NGHIÊN CỨU VÀ ỨNG DỤNG THUẬT TOÁN HỌC SÂU TRONG PHÁT HIỆN TÊN MIỀN BẤT THƯỜNG

Huỳnh Đệ Thủ, Trần Công Hùng,
Lê Võ Hoàng Anh, Huỳnh Thiên Lộc, Nguyễn Kim Thoa
Trường Đại học Tư thục Quốc tế Sài Gòn

Tóm Tắt — Gần đây, nhiều botnet sử dụng thuật toán tạo tên miền bất thường tự động (DGA) để sinh và đăng ký nhiều tên miền bất thường ngẫu nhiên khác nhau cho máy chủ lệnh và điều khiển của chúng nhằm chống lại việc bị kiểm soát và đưa vào danh sách đen. Do đó, việc phát hiện tên miền bất thường đang là mối quan tâm nghiên cứu của nhiều nhà nghiên cứu trên toàn thế giới vì tính lan rộng, độ tinh vi cao và hậu quả nghiêm trọng đối với nhiều tổ chức và người dùng. Bài báo này tập trung nghiên cứu xây dựng mô hình phát hiện tên miền bất thường dựa trên các kiến trúc Recurrent Neural Network (RNN), Long ShortTerm Memory (LSTM) và Convolutional Neural Network (CNN) của thuật toán học sâu và thử nghiệm đánh giá các mô hình này trên bộ dữ liệu gồm 16.7 triệu tên miền trong đó có 9 triệu tên miền bất thường và 7.7 tên miền lành tính. Kết quả thử nghiệm chỉ ra rằng mô hình học sâu sử dụng kiến trúc mạng học sâu LSTM cho tỷ lệ chính xác cao nhất. Ngoài ra, chúng tôi cũng đã đề xuất và triển khai thử nghiệm thành công giải pháp tích hợp mô hình phát hiện tên miền bất thường vào hệ thống giám sát cảnh báo sớm (SOC).

Từ khóa — Thuật toán tạo tên miền tự động, Học sâu, Botnets, An ninh mạng, Mạng Neural

I. GIỚI THIỆU

Trong cuộc sống hiện đại ngày nay, với việc mạng Internet ngày càng phát triển không ngừng, công nghệ thông tin được ứng dụng vào mọi mặt của đời sống, kinh tế, chính trị, xã hội đã giúp cho cá nhân, tổ chức, doanh nghiệp và các cơ quan hành chính nhà nước trên thế giới nói chung và Việt Nam nói riêng dễ dàng trao đổi thông tin và thực hiện các giao dịch được thuận lợi nhanh chóng. Cùng với sự phát triển của công nghệ thông tin và viễn thông đã mở ra không gian mới cho sự tấn công và đe dọa an ninh mạng. Trong số những mối đe dọa này, là kẻ tấn công sử dụng thuật toán tạo tên miền (Domain Generation Algorithm- DGA) [6, 8] để tạo ra một số lượng lớn các tên miền bất thường ngẫu nhiên để kết nối với các máy chủ điều khiển và lệnh độc hại, chủ yếu được sử dụng trong các hình thức chính như sau:

- Phân phối mã độc: Các loại mã độc phổ biến được phân phối bởi tên miền bất thường bao gồm ransomware, phần mềm gián điệp, trojan và botnet.
- Tấn công từ chối dịch vụ (DDoS): Tên tên miền bất thường cũng được sử dụng để thực hiện các cuộc tấn công DDoS quy mô lớn, nhằm làm sập các trang web hoặc dịch vụ quan trọng.
- Lừa đảo: Tên tên miền bất thường cũng được sử dụng để tạo ra các trang web giả mạo, nhằm lừa người dùng cung cấp thông tin cá nhân nhạy cảm.

Việc tấn công này đã gây ra những thiệt hại không nhỏ về hệ thống mạng và sự mất mát dữ liệu của người dùng, dẫn đến thiệt hại về kinh tế, xã hội của các cá nhân, tổ chức, doanh nghiệp và cơ quan hành chính nhà nước.

Do đó, phát hiện tên miền bất thường đang là mối quan tâm nghiên cứu của nhiều nhà nghiên cứu trên toàn thế giới vì mức độ lan rộng, độ tinh vi cao và hậu quả nghiêm trọng đối với nhiều tổ chức và người dùng. Trong nghiên cứu này, chúng tôi tập trung đánh giá hiệu quả của việc xây dựng mô hình phát hiện tên miền bất thường dựa trên kỹ thuật học sâu và học máy để xác định ra mô hình kiến trúc LSTM có hiệu quả cao trong bài toán phát hiện tên miền bất thường.

II. CÁC CÔNG TRÌNH LIÊN QUAN

Yadav và cộng sự [1] đề xuất phương pháp phân biệt các tên miền sinh tự động bằng thuật toán thường được sử dụng trong các botnet với tên miền hợp lệ dựa trên phân tích sự phân bố các nhóm ký tự (1-gram, hoặc 2-gram liền kề) trong tên miền. Các tác giả sử dụng độ đo phân kỳ Kullback-Leibler (K-L) để tính toán khoảng cách giữa các tên miền hợp lệ và các tên miền độc hại được sinh tự động. Hạn chế của phương pháp này là nó thường không phát hiện ra các họ DGA khác nhau.

Đức và cộng sự [2] đã sử dụng mạng bộ nhớ ngắn hạn dài Long Short-Term Memory Network - LSTM để giải quyết cả hai bài toán DGA Botnet. Nhóm nghiên cứu đề xuất một thuật toán mới với tên gọi là LSTM.MI, kết hợp cả hai mô hình phân loại, kế thừa ưu điểm của LSTM truyền thống và các cải tiến để tăng cường độ chính xác với nhiễu. Về dữ liệu thử nghiệm, các tên miền được nhóm tác giả thu thập từ thực tế, với 100.000 tên miền lành tính phổ biến nhất từ Alexa và 37 họ DGA Botnet được tổng hợp. Kết quả thử nghiệm cho thấy thuật toán đề xuất giúp nâng cao ít nhất 7% độ chính xác so với mô hình LSTM truyền thống. Nó cũng đạt độ chính xác cao trong bài toán phân

Tác giả liên hệ: Trần Công Hùng,
Email: tranconghung@siu.edu.vn
Đến tòa soạn: 23/4/2024, chỉnh sửa: 21/5/2024, chấp nhận đăng: 06/6/2024.

lớp nhị phân với F1-score đạt 98,49%, đồng thời có khả năng nhận ra 05 họ DGA Botnet bổ sung. Hạn chế của nghiên cứu là bộ dữ liệu chưa thực sự đầy đủ và một số họ DGA Botnet gần như không thể phát hiện được.

Curtin và cộng sự đã sử dụng mạng RNN để phát hiện và phân loại DGA Botnet [3]. Họ nhận ra rằng các tên miền được xây dựng dựa trên một không gian từ vừng, 32 có các nét đặc trưng khá giống với các tên miền lành tính, từ đó giúp tăng khả năng ẩn mình của các tên miền được sinh ra. Nhóm nghiên cứu đã đề xuất một khái niệm mới, gọi là thang điểm Smashword. Đây là thang đo lường mức độ giống nhau giữa tên miền của Botnet và các tên miền lành tính. Nhóm nghiên cứu áp dụng mô hình mạng Recurrent Neural Network và Side Information để áp dụng tiêu chuẩn đo lường trên. Bộ dữ liệu thử nghiệm với 1.000.000 tên miền Alexa, kết hợp với 41 họ DGA Botnet được nhóm nghiên cứu tổng hợp, với tổng cộng 2.300.000 tên miền cho cả hai nhãn. Thực nghiệm cho thấy mô hình mới có tiềm năng ứng dụng cao cải tiến được độ chính xác hơn so với các mô hình trước đó. Một số họ DGA Botnet phức tạp như matsnu, suppobiox hay rovnix cũng được phát hiện bởi giải pháp trên.

Qiao và cộng sự [4] đề xuất một kiến trúc học sâu phát triển dựa trên mạng LSTM và Attention. Kiến trúc mới có lõi gồm một lớp LSTM với một lớp Attention được bố trí tuần tự, kích thước đầu vào là 54 × 128. Nhóm tác giả đánh giá trên một bộ dữ liệu gồm 16 nhãn bao gồm cả nhãn lành tính. Kết quả thực nghiệm cho thấy mô hình đề xuất có F1-score đạt 94,58% cho bài toán phân loại. Ưu điểm của phương pháp là đạt độ chính xác cao và loại bỏ được quá trình trích chọn các đặc trưng. Tuy nhiên, phương pháp đề xuất cũng chỉ phát hiện tốt các tên miền character- based DGA. Ngoài ra, tỷ lệ cảnh báo sai tổng cũng còn tương đối cao, khoảng 5% tính theo độ đo F1.

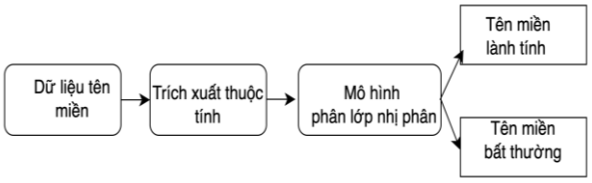
Vinayakumar và nhóm cộng sự nghiên cứu bài toán phát hiện tên miền độc hại được sinh bởi Botnet hay thư điện tử, URL độc hại [5]. Một số kỹ thuật biểu diễn đặc trưng dựa trên n-gram đã được sử dụng để mô hình hóa bài toán. Bộ dữ liệu để đánh giá bao gồm các tên miền lành tính và độc hại được thu thập từ OpenDNS, Alexa và OSINT Feeds. Kết quả so sánh thấy mô hình CNN-LSTM là hiệu quả nhất với F1- score đạt 96,3% cho bài toán phát hiện.

Liu và cộng sự [6] lập luận rằng biểu diễn đặc trưng bằng các giá trị vô hướng có thể dẫn đến mất mát thông tin. Nhóm tác giả đề xuất một Capsule Network tuần tự (mạng học sâu được cải tiến từ CNN) dựa trên thuật toán k-means, gọi là LSTMCapsNet. Mô hình sử dụng một đơn vị LSTM hai chiều để trích xuất các thuộc tính cơ bản và sử dụng thuật toán k-mean để phân cụm thành hai nhãn độc hại và lành tính. Đánh giá được thực hiện trên hai bộ dữ liệu, bao gồm bộ dữ liệu tên miền DGA từ mạng thực tế và tên miền DGA thu được thông qua thuật toán tạo tên miền tự động. Kết quả thực nghiệm cho thấy mô hình đề xuất đạt độ chính xác lần lượt là 99,17% và 97,75% trên hai bộ dữ liệu. Mô hình này không chỉ cải thiện khả năng nhận dạng 33 tên miền DGA và nhận dạng họ tên miền DGA trên lý thuyết mà còn thể hiện hiệu quả trong bộ dữ liệu thực tế.

III. ĐỀ XUẤT MÔ HÌNH HỖ TRỢ PHÁT HIỆN TÊN MIỀN BẤT THƯỜNG

A. Mô hình học máy

Các giai đoạn trong quá trình đánh giá thuật toán học máy đối với bài toán phân lớp nhị phân được thể hiện

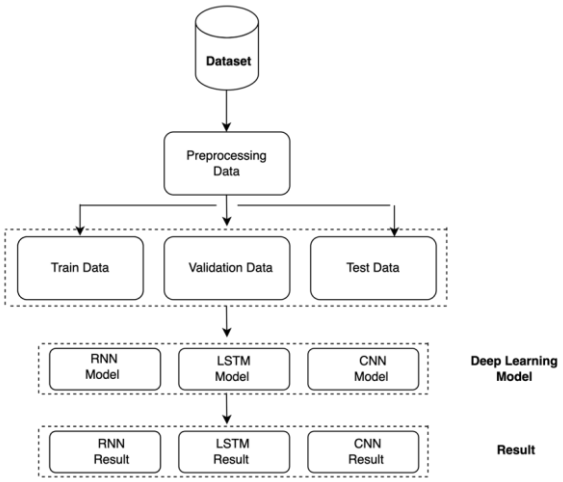


Hình 1. Sơ đồ mô hình huấn luyện, đánh giá

Theo đó, với dữ liệu đầu vào là các tên miền, bao gồm cả lành tính và độc hại đã được gán nhãn, sử dụng n-gram để tách các tên miền và kỹ thuật TF-IDF để biểu diễn các đặc trưng. Tiếp theo, các thuật toán học máy được sử dụng để huấn luyện, bao gồm: Support Vector Machines (SVM), Logistic Regression (LR), Naive Bayes (NB), Decision Trees (DT), Random Forests (RF), k-Nearest Neighbour (kNN), Adaptive Boosting (AB). Mô hình sau khi huấn luyện được dùng để giải quyết bài toán phân lớp nhị phân, với hai nhãn là độc hại và lành tính.

B. Mô hình học sâu

Trong nghiên cứu này sẽ tập trung xây dựng mô hình phát hiện bất thường mạng dựa trên các kiến trúc RNN, LSTM và CNN của mạng học sâu [9].



Hình 2: Mô hình tổng quan hệ thống

Bộ dữ liệu sẽ được tiền xử lý để có thể đáp ứng yêu cầu huấn luyện, cũng như cải thiện độ chính xác của mô hình về sau. Đây là bước hết sức quan trọng, vì dữ liệu càng có độ chi tiết cao thì kết quả mô hình đem lại càng tốt. Sau đó dữ liệu sẽ được tách thành 3 tập Train, Test và Validation Data người dùng. Train Data người dùng là tập dữ liệu sử dụng cho mục đích huấn luyện, chiếm tỉ lệ 80% so với tổng cả bộ dữ liệu. Test Data người dùng là tập kiểm thử kết quả của mô hình, Validation Data người dùng là tập sử dụng vào mục đích giám sát quá trình huấn luyện. Hai tập Test và Validation chiếm tỉ lệ 10%. Sau khi dữ liệu đã tách thành 3 tập trên, chúng được sử dụng để huấn luyện và đánh giá các mô hình RNN, LSTM và CNN.

C. Cấu trúc mô hình học sâu sử dụng

Các mô hình đều được xây dựng dựa trên kiến trúc của Neural Network Sau khi bộ dữ liệu đã được tiền xử lý, chúng được đưa qua lớp Reshape, sau đó đến lớp RNN/LSTM/CNN và đi đến các lớp Dense hay tên gọi khác là Fully- connected layer. Ngoài ra ở sau mỗi lớp Dense là một lớp Dropout Layer, thực hiện loại bỏ ngẫu nhiên tỉ lệ units ở lớp trước đó. Cuối cùng kết quả của các mô hình là dự đoán cho các tên miền trong tập kiểm tra. Sơ đồ các mô hình như trong hình 3:

Hình 3: Các mô hình RNN, LSTM và CNN

Người dùng cần đưa dữ liệu gốc qua lớp Reshape Layer (vì lớp RNN Layer/LSTM Layer/CNN Layer yêu cầu dữ liệu phải có dạng 3 chiều [batch, timesteps, feature]. Trong đó: batch: batch_size (số mẫu dữ liệu đưa vào tại một thời điểm); timesteps: số bước nhớ của mô hình; features: số đặc trưng của dữ liệu. Trước khi qua Reshape Layer, người dùng quy định dữ liệu sẽ đi vào theo từng mẫu một. Mỗi mẫu dữ liệu có 70 đặc trưng, vì vậy input_shape = (70). Vì vậy ở Reshape Layer người dùng cần dữ liệu vẫn đảm bảo có shape tương đương với input_shape (input_shape = timestep*features). Để dễ hình dung, người dùng đặt features = 70, timestep = 1. Sau khi qua lớp Reshape Layer, output sẽ được truyền vào lớp RNN/LSTM/GRU Layer, lớp này có chức năng phân tích các đặc trưng của dữ liệu. Output của lớp này có dạng 2 chiều (batch_size, units) với units là số unit có trong lớp. Ở các mô hình trên, output có dạng (None, 128). Hai lớp Dense Layer (tiếp theo (512 units và 64 units) tiếp tục phân tích các đặc trưng dữ liệu. Ngoài ra sau mỗi Dense Layer có sử dụng một lớp Dropout Layer. Lớp Dropout Layer thực hiện loại bỏ ngẫu nhiên k % số unit ở lớp Dense Layer phía trước, mục đích của việc loại bỏ này nhằm tránh hiện tượng overfitting xảy ra trong quá trình huấn luyện. Lớp Dense Layer cuối cùng có 15 units, có nhiệm vụ biến đổi đầu vào dạng logits thành softmax (dạng xác suất) Kết quả đầu ra trả về kết quả dự đoán cho các tên miền trong tập kiểm tra. Các độ đo và ma trận nhầm lẫn (confusion matrix) được tính toán để đánh giá hiệu suất của mô hình.

IV. ĐÁNH GIÁ MÔ HÌNH

A. Bộ dữ liệu

Trong bài nghiên cứu này sẽ sử dụng bộ dữ liệu có 16 triệu tên miền được AAYUSH V. SHAH xây dựng, cập nhật năm 2020 và lưu trữ công khai trên Kaggle [7].

Bảng 1: Chi tiết bộ dữ liệu

	Training	Validation	Testing
Benign	4.974.021	1.243.744	1.554.877
DGA	5.761.032	1.440.020	1.799.828
Total	10.735.053	2.683.764	3.354.705

B. Độ đo đánh giá

Bài toán phát hiện tên miền bất thường là bài toán phân lớp nhị phân, với nhãn 0 là tên miền lành tính, nhãn 1 là tên miền tên miền bất thường, được định nghĩa:

- TN: Số lượng mẫu tên miền là lành tính được phân loại đúng là lành tính.
- TP: Số lượng mẫu tên miền độc hại được phân loại đúng là độc hại.
- FP: Số lượng mẫu tên miền lành tính được phân loại sai thành độc hại.
- FN: Số lượng mẫu tên miền độc hại được phân loại sai thành lành tính.

Đánh giá kết quả thực nghiệm thông qua các tham số gồm Accuracy, Precision, Recall và F1-score, lần lượt được tính bằng các công thức (1), (2), (3) và (4) dưới đây:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (1)$$

Accuracy giúp đánh giá độ chính xác chung của mô hình, bao gồm việc phát hiện đúng tên miền là lành tính hoặc tên miền là độc hại.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (3)$$

Precision và Recall đánh giá riêng về khả năng phát hiện đúng một tên miền lành tính hay một tên miền độc hại trên tổng số lượng mẫu của tên miền đó. Nhìn chung, việc phát hiện nhầm một tên miền lành tính thành độc hại sẽ ít gây nguy hiểm hơn là việc chấp nhận nhầm một tên miền độc hại là lành tính.

$$\text{F1} - \text{Score} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision+Recall}} \quad (4)$$

F1-Score là đại lượng dùng chung để đánh giá một cách tổng thể giữa 1 Precison và Recall trong mô hình.

C. Kết quả

Bảng 2: Bảng tổng hợp kết quả các giá trị Precision, Recall, F1-Score

Thuật toán	Mô hình	Precision	Recall	F1-Score
Học máy	LR	0.74	0.58	0.46
	NB	0.75	0.57	0.46
	DT	0.76	0.58	0.46
	SVM	0.82	0.80	0.81
	RF	0.96	0.68	0.80

	kNN	0.93	0.05	0.11
	AB	0.64	0.89	0.75
Học sâu	CNN	0.9	0.89	0.89
	RNN	0.91	0.79	0.88
	LSTM	0.96	0.95	0.93

D. Đánh giá

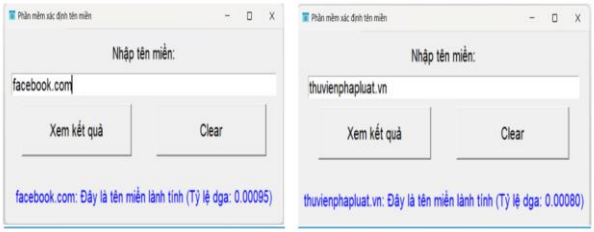
Từ số liệu trên ta thấy: mô hình LSTM đều có các kết quả Precision/Recall/F1-Score cao hơn so với CNN và RNN và mô hình học máy. Ta có thể khẳng định các thuật toán học sâu LSTM hiệu quả trong bài toán phát hiện tên miền bất thường nhờ vào sự phù hợp với bộ dữ liệu.

V. TRIỂN KHAI THỬ NGHIỆM VÀ ĐỀ XUẤT TÍCH HỢP MÔ HÌNH PHÁT HIỆN TÊN MIỀN BẤT THƯỜNG VÀO HỆ THỐNG GIÁM SÁT CẢNH BÁO SỚM (SOC)

A. Ứng dụng thử nghiệm

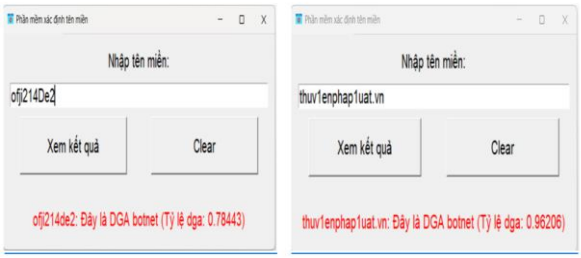
Mô hình thuật toán sau khi được huấn luyện sẽ được lưu lại để ứng dụng vào sản phẩm dự đoán tên miền. Để dự đoán tên miền, nhóm đã tạo một ứng dụng cơ bản bằng windown form với phần intput để nhập tên miền cần kiểm tra, "Xem kết quả" để gửi phân tích và "Clear" để xóa tên miền hiện tại.

- Tên miền lành tính:



Hình 4: Kết quả dự đoán tên miền lành tính

- Tên miền bất thường:



Hình 5: Kết quả dự đoán tên miền bất thường

B. Đề xuất ứng dụng thực tế

- Hiện trạng công việc giám sát SOC truyền thống:

- Người giám sát SOC thường không được trang bị đầy đủ để xử lý các mối đe dọa một cách hiệu quả, họ chỉ hiểu biết một phần nhất định, không thể nắm hết tất cả các lĩnh vực.

- Quá nhiều cảnh báo gây mệt mỏi cho người giám sát và khiến các cảnh báo quan trọng bị bỏ qua.

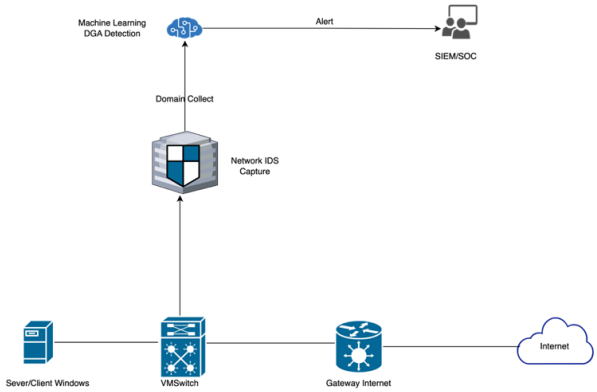
- Tốn thời gian bởi các cảnh báo được tạo cần có sự can thiệp của con người để phân tích và thực hiện hành động nên quá trình giảm thiểu mối đe dọa bị chậm trễ.

- Việc triển khai tích hợp mô hình thuật toán học sâu vào việc phát hiện tên miền bất thường trong các hệ thống giám sát cảnh báo sớm (SOC) trong thực tế tại các đơn vị đem lại các lợi ích như sau:

- Xác định các mối đe dọa bảo mật với độ tin cậy cao, đồng thời giảm thời gian và kinh nghiệm cần thiết trong SOC.

- Toàn bộ quy trình có thể được sắp xếp hợp lý, giúp xác định các sự kiện quan trọng và tự động thực hiện biện pháp khắc phục.

Tích hợp mô hình phát hiện tên miền bất thường vào các hệ thống giám sát cảnh báo sớm (SOC) trong thực tế tại các đơn vị theo mô hình tổng thể như sau:

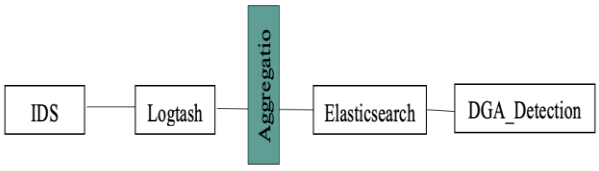


Hình 6: Tích hợp mô hình phát hiện tên miền bất thường vào hệ thống SOC

Các thành phần trong mô hình:

- Server/Client windows: một người dùng hoặc máy chủ kết nối đến một tên miền nào đó.
- Network IDS Capture: gói tin sẽ được bắt lại.
- Machine Learning DGA Detection [10-13]: trích xuất domain từ log ids và dự đoán.
- SIEM/SOC: kết quả được gửi lên SIEM/SOC cảnh báo cho đội ngũ giám sát hệ thống

Kiến trúc phần mềm:



Hình 7: Kiến trúc phần mềm của giải pháp tích hợp mô hình phát hiện tên miền bất thường vào hệ thống SOC

Các thành phần chính của kiến trúc:

- IDS: thu thập các log dữ liệu trên máy tính, máy chủ và trên mạng mà nó giám sát.
- Logtash: log sẽ được đưa đến logstash, logstash sẽ đọc những log này, thêm những thông tin như thời gian, IP, parse dữ liệu từ log (server nào, độ nghiêm trọng, nội dung log) ra, sau đó ghi xuống database là Elasticsearch. Mỗi dòng log của logstash được lưu trữ đưới dạng json.
- Aggregation query: để lấy danh sách domain trong vòng 24h từ logtash.

- Elasticsearch: cơ sở dữ liệu này được dùng để lưu trữ, tìm kiếm và query log.
- DGA_Detection & Alert: trích xuất domain từ log Elasticsearch chu kì 1 ngày, tiến hành đưa vào bộ phân loại để phát hiện domain bất thường, gửi lại kết quả phân loại lên cho đội ngũ giám sát.

VI. KẾT LUẬN

Trong bài báo này, chúng tôi đã đề xuất và thực hiện các thử nghiệm đánh giá hiệu suất của các mô hình phát hiện tên miền bất thường dựa trên kỹ thuật học sâu và học máy. Kết quả chỉ ra rằng, mô hình kiến trúc mạng LSTM có kết quả độ chính xác cao nhất trong bài toán phát hiện tên miền bất thường. Đây là một con số ấn tượng, đem đến tiềm năng cho việc xây dựng một ứng dụng hỗ trợ phát hiện sớm tên miền bất thường. Chúng tôi cũng đã đề xuất và triển khai thử nghiệm thành công giải pháp tích hợp mô hình phát hiện tên miền bất thường vào hệ thống giám sát cảnh báo sớm (SOC) thực tế tại đơn vị.

TÀI LIỆU THAM KHẢO

[1] Yadav S., Reddy A., Reddy A., and Ranjan S., "Detecting Algorithmically Generated Malicious Domain Names". In Proceedings of the 10th ACM SIGCOMM conference on Internet measurement (IMC '10). Association for Computing Machinery, New York, USA, 48–61.

[2] D. Tran, H. Mac, V. Tong, H. A. Tran, and L. G. Nguyen, "A LSTM based framework for handling multiclass imbalance in DGA botnet detection," Neurocomputing, vol. b 275, pp. 2401–2413, 2018, doi: 10.1016/j.neucom.2017.11.018.

[3] R. R. Curtin, A. B. Gardner, S. Grzonkowski, A. Kleymenov, and A. Mosquera, "Detecting DGA domains with recurrent neural networks and side information," ACM Int. Conf. Proceeding Ser., 2019, doi: 10.1145/3339252.3339258.

[4] Y. Qiao, B. Zhang, W. Zhang, A. K. Sangaiah, and H. Wu, "DGA domain name classification method based on Long Short-Term Memory with attention mechanism," Appl. Sci., vol. 9, no. 20, 2019, doi: 10.3390/app9204205.

[5] R. Vinayakumar, M. Alazab, S. Srinivasan, Q. V. Pham, S. K. Padannayil, and K. Simran, "A Visualized Botnet Detection System Based Deep Learning for the Internet of Things Networks of Smart Cities," IEEE Trans. Ind. Appl., vol. 56, no. 4, pp. 4436– 4456, 2020, doi: 10.1109/TIA.2020.2971952.

[6] X. Liu and J. Liu, "DGA botnet detection method based on capsule network and kmeans routing," Neural Comput. Appl., vol. 34, no. 11, pp. 8803–8821, 2022, doi: 10.1007/s00521-022-06904-3.

[7] https://www.kaggle.com/datasets/aayushah19/dga-or-benign-domain-names

[8] https://vi.wikipedia.org/wiki/H%E1%BB%8Dc_m%C3%A1y.

[9] https://blogs.nvidia.com/blog/supervised-unsupervised-learning/.

[10] ıttps://machinelearningcoban.com/2017/01/08/knn/#k-nearest-ıeighbor.

[11] H. S. Anderson, J. Woodbridge, and B. Filar, "DeepDGA: Adversarially-tuned domain generation and detection," AISec 2016 - Proc. 2016 ACM Work. Artif. Intell. Secur. co-located with CCS 2016, pp. 13–21, 2016, doi: 10.1145/2996758.2996767.

[12] M. Zago, "Mendeley Data," UMUDGA - University of Murcia Domain Generation Algorithm Dataset, 24/2/2020. [Online]. Available: https://data.mendeley.com/datasets/y8ph45msv8/1.

[13] H. S. Anderson, J. Woodbridge, and B. Filar, "DeepDGA: Adversarially-tuned domain generation and detection," AISec 2016 - Proc. 2016 ACM Work. Artif. Intell. Secur. co-located with CCS 2016, pp. 13–21, 2016, doi: 10.1145/2996758.2996767.

USING DEEP LEARNING ALGORITHM TO DETECTING ABNORMAL DOMAIN NAMES

Abstract: Recently, many botnets have been using Domain Generation Algorithm (DGA) to automatically generate and register various random, unusual domain names for their command and control servers in order to evade control and blacklisting. Therefore, detecting unusual domain names has become a research interest for many researchers worldwide due to the widespread nature, high sophistication, and serious consequences for many organizations and users. This paper focuses on building a model to detect unusual domain names based on Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN) architectures of deep learning algorithms and testing these models on a dataset consisting of 16.7 million domain names, of which 9 million are unusual domains and 7.7 million are benign. The experimental results indicate that the deep learning model using the LSTM network architecture achieves the highest accuracy. Additionally, we have proposed and successfully implemented a solution to integrate an anomalous domain detection model into the early warning monitoring system (SOC).

Keywords: Domain generation algorithm, Deep learning, Botnets, Network security, Neural networks

Huỳnh Đệ Thủ nhận bằng tiến sĩ về Khoa học máy tính từ Trường Đại học Khoa học và Công nghệ Hoa Trung, Trung Quốc vào năm 2018. Ông hiện đang giảng dạy và nghiên cứu tại Khoa Kỹ thuật & Khoa học máy tính, Đại học Quốc tế Sài Gòn. Lĩnh vực nghiên cứu chính: Mạng không dây, Mạng viễn thông, Internet vạn vật, An ninh mạng, An toàn thông tin và Khoa học dữ liệu.

Trần Công Hùng sinh năm 1961. Ông nhận bằng Kỹ sư điện tử viễn thông hạng Nhất của trường Đại học Bách khoa thành phố Hồ Chí Minh, Việt Nam năm 1987. Ông nhận bằng Kỹ sư tin học và máy tính của trường Đại học Bách khoa thành phố Hồ Chí Minh, Việt Nam, 1995. Ông nhận bằng thạc sĩ kỹ thuật viễn thông của Trường Đại học Bách Khoa Hà Nội, 1998. Ông nhận bằng Tiến Sỹ tại Trường Đại học Bách Khoa Hà Nội, Việt Nam, 2004. Lĩnh vực nghiên cứu chính: các tham số và phương pháp đo hiệu suất B - ISDN, QoS trong mạng tốc độ cao, MPLS, Mạng cảm biến không dây, Điện toán đám mây. Ông hiện là Phó Giáo sư Tiến sĩ, Trưởng Khoa Kỹ thuật và Khoa học máy tính, Trường Đại học Quốc tế Sài Gòn, thành phố Hồ Chí Minh, Việt Nam.

Lê Võ Hoàng Anh tốt nghiệp trường Đại học Công nghệ Thành phố Hồ Chí Minh, Việt Nam năm 2007. Cô hiện đang là học viên Cao học ngành Khoa học máy tính, Trường Đại học Quốc tế Sài Gòn, Việt Nam. Lĩnh vực nghiên cứu của cô: Mạng cảm biến không dây, Trí tuệ nhân tạo, Điện toán đám mây, Khoa học dữ liệu, Internet vạn vật, An ninh mạng và Khoa học dữ liệu.

Huỳnh Thiên Lộc là sinh viên ngành Khoa học máy tính chuyên ngành An ninh mạng tại Đại học Quốc tế Sài Gòn từ năm 2021. Các lĩnh vực mà anh quan tâm bao gồm: An ninh mạng, Bảo mật Web và an toàn thông tin.

Nguyễn Kim Thoa tốt nghiệp trường Đại học Văn Lang, Thành phố Hồ Chí Minh, Việt Nam năm 2003. Cô hiện đang là học viên Cao học ngành Khoa học máy tính, Trường Đại học Quốc tế Sài Gòn, Việt Nam. Lĩnh vực nghiên cứu của cô: Mạng cảm biến không dây, Trí tuệ nhân tạo, Điện toán đám mây, Khoa học dữ liệu, Internet vạn vật, An ninh mạng và Khoa học dữ liệu.