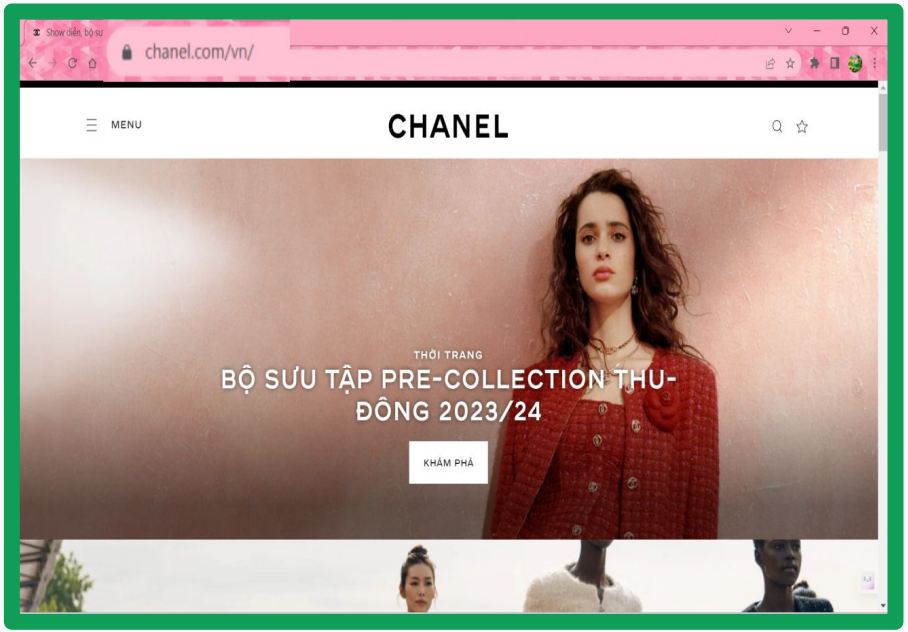
SHARK-EYES
A MULTIMODAL FUSION FRAMEWORK FOR MULTI-VIEW-BASED PHISHING WEBSITE DETECTION

Võ Quang Minh - 250101043

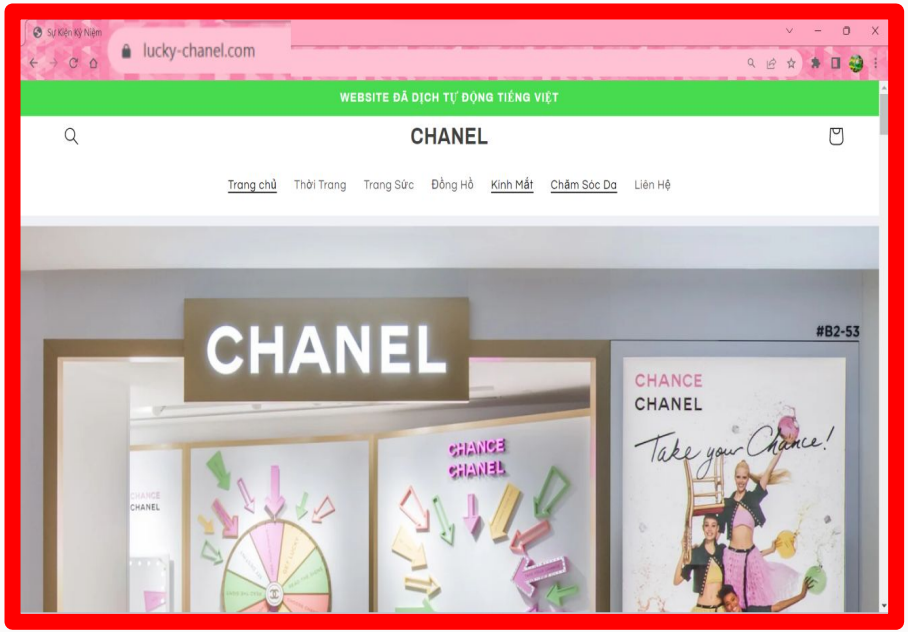
Tóm tắt

- Lớp: CS2205.CH201
- Link Github của nhóm:
- Link YouTube video: https://youtu.be/WeCvxeNr0o8
- Họ và Tên: Võ Quang Minh

Giới thiệu



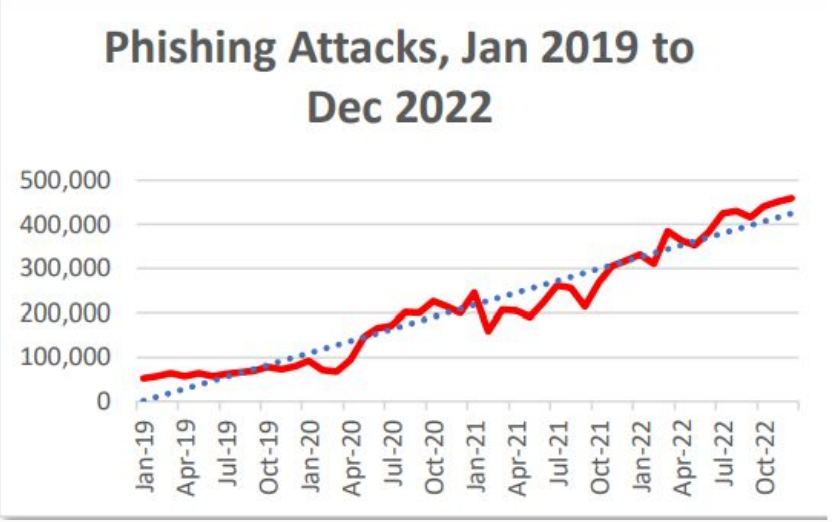
Trang web chính thống của Chanel



Trang web giả mạo của Chanel

Giới thiệu

- **Phishing** là một mối đe dọa trong lĩnh vực bảo mật thông tin.
- Mục tiêu là đánh cắp thông tin cá nhân và tài chính của cá nhân, tổ chức.
- Phương tiện lừa đảo: **URLs**.
- Không gian lan truyền: **mạng xã hội, email, SMS.**
- Mục tiêu nhắm đến:
  - ❖ Người thiếu kiến thức về công nghệ
  - ❖ Những lúc bất cẩn của người dùng



Hình 1: Biểu đồ thống kế số lượng các cuộc tấn công lừa đảo từ Anti-Phishing Working Group.

O Group, A.-P. W. Phishing attack trends report – 4q 2022, 2023. [accessed 15-July- 2023]

Mục tiêu

Triển khai đa mô hình	So sánh từng phần	Đánh giá tổng thể
Sử dụng kỹ thuật mô hình học sâu trích xuất thuộc tính từ các góc nhìn khác nhau.	So sánh và đánh giá từng phần của mô hình con. Kiểm tra hiệu quả khi triển khai đa mô hình (kết hợp các mô hình).	So sánh mô hình đề xuất với các công trình trước đó. Đánh giá hiệu suất và khả năng kháng nhiễu so với các phương pháp khác.

Nội dung và Phương pháp

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

Phương pháp thực hiện:

- Nghiên cứu hiện trạng các cuộc tấn công Phishing qua bài báo khoa học và thống kê (Anti Phishing Working Group).
- Tìm hiểu các phương pháp phát hiện Phishing được đề xuất gần đây qua các bài báo review.
- Tìm hiểu các mô hình học máy ứng dụng trong phát hiện Phishing.
- Tìm hiểu các tập dữ liệu phổ biến qua các bài báo review.

Nội dung và Phương pháp

Nội dung: 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 trong các bài báo khoa học.
- Tìm hiểu kiến trúc đa mô hình, điểm lợi và điểm hại.
- Nghiên cứu các phương pháp tấn công đối kháng mô hình học máy trong Phishing.

Nội dung và Phương pháp

Nội dung: 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:

- Thu thập dữ liệu các trang web lừa đảo và chính thống qua Internet.
- Triển khai mô hình Shark-eyes với nhiều nhánh đầu vào: tên miền và thẻ HTML (phân tích tên miền dưới dạng ảnh xám, cấu trúc tên miền, và cấu trúc thẻ HTML).
- Triển khai lại các mô hình học máy khác tập trung vào phân tích cấu trúc URL.
- Thực hiện tấn công đối kháng các mô hình đã triển khai.

Kết quả dự kiến

Kiến trúc và khả năng của mô hình	Mô hình học sâu đa phương thức, khắc phục hạn chế của mô hình cũ.
Phát hiện trang web lừa đảo	Triển khai mô hình phát hiện hiệu quả và chính xác.
Kết quả thực nghiệm	Hiệu năng cao, độ chính xác tốt, khả năng kháng nhiễu mạnh.

Tài liệu tham khảo

[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)