) Cài đặt, thử nghiệm và đánh giá các mô hình phát hiện botnet đã đề xuất sử dụng các tệp dữ liệu thực tế.

4. ĐỐI TƯỢNG NGHIÊN CỨU VÀ PHẠM VI NGHIÊN CỨU

Đối tượng nghiên cứu là botnet và đặc biệt là các họ DGA botnet.

Phạm vi nghiên cứu giới hạn trong các kỹ thuật, giải pháp phát hiện DGA botnet sử dụng dữ liệu truy vấn DNS.

5. PHƯƠNG PHÁP NGHIÊN CỨU

Luận án sử dụng phương pháp nghiên cứu lý thuyết kết hợp với phương pháp thực nghiệm. Trong đó, phương pháp nghiên cứu lý thuyết được sử dụng để thực hiện các phần việc sau: (1) Nghiên cứu nền tảng lý thuyết về botnet cho luận án, bao gồm khái quát về botnet, bot, phương thức hoạt động của botnet, vấn đề botnet khai thác hệ thống DNS trong quá trình hoạt động; (2) Nghiên cứu nền tảng lý thuyết về học máy cho luận án, bao gồm khái quát về học máy, một số giải thuật học máy có giám sát, phương pháp đánh giá và các độ đo đánh giá mô hình phát hiện dựa trên học máy; (3) Khảo sát, đánh giá các đề xuất, giải pháp đã có cho phát hiện botnet, DGA botnet, trên cơ sở đó tổng hợp các ưu điểm, nhược điểm làm cơ sở cho đề xuất của luận án; (4) Lựa chọn, đề xuất các đặc trưng mới, xây dựng các mô hình phát hiện DGA botnet dựa trên phân loại trên miền DGA với tên miền hợp lệ.

Phương pháp thực nghiệm được sử dụng trong luận án để thực hiện các phần việc sau: (1) Khảo sát các tập dữ liệu về botnet, DGA botnet và lựa chọn tập dữ liệu phù hợp cho thực nghiệm; (2) Thử nghiệm các mô hình phát hiện DGA botnet đề xuất trong luận án, đánh giá, so sánh các mô hình đề xuất với các mô hình, đề xuất đã có.

6. CÁC ĐÓNG GÓP CỦA LUẬN ÁN

Đóng góp thứ nhất của luận án là đề xuất mô hình phát hiện DGA botnet dựa trên học máy sử dụng các đặc trưng ký tự và các đặc trưng từ. Mô hình sử dụng các đặc trưng ký tự có khả năng phát hiện hiệu quả các character-based DGA botnet - là các botnet tự sinh tên miền sử dụng thuật toán ghép ngẫu nhiên các ký tự. Mô hình sử dụng các đặc trưng từ có khả năng phát hiện hiệu quả các word-based DGA botnet - là các botnet tự sinh tên miền sử dụng thuật toán ghép các từ theo từ điển.

Đóng góp thứ hai của luận án là đề xuất mô hình phát hiện DGA botnet dựa trên học kết hợp (*ensemble learning*). Mô hình này cho phép phát hiện hiệu quả cả character-based và word-based DGA botnet sử dụng thuật toán học kết hợp.

7. BỐ CỤC CỦA LUẬN ÁN

Luận án được bố cục thành ba chương với nội dung chính như sau:

Chương 1 giới thiệu tổng quan về botnet, khái quát về phát hiện botnet, các kỹ thuật phát hiện botnet và một số giải pháp, công cụ phát hiện botnet. Chương 1 cũng giới thiệu khái quát về học máy và mô tả một số giải thuật học máy có giám sát sử dụng trong các mô hình phát hiện botnet đề xuất trong các chương 2 và 3. Phần tiếp theo của chương mô tả các tập dữ liệu liên quan đến botnet được sử dụng trong luận án. Phần cuối của chương chỉ ra 2 vấn đề sẽ được giải quyết trong luận án.

Chương 2 trình bày khái quát về DGA botnet và cơ chế DGA botnet khai thác hệ thống DNS để duy trì hoạt động. Chương này cũng khảo sát các phương pháp, đề xuất hiện có cho phát hiện botnet nói chung và DGA botnet nói riêng. Phần tiếp theo của chương mô tả, thử nghiệm và đánh giá mô hình phát hiện character-based DGA botnet dựa trên học máy sử dụng các đặc trưng ký tự. Phần cuối của chương mô tả, thử nghiệm và đánh giá mô hình phát hiện character-based DGA botnet dựa trên học máy sử dụng các đặc trưng ký tự.

Chương 3 giới thiệu khái quát về học kết hợp (*ensemble learning*), khảo sát các kỹ thuật phát hiện DGA botnet dựa trên học kết hợp. Phần cuối của chương mô tả, thử nghiệm và đánh giá mô hình phát hiện DGA botnet đề xuất dựa trên học kết hợp.

Cuối cùng là Kết luận của luận án.

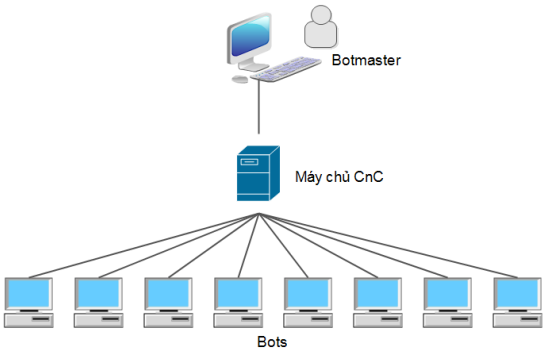
CHƯƠNG 1: TỔNG QUAN VỀ BOTNET VÀ PHÁT HIỆN BOTNET

1.1. TỔNG QUAN VỀ BOTNET

1.1.1. Khái quát về botnet và phương thức hoạt động

1.1.1.1. Giới thiệu về bot, botnet

Bot là một loại phần mềm độc hại cho phép kẻ tấn công giành quyền kiểm soát máy tính, hoặc thiết bị tính toán bị lây nhiễm. Máy tính bị nhiễm bot thường được gọi là *zombie* hay là *máy tính ma*. Trên thực tế có hàng ngàn, hàng trăm ngàn máy tính và thiết bị tính toán có kết nối Internet bị nhiễm một số loại bot mà người dùng không biết và không nhận ra chúng. Kẻ tấn công có thể truy cập các *zombie* và kích hoạt chúng thực thi các cuộc tấn công từ chối dịch vụ, hoặc gửi hàng loạt thư rác. Khi thực hiện truy vết ngược lại nguồn khởi phát các cuộc tấn công, người ta thường tìm thấy các *zombie* - cũng là nạn nhân chứ không phải là kẻ tấn công thực sự. Các bot do một hoặc một nhóm kẻ tấn công thông qua một hoặc một số máy tính (gọi là *botmaster*) kiểm soát và chúng được liên kết tạo thành một mạng lưới các máy bị kiểm soát được gọi là *botnet*. Botmaster thường điều khiển các bot trong botnet do mình kiểm soát thông qua hệ thống các máy chủ chỉ huy và kiểm soát (Command and Control, hoặc C&C, hoặc CnC), như minh họa trên Hình 1.1. Kênh giao tiếp giữa các bot và các máy chủ CnC trong một botnet có thể là IRC, HTTP hoặc giao thức truyền thông khác.



Hình 1.1: Mô hình botmaster kiểm soát các bot thông qua các máy chủ CnC

1.1.1.2. Phương thức hoạt động, vòng đời

Vòng đời của một mạng botnet bao gồm 7 bước. Theo đó, các bước trong vòng đời botnet bao gồm Khởi tạo, Đăng ký, Lây nhiễm sơ bộ, Xây dựng mạng bot, Tập hợp, Khởi động tấn công, và Nâng cấp và bảo trì.

1.1.2. Phân loại botnet

Các botnet có thể được phân loại theo 2 tiêu chí: (i) kiến trúc mạng và (ii) giao thức truyền thông. Botnet có thể được tổ chức theo nhiều mô hình mạng, chủ yếu theo mô hình tổ chức hệ thống các máy chủ CnC là trung gian giữa botmaster và các bot. Các giao thức truyền thông là các giao thức hỗ trợ giao tiếp giữa các máy chủ CnC và các bot trong botnet.

1.1.3. Lịch sử phát triển của botnet

1.1.4. Tác hại và các dạng khai thác botnet

Botnet có thể được sử dụng cho một loạt các hành động nguy hiểm, bao gồm tấn công DDoS, tạo và gửi thư rác (Spam), lừa đảo, lây lan phần mềm độc hại, quảng bá phần mềm quảng cáo, gián điệp, lưu trữ các trang web hoặc nội dung độc hại.

1.2. PHÁT HIỆN BOTNET

1.2.1. Khát quát về phát hiện botnet

Chính vì mối đe dọa từ botnet ngày một gia tăng, phát hiện botnet đề cập đến việc phát hiện các hoạt động nguy hiểm, hoặc bất thường được thực hiện trong môi trường mạng được kiểm soát. Phát hiện botnet hiện đang là một thách thức lớn đối với các nhà nghiên cứu và

các tổ chức do botnet được xem là mục tiêu di động nhờ tính phân tán cao và khả năng ẩn mình của các bot. Như vậy, tất cả các khía cạnh có liên quan đến phát hiện botnet bao gồm phát hiện, giảm thiểu và phản ứng phải luôn thay đổi theo thời gian. Để có thể phòng chống botnet hiệu quả cần sự phối hợp của nhiều bên liên quan. Các bên liên quan khác nhau, ví dụ như các cơ quan chính phủ, các doanh nghiệp, các nhà mạng và các nhà cung cấp dịch vụ Internet (ISP) có nhiều cách tiếp cận khác nhau để xử lý vấn đề botnet.

1.2.2. Các kỹ thuật phát hiện botnet

Có nhiều kỹ thuật phát hiện botnet đã được đề xuất và ứng dụng. Mục này trình bày 4 nhóm kỹ thuật phát hiện botnet được sử dụng phổ biến, bao gồm (i) phát hiện dựa trên honeynet, (ii) phát hiện dựa trên luật, dấu hiệu và (iii) phát hiện dựa trên bất thường.

1.2.2.1. Phát hiện dựa trên Honeynet

1.2.2.2. Phát hiện dựa trên luật, dấu hiệu

1.2.2.3. Phát hiện dựa trên bất thường

1.2.3. Một số giải pháp, công cụ phát hiện botnet

Có nhiều giải pháp, công cụ phát hiện botnet đã được phát triển và triển khai ứng dụng trên thực tế. Mục này mô tả 3 công cụ giám sát, phát hiện botnet điển hình, gồm BotHunter, BotSniffer và BotTrack.

1.3. KHÁI QUÁT VỀ HỌC MÁY VÀ CÁC THUẬT TOÁN SỬ DỤNG

Phân loại nhị phân là nhiệm vụ phân loại các phần tử của một tập hợp các đối tượng ra thành 2 nhóm dựa trên cơ sở là một số thuộc tính nào đó (còn gọi là đặc trưng). Đây là kỹ thuật rất phù hợp đối với các vấn đề phát hiện truy cập bất hợp pháp, tấn công mạng,..

1.3.1. Giới thiệu về học máy

1.3.2. Một số thuật toán học máy có giám sát

Mục này trình bày một số thuật toán học máy có giám sát truyền thống được sử dụng trong các mô hình phát hiện botnet đề xuất trong Chương 2 và Chương 3 của luận án, bao gồm: Naïve Bayes, Cây quyết định, Rừng ngẫu nhiên, SVM và Hồi quy Logistic.

1.3.3. Các độ đo đánh giá

Để đánh giá khả năng phát hiện của các mô hình đề xuất trong các Chương 2 và Chương 3, luận án sử dụng sáu độ đo bao gồm: PPV, TPR, FPR, FNR, F1 và ACC.

Ngoài ra, luận án sử dụng tỷ lệ phát hiện (DR-Detection Rate) để đo lường hiệu quả của mô hình phát hiện đề xuất khi dự đoán tên miền của các DGA botnet khác nhau trong quá trình kiểm thử mô hình trong giai đoạn phát hiện. DR cho mỗi loại botnet được tính như sau:

$$DR = \frac{NoDB}{NoTest} \qquad (1.1)$$

trong đó, NoDB là số tên miền của một DGA botnet được dự đoán đúng và NoTest là tổng số tên miền của DGA botnet đó khi đưa vào kiểm tra.

1.4. CÁC TẬP DỮ LIỆU CHO PHÁT HIỆN BOTNET SỬ DỤNG

1.4.1. Tập dữ liệu Netlab360

Netlab 360 là tập dữ liệu chủ yếu được sử dụng trong luận án. Đây là bộ dữ liệu do Network Security Research Lab at 360 cung cấp công khai với hàng triệu mẫu từ nhiều họ DGA được thu thập từ các hệ thống mạng thực tế. Hệ thống phát hiện DGA botnet của Netlab 360 sàng lọc lượng dữ liệu khổng lồ và các mẫu phần mềm độc hại để tìm các DGA botnet đáng ngờ, mới nhất theo thời gian thực. Nguồn dữ liệu về các họ DGA botnet liên tục được cập nhật từ các cá nhân cũng như các tổ chức nghiên cứu về DGA botnet.

1.4.2. Các tập dữ liệu khác được sử dụng

Ngoài các dữ liệu về botnet từ hai tập dữ liệu đã trình bày ở trên, dữ liệu DGA botnet được bổ sung từ bộ sưu tập 33 DGA botnet của tác giả Johannes Bader (*bao gồm cả mã nguồn các thuật toán sinh*).

Để có được kết quả đánh giá một cách tổng quát, trong luận án sử dụng bộ dữ liệu UMUDGA của Universidad de Murcia. Bộ dữ liệu có hơn 30 triệu tên miền được tạo theo thuật toán được gắn nhãn thủ công sẵn sàng sử dụng cho phân tích học máy. Từ bộ dữ liệu này sẽ chọn ra một số họ botnet chưa được công bố trên Netlab360 để thử nghiệm phát hiện dựa ra các mô hình sẽ được đề xuất ở chương 2 và chương 3.

Tập dữ liệu các tên miền lành tính được lấy top 1 triệu tên miền của Alexa. Các tên miền được lược bỏ TLD, chỉ lấy phần SLD và loại bỏ các tên miền trùng nhau (*có TLD khác nhau*). Luận án sử dụng 110,000 tên miền đầu tiên có thứ hạng cao nhất trong tệp dữ liệu này để xây dựng và kiểm thử các mô hình phát hiện DGA botnet đề xuất.

1.5. HƯỚNG NGHIÊN CỨU CỦA LUẬN ÁN

1.5.1. Ưu điểm và nhược điểm của các kỹ thuật phát hiện botnet

Bảng 1.1 tổng hợp các ưu điểm và nhược điểm của các kỹ thuật phát hiện botnet.

Bảng 1.1: Ưu nhược điểm của các kỹ thuật phát hiện botnet

Kỹ thuật	Ưu điểm	Nhược điểm
Phát hiện dựa trên Honeynet	Đơn giản trong triển khai, ít yêu cầu về nguồn lực, chi phí triển khai tối thiểu và hữu dụng với dữ liệu mã hoá.	Khó mở rộng, nhiều thách thức khi giám sát các dạng botnet và các dạng tấn công có liên quan, có khả năng bị vô hiệu hóa.
Phát hiện dựa trên luật, dấu hiệu	Có khả năng phát hiện nhanh và chính xác các bot và botnet đã biết.	Không có khả năng phát hiện các bot và botnet mới, cần thường xuyên cập nhật cơ sở dữ liệu dấu hiệu, chữ ký.
Phát hiện dựa trên bất thường	Có khả năng phát hiện các dạng bot, botnet mới, có khả năng tự động hóa việc xây dựng mô hình phát hiện.	Tỷ lệ cảnh báo sai thường cao hơn so với phát hiện dựa trên luật, dấu hiệu; đòi hỏi tài nguyên tính toán lớn hơn cho xây dựng mô hình và giám sát phát hiện.

1.5.2. Các vấn đề giải quyết trong luận án

Từ việc phân tích mô hình hoạt động và các tác hại của các dạng botnet nói chung và DGA botnet nói riêng, việc nghiên cứu các giải pháp, kỹ thuật phát hiện botnet, DGA botnet là rất cấp thiết. Luận án cũng đã nghiên cứu, khảo sát các kỹ thuật phát hiện botnet dựa trên Honeynet, dựa trên dấu hiệu, luật, dựa trên bất thường và một số giải pháp, công cụ cho giám sát và phát hiện các dạng botnet. Mỗi phương pháp, giải pháp có các ưu điểm và nhược điểm riêng như đã chỉ ra trong mục 1.5.1.

Hướng nghiên cứu của luận án là sử dụng phương pháp phát hiện bot, botnet dựa trên bất thường do phương pháp này có khả năng phát hiện các dạng bot, botnet mới, đồng thời có khả năng tự động hóa việc xây dựng mô hình phát hiện. Trên cơ sở khảo sát, phân tích các ưu điểm và hạn chế của các đề xuất đã có, luận án tập trung nghiên cứu, giải quyết các vấn đề sau: (1) nghiên cứu, đề xuất tập đặc trưng phân loại tên miền mới phù hợp hơn cho xây dựng các mô hình phát hiện DGA botnet, nhằm tăng tỷ lệ phát hiện đúng và giảm tỷ lệ cảnh báo sai và (2) nghiên cứu, lựa chọn sử dụng phương pháp học máy phù hợp cho xây dựng các mô hình phát hiện DGA botnet, nhằm xây dựng một mô hình phát hiện thống nhất cho phép phát hiện hiệu quả nhiều dạng DGA botnet. Vấn đề (1) là do tập đặc trưng phân loại tên miền sử dụng trong các đề xuất đã có chưa thực sự phù hợp để phân biệt các tên miền DGA với các tên miền lành tính dẫn đến tỷ lệ cảnh báo sai còn tương đối cao. Vấn đề (2) xuất phát từ thực tế là mỗi đề xuất đã có chỉ có khả năng phát hiện hiệu quả một số họ DGA botnet, hoặc trên một tập dữ liệu cụ thể, mà không thể phát hiện hiệu quả nhiều dạng DGA botnet.

1.6. KẾT LUẬN CHƯƠNG

Botnet đã và đang trở thành một trong những mối đe dọa an ninh chính cho các cơ quan, tổ chức, doanh nghiệp và người dùng Internet. Do vậy, nghiên cứu phát triển các kỹ thuật và giải pháp hiệu quả cho giám sát, phát hiện botnet là việc cấp thiết. Chương 1 giới thiệu tổng quan về botnet, vấn đề phát hiện botnet, khái quát về học máy và các giải thuật học máy sử dụng cho phát hiện botnet và các tập dữ liệu sử dụng trong luận án. Cụ thể, trong phần đầu chương trình bày khái quát về botnet và phương thức hoạt động của chúng, phân loại botnet dựa trên kiến trúc mạng và giao thức truyền thông, vắn tắt về lịch sử phát triển của botnet và tác hại cũng như các dạng khai thác botnet.

Một trong các nội dung chính được trình bày trong chương này là vấn đề phát hiện botnet. Luận án phân tích 3 hướng phát hiện botnet được sử dụng phổ biến bao gồm: phát hiện dựa trên honeynet, phát hiện dựa trên luật, dấu hiệu và phát hiện dựa trên dựa trên bất thường, đồng thời tổng hợp các ưu và nhược điểm của 3 hướng trên làm cơ sở cho hướng nghiên cứu của luận án.

Trong hướng phát hiện botnet dựa trên dựa trên bất thường, việc ứng dụng học máy trong xây dựng các mô hình và giải pháp phát hiện botnet ngày càng được quan tâm do học máy có thể ứng dụng để tự động hóa việc xây dựng mô hình hoặc hồ sơ phát hiện. Điều này giúp giảm đáng kể yêu cầu nhân lực chuyên gia cho xây dựng tập luật, dấu hiệu theo phương pháp thủ công. Để phục vụ cho việc ứng dụng học máy trong các mô hình phát hiện botnet đề xuất trong chương 2 và chương 3, chương này trình bày khái quát về học máy, tập trung mô tả các thuật toán học máy có giám sát truyền thống. Chương cũng mô tả các độ đo đánh giá các mô hình phát hiện DGA botnet dựa trên học máy đề xuất trong luận án.

Phần tiếp theo của chương 1 trình bày về các tập dữ liệu sử dụng trong luận án, bao gồm tập dữ liệu Netlab 360, CTU-13 và tập dữ liệu tên miền lành tính từ nguồn Alexa. Đây là các tập dữ liệu tên miền do các DGA botnet sinh ra, được thu thập từ nhiều nguồn. Từ các tập dữ liệu gốc, luận án xây dựng tập dữ liệu chung, gồm tập tên miền DGA và tập tên miền lành tính sử dụng trong các mô hình phát hiện botnet dựa trên học máy đề xuất trong luận án ở các chương 2 và chương 3.

Phần cuối của chương 1 nêu 2 vấn đề chính được tập trung giải quyết trong các chương 2 và 3 của luận án.

CHƯƠNG 2: PHÁT HIỆN DGA BOTNET DỰA TRÊN HỌC MÁY SỬ DỤNG CÁC ĐẶC TRƯNG KÝ TỰ VÀ TỪ

2.1. DGA BOTNET VÀ CƠ CHẾ KHAI THÁC HỆ THỐNG DNS

2.1.1. Khái quát về DGA botnet

2.1.1.1. Giới thiệu về DGA botnet

DGA botnet là các họ botnet sử dụng kỹ thuật DGA (*Domain Generation Algorithm*) để sinh và đăng ký nhiều tên miền ngẫu nhiên khác nhau cho máy chủ chỉ huy và điều khiển CnC của chúng nhằm chống lại việc bị kiểm soát và đưa vào danh sách đen. Các botnet dạng này còn được gọi là DGA-based botnet, hay ngắn gọn là DGA botnet. Các DGA botnet sử dụng thuật toán DGA để định kỳ sinh và đăng ký một lượng lớn tên miền giả ngẫu nhiên mà chúng được phân giải thành địa chỉ IP của máy chủ CnC của botnet. Lý do chính của việc sử dụng DGA là làm phức tạp việc kiểm soát và thu hồi tên miền. Nếu botnet sử dụng một tên miền tĩnh cho máy chủ CnC của nó, việc kiểm soát và thu hồi tên miền có thể được thực hiện dễ dàng thông qua việc phối hợp với bên quản lý tên miền gốc để chỉnh sửa các bản ghi tên miền trên máy chủ DNS. Tuy nhiên, khi DGA được sử dụng để sinh các tên miền động, việc

kiểm soát và thu hồi các tên miền sẽ trở nên rất khó khăn. Do các bot sử dụng một tên miền mới được tạo ra sau một giai đoạn để kết nối đến máy chủ CnC, việc kiểm soát các tên miền đã hết hạn sử dụng không có ý nghĩa.

2.1.1.2. Các loại DGA botnet

Tập ký tự được sử dụng để tự động sinh tên miền DGA botnet có thể được chia thành 3 dạng chính: character-based DGA botnet, word-based DGA botnet và mixed DGA botnet.

2.1.2. Cơ chế DGA botnet khai thác hệ thống DNS

2.1.2.1. Giới thiệu hệ thống DNS

2.1.2.2. Cơ chế DGA botnet khai thác hệ thống DNS

2.2. PHÁT HIỆN DGA BOTNET DỰA TRÊN CÁC ĐẶC TRƯNG KÝ TỰ

2.2.1. Các phương pháp phát hiện DGA botnet

Có thể thấy các họ DGA botnet chiếm tỷ lệ lớn trong số các họ botnet đã và đang hoạt động và do vậy các phương phát hiện DGA botnet đã được đề xuất trong thời gian qua cũng rất phong phú. Có thể chia các phương phát hiện DGA botnet đã được đề xuất trong những năm qua thành 3 nhóm: phát hiện dựa trên phân tích truy vấn DNS, phát hiện dựa trên thống kê và phát hiện dựa trên học máy. Mục này phân tích các đề xuất nổi bật của từng nhóm và các ưu, nhược điểm của chúng.

2.2.1.1. Phát hiện DGA botnet dựa trên phân tích truy vấn DNS

2.2.1.2. Phát hiện DGA botnet dựa trên thống kê

2.2.1.3. Phát hiện DGA botnet dựa trên học máy

Bên cạnh thống kê, các kỹ thuật học máy đã và đang được ứng dụng rộng rãi trong phát hiện tấn công, xâm nhập nói chung và phát hiện botnet nói riêng. Ưu điểm của phát hiện DGA botnet dựa trên học máy là độ chính xác ngày càng được cải thiện cao và khả năng tự động xây dựng mô hình phát hiện từ tập dữ liệu huấn luyện. Mục này khảo sát một số đề xuất tiêu biểu cho phát hiện DGA botnet dựa trên các kỹ thuật học máy truyền thống, các kỹ thuật học sâu, cũng như phương pháp kết hợp giữa học máy truyền thống và học sâu.

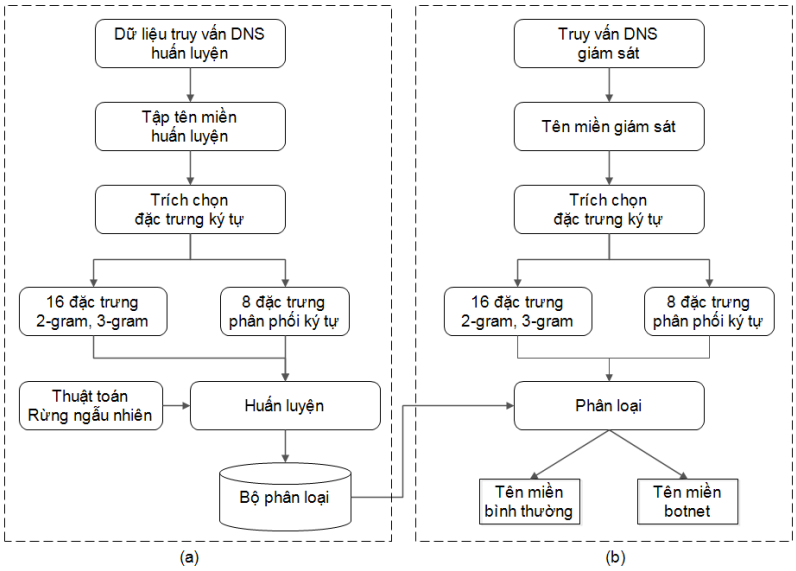
2.2.1.4. Ưu điểm và hạn chế của các đề xuất phát hiện DGA botnet

Kết quả khảo sát cho thấy, các giải pháp đề xuất phát hiện botnet dựa trên phân loại các tên miền hợp lệ và các tên miền được sinh tự động cho các máy chủ CnC của botnet là một trong các hướng đi hiệu quả trong phát hiện botnet do độ chính xác phát hiện cao, tỷ lệ cảnh báo sai thấp và yêu cầu chi phí tính toán không quá lớn. Đặc biệt, các đề xuất phát hiện DGA botnet dựa trên học máy là nhánh nghiên cứu rất có triển vọng do mô hình phát hiện có thể được xây dựng tự động từ dữ liệu huấn luyện. Đây cũng chính là nhánh nghiên cứu mà luận án lựa chọn thực hiện.

Mặc dù các nghiên cứu đã có cho phát hiện DGA botnet đã đạt được nhiều kết quả hứa hẹn, vẫn còn một số vấn đề cần tiếp tục nghiên cứu như: (1) do các tập đặc trưng lựa chọn cho phân loại tên miền và/hoặc giải thuật học máy sử dụng chưa thực sự phù hợp nên độ chính xác phát hiện chung đạt khoảng 90% và do vậy tỷ lệ phát hiện sai khoảng 10% vẫn là tương đối cao, điều này làm giảm khả năng triển khai ứng dụng trên thực tế; (2) một số đề xuất chỉ cho độ chính xác phát hiện cao với một tập dữ liệu cụ thể, hoặc một số họ DGA botnet, nhưng có khả năng phát hiện kém, hoặc không thể phát hiện một số họ DGA botnet khác; và (3) một số đề xuất đòi hỏi chi phí tính toán rất lớn cho quá trình xây dựng mô hình, cũng như quá trình giám sát phát hiện và điều này làm giảm khả năng triển khai ứng dụng trên các hệ thống mạng có lưu lượng lớn. Mô hình đề xuất phát hiện CDM trong mục này nhằm giải quyết vấn đề (1) nêu trên.

2.2.2. Giới thiệu mô hình phát hiện CDM

Hình 2.1 biểu diễn mô hình phát hiện character-based DGA botnet đề xuất (*CDM – Character-based DGA Botnet Detection Model*) dựa trên phân loại tên miền trích xuất từ dữ liệu truy vấn DNS. Mô hình CDM sử dụng tập đặc trưng ký tự có khả năng phân biệt hiệu quả giữa các tên miền lành tính và các tên miền character-based DGA - là các tên miền được sinh tự động sử dụng tổ hợp ngẫu nhiên các ký tự. Các đặc trưng ký tự gồm các đặc trưng được được trích xuất dựa trên quan sát, phân tích sự khác biệt giữa các tên miền lành tính và các tên miền character-based DGA ở mức ký tự mà không xem xét đến ngữ nghĩa của các cụm ký tự trong tên miền. Mô hình CDM được xây dựng trên cơ sở phân tích hoạt động của các botnet: botmaster sử dụng thuật toán DGA để sinh và đăng ký tự động các tên miền cho các máy chủ CnC nhằm lẩn trách bị phát hiện và đưa vào danh sách đen, đồng thời các bot định kỳ sử dụng thuật toán DGA để sinh tự động tên miền và sau đó truy vấn hệ thống DNS để tìm địa chỉ IP của máy chủ CnC theo tên miền sinh tự động. Từ dữ liệu truy vấn DNS các tên miền truy vấn được bóc tách và phân loại nhằm phát hiện các hoạt động của botnet trong hệ thống. Mô hình phát hiện CDM được thực hiện thành 2 giai đoạn: (a) giai đoạn huấn luyện và (b) giai đoạn phát hiện.



Hình 2.1: Mô hình phát hiện Character-based DGA botnet

2.2.3. Tập dữ liệu huấn luyện và kiểm thử

Để đánh giá hiệu năng của mô hình CDM, luận án sử dụng các tập dữ liệu tên miền đã được bóc tách và gán nhãn, bao gồm tập các tên miền lành tính và tập các tên miền độc hại do DGA botnet sinh và sử dụng. Danh sách các tên miền lành tính gồm 100,000 tên miền có thứ hạng cao nhất trong xếp hạng của Alexa. Danh sách các tên miền độc hại được thu thập tại, bao gồm 171.393 tên miền được sinh và sử dụng bởi 39 họ DGA botnet. Từ tập dữ liệu ban đầu, các tên miền sẽ được xử lý để bỏ đi phần tên miền mức cao nhất (TLD) chỉ lấy phần tên miền thứ cấp (SLD). Ví dụ, với tên miền "example.com", sau khi qua xử lý, dữ liệu thu được sẽ là "example". Bảng 2.3 là danh sách 39 họ DGA botnet được lựa chọn để huấn luyện và kiểm thử, trong đó: (1) tập huấn luyện: bao gồm 100,000 tên miền lành tính và 100.000 tên miền của 13 họ DGA botnet; (2) tệp kiểm thử: bao gồm 71.393 tên miền của 39 họ DGA botnet. Ngoài ra, để đánh giá khả năng phát hiện botnet mới của mô hình CDM, luận án sử dụng tập dữ liệu UMUDGA gồm 7 họ DGA botnet, không xuất hiện trong tập huấn luyện.

2.2.4. Tiền xử lý dữ liệu

2.2.4.1. Giới thiệu

Các tên miền do botnet sinh tự động thường có các đặc trưng DNS, đặc trưng mạng và đặc trưng ngữ nghĩa khác biệt so với các tên miền thông thường. Nghiên cứu đề xuất sử dụng

kỹ thuật phân tích phân bố các nguyên âm, chữ số và các ký tự khác để phân biệt các tên miền hợp lệ và các tên miền sinh bằng thuật toán của botnet. Mở rộng hơn, đề xuất sử dụng 2 nhóm đặc trưng của tên miền, gồm các đặc trưng DNS (địa chỉ IP, địa chỉ mạng, quốc gia, TTL,...) và các đặc trưng từ vựng (phân bố các ký tự của tên miền). Trong khi đó, đề xuất sử dụng 36 đặc trưng trong 2 nhóm, gồm 18 đặc trưng từ vựng (trung bình, phương sai và độ lệch chuẩn của 1-gram, 2-gram, 3-gram và 4-gram, entropy, các đặc trưng ký tự, số, nguyên âm, phụ âm) và 18 đặc trưng mạng (TTL, số lượng địa chỉ mạng,...). Kế thừa từ công bố trước đây của nhóm nghiên cứu [24], luận án tập trung khai thác các đặc trưng thống kê từ vựng dựa trên các cụm 2-gram và 3-gram, và các đặc trưng phân bố các dạng ký tự trong tên miền. Cụ thể, mô hình CDM đề xuất sử dụng 24 đặc trưng mức ký tự cho mỗi tên miền, bao gồm:

- Đặc trưng n-gram gồm 16 đặc trưng thống kê cho các cụm 2-gram và 3-gram;
- Đặc trưng loại ký tự gồm 6 đặc trưng phân bố nguyên âm, ký tự, chữ số;
- Đặc trưng thống kê gồm 2 đặc trưng entropy theo ký tự và giá trị kỳ vọng của tên miền.

2.2.4.2. Các đặc trưng n-gram

2.2.4.3. Các đặc trưng loại ký tự

2.2.4.4. Các đặc trưng thống kê

2.2.5. Thử nghiệm và kết quả

2.2.5.1. Kịch bản thử nghiệm

Tập dữ liệu huấn luyện gồm 200,000 tên miền được sử dụng để xây dựng và kiểm tra hiệu suất của mô hình CDM sử dụng thuật toán máy học rừng ngẫu nhiên (37-trees). Luận án sử dụng phương pháp kiểm tra chéo 10 lần (10-fold cross-validation) với 80% tập dữ liệu lấy ngẫu nhiên cho huấn luyện và 20% còn lại cho kiểm tra để tính kết quả trung bình hiệu suất phát hiện của mô hình đề xuất. Kết quả sẽ được so sánh với các đề xuất trước đây để đánh giá hiệu suất của mô hình.

Để so sánh và chứng minh hiệu quả của mô hình CDM với 24 đặc trưng ký tự, tập huấn luyện sẽ được sử dụng để kiểm tra hiệu suất của mô hình sử dụng 18 đặc trưng được đề xuất bởi Hoang và cộng sự.

Mô hình CDM sau huấn luyện được sử dụng cho thử nghiệm phát hiện sử dụng tập 71,393 tên miền DGA boetnet sinh bởi 39 họ lấy tại Netlab360 và 31,000 tên miền DGA botnet sinh ra bởi 7 họ tại UMUDGA (7 họ này không được công bố trong Netlab 360 - Để tính toán tỷ lệ phát hiện (DR) cho mỗi họ botnet cũng như tỷ lệ phát hiện chung trên toàn tập kiểm thử. Ngoài giá trị DR của mỗi họ tên miền, giá trị DR chung là tỷ lệ của tổng số tên miền phát hiện chính xác và tổng số tên miền thử nghiệm.

2.2.5.2. Kết quả thử nghiệm

Bảng 2.1 cho thấy, hiệu suất của CDM tốt hơn so với hiệu suất của Hoang và cộng sự khi sử dụng cùng tập huấn luyện với F1 và ACC lần lượt là 99.60% và 99.60% so với 94.60% và 94.61%. Tỷ lệ dương tính giả và âm tính giả của CDM cũng giảm đáng kể, cụ thể là tỷ lệ dương tính giả và âm tính giả của CDM lần lượt là 0.43% và 0.38% so với 5.13% và 5.67% của mô hình đề xuất bởi Hoang và cộng sự. Như vậy, có thể khẳng định tập 24 đặc trưng ký tự sử dụng trong CDM có khả năng phân loại các tên miền DGA tốt hơn so với tập 18 đặc trưng trong Hoang và cộng sự.

Bảng 2.1: Hiệu suất của mô hình CDM so với Hoang và cộng sự [24]

Mô hình phát hiện	PPV	TPR	FPR	FNR	ACC	F1
Hoang và cộng sự	94.87	94.33	5.13	5.67	94.60	94.61
CDM	99.57	99.62	0.43	0.38	99.60	99.60

Bảng 2.2 so sánh hiệu suất của mô hình CDM với hiệu suất của các mô hình phát hiện đã có. Có thể thấy mô hình CDM cho hiệu suất tốt hơn đáng kể so với các đề xuất của Truong và cộng sự, Hoang và cộng sự, Qiao và cộng sự, Zhao và cộng sự.

Các Bảng 2.3, Bảng 2.4 và

Bảng 2.5 cung cấp tỷ lệ phát hiện (DR) của mô hình CDM trong giai đoạn phát hiện trên tập tên miền DGA botnet sinh bởi 39 họ DGA botnet chia tương ứng thành 3 nhóm: nhóm có DR>=90%, nhóm có 90% > DR>=50% và nhóm có DR < 50%.

Bảng 2.2: Hiệu suất của mô hình CDM so với các mô hình trước đó

Mô hình phát hiện	**PPV**	**TPR**	**FPR**	**FNR**	**ACC**	**F1**
Truong và cộng sự	94.70		4.80		92.30	
Hoang và cộng sự	90.70	91.00	9.30		90.90	90.90
Qiao và cộng sự	95.05	85.14				94.58
Zhao và cộng sự			6.14	7.42	94.04	
Mô hình đề xuất CDM	99.57	99.62	0.43	0.38	99.60	99.60

Bảng 2.3: Các họ botnet có tỷ lệ phát hiện (DR) lớn hơn 90%

STT	**Họ DGA botnet**	**Tổng số tên miền**	**Phát hiện chính xác**	**DR%**
1	emotet	4000	3987	99.68
2	gameover	4000	4000	100.00
3	murofet	4000	3992	99.80
4	necurs	4000	3974	99.35
5	pykspa_v1	4000	3988	99.70
6	ramnit	4000	3982	99.55
7	ranbyus	4000	3983	99.58
8	rovnix	4000	4000	100.00
9	shiotob	4000	3987	99.68
10	symmi	1200	1159	96.58
11	tinba	4000	3999	99.98
12	simda	4000	3986	99.65
13	virut	4000	3990	99.75
14	proslikefan	100	98	98.00
15	tempedreve	195	190	97.44
16	tinynuke	32	32	100.00
17	vidro	100	100	100.00
18	pykspa_v2_real	199	197	98.99
19	pykspa_v2_fake	799	790	98.87
20	padcrypt	168	165	98.21
21	nymaim	480	455	94.79
22	vawtrak	827	799	96.61
23	shifu	2546	2510	98.59
24	fobber_v1	298	298	100.00
25	fobber_v2	299	299	100.00
26	dircrypt	762	757	99.34
27	cryptolocker	1000	997	99.70
28	locky	1158	1147	99.05
29	chinad	1000	1000	100.00
30	qadars	2000	1981	99.05
31	dyre	1000	1000	100.00
Tổng		**62163**	**61842**	**99.48**

Bảng 2.4: Các họ botnet có tỷ lệ phát hiện (DR) từ 50%-90%

STT	**Họ DGA botnet**	**Tổng số tên miền**	**Phát hiện chính xác**	**DR%**
1	mydoom	50	44	88.00
2	gspy	100	76	76.00
3	enviserv	500	252	50.40
4	conficker	495	442	89.29
	Tổng cộng	**1145**	**814**	**71.09**

Bảng 2.5: Các họ botnet có tỷ lệ phát hiện thấp

STT	Họ DGA botnet	Tổng số tên miền	Phát hiện chính xác	DR%
1	banjori	4000	0	0
2	matsnu	881	107	12.15
3	bigviktor	999	111	11.11
4	suppobox	2205	425	19.27
	Tổng cộng	**8085**	**643**	**7.95**

Bảng 2.6: Tỷ lệ phát hiện của CDM trên tập dữ liệu UMUDGA

STT	Họ DGA botnet	Tổng số tên miền	Phát hiện chính xác	DR%
1	alureon	5000	4911	98.22
2	bedep	5000	4991	99.82
3	corebot	5000	4988	99.76
4	kraken	2000	1968	98.40
5	pushdo	5000	4718	94.40
6	zeus	5000	5000	100.00
		27000	**26576**	**98.43**
7	pizd	4000	642	16.05
	Tổng cộng	**31000**	**27218**	**87.85**

2.2.6. Đánh giá

Dựa vào kết quả thử nghiệm ở các Bảng 2.2, Bảng 2.3, Bảng 2.4 và

Bảng 2.5, có thể rút ra những nhận xét sau: Mô hình phát hiện CDM hoạt động tốt hơn các đề xuất trước đó với tất cả các độ đo, trong đó mô hình đề xuất cho độ chính xác và độ đo F1 cao hơn đáng kể so với các mô hình trước đó. Chẳng hạn, độ đo F1 của Hoang và cộng sự [24], Qiao và cộng sự [69] và mô hình phát hiện CDM đề xuất tương ứng là 90.90%, 94.58% và 99.59%. Ngoài ra, tỷ lệ dương tính giả (FPR) và tỷ lệ âm tính giả (FNR) của mô hình phát hiện CDM đề xuất cũng thấp hơn đáng kể so với các mô hình trước đó, như thể hiện trên Bảng 2.2.

Thông qua phát hiện thử nghiệm trên 39 họ botnet cho thấy, mô hình CDM có khả năng phát hiện hiệu quả hầu hết các họ DGA botnet. Trong số 39 DGA botnet, 31 họ DGA botnet được phát hiện với tỷ lệ phát hiện trên 90%, như trình bày trong Bảng 2.3. Tỷ lệ phát hiện trung bình của nhóm DGA botnet này là 99.48%. Bốn DGA botnet trong nhóm thứ hai, như trong Bảng 2.4 cũng có DR trung bình tương đối cao là 71.09%. Lý do mà mô hình CDM đề xuất hoạt động tốt trong phát hiện tên miền DGA botnet thuộc các nhóm này là bởi tập 24 đặc trưng ký tự đề xuất trong mô hình là phù hợp cho phân biệt các tên miền character-based DGA và các tên miền lành tính.Bảng 2.6 thể hiện tỷ lệ phát hiện của mô hình CDM với bộ dữ liệu UMUDGA. Theo đó, có thể thấy rằng với 06 họ character-based DGA botnet, tỷ lệ phát hiện đạt 98.43%. Đây là những botnet không được công bố trong tập dữ liệu Netlab 360, không được sử dụng để huấn luyện mô hình, điều này khẳng định CDM có thể phát hiện hiệu quả các character-based DGA botnet mới. Mô hình CDM không thể phát hiện bất kỳ tên miền nào được tạo bởi botnet ‘banjori’ và chỉ có thể phát hiện một số tên miền được tạo bởi botnet ‘matsnu’, ‘bigviktor’, ‘suppobox’ và ‘pizd’. Điều này là do các DGA botnet này sử dụng các thuật toán DGA có thể tạo ra các tên miền rất giống với các tên miền lành tính. Đây cũng là vấn đề sẽ được giải quyết trong các mô hình phát hiện trong mục tiếp theo của luận án.

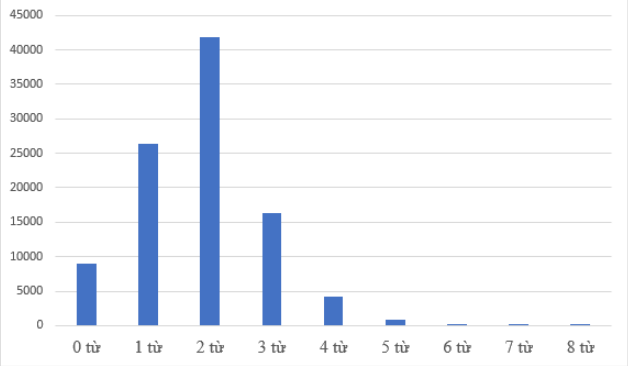
2.3. PHÁT HIỆN WORD-BASED DGA BOTNET

2.3.1. Đặt vấn đề

Mục 2.1.1.2 đã trình bày sơ lược về các dạng DGA botnet, trong đó có word-based và mixed DGA botnet (từ đây luận án gọi chung là word-based DGA botnet). Mục này đi sâu phân tích các đặc điểm của hai dạng botnet này. Khác với character-based DGA botnet, word-

based DGA botnet sinh các tên miền bằng cách tổ hợp các từ tiếng Anh lấy từ các danh sách các từ được lập sẵn. Các tên miền DGA dạng này thường chứa hai hoặc ba từ được lấy các danh sách từ khác nhau được chọn và nối ngẫu nhiên. Cuối cùng, một TLD được thêm vào cuối giống như một tên miền thông thường. Theo các nhà nghiên cứu bảo mật, các họ DGA botnet tiến hóa, như Matsnu sử dụng một kỹ thuật thông minh để tránh các cơ chế kiểm tra thông thường. Tên miền do Matsnu tạo ra có thể sử dụng các ký tự như "-" để nối các từ (như world-bite-care.com), hoặc không dùng ký tự nối (như activitypossess.com).

Khó khăn lớn nhất cho phát hiện các word-based DGA botnet là chúng có khả năng sinh các tên miền được tổ hợp từ các từ tiếng Anh có nghĩa và *các tên miền này rất giống so với các tên miền lành tính đang được sử dụng rộng rãi*. Điều này được chứng minh trong Yang và cộng sự khi phân tích một triệu tên miền hàng đầu, đã phát hiện ra rằng hơn 67% tên miền chứa ít nhất một từ tiếng Anh và gần 30% tên miền hoàn toàn bao gồm các từ tiếng Anh. Và theo thống kê trong dữ liệu thực nghiệm với 98,866 tên miền lành tính nguyên tố có thứ hạng cao nhất, thì số tên miền lành tính không có từ tiếng Anh chỉ chiếm 9.05%, có 1 từ chiếm 26.70% và có 2 từ chiếm tới 42.34%, như minh họa trên Hình 2.2.



Hình 2.2: Biểu đồ phân bố các tên miền với số lượng từ tương ứng

Do các tên miền do các word-based DGA botnet, như bigviktor, matsnu, ngioweb và suppobox sinh ra rất giống các tên miền lành tính, nhiều phương pháp phát hiện DGA botnet dựa trên phân loại tên miền, như Hoang và cộng sự và mô hình CDM đề xuất ở mục 2.2 của luận án hầu như không thể phát hiện các DGA botnet này. Cụ thể như, mô hình CDM đề xuất sử dụng tập đặc trưng ký tự có khả năng phát hiện các character-based DGA botnet với tỷ lệ phát hiện đúng bình quân là 99.48%, nhưng không thể phát hiện tên miền nào của banjori và chỉ có thể phát hiện một số tên miền được tạo bởi matsnu, bigviktor và suppobox. Như vậy, có thể thấy cần nghiên cứu phát triển các mô hình sử dụng các tập đặc trưng phù hợp hơn cho phép phát hiện hiệu quả các word-based DGA botnet.

2.3.2. Các phương pháp phát hiện word-based DGA botnet

Mục này khảo sát một số nghiên cứu phát hiện word-based và mixed DGA botnet, bao gồm các đề xuất phát hiện dựa trên học máy có giám sát truyền thống và các đề xuất phát hiện dựa trên học sâu.

2.3.2.1. Phát hiện dựa trên học máy có giám sát truyền thống

2.3.2.2. Phát hiện dựa trên học sâu

2.3.2.3. Ưu điểm và hạn chế của các đề xuất phát hiện word-based DGA botnet

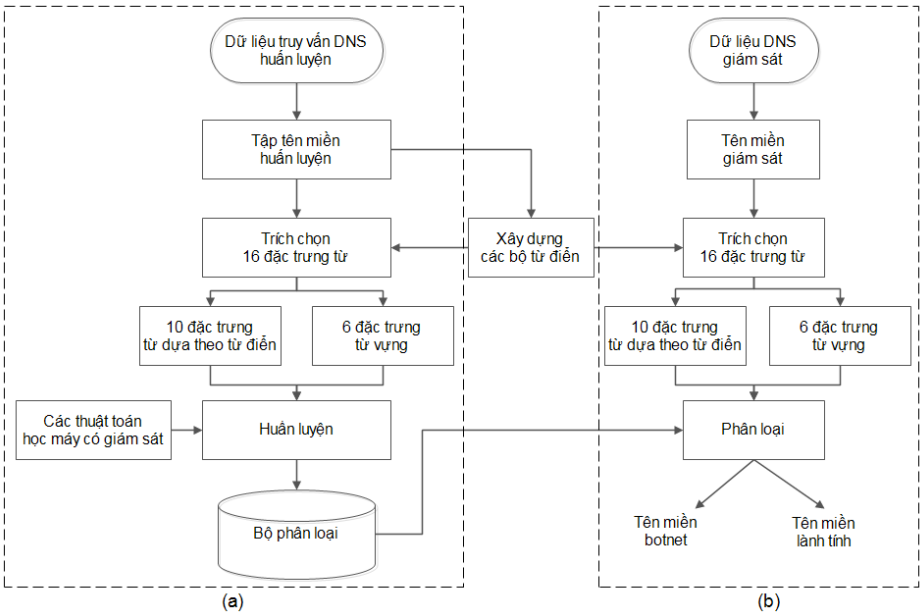
Có thể thấy, hầu hết các đề xuất phát hiện DGA botnet dựa trên các kỹ thuật học máy truyền thống chủ yếu tập trung phát hiện các họ character-based DGA botnet. Chúng không có khả năng phát hiện hoặc không hiệu quả trong phát hiện các họ word-based DGA botnet.

Với các hướng phát hiện DGA botnet dựa trên học sâu, ưu điểm chính của các nghiên cứu theo hướng này là có độ chính xác phát hiện DGA botnet và tính linh hoạt cao. Một số nghiên cứu đi sâu phân tích các đặc điểm ngữ nghĩa của tên miền, bao gồm phân phối theo

từ, phân bố theo ký tự và mối tương quan đã cải thiện đáng kể kết quả phân loại các tên miền thuộc họ word-based DGA. Hạn chế chủ yếu của các bộ phân loại dựa trên học sâu là phần lớn không thực sự hiệu quả trong phân biệt các tên miền word-based DGA, mặc dù các đề xuất này khá thành công đối với các tên miền character-based DGA. Một số mô hình có khả năng hoạt động tốt trên các bộ dữ liệu thử nghiệm khác nhau, nhưng hiệu suất có thể bị ảnh hưởng khi cố gắng tổng quát hóa thành các họ DGA mới hoặc các phiên bản mới của các họ đã biết. Ngoài ra, các đề xuất dựa trên học sâu thường yêu cầu lớn hơn về tài nguyên tính toán, thời gian huấn luyện và phát hiện đều dài hơn so với các mô hình dựa trên học máy truyền thống. Các mục tiếp theo trình bày mô hình phát hiện WDM sử dụng tập 16 đặc trưng từ mới, cho phép phát hiện hiệu quả các họ word-based DGA botnet.

2.3.3. Giới thiệu mô hình WDM

Mô hình phát hiện word-based DGA botnet (WDM), như biểu diễn trên Hình 2.3 bao gồm hai giai đoạn: (a) giai đoạn huấn luyện và (b) giai đoạn phát hiện. Trong giai đoạn huấn luyện, mô hình được xây dựng từ dữ liệu huấn luyện. Trong giai đoạn phát hiện, mô hình đã xây dựng được sử dụng để phân loại từng tên miền giám sát nếu nó là tên miền lành tính hoặc tên miền DGA botnet.



Hình 2.3:Mô hình phát hiện word-based DGA botnet

2.3.4. Tập dữ liệu thử nghiệm

Tập dữ liệu được sử dụng bao gồm ba tập con như sau: (i) tập con gồm 48,000 tên miền lành tính được trích xuất từ một triệu tên miền top Alexa; (ii) Tập con gồm 64,000 tên miền word-based DGA botnet được tạo bằng tập lệnh DGA cho 4 họ word-based DGA botnet điển hình, bao gồm bigviktor, matsnu, suppobox và pizd. Trong đó 48,000 tên miền của tập con này được sử dụng để huấn luyện và kiểm tra chéo các mô hình phát hiện và 16,000 tên miền được sử dụng để kiểm thử phát hiện; và (iii) Tập con gồm 63,905 tên miền DGA được tạo bởi 16 họ botnet DGA thu thập từ Netlab360. Trong đó 48,000 tên miền của tập con này được sử dụng để huấn luyện và kiểm tra chéo các mô hình phát hiện và 15,905 tên miền được sử dụng để kiểm thử phát hiện. Ngoài ra, sử dụng thêm tập dữ liệu gồm 31,000 tên miền DGA botnet thu thập từ UMUDGA.

Từ 3 tập dữ liệu con trên, tạo 2 tập dữ liệu DATASET-01 và DATASET-02 cho các kịch bản thử nghiệm khác nhau. Tập DATASET-01 được sử dụng để đánh giá khả năng phát hiện các word-based DGA botnet của mô hình WDM. Tập DATASET-01 bao gồm (i) tập huấn luyện gồm 48,000 tên miền lành tính và 48,000 tên miền word-based DGA, và (ii) tập kiểm thử phát hiện gồm 16,000 tên miền word-based DGA.

DATASET-02 được sử dụng để đánh giá khả năng phát hiện các loại DGA botnet, bao gồm cả word-based DGA botnet và character-based DGA botnet của mô hình WDM. Tập DATASET-02 bao gồm (i) tập huấn luyện gồm 48,000 tên miền lành tính và 48,000 tên miền DGA và (ii) tập kiểm thử phát hiện gồm 15,905 tên miền DGA..

2.3.5. Tiền xử lý dữ liệu

2.3.5.1. Giới thiệu

Dựa vào các đặc điểm của các họ word-based DGA botnet, 16 đặc trưng từ được trích xuất cho mỗi tên miền trong cả giai đoạn huấn luyện và phát hiện. Các đặc trưng này được đặt tên là f1, f2, f3, ..., f16. 10 đặc trưng được đề xuất mới trong mô hình WDM.

2.3.5.2. Trích chọn các đặc trưng

2.3.6. Thử nghiệm và kết quả

2.3.6.1. Kịch bản thử nghiệm

Các thử nghiệm được thực hiện theo các kịch bản sau:

Kịch bản 1: Huấn luyện và kiểm tra chéo mô hình phát hiện sử dụng "Tập huấn luyện" thuộc tập DATASET-01. Các thuật toán học máy có giám sát, bao gồm Naïve Bayes (*NB*), cây quyết định, rừng ngẫu nhiên, hồi quy Logistic và SVM được sử dụng theo trình tự để xây dựng các mô hình phát hiện, trong đó 80% dữ liệu được sử dụng để huấn luyện mô hình và 20% dữ liệu được sử dụng để kiểm tra chéo. Thuật toán rừng ngẫu nhiên sử dụng với 35 cây

Kịch bản 2: Kiểm thử mô hình phát hiện được xây dựng trong kịch bản 1 bằng cách sử dụng "Tập kiểm thử" của tập DATASET-01. Mục đích của kịch bản này là tìm tỷ lệ phát hiện (*DR*) của mô hình trên một số word-based DGAbotnet.

Kịch bản 3: Huấn luyện và kiểm tra chéo mô hình phát hiện bằng cách sử dụng 'Tập huấn luyện' của DATASET-02. Các thuật toán NB, J48 tree, RF-35, hồi quy logistic và SVM được sử dụng theo trình tự để xây dựng các mô hình phát hiện, trong đó 80% dữ liệu được sử dụng để huấn luyện xây dựng mô hình và 20% dữ liệu được sử dụng để kiểm tra chéo.

Kịch bản 4: Kiểm thử các mô hình phát hiện được xây dựng trong kịch bản 3 bằng cách sử dụng 'Tập thử nghiệm' của DATASET-02. Mục đích của kịch bản này là để tìm tỷ lệ phát hiện (*DR*) của mô hình trên các DGA botnet điển hình, bao gồm cả word-based DGA botnet và character-based DGA botnet.

2.3.6.2. Kết quả

Bảng 2.7 trình bày hiệu suất phát hiện của mô hình đề xuất dựa trên 5 thuật toán học máy sử dụng "Tập huấn luyện" của DATASET-01. Các độ đo hiệu suất trên bảng này xác nhận rằng mô hình đề xuất hoạt động rất tốt trên DATASET-01 với tất cả 5 thuật toán học máy. Mô hình được xây dựng từ 'Tập huấn luyện' của DATADET-01 cũng cho tỷ lệ phát hiện cao đối với tất cả 4 word-based DGAbotnet, như được hiển thị trong

Bảng 2.7: Hiệu suất phát hiện của mô hình sử dụng DATASET-01 (%)

Thuật toán	PPV	TPR	FPR	FNR	ACC	F1
NB	98.47	91.16	1.64	8.84	94.48	94.67
J48	*98.25*	*95.81*	*1.78*	*4.19*	*96.99*	*97.01*
RF-35	97.27	95.95	2.74	4.05	96.60	96.61
Logistic	98.63	92.97	1.45	7.03	95.60	95.71
SVM	98.70	93.73	1.36	6.27	96.07	96.15

Bảng 2.8: Tỷ lệ phát hiện (DR) của mô hình sử dụng DATASET-01 (%)

Thuật toán / Botnet	NB	J48	RF-35	Logistic	SVM
Bigviktor	96.35	96.78	95.28	96.88	97.08
Matsnu	99.13	97.78	97.55	99.10	99.03
Pizd	98.98	98.63	97.50	98.98	98.98
Suppobox	99.48	99.30	96.93	99.48	99.48

Trung bình	**98.51**	**98.19**	**96.81**	**98.63**	**98.66**

Bảng 2.9: Hiệu suất phát hiện của mô hình sử dụng DATASET-02 (%)

Thuật toán	PPV	TPR	FPR	FNR	ACC	F1
NB	65.30	89.13	27.18	10.78	78.77	75.38
J48	96.89	94.62	3.15	5.38	95.71	95.75
RF-35	96.02	94.78	3.99	5.22	95.39	95.40
Logistic	88.34	90.47	11.29	9.53	89.57	89.40
SVM	88.79	90.15	10.94	9.85	89.59	89.47

Bảng 2.10: Tỷ lệ phát hiện (DR) của mô hình (%) sử dụng DATASET-02

Thuật toán Botnet	NB	J48	RF-35	Logistic	SVM
Bigviktor	77.80	70.70	67.70	88.60	90.00
Matsnu	60.33	98.34	94.59	78.01	81.99
Pizd	8.40	97.90	99.40	73.60	75.70
Suppobox	7.30	99.10	97.70	94.10	97.30
Flubot	73.90	99.20	99.10	96.00	96.10
Necurs	53.40	91.70	90.20	83.10	83.10
Ramnit	51.30	92.10	91.20	84.50	84.50
Ranbyus	72.80	98.00	97.20	94.60	94.90
Rovnix	100.00	99.30	99.60	99.30	99.40
Tinba	27.40	98.90	97.60	61.50	91.40
Cryptolocker	48.50	96.70	95.80	91.80	92.20
Dyre	100.00	100.00	100.00	100.00	100.00
Emotet	96.50	99.40	99.10	97.40	97.70
Gameover	100.00	99.80	99.80	99.90	99.90
Murofet	84.00	99.50	99.70	99.00	99.00
Shiotob	74.60	95.20	94.80	84.00	85.00
Trung bình	**64.82**	**95.98**	**95.32**	**91.14**	**91.92**

Bảng 2.10 thể hiện hiệu suất phát hiện của mô hình đề xuất dựa trên 5 thuật toán học máy sử dụng ‘Tập huấn luyện’ của DATASET-02. Các độ đo hiệu suất trên bảng này cũng xác nhận rằng mô hình đề xuất hoạt động tương đối tốt trên DATASET-02 với tất cả 5 thuật toán học máy. Mô hình được xây dựng từ ‘Tập huấn luyện’ của DATASET-02 cũng tạo ra tỷ lệ phát hiện khá tốt đối với hầu hết các DGA botnet, như trình bày trong Bảng 2.10.

2.3.7. Đánh giá

Từ kết quả thực nghiệm cho trong Bảng 2.7, Bảng 2.8, Bảng 2.9, Bảng 2.10, có thể rút ra các nhận xét sau: Mô hình phát hiện WDM mang lại hiệu suất cao trên DATASET-01 với độ chính xác phát hiện tổng thể (ACC) và độ đo F1 trên 95% sử dụng 5 thuật toán học máy. Trong đó, thuật toán cây quyết định J48 hoạt động tốt nhất với tỷ lệ phát hiện cao nhất và tỷ lệ cảnh báo sai thấp nhất, như được hiển thị trong Bảng 2.7. Tỷ lệ phát hiện của 4 word-based DGA botnet điển hình biểu diễn trong Bảng 2.8 cũng xác nhận rằng mô hình WDM có khả năng phát hiện hiệu quả các word-based DGA botnet. Điều này có nghĩa là 16 đặc trưng từ sử dụng là phù hợp cho việc phân loại tên miền word-based DGA và tên miền lành tính.

Mô hình phát hiện WDM cũng tạo ra hiệu suất khá tốt trên DATASET-02 với độ chính xác phát hiện tổng thể (ACC) và độ đo F1 trên 95% sử dụng thuật toán cây quyết định và rừng ngẫu nhiên. Trong khi các mô hình dựa trên hồi quy logistic và SVM đạt được độ chính xác phát hiện tổng thể (ACC) và độ đo F1 trên 89%, thì mô hình dựa trên Naïve Bayes chỉ đạt độ đo F1 khoảng 75%, như trình bày trong

Bảng 2.9. Tỷ lệ phát hiện 4 word-based DGA botnet và 12 character-based DGA botnet được hiển thị trong Bảng 2.10 xác nhận rằng mô hình dựa trên cây quyết định J48 hoạt động

tốt trên hầu hết các botnet thử nghiệm, ngoại trừ ‘Bigviktor’. Mặc dù mô hình dựa trên SVM có tỷ lệ phát hiện trên ‘Bigviktor’ cao hơn so với mô hình dựa trên J48, tuy nhiên mô hình dựa trên J48 có tỷ lệ phát hiện trên hầu hết các botnet tốt hơn đáng kể so với mô hình dựa trên SVM.

Bảng 2.11 hiển thị so sánh hiệu suất phát hiện giữa mô hình WDM và các đề xuất phát hiện DGA botnet khác và Bảng 2.12 so sánh tỷ lệ phát hiện của 16 word-based DGA botnet và character-based DGA botnet giữa mô hình WDM dựa trên cây quyết định J48 và mô hình phát hiện CDM. Từ kết quả trình bày trong hai bảng này, có thể rút ra các nhận xét sau: (i) Mô hình WDM hoạt động tốt hơn nhiều so với các đề xuất phát hiện DGA botnet khác; (ii) Mô hình WDM có khả năng phát hiện hiệu quả các character-based DGA botnet, mặc dù có tỷ lệ phát hiện thấp hơn so với mô hình CDM; (iii) Mặc dù mô hình CDM đạt hiệu suất tốt hơn so với mô hình WDM trên các character-based DGA botnet, nhưng mô hình CDM lại hầu như không thể phát hiện các word-based DGA botnet. Trong khi đó, mô hình WDM có khả năng phát hiện hiệu quả các word-based DGA botnet, gồm Matsnu, Pizd và Suppobox.

Bảng 2.11: Hiệu suất phát hiện của WDM so với các đề xuất khác (%)

Mô hình phát hiện	PPV	TPR	FPR	FNR	ACC	F1
Truong và cộng sự	94.70		4.80		92.30	
Hoang và cộng sự	90.70	91.00	9.30		90.90	90.90
Qiao và cộng sự	95.05	95.14				94.58
Zhao và cộng sự			6.14	7.42	94.04	
CDM	99.57	99.62	0.43	0.38	99.60	99.60
WDM (DATASET-01)	98.25	95.81	1.78	4.19	96.99	97.01
WDM (DATASET-02)	96.89	94.62	3.15	5.38	95.71	95.75

Bảng 2.12: So sánh tỷ lệ phát hiện của 2 mô hình WDM và CDM

Bonet	WDM	CDM
Bigviktor	70.70	3.00
Matsnu	98.34	1.14
Pizd	97.90	
Suppobox	99.10	0.95
Flubot	99.20	
Necurs	91.70	98.67
Ramnit	92.10	97.20
Ranbyus	98.00	99.82
Rovnix	99.30	100.00
Tinba	98.90	98.77
Cryptolocker	96.70	99.00
Dyre	100.00	98.00
Emotet	99.40	99.85
Gameover	99.80	100.00
Murofet	99.50	99.85
Shiotob	95.20	99.55

2.4. KẾT LUẬN CHƯƠNG

Chương này giới thiệu chi tiết về DGA botnet, các loại DGA botnet và cơ chế các DGA botnet khai thác hệ thống DNS để duy trì hoạt động. Nhờ khả năng sinh và gán tên miền, địa chỉ IP tự động cho máy chủ CnC, đồng thời với khả năng tự động sinh và truy vấn các tên miền bởi các bot, các DGA botnet có khả năng lẩn tránh rà quét và kéo dài thời gian tồn tại của mình. Mặc dù vậy, do các bot trong botnet thường xuyên tương tác với hệ thống DNS, việc phân tích các truy vấn DNS có thể giúp phát hiện sự tồn tại của các bot và botnet.

Phần tiếp theo của chương 2 khảo sát các phương pháp phát hiện DGA botnet theo 3 nhóm: phát hiện dựa trên phân tích truy vấn DNS, phát hiện dựa trên thống kê và phát hiện dựa trên học máy. Mặc dù có nhiều ưu điểm, nhưng hạn chế lớn nhất của các phương pháp

đã có là tập đặc trưng phân loại tên miền được lựa chọn chưa thực sự phù hợp dẫn đến tỷ lệ phát hiện sai còn tương đối cao (đến 10%). Để khắc phục vấn đề này, chương 2 phát triển mô hình CDM cho phát hiện DGA botnet với mục tiêu là cải thiện tỷ lệ phát hiện đúng và giảm cảnh báo sai. Mô hình CDM đề xuất sử dụng 24 đặc trưng ký tự cho phân loại tên miền DGA với tên miền lành tính, trong đó kế thừa 18 đặc trưng từ và bổ sung 6 đặc trưng mới. Thuật toán học máy Rừng ngẫu nhiên được lựa chọn để xây dựng mô hình phát hiện. Các thử nghiệm trên tập dữ liệu gồm 100,000 tên miền lành tính và 153,000 tên miền DGA cho thấy, mô hình CDM đạt các độ đo đánh giá vượt trội so với các mô hình đã có. Cụ thể, mô hình CDM đạt độ chính xác chung trên 99.60% và tỷ lệ cảnh báo sai khoảng 0.4%. Mô hình phát hiện CDM và các kết quả thử nghiệm đã được công bố trong bài báo "An Improved Model for Detecting DGA botnets Using Random Forest Machine Learning Algorithm", đăng trên tạp chí Information Security Journal, 2021, ESCI Scopus Q2 [CT1].

Phần cuối của chương đề xuất mô hình WDM cho phép phát hiện hiệu quả các word-based DGA botnet. Mô hình WDM cải thiện hiệu suất phát hiện các word-based DGA botnet bằng cách sử dụng một bộ 16 đặc trưng mới dựa trên từ để phân biệt các tên miền word-based DGA và tên miền lành tính. Kết quả thử nghiệm xác nhận rằng mô hình WDM đạt độ đo F1 là 97,01% đối với bộ dữ liệu word-based DGA botnet (*DATASET-01*). Hơn nữa, mô hình dựa trên cây quyết định J48 cũng hoạt động tương đối tốt trên tập dữ liệu kết hợp (*DATASET-02*) gồm các tên miền word-based DGA botnet và character-based DGA botnet với độ đo F1 là 95,75%. Mô hình WDM và các kết quả thử nghiệm đã được công bố trong bài báo "An Novel Machine Learning-based Approach for Detecting Word-based Botnets", đăng trên tạp chí Journal of Theoretical and Applied Information Technology, 2021, Scopus Q4 [CT2].

CHƯƠNG 3: PHÁT HIỆN DGA BOTNET DỰA TRÊN HỌC KẾT HỢP

3.1. KHÁI QUÁT VỀ HỌC KẾT HỢP

3.1.1. Giới thiệu

Học kết hợp (*ensemble learning*) hay còn gọi là học theo nhóm là một cách tiếp cận cho học máy nhằm tìm kiếm hiệu suất dự đoán tốt hơn bằng cách kết hợp các dự đoán từ nhiều mô hình thành phần. Có nhiều phương pháp kết hợp các dự đoán từ các mô hình thành phần, nhưng có 3 phương pháp học kết hợp được thừa nhận và sử dụng rộng rãi bao gồm: (i) đóng gói (*bagging*) là việc kết hợp nhiều cây quyết định trên các mẫu khác nhau của cùng một tập dữ liệu và tính trung bình các dự đoán; (ii) xếp chồng *(stacking*) là việc kết hợp nhiều loại mô hình khác nhau trên cùng một tập dữ liệu và sử dụng một mô hình khác để kết hợp tốt nhất các dự đoán; (iii) tăng cường (*boosting*) liên quan đến việc thêm các thành viên tổng hợp một cách tuần tự để sửa các dự đoán được thực hiện bởi các mô hình trước đó và xuất ra giá trị trung bình có trọng số của các dự đoán.

Phần tiếp theo của mục này giới thiệu về một số kỹ thuật học kết hợp đơn giản, bao gồm max voting, averaging và weighted averaging, và một số kỹ thuật học kết hợp nâng cao, bao gồm bagging, stacking và boosting.

3.1.2. Kỹ thuật học kết hợp đơn giản

3.1.3. Kỹ thuật học kết hợp nâng cao

3.2. CÁC PHƯƠNG PHÁP PHÁT HIỆN BOTNET DỰA TRÊN HỌC KẾT HỢP

Mặc dù học kết hợp đã được sử dụng trong nhiều lĩnh vực, nhưng trong lĩnh vực phát hiện botnet nói chung và phát hiện DGA botnet nói riêng, số lượng các các đề xuất phát hiện botnet sử dụng học kết hợp còn chưa nhiều. Mục này thực hiện khảo sát một số đề xuất tiêu biểu, gồm Bijalwan và cộng sự, Zahraa và cộng sự và Charan và cộng sự.

3.2.1. Các phương pháp phát hiện DGA botnet dựa trên học kết hợp

3.2.2. Ưu và nhược điểm của các đề xuất phát hiện botnet dựa trên học kết hợp

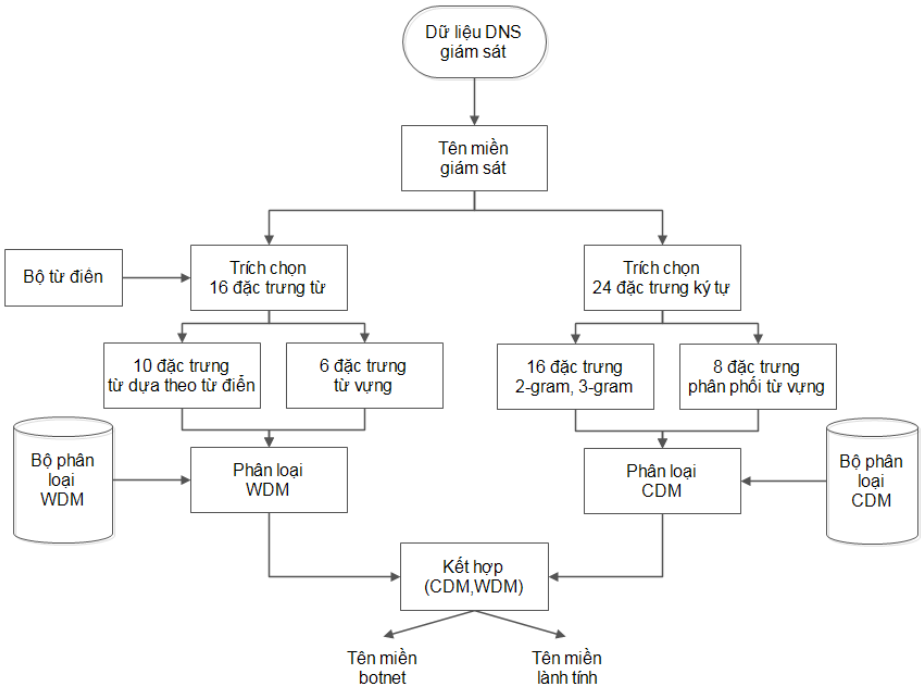
Các phương pháp đã sử dụng các thuật toán học kết hợp có thể kể đến như Adaboost, Stacking và Bagging làm tăng hiệu quả phát hiện đáng kể.

Trong các phương pháp phát hiện botnet dựa trên kết hợp nêu trên đều có sử dụng đến các đặc trưng mạng. Khi sử dụng các đặc trưng mạng đòi hỏi nhiều các chi phí liên quan đến lưu trữ, xử lý tài nguyên mạng lớn. Cũng với phương pháp học kết hợp được nêu ở trên, phần còn lại của chương này đề xuất mô hình phát hiện DGA botnet dựa trên học kết hợp giữa 2 mô hình đã đề xuất ở chương 2 đó là: CDM và WDM.

3.3. MÔ HÌNH PHÁT HIỆN DGA BOTNET DỰA TRÊN HỌC KẾT HỢP

3.3.1. Giới thiệu mô hình

Từ kết quả 2 mô hình phát hiện character-based DGA botnet (CDM) và word-based DGA botnet (WDM) đã được trình bày tại mục 2.2 và mục 2.3, có thể thấy mỗi mô hình đều có điểm mạnh và điểm yếu riêng. Mô hình CDM có khả năng phát hiện hiệu quả các character-based DGA botnet, nhưng hầu như không thể phát hiện các word-based DGA botnet. Ngược lại, mô hình WDM có khả năng phát hiện hiệu quả các word-based DGA botnet, nhưng tỷ lệ phát hiện các character-based DGA botnet của WDM nói chung thấp hơn CDM. Do vậy, phần này đề xuất mô hình dựa trên học kết hợp để hợp nhất 2 mô hình CDM và WDM nhằm phát huy ưu điểm của cả hai mô hình trong một mô hình phát hiện DGA botnet thống nhất.



Hình 3.1: Giai đoạn phát hiện của mô hình học kết hợp đề xuất

Mô hình kết hợp có giai đoạn huấn luyện được kế thừa từ hai mô hình thành phần CDM và WDM đã được mô tả ở chương 2. Theo đó, các mô hình CDM và WDM được huấn luyện riêng và kết quả là hai bộ phân loại CDM và WDM. Trong giai đoạn phát hiện của mô hình kết hợp như biểu diễn trên Hình 3.1, 2 mô hình CDM và WDM được sử dụng để xử lý tên miền song song và kết quả từ 2 mô hình sẽ được kết hợp để tăng hiệu suất phát hiện chung. *Kết hợp (CDM, WDM)* là phép hợp các phát hiện chính xác từ 2 mô hình. Việc kết hợp này xuất phát từ ý tưởng lấy kết quả phát hiện chính xác của từng mô hình, xếp chồng cho kết quả là số lượng tối đa các DGA bonet được phát hiện. Vì sử dụng chung dữ liệu kiểm thử đối với cả 2 mô hình nên trọng số khi kết hợp được tính theo tỷ lệ 1:1.

3.3.2. Tập dữ liệu huấn luyện và kiểm thử

Dữ liệu huấn luyện: Đối với mô hình học kết hợp đề xuất, sử dụng bộ phân loại đã được huấn luyện từ hai mô hình CDM và WDM đề xuất tại chương 2.

Dữ liệu kiểm thử: như đã đề cập ở trên, để so sánh hiệu suất của các mô hình, tập dữ liệu kiểm thử sẽ được lấy từ tệp dữ liệu kiểm thử trong mô hình phát hiện character-based DGA botnet được trình bày tại mục 2.2.3. Dữ liệu kiểm thử gồm các tên miền của 39 họ DGA botnet với 71,393 mẫu tên miền DGA botnet. Ngoài ra, tập dữ liệu UMUDGA cũng được sử dụng để đánh giá khả năng phát hiện các tên miền DGA botnet mới, không tồn tại trong tập dữ liệu huấn luyện.

3.3.3. Tiền xử lý, huấn luyện và phát hiện

3.3.4. Các kết quả

Bảng 3.1 biểu diễn tỷ lệ phát hiện (DR) của các mô hình CDM, WDM và mô hình kết hợp với các DGA botnet có $DR_{\text{kết hợp}} > 90\%$. Bảng 3.10 biểu diễn tỷ lệ phát hiện (DR) của các mô hình CDM, WDM và mô hình kết hợp với các DGA botnet có $DR_{\text{kết hợp}} < 90\%$. Từ các kết quả trên có thể thấy, mô hình kết hợp đã kết hợp được các ưu điểm của các mô hình CDM và WDM: phát hiện hiệu quả cả các character-based và word-based DGA botnet.

Bảng 3.1: Các DGA botnet có tỷ lệ DR lớn hơn 90% với mô hình đề xuất

STT	Họ botnet	Số lượng	CDM %	WDM %	Kết hợp %
1	Rovnix	4000	100.00	99.50	100.00
2	Dyre	1000	100.00	99.90	100.00
3	Chinad	1000	100.00	97.90	100.00
4	Fobber_v1	298	100.00	100.00	100.00
5	Tinynuke	32	100.00	100.00	100.00
6	Gameover	4000	100.00	99.98	100.00
7	Murofet	4000	99.80	99.78	100.00
8	Cryptolocker	1000	99.70	96.20	100.00
9	Padcrypt	168	98.21	98.21	100.00
10	Dircrypt	762	99.34	93.31	100.00
11	Fobber_v2	299	100.00	89.30	100.00
12	Vidro	100	100.00	49.00	100.00
13	Emotet	4000	99.68	99.55	99.98
14	Tinba	4000	99.98	99.08	99.98
15	Ranbyus	4000	99.58	99.30	99.93
16	Shiotob	4000	99.68	95.95	99.88
17	Pykspa_v1	4000	99.70	58.90	99.83
18	Necurs	4000	99.35	87.75	99.78
19	Ramnit	4000	99.55	91.45	99.75
20	Virut	4000	99.75	0.00	99.75
21	Qadars	2000	99.05	95.40	99.65
22	Simda	4000	99.65	0.00	99.65
23	Pykspa_v2_fake	799	98.87	61.08	99.25
24	Locky	1158	99.05	83.16	99.14
25	Pykspa_v2_real	199	98.99	63.32	98.99
26	Shifu	2546	98.59	34.92	98.82
27	Matsnu	881	12.15	98.41	98.64
28	Proslikefan	100	98.00	50.00	98.00
29	Tempedreve	195	97.44	64.62	97.44
30	Vawtrak	827	96.61	61.67	97.10
31	Symmi	1200	96.58	31.67	96.83
32	Suppobox	2205	19.27	92.83	96.05
33	Nymaim	480	94.79	61.25	95.21

34	Mydoom	50	88.00	74.00	94.00
	Tổng cộng	**65299**	**95.59**	**77.18**	**99.53**

Bảng 3.2: Các DGA botnet có tỷ lệ DR nhỏ hơn 90% với mô hình đề xuất

STT	Họ botnet	Số lượng	CDM %	WDM %	Kết hợp %
1	Conficker	495	89.29	52.93	89.49
2	Bigviktor	999	11.11	70.97	76.18
3	Gspy	100	76.00	8.00	76.00
4	Enviserv	500	50.40	19.40	52.00
5	Banjori	4000	0.00	0.00	0.00
	Tổng cộng	**6094**	**14.46**	**17.66**	**25.24**

Bảng 3.3: Tỷ lệ phát hiện đối với tập dữ liệu UMUDGA

STT	Họ botnet	Số lượng	CDM %	WDM %	Kết hợp %
1	Alureon	5,000	98.22	85.32	98.94
2	Bedep	5,000	99.82	97.80	99.92
3	Corebot	5,000	99.76	98.24	99.94
4	Kraken	2,000	98.40	69.50	99.10
5	Pushdo	5,000	94.36	35.90	95.08
6	Zeus	5,000	100.00	99.96	100.00
		27,000	**98.43**	**82.41**	**98.80**
7	Pizd	4,000	16.05	97.93	98.05
	Tổng cộng	**31,000**	**87.80**	**84.41**	**98.70**

3.3.5. Đánh giá

Bảng 3.1 thống kê các botnet có tỷ lệ DR lớn hơn 90% khi sử dụng mô hình kết hợp đề xuất, trong tổng số 39 họ DGA botnet thì có 34 họ cho DR lớn hơn 94.00%. 12 họ DGA botnet có DR đạt 100%, gồm Rovnix, Dyre, Chinad, Fobber_v1, Tinynuke, Gameover, Murofet, Cryptolocker, Padcrypt, Dircrypt, Fobber_v2 và Vidro. Tỷ lệ phát hiện trung bình của nhóm này đạt tới 99.53%.

Bảng 3.2 liệt kê 5 họ DGA botnet có DR không cao khi sử dụng mô hình kết hợp. Trong đó, Conficker có DR đạt 89.49%, Bigviktor và Gspy có DR đạt khoảng 76%, Enviserv có DR đạt 52%. Đặc biệt, mô hình kết hợp không thể phát hiện Banjori botnet do các mô hình thành phần cũng không thể phát hiện botnet này.

Bảng 3.3 thể hiện tỷ lệ phát hiện của các mô hình CDM, WDM và kết hợp đối với tệp dữ liệu kiểm thử lấy từ tập UMUDGA. Kết quả cho thấy, đối với các character-based DGA botnet (6 họ botnet đầu danh sách), mô hình CDM cho tỷ lệ phát hiện đạt 98.43%. Đối với 'Pizd' là word-based DGA botnet, mô hình kết hợp cho tỷ lệ phát hiện đạt 98.05%. Tỷ lệ phát hiện tổng thể của mô hình kết hợp đạt 98.70% đối với 31,000 tên miền DGA botnet, bao gồm cả character-based và word-based DGA botnet.

Mô hình phát hiện botnet DGA được đề xuất dựa trên học kết hợp "muộn" có thể phát hiện hiệu quả hầu hết các botnet DGA, bao gồm character-based và word-based DGA botnet vì nó có thể tận dụng lợi thế của cả mô hình CDM và WDM thành phần. Kết quả thực nghiệm đưa ra trong Bảng 3.4 cho thấy mô hình kết hợp được đề xuất có khả năng phát hiện hiệu quả 37 trong số 39 họ botnet DGA có DR> 89%, trong đó 12 họ botnet DGA có DR = 100% và 31 botnet có DR> 97%. Cũng rõ ràng rằng mô hình kết hợp "muộn" được đề xuất hoạt động tốt hơn nhiều so với mô hình kết hợp "sớm". Tỷ lệ phát hiện (DR) của mô hình kết hợp "muộn" được đề xuất cao hơn đáng kể so với mô hình kết hợp "sớm" cho tất cả các họ botnet, ngoại trừ gspy. Mô hình kết hợp "sớm" thậm chí không phát hiện được một số botnet DGA, chẳng hạn như virus, simda, conficker và enviserv.

Tóm lại, mô hình phát hiện kết hợp đề xuất đã khai thác được điểm mạnh của cả 2 mô hình thành phần là CDM và WDM: mô hình phát hiện kết hợp có khả năng phát hiện hiệu quả hầu hết các character-based DGA botnet và word-based DGA botnet. Theo đó, mô hình

phát hiện kết hợp có DR cao hơn CDM với character-based DGA botnet và mô hình phát hiện kết hợp có DR cao hơn WDM với word-based DGA botnet.

Hạn chế của mô hình kết hợp là thời gian huấn luyện và phát hiện dài hơn các mô hình thành phần, nhưng với hiệu quả phát hiện vượt trội, mô hình kết hợp cho hiệu quả tổng hợp tốt hơn. Ngoài ra, quá trình huấn luyện các mô hình thành phần có thể thực hiện theo mẻ trong chế độ offline. Một hạn chế khác của mô hình kết hợp là không thể phát hiện một số DGA botnet có phương pháp tạo tên miền đặc biệt như Banjori.

3.4. KẾT LUẬN CHƯƠNG

Chương 3 giới thiệu khái quát về học kết hợp, các phương pháp học kết hợp, bao gồm các phương pháp học kết hợp đơn giản và học kết hợp nâng cao. Các phương pháp học kết hợp đơn giản bao gồm max voting, averaging và weighted averaging, và các phương pháp học kết hợp nâng cao bao gồm bagging, stacking và boosting. Phần tiếp theo của chương này khảo sát một số nghiên cứu phát hiện botnet dựa trên học kết hợp. Nhìn chung, các đề xuất phát hiện botnet dựa trên học kết hợp có số lượng khá ít và hiệu quả của học kết hợp chưa thực sự rõ ràng.

Phần cuối của chương tập trung giải quyết vấn đề phát hiện cả character-based và word-based DGA botnet trong một mô hình thống nhất bằng cách đề xuất mô hình phát hiện DGA botnet dựa trên học kết hợp. Mô hình phát hiện kết hợp sử dụng hai mô hình thành phần là CDM và WDM đã được thử nghiệm và đánh giá ở chương 2. Mô hình kết hợp đã khai thác được điểm mạnh của cả 2 mô hình thành phần là CDM và WDM: mô hình kết hợp có khả năng phát hiện hiệu quả hầu hết các DGA botnet, bao gồm cả character-based DGA botnet và word-based DGA botnet. Kết quả thử nghiệm cho thấy, trong số 39 họ DGA botnet thử nghiệm, mô hình kết hợp có tỷ lệ phát hiện DR từ 94% trở lên với 34 họ DGA botnet, trong đó 12 họ botnet có DR đạt 100%. Mô hình kết hợp chỉ không thể phát hiện 1 họ botnet là Banjori do các mô hình thành phần cũng không thể phát hiện botnet này. Mô hình kết hợp đề xuất và kết quả thử nghiệm, đánh giá được đăng trên bài báo “Một mô hình phát hiện DGA botnet dựa trên học kết hợp”, tạp chí Khoa học Công nghệ Thông tin và Truyền trông, ISSN: 2525-2224, Vol. 1, No. 1, 2022 [CT3].

KẾT LUẬN

Botnet đã và đang trở thành một trong các nguy cơ gây mất an toàn thông tin hàng đầu do chúng không ngừng phát triển về cả quy mô và mức độ tinh vi trong các kỹ thuật chỉ huy và kiểm soát. Nhiều dạng botnet sử dụng kỹ thuật DGA để sinh và đăng ký nhiều tên miền ngẫu nhiên khác nhau cho máy chủ CnC của chúng nhằm chống lại việc bị kiểm soát và vô hiệu hóa. Các DGA botnet thường khai thác hệ thống DNS để duy trì hoạt động, do vậy việc phân tích phát hiện các tên miền truy vấn hệ thống DNS có thể giúp phát hiện các hoạt động của botnet. Luận án này tập trung giải quyết hai vấn đề: (1) nghiên cứu, đề xuất tập đặc trưng phân loại tên miền mới phù hợp hơn cho xây dựng các mô hình phát hiện DGA botnet, nhằm tăng tỷ lệ phát hiện đúng và giảm tỷ lệ cảnh báo sai và (2) nghiên cứu, lựa chọn sử dụng phương pháp học máy phù hợp cho xây dựng các mô hình phát hiện DGA botnet, nhằm xây dựng một mô hình phát hiện thống nhất cho phép phát hiện hiệu quả nhiều dạng DGA botnet.

Với vấn đề (1) nghiên cứu, đề xuất tập đặc trưng phân loại tên miền mới phù hợp hơn cho xây dựng các mô hình phát hiện DGA botnet, nhằm tăng tỷ lệ phát hiện đúng và giảm tỷ lệ cảnh báo sai, luận án đề xuất mô hình CDM cho phát hiện character-based DGA botnet và mô hình WDM cho phát hiện word-based DGA botnet. Mô hình phát hiện CDM đề xuất sử dụng 24 đặc trưng ký tự để phân loại tên miền lành tính với tên miền sinh bởi các DGA botnet, gồm 16 đặc trưng thống kê n-gram, 6 đặc trưng phân bố nguyên âm, ký tự, chữ số, và 2 đặc trưng entropy theo ký tự và giá trị kỳ vọng của tên miền. Các thử nghiệm trên tập dữ liệu gồm 100,000 tên miền lành tính và 153,000 tên miền DGA cho thấy, mô hình CDM đề đạt các độ

đo đánh giá vượt trội so với các mô hình đã có. Cụ thể, mô hình CDM đạt độ chính xác chung trên 99.60% và tỷ lệ cảnh báo sai rất thấp, chỉ khoảng 0.4%. Như vậy có thể khẳng định tập 24 đặc trưng ký tự sử dụng trong mô hình CMD là phù hợp cho phát hiện các họ character-based DGA botnet.

Mặc dù mô hình CDM đạt hiệu xuất phát hiện tốt cho hầu hết các character-based DGA botnet, CDM không có khả năng phát hiện hiệu quả các họ word-based DGA botnet, như 'banjori', 'matsnu', 'bigviktor' và 'suppobox'. Điều này là do các word-based DGA botnet có khả năng sinh các tên miền rất giống các tên miền lành tính sử dụng tổ hợp các từ tiếng Anh lấy từ các danh sách dựng sẵn. Để giải quyết vấn đề này, luận án đề xuất mô hình WDM cho phép phát hiện hiệu quả các họ word-based DGA botnet. Mô hình WDM đề xuất sử dụng 16 đặc trưng từ cho phân loại tên miền word-based DGA botnet với các tên miền lành tính, bao gồm 10 đặc trưng từ dựa trên từ điển và 6 đặc trưng từ vựng. Luận án sử dụng 5 thuật toán học máy có giám sát, bao gồm Naïve Bayes, cây quyết định, rừng ngẫu nhiên, hồi quy logistic và SVM để xây dựng và kiểm thử mô hình phát hiện. Các kết quả thử nghiệm trên các tập dữ liệu DATASET-01 và DATASET-02 với 4 kịch bản cho thấy mô hình WDM có khả năng phát hiện hiệu quả các word-based DGA botnet, cũng như có khả năng phát hiện tốt nhiều character-based DGA botnet với độ đo F1 đạt trên 95%. Trong các thuật toán học máy sử dụng, thuật toán học máy cây quyết định J48 cho hiệu xuất phát hiện tổng thể tốt nhất trong các thuật toán thử nghiệm. Như vậy có thể khẳng định tập 16 đặc trưng từ sử dụng trong mô hình WMD là phù hợp cho phát hiện các họ word-based DGA botnet.

Với vấn đề (2) nghiên cứu, lựa chọn sử dụng phương pháp học máy phù hợp nhằm xây dựng một mô hình phát hiện thống nhất cho phép phát hiện hiệu quả nhiều dạng DGA botnet, luận án đề xuất mô hình phát hiện DGA botnet dựa trên học kết hợp. Mô hình phát hiện kết hợp đề xuất nhằm khai thác điểm mạnh của 2 mô hình thành phần là CDM và WDM: mô hình phát hiện kết hợp có khả năng phát hiện hiệu quả hầu hết các DGA botnet, bao gồm cả character-based DGA botnet và word-based DGA botnet. Các kết quả thử nghiệm cho thấy, mô hình phát hiện dựa trên học kết hợp đạt tỷ lệ phát hiện trung bình là 99.53% trên 39 họ DGA botnet thử nghiệm. Cụ thể, mô hình kết hợp có tỷ lệ phát hiện đạt từ 94% trở lên với 34 họ DGA botnet, trong đó 12 họ botnet có tỷ lệ phát hiện đạt 100%. Trong số 39 họ DGA botnet, chỉ có 5 họ DGA botnet có tỷ lệ phát hiện dưới 90%. Ngoài ra, mô hình kết hợp cũng có khả năng phát hiện hiệu quả các character-based và word-based DGA botnet mới trong tập dữ liệu UMUDGA với tỷ lệ phát hiện trung bình đạt 98,70%.

Các đề xuất phát hiện DGA botnet dựa trên tên miền thực thi hiệu quả hơn so với các phương pháp dựa trên lưu lượng mạng bởi giảm thiểu các đặc trưng, xử lý dữ liệu luồng và gói tin, do đó sẽ nhanh hơn, chi phí đỡ tốn kém hơn. Các mô hình khi đưa vào ứng sẽ được cài đặt tại DNS server nhằm ngăn chặn các bot có thể liên lạc được với CnC server hoặc trước firewall trong các hệ thống đơn lẻ nhằm phát hiện máy tính nào là bot.

Các hạn chế của mô hình kết hợp bao gồm: (1) thời gian huấn luyện và phát hiện dài hơn so với mô hình thành phần và (2) mô hình kết hợp không có khả năng phát hiện một số DGA botnet thuộc họ mixed DGA, như Banjori. Đây cũng là các vấn đề cần giải quyết cho hướng phát triển trong tương lai của luận án. Ngoài ra, việc phát triển hệ thống phát hiện DGA botnet dựa trên các mô hình phát hiện đề xuất cũng là một hướng mở của đề tài luận án.

DANH MỤC CÁC CÔNG TRÌNH CÔNG BỐ

TẠP CHÍ KHOA HỌC

[CT1] Xuan Dau Hoang, Xuan Hanh Vu, 2021: An improved model for detecting DGA botnets using random forest algorithm, Information Security Journal: A Global Perspective, DOI: 10.1080/19393555.2021.1934198. ESCI Scopus Q2.

[CT2] Xuan Hanh Vu, Xuan Dau Hoang, 2021: An Novel Machine Learning-based Approach for Detecting Word-based Botnets, Journal of Theoretical and Applied Information Technology, Vol 99 – 24. Scopus Q4.

[CT3] Vũ Xuân Hạnh, Hoàng Xuân Dậu, Đinh Trường Duy, 2022, "Một mô hình phát hiện DGA botnet dựa trên học kết hợp", tạp chí Khoa học Công nghệ Thông tin và Truyền thông, ISSN: 2525-2224, Vol. 1, No. 1, 2022.

HỘI THẢO KHOA HỌC

[CT4] Hoang X.D., Vu X.H, 2021. An Enhanced Model for DGA Botnet Detection Using Supervised Machine Learning. Intelligent Systems and Networks, ICISN 2021. Lecture Notes in Networks and Systems, vol 243. Springer, Singapore. DOI: 10.1007/978-981-16-2094-2_6. Scopus Q4.

[CT5] Vũ Xuân Hạnh, Hoàng Xuân Dậu, 2019. Phát hiện DGA Botnet sử dụng kết hợp nhiều nhóm đặc trưng phân loại tên miền. Hội nghị KHCN Quốc gia lần thứ XII (FAIR); Huế, ngày 07-08/6/2019. DOI: 10.15625/vap.2019.00047.

[CT6] Nguyễn Trọng Hưng, Hoàng Xuân Dậu, Vũ Xuân Hạnh, 2018 "Phát hiện botnet dựa trên phân loại tên miền sử dụng kỹ thuật học máy", Hội thảo lần III: Một số vấn đề lựa chọn về an toàn an ninh thông tin, Tạp chí Thông tin và truyền thông 12/2018, ISSN: 1859 – 3550.