**BÁO CÁO ĐỒ ÁN CUỐI KỲ**

**Môn học**

**CS519 - PHƯƠNG PHÁP LUẬN NGHIÊN CỨU KHOA HỌC**

**Lớp học**

**CS519.O21.KHTN**

**Giảng viên**

**PGS.TS. LÊ ĐÌNH DUY**  
**Thời gian**

**02/2024 - 06/2024**

*----- Trang này cố tình để trống -----*

# **THÔNG TIN CHUNG CỦA NHÓM**

* Link YouTube video của báo cáo (tối đa 5 phút):   
  *(ví dụ: https://www.youtube.com/watch?v=AWq7uw-36Ng)*
* Link slides (dạng .pdf đặt trên Github của nhóm):   
  *(ví dụ: https://github.com/mynameuit/CS519.O21.KHTN/TenDeTai.pdf)*
* *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: Hoàng Sơn Kim * MSSV: 12345   A person in a suit and tie  Description automatically generated | * Lớp: CS519.O21.KHTN * Tự đánh giá (điểm tổng kết môn): 7.5/10 * Số buổi vắng: 1 * Số câu hỏi QT cá nhân: 3 * Số câu hỏi QT của cả nhóm: 15 * Link Github: https://github.com/mynameuit/CS519.O21.KHTN/ * Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm:   + Lên ý tưởng XYZ   + Viết phần ABC   + Làm video YouTube |

# **ĐỀ CƯƠNG NGHIÊN CỨU**

|  |
| --- |
| **TÊN ĐỀ TÀI (IN HOA)**  ĐẾM SỐ NGƯỜI ĐEO KHẨU TRANG DÙNG DEEP LEARNING |
| **TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)** |
| **TÓM TẮT** *(Tối đa 400 từ)* |
| **GIỚI THIỆU** *(Tối đa 1 trang A4)* |
| **MỤC TIÊU** *(Viết trong vòng 3 mục tiêu)* |
| **NỘI DUNG VÀ PHƯƠNG PHÁP** |
| **KẾT QUẢ MONG ĐỢI** |
| **TÀI LIỆU THAM KHẢO** *(Định dạng DBLP)*  [1]. Tianyu Wang, Xiaowei Hu, Chi-Wing Fu, Pheng-Ann Heng:  Single-Stage Instance Shadow Detection With Bidirectional Relation Learning. CVPR 2021: 1-11 |

*----- Trang này cố tình để trống - Các nhóm copy & paste bài làm của mình vào trang tiếp theo -----*

# **THÔNG TIN CHUNG CỦA BÁO CÁO**

* Link YouTube video của báo cáo (tối đa 5 phút): <https://youtu.be/CdB5sbcwIQ8>
* Link slides (dạng .pdf đặt trên Github của nhóm): https://github.com/PNg-HA/CS519.O21.KHTN/blob/270b902d4d8fb081f70eda59bccce910dac98800/CS519.O21.KHTN.DeCuong.FinalReport.AIO.Slide.pdf

|  |  |
| --- | --- |
| * Họ và Tên: Phạm Nguyễn Hải Anh * MSSV: 21520586 | * Lớp: CS519.O21.KHTN * Tự đánh giá (điểm tổng kết môn): 7.5/10 * Số buổi vắng: 2 * Số câu hỏi QT cá nhân: 3 * Số câu hỏi QT của cả nhóm: 15 * Link Github: https://github.com/PNg-HA/CS519.O21.KHTN * Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm:   + Lên ý tưởng   + Viết nội dung |
| * Họ và Tên: Ngô Thanh Sang * MSSV: 21522543 | * Lớp: CS519.O21.KHTN * Tự đánh giá (điểm tổng kết môn): 7.5/10 * Số buổi vắng: 1 * Số câu hỏi QT cá nhân: 3 * Số câu hỏi QT của cả nhóm: 15 * Link Github: https://github.com/5hry/CS519.O21.KHTN/ * Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm:   + Làm slide   + Làm poster   + Làm video YouTube |

# **ĐỀ CƯƠNG NGHIÊN CỨU**

|  |
| --- |
| **TÊN ĐỀ TÀI (IN HOA)**  NGHIÊN CỨU, TRIỂN KHAI HỆ THỐNG PHÁT HIỆN XÂM NHẬP MẠNG DỰA TRÊN HỌC BẦY ĐÀN CHO IOT |
| **TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)**  RESEARCHING AND EVALUATION OF NETWORK-BASED INTRUSION DETECTION SYSTEM BASED ON SWARM LEARNING FOR IOT |
| **TÓM TẮT** *(Tối đa 400 từ)*  Nhu cầu sử dụng thiết bị IoT tại Việt Nam không ngừng tăng đi kèm với các thách thức về an ninh thông tin trong hệ thống IoT. Hệ thống phát hiện xâm nhập trên nền học liên kết (FL) được đề xuất nhằm đảm bảo bảo mật và sự riêng tư cho dữ liệu huấn luyện. Tuy nhiên, mô hình học liên kết hiện tại vẫn có một số vấn đề như sự phụ thuộc vào một trung tâm chuyên biệt, nguy cơ sự cố trung tâm, và tính riêng tư chưa được đảm bảo hoàn toàn do sự tồn tại của bên thứ ba. Các vấn đề an toàn khác bao gồm máy khách độc hại và các cuộc tấn công vào máy chủ trung tâm, cùng với nguy cơ bị đánh cắp hoặc chỉnh sửa file trọng số trong quá trình trao đổi.  Đề tài này đưa ra giải pháp cho hệ thống phát hiện xâm nhập mạng (NIDS) cho IoT, đảm bảo quyền riêng tư và an toàn bằng cách sử dụng khung học bầy đàn. Hệ thống này loại bỏ bên trung gian xử lý tham số và xác thực các máy trong mạng lưới bằng công nghệ Blockchain. Các mô hình học sâu phi tập trung được triển khai để đảm bảo hiệu suất khi huấn luyện trên các tập dữ liệu lớn. Hệ thống NIDS được thực nghiệm và đánh giá qua các dataset ToN-IoT và BoT-IoT. Đề tài cũng triển khai thuật toán mã hóa đồng cấu để bảo vệ tính riêng tư của dữ liệu. |
| **GIỚI THIỆU** *(Tối đa 1 trang A4)*  Nhu cầu sử dụng thiết bị IoT tại Việt Nam không ngừng tăng qua từng năm. Theo thống kê của Research and Market, năm 2021, quy mô thị trường IoT Việt Nam đạt khoảng 2,5 tỷ USD và dự kiến sẽ là 8,5 tỷ USD vào năm 2027[[1]](#footnote-2). Đi kèm với sự tăng trưởng này là những thách thức về an ninh thông tin trong hệ thống IoT[[2]](#footnote-3). Hệ thống phát hiện xâm nhập trên nền học liên kết (FL) được đề xuất như một giải pháp nhằm đảm bảo tính bảo mật cho hệ thống IoT cũng như sự riêng tư cho dữ liệu huấn luyện. Tuy nhiên, mô hình học liên kết vẫn tồn đọng nhiều vấn đề. Đó là vẫn tồn tại một trung tâm chuyên biệt nhằm xử lý các tham số từ các mô hình huấn luyện tại biên. Nếu trung tâm gặp sự cố sẽ ảnh hướng đến toàn bộ mô hình. Ngoài ra, tính riêng tư vẫn chưa được đảm bảo hoàn toàn khi vẫn có một bên thứ ba. Bên cạnh đó, tính an toàn của hệ thống học liên kết đang bị nghi vấn, khi trong hệ thống tồn tại những máy khách độc hại hay những cuộc tấn công lên máy chủ tập trung làm ảnh hướng tới mô hình huấn luyện được tổng hợp [1]. Ngoài ra, file trọng số (weight) có thể bị đánh cắp hoặc chỉnh sửa trong quá trình trao đổi giữa client và server trong quá trình huấn luyện của FL [2], từ đó có thể dịch ngược lại thành các dữ liệu huấn luyện.  Đề tài này hướng đến một giải pháp cho hệ thống phát hiện xâm nhập mạng (NIDS) cho IoT, đảm bảo quyền riêng tư và an toàn. Cụ thể, giải pháp NIDS này sử dụng khung học bầy đàn (Swarm Learning framework) [3] – đảm bảo tính phân tán khi không còn một bên trung gian xử lý các tham số, và các máy cùng mạng lưới được xác thực bằng việc sử dụng công nghệ Blockchain. Các mô hình học sâu phi tập trung được triển khai trên kiến trúc này nhằm đảm bảo hiệu suất khi huấn luyện trên các tập dữ liệu lớn. Tác giả thực nghiệm và đánh giá hệ thống NIDS thông qua các dataset ToN-IoT [5], BoT-IoT [6]. Đồng thời đề tài này cũng sẽ triển khai thuật toán mã hóa đồng cấu để thay thế các trọng số được gửi đi, nhằm đảm bảo tính riêng tư của dữ liệu. |
| **MỤC TIÊU** *(Viết trong vòng 3 mục tiêu)*   * Nghiên cứu triển khai framework học bầy đàn để đảm bảo tính an toàn trong mạng lưới cũng như quyền riêng tư dữ liệu của người dùng. * Tích hợp mô hình học sâu cho hệ thống phát hiện xâm nhập mạng cho IoT trên framework học bầy đàn. * Đánh giá hiệu suất của hệ thống NIDS được đề xuất. |
| **NỘI DUNG VÀ PHƯƠNG PHÁP**  **Nội dung 1:** Nghiên cứu và tìm hiểu các công nghệ liên quan  Nội dung:   * Tìm hiểu và triển khai CNN, RNN trên FL cho phát hiện xâm nhập cho IoT. * Tìm hiểu về Blockchain để xác thực máy khách và trao đổi trạng thái một cách bảo mật. * Tìm hiểu về mã hóa đồng cấu trong mô hình máy học.   Phương pháp: Đọc tài liệu, sách báo khoa học và các khóa học về các công nghệ liên quan.  Kết quả dự kiến: Nắm được cấu trúc, nguyên lý hoạt động và khả năng mở rộng của các công nghệ liên quan trong hệ thống. Đồng thời triển khai hệ thống học sâu phân tán.  **Nội dung 2**: Nghiên cứu, triển khai hệ thống trên framework học bầy đàn  Nội dung: Tìm hiểu, xây dựng hệ thống, triển khai framework học bầy đàn  Phương pháp:    **Hình 2:** Mô hình hoạt động của khung học bầy đàn   * Triển khai trên máy tính cá nhân và hệ thống cloud của khoa Mạng máy tính và Truyền thông – UIT. * Các nút SN (Swarm Network) trao đổi trạng thái nội bộ của các máy khách với nhau qua hợp đồng thông minh Etherum. Có 2 loại hợp đồng thông mình là hợp đồng cài đặt môi trường và hợp đồng thực hiện chương trình máy học đã được chuyển đổi tương thích SwarmCallback API. * Nút ML sẽ lấy tập dữ liệu của máy khách để huấn luyện, rồi gửi trọng số cho nút SL. Các nút SL sẽ chia sẻ trọng số này cho mạng lưới để tổng hợp, sau đó gửi lại trọng số đã tổng hợp cho nút ML để tiếp tục huấn luyện trong vòng lặp kế tiếp. * Nghiên cứu và triển khai các thuật toán bầu cử nút leader và thuật toán tổng hợp trọng số của framework Swarm Learning.   Kết quả dự kiến: Nắm bắt nguyên lý hoạt động và triển khai hệ thống trên framework Swarm Learning.  **Nội dung 3**: Xây dựng mô hình học sâu tích hợp mã hóa đồng cấu trên khung học bầy đàn  Nội dung:   * Sử dụng cho việc phát hiện xâm nhập mạng dựa trên framework Swarm Learning.   Phương pháp:   * Sử dụng mô hình CNN, RNN tích hợp mã hóa đồng cấu vào bài toán phát hiện xâm nhập mạng cho IoT [4], chỉnh sửa cho tương thích với framework Swarm Learning. * Sử dụng bộ dataset ToN-IoT [5], BoT-IoT [6].   Kết quả dự kiến: Hoàn tất việc triển khai hệ thống phát hiện xâm nhập mạng cho IoT sử dụng các mô hình học sâu trên khung học bầy đàn.  **Nội dung 4:** Thực nghiệm, đánh giá hiệu suất  Nội dung: Tổng hợp các kết quả thực nghiệm, đánh giá kết quả hiệu suất của hệ thống và viết báo cáo.  Phương pháp:   * Đo đạt các thông số liên quan khi thực nghiệm hệ thống * Tổng hợp các số liệu dựa trên các tiêu chí đánh giá. * Sử dụng công cụ để soạn báo cáo.   Kết quả dự kiến: Hoàn thành triển khai và đánh giá hệ thống phát hiện xâm nhập cho IoT dựa trên các tiêu chí Accuracy, Precision, Recall, F1-Score, True Positive Rate (TPR), False Positive Rate (FPR). |
| **KẾT QUẢ MONG ĐỢI**  Đề xuất, triển khai hệ thống phát hiện xâm nhập mạng cho hạ tầng IoT cloud của khoa Mạng máy tính và Truyền thông – UIT trên nền học sâu phân tán với framework Swarm Learning. Qua đó đảm bảo tính bảo mật cho hệ thống IoT và đánh giá hiệu suất khi huấn luyện trên các tập dữ liệu lớn. |
| **TÀI LIỆU THAM KHẢO** *(Định dạng DBLP)*  1. Li, Y., Chen, C., Liu, N., Huang, H., Zheng, Z., & Yan, Q: A blockchain-based decentralized federated learning framework with committee consensus. IEEE Network 2020: 234-241.  2. Sun, Y., Ochiai, H., & Esaki, H: Decentralized deep learning for multi-access edge computing: A survey on communication efficiency and trustworthiness. IEEE Transactions on Artificial Intelligence 2021: 963-972.  3. Warnat-Herresthal, S., Schultze, H., Shastry, K. L., Manamohan, S., Mukherjee, S., Garg, V., ... & Schultze, J. L.: Swarm learning for decentralized and confidential clinical machine learning. Nature 2021: 265-270.  4. Rashid, M. M., Khan, S. U., Eusufzai, F., Redwan, M. A., Sabuj, S. R., & Elsharief, M.: A Federated Learning-Based Approach for Improving Intrusion Detection in Industrial Internet of Things Networks. Network 2023: 158-179.  5. Moustafa, N.: A new distributed architecture for evaluating AI-based security systems at the edge: Network TON\_IoT datasets. Sustainable Cities and Society 2021: 102994.  6. Koroniotis, N., Moustafa, N., Sitnikova, E., & Turnbull, B.: Towards the development of realistic botnet dataset in the internet of things for network forensic analytics: Bot-iot dataset. Future Generation Computer Systems 2019: 779-796. |

1

1. https://vietnamnet.vn/thi-truong-iot-ty-usd-viet-nam-nen-bat-dau-tu-o-to-thiet-bi-dien-2154732.html [↑](#footnote-ref-2)
2. https://tapchicongthuong.vn/bai-viet/nhung-thach-thuc-trong-an-toan-thong-tin-he-thong-iot-va-mot-so-de-xuat-giai-phap-bao-mat-78403.htm [↑](#footnote-ref-3)