ĐẠI HỌC BÁCH KHOA HÀ NỘI

ĐỒ ÁN TỐT NGHIỆP

Tích hợp và chuẩn hóa dữ liệu đầu vào cho chức năng đánh giá rủi ro bị tấn công APT của hệ thống RiDX

PHẠM HUY HOÀNG
hoang.ph204653@sis.hust.edu.vn

Ngành: Khoa học máy tính

Giảng viên hướng dẫn: TS. Vũ Thị Hương Giang _____

ThS. Nguyễn Mạnh Tuấn _____

Khoa: Khoa học máy tính

Trường: Công nghệ Thông tin và Truyền thông

HÀ NỘI, 07/2024

LỜI CẢM ƠN

Đầu tiên em xin phép gửi lời cảm ơn chân thành đến TS. Vũ Thị Hương Giang, người trực tiếp hướng dẫn em trong quá trình thực hiện đồ án. Cô đã tận tình chỉ bảo, giải đáp những thắc mắc và tiếp thêm động lực để em hoàn thiện các công việc được giao. Nhờ sự hướng dẫn của cô em đã tự tin trong việc giải quyết các vấn đề khó khăn mà đề tài đặt ra. Em cũng xin gửi lời cảm ơn đến ThS. Nguyễn Mạnh Tuấn, giáo viên đồng hướng dẫn của em trong suốt thời gian làm đồ án. Tuy không phải là giáo viên hướng dẫn trực tiếp nhưng thầy vẫn luôn nhiệt tình góp ý giúp cho đồ án của em được hoàn thiện hơn. Nhờ có cô và thầy em đã vượt qua những khó khăn trong quá trình thực hiện đề tài. Tuy chỉ mới đồng hành cùng thầy cô trong một kỳ ngắn ngủi nhưng em đã được học rất nhiều điều về kiến thức và kỹ năng, giúp bản thân trưởng thành hơn.

Em muốn gửi lời cảm ơn chân thành đến bố mẹ, những người luôn ủng hộ, động viên em trong quá trình học tập tại trường Đại học Bách khoa Hà Nội. Bố mẹ đã hy sinh nhiều thứ để em có cơ hội được học tập, rèn luyện tại trường qua đó trở thành công dân có ích cho xã hội. Em cũng xin gửi lời cảm ơn đến anh chị, bạn bè những người luôn đồng hành, hỗ trơ cùng em trong những lúc khó khăn để hoàn thiện đồ án này.

LỜI CAM KẾT

Họ và tên sinh viên:	Phạm Huy Hoàng
Điện thoại liên lạc:	0327509130
Email:	hoang.ph204653@sis.hust.edu.vn
Lớp:	Khoa học máy tính 04 - K65
Hệ đào tạo:	Cử nhân

Tôi – *Phạm Huy Hoàng* – cam kết Đồ án Tốt nghiệp (ĐATN) là công trình nghiên cứu của bản thân tôi dưới sự hướng dẫn của *TS. Vũ Thị Hương Giang*. Các kết quả nêu trong ĐATN là trung thực, là thành quả của riêng tôi, không sao chép theo bất kỳ công trình nào khác. Tất cả những tham khảo trong ĐATN – bao gồm hình ảnh, bảng biểu, số liệu, và các câu từ trích dẫn – đều được ghi rõ ràng và đầy đủ nguồn gốc trong danh mục tài liệu tham khảo. Tôi xin hoàn toàn chịu trách nhiệm với dù chỉ một sao chép vi phạm quy chế của nhà trường.

Hà Nội, ngày tháng năm

Tác giả ĐATN

Phạm Huy Hoàng

TÓM TẮT NỘI DUNG ĐỒ ÁN

Các hệ thống web đã và đang hiện diện trong nhiều lĩnh vực đời sống, đóng góp vai trò quan trọng thúc đẩy phát triển kinh tế xã hội. Đảm bảo an ninh an toàn thông tin luôn là ưu tiên hàng đầu của các tổ chức khi đưa vào vận hành hệ thống CNTT. Với sự xuất hiện ngày càng nhiều của các cuộc tấn công APT (Advance Persistent Threat) gây thiệt hại lớn về tài sản và cơ sở hạ tầng, cần phải có những công cụ giám sát và đánh giá rủi ro ATTT hiệu quả nhằm đưa ra các biện pháp kịp thời giúp hạn chế và giảm thiểu thiệt hại.

Hệ thống RiDX ra đời với mục đích tạo ra một công cụ đánh giá rủi ro ATTT sử dụng mạng Bayes với đầu vào là kịch bản rủi ro của hệ thống. RiDX gồm các chức năng chính như quản lý tài sản, xây dựng kịch bản triển khai hệ thống, đánh giá rủi ro hệ thống trước khi đưa vào vận hành. Tuy nhiên, RiDX chưa đánh giá được mức độ rủi ro và đưa ra những cảnh báo kịp thời về khả năng xảy ra tấn công APT từ các dữ liệu giám sát lưu lượng mạng thời gian thực thu được khi vận hành kịch bản triển khai hệ thống. Lý do chính là do thiếu cơ chế ánh xạ thời gian thực các dữ liệu này vào đầu vào của mạng Bayes đánh giá rủi ro.

Trong phạm vi DATN, em đề xuất 2 giải pháp khắc phục các hạn chế trên.

Một là cơ chế giám sát dữ liệu thời gian thực và cảnh báo rủi ro tấn công APT trong giai đoạn vận hành kịch bản triển khai. Cơ chế này được cài đặt bằng cách mở rộng microservice đánh giá rủi ro ATTT của hệ thống RiDX.

Hai là cơ chế chuẩn hóa và tích hợp dữ liệu giám sát lưu lượng mạng thời gian thực khi vận hành kịch bản triển khai hệ thống thành đầu vào của mạng Bayes đánh giá rủi ro của RiDX. Ý tưởng là sử dụng các mô hình học máy nhằm phán đoán thông tin về cuộc tấn công xảy ra như giai đoạn tấn công, năng lực tấn công làm đầu vào cho mạng đánh giá rủi ro. Cơ chế này được cài đặt tích hợp với chức năng trên.

Hệ thống RiDX được thử nghiệm với 2 bộ dữ liệu mô phỏng tấn công APT thời gian thực là Unraveled và SCVIC-APT-2021. So sánh kết quả đánh giá rủi ro bị tấn công APT với các nhãn đánh dấu giai đoạn tấn công, mức độ ảnh hưởng tấn công cho thấy hệ thống đưa ra cảnh báo chính xác đạt 90% về mức độ rủi ro của hệ thống khi xảy ra các cuộc tấn công APT, chứng minh được tính hiệu quả trong việc phát hiện rủi ro về tấn công APT trong quá trình giám sát.

MỤC LỤC

DANH MỤC HÌNH VẼ

DANH MỤC BẢNG BIỂU

DANH MỤC THUẬT NGỮ VÀ TỪ VIẾT TẮT

Thuật ngữ	Ý nghĩa
ĐATN	Đồ án tốt nghiệp
ĐHBKHN	Đại học Bách khoa Hà Nội
API	Giao diện lập trình ứng dụng (Application Programming Interface
APT	Advance Persistent Threat
ATTT	An toàn thông tin
CNTT	Công nghệ thông tin
CPE	Lược đồ đặt tên có cấu trúc cho các hệ thống, phần mềm và gói công nghệ thông tin (Common Platform Enumeration)
CPU	Bộ vi xử lý trung tâm (central processing unit)
CSDL	Cơ sở dữ liệu
CSV	comma-separated values
CVE	Danh sách các lỗ hổng bảo mật được tiết lộ và phơi bày công khai (Common Vulnerabilities and Exposures)
HTTP	Giao thức truyền tải siêu văn bản (Hypertext Transfer Protocol)
SV	Sinh viên

CHƯƠNG 1. GIỚI THIỆU ĐỀ TÀI

1.1 Đặt vấn đề

An toàn thông tin luôn là vấn đề quan trọng hàng đầu trong quá trình triển khai, vận hành một hệ thống CNTT. Với nhu cầu tăng tốc quá trình phát triển và phát hành phần mềm, các tổ chức ngày càng sử dụng các phương pháp phát triển linh hoạt CI/CD (Continuous Integration and Continuous Delivery - Tích hợp và triển khai liên tục) cùng với các quy trình phát triển phần mềm như Agile và Scrum. Hệ thống càng phức tạp dẫn đến nguy cơ về khai thác rủi ro an ninh thông tin ngày càng lớn. Điều này đòi hỏi phải có các cơ chế quản lý, bảo trì và giám sát rủi ro hiệu quả nhằm giảm thiểu các mối đe dọa cho hệ thống.

Các tổ chức ngày càng chú tâm trong việc đảm bảo an toàn vận hành hệ thống phần mềm của họ. Trước khi đưa vào vận hành, hệ thống thường được xây dựng theo một kịch bản triển khai (**deployment scenario**)[1]. Đây là cơ sở để xác định các tiêu chí và yêu cầu về cơ sở hạ tầng phục vụ hoạt động của hệ thống web. Kịch bản triển khai bao gồm thông tin về danh sách các tài sản (phần cứng, phần mềm, ...), mối quan hệ giữa các tài sản, các mối đe dọa (CVE) tồn tại trên tài sản và biện pháp phòng vệ. Điều này cho phép các tổ chức tiến hành đánh giá rủi ro, phát hiện các lỗ hổng dễ bị khai thác của hệ thống qua đó có thể đưa ra những biện pháp ngăn chặn rủi ro kịp thời trước khi đưa vào vận hành.

Hiện nay xuất hiện ngày nhiều báo cáo về các cuộc tấn công mạng nhắm đến các hệ thống CNTT trọng yếu. Các cuộc tấn công Advanced Persistent Threat (APT) đang là vấn đề nhức nhối đối với các hệ thống cơ sở hạ tầng thông tin thế giới nói chung và Việt Nam nói riêng. Đây là chiến dịch tấn công phức tạp và dai dẳng, sử dụng các kỹ thuật tiên tiến nhất để nhắm đến điểm yếu của hệ thống do một nhóm kẻ tấn công thực hiện. Mục tiêu của các cuộc tấn công APT được các đối tượng lựa chọn kỹ lưỡng và thường là các doanh nghiệp lớn, các cơ quan an ninh chính phủ. Những cuộc tấn công này thường gây ra những hậu quả nghiêm trọng đối với hệ thống như đánh cắp dữ liệu nhạy cảm, gián đoạn dịch vụ và phá hoại cơ sở hạ tầng. Trong bối cảnh bùng bổ về trí tuệ nhân tạo và dữ liệu lớn hiện nay, các cuộc tấn công APT còn được hỗ trợ bởi các mô hình ngôn ngữ lớn mạnh mẽ khiến cho các tổ chức ngày càng khó phát hiện và phòng ngừa.

Trước những vấn đề trên, việc áp dụng những cơ chế giám sát và theo dõi rủi ro về tấn công APT cho hệ thống CNTT trong quá trình vận hành là cực kì quan trọng. Hiện nay việc đánh giá rủi ro an ninh thông tin đã và đang được nghiên cứu và phát triển rộng rãi. Ngày càng nhiều phương pháp và công cụ được các doạn

nghiệp và tổ chức sử dụng để đáp ứng nhu cầu ngày càng tăng trong việc cảnh báo rủi ro ATTT. RiDX là hệ thống đánh giá và giám sát rủi ro ATTT đối với kịch bản triển khai dựa trên các nghiên cứu sử dụng mạng Bayes động [1][2]. Hệ thống cung cấp các cơ chế đánh giá mức độ nghiêm trọng của hệ thống theo các giai đoạn phân tích yêu cầu, triển khai và vận hành căn cứ vào các mối đe dọa tồn tại trong kịch bản triển khai. Phiên bản hiện tại của hệ thống RiDX đã hoàn thiện các chức năng cơ bản như xây dựng kịch bản triển khai, xây dựng đồ thị tấn công và đánh giá rủi ro kịch bản trước khi đưa vào triển khai. Tuy nhiên hệ thống chưa có cơ chế theo dõi, giám sát dữ liệu hoạt động của kịch bản triển khai trong giai đoạn vận hành nhằm phát hiện và cảnh báo kịp thời về dấu hiệu của cuộc tấn công APT có thể xảy ra. Xây dựng một cơ chế giám sát, cảnh báo rủi ro ATTT với các cuộc tấn công APT là mục tiêu đồ án hướng tới.

Dữ liệu được sử dụng để giám sát rủi ro là dữ liệu được thu thập từ lưu lượng mạng, phần cứng, hệ thống bảo mật, môi trường ảo hóa,... của hệ thống CNTT đang được vận hành. Để xây dựng cơ chế phát hiện rủi ro tấn công APT, đồ án sử dụng các tập dữ liệu đo lường lưu lượng truy cập mạng dưới dạng các tệp tin pcap trong quá trình mô phỏng cuộc tấn công APT xảy ra đối với hệ thống thử nghiệm và được chuyên gia gán nhãn tương ứng với giai đoạn tấn công, năng lực tấn công của attacker [3] [4]. Dữ liệu sẽ được tích hợp với đầu vào mạng đánh giá rủi ro của hệ thống RiDX nhằm cố gắng đưa ra kết quả đánh giá rủi ro khớp với thông tin về cuộc tấn công tại thời điểm giám sát. Đầu vào mạng đánh giá rủi ro được xây dựng bao gồm 3 subnet là: năng lực tấn công của attacker, thông tin kịch bản triển khai và năng lực phòng thủ. Tuy nhiên chưa có cơ chế ánh xạ các trường thông tin về lưu lượng mạng với 3 subnet được đưa ra ở trên. Đồ án sẽ tập trung giải quyết vấn đề này.

1.2 Mục tiêu và phạm vi đề tài

Phần 1.1 đã chỉ ra những hạn chế của hệ thống RiDX. Phiên bản hiện tại của hệ thống chưa có chức năng giám sát rủi ro tấn công APT thời gian thực cho kịch bản triển khai. Bên cạnh đó cần tích hợp dữ liệu thu thập thực tế hoạt động của hệ thống trong giai đoạn vận hành khi xuất hiện tấn công APT cho chức năng giám sát rủi ro tấn công APT của hệ thống RiDX ở trên. Vì vậy đồ án sẽ gồm 2 mục tiêu chính: (i) Xây dựng chơ chế giám sát tấn công APT đối với dữ liệu thời gian thực (ii) Tích hợp và chuẩn hóa dữ liệu thu thập thực tế khi xảy ra tấn công APT với đầu vào mạng đánh giá rủi ro RiDX.

1.3 Định hướng giải pháp

Các giải pháp được đưa ra bao gồm:

- Xây dựng cơ chế giám sát rủi ro tấn công APT làm việc với dữ liệu thời gian thực. Chức năng sẽ được tích hợp vào hệ thống RiDX và được áp dụng với giai đoạn triển khai hệ thống.
- Tích hợp dữ liệu thực tế về tấn công APT làm đầu vào cho mạng đánh giá rủi ro RiDX. Sử dụng các phương pháp học máy để phán đoán nhãn của dữ liệu đầu vào và ánh xạ với các subnet đầu vào của mạng đánh giá rủi ro.
- Xây dựng lại các kịch bản triển khai tương ứng với bộ dữ liệu đầu vào. Thông tin chi tiết về các bộ dữ liệu được trình bày tại chương 3 và kịch bản triển khai chi tiết được trình bày tại chương 5.

1.4 Bố cục đồ án

Bố cục của báo cáo đồ án được tổ chức như sau.

Chương 2 trình bày về khảo sát và phân tích yêu cầu của đồ án. Chương này sẽ trình bày tổng quan các yêu cầu về chức năng giám sát rủi ro tấn công APT được tích hợp vào hệ thống RiDX.

Chương 3 trình bày về nền tảng lý thuyết và công nghệ sử dụng. Chương này sẽ nêu ra nền tảng lý thuyết cốt lõi trong việc đánh giá rủi ro ATTT sử dụng mạng Bayes động, thông tin về các bộ dữ liệu được tích hợp và các công nghệ được sử dụng để xây dựng hệ thống.

Chương 4 sẽ trình bày về kết quả thiết kế kiến trúc của hệ thống và thiết kế chi tiết chức năng giám sát rủi ro tấn công APT. Các kịch bản kiểm thử sẽ được sử dụng để đánh giá tính hiệu quả của hệ thống.

Chương 5 sẽ trình bày về các đóng góp chính của đồ án nhằm giải thích các phương pháp sử dụng và kết quả thu được.

Chương 6 trình bày về các kết quả của đồ án và định hướng phát triển cho sản phẩm.

CHƯƠNG 2. KHẢO SÁT VÀ PHÂN TÍCH YÊU CẦU

2.1 Khảo sát hiện trạng

Bảo mật thông tin luôn là ưu tiên hàng đầu của các tổ chức CNTT khi phát triển và đưa vào vận hành sản phẩm. Thống kê trên thế giới cho thấy giới mỗi giây có tới 900 cuộc tấn công mạng và 5 mã độc mới sinh ra, phát hiện 40 điểm yếu lỗ hổng mỗi ngày. Riêng tại Việt Nam thống kê năm 2023 có 13.900 vụ tấn công an ninh mạng vào các hệ thống, tăng 9,5% so với năm 2022. Đặc biệt ngày càng có nhiều báo cáo về các cuộc tấn công có chủ đích APT nhắm vào các cơ quan CNTT trọng yếu. Các cuộc tấn công ngày càng trở nên phức tạp, sử dụng các kỹ thuật tiên tiến đã gây ra nhiều hậu quả nghiêm trọng đối với hệ thống như đánh cắp dữ liệu nhạy cảm, gián đoạn dịch vụ và phá hoại cơ sở hạ tầng.

Trong bối cảnh đó, việc giám sát và đánh giá rủi ro an toàn thông tin ngày càng trở nên quan trọng. Yêu cầu được đặt ra là có một cơ chế giám sát thời gian thực đánh giá mức độ rủi ro, cảnh báo sớm rủi ro tấn công APT một cách hiệu quả cho hệ thống CNTT trong giai đoạn vận hành qua đó có thể đưa ra các biện pháp ngăn chặn kịp thời.

2.2 Tổng quan chức năng

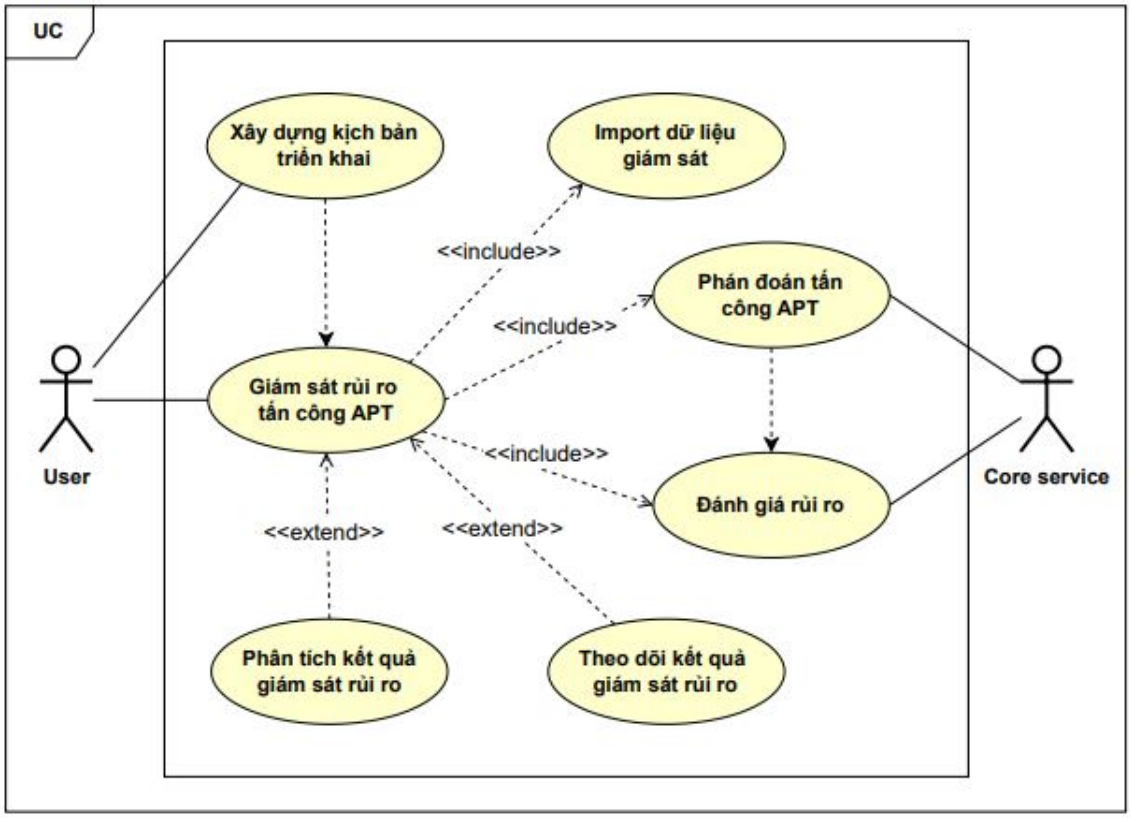
Hệ thống RiDX được phát triển bởi các sinh viên ĐHBKHN với mục đích xây dựng mô hình đánh giá rủi ro ATTT cho kịch bản triển khai của hệ thống CNTT. Hệ thống cung cấp các dịch vụ xây dựng kịch bản triển khai, đánh giá rủi ro ATTT trước khi đưa vào vận hành. Các chức năng quan trọng của hệ thống được liệt kê như sau:

- **Quản lý danh mục tài sản và danh mục lỗ hổng bảo mật**: Hệ thống quản lý danh mục các loại tài sản như phần cứng, phần mềm, hệ điều hành,... cho phép người dùng khởi tạo, thống kê tình trạng tài sản trong kịch bản của mình. Hệ thống còn quản lý và cập nhật liên tục danh mục về CPE, CVE là kênh tham khảo giúp người dùng xây dựng kịch bản triển khai phù hợp.
- **Xây dựng kịch bản triển khai**: Cho phép người sử dụng xây dựng kịch bản triển khai bao gồm tài sản, mục tiêu bảo mật và biện pháp phòng ngừa. Từ kịch bản triển khai có thể xác định được các lỗ hổng tồn tại trên tài sản, đồ thị tấn công dựa trên mối quan hệ về quyền truy cập giữa các tài sản.
- **Đánh giá rủi ro ATTT của kịch bản triển khai**: Đánh giá rủi ro trên 1 kịch bản triển khai sử dụng mạng Bayes động với đầu vào là 3 subnet bao gồm kịch bản triển khai, khả năng tấn công và khả năng phòng thủ.

- **Thu thập dữ liệu tự động**: Tự động thu thập, cập nhật dữ liệu từ các nguồn chuyên gia về danh mục CVE, CPE, CWE và cấu hình thông số khả năng tấn công APT của attacker và khả năng phòng thủ của hệ thống.

Chức năng giám sát rủi ro tấn công APT được xây dựng tích hợp vào hệ thống RiDX với mục tiêu cung cấp một cơ chế giám sát đánh giá rủi ro kịch bản triển khai trong giai đoạn vận hành từ dữ liệu thu thập lưu lượng mạng và cảnh báo nguy cơ về một cuộc tấn công APT có thể xảy ra. Dưới đây là mô tả chi tiết về yêu cầu chức năng.

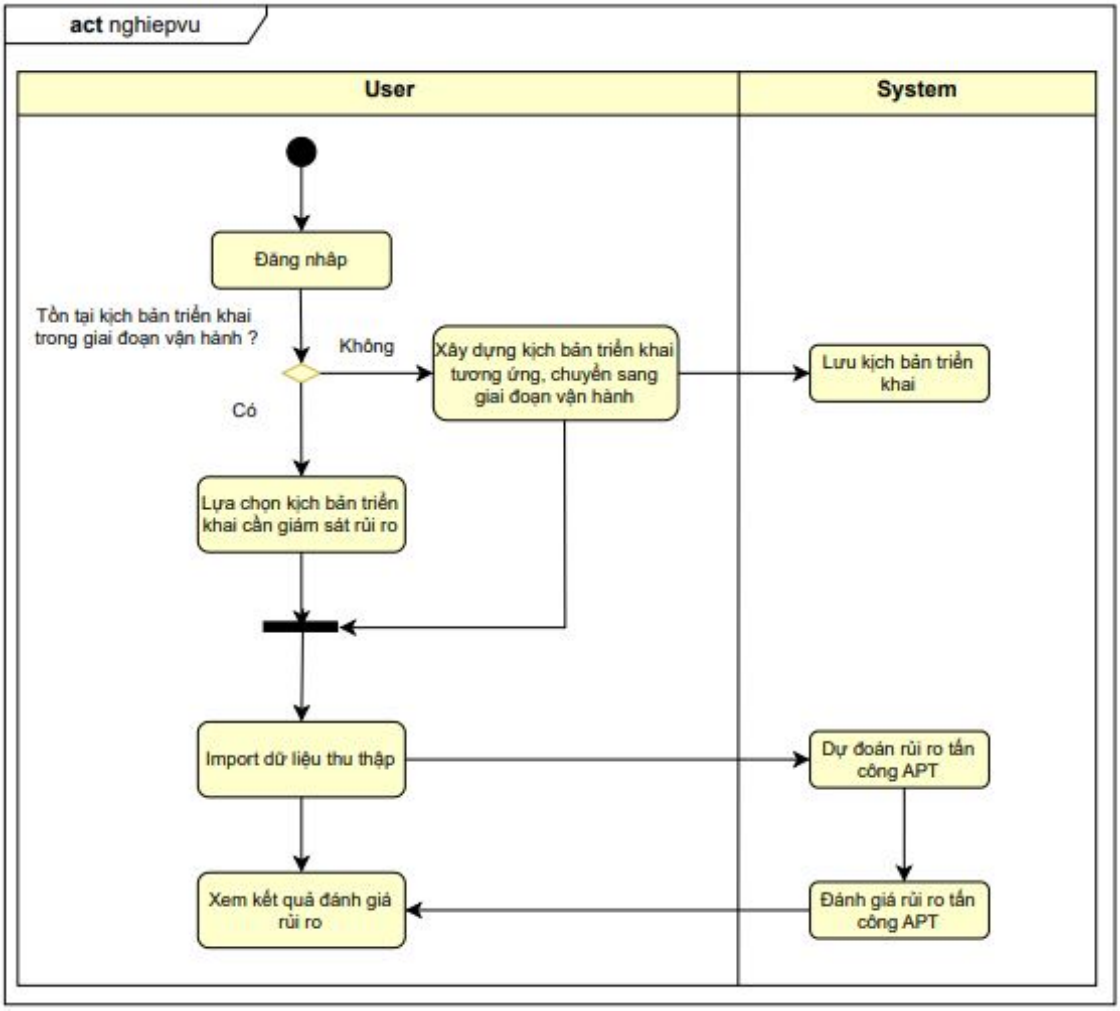
2.2.1 Biểu đồ usecase chức năng giám sát rủi ro tấn công APT



Hình 2.1: Biểu đồ usecase chức năng giám sát rủi ro tấn công APT

Biểu đồ usecase chức năng giám sát rủi ro tấn công APT được biểu diễn tại hình 2.1. Người dùng có thể giám sát rủi ro đối với kịch bản triển khai sau khi được xây dựng và đưa vào vận hành. Dữ liệu giám sát được đẩy lên cần tương ứng với kịch bản triển khai. Core-service sẽ tiến hành phán đoán thông tin về tấn công APT từ dữ liệu đầu vào bao gồm giai đoạn tấn công, năng lực tấn công,... từ đó ánh xạ với mạng đánh giá rủi ro và phản hồi kết quả. Kết quả giám sát được phân tích thành nhiều thành phần bao gồm trạng thái hệ thống, chỉ số severity, chỉ số likelihood, giai đoạn tấn công,...

2.2.2 Quy trình nghiệp vụ chức năng Giám sát rủi ro tấn công APT



Hình 2.2: Quy trình nghiệp vụ chức năng Giám sát rủi ro tấn công APT

Hình 2.2 mô tả quy trình nghiệp vụ của chức năng giám sát rủi ro tấn công APT. Sau khi đăng nhập vào hệ thống, người dùng cần kiểm tra xem có tồn tại kịch bản triển khai trong giai đoạn vận hành, nếu không thì cần xây dựng kịch bản triển khai tương ứng với dữ liệu giám sát đầu vào. Sau các bước đánh giá rủi ro, người dùng chuyển trạng thái kịch bản triển khai sang giai đoạn vận hành. Bước tiếp theo người dùng có thể import dữ liệu giám sát và theo dõi kết quả giám sát rủi ro tấn công APT. Hệ thống sẽ tiến hành đánh giá rủi ro tấn công từ dữ liệu đầu vào và phản hồi kết quả đánh giá cho người dùng.

2.3 Đặc tả chức năng

2.3.1 Đặc tả usecase Giám sát rủi ro tấn công APT

<table>
<tr><td>Tên usecase</td><td colspan="3">Giám sát rủi ro tấn công APT</td></tr>
<tr><td>Mục đích sử dụng</td><td colspan="3">Giám sát rủi ro tấn công APT cho kịch bản triển khai ở giai đoạn vận hành</td></tr>
<tr><td>Tác nhân</td><td colspan="3">User, System</td></tr>
<tr><td>Sự kiện kích hoạt</td><td colspan="3">Nhấn vào nút ’submit’ sau khi import dữ liệu giám sát</td></tr>
<tr><td>Tiền điều kiện</td><td colspan="3">User khởi tạo kịch bản triển khai thành công, chuyển từ giai đoạn triển khai sang giai đoạn vận hành</td></tr>
<tr><td>Hậu điều kiện</td><td colspan="3">Hiển thị kết quả giám sát, trạng thái hệ thống và thông tin về giai đoạn tấn công APT (nếu có)</td></tr>
<tr><td rowspan="6">Luồng sự kiện</td><td>STT</td><td>Thực hiện bởi</td><td>Hành động</td></tr>
<tr><td>1</td><td>User</td><td>Chọn chức năng giám sát rủi ro tấn công APT</td></tr>
<tr><td>2</td><td>User</td><td>Đẩy file dữ liệu thu thập lưu lượng mạng của kịch bản triển khai, ấn nút ’Submit’</td></tr>
<tr><td>3</td><td>System</td><td>Chuẩn hóa dữ liệu đầu vào, sử dụng các mô hình học máy để dự đoán thông tin về cuộc tấn công.</td></tr>
<tr><td>4</td><td>System</td><td>Ánh xạ kết quả thu được với đầu vào mạng đánh giá rủi ro, phản hồi kết quả đánh giá qua mỗi bước giám sát</td></tr>
<tr><td>5</td><td>System</td><td>Hiển thị kết quả giám sát rủi ro, cảnh báo rủi ro xuất hiện tấn công APT (nếu có) cho người dùng</td></tr>
<tr><td>Luồng sự kiện thay thế</td><td colspan="3">Không</td></tr>
</table>

Bảng 2.1: Đặc tả usecase giám sát rủi ro tấn công APT

2.3.2 Đặc tả usecase xây dựng kịch bản triển khai

<table>
<tr><td>Tên usecase</td><td colspan="3">Xây dựng kịch bản triển khai</td></tr>
<tr><td>Mục đích sử dụng</td><td colspan="3">Khởi tạo kịch bản triển khai phục vụ đánh giá rủi ro ATTT</td></tr>
<tr><td>Tác nhân</td><td colspan="3">User, System</td></tr>
<tr><td>Sự kiện kích hoạt</td><td colspan="3">User đăng nhập thành công, import dữ liệu kịch bản triển khai</td></tr>
<tr><td>Tiền điều kiện</td><td colspan="3">User khởi tạo các tài sản và biện pháp phòng vệ cần thiết trên hệ thống</td></tr>
<tr><td>Hậu điều kiện</td><td colspan="3">Không</td></tr>
<tr><td rowspan="6">Luồng sự kiện</td><td>STT</td><td>Thực hiện bởi</td><td>Hành động</td></tr>
<tr><td>1</td><td>User</td><td>User nhấn nút "Download template"</td></tr>
<tr><td>2</td><td>System</td><td>Hiển thị popup "Download template deployment scenario"</td></tr>
<tr><td>3</td><td>User</td><td>Lựa chọn system-profile, nhấn nút "Download"</td></tr>
<tr><td>4</td><td>User</td><td>Chỉnh sửa thông tin tại tệp tin template, nhấn nút "Import" và đẩy tệp tin kịch bản triển khai lên</td></tr>
<tr><td>5</td><td>System</td><td>Hiển thị thông tin về kịch bản triển khai, hiển thị nút để lưu vào CSDL</td></tr>
<tr><td>Luồng sự kiện thay thế</td><td colspan="3">- System hiển thị thông báo định dạng file không hợp lệ
- System hiển thị thông báo dữ liệu file không hợp lệ</td></tr>
</table>

Bảng 2.2: Đặc tả usecase xây dựng kịch bản triển khai

2.4 Yêu cầu phi chức năng

2.4.1 Yêu cầu về hệ thống

Hệ thống RiDX được thiết kế theo kiến trúc microservice. Một số yêu cầu phi chức năng của một hệ thống microservice bao gồm:

- **Thiết kế gateway**: Các service trong hệ thống nên được thiết kế đi qua một cổng trung gian (API-gateway). API-gateway có thể đảm nhận nhiều vai trò khác nhau như bảo mật API, giám sát và quản lý các request từ phía client. Điều này giúp ích rất lớn trong việc kiểm soát tất cả các API của các service

trong hệ thống từ một nơi duy nhất. Phiên bản hiện tại của hệ thống RiDX đã đáp ứng 1 phần yêu cầu này với với hệ thống bao gồm 1 service API-gateway và 8 service giao tiếp qua nó. Trong tương lai hệ thống cần tích hợp 3 service còn lại đang hoạt động độc lập đi qua API-gateway.

- **Cân bằng tải**: Cân bằng tải là phương pháp phân phối lưu lượng mạng một cách đồng đều trên một nhóm tài nguyên hỗ trợ ứng dụng nhằm trả về dữ liệu chính xác cho người dùng một cách nhanh chóng và đáng tin cậy. Một trong những khía cạnh quan trọng khi thiết kế microservice là phân bổ tải trên các nút. Thuật toán cân bằng tải của hệ thống RiDX nên được triển khai trên ít nhất 2 nút để phục vụ lưu lượng truy cập.

2.4.2 Yêu cầu về hiệu năng

Đầu vào của chức năng giám sát rủi ro tấn công APT là dữ liệu thời gian thực. Yêu cầu hiệu suất của chức năng tập trung vào việc đảm bảo độ trễ thấp - tính toán mạng Bayes hợp lý để trả về kết quả giám sát rủi ro thời gian thực. Điều này rất quan trọng trong việc phát hiện cuộc tấn công APT trong quá trình giám sát từ đó đưa ra phương án xử lý kịp thời. Ngoài ra chức năng cần có độ tin cậy cao - có tính sẵn sàng cao và khả năng phục hồi trước các lỗi xảy ra.

2.4.3 Yêu cầu về bảo mật

Chức năng cần đáp ứng yêu cầu chung của toàn bộ hệ thống về an ninh thông tin: (i) cơ chế xác thực người dùng hiệu quả, (ii) dữ liệu truyền đi qua internet cần được mã hóa, (iii) tránh để lộ các thông tin quan trọng của người dùng và hệ thống. Bên cạnh đó dữ liệu đầu vào của chức năng giám sát rủi ro tấn công APT được gửi lên từ người dùng bằng cách import file, vì vậy cần có cơ chế xác thực định dạng file hợp lý, xử lý dữ liệu ngoại lệ tránh gặp phải các trường hợp bị tấn công thông qua mã độc chèn bên trong nội dung file.

CHƯƠNG 3. CƠ SỞ LÝ THUYẾT VÀ CÔNG NGHỆ SỬ DỤNG

3.1 Nền tảng lý thuyết

Trong chương trước đồ án đã trình bày tổng quan về các chức năng chính của hệ thống RiDX hiện tại và giới thiệu về các bộ dữ liệu được sử dụng cho chức năng đánh giá rủi ro tấn công APT. Chương 3 sẽ trình bày về nền tảng lý thuyết và công nghệ được sử dụng của hệ thống.

Hệ thống RiDX được phát triển bởi các sinh viên ĐHBKHN từ những nghiên cứu về đánh giá rủi ro ATTT sử dụng mạng Bayes động của nhóm tác giả Giang V.T.H và Tuan N.M [1][5] [2]. Hệ thống hỗ trợ đánh giá rủi ro ATTT trên kịch bản triển khai của hệ thống web trong ba giai đoạn: phân tích yêu cầu, triển khai và vận hành; giúp người dùng có cái nhìn khách quan về tình trạng bảo mật của hệ thống từ đó lựa chọn kịch bản triển khai hợp lý cho hệ thống của mình. Lý thuyết cốt lõi của hệ thống sẽ được trình bày dưới đây.

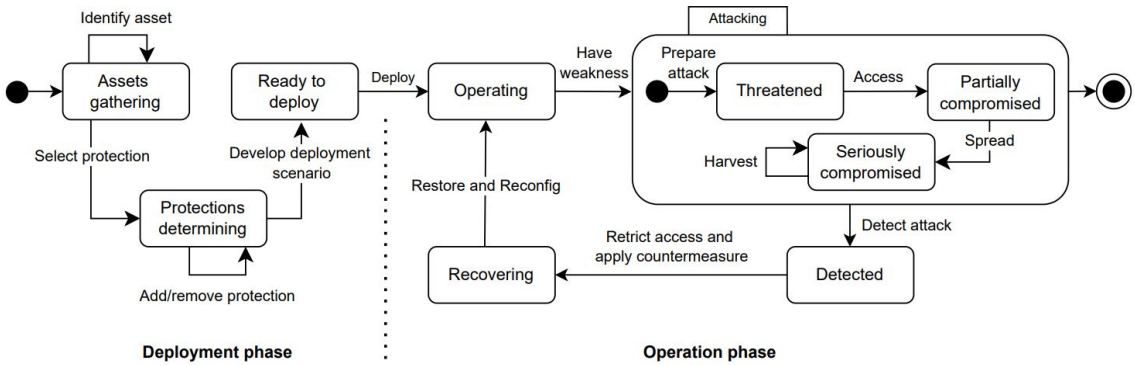
3.1.1 Mô hình kịch bản triển khai và đồ thị tấn công

a, Kịch bản triển khai

Kịch bản triển khai là thành phần cốt lõi của hệ thống RiDX. Mô hình xây dựng kịch bản triển khai đã được giới thiệu trong đồ án của anh Hoàng Trung Hiếu K62 ĐHBKHN. Một kịch bản triển khai bao gồm các thông tin:

- Danh sách tài sản: Tài sản sử dụng trong hệ thống triển khai. Tài sản có thể là phần cứng, phần mềm, hệ điều hành, mã nguồn,...
- Ma trận giao tiếp giữa các tài sản: Thông tin về mối quan hệ giữa các tài sản xác định bởi quyền giao tiếp và vector truy cập.
- Biện pháp phòng vệ: Danh sách biện pháp phòng vệ áp dụng cho kịch bản triển khai.
- Yêu cầu bảo mật: Tập yêu cầu bảo mật đối với hệ thống triển khai.
- Giai đoạn triển khai hệ thống: Giai đoạn triển khai của hệ thống web, gồm 3 giai đoạn Requirements Analysis, Deployments, Operations.

Kịch bản triển khai của hệ thống sẽ bao gồm 3 giai đoạn: giai đoạn phân tích yêu cầu, giai đoạn triển khai và giai đoạn vận hành. Trong đó tại giai đoạn triển khai và giai đoạn vận hành trạng thái hệ thống được mô tả như sau.



Hình 3.1: Biểu đồ trạng thái của hệ thống

Hình 3.1 mô tả các trạng thái của hệ thống tại 2 giai đoạn là triển khai và vận hành. Trong đó tại giai đoạn triển khai bao gồm các trạng thái thiết lập tài sản (Assets-gathering), thiết lập biện pháp phòng vệ (Protections-determining) và sẵn sàng triển khai (Ready-to-deploy). Giai đoạn này cho phép người dùng xây dựng kịch bản triển khai, bao gồm thiết lập tài sản, lựa chọn các biện pháp phòng vệ và tiến hành đánh giá rủi ro. Từ đó người dùng có thể lựa chọn kịch bản hợp lý cho hệ thống để đưa vào vận hành. Giai đoạn vận hành bao gồm các trạng thái: vận hành (Operating), bị tấn công APT (APT attack), phát hiện tấn công (Detected) và phục hồi (Recovering). Giai đoạn bị tấn công APT xuất hiện khi có dấu hiệu khai thác các lỗ hổng bảo mật xuất hiện trên kịch bản triển khai. Hệ thống trong quá trình xảy ra tấn công APT gồm 3 trạng thái tương ứng với 3 mức độ đe dọa tương ứng với các giai đoạn tấn công APT: xâm nhập, lan rộng và khai thác dữ liệu. Sau khi phát hiện có dấu hiệu tấn công, hệ thống chuyển sang trạng thái phát hiện và phục hồi với việc áp dụng các biện pháp phòng vệ. Sau giai đoạn này hệ thống quay trở lại trạng thái vận hành.

b, Ánh xạ tài sản với lỗ hổng bảo mật

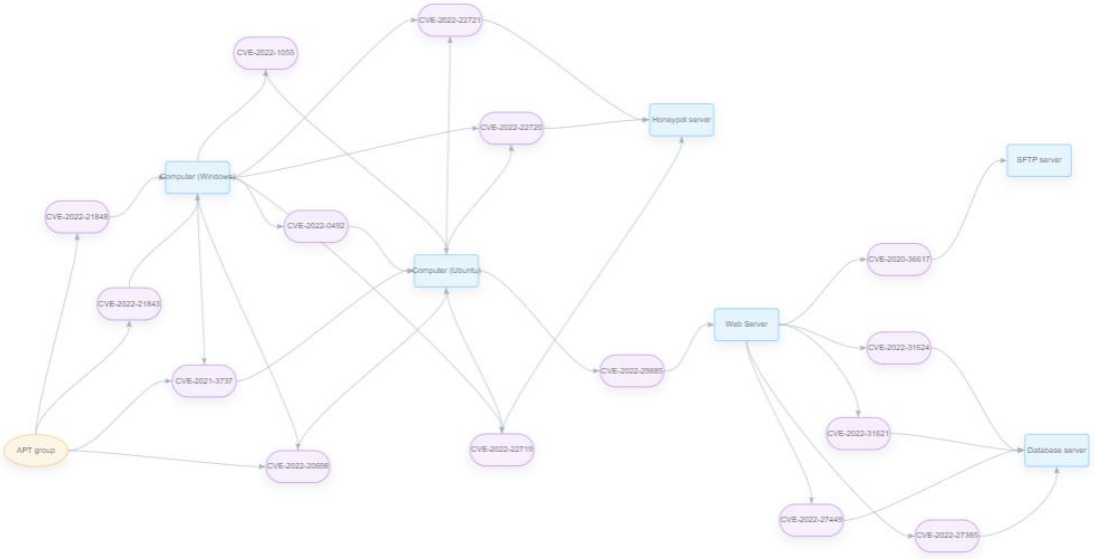
Sau khi xây dựng kịch bản triển khai, tài sản sẽ được ánh xạ với định danh CPE tương ứng. Căn cứ vào thông tin về tên sản phẩm, nhà sản xuất và phiên bản, mỗi tài sản được xác định với một định danh CPE duy nhất. Điều này nhằm cung cấp một phương thức mã hóa tiêu chuẩn các tài sản trong kịch bản triển khai và làm cơ sở để xác định chính xác các lỗ hổng bảo mật tồn tại bên trong tài sản.

Bước tiếp theo căn cứ vào CPE, người dùng có thể xác định các lỗ hổng bảo mật CVE tồn tại bên trong tài sản. Đây là yếu tố quyết định để xây dựng mô hình đánh giá rủi ro sử dụng mạng Bayes. Mỗi tài sản có thể tồn tại 1 hoặc nhiều lỗ hổng bảo mật CVE. Thông tin về lỗ hổng bảo mật bao gồm điểm Exploitability (khả năng khai thác), điều kiện tiên quyết (quyền cần thiết để khai thác), hậu điều kiện (quyền đạt được sau khi khai thác) và attack vector (vector tấn công). Mô hình này phù hợp

với cơ chế tấn công leo thang đặc quyền của các cuộc tấn công APT trong thực tế.

c, Xây dựng đồ thị tấn công

Đồ thị tấn công là mô hình đồ thị biểu diễn thông tin về quan hệ giữa các tài sản và điểm yếu với nhau bên trong kịch bản triển khai. Nó là cơ cở để sinh cấu trúc mạng Bayes động đánh giá rủi ro ATTT. Đồ thị tấn công được xây dựng dựa trên thông tin về ma trận giao tiếp giữa các tài sản và các lỗ hổng bảo mật tồn tại trên mỗi tài sản, được xây dựng với 3 giai đoạn tương ứng với 3 giai đoạn tấn công APT: Preparing, Access, Resident. Đồ thị bao gồm các nút biểu diễn tài sản, CVE và attacker; các cạnh của đồ thị được xác định dựa vào ma trận giao tiếp tài sản. Hình 3.2 mô tả đồ thị tấn công của kịch bản triển khai hệ thống Unraveled xây dựng tại phase 1 - Thăm dò.



Hình 3.2: Đồ thị tấn công của hệ thống Unraveled tại phase 2 - Access

3.1.2 Đánh giá rủi ro sử dụng mạng Bayes động

a, Mạng Bayes và mạng Bayes động

Mạng Bayes là một mô hình xác suất đồ thị được sử dụng để mô hình hóa các phụ thuộc có điều kiện giữa các biến ngẫu nhiên. Sự phụ thuộc giữa các biến được biểu diễn dưới dạng cạnh của đồ thị. Đồ thị của mạng Bayes thể hiện các mối quan hệ xác suất mà chúng ta có thể sử dụng để suy ra và tính toán một cách hiệu quả các biến ngẫu nhiên trong đồ thị. Mạng Bayes được xây dựng dựa trên định lý Bayes được phát triển bởi nhà toán học Thomas Bayes theo công thức:

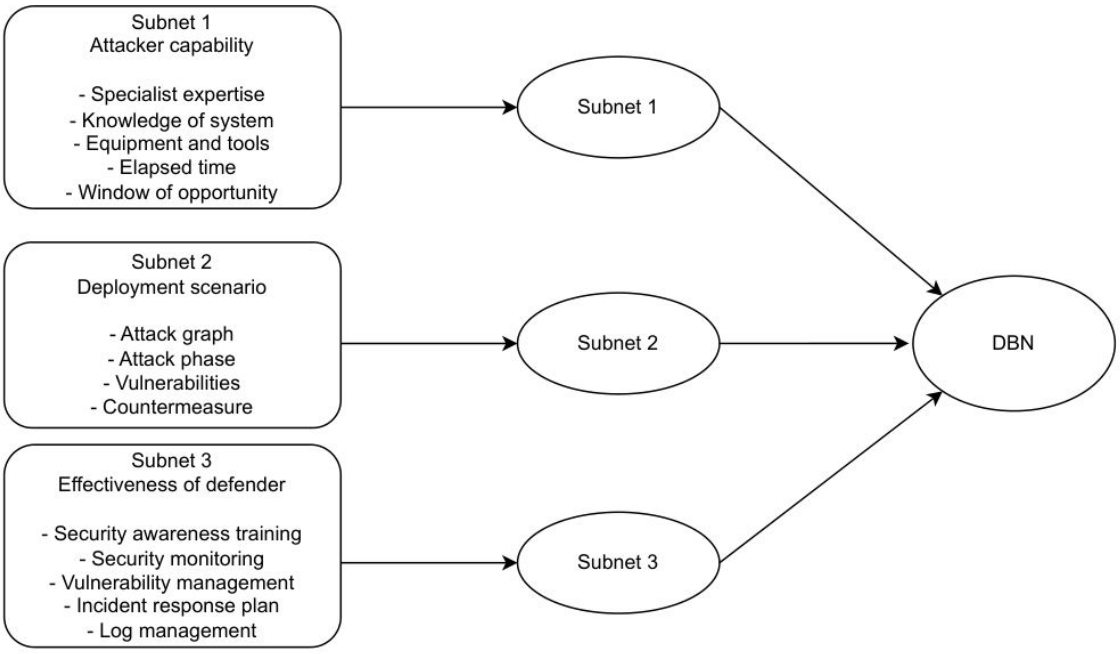
$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Mạng Bayes bao gồm 2 thành phần: cấu trúc mạng và tham số mạng. Cấu trúc mạng biểu diễn quan hệ giữa các biến với một DAG, tham số mạng tương ứng với phân phối xác suất tại mỗi nút của đồ thị. Mạng Bayes sẽ được cập nhật khi có những thay đổi trong các thành phần của CPT giữa các nút hoặc khi mối quan hệ nhân quả giữa các nút bị thay đổi.

Mạng Bayes là một mô hình tĩnh, chỉ phù hợp với những cơ chế ổn định, cân bằng như đánh giá rủi ro ATTT trước khi hệ thống đưa vào triển khai. Đối với những hệ thống triển khai trong thực tế, luôn có sự tác động giữa trạng thái lỗ hổng trong quá khứ và hiện tại, cần có một cơ chế động như mô hình mạng Bayes động. Mạng Bayes động [6] giữ nguyên cấu trúc và số lượng tham số so với mạng Bayes, tuy nhiên bổ sung các cơ chế có thể mô hình hóa ảnh hưởng thời gian. Mạng Bayes động được sử dụng để đánh giá và giám sát rủi ro ATTT của hệ thống RiDX.

b, Đầu vào mạng đánh giá rủi ro Bayes

Mô hình đánh giá rủi ro áp dụng mạng Bayes động ban đầu được phát triển với đầu vào là kịch bản triển khai. Điểm yếu của mô hình này là được thiết kế để đánh giá rủi ro liên quan đến các cuộc tấn công STRIDE thông thường, thiếu khả năng xử lý các đặc điểm rủi ro của tấn công APT. Để phù hợp với kịch bản tấn công APT liên quan đến tấn công leo thang đặc quyền hay chuyển động ngang, anh Nguyễn Việt Thắng K63 ĐHBKHN đã bổ sung thêm 2 subnet là khả năng tấn công và khả năng phòng thủ cho đầu vào mạng đánh giá rủi ro. Đây là sơ đồ thể hiện đầy đủ đầu vào của mạng rủi ro Bayes với 3 subnet.



Hình 3.3: Đầu vào mạng đánh giá rủi ro

Hình 3.3 mô tả đầu vào của mạng đánh giá rủi ro Bayes động với 3 subnet là kịch bản triển khai, năng lực tấn công và khả năng phòng thủ. Subnet 1 tương ứng với khả năng tấn công của attacker bao gồm các thành phần: Kỹ năng chuyên môn (Specialist expertise), kiến thức về hệ thống (Knowledge of system), công cụ cần thiết (Equipment and tools), thời gian (Elapsed time) và cơ hội (Window of opportunity). Subnet 3 tương ứng với khả năng phòng thủ của hệ thống, bao gồm các thành phần: nhận thức về an ninh (Security awareness training), giám sát an ninh (Security monitoring), quản lý lỗ hổng (Vulnerability management), kế hoạch ứng phó sự cố (Incident response plan) và quản lý nhật ký (Log management). Subnet 2 là kịch bản triển khai do người dùng xây dựng, gồm các thành phần chính là đồ thị tấn công, biện pháp phòng thủ, các lỗ hổng bảo mật,... Trong quá trình giám sát rủi ro, subnet 2 của mạng đánh giá rủi ro bổ sung thông tin về giai đoạn tấn công, mức độ rủi ro của hệ thống càng nghiêm trọng khi cuộc tấn công APT ngày càng phát triển.

Các thông tin đầu vào của subnet 1 và subnet 3 được xác định từ các biến/các độ đo như Security-Control-Coverage (phạm vi kiểm soát an ninh), Frequency-of-Attacks (tần suất tấn công), Lateral-Movement-Rate (tốc độ chuyển động ngang)... Tuy nhiên hệ thống RiDX chưa có cơ chế tích hợp tự động cập nhật các độ đo trên trong một kịch bản triển khai cụ thể, vì vậy hiện tại các subnet 1 và subnet 3 được người dùng cấu hình thủ công.

3.2 Phân loại đa lớp

3.2.1 Giới thiệu

Trong lĩnh vực học máy (machine learning), phân loại đa lớp hay phân loại đa thức là một bài toán phân loại với từ 3 lớp trở lên. Mỗi thực thể được gán cho một lớp mà không có bất kỳ sự chồng chéo nào. Ví dụ: sử dụng mô hình để xác định các loại động vật trong hình ảnh từ bách khoa toàn thư. Trong quá trình huấn luyện, mô hình sẽ tìm hiểu các mẫu cụ thể cho từng lớp và sử dụng các mẫu đó để dự đoán tư cách thành viên của dữ liệu trong tương lai. Bên cạnh hồi quy, phân loại đa lớp là một trong những tác vụ học máy phổ biến nhất.

3.2.2 Các chiến lược phân loại đa lớp

Phân loại đa lớp được chia thành 3 chiến lược cụ thể như sau.

a, Đưa về phân loại nhị phân

Đưa về phân loại nhị phân là chiến lược nhằm giảm bớt vấn đề phân loại đa lớp bằng cách biến đổi thành nhiều bài toán phân loại nhị phân. Các phương pháp được chia thành hai loại one-vs-rest (OvR) và one-vs-one (OvO).

One-vs-rest là phương pháp heuristic đào tạo một mô hình phân loại duy nhất cho mỗi lớp với các mẫu của lớp đó là dương tính mẫu và tất cả các mẫu khác là âm tính. Kỹ thuật này yêu cầu các bộ phân loại cơ sở tạo ra điểm có giá trị thực cho quyết định của nó thay vì chỉ nhãn lớp; chỉ riêng nhãn lớp riêng biệt có thể dẫn đến sự mơ hồ, trong đó nhiều lớp được dự đoán cho một mẫu. OvR là phương pháp rất phổ biến, tuy nhiên nó gặp phải một số vấn đề. Thứ nhất, thang đo của các giá trị độ tin cậy có thể khác nhau giữa các phân loại nhị phân. Thứ hai, dữ liệu trong thực tế hầu hết mất cân bằng giữa 2 lớp.

One-vs-one là một phương pháp heuristic khác để sử dụng thuật toán phân loại nhị phân để phân loại nhiều lớp. Giống như OvR, OvO chia tập dữ liệu phân loại nhiều lớp thành các vấn đề phân loại nhị phân. Khác với kĩ thuật OvR chia tập dữ liệu thành một tập dữ liệu nhị phân cho mỗi lớp, phương pháp OvO chia tập dữ liệu thành một tập dữ liệu cho mỗi lớp so với các lớp khác. OvO có sự mơ hồ ở chỗ một số vùng trong không gian đầu vào của nó có thể nhận được cùng số phiếu bầu.

b, Mở rộng phân loại nhị phân

Để giải quyết các vấn đề gặp phải khi đưa bài toán phân loại đa lớp về bài toán phân loại nhị phân, rất nhiều phương pháp được đưa ra mở rộng từ phân loại nhị phân. Một số thuật toán đã được phát triển dựa trên neural-network, desicion-tree, k-nearest-neighbor, naive-bayes, SVM để giải quyết các vấn đề phân loại nhiều lớp. Những loại kỹ thuật này còn có thể được gọi là kỹ thuật thích ứng thuật toán.

c, Phân loại theo thứ bậc

Phân loại theo cấp bậc giải quyết vấn đề phân loại nhiều lớp bằng cách chia không gian đầu ra thành một cây. Mỗi nút cha được chia thành nhiều nút con và quá trình này được tiếp tục cho đến khi mỗi nút con chỉ đại diện cho một lớp. Một số phương pháp đã được đề xuất dựa trên phân loại theo thứ bậc. Ưu điểm của chiến lược phân loại theo thứ bậc là giúp phân loại hiệu quả hơn bằng cách giảm số lượng lớp cần xem xét tại mỗi bước. Tuy nhiên chiến lược này yêu cầu kiến thức về cấu trúc phân cấp của các lớp.

3.2.3 Một số mô hình phân loại đa lớp tiêu biểu

Các mô hình Random Forest và XGBoost được sử dụng để xây dựng mô hình học máy dự đoán các thông tin về giai đoạn tấn công APT, năng lực tấn công của attacker từ dữ liệu đầu vào cho chức năng giám sát rủi ro ATTT của hệ thống RiDX.

a, Random forest

Random Forest là thuật toán sử dụng trong việc xây dựng mô hình dự đoán và phân loại, là một phương pháp học có giám sát dựa trên việc kết hợp nhiều cây

quyết định (decision trees) trong quá trình huấn luyện và dự đoán.

Ý tưởng cơ bản của Random Forest là xây dựng một tập hợp (ensemble) của cây quyết định độc lập, mỗi cây được huấn luyện trên một tập con dữ liệu ngẫu nhiên và một tập con ngẫu nhiên của các thuộc tính. Các cây quyết định trong Rừng ngẫu nhiên hoạt động đồng thuận (voting) để đưa ra dự đoán cuối cùng.

Các bước của thuật toán Random Forest bao gồm:

1. Lấy mẫu ngẫu nhiên: Từ tập dữ liệu huấn luyện, lấy ngẫu nhiên một số mẫu để tạo thành tập dữ liệu con cho mỗi cây quyết định.
2. Lấy thuộc tính ngẫu nhiên: Chọn ngẫu nhiên một số thuộc tính từ tập thuộc tính để tạo thành tập thuộc tính con cho mỗi cây quyết định.
3. Xây dựng cây quyết định: Đối với mỗi cây, xây dựng cây quyết định bằng cách sử dụng thuật toán như ID3 hoặc CART trên tập dữ liệu con và tập thuộc tính con tương ứng.
4. Dự đoán: Khi cây quyết định đã được xây dựng, các dự đoán cuối cùng được đưa ra bằng cách lấy phương án đa số của các cây trong rừng.

b, XGBoost

XGBoost (Extreme Gradient Boosting) là giải thuật tăng cường độ dốc phân tán được tối ưu hóa được thiết kế để đào tạo các mô hình học máy hiệu quả và có thể mở rộng. Đây là một phương pháp học tập tổng hợp kết hợp các dự đoán của nhiều mô hình yếu để tạo ra dự đoán mạnh hơn. XGBoost nổi tiếng nhờ hiệu quả tính toán, cung cấp khả năng xử lý hiệu quả, phân tích tầm quan trọng của tính năng sâu sắc và xử lý liền mạch các giá trị bị thiếu. Đây là thuật toán phù hợp cho nhiều nhiệm vụ, bao gồm hồi quy, phân loại và xếp hạng.

3.3 Cơ chế giám sát rủi ro tấn công APT đối với dữ liệu thời gian thực

3.3.1 Giám sát rủi ro ATTT thời gian thực

Giám sát thời gian thực (Real-time monitoring) là việc cung cấp dữ liệu được cập nhật liên tục về hệ thống, quy trình hoặc sự kiện. Việc giám sát như vậy cung cấp luồng thông tin ở độ trễ bằng 0 hoặc thấp, do đó có độ trễ tối thiểu giữa việc thu thập và phân tích dữ liệu. Nó cho phép phát hiện nhanh các điểm bất thường, các vấn đề về hiệu suất và các sự kiện quan trọng. Giám sát thời gian thực là một loại giám sát CNTT trong đó dữ liệu được thu thập từ phần cứng, mạng, hệ thống bảo mật, môi trường ảo hóa và ngăn xếp ứng dụng tại chỗ - bao gồm cả dữ liệu trên đám mây - và hiển thị kết quả tại giao diện người dùng phần mềm. Từ dữ liệu này, nhân viên CNTT sẽ phân tích hiệu suất hệ thống, gắn cờ các điểm bất thường và giải quyết vấn đề. Đây là một phương pháp quan trọng để duy trì an ninh mạng và

đảm bảo trải nghiệm người dùng cuối cũng như hiệu suất mạng tốt.

Giám sát rủi ro ATTT là quá trình theo dõi, đánh giá và quản lý các nguy cơ có thể ảnh hưởng đến tính bảo mật, tính toàn vẹn và tính sẵn sàng của thông tin trong một tổ chức. Việc giám sát rủi ro ATTT thời gian thực được thực hiện qua việc thu thập dữ liệu từ các nguồn như mức độ sử dụng CPU và bộ nhớ, thời gian phản hồi của ứng dụng, độ trễ mạng, phân tích luồng lưu lượng mạng,... từ đó có thể xác định các mối đe dọa tiềm ẩn và lỗ hổng bảo mật có thể ảnh hưởng đến thông tin qua đó đưa ra các biện pháp phòng vệ hoặc loại bỏ các rủi ro kịp thời.

3.3.2 Giải pháp tổng thể

Đối với hệ thống đánh giá rủi ro RiDX, chức năng giám sát rủi ro tấn công APT thời gian thực được tích hợp với dữ liệu thực tế về các cuộc tấn công APT (Unraveled, SCVIC-APT-2021). Kết quả giám sát rủi ro bao gồm các thông tin được mô tả tại bảng 3.1.

Thông tin giám sát	**Mô tả**
Severity	Mức độ nghiêm trọng của hệ thống
Likelihood	Xác suất xảy ra của rủi ro của hệ thống
Risk level	Mức độ rủi ro, tổng hợp từ 2 chỉ số Severity và Likelihood
Asset severity	Mức độ nghiêm trọng trên từng tài sản
Attack capability & effectiveness of defender	Chỉ số đánh giá mức độ tấn công (subnet1) và khả năng phòng thủ (subnet 3)
State machinery	Biểu diễn trạng thái của hệ thống qua mỗi timestep trong quá trình giám sát rủi ro tấn công APT

Bảng 3.1: Thông tin giám sát rủi ro

Các giải pháp tổng thể được đưa ra như sau:

1. Xây dựng các kịch bản triển khai tương ứng với các bộ dữ liệu Unraveled và SCVIC-APT-2021.
2. Xây dựng cơ chế giám sát dữ liệu thời gian thực từ tập dữ liệu đầu vào, mỗi dòng thống kê lưu lượng mạng tương ứng với một lần hệ thống giám sát dữ liệu thu thập.
3. Xây dựng mô hình học máy từ các tập dữ liệu có gán nhãn Unraveled và SCVIC-APT-2021 nhằm phán đoán giai đoạn tấn công và khả năng phát hiện tấn công.

4. Sử dụng các mô hình học máy phán đoán giai đoạn tấn công APT từ dữ liệu đầu vào, ánh xạ kết quả với đầu vào mạng đánh giá rủi ro.
5. Xây dựng màn hình hiển thị kết quả giám sát rủi ro thời gian thực.

Các giải pháp và kết quả sẽ được trình bày chi tiết tại chương 5.

3.4 Công nghệ sử dụng

3.4.1 Frontend

a, ReactJS

Hệ thống RiDX sử dụng framework ReactJS cho phần giao diện người dùng. ReactJS [7] là một thư viện JavaScript mã nguồn mở nhằm hỗ trợ thiết kế giao diện web nhanh và hiệu quả hơn. Được phát triển bởi Facebook và ra mắt vào năm 2013, nó nhanh chóng trở thành một trong những thư viện phổ biến nhất trong việc phát triển giao diện người dùng. ReactJS được xây dựng theo kiến trúc dựa trên thành phần, cho phép người dùng phân chia giao diện web thành các thành phần nhỏ hơn và có thể tái sử dụng. Điều này giúp cho nhà phát triển có thể xây dựng những ứng dụng phức tạp và bảo trì theo thời gian. Với ReactJS, dữ liệu không chỉ được kết xuất ở tầng Server mà còn ở tầng Client. Khi có trạng thái thay đổi, ReactJS chỉ cập nhật những phần tử cần thiết thông qua Virtual DOM (DOM ảo) giúp cải thiện hiệu suất và trải nghiệm người sử dụng.

3.4.2 Backend

a, NodeJS

NodeJS [8] là một môi trường chạy JavaScript đa nền tảng và mã nguồn mở. Được phát hành vào năm 2009, NodeJS cho phép lập trình viên xây dựng các ứng dụng web hiệu năng cao và có thể mở rộng. NodeJS có thể hoạt động trên nhiều nền tảng hệ điều hành khác nhau, từ Windows, Linux đến OS X. Bên cạnh đó NodeJS cung cấp một thư viện phong phú dưới dạng các module Javascript khác nhau giúp đơn giản hóa việc lập trình và giảm thiểu thời gian cần thiết.

b, NestJS

Kiến trúc Microservice của hệ thống RiDX được xây dựng với framework NestJS. NestJS [9] là một NodeJS framework mã nguồn mở mới mẻ dùng để phát triển ứng dụng phía máy chủ bằng ngôn ngữ lập trình TypeScript trên cơ sở của NodeJS. NestJS được xây dựng dựa trên nguyên tắc của Angular [10] nhằm tạo ra một môi trường phát triển hiện đại và mạnh mẽ cho việc xây dựng các ứng dụng web và API. Ngoài ra, NestJS cũng cung cấp những tính năng thiết kế đặc biệt cho việc xây dựng ứng dụng Microservices và các ứng dụng Module. NestJS còn hỗ trợ các cơ sở dữ liệu như: MongoDB, MySQL, PostgreSQL.

Mục tiêu của NestJS là cung cấp một cấu trúc ứng dụng rõ ràng, dễ quản lý, giúp tăng tính bảo trì và sự tổ chức trong mã nguồn. Để đạt được điều này, NestJS triển khai mô hình kiến trúc lõi (core architecture) dựa trên các nguyên tắc của Angular, đặc biệt là sử dụng Dependency Injection (DI) và Modules (Các module). Đối với hệ thống các hệ thống sử dụng kiến trúc microservice như RiDX, NestJS hỗ trợ đầy đủ các thành phần như kết nối: RabbitMQ, gRPC, Kafka,... khiến cho việc lập trình trở nên đơn giản và có thể dễ dàng bảo trì.

c, FastAPI

Bên cạnh NestJS, RiDX còn sử dụng FastAPI cho module đánh giá rủi ro với mạng Bayes. FastAPI [11] là một web framework hiện đại, hiệu năng cao để xây dựng APIs với ngôn ngữ lập trình Python [12]. FastAPI được thiết kế với mục đích dễ dàng sử dụng, giúp quá trình phát triển sản phẩm nhanh chóng. FastAPI sử dụng phiên bản mới nhất của Python như cấu trúc async/await giúp lập trình viên xây dựng những ứng dụng mạnh mẽ và dễ dàng bảo trì. Ngoài ra FastAPI còn hỗ trợ xây dựng tài liệu API tự động và xác thực hỗ trợ nhà phát triển quản lý API thuận tiện.

Một trong những tính năng chính của FastAPI là hiệu suất cao. Nó được xây dựng dựa trên framework Starlette cung cấp máy chủ ASGI hiệu suất cao và hệ thống định tuyến. FastAPI cũng sử dụng thư viện Pydantic để xác thực dữ liệu nhanh chóng và hiệu quả. Ngoài ra, FastAPI có tích hợp nhiều chức năng trong việc định nghĩa type như cơ chế bảo mật của API, rule validate schema phức tạp của request, response,... Những điều này khiến cho FastAPI trở thành một framework rất được ưa chuộng ở thời điểm hiện tại.

3.4.3 MongoDB Database

Dữ liệu của hệ thống RiDX được lưu trữ bởi MongoDB [13]. MongoDB là hệ quản trị cơ sở dữ liệu NoSQL mã nguồn mở phổ biến hướng tài liệu (document). Nó được thiết kế để lưu trữ dữ liệu không cấu trúc dưới dạng JSON thay vì dạng bảng như CSDL truyền thống, điều này giúp cho việc truy vấn dữ liệu trở nên nhanh chóng. Cấu trúc của MongoDB rất linh hoạt cho phép dữ liệu lưu trữ không cần tuân theo một cấu trúc nhất định và không có sự ràng buộc lẫn nhau.

MongoDB hỗ trợ đa dạng các loại dữ liệu như số, chuỗi ký tự, thời gian, mảng và tài liệu. Nó còn cung cấp cơ chế tự động đánh chỉ mục nhằm giúp cho tốc độ truy vấn thông tin đạt hiệu suất cao. MongoDB còn có ưu điểm là rất dễ mở rộng, xử lý được lượng lớn dữ liệu. MongoDB hỗ trợ hầu hết các framework phổ biến như NestJS, Springboot và FastAPI.

3.4.4 Thư viện Sckit-learn

Scikit-learn là thư viện mạnh mẽ cung cấp các thuật toán học máy được viết trên ngôn ngữ Python. Thư viện tích hợp rất nhiều thuật toán xử lý các bài toán học máy và thống kê gồm: phân loại (classification), hồi quy (regression), phân cụm (clustering) và giảm chiều dữ liệu (dimensionality-reduction). Scikit-learn còn cung cấp các công cụ hỗ trợ trực quan hóa mô hình học máy, tiền xử lý, điều chỉnh, lựa chọn và đánh giá mô hình. Thư viện được biết đến với tính dễ phát triển tương đối nhờ các API được thiết kế nhất quán và hiệu quả, tài liệu mở rộng cho hầu hết các thuật toán và cộng đồng hỗ trợ đông đảo.

Scikit-learn được viết chủ yếu bằng Python và sử dụng NumPy cho đại số tuyến tính hiệu suất cao cũng như cho các phép tính array. Một số thuật toán Scikit-learn cốt lõi được viết bằng Cython để tăng hiệu suất tổng thể. Là một thư viện cấp cao bao gồm một số triển khai các thuật toán học máy khác nhau, Scikit-learn cho phép người dùng xây dựng, đào tạo và đánh giá mô hình bằng một vài dòng mã. Scikit-learn cung cấp một bộ API cấp cao thống nhất để xây dựng quy trình hoặc quy trình học máy.

Đối với hệ thống RiDX sử dụng module học máy trong việc phán đoán rủi ro tấn công APT, Scikit-learn là một thư viện thích hợp khi có thể dễ dàng tích hợp với các hệ thống web phổ biến hiện nay như FastAPI. Việc sử dụng Scikit-learn đem lại hiệu quả cao khi cần giải quyết các bài toán đơn giản sử dụng học máy cho hệ thống CNTT.

3.5 Hệ thống RiDX

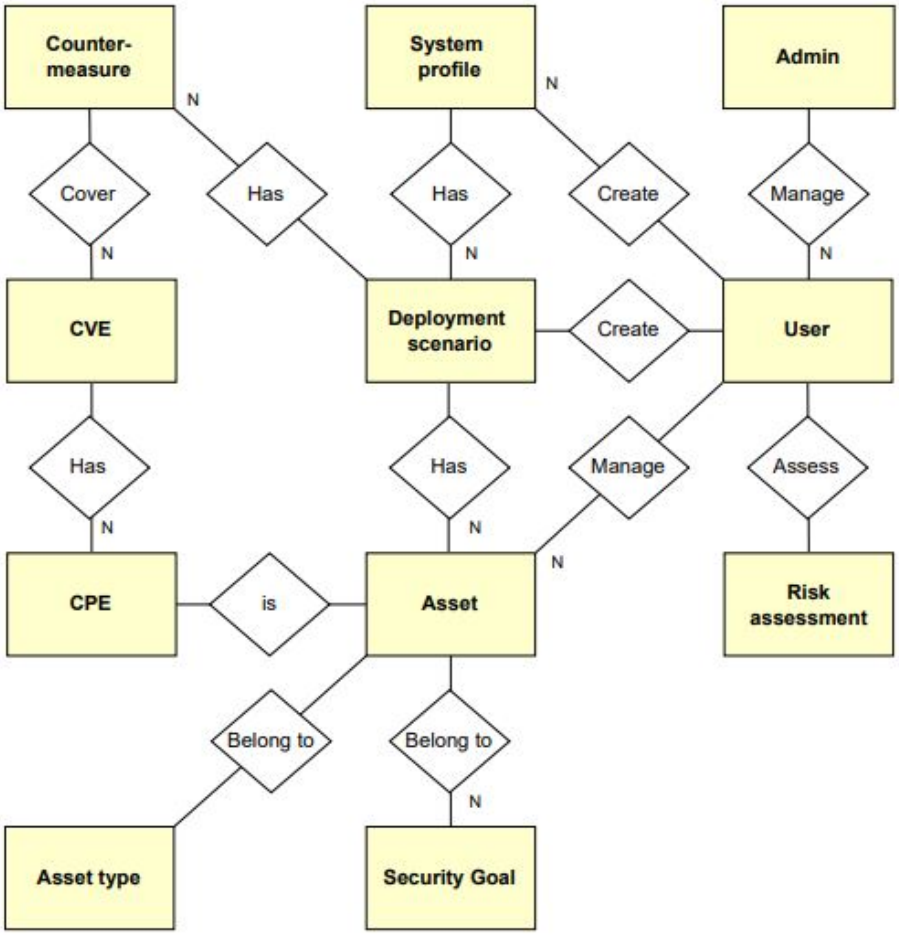
Hệ thống RiDX được phát triển bởi các sinh viên ĐHBKHN từ nghiên cứu về đánh giá rủi ro ATTT cho kịch bản triển khai sử dụng mạng Bayes động của nhóm tác giả Giang V.T.H và Tuan N.M [1][5]. Dưới đây là các chức năng chính của hệ thống:

- Quản lý tài sản: cho phép thêm, sửa, xóa tài sản - thành phần cốt lõi của kịch bản triển khai.
- Xây dựng kịch bản triển khai: bao gồm thông tin về danh sách tài sản, mối quan hệ giữa các tài sản, biện pháp phòng vệ và mục tiêu bảo mật.
- Ánh xạ tài sản với CPE, CVE tương ứng: xác định định danh và các lỗ hổng bảo mật của tài sản từ thông tin về tên, nhà sản xuất, phiên bản.
- Xác định nguy cơ tấn công: attacker, tài sản nhắm đến, quyền truy cập và vector truy cập.
- Xây dựng đồ thị tấn công: đồ thị tấn công được xây dựng từ mối quan hệ giữa

tài sản với nhau và với các lỗ hổng bảo mật, là cơ sở để xây dựng mạng Bayes đánh giá rủi ro.

- Áp dụng biện pháp phòng vệ lên lỗ hổng thuộc kịch bản triển khai: ngăn chặn cuộc tấn công lan rộng.
- Cấu hình tấn công APT: cấu hình subnet 1 và subnet 3 các chỉ số về năng lực tấn công và khả năng phòng thủ.
- Đánh giá rủi ro ATTT: đánh giá rủi ro ATTT của kịch bản triển khai trước khi đưa vào vận hành, gồm các thông tin severity và likelihood.
- Giám sát rủi ro tấn công APT: giám sát, cảnh báo dấu hiệu về cuộc tấn công APT trong quá trình đưa vào vận hành kịch bản triển khai.

Hình vẽ 3.4 dưới đây là biểu đồ thực thể-quan hệ của hệ thống RiDX. Biểu đồ gồm 11 thực thể: (i) Admin, (ii) User, (iii) System-profile, (iv) Deployment-scenario, (v) Risk-assessment, (vi) Asset, (vii) Security-goal, (viii) Counter-measure, (ix) CPE, (x) CVE, (xi) Asset-type.



Hình 3.4: Biểu đồ thực thể-quan hệ của hệ thống RiDX

Kiến trúc tổng quan các thành phần của hệ thống được trình bày tại chương 4.1.

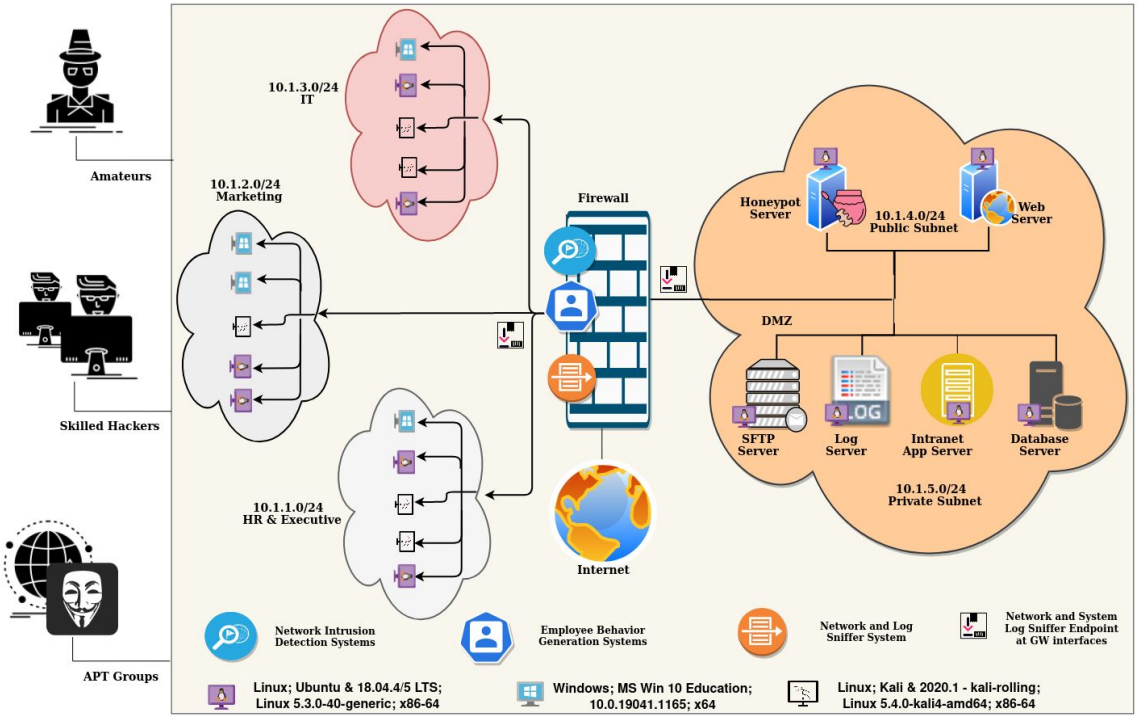
3.6 Dữ liệu sử dụng

3.6.1 Bộ dữ liệu Unraveled

a, Giới thiệu

Unraveled [4] là tập dữ liệu bán tổng hợp nhằm phát hiện các cuộc tấn công APT. Tập dữ liệu được tạo ra bằng cách mô phỏng hành vi bình thường của nhân viên trong 1 tổ chức nhất định, nhằm mô phỏng cuộc tấn công APT xảy ra sát với thực tế nhất.

b, Kiến trúc hệ thống



Hình 3.5: Kiến trúc hệ thống Unraveled

Hình 3.5 mô tả kiến trúc hệ thống được sử dụng trong tập dữ liệu Unraveled. Kiến trúc hệ thống chia thành 2 phần: mạng công ty và mạng sản xuất ngăn cách bởi tường lửa. Mạng công ty bao gồm 3 mạng con mô phỏng 3 phòng ban quản lý, mỗi phòng ban có 5 nhân viên. Mỗi nhân viên tương ứng với 1 Asset là máy tính cài đặt hệ điều hành Window, Ubuntu hoặc Kali. Mạng sản xuất bao gồm 2 mạng con: mạng con công cộng và mạng con riêng tư. Mạng con công cộng bao gồm một máy chủ web công cộng và 1 honeypot. Mạng con riêng tư bao gồm các máy chủ: database server, log server, SFTP server, Intranet App server. Tất cả lưu lượng truy cập từ mạng công cộng đến mạng riêng tư đều bị chặn. Ngược lại, bất kỳ máy chủ nào ở mạng riêng tư có thể tiếp cận mạng công cộng.

c, Dữ liệu

Dữ liệu được thống kê tương đương 6,8 triệu lưu lượng mạng trong 6 tuần. Với mỗi ngày trong tuần, dữ liệu lưu lượng được thống kê tại 5 mạng con riêng biệt. Từ tuần thứ 2 đến tuần thứ 6 ngoài 5 mạng con dữ liệu còn bổ sung thống kê lưu lượng mạng trên gateway. Dữ liệu thu thập bao gồm các file định dạng CSV chứa thông tin pcap về các luồng lưu lượng mạng. Thống kê lưu lượng mạng trên từng bộ phận được mô tả tại bảng 3.2 dưới đây.

Traffic through at	Number of Flows	Benign Flows	Attack Flows
HR & Executive	1,792,784	1,792,756 (99.9%)	28 (0.1%)
Marketing	734360	734,331 (99.9%)	29 (0.1%)
IT	1,072,398	1,038,246 (96.8%)	34,152 (3.2%)
Production-public	210,401	207,423 (98.5%)	2,978 (1.41%)
Production-private	175,692	174,506 (99.32%)	1,186 (.61%)
Gateway	2,891,695	2,832,652 (97.95%)	59,043 (2.04%)

Bảng 3.2: Thống kê thu thập dữ liệu lưu lượng mạng

Tập dữ liệu Unraveled cung cấp 4 loại nhãn được mô tả dưới đây. Đây là cơ sở để đánh giá mức độ hiệu quả của việc ánh xạ dữ liệu đầu vào với mạng đánh giá rủi ro RiDX.

- **Activity**: Các kỹ thuật sử dụng bởi acttacker theo từng giai đoạn tấn công dựa vào bộ dữ liệu MITRE [14].
- **Stage**: Các giai đoạn tấn công: Benign (bình thường), Reconnaissance (giai đoạn trinh sát), Establish-Foothold (thiết lập chỗ đứng), Lateral-Movement (chuyển động ngang), Data-Exfiltration (Khai thác dữ liệu) và Cover-Up (Che đậy).
- **Defender response**: Phản ứng phòng vệ: Benign (không phát hiện rủi ro), Detected (phát hiện rủi ro)
- **Signature**: Phân loại tổ chức tấn công: AA (Amatuer-hacker), SH (Skilled-hacker) và APT (tổ chức APT).

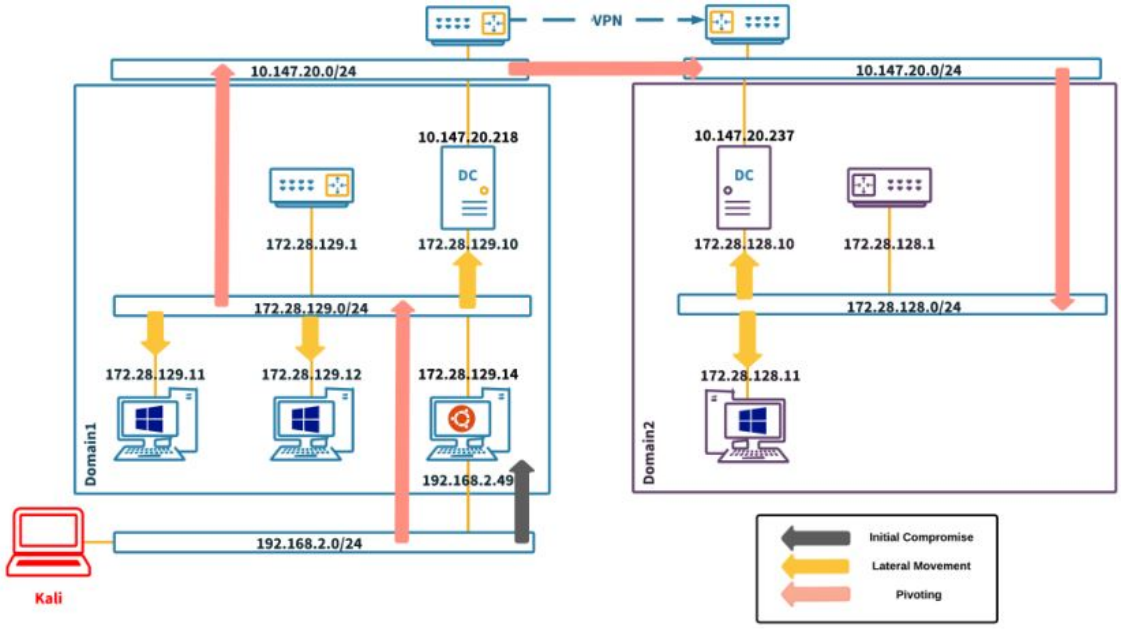
3.6.2 Bộ dữ liệu SCVIC-APT-2021

a, Giới thiệu

SCVIC-APT-2021 là bộ dữ liệu do nhóm tác giả [3] tạo ra nhằm xác định một điểm chuẩn cho việc phát hiện mối đe dọa tấn công APT dựa trên ML trong lưu

lượng truy cập mạng. Tập dữ liệu thể hiện một cách thực tế kiến trúc mạng hiện đại và các đặc điểm của cuộc tấn công APT.

b, Kiến trúc hệ thống



Hình 3.6: Kiến trúc hệ thống SCVIC-APT-2021

Hình 3.6 mô tả kiến trúc hệ thống được sử dụng trong tập dữ liệu SCVIC-APT-2021. Kiến trúc được thiết kế với 4 mạng con chia thành 2 miền để kích hoạt các cuộc tấn công APT như chuyển động ngang và xoay vòng. Miền thứ nhất bao gồm 4 thiết bị: 1 thiết bị là bộ điều khiển miền (DC) và 3 PC thông thường. Miền thứ hai bao gồm 1 bộ điều khiển miền (DC) và 1 PC thông thường - mục tiêu cuối cùng của kẻ tấn công. Hai miền được kết nối thông qua VPN. Cuộc tấn công được thiết lập bởi máy chủ giả lập hệ điều hành Kali nhắm vào miền thứ nhất.

Quy trình tấn công APT hoàn chỉnh của bộ dữ liệu SCVIC-APT-2021 được mô tả như sau. Thông qua bước thỏa hiệp ban đầu kẻ tấn công xâm nhập vào máy đầu tiên và xoay vòng sang miền thứ nhất. Ngoài ra kẻ tấn công chuyển động ngang bằng nhiều cách tiếp cận khác nhau và thu thập tất cả các thông tin sẵn có. Kẻ tấn công chuyển sang miền thứ hai thông qua DC trong miền thứ nhất bằng VPN. Sau khi chuyển động ngang trong miền thứ hai, kẻ tấn công xâm phạm và lấy cắp dữ liệu từ PC cuối cùng.

Các kỹ thuật tấn công được sử dụng trong bộ dữ liệu SCVIC-APT-2021 được mô tả tại bảng 3.3:

Attack Stage	**Reconnaisance**	**Initial Compromise**	**Lateral Movement**	**Pivoting**	**Data Exfiltration**
Attack Techniques	Active Scanning	VSFTPD	Pass the Hash/Ticket	AutoRoute	DNS Tunneling
	Gathering Victim Host Information		Remote Desktop Protocol	Socks4a	C2 Tunnelling
	Gather Victim Network Information		WMI	Proxy chain	Encode and Encrypt

Bảng 3.3: Các kỹ thuật tấn công sử dụng trong bộ dữ liệu SCVIC-APT-2021

c, Dữ liệu

Dữ liệu thu thập dưới định dạng CSV bao gồm lưu lượng mạng của hệ thống của 4 vòng quy trình APT. Mỗi vòng tương ứng với mỗi đợt xâm nhập APT sử dụng một chiến thuật khác nhau ngẫu nhiên. Bộ dữ liệu cung cấp nhãn về các giai đoạn tấn công APT gồm các giai đoạn được mô tả dưới đây.

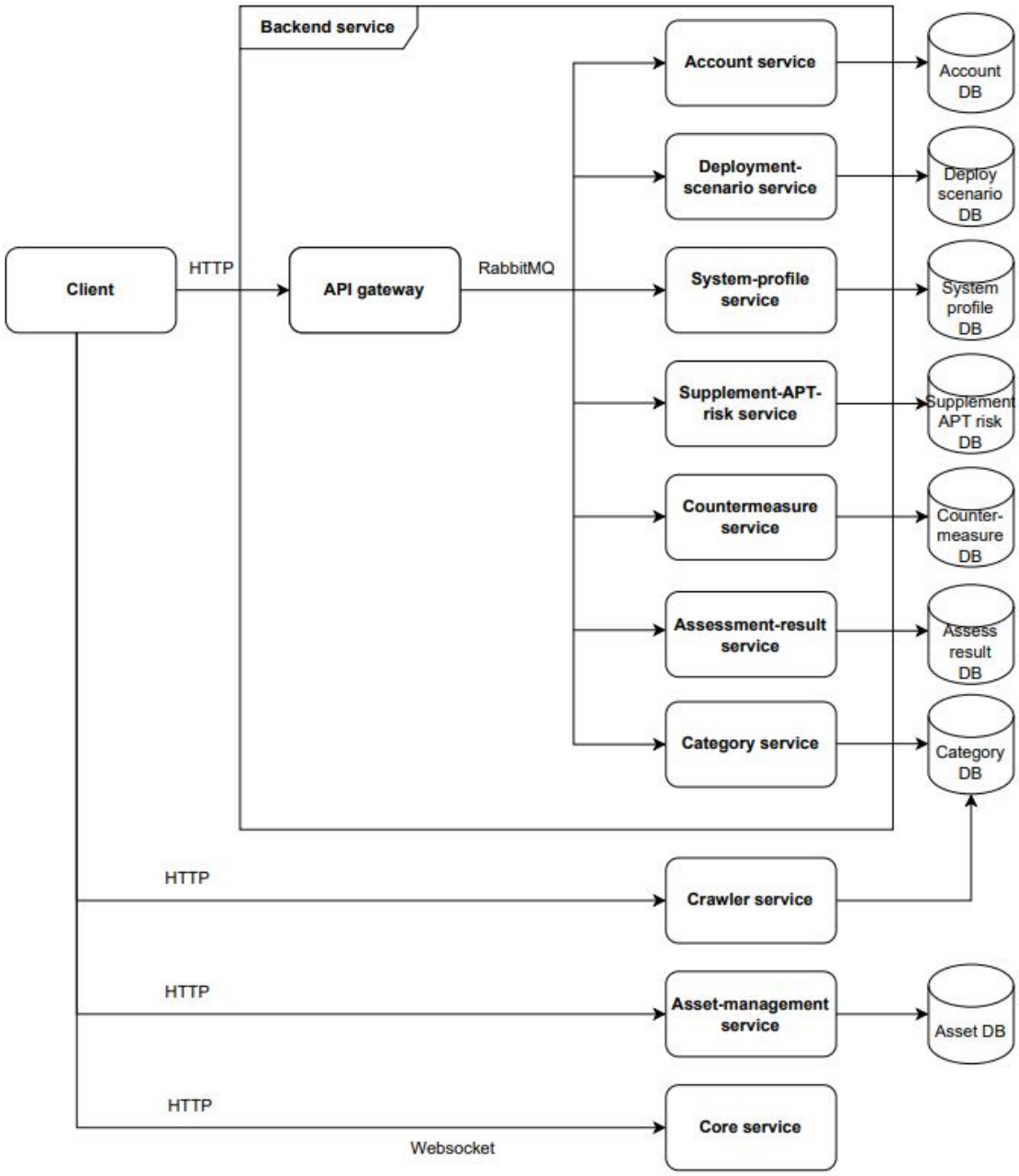
1. **Normal Traffic**: Chưa xảy ra tấn công APT.
2. **Reconnaissance**: Thu thập thông tin xác thực nhằm hỗ trợ các giai đoạn APT tiếp theo
3. **Initial Compromise**: Thiết lập chỗ đứng trong mạng.
4. **Lateral Movement**: Chuyển động ngang để truy cập vào các máy trạm khác nhằm có quyền truy cập vào máy chủ chính hoặc chuyển hướng sang mạng khác.
5. **Pivoting**: Thiết lập các đường hầm hoặc kênh để chuyển hướng sang mạng khác.
6. **Data Exfiltration**: Sau khi xâm phạm mục tiêu mạng và có quyền truy cập vào các mục tiêu có giá trị, kẻ tấn công đánh cắp tài sản và chuyển về máy chủ của mình.

CHƯƠNG 4. THIẾT KẾ, TRIỂN KHAI VÀ ĐÁNH GIÁ HỆ THỐNG

Trong chương trước đồ án đã trình bày đầy đủ về lý thuyết cốt lõi và các công nghệ được sử dụng của hệ thống đánh giá rủi ro ATTT RiDX nói chung và chức năng giám sát rủi ro tấn công APT nói riêng. Chương này sẽ đi sâu trình bày về khía cạnh thiết kế kiến trúc, thiết kế chi tiết, xây dựng phần mềm, kiểm tra và triển khai.

4.1 Thiết kế kiến trúc

4.1.1 Lựa chọn kiến trúc phần mềm



Hình 4.1: Kiến trúc của hệ thống

Hình 4.1 mô tả các thành phần của hệ thống RiDX được xây dựng theo kiến trúc microservice, trong đó chia ứng dụng thành các dịch vụ riêng biệt chạy độc lập. Mỗi dịch vụ sẽ đại diện cho một chức năng khác nhau của hệ thống và kết nối với CSDL riêng của nó. Người dùng sẽ giao tiếp với toàn bộ backend-service thông qua một dịch vụ trung gian duy nhất là API-gateway. API-gateway có tác dụng định tuyến các request của người dùng đến các service phù hợp. Ngoài ra API-gateway còn thực hiện một số nhiệm vụ khác như xác thực và phân quyền, cân bằng tải, caching,... Điều này giúp cho việc quản lý các API bên trong hệ thống trở nên đơn giản hơn và có thể che dấu phần cấu trúc của hệ thống Microservices với bên ngoài.

Các thành phần service của hệ thống như sau:

- API-gateway: Điểm truy cập duy nhất cho các yêu cầu của người dùng, chịu trách nhiệm xác thực bảo mật, phân quyền và định tuyến request đến service thích hợp.
- Account-service: Xử lý xác thực và quản lý người dùng
- Deployment-scenario service: Dịch vụ quản lý thông tin về kịch bản triển khai. Hỗ trợ người dùng khởi tạo, chỉnh sửa kịch bản triển khai và đồ thị tấn công.
- System-profile-service: Quản lý thông tin về hồ sơ hệ thống. Hồ sơ sẽ gồm danh sách các kịch bản triển khai có thể xây dựng của một hệ thống.
- Category-service: Quản lý, hỗ trợ người dùng truy vấn dữ cấc loại dữ liệu CVE, CPE, CWE. Đây là yếu tố tham khảo giúp người dùng xây dựng kịch bản triển khai hợp lý.
- Countermeasure-service: Quản lý thông tin về các biện pháp phòng ngừa rủi ro. Người dùng có thể bổ sung các biện pháp phòng vệ lên kịch bản triển khai để tiến hành so sánh kết quả đánh giá rủi ro so với việc không áp dụng biện pháp phòng vệ.
- Assessment-result service: Xử lý kết quả đánh giá rủi ro đối với kịch bản triển khai giúp người dùng nắm bắt được tình trạng bảo mật của hệ thống.
- Supplement-APT-risk service: Quản lý cấu hình thông số về một cuộc tấn công APT bao gồm khả năng tấn công và khả năng phòng thủ.

Ngoài các service trình bày ở trên, hệ thống RiDX còn bao gồm 3 dịch vụ chưa được tích hợp với API-gateway:

- Assess-management-service: Dịch vụ quản lý tài sản. Gồm các chức năng thêm, bớt, cấu hình các loại tài sản như phần cứng, phần mềm, hệ điều hành,...

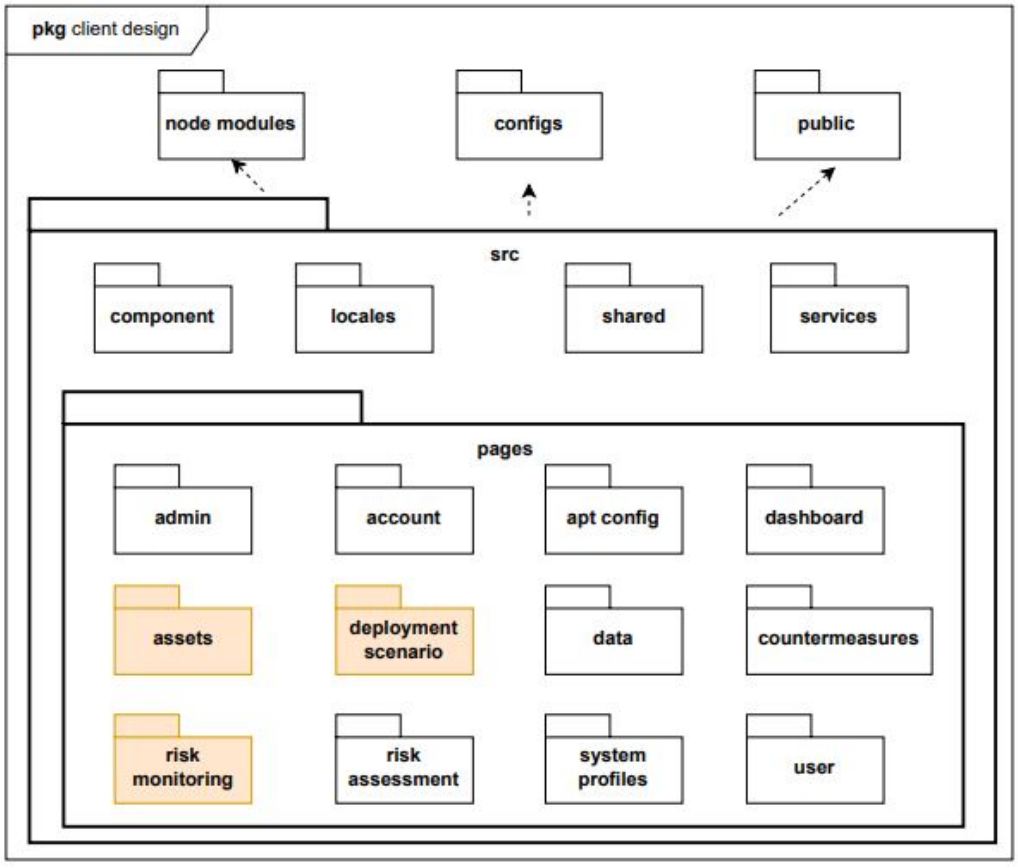
- Core-service: Dịch vụ cốt lõi có chức năng đánh giá và giám sát rủi ro ATTT, phát hiện mối đe dọa và quét lỗ hổng bảo mật.
- Crawler-service: Dịch vụ thu thập tự động dữ liệu về CPE, CVE, CWE và cập nhật cấu hình subnet 1, subnet 3 từ các nguồn dữ liệu chuyên gia.

Với kiến trúc microservice của hệ thống được trình bày như trên, Client sẽ giao tiếp với API-gateway thông qua giao thức HTTP (Restful API). API-gateway sẽ định tuyến request đến các service tương ứng và giao tiếp sử dụng hệ thống hàng đợi message RabbittMQ. Ngoài ra đồ án còn bổ sung một API với giao thức websocket phục vụ chức năng giám sát rủi ro tấn công APT thời gian thực.

Mỗi service bên trong hệ thống microservice sẽ được thiết kế gồm 3 tầng: API, tầng nghiệp vụ và tầng lưu trữ. Tầng API chịu trách nhiệm tiếp nhận các request từ phía client và chuyển request đến tầng nghiệp vụ xử lý. Tầng nghiệp vụ là nơi request được xử lý logic, tổ chức các luồng dữ liệu và gửi yêu cầu truy xuất dữ liệu đến tầng lưu trữ. Tầng lưu trữ dữ liệu là thành phần tương tác chính với CSDL, có chức năng truy xuất, cập nhật dữ liệu theo yêu cầu xử lý từ tầng nghiệp vụ.

4.1.2 Thiết kế tổng quan

a, Thiết kế tổng quan thành phần Client

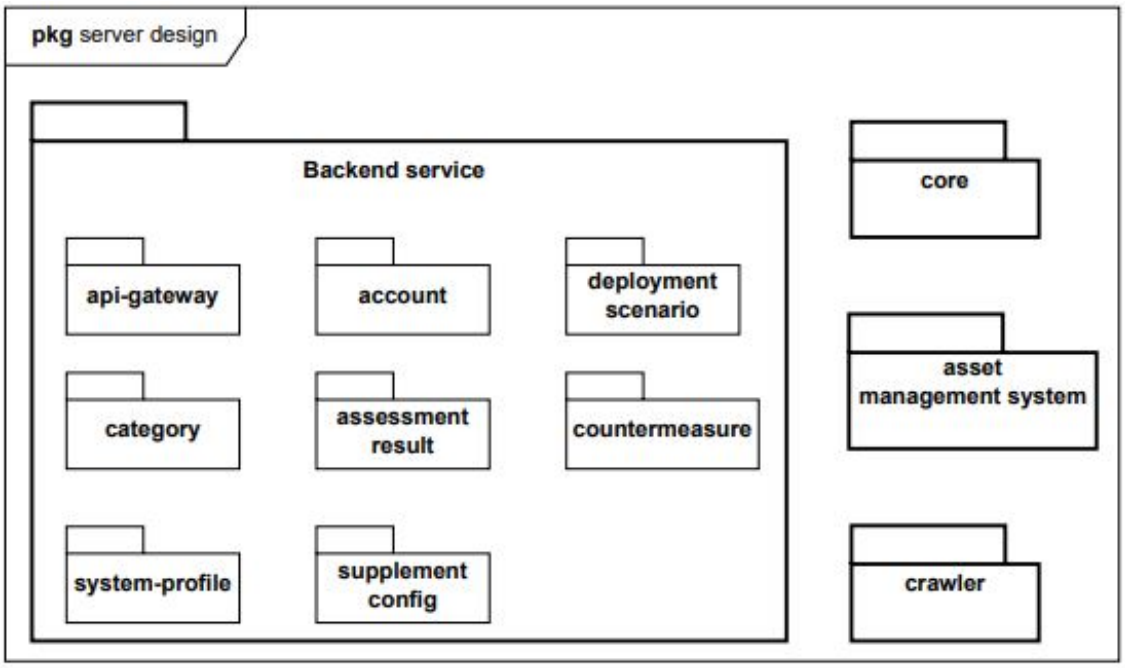


Hình 4.2: Thiết kế tổng quan thành phần client

Hình 4.2 mô tả cấu trúc mã nguồn thành phần Client. Sơ đồ thể hiện các thành phần giao diện chính bao gồm giao diện admin, giao diện quản lý tài sản, giao diện quản lý kịch bản triển khai,... Các thành bôi cam là các thành phần được chỉnh sửa để tích hợp chức năng đánh giá rủi ro tấn công APT. Các gói của thành phần Client bao gồm:

- node-modules: lưu trữ các thư viện được sử dụng bởi ứng dụng.
- locale: hỗ trợ đa ngôn ngữ trong ứng dụng.
- components: chứa các thành phần chung được chia sẻ trên các phần khác nhau của ứng dụng.
- shared: chứa mã có thể tái sử dụng trong toàn bộ ứng dụng.
- service: bao gồm dịch vụ API chịu trách nhiệm gửi yêu cầu đến máy chủ.
- pages: bao gồm các trang giao diện người dùng cho ứng dụng.

b, Thiết kế tổng quan thành phần server



Hình 4.3: Thiết kế tổng quan thành phần server

Hình 4.3 biểu diễn cấu trúc mã nguồn của máy chủ. Sơ đồ bao gồm thành phần backend-service thiết kế theo kiến trúc microservice và ba thành phần kèm theo. Các thành phần của hệ thống server bao gồm:

- Backend-service: được thiết kế theo kiến trúc microservice sử dụng khung NestJs, bao gồm các thành phần chính là apps - chứa các dịch vụ xử lý yêu cầu từ client, libs - các thành phần được tái sử dụng nhiều lần trong mã nguồn và

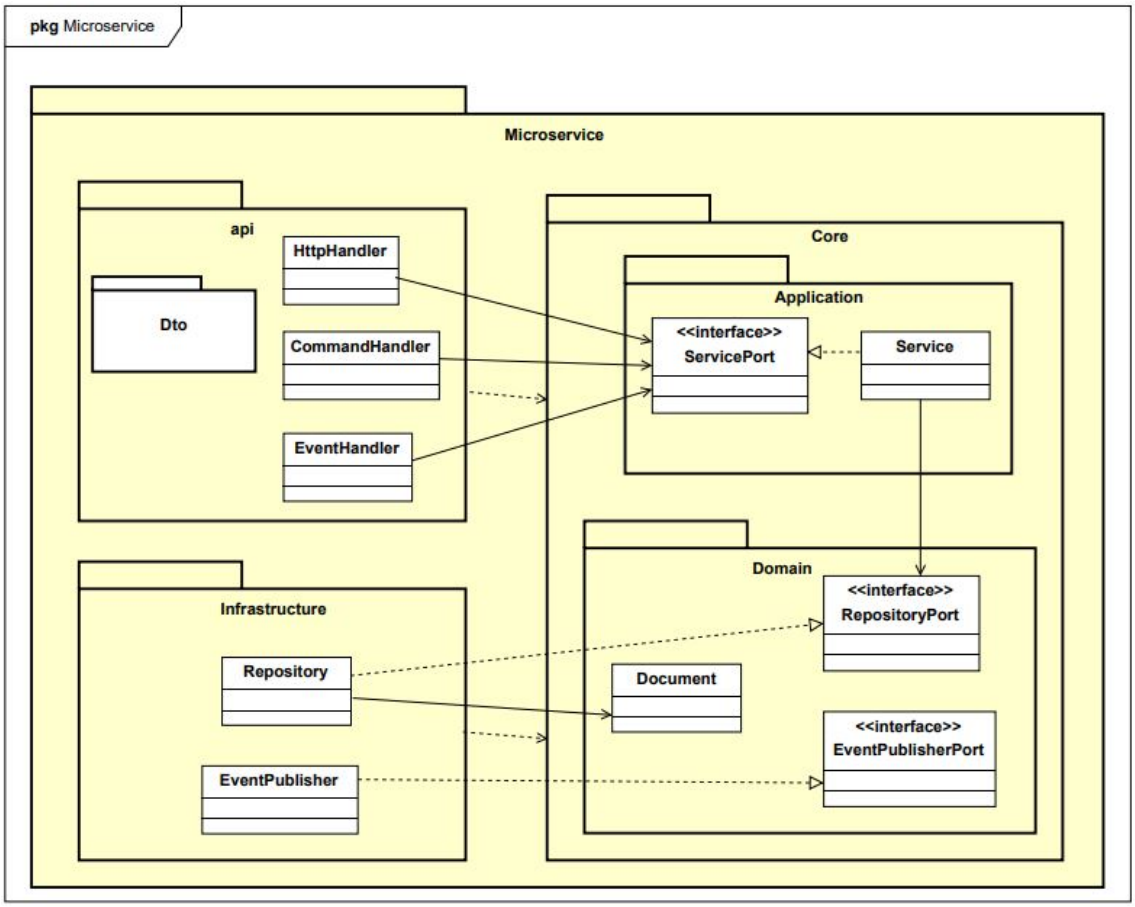
node-modules - thư viện được sử dụng bởi ứng dụng.

- Core-service: dịch vụ cốt lõi đánh giá rủi ro sử dụng mạng Bayes, được đóng gói thành một service riêng viết bằng ngôn ngữ Python.
- Asset-management-service: dịch vụ quản lý tài sản, được đóng gói thành một service riêng viết bằng khung NestJs.
- Crawler-service: dịch vụ thu thập tự động dữ liệu danh mục và cập nhật cấu hình APT.

4.1.3 Thiết kế chi tiết gói

Trong phần này đồ án sẽ trình bày thiết kế chi tiết gói chung của microservice, API-gateway và Core-service.

a, Thiết kế chi tiết gói chung của microservice

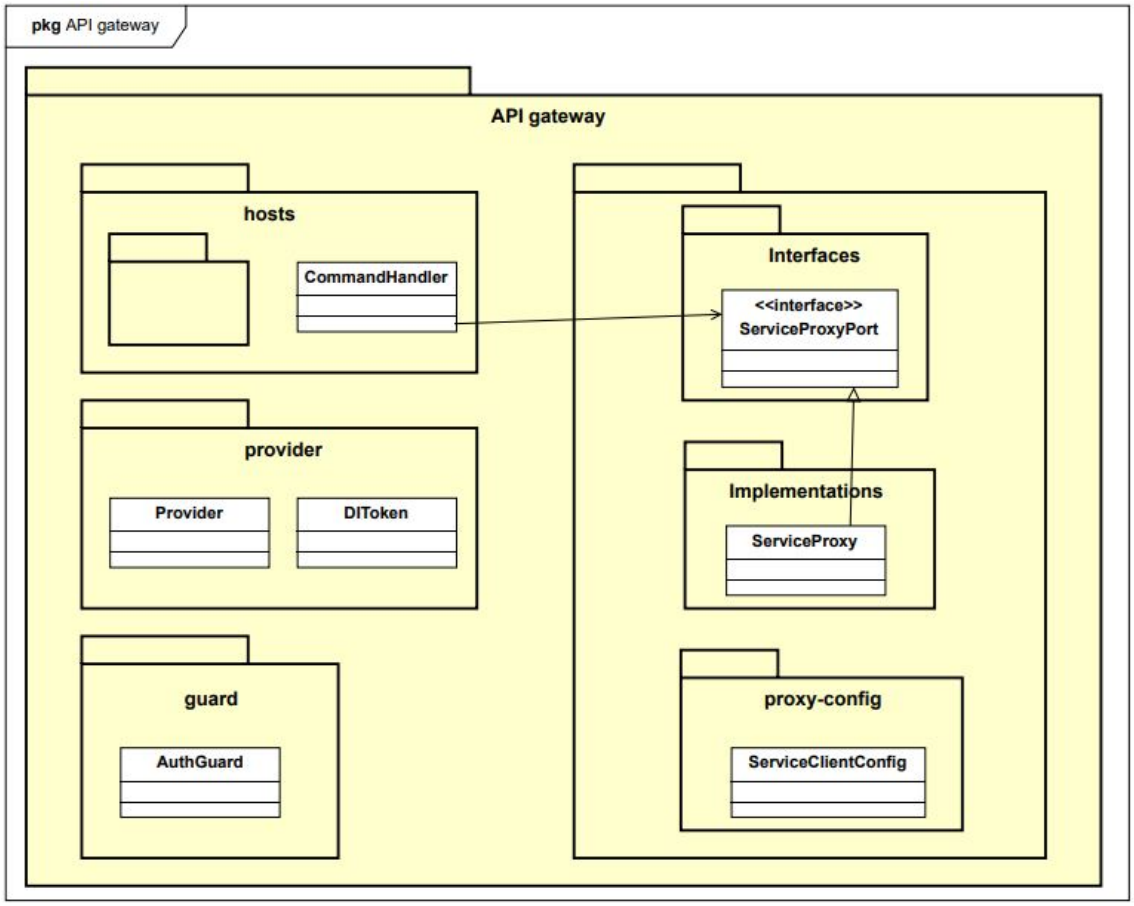


Hình 4.4: Thiết kế chi tiết gói chung mỗi microservice

Hình 4.4 là thiết kế chi tiết gói chung cho mỗi microservice. Các thành phần HttpHandler, EvenHandler và CommandHandler có chức năng là các endpoint của API, phụ thuộc vào giao diện ServicePort được định nghĩa trong gói core. Lớp Service thực hiện từ giao diện ServicePort có chức năng xử lý nghiệp vụ trong

máy chủ. Tương tự lớp Repository thực hiện từ giao diện RepositoryPort có chức năng CRUD dữ liệu. Thiết kế này đảm bảo tiêu chí DI(Dependency Injection) - phụ thuộc lỏng lẻo, các lớp không bị phụ thuộc lẫn nhau mà phụ thuộc vào triển khai của chúng. Điều này giúp ứng dụng có thể dễ dàng thay thế các tính năng và mở rộng.

b, Thiết kế chi tiết gói API-gateway



Hình 4.5: Thiết kế chi tiết gói API-gateway

Hình 4.5 mô tả thiết kế đóng gói chi tiết của dịch vụ API-gateway - dịch vụ trung gian xử lý tất cả các yêu cầu từ Client. Các thành phần ở đây bao gồm:

- hosts: Chứa toàn bộ các bộ điều khiển chịu trách nhiệm xử lý yêu cầu từ phía client. Nó có chức năng định tuyến các yêu cầu đến các proxy dịch vụ tương ứng trong hệ thống microservice.
- service: Bao gồm các giao diện triển khai bởi các proxy dịch vụ mà các bộ điều khiển phụ thuộc.
- provider: Chứa các phần phụ thuộc cho các proxy dịch vụ.
- auth: Gồm các lớp bảo mật, tích hợp cơ chế xác thực và phân quyền đối với

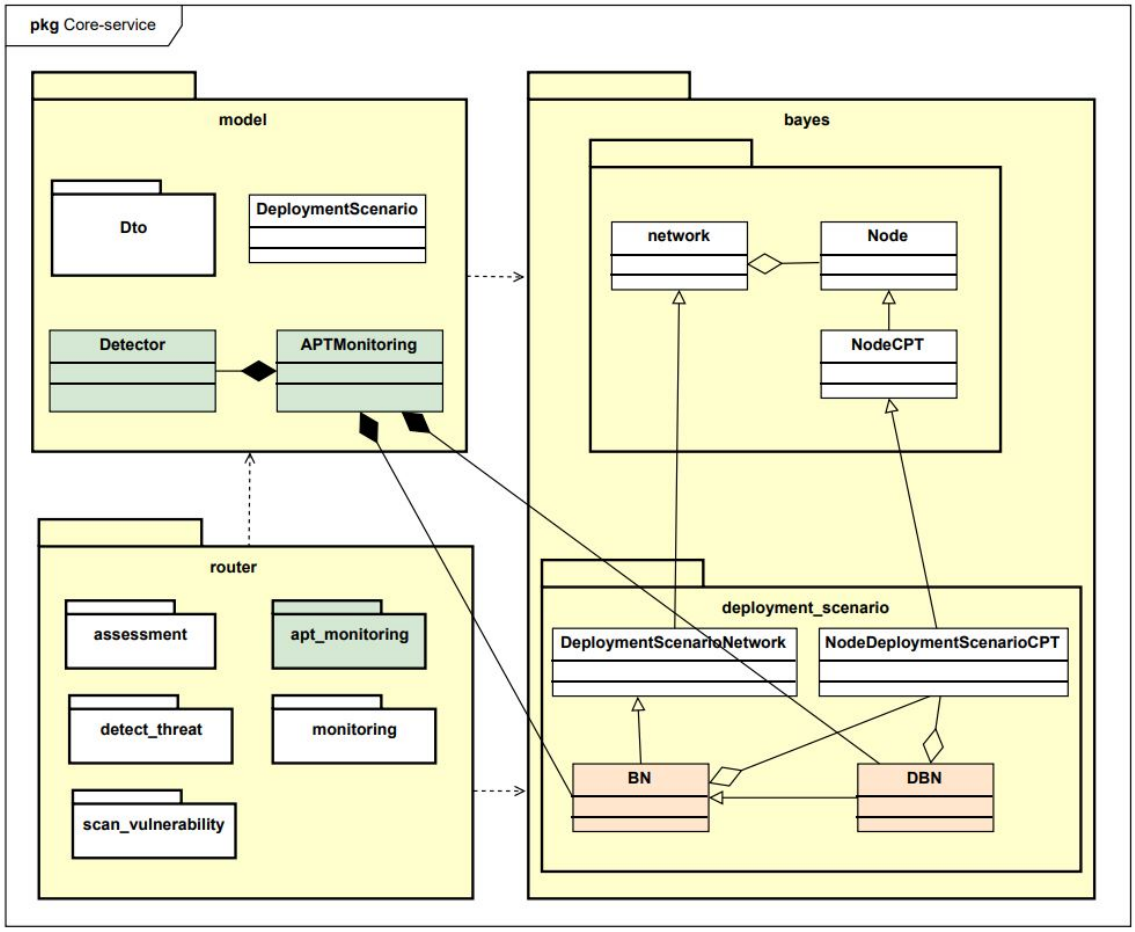
các yêu cầu từ phía client.

c, Thiết kế chi tiết gói Core-service

Hình 4.6 cung cấp mô tả chi tiết về các gói và mối quan hệ phụ thuộc giữa các lớp bên trong thành phần Core-service. Các gói bên trong Core-service bao gồm:

- router: gói router gồm các endpoint của các API bao gồm đánh giá rủi ro, giám sát rủi ro tấn công APT, phát hiện mối đe dọa và rà soát lỗ hổng.
- model: bao gồm các lớp DeploymentScenario, APTMonitoring, Detector và gói Dto.
- bayes: gói bayes bao gồm các lớp cốt lõi của hệ thống mạng Bayes được sử dụng để đánh giá kết quả rủi ro kịch bản triển khai.

Các lớp màu xanh được thêm mới phục vụ chức năng giám sát rủi ro tấn công APT, trong đó lớp Detector có nhiệm vụ phán đoán thông tin về cuộc tấn công từ dữ liệu đầu vào, lớp APTMonitoring tổng hợp dữ liệu và đánh giá rủi ro cuộc tấn công. Các lớp màu cam được chỉnh sửa để phục vụ chức năng mới bổ sung trên.



Hình 4.6: Thiết kế chi tiết gói Core-service

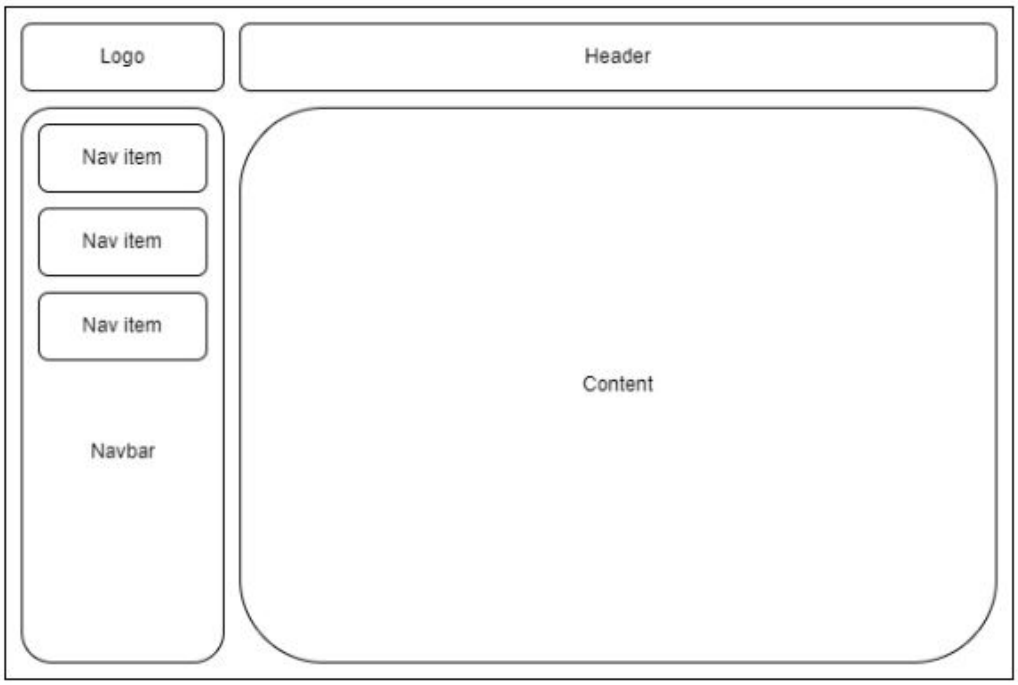
4.2 Thiết kế chi tiết

4.2.1 Thiết kế giao diện

a, Thiết kế giao diện tổng quan

Giao diện tổng quan hệ thống RiDX được thiết kế đáp ứng các tiêu chí về giao diện thân thiện, dễ sử dụng và hiệu quả. Các yếu tố thiết kế giúp đảm bảo xử lý kết xuất hình ảnh hiệu quả, thời gian phản hồi kịp thời và hiệu suất tối ưu trên màn hình desktop và laptop. Giao diện cần cung cấp các hiệu ứng và gợi ý bằng biểu tượng hoặc thanh điều hướng giúp người dùng tương tác một cách dễ dàng.

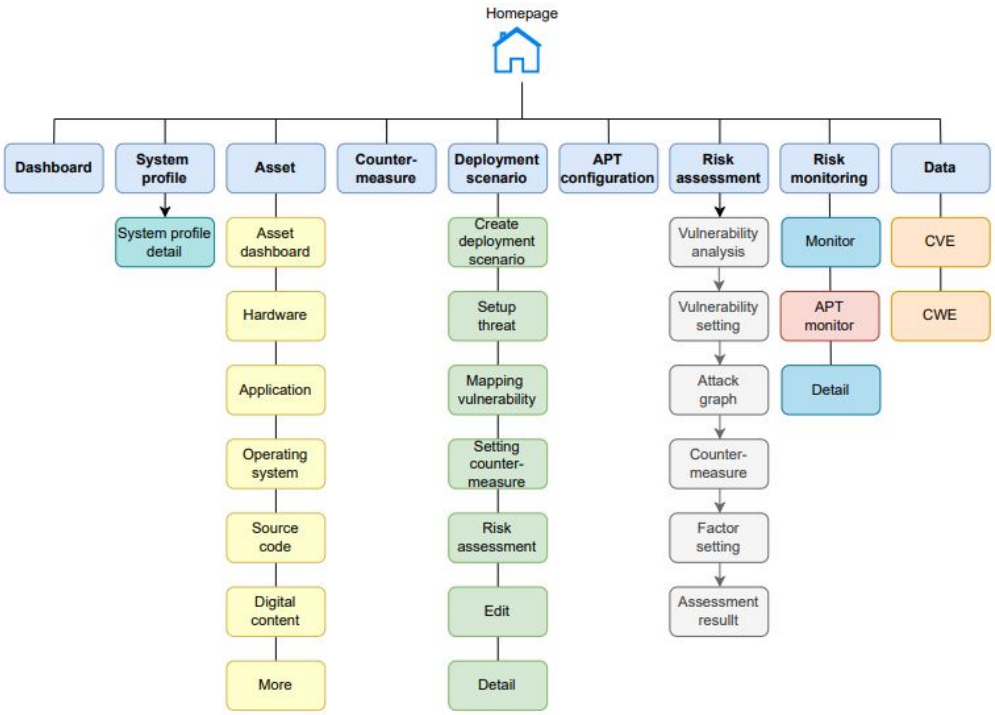
Màu sắc chủ đạo của giao diện là trắng và xanh dương. Đối với mỗi chức năng, màu sắc được điều chỉnh để phù hợp với yêu cầu nghiệp vụ tương ứng. Ví dụ màu sắc biểu thị mức độ rủi ro sẽ gồm bốn màu theo thứ tự xanh lá cây, vàng, cam và đỏ.



Hình 4.7: Bố cục chung của giao diện hệ thống

Về thiết kế giao diện tổng thể, tất cả các trang đều tuân theo một bố cục nhất quán, bao gồm bốn thành phần chính: (i) Header - đặt ở đầu trang, (ii) Logo - biểu tượng của hệ thống, (iii) Thanh navbar bao gồm các navitem - là menu điều hướng đến các chức năng dịch vụ khác nhau, (iv) Nội dung - thành phần nội dung chính của trang. Bố cục giao diện chung của hệ thống được minh họa trong hình 4.7

b, Thiết kế sitemap giao diện hệ thống RiDX

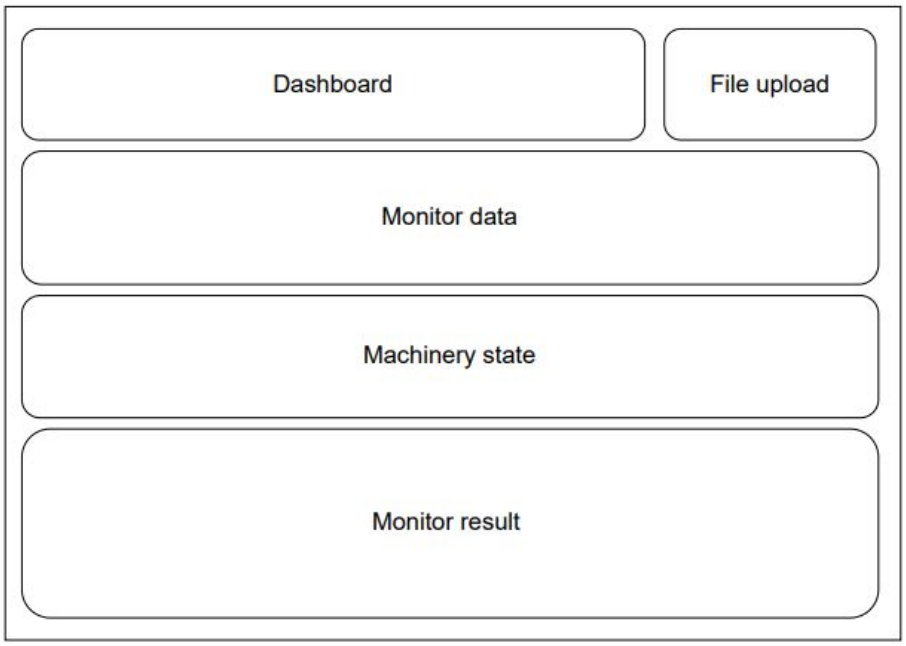


Hình 4.8: Biểu đồ sitemap giao diện hệ thống RiDX

Biểu đồ sitemap giao diện hệ thống RiDX được thiết kế tại hình 4.8. Các thành phần giao diện chính bao gồm (i) Dashboard, (ii) System-profile, (iii) Asset, (iv) Countermeasure, (v) Deployment-scenario, (vi) APT-config, (vii) Risk-assessment, (viii) Risk-monitoring, (ix) Data. Thành phần giao diện APT-monitor phục vụ chức năng giám sát rủi ro tấn công APT được bổ sung vào trang Risk-monitor.

c, Thiết kế giao diện chức năng giám sát rủi ro tấn công APT

Giao diện chức năng giám sát rủi ro tấn công APT được thiết kế nhằm đáp ứng các yêu cầu hiển thị các thông tin về dữ liệu thu thập, kết quả giám sát và trạng thái hệ thống. Màu sắc của các thành phần giám sát được cập nhật trong quá trình giám sát rủi ro nhằm thể hiện tương ứng mức bộ rủi ro của hệ thống. Bố cục giao diện chức năng giám sát rủi ro tấn công APT được minh họa tại hình 4.9. Các thành phần chính bao gồm (i) Dashboard - thông tin giám sát chung, (ii) Đẩy file - cho phép đẩy file dữ liệu đầu vào, (iii) Dữ liệu giám sát - hiển thị dữ liệu giám sát thu thập (các thông tin lưu lượng mạng), (iv) Biểu đồ trạng thái máy - biểu diễn trạng thái của hệ thống trong quá trình giám sát và (v) Kết quả giám sát rủi ro - các biểu đồ hiển thị thông tin giám sát rủi ro qua từng bước.



Hình 4.9: Bố cục của giao diện chức năng giám sát sủi ro tấn công APT

d, Thiết kế giao diện dịch vụ chức năng giám sát rủi ro tấn công APT

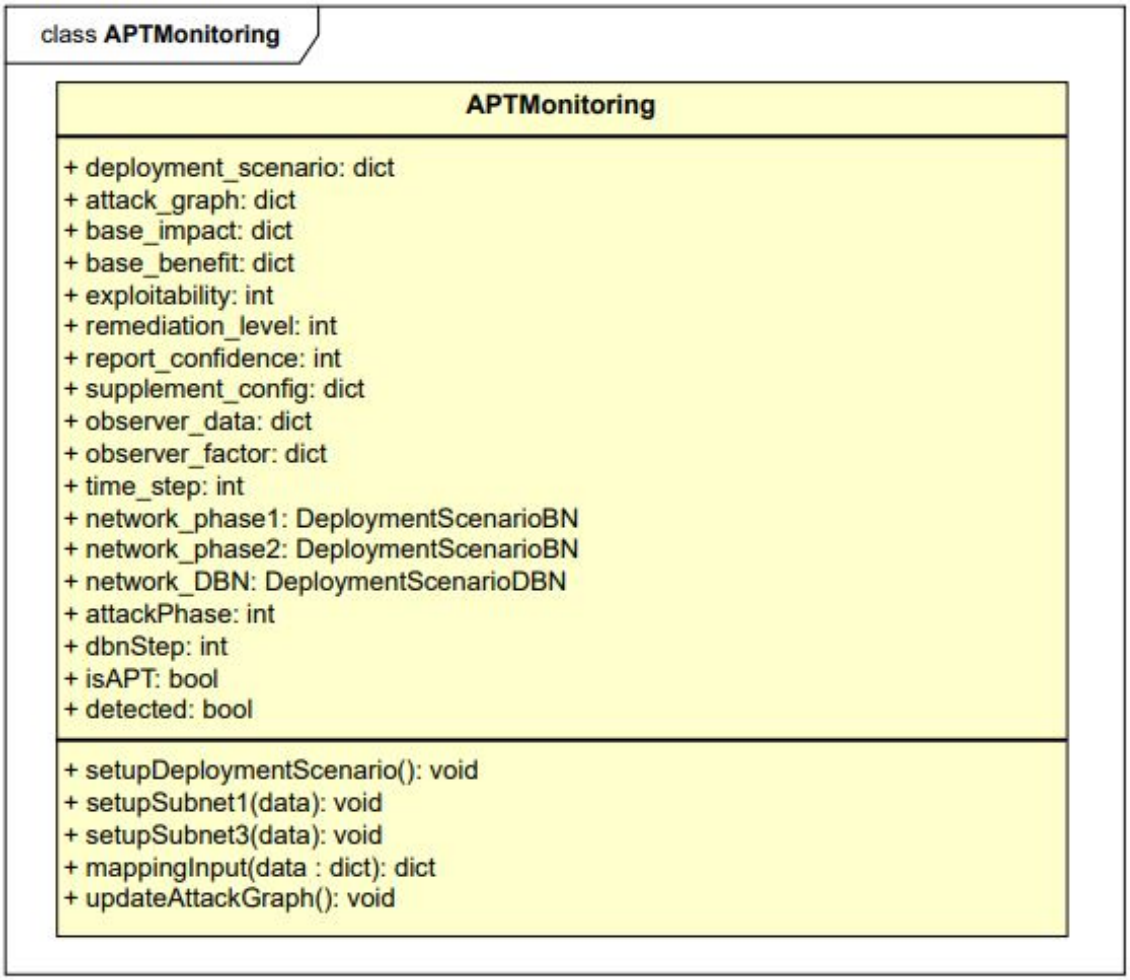
Tên mẫu	Đầu vào	Đầu ra	Chức năng
get-deployment-scenario	deploymentScenarioId	response	trích xuất kịch bản triển khai theo id
get-attack-graph	deploymentScenarioId	response	trích xuất đồ thị tấn công từ kịch bản triển khai
get-supplement-config	deploymentScenarioId	response	trích xuất cấu hình bổ sung từ kịch bản triển khai
apt-monitoring	- deploymentScenario - supplementConfig - monitorData	response	giám sát rủi ro tấn công APT
get-system-profile	systemProfileId	response	trích xuất hồ sơ hệ thống theo id

Bảng 4.1: Giao diện dịch vụ chức năng giám sát rủi ro tấn công APT

Bảng 4.1 mô tả danh sách các chức năng được sử dụng tại giao diện giám sát rủi ro tấn công APT. Đầu tiên phía server cung cấp dữ liệu kịch bản triển khai và cấu hình bổ sung. Client sử dụng các thông tin trên cùng với dữ liệu đầu vào để tiến hành giám sát rủi ro hệ thống. Ngoài ra một số chức năng được cung cấp bao gồm trích xuất và hiển thị thông tin hồ sơ hệ thống và đồ thị tấn công.

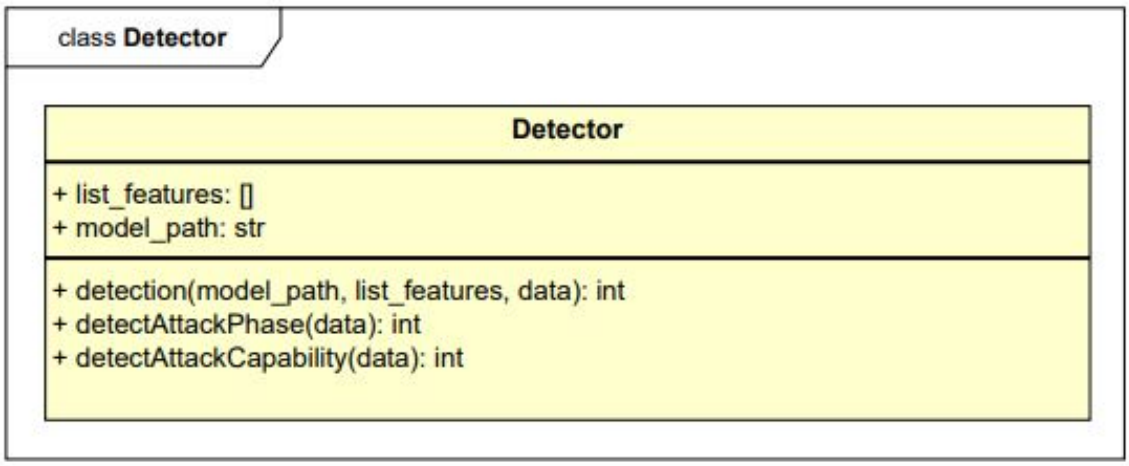
4.2.2 Thiết kế chi tiết lớp

Phần này sẽ trình bày thiết kế chi tiết lớp AptMonitoring và Detector triển khai tại Core-service phục vụ chức năng giám sát rủi ro tấn công APT. Lớp APTMonitoring có chức năng đánh giá rủi ro ATTT trong quá trình giám sát, bao gồm các thành phần như kịch bản triển khai (deployment-scenario), cấu hình bổ sung (supplement-config), mạng Bayes động (network-DBN)... Một số chức năng quan trọng mà lớp cung cấp là setupSubnet1, setupSubnet3, mappingInput - đánh giá kết quả giám sát. Các tham số và chức năng của lớp APTMonitoring được mô tả tại hình 4.10.



Hình 4.10: Thiết kế chi tiết lớp APTMonitoring

Lớp Detector có chức năng phán đoán dữ liệu đầu vào cho mạng đánh giá rủi ro bao gồm thông tin về giai đoạn tấn công, năng lực tấn công. Lớp sử dụng các mô hình học máy được huấn luyện sẵn để tiến hành phán đoán từ dữ liệu giám sát đầu vào. Hình 4.11 dưới đây là thiết kế chi tiết lớp Detector.

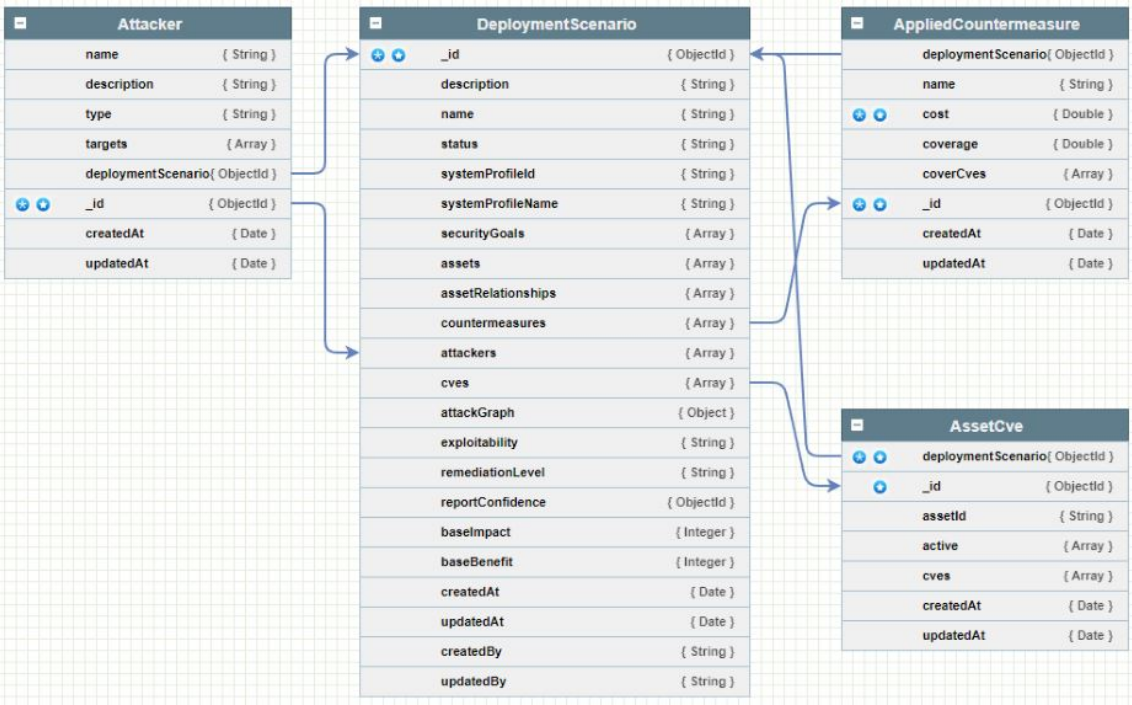


Hình 4.11: Thiết kế chi tiết lớp Detector

4.2.3 Thiết kế cơ sở dữ liệu

Cơ sở dữ liệu của hệ thống RiDX được tổ chức không cấu trúc (NoSQL) lưu trữ dưới dạng document bởi MongoDB. Đối với kiến trúc microservice, mỗi dịch vụ có một cơ sở dữ liệu riêng. Dưới đây là thiết kế cơ sở dữ liệu một số dịch vụ phục vụ giám sát rủi ro tấn công APT.

a, Cơ sở dữ liệu dịch vụ kịch bản triển khai



Hình 4.12: Thiết kế cơ sở dữ liệu kịch bản triển khai

Hình 4.12 mô tả thiết kế cơ sở dữ liệu của dịch vụ kịch bản triển khai bao gồm 4 collections: DeploymentScenario, Attacker, AppliedCountermeasure và AssetCVE.

b, Cơ sở dữ liệu dịch vụ cấu hình APT bổ sung

Hình 4.13: Thiết kế cơ sở dữ liệu dịch vụ cấu hình APT bổ sung

Cơ cở dữ liệu dịch vụ cấu hình APT bổ sung được thiết kế theo hình 4.13 gồm 2 collection là FactorAssessment và AptConfig. FactorAssessment là kết quả đánh giá rủi ro về subnet 1 và subnet 3 của hệ thống còn AptConfig là dữ liệu về subnet1 và subnet 3 do người dùng cấu hình.

4.3 Xây dựng ứng dụng

4.3.1 Thư viện và công cụ sử dụng

Hình 4.2 liệt kê các công cụ và thư viện được sử dụng trong quá trình xây dựng hệ thống RiDX.

Mục đích	Công cụ	Địa chỉ URL
Code editor	Visual Studio Code 1.89.1	https://code.visualstudio .com/
IDE lập trình	PyCharm Community Edition 2021.3.3	https://www.jetbrains.com /pycharm/
Lập trình backend	FastAPI	https://fastapi.tiangolo.com
Thư viện hỗ trợ tính toán mạng Bayes	PySmile	https://www.bayesfusion.com /smile
Môi trường lập trình backend	NodeJS	https://nodejs.org
Lập trình backend	NestJS	https://nestjs.com
Lập trình giao diện	ReactJS	https://reactjs.org
Database	MongoDB	https://www.mongodb.com
Nền tảng triển khai ứng dụng ảo hóa	Docker	https://www.docker.com

Bảng 4.2: Thư viện và công cụ sử dụng

4.3.2 Kết quả đạt được

Sau quá trình xây dựng các chức năng phục vụ nghiệp vụ đồ án, kết quả đạt được như sau: ứng dụng bổ sung thêm chức năng giám sát rủi ro tấn công APT tích hợp dữ liệu thời gian thực tích hợp với các chức năng hiện có. Toàn bộ mã nguồn được đẩy lên thư mục Github của đề tài cùng với tài liệu kỹ thuật, hướng dẫn sử dụng hệ thống. Thống kê chi tiết phần mềm được đưa ra tại bảng 4.3.

Information	Statistic
Số lượng ứng dụng	13
Số lượng lớp tại server	332
Số lượng màn hình thêm mới và cập nhật	3
Số lượng gói tại server	101
Tổng dung lượng mã nguồn toàn bộ hệ thống	1.2GB

Bảng 4.3: Thống kê ứng dụng

4.4 Kiểm thử

4.4.1 Kiểm thử tương thích

Bảng 4.4 biểu diễn kết quả kiểm thử tương thích ứng dụng trên một số thiết bị.

Device	Specifications	UI	Function
Dell Vostro	Screen: 15.6 inches FHD, 16GB RAM	Pass	Pass
Laptop Lenovo	Screen: 13 inches HD, 8GB RAM	Pass	Pass

Bảng 4.4: Kết quả kiểm thử tương thích

4.4.2 Kiểm thử chức năng

a, Chức năng import file dữ liệu đầu vào

STT	Input	Output	Exception	Kết quả
1	Import file	Màn hình hiển thị tên file import	Không xử lý	OK
2	Không import file, nhấn submit	Màn hình hiển thị cảnh báo chưa nhập file	Không xử lý	OK

b, Chức năng giám sát rủi ro tấn công APT

STT	Input	Output	Exception	Kết quả
1	Import file CSV lỗi	Màn hình hiển thị cảnh báo dữ lỗi dữ liệu đầu vào	Có	OK
2	Dữ liệu đầu vào thuộc kịch bản triển khai khác	Màn hình trả về kết quả giám sát mặc định với kịch bản triển khai	Không	OK

4.5 Triển khai

Hệ thống RiDX với bản cập nhật bổ sung chức năng giám sát rủi ro ATTT với dữ liệu thời gian thực đã được triển khai tại địa chỉ https://ridx.id.vn. Ngoài ra có thể chạy hệ thống trên localhost theo hướng dẫn đã được ghi chú trong file README của mã nguồn hệ thống.

CHƯƠNG 5. CÁC GIẢI PHÁP VÀ ĐÓNG GÓP NỔI BẬT

Trong chương trước đồ án đã trình bày về thiết kế kiến trúc của hệ thống, quá trình thiết kế, kiểm thử và triển khai sản phẩm hệ thống RiDX. Chương này sẽ lần lượt trình bày về những đóng góp chính của đồ án: (i) Xây dựng kịch bản triển khai hệ thống SCVIC-APT-2021, (ii) Ánh xạ dữ liệu tấn công APT với đầu vào mạng đánh giá rủi ro, (iii) Xây dựng chức năng giám sát rủi ro tấn công APT thời gian thực.

5.1 Xây dựng kịch bản triển khai cho bộ dữ liệu SCVIC-APT-2021

5.1.1 Vấn đề

Để xây dựng cơ chế giám sát rủi ro tấn công APT hiệu quả tích hợp với dữ liệu thực tế cần có một kịch bản triển khai hệ thống phù hợp với hệ thống xây dựng của tập dữ liệu. Dữ liệu đầu vào là 2 tập dữ liệu Unraveled và SCVIC-APT-2021 trong đó kịch bản triển khai của tập dữ liệu Unraveled đã được xây dựng chi tiết trong đồ án của anh Nguyễn Việt Thắng CNTT K63. Cần xây dựng kịch bản triển khai cho tập dữ liệu SCVIC-APT-2021 trong đó các yếu tố về tài sản, mối quan hệ, lỗ hổng bảo mật và mục tiêu phù hợp với thông tin về hệ thống do tác giả của tập dữ liệu đưa ra [3].

5.1.2 Giải pháp

Kịch bản triển khai thử nghiệm của hệ thống SCVIC-APT-2021 được xây dựng từ những phân tích về thông tin của hệ thống được sử dụng để xây dựng tập dữ liệu. Thông tin về bộ dữ liệu đã được đề cập tại chương 3.6.2.

5.1.3 Kết quả

Dưới đây là kịch bản triển khai xây dựng của hệ thống từ tập dữ liệu SCVIC-APT-2021. Bảng 5.1 liệt kê danh sách các tài sản của hệ thống. Thông tin về tài sản bao gồm phiên bản tài sản và thông tin về định danh CPE tương ứng với mỗi tài sản.

ID	Asset name	Product name	CPE
Win-1-D1	Computer (Windows)	Windows 10 1709	cpe:2.3:o:microsoft:windows_10 :1709:*:*:*:*:*:x64:*
Win-2-D1	Computer (Windows)	Windows 10 1709	cpe:2.3:o:microsoft:windows_10 :1709:*:*:*:*:*:x64:*
FTP-D1	FTP server (Domain 1)	VSFTPD 2.3.4	cpe:2.3:a:vsftpd_project:vsftpd :2.3.4:*:*:*:*:*:*:*
DC-1	Domain Controller 1	Windows server 1709	cpe:2.3:o:microsoft:windows_server :1709:*:*:*:*:*:*:*
Win-D2	Computer (Windows)	Windows 10 1709	cpe:2.3:o:microsoft:windows_10 :1709:*:*:*:*:*:x64:*
DC-2	Domain Controller 2	Windows server 1709	cpe:2.3:o:microsoft:windows_server :1709:*:*:*:*:*:*:*

Bảng 5.1: Danh sách tài sản hệ thống SCVIC-APT-2021

Bảng 5.2 mô tả quan hệ giữa các tài sản, bao gồm thông tin về tài sản nguồn, tài sản đích, vector truy cập và quyền truy cập.

No	ID source target	ID target asset	Access vector	Privilege
1	FTP-D1	Win-1-D1	NETWORK	NONE
2	FTP-D1	Win-2-D1	NETWORK	NONE
3	FTP-D1	DC-1	NETWORK	OS_USER
4	Win-1-D1	DC-1	NETWORK	OS_USER
5	Win-2-D1	DC-1	NETWORK	OS_USER
6	DC-1	DC-2	ADJACENT_NETWORK	OS_ADMIN
7	DC-2	Win-D2	NETWORK	OS_ADMIN

Bảng 5.2: Mối quan hệ giữa các tài sản hệ thống SCVIC-APT-2021

Asset	Security goal	Confidentialy	Intergrity	Availability
Win-1-D1	Hệ điều hành luôn hoạt động ổn định	LOW	HIGH	HIGH
	Quyền root trên hệ điều hành được bảo vệ	HIGH	MEDIUM	HIGH
Win-2-D1	Hệ điều hành luôn hoạt động ổn định	LOW	HIGH	HIGH
	Quyền root trên hệ điều hành được bảo vệ	HIGH	MEDIUM	HIGH
FTP-D1	FTP server hoạt động ổn định	HIGH	HIGH	HIGH
DC-1	Ngăn chặn truy cập trái phép	HIGH	MEDIUM	HIGH
Win-D2	Hệ điều hành luôn hoạt động ổn định	LOW	HIGH	HIGH
	Quyền root trên hệ điều hành được bảo vệ	HIGH	MEDIUM	HIGH
DC-2	Ngăn chặn truy cập trái phép	HIGH	MEDIUM	HIGH

Bảng 5.3: Danh sách mục tiêu bảo mật hệ thống SCVIC-APT-2021

Bảng 5.3 cho biết thông tin về mục tiêu bảo mật cần đạt được trên mỗi tài sản. Mục tiêu bảo mật bao gồm các khía cạnh C (Confidentally), I (Intergrity), A (Availability). Các chỉ số CIA bao gồm 3 mức độ LOW, MEDIUM, HIGH.

ID	Countermeasure	Covered CVE	Coverage	Cost
1	Áp dụng tiện ích APLN và SNI trong TLS	CVE-2021-3618	0.8	100

Bảng 5.4: Danh sách biện pháp phòng vệ hệ thống SCVIC-APT-2021

Bảng 5.4 liệt kê danh sách các biện pháp phòng vệ được áp dụng để phòng ngừa lỗ hổng xuất hiện bên trong hệ thống. Việc áp dụng biện pháp phòng vệ yêu cầu mức chi phí nhất định tương ứng với mức độ bao phủ trong việc khai thác lỗ hổng.

Asset	CVE ID	Attack vector	Pre-condition	Post-condition	C	I	A
FTP -D1	CVE-2021 -3618	NETWORK	NONE	OS_USER	Comp-lete	Complete	Comp-lete
Win-1 -D1	CVE-2022 -24502	NETWORK	OS_ USER	NONE	None	None	Partial
	CVE-2022 -24503	NETWORK	OS_ USER	NONE	None	None	Partial
	CVE-2020 -0659	NETWORK	OS_ USER	NONE	None	None	Partial
	CVE-2020 -0660	NETWORK	OS_ USER	NONE	Partial	Partial	Partial
Win-2 -D1	CVE-2022 -24502	NETWORK	OS_ USER	NONE	None	None	Partial
	CVE-2022 -24503	NETWORK	OS_ USER	NONE	None	None	Partial
	CVE-2020 -0659	NETWORK	OS_ USER	NONE	None	None	Partial
	CVE-2020 -0660	NETWORK	OS_ USER	NONE	Partial	Partial	Partial
Win -D2	CVE-2022 -24502	NETWORK	OS_ ADMIN	OS_USER	None	None	Partial
	CVE-2022 -24503	NETWORK	OS_ ADMIN	OS_USER	None	None	Partial
	CVE-2020 -0659	NETWORK	OS_ ADMIN	OS_USER	None	None	Partial
	CVE-2020 -0660	NETWORK	OS_ ADMIN	OS_USER	Partial	Partial	Partial
DC-1	CVE-2018 -11788	NETWORK	OS_ USER	OS _ADMIN	None	None	Partial
	CVE-2017 -11850	NETWORK	OS_ USER	OS_USER	None	None	Partial
	CVE-2017 -11927	NETWORK	OS_ USER	OS_USER	None	None	Partial
DC-2	CVE-2018 -11788	ADJACENT-NETWORK	OS_ ADMIN	OS_USER	Comp-lete	Complete	Comp-lete
	CVE-2017 -11850	ADJACENT-NETWORK	OS_ ADMIN	OS_USER	None	None	Partial
	CVE-2017 -11927	ADJACENT-NETWORK	OS_ ADMIN	OS_USER	None	Partial	Partial

Bảng 5.5: Thông tin về các lỗ hổng CVE liên quan của hệ thống SCVIC-APT-2021

Bảng 5.5 mô tả thông tin về các lỗ hổng CVE liên quan của hệ thống SCVIC-APT-2021. Thông tin về CVE bao gồm định danh CVE, tiền điều kiện khai thác, hậu điều kiện (quyền đạt được sau khi khai thác) và các khía cạnh CIA tương ứng.

5.2 Tích hợp dữ liệu thực tế với mạng đánh giá rủi ro RiDX

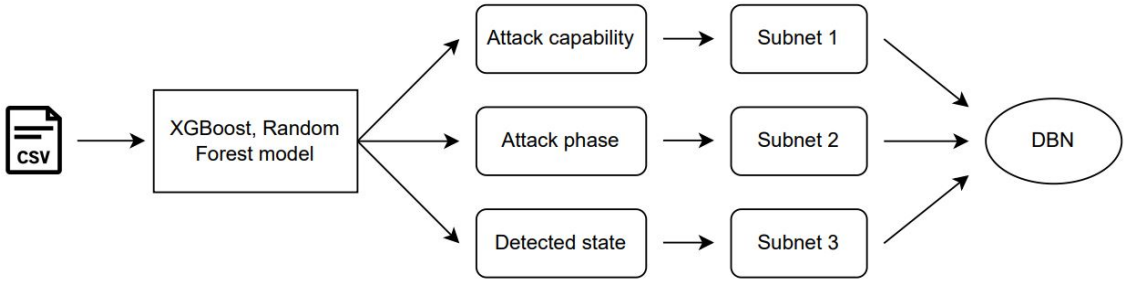
5.2.1 Vấn đề

Việc tích hợp hiệu quả dữ liệu thực tế về sự xuất hiện của tấn công APT với đầu vào mạng đánh giá rủi ro Bayes của hệ thống RiDX là rất quan trọng trong việc giám sát và phát hiện rủi ro ATTT. Dữ liệu đầu vào là dữ liệu đo lường luồng lưu lượng mạng gồm một số trường thông tin như thời lượng luồng, tổng số gói tin gửi đi, tổng số gói tin nhận về, tổng số bytes gửi đi,... Những trường thông tin này không có cơ sở kỹ thuật để ánh xạ trực tiếp với các biến/các công cụ đo xác định đầu vào cho subnet 1 và subnet 3 vào mạng đánh giá rủi ro được đề cập tại chương 3. Hơn nữa từ dữ liệu lưu lượng mạng đơn thuần cũng không thể ánh xạ trực tiếp với giai đoạn tấn công APT khi cuộc tấn công xảy ra. Bên cạnh đó bộ dữ liệu được xây dựng và gán nhãn về giai đoạn tấn công APT dựa trên góc nhìn của attacker, vì vậy việc dựa vào góc nhìn hệ thống để ánh xạ các trường thông tin trên với mạng đánh giá rủi ro là không hợp lý.

5.2.2 Giải pháp

Các tập dữ liệu Unraveled và SCVIC-APT-2021 là dữ liệu mô phỏng cuộc tấn công APT xảy ra đối với hệ thống và được gán nhãn tương ứng với giai đoạn tấn công và khả năng phát hiện tấn công của hệ thống. Vì vậy giải pháp được đưa ra nhằm tích hợp dữ liệu tấn công APT với mạng đánh giá rủi ro RiDX là sử dụng các mô hình học máy để phán đoán thông tin về các subnet đầu vào của mạng đánh giá rủi ro từ dữ liệu luồng lưu lượng mạng. Điều này phụ thuộc vào số lượng nhãn cung cấp bởi bộ dữ liệu đầu vào. Số lượng nhãn lớp của các tập dữ liệu hầu hết đều lớn hơn 2, vì vậy mô hình phân loại ở đây là phân loại đa lớp. Một số mô hình được sử dụng bao gồm Random-forest, XgBoost, GradientBoost, Decision-tree.

Luồng tích hợp dữ liệu thực tế với mạng đánh giá rủi ro được mô tả tại hình 5.1. Các mô hình học máy được sử dụng để phán đoán thông tin về giai đoạn tấn công, khả năng attacker và khả năng phát hiện tấn công của hệ thống từ dữ liệu giám sát đầu vào, kết quả thu được sẽ làm đầu vào cho mạng Bayes động và trả về cho người dùng các chỉ số giám sát rủi ro của hệ thống. Các mô hình sau khi huấn luyện được tích hợp với core-service tương ứng với một phần của quá trình đánh giá rủi ro.



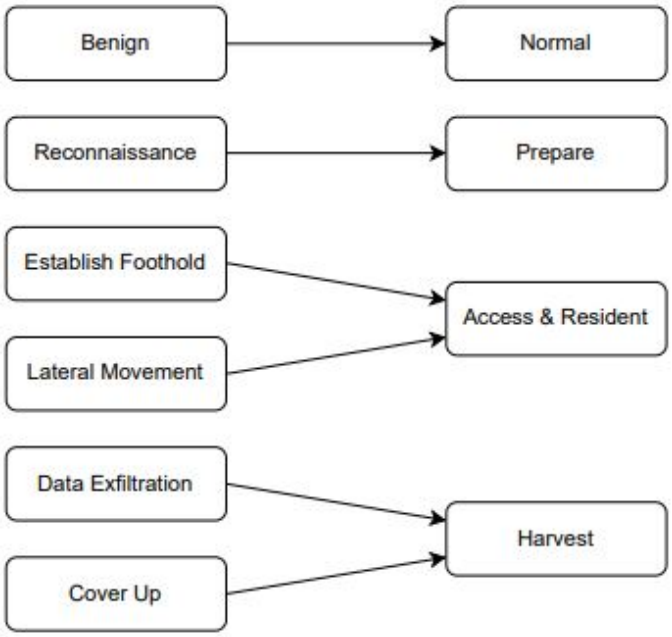
Hình 5.1: Luồng ánh xạ dữ liệu

5.2.3 Tích hợp bộ dữ liệu Unraveled

Bộ dữ liệu Unraveled cung cấp 4 loại nhãn là giai đoạn tấn công APT, kĩ thuật tấn công, khả năng phát hiện tấn công và tổ chức tấn công được đề cập tại chương 2. Trường dữ liệu kĩ thuật tấn công được gán nhãn dựa trên giai đoạn tấn công APT nên sẽ không được sử dụng. Ba loại nhãn còn lại được sử dụng để xây dựng các mô hình học máy phán đoán đầu vào của mạng đánh giá rủi ro. Việc mất cân bằng dữ liệu đã được đề cập khi nhãn ở trạng thái bình thường chiếm tỷ lệ rất lớn, lên đến 95% trên toàn bộ tập dữ liệu, vì vậy chỉ số F1 được sử dụng để đánh giá tính hiệu quả của các mô hình học máy. Trong quá trình xây dựng các mô hình học máy, một số bước tiền xử lý dữ liệu được áp dụng như xóa trùng lặp, thay thế các giá trị lỗi, bổ sung các giá trị còn thiếu và giảm chiều dữ liệu.

a, Giai đoạn tấn công

Đối với subnet 2, các thành phần về kịch bản triển khai, lỗ hổng bảo mật, biện pháp phòng vệ được người dùng xây dựng từ đầu, không thay đổi trong quá trình giám sát rủi ro. Yếu tố thay đổi thời gian thực ở đây là giai đoạn tấn công APT, cần xác định giai đoạn tấn công đầu vào cho subnet 2 tại mỗi bước giám sát. Bộ dữ liệu Unraveled cung cấp nhãn tương ứng với 6 giai đoạn tấn công bao gồm Benign, Reconnaissance, Establish-Foothold, Lateral-Movement, Data-Exfiltration và Cover-Up; hệ thống RiDX được thiết kế với 4 giai đoạn tấn công APT là Normal(Bình thường), Preparing (Chuẩn bị), Access & Resident (Xâm nhập và cư ngụ) và Harvest (Khai thác). Cần ánh xạ 6 giai đoạn tấn công của tập dữ liệu Unraveled với 4 giai đoạn tấn công APT của hệ thống RiDX. Chi tiết ánh xạ được mô tả tại hình 5.2.



Hình 5.2: Ánh xạ giai đoạn tấn công Unraveled

Sau bước ánh xạ các lớp giai đoạn tấn công như trên, bộ dữ liệu Unraveled sẽ được sử dụng để huấn luyện mô hình học máy phán đoán giai đoạn tấn công xảy ra đối với hệ thống. Quá trình huấn luyện thu được kết quả tai bảng 5.6. Kết quả cho thấy chỉ số macro-average-f1 của mô hình XGBoost đạt 89%, cao nhất trong số các mô hình sử dụng.

	GradientBoost	Decision Tree	Random Forest	XGBoost
Normal	0.99	0.99	1.00	1.00
Prepare	0.90	0.94	0.99	0.99
Access and Resident	0.64	0.81	0.9	0.94
Harvest	0.17	0.35	0.26	0.61
Weight average F1	0.99	0.99	0.99	0.99
Macro average F1	0.52	0.60	0.79	0.89

Bảng 5.6: Kết quả huấn luyện tập dữ liệu Unraveled với nhãn giai đoạn tấn công

b, Khả năng của attacker

Tương tự giai đoạn tấn công, bộ dữ liệu Unraveled cũng được sử dụng để huấn luyện mô hình phán đoán năng lực tấn công. Bộ dữ liệu Unraveled cung cấp nhãn về tổ chức tấn công bao gồm Benign (bình thường), AA (Amatuer), SH (Skilled-hacker) và APT (APT-group). Nhãn Benign tương ứng với trạng thái hệ thống chưa bị tấn công. Các nhãn AA, SH, APT tương ứng với các tổ chức tấn công được ghi

nhận, điều này có ý nghĩa trong việc xác định năng lực tấn công tương ứng với subnet 1 của mạng đánh giá rủi ro. Nhãn AA tương đương với năng lực tấn công ở mức thấp còn nhãn APT tương đương với năng lực tấn công của attacker ở mức cao nhất. Kết quả huấn luyện được hiển thị tại bảng 5.7. Từ bảng kết quả huấn luyện có thể thấy mô hình XGBoost đạt độ chính xác cao nhất với chỉ số macro-average-f1 đạt 91%.

	GradientBoost	Decision Tree	Random Forest	XGBoost
Benign	0.99	1.00	1.00	1.00
AA	0.60	0.68	0.82	0.85
APT	0.51	0,77	0.86	0.88
Weight average F1	0.99	0.99	0.99	0.99
Macro average F1	0.70	0.60	0.89	0.91

Bảng 5.7: Kết quả huấn luyện mô hình Unraveled với nhãn khả năng tấn công

Sau bước phán đoán năng lực chỉ số năng lực tấn công, việc ánh xạ với 5 nhân tố của mạng đánh giá rủi ro 1 vẫn là chưa có cơ sở. Vì vậy đồ án đề xuất mức điểm của từng tác nhân tương ứng với năng lực tấn công phán đoán. Chi tiết về điểm số của các nhân tố subnet 1 tương ứng với từng khả năng attacker được mô tả tại bảng 5.8. Điểm số của từng tác nhân càng cao thì khả năng của tác nhân đe dọa trong việc đáp ứng các yêu cầu của yếu tố đó càng lớn. Kết quả chỉ số năng lực tấn công của từng loại nhãn được mô tả dưới bảng.

No	Factor	Score		
		Normal	AA	APT
1	Specialist expertise	0	4.5	5.5
2	Knowledge of system	0	4.5	6.5
3	Required equipment and tools	0	3.5	6.5
4	Elapsed time	0	4.5	5.5
5	Window of opportunity	0	4	5
	Attack capability	**0%**	**11.04%**	**99.07%**

Bảng 5.8: Điểm số các nhân tố khả năng attacker tương ứng với các mức độ tấn công

c, Khả năng phát hiện tấn công của hệ thống

Subnet 3 của mạng đánh giá rủi ro bao gồm 5 node về khả năng phòng thủ của hệ thống. Tuy nhiên bộ dữ liệu Unraveled chỉ cung cấp nhãn phát hiện tấn công với 2 lớp là Normal và Detected. Vì vậy cần thiết lập 2 trạng thái subnet tương ứng với 2 khả năng phòng thủ của hệ thống. Các chỉ số về subnet 3 của hai trạng thái phát hiện tấn công của hệ thống được xác định tại bảng 5.9. Điểm của các nhân tố tại trạng thái Deteted cao hơn trạng thái Normal.

No	Factor	Score	
		Normal	Detected
1	Security Monitoring	5	5.5
2	Log Management	2.5	4.5
3	Vulnerability Management	5	7
4	Security Awareness Training	7.5	8.5
5	Incident Response Plan	5.5	6

Bảng 5.9: Điểm số các nhân tố khả năng phòng thủ của 2 trạng thái phát hiện và không phát hiện tấn công

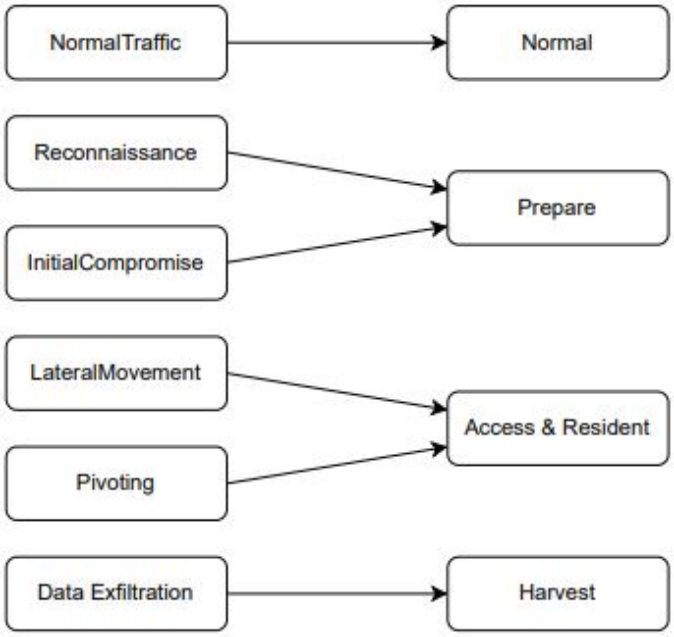
Kết quả huấn luyện mô hình phán đoán khả năng phát hiện tấn công cho tập dữ liệu Unraveled được hiển thị tại bảng 5.10. Giá trị độ đo weight average F1 là rất lớn, gần như đạt 100%. Điều này được lý giải là do hệ thống mô phỏng bộ dữ liệu Unraveled chỉ có thể phát hiện tấn công do attacker năng lực thấp (AA) gây ra nhờ các kỹ thuật bẫy được sử dụng như honeypot-server,... Mặc dù vậy, yếu tố về khả năng phát hiện tấn công của hệ thống vẫn là một yếu tố có giá trị tham khảo khi giám sát rủi ro sử dụng mạng Bayes động từ dữ liệu lưu lượng mạng.

	Random Forest	XGBoost
Normal	1.00	1.00
Detected	0.99	0.99
Weight average F1	0.99	0.99
Macro average F1	0.98	0.99

Bảng 5.10: Kết quả huấn luyện mô hình Unraveled với nhãn khả năng phát hiện tấn công

5.2.4 Tích hợp bộ dữ liệu SCVIC-APT-2021

Bộ dữ liệu SCVIC-APT-2021 chỉ cung cấp nhãn về giai đoạn tấn công APT, vì vậy chỉ có thể tích hợp vào subnet 2 của mạng đánh giá rủi ro. Tương tự bộ dữ liệu Unraveled, SCVIC-APT-2021 cũng bao gồm 6 giai đoạn tấn công APT và cần ánh xạ với 4 giai đoạn tấn công APT của hệ thống RiDX. Chi tiết ánh xạ được mô tả tại hình 5.3. Hai giai đoạn trinh sát (Reconnaissance) và thỏa hiệp (InitialCompromise) được ánh xạ với giai đoạn chuẩn bị tấn công (Prepare). Các giai đoạn chuyển động ngang (LateralMovement) và xoay vòng (Pivoting) tương ứng với giai đoạn xâm nhập (Access) và giai đoạn khai thác dữ liệu (Data exfiltration) tương ứng với cư ngụ (Resident).



Hình 5.3: Ánh xạ giai đoạn tấn công SCVIC-APT-2021

Sau bước ánh xạ trên, nhãn của bộ dữ liệu SCVIC-APT-2021 bao gồm 4 lớp tương ứng với 4 giai đoạn tấn công APT của RiDX. Một số mô hình học máy như Decision-tree, Gradient-boost, XGBoost và Random forest được sử dụng để huấn luyện mô hình phán đoán giai đoạn tấn công APT cho tập dữ liệu SCVIC-APT-2021. Tập dữ liệu Training được sử dụng trong quá trình huấn luyện và tập dữ liệu Testing được sử dụng để đánh giá kết quả. Kết quả kiểm tra với tập dữ liệu testing của SCVIC-APT-2021 được mô tả tại bảng 5.11. Trong số các mô hình được sử dụng, mô hình Random-forest cho kết quả tốt nhất với độ đo macro-average-f1 đạt 75%.

	GradientBoost	Decision Tree	XGBoost	Random Forest
Normal	0.99	0.99	1.00	1.00
Prepare	0.56	0.71	0.76	0.77
Access and Resident	0.73	0.78	0.85	0.82
Harvest	0.12	0.28	0.31	0.40
Weight average F1	0.99	0.99	0.99	0.99
Macro average F1	0.64	0.66	0.70	0.75

Bảng 5.11: Kết quả huấn luyện tập dữ liệu SCVIC-APT-2021 với nhãn giai đoạn tấn công

5.2.5 Kết quả

Đối với các mô hình huấn luyện của bộ dữ liệu Unraveled, sử dụng mô hình XGBoost cho kết quả tốt nhất trong cả 3 mô hình huấn luyện. Còn với bộ dữ liệu SCVIC-APT-2021 mô hình Random-forest cho kết quả tốt nhất. Các mô hình sẽ được lưu dưới dạng file pkl và tích hợp vào core-service để xây dựng bước phán đoán dữ liệu đầu vào cho mạng đánh giá rủi ro.

5.3 Xây dựng chức năng giám sát rủi ro tấn công APT

5.3.1 Vấn đề

Chức năng giám sát rủi ro tấn công APT cần đáp ứng yêu cầu tích hợp và giám sát dữ liệu thời gian thực. Yêu cầu về chức năng là cần có độ trễ thấp, api phải có khả năng xử lý kết quả đánh giá và phản hồi thông tin trong thời gian ngắn nhất có thể để giám sát thông tin rủi ro một cách chính xác. Bên cạnh đó chức năng cần phải có khả năng xử lý đồng thời mạnh mẽ nhằm đáp ứng số lượng yêu cầu giám sát cao.

Màn hình chức năng giám sát rủi ro cần hiển thị đầy đủ thông tin về trạng thái hệ thống và có cơ chế cảnh báo khi phát hiện rủi ro tại mỗi bước giám sát. Các thông tin cần hiển thị bao gồm thời gian giám sát, dữ liệu giám sát đầu vào tại mỗi bước, trạng thái hệ thống, giai đoạn tấn công APT (nếu có) và khả năng phát hiện tấn công của hệ thống. Kết quả giám sát còn cần biểu diễn thông tin về mức độ rủi ro của hệ thống, mức độ rủi ro trên mỗi tài sản và khả năng xảy ra tấn công đối với hệ thống.

5.3.2 Giải pháp

Giải pháp tổng thể cho chức năng giám sát rủi ro tấn công APT thời gian thực được trình bày như sau. Chức năng cho phép người dùng nhập file dữ liệu giám sát từ các tập dữ liệu tấn công APT (Unraveled, SCVIC-APT-2021) với định dạng

CSV. Hệ thống đọc từng dòng dữ liệu của file để tiến hành giám sát rủi ro, dữ liệu mỗi dòng tương ứng với dữ liệu thu thập từ nhật ký lưu lượng mạng trong quá trình vận hành hệ thống. Với mỗi bước giám sát, đầu tiên hệ thống áp dụng các mô hình học máy để phát hiện các yếu tố về xuất hiện tấn công APT như giai đoạn tấn công, khả năng tấn công và khả năng phát hiện tấn công. Sau đó kết quả thu được sẽ được sử dụng làm đầu vào mạng đánh giá rủi ro và trả về kết quả giám sát. Màn hình người dùng hiển thị kết quả giám sát bao gồm thông tin về mức độ rủi ro, khả năng xảy ra rủi ro và trạng thái hệ thống tại mỗi bước giám sát được đề cập tại chương 3.

a, Xây dựng API websocket phục vụ giám sát dữ liệu thời gian thực

Việc xử lý dữ liệu thời gian thực cần đáp ứng các yêu cầu về kết nối và độ trễ. Websocket là giao thức cho phép thiết lập kết nối TCP hai chiều liên tục giữa client và server. Nó rất phù hợp trong việc xử lý dữ liệu cập nhật thời gian thực. Một số ưu điểm của websocket bao gồm:

1. Kết nối hai chiều (Bidirectional): Client và server đều có thể gửi tin nhắn đến nhau bất cứ lúc nào mà không cần phải đợi yêu cầu từ bên kia.
2. Hiệu suất cao (Performance): WebSocket duy trì một kết nối duy nhất mở, điều này giảm thiểu chi phí thiết lập kết nối nhiều lần và giảm độ trễ so với HTTP.
3. Tiết kiệm băng thông (Bandwidth Efficiency): Bởi vì không cần gửi các header HTTP trong mỗi tin nhắn, WebSocket tiết kiệm băng thông hơn so với các giao thức HTTP thông thường.

Chức năng giám sát rủi ro tấn công APT sử dụng giao thức websocket kết nối từ client đến core-service. Dữ liệu giám sát được truyền tải đến core-service thông qua giao thức sẽ làm đầu vào cho các mô hình học máy để dự đoán các thông tin đầu vào mạng đánh giá rủi ro. Quá trình thử nghiệm đo được thời gian tính toán mạng đánh giá rủi ro và phản hồi kết quả đánh giá rủi ro trung bình là 0.532s, vì vây đồ án cấu hình khoảng thời gian giữa 2 bước giám sát là 2s để kết quả đánh giá rủi ro là chính xác nhất.

b, Trực quan thông tin giám sát rủi ro

Màn hình giám sát rủi ro tấn công APT bao gồm:

1. Thông tin chung: thông tin về thời gian giám sát, số lần giám sát, trạng thái tấn công (nếu có) và trạng thái phát hiện tấn công
2. Thông tin dữ liệu thu thập: thông số đo lường lưu lượng mạng bao gồm IP nguồn, IP đích, tổng số gói tin gửi đi, tổng số gói tin nhận về, tổng số bytes gửi đi,...

3. Biểu đồ trạng thái hệ thống: Biểu đồ biểu diễn trạng thái hệ thống tại mỗi bước giám sát, được đề cập tại chương 3.1.1.
4. Các biểu đồ giám sát rủi ro hệ thống: Các biểu đồ bao gồm biểu đồ chỉ số severity, biểu đồ chỉ số likelihood, biểu đồ rủi ro trên mỗi tài sản và biểu đồ năng lực attacker - khả năng phòng thủ hệ thống.

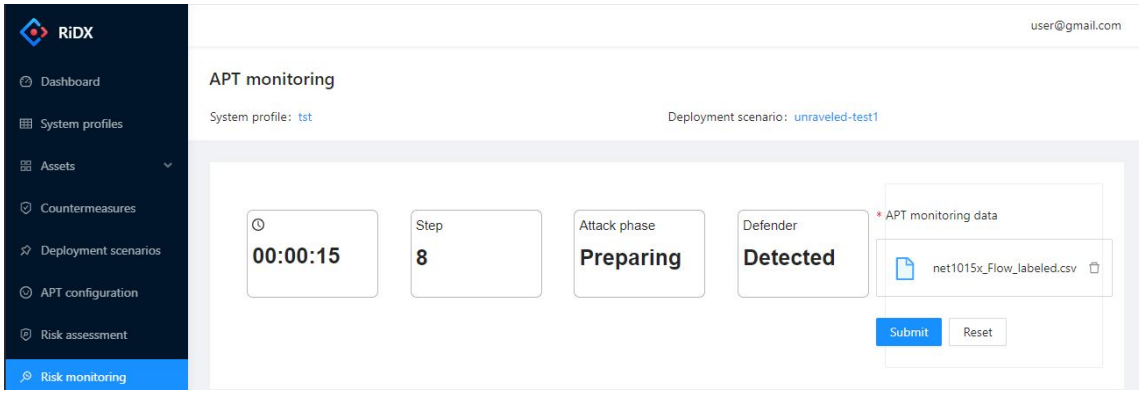
Dữ liệu tại các biểu đồ thông tin sẽ được cập nhật thời gian thực theo kết quả giám sát rủi ro

5.3.3 Kết quả

Dưới đây là hình ảnh về màn hình chức năng giám sát rủi ro tấn công APT.

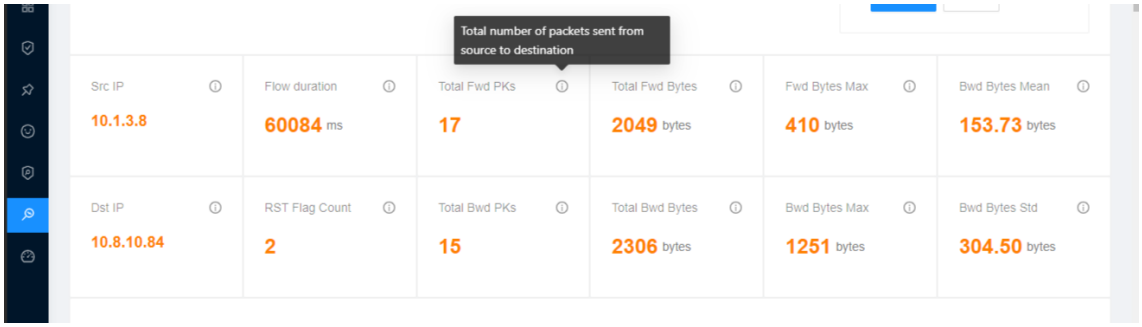
a, Các màn hình dashboard

Màn hình 5.4 cho phép người dùng import dữ liệu đầu vào dưới dạng CSV. Hệ thống đọc từng dòng tương ứng với dữ liệu tại từng bước giám sát và tiến hành đánh giá rủi ro. Màn hình hiển thị thời gian giám sát, số lần giám sát, trạng thái tấn công xuất hiện và trạng thái phòng thủ của hệ thống. Trạng thái tấn công thay đổi với 4 trạng thái tấn công của hệ thống RiDX còn trạng thái phòng thủ hiển thị thông tin về trạng thái phát hiện tấn công (Normal, Detected).



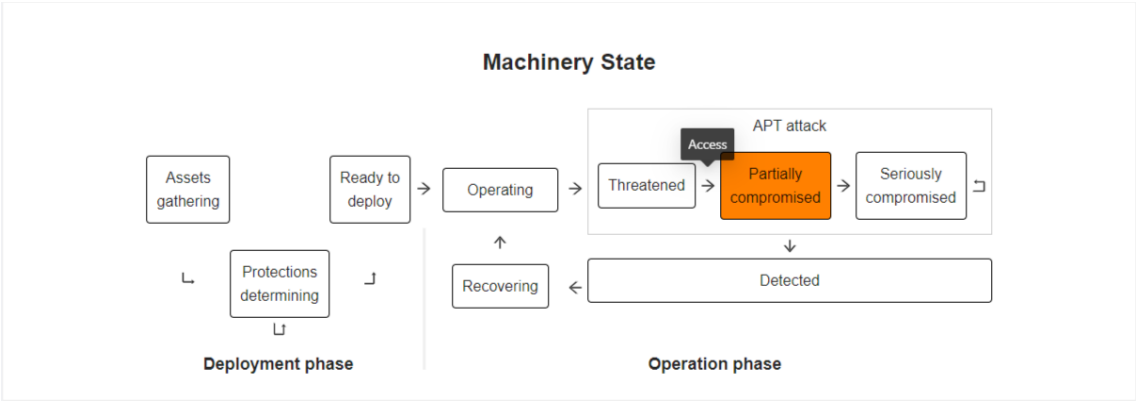
Hình 5.4: Màn hình dashboard giám sát rủi ro

Màn hình 5.5 hiển thị các thông tin về dữ liệu đầu vào tại mỗi bước giám sát. Dữ liệu bao gồm một số trường thông tin quan trọng của dữ liệu lưu lượng mạng như thời lượng lưu lượng (flow-durations), số packet gửi đi, số packet nhận về, tổng số bytes gửi đi... Màu sắc của các giá trị thay đổi tương ứng với mức độ rủi ro của hệ thống.



Hình 5.5: Màn hình logging dữ liệu đầu vào

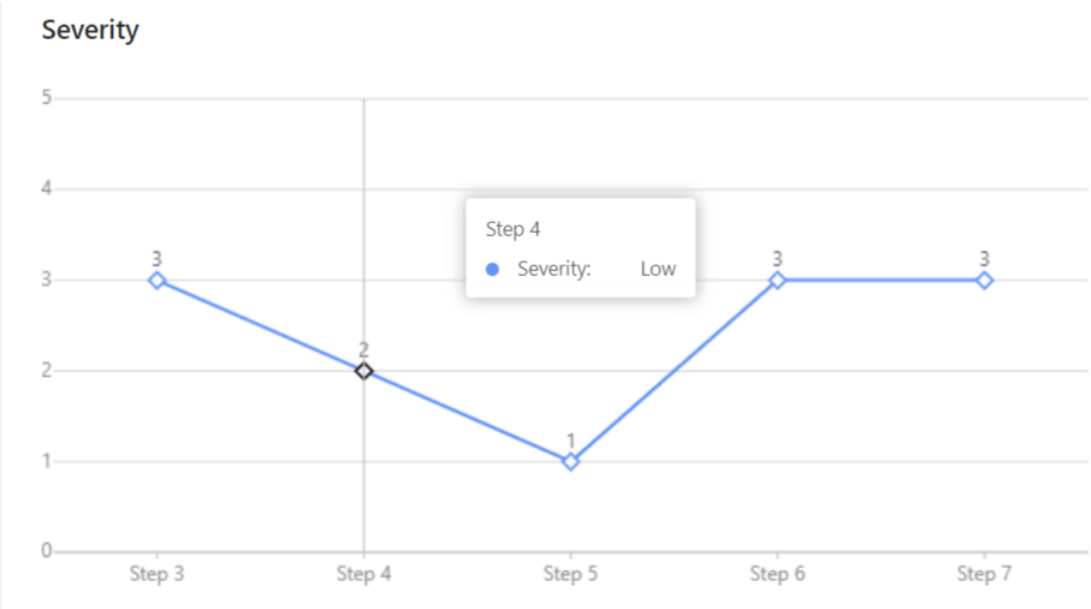
Màn hình 5.6 biểu diễn trạng thái của hệ thống tại mỗi bước giám sát. Trạng thái mặc định của hệ thống là Operating (Triển khai). 3 trạng thái của hệ thống dưới cuộc tấn công APT tương ứng với 3 giai đoạn tấn công APT. Trạng thái Detected sẽ được hiển thị nếu hệ thống phát hiện tấn công. Nếu hệ thống không phát hiện có dấu hiệu tấn công, trạng thái sẽ quay trở lại là Operating.



Hình 5.6: Màn hình biểu diễn trạng thái máy

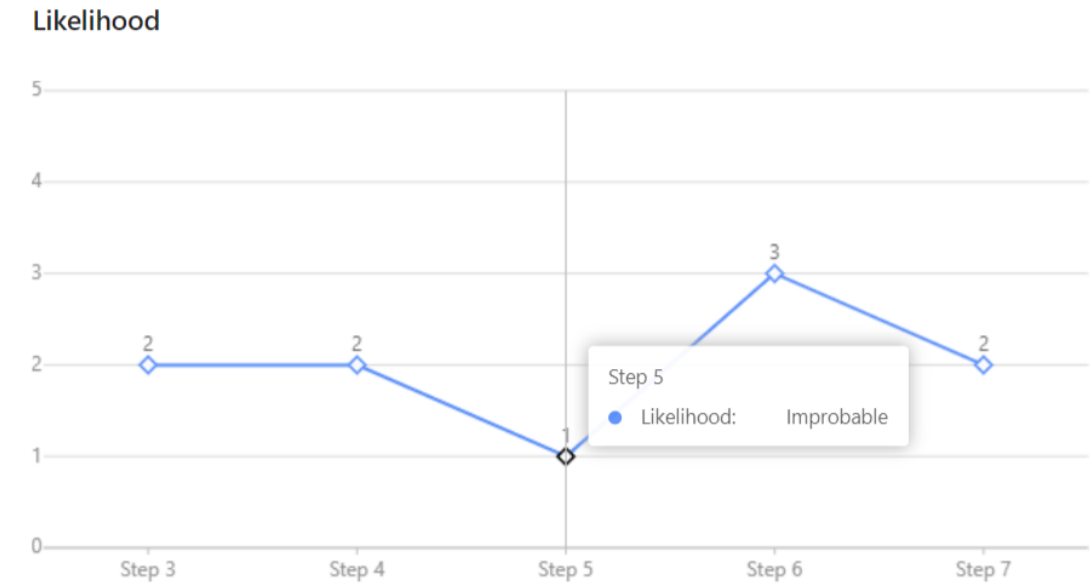
b, Các màn hình giám sát rủi ro

Màn hình 5.7 biểu diễn thông tin về mức độ rủi ro của hệ thống tại mỗi bước giám sát. Các mức độ rủi ro bao gồm: 1 - Negligible, 2 - Low, 3 - Moderate, 4 - Significant và 5 - Catastrophic.



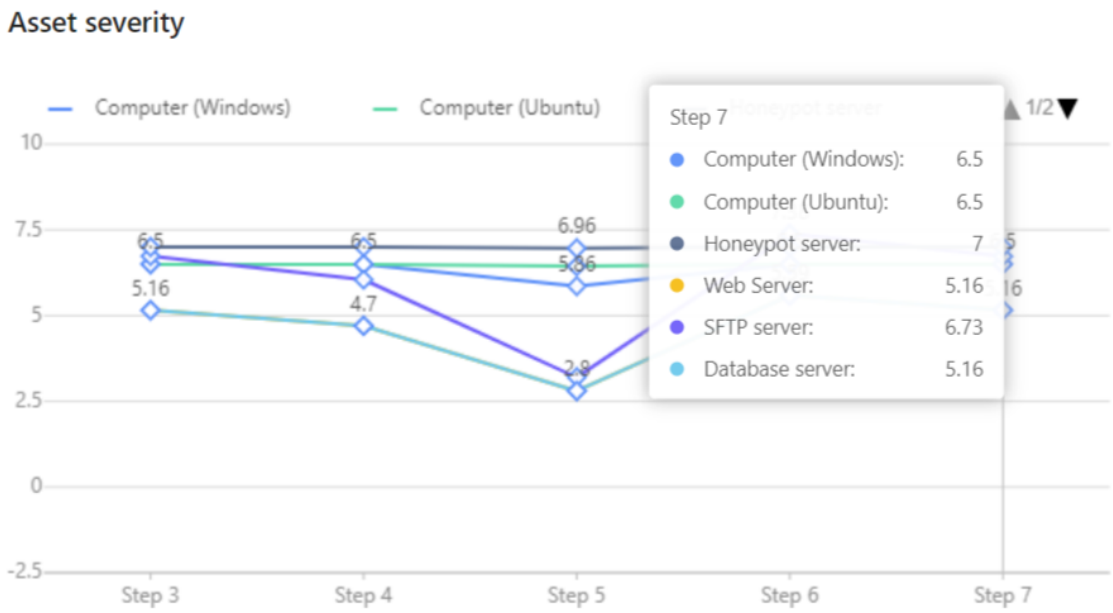
Hình 5.7: Màn hình giám sát mức độ rủi ro - Severity

Màn hình 5.8 biểu diễn thông tin về khả năng xảy ra rủi ro của hệ thống tại mỗi bước giám sát. Các mức độ khả năng xảy ra rủi ro bao gồm: 1 - Improbable, 2 - Remote, 3 - Occasional, 4 - Proable và 5 - Frequent.



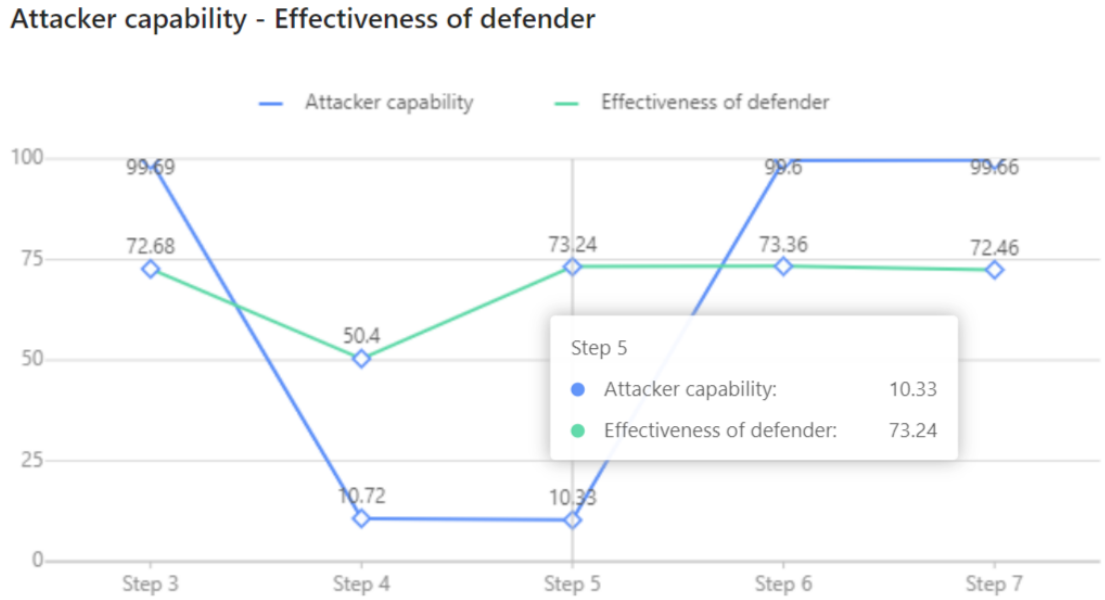
Hình 5.8: Màn hình giám sát khả năng xảy ra rủi ro - Likelihood

Màn hình 5.9 biểu diễn thông tin về mức độ rủi ro trên mỗi tài sản hệ thống tại mỗi bước giám sát.



Hình 5.9: Màn hình giám sát mức độ rủi ro trên mỗi tài sản hệ thống

Màn hình 5.10 biểu diễn thông tin về năng lực attacker và khả năng phòng thủ của hệ thống thông qua tỷ lệ phần trăm mức độ hiệu quả tấn công và phòng thủ.



Hình 5.10: Màn hình giám sát năng lực tấn công và khả năng phòng thủ

CHƯƠNG 6. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

6.1 Kết luận

Trong quá trình xây dựng đồ án, tôi đã xây dựng chức năng giám sát rủi ro tấn công APT tích hợp với dữ liệu thời gian thực. Tuy còn một số điểm chưa hoàn thiện, chức năng đã đáp ứng được yêu cầu xử lý dữ liệu thời gian thực và tích hợp dữ liệu đầu vào với mạng đánh giá rủi ro và hiển thị kết quả giám sát thời gian thực với người dùng. Một số công việc chưa hoàn thiện là cập nhật đồ thị tấn công thời gian thực với attacker và tài sản bị tấn công.

Qua quá trình thực hiện đồ án, tôi đã được tiếp thu nhiều kiến thức bổ ích về ATTT, triển khai hệ thống,... Bên cạnh đó tôi cũng nâng cao được kĩ năng về lập trình và tìm hiểu vấn đề.

6.2 Hướng phát triển

Các công việc được thực hiện trong đồ án:

- Xây dựng chức năng giám sát dữ liệu thời gian thực với dữ liệu đầu vào là dữ liệu chuyên gia về tấn công APT.
- Ánh xạ dữ liệu đầu vào với mạng đánh giá rủi ro: Sử dụng các mô hình học máy để phán đoán đầu vào mạng đánh giá rủi ro.
- Trực quan hóa thông tin giám sát: Hiển thị thông tin giám sát rủi ro tấn công APT bao gồm trạng thái hệ thống, chỉ số Severity và Likelihood, giai đoạn tấn công và khả năng phòng thủ.

Lỗ hổng bảo mật của hệ thống không chỉ là các lỗ hổng CVE được phát hiện mà còn từ hành vi con người (truy cập link độc hại, đặt mật khẩu dễ đoán,...). Vì vậy RiDX được kỳ vong xây dựng bổ sung các cơ chế đánh giá rủi ro từ các mối nguy cơ chưa được xác định đó. Bên cạnh đó cần bổ sung các cơ chế về kĩ thuật tấn công trong việc xác định rủi ro khi có tấn công APT xảy ra. Ngoài ra kịch bản triển khai cần được xây dựng chi tiết hơn đáp ứng nhu cầu với những hệ thống lớn phức tạp hiện nay.

TÀI LIỆU THAM KHẢO

[1] V. T. H. Giang and N. M. Tuan, "Application of bayesian network in risk assessment for website deployment scenarios," *Journal of Science and Technology on Information security*, vol. 2, no. 14, pp. 3–17, 2021.

[2] M. T. Nguyen and T. H. G. Vu, "A review of cyber security risk assessment for web system during its deployment and operation," *Journal on Information Technologies & Communications*, pp. 32–45, 2023.

[3] J. Liu, Y. Shen, M. Simsek, *et al.*, "A new realistic benchmark for advanced persistent threats in network traffic," *IEEE Networking Letters*, vol. 4, no. 3, pp. 162–166, 2022.

[4] S. Myneni, K. Jha, A. Sabur, *et al.*, "Unraveled—a semi-synthetic dataset for advanced persistent threats," *Computer Networks*, vol. 227, p. 109 688, 2023.

[5] T.-H.-G. Vu, T.-H. Hoang, and M.-T. Nguyen, "Assessing web security risks using dynamic bayesian network," in *Proceedings of the 11th International Symposium on Information and Communication Technology*, 2022, pp. 165–172.

[6] K. P. Murphy, "Dynamic bayesian networks: Representation, inference and learning, dissertation," *PhD thesis, UC Berkley, Dept. Comp. Sci*, 2002.

[7] *React — reactjs.org*, `https://reactjs.org`.

[8] *Node.js — Run JavaScript Everywhere — nodejs.org*, `https://nodejs.org/en`.

[9] *Documentation | NestJS - A progressive Node.js framework — docs.nestjs.com*, `https://docs.nestjs.com/`.

[10] *Angular — angular.dev*, `https://angular.dev/`.

[11] *FastAPI — fastapi.tiangolo.com*, `https://fastapi.tiangolo.com/`.

[12] *Welcome to Python.org — python.org*, `https://www.python.org/`.

[13] *MongoDB: The Developer Data Platform — mongodb.com*, `https://www.mongodb.com/`.

[14] *MITRE ATT&CK&xAE; — attack.mitre.org*, `https://attack.mitre.org/`.