# MÔ HÌNH PHÁT HIỆN TẤN CÔNG DDOS SỬ DỤNG MACHINE LEARNING

*VÕ HỒ THU SANG*[*]
*NGUYỄN ĐỨC NHUẬN, PHAN HOÀNG HẢI*
*Khoa Tin học, Trường Đại học Sư phạm, Đại học Huế*
[*]*Email: vohothusang@dhsphue.edu.vn*

**Tóm tắt:** Tấn công DDoS trên Internet đã và đang gây tổn thất, ảnh hưởng lớn đến vấn đề an ninh cũng như hiệu suất mạng. Bên cạnh đề xuất, cải tiến các mô hình phân lớp lưu lượng tấn công DDoS, rút gọn và chỉ ra tập đặc trưng liên quan đến lưu lượng tấn công DDoS là một bài toán mở cần được quan tâm nghiên cứu để tăng hiệu quả dự báo, giảm độ phức tạp tính toán, giảm khả năng overfitting của mô hình. Trong bài báo này, chúng tôi đề xuất mô hình phát hiện tấn công DDoS sử dụng kết hợp ba mô hình thực hiện rút gọn tập đặc trưng từ tập đặc trưng đầu vào thay vì sử dụng các mô hình/ phương pháp riêng lẻ được sử dụng trong một số các nghiên cứu tấn công DDoS gần đây. Với những đặc trưng được lựa chọn, các mô hình học có giám sát phổ biến như SVC, Kneighbor, Naïve Bayes, Random Forest được triển khai để phát hiện tấn công DDoS, qua các chỉ số đánh giá gồm Accuracy, F1 score, AUC mô hình đề xuất có hiệu quả tốt nhất với Random Forest.

**Từ khóa:** DDoS, machine learning, SVC, Kneighbor, Naïve Bayes, Random Forest, rút gọn tập đặc trưng.

## 1. MỞ ĐẦU

Tấn công từ chối dịch vụ phân tán DDoS (là một biến thể của tấn công DoS) được xem là một trong những kiểu tấn công phổ biến trên Internet. Cùng với sự đa dạng của loại thiết bị, dịch vụ cũng như sự phát triển nhanh chóng của các mô hình tấn công mạng trong những thập kỉ qua, đã đặt ra yêu cầu cấp bách cho nhà nghiên cứu trong việc phân lớp giữa lưu lượng bình thường và lưu lượng tấn công DDoS trên mạng. Các mô hình ML được áp dụng vào lớp bài toán phát hiện tấn công DDoS trong nghiên cứu [1-14] đều có những ưu nhược điểm riêng, nhưng tồn tại những vấn đề chung cần được cải thiện:

- Tập dữ liệu được sử dụng đã lỗi thời và không được cập nhật cho những tấn công mới.

- Chỉ có thể phát hiện tấn công từ một host cụ thể, mà không thể phát hiện tấn công từ Bonet.

- Trong việc lựa chọn, rút gọn tập đặc trưng, việc sử dụng một mô hình/biện pháp đơn lẻ có thể dẫn tới việc làm mất thông tin, không chính xác hoặc phải mất nhiều thời gian điều chỉnh tham số mô hình để tối ưu kết quả. Ngoài ra, việc sử dụng các phương pháp trích chọn đặc trưng thường được sử dụng phổ biến như PCA lại không chỉ ra được được tập đặc trưng liên quan với tấn công DDoS.

- Thời gian thực hiện của mô hình.

Tạp chí Khoa học, Trường Đại học Sư phạm, Đại học Huế
ISSN 1859-1612, Số 3(59)/2021: tr.161-171
Ngày nhận bài: 16/3/2021; Hoàn thành phản biện: 23/3/2021; Ngày nhận đăng: 26/3/2021

Trong bài báo này, chúng tôi đề xuất mô hình phát hiện tấn công DDoS tập trung vào việc cải tiến bước rút gọn tập đặc trưng trên tập dữ liệu UNSW-NB15. Để rút gọn tập đặc trưng, thay vì sử dụng 1 phương pháp riêng lẻ (như đánh giá tương quan) hay sử dụng 1 mô hình ML (như PCA) như trong các nghiên cứu phát hiện DDoS được sử dụng gần đây, chúng tôi đề xuất sử dụng kết hợp nhiều mô hình với quan điểm: 1) Thay vì mất nhiều thời gian để tối ưu hóa các tham số của 1 mô hình, sự kết hợp của nhiều mô hình yếu sẽ cho kết quả tin cậy và tốt hơn( điều này được khẳng định qua các phương thức ensemble của ML), 2) Kết quả của mô hình đề trả lời được tập đặc trưng liên quan đến tấn công DDoS (điều mà PCA không làm được). Trên tập đặc trưng lựa chọn được, chúng tôi sử dụng lần lượt các mô hình phân lớp đơn giản được sử dụng phổ biến trong các nghiên cứu về tấn công DDoS như SVC, Kneighbor, Naïve Bayes, Random Forest để phát hiện tấn công DDoS. Đóng góp của nghiên cứu thể hiện ở những điểm sau:

- Sử dụng tập dữ liệu UNSW-NB15 thay vì sử dụng các tập dữ liệu được xem là lỗi thời như KDD 99. Phân tích đặc điểm lưu lượng tấn công từ Botnet để đưa ra tập gồm 14 đặc trưng  để làm đầu vào cho các bước tính toán và xử lý tiếp theo.

- Đề xuất mô hình rút gọn tập đặc trưng sử dụng kết hợp kết quả bầu chọn của 03 mô hình. Phương pháp này chỉ ra được tập đặc trưng nào liên quan nhất với tấn công DDoS và việc sử dụng kết hợp của nhiều mô hình cho kết quả đáng tin cậy hơn việc sử dụng một mô hình, phương pháp riêng lẻ.

Phần còn lại của bài báo được tổ chức như sau: phần II tóm tắt các nghiên cứu liên quan trong lĩnh vực phát hiện tấn công DDoS trong thời gian gần đây. Phần III sơ lược về tấn công DDoS và tập dữ liệu tấn công DDoS, phần IV trình bày về mô hình đề xuất phát hiện tấn công DDoS sử dụng ML. Đoạn V giới thiệu các kết quả kiểm nghiệm, đánh giá của mô hình. Cuối cùng, phần VI là kết luận của nghiên cứu cùng hướng nghiên cứu trong tương lai.

## 2. CÁC NGHIÊN CỨU LIÊN QUAN

Với sự triển nhanh chóng của các biến thể tấn công DDoS, nên các phương pháp tiếp cận truyền thống như sử dụng chữ kí (signature) đã bộc lộ nhược điểm không thể linh hoạt trong việc phát hiện tấn công, phản ứng trước các dạng tấn công mới [4,8]. Những nghiên cứu gần đây, sử dụng phương pháp ML cho phép các bộ lọc tự học trên các dữ liệu lịch sử đã có để nhận diện các lưu lượng bất thường trên mạng, đã và đang là hướng nghiên cứu được quan tâm với nhiều đề xuất, cải tiến.

Nhóm tác giả [1] đề suất mô hình học giám sát sử dụng bộ phân lớp tấn công Random Forest Classifier để phát hiện tấn công DDoS với độ chính xác 96%. Tuy nhiên mô hình chỉ thiện tiền xử lý dữ liệu và trích chọn đặc trưng bằng một kỹ thuật đơn lẻ mà không thông qua kết hợp của đồng thời của các mô hình để nâng cao mức độ đánh giá của các đặc trưng được trích chọn. Nhóm tác giả [3] đề xuất mô hình sử dụng ML để nhận diện tấn công DDoS từ các thiết bị IoT. Thay vì sử dụng tập dữ liệu công cộng, nhóm tác giả tự xây dựng kịch bản để thu thập bộ dữ liệu huấn luyện và kiểm thử. Từ bộ dữ liệu này,

nhóm tác giả đề xuất phương án lựa chọn đặc trưng thông qua việc đánh giá đặc tính lưu lượng mà chưa có cơ sở đánh giá kết hợp để thấy được độ quan trọng của các đặc trưng trong mô hình cũng như khả năng mất thông tin khi loại bỏ các đặc trưng còn lại trong tập dữ liệu ban đầu. Nhóm tác giả [4] sử dụng mô hình kết hợp PCA-RNN bao gồm trích chọn đặc trưng với PCA và sử dụng mạng Neuron hồi qui để phát hiện tấn công. Đóng góp nhóm tác giả là đề xuất tập đặc trưng của lưu lượng tấn công DDoS trước khi thực hiện trích chọn đặc trưng với PCA, với các pương pháp trích chọn đặc trưng, các biến độc lập mới sau khi biến đổi trở nên khó hiểu và cũng không chỉ ra được các đặc trưng nào liên quan đến tấn công DDoS. Ngoài ra, việc sử dụng RNN cũng tăng độ phức tạp của mô hình với tập dữ liệu lớn. Nhóm nghiên cứu [7] sử dụng kết hợp phương pháp đánh giá độ tương quan để rút gọn tập đặc trưng sau đó sử dụng mạng Neuron để phát hiện tấn công DDoS. Việc sử dụng phương pháp đánh giá độ tương quan để loại bỏ các đặc trưng có độ tương quan lớn (thường sử dụng ngưỡng > 0.75) tuy đơn giản, không tốn tài nguyên xử lý, nhưng đối với những đặc trưng có mối quan hệ không tuyến tính thì giá trị độ tương quan này không đủ cơ sở để loại bỏ đặc trưng, ngoài ra việc loại bỏ một đặc trưng sẽ ảnh hưởng đến độ quan trọng của các đặc trưng khác nên việc đánh rớt một nhóm các đặc trưng dựa vào độ tương quan dẫn tới mất thông tin cũng như hiệu suất của mô hình. Nhóm nghiên cứu [8] sử dụng phương Chi-bình phương và information gain để lựa chọn đặc trưng, Với các đặc trưng đó, nhóm nghiên cứu lần lượt thử nghiệm với các mô hình ML để phát hiện tấn công như Navies Bayes, C4.5, SVM, KNN, K-mean, Fuzzy_C means, trong đó Fuzzy_C mean cho kết quả chính xác nhất so với các mô hình còn lại. Bên cạnh các giải pháp phát hiện tấn công sử dụng ML theo tiếp cận học giám sát, nhóm nghiên cứu [9,10.12] sử dụng tiếp cận bán giám sát để tận dụng ưu điểm của cả các mô hình có giám sát và không giám sát, nhưng đồng thời sự phức tạp của mô hình dự báo cũng tăng lên.

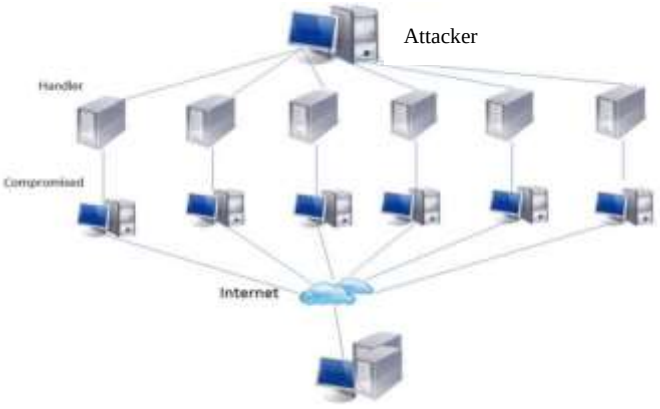
Việc rút gọn tập đặc trưng quyết định sự thành công của mô hình phát hiện tấn công DDoS sử dụng ML bởi điều này không chỉ làm giảm độ phức tạp tính toán của mô hình mà còn giảm hiện tượng overfiting của mô hình. Trong phạm vi bài báo này, chúng tôi sử dụng thuật ngữ rút gọn tập đặc trưng với ý nghĩa bao gồm: 1)lựa chọn tập đặc trưng đầu vào gồm 14 đặc trưng liên quan tới lưu lượng tấn công DDoS từ tập dữ liệu ban đầu và 2) sử dụng kết hợp 3 mô hình để giảm số chiều của tập đặc trưng đầu vào bằng phương pháp lựa chọn đặc trưng qua đó đưa ra tập đặc trưng liên quan đến tấn công DDoS.

## 3. TẤN CÔNG DDOS VÀ TẬP DỮ LIỆU TẤN CÔNG DDOS

### 3.1. Tấn công DDoS

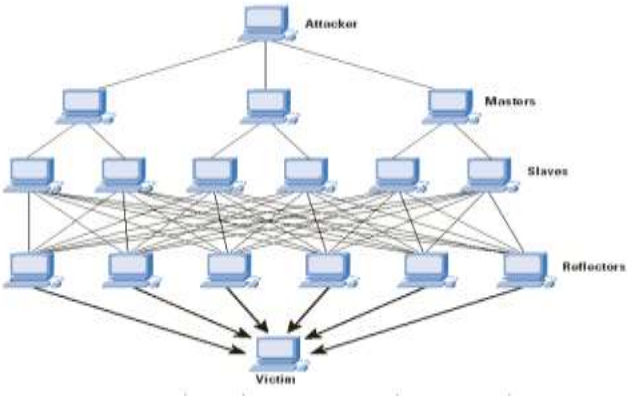
DDoS là tấn công từ chối dịch vụ phân tán, trong đó, kẻ tấn công tập hợp các máy tính đã bị tấn công trước đó thành một mạng lưới được gọi là Botnet và điều khiển chúng tấn công đồng thời vào một hoặc nhiều máy, dịch vụ, mạng đích [14]. Về kiến trúc, tấn công DDoS được chia thành 02 loại, đó là tấn công trực tiếp và tấn công gián tiếp.

- Trong kiến trúc tấn công trực tiếp (hình 1) kẻ tấn công sẽ điều khiện hệ thống máy tính ma (Bonet thông qua các máy trung gian (Handlerđ) để đồng loạt tạo và gởi các yêu cầu truy cập giả mạo đến hệ thống nạn nhân, gây ngập đường truyền mạng, khả năng xử lý của máy nạn nhân dẫn đến tình trạng gián đoạn hoặc ngừng dịch vụ cung cấp cho các người dùng khác.




Hình 1. *Mô hình tấn công DDoS trực tiếp*

- Trong kiến trúc tấn công DDoS gián tiếp (hình 2), kẻ tấn công điều khiển hệ thống các máy tính bị tấn công trước đó (Slave) để gởi đồng thời các yêu cầu truy cập giả mạo với địa chỉ nguồn là địa chỉ của máy nạn nhân đến một số các máy khác (gọi là Reflector _thường là các máy chủ có công suất lớn trên mạng Internet mà không chịu sự điều khiển của tin tặc) trên mạng Internet. Khi các Reflector có số lượng lớn, số phản hồi tạo ra có thể gây ngập đường truyền mạng hoặc làm cạn kiệt tài nguyên của máy nạn nhân, dẫn đến gián đoạn hoặc ngừng dịch vụ cung cấp cho người dùng.



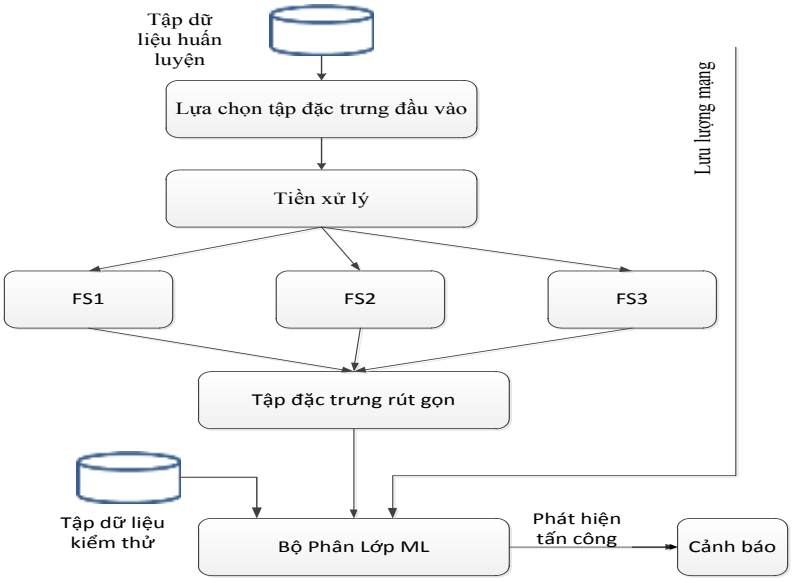

Hình 2. *Mô hình tấn DdoS gián tiếp*

### 3.2. Tập dữ liệu tấn công DDoS

Có 2 tập dữ liệu thường được sử dụng trong các nghiên cứu phát hiện tấn công DDoS là KKD 99 và NUSW-NB15. KKD 99 có 2 phiên bản DARAP98 và NSLKDD. Các nghiên cứu về hệ thống NIDS sử dụng tập dữ liệu KKD99 đã chỉ ra tập dữ liệu tồn tại những hạn chế bao gồm: dữ liệu trong tập dữ liệu lỗi thời vì vậy không có tính cập nhật với các loại lưu lượng mạng thông thường và tấn công hiện nay; tập dữ liệu tồn tại nhiều nhiều dữ liệu lặp, trống; phân bố xác suất của dữ liệu huấn luyện khác với phân

bố xác suất của lớp dữ liệu thử nghiệm. Việc sử dụng tập dữ liệu NUSW-NB15 có thể khắc phục những nhược điểm này [7,13].

Dữ liệu UNSW-NB15 gồm 49 đặc trưng và chứa tổng 2.540.044 bảng ghi được lưu trữ trong 4 file CSV và một phần dữ liệu trong tập này được chia thành tập dữ liệu huấn luyện và kiểm thử. Tập dữ liệu huấn luyện gồm 175.341 bản ghi, và tập dữ liệu kiểm thử gồm 82,332 bản ghi lưu trữ thông tin của 9 loại lưu lượng tấn công và bình thường. 9 loại lưu lượng tấn công gồm: Fuzzers, Analysis, Backdoor, DoS, Exploit, Generic, Reconnaissance, Shellcode, Worm. 49 đặc trưng của tập dữ liệu này được chia thành 6 nhóm đặc trưng [13].




Hình 3. *Mô hình đề xuất phát hiện tấn công DDoS*

## 4. MÔ HÌNH ĐỀ XUẤT PHÁT HIỆN TẤN CÔNG DDOS

Mô hình dò tìm/phát hiện tấn công sử dụng ML thực hiện phân lớp lưu lượng mạng là tấn công hay lưu lượng bình thường. Nếu có tấn công xảy ra, sẽ tạo cảnh báo cho hệ thống để chặn lưu lượng từ nguồn phát tương ứng hoặc/và đánh rớt gói tin. Mô hình gồm các pha: tiền xử lý dữ liệu, lựa chọn đặc trưng và dò tìm/phát hiện tấn công (hình 3).

### 4.1. Lựa chọn tập đặc trưng đầu vào

Trên cơ sở phân tích đặc điểm lưu lượng tấn công DDoS là chúng tôi đề xuất sử dụng 14 đặc trưng sau:

- Nhóm các đặc trưng về số lượng lưu lượng gồm 06 đặc trưng (sloat; dload; spkts, dpkts). Chúng tôi đề xuất sử dụng các đặc trưng này bởi khi các máy tính bị nhiễm và bị điều khiển bởi bot, chúng sẽ gởi lưu lượng lớn các gói tin vào mạng và máy đích làm cho mạng và máy đích trở nên quá tải. Ý nghĩa các đặc trưng này được cho ở bảng 1.

Bảng 1. *Các đặc trưng về lưu lượng*

| Stt | Tên | Giải thích |
|---|---|---|
| 1 | sload | Các bit được gởi từ nguồn trong thời gian 1 giây |
| 2 | dload | Các bit đến đích trong thời gian 1 giây |
| 3 | spkts | Số lượng gói tin từ nguồn đến đích |
| 4 | dpkts | Số lượng gói tin từ nguồn đến đích |
| 5 | sbytes | Số bytes gởi từ nguồn tới đích của |
| 6 | dbytes | Số bytes gởi từ đích tới nguồn |

- Nhóm các đặc trưng làm giảm chất lượng phục vụ (dur, ct_ftp_cmd, ct_srv_src, ct_srv_dst, ct_src_ltm, ct_src_dport_ltm, ct_dst_sport_ltm, ct_dst_src_ltm). Chúng tôi đề xuất sử dụng nhóm các đặc trưng này bởi đặc tính của tấn công DDoS là làm chậm hoặc làm ngập khả năng phục vụ của các đối tượng đích bằng cách chiếm dụng các kết nối trong thời gian dài làm cho đối tượng đích không thể phục vụ những người dùng hợp pháp khác.

Bảng 2. *Các đặc trưng chất lượng dịch vụ*

| Stt | Tên | Giải thích |
|---|---|---|
| 1 | dur | Chiều dài tính theo giây của các kết nối |
| 2 | ct_ftp_cmd | Số các luồng có thực hiện các lệnh command trong phiên ftp |
| 3 | ct_srv_src | Số lượng các kết nối chứa cùng dịch vụ và địa chỉ nguồn trong 100 kết nối |
| 4 | ct_srv_dst | Số lượng các kết nối chứa cùng dịch vụ và địa chỉ trong 100 kết nối |
| 5 | ct_src_ltm | Số lượng kết nối có cùng địa chỉ nguồn trong 100 kết nối |
| 6 | ct_src_dport_ltm | Số lượng các kết nối có cùng địa chỉ nguồn và port đích trong 100 kết nối |
| 7 | ct_dst_sport_ltm | Số lượng các kết nối có cùng địa chỉ đích và port nguồn trong 100 kết nối |
| 8 | ct_dst_src_ltm | Số lượng các kết nối có cùng địa chỉ nguồn và địa chỉ đích trong 100 kết nối. |

- Nhãn/đích: attack_cat là thuộc tính phân loại cho biết luồng đó là lưu lượng bình thường (normal) hay tấn công (9 loại tấn công kể trên). Với mục đích nghiên cứu mô hình nhận dạng tấn công DDoS, chúng tôi trích chọn các bản ghi tương ứng với lưu lượng bình thường và lưu lượng tấn công DoS, đồng thời chuyển kiểu dữ liệu của thuộc tính này: tấn công DDoS – 0; và các lưu lượng bình thường – 1.

Tập dữ liệu sau khi lựa chọn có 15 đặc trưng gồm 14 đặc trưng và 1 nhãn với 31,283 quan sát gồm lưu lượng tấn công DoS và lưu lượng thông thường.

### 4.2. Tiền xử lý dữ liệu

- Do dữ liệu của các đặc trưng trong tập dữ liệu không phải là phân bố chuẩn, việc chuẩn hóa cho tập các đặc trưng được áp dụng nhằm đưa phân bố dữ liệu của các đặc trưng về phân phối chuẩn. Công thức tính chuẩn hóa:

$$x^{(new)} = \frac{x - \mu}{\sigma}$$

Trong đó μ là giá trị trung bình; σ là phương sai, được tính bởi công thức sau:

$$\mu = \frac{1}{N}\sum_{i=1}^{N}(x_i);\ \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

### 4.3. Rút gọn tập đặc trưng đầu vào

Mượn ý tưởng của phương thức ensemble trong ML, chúng tôi đề xuất mô hình rút gọn tập đặc trưng sử dụng kết hợp nhiều mô hình để lựa chọn đặc trưng (trong bài này chúng tôi sử dụng 03 mô hình). Mỗi mô hình thực hiện bình chọn cho danh sách các đặc trưng có thể được rút gọn dựa trên độ quan trọng của đặc trưng đó với với mô hình. Kết quả cuối cùng, các đặc trưng được rút gọn là những đặc trưng được bầu chọn đồng thời bởi 3 mô hình. Chi tiết như sau:

**- Mô hình lựa chọn đặc trưng FS1: LassoCV()**

LassoCV được xếp vào nhóm thuật toán chuẩn hóa nhằm hạn chế sự khác biệt, chênh lệch giữa kết quả dự báo và kết quả thực tế của mô hình hồi quy tuyến tính. Kết quả của mô hình là danh sách giá trị độ quan trọng của các đặc trưng (đánh giá dựa vào giá trị coefficient), những giá trị thấp tương ứng với những đặc trưng có khả năng cao bị loại khỏi mô hình.

Đoạn mã thực thi sử dụng thư việc Sklearn trong Python:

```
from sklearn.linear_model import LassoCV
lcv = LassoCV()
lcv.fit(X_train, Y)
lcv_ mask = lcv.coef_!=0
```

Trong đó *X_train* là tập đặc trưng đầu vào*; Y là nhãn/biến phụ thuộc mask_lcv* là danh sách giá trị độ quan trọng của các đặc trưng trong X_train.

**- Mô hình lựa chọn đặc trưng FS2: RandomForestRegressor()**

RandomForest Là thuật toán học có giám sát tạo ra cây quyết định trên các mẫu dư xlieeuj được chọn ngẫu nhiên được dự đoán từ mỗi cây và chọn giải pháp tốt nhất bằng cách bỏ phiếu. Những đặc trưng được RandomForest đánh giá độ quan trọng thấp dựa vào giá trị đặc trưng độ quan trọng (feature_importance) là những đặc trưng ưu tiên được loại khỏi tập đặc trưng đầu vào.

Đoạn mã thực thi sử dụng thư việc Sklearn trong Python:

```
from sklearn.feature_selection import RFE
from sklearn.ensemble import GradientBoostingRegressor
rfe = RFE(estimator=GradientBoostingRegressor(), n_features_to_select=
sum(mask_lcv), step = 1, verbose=1)
rfe.fit(X_train,Y)
gbr_mask = rfe.support_
```

Trong đó *X_train* là tập dữ huấn luyện đã loại bỏ đặc nhãn*; Y là nhãn/ biến phụ thuộc; grb_mask* là danh sách bầu chọn các đặc trưng đầu vào.

**- Mô hình lựa chọn đặc trưng FS3: GradientBoostingRegressor()**

Gradient Boosting là thuật toán học có giám sát được sử dụng rộng rãi cho lớp bài toán hồi qui và phân lớp; sử dụng dụng phương thức ensemble để đưa ra kết quả dự báo từ kết quả tổng hợp của nhiều mô hình. Sử dụng kết quả của mô hình này, chúng ta thu được bầu chọn các đặc trưng có độ quan trọng thấp tương ứng là những đặc trưng ưu tiên được rút gọn từ tập đặc trưng đầu vào.

Đoạn mã thực thi sử dụng thư việc Sklearn trong Python:

```
rfe2 = RFE(estimator=RandomForestClassifier(),
n_features_to_select=sum(mask_lcv), step=1, verbose=1)
rfe2.fit(X_train,Y)
rfc_mask = rfe2.support_
```

Trong đó *X_train* là tập dữ huấn luyện đã loại bỏ nhãn*; Y là nhãn/ biến phụ thuộc , rfc_mask* là danh sách bầu chọn các đặc trưng đầu vào.

Kết quả của 3 mô hình FS1, FS2, FS3 được tổng hợp để bình chọn cho những đặc trưng nào có khả năng được loại bỏ khỏi tập đặc trưng đầu vào cao nhất. Đoạn mã thực thi trong Python như sau:

```
votes = np.sum([lcv_mask, rfc_mask,gbr_mask],axis=0))
mask = votes==3
X_train_reduced = df_X.loc[:, mask]
```

Trong đó *np* là thư viện numpy, *X_train_reduced* là tập dữ liệu X_train sau khi được rút rút gọn số chiều theo kết quả bầu chọn của 3 mô hình.

## 5. MÔ PHỎNG VÀ ĐÁNH GIÁ HIỆU QUẢ MÔ HÌNH

Để cài đặt, đánh giá hiệu quả của các mô hình, chúng tôi thực hiện xây dựng các mô hình bằng ngôn ngữ Python và chạy trên máy Windown 10, RAM 16 GB, Chip Intel® Core i5. Các thư viện được sử dụng cho ở bảng 3; Các mô hình ML được triển khai với các tham số mặc định.

Bảng 3. *Các thư viện trong Python được sử dụng trong chương trình*

| Stt | Thư viện | Chú thích |
|---|---|---|
| 1. | Pandas | Phân tích dữ liệu |
| 2. | Numpy | Xử lý mảng đa chiều, ma trận |
| 3. | Seaborn | Trực quan hóa dữ liệu |
| 4. | Matplotlib | Vẽ đồ thị 2D |
| 5. | Scikit-learn | Phân tích và khai phá dữ liệu |

**Các nhận định và so sánh**

Từ kết quả thực nghiệm ở bảng 4:

- So với trước khi thực hiện rút gọn tập đặc trưng, mô hình đề xuất cho kết quả tương đương đã cho thấy hiệu quả của mô hình đề xuất (giảm chiều tập đặc trưng, còn lại 05 đặc trưng so với 14 đặc trưng ban đầu) sẽ cải thiện độ phức tạp tính toán của mô hình phân lớp.

- Đối với phương pháp rút gọn sử dụng phương pháp đánh giá độ tương quan, mô hình đề xuất có kết quả tương đương tuy nhiên nếu với ngưỡng >=0.75, tập đặc trưng sau khi rút gọn là 10 đặc trưng. Ngoài ra, với phương pháp đánh giá này, nếu muốn giảm số chiều xuống 5 thì phải sử dụng ngưỡng >= 0.05, một giá trị ngưỡng không hợp lý trong phương pháp đánh giá tương quan.

- Đối với PCA với n_components=5, kết quả dự đoán của mô hình đề xuất cho kết quả tốt hơn. PCA sẽ cho kết quả tương đương với mô hình đề xuất khi n_components =10. Ngoài ra, như đã nhận định ở trên, với kết quả của PCA, không thể chỉ ra tập thuộc tính liên quan nhất với dạng lưu lượng tấn công DDoS.

- Với kết quả thực nghiệm có thể thấy mô hình đề xuất cho kết quả tốt nhất với bộ phân lớp ML Random Forest để phân lớp lưu lượng tấn công DDoS

Kết quả thử nghiệm của mô hình:

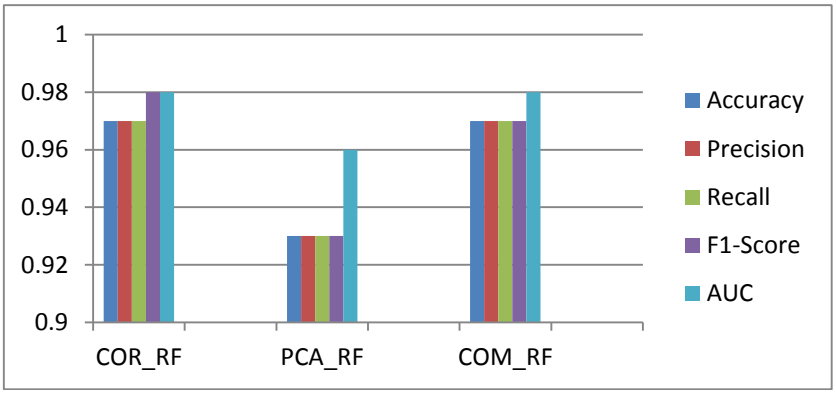
Bảng 4. *Kết quả đánh giá và so sánh của các mô hình*

| STT | Phương pháp rút gọn | Mô hình phân lớp ML | Accuracy | Precision | Recall | F1-Score | AUC |
|---|---|---|---|---|---|---|---|
| 1. | Không rút gọn | Random Forest | 0.97 | 0.97 | 0.97 | 0.97 | 0.99 |
| | | SVC | 0.91 | 0.92 | 0.91 | 0.90 | 0.96 |
| | | Kneighbors | 0.94 | 0.94 | 0.94 | 0.93 | 0.91 |
| | | Naïve Bayes | 0.84 | 0.83 | 8.84 | 0.82 | 0.84 |
| 2. | Đánh giá tương quan >=0.75 | Random Forest | 0.97 | 0.97 | 0.97 | 0.98 | 0.985 |
| | | SVC | 0.91 | 0.91 | 0.90 | 0.91 | 0.93 |
| | | Kneighbors | 0.93 | 0.93 | 0.93 | 0.93 | 0.89 |
| | | Naïve Bayes | 0.85 | 0.83 | 0.85 | 0.83 | 0.88 |
| 3. | PCA (n_component = 5) | Random Forest | 0.93 | 0.93 | 0.93 | 0.93 | 0.96 |
| | | SVC | 0.87 | 0.88 | 0.87 | 0.85 | 0.92 |
| | | Kneighbors | 0.93 | 0.93 | 0.93 | 0.93 | 0.89 |
| | | Naïve Bayes | 0.80 | 0.73 | 0.80 | 0.75 | 0.83 |
| 4. | Mô hình đề xuất | Random Forest | 0.97 | 0.97 | 0.97 | 0.97 | 0.98 |
| | | SVC | 0.90 | 0.91 | 0.90 | 0.89 | 0.85 |
| | | Kneighbors | 0.94 | 0.94 | 0.94 | 0.94 | 0.91 |
| | | Naïve Bayes | 0.84 | 0.82 | 0.84 | 0.82 | 0.81 |

Mô hình cho kết quả tập đặc trưng liên quan nhất của tấn công DDoS được tính toán bởi mô hình là tập gồm 5 đặc trưng: [‘dur’, ‘sload’, ‘sbytes’, ‘dbytes’, ‘ctt_srv_src’].

Bảng 5. *Kết quả đánh giá của mô hình phân lớp với Random Forest*

| STT | Phương pháp rút gọn | Mô hình phân lớp ML | Accuracy | Precision | Recall | F1-Score | AUC |
|---|---|---|---|---|---|---|---|
| 1. | Đánh giá tương quan >=75 | Random Forest (COR_RF) | 0.97 | 0.97 | 0.97 | 0.98 | 0.98 |
| 2. | PCA (n_component = 5) | Random Forest (PCA_RF) | 0.93 | 0.93 | 0.93 | 0.93 | 0.96 |
| 3. | Mô hình đề xuất | Random Forest COM_RF | 0.97 | 0.97 | 0.97 | 0.97 | 0.98 |




Hình 4. *Kết quả đánh giá của mô hình phân lớp với Random Forest*

## 6. KẾT LUẬN

Bài báo này đánh giá và so sánh tính hiệu quả của mô hình đề xuất sử dụng kết hợp 03 mô hình rút gọn tập đặc trưng trên tập dữ liệu NUSW-NB15. Kết quả kiểm nghiệm cho thấy mô hình đề xuất cho kết quả dự đoán tốt hơn, tập đặc trưng nhỏ hơn so với các phương pháp PCA và đánh giá độ tương quan. Mượn ý tưởng đoàn của phương thức ensemble trong machine learning, việc sử dụng kết quả tổng hợp từ 03 mô hình sẽ cho kết quả đáng tin cậy hơn so với việc sử dụng kết quả riêng lẻ của một mô hình. Kết quả kiểm nghiệm cũng cho thấy mô hình đề xuất cho kết quả tốt nhất với bộ phân lớp sử dụng Random Forest.

## TÀI LIỆU THAM KHẢO

[1] Saied, et al (2015). Detection of known and unknown DDoS attacks using Artificial Neural Networks, *Neurocomputing http://dx.doi.org/10.1016/j.neucom.2015.04.101i.*

[2] Andrew W. Moore and ndrew W. Moore (2005). Internet Traffic Classification Using Bayesian Analysis Techniques; *SIGMETRICS’05.*

[3] Jungtaek Seo1, Cheolho Lee1, Taeshik Shon2, Kyu-Hyung Cho2 (2005), A New DDoS Detection Model Using Multiple SVMs and TRA, *International Federation for Information Processing*, LNCS 3823, pp. 976 – 985.

[4] Kokila RT', Thamarai Selvi, Kannan Govindarajan (2014). DDoS Detection and Analysis in SDN-based Environment Using Support Vector Machine Classifier; *Sixth Interational Conference on Advanced Computingv(ICoAC).*

[5] *LuanPM –Adminvietnam (2015),* Kiến thức về DDOS, https://adminvietnam.org/kien-thuc-ve-ddos-phan-2-phan-loai/1031/.

[6] Manjula Suresh and R. Anitha (2011). Evaluating Machine Learning Algorithms for Detecting DDoS Attacks; *Springer-Verlag Berlin Heidelberg*, CNSA 2011, CCIS 196, pp. 441–452.

[7] Marwane Zekri and Youssef Saadi (2017). DDoS attack detection using machine learning techniques in cloud computing environments; *IEEE.*

[8] Mohamed Idhammad, Karim Afdel and Mustapha Belouch (2018). Semi-supervised machine learning, *Springer Science+Business Media, LLC, part of Springer Nature.*

[9] Mohamed Idhammad, Karim Afdel(*2017*). DoS Detection Method based on Artificial Neural Networks, *(IJACSA) International Journal of Advanced Computer Science and Applications,Vol. 8, No. 4.*

[10] Naveen Bindra*a*, and Manu Sood (2019). Detecting DDoS Attacks Using Machine Learning Techniques and Contemporary Intrusion Detection Dataset, *Automatic Control and Computer Sciences*, Vol. 53, No. 5, pp. 419–428.

[11] Nour Moustafa, Jill Slay (2016), The significant features of the UNSW-NB15 and the KDD99 data sets for Network Intrusion Detection Systems, *IEEE.*

[12] Qian Li, Linhai Meng, Yuan Zhang and Jinyao Yan (2019). DDoS Attacks Detection Using Machine Learning Algorithms, *Springer Nature Singapore Pte Ltd. 2019 G. Zhai et al. (Eds.),* IFTC 2018, CCIS 1009, pp. 205–216.

[13] Rohan Doshi, Rohan Doshi and Nick Feamster (2018). Machine Learning DDoS Detection for Consumer Internet of Things Devices, *IEEE Symposium on Security and Privacy Workshops.*

[14] Xiaoyong Yuan∗, Chuanhuang Li†, Xiaolin Li *(2017),* DeepDefense: Identifying DDoS Attack via Deep Learning; *IEEE.*

**Title:** A PROPOSED MACHINE LEARNING MODEL FOR DETECTING DDoS ATTACK

**Abstract:** DDoS attacks on the Internet have been badly affecting security and performance of the network. In addition to proposing and improving DDoS traffic classifier models, reducing the feature set is one of the key issues to increase predictor's efficiency as well as reducing the complexity of the model. In this paper, we proposed a DDoS attack detection model that used a combination of three models of feature selection from the input feauture set instead of using an individual model/ method used in recent approaches of detecting DDoS attack. With selected features, popular supervised learning models such as the SVC, Kneighbor, Naïve Bayes, Random Forest was used to detect DDoS attacks. Being evaluated with accuracy, F1score, AUC, our experiment show that the proposed method has better results.

**Keywords:** DDoS, SVC, Kneighbor, Naïve Bayes, Random Forest, Feature selection, Dimensionality reduction, Machine learning.