

NÂNG CAO HIỆU QUẢ PHÁT HIỆN TẤN CÔNG APT DỰA TRÊN HỌC SÂU KẾT HỢP

Nguyễn Trung Thành*, Nguyễn Hoa Cường†

*Khoa An toàn thông tin, Học viện Công nghệ Bưu chính Viễn thông

† Khoa An toàn thông tin, Học viện Công nghệ Bưu chính Viễn thông

Tóm tắt: Tấn công APT là kỹ thuật tấn công cao cấp, dai dẳng, có chủ đích. Kỹ thuật tấn công này gây ra nhiều thiệt hại lớn cho tổ chức, cơ quan nhà nước. Do đó, nhiệm vụ phát hiện và cảnh báo sớm kỹ thuật tấn công này rất cần thiết hiện nay. Để nâng cao hiệu quả của quá trình phát hiện tấn công APT, bài báo này đề xuất một hướng tiếp cận mới dựa trên các mô hình học sâu kết hợp. Quy trình phát hiện tấn công APT-IP được đề xuất trong nghiên cứu này: Toàn bộ dữ liệu về lưu lượng mạng sẽ tiền xử lý, phân tích thông qua mạng học sâu kết hợp mạng CNN¹, sau đó các dữ liệu này không dùng phân loại mà tiếp tục được phân tích và đánh giá thông qua mạng BiLSTM². Cuối cùng, dữ liệu sau khi xử lý bởi mạng BiLSTM sẽ tiến hành phân loại để xác định tấn công APT-IP.

Từ khóa: APT; Phát hiện tấn công APT; Học sâu kết hợp; Hành vi bất thường.

I. MỞ ĐẦU

A. Tấn công APT và vấn đề phát hiện

Tấn công APT là kỹ thuật tấn công cao cấp, giai đoạn và có chủ đích [1, 2, 3]. Phần lớn các chiến dịch của tấn công APT đều được các tổ chức chính phủ đứng đầu sau để hỗ trợ về nhân lực và vật chất. Chính vì vậy, vấn đề phát hiện và cảnh báo sớm chiến dịch tấn công này đang rất cần thiết hiện nay. Trong các nghiên cứu [1, 2, 3] đã liệt kê một số hướng tiếp cận cho phát hiện tấn công APT dựa trên đặc điểm, quy trình và vòng đời của chúng. Đặc biệt, để phát hiện dấu hiệu tấn công APT thì các hướng tiếp cận thường tập trung vào tìm kiếm, phân tích và đánh giá các dấu hiệu và hành vi của mã độc APT dựa trên các bộ dữ liệu về mạng sử dụng các kỹ thuật học máy. Bên cạnh đó, trong nghiên cứu [1, 4] đã chứng minh được rằng bộ dữ liệu về lưu lượng mạng được đánh giá là bộ dữ liệu tốt cho phát hiện tấn công APT vì bộ dữ liệu này ghi nhận nhiều dấu hiệu và hành vi của chúng.

Trong nghiên cứu [1] đã đề xuất một số phương pháp và cách thức khác nhau cho phát hiện tấn công APT bao gồm: phát hiện hành vi bất thường, phân tích đồ thị.... Chúng tôi nhận thấy rằng, các hướng tiếp cận và đề xuất như vậy đã mang lại hiệu quả tốt, tuy nhiên, trước thực trạng dữ liệu giám sát ngày càng phong phú và đa dạng thì

cần phải có những thay đổi về mặt xử lý và phân tích mới có thể phù hợp với nhiệm vụ phát hiện APT hiện nay. Chính vì vậy, trong bài báo này, chúng tôi đề xuất một hướng tiếp cận mới cho phát hiện tấn công APT dựa trên phân tích hành vi bất thường của lưu lượng mạng sử dụng mô hình học sâu kết hợp CNN-BiLSTM. Để thực hiện nhiệm vụ phát hiện dấu hiệu tấn công APT, các mô hình học sâu kết hợp do chúng tôi xây dựng sẽ thực hiện 2 giai đoạn chính bao gồm: i) trích xuất thuộc tính IP dựa trên flow: đối với giai đoạn này, chúng tôi sẽ tiến hành phân tích lưu lượng mạng thành các mạng flow theo từng địa chỉ IP rồi sử dụng mô hình học sâu kết hợp để trích xuất thuộc tính IP dựa trên hành vi của mạng flow; ii) phân loại APT IP: dựa trên các thuộc tính của IP được trích xuất trong nhiệm vụ (i), chúng tôi sẽ tiến hành phân loại để xác định các IP tấn công APT và IP bình thường.

B. Tính mới và khoa học trong bài báo

Đóng góp cơ bản và quan trọng trong bài báo này chính là mô hình phân tích, trích xuất và phân loại APT IP. Theo đó, các tác giả đã đề xuất trong bài báo này một mô hình học sâu kết hợp mới, mang lại hiệu quả cao cho quá trình trích xuất hành vi bất thường của IP để từ đó hỗ trợ nâng cao hiệu quả quá trình phân loại APT IP.

II. CÁC NGHIÊN CỨU LIÊN QUAN

Chu và các cộng sự [5] đã tiến hành thực nghiệm phát hiện tấn công APT dựa trên bộ dữ liệu NSL-KDD sử dụng mạng nhiều tầng truyền thẳng (Multilayer Perceptron - MLP) và thuật toán hỗ trợ máy học (Support vector Machine - SVM). Trong quá trình thực nghiệm các tác giả đã tiến hành thay đổi một số tham số trong nhân của 2 thuật toán MLP và SVM. Kết quả phát hiện tấn công APT dựa trên bộ dữ liệu thực nghiệm bằng mô hình MLP đã mang lại kết quả tốt hơn trong một số trường hợp cụ thể. Tương tự cách tiếp cận trong [5] Nkiruka Eke [6] đã sử dụng các bộ dữ liệu KDD 99 cho phát hiện tấn công APT dựa trên thuật toán học sâu có bộ nhớ ngắn hạn (Long Short-Term Memory- LSTM). Ngoài ra Joloudari [7] đã đề sử dụng thuật toán cây quyết định C5.0 cho phát hiện tấn công APT dựa trên bộ dữ liệu NSL-KDD. Kết quả thực nghiệm cho thấy thuật toán học sâu 6 lớp mang lại kết quả tốt hơn so với các thuật toán cây quyết định C5.0. Độ chính xác của các phương pháp này tương ứng là 98.85%, 95.64%, 88.37%. Bên cạnh đó, thuật toán học sâu 6 lớp do tác giả đề xuất cũng mang lại hiệu quả cao khi sai số chỉ là 1.13%. Tương tự như vậy, Cosimo [8] đã đề xuất mạng Autoencoder cho phát hiện tấn công mạng dựa trên bộ dữ

Tác giả liên hệ: Nguyễn Hoa Cường,

Email: cuongnh@ptit.edu.vn

Đến tòa soạn: 9/2023, chỉnh sửa: 10/2023, chấp nhận đăng: 11/2023.

¹ Mạng Neural tích chập (Convolutional Neural Network)

² Mạng Mạng nơ ron hồi tiếp hai chiều (Bidirectional recurrent neural network)

liệu NSL-KDD. Trong phần thực nghiệm mạng Autoencoder đã thể hiện sự vượt trội cho nhiệm vụ phát hiện tấn công APT so với các mô hình và thuật toán khác. Tuy nhiên, chúng tôi nhận thấy rằng, bộ dữ liệu NSL-KDD và KDD 99 là những bộ dữ liệu đã được chuẩn hóa và cân bằng về mặt dữ liệu sạch và dữ liệu tấn công, nếu tác giả sử dụng thuật toán rừng ngẫu nhiên thì kết quả có thể cao hơn thuật toán học sâu của tác giả đề xuất. Bên cạnh đó trong nghiên cứu của Branka [2] đã chứng minh việc sử dụng bộ dữ liệu KDD 99 cho nhiệm vụ phát hiện tấn công APT không còn thích hợp với nhu cầu thực tế. Ngoài ra, trong nghiên cứu [9], Sai Charan và các cộng sự cũng đề xuất sử dụng mô hình mạng LSTM cho phát hiện tấn công APT trong các hệ thống ngân hàng. Cụ thể, dựa trên nền tảng công nghệ dữ liệu lớn, các tác giả thực hiện ý tưởng sử dụng mô hình LSTM cho phát hiện tấn công APT dựa trên từng giai đoạn phát triển của chúng. Trong phần thực nghiệm, các tác giả đã tiến hành đánh giá sự hiệu quả của mô hình LSTM cho phát hiện tấn công APT bằng cách tính thời gian xử lý và phát hiện tấn công dựa trên quá trình chia bộ dữ liệu thực nghiệm.

Ngoài ra, Hofer-Schmitz và các cộng sự [10] đã trình bày phương pháp tối ưu cho phát hiện IP APT dựa trên bộ dữ liệu CICIDS2017. Cụ thể, trong nghiên cứu của mình các tác giả đã sử dụng thuật toán PCA nhằm giảm chiều thuộc tính của Flow từ 76 xuống 66 thuộc tính rồi tìm cách chi bộ dữ liệu thành 3 nhóm để xử lý và phân tích trên từng nhóm dữ liệu. Trong nghiên cứu [11], Bodström và các cộng sự đã đề xuất mô hình học sâu kết nhiều mô hình và thuật toán khác nhau. Mỗi lớp có chức năng tổng hợp và phân tích thuộc tính làm đầu vào cho các lớp sau. Với sự kết hợp này theo phân tích của nhóm nghiên cứu thì sẽ tận dụng được ưu điểm của các thuật toán học sâu trong nhiệm vụ phân tích và tổng hợp thuộc tính. Cho và các cộng sự [12] đã đề xuất phương pháp phát hiện tấn công APT sử dụng học sâu kết hợp. Theo đó, trong nghiên cứu của mình, các tác giả đề xuất một mô hình học sâu dựa trên sự kết hợp giữa mô hình CNN-LSTM để giám sát và phát hiện tấn công APT dựa trên lưu lượng mạng.

III. ĐỀ XUẤT MÔ HÌNH HỌC SÂU KẾT HỢP CNN-BILSTM

A. Giới thiệu về CNN

CNN là một mạng học sâu phổ biến hiện nay. CNN bao gồm tập hợp các lớp cơ bản như: lớp tích chập + lớp phi tuyến, lớp kết nối đầy đủ. Các lớp này liên kết với nhau theo một thứ tự nhất định phụ thuộc vào kiến trúc thiết kế và mục đích sử dụng của mô hình. Cấu trúc chi tiết của CNN cũng như các thuật ngữ (bước tiến, phân đệm, tổng hợp) đã được trình bày chi tiết trong bài báo [13, 14]. Trong bài báo này, chúng tôi đề xuất ứng dụng mô hình CNN trích xuất các đặc trưng của nhiều flow và kết hợp chúng lại với nhau thay vì chỉ sử dụng đối tượng là các flow riêng lẻ để nghiên cứu. Trong đó hàm kích hoạt chúng tôi sử dụng là ReLU (1).

$$f(x) = \max(0, x) \quad (1)$$

B. Giới thiệu về BiLSTM

Mặc dù mạng LSTM đã khắc phục được một số nhược điểm về khả năng ghi nhớ của các mạng nơ-ron hồi quy truyền thống. Nhưng nó vẫn có một số hạn chế là chỉ có khả năng ghi nhớ và học một chiều. Điều này dẫn đến mạng LSTM có thể làm mất các thông tin quan trọng trong một

số trường hợp. Để khắc phục tình trạng này thì trong nghiên cứu [15] đã đề xuất mạng học sâu BiLSTM. Trong nghiên cứu đã giới thiệu mô hình BiLSTM gồm 2 phần forward LSTM và backward LSTM. Điều này cho phép mô hình không chỉ kế thừa khả năng ghi nhớ xa của LSTM mà còn có khả năng ghi nhớ 2 chiều thông tin. Hai lớp LSTM tạo ra 2 trạng thái ẩn tương ứng, h_i^f từ forward LSTM và h_i^b từ backward LSTM. Trong đó, h_i^f tích hợp các thông tin trước và h_i^b tích hợp các thông tin từ sau. Xác định trạng thái cuối cùng h_i bằng cách ghép 2 trạng thái bằng công thức (2) dưới đây.

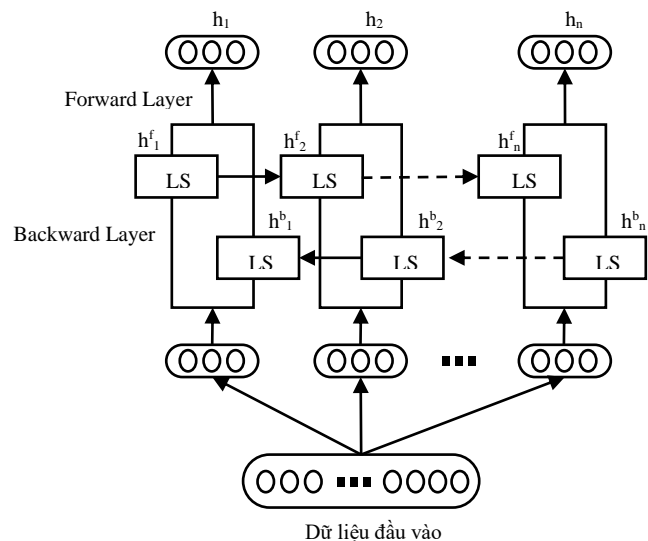
$$h_i = h_i^f || h_i^b \quad (2)$$

Trong đó :

h_i là trạng thái tại state i (chứa thông tin từ cả 2 hướng)

$||$: là phép nối

Hình 1 dưới đây thể hiện kiến trúc của mạng BiLSTM gồm 2 thành phần forward LSTM và backward LSTM.



Hình 1. Mô hình mô tả 2 thành phần LSTM trong BiLSTM

Từ kiến trúc của mạng BiLSTM được thể hiện qua hình 1 thì rõ ràng mô hình này có khả năng học được thêm một chiều thông tin. Điều này sẽ cải thiện đáng kể khả năng ghi nhớ trong quá trình training từ đó làm tăng hiệu quả của quá trình phân loại. Với cách tiếp cận sử dụng mạng BiLSTM để phát hiện APT IP, chúng tôi kỳ vọng mạng BiLSTM này có thể ghi nhớ vị trí nhóm flow trong không gian nhúng để từ đó nâng cao khả năng phân loại APT IP và IP bình thường.

C. Áp dụng cnn-bilstm cho phát hiện APT IP

Quy trình phát hiện IP tấn công APT bao gồm các bước như sau:

- **Bước 1:** xây dựng khung (frame) dựa trên mạng flow: Theo đó, thay vì nghiên cứu đối tượng là các flow riêng lẻ, chúng tôi đề xuất sử dụng các đối tượng nghiên cứu là frame. Trong đó mỗi frame được tạo ra theo quá trình xử lý flow gồm các giai đoạn:

Giai đoạn chia kích thước: trước tiên các flow của cùng một IP sẽ được nhóm vào một nhóm, sau đó chúng được

chia thành các nhóm nhỏ có kích thước như nhau, trong bài báo này qua quá trình thực nghiệm chúng tôi nhận thấy mỗi frame có kích thước là 50 flow cho kết quả tốt nhất.

Giai đoạn đệm: Nếu các frame bị lẻ do số lượng flow không chia hết cho 50 sẽ được sử dụng kỹ thuật zeros post-padding để xử lý.

Bước 2: Trích xuất đặc trưng IP dựa trên frame: các frame đã được xây dựng ở bước 1 sẽ được sử dụng làm đầu vào cho mạng CNN gồm các lớp tích chập, lớp tổng hợp cùng hàm kích hoạt là ReLU để trích xuất các đặc trưng của các flow lân cận nhau. Các đặc trưng sau khi được CNN trích xuất sẽ được làm phẳng thành các véc tơ và đi qua mạng BiLSTM tiếp tục học được các đặc trưng mới. Kết quả của BiLSTM là một véc tơ bao hàm các đặc tính nổi bật của các frame sau khi đã được trích xuất và chọn lọc bởi CNN- BiLSTM.

Bước 3: Phân loại frame: Cuối cùng, một hàm phân loại (Softmax Classifier) sẽ được sử dụng nhưng đối tượng được thay đổi từ flow đơn lẻ thành các frame để quyết định xem frame đó rơi vào lớp nào trong hai lớp APT và bình thường. Kết quả phân loại các frame này sẽ được sử dụng để phân loại IP.

IV. THỰC NGHIỆM ĐÁNH GIÁ

A. Dữ liệu thực nghiệm

Bộ dữ liệu thực nghiệm trong bài báo bao gồm:

- Dữ liệu thực nghiệm được thu thập và phân tích từ 29 file network traffic trong bộ dữ liệu Malware Capture CTU-13 của 6 loại mã độc từ các cuộc tấn công APT gồm: Andromeda, Colbalt, Cridex, Dridex, Emotet và Gh0stRAT [16].

Bộ dữ liệu network traffic sạch trong bài báo được trích xuất từ máy chủ chính phủ điện tử của tỉnh Quảng Nam [17] theo đề tài nghiên cứu khoa học KC.01.05/16-20 của bộ khoa học và công nghệ Việt Nam. Bộ dữ liệu này được thu thập trong 1 ngày 27/7/2019.

Bảng 1 dưới đây mô tả chi tiết bộ dữ liệu thực nghiệm trong bài báo

Bảng 1. Thành phần bộ dữ liệu thực nghiệm

Kiểu dữ liệu	Tổng	APT	Bình thường
Flows	3.068.915	871.914	2.197.001
IP	1690	118	1572

B. Kịch bản thực nghiệm

Bài báo sẽ bao gồm 2 kịch bản thực nghiệm như sau:

Kịch bản 1: Thực nghiệm phát hiện APT IP trên một số mạng riêng lẻ và hướng tiếp cận khác.

Kịch bản 2: Thực nghiệm đánh giá sự hiệu quả của mô hình học sâu kết hợp CNN-BiLSTM. Trong quá trình thực nghiệm chúng tôi sẽ tiến hành thay đổi tham số của từng mạng để tìm ra mô hình tối ưu nhất.

C. Kết quả thực nghiệm kịch bản 1

a) Sử dụng LSTM

Bảng 2 dưới đây thể hiện kết quả thực nghiệm khi chúng tôi áp dụng mô hình LSTM [18] cho phát hiện tấn công APT.

Bảng 2. Kết quả phát hiện tấn công APT sử dụng thuật toán LSTM [18]

Tham số LSTM	Đánh giá frame				Đánh giá IP			
	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1
2	0.857	0.957	0.475	0.635	0.966	0.714	0.428	0.535
3	0.874	0.941	0.557	0.620	0.973	0.937	0.428	0.588
4	0.858	0.976	0.475	0.635	0.970	0.833	0.428	0.566

Với các kết quả được báo cáo trong bảng 2 có thể nhận thấy rằng độ chính xác của nhiệm vụ phân loại flow và IP đã thay đổi liên tục trên hầu hết các độ đo và không theo quy luật khi tăng độ phức tạp của mô hình LSTM. Đối với kết quả phân loại frame, với điểm Accuracy tổng thể đạt từ 85,7% đến 87,4%, có thể nhận thấy mô hình LSTM có vẻ đã mang lại hiệu quả cho quá trình phân loại flow. Tuy nhiên, trong bối cảnh có sự mất cân bằng về bộ dữ liệu, độ đo Accuracy có thể được đẩy lên rất cao nếu mô hình dự đoán tất cả các IP đều thuộc vào lớp bình thường (do số IP bình thường chiếm tỉ lệ rất cao trên tổng số IPs). Với điểm F1, mạng LSTM cho các kết quả trên dao động từ 47% đến 56% trên kết quả phân loại flow, đây là các kết quả tương đối thấp và không đem lại nhiều ý nghĩa trong quá trình ứng dụng vào thực tế.

b) Kết quả thực nghiệm phát hiện tấn công APT sử dụng mạng CNN-LSTM

Bảng 3 dưới đây thể hiện kết quả thực nghiệm khi chúng tôi áp dụng mô hình CNN-LSTM [12] cho phát hiện tấn công APT.

Bảng 3. Kết quả thực nghiệm phát hiện tấn công APT sử dụng mô hình CNN- LSTM [12]

Tham số CNN	Tham số LSTM	Đánh giá frame				Đánh giá IP			
		Acc	Pre	Rec	F1	Acc	Pre	Rec	F1
3	2	0.86	0.99	0.49	0.66	0.975	0.94	0.45	0.61
4	3	0.871	0.985	0.515	0.67	0.976	0.974	0.514	0.66
4	4	0.873	0.995	0.517	0.681	0.978	0.95	0.54	0.69

Nhìn chung từ bảng 3 thấy được kết quả thực nghiệm giữa các mô hình với số lượng tham số khác nhau có sự khác biệt rõ rệt. Bên cạnh đó, các số liệu trong bảng 3 cũng cho thấy rõ nếu kết quả phân loại frame tốt sẽ cho kết quả phân loại IP tốt. Tuy nhiên, không phải cứ nhiều mạng CNN kết hợp với nhiều mạng LSTM thì sẽ cho kết quả tốt nhất. Nguyên nhân của vấn đề này có thể giải thích do cấu trúc các mạng có sự khác nhau và các tham số được khởi tạo một cách ngẫu nhiên dẫn đến cách học của mỗi mô hình là khác nhau, và các pattern được các mô hình học được cũng khác nhau dẫn đến sự sai khác trong kết quả phân loại. Trong bảng 3 thì mô hình [4 CNN – 2 LSTM] cho kết quả tốt nhất đối với cả quá trình phân loại frame và IP APT với điểm F1-score tăng đáng kể (khoảng 12% so với mô hình LSTM). Từ kết quả thực nghiệm này, có thể thấy rằng các mô hình học sâu kết hợp sẽ là hướng tiếp cận phù hợp cho các bài toán phát hiện tấn công dựa trên bất thường không có sự mất cân bằng quá lớn trong dữ liệu.

D. Kết quả thực nghiệm kịch bản 2

Bảng 4 dưới đây thể hiện kết quả thực nghiệm khi tiến hành thực nghiệm mô hình CNN-BiLSTM cho nhiệm vụ phát hiện tấn công APT.

Bảng IV. Kết quả thực nghiệm mô hình CNN-BiLSTM

Tham số CNN	Tham số BiLSTM	Đánh giá frame				Đánh giá IP			
		Acc	Pre	Rec	F1	Acc	Pre	Rec	F1
3	3	0.86	0.99	0.49	0.65	0.97	0.94	0.43	0.59
4	4	0.87	0.99	0.51	0.67	0.97	0.95	0.51	0.61
4	2	0.89	0.99	0.57	0.73	0.98	0.95	0.54	0.69

Dựa trên kết quả thực nghiệm của bảng 2 thì thấy được mô hình CNN-BiLSTM cho kết quả phân loại IP tốt nhất tại [4CNN-2 BiLSTM]. So sánh kết quả thực nghiệm của bảng 2, 3, 4 thì thấy được mô hình CNN-BiLSTM mang lại hiệu quả cao hơn so với mạng LSTM và mô hình CNN-LSTM. Nguyên nhân của vấn đề này là do mặc dù mạng LSTM có khả năng nhớ, nhưng trong quá trình tổng hợp thuộc tính thì mạng LSTM coi các mạng này có vai trò ngang nhau điều này dẫn đến nhiều thuộc tính quan trọng đã bị những thuộc tính ít quan trọng làm nhiễu và bị lu mờ. Trong khi đó mạng BiLSTM có khả năng ghi nhớ 2 chiều nên đã học được những đặc trưng từ xa tốt hơn.

V. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Trong bài báo này, với mục tiêu nâng cao hiệu quả của hệ thống phát hiện tấn công APT chúng tôi đã đề xuất một mô hình học sâu kết hợp CNN-BiLSTM. Đây là một mô hình học sâu kết hợp mới, chưa có nghiên cứu nào đề xuất. Dựa trên mô hình này, các thuộc tính và đặc trưng của IP trong lưu lượng mạng đã được làm nổi bật từ đó hỗ trợ phân loại chính xác các APT IP. Các kết quả thực nghiệm trong bài báo cho thấy mô hình này đã mang lại hiệu quả tốt hơn so với các mạng học sâu riêng lẻ LSTM hoặc học sâu kết hợp CNN-LSTM. Kết quả thực nghiệm này đã chứng minh được tính đúng đắn và hiệu quả của mô hình đề xuất. Trong tương lai, để tiếp tục cải thiện mô hình này chúng tôi cho rằng cần cải thiện 2 vấn đề: thứ nhất là khả năng tổng hợp thuộc tính của IP trong lưu lượng mạng. Thứ 2, cần phải có các phương pháp cải tiến để nâng cao quá trình phân loại.

LỜI CẢM ƠN

Bài báo này được thực hiện bởi sự tài trợ của quỹ phát triển nghiên cứu khoa học tại Học Viện Công Nghệ Bưu Chính Viễn Thông theo quyết định số 202/QĐ-HV năm 2023.

TÀI LIỆU THAM KHẢO

[1] Adel Alshamrani; Ankur Chowdhary; Sowmya Myneni; Dijiang Huang (2019) A Survey on Advanced Persistent Threats: Techniques, Solutions, Challenges, and Research Opportunities. IEEE Comm Surveys & Tutorials 21(2):1851-1877.
 [2] Do Xuan, Cho and Duong, Duc. ‘Optimization of APT Attack Detection Based on a Model Combining ATTENTION and Deep Learning’. Journal of Intelligent & Fuzzy Systems, vol. 42, no. 4, pp. 4135-4151, 2022. 10.3233/JIFS-212570.
 [3] Cho Do Xuan, Hoang Thanh, Nguyen Tung Lam. International Journal of Electrical and Computer Engineering (IJECE). Vol. 11, No. 3, 2021, pp. 2360-2370. DOI: 10.11591/ijece.v11i3.pp2360-2370.

[4] Xuan, Cho Do, Huong, D.T., and Nguyen, Toan. ‘A Novel Intelligent Cognitive Computing-based APT Malware Detection for Endpoint Systems’. Journal of Intelligent & Fuzzy Systems, vol. 43, no. 3, pp. 3527-3547, 2022. 10.3233/JIFS-220233.
 [5] Wen-Lin Chu, Chih-Jer Lin, Ke-Neng Chang (2019) Detection and Classification of Advanced Persistent Threats and Attacks Using the Support Vector Machine. Applied Sciences 9 (21). <https://doi.org/10.3390/app9214579>
 [6] Hope Nkiruka Eke, Andrei Petrovski, Hatem Ahriz (2019). The use of machine learning algorithms for detecting advanced persistent threats. In: proceedings of the 12th International on security of information and networks conference 2019 (SINCONF 2019), Sochi, pp.1-8.
 [7] Hassannataj Joloudari, M. Haderbadi, A. Mashmool, M. Ghasemigol, S. S. Band and A. Mosavi, "Early Detection of the Advanced Persistent Threat Attack Using Performance Analysis of Deep Learning," in IEEE Access, vol. 8, pp. 186125-186137, 2020, doi: 10.1109/ACCESS.2020.3029202
 [8] Cosimo Ieracitano, Ahsan Adeel, Francesco Carlo Morabito, Amir Hussain. A novel statistical analysis and autoencoder driven intelligent intrusion detection approach. Neurocomputing, Volume 387, 2020, Pages 51-62
 [9] Sai Charan P.V., Gireesh Kumar T., Mohan Anand P. (2019) Advance Persistent Threat Detection Using Long Short Term Memory (LSTM) Neural Networks. In: Somani A., Ramakrishna S., Chaudhary A., Choudhary C., Agarwal B. (eds) Emerging Technologies in Computer Engineering: Microservices in Big Data Analytics. ICETCE 2019. Communications in Computer and Information Science, vol 985. Springer, Singapore. https://doi.org/10.1007/978-981-13-8300-7_5
 [10] Aaron Tuor, Samuel Kaplan, Brian Hutchinson, Nicole Nichols, Sean Robinson (2017) Deep Learning for Unsupervised Insider Threat Detection in Structured Cybersecurity Data Streams. In: proceedings of the AAAI-17 Workshop on Artificial Intelligence for Cyber Security WS-17-04. San Francisco, pp. 9-21
 [11] Tero Bodström; Timo Hämäläinen (2019) A Novel Deep Learning Stack for APT Detection. Applied Sciences 9 (6). <https://doi.org/10.3390/app9061055>
 [12] Cho Do Xuan, Hoa Dinh Nguyen, Hoang Mai Dao (2020). APT attack detection based on flow network analysis techniques using deep learning. Journal of Intelligent & Fuzzy Systems 39 (3): 4785-4801
 [13] Zewen Li, Wenjie Yang, Shouheng Peng, Fan Liu (2020). A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. arXiv, arXiv:2004.02806.
 [14] Keiron O’Shea, Ryan Nash (2015) An Introduction to Convolutional Neural Networks. arXiv, arXiv:1511.08458
 [15] Savelie Cornegruta, Robert Bakewell, Samuel Withey, Giovanni Montana. Modelling Radiological Language with Bidirectional Long Short-Term Memory Networks. arXiv, 2017, arXiv:1609.08409
 [16] Malware Capture Facility Project. Available online: <https://www.stratosphereips.org/datasets-malware>. (accessed on 8 June 2023).
 [17] Quang Nam Portal. Available online: <http://english.quangnam.gov.vn/default.aspx> (accessed on 8 June 2023)
 [18] Kaibo Duan; Sathiya Keerthi, S.; Wei Chu; Shirish Krishnaj Shevade; Aun Neow Poo. Multi-category Classification by Soft-Max Combination of Binary Classifiers. In proceedings of the 4th International Workshop, MCS 2003 Guildford, UK, 11–13 June 2003; pp 125–134.

BASED ON HYBRID DEEP LEARNING MODELS TO ENHANCE THE EFFECTIVENESS OF APT ATTACK DETECTION

Abstract: APT (Advanced Persistent Threat) attacks are sophisticated, persistent, and purposeful cyberattacks.

These attacks cause significant damage to organizations and government agencies. Therefore, the task of early detection and warning of this type of attack is crucial today. To enhance the effectiveness of APT attack detection, this research paper proposes a new approach based on hybrid deep learning models. The APT-IP attack detection process outlined in this research involves preprocessing the entire network traffic data, analyzing it using a combination of deep learning network and Convolutional Neural Network (CNN), and then further analyzing and evaluating the data using a Bidirectional Long Short-Term Memory (BiLSTM) network. Finally, the data processed by the BiLSTM network is classified to identify APT-IP attacks.

Keywords: APT; APT detection; deep learning model; abnormal behaviors.



Nguyễn Trung Thành, Nhận học vị Tiến sỹ năm 2017. Hiện công tác tại Học viện Công nghệ Bưu chính Viễn thông. Lĩnh vực nghiên cứu: Blockchain, Hệ thống lưu trữ hiệu năng cao (NoSQL/NewSQL), An toàn thông tin, điện toán đám mây, Công nghệ & ứng dụng Web3, NFT

Email: thanhnt3@ptit.edu.vn



Nguyễn Hoa Cường, Nhận học vị Thạc sỹ năm 2013. Hiện công tác tại Học viện Công nghệ Bưu chính Viễn thông. Lĩnh vực nghiên cứu: An toàn và bảo mật thông tin, đạo đức và chính sách an toàn mạng,

Email: cuongnh@ptit.edu.vn