HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Ngô Văn Nhận

NGHIÊN CỨU PHƯƠNG PHÁP PHÁT HIỆN TẤN CÔNG DDOS TRONG MẠNG DI ĐỘNG 5G

CHUYÊN NGÀNH : HỆ THỐNG THÔNG TIN
MÃ SỐ: 8.48.01.04

TÓM TẮT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

HÀ NỘI – 2025

Đề án tốt nghiệp được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: TS. Nguyễn Đình Hóa ……

(Ghi rõ học hàm, học vị)

Phản biện 1: ……

Phản biện 2: ...……..

Đề án tốt nghiệp sẽ được bảo vệ trước Hội đồng chấm đề án tốt nghiệp thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: ..... giờ ..... ngày ..... tháng .... năm .....

Có thể tìm hiểu đề án tốt nghiệp tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông.

MỞ ĐẦU

1. Lý do chọn đề tài

Trong thời đại hiện nay, khi công nghệ thông tin và truyền thông phát triển mạnh mẽ, khái niệm Mạng di động thế hệ mới, mạng 4G, 5G ngày càng trở lên quan trọng và phổ biến. Mạng di động thế hệ mới 4G, 5G mang đến rất nhiều lợi ích và là yêu cầu tất yếu trong công cuộc Chuyển đổi số và sự phát triển mạnh mẽ của khoa học công nghệ hiện đại.

Để nhanh chóng áp dụng các yêu cầu đảm bảo an toàn thông tin cho mạng di động 5G vào thực tế, tạo điều kiện thuận lợi tối đa cho việc thương mại hóa công nghệ 5G một cách an toàn, tin cậy tại Việt Nam, Bộ Khoa học và Công nghệ yêu cầu các doanh nghiệp nghiên cứu, áp dụng các yêu cầu bảo đảm an toàn thông tin cho mạng 5G trong quá trình lựa chọn thiết bị 5G; xây dựng và thiết lập mạng 5G; tiếp tục nghiên cứu, phát triển các yêu cầu đảm bảo an toàn thông tin cho mạng 5G phù hợp với tình hình thực tế và sự phát triển của khoa học và công nghệ. Do đó, việc nghiên cứu kiến trúc mạng 5G, nguy cơ đe dọa an toàn bảo mật và các giải pháp bảo mật trong mạng di động 5G là cấp thiết giúp bắt kịp xu thế công nghệ mới, sớm đưa mạng 5G vào các hoạt động thực tiễn.

Để bảo vệ mạng 5G, một số cuộc tấn công nhất định phải được giải quyết; đáng chú ý nhất là các cuộc tấn công DDoS. Một cuộc tấn công DDoS trong mạng di động hiện tại (ví dụ: 4G) chỉ có thể xâm phạm một dịch vụ. Tuy nhiên, trong mạng 5G, nếu một hacker độc hại nắm quyền kiểm soát một phần và thực hiện một cuộc tấn công DDoS, điều này có thể làm tổn hại đến các dịch vụ thuộc cùng một mạng ảo. Các hệ thống IoT đang phải đối mặt với sự gia tăng bề mặt tấn công theo cấp số nhân do số lượng lớn các thiết bị IoT có thể bị xâm phạm [2]. Việc theo dõi, phát hiện sớm các truy cập bất thường sẽ giúp ngăn chặn các cuộc tấn công DDoS và hạn chế hậu quả do chúng gây ra.

2. Mục đích nghiên cứu

Mục tiêu của đề án là nghiên cứu tìm hiểu xoay quanh việc xây dựng các giải pháp kỹ thuật nhằm đảm bảo an toàn mạng trước các cuộc tấn công từ chối dịch vụ phân tán (DDoS), đặc biệt trong môi trường mạng hiện đại như mạng di động 5G. Tìm ra các biện pháp tối ưu trong việc phát hiện sớm và chính xác các cuộc tấn công DDoS, tối ưu hóa hiệu suất phát hiện trong môi trường 5G, ứng dụng trí tuệ nhân tạo và học máy.

3. Đối tượng và phạm vi nghiên cứu

Đề án tập trung nghiên cứu xoay quanh các đối tượng:

- Mạng di động 5G và các quy cơ đe dọa về tấn công mạng di động; tấn công DDoS trong mạng di động 5G;

- Các phương pháp học máy, học sâu đã được công bố, có thể nhận diện các truy cập bất thường;

- Các công cụ giám sát, ngăn chặn tấn công mạng, đảm bảo an toàn bảo mật cho hệ thống mạng di động.

Phạm vi nghiên cứu:

- Nghiên cứu kiến trúc tổng quan Core network của mạng di động 5G trong phạm vi thuộc lĩnh vực hệ thống thông tin, các thiết bị mạng tham gia và mô hình hoạt động trong mạng di động;

- Các mô hình, giải pháp, thuật toán đã được công bố: KNN, SVM, MLP... sử dụng tập dữ liệu đối sánh 5G-NIDD [3].

4. Phương pháp nghiên cứu

Đề án sử dụng kết hợp các phương pháp nghiên cứu sau:

- Nghiên cứu lý thuyết: Tiến hành khảo sát, tìm hiểu các nghiên cứu về kiến trúc mạng di động nói chung và mạng di động 5G nói riêng, từ đó phát hiện, đánh giá các nguy cơ về an toàn bảo mật trong mạng di động 5G. Tìm hiểu các mô hình học máy, học sâu có thể áp dụng vào nhận diện các truy cập bất thường, cụ thể là phát hiện các cuộc tấn công DDoS.

- Nghiên cứu thực nghiệm: Sử dụng các công cụ giám sát mạng để theo dõi, thu thập các gói tin lan truyền trong mạng di động private 5G, áp dụng các mô hình học máy trên tập dữ liệu đối sánh để phân lớp các gói tin bình thường và gói tin bất thường. Kết luận về việc tấn công, đánh giá kết quả thu được.

CHƯƠNG 1: KIẾN TRÚC MẠNG DI ĐỘNG 5G VÀ AN TOÀN BẢO MẬT TRONG MẠNG DI ĐỘNG 5G

1.1. Tổng quan về mạng di động 5G

1.1.1. Kiến trúc mạng 5G

Mạng di động thế hệ thứ năm – mạng di động 5G là thế hệ tiếp theo của công nghệ truyền thông di động sau thế hệ thứ bốn (4G). Theo các nhà phát triển, mạng 5G có tốc độ nhanh hơn khoảng 100 lần so với mạng 4G hiện tại, giúp mở ra nhiều hứa hẹn về các dịch vụ mới và tiện ích. So sánh với công nghệ 3G, 4G, mạng 5G có độ trễ cực thấp (trong khoảng 5-20 ms) và có tham vọng cải tiến tối ưu, đẩy độ trễ xuống chỉ còn 1 ms [1]. Điều này có ý nghĩa rất quan trọng trong việc ứng dụng công nghệ 5G vào một số lĩnh vực đòi hỏi độ trễ của tín hiệu cực thấp như xe tự lái, máy bay tự lái, phẫu thuật từ xa… Sự phát triển của mạng 5G với hệ thống Internet vạn vật (IoT) được thiết lập để cải thiện khả năng liên lạc đáng tin cậy và kết nối ổn định. Công nghệ truy cập vô tuyến mới 5G có độ trễ thấp, tính sẵn sàng cao và tốc độ vượt trội; tất cả đều cần thiết cho các hệ thống IoT [4, 5].

Kiến trúc mạng 5G là một sự khác biệt căn bản so với các thế hệ trước, ở chỗ nó được đặc trưng bởi tính linh hoạt và khả năng thích ứng cao hơn. Không giống như các thế hệ mạng di động trước đó, chủ yếu được xây dựng xung quanh một cấu trúc mạng cứng nhắc, phân cấp, kiến trúc 5G được thiết kế để có tính linh hoạt cao, phân tán và được xác định bằng phần mềm. Nó kết hợp một loạt các khái niệm và kỹ thuật, chẳng hạn như phân chia mạng, điện toán biên và SDN, cùng nhau góp phần nâng cao hiệu suất, chức năng và trải nghiệm người dùng. Thiết kế của kiến trúc mạng 5G giải quyết một số yêu cầu và thách thức do một loạt các ứng dụng mang lại, từ băng thông rộng di động tốc độ cao đến các giao tiếp cực kỳ đáng tin cậy, độ trễ thấp. Tính linh hoạt và khả năng thích ứng của cấu trúc 5G bắt nguồn từ sự chuyển đổi sang một phương pháp tiếp cận hướng đến phần mềm hơn từ phương pháp tiếp cận phần cứng truyền thống, cho phép mức độ linh hoạt, khả năng mở rộng và hiệu quả chưa từng thấy ở các thế hệ trước.

Đặc điểm cốt lõi của cấu trúc mạng 5G là việc sử dụng kiến trúc phẳng, phi tập trung thay vì kiến trúc phân cấp truyền thống, để giảm độ trễ và tối ưu hóa định tuyến lưu lượng. Việc sử dụng kiến trúc phi tập trung cũng cho phép tốc độ thông lượng cao hơn, cho phép truyền dữ liệu nhanh hơn [6]. Để đối phó với nhu cầu về tốc độ dữ liệu cao và khả năng kết

nối lớn cần thiết cho các thiết bị IoT, 5G cũng bao gồm việc sử dụng các công nghệ ăng-ten tiên tiến, chẳng hạn như MIMO. Massive MIMO tăng dung lượng của một ô bằng cách sử dụng một số lượng lớn ăng-ten truyền tại trạm gốc để phục vụ một số lượng đáng kể người dùng trong cùng một tài nguyên tần số thời gian. Điều quan trọng là mạng 5G triển khai kiến trúc dựa trên đám mây, nơi các chức năng và dịch vụ có thể được khởi tạo theo cách linh hoạt và năng động. Điều này dẫn đến một sự thay đổi mô hình từ cơ sở hạ tầng phần cứng chuyên dụng sang một môi trường phần mềm linh hoạt hơn, nâng cao khả năng thích ứng và khả năng mở rộng của các dịch vụ mạng.

1.1.2. Các tính năng và công nghệ hỗ trợ

- Giao tiếp D2D (Device-to-Device):
- MIMO khổng lồ tiên tiến (Advance Massive MIMO):
- Mạng được xác định bằng phần mềm (Software Defined Network - SDN):
- Ảo hóa chức năng mạng (Network Function Virtualization - NFV):
- Điện toán biên (Edge computing):
- Truyền thông kiểu máy quy mô lớn (mMTC):

1.2. Nguy cơ đe dọa an toàn bảo mật trong mạng di động 5G

Dù có những tiến bộ về kiến trúc, 5G cũng bộc lộ những bề mặt tấn công mới. Các giao diện lập trình động (API) và chức năng ảo hóa như NFV và SDN gia tăng độ linh hoạt nhưng đồng thời cũng là mục tiêu hấp dẫn cho kẻ tấn công. Các tín hiệu truy cập ban đầu như SSB, MIB và SIB1 không được mã hóa, dễ bị giả mạo hoặc can thiệp, tạo điều kiện cho các cuộc tấn công chiếm quyền truy cập mạng ngay từ giai đoạn đồng bộ hóa. Ngoài ra, mô hình phân chia mạng (network slicing) cũng tiềm ẩn rủi ro khi cấu hình lát mạng không chuẩn hoặc thiếu cách ly nghiêm ngặt, tạo điều kiện cho các tấn công chéo lát (cross-slice attack), nơi tin tặc có thể xâm nhập một lát có bảo mật yếu như lát IoT rồi di chuyển sang các lát mạng khác.

Ngoài việc cung cấp hiệu suất mạng, các hệ thống IoT hỗ trợ 5G cũng cần duy trì tính bảo mật và cải thiện độ tin cậy của dịch vụ. Một cuộc tấn công được thực hiện thành công vào mạng 5G có thể dẫn đến các hoạt động không mong muốn và gây ra hậu quả nghiêm trọng. Những tin tặc đang sử dụng các chiến thuật mới để kiếm tiền từ các cuộc tấn công của chúng bằng cách kiểm soát dữ liệu nhạy cảm, yêu cầu tiền chuộc hoặc khiến mạng không hoạt động.

Mối đe dọa về an toàn bảo mật trong mạng di động 5G cần chú ý:

1.2.1. Xác thực (Authentication)

Xác thực là một khái niệm quan trọng trong bảo mật mạng 5G vì nó cho phép xác minh danh tính người dùng trong mạng. Nhiều cơ chế được sử dụng trong mạng để xác thực dữ liệu. Nó được chia thành hai phần: xác thực chính và xác thực phụ. Xác thực chính cho phép xác thực lẫn nhau của các thiết bị và mạng trong cả mạng 5G và 4G. Tuy nhiên, xác thực chính trong mạng dựa trên 5G có một số vấn đề, bao gồm kiểm soát kiến thức và việc gọi xác thực thiết bị không được hỗ trợ đầy đủ.

1.2.2. Tính bảo mật (Confidentiality)

Tính bảo mật đảm bảo rằng chỉ người được ủy quyền mới có thể truy cập thông tin về người gửi. Khóa cần thiết cho nút gốc thế hệ tiếp theo thứ cấp được tạo và cung cấp bởi nút gốc chính trước bất kỳ quá trình truyền vô tuyến an toàn nào. Tín hiệu cho kiểm soát tài nguyên vô tuyến có thể được gửi giữa nút gốc thế hệ tiếp theo thứ cấp và thiết bị người dùng.

1.2.3. Tính sẵn sàng (Availability)

Các dịch vụ thông minh dựa trên mạng di động 5G được hưởng lợi từ các tài nguyên đám mây, giúp phát triển cơ sở hạ tầng hiệu quả về chi phí. Tuy nhiên, chúng tiềm ẩn các mối đe dọa bảo mật như tấn công mạng gây ảnh hưởng tới độ tin cậy của mạng. Các cuộc tấn công DDoS yêu cầu tài nguyên vật lý và logic ở cấp độ biên và đám mây, điều này có ảnh hưởng đến các quá trình phân chia mạng. Các cuộc tấn công gây nhiễu gây ra sự cố trên các cơ sở truy cập vô tuyến, ngăn người dùng truy cập các dịch vụ di động. Các cuộc tấn công vào tài nguyên 5G, bao gồm hệ thống hỗ trợ, mặt phẳng điều khiển và vô tuyến, có thể làm gián đoạn các dịch vụ mạng thông minh.

1.2.4. Tính chống chối bỏ (Non-repudiation)

Khả năng chứng minh tính hợp lệ và độ tin cậy của một thông điệp hoặc giao dịch và ngăn người gửi từ chối sự tham gia của họ vào giao tiếp được gọi là tính chống chối bỏ. Việc người dùng từ chối không thể bị ngăn chặn chỉ bằng xác thực. Tuy nhiên, việc phân biệt giữa những người dùng hoặc thiết bị khác nhau là rất quan trọng để tạo ra dữ liệu an toàn, nên xác thực là cần thiết để đảm bảo tính chống chối bỏ. Các mạng 5G có thể sử dụng chữ ký số để đảm bảo tính chống chối bỏ. Chữ ký số sử dụng một kỹ thuật mật mã để tạo ra một mã duy nhất được thêm vào thông điệp, có thể được sử dụng để chứng minh tính xác thực và tính toàn vẹn của tin nhắn. Điều này gây khó khăn cho người gửi trong việc từ chối sự tham gia của họ vào giao tiếp.

1.2.5. Tính toàn vẹn (Integrity)

Một thước đo bảo mật trong mạng 5G là bảo vệ tính toàn vẹn của mặt phẳng người dùng giữa nút mạng và các thiết bị. Bảo vệ tính toàn vẹn là một tính năng tốn nhiều tài nguyên, các thiết bị IoT không thể triển khai với tỉ lệ dữ liệu cao do những hạn chế của chúng. Do đó, kiến trúc mạng thông minh dựa trên 5G phải bao gồm các giao thức đảm bảo tính toàn vẹn của mạng. Ví dụ, các mạng 5G có thể sử dụng các thuật toán mật mã như Tiêu chuẩn Mã hóa Tiên tiến (AES) và Thuật toán Băm An toàn (SHA) để tạo ra một mã duy nhất được thêm vào dữ liệu, có thể được sử dụng để xác minh tính toàn vẹn của dữ liệu. Điều này đảm bảo rằng dữ liệu không bị sửa đổi hoặc giả mạo trong quá trình truyền.

1.3. Một số hình thức tấn công

1.3.1. Tấn công DoS (Denial of Service - Từ chối dịch vụ)

Trong một cuộc tấn công từ chối dịch vụ (DoS), kẻ tấn công cố gắng gửi một lượng lớn dữ liệu giả mạo nhắm vào các nút của mạng. Số lượng lớn các yêu cầu này có thể tiêu tốn năng lượng, băng thông và thời gian,…, khiến nút bị treo hoặc không thể phản hồi các yêu cầu hợp lệ. Trong mạng 5G, kẻ tấn công còn khai thác các lỗ hổng trong giao thức mạng, hoặc lạm dụng các lỗ hổng bảo mật của hệ điều hành và các ứng dụng được tải xuống từ cửa hàng ứng dụng để can thiệp vào các thiết bị di động.

1.3.2. Tấn công MitM (Man-in-the-Middle)

1.3.3. Tấn công chiếm quyền điều khiển (Hijacking attacks)

1.3.4. Tấn công giả mạo (Spoofing attacks)

1.3.5. Tấn công gây nhiễu (Jamming attack)

1.3.6. Tấn công trạm gốc giả mạo (Rogue base station attacks)

1.3.7. Tấn công kênh bên (Side-channel attack)

1.4. Các phương pháp bảo mật hiện có

Để củng cố bảo mật của mạng 5G, vô số các giải pháp hiện tại và tiềm năng đang được khám phá. Các giải pháp này có nhiều lớp và hoạt động trong các khía cạnh khác nhau của mạng, chẳng hạn như các cấp độ kỹ thuật, tổ chức, quy định và người dùng. Các giải pháp kỹ

thuật tạo thành nền tảng của bảo mật 5G, vì chúng trực tiếp giải quyết các mối đe dọa và lỗ hổng khác nhau.

1.4.1. ENDER

ENDER là viết tắt của hệ thống tiền quyết định (*pre-dEcisioN*), quyết định nâng cao (*advance Decision*), và học hỏi (*lEaRning*) [8]. Hai phương pháp lý thuyết quyết định được sử dụng trong chiến lược CPS để xác định lưu lượng tấn công trên đám mây nhằm giảm thiểu các cuộc tấn công lưu lượng HX-DoS và SIPDAS. Tiếp theo là một quy trình tương tự như của một hệ thống phát hiện xâm nhập tiêu chuẩn. Do đó, nó có khả năng xác định và đánh dấu một tin nhắn tấn công. Khi một tin nhắn tấn công SIPDAS hoặc HX-DoS được phát hiện, tin nhắn đó được đánh dấu bằng một bit duy nhất. Các thuật toán sẽ xóa những tin nhắn này khỏi hệ thống.

1.4.2. Thuật toán dựa trên học máy (Machine learning)

Việc triển khai một hệ thống xác thực dựa trên học máy có thể làm giảm các mối đe dọa giả mạo trong mạng 5G. Mô hình xác thực được củng cố bằng cách áp dụng một đặc điểm hai chiều cho phương pháp một chiều vì phương pháp hai chiều hiệu quả hơn trong việc xác định những kẻ tấn công. Bộ phân loại được trình bày có khả năng xác thực nhanh chóng vì nó được huấn luyện ngoại tuyến.

1.4.3. SDN-5G

Để làm cho mạng trở nên an toàn hơn nữa, kiến trúc bảo mật SDN-5G đã được phát triển. Phương pháp này dựa trên một khóa bí mật được đồng bộ hóa. Khóa này được tạo ra bởi một thuật toán mã hóa, sau đó chúng được lưu trữ trong các hệ thống hỗ trợ và thiết bị 5G. Mạng sẽ có thể biết liệu một cuộc tấn công có đến từ một cuộc tấn công giả mạo IP hay không vì khóa bí mật thay đổi theo thời gian. Bên cạnh đó, cuộc tấn công sẽ được phát hiện và ngăn chặn vì hệ thống hỗ trợ không thể cập nhật khóa bí mật được lưu trữ trên thiết bị. Liên quan đến kiểu tấn công phát lại, khả năng điều này xảy ra trong mạng 5G theo thiết kế này gần như bằng 0, vì khóa bí mật sẽ thay đổi với mỗi lần truyền, khiến tin nhắn được phát lại với một khóa đã lỗi thời.

1.4.4. Khung học sâu (Deep learning framework)

Bằng cách sử dụng các thuật toán học sâu dựa trên AI, các cuộc tấn công gây nhiễu được giảm thiểu. Phương pháp học sâu được đề xuất sử dụng một mạng lưới các thiết bị không

dây không đồng nhất được kết nối với nhau để thu thập dữ liệu cần thiết. Việc mở rộng khám phá các mối đe dọa mạng đòi hỏi phải sử dụng rộng rãi kỹ thuật học không giám sát.

1.4.5. Phương pháp NFV

Ảo hóa phân tán cải thiện đáng kể tính bảo mật của mạng; nó cũng tăng tính linh hoạt và độ tin cậy của mạng và giải quyết nhiều vấn đề tấn công ở nhiều tầng của mạng dựa trên 5G, bao gồm giả mạo dữ liệu, nghe lén, v.v.. Truyền dữ liệu được mã hóa ngăn chặn các cuộc tấn công giả mạo dữ liệu trong quá trình truyền.

1.4.6. SDN-guard

1.4.7. SDN-SC

1.4.8. Học Q-network sâu (Deep Q-network learning)

1.5. Học máy và bảo mật 5G

1.5.1. Vai trò của Học máy trong việc tăng cường bảo mật 5G

Việc sử dụng các thuật toán Học máy (ML) cung cấp một giải pháp tiên tiến và hiệu quả để tăng cường và giảm các thách thức bảo mật liên quan đến sự ra đời của mạng 5G. Các thuộc tính vốn có của ML, chẳng hạn như khả năng thích ứng, khả năng dự đoán và xử lý dữ liệu quy mô lớn, cung cấp các giải pháp đầy hứa hẹn để xử lý quy mô và độ phức tạp chưa từng có của hệ sinh thái 5G. Sự ra đời của mạng 5G đã mang lại một sự thay đổi mô hình trong thế giới viễn thông, hứa hẹn khả năng kết nối tốc độ cao, độ trễ thấp và tích hợp liền mạch hàng tỷ thiết bị trên toàn cầu. Tuy nhiên, sự tiến bộ to lớn này cũng mang lại một loạt các thách thức bảo mật mới đòi hỏi các giải pháp sáng tạo. Các thuật toán ML có khả năng xử lý khối lượng lớn dữ liệu được tạo bởi mạng 5G, trích xuất thông tin chi tiết có giá trị về các mối đe dọa có thể xảy ra và cung cấp các biện pháp bảo mật chủ động để ngăn chặn các cuộc tấn công tiềm ẩn. ML đã nổi lên như một công cụ mạnh mẽ trong bối cảnh này, cung cấp các khả năng mạnh mẽ để cải thiện khuôn khổ bảo mật trong mạng 5G.

1.5.2. Ứng dụng của Học máy trong bảo mật 5G

Học máy, với các khả năng phân tích nâng cao, đóng vai trò quan trọng trong các ứng dụng khác nhau trong bảo mật 5G. Các ứng dụng này trải rộng trên các lĩnh vực đa dạng như phát hiện xâm nhập, giảm thiểu phần mềm độc hại và mô hình hóa bảo mật dự đoán. Trong

ngữ cảnh phát hiện xâm nhập, các thuật toán học máy phân tích dữ liệu lưu lượng mạng để xác định các mẫu lệch khỏi chuẩn mực thông thường. Các kỹ thuật như SVM và Cây quyết định giúp tăng cường khả năng phản ứng và độ chính xác của hệ thống bảo mật, do đó giảm thiểu rủi ro tấn công mạng thành công.

Một số công trình nghiên cứu đã được công bố về việc sử dụng các thuật toán học máy trong bảo mật 5G:

- **"A Comprehensive Study on the Role of Machine Learning in 5G Security: Challenges, Technologies, and Solutions"**

- **"Cyber-Attack Detection and Mitigation Using SVM for 5G Network"**

1.5.3. Tối ưu hóa giao thức bảo mật với Học máy

ML cũng có thể được tận dụng để tối ưu hóa hoạt động của các giao thức bảo mật trong mạng 5G. Mục tiêu của tối ưu hóa giao thức bảo mật là tăng cường hiệu suất bảo mật của mạng mà không tiêu tốn quá nhiều tài nguyên tính toán hoặc ảnh hưởng đến hiệu quả hoạt động của mạng. Một lĩnh vực quan trọng mà ML có thể được sử dụng là trong việc tối ưu hóa các giao thức quản lý khóa, một nền tảng của giao tiếp an toàn trong mạng 5G. Các thuật toán ML có thể học và dự đoán thời gian tối ưu cho vòng quay khóa, do đó tăng cường bảo mật của giao tiếp được mã hóa trong khi giảm thiểu chi phí liên quan đến các thay đổi khóa thường xuyên.

CHƯƠNG 2: PHÁT HIỆN TẤN CÔNG DDOS TRONG MẠNG DI ĐỘNG 5G

2.1. Tấn công DDoS mạng di động 5G

DoS (*Denial-of-Service*) hay tấn công từ chối dịch vụ là hình thức tấn công mạng mà kẻ tấn công tìm cách ngăn cản người dùng máy tính sử dụng mạng truy cập đến máy chủ. Hình thức tấn công DoS phổ biến chính là tin tặc sẽ gửi một lưu lượng lớn truy cập vào hệ thống máy chủ làm cạn kiệt tài nguyên của nạn nhân. Có nhiều phương pháp tấn công từ chối dịch vụ như SYN Flood, UDP Flood, TCP Flood, Slowloris, Ping Flood, Ip NULL,…

Một cuộc tấn công 5G DoS nhằm mục đích làm gián đoạn chức năng của một thành phần cụ thể của mạng 5G, chẳng hạn như các trạm gốc hoặc các chức năng mạng ảo lõi, bằng cách áp đảo nó bằng lưu lượng hoặc yêu cầu quá mức, khiến các dịch vụ tạm thời hoặc vĩnh viễn không khả dụng. Một cuộc tấn công 5G DDoS là một hình thức phân tán của phương pháp này, trong đó nhiều thiết bị mạng bị xâm nhập bằng số lượng lớn lưu lượng độc hại một cách đồng thời và nhanh hơn. Mạng 5G đặc biệt dễ bị tác động bởi các cuộc tấn công này do sự hỗ trợ của chúng đối với việc kết nối cùng lúc nhiều thiết bị bởi băng thông mở rộng. Ngoài ra, với kiến trúc di động phân tán thế hệ mới, kết hợp các tính năng nâng cao như phân chia mạng và điện toán biên, tiềm ẩn nhiều bề mặt tấn công mới.

2.1.1. Mạng Botnet IoT

Một số lượng lớn các thiết bị IoT dự kiến kết nối qua đặt ra một thách thức đáng kể trong nghiên cứu bảo mật mạng di động 5G. Kẻ tấn công khai thác các thiết bị IoT được kết nối qua 5G để tiêm mã độc hại, tạo ra các botnet lớn có thể được điều khiển từ xa thông qua máy chủ Điều khiển và chỉ huy (C&C) [13]. Do việc cài đặt chức năng bảo mật cấp cao cho các thiết bị IoT cấp thấp là khá khó khăn hoặc không khả thi, nhiều thiết bị IoT có bảo mật tích hợp tối thiểu với mật khẩu yếu và bảo vệ cũ. Vì vậy, chúng có nhiều khả năng bị khai thác.

2.1.2. Bão tín hiệu (Signaling Storm)

Bão tín hiệu 5G là một loại tấn công Từ chối dịch vụ phân tán (DDoS) khai thác *Mặt phẳng điều khiển* của mạng 5G bằng cách áp đảo nó với một lượng lớn các thông báo tín hiệu. Không giống như các cuộc tấn công truyền thống tràn ngập mặt phẳng dữ liệu bằng lưu lượng truy cập, các cơn bão tín hiệu nhắm vào các quy trình chịu trách nhiệm quản lý kết nối, tính

di động và kiểm soát phiên, tạo ra tắc nghẽn tài nguyên trong mạng lõi. Khi các thiết bị mạng bị xâm nhập hoặc được cấu hình kém, liên tục gửi các yêu cầu kết nối, thiết lập phiên hoặc các tín hiệu liên quan đến tính di động, nó buộc mạng phải xử lý các tác vụ này liên tục.

2.2. SVM trong phát hiện tấn công DDoS

2.2.1. Cơ bản về thuật toán SVM

Với sự phát triển vượt trội của mạng di động 5G, nảy sinh nhu cầu cấp thiết trong việc phát triển các kỹ thuật học máy để phát hiện tấn công mạng. Các phương pháp phân loại như Rừng ngẫu nhiên (RF), K-láng giềng gần nhất (KNN) và Hồi quy Logistic (LR) có thể là những công cụ hữu ích để phát hiện một cuộc tấn công từ chối dịch vụ phân tán (DDoS).

Máy vectơ hỗ trợ (SVM) là một thuật toán học máy có giám sát, nó có thể sử dụng trong cả các tác vụ phân loại hoặc hồi quy. Tuy nhiên, trên thực tế SVM thường được sử dụng cho việc phân loại. SVM hướng đến việc tìm siêu mặt phẳng (hyper-plane) tối ưu trong không gian n-chiều để phân tách các điểm dữ liệu thành các lớp khác nhau. Thuật toán tối đa hóa biên độ giữa các điểm gần nhất của các lớp khác nhau. Khoảng cách giữa siêu mặt phẳng tới các vectơ hỗ trợ được gọi là biên độ (margin). SVM hướng đến mục tiêu tối đa hóa khoảng cách này để có hiệu suất phân loại tốt nhất. Ngoài phân loại tuyến tính, SVM cũng có thể sử dụng các hạt nhân (kernel) để thực hiện phân loại phi tuyến tính. Kernel là một hàm ánh xạ các điểm dữ liệu vào không gian có nhiều chiều hơn mà không cần tính toán rõ ràng tọa độ trong không gian đó. Điều này cho phép SVM hoạt động hiệu quả với dữ liệu phi tuyến tính bằng cách thực hiện ánh xạ ngầm định.

2.2.2. Áp dụng vào phát hiện tấn công DDoS

Trong khuôn khổ phát hiện các cuộc tấn công DDoS, SVM phân loại lưu lượng truy cập giữa các gói tin bình thường và bất thường, đồng thời các cuộc tấn công có thể được phân loại dựa trên các đặc điểm lưu lượng truy cập như tốc độ cổng nguồn và độ lệch chuẩn của các gói tin luồng. SVM đã được sử dụng vì một lý do rất đơn giản nhưng hiển nhiên, đó là trên thực tế nó có cơ chế phát triển khi nhắc đến các thuật toán và các kỹ thuật dự đoán tĩnh của nó vượt trội trên hầu hết các khuôn khổ, chưa kể đến việc nó có thể đào tạo và học các thuật toán nhanh chóng và thông minh như thế nào trong cả năng lực tuyến tính và phi tuyến tính.

Tỷ lệ phát hiện các cuộc tấn công từ chối dịch vụ phân tán (DDoS) sử dụng thuật toán SVM nằm trong khoảng từ 93% đến 98% [12]. Vì vậy, xuất phát từ mức độ hiệu quả cao của

các cuộc tấn công DDoS và khả năng cản trở các chức năng mạng thông thường của chúng là động lực chính để bắt đầu nghiên cứu này bằng cách sử dụng SVM. 5G tạo ra mức độ đe dọa bảo mật cao hơn chủ yếu là do có thêm các vectơ mà kẻ tấn công có thể lợi dụng. Một yếu tố khác cần xem xét ngữ cảnh này là khả năng của 5G trong việc cho phép nhiều thiết bị được kết nối với nhau, được gọi là Internet vạn vật (IoT).

2.2.3. Ảnh hưởng của việc chọn Kernel và Tinh chỉnh tham số

Việc lựa chọn hàm kernel phù hợp và tinh chỉnh các tham số của thuật toán SVM đóng vai trò quan trọng, ảnh hưởng trực tiếp đến hiệu suất của mô hình, đặc biệt trong bài toán phát hiện tấn công DDoS.

Lựa chọn Kernel:

Tinh chỉnh tham số SVM:

2.3. Áp dụng SVM vào thực tế với dữ liệu cực lớn của mạng 5G

2.3.1. Thách thức

+ Khối lượng dữ liệu (Big Data):

+ **Yêu cầu xử lý thời gian thực:**

+ **Tính đa dạng và động của lưu lượng mạng 5G:**

+ **Tài nguyên tính toán:**

+ **Chọn lọc đặc trưng (Feature Selection**

2.3.2. Giải pháp và hướng tiếp cận

+ Kỹ thuật giảm chiều dữ liệu: Sử dụng các kỹ thuật như Phân tích thành phần chính (PCA), Kernel PCA (KPCA) để giảm số lượng đặc trưng mà vẫn giữ lại được thông tin quan trọng, từ đó có thể giảm thời gian huấn luyện và độ phức tạp của mô hình;

+ **Học tăng cường (Incremental Learning) và Học trực tuyến (Online Learning):** Cho phép mô hình SVM cập nhật và học hỏi từ các luồng dữ liệu mới mà không cần phải huấn luyện lại toàn bộ mô hình từ đầu. Điều này rất quan trọng để thích ứng với sự thay đổi của các mẫu tấn công và lưu lượng mạng.

+ **Kiến trúc phân tán và tính toán song song:** Xử lý dữ liệu và huấn luyện mô hình trên các hệ thống phân tán (ví dụ: sử dụng *Apache Spark*) để tăng tốc độ;

+ **SVM tối ưu hóa:** Nghiên cứu và áp dụng các biến thể SVM được tối ưu hóa cho dữ liệu lớn hoặc các phương pháp giải gần đúng (*Approximate solvers*);

+ **Kết hợp với các công nghệ khác:** Tích hợp SVM với các công nghệ xử lý dữ liệu lớn (Big Data platforms) và các hệ thống giám sát mạng tiên tiến. Một số nghiên cứu đề xuất kết hợp SVM với các thuật toán học máy khác hoặc các mô hình học sâu để tăng cường khả năng phát hiện;

+ **Sử dụng phần cứng chuyên dụng:** Xem xét việc sử dụng các giải pháp tăng tốc phần cứng (ví dụ: GPU, FPGA) cho các tác vụ tính toán nặng của SVM;

+ **Lấy mẫu dữ liệu thông minh (Intelligent Data Sampling):** Thay vì sử dụng toàn bộ dữ liệu lớn để huấn luyện, có thể áp dụng các kỹ thuật lấy mẫu thông minh để chọn ra một tập dữ liệu đại diện, giúp giảm thời gian huấn luyện mà vẫn đảm bảo hiệu suất.

2.4. Phương án triển khai phát hiện xâm nhập bằng SVM

Như đã trình bày ở trên, việc lựa chọn kernel, tinh chỉnh tham số của SVM hay kết hợp với các kỹ thuật khác ảnh hưởng rất nhiều tới hiệu quả của việc phát hiện tấn công, đặc biệt với lượng dữ liệu vô cùng lớn của 5G. Trong đề án này, tác giả thực hiện nghiên cứu trên 2 phương pháp phát hiện xâm nhập mạng:

- Phương pháp 1: Sử dụng học máy truyền thống (SVM) và tối ưu hóa bằng cách loại bỏ các đặc trưng không cần thiết thông qua RFE (*Recursive Feature Elimination*);

- Phương pháp 2: Sử dụng học máy truyền thống (SVM) kết hợp lựa chọn đặc trưng dựa trên phân tích thống kê (SelectKBest) và xử lý mất cân bằng dữ liệu (SMOTE - *Synthetic Minority Over-sampling Technique*).

Phương pháp 1 (SVM + RFE) đã chứng tỏ sự vượt trội về mặt hiệu suất. Sự thành công này có thể được quy cho hai yếu tố chính:

- Lựa chọn đặc trưng bằng RFE: Phương pháp này đã chọn ra một bộ 25 đặc trưng chất lượng cao, có khả năng phân loại tốt khi được xem xét cùng nhau;

- Sự đơn giản và hiệu quả: Chỉ cần sử dụng class_weight='balanced' đã đủ để xử lý mất cân bằng khi kết hợp với bộ đặc trưng tốt, mà không cần đến kỹ thuật phức tạp hơn như SMOTE.

Phương pháp 2 (SVM + SelectKBest + SMOTE) là một cách tiếp cận hợp lý và nhanh hơn về mặt tính toán, nhưng bộ đặc trưng được chọn bởi SelectKBest có thể không tối ưu

bằng, dẫn đến hiệu suất thấp hơn một chút, đặc biệt là ở khả năng nhận diện một số lớp cụ thể như SYNFlood.

Các công trình được trích dẫn trong đề án cũng tập trung vào việc sử dụng học máy để giải quyết các vấn đề bảo mật 5G, nhưng có thể có những điểm khác biệt về phương pháp tiếp cận chi tiết hoặc phạm vi:

- Hai phương án này tập trung sâu hơn vào việc áp dụng cụ thể thuật toán SVM cho bài toán phát hiện tấn công DDoS và đi sâu vào các kỹ thuật tiền xử lý dữ liệu (RFE, SelectKBest, SMOTE) để tối ưu hóa hiệu suất trên bộ dữ liệu 5G-NIDD. Trong khi nghiên cứu của Fakhouri [6] cung cấp cái nhìn tổng thể về nhiều ứng dụng và thuật toán ML, đề án này đi vào chi tiết hơn về các phương pháp tối ưu hóa cho một thuật toán cụ thể (SVM) và một loại tấn công cụ thể (DDoS) trong môi trường 5G;

- Việc giới thiệu hai phương pháp (SVM + RFE và SVM + SelectKBest + SMOTE) thể hiện sự nỗ lực trong việc tìm ra cách tiếp cận tối ưu nhất, đặc biệt là trong việc xử lý mất cân bằng dữ liệu và lựa chọn đặc trưng, điều mà công trình của Alshunaifi [12] có thể chưa đi sâu vào chi tiết các kỹ thuật tối ưu hóa này.

CHƯƠNG 3: THỰC NGHIỆM, ĐÁNH GIÁ KẾT QUẢ

3.1. Mô hình triển khai

3.1.1. Các chỉ số đánh giá

3.1.2. Bộ dữ liệu mẫu

Quá trình thu thập dữ liệu có tầm quan trọng hàng đầu trong thiết kế của một hệ thống phát hiện xâm nhập. Trên thực tế, phân tích dữ liệu tiết lộ hành vi bất thường và cung cấp các chỉ số liên quan có thể giúp thiết kế các hệ thống phát hiện xâm nhập. Ngoài ra, chất lượng của dữ liệu có ảnh hưởng trực tiếp đến hiệu suất của các thuật toán phát hiện, đặc biệt là các thuật toán dựa trên học máy. Dữ liệu này có thể được lấy từ các mạng chuyên dụng trong thế giới thực hoạt động với quyền truy cập hạn chế hoặc các mạng tạm thời và hệ thống thử nghiệm được triển khai đặc biệt để thu thập bộ dữ liệu.

Trong phạm vi nghiên cứu của đề án, tác giả đề xuất sử dụng bộ dữ liệu 5G-NIDD làm bộ dữ liệu mẫu, dùng để thử nghiệm các phương pháp đã trình bày. Bộ dữ liệu 5G-NIDD là bộ dữ liệu xâm nhập mẫu mới nhất, có nguồn gốc từ một mạng thử nghiệm 5G thực [3]. Nó được tạo ra bằng cách sử dụng một mạng 5G đầy đủ chức năng được tạo ra cho các mục đích thử nghiệm 5G, giống với một mạng thực tế. 5G-NIDD bao gồm các tính năng duy nhất, đặc trưng cho các luồng mạng 5G.

Quá trình thu thập và xử lý dữ liệu của 5G-NIDD: Dữ liệu được thu thập ghi lại lưu lượng tấn công và lưu lượng bình thường khi đi qua 2 trạm RAN. Mỗi loại tấn công được ghi lại trong các phiên riêng biệt dưới dạng file .pcap, sau đó các file .pcap được đưa vào trong các quy trình xử lý.

1, Loại bỏ lớp GTP: Vì dữ liệu đi qua giao diện vô tuyến nên chứa lớp GTP-U. Loại bỏ lớp GTP là cần thiết để phân tích chính xác. Sử dụng Tracewrangler;

2, Chuyển đổi sang luồng mạng: Giao thức Argus được dùng để chuyển dữ liệu từ dạng gói sang luồng mạng. Mỗi luồng có IP, port, protocol,... Điều này làm giảm kích thước dữ liệu và thuận tiện hơn cho mô hình học máy;

3, Gộp dữ liệu và dán nhãn: Dữ liệu được chuyển sang file CSV và án nhãn. Gộp dữ liệu từ các phiên theo từng trạm RAN sau đó kết hợp lại. Kết quả cuối cùng: ***1.215.890 dòng*** dữ liệu luồng mạng.

4, Mã hóa dữ liệu: Một số đặc trưng để là dữ liệu phân loại (categorical), không thể trực tiếp đưa vào mô hình ML. Vì vậy, cần dùng kỹ thuật One-hot Encoding để chuyển thành số nhị phân.

3.1.3. Kiến trúc hệ thống

Bộ dữ liệu mẫu 5G-NIDD cung cấp rất nhiều đặc trưng khác nhau trong các dữ liệu mạng (Bảng 3.3). Tuy nhiên, trong phạm vi nghiên cứu thực nghiệm, tác giả chỉ tập trung vào một số đặc trưng cơ bản:

- Thông số gói tin (TotPkts, SrcPkts, DstPkts): Lưu lượng gói tin;
- Thông số byte (TotBytes, SrcBytes, DstBytes): Kích thước dữ liệu;
- Tốc độ truyền (Rate, SrcRate, DstRate);
- Tải mạng (Load, SrcLoad, DstLoad);
- Thời gian (Mean, Dur);
- Thông tin giao thức (Proto, Sport, Dport).

Các đặc trưng được chọn dựa trên tần suất xuất hiện lớn trong bộ dữ liệu. Việc chọn ra các đặc trưng là để tập trung vào những đặc trưng xuất hiện liên tục và liên quan đến lưu lượng, kích thước, thông tin, thời gian, tốc độ, tỷ lệ truyền tải mạng, giúp mô hình học máy phán đoán chính xác hơn và tránh overfitting.

3.2. Tiến hành thực nghiệm

3.2.1. Cài đặt môi trường mô phỏng

Để tiến hành các thực nghiệm đánh giá hiệu quả của thuật toán SVM trong việc phát hiện tấn công DDoS trên mạng 5G, một môi trường mô phỏng đã được thiết lập với các thành phần phần cứng và phần mềm cụ thể. Môi trường này đảm bảo tính nhất quán và khả năng tái tạo của các kết quả thực nghiệm.

- **Phần cứng (Hardware):**
- **Phần mềm (Software):**

3.2.2. Các bước thực hiện

- **Bước 1: Tiền xử lý và phân tích dữ liệu** thông qua các bước sau:
- **Bước 2: Huấn luyện và Tối ưu hóa mô hình:**
- **Bước 3: Lưu trữ và triển khai mô hình:**

- Bước 4: Sử dụng mô hình để phát hiện tấn công:

3.3. Kết quả thực nghiệm

3.3.1. Kết quả thu được

Cả 2 phương pháp thực nghiệm đều cho thấy hiệu suất cao, nhưng có sự khác biệt về kiến trúc, kỹ thuật và kết quả cuối cùng:

- Phương pháp 1: Độ chính xác 99.87%. Đạt điểm F1-score 1.00 trên hầu hết các loại tấn công như Benign, HTTPFlood, ICMPFlood, SYNFlood, SYNScan, TCPConnectScan, UDPFlood, UDPScan.

- Phương pháp 2 (SVM + SelectKBest + SMOTE): Độ chính xác 97.99%. Hiệu suất thấp hơn một chút ở lớp SYNFlood với recall là 0.86. Mặc dù hiệu suất rất tốt, nhưng vẫn thấp hơn phương pháp 1.

3.3.2. Đánh giá kết quả

Kết quả thực nghiệm cho thấy cả hai phương pháp tiếp cận đều xây dựng được mô hình SVM hiệu quả để phát hiện tấn công trên mạng 5G. Tuy nhiên, có sự khác biệt rõ rệt về hiệu suất:

- Phương pháp 1 (SVM + RFE) tỏ ra vượt trội, cân bằng tốt giữa hiệu suất cao và sự đơn giản. RFE là một phương pháp lựa chọn đặc trưng mạnh mẽ. Việc sử dụng RFE để lựa chọn một tập hợp gồm 25 đặc trưng chất lượng cao đã giúp mô hình loại bỏ nhiễu và tập trung vào các tín hiệu quan trọng nhất. Điều này dẫn đến độ chính xác cao hơn và khả năng nhận diện cân bằng trên tất cả các loại tấn công. Hơn nữa, phương pháp này chứng minh rằng việc xử lý mất cân bằng bằng cách điều chỉnh trọng số (class_weight='balanced') là đủ hiệu quả khi kết hợp với một bộ đặc trưng tốt;

- Phương pháp 2 (SVM + SelectKBest + SMOTE) là một phương pháp hợp lý nhưng hiệu suất thấp hơn. Xử lý tường minh vấn đề mất cân bằng dữ liệu với SMOTE. SelectKBest là một kỹ thuật lựa chọn đặc trưng nhanh. Mặc dù đã nỗ lực xử lý mất cân bằng dữ liệu một cách tường minh bằng SMOTE, nhưng bộ 20 đặc trưng được chọn bởi SelectKBest có thể không tối ưu bằng. SelectKBest đánh giá các đặc trưng một cách độc lập và có thể đã bỏ lỡ các mối quan hệ tương tác phức tạp, dẫn đến việc mô hình nhận diện một số lớp tấn công (như SYNFlood) kém hơn.

Ưu điểm của các phương pháp thực nghiệm so với phương pháp cũ:

- Hiệu suất cao hơn: Cả hai phương pháp đều cho độ chính xác cao hơn đáng kể so với một cách tiếp cận cơ bản không có lựa chọn đặc trưng hay tối ưu hóa tham số;

- Mô hình gọn nhẹ, dễ kiểm soát nhưng vẫn hiệu quả: Việc giảm số lượng đặc trưng (từ hơn 90 xuống còn 25 hoặc 20) giúp mô hình hoạt động nhanh hơn ở cả giai đoạn huấn luyện và dự đoán;

- Khả năng khái quát hóa tốt: Quy trình tối ưu hóa tham số có hệ thống (GridSearchCV/RandomizedSearchCV) giúp tìm ra mô hình có khả năng hoạt động tốt trên dữ liệu mới, tránh overfitting;

- Xử lý được vấn đề mất cân bằng: Cả hai phương pháp đều có cơ chế riêng để xử lý vấn đề mất cân bằng dữ liệu, giúp cải thiện khả năng phát hiện các cuộc tấn công hiếm gặp.

Các loại tấn công có thể phát hiện:

Nhược điểm:

- Phương pháp 1: RFE có thể tốn nhiều thời gian tính toán hơn so với các phương pháp khác;

- Phương pháp 2: Hiệu suất thấp hơn hai phương pháp còn lại. SelectKBest có thể bỏ lỡ sự tương tác giữa các đặc trưng.

Hướng phát triển:

- Thu thập thêm dữ liệu về các loại tấn công mới và biến thể;

- Tối ưu hóa các tham số của thuật toán và cải tiến các bước tiền xử lý;

- Phát triển khả năng học liên tục hoặc học online để mô hình tự cập nhật;

- Tích hợp với các hệ thống bảo mật khác như SIEM, Firewall;

- Cải thiện khả năng xử lý dữ liệu phân tán để tăng tốc độ;

- Phát triển giao diện quản lý và giám sát trực quan hơn;

- Tự động hóa quy trình cập nhật mô hình;

- Thử nghiệm tích hợp các phương pháp học sâu hoặc mô hình ensemble để cải thiện độ chính xác;

- Thực hiện các Feature Engineering nâng cao.

KẾT LUẬN VÀ ĐỀ XUẤT

Đề án đã tìm hiểu tổng quan kiến trúc mạng 5G và các nguy cơ bảo mật, đặc biệt là tấn công DDoS. SVM được nghiên cứu và áp dụng để phát hiện tấn công DDoS, phân loại lưu lượng mạng thành bình thường hoặc độc hại. Quá trình thực nghiệm trên bộ dữ liệu 5G-NIDD cho thấy SVM có độ chính xác cao trong phát hiện các loại tấn công DDoS phổ biến, phù hợp cho giám sát thời gian thực. Tuy nhiên, việc triển khai thực tế đối mặt với những thách thức về khối lượng dữ liệu lớn, yêu cầu xử lý thời gian thực, tính đa dạng của lưu lượng mạng và tài nguyên tính toán trên hệ thống. Hiệu quả của mô hình phụ thuộc rất nhiều vào chất lượng dữ liệu huấn luyện và việc tinh chỉnh các tham số của SVM.

Để nâng cao hiệu quả phát hiện tấn công DDoS trong mạng 5G, tác giả đề xuất cần tập trung nghiên cứu chuyên sâu vào các hướng sau:

Dữ liệu và đặc trưng: Thu thập thêm dữ liệu về các loại tấn công mới và biến thể. Nghiên cứu sâu hơn về kỹ thuật Feature Engineering nâng cao để trích xuất đặc trưng có ý nghĩa từ dữ liệu mạng 5G.

Tối ưu hóa mô hình: Áp dụng các kỹ thuật tối ưu hóa tham số nâng cao (ví dụ: *tối ưu hóa Bayes*) và phát triển khả năng học liên tục (*Continual learning*) hoặc học trực tuyến (*Online learning*) để mô hình tự động cập nhật và thích nghi với các mẫu tấn công thay đổi. Thử nghiệm kết hợp học sâu (*Deep Learning*) để cải thiện độ chính xác và khả năng phát hiện các cuộc tấn công phức tạp hơn.

Hiệu suất và triển khai: Nghiên cứu kiến trúc phân tán và tính toán song song trên các nền tảng Big Data (như *Apache Spark*) để tăng tốc độ xử lý dữ liệu. Tích hợp hệ thống phát hiện với các giải pháp bảo mật hiện có (*SIEM, Firewall*) để tạo ra giải pháp toàn diện cho mạng 5G. Phát triển giao diện quản lý và giám sát trực quan, đồng thời tự động hóa quy trình cập nhật mô hình.