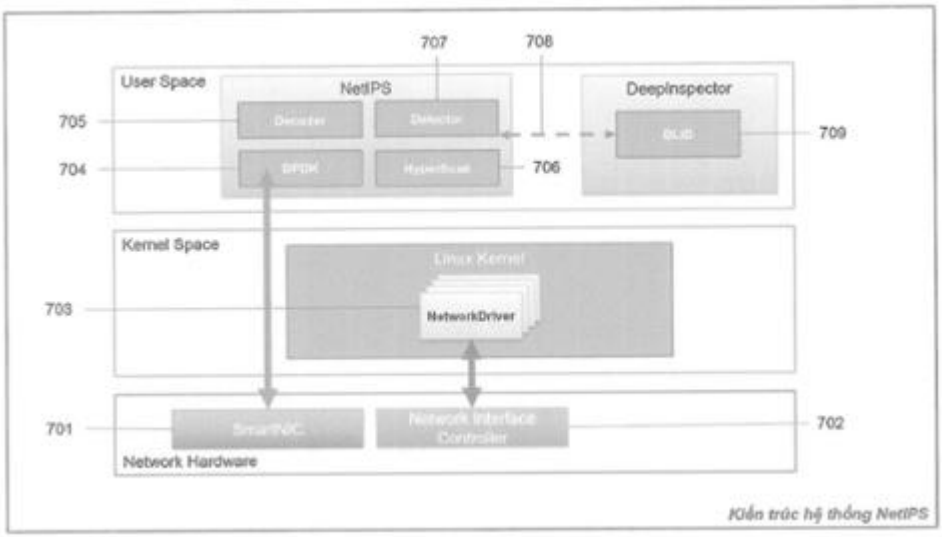
(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ
(19) **Cộng hòa xã hội chủ nghĩa Việt Nam (VN)**
CỤC SỞ HỮU TRÍ TUỆ
(11) **1-0042196**
(51)[2020.01] **G06F 21/00**
(13) **B**

(21) 1-2021-05291
(22) 26/08/2021
(45) 25/12/2024 441
(43) 25/11/2021 404
(73) 1. Trường Đại học Công nghệ - Đại học Quốc Gia Hà Nội (VN)
E3, 144 Xuân Thủy, Phường Dịch Vọng Hậu, Quận Cầu Giấy, Thành Phố Hà Nội
2. Nguyễn Ngọc Hoá (VN)
P311-E3, 144 Xuân Thuỷ, phường Dịch Vọng Hậu, quận cầu Giấy, Hà Nội
(72) Nguyễn Ngọc Hoá (VN); Dư Phương Hạnh (VN).

(54) PHƯƠNG PHÁP VÀ HỆ THỐNG PHÁT HIỆN, NGĂN CHẶN XÂM NHẬP MẠNG SỬ DỤNG LUẬT VÀ MÔ HÌNH HỌC SÂU

(57) Sáng chế đề cập đến cả hệ thống và phương pháp phát hiện và ngăn chặn xâm nhập mạng sử dụng kết hợp cả phân tích theo luật và phân tích dựa trên mô hình học sâu áp dụng kèm chiến lược lấy mẫu theo tần suất và thời khoảng. Theo phương pháp này, những luồng mạng không được thẩm định bởi bất kỳ luật nào sẽ được lấy mẫu theo tần suất và thời khoảng đã được xác lập trong tệp cấu hình, từ đó tiến hành phân tích sâu với mô hình mạng nơ ron sâu DNN để phát hiện các luồng mạng bất thường, bị tấn công xâm nhập mạng để từ đó có thể phòng chống và ngăn chặn xâm nhập. Phương pháp kết hợp đề xuất bao gồm các bước chính: (i) phát hiện xâm nhập với kỹ thuật siêu so khớp mẫu dựa trên tập luật; (ii) lấy mẫu theo chiến lược dựa trên tần suất và thời khoảng những luồng dữ liệu mạng không được thẩm tra bởi bất kỳ luật nào; (iii) phân tích sâu, phát hiện xâm nhập sử dụng mô hình học sâu DLID những luồng mạng đã được lấy mẫu; (iv) tự động sinh và cập nhật luật từ các luồng dữ liệu mạng xâm nhập bị phát hiện bởi DLID cho hệ thống phát hiện và ngăn chặn xâm nhập mạng. Dựa trên phương pháp này, chúng tôi đã xây dựng hệ thống phát hiện và ngăn chặn xâm nhập NetIPS, trong đó bao gồm cả thành phần DeepInspector đảm nhiệm vai trò phân tích sâu với mô hình học sâu DLID, cho phép triển khai được với cả những mạng có băng thông lớn.



Kiến trúc hệ thống NetIPS

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp phân tích luồng mạng kết hợp sử dụng phân tích theo luật và phân tích dựa trên mô hình học sâu để xây dựng hệ thống phát hiện và phòng chống xâm nhập mạng máy tính. Cụ thể, sáng chế đề cập đến phương pháp phát hiện và phòng chống xâm nhập mạng dựa vào phân tích, phát hiện các luồng mạng bất thường, có dấu hiệu tấn công xâm nhập mạng, từ đó có thể ngăn chặn và phòng chống tấn công. Phương pháp phát hiện xâm nhập theo sáng chế kết hợp cả kỹ thuật phát hiện dựa trên tập luật (thể hiện qua các dấu hiệu, chữ ký, …) và phát hiện dựa trên phân tích phát hiện luồng mạng bất thường sử dụng mô hình học sâu. Để tránh tình trạng thắt cổ chai, gây ngẽn mạng khi phân tích sử dụng mô hình học sâu, sáng chế đề xuất chiến lược lấy mẫu theo tần xuất và thời khoảng đã được xác định trước chỉ đối với những luồng mạng không được thẩm tra bởi bất kỳ luật nào. Hơn nữa, để xử lý được những luồng mạng quy mô lớn, sáng chế đề xuất xây dựng hệ thống tích hợp thiết bị mạng chuyên dụng và phương pháp xử lý luồng mạng ở không gian người dùng (user-space level) để không bị mất gói tin và xử lý được những luồng mạng có quy mô rất lớn.

Tình trạng kỹ thuật của sáng chế

Hiện nay, trong môi trường Internet, nguy cơ phân vùng mạng nội bộ của một tổ chức có thể bị tấn công xâm nhập, chiếm quyền sử dụng máy chủ hay thực hiện triển khai phần mềm mã độc có chủ đích APT (Advanced Persistent Threat), chẳng hạn như được lập lịch để tấn công từ chối dịch vụ, tấn công vào thời điểm đã định sẵn, …, thậm chí định kỳ gửi dữ liệu người dùng cho kẻ tấn công. Để có thể giảm thiểu những nguy cơ, thiệt hại có thể có từ những tấn công xâm nhập hệ thống mạng như thế, giải pháp triển khai hệ thống thiết bị phòng chống xâm nhập IDPS có thể đảm nhiệm được cả vai trò phát hiện xâm nhập IDS (intrusion detection system) cũng như ngăn chặn xâm nhập IPS (intrusion prevention system) được xem như một trong những giải pháp hiệu quả nhất hiện nay.

Hệ thống NIPS (Network-based IPS) đảm nhiệm việc phòng chống mối đe dọa an ninh mạng dựa trên việc thẩm tra các luồng lưu lượng mạng để phát hiện nguy cơ xâm nhập và từ đó ngăn chặn việc khai thác lỗ hổng bảo mật trong các hệ thống bên trong của phân vùng mạng nội bộ của tổ chức. Quy trình chung thẩm tra, phân tích luồng dữ liệu mạng để phát hiện xâm nhập được minh hoạ như hình 1. NIPS thường nằm ngay sau hệ thống tường lửa và cung cấp lớp phân tích gói tin sâu hơn để phát hiện và phòng chống

xâm nhập mạng, minh hoạ như hình 2. Thông thường, NIPS được triển khai nội tuyến (inline; trên đường giao tiếp trực tiếp giữa nguồn và đích), chủ động phân tích và thực hiện các hành động tự động trên tất cả các luồng lưu lượng truy cập vào mạng. Cụ thể, những hành động này bao gồm:

- Gửi cảnh báo cho quản trị viên (tương tự như IDS),
- Loại bỏ các gói độc hại,
- Chặn lưu lượng truy cập từ địa chỉ nguồn,
- Thiết lập lại kết nối.

Với cơ chế bảo mật triển khai nội tuyến, NIPS phải hoạt động hiệu quả để tránh làm giảm hiệu suất mạng. Nó cũng phải hoạt động nhanh chóng vì việc khai thác có thể xảy ra trong thời gian gần thực. NIPS cũng phải phát hiện và phản hồi chính xác, để loại bỏ các mối đe dọa và dương tính giả (các gói tin hợp pháp bị phân tích và xác định nhầm thành mối đe dọa).

Để phát hiện được nguy cơ xâm nhập mạng, NIPS thường sử dụng hai phương pháp phân tích điển hình: (i) phương pháp phát hiện dựa trên tập luật và (ii) phương pháp phát hiện dựa trên thống kê bất thường. Phát hiện dựa trên tập luật dựa trên một từ điển gồm các mẫu hoặc chữ ký có thể nhận dạng duy nhất trong mỗi luồng dữ liệu mạng. Khi một lỗ hổng bị khai thác được phát hiện, chữ ký của nó được ghi lại và lưu trong tập luật được sử dụng để phát hiện xâm nhập. Phát hiện dựa trên tập luật cho NIPS được chia thành hai loại:

- Dấu hiệu bị khai thác: xác định kẻ tấn công khai thác xâm nhập dựa vào mẫu minh chứng nỗ lực khai thác cụ thể.
- Dấu hiệu lỗ hổng bảo mật: cho phép phát hiện xâm nhập mạng khi có dấu hiệu nhắm đến lỗ hổng của hệ thống đích. Phương pháp này cho phép phát hiện không phụ thuộc vào mẫu khai thác đã biết; tuy nhiên lại tiềm ẩn tăng nguy cơ cơ dương tính giả khi người dùng thông tin sử dụng các chức năng hệ thống đích có chứa lỗ hổng bảo mật.

Tính năng phát hiện bất thường thống kê lấy ngẫu nhiên các mẫu lưu lượng mạng và so sánh chúng với mức hiệu suất cơ bản được tính toán trước. Khi mẫu hoạt động lưu lượng mạng nằm ngoài các tham số của hiệu suất cơ sở, NIPS sẽ thực hiện hành động để xử lý tình huống. Đây chính là hướng nghiên cứu trong nhiều sáng chế cũng như công trình khoa học nhằm nâng cao hiệu năng phát hiện và ngăn chặn xâm nhập mạng. Hiện để phòng chống xâm nhập mạng, đã có một số sáng chế đề cập đến nhiều phương pháp chuyên biệt khác nhau. Cụ thể như sau:

Sáng chế US9319425B2, công bố năm 2016, đề xuất hệ thống và phương pháp xây dựng công cụ an ninh mạng, có khả năng xác định lỗ hổng bảo mật hiện tại của mạng riêng, phát hiện dấu hiệu tấn công khai thác lỗ hổng bảo mật. Trong sáng chế này, nhóm tác giả đã xây dựng kỹ thuật để có thể phát hiện những luồng dữ liệu mạng có dấu hiệu bị tấn công dựa trên các lỗ hổng bảo mật, tuy đã được xác định trong tập luật nhưng chưa được kích hoạt. Trong sáng chế số US9565208B2, được xuất bản năm 2017, các tác giả đã xây dựng phương pháp phát hiện xâm nhập mạng dựa trên việc (i) xác định một hoặc nhiều mối đe dọa trong các phiên dữ liệu, được xác định tương ứng với lưu lượng mạng từ nhiều máy chủ; (ii) kỹ thuật xác định điểm phát hiện phân lớp tương ứng với điểm chắc chắn và điểm đe dọa; (iii) xác định điểm máy chủ phân lớp tương ứng với điểm chắc chắn và điểm đe dọa; và (iv) tạo ra dữ liệu cảnh báo bao gồm điểm phát hiện phân lớp và điểm máy chủ phân lớp. Khi triển khai, lưu lượng mạng có thể được nhận một cách thụ động thông qua sử dụng cổng chuyên dụng của một bộ chuyển mạch mạng; chẳng hạn sử dụng cổng SPAN.

Trong tài liệu sáng chế số US10454955B2, được xuất bản năm 2019, đề cập đến phương pháp giám sát ngữ cảnh thời gian thực để phát hiện và ngăn chặn xâm nhập mạng. Nhóm tác giả đã sử dụng một tác nhân để giám sát các giao dịch ứng dụng đến và đi; từ đó truyền dữ liệu giao dịch được giám sát đến một máy chủ giám sát để phân tích, phát hiện giao dịch bất thường. Tác nhân nhận được hướng dẫn từ máy chủ giám sát để thực hiện một hành động dựa trên dữ liệu giao dịch được giám sát đã truyền và báo cáo trạng thái của hành động đã thực hiện cho máy chủ giám sát.

Tài liệu sáng chế CN108347430B, công bố năm 2021, đề cập phương phát hiện xâm nhập mạng và quét lỗ hổng bảo mật dựa trên học sâu. Phương pháp này bao gồm các bước chính: (i) thu thập tệp mẫu độc hại và thiết lập cơ sở dữ liệu tệp độc hại; (ii) huấn luyện và mô hình hóa dựa trên hành vi của tệp độc hại trong cơ sở dữ liệu tệp độc hại bằng cách sử dụng thuật toán học sâu và đào tạo gia tăng mô hình về giám sát thời gian thực được thực hiện theo tệp mẫu độc hại mới đã nhận để có được mô hình phân loại ; (ii) mô phỏng các tệp mẫu độc hại trong cơ sở dữ liệu tệp độc hại để chạy trong các môi trường khác nhau và phát hiện các đặc điểm tấn công của tệp mẫu độc hại bằng cách sử dụng IDS (hệ thống phát hiện xâm nhập); và (iv) phân tích cơ sở dữ liệu tệp độc hại bằng cách sử dụng thuật toán khai thác dữ liệu, xây dựng thư viện tính năng chế độ tấn công lỗ hổng, tạo gói tấn công mạng và quét lỗ hổng mạng.

Sáng chế WO/2021/082339, công bố năm 2021, mô tả phương pháp sử dụng học máy kết hợp luật để phát hiện xâm nhập và triển khai trên thiết bị phần cứng với thư viện DPDK (Data Plane Development Kit). Sáng chế này cũng đề xuất nâng cao hiệu năng so khớp luật với giải thuật PFAC (Parallel Failureless Aho-Corasick) triển khai trên GPU. Tuy nhiên, sáng chế này chưa làm rõ tập đặc trưng luồng mạng cũng như giải thuật học máy được áp

dụng để phân tích, phát hiện xâm nhập từ phân tích sâu luồng mạng. Sáng chế CN114357445B, công bố năm 2021, đề cập đến phương pháp và thiết bị rò quét lỗ hổng và phát hiện xâm nhập mạng dựa trên phương pháp học sâu sử dụng mạng nơ ron và học tăng cường.Nhìn chung, cả hai sáng chế này cũng đã đề cập đến phương pháp học máy để phát hiện xâm nhập theo thời gian thực, nhưng cũng chưa có sáng chế nào làm rõ chiến lược kiểm soát luồng mạng theo tần xuất và thời gian lấy mẫu để phân tích sâu hơn với mô hình mạng DNN (Deep Neural Network) để có thể triển khai ở chế độ nội tuyến.

Bản chất kỹ thuật của sáng chế

Mục tiêu của sáng chế là đề xuất hệ thống và phương pháp phát hiện, ngăn chặn xâm nhập mạng dựa trên phân tích luồng dữ liệu ở chế độ nội tuyến để phát hiện xâm nhập thông qua kỹ thuật kết hợp sử dụng tập luật và với chiến lược lấy mẫu theo tần xuất và thời khoảng để phân tích sâu hơn với mô hình DNN.

Tình hình mất an toàn thông tin ngày càng phức tạp, nhiều tổ chức tội phạm mạng hình thành, với kỹ thuật tấn công tinh vi, phức tạp, mục tiêu nhắm đến đa dạng và gây nên các tổn thất khó lường. Do vậy, việc phát hiện và ngăn chặn xâm nhập mạng là một vấn đề cần thiết, đòi hỏi giải pháp phải được cải tiến, thay đổi liên tục nhằm kịp thời phát hiện các dấu hiệu xâm nhập mạng mới nhất và đáp ứng với hiệu năng cao nhất. Đặc biệt, khi triển khai NIPS với những phân vùng mạng có lưu lượng lớn (từ 10Gps trở lên), việc sử dụng mô hình học máy nói chung cũng như mô hình học sâu nói riêng để phát hiện xâm nhập đối mặt với những thách thức rất lớn về hiệu năng xử lý của quá trình phát hiện xâm nhập. Hình 1 minh hoạ quy trình chung phân tích luồng dữ liệu mạng để phát hiện xâm nhập mạng. Thông thường, với quy mô luồng dữ liệu mạng lớn, người quản trị mạng phải vô hiệu hoá chức năng phát hiện dựa trên mô hình học máy. Đây chính là những động lực để chúng tôi nghiên cứu, đề xuất sáng chế này. Ý tưởng chính của chúng tôi để triển khai NIPS với luồng dữ liệu mạng quy mô lớn dựa trên phương pháp kết hợp sử dụng cả tập luật và mô hình học sâu sử dụng mạng nơ ron sâu DNN (Deep Neural Network). Chi tiết của phương pháp này sẽ được trình bày cụ thể ở mục sau.

Sáng chế đề cập đến các phương pháp phát hiện xâm nhập mạng được triển khai trên hệ thống tính toán hiệu năng cao và sử dụng bộ tăng tốc mạng SmartNIC và bộ xử lý đồ hoạ GPU, triển khai theo mô hình nội tuyến (inline), để phân tích luồng mạng, phát hiện và ngăn chặn xâm nhập mạng theo thời gian thực. Phương pháp này được xây dựng trên cơ sở kết hợp sử dụng các bộ luật (có các chữ ký và mẫu dấu hiệu xâm nhập biết trước) và chiến lược lấy mẫu theo tần suất và thời gian xác lập để phân tích sâu sử dụng mạng nơ ron DNN.

Để có thể xử lý được những luồng dữ liệu mạng quy mô lớn (với những cổng mạng băng thông từ 10Gbps trở lên), chúng tôi đề xuất xây dựng thiết bị NetIPS với việc sử dụng các linh kiện cấu thành nền tảng tính toán hiệu năng cao kết hợp cả bộ xử lý nhiều lõi và bộ xử lý đồ hoạ GPU. Ngoài ra, để tránh làm rớt gói tin với những luồng dữ liệu mạng thông lượng lớn, bộ tăng tốc xử lý mạng (network accelerator) cũng được lựa chọn tích hợp trong NetIPS để có thể thu nhận được luồng dữ liệu mạng này và đẩy lên mức người dùng để phân tích, phát hiện xâm nhập.

Phương pháp phát hiện và ngăn chặn xâm nhập mạng bao gồm các bước chính như sau:

- Thu thập dữ liệu mạng đáp ứng hiệu năng cao >=10Gb: với việc sử dụng thiết bị card mạng chuyên dụng hay còn được gọi là bộ tăng tốc mạng SmartNIC cho phép bằng thông từ 10Gbps trở lên chúng tôi còn sử dụng phương pháp xử lý gói tin của hệ thống NetIPS ở không gian người dùng trong hệ điều hành Linux. Theo đó, các gói tin được thu nhận từ card mạng chuyên dụng được chuyển thẳng đến phần mềm hệ thống NetIPS thông qua thư viện DPDK.

- Phát hiện và ngăn chặn xâm nhập mạng dựa trên các bộ luật: với thông lượng từ 10Gb/s, phân loại và phân tích so khớp các dữ liệu đi vào (input) với các tập luật tương ứng (được cập nhật liên tục, từ nhiều nguồn khác nhau trong đó bao gồm cả kết quả đầu ra của mô hình học sâu), đối với các dữ liệu đi vào trùng khớp với các dấu hiệu của luật bất kỳ trong cơ sở dữ liệu, sẽ được xử lý Drop/Reject không cho phép đi qua để vào trong hệ thống mạng, đồng thời tự động sinh ra cảnh báo (alert) đến cho quản trị viên để cung cấp thông tin về quá trình vừa xử lý. Cơ chế gửi Alert cho quản trị viên được hệ thống thực hiện tự động thực hiện trên NetIPS hiển thị giao diện quản trị, hoặc qua hệ thống thư điện tử (email), hay qua hệ thống SMS tuỳ thuộc mức độ nguy hiểm của từng loại xâm nhập mạng được hệ thống phát hiện và tuỳ thuộc vào thiết lập của quản trị viên.

- Trích xuất dữ liệu mạng theo chuỗi thời gian thực: Các dữ liệu không trùng khớp với các luật được chuyển tiếp đi vào trong hệ thống và được trích xuất (capture) thành các tệp định dạng ".pcap" theo chu kỳ và thời khoảng nhất định (tuỳ thuộc vào cấu hình của quản trị viên). Đây là kỹ thuật sử dụng dựa trên ý tưởng cửa sổ trượt (sliding windows) và lấy mẫu ngẫu nhiên (random). Các file .pcap này là dữ liệu đầu vào của mô hình học máy sâu sử dụng mạng nơ ron DNN cho thành phần tiếp theo.

- Trích rút đặc trưng trên dữ liệu được thu thập: tệp .pcap được trích xuất ở thành phần trên không thể sử dụng ngay trên mô hình học học sâu, mà cần thực hiện trích rút các đặc trưng của nó. Đây là một trong những bước cần thiết và quan trọng để đưa vào mô hình học sâu nhằm phát hiện dấu hiệu xâm nhập mạng. Trong phương pháp của chúng tôi CICFlowmeter được sử dụng để phân tích tệp .pcap, mô hình hoá thành 83 đặc trưng (xem

chi tiết tại https://github.com/ahlashkari/CICFlowMeter) và lưu trữ dưới định dạng file .csv để sử dụng trong mô hình học sâu phát hiện xâm nhập mạng.

- Mô hình mạng nơ ron DNN được huấn luyện và xây dựng thông qua các giai đoạn sau:

+ Xây dựng testbed và thu thập dữ liệu huấn luyện: Một hệ thống mạng mô phỏng được xây dựng bao gồm các thành phần chính như: các máy chủ phục vụ các hệ thống thư điện tử (email), website (web app, web server, web database), reverse Proxy, ... Bên cạnh đó là các máy tính để tiến hành tấn công xâm nhập mạng với 11 hình thức, kiểu tấn công khác nhau (SQL Injection, Brute Force-Web, Brute Force-XSS, BruteForce-SSH, BruteForce-FTP, DoS-SlowHTTPTest, DoS-Hulk, DoS-Slowloris, DoS-GoldenEye, Infilteration, BotAttack) và máy tính người dùng để tạo ra các luồng dữ liệu lành tính (Begnin). Đối với mỗi loại tấn công xâm nhập, được chúng tôi xây dựng thành các kịch bản để thực hiện và thu thập dữ liệu và đánh dấu thành các file .pcap tương ứng đối với mỗi loại.

+ Trích rút và gán nhãn các dữ liệu: đối với mỗi file .pcap được đánh dấu cho các loại dữ liệu, chúng tôi sử dụng CICFlowmeter để trích rút đặc trưng, sau đó với các dữ liệu sẽ được gán nhãn tương ứng với từng thể loại như: đối với dữ liệu lành tính là Benign; đối với các dạng xâm nhập mạng sẽ được gán nhãn tương ứng như: Brute force; Buffer Overflow; SQL injection...Kết quả đầu ra của quá trình này là tập hợp các file định dạng .csv tương ứng với các loại dữ liệu đã được gán nhãn.

+ Huấn luyện mô hình mạng nơ ron DNN: các file .csv được chúng tôi sử dụng kỹ thuật tự động lựa chọn theo thứ tự ngẫu nhiên (random) các hàng (raw) trên các file .csv và kết hợp chúng lại (merge) thành 1 file .csv duy nhất, trên đó bao gồm tập hợp tất cả các đặc trưng của các loại dữ liệu lành tính và dữ liệu xâm nhập mạng theo một thứ tự ngẫu nhiên. Chúng tôi sử dụng mô hình học sâu DNN mà đầu vào là file .csv vừa được sinh ra để bắt đầu quá trình huấn luyện theo phương pháp lựa chọn ngẫu nhiên theo tỉ lệ 3/7. Sau đó, sử dụng FastAI để hiện thực hoá mô hình học của DNN và đưa ra kết quả học sâu về phát hiện xâm nhập mạng. Mô hình này đã được chúng tôi tối ưu bằng cách lựa chọn các tham số (hyperparameter) từ 4 thành phần làm trọng tâm bao gồm (categorical embeddings, continuous variables, hidden layers, và output hidden layer) tương ứng với các tham số (Batch Normalization và Rectified Linear Unit - ReLU) để mang lại hiệu quả cao nhất.

- Phát hiện và ngăn chặn xâm nhập mạng sử dụng mô hình học sâu DLID: công cụ phần mềm DeepInspector được chúng tôi xây dựng và triển khai dịch vụ chạy ở chế độ nền (deamon), để tiếp nhận các luồng dữ liệu mạng theo chuẩn PCAP và phân tích, phát hiện xâm nhập sử dụng mô hình DLID đã được huấn luyện và lưu trữ. Tại đây các luồng dữ liệu mạng dưới dạng PCAP được gửi đến DeepInspector theo cơ chế truyền thông đa tiến trình

sử dụng Unix socket. Thông qua CICFlowmeter, dữ liệu PCAP sẽ được trích rút và chuẩn hóa thành tập các đặc trưng, mỗi luồng dữ liệu mạng sẽ được mô hình hoá bởi 79 đặc trưng và 1 nhãn là một trong 11 loại tấn công xâm nhập hoặc luồng sạch. Các đặc trưng và nhãn đã phân loại sẽ được sử dụng trong quá trình huấn luyện mô hình DLID ở trên. Dựa trên mô hình đã được huấn luyện, khi có luồng dữ liệu mạng mới cần phân tích, chúng tôi sẽ tiến hành trích rút 79 đặc trưng, như bước huấn luyện, đưa vào mô hình để tiến hành phát hiện luồng mạng có bị tấn công xâm nhập nào hay không.

- Cập nhật tự động các luật (kết quả phát hiện của mô hình học sâu mạng nơ ron DNN) để làm giàu cho bộ luật có sẵn: sau khi mô hình học sâu DLID phân tích tập các luồng mạng, nếu được xác định có xâm nhập thuộc một 1 trong 11 lớp tấn công xâm nhập, DeepInspector sẽ tự động tiến hành sinh luật và cảnh báo đến người quản trị mạng thông qua hệ thống NetIPS bằng cơ chế Unix socket để NetIPS cập nhật và ghi nhận những luật mới, phục vụ ngăn chặn xâm nhập khi có những luồng dữ liệu mạng tấn công tương tự.

Mô tả vắn tắt các hình vẽ

Hình 1: hình vẽ minh hoạ quy trình chung thu nhận, phân tích luồng dữ liệu mạng để phát hiện và ngăn chặn xâm nhập mạng.

Hình 2: hình vẽ minh hoạ mô hình triển khai thiết bị và hệ thống phòng chống xâm nhập mạng NetIPS.

Hình 3a: hình vẽ minh hoạ phương pháp phát hiện và ngăn chặn xâm nhập ở chế độ nội tuyến sử dụng kết hợp cả luật và mô hình học máy DNN.

Hình 3b: hình vẽ minh hoạ phương pháp phát hiện xâm nhập sử dụng kết hợp cả luật và mô hình học máy DNN.

Hình 4: hình vẽ minh hoạ kiến trúc mô hình phát hiện xâm nhập dựa trên học sâu DLID (Deeplearning based Intrusion Detection).

Hình 5: hình vẽ minh hoạ quy trình chuẩn hoá bộ dữ liệu để huấn luyện và xây dựng mô hình học sâu DLID.

Hình 6: hình vẽ minh hoạ quy trình thu nhận và phân tích các luồng dữ liệu mạng dựa trên mô hình học sâu DLID.

Hình 7: hình vẽ minh hoạ kiến trúc hệ thống phát hiện và ngăn chặn xâm nhập NetIPS dựa trên phương pháp kết hợp sử dụng cả luật và phân tích dựa trên mô hình học sâu áp dụng kèm chiến lược lấy mẫu theo tần suất và thời khoảng.

Hình 8: hình vẽ minh hoạ mô hình kiến trúc phần cứng thiết bị NetIPS.

Hình 9: hình vẽ minh hoạ quy trình thiết lập và triển khai phần mềm NetIPS và phần mềm DeepInspector.

Mô tả chi tiết sáng chế

Sáng chế sẽ được hiểu rõ hơn, và các mục đích, các chi tiết, các dấu hiệu và các ưu điểm khác của sáng chế sẽ rõ ràng hơn từ phần mô tả chi tiết dưới đây. Các phương án thực hiện cụ thể của sáng chế chỉ được đưa ra theo cách minh hoạ không nhằm mục đích hạn chế sáng chế và có dựa vào các hình vẽ kèm theo.

(a) Phương pháp phát hiện xâm nhập kết hợp so khớp luật và sử dụng mô hình học sâu kết hợp chiến lược lấy mẫu theo tần suất và thời khoảng:

Giải thuật thể hiện phương pháp phát hiện xâm nhập kết hợp so khớp luật và sử dụng mô hình học sâu được minh hoạ ở hai Hình 3A và 3B. Trong phương pháp này, khi thiết bị hệ thống NetIPS được triển khai theo mô hình nội tuyến (tương ứng với trường hợp minh hoạ ở hình 3a), luồng dữ liệu mạng từ bên ngoài sẽ được hệ thống NetIPS thu nhận. Sau khi thực hiện các bước giải mã 301, chúng tôi sẽ tiến hành quá trình phát hiện xâm nhập dựa theo hai bộ phát hiện. Bộ phát hiện đầu tiên 302 sử dụng tập luật để phát hiện xâm nhập. Tập luật này được lưu trữ trong các tệp 309 có định dạng luật giống như các luật của phần mềm công cụ Suricata (https://suricata.io/) hay Snort (https://www.snort.org/). Mỗi luật sẽ có một mẫu hoặc một chữ ký riêng cho phép định danh được luồng dữ liệu mạng có phải là luồng bị tấn công xâm nhập hay không. Dựa vào tập luật, trong mô hình nội tuyến là, mỗi luồng dữ liệu mạng sẽ được xác định hệ thống NetIPS thi hành một trong bốn hành động: (i) loại bỏ gói tin (drop); (ii) từ chối gói tin (reject – loại bỏ gói tin và thông báo đến nguồn gửi gói tin); (iii) cho qua nhưng cảnh báo (alert); (iv) và cho qua không cảnh báo (pass).

Với những luồng mạng được bộ so khớp luật xác định các hành động là loại bỏ Drop hay từ chối Reject, hệ thống NetIPS sẽ tiến hành ngăn chặn luồng dữ liệu đầu vào 303. Khi đó, luồng dữ liệu mạng sẽ không được phép đi vào mạng bên trong hệ thống NetIPS.

Trong trường hợp kết quả của bộ so khớp là hành động cảnh báo Alert, hệ thống sẽ ghi nhận cảnh báo 304 và vẫn cho luồng dữ liệu đi vào mạng bên trong của NetIPS. Tương tự, nếu luồng dữ liệu mạng vào trùng với luật có hành động cho qua Pass 305, hệ thống sẽ để luồng đó đi vào mạng bên trong NetIPS.

Với trường hợp không có luật nào khớp với luồng dữ liệu mạng vào, tương ứng với trường hợp còn lại (other), luồng này có thể sẽ được thẩm tra sâu hơn trong bộ phát hiện xâm nhập dựa vào mô hình học sâu DeepInspector của chúng tôi. Việc kích hoạt bộ phân tích sâu này cũng được chúng tôi thiết lập thông qua tham số lựa chọn ***deep_inspecting***

trong cấu hình hệ thống NetIPS: ***deep_inspecting*** có giá trị *"yes"* thì NetIPS sẽ thẩm tra sâu hơn với các luồng dữ liệu nằm ngoài tập luật; giá trị "*no*" thì hệ thống sẽ không tích hợp quá trình thẩm tra sâu 306.

Để thẩm tra sâu với mô hình học máy DNN, chúng tôi sử dụng sáu tham số chính để điều khiển quá trình thẩm tra sâu: ***inspection_frequency, frequency_min, frequency_max*** và ***inspection_interval, interval_min và interval_max***. Các tham số này đều là số tự nhiên có đơn vị đo là mili giây và có ý nghĩa như sau:

- ***inspection_frequency***: tần suất lấy mẫu để phân tích sâu khi giá trị này được xác định lớn hơn 0; khi tham số này có giá trị 0, hệ thống NetIPS sẽ lấy mẫu phân tích sâu với tần suất ngẫu nhiên trong khoảng được xác định từ ***[frequency_min, frequency_max]***. Giá trị mặc định của tần suất lấy mẫu là 2 phút và khoảng mặc định là từ 1 đến 5 phút.
- ***inspection_interval***: khoảng thời gian lấy mẫu để phân tích sâu khi giá trị này lớn hơn 0. Trong trường hợp tham số này có giá trị 0, hệ thống sẽ lấy mẫu để phân tích sâu trong khoảng ngẫu nhiên từ ***[interval_min, interval_max]***. Giá trị mặc định của thời gian lấy mẫu là 20 giây và khoảng mặc định là từ 10 đến 30 giây.

Các tham số này sẽ được lựa chọn và cấu hình trong hệ thống NetIPS. Với quy mô số lượng lớn các luồng dữ liệu mạng (thông lượng từ 10Gbps trở lên), tần suất và thời gian lấy mẫu sẽ quyết định đến hiệu năng của hệ thống NetIPS. Tần suất lấy mẫu quá nhanh sẽ dẫn đến việc hệ thống phải thi hành nhiều lần phát hiện xâm nhập dựa trên mô hình học sâu. Do đó, các tham số này cũng phải được lựa chọn theo năng lực của hệ thống tính toán NetIPS.

Bộ phát hiện xâm nhập DeepInspector được xây dựng từ mô hình học sâu sử dụng mạng nơ ron sâu DNN. Trong bộ phân tích sâu này, luồng dữ liệu mạng được lấy mẫu với định dạng PCAP sẽ được tiến hành trích rút đặc trưng 307. Mỗi luồng mạng sẽ được trích rút 79 đặc trưng thông qua công cụ CICFlowMeter (trong 83 đặc trưng từ luồng mạng được CICFlowmeter xác định, chúng tôi không sử dụng 4 đặc trưng là "Flow ID,Src IP,Src Port, DstIP". Từ 79 đặc trưng còn lại, chúng tôi sử dụng 2 đặc trưng phục vụ cho phân lớp là "DstPort,Protocol", tức biến phân loại; 77 đặc trưng còn lại sẽ được sử dụng làm các biến liên tục trong mô hình mạng DNN.

Tập đặc trưng của các luồng mạng được lấy mẫu sẽ được chuyển đến bộ phát hiện xâm nhập mạng DNN 308. Dựa vào mô hình DNN đã được huấn luyện từ tập dữ liệu chuyên biệt được chúng tôi thu thập và xây dựng, DeepInspector sẽ tiến hành phân lớp từng luồng dữ liệu mạng vào một trong 12 lớp, trong đó 1 lớp là luồng dữ liệu sạch và 11 lớp còn lại tương ứng với 11 kiểu tấn công xâm nhập phổ biến. Nếu luồng dữ liệu mạng nào được phân vào một trong 11 lớp tấn công xâm nhập, DeepInspector sẽ tiến hành gửi

cảnh báo đến bộ sinh cảnh báo 304 và sinh luật mới bổ sung vào tập luật 309 của hệ thống NetIPS.

Trong trường hợp hệ thống NetIPS được triển khai thụ động để phát hiện xâm nhập mạng (hình 3B, mô hình IDS,) các luồng dữ liệu mạng sẽ được bẻ ghi đến thiết bị NetIPS thông qua các cổng SPAN/Mirror của bộ chuyển mạch chính. Lúc đó, toàn bộ các hành động Drop/Reject không thể tác động đến luồng dữ liệu mạng từ bên ngoài vào bên trong. Do đó, phần ngăn chặn xâm nhập 303 (hình 3A) không tham gia vào quá trình hoạt động của NetIPS. Toàn bộ các luật có hành động Drop/Reject đều chuyển thành các cảnh báo gửi đến người quản trị và lưu lại trong hệ thống quản trị NetIPS.

(b) Mô hình học sâu DNN phát hiện xâm nhập mạng:

Trong phương pháp kết hợp mà chúng tôi đề xuất trong sáng chế này, mô hình học sâu sử dụng mạng nơ ron sâu (Deep Neural Network – DNN) được chúng tôi lựa chọn sử dụng. Dựa trên các kết quả thực nghiệm của chúng tôi, việc phát hiện xâm nhập mạng với mô hình học sâu DNN sử dụng kỹ thuật ***tabular_learner*** của khung phát triển fastAI cho phép đạt được độ chính xác tốt nhất, giảm thiểu được tỷ lệ phát hiện nhầm so với các mô hình học sâu khác cũng như các khung phát triển khác như Keras, TensorFlow, Theano, …

Trong mô hình DNN của chúng tôi, có bốn thành phần quan trọng, bao gồm: biến loại (categorical variables), biến liên tục (continuous variables), lớp ẩn (hidden layers) và lớp đầu ra (output layers). Trong đó biến loại là các biến mà có giá trị hữu hạn không phải là số ví dụ như: địa chỉ IP, giao thức, ... Trong khi đó biến liên tục là các biến có giá trị bất kỳ nằm trong một khoảng giá trị xác lập trước.

Bên cạnh các thành phần trên, mô hình kiến trúc có các tham số sau:

- Batch Normalization (chuẩn hóa lô): Chuẩn hóa lô là một trong những phương pháp chuẩn hóa phổ biến nhất trong mô hình học sâu. Nó cho phép huấn luyện mạng nơ-ron sâu nhanh hơn và ổn định hơn bằng cách ổn định sự phân bố của các lớp đầu vào trong quá trình huấn luyện. Chuẩn hóa lô cũng đóng vai trò như một quy trình hóa để giúp giảm tình trạng quá tải. Sử dụng chuẩn hóa theo lô, quá trình huấn luyện mô hình không cần phải sử dụng tính năng loại bỏ bớt thông tin (dropout) do đó không bị mất mát thông tin.
- ReLU: Hàm ReLU thường được sử dụng khi huấn luyện mạng nơ-ron. ReLU lọc các giá trị nhỏ hơn 0 theo công thức $f(x) = \max(0, x)$. Hàm này có nhiều ưu điểm so với $Sigmoid$ và $Tanh$ như tốc độ hội tụ và tính toán nhanh hơn rất nhiều.

Kiến trúc mô hình DNN áp dụng trong phương pháp đề xuất của chúng tôi được xác lập như sau:

- Hai (02) đặc trưng luồng dữ liệu là cổng đích (DstPort) và giao thức (Protocol) được coi như hai biến phân loại 401. Mỗi thuộc tính sẽ đi qua bộ nhúng phân loại (categorical embedding) và loại bớt dữ liệu dư thừa (dropout).
- Bảy bảy (77) đặc trưng liên quan đến luồng dữ liệu mạng (cụ thể gồm: 'Timestamp', 'Flow Duration', 'Tot Fwd Pkts', 'Tot Bwd Pkts', 'TotLen Fwd Pkts', 'TotLen Bwd Pkts', 'Fwd Pkt Len Max', 'Fwd Pkt Len Min', 'Fwd Pkt Len Mean', 'Fwd Pkt Len Std', 'Bwd Pkt Len Max', 'Bwd Pkt Len Min', 'Bwd Pkt Len Mean', 'Bwd Pkt Len Std', 'Flow Byts/s', 'Flow Pkts/s', 'Flow IAT Mean', 'Flow IAT Std', 'Flow IAT Max', 'Flow IAT Min', 'Fwd IAT Tot', 'Fwd IAT Mean', 'Fwd IAT Std', 'Fwd IAT Max', 'Fwd IAT Min', 'Bwd IAT Tot', 'Bwd IAT Mean', 'Bwd IAT Std', 'Bwd IAT Max', 'Bwd IAT Min', 'Fwd PSH Flags', 'Bwd PSH Flags', 'Fwd URG Flags', 'Bwd URG Flags', 'Fwd Header Len', 'Bwd Header Len', 'Fwd Pkts/s', 'Bwd Pkts/s', 'Pkt Len Min', 'Pkt Len Max', 'Pkt Len Mean', 'Pkt Len Std', 'Pkt Len Var', 'FIN Flag Cnt', 'SYN Flag Cnt', 'RST Flag Cnt', 'PSH Flag Cnt', 'ACK Flag Cnt', 'URG Flag Cnt', 'CWE Flag Count', 'ECE Flag Cnt', 'Down/Up Ratio', 'Pkt Size Avg', 'Fwd Seg Size Avg', 'Bwd Seg Size Avg', 'Fwd Byts/b Avg', 'Fwd Pkts/b Avg', 'Fwd Blk Rate Avg', 'Bwd Byts/b Avg', 'Bwd Pkts/b Avg', 'Bwd Blk Rate Avg', 'Subflow Fwd Pkts', 'Subflow Fwd Byts', 'Subflow Bwd Pkts', 'Subflow Bwd Byts', 'Init Fwd Win Byts', 'Init Bwd Win Byts', 'Fwd Act Data Pkts', 'Fwd Seg Size Min', 'Active Mean', 'Active Std', 'Active Max', 'Active Min', 'Idle Mean', 'Idle Std', 'Idle Max', 'Idle Min') được sử dụng như các biến liên tục 402.
- Nhãn lớp phân loại của luồng dữ liệu mạng cũng được đưa cùng với các thuộc tính vào lớp ẩn đầu tiên để huấn luyện 403. Lớp này được cấu thành từ ba khối chuẩn "Linear", "ReLU" và "BatchNorm1D". Từ 80 thuộc tính đầu vào (kể cả thuộc tính nhãn lớp), đầu ra của lớp ẩn thứ nhất được chúng tôi thiết lập thành 400 đặc trưng.
- Đối với lớp ẩn thứ hai 404, cấu trúc lớp này giống như lớp ẩn thứ nhất. Tuy nhiên, từ 400 thuộc tính đầu vào, đầu ra sẽ được chuẩn hoá thành 100.
- Lớp đầu ra cuối cùng 405 đảm nhiệm vai trò phân lớp thông qua bộ lọc tuyến tính *Linear*. Bộ này ánh xạ từ 100 thuộc tính đầu vào thành 1 giá trị duy nhất đại diện cho 1 lớp trong tổng số 12 lớp: 1 lớp luồng sạch (Benign) và 11 lớp tấn công xâm nhập (SQL Injection, Brute Force-Web, Brute Force-XSS, BruteForce-SSH, BruteForce-FTP, DoS-SlowHTTPTest, DoS-Hulk, DoS-Slowloris, DoS-GoldenEye, Infilteration, BotAttack).

(c) Huấn luyện mô hình DLID:

Quy trình huấn luyện mô hình DNN được chúng tôi thể hiện qua hình 5. Dựa trên tập dữ liệu đã được các tổ chức uy tín trên thế giới công bố, chúng tôi đã thu thập được tập dữ

liệu PCAP quy mô lớn phân chia theo 11 loại tấn công xâm nhập và luồng sạch 501. Dữ liệu PCAP mỗi loại tấn công xâm nhập được chúng tôi chia theo thư mục, tên thư mục tương ứng với nhãn của loại tấn công hay luồng sạch 502. Từ bộ dữ liệu đó, chúng tôi sử dụng công cụ CICFlowmeter để tiến hành sinh tập các thuộc tính của mỗi luồng mạng. Nhãn phân loại mỗi luồng mạng được lấy giá trị tương ứng với kiểu tấn công xác định qua tên thư mục chứa dữ liệu PCAP 503. Sau khi chuẩn hoá, loại bỏ đi những thuộc tính dư thừa, chúng tôi được 12 tệp cvs tương ứng với 12 lớp, trong đó mỗi lớp có luồng dữ liệu mạng bị tấn công xâm nhập đều từ 100.000 mẫu; tổng số luồng bị tấn công trên 1.500.00 mẫu; tổng số luồng sạch trên 1.000.000 mẫu. Các tệp được được gộp lại và xáo trộn để hình thành nên bộ dữ liệu phục vụ huấn luyện mô hình DNN của chúng tôi 504. Trong bước tiếp theo, chúng ta tiến hành huấn luyện mô hình DNN, đã được xây dựng như minh hoạ ở Hình 4, dựa trên khung phát triển học sâu FastAI 505. Kết thúc quá trình huấn luyện, chúng tôi lưu lại mô hình đã huấn luyện DLID trong tệp lưu trên thiết bị lưu trữ để có thể tái sử dụng trong công cụ phân tích, phát hiện xâm nhập sử dụng học sâu.

Mô hình DLID sau khi huấn luyện với bộ dữ liệu mà chúng tôi thu thập được đã có khả năng phát hiện và phân loại luồng dữ liệu mạng với độ chính xác (accuracy) đạt 98,47%, độ đo ROC-AUC đạt 98,18% theo phương pháp kiểm thử chéo 5-fold.

(d) Phát hiện và ngăn chặn xâm nhập mạng sử dụng mô hình học sâu DLID:

Từ mô hình học sâu DLID đã được huấn luyện và lưu trữ, chúng tôi đã xây dựng công cụ phần mềm DeepInspector, triển khai dịch vụ chạy ở chế độ nhân (deamon), để tiếp nhận các luồng dữ liệu mạng theo chuẩn PCAP và phân tích, phát hiện xâm nhập. Quy trình thực hiện phân tích, phát hiện xâm nhập được minh hoạ như hình 6. Theo đó, khi hệ thống NetIPS được cấu hình để bật chế độ phát hiện xâm nhập kết hợp cả luật và học máy, các luồng dữ liệu mạng sẽ được thu nhận và gửi đến phần mềm hệ thống DeepInspector theo cơ chế truyền thông đa tiến trình sử dụng Unix socket 601. Dữ liệu PCAP mà DeepInspector thu nhận được sẽ được phân tích, chuyển thành tập các đặc trưng thông qua CICFlowmeter, mỗi luồng dữ liệu mạng sẽ được mô hình hoá thông qua 83 đặc trưng. Tiếp đó, chúng tôi sẽ chuẩn hoá tập đặc trưng này và chỉ giữ lại đúng 79 đặc trưng sử dụng trong quá trình huấn luyện mô hình DLID ở trên 602. Tập đặc trưng này sẽ được phân tích để dự đoán liệu có luồng mạng bị tấn công hay không với mô hình học sâu DLID đã được huấn luyện 603. Kết quả của bước này ta thu được tập các luồng mạng đã được phân loại. Sau đó, nếu có luồng mạng được xác định có xâm nhập với kiểu là một 1 trong 11 lớp đã nêu trên, DeepInspector sẽ tiến hành sinh luật và cảnh báo đến người quản trị mạng thông qua hệ thống NetIPS 604. Luật mới được sinh từ DeepInspector sẽ được lưu vào trong thư mục */etc/netips/rules/* chứa các tập luật của hệ thống NetIPS. Từ đó, DeepInspector sẽ gửi thông tin đến hệ thống NetIPS qua cơ chế Unix socket để NetIPS cập nhật và ghi nhận những

luật mới, phục vụ ngăn chặn xâm nhập khi có những luồng dữ liệu mạng tấn công tương tự sau này 605.

(e) Phương pháp xử lý gói tin ở không gian người dùng:

Với những hệ thống phần mềm truyền thông sử dụng card mạng, gói tin đến sẽ được thu nhận thông qua hàng đợi nhận RX (receive). Từ đó, nó được sao chép vào bộ nhớ chính thông qua cơ chế truy cập bộ nhớ trực tiếp (Direct Memory Access - DMA). Sau đó, hệ thống cần được thông báo về gói tin mới và chuyển dữ liệu vào một bộ đệm được cấp phát đặc biệt (Linux phân bổ các bộ đệm này cho mọi gói tin). Để làm điều này, Linux sử dụng cơ chế ngắt: một ngắt được tạo ra nhiều lần khi một gói tin mới xâm nhập vào hệ thống. Sau đó gói tin cần được chuyển đến không gian người dùng.

Cơ chế trên dẫn đến tình trạng nghẽn khi phải xử lý nhiều gói hơn, do phải sử dụng nhiều tài nguyên hơn và dẫn đến giảm hiệu suất tổng thể của hệ thống. Cụ thể, cơ chế phân bổ gói tin trong nhân cần nhiều chu kỳ truyền dữ liệu từ CPU đến bộ nhớ chính. Cấu trúc dữ liệu ngăn xếp mạng Linux được thiết kế để tương thích với càng nhiều giao thức càng tốt. Cấu trúc quá phức tạp đó cũng dẫn đến quá trình xử lý diễn ra chậm so với việc chỉ xử lý gói tin mạng.

Một yếu tố khác ảnh hưởng tiêu cực đến hiệu suất là xử lý gói tin trong nhân cần nhiều quá trình chuyển đổi ngữ cảnh - context switching. Khi một ứng dụng trong không gian người dùng cần gửi hoặc nhận một gói, nó sẽ thực hiện một cuộc gọi hệ thống. Ngữ cảnh được chuyển sang chế độ nhân và sau đó trở lại chế độ người dùng. Điều này tiêu tốn một lượng đáng kể tài nguyên hệ thống.

Hình 7 minh hoạ phương pháp xử lý gói tin của hệ thống NetIPS ở không gian người dùng trong hệ điều hành Linux. Trong phương pháp này, các gói tin được thu nhận từ card mạng chuyên dụng 701 (thường được gọi là bộ tăng tốc mạng SmartNIC) sẽ được chuyển thẳng đến phần mềm hệ thống NetIPS thông qua thư viện DPDK 704 (không thông qua nhân hệ điều hành). Các cổng mạng trong SmartNIC, lúc bấy giờ, sẽ không còn được quản lý bởi nhân Linux (người dùng sẽ không còn có thể cấu hình, xem thông tin các cổng mạng bằng các công cụ truyền thống như *ifconfig*). Tất cả các giao tiếp giữa NetIPS và card mạng SmartNIC được quản lý bởi bộ điều khiển chế độ thăm dò DPDK (Poll Mode Drivers - PMD). Các trang *hugepages* được cấu hình và phải được cấp phát trước nhằm thay thế cơ chế DMA trong xử lý gói truyền thống.

Từ đó, các giai đoạn chính của quá trình xử lý gói trong NetIPS gồm:

1. Các gói đến sẽ chuyển đến một bộ đệm vòng (ring buffer). NetIPS (thông qua thư viện DPDK) định kỳ kiểm tra vùng đệm này để tìm các gói tin mới.

2. Nếu bộ đệm chứa các bộ mô tả gói mới, NetIPS sẽ tham chiếu đến bộ đệm gói DPDK trong vùng bộ nhớ được cấp phát đặc biệt bằng cách sử dụng các con trỏ trong bộ mô tả gói.
3. Nếu bộ đệm vòng không chứa bất kỳ gói nào, NetIPS sẽ xếp hàng các thiết bị mạng dưới DPDK và sau đó lại tham chiếu đến bộ đệm.

Trong kiến trúc NetIPS, các gói tin thu nhận được từ thư viện DPDK sẽ được chuyển đến bộ giải mã 705. Tại đây, gọi tin sẽ được phân tích sâu để xác định các giao thức, dữ liệu mô tả và gửi đến các bộ phát hiện xâm nhập 707 theo phương pháp kết hợp đã mô tả ở hình 3. Trong bộ phát hiện dựa theo luật, chúng tôi sử dụng kỹ thuật siêu so khớp Hyperscan 706 với định hướng cải thiện hiệu năng quá trình so khớp tập luật (Hyperscan cho hiệu năng so khớp tốt hơn so với một số phương pháp so khớp khác như Aho-Corasick, Boyer-Moore, ...). Trong quá trình phân tích, phát hiện gói tin có dấu hiệu tân công xâm nhập mạng theo phương pháp kết hợp sử dụng so khớp theo luật và mô hình học sâu DNN, phân hệ Detector 707 sẽ triệu gọi đến thư viện Hyperscan và tương tác với công cụ *DeepInspector* thông qua cơ chế truyền thông đa tiến trình sử dụng Unix socket 708. Dữ liệu gói tin theo chuẩn PCAP sẽ được gửi đến DeepInspector để phân tích sâu hơn theo mô hình DLID 709 xem liệu có phải có tấn công xâm nhập mạng hay không. Cơ chế lấy mẫu, tần suất và thời gian lấy mẫu gói tin sẽ được chúng tôi mô tả chi tiết ở mục sau.

Ngoài việc sử dụng SmartNIC để tăng tốc quá trình thu nhận, xử lý gói tin ở không gian người dùng, hệ thống NetIPS còn sử dụng card mạng truyền thống để đảm nhiệm vai trò quản trị hệ thống 702. Card mạng này vẫn sẽ được quản lý bởi nhân hệ điều hành Linux thông qua các bộ điều khiển card mạng truyền thống 703.

Thiết bị phần cứng để hình thành hệ thống NetIPS của chúng tôi được minh hoạ thông qua hình 8. Trong mô hình thiết kế phần cứng NetIPS, chúng tôi đã trang bị bộ tăng tốc mạng thông minh Napatech SmartNIC NT40E3 4x10Gbps 801. Đây là thành phần quan trọng cho phép tiếp nhận được đồng thời 4 luồng dữ liệu mạng có thông lượng đến 10Gbps. Cổng mạng truyền thống được chúng tôi sử dụng card mạng thông thường được trang bị kèm bo mạch chủ với thông lượng 1Gbps 802. Để tăng tốc quá trình phân tích sâu các gói tin thông qua công cụ DeepInspector, chúng tôi sử dụng thêm một card đồ hoạ Testla T4 để tăng tốc độ dự đoán luồng dữ liệu mạng 803. Để có thể có đủ không gian bộ nhớ đệm cho việc xử lý và phân tích các luồng gói tin quy mô lớn, không gian bộ nhớ chính 804 cũng như năng lực xử lý của bộ vi xử lý 805 cũng phải được chú trọng trang bị. Để có thể đạt được năng lực xử lý đến 160 triệu kết nối đồng thời, 1.200.000 kết nối mới mỗi giây, thông lượng IPS đến 40Gbps, hệ thống phần cứng NetIPS đã được trang bị 2 bộ vi xử lý Intel® Xeon Platinum 8160 (24 lõi/48 luồng); 384GB bộ nhớ chính.

(f) Triển khai hệ thống NetIPS phát hiện và ngăn chặn xâm nhập mạng theo phương pháp sử dụng kết hợp cả luật và học sâu kèm chiến lược lấy mẫu theo tần suất và thời khoảng:

Từ hệ thống phần cứng và phần mềm NetIPS đã xây dựng thông qua các bước trên, chúng tôi đã hình thành được hệ thống phát hiện và ngăn chặn xâm nhập mạng NetIPS theo phương pháp kết hợp so khớp dựa trên tập luật và dự báo xâm nhập sử dụng mô hình học sâu DLID. Hình 9 minh hoạ quy trình thiết lập và triển khai phần mềm NetIPS và phần mềm DeepInspector. Khi triển khai hệ thống NetIPS ở chế độ nội tuyến, người quản trị mạng cần phải triển khai thiết bị NetIPS theo mô hình như minh hoạ ở hình 2. Đối với phần mềm NetIPS, người quản trị cần phải cấu hình thiết bị NetIPS trong tệp */etc/netips/netips.conf*. Tập luật được chúng tôi thu thập, phân loại theo từng tệp và lưu trong thư mục */etc/netips/rules/* 901. Các luật được sinh từ phần mềm hệ thống DeepInspector được lưu trong tệp với tên *hunted_threats.rules* trong thư mục này. Mô hình học sâu DLID đã huấn luyện sẽ được chúng tôi lưu trong thư mục */etc/netips/models/* 902. Người quản trị mạng cần lựa chọn mô hình mong muốn thông qua cấu hình các tham số nhóm *deep-inspection:* trong tệp cấu hình *netips.conf*.

```
#Deep Inspection by DLID
deep-inspection:
  deep-inspecting: yes # yes/no
  default-model-path: /etc/netips/models
  model-name: NetIPS-2021
  inspection-frequency: 120000 # milliseconds; 0: random in range [frequency-
min - frequency-max]
  frequency-min: 60000 # milliseconds
  frequency-max: 300000  # milliseconds
  inspection-interval: 20000  # milliseconds
  interval-min: 10000 # milliseconds
  interval-max: 30000 # milliseconds
```

Các tham số quan trọng liên quan đến tần suất và thời gian lấy mẫu để phân tích sâu cũng được xác lập trong nhóm tham số *deep-inspection* của tệp cấu hình NetIPS 903. Sau khi thiết lập xong các tham số này và cấu hình những tham số cần thiết cho phần mềm phát hiện và ngăn chặn xâm nhập mạng NetIPS, chúng ta cần tiến hành khởi chạy NetIPS 904 và dịch vụ phân tích sâu *DeepInspector* 905.

Yêu cầu bảo hộ

1. Phương pháp phát hiện, ngăn chặn xâm nhập mạng sử dụng luật và mô hình học sâu kết hợp chiến lược lấy mẫu theo tần suất và thời khoảng, bao gồm các bước sau:

(a) sử dụng card mạng chuyên dụng băng thông lớn (701) để thu nhận luồng dữ liệu mạng và chuyển thẳng đến phần mềm hệ thống phát hiện và ngăn chặn xâm nhập mạng để xử lý trong không gian người dùng thông qua kỹ thuật DPDK (704),

(b) luồng dữ liệu mạng, sau khi qua bộ giải mã, sẽ được chuyển đến hai bộ phát hiện xâm nhập để phân tích:

- bộ phát hiện dựa trên luật, với tập luật gồm các dấu hiệu và chữ ký đặc trưng của các tấn công xâm nhập, và phương pháp siêu so khớp Hyperscan để kiểm tra luồng mạng có bị tấn công xâm nhập hay không (706);

- bộ phát hiện dựa trên mô hình học sâu DLID (709), được gọi là DeepInspector, với mô hình kiến trúc gồm 2 thuộc tính phân loại [DstPort, Protocol] và 77 thuộc tính đặc trưng luồng mạng; hai lớp ẩn trong đó mỗi lớp đều có ba khối chuẩn "Linear", "ReLU" và "BatchNorm1D" với đầu vào và đầu ra của mỗi lớp lần lượt là [80, 400] và [400, 100]; lớp đầu ra đảm nhiệm vai trò phân lớp thông qua bộ lọc tuyến tính *Linear* ánh xạ từ 100 thuộc tính đầu vào thành giá trị duy nhất đại diện cho 1 trong 12 lớp: 1 lớp luồng sạch [Benign] và 11 lớp tấn công xâm nhập [SQL Injection, Brute Force-Web, Brute Force-XSS, BruteForce-SSH, BruteForce-FTP, DoS-SlowHTTPTest, DoS-Hulk, DoS-Slowloris, DoS-GoldenEye, Infilteration, BotAttack],

(c) xử lý kết quả phát hiện xâm nhập từ cả hai bộ phát hiện (304): loại bỏ/từ chối luồng nếu so khớp trùng với luật tương ứng loại bỏ/từ chối; cảnh báo đến quản trị mạng nếu luồng mạng trùng với luật cảnh báo hoặc bộ phát hiện dựa trên mô hình học sâu DLID trả về kết quả phân tích nằm trong lớp có tấn công xâm nhập.

2. Phương pháp phát hiện, ngăn chặn xâm nhập mạng sử dụng luật và mô hình học sâu kết hợp chiến lược lấy mẫu theo tần suất và thời khoảng theo điểm 1, trong đó bước xác định phân tích sâu luồng dữ liệu mạng bởi bộ phát hiện DeepInspector được triển khai theo các bước như sau:

(a) chỉ những luồng mạng không bị loại bỏ, từ chối, cảnh báo và cho qua (drop/reject/alert/pass) mới được gửi đến bộ thẩm tra sâu hơn (306) trong bộ phát hiện xâm nhập dựa vào mô hình học sâu DeepInspector;

(b) để nâng cao hiệu năng quá trình phân tích sâu luồng mạng với thông lượng mạng quy mô lớn, những luồng mạng được gửi đến DeepInspector theo chiến lược lấy mẫu dựa trên tần xuất và thời khoảng lấy mẫu. Chiến lược này được thiết lập thông qua sáu tham số chính trong hệ thống phát hiện và ngăn chặn xâm nhập: *inspection_frequency, frequency_min, frequency_max* và *inspection_interval, interval_min, interval_max*. Các tham số này đều là số tự nhiên có đơn vị đo là mili giây, trong đó inspection_frequency là tần suất lấy mẫu để phân tích sâu khi giá trị này được xác định lớn hơn 0; khi tham số này có giá trị 0, hệ thống sẽ lấy mẫu phân tích sâu với tần suất ngẫu nhiên trong khoảng được xác định từ [frequency_min, frequency_max]; giá trị mặc định của tần suất lấy mẫu là 2 phút và khoảng mặc định là từ 1 đến 5 phút. Thời khoảng lấy mẫu được xác lập trong tham số inspection_interval; trong trường hợp tham số này có giá trị 0, hệ thống sẽ lấy mẫu để phân tích sâu trong khoảng ngẫu nhiên từ [interval_min, interval_max]; giá trị mặc định của thời gian lấy mẫu là 20 giây và khoảng mặc định là từ 10 đến 30 giây.

3. Phương pháp phát hiện, ngăn chặn xâm nhập mạng sử dụng luật và mô hình học sâu theo điểm 1, trong đó còn bao gồm bước huấn luyện mô hình mạng DLID được tiến hành theo các bước như sau:

(a) xây dựng và thu thập dữ liệu huấn luyện từ hệ thống mạng mô phỏng bằng cách sử dụng tấn công xâm nhập mạng theo nhiều 11 loại tấn công khác nhau và máy tính người dùng để tạo ra các kết nối, dữ liệu lành tính đến các máy chủ của hệ thống mạng mô phỏng như: hệ thống thư điện tử, hệ thống website, ….; mỗi loại, hình thức, dạng dữ liệu được xây dựng thành các kịch bản để thu thập dữ liệu và lưu các tệp theo định dạng PCAP trong thư mục tương ứng với loại tấn công xâm nhập;

(b) thu thập các tập dữ liệu PCAP phục vụ huấn luyện từ các nguồn tin cậy của các tổ chức nghiên cứu trên thế giới; phân nhóm các tệp theo 11 loại tấn công xâm nhập có thể có và lưu trong thư mục tương ứng với tên loại tấn công;

(c) trích rút 79 đặc trưng của các luồng dữ liệu mạng trong toàn bộ tập dữ liệu PCAP đã thu thập được từ hai bước trên với công cụ CICFlowmeter; lưu các tệp kết quả trích rút đặc trưng và gán nhãn tương ứng với tên loại tấn công hay luồng sạch;

(d) chuẩn hoá để đảm bảo có được 02 thuộc tính phân loại và 77 thuộc tính còn lại là liên tục; sau đó, gộp các tệp và tiến hành xáo trộn ngẫu nhiên và lưu kết quả vào trong tệp có định dạng CSV;

(e) sử dụng tệp CSV đã lưu đến tiến hành huấn luyện mô hình học sâu phát hiện xâm nhập DLID;

(f) lưu lại mô hình học sâu DLID đã được huấn luyện để phục vụ cho bộ phân tích sâu DeepInspector.

4. Phương pháp phát hiện, ngăn chặn xâm nhập mạng sử dụng luật và mô hình học sâu kết hợp chiến lược lấy mẫu theo tần suất và thời khoảng theo điểm 1, trong đó bước phân tích sâu luồng dữ liệu mạng ở DeepInspector như sau:

(a) hệ thống phát hiện và ngăn chặn xâm nhập mạng, theo tần suất và thời khoảng lấy mẫu, thu nhận luồng mạng theo chuẩn PCAP và gửi đến bộ DeepInspector sử dụng cơ chế truyền thông liên tiến trình Unix socket (708);

(b) công cụ DeepInspector triệu gọi CICFlowmeter (307) để trích 79 đặc trưng của mỗi luồng dữ liệu mạng PCAP đã nhận được;

(d) tiến hành dự đoán xâm nhập mạng với tập đặc trưng luồng mạng dựa trên mô hình DLID đã huấn luyện (308);

(e) sinh luật xử lý xâm nhập và gửi cảnh báo đến người quản trị mạng tương ứng với kết quả dự đoán xâm nhập;

(f) ghi nhận và tiến hành cập nhật tập luật của hệ thống phát hiện và ngăn chặn xâm nhập (309).

5. Hệ thống phát hiện, ngăn chặn xâm nhập mạng sử dụng luật và mô hình học sâu kết hợp chiến lược lấy mẫu theo tần suất và thời khoảng, bao gồm:

(a) bộ tăng tốc mạng SmartNIC (701) cho phép tiếp nhận luồng mạng có băng thông lớn từ 10Gbps kết hợp và đảm bảo các gói tin được thu nhận được chuyển thẳng đến phần mềm hệ thống NetIPS thông qua thư viện DPDK (704);

(b) bộ giải mã bao gồm:

- bộ xử lý đồ hoạ GPU để tăng tốc quá trình phân tích sâu luồng mạng,

- bộ phát hiện dựa trên luật, với tập luật gồm các dấu hiệu và chữ ký đặc trưng của các tấn công xâm nhập, và phương pháp siêu so khớp Hyperscan để kiểm tra luồng mạng có bị tấn công xâm nhập hay không (706);

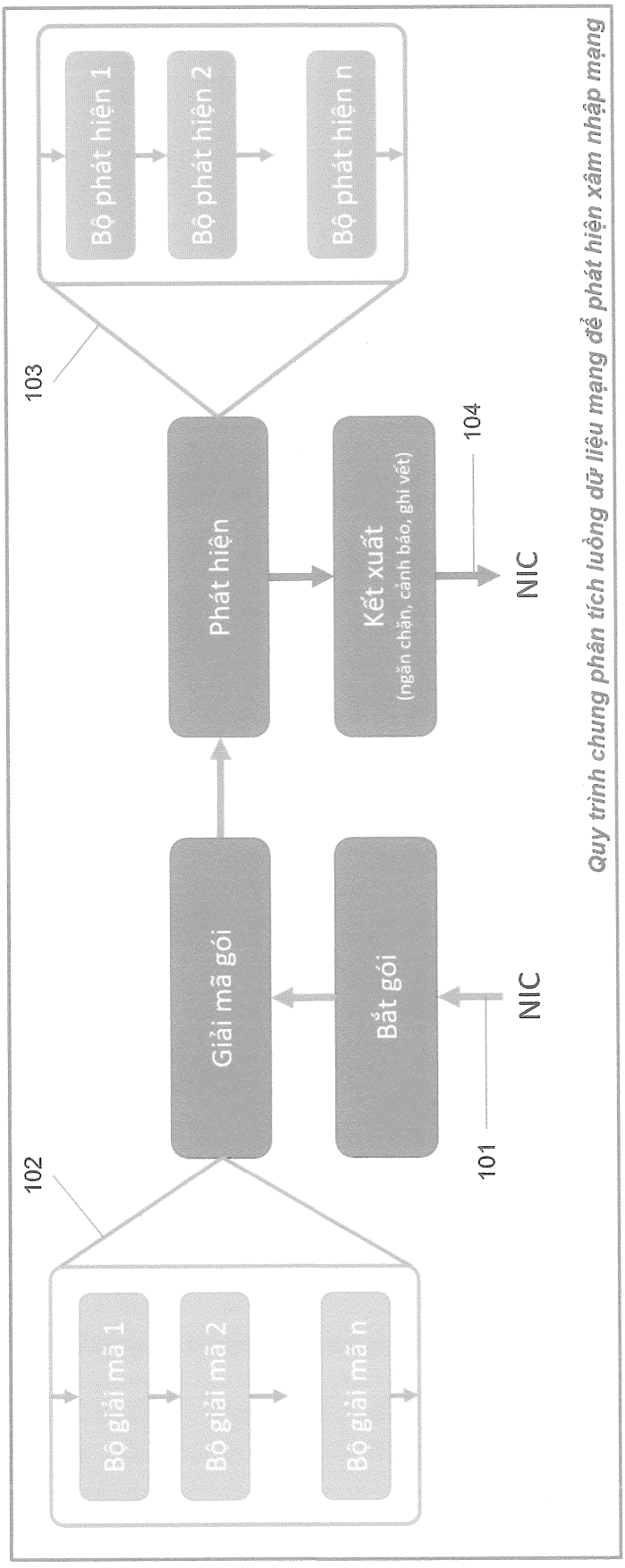
- bộ phát hiện dựa trên mô hình học sâu DLID (709), được gọi là DeepInspector, với mô hình kiến trúc gồm 2 thuộc tính phân loại [DstPort, Protocol] và 77 thuộc tính đặc trưng luồng mạng; hai lớp ẩn trong đó mỗi lớp đều có ba khối chuẩn "Linear", "ReLU" và "BatchNorm1D" với đầu vào và đầu ra của mỗi lớp lần lượt là [80, 400] và [400, 100]; lớp đầu ra đảm nhiệm vai trò phân lớp thông qua bộ lọc

tuyến tính *Linear* ánh xạ từ 100 thuộc tính đầu vào thành giá trị duy nhất đại diện cho 1 trong 12 lớp: 1 lớp luồng sạch [Benign] và 11 lớp tấn công xâm nhập [SQL Injection, Brute Force-Web, Brute Force-XSS, BruteForce-SSH, BruteForce-FTP, DoS-SlowHTTPTest, DoS-Hulk, DoS-Slowloris, DoS-GoldenEye, Infilteration, BotAttack],

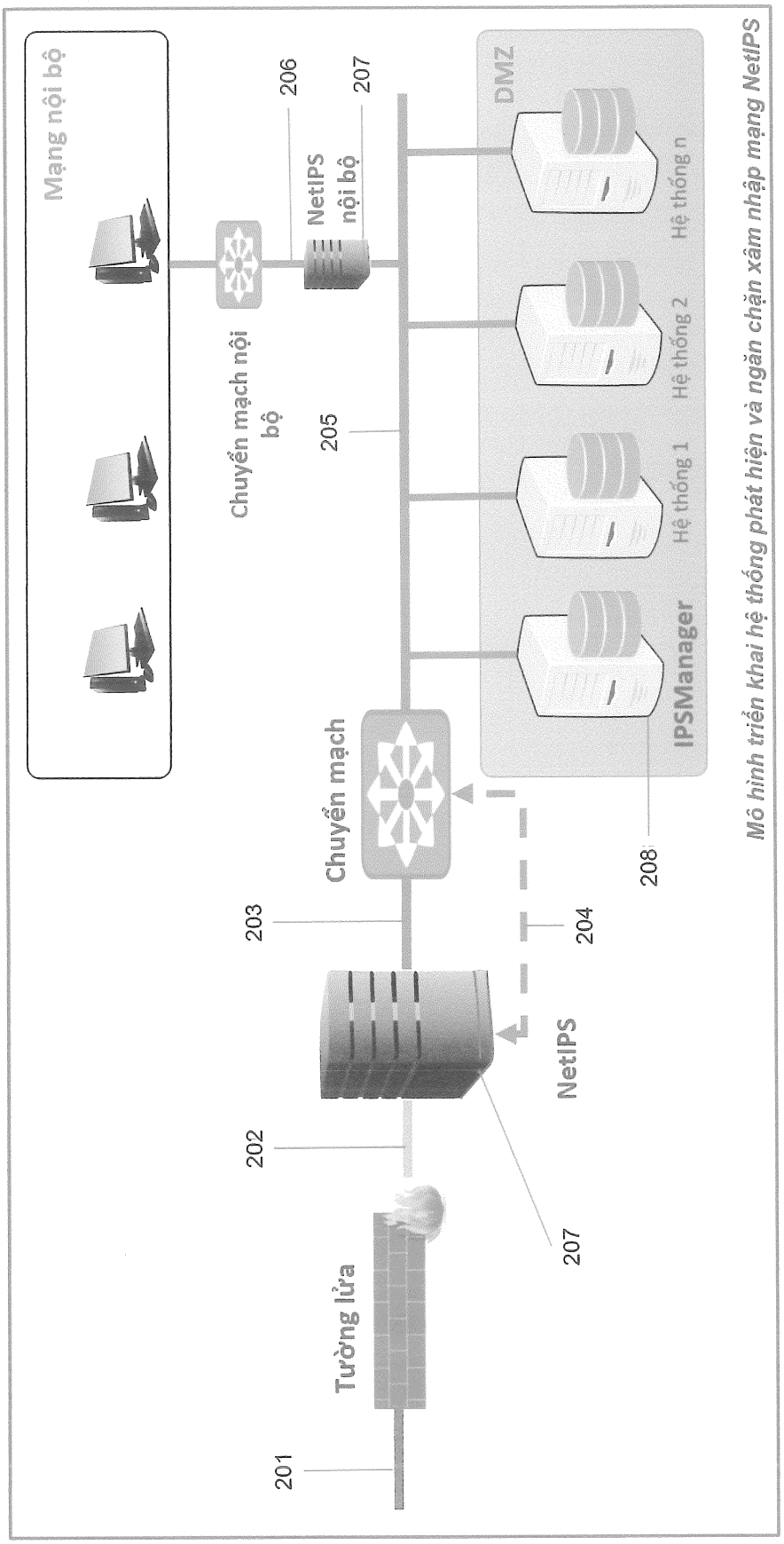
(c) bộ vi xử lý hiệu năng cao (2 chip, mỗi chip có tối thiểu 12 lõi/24 luồng) để xử lý kết quả phát hiện xâm nhập từ cả hai bộ phát hiện (304): loại bỏ/từ chối luồng nếu so khớp trùng với luật tương ứng loại bỏ/từ chối; cảnh báo đến quản trị mạng nếu luồng mạng trùng với luật cảnh báo hoặc bộ phát hiện dựa trên mô hình học sâu DLID trả về kết quả phân tích nằm trong lớp có tấn công xâm nhập, (c) bộ nhớ chính có dung lượng lớn (từ 256GB trở lên),

(d) bộ nhớ chính có dung lượng lớn (từ 256GB trở lên).

Hình 1

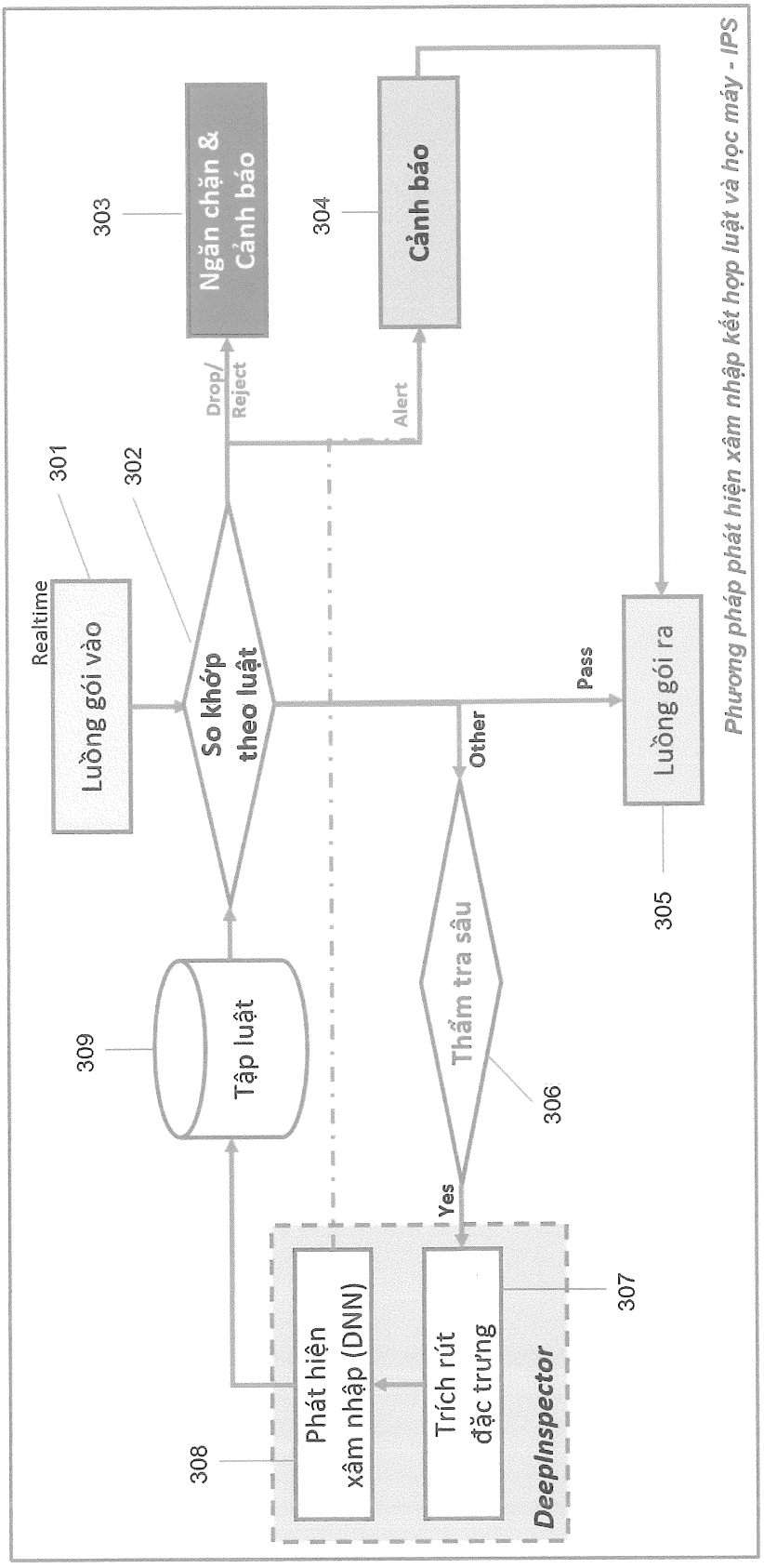

Quy trình chung phân tích luồng dữ liệu mạng để phát hiện xâm nhập mạng
NIC
101
Bắt gói
Giải mã gói
102
Bộ giải mã 1
Bộ giải mã 2
Bộ giải mã n
Phát hiện
103
Bộ phát hiện 1
Bộ phát hiện 2
Bộ phát hiện n
Kết xuất
(ngăn chặn, cảnh báo, ghi vết)
104
NIC

Hình 2

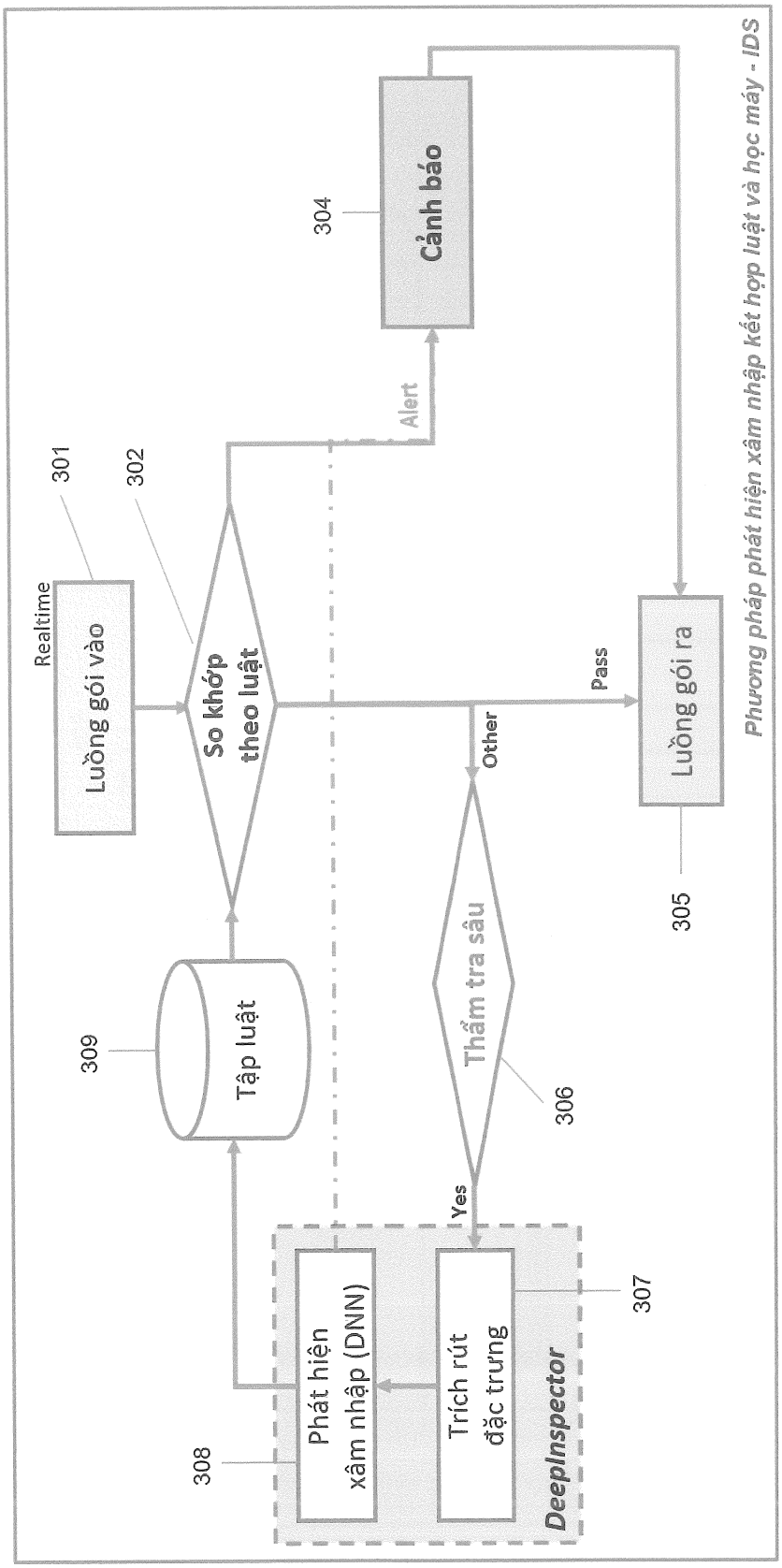


Mô hình triển khai hệ thống phát hiện và ngăn chặn xâm nhập mạng NetIPS

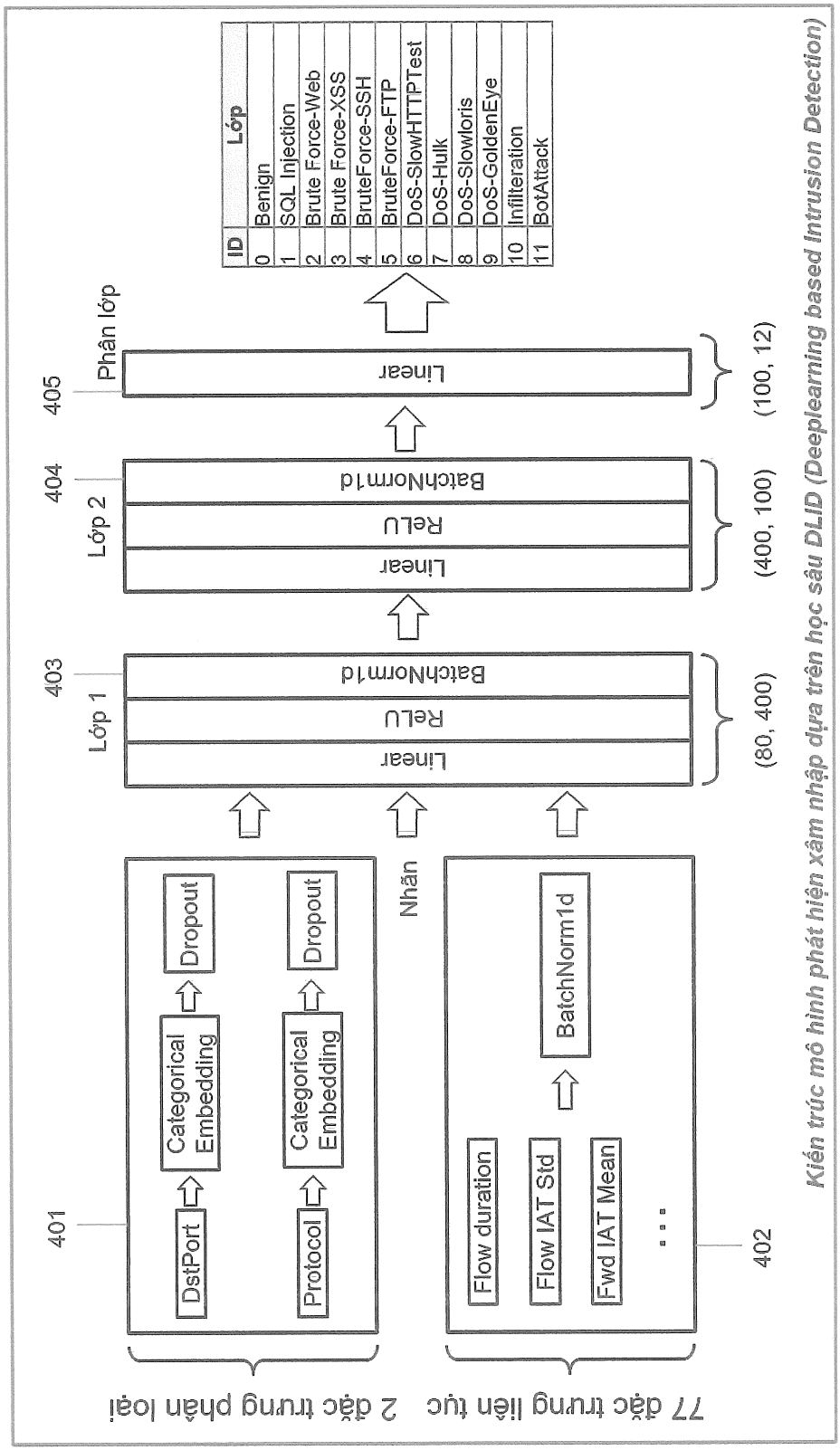
Hình 3A


Phương pháp phát hiện xâm nhập kết hợp luật và học máy - IPS
Realtime
Luồng gói vào
301
So khớp theo luật
302
Drop/ Reject
Alert
Ngăn chặn & Cảnh báo
303
Cảnh báo
304
Pass
Luồng gói ra
305
Other
Thẩm tra sâu
306
Yes
Tập luật
309
DeepInspector
Phát hiện xâm nhập (DNN)
Trích rút đặc trưng
307
308

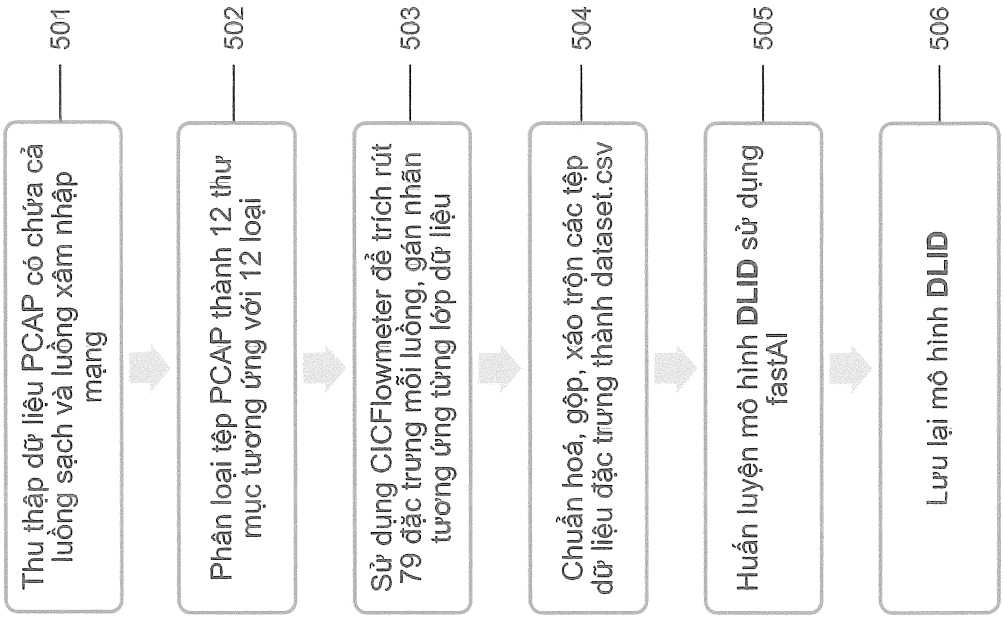
Hình 3B


Phương pháp phát hiện xâm nhập kết hợp luật và học máy - IDS
Realtime
Luồng gói vào
301
So khớp theo luật
302
Alert
Cảnh báo
304
Pass
Luồng gói ra
305
Other
Thẩm tra sâu
306
Yes
Tập luật
309
DeepInspector
Phát hiện xâm nhập (DNN)
Trích rút đặc trưng
307
308

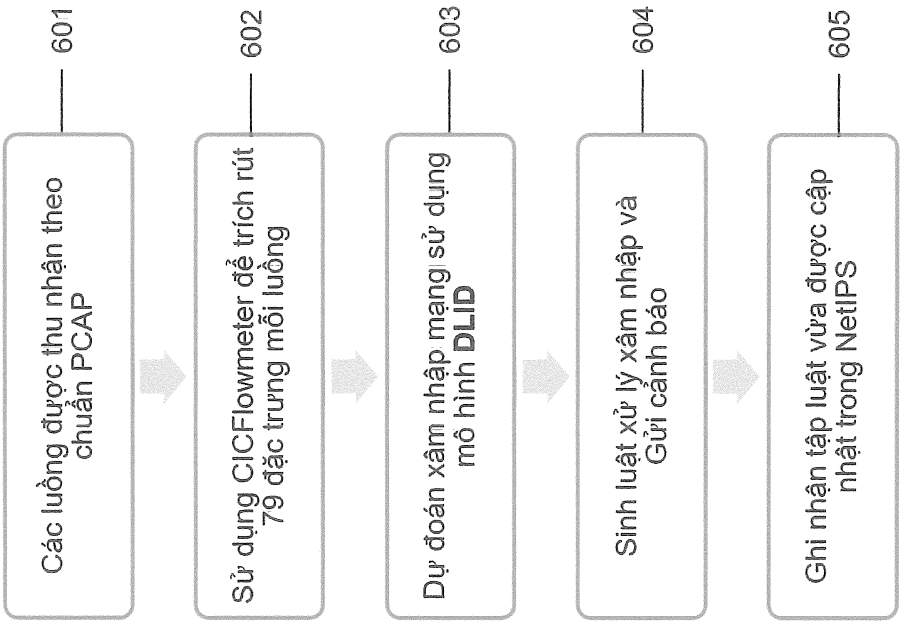
Hình 4


77 đặc trưng liên tục
2 đặc trưng phân loại
Flow duration
Flow IAT Std
Fwd IAT Mean
...
BatchNorm1d
Protocol
DstPort
Categorical Embedding
Categorical Embedding
Dropout
Dropout
Nhãn
Lớp 1
Linear
ReLU
BatchNorm1d
Lớp 2
Linear
ReLU
BatchNorm1d
Phân lớp
Linear
401
402
403
404
405
(80, 400)
(400, 100)
(100, 12)
ID
Lớp
0 Benign
1 SQL Injection
2 Brute Force-Web
3 Brute Force-XSS
4 BruteForce-SSH
5 BruteForce-FTP
6 DoS-SlowHTTPTest
7 DoS-Hulk
8 DoS-Slowloris
9 DoS-GoldenEye
10 Infilteration
11 BotAttack
Kiến trúc mô hình phát hiện xâm nhập dựa trên học sâu DLID (Deeplearning based Intrusion Detection)

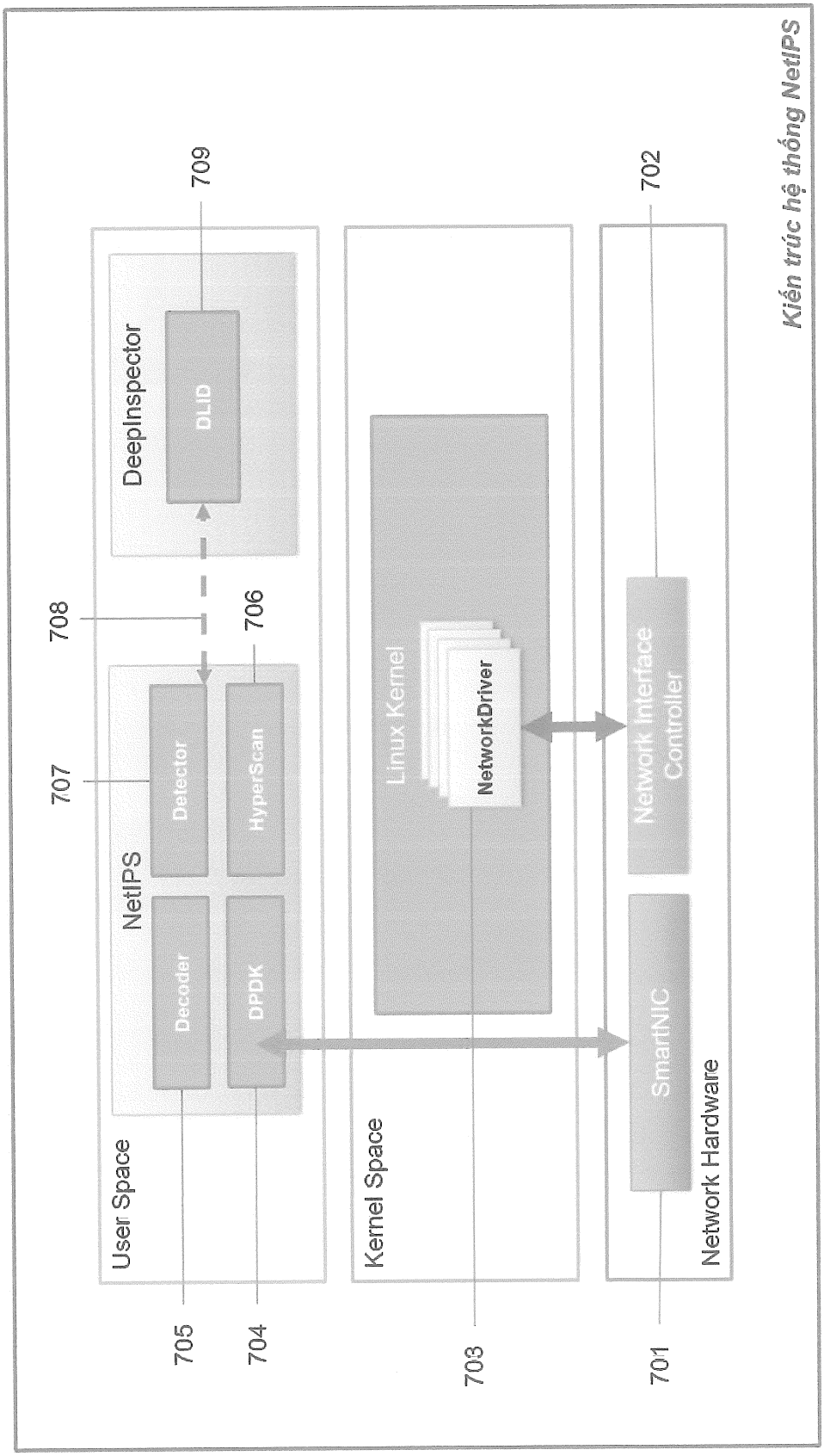
Hình 5


Thu thập dữ liệu PCAP có chứa cả luồng sạch và luồng xâm nhập mạng
501
Phân loại tệp PCAP thành 12 thư mục tương ứng với 12 loại
502
Sử dụng CICFlowmeter để trích rút 79 đặc trưng mỗi luồng, gán nhãn tương ứng từng lớp dữ liệu
503
Chuẩn hoá, gộp, xáo trộn các tệp dữ liệu đặc trưng thành dataset.csv
504
Huấn luyện mô hình DLID sử dụng fastAI
505
Lưu lại mô hình DLID
506

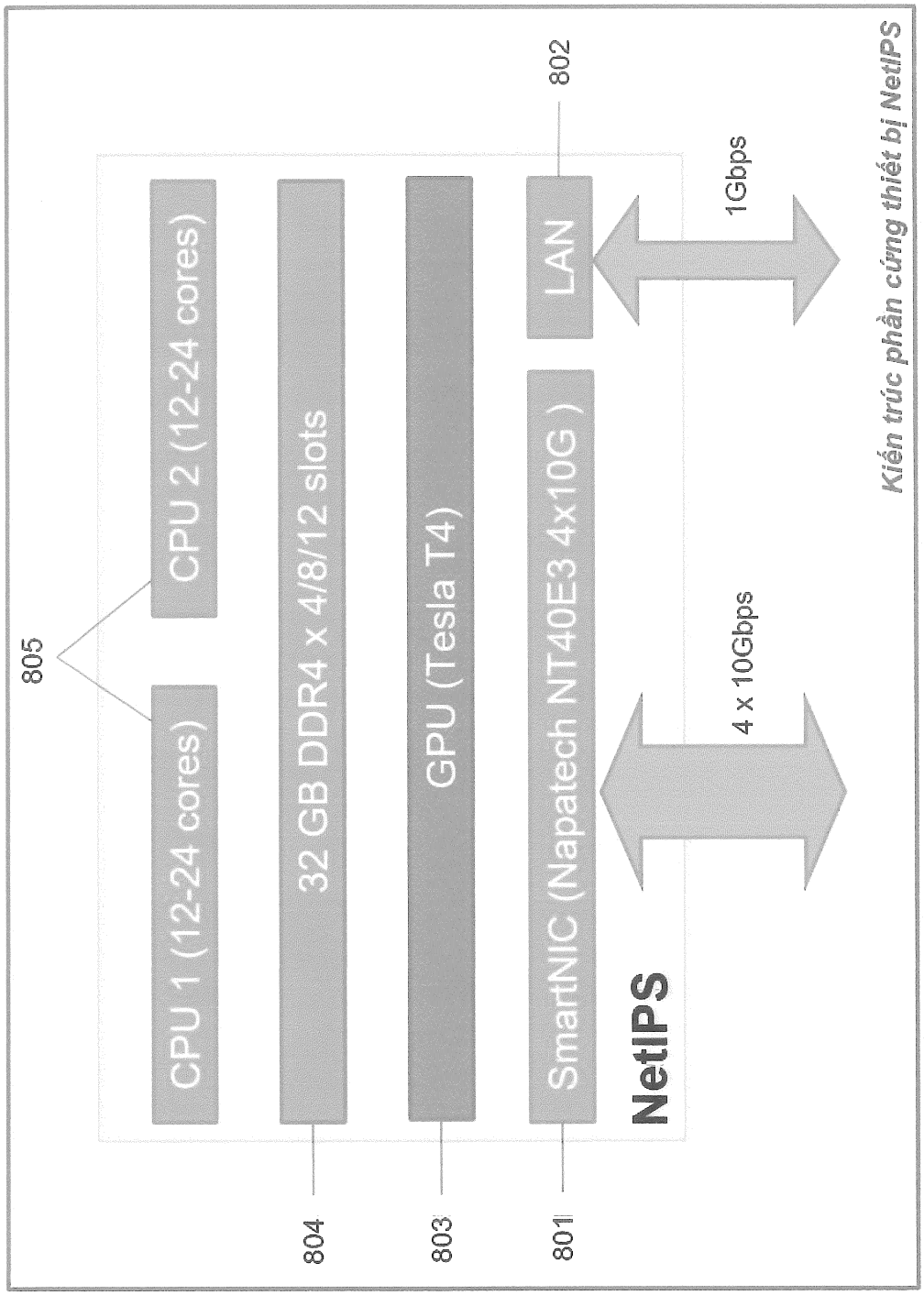
Hình 6


Các luồng được thu nhận theo chuẩn PCAP
601
Sử dụng CICFlowmeter để trích rút 79 đặc trưng mỗi luồng
602
Dự đoán xâm nhập mạng sử dụng mô hình DLID
603
Sinh luật xử lý xâm nhập và Gửi cảnh báo
604
Ghi nhận tập luật vừa được cập nhật trong NetIPS
605

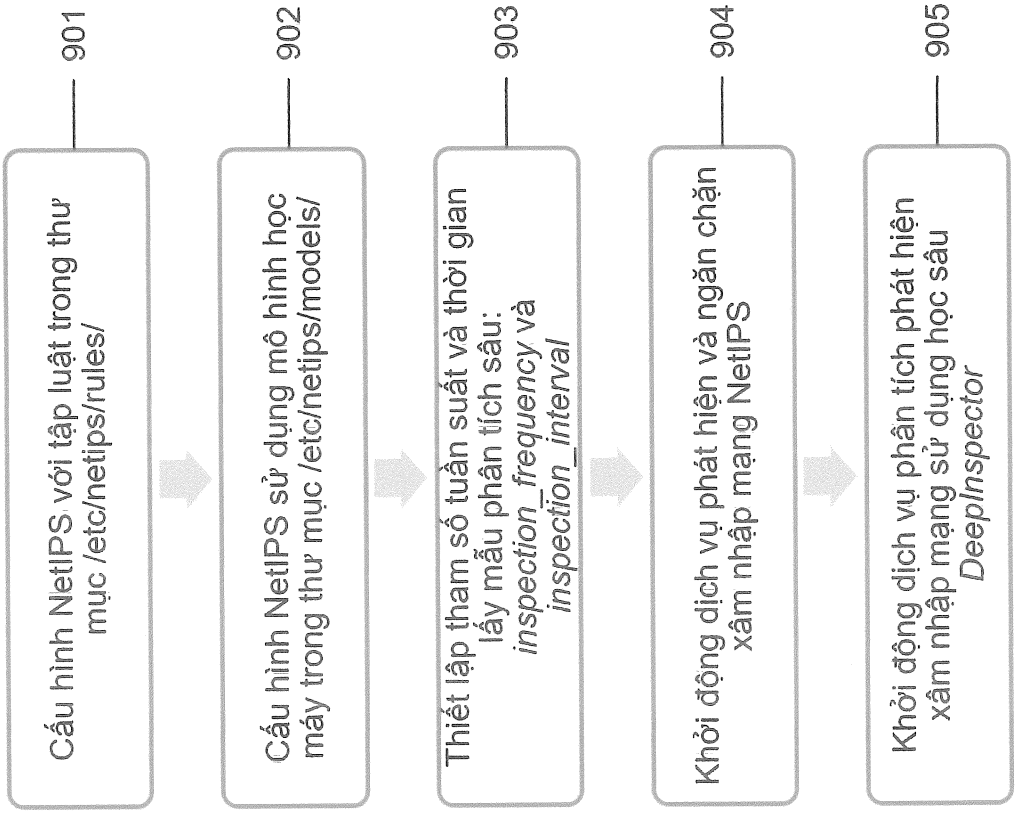
Hình 7


Kiến trúc hệ thống NetIPS
User Space
NetIPS
Decoder
DPDK
Detector
HyperScan
DeepInspector
DL ID
Kernel Space
Linux Kernel
NetworkDriver
Network Hardware
SmartNIC
Network Interface Controller
701
702
703
704
705
706
707
708
709

Hình 8


805
CPU 1 (12-24 cores)
CPU 2 (12-24 cores)
32 GB DDR4 x 4/8/12 slots
GPU (Tesla T4)
SmartNIC (Napatech NT40E3 4x10G)
LAN
802
1Gbps
4 x 10Gbps
NetIPS
804
803
801
Kiến trúc phần cứng thiết bị NetIPS

Hình 9


901
Cấu hình NetIPS với tập luật trong thư mục /etc/netips/rules/
902
Cấu hình NetIPS sử dụng mô hình học máy trong thư mục /etc/netips/models/
903
Thiết lập tham số tuần suất và thời gian lấy mẫu phân tích sâu: inspection_frequency và inspection_interval
904
Khởi động dịch vụ phát hiện và ngăn chặn xâm nhập mạng NetIPS
905
Khởi động dịch vụ phân tích phát hiện xâm nhập mạng sử dụng học sâu DeepInspector