THÔNG TIN CHUNG CỦA NHÓM

- Link YouTube video của báo cáo (tối đa 5 phút): *https://youtu.be/WeCvxeNr0o8*
- Link slides (dạng .pdf đặt trên Github của nhóm): *https://github.com/minhvq20-droid/CS2205.CH201/blob/main/SHARK-EYESSHARK-EYES%20A%20MULTIMODAL%20FUSION%20FRAMEWORK%20FOR%20MULTI-VIEW-BASED%20PHISHING%20WEBSITE%20DETECTION.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*
- *Lớp Cao học, mỗi nhóm một thành viên*

- Họ và Tên: Võ Quang Minh
- MSSV: 250101043

- Lớp: CS2205.CH201
- Tự đánh giá (điểm tổng kết môn): 8/10
- Số buổi vắng: 1
- Số câu hỏi QT cá nhân: 6
- Số câu hỏi QT của cả nhóm: 6
- Link Github: https://github.com/minhvq20-droid

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

PHƯƠNG PHÁP PHÁT HIỆN WEBSITE LỪA ĐẢO DỰA TRÊN HỌC SÂU ĐA PHƯƠNG THỨC KHÁNG MẪU TRỐN TRÁNH

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

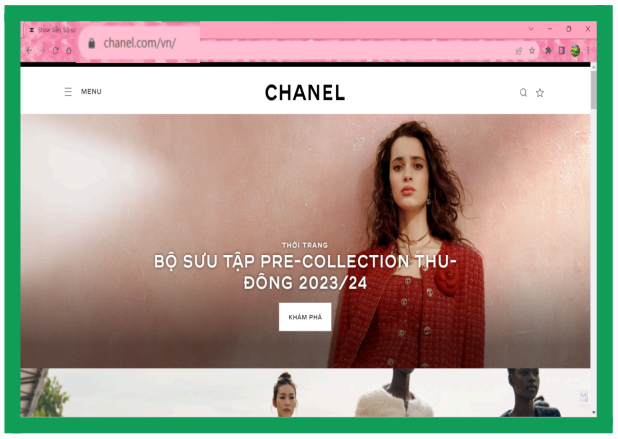
SHARK-EYES: A MULTIMODAL FUSION FRAMEWORK FOR MULTI-VIEW-BASED PHISHING WEBSITE DETECTION

TÓM TẮT *(Tối đa 400 từ)*

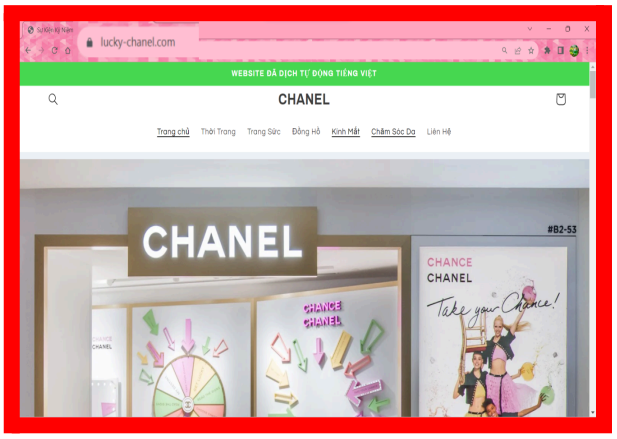
Trong thời kỳ số hóa hiện nay, công nghệ thông tin đã lan rộng đến nhiều lĩnh vực cuộc sống, mang đến nhiều tiện ích quan trọng cho con người. Đặc biệt, trong bối cảnh đại dịch Covid-19 kéo dài, nhu cầu sử dụng Internet để làm việc, giải trí, kết nối và học tập đã tăng đáng kể. Tuy nhiên, điều này cũng đồng nghĩa với việc gia tăng các nguy cơ mạng từ các hoạt động tội phạm, đặc biệt là tấn công Phishing. Phishing là một hình thức lừa đảo phổ biến, tận dụng sự thiếu hiểu biết về an ninh mạng của nạn nhân. Kẻ tấn công thường tạo ra các trang web giả mạo, sao chép giao diện của các trang chính thức để chiếm đoạt thông tin cá nhân, tài sản như số điện thoại, tài khoản ngân hàng, mật khẩu và thực hiện các giao dịch gian lận. Điều này không chỉ gây thiệt hại cho người dùng cá nhân mà còn đe dọa danh tiếng và tài chính của các tổ chức doanh nghiệp khi bị mạo danh. Để giải quyết vấn đề này, các nhà nghiên cứu đã áp dụng các phương pháp như thêm các trang web giả mạo vào danh sách đen hoặc sử dụng nhận diện dựa trên các chữ ký đặc trưng. Tuy nhiên, với sự tinh vi của tội phạm mạng ngày càng gia tăng, các phương pháp truyền thống này không còn đủ hiệu quả và đôi khi tốn thời gian. Do đó, các nhà nghiên cứu đang tích cực áp dụng các kỹ thuật học máy và công nghệ học sâu đa phương thức để cải thiện khả năng phát hiện giúp tiết kiệm thời gian và nâng cao hiệu quả. Vì vậy, chúng tôi nghiên cứu Shark-eyes phương pháp phát hiện website lừa đảo dựa trên học sâu đa phương thức kháng mẫu trốn tránh.

GIỚI THIỆU *(Tối đa 1 trang A4)*

Tấn công Phishing là một mối đe dọa trong lĩnh vực bảo mật thông tin và có ảnh hưởng nghiêm trọng đến các cá nhân, tổ chức. Dù đã xuất hiện từ lâu, tuy nhiên Phishing vẫn là một thách thức lớn trong bảo mật thông tin, gây thiệt hại tài chính và danh tiếng cho cá nhân và tổ chức. Các cuộc tấn công ngày càng tinh vi, kéo dài và khó phát hiện, sử dụng nhiều phương tiện như URL, websites, email, và SMS để đánh cắp thông tin nhạy cảm. Kẻ tấn công tạo ra các trang web giả mạo giống hệt trang thật để lừa người dùng. Theo APWG [1], số lượng tấn công Phishing đã tăng mạnh, với 250,000 cuộc tấn công chỉ trong tháng 1/2021 và đạt kỷ lục 4.7 triệu cuộc năm 2022.



***Hình 1**: Trang web chính thống của Chanel*



***Hình 2:** Trang web giả mạo của Chanel*

Trước thách thức này, nhiều giải pháp đã và đang được nghiên cứu và triển khai, chia thành ba nhóm chính: dựa trên danh sách (List-based), dựa trên trang tương đồng (Page similarity-based), và dựa trên máy học (Machine Learning-based) [2] [3]. Kỹ thuật List-based kiểm tra quyền truy cập dựa trên danh sách tên miền, URL nhưng gặp khó khăn trong việc cập nhật và duy trì danh sách. Kỹ thuật Page similarity-based so sánh giao diện trang web nhưng gặp thách thức khi trang lừa đảo sử dụng kỹ thuật gây nhiễu. Kỹ thuật Machine Learning-based sử dụng các đặc trưng liên quan đến nội dung, giao diện và URL để phân loại trang web, nhưng cần liên tục cập nhật mô hình để đối phó với kỹ thuật lẩn tránh mới của kẻ tấn công.

Trong thời gian gần đây, kỹ thuật học sâu (Deep Learning) đã được ứng dụng rộng rãi trong việc phát hiện các trang web lừa đảo [4] [5]. Với khả năng tự động trích xuất đặc trưng, các mô hình học máy sử dụng học sâu không cần dựa vào các kỹ thuật trích

xuất đặc trưng truyền thống và có khả năng thích ứng với các kỹ thuật lẩn trốn của kẻ tấn công thông qua dữ liệu huấn luyện. Đồng thời, các mô hình đa phương thức (multimodal models) cũng đã được triển khai [6]. Các mô hình này kết hợp nhiều mô hình riêng biệt, mỗi mô hình đảm nhận một nhiệm vụ cụ thể nhằm đạt được hiệu quả cao hơn trong việc phát hiện trang web lừa đảo. Đa mô hình là một lĩnh vực nghiên cứu mới, mang lại nhiều tiềm năng khai thác, với nhiều nghiên cứu đã triển khai thành công và thu được kết quả tích cực, vượt trội so với các mô hình học máy truyền thống.

Vì vậy, nghiên cứu này đề xuất một đa mô hình học sâu dựa trên cấu trúc DOM của HTML và tên miền URL để trích xuất và phân loại trang web, với cơ chế Attention giúp nhấn mạnh các thuộc tính quan trọng, cải thiện hiệu quả phát hiện trang lừa đảo, đặc biệt là các trang 0-day. Bên cạnh đó, chúng tôi kỳ vọng khi thực nghiệm mô hình đạt hiệu suất cao trong phát hiện và kháng nhiễu trước các cuộc tấn công đối kháng, không phụ thuộc vào thông tin bên thứ ba.

MỤC TIÊU

- Triển khai đa mô hình Shark-eyes phát hiện trang web lừa đảo bằng các thuộc tính được trích xuất bằng kỹ thuật hình học sâu từ các góc nhìn khác nhau.
- Đánh giá và so sánh từng phần của đa mô hình với nhau và so với tổng thể để kiểm tra độ hiệu quả khi triển khai đa mô hình.
- So sánh mô hình học sâu đa phương thức được đề xuất với các công trình nghiên cứu khác để kiểm tra hiệu suất và khả năng kháng nhiễu của mô hình Shark-eyes.

NỘI DUNG VÀ PHƯƠNG PHÁP

Nội dung 1: Tìm hiểu tổng quan về xu thế triển khai Phishing và các giải pháp ngăn chặn.

Phương pháp thực hiện:

- Tìm hiểu hiện trạng của các cuộc tấn công Phishing hiện nay thông qua các bài báo khoa học và các trang thống kê tấn công Phishing.

- Tìm hiểu về các phương pháp phát hiện Phishing được đề xuất trong thời gian gần đây thông qua các bài báo khoa học review.
- Tìm hiểu các mô hình học máy được ứng dụng triển khai trong lĩnh vực phát hiện Phishing.
- Tìm hiểu về các tập dữ liệu thường được sử dụng trong các công trình nghiên cứu thông qua các bài báo khoa học review.

Nội dung 2: Nghiên cứu và xây dựng mô hình học máy phát hiện trang web lừa đảo.

Phương pháp thực hiện:

- Nghiên cứu các mô hình học máy được đề xuất, công bố trong các bài báo khoa học để lấy ý tưởng.
- Tìm hiểu, nghiên cứu kiến trúc đa mô hình, điểm lợi, điểm hại của đa mô hình đem lại.
- Tìm hiểu các phương pháp tấn công đối kháng mô hình học máy trong Phishing.

Nội dung 3: Nghiên cứu mô hình học sâu đa phương thức ứng dụng vào trong các mô hình phát hiện trang web lừa đảo.

Phương pháp nghiên cứu:

- Triển khai thu thập dữ liệu các trang web lừa đảo và chính thống thông qua Internet.
- Triển khai kiến trúc mô hình học sâu đa phương thức Shark-eyes đã được thiết kế. Với nhiều nhánh đầu vào tập trung chủ yếu vào tên miền và danh sách các thẻ HTML. Cụ thể là phân tích tên miền khi chuyển thành ảnh xám, cấu trúc tên miền và cấu trúc các thẻ HTML được sử dụng.
- Triển khai lại các mô hình học máy phát hiện trang web lừa đảo ở các công trình nghiên cứu khác tập trung chủ yếu vào một nhóm đối tượng là URL, chủ yếu là phân tích cấu trúc URL.
- Thực hiện tấn công đối kháng các mô hình đã triển khai.

KẾT QUẢ MONG ĐỢI

- Tài liệu mô tả kiến trúc, khả năng của mô hình học sâu đa phương thức Shark-eyes.
- Giải quyết được các điểm hạn chế của các mô hình học máy trước đây.
- Triển khai mô hình học sâu đa phương thức Shark-eyes có khả năng phát hiện các trang web lừa đảo hiệu quả và chính xác.
- Kết quả thực nghiệm khả quan, đưa ra các đánh giá về hiệu năng, độ chính xác, khả năng kháng nhiễu của mô hình.

TÀI LIỆU THAM KHẢO *(Định dạng DBLP)*

[1]. Anti-Phishing Working Group. Phishing Attack Trends Report – 4Q 2022. [accessed 15-July-2023]. 2023. url: https://apwg.org/trendsreports/.

[2]. Rasha Zieni, Luisa Massari, Maria Carla Calzarossa:
Phishing or Not Phishing? A Survey on the Detection of Phishing Websites. IEEE Access 11: 18499-18519 (2023)

[3]. Wenhao Li, Selvakumar Manickam, Sham-ul-Arfeen Laghari, Yung-Wey Chong: Uncovering the Cloak: A Systematic Review of Techniques Used to Conceal Phishing Websites. IEEE Access 11: 71925-71939 (2023)

[4]. Manoj Kumar Prabakaran, Parvathy Meenakshi Sundaram, Abinaya Devi Chandrasekar:
An enhanced deep learning-based phishing detection mechanism to effectively identify malicious URLs using variational autoencoders. IET Inf. Secur. 17(3): 423-440 (2023)

[5]. Eman Abdullah Aldakheel, Mohammed Zakariah, Ghada Abdalaziz Gashgari, Fahdah A. Almarshad, Abdullah I. A. Alzahrani:
A Deep Learning-Based Innovative Technique for Phishing Detection in Modern Security with Uniform Resource Locators. Sensors 23(9): 4403 (2023)

[6]. Tadas Baltrusaitis, Chaitanya Ahuja, Louis-Philippe Morency:
Multimodal Machine Learning: A Survey and Taxonomy. IEEE Trans. Pattern Anal. Mach. Intell. 41(2): 423-443 (2019)

[7]. Jian Feng, Lianyang Zou, Ou Ye, Jingzhou Han:
Web2Vec: Phishing Webpage Detection Method Based on Multidimensional Features Driven by Deep Learning. IEEE Access 8: 221214-221224 (2020)
[8]. Jianting Yuan, Guanxin Chen, Shengwei Tian, Xinjun Pei:
Malicious URL Detection Based on a Parallel Neural Joint Model. IEEE Access 9: 9464-9472 (2021)
[9]. Peng Yang, Guangzhen Zhao, Peng Zeng:
Phishing Website Detection Based on Multidimensional Features Driven by Deep Learning. IEEE Access 7: 15196-15209 (2019)