HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

BÙI ĐIỀN PHONG

XÂY DỰNG CÔNG CỤ PHÁT HIỆN XÂM NHẬP MẠNG MÁY TÍNH

LUẬN VĂN THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

TP HỒ CHÍ MINH – NĂM 2022

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

BÙI ĐIỀN PHONG

XÂY DỰNG CÔNG CỤ PHÁT HIỆN XÂM NHẬP MẠNG MÁY TÍNH

CHUYÊN NGÀNH : HỆ THỐNG THÔNG TIN

MÃ SỐ : 08.48.01.04

ĐỀ CƯƠNG LUẬN VĂN THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

NGƯỜI HƯỚNG DẪN KHOA HỌC

TS. NGUYỄN ĐỨC THÁI

TP HỒ CHÍ MINH – NĂM 2022

LỜI CAM ĐOAN

Tôi cam đoan rằng luận văn ***"Xây dựng công cụ phát hiện xâm nhập mạng máy tính"*** là công trình nghiên cứu của chính tôi.

Tôi cam đoan các số liệu, kết quả nêu trong luận văn là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác.

Không có sản phẩm/nghiên cứu nào của người khác được sử dụng trong luận văn này mà không được trích dẫn theo đúng quy định.

TP. Hồ Chí Minh, ngày 15 tháng 7 năm 2022

Học viên thực hiện luận văn

Bùi Điền Phong

LỜI CẢM ƠN

Trong suốt quá trình học tập và nghiên cứu thực hiện luận văn, ngoài nỗ lực của bản thân, tôi đã nhận được sự hướng dẫn nhiệt tình quý báu của quý Thầy Cô, cùng với sự động viên và ủng hộ của gia đình, bạn bè và đồng nghiệp. Với lòng kính trọng và biết ơn sâu sắc, tôi xin gửi lời cảm ơn chân thành tới:

Ban Giám Đốc, Phòng đào tạo sau đại học, Học viện Công nghệ Bưu chính Viễn thông cơ sở TPHCM và quý Thầy Cô đã tạo mọi điều kiện thuận lợi giúp tôi hoàn thành luận văn.

Tôi xin chân thành cảm ơn Thầy **TS. Nguyễn Đức Thái**, người thầy kính yêu đã hết lòng giúp đỡ, hướng dẫn, động viên, tạo điều kiện cho tôi trong suốt quá trình thực hiện và hoàn thành luận văn.

Tôi xin chân thành cảm ơn gia đình, bạn bè, đồng nghiệp trong cơ quan đã động viên, hỗ trợ tôi trong lúc khó khăn để tôi có thể học tập và hoàn thành luận văn.

Mặc dù đã có nhiều cố gắng, nỗ lực, nhưng do thời gian và kinh nghiệm nghiên cứu khoa học còn hạn chế nên không thể tránh khỏi những thiếu sót. Tôi rất mong nhận được sự góp ý của quý Thầy Cô cùng bạn bè đồng nghiệp để kiến thức của tôi ngày một hoàn thiện hơn.

Xin chân thành cảm ơn!

TP. Hồ Chí Minh, ngày 15 tháng 7 năm 2022

Học viên thực hiện luận văn

Bùi Điền Phong

DANH SÁCH HÌNH VẼ

DANH SÁCH BẢNG

DANH MỤC CÁC THUẬT NGỮ, CHỮ VIẾT TẮT

Viết tắt	Tiếng Anh	Tiếng Việt
IDS	Intrusion Detection System	Hệ thống phát hiện xâm nhập
NIDS	Network Intrusion Detection System	Hệ thống phát hiện xâm nhập mạng
HIDS	Host-based intrusion detection system	Hệ thống phát hiện xâm nhập dựa trên máy chủ
NAT	Network Address Translation	
KNN	K Nearest Neighbor	
DT	Decision Tree	
SVM	Support Vector Machine	

MỤC LỤC

PHẦN MỞ ĐẦU

1. Lý do chọn đề tài

Mạng NSFnet được thành lập và đã kết nối năm trung tâm máy tính vào năm 1986. Cũng vì thế, nó đã đem đến sự bùng nổ trong việc kết nối, đặc biệt ở khu vực các trường đại học. Từ đó, ARPANET và NSF cùng tồn tại song song, sử dụng chung 1 giao thức, và chúng có sự kết nối lẫn nhau. Tiếp đến năm 1990, ARPANET giờ đây là một dự án dừng hoạt động nhưng vì mạng do ARPANET cùng NSF sinh ra đã được áp dụng vào mục đích là dân dụng, đây cũng là tiền thân của Internet ngày nay. Cho đến lúc này mạng Internet được sử dụng chủ yếu bởi đối tượng là những nhà nghiên cứu, đồng thời dịch vụ được dùng phổ biến nhất là FTP và email. Thời điểm này Internet đã được coi như là một phương tiện đại chúng.

Ngày nay Internet được sử dụng phổ biến mọi lúc mọi nơi, từ máy tính để bàn, cá nhân đến điện thoại cũng như các thiết bị thông minh, ngoài ra nó được áp dụng vào nhiều lĩnh vực của đời sống như giáo dục, y tế, kinh tế, quốc phòng. Từ đó cũng phát sinh nên nhiều vấn đề liên quan đến bảo mật cũng như an ninh trên mạng Internet. An ninh mạng là một vấn đề lớn và rất quan trọng với mục tiêu là đảm bảo an ninh môi trường làm việc của cá nhân hay tổ chức. Sự mất mát về thông tin và bảo mật của cho một người dùng có thể sẽ không quá lớn hoặc với một số trường hợp là không đáng để truy cứu, nhưng đối với các doanh nghiệp hay các tổ chức lớn tổn thất có thể lên tới hàng triệu đô la. Hoặc các cơ quan tổ chức thuộc nhà nước sẽ dẫn đến các nguy cơ lộ các thông tin bí mật dẫn đến mối nguy hại cho quốc gia. Hàng loạt các cuộc tấn công có thể nhắm vào mọi thứ như là dữ liệu cá nhân hoặc tổ chức, tài khoản ngân hàng, phần mềm, tài khoản người dùng, mạng cục bộ, … Đó là lý do tại sao các công cụ bảo mật phát triển ngày càng nhiều để đáp ứng các dạng phần mềm nguy hiểm, phần mềm độc hại và tin tặc ngày nay.

Trong thời buổi công nghệ thông tin hiện nay, một trong những yếu tố mà các doanh nghiệp phải ưu tiên xem xét đến là bảo mật thông tin hay an toàn an ninh mạng. Không ít các doanh nghiệp phải thuê đối tác thứ ba với mục đích bảo vệ hệ thống mạng và bảo mật thông tin, ngoài ra cũng có những doanh nghiệp lập nên các kế hoạch về việc tính toán chi phí để mua sản phẩm phần cứng hay phần mềm nhằm đáp

ứng việc đảm bảo an toàn dữ liệu của công ty. Tuy nhiên, những giải pháp này khiến các cơ quan doanh nghiệp luôn phải đau đầu vì phải thực hiện cân đối về chính sách và tài chính hằng năm để đáp ứng mục tiêu là có được giải pháp an toàn tối ưu và chi phí rẻ đảm bảo việc trao đổi thông tin được an toàn, bảo vệ thông tin của công ty trước những mối nguy cơ tấn công của các tội phạm công nghệ.

Do đó đề tài **"Xây dựng công cụ phát hiện xâm nhập mạng máy tính"** được phát triển nhằm đáp ứng một phần nào yêu cầu của các cơ quan, tổ chức, doanh nghiệp đảm bảo được an toàn thông tin và bảo mật hệ thống mạng của đơn vị mình.

Luận văn nghiên cứu về phương pháp phát hiện và phòng chống xâm nhập mạng máy tính và xây dựng công cụ phát hiện xâm nhập mạng máy tính.

Mục đích đề tài nhằm nhằm xây dựng một hệ thống phát hiện xâm nhập và phòng chống các cuộc tấn công từ internet và áp dụng giải pháp vào trong thực tiễn công việc tại Trung tâm Y tế huyện Gò Dầu.

2. Tổng quan về vấn đề nghiên cứu

Xây dựng hệ thống phát hiện xâm nhập theo thời gian thực. Thêm vào đó, hệ thống có chức năng nhận dạng và phân tích các cuộc xâm nhập trái phép vào hệ thống. Sau khi thu được kết quả sẽ tổng hợp và kết xuất dữ liệu báo cáo ra file để tổng hợp.

3. Mục đích nghiên cứu

Tập trung nghiên cứu các loại xâm nhập mạng, phân tích và phân loại thành các mức độ nguy hiểm khác nhau.

Nghiên cứu cơ chế hoạt động của một hệ thống chống xâm nhập mạng, từ đó đưa ra tiếp thu và đưa ra các giải pháp phù hợp trong việc xây dựng hệ thống.

Nghiên cứu các công cụ hỗ trợ phân tích các luồng thông tin ra vào mạng máy tính kết hợp với những thuật toán phân lớp hoặc phân cụm để theo dõi và truy vết dấu hiệu hợp pháp và bất hợp pháp, sau đó gửi tính hiệu cảnh báo cho quản trị mạng biết những dấu hiệu xâm nhập trái phép.

4. Đối tượng và phạm vi nghiên cứu

Đề tài nghiên cứu đưa ra cách nhìn tổng quan nhất về một hệ thống phát hiện xâm nhập mạng máy tính, các phương thức tấn công mạng và các giải pháp bảo mật hệ thống mạng. Bên cạnh đó xây dựng một hệ thống giám sát và cảnh báo bằng email

đến quản trị viên, giúp cho việc quản trị hệ thống mạng trở nên cơ động và an toàn hơn.

Đề tài sẽ được ứng dụng ngay tại Trung tâm Y tế huyện Gò Dầu của học viên hoặc có thể mở rộng trong toàn ngành nơi mà học viên đang công tác và đáp ứng nhu cầu của cơ quan.

5. Phương pháp nghiên cứu

- Tìm hiểu về cách thức xâm nhập trái phép trong mạng máy tính.
- Nghiên cứu lý thuyết về các khả năng tấn công mạng.
- Phân tích các khả năng phát hiện xâm nhập và phòng chống tấn công.
- Nghiên cứu các thuật toán, đặt biệt là các thuật toán Support Vector Machine (SVM), Decision Tree (DT), K Nearest Neighbor để phát hiện xâm nhập trái phép.
- Nghiên cứu các ứng dụng đang sử dụng hiện nay để phát hiện xâm nhập trái phép.
- Tiến hành hệ thống thử nghiệm.

6. Bố cục luận văn

Ngoài phần mở đầu, mục lục, kết luận và tài liệu tham khảo, nội dung chính của luận án được chia thành 4 chương, cụ thể như sau:

Chương 1: Tổng quan các phương pháp phát hiện và phòng chống xâm nhập mạng.

Chương 2: Các công trình liên quan.

Chương 3: Hệ thống phát hiện và phòng chống xâm nhập mạng.

Chương 4: Xây dựng và triển khai hệ thống phát hiện xâm nhập mạng dựa vào học máy cho hệ thống mạng trung tâm y tế huyện Gò Dầu.

PHẦN NỘI DUNG

CHƯƠNG 1. TỔNG QUAN CÁC PHƯƠNG PHÁP PHÁT HIỆN VÀ PHÒNG CHỐNG XÂM NHẬP MẠNG

1.1 Tổng quan đề tài

1.1.1 Xâm nhập mạng là gì?

Xâm nhập mạng [1] đề cập đến bất kỳ hoạt động trái phép nào trên mạng. Xâm nhập mạng có hai loại, một là xâm nhập mạng để kiểm thử hay còn gọi là pentest, hai là tấn công mạng. Kiểm thử xâm nhập là hành động xâm nhập nhằm tìm ra lỗ hổng bảo mật trong hệ thống để khắc phục. Sau khi tìm ra các lỗ hổng nghiêm trọng hoặc nguy hiểm, người kiểm thử sẽ đưa ra giải pháp để vá lỗ hổng nhằm bảo vệ tài nguyên của các tổ chức được bảo mật. Ngược lại, tấn công mạng là hành động lợi dụng mạng Internet để xâm nhập trái phép vào hệ thống website của cá nhân, tổ chức hay máy tính hoặc nguy hiểm hơn là cơ sở dữ liệu, … Mục đích chính của việc này là để đánh cắp thông tin, mã hóa dữ liệu hoặc làm gián đoạn các hoạt động vận hành của tổ chức cũng như doanh nghiệp. Những đối tượng của các cuộc tấn công này có thể là cá nhân, tổ chức, doanh nghiệp lớn nhỏ hoặc các cơ quan nhà nước. Đặc biệt, đối tượng phổ biến nhất bị nhắm đến là các doanh nghiệp bởi vì mục tiêu chính của những kẻ tấn công này là lợi nhuận. Những tin tặc tấn công chủ yếu nhằm trục lợi phi pháp, tống tiền các doanh nghiệp. Ngoài ra, cũng tồn tại một số doanh nghiệp cạnh tranh không lành mạnh tấn công mạng đối thủ của mình bằng cách đánh sập website của họ. Điều này gây ảnh hưởng ít nhiều đến hình ảnh và uy tín của công ty, thương hiệu.

1.1.2 Phát hiện xâm nhập mạng

Phát hiện xâm nhập mạng [1] là tập hợp các phương pháp và kỹ thuật dùng để dò tìm những hoạt động có dấu hiệu bất thường, đáng nghi ngờ trên mạng. Một tập hợp các phương thức, công cụ, và tài nguyên giúp cho người quản trị xác định, đánh giá, và báo cáo các hoạt động không được cho phép trên mạng được định nghĩa là hệ thống phát hiện xâm nhập (IDS). Phát hiện xâm nhập được hiểu là một tiến trình mà nó được quyết định khi một người dùng chưa được xác thực cố gắng xâm nhập trái phép vào hệ thống mạng. IDS với hàng triệu tình huống được trang bị để nhận dạng

tấn công cũng như cập nhật thường xuyên sẽ kiểm tra tất cả các gói tin đi qua nó và quyết định liệu gói tin này có vấn đề gì khả nghi hay không. Vì thế, hệ thống này đã trở nên thực sự quan trọng, đồng thời là lựa chọn hàng đầu trong việc phát hiện và phòng chống xâm nhập mạng.

Công cuộc nghiên cứu và xây dựng hệ thống ngăn chặn và phát hiện xâm nhập (IDS/IPS) đang được nhiều người quan tâm cũng như phát triển và sẽ không ngừng phát triển mạnh mẽ trong thời gian tới. Những sản phẩm thương mại hiện tại trên thị trường đều có chi phí cao, vượt ngoài khả năng đầu tư của nhiều doanh nghiệp lớn nhỏ dẫn đến hệ quả là những nghiên cứu mã nguồn mở về IDS/IPS được đầu tư nghiên cứu và triển khai. Nhìn chung, những nghiên cứu trong nước liên quan đến IDS/IPS bằng mã nguồn mở tuy nhiều nhưng lại tập trung vào Snort. Ngoài ra, các nghiên cứu này vẫn chưa được áp dụng rộng rãi, tồn đọng nhiều hạn chế như: không có giao diện thân thiện; chưa được tích hợp sẵn thành phần báo động, hoặc có thì cũng chỉ giới hạn qua giao diện console hay qua giao diện Web, chính vì thế nó chưa tạo được sự linh động và tiện dụng cho người dùng; do chỉ tập trung nghiên cứu về Snort làm cho các phần mềm mang tính chất đơn lẻ mặc dù nhu cầu tích hợp các tính năng giám sát khác nhằm đạt được hiệu quả sử dụng cao hơn chưa thật sự được chú trọng và phát triển.

Mặt khác, IDS/IPS phải luôn trong tình trạng cập nhật những dấu hiệu tấn công mới vì các kiểu tấn công mạng đang ngày một tinh vi và phức tạp kéo theo những dấu hiệu của chúng liên tục thay đổi. Chuyên viên quản trị mạng có thể dựa vào các phân tích khác chẳng hạn như dấu hiệu bất thường của lưu lượng vào hoặc ra khỏi hệ thống, hoạt động của RAM, CPU, ... để phản ứng kịp thời. Thêm vào đó, hệ thống báo động cần triển khai phải mang tính chất đa dạng với nhiều hình thức, tiện dụng và linh động sẽ mang lại sự hỗ trợ thiết thực cho nhân viên quản trị mạng. Hầu hết các hệ thống có hai đặc điểm chung là tính đa dạng và thay đổi, điều này đã được các nhà nghiên cứu tìm hiểu và chứng minh. Công việc nghiên cứu và triển khai hệ thống giám sát, phát hiện và phòng chống xâm nhập mạng tồn tại các yếu tố: nhanh chóng, chính xác, trực quan, tiện lợi và linh động là một vấn đề cấp thiết trong thực tế.

Phát triển một hệ thống giám sát trực quan để theo dõi các diễn biến trên mạng như: lưu lượng ra/vào một Server, Switch hoặc hoạt động của CPU, bộ nhớ,… sẽ giúp

người quản trị mạng có những phân tích rõ ràng, nhanh chóng để đưa ra quyết định ứng phó kịp thời khi bị hệ thống bị xâm nhập.

IDS căn cứ vào các dấu hiệu tấn công được triển khai có thể giúp phát hiện một cách nhanh chóng các cuộc tấn công mạng. Khi kết hợp hệ thống với tường lửa sẽ có thể chống lại các cuộc tấn công xâm nhập. Dù vậy, dấu hiệu của từng kiểu tấn công mỗi ngày một tinh vi phức tạp hơn thì hệ thống càng phải được cập nhật thường xuyên những dấu hiệu mới để có thể phát hiện nhanh chóng các bất thường trên mạng, người quản trị mạng có thể dựa vào những đồ thị trực quan về lưu lượng ra vào hệ thống để có những phản ứng kịp thời.

Hệ thống báo động cũng cần thiết phải triển khai để gửi thông báo cho người quản trị trong một số trường hợp đặc biệt nghiêm trọng như: Server bỗng nhiên ngưng hoạt động, một dịch vụ mạng nào đó ngưng hoạt động hay có tấn công. Hệ thống có thể được triển khai qua nhiều hình thức khác nhau để phát thông báo như: qua Web, E-mail, tin nhắn SMS đến người quản trị mạng.

1.2 Các hình thức tấn công mạng

1.2.1 Phân loại tấn công mạng

Có nhiều loại tấn công mạng khác nhau theo đó mức độ nguy hiểm cũng khác nhau với từng loại, tuy nhiên có bốn loại tấn công nguy hiểm nhất mang sức mạnh phá hủy hạ tầng mạng một doanh nghiệp như [2]:

1. **Tấn công từ chối dịch vụ (DDoS)**

Đây là loại tấn công rất phổ biến và cũng là một trong những tấn công mạng có khả năng phá hoại website doanh nghiệp lớn nhất. Những cuộc tấn công bằng DDoS đe dọa trực tiếp và ngay lập tức đến sự hoạt động của các trung tâm dữ liệu của các doanh nghiệp. Những kẻ tấn công mạng sẽ có nhiều cơ hội để xây dựng các mạng botnet rộng lớn và sau đó kích hoạt các tấn công DDoS toàn cầu bởi sự gia tăng không ngừng của các thiết bị có kết nối IoT với bảo mật kém. Một chứng minh rõ ràng của điều này là sự cố "tắc nghẽn mạng" vào tháng 6, 2019 đã làm cho Google Cloud phải ngừng hoạt động và nó cũng làm ít nhất 16 sản phẩm vệ tinh của Google phải ngừng hoạt động theo, các sản phẩm này chính là: Gmail, G-Suite, Google Cloud, Google Drive, Google Docs và YouTube…

2. Mã hóa dữ liệu tống tiền (tấn công Ransomeware)

Cơ sở hạ tầng doanh nghiệp là mục tiêu được những tội phạm mạng nhắm vào bằng ransomware vì thiệt hại của nó có thể ảnh hưởng lâu dài trên diện rộng. Dẫn chứng cho sự tấn công này là công ty lưu trữ Nayana (Hàn Quốc) bị tấn công bằng ransomware làm cho hàng nghìn website của khách hàng lưu trữ trên máy chủ của công ty bị đình trệ hoạt động suốt nhiều tuần liền. Mặc dù đã chi trả một triệu đô la Mỹ tiền chuộc, song vẫn còn nhiều dịch vụ chưa thật sự được phục hồi. Ngoài việc đe dọa dữ liệu của khách hàng trên máy chủ nhà cung cấp dịch vụ, loại tấn công này còn làm giảm sút lòng tin của khách hàng vào doanh nghiệp kinh doanh. Không chỉ có nguy cơ bị phát tán dữ liệu một cách công khai ở khắp mọi nơi, dữ liệu còn bị thay đổi, làm đe dọa đến tính toàn vẹn vốn có của dữ liệu.

3. Tấn công từ bên thứ ba

Vì có thêm một bên thứ ba khác mà các nhà cung cấp dịch vụ bảo mật thường hay gặp phải sự cố khi quản lý bảo mật cho trung tâm dữ liệu của khách hàng. Với mục đích gây hại cho các doanh nghiệp, những kẻ tấn công thường nhắm đến mục tiêu là những nhà cung cấp dịch vụ bảo mật. Cụ thể là vào năm 2019, tổ chức của NordVPN (công ty cung cấp dịch vụ mạng riêng ảo có tính năng bảo mật hàng đầu được đông đảo các doanh nghiệp sử dụng để bảo vệ những dữ liệu nhạy cảm) đã thừa nhận việc một trong số những trung tâm dữ liệu của họ bị tấn công vào năm 2018. Sự việc xảy ra khi họ cài đặt hệ thống truy cập từ xa bởi một bên thứ ba mà chưa thông báo nó cho khách hàng làm dẫn đến tình trạng máy chủ thiếu an toàn.

4. Tấn công vào ứng dụng web do mật khẩu yếu

Tuy không gây ảnh hưởng trực tiếp đến những dịch vụ của trung tâm dữ liệu nhưng bù lại những tấn công vào các ứng dụng web hoặc máy chủ, cụ thể như các form đăng nhập vào ứng dụng web vẫn có thể gây ra gián đoạn các dịch vụ web. Việc đặt mật khẩu yếu, kém hoặc mật khẩu dễ đoán của người dùng là nguyên nhân cho những loại tấn công này. Chúng có thể được thực hiện thông qua các cuộc tấn công đến từ ứng dụng web. Với những mục được nhắm từ trước việc tấn công tốn ít băng thông hơn nhưng nó vẫn có thể khiến các dịch vụ web ngừng hoạt động một cách dễ dàng.

Ngoài ra vẫn còn một số loại tấn công khác được sử dụng đến thời điểm hiện nay như:

- **Tấn công điểm cuối (Endpoint attacks)**: Các cuộc tấn công điểm cuối đã đạt được quyền truy cập trái phép vào các thiết bị người dùng, máy chủ hoặc các điểm cuối khác. Loại hình tấn công này gây ảnh hưởng đến các thiết bị khác bằng cách lây nhiễm bằng phần mềm độc hại.
- **Tấn công bằng các phần mềm độc hại (Malware attacks)**: loại tấn công này lây nhiễm tài nguyên CNTT với phần mềm độc hại, cho phép kẻ tấn công thỏa hiệp với các hệ thống, đánh cắp dữ liệu và gây tổn hại đến hại thống. Chúng bao gồm các cuộc tấn công Ransomware (mã độc tống tiền), Spyware (phần mềm gián điệp), Virus, Worm (phần mềm độc hại lây lan với tốc độ nhanh).
- **Các lỗ hổng, khai thác và tấn công (Vulnerabilities, exploits and attacks)**: các lỗ hổng khai thác trong phần mềm được sử dụng trong tổ chức, để có quyền truy cập trái phép, thỏa hiệp hoặc hệ thống phá hoại.
- **Chiến dịch tấn công sử dụng kỹ thuật cao (Advanced persistent threats)**, đây là những mối đe dọa nhiều lớp phức tạp, không chỉ bao gồm các cuộc tấn công mạng mà còn cả các loại tấn công khác.

Trong một cuộc tấn công mạng, những kẻ tấn công tập trung vào việc thâm nhập chu vi mạng của công ty và đạt được quyền truy cập vào các hệ thống nội bộ. Thông thường, khi kẻ tấn công vào được bên trong hệ thống chúng sẽ kết hợp các loại tấn công khác, ví dụ như ảnh hưởng đến điểm cuối, lan truyền phần mềm độc hại hoặc khai thác lỗ hổng trong một hệ thống trong mạng.

1.2.2 Các kỹ thuật tấn công

Bên cạnh các loại tấn công trên, các kỹ thuật tấn công mạng cũng chuyển biến khôn lường từ các kỹ thuật cơ bản, tuy nhiên mức độ nguy hiểm khi bị tấn công khó lường trước được và gây ra những hậu quả nghiêm trọng nếu không được phát hiện kịp thời. Một số các kỹ thuật tấn công mạng phổ biến như [3]:

- **Truy cập trái phép (Unauthorized access)**: Truy cập trái phép đề cập đến những kẻ tấn công truy cập mạng mà không nhận được sự cho phép. Các nguyên nhân của cuộc tấn công truy cập trái phép thường là do mật khẩu yếu,

thiếu bảo vệ các tài khoản xã hội, các tài khoản bị xâm nhập trước đó và các mối đe dọa nội bộ.

- **Từ chối dịch vụ (Distributed Denial of Service - DDoS):** Những kẻ tấn công xây dựng botnet và sử dụng chúng để hướng lưu lượng truy cập sai vào mạng hoặc máy chủ. DDoS có thể xảy ra ở cấp độ mạng, ví dụ bằng cách gửi khối lượng khổng lồ của các gói SYN / ACK có thể áp đảo một máy chủ hoặc ở cấp độ ứng dụng.
- **Tấn công Man-in-the-Middle (MitM)**: loại tấn công này liên quan đến những kẻ tấn công chặn lưu lượng, giữa các trang web và trang bên ngoài hoặc bên trong mạng. Nếu các giao thức truyền thông không được bảo mật hoặc kẻ tấn công tìm cách vượt qua bảo mật, họ có thể đánh cắp dữ liệu đang được truyền đi, có được thông tin đăng nhập của người dùng và chiếm quyền điều khiển tài khoản của họ.
- **Tấn công cơ sở dữ liệu (SQL Injection)**: Nhiều trang web chấp nhận đầu vào của người dùng và không xác thực các đầu vào đó. Những kẻ tấn công sau đó có thể điền vào một biểu mẫu hoặc thực hiện cuộc gọi API, truyền mã độc thay vì các giá trị dữ liệu dự kiến. Mã được thực thi trên máy chủ và cho phép kẻ tấn công lợi dụng nó.
- **Tấn công leo thang đặc quyền (Privilege Escalation)**: Khi kẻ tấn công xâm nhập mạng, họ có thể sử dụng sự leo thang đặc quyền để mở rộng tầm với của họ. Sự leo thang đặc quyền ngang liên quan đến những kẻ tấn công đạt được quyền truy cập vào các hệ thống bổ sung, liền kề và leo thang thẳng đứng có nghĩa là những kẻ tấn công có được một mức đặc quyền cao hơn cho cùng một hệ thống.
- **Tấn công nội bộ (Insider threats)**: hay còn gọi là mối đe dọa nội bộ. Đây là mối đe dọa đối với tổ chức đến từ những người thuộc tổ chức đó, chẳng hạn như nhân viên, nhân viên cũ, nhà thầu hoặc cộng sự kinh doanh, những người có thông tin nội bộ liên quan đến các phương thức bảo mật, dữ liệu và hệ thống máy tính của tổ chức. Các mối đe dọa nội bộ có thể khó phát hiện và bảo vệ chống lại, vì người trong cuộc không cần thâm nhập mạng để gây hại. Các công nghệ mới như phân tích hành vi người dùng (UEBA) có thể giúp xác

định hành vi đáng ngờ hoặc bất thường của người dùng nội bộ, có thể giúp xác định các cuộc tấn công nội bộ.

1.3 Dấu hiệu nhận diện một cuộc tấn công mạng

Theo tài liệu [3] một cuộc tấn công mạng có thể được nhận biết thông qua các dấu hiệu sau

1.3.1 Dựa vào các gói tin (packets)

Nguồn gốc và biểu hiện của một gói tin có thể giúp dự đoán những kết quả của một gói tin được phân phối. Một gói có thể mang một đoạn mã không giống phần mềm độc hại, nhưng khi được lắp lại với các thành phần khác, tạo thành một chương trình phần mềm độc hại. Hoặc một gói có thể chứa dữ liệu để đột nhập vào hệ thống với ID giả có thể không bị phát hiện trong nhiều tuần hoặc vài tháng. Người sở hữu phần mềm độc hại sử dụng tất cả các loại thủ thuật để tránh bị phát hiện.

Bằng cách nhìn vào lưu lượng gói, có thể theo dõi hành vi của gói đó. Ngoài ra, có thể xác định một số loại gói nhất định đang "đánh hơi xung quanh" các khu vực cụ thể vài phút mỗi ngày hoặc vài giờ một tuần. Nếu để ý và quan sát các gói tin, sẽ có thể dễ dàng nhận thấy một cuộc tấn công sắp xảy ra. Thêm vào đó, có thể xác định một gói có khả năng đáng ngờ nếu nó đang giao tiếp với một máy chủ proxy ẩn danh hoặc với các máy chủ nằm ở các quốc gia giả mạo. Lĩnh vực nghiên cứu này, được gọi là phân tích hành vi, đang ngày càng trở nên quan trọng trong an ninh mạng. Các chuyên gia an ninh mạng phải xem xét các kỹ thuật mới để bảo vệ dữ liệu của tổ chức, các nhà cung cấp dịch vụ mạng toàn cầu cần có sự nhìn nhận về chuyển động của một lượng lưu lượng truy cập khổng lồ.

Nghiên cứu các gói tin và lưu lượng truy cập đang trở nên quan trọng hơn khi ngày càng có nhiều các đối tượng, thiết bị và máy móc được kết nối thông qua Internet of Things (IoT). Ngoài máy tính, những kẻ tấn công có thể tìm cách chiếm quyền điều khiển các bộ định tuyến, thiết bị và các thiết bị khác để thực hiện các cuộc tấn công mạng.

Bất cứ ai cũng có thể là một đồng phạm vô tình trong một cuộc tấn công. Một người dùng ngồi vui vẻ trên một tài khoản FaceTime hoặc Skype với đồng nghiệp thậm chí có thể không nhận ra điện thoại thông minh hoặc máy tính xách tay của họ đã bị biến thành một bot tấn công. Người dùng có thể nhận thấy rằng đột nhiên hình

ảnh hoặc video bị mờ đi hoặc dừng lại, hoặc âm thanh bị cắt xén. Trong một số trường hợp, những trục trặc này là tác động của một cuộc tấn công đang được tiến hành.

Tin tặc là chuyên gia tìm kiếm các lỗ hổng trong các thiết bị hàng ngày thậm chí trước khi các nhà sản xuất sản phẩm nhận thấy những sai sót. Và một khi tin tặc tìm thấy lỗ hổng, họ sẽ tìm cách khai thác nó. Khi các nhà phân tích an ninh mạng phát hiện các chỉ số về thỏa hiệp (IOC), họ thông báo cho các mục tiêu - bao gồm các tổ chức không phải là khách hàng. Các công ty không biết họ đang bị tấn công cho đến khi các nhà phân tích an ninh mạng nói với họ rằng, giả sử dữ liệu đang được chuyển từ mạng của họ đến một máy chủ ở Đông Âu hoặc một số điểm đến ngoài ý muốn khác.

1.3.2 Dựa trên các cảnh báo từ hệ thống IDS

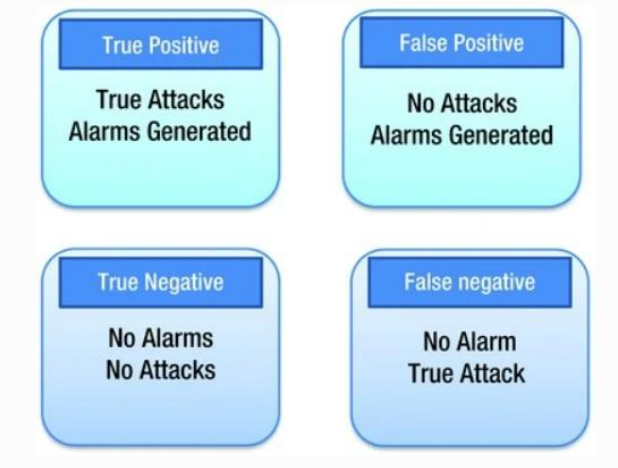
Ngay cả khi có một hệ thống tốt trong việc ghi nhận lịch sử tất cả các lưu lượng truy cập vào mạng nội bộ, thì việc kiểm tra một loạt các lịch sử này chưa mang lại hiệu quả cao. Không thể phân biệt thủ công giữa một gói độc hại và gói mạng tốt. Ngay cả với sự giúp đỡ của máy tính, đây là một công việc chuyên sâu đòi hỏi nhiều sức mạnh xử lý. Trong những năm qua, trong thế giới được kết nối chủ yếu thông qua nhiều phương tiện khác nhau bao gồm máy tính bảng và điện thoại di động, những người xấu với ý định xấu nhắm vào các tập đoàn khác nhau cũng như các cá nhân. Vì không thể phát hiện các cuộc tấn công như vậy theo cách thủ công để ngăn chặn hoặc giảm thiểu chúng, nên bắt buộc phải có một công cụ tự động để giúp giám sát hệ thống cho các cuộc tấn công. IDS đã trở thành một công cụ hữu ích để cung cấp giám sát này. Một số các thuật ngữ cảnh báo trong hệ thống IDS:

Dương tính thật (True Positives): Đây là những cảnh báo rằng một cái gì đó đúng và nó thật sự đúng. Ví dụ: IDS tìm thấy một gói có chứa mã độc và thực sự đúng là gói có mã độc, như được xác nhận bởi điều tra.

Âm tính thật (True Negatives): Đây là những cảnh báo rằng một cái gì đó là không đúng và nó thật sự không đúng. Ví dụ: IDS tìm thấy một gói là không có vấn đề gì và nó thực sự không có vấn đề gì.

Dương tính giả (False Positives): Đây là những cảnh báo chỉ ra rằng một cái gì đó đúng với một gói tin nhưng thực chất nó là không đúng. Ví dụ: IDS tìm thấy một gói có mã độc nhưng nó thực sự là một mã không độc.

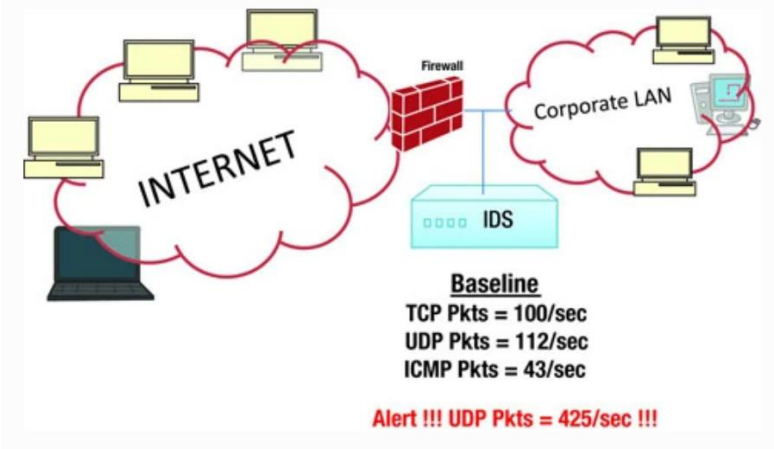
Âm tính giả (False Negatives): Đây là những cảnh báo rằng một cái gì đó là không đúng khi thực chất nó là đúng. Ví dụ: IDS thấy rằng một gói không có bất kỳ mã độc nào nhưng nó thực sự chứa một mã độc, như được tìm thấy thông qua điều tra.



Hình 1.1: Định nghĩa các cảnh báo trong hệ thống IDS [1]

1.3.3 Phát hiện dựa trên dòng dữ liệu bất thường

Phát hiện dựa trên bất thường bảo vệ chống lại các mối đe dọa chưa biết. Nếu bất kỳ lưu lượng truy cập nào được tìm thấy là bất thường từ đường cơ sở, thì một cảnh báo được kích hoạt bởi các ID bị nghi ngờ là sự xâm nhập. IDP trước tiên tạo ra một cấu hình cơ sở thể hiện hành vi bình thường của lưu lượng truy cập. Hồ sơ cơ sở được tạo bằng cách cho phép hệ thống IDS tìm hiểu lưu lượng truy cập trong một khoảng thời gian để IDP có thể nghiên cứu hành vi di chuyển trong giờ cao điểm, giờ không cao điểm, giờ đêm, giờ đầu kinh doanh. Sau khi học, sự di chuyển của lưu lượng được thu thập trong khoảng thời gian được nghiên cứu thống kê và một hồ sơ cơ sở được tạo ra. Khi IDS được thay đổi từ chế độ học tập sang chế độ phát hiện/phòng ngừa, nó bắt đầu so sánh lưu lượng thông thường với cấu hình được tạo và nếu tìm thấy bất kỳ sự bất thường hoặc sai lệch nào so với hồ sơ cơ sở sự xâm nhập sẽ được ngăn chặn nếu nó được cấu hình cho chế độ phòng ngừa. Cấu hình tùy chỉnh cũng có thể được tạo cho hành vi lưu lượng truy cập cụ thể, chẳng hạn như số lượng e-mail được gửi bởi các nỗ lực truy cập của người dùng. Hình 1.2 dưới đây mô tả sự bất thường được phát hiện trên dòng dữ liệu.



Hình 1.2: Phát hiện bất thường dựa trên dòng dữ liệu [1]

Một số dấu hiệu về các hành vi đáng ngờ: quá nhiều phiên đăng nhập Telnet trong một ngày, lưu lượng truy cập HTTP trên một cổng không chuẩn, lưu lượng SNMP bị quá tải.

Để phát hiện xâm nhập hiệu quả, IDS phải có hồ sơ cơ sở mạnh mẽ bao gồm toàn bộ mạng của tổ chức và các phân khúc. Nó sẽ bao gồm hành vi di chuyển bình thường của tất cả các thành phần nhằm mục đích được bao phủ bởi hệ thống phát hiện và phòng ngừa xâm nhập. Hồ sơ cơ sở có thể thay đổi về độ phức tạp từ một nội dung đơn giản đến toàn diện, tùy thuộc vào các đặc điểm của mạng và các thành phần của nó. Ví dụ: một hồ sơ có thể bao gồm các dữ liệu sau:

- Một ứng dụng web đã đăng nhập từ xa bởi một bộ người dùng cụ thể
- Một ứng dụng có thiết kế mật khẩu được chấp nhận cụ thể
- Lưu lượng truy cập trong giờ cao điểm và giờ không cao theo quy định của tổ chức
- Mẫu kết nối từ mạng đối tác bên ngoài
- Kết nối từ một bộ thiết bị di động với máy chủ cơ sở dữ liệu

Thách thức của phương pháp phát hiện dựa trên sự bất thường là tạo ra một hồ sơ hiệu quả. Hồ sơ ban đầu, đôi khi được gọi là "hồ sơ đào tạo", được tạo bằng cách nghiên cứu mô hình lưu lượng trong một khoảng thời gian. Yếu tố thời gian có thể thay đổi từ tổ chức này sang tổ chức khác. Nó có thể là một vài giờ đến vài ngày. Khi cấu hình này được tạo, ID được đưa vào chế độ phát hiện và mỗi khi có một gói tin, một mẫu được khớp với cấu hình cơ sở. Đường cơ sở này có thể được thay đổi khi được yêu cầu dựa trên hành vi di chuyển. Nếu bất kỳ hoạt động độc hại nào đã tồn tại

ngay từ đầu, trong khi xây dựng hồ sơ cơ bản, hoạt động này cũng sẽ trở thành một phần của hồ sơ cơ sở và loại hoạt động độc hại như vậy sẽ không bị phát hiện. Do đó, phát hiện bất thường không nhất thiết phải phát hiện từng cuộc tấn công chưa biết. Giới hạn dựa trên hồ sơ cơ sở được tạo.

Các loại bất thường

Hệ thống phòng ngừa và phát hiện xâm nhập dựa trên dị thường (IDP) bảo vệ khỏi sự bất thường gây ra do vi phạm các giao thức và tải trọng ứng dụng. Nó cũng bảo vệ chống lại các cuộc tấn công từ chối dịch vụ và các cuộc tấn công tràn bộ đệm.

Giao thức bất thường

Giao thức bất thường đề cập đến sự bất thường trong định dạng giao thức và hành vi giao thức liên quan đến các tiêu chuẩn và thông số kỹ thuật của Internet. Có nhiều khía cạnh trong giao thức TCP và IP cần được theo dõi, ví dụ, các cờ khác nhau, SYN, ACK, FIN, sự kết hợp của chúng và các cờ dự trữ. Cách phân mảnh IP khi lắp ráp lại được thực hiện theo tiêu chuẩn. Nếu sự bất thường này không được phát hiện bởi ID, máy chủ cuối có thể không xử lý được các gói này và điều này có thể dẫn đến sự cố của hệ thống. Ở cấp độ ứng dụng, IDP phải có khả năng thực hiện phân tích giao thức sâu hơn để hiểu sự bất thường về giao thức ứng dụng. Nó cũng đòi hỏi sự hiểu biết sâu sắc về ngữ nghĩa ứng dụng để phát hiện sự bất thường của tải trọng ứng dụng. Một số ví dụ khác bao gồm: Phân đoạn TCP bất thường và các cờ TCP kết hợp, kiểm tra hoạt động bất thường, cờ phân mảnh IP không chính xác, các lệnh giao thức bất hợp pháp và việc sử dụng nó, chạy giao thức trên cổng không chuẩn, sử dụng sai các dịch vụ giao thức.

Phát hiện bất thường thống kê

Từ chối dịch vụ (DoS) và từ chối dịch vụ phân tán (DDoS) dẫn đến một loạt lưu lượng truy cập trên mạng diễn ra bất bình. Để khắc phục loại tấn công này, các cấu hình cơ bản được tạo ra trên lưu lượng di chuyển bình thường như được mô tả trước đó là dựa trên mô hình thống kê như Naïve Bayes để xác định các gói bất thường trên mạng. Trong khi học hành vi của lưu lượng mạng, chức năng của mô hình thống kê là tính toán điểm xác suất cho từng gói dữ liệu được coi là lưu lượng bình thường. Điểm số được tính toán dựa trên dữ liệu được lấy mẫu trong một khoảng thời gian và được lưu trữ trong một hồ sơ cơ sở. Một ngưỡng được đặt cho mỗi bộ giao thức và

người dùng. Khi IDS ở chế độ giám sát, dữ liệu được kiểm tra so với đường cơ sở và ngưỡng. Bất cứ khi nào một gói bất thường được phát hiện và điểm số trên ngưỡng, thì một cảnh báo được kích hoạt. Quá trình cảnh cáo sẽ chỉ báo khi dữ liệu được tìm thấy là bất thường trong một khoảng thời gian đủ, nếu không IDP sẽ chỉ đơn giản bỏ qua dấu vết. Ngưỡng có thể được đặt cho các cấu hình khác nhau, cho các giao thức khác nhau và cho người dùng khác nhau.

Khi IDP ở chế độ giám sát, nếu có bất cứ điều gì bất thường đối với đường cơ sở, hệ thống sẽ tạo ra một cảnh báo nhưng kết quả phân tích xác nhận rằng cảnh báo được tìm thấy là dương tính giả. Là một quản trị viên bảo mật, người ta mong đợi một loại hành vi tương tự xuất hiện mỗi ngày và để giảm thiểu chi tiêu. Lưu lượng vẫn được coi là bình thường và bất cứ điều gì vượt quá ngưỡng được thiết lập sẽ được xem là một sự xâm nhập. Ngưỡng cũng có thể được đặt cho một tập hợp người dùng hoặc tập hợp các giao thức.

Hồ sơ dựa trên các biện pháp thống kê có thể phát hiện một số dị thường DoS dựa trên các phân phối hoặc các vụ nổ ngắn hạn của lưu lượng đỉnh (tức là cao). Các cấu hình cơ sở liên tục được học trong khi hệ thống ở chế độ phát hiện và đường cơ sở được tạo lại để điều chỉnh mô hình lưu lượng thay đổi để tránh dương tính giả.

Bằng cách tạo các hồ sơ khác nhau, các cuộc tấn công DoS có thể được ngăn chặn. Ví dụ [1], đối với mỗi cuộc tấn công DoS, một hồ sơ có thể được tạo. Lỗ hổng hệ thống được nhắm mục tiêu bởi các phân đoạn TCP khi cờ SYN được bật sẽ tạo không gian cho cuộc tấn công DoS, khi đó những tấn công dạng này sẽ được gọi là cuộc tấn công SYN flood. Bất cứ khi nào có lưu lượng truy cập SYN flood trên mạng, các cảm biến IDS có thể phát hiện cuộc tấn công SYN flood bằng cách so sánh lưu lượng mạng với hồ sơ SYN flood do đó cảnh báo một cuộc tấn công SYN flood. Tương tự, hồ sơ UDP flood, hồ sơ phân đoạn dữ liệu TCP hoặc hồ sơ ICMP flood có thể được phát hiện và cảnh báo.

Mặc dù IDS dựa trên bất thường có lợi thế là phát hiện các cuộc tấn công chưa biết, việc xác định các quy tắc cho nó là lại trở nên khó khăn. Mỗi giao thức phải được phân tích, xử lý và so sánh với đường cơ sở. Bất kỳ giao thức tùy chỉnh nào đều làm cho nó trở nên không phù hợp.

Một cạm bẫy lớn khác của phát hiện bất thường là xác định lưu lượng bình thường trong khi tạo ra một đường cơ sở. Lưu lượng bình thường phải được sạch sẽ và không nên có bất kỳ hoạt động độc hại nào trong mạng. Trong trường hợp có bất kỳ hoạt động độc hại nào trong quá trình học tập, thì hồ sơ cơ sở tìm hiểu điều này và khiến việc phát hiện sự xâm nhập này khó khăn hơn hoặc thậm chí có thể không phát hiện sự xâm nhập của lưu lượng độc hại như vậy. Ví dụ, các cuộc tấn công trinh sát như dấu vân tay hoặc thư mục, tuân thủ giao thức mạng, dễ dàng không được chú ý vì nó tuân thủ các giới hạn giao thức và tải trọng.

Phát hiện phân tích trạng thái giao thức

Phương pháp này tương tự như phát hiện dựa trên bất thường, ngoại trừ các hồ sơ được tạo bởi các nhà cung cấp cung cấp thiết bị cảm biến (IDP). Các hồ sơ được xác định trước và được tạo thành từ hoạt động lưu lượng mạng an toàn được chấp nhận chung theo quy định của các tiêu chuẩn. "Trạng thái" có nghĩa là IDP có khả năng theo dõi trạng thái của giao thức cả trong lớp mạng và lớp ứng dụng. Ví dụ, trong trường hợp trạng thái thiết lập kết nối TCP, ID nên nhớ tất cả các trạng thái kết nối. Tương tự, trong trường hợp xác thực, phiên kết nối ban đầu ở trạng thái trái phép và IDS nên nhớ các trạng thái này. Sau khi trao đổi một số thông tin giữa hai bên máy khách và máy chủ, người dùng được xác thực và được cho phép truy cập vào mạng. Trong giai đoạn này, việc di chuyển là an toàn và IDP nên nhớ trạng thái hoặc nó sẽ dẫn đến dương tính thật.

Phương pháp phát hiện bất thường trong giao thức trạng thái sử dụng các cấu hình đã được tạo dựa trên các tiêu chuẩn và thông số kỹ thuật được chỉ định bởi nhà cung cấp thường tuân thủ hầu hết các giao thức từ các cơ quan tiêu chuẩn (Lực lượng đặc nhiệm kỹ thuật Internet). Nếu bất kỳ nhà cung cấp nào đã thực hiện các giao thức, với sự thay đổi của các tiêu chuẩn, nó sẽ gây khó khăn cho IDP trong việc phát hiện và phân tích các trạng thái. Trong các trường hợp như vậy, các mô hình giao thức IDPS cũng cần được cập nhật cho các thay đổi giao thức tùy chỉnh.

Hạn chế chính của phương pháp này là chúng rất tốn kém về quá trình và bộ nhớ như nhiều giao thức và IDP phải theo dõi đồng thời trạng thái của chúng. Một vấn đề khác là nếu một cuộc tấn công nằm trong hành vi giao thức thường được chấp

nhận, thì nó có thể đi qua. Nếu việc triển khai giao thức thay đổi từ hệ điều hành thì IDP có thể không hoạt động tốt trong việc phát hiện các xâm nhập.

1.4 Giải pháp phát hiện và phòng chống xâm nhập

Ngoài việc nắm rõ các dấu hiệu của cuộc tấn công, ta cũng cần phải trang bị những giải pháp phát hiện và phòng chống chống xâm nhập mạng [3] như:

1.4.1 Phân chia mạng

Một phần cơ bản của việc tránh các mối đe dọa bảo mật mạng là phân chia mạng thành các khu vực dựa trên các yêu cầu bảo mật. Điều này có thể được thực hiện bằng cách sử dụng các mạng con trong cùng một mạng hoặc bằng cách tạo các mạng cục bộ ảo (VLANS), mỗi loại hoạt động giống như một mạng riêng biệt. Phân khúc giới hạn tác động tiềm ẩn của một cuộc tấn công vào một khu vực và yêu cầu kẻ tấn công thực hiện các biện pháp đặc biệt để thâm nhập và có quyền truy cập vào các khu vực mạng khác.

1.4.2 Điều chỉnh quyền truy cập Internet qua máy chủ proxy

Không cho phép người dùng mạng truy cập Internet không được kiểm tra. Vượt qua tất cả các yêu cầu thông qua một proxy trong suốt và sử dụng nó để kiểm soát và giám sát hành vi của người dùng. Đảm bảo rằng các kết nối bên ngoài thực sự được thực hiện bởi một con người và không phải là một bot hoặc cơ chế tự động khác. Các tên miền danh sách trắng để đảm bảo người dùng công ty chỉ có thể truy cập các trang web đã được phê duyệt một cách rõ ràng.

1.4.3 Đặt thiết bị bảo mật chính xác

Tường lửa thường đặt tại vùng biên giới của hệ thống máy tính hay hệ thống mạng. Hoặc tùy vào loại tường lửa như tường lửa cá nhân (Internal Firewall) hay tường lửa hệ thống (System Firewall) mà vị trí đặt cũng như chức năng sẽ có sự khác nhau. Với Internal Firewall thì sau khi được cài đặt thì nó sẽ chiếm giữ việc quản lý tất cả các thông tin đi ra hay đi vào máy tính cá nhân của người dùng. System Firewall sẽ được lắp đặt ngay sau thiết bị kết nối WAN, như Router sử dụng các kênh thuê riêng (leased-line), hay Rounter ADSL … Nếu không thể triển khai tường lửa chính thức ở khắp mọi nơi, hãy sử dụng chức năng tường lửa tích hợp của các thiết bị chuyển mạch và bộ định tuyến, sử dụng hardware firewall hoặc thiết bị chống DDoS

để chống tấn công DDoS… Xem xét cẩn thận nơi đặt các thiết bị chiến lược như bộ cân bằng tải - nếu chúng nằm ngoài vùng Demilitarized Zone (DMZ) – vùng nằm giữa LAN và internet (chứa các server và cung cấp các dịch vụ cho các host trong LAN cũng như các host từ các LAN ở bên ngoài), chúng sẽ không được bảo vệ bởi bộ máy bảo mật mạng.

1.4.4 Sử dụng NAT (Network Address Translation)

NAT cho phép dịch địa chỉ IP nội bộ thành địa chỉ có thể truy cập trên các mạng công cộng. Có thể sử dụng nó để kết nối nhiều máy tính với Internet bằng một địa chỉ IP duy nhất. Điều này cung cấp thêm một lớp bảo mật, bởi vì bất kỳ lưu lượng truy cập nào cũng phải đi qua thiết bị NAT.

1.4.5 Giám sát lưu lượng mạng

Đảm bảo có khả năng hiển thị hoàn toàn lưu lượng truy cập đến, đi và mạng nội bộ, với khả năng tự động phát hiện các mối đe dọa và hiểu bối cảnh và tác động của họ. Kết hợp dữ liệu từ các công cụ bảo mật khác nhau để có được một hình ảnh rõ ràng về những gì đang xảy ra trên mạng, nhận ra rằng nhiều cuộc tấn công trải rộng trên nhiều hệ thống CNTT, tài khoản người dùng.

Đạt được mức độ hiển thị này có thể khó khăn với các công cụ bảo mật truyền thống. CYNET 360 là một giải pháp bảo mật tích hợp cung cấp các phân tích mạng tiên tiến, liên tục theo dõi lưu lượng mạng, tự động phát hiện hoạt động độc hại và phản ứng với nó tự động hoặc vượt qua thông tin về ngữ cảnh cho nhân viên bảo mật.

1.4.6 Sử dụng công nghệ "đánh lừa"

Không có biện pháp bảo vệ mạng nào thành công 100% và những kẻ tấn công cuối cùng sẽ thành công trong việc thâm nhập mạng. Nhận ra điều này và đặt công nghệ đánh lừa để tạo ra những mồi nhử trên mạng, những kẻ tấn công bị cám dỗ sẽ "tấn công" và nó cho phép chuyên viên quản trị quan sát kế hoạch và kỹ thuật của kẻ tấn công. Có thể sử dụng mồi nhử để phát hiện các mối đe dọa trong tất cả các giai đoạn của vòng đời tấn công: Tệp dữ liệu, thông tin xác thực và kết nối mạng.

1.5 Các thuật toán học máy trong hệ thống phát hiện xâm nhập mạng

Học máy (Machine Learning - ML) là một tập hợp con của AI bao gồm tất cả các phương thức và thuật toán cho phép các máy để tìm hiểu tự động bằng các mô

hình toán học để trích xuất thông tin hữu ích từ các bộ dữ liệu lớn. ML phổ biến nhất (còn gọi là học tập nông) được sử dụng cho IDS là [4] cây quyết định, K-nearest neighbors (KNN), mạng thần kinh nhân tạo (ANN), máy vectơ hỗ trợ (SVM), cụm K-Mean, và mạng học tập nhanh.

1.5.1 Decision Tree (DT)

DT là một trong những thuật toán ML có giám sát cơ bản được sử dụng để phân loại và hồi quy bộ dữ liệu đã cho bằng cách áp dụng hàng loạt các quyết định (quy tắc). Mô hình có cấu trúc cây thông thường với các nút, cành và lá. Mỗi nút biểu thị một thuộc tính hoặc một tính năng. Một nhánh đại diện cho một quyết định hoặc quy tắc trong khi mỗi lá đại diện cho một kết quả hoặc nhãn lớp. Thuật toán DT sẽ tự động chọn các tính năng tốt nhất để xây dựng nên cây quyết định và sau đó thực hiện thao tác cắt tỉa để loại bỏ các nhánh không liên quan từ cây để tránh sự phù hợp. Các mô hình DT phổ biến nhất là Cart, C4.5 và ID3. Nhiều thuật toán học tập nâng cao như rừng ngẫu nhiên (Random Forest) và XGBoost (eXtreme Gradient Boosting) được xây dựng từ nhiều cây quyết định.

1.5.2 **K-***Nearest Neighbor (KNN)*

KNN là một trong những thuật toán ML có giám sát đơn giản nhất sử dụng ý tưởng về "điểm dữ liệu tương tự tồn tại gần nhau trong một không gian" để dự đoán lớp của một mẫu dữ liệu nhất định. Nó xác định đầu ra một mẫu dựa trên thông tin của K điểm dữ liệu trong tập huấn luyện gần nó nhất (hay còn gọi là K-lân cận). Trong thuật toán KNN, tham số K ảnh hưởng đến hiệu suất của mô hình. Nếu giá trị của K nhỏ, mô hình có thể dễ bị overfitting khi đó mô hình không học được gì từ dữ liệu. Trong khi, khi lựa chọn giá trị K lớn có thể dẫn đến việc phân loại sai đối tượng mẫu. Karatas đã so sánh hiệu suất của các thuật toán ML khác nhau bằng cách sử dụng tập dữ liệu CSE-CIC-IDS2018. Họ giải quyết vấn đề về sự mất cân bằng dữ liệu bằng bằng cách sử dụng kỹ thuật tăng kích thước mẫu (Smote), dẫn đến việc cải thiện tỷ lệ phát hiện cho các mẫu thuộc lớp thiểu số.

1.5.3 Support vector machine

SVM là một thuật toán ML có giám sát dựa trên ý tưởng về một siêu mặt phẳng (hay còn gọi là hyper lane) để phân tách các điểm của dữ liệu. Siêu mặt phẳng này sẽ

chịu trách nhiệm chia không gian thành các miền khác nhau và mỗi miền này sẽ chứa một loại dữ liệu. SVM được sử dụng làm giải pháp của cả các vấn đề tuyến tính và phi tuyến tính. Đối với các vấn đề phi tuyến tính, các hàm kernel được sử dụng. Ý tưởng là đầu tiên ánh xạ một vectơ đầu vào có kích thước thấp thành không gian tính năng có kích thước cao bằng hàm kernel. Tiếp theo, một mặt phẳng cận biên tối đa và tối ưu thu được sẽ hoạt động như một ranh giới quyết định sử dụng các vectơ hỗ trợ.

1.5.4 K-*mean clustering*

Việc phân cụm là một ý tưởng phân chia dữ liệu thành các cụm có ý nghĩa (hoặc nhóm), bằng cách đặt dữ liệu có tính chất giống nhau vào cùng một cụm. K-Mean Clustering là một trong những thuật toán ML không giám sát phổ biến dựa trên centroid (tâm của cụm dữ liệu). K đại diện cho số lượng centroid trong một tập dữ liệu. Khoảng cách từ mỗi điểm dữ liệu tới tâm được tính toán để chỉ định một số điểm dữ liệu nhất định sẽ vào cùng một cụm nào. Ý tưởng chính của thuật toán là làm giảm tổng khoảng cách giữa các điểm dữ liệu và centroid tương ứng của chúng trong một cụm.

Yao [5] đã đề xuất một khung mô hình phát hiện xâm nhập nhiều cấp độ có tên multilevel semi-supervised ML (MSML) để vấn đề về sự mất cân bằng nghiêm trọng của lưu lượng mạng trong các danh mục khác nhau và sự phân bổ không đồng nhất giữa tập huấn luyện và tập kiểm tra trong không gian đặc trưng. Giải pháp được đề xuất bao gồm bốn mô-đun như trích xuất cụm thuần túy, phát hiện mẫu, phân loại fine-grained và cập nhật mô hình. Ý tưởng là nếu một cuộc tấn công không được dán nhãn trong một mô-đun thì nó được chuyển tiếp đến phần tiếp theo để phát hiện. Phương pháp đề xuất đã được thử nghiệm bằng cách sử dụng bộ dữ liệu KDD Cup'99. Kết quả thử nghiệm cho thấy sự vượt trội của mô hình khi dùng để phát hiện các cuộc tấn công ngay cả với các trường hợp nhãn có tỉ lệ thấp trong bộ dữ liệu.

1.5.5 Artificial neural network

ANN cũng là một thuật toán ML có giám sát và được truyền cảm hứng từ hệ thống thần kinh của bộ não con người. ANN được tạo thành từ các yếu tố xử lý gọi là tế bào thần kinh (nút) và các kết nối giữa chúng (cạnh). Các nút này được sắp xếp trong một lớp đầu vào, nhiều lớp ẩn và một lớp đầu ra. Thuật toán lan truyền ngược

hay còn gọi là Backpropagation được sử dụng như một kỹ thuật học tập cho ANN. Ưu điểm chính của việc sử dụng kỹ thuật ANN là khả năng thực hiện mô hình hóa phi tuyến tính bằng cách học tập từ các bộ dữ liệu lớn. Tuy nhiên, vấn đề chính khi đào tạo mô hình với ANN là mức tiêu thụ thời gian cao do tính chất phức tạp của nó, làm chậm quá trình học tập và để đạt được một giải pháp dưới mức tối ưu.

Để khắc phục những hạn chế của ANN, Huang đã đề xuất một ANN mới gọi là một máy học tập cực đoan (Extreme learning machine - ELM). ELM là một mạng thần kinh truyền thẳng với một lớp ẩn duy nhất, sử dụng ngẫu nhiên các trọng lượng (weight) đầu vào và độ lệch (bias) lớp ẩn mà không điều chỉnh và xác định trọng lượng đầu ra theo cách phân tích. Dựa trên ý tưởng của ELM, Li đã đề xuất một mạng học nhanh (Fast learning network - FLN). FLN dựa trên việc kết nối mạng thần kinh chuyển tiếp nhiều lớp và một mạng thần kinh chuyển tiếp dữ liệu một lớp song song. FLN cho thấy hiệu suất và sự ổn định khá hợp lý bằng cách sử dụng một số ít các nút ẩn và sử dụng ít thời gian hơn.

Ali [6] đã đề cập đến vấn đề IDS bằng cách đề xuất một mô hình dựa trên mạng học nhanh và tối ưu hóa bầy hạt được đặt tên là PSO-FLN, mô hình đã thử nghiệm bằng cách sử dụng bộ dữ liệu KDD Cup'99. Thử nghiệm được diễn ra bằng cách so sánh FLN với các thuật toán tối ưu hóa khác nhau. Kết quả cho thấy PSO-FLN vượt trội hơn các mô hình FLN khác với các thuật toán tối ưu hóa khác nhau làm thuật toán di truyền, tối ưu hóa tìm kiếm Harmoney và tối ưu hóa dựa trên việc học. Họ cũng chứng minh rằng việc tăng số lượng tế bào thần kinh trong lớp ẩn làm tăng độ chính xác.

1.5.6 Ensemble methods

Ý tưởng chính đằng sau các phương thức Ensemble là tổng hợp các kết quả dự đoán của nhiều model thành một model cuối cùng. Vì mỗi bộ phân loại có một số điểm mạnh và điểm yếu. Một số người có thể thực hiện tốt để phát hiện một loại tấn công cụ thể và cho thấy hiệu suất kém về các loại tấn công khác. Cách tiếp cận Ensemble là kết hợp các phân loại yếu bằng cách đào tạo nhiều phân loại và sau đó tạo thành một bộ phân loại mạnh hơn bằng cách chọn bằng thuật toán đã được bỏ phiếu.

Shen đã đề xuất nghiên cứu sử dụng phương pháp hòa đồng bằng cách chọn Elm làm bộ phân loại cơ sở. Để tối ưu hóa phương pháp đề xuất, thuật toán tối ưu hóa BAT được sử dụng trong giai đoạn cắt tỉa. Mô hình đã được thử nghiệm bằng cách sử dụng bộ dữ liệu KDD Cup'99, NSL-KDD và Kyoto. Kết quả thử nghiệm cho thấy nhiều Elms kết hợp vượt trội hơn về mặt hiệu suất so với các phương pháp khác.

Gao đã đề xuất một mô hình nghiên cứu của mình bằng cách sử dụng một số phân loại cơ sở dưới dạng DT, RF, KNN, mạng thần kinh sâu (DNN) và chọn ra bộ phân loại tốt nhất bằng thuật toán bỏ phiếu. Phương pháp được đề xuất đã được xác minh bằng cách thực hiện các thí nghiệm trên bộ dữ liệu NSL-KDD. Kết quả thí nghiệm đã chứng minh hiệu suất cao của mô hình đề xuất bằng cách so sánh với các mô hình khác.

CHƯƠNG 2. CÁC CÔNG TRÌNH LIÊN QUAN

2.1 Một số công trình nghiên cứu tại Việt Nam

Sự phổ biến ngày càng tăng của Internet of Things (IoT) đã tác động đáng kể đến cuộc sống hàng ngày của chúng ta trong vài năm qua. Một mặt, nó mang lại sự tiện lợi, đơn giản và hiệu quả cho chúng ta; mặt khác, các thiết bị dễ bị tấn công mạng do thiếu cơ chế bảo mật vững chắc và hỗ trợ bảo mật phần cứng. Trong bài báo này, nhóm tác giả trình bày IMIDS, một IDS thông minh để bảo vệ các thiết bị IoT. Cốt lõi của IMIDS [7] là một mô hình mạng nơ-ron tích hợp nhẹ để phân loại nhiều mối đe dọa mạng. Để giảm thiểu vấn đề thiếu dữ liệu đào tạo, nhóm tác giả cũng đề xuất một trình tạo dữ liệu tấn công được cung cấp bởi một mạng đối thủ chung có điều kiện. Trong thử nghiệm, Nhóm tác giả chứng minh rằng IMIDS có thể phát hiện 9 kiểu tấn công mạng (ví dụ: backdoor, shellcode, worm) với chỉ số F trung bình là 97,22% và vượt trội so với các đối thủ cạnh tranh. Hơn nữa, hiệu suất phát hiện của IMIDS được cải thiện đáng kể sau khi được đào tạo thêm bởi dữ liệu do trình tạo dữ liệu tấn công của Nhóm tác giả tạo ra. Những kết quả này chứng minh rằng IMIDS có thể là một IDS thực tế cho kịch bản IoT.

An ninh mạng ngày càng trở nên thách thức do sự gia tăng của Internet vạn vật (IoT), nơi một số lượng lớn các thiết bị thông minh, nhỏ bé đẩy hàng nghìn tỷ byte dữ liệu lên Internet. Tuy nhiên, các thiết bị này có nhiều lỗi bảo mật khác nhau do thiếu cơ chế phòng vệ và hỗ trợ bảo mật phần cứng, do đó khiến chúng dễ bị tấn công mạng. Ngoài ra, các cổng kết nối IoT cung cấp các tính năng bảo mật rất hạn chế để phát hiện các mối đe dọa như vậy, đặc biệt là sự vắng mặt của các phương pháp phát hiện xâm nhập được hỗ trợ bởi học sâu. Thật vậy, các mô hình học sâu đòi hỏi sức mạnh tính toán cao vượt quá khả năng của các cổng này. Trong bài báo này [8], Nhóm tác giả giới thiệu Realguard, một NIDS dựa trên DNN hoạt động trực tiếp trên các cổng cục bộ để bảo vệ các thiết bị IoT trong mạng. Tính ưu việt của đề xuất của Nhóm tác giả là nó có thể phát hiện chính xác nhiều cuộc tấn công mạng trong thời gian thực với một dấu vết tính toán nhỏ. Điều này đạt được nhờ cơ chế trích xuất tính năng nhẹ và mô hình phát hiện tấn công hiệu quả được hỗ trợ bởi mạng nơ-ron sâu. Các đánh giá của Nhóm tác giả về bộ dữ liệu thực tế chỉ ra rằng Realguard có thể phát hiện mười kiểu tấn công (ví dụ: quét cổng, Botnet và FTP-Patator) trong thời

gian thực với độ chính xác trung bình là 99,57%, trong khi tốt nhất của các đối thủ cạnh tranh của Nhóm tác giả là 98,85%. Hơn nữa, đề xuất của Nhóm tác giả hoạt động hiệu quả trên các cổng hạn chế tài nguyên (Raspberry PI) với tốc độ xử lý gói cao được báo cáo khoảng 10.600 gói mỗi giây.

Trong bài báo này [9], Nhóm tác giả đề xuất phương pháp tiếp cận mạng nơ-ron tổng hợp để nhiều mô hình mạng nơ-ron có thể được triển khai đồng thời để phát hiện các điểm bất thường trong hệ thống. Ngoài ra, kiến trúc hiệu suất cao tích hợp mạng nơ-ron cho phần cứng có thể cấu hình lại cũng được thiết kế để khai thác tính song song của công nghệ. Nhóm tác giả cẩn thận triển khai bốn mô hình khác nhau để phát hiện các cuộc tấn công SYN, DNS, UNP và ICMP với cùng một cấu hình mạng nơ-ron trên nền tảng NetFPGA. Mạng nơ-ron có thể được thực thi trong đường ống để cải thiện hiệu suất. Các thử nghiệm với tập dữ liệu đã tạo được tiến hành để kiểm tra hiệu suất cả về thông lượng và độ chính xác. Kết quả thử nghiệm của Nhóm tác giả cho thấy hệ thống có thể xử lý các gói với thông lượng lên đến 30,48Gbps và có thể nhận ra 99,98% các gói tấn công với chỉ 0,98% cảnh báo sai.

2.2 Một số công trình nghiên cứu trên thế giới

Ngày nay các IDS đóng một vai trò quan trọng trong các tổ chức vì có rất nhiều cuộc tấn công mạng ảnh hưởng đến các vấn đề bảo mật: bí mật, tính toàn vẹn, tính khả dụng. Hiện nay, có rất nhiều công cụ mã nguồn mở để phát hiện xâm nhập nhưng chúng có cú pháp quy tắc và signature khác nhau, không thể sử dụng trên các công cụ khác nhau. Trong bài báo này [10], nhóm tác giả đề xuất một kỹ thuật phát hiện xâm nhập bằng cách sử dụng mô hình học sâu có thể phân loại các kiểu tấn công khác nhau mà không cần quy tắc do con người tạo ra hoặc ánh xạ signature. Nhóm tác giả áp dụng công nghệ học sâu có giám sát là RNN, Stacked RNN và CNN để phân loại năm loại tấn công phổ biến bằng cách sử dụng Keras trên đầu TensorFlow. Kỹ thuật của Nhóm tác giả chỉ yêu cầu thông tin tiêu đề gói và không cần bất kỳ tải trọng nào của người dùng. Để xác minh hiệu suất, Nhóm tác giả sử dụng tập dữ liệu MAWI là các tệp pcap và so sánh kết quả của Nhóm tác giả với Snort IDS. Do thiếu tải trọng của người dùng, kết quả cho thấy Snort không thể phát hiện cuộc tấn công quét mạng thông qua ICMP và UDP. Trong khi đó, Nhóm tác giả chứng minh rằng RNN, Stacked RNN và CNN có thể được sử dụng để phân loại tấn công quét cổng, quét

mạng qua ICMP, quét mạng qua UDP, quét mạng qua TCP và tấn công DoS với độ chính xác cao. RNN mang lại độ chính xác cao nhất.

IDS [11] có thể xác định hiệu quả các hành vi bất thường trong mạng; tuy nhiên, nó vẫn có tỷ lệ phát hiện thấp và tỷ lệ báo động sai cao, đặc biệt là đối với các trường hợp dị thường với ít bản ghi hơn. Trong bài báo này, Nhóm tác giả đề xuất một IDS hiệu quả bằng cách sử dụng tối ưu hóa dữ liệu kết hợp bao gồm hai phần: lấy mẫu dữ liệu và lựa chọn tính năng, được gọi là DO IDS. Trong lấy mẫu dữ liệu, Rừng cách ly (iForest) được sử dụng để loại bỏ các giá trị ngoại lai, thuật toán di truyền (GA) để tối ưu hóa tỷ lệ lấy mẫu và bộ phân loại Rừng ngẫu nhiên (RF) làm tiêu chí đánh giá để có được tập dữ liệu đào tạo tối ưu. Trong lựa chọn tính năng, GA và RF được sử dụng lại để thu được tập hợp con tính năng tối ưu. Cuối cùng, một IDS dựa trên RF được xây dựng bằng cách sử dụng bộ dữ liệu đào tạo tối ưu thu được bằng cách lấy mẫu dữ liệu và các tính năng được lựa chọn bằng cách chọn tính năng. Thử nghiệm sẽ được thực hiện trên bộ dữ liệu UNSW-NB15. So với các thuật toán khác, mô hình có lợi thế rõ ràng trong việc phát hiện các hành vi bất thường hiếm gặp.

Các mối đe dọa an ninh mạng đang là mối quan tâm ngày càng tăng trong các môi trường có mạng. Sự phát triển của IDS là cơ bản để cung cấp thêm mức độ bảo mật. Nhóm tác giả đã phát triển một IDS dựa trên sự bất thường không được giám sát sử dụng các kỹ thuật thống kê để tiến hành quá trình phát hiện. Mặc dù cung cấp nhiều lợi thế, IDS dựa trên sự bất thường có xu hướng tạo ra một số lượng lớn các cảnh báo sai. Các kỹ thuật Máy học (ML) đã được quan tâm rộng rãi trong các nhiệm vụ phát hiện xâm nhập. Trong bài báo này [12], Máy vectơ hỗ trợ (SVM) được coi là một kỹ thuật ML có thể bổ sung cho hiệu suất của IDS của Nhóm tác giả, cung cấp dòng phát hiện thứ hai để giảm số lượng cảnh báo sai hoặc như một kỹ thuật phát hiện thay thế. Nhóm tác giả đánh giá hiệu suất của IDS của mình so với SVM một lớp và hai lớp, sử dụng các dạng tuyến tính và phi tuyến tính. Kết quả mà Nhóm tác giả trình bày cho thấy SVM hai lớp tuyến tính tạo ra kết quả chính xác cao và độ chính xác của SVM một lớp tuyến tính là rất tương đương và nó không cần tập dữ liệu đào tạo liên quan đến dữ liệu độc hại. Tương tự, kết quả chứng minh rằng IDS của Nhóm tác giả có thể được hưởng lợi từ việc sử dụng các kỹ thuật ML để tăng độ

chính xác của nó khi phân tích tập dữ liệu bao gồm các đối tượng địa lý không đồng nhất.

Việc giới thiệu hệ thống IoT vào các ứng dụng chăm sóc sức khỏe đã giúp có thể theo dõi từ xa thông tin của bệnh nhân và đưa ra chẩn đoán thích hợp bất cứ khi nào cần. Tuy nhiên, việc cung cấp các tính năng bảo mật cao đảm bảo tính chính xác và tính bảo mật của dữ liệu bệnh nhân là một thách thức đáng kể. Bất kỳ thay đổi nào đối với dữ liệu có thể ảnh hưởng đến việc điều trị của bệnh nhân, dẫn đến thương vong về người trong điều kiện khẩn cấp. Do tính năng động cao và tính năng động nổi bật của dữ liệu liên quan đến các hệ thống như vậy, học máy hứa hẹn cung cấp một giải pháp hiệu quả khi phát hiện xâm nhập. Tuy nhiên, hầu hết các IDS chăm sóc sức khỏe hiện có đều sử dụng số liệu lưu lượng mạng hoặc dữ liệu sinh trắc học của bệnh nhân để xây dựng bộ dữ liệu của họ. Bài báo này [13] nhằm mục đích chỉ ra rằng việc kết hợp cả số liệu mạng và sinh trắc học khi các tính năng hoạt động tốt hơn so với việc chỉ sử dụng một trong hai loại tính năng. Nhóm tác giả đã xây dựng thử nghiệm Hệ thống Giám sát Chăm sóc Sức khỏe Nâng cao (EHMS) theo thời gian thực để theo dõi sinh trắc học của bệnh nhân và thu thập các chỉ số lưu lượng mạng. Dữ liệu được giám sát sẽ được gửi đến một máy chủ từ xa để có thêm các quyết định chẩn đoán và điều trị. Các cuộc tấn công mạng từ người trung gian đã được sử dụng và một tập dữ liệu gồm hơn 16 nghìn bản ghi dữ liệu chăm sóc sức khỏe bình thường và tấn công đã được tạo ra. Sau đó, hệ thống áp dụng các phương pháp học máy khác nhau để đào tạo và kiểm tra tập dữ liệu chống lại các cuộc tấn công này. Kết quả chứng minh rằng hiệu suất đã được cải thiện từ 7% đến 25% trong một số trường hợp và điều này cho thấy sự mạnh mẽ của hệ thống được đề xuất trong việc cung cấp khả năng phát hiện xâm nhập thích hợp.

An ninh mạng bảo vệ và phục hồi hệ thống máy tính và mạng khỏi các cuộc tấn công mạng. Tầm quan trọng của an ninh mạng đang tăng lên tương ứng với sự phụ thuộc ngày càng nhiều của mọi người vào công nghệ. NIDS dựa trên phát hiện bất thường là điều cần thiết đối với bất kỳ khuôn khổ bảo mật nào trong mạng máy tính. Trong bài báo này [14], Nhóm tác giả đề xuất hai mô hình dựa trên học sâu để giải quyết sự phân loại nhị phân và đa lớp của các cuộc tấn công mạng. Nhóm tác giả sử dụng kiến trúc mạng nơ-ron phức hợp cho các mô hình của mình. Ngoài ra, một cách

tiếp cận tiền xử lý hai bước kết hợp được đề xuất để tạo ra các tính năng có ý nghĩa. Cách tiếp cận được đề xuất kết hợp giảm kích thước và kỹ thuật tính năng bằng cách sử dụng tổng hợp tính năng sâu. Hiệu suất của các mô hình của Nhóm tác giả được đánh giá bằng cách sử dụng hai tập dữ liệu điểm chuẩn, đó là khám phá kiến thức trong phòng thí nghiệm an ninh mạng trong tập dữ liệu cơ sở dữ liệu và tập dữ liệu dựa trên mạng năm 2015 của Đại học New South Wales. Hiệu suất được so sánh với các phương pháp học sâu tương tự trong tài liệu, cũng như các mô hình phân loại hiện đại. Kết quả thử nghiệm cho thấy các mô hình của Nhóm tác giả đạt được hiệu quả tốt về độ chính xác và khả năng thu hồi, vượt trội so với các mô hình tương tự trong tài liệu.

Với những tiến bộ mới nhất trong công nghệ thông tin và truyền thông, lượng lớn thông tin nhạy cảm của người dùng và công ty được chia sẻ liên tục trên mạng, khiến nó dễ bị tấn công có thể xâm phạm tính bảo mật, tính toàn vẹn và tính khả dụng của dữ liệu. IDS [15] là cơ chế bảo mật quan trọng có thể thực hiện phát hiện kịp thời các sự kiện độc hại thông qua việc kiểm tra lưu lượng mạng hoặc nhật ký dựa trên máy chủ. Nhiều kỹ thuật học máy đã được chứng minh là thành công trong việc phát hiện sự bất thường trong suốt nhiều năm, nhưng chỉ một số ít được coi là bản chất tuần tự của dữ liệu. Công việc này đề xuất phương pháp tiếp cận tuần tự và đánh giá hiệu suất của Rừng ngẫu nhiên (RF), Perceptron nhiều lớp (MLP) và Bộ nhớ dài hạn (LSTM) trên tập dữ liệu CIDDS-001. Các thước đo hiệu suất kết quả của phương pháp tiếp cận cụ thể này được so sánh với các thước đo thu được từ phương pháp truyền thống hơn, chỉ xem xét thông tin luồng riêng lẻ, để xác định phương pháp luận nào phù hợp nhất với tình huống liên quan. Các kết quả thử nghiệm cho thấy rằng việc phát hiện bất thường có thể được giải quyết tốt hơn từ quan điểm tuần tự. LSTM là một mô hình có độ tin cậy cao để thu thập các mẫu tuần tự trong dữ liệu lưu lượng mạng, đạt độ chính xác 99,94% và điểm f1 là 91,66%.

Trong vài năm gần đây, các công ty lớn ngày càng phụ thuộc nhiều hơn vào Mạng được xác định bằng phần mềm “SDN” để đáp ứng nhu cầu của họ đối với mạng có thể lập trình. Nhưng giống như các mạng khác, SDN có một số vấn đề về bảo mật. Nhiều công nghệ được sử dụng để giải quyết những vấn đề như vậy và học máy được coi là một trong những công nghệ tốt nhất. Máy học đã chứng tỏ khả năng tìm ra các

mẫu dữ liệu khi các công nghệ khác không thành công. Điều này làm cho nó trở thành một lựa chọn hoàn hảo cho IDS và phát hiện dựa trên bất thường nói chung. Trong nghiên cứu này [16], Nhóm tác giả đề xuất một IDS dựa trên bất thường mới được hưởng lợi từ khả năng của SDN cung cấp các tính năng thống kê về mọi luồng đi qua mạng và chuyển các tính năng này đến hệ thống bỏ phiếu bao gồm một số thuật toán học máy, mang lại hệ thống có khả năng nghiên cứu hành vi của người dùng và dự đoán bất kỳ sự xâm nhập nào có thể xảy ra. Hệ thống bỏ phiếu được đào tạo và thử nghiệm bằng cách sử dụng bộ dữ liệu NSL-KDD và KDDCup99 và kết quả cho thấy độ chính xác ngày càng tăng và giảm tỷ lệ dương tính giả.

Trong những thập kỷ qua, Internet và công nghệ thông tin đã nâng cao các vấn đề bảo mật do việc sử dụng mạng rất lớn. Bởi vì thông tin tiên tiến này và giao tiếp và chia sẻ thông tin, các mối đe dọa của an ninh mạng đang gia tăng hàng ngày. IDS [17] được coi là một trong những thành phần bảo mật quan trọng nhất giúp phát hiện các vi phạm an ninh mạng trong các tổ chức. Tuy nhiên, có rất nhiều thách thức đặt ra trong khi thực hiện các động lực và NIDS hiệu quả cho các cuộc tấn công chưa biết và không thể đoán trước. Xem xét cách tiếp cận máy học để phát triển một IDS hiệu quả và linh hoạt. Một mô hình mạng nơ-ron sâu được đề xuất để tăng hiệu quả của IDS.

Mạng là những thiết bị dễ bị tấn công do tính năng cơ bản của chúng là hỗ trợ truy cập từ xa và giao tiếp dữ liệu. Thông tin trong mạng cần được bảo mật và an toàn để cung cấp phương tiện giao tiếp và chia sẻ hiệu quả trong mạng dữ liệu. Do những thách thức và mối đe dọa của dữ liệu trong mạng, an ninh mạng là một trong những vấn đề quan trọng nhất trong cơ sở hạ tầng công nghệ thông tin. Do đó, các biện pháp bảo mật được xem xét trong mạng để giảm xác suất truy cập vào dữ liệu bảo mật của tin tặc. Mục đích của an ninh mạng là bảo vệ mạng và các thành phần của mạng khỏi bị truy cập và lạm dụng trái phép nhằm cung cấp một thiết bị liên lạc an toàn và bảo mật cho người dùng. Trong công trình nghiên cứu hiện tại [18], việc xem xét lại sự phát triển gần đây của các mối đe dọa mạng và các biện pháp bảo mật đã được trình bày và các công trình nghiên cứu trong tương lai cũng được đề xuất. Các cuộc tấn công khác nhau vào mạng và bảo mật được đo lường chống lại chúng được thảo luận để tăng cường bảo mật trong web dữ liệu. Vì vậy, các ý tưởng mới trong hệ thống an

ninh mạng có thể được trình bày bằng cách phân tích các bài báo đã xuất bản để tiến tới lĩnh vực nghiên cứu.

Bài báo này [4] đề cập đến các phương pháp và công nghệ hiện đại để phát hiện các hành vi xâm nhập mạng. Ưu điểm và nhược điểm của các thiết bị phát hiện xâm nhập hiện đại được xem xét. Vấn đề dương tính giả của IDS và hậu quả của các hành động đó được mô tả. Đã có phân tích các cuộc tấn công nam nhân tố vào IDS Snort.

2.3 Kết luận chương

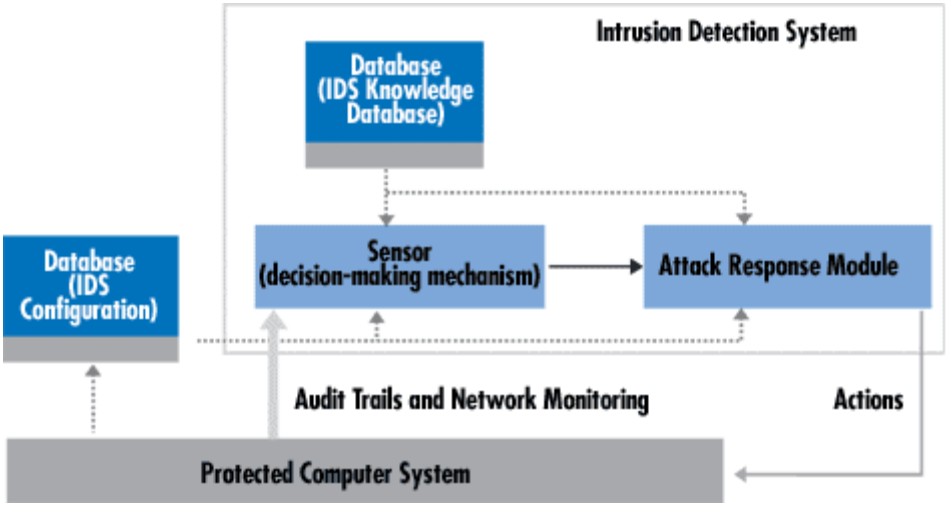
Trong chương này thông qua việc nghiên cứu tìm hiểu được một số thuật toán và những công trình liên quan tới mô hình học máy, giải quyết các bài toán liên quan về dữ liệu mạng, từ đó hiểu được những ưu nhược điểm của các thuật toán, tạo tiền đề và cơ sở vững chắc cho nghiên cứu của đề tài luận văn này.

CHƯƠNG 3. HỆ THỐNG PHÁT HIỆN VÀ PHÒNG CHỐNG XÂM NHẬP MẠNG

3.1 Tổng quan về IDS

IDS hay còn gọi là hệ thống phát hiện xâm nhập là hệ thống phần cứng hoặc phần mềm có chức năng giám sát lưu thông mạng, tự động theo dõi các sự kiện xảy ra trên hệ thống máy tính, phân tích để phát hiện ra các vấn đề liên quan đến an ninh, bảo mật và đưa ra cảnh báo cho nhà quản trị.

Kiến trúc IDS [19]

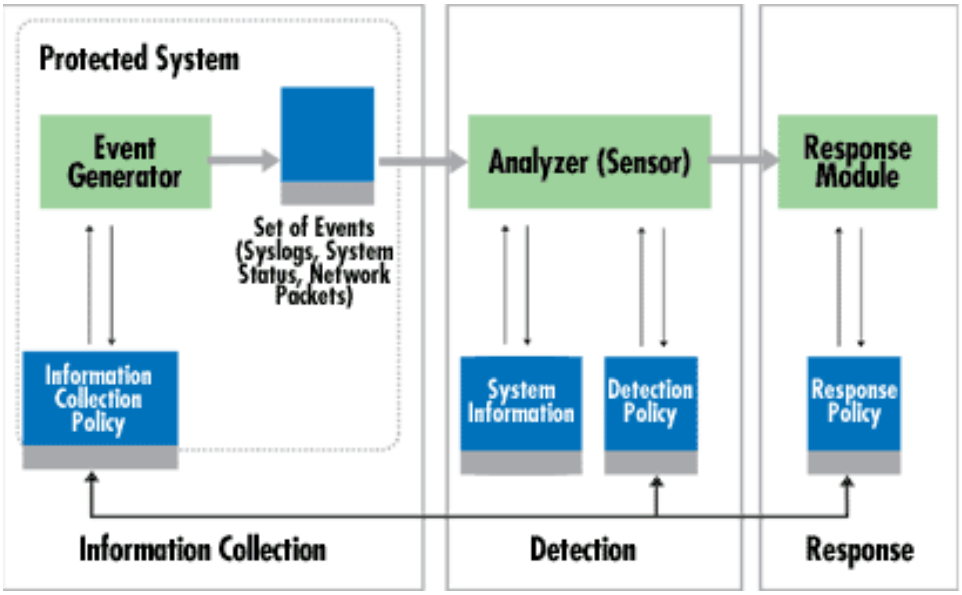


Hình 3.1: Một IDS mẫu. Độ rộng mũi tên tỷ lệ với lượng thông tin giữa các thành phần trong hệ thống [19]

Cảm biến được tích hợp với thành phần chịu trách nhiệm thu thập dữ liệu (Hình 3.1) - một bộ tạo sự kiện. Cách thức thu thập được xác định bởi chính sách tạo sự kiện để xác định chế độ lọc thông tin sự kiện. Bộ tạo sự kiện (hệ điều hành, ứng dụng, mạng) tạo ra một tập hợp các sự kiện nhất quán về chính sách có thể là một bản ghi các sự kiện hệ thống hoặc các gói mạng. Số chính xác này được thiết lập cùng với thông tin chính sách có thể được lưu trữ trong hệ thống được bảo vệ hoặc bên ngoài. Trong một số trường hợp nhất định, ví dụ, khi các luồng dữ liệu sự kiện được chuyển trực tiếp đến máy phân tích mà không có sự lưu dữ liệu nào được thực hiện. Điều này liên quan đến các gói mạng nói riêng.

Thành phần của IDS [19]

IDS gồm có ba thành phần chính [19] bao gồm: thành phần thu thập gói tin, phát hiện gói tin, phản hồi gói tin được mô tả trong hình 3.1 dưới đây:



Hình 3.2: Các thành phần của IDS [19]

- Thành phần thu thập gói tin (Information Collection)

Thành phần này có nhiệm vụ lấy tất cả các gói tin đi đến mạng. Thông thường, các gói tin có địa chỉ không phải của một card mạng thì sẽ bị card mạng đó huỷ bỏ nhưng card mạng của IDS được đặt ở chế độ thu nhận tất cả. Tất cả các gói tin qua chúng đều được sao chụp, xử lý, phân tích đến từng trường thông tin. Bộ phận thu thập gói tin sẽ đọc thông tin từng trường trong gói tin, xác định chúng thuộc kiểu gói tin nào, dịch vụ gì... Các thông tin này được chuyển đến thành phần phát hiện tấn công.

- Thành phần phát hiện gói tin (Detection)

Ở thành phần này, các bộ cảm biến đóng vai trò quyết định. Vai trò của bộ cảm biến là dùng để lọc thông tin và loại bỏ những thông tin dữ liệu không tương thích đạt được từ các sự kiện liên quan tới hệ thống bảo vệ, vì vậy có thể phát hiện được các hành động nghi ngờ.

- Thành phần phản hồi (Response)

Khi có dấu hiệu của sự tấn công hoặc thâm nhập, thành phần phát hiện tấn công sẽ gửi tín hiệu báo hiệu (alert) có sự tấn công hoặc thâm nhập đến thành phần phản ứng. Lúc đó thành phần phản ứng sẽ kích hoạt tường lửa thực hiện chức nang

ngan chặn cuộc tấn công hay cảnh báo tới người quản trị. Dưới đây là một số kỹ thuật ngăn chặn:

- Cảnh báo thời gian thực: gửi các cảnh báo thời gian thực đến người quản trị để họ nắm được chi tiết các cuộc tấn công, các đặc điểm và thông tin về chúng.
- Ghi lại vào tập tin: các dữ liệu của các gói tin sẽ được lưu trữ trong hệ thống các tập tin log. Mục đích là để những người quản trị có thể theo dõi các luồng thông tin và là nguồn thông tin giúp cho module phát hiện tấn công hoạt động.
- Ngăn chặn, thay đổi gói tin: khi một gói tin khớp với dấu hiệu tấn công thì IDS sẽ phản hồi bằng cách xóa bỏ, từ chối hay thay đổi nội dung của gói tin, làm cho gói tin trở nên không bình thường.

Cách thức hoạt động của IDS [1]

Sau khi thu thập dữ liệu, IDS sẽ quan sát lưu lượng mạng và kết hợp các mẫu lưu lượng truy cập đến các cuộc tấn công đã biết. Phương pháp này thường được gọi là tương quan mẫu. Khi hoạt động đáng ngờ hoặc độc hại được phát hiện, IDS sẽ gửi tín hiệu đến các kỹ thuật viên hoặc quản trị viên CNTT được chỉ định. Các báo động của IDS cho phép nhanh chóng bắt đầu khắc phục sự cố và xác định các nguồn gốc gốc hoặc truy vết và ngăn chặn các tác nhân có hại.

Các IDS chủ yếu sử dụng hai phương pháp phát hiện xâm nhập chính: phát hiện xâm nhập dựa trên signature và phát hiện xâm nhập dựa trên các đặc điểm bất thường. Phát hiện xâm nhập dựa trên signature được thiết kế để phát hiện các mối đe dọa có thể xảy ra bằng cách so sánh lưu lượng truy cập mạng và dữ liệu nhật ký cho các mẫu tấn công hiện có. Phát hiện dựa trên signature cho phép phát hiện chính xác và xác định các cuộc tấn công đã biết có thể.

Ngược lại với phương pháp phát hiện xâm nhập dựa trên chữ kí là phát hiện xâm nhập dựa trên đặc điểm bất thường - nó được thiết kế để xác định các cuộc tấn công không xác định, chẳng hạn như phần mềm độc hại mới và thích ứng với chúng khi đang bay bằng cách sử dụng máy học. Các kỹ thuật học máy cho phép IDS tạo ra các hoạt động đáng tin cậy được gọi là mô hình tin cậy, sau đó so sánh hành vi mới với các mô hình tin cậy đã được xác minh. Báo động sai có thể xảy ra khi sử dụng phương pháp dựa trên sự bất thường, vì lưu lượng mạng không xác định trước đây nhưng hợp lệ, có thể được xác định nhầm là hoạt động độc hại.

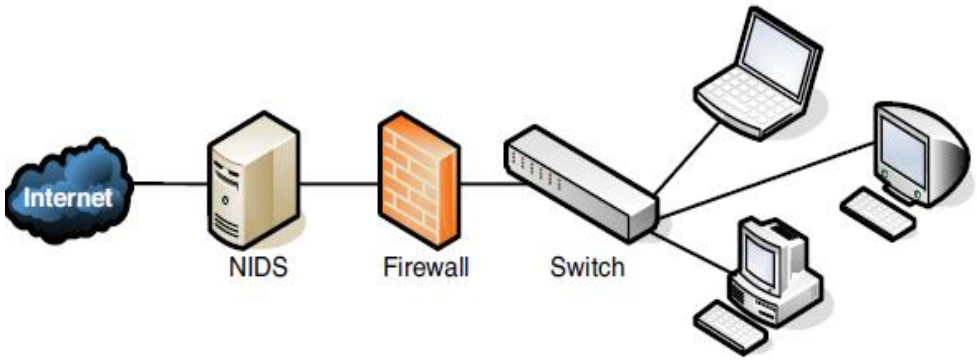
Các IDS dựa trên đặc trưng và điểm bất thường giúp cho hệ thống tăng phạm vi phòng chống xâm nhập. Điều này thể hiện việc xác định càng nhiều mối đe dọa sẽ càng làm tăng tính hiệu quả cho phương pháp này. IDS toàn diện có thể phát hiện ra các kỹ thuật tránh né các mạng loại mạng ẩn, sử dụng để đánh lừa hệ thống phòng chống xâm nhập vào suy nghĩ không có một cuộc tấn công diễn ra. Những kỹ thuật này có thể bao gồm phân mảnh, các cuộc tấn công băng thông thấp, giả mạo địa chỉ hoặc ủy quyền,…

3.2 Vai trò và chức năng của hệ thống phát hiện và phòng chống xâm nhập

3.2.1 Vai trò và chức năng của IDS

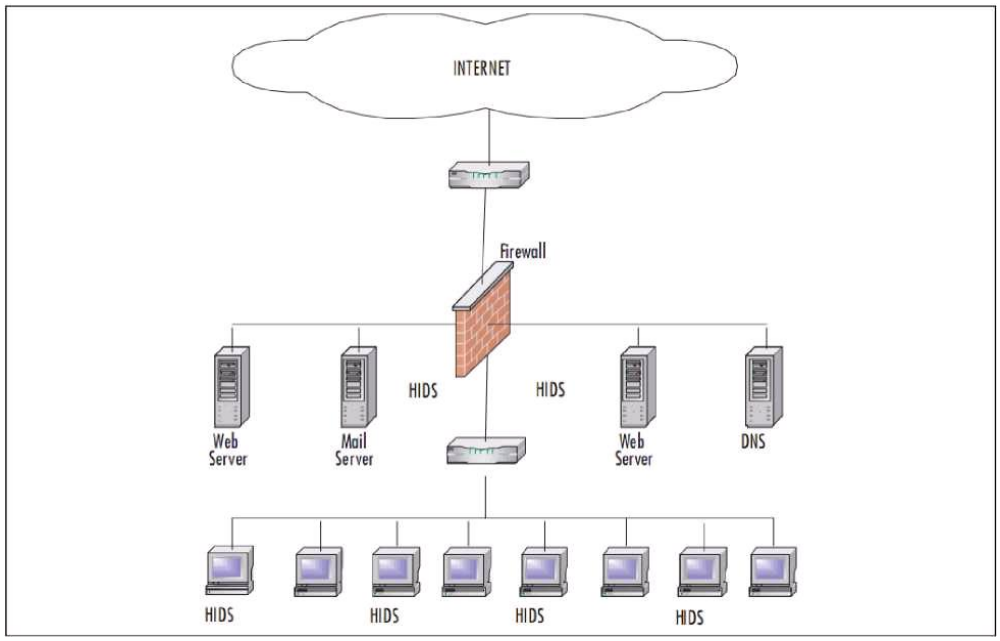
Các IDS được thiết kế để xác định hoạt động đáng ngờ và độc hại thông qua lưu lượng mạng và IDS cho phép khai thác và kiểm tra xem hệ thống mạng có đang bị tấn công hay không. Có hai loại IDS chính [1]:

- NIDS sử dụng bộ dò và bộ cảm biến được cài đặt trên toàn mạng để theo dõi mạng nhằm mục đích tìm kiếm những lưu lượng trùng với những mô tả được định nghĩa hay là những dấu hiệu. Một số lợi thế của NIDS có thể kể đến như: quản lý được cả một network segment (gồm nhiều host); cài đặt và bảo trì đơn giản; tránh DoS ảnh hưởng đến một host nào đó; có khả năng xác định lỗi ở tầng Network (ở mô hình OSI) và độc lập với hệ điều hành,... Ngoài ra, cũng tồn tại một số hạn chế: có thể xảy ra báo động giả; không thể phân tích các gói tin được mã hóa; đòi hỏi việc cập nhật các signature mới nhất để đảm bảo an toàn; không có thông báo về việc tấn công có thành công hay không,…



Hình 3.3: Mô hình mạng NIDS

- HIDS được sử dụng để phân tích và phát hiện xâm nhập. Đồng thời, hệ thống báo cáo chính xác các hoạt động mà kẻ tấn công đã thực hiện, bao gồm các câu lệnh, hoặc các tệp tin đã được mở. Lợi thế khi sử dụng HIDS như: xác định được người dùng liên quan tới sự kiện; phát hiện các cuộc tấn công diễn ra trên một máy; phân tích được các dữ liệu mã hoá; cung cấp thông tin về host khi diễn ra tấn công. Và một số hạn chế như: thông tin không đáng tin cậy; phải được thiết lập trên từng host cần giám sát; không có khả năng phát hiện các cuộc dò quét mạng; cần tài nguyên trên host để hoạt động; không hiệu quả khi bị DoS,…

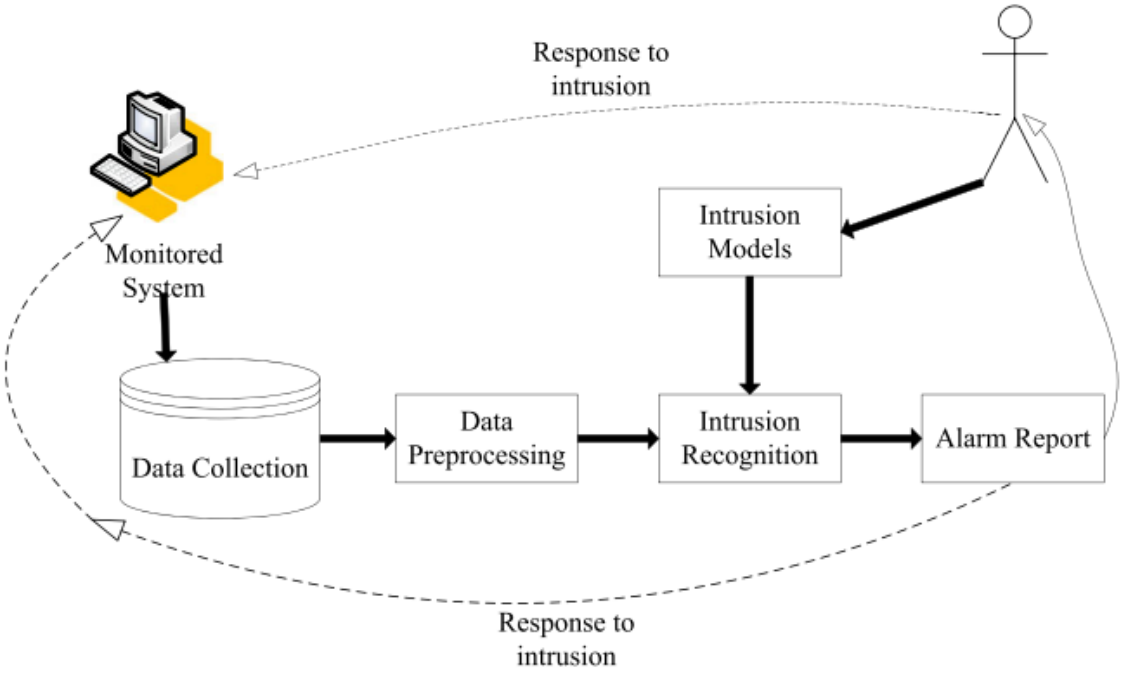


Hình 3.4: Mô hình mạng HIDS

Các NIDS thu thập lưu lượng mạng từ tất cả các thiết bị thông qua NIDS và HIDS, do đó việc phát hiện xâm nhập trên cơ sở hạ tầng CNTT được đảm bảo và nâng cao.

3.2.2 Chức năng IDS

Hệ thống IDS bao gồm bốn chức năng chính [20] là thu thập dữ liệu, chọn lọc đặc trưng, phân tích và thực thi (hình 3.5).



Hình 3.5: Chức năng của IDS [20]

Ở quá trình thu thập dữ liệu, các loại dữ liệu thu thập được ghi nhận lại thành từng tập tin, sau đó được sử dụng để phân tích. Các đặc trưng cụ thể từ tập dữ liệu lớn thường có sẵn trong mạng và chúng được đánh giá cho việc phát hiện xâm nhập. Ví dụ như địa chỉ IP của một nguồn và hệ thống, các loại giao thức, độ dài và kích thước của header cũng có thể được sử dụng làm yếu tố phục vụ cho sự phát hiện xâm nhập. Ở giai đoạn phân tích, dữ liệu sẽ được khai tác để tìm ra các dữ liệu phù hợp. Hệ thống IDS sẽ thiết lập các qui tắc để phân tích dữ liệu, tại đây lưu lượng mạng được phân tích lại một lần nữa bằng các mẫu và signature đã được xác định trước. Ngoài ra, hệ thống IDS cũng có thể dựa trên sự bất thường từ tín hiệu của hệ thống để từ đó hình thành các mô hình toán học sử dụng cho mục tiêu phát hiện xâm nhập.

Ở phần thực thi, hệ thống IDS sẽ xác định về các hoạt động tấn công và phản ứng của hệ thống. Từ đó báo tín hiệu đến nhân viên quản trị hệ thống tất cả các dữ liệu được yêu cầu thông qua email, hoặc IDS có thể chủ động loại bỏ các gói tin đáng nghi để nó không gây hại cho hệ thống.

3.3 Công cụ giám sát mạng Snort

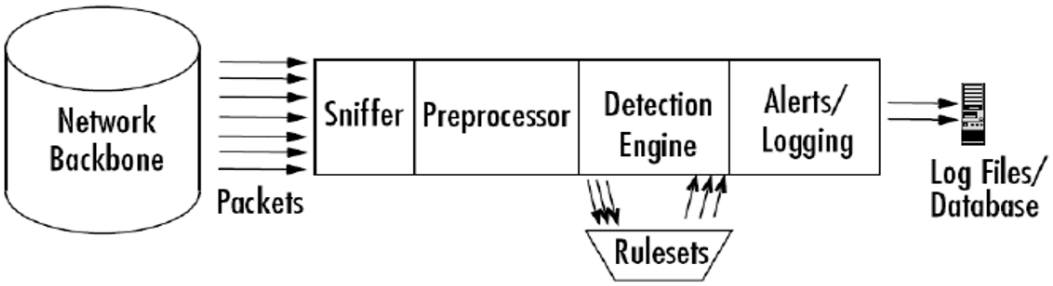
3.3.1 Giới thiệu Snort

Snort [21] là một NIDS được Martin Roesh phát triển dưới mô hình mã nguồn mở. Snort không những miễn phí mà nó còn có rất nhiều tính năng tuyệt vời mà không

phải sản phẩm thương mại nào cũng có thể có được. Snort được xây dựng để phát hiện và chống xâm nhập. Được thiết kế trên module để kiểm tra các gói dữ liệu vào/ra bằng cách tạo nên các rule để phát hiện các gói dữ liệu bất thường. Snort có thể chạy trên nhiều hệ thống nền như Windows, MacOS, Linux, OpenBSD, NetBSD, Solaris, HP-UX, ... Snort còn có thể được cấu hình để chạy như một NIDS. Snort hỗ trợ khả năng hoạt động trên các giao thức sau: Ethernet, 802.11,Token Ring, FDDI, Cisco HDLC, SLIP, PPP, và PF của OpenBSD.

Kiến trúc Snort bao gồm nhiều thành phần, mỗi phần có một chức năng riêng biệt. Các thành phần đó bao gồm:

- Mô-đun giải mã gói tin (Packet Decoder)
- Mô-đun tiền xử lý (Preprocessors)
- Mô-đun phát hiện (Detection Engine)
- Mô-đun log và cảnh báo (Logging and Alerting System)
- Mô-đun kết xuất thông tin (Output Module)



Hình 3.6: Kiến trúc của Snort [21]

Khi hoạt động Snort sẽ thực hiện việc lắng nghe và thu bắt tất cả các gói tin di chuyển qua nó. Các gói tin sau khi thu sẽ được đưa vào Mô-đun giải mã gói tin. Tiếp theo gói tin sẽ được đưa vào mô-đun tiền xử lý, cuối cùng là mô-đun phát hiện. Tại đây tùy vào việc có phát hiện được xâm nhập hay không mà gói tin có thể được bỏ qua để lưu thông tiếp hoặc sẽ được đưa vào mô-đun log và cảnh báo để tiếp tục xử lý. Ngay khi các cảnh báo được xác định mô-đun kết xuất thông tin sẽ thực hiện việc đưa cảnh báo ra theo đúng định dạng mong muốn.

3.3.2 Bộ luật Snort

Tương tự như virus, hầu hết các hoạt động tấn công hay xâm nhập đều có các dấu hiệu riêng. Các thông tin về các dấu hiệu này sẽ được sử dụng để tạo nên các luật

cho Snort. Thông thường, các bẫy (honey pots) được tạo ra để tìm hiểu xem các kẻ tấn công làm gì cũng như các thông tin về công cụ và công nghệ chúng sử dụng. Và ngược lại, cũng có các cơ sở dữ liệu về các lỗ hổng bảo mật mà những kẻ tấn công muốn khai thác. Các dạng tấn công đã biết này được dùng như các dấu hiệu để phát hiện tấn công xâm nhập. Các dấu hiệu đó có thể xuất hiện trong phần header của các gói tin hoặc nằm trong phần nội dung của chúng. Hệ thống phát hiện của Snort hoạt động dựa trên các luật (rules) và các luật này lại được dựa trên các dấu hiệu nhận dạng tấn công. Các luật có thể được áp dụng cho tất cả các phần khác nhau của một gói tin dữ liệu . Một luật có thể được sử dụng để tạo nên một thông điệp cảnh báo, log một thông điệp hay có thể bỏ qua một gói tin.

a. Cấu trúc luật của Snort

Xem xét một ví dụ đơn giản : alert tcp 192.168.2.0/24 23 -> any any (content:”confidential”; msg: “Detected confidential”) Ta thấy cấu trúc của một luật có dạng như sau:

Rule Header	Rule Option

- Tất cả các luật của Snort về logic đều gồm 2 phần: Phần header và phần Option.
- Phần Header chứa thông tin về hành động mà luật đó sẽ thực hiện khi phát hiện ra có xâm nhập nằm trong gói tin và nó cũng chứa các tiêu chuẩn để áp dụng luật với gói tin đó.
- Phần Option chứa một thông điệp cảnh báo và các thông tin về các phần của gói tin dùng để tạo nên cảnh báo. Phần Option chứa các tiêu chuẩn phụ thêm để đối sánh luật với gói tin. Một luật có thể phát hiện được một hay nhiều hoạt động thăm dò hay tấn công. Các luật thông minh có khả năng áp dụng cho nhiều dấu hiệu xâm nhập. Dưới đây là cấu trúc chung của phần Header của một luật Snort:

Action	Protocol	Address	Port	Direction	Address	Port

- Action: là phần quy định loại hành động nào được thực thi khi các dấu hiệu của gói tin được nhận dạng chính xác bằng luật đó. Thông thường các hành động tạo ra một cảnh báo hoặc log thông điệp hoặc kích hoạt một luật khác.

- Protocol: là phần qui định việc áp dụng luật cho các packet chỉ thuộc một giao thức cụ thể nào đó. Ví dụ như IP, TCP, UDP …
- Address: là phần địa chỉ nguồn và địa chỉ đích. Các địa chỉ có thể là một máy đơn, nhiều máy hoặc của một mạng nào đó. Trong hai phần địa chỉ trên thì một sẽ là địa chỉ nguồn, một sẽ là địa chỉ đích và địa chỉ nào thuộc loại nào sẽ do phần Direction "->" qui định. • Port: xác định các cổng nguồn và đích của một gói tin mà trên đó luật được áp dụng.
- Direction: phần này sẽ chỉ ra đâu là địa chỉ nguồn, đâu là địa chỉ đích. Ví dụ: alert icmp any any -> any any (msg: "Ping with TTL=100";ttl: 100;) Phần đứng trước dấu mở ngoặc là phần Header của luật còn phần còn lại là phần Option. Chi tiết của phần Header như sau:
- Hành động của luật ở đây là "alert" : một cảnh báo sẽ được tạo ra nếu như các điều kiện của gói tin là phù hợp với luật(gói tin luôn được log lại mỗi khi cảnh báo được tạo ra).
- Protocol của luật ở đây là ICMP tức là luật chỉ áp dụng cho các gói tin thuộc loại ICMP. Bởi vậy, nếu như một gói tin không thuộc loại ICMP thì phần còn lại của luật sẽ không cần đối chiếu.
- Địa chỉ nguồn ở đây là "any": tức là luật sẽ áp dụng cho tất cả các gói tin đến từ mọi nguồn còn cổng thì cũng là "any" vì đối với loại gói tin ICMP thì cổng không có ý nghĩa. Số hiệu cổng chỉ có ý nghĩa với các gói tin thuộc loại TCP hoặc UDP thôi.

Còn phần Option trong dấu đóng ngoặc chỉ ra một cảnh báo chứa dòng "Ping with TTL=100" sẽ được tạo khi tìm thấy điều kiện TTL=100. TTL là Time To Live là một trường trong Header IP.

3.4 Các mô hình sử dụng cho hệ thống IDS

3.4.1 Mô hình Decision Tree

Cây quyết định [22] là một phương pháp học có giám sát được sử dụng để phân loại và hồi quy. Mục tiêu là tạo một mô hình dự đoán giá trị của biến mục tiêu bằng cách học các quy tắc quyết định đơn giản được suy ra từ các tính năng dữ liệu. Một cây có thể được xem là một xấp xỉ không đổi.

Một số lợi thế của cây quyết định là:

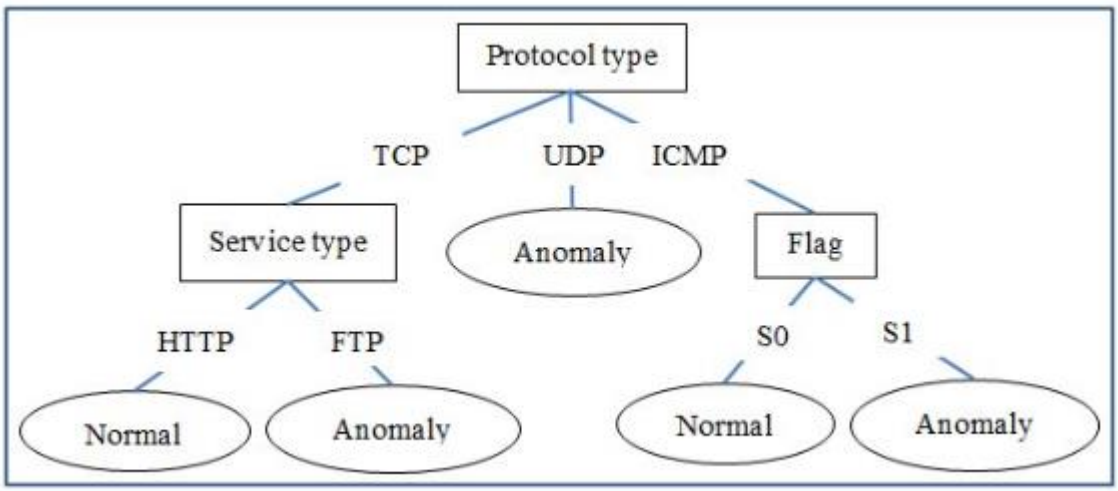
- Đơn giản để hiểu và dễ dàng thực hiện.
- Yêu cầu chuẩn bị ít dữ liệu. Các kỹ thuật khác thường yêu cầu chuẩn hóa dữ liệu, các biến giả cần được tạo và loại bỏ các giá trị trống.
- Chi phí sử dụng cây (tức là dữ liệu dự đoán) là logarit trong số điểm dữ liệu được sử dụng để đào tạo cây.
- Có thể xử lý cả dữ liệu số và dữ liệu phân loại. Tuy nhiên, hiện tại việc triển khai scikit-learning không hỗ trợ các biến phân loại. Các kỹ thuật khác thường chuyên về phân tích tập dữ liệu chỉ có một loại biến.
- Có khả năng xử lý các vấn đề đa đầu ra.
- Sử dụng mô hình hộp trắng. Nếu một tình huống nhất định có thể quan sát được trong một mô hình, thì lời giải thích cho điều kiện đó dễ dàng được giải thích bằng logic boolean. Ngược lại, trong mô hình hộp đen (ví dụ: trong mạng nơ-ron nhân tạo), kết quả có thể khó giải thích hơn.
- Có thể xác nhận một mô hình bằng cách sử dụng các bài kiểm tra thống kê. Điều đó làm cho nó có thể tính đến độ tin cậy của mô hình.
- Hoạt động tốt ngay cả khi các giả định của nó phần nào bị vi phạm bởi mô hình thực mà từ đó dữ liệu được tạo ra.

Những điểm yếu của cây quyết định bao gồm:

- Cây quyết định có thể tạo ra cây quá phức tạp, không tổng quát hóa dữ liệu tốt. Các cơ chế như thiết lập số lượng mẫu tối thiểu cần thiết tại một nút lá hoặc thiết lập độ sâu tối đa của cây là cần thiết để tránh vấn đề này.
- Cây quyết định có thể không ổn định vì các biến thể nhỏ trong dữ liệu có thể dẫn đến việc tạo ra một cây hoàn toàn khác. Vấn đề này được giảm thiểu bằng cách sử dụng cây quyết định trong một tập hợp.
- Các dự đoán của cây quyết định không trơn tru cũng không liên tục, mà là các phép gần đúng không đổi.
- Một cây quyết định tối ưu khi hoàn chỉnh NP dưới một số khía cạnh của tính tối ưu và ngay cả đối với các khái niệm đơn giản. Do đó, các thuật toán học cây quyết định thực tế dựa trên các thuật toán heuristic như thuật toán greedy, trong đó các quyết định tối ưu cục bộ được thực hiện tại mỗi nút với mục tiêu tìm được sự tối ưu trên toàn cục. Điều này có thể được giảm thiểu bằng cách

huấn luyện nhiều cây theo nhóm, trong đó các tính năng và mẫu được lấy ngẫu nhiên để thay thế.

- Có những khái niệm phức tạp, vì cây quyết định không thể hiện chúng một cách dễ dàng, chẳng hạn như vấn đề XOR, chẵn lẻ hoặc bộ ghép kênh.



Hình 3.7: Sơ đồ cây quyết định [23]

3.4.2 Mô hình KNN

K-Nearest Neighbor [24] là một thuật toán khai phá dữ liệu và đồng thời là thuật toán học có giám sát. Đầu ra của biến mục tiêu được dự đoán bằng cách tìm k lân cận gần nhất, bằng cách tính khoảng cách Euclide [24] [25]. Đây là một kỹ thuật phân loại phi tham số không đưa ra bất kỳ giả định nào về dữ liệu cơ bản. Một số ưu điểm của KNN là:

- Dễ thực hiện và dễ hiểu.
- Sẽ rất hiệu quả và hiệu quả nếu dữ liệu huấn luyện rất lớn.
- Nó mạnh mẽ đối với dữ liệu nhiễu.
- Nó liên tục phát triển và dễ dàng thích nghi với môi trường mới.
- Dễ thực hiện cho bài toán nhiều lớp

Thuật toán KNN có công thức tổng quát như sau:

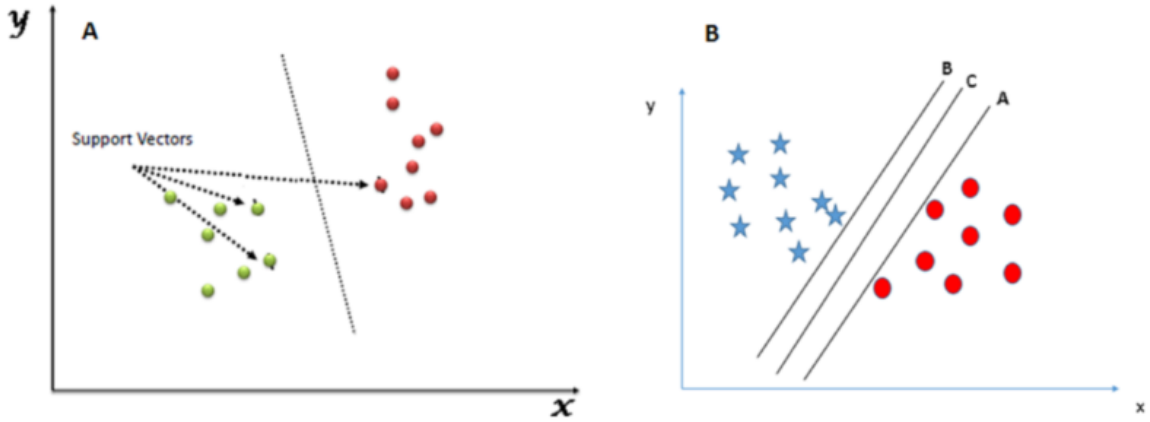
$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

Trong đó, Xi là phần tử đặc trưng thứ i của đối tượng x, Yi là phần tử đặc trưng thứ I của đối tượng y và n là tổng số đặc trưng của bộ dữ liệu.

3.4.3 Mô hình máy Vector hỗ trợ (SVM)

Máy hỗ trợ vectơ (SVM) [26] là một tập hợp các phương pháp học được giám sát được sử dụng để phân loại, hồi quy và phát hiện ngoại lệ. Ngoài ra, SVM cũng là một trong những thuật toán phân loại đáng tin cậy nhất trong máy học tập vì nó hỗ trợ quá trình dự đoán nhanh chóng và đơn giản so với các thuật toán khác. SVM phân loại các điểm dữ liệu dựa trên các vectơ hỗ trợ trong kho dữ liệu để tạo một siêu phẳng phân chia các nhãn lớp thành các lớp liên quan của chúng.

SVM phân loại các vectơ hỗ trợ dựa trên giá trị "gamma" được cung cấp cho thuật toán. Ví dụ: nếu gamma = 0, SVM dự đoán siêu phẳng là một đường cong. Nếu gamma = auto, SVM chỉ đơn giản là dự đoán siêu phẳng theo đầu vào dữ liệu đã cho.



Hình 3.8: Phân lớp với SVM. (A) Kỹ thuật phân lớp SVM. (B) Kỹ thuật lựa chọn siêu phẳng SVM

Hình 3.3 thể hiện tập hợp dữ liệu gồm các ngôi sao xanh và vòng tròn đỏ, nơi cần tìm ra siêu mặt phẳng phù hợp để phân loại chúng một cách chính xác. Ở đây, có ba siêu phẳng A, B và C. Nhưng siêu phẳng thực tế cần xem xét là siêu phẳng C vì nó có cùng khoảng cách từ cả hai điểm dữ liệu.

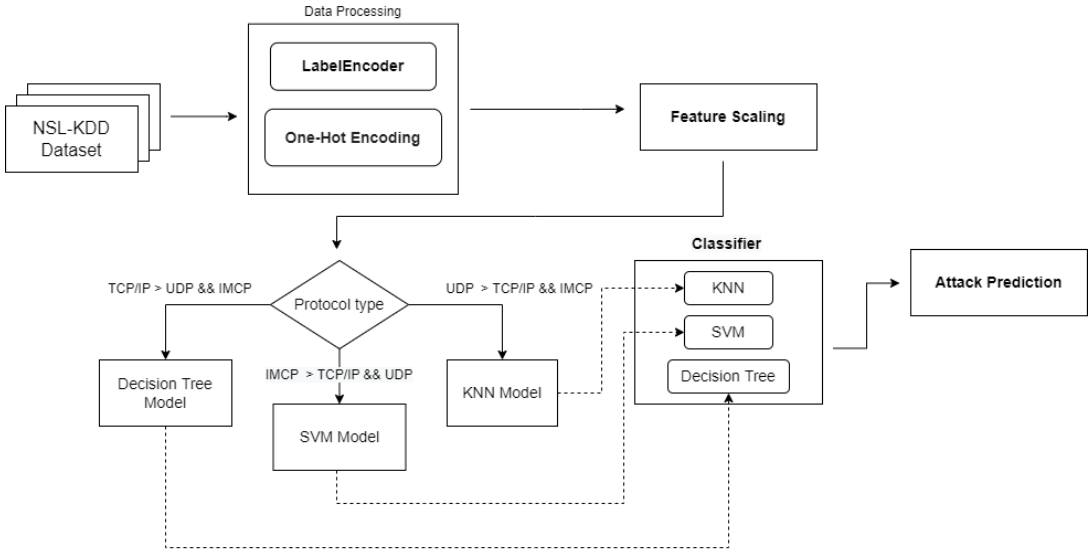
Một số ưu điểm của thuật toán SVM là:

- Hiệu quả trong không gian chiều cao.
- Vẫn có hiệu quả trong trường hợp số thứ nguyên lớn hơn số lượng mẫu.
- Sử dụng một tập hợp con các điểm huấn luyện trong hàm quyết định (được gọi là vectơ hỗ trợ), vì vậy nó cũng hiệu quả về bộ nhớ.

- Đa năng: các chức năng Kernel khác nhau có thể được chỉ định cho chức năng quyết định. Các nhân chung được cung cấp, nhưng cũng có thể chỉ định các nhân tùy chỉnh.
- Những nhược điểm của máy vector hỗ trợ bao gồm:
- Nếu số lượng tính năng lớn hơn nhiều so với số lượng mẫu, hãy tránh việc lựa chọn các chức năng của Kernel và thuật ngữ chính quy hóa là rất quan trọng.
- SVM không trực tiếp cung cấp các ước tính xác suất, những ước tính này được tính toán bằng cách sử dụng xác thực chéo năm lần đắt tiền.

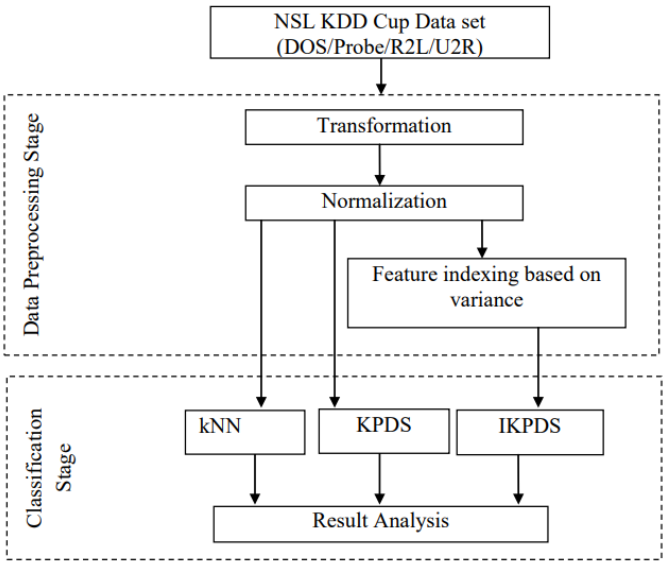
3.5 Mô hình IDS đề xuất

Dựa vào tài liệu [22, 25], luận văn này xin đề xuất mô hình phát hiện xâm nhập mạng sử dụng học máy gồm các quá trình xử lý như sau:



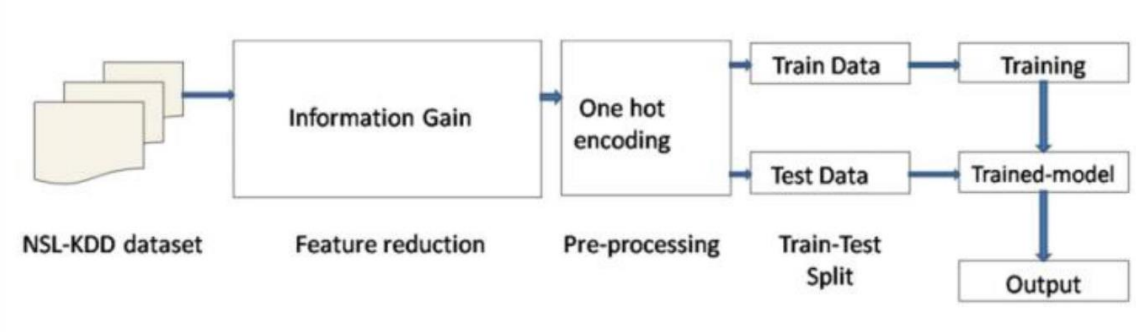
Hình 3.9: Mô hình đề xuất của luận văn

Điểm khác biệt của mô hình đề xuất với hai tài liệu trên là với mô hình đề xuất từ tài liệu [21], ngoài thuật toán KNN ra thì luận văn có sử dụng 2 thuật toán khác là Decision Tree và SVM (ở tài liệu [21] sử dụng KNN, KPDS, IKPDS). Xét về mô hình, ở giai đoạn áp dụng thuật toán có sự khác nhau về điều kiện để áp dụng thuật toán cho mô hình. Theo đó, luận văn áp dụng các thuật toán dựa trên số lượng giao thức trong tập dữ liệu, còn tài liệu [21] không có ràng buộc về điều kiện khi áp dụng thuật toán.



Hình 3.10: Mô hình đề xuất của [21]

Tương tự như vậy, đối với tài liệu [24], mô hình đề xuất chỉ thiên về việc xử lí dữ liệu, đồng thời nhóm tác giả đã khẳng định là mô hình đề xuất dựa trên thuật toán Decision Tree, và các kết quả thu được sẽ so sánh với các kết quả của thuật toán khác.



Hình 3.11: Mô hình đề xuất của [24]

Sau đây luận văn xin mô tả cụ thể các quá trình xảy ra trong mô hình đề xuất. Cụ thể quá trình bao gồm ba bước như sau:

3.5.1 Đọc, lưu dữ liệu Log từ SNORT và xử lý dữ liệu

Module này sẽ thực hiện nhiệm vụ đọc log của SNORT, log này đã được hiệu chỉnh và chỉ lưu lại những luồng dữ liệu đã được lọc trên SNORT, tức là phải có dấu hiệu cần lưu lại mới đưa vào Log. File log sẽ được lưu lại dưới dạng CSV. Việc điều chỉnh thời gian đọc log và lưu log là do ta chọn và xử lý, trong luận văn này đề xuất

là 10 phút đọc 1 lần, khoảng 10.000 dòng dữ liệu lấy từ Log. Tiếp theo đó tiến hành quá trình xử lý dữ liệu bằng LabelEncoder và One-Hot Encoding.

Một trong các bước quan trọng để có đầu ra tốt nhất là làm sạch và xử lý dữ liệu. Mỗi mô hình học máy khác nhau thường có các cách tiếp cận khác nhau cho các loại dữ liệu khác nhau, vì vậy trong luận văn này, sẽ sử dụng kĩ thuật One hot Encoding và Label Encoding trong quá trình này. Đây là hai trong số các kĩ thuật phân loại dữ liệu cho các mô hình học máy.

Label Encoding [27]

Trong học máy, quá trình xử lý các bộ dữ liệu có chứa nhiều nhãn trong một hoặc nhiều hơn một cột. Những nhãn này có thể ở dạng từ hoặc số. Để làm cho dữ liệu dễ hiểu hoặc ở dạng người có thể đọc được, dữ liệu huấn luyện thường được dán nhãn bằng lời.

Để giúp cho dữ liệu dễ dàng xử lý trong các mô hình, Label Encoding hỗ trợ chuyển đổi nhãn thành một dạng số để chuyển đổi chúng thành dạng có thể đọc được. Các thuật toán học máy sau đó có thể quyết định theo cách tốt hơn cách các nhãn đó phải được vận hành. Đây là một bước xử lý trước quan trọng cho bộ dữ liệu có cấu trúc trong việc học có giám sát.

Hình 3.12: Cách hoạt động của Label Encoding

One hot encoding [28]

Dữ liệu phân loại đề cập đến các biến được tạo thành từ các giá trị nhãn, ví dụ: biến "màu" có thể có các giá trị "đỏ", "xanh lam" và "xanh lá cây". Hãy nghĩ về các giá trị như các danh mục khác nhau mà đôi khi có thứ tự tự nhiên đối với chúng.

Một số thuật toán học máy có thể hoạt động trực tiếp với dữ liệu phân loại tùy thuộc vào việc triển khai, chẳng hạn như cây quyết định, nhưng hầu hết đều yêu cầu bất kỳ biến đầu vào hoặc đầu ra nào phải là một số hoặc số có giá trị. Điều này có nghĩa là mọi dữ liệu phân loại phải được ánh xạ tới các số nguyên.

One hot encoding là một phương pháp chuyển đổi dữ liệu để chuẩn bị cho một thuật toán và nhận được dự đoán tốt hơn. One-hot chuyển đổi từng giá trị phân loại thành một cột phân loại mới và gán giá trị nhị phân 1 hoặc 0 cho các cột đó. Mỗi giá trị số nguyên được biểu diễn dưới dạng véc tơ nhị phân. Tất cả các giá trị bằng 0 và chỉ mục được đánh dấu bằng 1.

Type
AA
AB
CD
AA

Onehot encoding →

Type	AA_Onehot	AB_Onehot	CD_Onehot
AA	1	0	0
AB	0	1	0
CD	0	0	1
AA	0	0	0

Hình 3.13: Cách One hot encoding biến đổi dữ liệu

One hot encoding rất hữu ích cho dữ liệu không có mối quan hệ với nhau. Các thuật toán học máy coi thứ tự của các số như một thuộc tính có ý nghĩa. Nói cách khác, chúng sẽ đọc một số cao hơn là tốt hơn hoặc quan trọng hơn một số thấp hơn.

Mặc dù điều này hữu ích cho một số tình huống thứ tự, nhưng một số dữ liệu đầu vào không có bất kỳ xếp hạng nào cho giá trị danh mục và điều này có thể dẫn đến các vấn đề với dự đoán và hiệu suất kém. Đó là khi One hot encoding giúp tiết kiệm thời gian.

One hot encoding làm cho dữ liệu huấn luyện trở nên hữu ích và biểu đạt hơn, đồng thời nó có thể được thay đổi tỷ lệ một cách dễ dàng. Bằng cách sử dụng các giá trị số sẽ giúp dễ dàng xác định xác suất cho các giá trị của mình hơn. Đặc biệt, One hot encoding được sử dụng cho các giá trị đầu ra, vì nó cung cấp nhiều dự đoán hơn so với các nhãn đơn lẻ.

3.5.2 Chuẩn hóa và trích xuất đặc trưng

Dữ liệu sau khi được xử lý sẽ được chuẩn hóa bằng kĩ thuật Feature Scaling – một trong các kỹ thuật dữ liệu giúp chuẩn hóa các đặc trưng độc lập có trong dữ liệu trong một phạm vi cố định. Nếu tỷ lệ đặc trưng không được thực hiện, thì một thuật toán học máy có xu hướng cân nặng các giá trị lớn hơn, cao hơn và xem xét các giá trị nhỏ hơn là giá trị thấp hơn, bất kể đơn vị của các giá trị. Ví dụ: Nếu một thuật toán không sử dụng phương pháp chia tỷ lệ tính năng thì có thể xem xét giá trị 3000 mét lớn hơn 5 km nhưng điều đó thực sự không đúng và trong trường hợp này, thuật toán

sẽ đưa ra dự đoán sai. Vì vậy, việc sử dụng quy mô tính năng để mang tất cả các giá trị đến cùng một cường độ. Đối với bộ dữ liệu NSL-KDD sẽ sử dụng kỹ thuật chuẩn hóa Standardization - là một kỹ thuật rất hiệu quả, giúp thu nhỏ lại một giá trị đặc trưng để nó có phân phối với 0 giá trị trung bình và phương sai bằng 1. Công thức tổng quát được thể hiện như sau:

$$X_{new} = \frac{X_i - X_{mean}}{Standard\ Deviation}$$

Đề xuất của luận văn là dựa vào số lượng các giao thức trong bộ dữ liệu mà chọn mô hình phù hợp, cụ thể là trường dữ liệu protocol_type với tổng cộng ba loại giao thức là TCP, UDP và IMCP. Nếu số lượng giao thức TCP lớn hơn tổng số lượng của hai loại giao thức còn lại thì mô hình Decision Tree sẽ là mô hình thích hợp cho bộ dữ liệu. Ngược lại, nếu số lượng giao thức IMCP lớn hơn tổng hai giao thức còn lại thì SVM sẽ là mô hình phù hợp. Trường hợp cuối cùng sẽ là KNN.

3.5.3 Phân lớp và dự đoán

Mô-đun này sẽ sử dụng các mô hình tương ứng bao gồm KNN, SVM và cây quyết định, cùng với đặc trưng đã được chọn từ mô-đun 2 để để tiến hành phân lớp. Kết quả dự đoán sau khi áp dụng các mô hình sẽ bao gồm hai phân lớp: bị tấn công hoặc không bị tấn công. Cụ thể, sau khi quá trình xử lý dữ liệu kết thúc, từng mô hình KNN, Decision Tree, và SVM sẽ được áp dụng vào dữ liệu, kết quả thu được ở mỗi mô hình sẽ được phân tích dựa trên các độ đo hiệu quả mô hình, cụ thể là: Accuracy, Precision, F - measure, Recall.

Độ chính xác (Accuracy): Là tỷ lệ phần trăm của các phiên bản được phân loại chính xác trên tổng số phiên bản.

$$AC = \frac{TP + TN}{TP + TN + FP + FN}$$

Độ chính xác (Precision): Là tỷ lệ giữa các cá thể có liên quan theo các phiên bản được truy xuất.

$$Precision = \frac{TP}{TP + FP}$$

Recall là tỷ lệ của các phiên bản có liên quan đã được truy xuất trên tổng số lượng các phiên bản có liên quan.

$$Precision = \frac{TP}{TP + FP}$$

Độ đo F1: Là giá trị trung bình có trọng số của độ chính xác và độ thu hồi.

$$F1\ measure = \frac{2 * (Recall * Precision)}{Recall + Precision}$$

- TP (True positive): là số lượng các trường hợp tích cực được phân loại sửa chữa.
- TN (True negative): là số lượng các trường hợp âm được phân loại chính xác.
- FP (False positive): là số thực thể dương bị phân loại sai.
- FN (False negative): là số lượng các trường hợp phủ định bị phân loại sai.

3.6 Kết luận chương

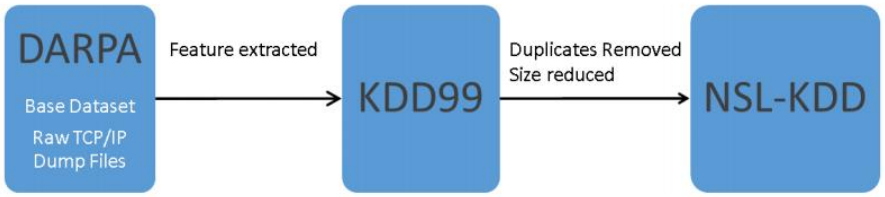
Chương 3 đã nêu khái quát về các đặc điểm của một hệ thống IDS. Ngoài ra, chương 3 cũng đề cập đến các thuật toán được sử dụng trong luận văn và mô hình đề xuất cho NIDS, các bước cụ thể xây hình mô hình đã được đề xuất sẽ được trình bày trong chương 4 của luận văn.

CHƯƠNG 4. XÂY DỰNG VÀ TRIỂN KHAI HỆ THỐNG PHÁT HIỆN XÂM NHẬP MẠNG DỰA VÀO HỌC MÁY CHO HỆ THỐNG MẠNG TRUNG TÂM Y TẾ HUYỆN GÒ DẦU

4.1 Mô tả bộ dữ liệu sử dụng NSL-KDD

Quá trình hình thành

Trong quá trình nghiên cứu về lĩnh vực bảo mật thì cái tên DARPA không quá xa lạ với mọi người. DARPA là một bộ dữ liệu thô cơ sở. Trong khi đó, KDD99 là phiên bản trích xuất tính năng của bộ dữ liệu DARPA. Tiếp đó, NSL-KDD là phiên bản được loại bỏ và giảm kích thước của bộ dữ liệu KDD99, đồng thời là bộ dữ liệu chuẩn cho dữ liệu Internet hiện nay.



Hình 4.1: Mối quan hệ giữa các bộ dữ liệu cho hệ thống IDS: DARPA, KDD99, NSL-KDD

NSL-KDD không phải là tập dữ liệu đầu tiên dành cho các IDS. Đã từng có một cuộc thi là KDD Cup, một cuộc thi quốc tế về các công cụ Khai thác tri thức và khai phá dữ liệu. Năm 1999, cuộc thi này được tổ chức với mục đích thu thập các bản ghi lưu lượng mạng. Nhiệm vụ của cuộc thi là xây dựng một hệ thống phát hiện xâm nhập mạng, một mô hình dự đoán có thể phân biệt được các kết nối "xấu" – gọi là xâm nhập hoặc tấn công – và các kết nối thông thường. Kết quả sau cuộc thi đã thu thập được một lượng bản ghi lưu lượng mạng và gom thành tập dữ liệu gọi là KDD'99, và từ đó, tập dữ liệu NSL-KDD được tạo ra, như là một phiên bản đã sửa đổi, tối ưu hóa của KDD'99 từ Đại học New Brunswick.

Tập dữ liệu này gồm 4 tập dữ liệu con: KDDTest+, KDDTest-21, KDDTrain+, KDDTrain+_20Percent, mặc dù KDDTest-21 và KDDTrain+_20Percent là các tập con của KDDTest+ và KDDTrain+. Từ đây, KDDTrain+ sẽ được xem là tập huấn luyện và KDDTest+ sẽ được xem là tập kiểm tra. Tập KDDTest-21 là một tập con của tập kiểm tra, loại bỏ những bản ghi dữ liệu khó nhất (điểm 21) và tập

KDDTrain+_20Percent là tập con của tập huấn luyện, với số bản ghi bằng 20% tổng số bản ghi có trong tập huấn luyện. Nói cách khác, các bản ghi lưu lượng mạng có trong KDDTest-21 và KDDTrain+_20Percent đã lần lượt có trong các tập kiểm tra và tập huấn luyện, đồng thời không có bản ghi nào đồng thời tồn tại trong cả 2 tập dữ liệu.

Tập dữ liệu gồm các bản ghi lưu lượng mạng Internet được quan sát bởi một mạng phát hiện xâm nhập đơn giản và là những lưu lượng một IDS có thể gặp phải, là những dấu vết còn sót lại. Tập dữ liệu gồm 43 thuộc tính trong mỗi bản ghi, với 41 thuộc tính liên quan đến chính lưu lượng, 2 thuộc tính cuối là nhãn (tấn công hoặc không tấn công) và điểm (mức độ nghiêm trọng của lưu lượng đầu vào).

Mô tả bộ dữ liệu

NSL-KDD

Trong tập dữ liệu NSL-KDD có 4 lớp tấn công bao gồm: Tấn công từ chối dịch vụ (Denial of Services – DoS), Do thám (Probe), User to Root (U2R) và Remote to Local (R2L). Mô tả ngắn gọn của mỗi lớp tấn công như sau:

- DoS là kiểu tấn công hướng đến việc gián đoạn lưu lượng mạng được gửi đến hoặc đi từ hệ thống mục tiêu. IDS bị tấn công với một lượng lưu lượng không bình thường mà nó không thể xử lý được và tự ngưng hoạt động để bảo vệ chính nó. Điều này làm cho lưu lượng bình thường không đi đến được mạng. Cho ví dụ, một đơn vị bán hàng trực tuyến bị tràn ngập các đơn hàng online trong một ngày có khuyến mãi lớn, và do mạng không thể xử lý tất cả các yêu cầu nên nó sẽ ngưng hoạt động, khiến cho các khách hàng không thể thực hiện các giao dịch thanh toán hàng. Đây là tấn công phổ biến nhất trong tập dữ liệu.
- Do thám (Probe) là kiểu tấn công cố gắng thu thập các thông tin từ một mạng nào đó. Mục tiêu giống như một kẻ trộm đánh cánh các thông tin quan trọng, bất kể là thông tin cá nhân của người dùng hoặc thông tin tài khoản ngân hàng.
- U2R là kiểu tấn công bắt đầu từ một tài khoản người dùng thông thường và cố gắng chiếm quyền truy cập vào hệ thống hoặc mạng như người dùng root. Kẻ tấn công cố tấn công vào các điểm yếu trong hệ thống để chiếm quyền root.

- R2L là kiểu tấn công cố gắng chiếm quyền truy cập vào hệ thống từ một máy tính từ xa. Một kẻ tán cong không có quền truy cập nội bộ vào hệ thống hoặc mạng và cố gắng tấn công để tìm cách truy cập vào mạng.

4.2 Môi trường mô phỏng quá trình thực nghiệm

Dựa vào dữ liệu và các thuật toán có sẵn được sử dụng cho hệ thống IDS, luận văn đề xuất mô hình phân loại các thuật toán dựa trên các giao thức truyền dẫn thông tin. Tiến hành áp dụng các thuật toán học máy trên bộ dữ liệu NSL-KDD, đánh giá các kết quả đạt được và từ đó đưa ra sự phân loại phù hợp cho các tập dữ liệu.

Cài đặt thuật toán SVM, KNN, Decision Tree trên môi trường Google Collab và kiểm nghiệm kết quả.

4.3 Kết quả thực nghiệm

Như đã trình bày ở trên, luận văn sẽ sử dụng ba thuật toán phân lớp là KNN, Decision Tree và SVM để tiến hành thực nghiệm trên bộ dữ liệu NSL-KDD. Ngoài ra, kết quả thu được sẽ tiến hành phân tích các độ đo đánh giá hiệu quả mô hình, từ đó kết hợp với quá trình chọn mô hình thích hợp cho bộ dữ liệu dựa vào số lượng của từng loại giao thức trong bộ dữ liệu, cụ thể là trường protocol_type.

Trước tiên, ta chạy thực nghiệm thuật toán KNN với bộ dữ liệu NSL-KDD, kết quả được ghi nhận như sau:

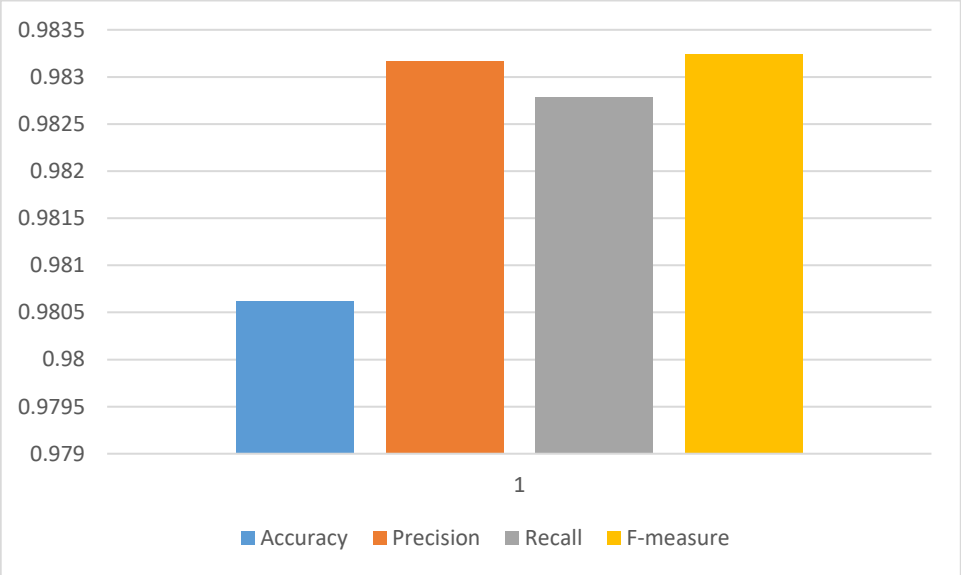
Bảng 4.1: Các độ đo accuracy, precision, recall, f-measure, traning time, testing time của thuật toán KNN trên tập dữ liệu NSL-KDD

k value	Accuracy	Precision	Recall	F-measure	Training time (s)	Testing time (s)
k = 3	0.95968	0.95506	0.97506	0.96495	0.380	11.477
k = 5	0.95755	0.95021	0.97662	0.96322	0.383	11.191
k = 10	0.95484	0.94720	0.97506	0.96091	0.390	12.849

Bảng 4.1 ghi nhận các kết quả thu được sau khi thực nghiệm tập dữ liệu với thuật toán KNN. Kết quả cho thấy thuật toán KNN có đầu ra tốt nhất tại n = 3 với độ chính xác 95.97%. Các độ đo hiệu quả mô hình còn lại cũng đạt các giá trị cao tương tự Accuracy.

Bảng 4.2 Các độ đo accuracy, precision, recall, f-measure, traning time, testing time của thuật toán Decision Tree trên tập dữ liệu NSL-KDD

Accuracy	Precision	Recall	F-measure	Training time (s)	Testing time (s)
0.98062	0.98317	0.98278	0.98324	0.557	0.004



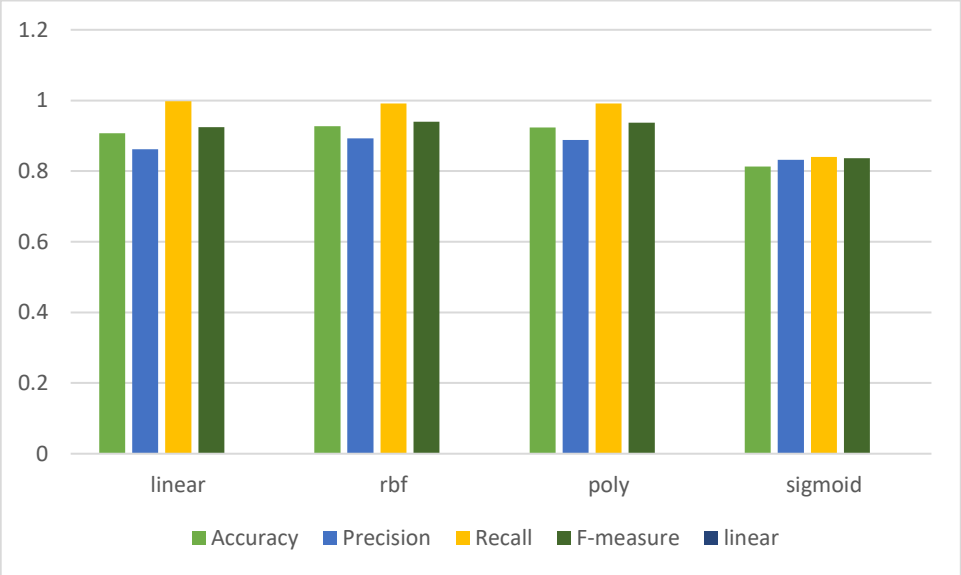
Hình 4.2: Biểu đồ thể hiện độ đo khi áp dụng thuật toán Decision Tree trên tập dữ liệu NSL-KDD

Thuật toán Decision Tree ghi nhận các kết quả vượt trội, với độ chính xác (Accuracy) đạt 98.06%, hơn khoảng 2.09% so với thuật toán KNN. Ngoài ra, thời thử nghiệm (Testing time) của thuật toán cũng ghi nhận kết quả ấn tượng chỉ với 40 mili giây.

Bảng 4.3 Các độ đo accuracy, precision, recall, f-measure, traning time, testing time của thuật toán SVM trên tập dữ liệu NSL-KDD

Measure	Kernel			
	linear	rbf	poly	sigmoid
Accuracy	0.90751	0.92716	0.92366	0.81263
Precision	0.86203	0.89232	0.88810	0.83241
Recall	0.99735	0.99182	0.99182	0.84018
F-measure	0.92472	0.93942	0.93663	0.83622

Training time (s)	449.304s	91.381	87.186	465.570s
Testing time (s)	11.112s	11.467	6.117	24.432s

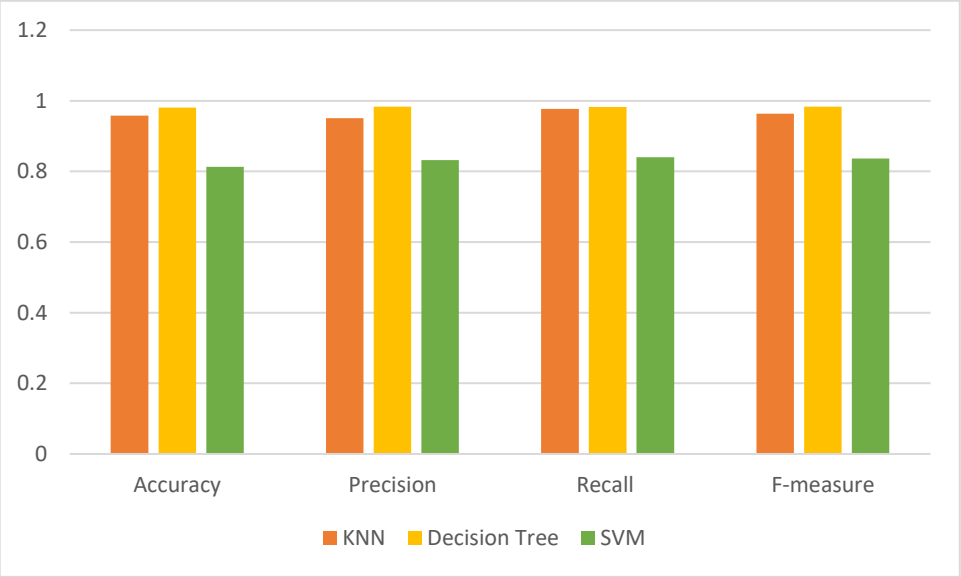


Hình 4.3: Biểu đồ thể hiện độ đo khi áp dụng thuật toán SVM trên tập dữ liệu NSL-KDD

Bảng 4.3 mô tả các kết quả thu được với từng loại kernel khác nhau của thuật toán SVM. Qua đó có thể thấy rõ được rằng ở kernel rbf ghi nhận kết quả tốt nhất với độ chính xác đạt 92.71%. Tuy nhiên, nếu vừa xét về thời gian và độ chính xác thì với kernel là poly vượt trội hơn so với rbf. Thời gian training và testing ở kernel poly lần lượt là 87.186 và 6.117 giây, đồng thời độ chính xác chỉ kém 0.35% so với kernel rbf.

Bảng 4.4: Các độ đo accuracy, precision, recall, f-measure, traning time, testing time tổng hợp từ ba thuật toán trên tập dữ liệu NSL-KDD

Classifier	**Accuracy**	**Precision**	**Recall**	**F-measure**
KNN	0.95755	0.95021	0.97662	0.96322
Decision Tree	0.98062	0.98317	0.98278	0.98324
SVM	0.81263	0.83241	0.84018	0.83622



Hình 4.4: Biểu đồ thể hiện tổng hợp độ đo khi áp dụng các thuật toán trên tập dữ liệu NSL-KDD

Bảng 4.4 cho thấy kết quả tổng hợp của ba thuật toán KNN, Decision Tree và SVM. Kết quả tổng hợp được lấy kết quả tốt nhất mỗi lần chạy của từng thuật toán. Dựa vào bảng tổng hợp kết quả, có thể thấy được Decision Tree ghi nhận các thông số đầu ra vượt trội hơn so với hai thuật toán còn lại. Dựa các kết quả này, mô hình đề xuất sẽ có thể triển khai và áp dụng vào việc lựa chọn mô hình cho các bộ dữ liệu sau.

Xét bộ dữ liệu NSL-KDD, trường dữ liệu protocol_type có tổng cộng ba loại giao thức là TCP, UDP và IMCP. Trong đó, số lượng của từng loại giao thức được mô tả cụ thể như sau: TCP, UDP, IMCP có số lượng lần lượt là 18880, 2621 và 1043. Theo kết quả có được, với dữ liệu có số lượng giao thức TCP lớn hơn số lượng các loại giao thức còn lại thì Decision Tree sẽ là mô hình thích hợp cho bộ dữ liệu. Tương tự như vậy, đối với dữ liệu có số lượng UDP lớn hơn số lượng các giao thức còn lại thì KNN sẽ là mô hình phù hợp. Và cuối cùng là ứng với dữ liệu có số lượng IMCP lớn hơn số lượng các giao thức còn lại thì SVM sẽ là mô hình phù hợp.

4.4 Kết luận chương

Chương 4 đã giới thiệu và mô tả quá trình thực nghiệm trên bộ dữ liệu NSL-KDD - bộ dữ liệu được lấy làm chuẩn để giúp các nhà nghiên cứu so sánh các phương thức phát hiện xâm nhập khác nhau. Các kết quả được nhận xét, phân tích và đánh giá. Đồng thời, chương 4 cũng đã đề xuất phương pháp chọn lọc hệ thống phù hợp với trường dữ liệu protocol_type của mỗi dịch vụ, xem xét chọn trường dữ liệu phù hợp, từ đó chọn mô hình tốt nhất cho bộ dữ liệu.

KẾT LUẬN

1. Kết quả nghiên cứu của đề tài

Trên thế giới, các công trình nghiên cứu về lĩnh vực an ninh, an toàn thông tin được tiến hành bởi các trường đại học về khoa học máy tính, công nghệ thông tin, nhưng những công trình nổi bật và được ứng dụng rộng rãi là những công trình nghiên cứu được tiến hành bởi các công ty bảo mật. Một số công ty bảo mật hàng đầu phát triển các bộ giải pháp về an ninh mạng có thể kể đến là: Cisco System, Juniper Network, TrendMicro.. Sản phẩm của nghiên cứu mà các công ty bảo mật tiến hành có thể là những sản phẩm riêng rẽ như Cisco IDS, Juniper IPS, TrendMicro Server Protect, hoặc cũng có thể là cả một bộ giải pháp với nhiều sản phẩm phần cứng phần mềm khác nhau.

Về việc nghiên cứu phát triển các hệ thống phát hiện xâm nhập, có một số xu hướng nghiên cứu về lĩnh vực này đã được thực hiện. Đó là phát hiện xâm nhập dựa trên dấu hiệu xâm nhập; phát hiện xâm nhập dựa vào khả năng tự học của hệthống; phát hiện xâm nhập bằng cách kết hợp cả hai phương pháp trên. Đối với mỗi phương pháp phát hiện xâm nhập sẽ dẫn đến rất nhiều nghiên cứu, như những nghiên cứu về trí tuệ nhân tạo, hệ chuyên gia, nghiên cứu về phương pháp đặt luật so sánh trong phân tích lưu thông mạng.

Như vậy, an ninh thông tin là lĩnh vực ngày càng có thêm nhiều công trình nghiên cứu cũng như các sản phẩm được tạo ra từ các công trình đó, bởi trình độ của các hacker cũng ngày càng tiến bộ. Tuy rằng các sản phẩm an ninh thông tin mà các công ty nước ngoài phát triển hoạt động hiệu quả, nhưng các sản phẩm này có mức giá cao và hàng năm phải tốn chi phí để cập nhật. Việc áp dụng các sản phẩm này một cách thụ động cũng là một nhược điểm. Vì khi đó đội ngũ quản trị mạng sẽ không thực sự hiểu bản chất hệ thống hoạt động như thế nào, hệ thống phân tích những gì ở mức dưới của hệ thống thông tin, dẫn đến việc không linh hoạt trong nghiệp vụ quản trị bảo mật.

Với việc hoàn thành các sản phẩm của luận văn là hệ thống phát hiện xâm nhập mạng dựa trên nền tảng mã nguồn mở và bộ quy trình phòng ngừa và ngăn chặn xâm nhập mạng, hy vọng rằng việc ứng dụng các sản phẩm này sẽ góp phần cải thiện những điểm yếu trong hệ thống an ninh thông tin của Trung tâm Y tế huyện Gò Dầu.

Bên cạnh đó, nó sẽ mở ra những hướng phát triển tiếp theo trong nghiên cứu và ứng dụng hệ thống phát hiện xâm nhập mạng, giúp nền công nghệ thông tin nước ta có những bước tiến trong ứng dụng và làm chủ những sản phẩm công nghệ về an ninh thông tin.

2. Hạn chế của luận văn

Bên cạnh các kết quả đạt được, luận văn cũng còn một số hạn chế nhất định. Trong đó, dữ liệu sử dụng chưa phải là dữ liệu thực tế từ hệ thống IDS, từ đó chưa thể áp dụng được mô hình đề xuất vào bộ dữ liệu thực tế để kiểm tra cũng như đánh giá kết quả một cách trực quan nhất. Ngoài ra, hiện nay một số hệ thống IDS đã áp dụng các thuật toán học sâu (một hướng nghiên cứu phát triển từ học máy) tối ưu hơn, nhằm mục đích cải thiện các tốc độ xử lý dữ liệu cũng như phát hiện được những loại tấn công mạng nguy hiểm hơn.

3. Hướng phát triển của luận văn

Áp dụng các kỹ thuật cũng như thuật toán tối ưu hơn về tốc độ xử lý, thời gian huấn luyện và thử nghiệm cho mô hình. Ngoài ra, dữ liệu sử dụng sẽ thay thế bằng các tập dữ liệu thực tế được lấy từ một hệ thống IDS đang hoạt động bất kỳ, từ đó có thể đưa ra các kết quả cũng như đánh giá khách quan nhất về mức độ hiệu quả của hệ thống phát hiện xâm nhập mạng.

DANH MỤC TÀI LIỆU THAM KHẢO

[1] Umesh Hodeghatta Rao & Umesha Nayak, Intrusion Detection and Prevention Systems, 2014.

[2] Vnetwork, 3 November 2020. [Online]. Available: https://vnetwork.vn/news/4-loai-tan-cong-mang-nguy-hiem-nhat-hien-nay.

[3] R. Daş, A. Karabade and G. Tuna, "Common network attack types and defense mechanisms,"," *23nd Signal Processing and Communications Applications Conference (SIU),* 2015.

[4] Aleksey A. Titorenko, Alexey A. Frolov, “Analysis of Modern Intrusion Detection System,” 2018.

[5] H. Yao, D. Fu, P. Zhang, M. Li and Y. Liu, "MSML: A Novel Multilevel Semi-Supervised Machine Learning Framework for Intrusion Detection System," *IEEE Internet of Things Journal,* vol. 6, pp. 1949-1959, 2019.

[6] M. H. Ali, B. A. D. Al Mohammed, A. Ismail and M. F. Zolkipli, "A New Intrusion Detection System Based on Fast Learning Network and Particle Swarm Optimization," *IEEE Access,* vol. 6, pp. 20255-20261, 2018.

[7] Kim-Hung Le, Minh-Huy Nguyen, Trong-Dat Tran, Ngoc-Duan Tran, “IMIDS: An Intelligent Intrusion Detection System against Cyber Threats in IoT,” 10 February 2022.

[8] Xuan-Ha Nguyen, Xuan-Duong Nguyen, Hoang-Hai Huynh, Kim-Hung Le, “Realguard: A Lightweight Network Intrusion Detection System for IoT Gateways,” 7 January 2022.

[9] Tran Ngoc Thinh, Tran Hoang Quoc Bao, Duc-Minh Ngo, Cuong Pham-Quoc, “High-performance anomaly intrusion detection system with ensemble neural networks on reconfigurable hardware,” 26 April 2021.

[10] Navaporn Chockwanich, Vasaka Visoottiviseth, “Intrusion Detection by Deep Learning with TensorFlow,” trong *International Conference on Advanced Communications Technology(ICACT),* 2019.

[11] Jiadong Ren, Jiawei Guo, Wang Qian, Huang Yuan, Xiaobing Hao , and Hu Jingjing, “Building an Effective Intrusion Detection System by Using Hybrid Data Optimization Based on Machine Learning Algorithms,” 16 June 2019.

[12] Kinan Ghanem, Francisco J. Aparicio-Navarro, Konstantinos G. Kyriakopoulos, Sangarapillai Lambotharan, Jonathon A. Chambers, “Support Vector Machine for Network Intrusion and Cyber-Attack Detection,” trong *Sensor Signal Processing for Defence (SSPD)*, London, 2017.

[13] Anar Ahady, Ali Ghubaish, Tara Salman, Devrim Ünal, “Intrusion Detection System for Healthcare Systems Using Medical and Network Data: A Comparison Study,” February 2017.

[14] Isra Al-Turaiki, Najwa Altwaijry, “A Convolutional Neural Network for Improved Anomaly-Based Network Intrusion Detection,” *Big Data,* tập 9, p. 233–252, 2021.

[15] Nuno Oliveira, Isabel Praça, Eva Maia, Orlando Sousa , “Intelligent Cyber Attack Detection and Classification for Network-Based Intrusion Detection Systems,” 23 November 2021.

[16] Oqbah Ghassan Abbas, Khaldoun Khorzom, Mohammed Assora, “Machine Learning based Intrusion Detection System for Software Defined Networks,” *International Journal of Engineering Research & Technology (IJERT),* tập 9, số 09, 2020.

[17] Mohannad Zead Khairallah, “Network Attacks Detection using Deep neural network,” 2021.

[18] Roza Dastres, Mohsen Soori, “A Review in Recent Development of Network Threats and Security Measures,” *International Journal of Computer and Information Engineering,* tập 15, 2021.

[19] E. Lundin, E. Jonsson, "Survey of research in the intrusion detection area," Department of Computer Engineering, Goteborg, 2002.

[20] Ahmadian Ramaki, Ali & Ebrahimi Atani, Reza, "A survey of IT early warning systems: architectures, challenges, and solutions," *Security and Communication Networks,* 2016.

[21] Jack Koziol, Intrusion Detection with Snort, 2003.

[22] Ratul Chowdhury, Pallabi Banerjee, Soumya Deep Dey, Banani Saha, Samir Kumar Bandyopadhyay, “A Decision Tree Based Intrusion Detection System for Identification of Malicious Web Attacks,” 9 July 2020.

[23] D. Singh, “A Collaborative IDS Framework for Cloud,” Research Gate, 2013.

[24] M.Nikhitha, M.A.Jabbar, “K Nearest Neighbor Based Model for Intrusion Detection System,” *International Journal of Recent Technology and Engineering (IJRTE)*, tập 8, số 2, July 2019.

[25] B. Basaveswara Rao, K.Swathi, “Fast kNN Classifiers for Network Intrusion Detection System,” *Indian Journal of Science and Technology*, tập 10, 2017.

[26] Arushi Agarwal, Purushottam Sharma, Mohammed Alshehri, Ahmed A. Mohamed, , Osama Alfarraj, “Classification model for accuracy and intrusion detection using machine learning approach,” p. 20, 7 April 2021.

[27] A. Chugh, “GreeksforGreeks,” 24 Sep 2021. [Trực tuyến]. Available: https://www.geeksforgeeks.org/ml-label-encoding-of-datasets-in-python/.

[28] A. Fawcett, “Educative,” 11 Feb 2021. [Trực tuyến]. Available: https://www.educative.io/blog/one-hot-encoding.

[29] NM Shanono, Zulkiflee M, NA Abu, W. Yassin, MA Faizal, “Intrusion Detection System Architecture: Issues and Challenges,” tập 62, August 2020.

BẢNG CAM ĐOAN

Tôi cam đoan đã thực hiện việc kiểm tra mức độ tương đồng nội dung luận văn/luận án qua phần mềm DoIT một cách trung thực và đạt kết quả mức độ tương đồng **13%** toàn bộ nội dung luận văn/luận án. Bản luận văn/luận án kiểm tra qua phần mềm là bản cứng luận văn đã nộp để bảo vệ trước hội đồng. Nếu sai tôi xin chịu các hình thức kỷ luật theo quy định hiện hành của Học viện.

TPHCM, ngày 15 tháng 7 năm 2022

HỌC VIÊN CAO HỌC

Bùi Điền Phong