# THÔNG TIN CHUNG CỦA BÁO CÁO

- Link YouTube video của báo cáo (tối đa 5 phút):
  *https://youtu.be/BDUu3fha5ow*
- Link slides (dạng .pdf đặt trên Github):
  *https://github.com/BinhTruongCong/CS2205.APR2023*
- *Mỗi thành viên của nhóm điền thông tin vào một dòng theo mẫu bên dưới*
- *Sau đó điền vào Đề cương nghiên cứu (tối đa 5 trang), rồi chọn Turn in*

| | |
|---|---|
| ● Họ và Tên: Trương Công Bình<br>● MSSV: 230201039<br>  | ● Lớp: CS2205.APR2023<br>● Tự đánh giá (điểm tổng kết môn): 6/10<br>● Số buổi vắng: 0<br>● Số câu hỏi QT cá nhân:<br>● Link Github:<br>https://github.com/BinhTruongCong/CS2205.APR2023 |

# ĐỀ CƯƠNG NGHIÊN CỨU

**TÊN ĐỀ TÀI (IN HOA)**

ỨNG DỤNG MÔ HÌNH LSTM ĐƯỢC SỬ DỤNG ĐỂ PHÂN LOẠI TẤN CÔNG DDOS

**TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)**

USING THE LSTM MODEL FOR DDOS ATTACKS CLASSIFICATION

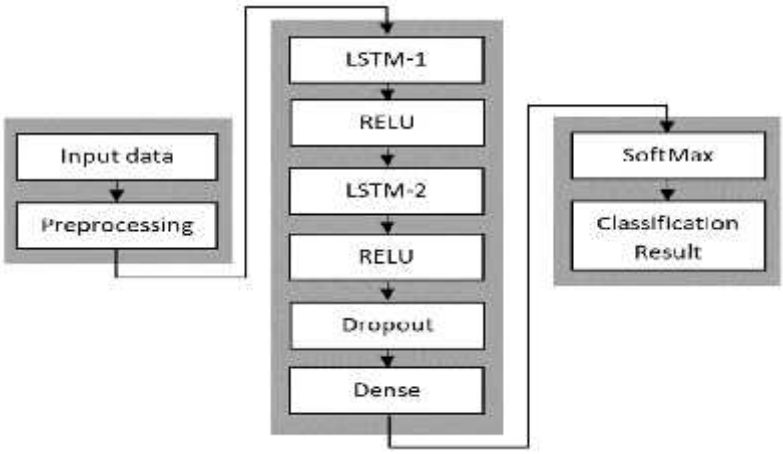
**TÓM TẮT** *(Tối đa 400 từ)*


- Kể từ khi các cuộc tấn công DDoS xuất hiện, cộng đồng nghiên cứu đã giải quyết mối đe dọa này thông qua một số kỹ thuật phát hiện, bao gồm: sơ đồ theo dõi ngược, hệ thống tự động lọc lưu lượng, phát hiện dựa trên chữ ký và phát hiện dựa trên sự bất thường. Trong đó mô hình Machine Learning (ML) có thể được sử dụng để phát hiện sự xâm nhập vào lưu lượng mạng như một trong những kỹ thuật phát hiện hiệu quả nhất. Đặc biệt, mô hình học sâu Deep Learning (DL) đã cho thấy hiệu quả xuất sắc trong những năm gần đây. Vì dữ liệu thực là phi tuyến, phức tạp và có nhiều chiều nên việc xây dựng mô hình (DL) có một số nơ-ron ẩn và mỗi nơ-ron có chức năng phi tuyến. Cấu trúc phức tạp của các mô hình (DL) giúp chúng hiểu rõ hơn các dữ liệu phức tạp và phi tuyến nhất định trong miền mục tiêu. Qua đó chúng tôi nghiên cứu việc sử dụng mô hình bộ nhớ ngắn hạn (LSTM) trong phân loại các cuộc tấn công DDoS đồng thời tập trung giải thích các dự đoán của mô hình (DL) bằng các phương pháp LIME, SHAP, Anchor và LORE


**GIỚI THIỆU** *(Tối đa 1 trang A4)*

- Các cuộc tấn công từ chối dịch vụ phân tán (DDoS) gây ra mối đe dọa đáng kể đối với an ninh mạng, nhằm mục đích lầm quá tải hệ thống bằng vô số lưu lượng truy cập bất hợp pháp, khiến hệ thống này không khả dụng đối với người dùng. Việc phát hiện sớm và chính xác là rất quan trọng để giảm thiểu các cuộc

tấn công này. Mạng Bộ nhớ ngắn hạn (LSTM), một loại mạng thần kinh nhân tạo (ANN), đã nổi lên như một công cụ mạnh mẽ để ngăn chặn DDoS nhờ khả năng xử lý dữ liệu chuỗi thời gian phức tạp như lưu lượng truy cập mạng.




Kiến trúc mô hình phân loại LSTM

- Đánh giá hiệu suất LSTM trong việc phát hiện các loại tấn công DDoS khác nhau bằng cách sử dụng các bộ dữ liệu CIC được sử dụng phổ biến và công khai
- Việc phân loại chính xác lưu lượng mạng giúp phát hiện sớm các cuộc tấn công DDoS và thực hiện các biện pháp phòng thủ kịp thời
- Một nghiên cứu đã sử dụng mô hình LSTM để phân loại lưu lượng mạng trong mạng máy tính và đạt được độ chính xác cao hơn 98% trong việc phát hiện các cuộc tấn công DDoS.
- Một nghiên cứu khác đã sử dụng mô hình LSTM để phân loại lưu lượng mạng trong mạng điện thoại di động và đạt được độ chính xác cao hơn 95% trong việc phát hiện các cuộc tấn công DDoS

**MỤC TIÊU**

*(Viết trong vòng 3 mục tiêu, lưu ý về tính khả thi và có thể đánh giá được)*

- Tăng cường khả năng thích ứng với các loại tấn công mới
- Hỗ trợ việc phân tích dữ liệu và ra quyết định
- Phát triển các hệ thống phòng chống DDoS hiệu quả hơn

**NỘI DUNG VÀ PHƯƠNG PHÁP**

*(Viết nội dung và phương pháp thực hiện để đạt được các mục tiêu đã nêu)*

***Nội dung:***

- Mô hình mạng nơ-ron dài ngắn hạn (LSTM) là một loại mạng nơ-ron nhân tạo (ANN) được sử dụng phổ biến trong xử lý ngôn ngữ tự nhiên, nhận dạng giọng nói, và phân tích chuỗi thời gian. Nó được thiết kế để giải quyết vấn đề biến mất gradient, một vấn đề thường gặp trong các mạng nơ-ron truyền thống khi xử lý các chuỗi dữ liệu dài.
- Mô hình LSTM cho phân loại tấn công DDoS thường bao gồm các thành phần sau:
- Lớp thu thập dữ liệu: Dữ liệu lưu lượng truy cập mạng được ghi lại, bao gồm các tính năng như địa chỉ IP nguồn và đích, cổng, giao thức, byte được truyền.
- Lớp tiền xử lý dữ liệu: Xử lý trước dữ liệu: Dữ liệu được thu thập được làm sạch, chuẩn hóa và định dạng để tương thích với LSTM.
- Lớp LSTM: Bao gồm nhiều lớp nơ-ron LSTM được kết nối với nhau. Các nơ-ron LSTM có khả năng học hỏi các phụ thuộc thời gian trong dữ liệu lưu lượng mạng.
- Lớp phân loại: Phân loại dữ liệu đầu vào là lưu lượng truy cập bình thường hay tấn công DDoS.
- Sử dụng mô hình LSTM để phân loại 17 loại tấn công DDoS.

***Phương Pháp***

1. **LIME (Local Interpretable Model-Agnostic Explanations)**
   - LIME tạo ra các mô hình giải thích cục bộ xấp xỉ mô hình LSTM bằng một mô hình tuyến tính đơn giản trong vùng lân cận của điểm dữ liệu được giải thích.

- Mô hình tuyến tính này được đào tạo trên một tập dữ liệu nhỏ được tạo ra từ điểm dữ liệu được giải thích và các điểm dữ liệu lân cận.
- Các trọng số của mô hình tuyến tính được sử dụng để xác định các đặc điểm quan trọng nhất cho dự đoán của LSTM tại điểm dữ liệu được giải thích.

2. **SHAP (SHapley Additive exPlanations):**
   - SHAP sử dụng phân bổ Shapley để phân bổ mức độ ảnh hưởng của mỗi đặc điểm đối với dự đoán của LSTM.
   - Phân bổ Shapley là một phương pháp phân bổ giá trị công bằng cho mỗi người chơi trong một trò chơi hợp tác.
   - Trong ngữ cảnh giải thích mô hình, mỗi đặc điểm được coi là một người chơi và giá trị Shapley của nó đại diện cho mức độ ảnh hưởng của nó đối với dự đoán.
3. **Anchor:**
   - Anchor tìm kiếm các điểm dữ liệu lân cận với điểm dữ liệu được giải thích có cùng dự đoán với LSTM.
   - Các điểm dữ liệu này được gọi là "anchor" và được sử dụng để giải thích dự đoán của LSTM tại điểm dữ liệu được giải thích.
   - Anchor có thể được sử dụng để so sánh điểm dữ liệu được giải thích với các điểm dữ liệu tương tự khác và xác định các đặc điểm khác biệt.
4. **LORE (Local Optimal Reconstruction Explanation):**
   - LORE sử dụng phương pháp tối ưu hóa cục bộ để tìm kiếm một tập con nhỏ các đặc điểm có thể tái tạo dự đoán của LSTM tại điểm dữ liệu được giải thích.
   - Tập con này được gọi là "LORE" và được sử dụng để giải thích dự đoán của LSTM.
   - LORE có thể được sử dụng để xác định các đặc điểm quan trọng nhất cho dự đoán của LSTM tại điểm dữ liệu được giải thích.

**KẾT QUẢ MONG ĐỢI**

*(Viết kết quả phù hợp với mục tiêu đặt ra, trên cơ sở nội dung nghiên cứu ở trên)*

- Mô hình LSTM đạt được độ chính xác cao trong việc phân loại các cuộc tấn công DDoS.
- Các phương pháp giải thích giúp hiểu rõ cách thức hoạt động của mô hình LSTM.
- 51 đặc điểm quan trọng được xác định để phân loại tấn công DDoS.
- Phương pháp LIME đạt hiệu suất tốt nhất về độ chính xác mô tả (DA) và độ thưa thớt mô tả (DS).

**TÀI LIỆU THAM KHẢO** *(Định dạng DBLP)*


1. Almaiah, M.A. Almaiah, M.A. A New Scheme for Detecting Malicious Attacks in Wireless Sensor Networks Based on Blockchain Technology. In Artificial Intelligence and Blockchain for Future Cybersecurity Applications; Springer: Berlin/Heidelberg, Germany, 2021; pp. 217–234. [Google Scholar]
2. Zargar, S.T.; Joshi, J.; Tipper, D. A Survey of Defense Mechanisms Against Distributed Denial of Service (DDoS) Flooding Attacks. IEEE Commun. Surv. Tutor. 2013, 15, 2046–2069. [Google Scholar]
3. Hou, J.; Fu, P.; Cao, Z.; Xu, A. Machine Learning Based DDos Detection Through NetFlow Analysis. In Proceedings of the IEEE Military Communications Conference MILCOM, Los Angeles, CA, USA, 29 October 2018. [Google Scholar]
4. DDoS Attacks History. Radware. Available online: https://www.radware.com/security/ddos-knowledge-center/ddos-chronicles/ddos-attacks-history/ (accessed on 17 July 2023).
5. Choi, H.; Lee, H. Identifying Botnets by Capturing Group Activities in DNS Traffic. Comput. Netw. 2012, 56, 20–33. [Google Scholar]
6. Suresh, S.; Ram, N. A Review on Various DPM Traceback Schemes to Detect DDoS Attacks. Indian J. Sci. Technol. 2016, 9, 1–8. [Google Scholar] [CrossRef] [Green Version]
7. Argyraki, K.; Cheriton, D. Active Internet Traffic Filtering: Real-Time Response to Denial of Service Attacks. arXiv 2003, arXiv:cs/0309054. [Google Scholar]
8. Anjum, F.; Subhadrabandhu, D.; Sarkar, S. Signature Based Intrusion Detection for Wireless Ad-Hoc Networks: A Comparative Study of Various Routing Protocols. In Proceedings of the IEEE 58th Vehicular Technology Conference, Orlando, FL, USA, 6 October 2003. [Google Scholar]

9. Cloudflare DDoS Threat Report 2022 Q3. Cloudflare. Available online: https://blog.cloudflare.com/cloudflare-ddos-threat-report-2022-q3/ (accessed on 17 July 2023).
10. Hoque, N.; Kashyap, H.; Bhattacharyya, D.K. Real-Time DDoS Attack Detection Using FPGA. Comput. Commun. 2017, 110, 48–58. [Google Scholar] [CrossRef]