Archives available at journals.mriindia.com

International Journal on Advanced Computer Engineering and Communication Technology

ISSN: 2278-5140
Volume 14 Issue 01, 2025

Prevention Of Phishing Attack On Various Applications

Prof. V. S. Nalawade[1], Bankar Nandini Sanjay[2], Mohite Pooja Nanasaheb[3], Saykar Vaibhav Vikram[4], Padhar Tejas Khandeshwar[5]

[1]Dean of Academics and HOD of AIDS Dept, S. B. Patil College of Engineering
[2-5]Department of computer Engineering, Savitribai Phule Pune University
deanacademicssbpcoe@gmail.com[1], nandinibankar02@gmail.com[2], mohitepooja2004@gmail.com[3], saykarvaibhav124@gmail.com[4], tejaspadhar5252@gmail.com[5]

Peer Review Information

Submission: 15 Feb 2025
Revision: 23 March 2025
Acceptance: 27 April 2025

Keywords

Phishing Attacks
Cyber Fraud
Online Scams
Detection Techniques

Abstract

Phish-bowl Solutions is an emerging startup dedicated to tackling the escalating issue of online scams and cyber fraud. The company addresses a wide spectrum of digital deceit, including fraudulent donation solicitations, fake lottery schemes, impersonation tactics, counterfeit websites, phishing scams, and sextortion threats. In India, the scale of this problem is significant, with more than 30 percentage of the population falling victim to financial losses due to such scams in 2021 alone. Phish-bowl Solutions offers a comprehensive suite of tools and services designed to help individuals recognize, prevent, and mitigate the impact of online fraud. These offerings include advanced detection algorithms, educational resources, real-time scam alerts, and user-friendly interfaces for reporting suspicious activities. By leveraging cutting-edge technology and expertise, Phish-bowl Solutions aims to enhance digital security and empower users to safeguard their personal and financial information from malicious threats in the ever-evolving online landscape.

INTRODUCTION

Phishing attacks have become a significant threat in India, with over 30% of individuals experiencing financial losses due to online scams in 2021. These scams encompass various forms of cyber fraud, including fake donation requests, fraudulent lotteries, impersonation schemes, counterfeit websites, phishing attacks, and sextortion. Despite the widespread impact, many people remain vulnerable due to a lack of effective tools and guidance for identifying and preventing such threats. Traditional detection methods, such as list-based and similarity-based approaches, often fall short in addressing the evolving and sophisticated nature of phishing scams. This project aims to develop a comprehensive solution to help users recognize and avoid online scams, thereby protecting their financial and personal information from cyber fraud. By leveraging advanced detection techniques such as machine learning, deep learning, and natural language processing, the project seeks to enhance the accuracy and efficiency of phishing detection and prevention methods. Additionally, the integration of real-time monitoring systems and hybrid machine learning models will provide faster and more accurate detection, ensuring better security and resilience against online scams. This initiative is crucial for safeguarding users' financial and personal information in an increasingly digital world.

LITERATURE SURVEY

Here are the some limitations of previous paper,

1.A systematic literature review on phishing website detection techniques, Safi, Asadullah, and Satwinder Singh, 2023: Phishing involves tricking internet user s to divulge sensitive data by posing as a trus ted entity. A Systematic Literature Survey (S LR) analyzed various phishing detection met hods such as Lists-Based, Visual Similarity, Heuristic, Learning, and Deep Learning. Th e study reviewed 80 scientific papers from th e past five years. Machine Learning was the most common approach, with 57 studies emp loying it, particularly the Random Forest Cla ssifier used in 31 studies. The PhishTank and Alexa websites were primary data sources. Convolutional Neural Network (CNN) achie ved the highest accuracy at 99.98%. This SL R aims to provide an updated overview of th e latest trends and performance in phishing d etection techniques.

2. Detection of phishing websites using machine learning, Razaque, Abdul, Mohamed Ben Haj Frej, Dauren Sabyrov, Aidana Shaikhyn, Fathi Amsaad, and Ahmed Oun,2020: Phishing sends malicious links or attachments through emails that can perform various functions, including capturing the victim's login credentials or account information. These emails harm the victims, cause money loss, and identity theft. In this paper, we contribute to solving the phishing problem by developing an extension for the Google Chrome web browser. In the development of this feature, we used JavaScript PL. To be able to identify and prevent the fishing attack, a combination of Blacklisting and semantic analysis methods was used.

Furthermore, a database for phishing sites is generated, and the text, links, images, and other data on-site are analyzed for pattern recognition. Finally, our proposed solution was tested and compared to existing approaches. The results validate that our proposed method is capable of handling the phishing issue substantially.

3. Phishing or not phishing? A survey on the detection of phishing websites, Zieni, Rasha, Luisa Massari, and Maria Carla Calzarossa, 2023: Phishing is a persistent threat with serious effects on individuals and brands, continuously evolving to become more convincing and effective. This paper reviews the state of the art in phishing detection, discussing the main challenges and findings. It focuses on three important categories of detection approaches: list-based, similarity-based, and machine learning based. For each category, it describes the detection methods proposed in the literature, the datasets used for assessment, and identifies research gaps that need to be filled. This comprehensive review underscores the importance of ongoing innovation in phishing detection.

4. Phishing Detection Using Hybrid Machine learning Techniques , Helali, Rasha Gaffer M 2024: This research aims to develop a model to identify phishing websites by analyzing their common characteristics. A dataset of 549,346 entries was used to train various machine learning models, with the Random Forest Classifier achieving the highest accuracy of 94%. Feature selection and clustering techniques were also employed to enhance the model's ability to detect unknown attacks. The findings of this research can contribute to improving cybersecurity by providing a robust tool for identifying and preventing phishing attacks.

5. Real-time phishing detection using deep learning methods by extensions, Linh, Dam Minh, Ha Duy Hung, Han Minh Chau, Quang Sy Vu, and Thanh-Nam Tran 2024: Phishing is an attack method that relies on a user's insufficient vigilance and understanding of the internet. For example, an attacker creates an online transaction website and tricks users into logging into the fake website to steal their personal information, such as credit card numbers, email addresses, phone numbers, and physical addresses. This paper proposes implementing an extension to prevent phishing for internet users. In particular, this study develops a smart warning feature for the proposed extension using deep learning models. The proposed extension installed in the web browser protects users by checking for, warning about, and preventing untrusted connections. This study evaluated and compared the performance of machine learning models using a malicious uniform resource locator (URL) dataset containing 651,191 data samples. The results of the investigation confirm that the proposed extension using a convolutional neural network (CNN) achieved a high accuracy of 98.4%.

6.Phishing detection using natural language processing and machine learning, Mittal A, Engels DD, Kommanapalli H, Sivaraman R, Chowdhury T 2022: Phishing emails threaten organizations by mimicking legitimate communications. The DARTH framework uses machine learning to identify these threats

accurately. It characterizes and combines multiple models, each focusing on different features using Natural Language Processing (NLP) and neural networks. Tested on over 150,000 emails from various sources, including phishtank.com, DARTH achieved a 99.97% precision and 99.98% f- score, effectively identifying phishing emails 99.98% of the time. This multi-faceted, ensemble approach leverages machine learning techniques across a range of features to deliver highly accurate phishing email detection.

7.Machine learning for e-mail spam filtering: review, techniques and trends, Bhowmick, Alexy, and Shyamanta M. Hazarika, 2022: We present a comprehensive review of the most effective content-based e-mail spam filtering techniques. We focus primarily on Machine Learning-based spam filters and their variants, and report on a broad review ranging from surveying the relevant ideas, efforts, effectiveness, and the current progress. The initial exposition of the background examines the basics of e-mail spam filtering, the evolving nature of spam, spammers playing cat-andmouse with e-mail service providers (ESPs), and the Machine Learning front in fighting spam. We conclude by measuring the impact of Machine Learning-based filters and explore the promising offshoots of latest developments.

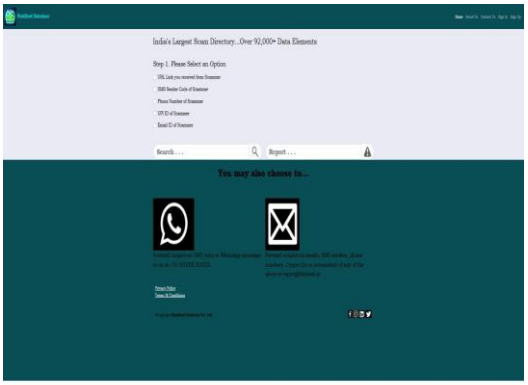
8. A novel ensemble machine learning method to detect phishing attack, Basit, Abdul, Maham Zafar, Abdul Rehman Javed, and Zunera Jalil, 2020: Phishing attacks have surged during the COVID-19 era, posing a major threat to internet users and organizations. Attackers use spoofed emails and fake websites to steal sensitive data. The sophistication of cybercriminals requires timely detection of phishing attempts. This research introduces a novel ensemble model to enhance phishing detection. It combines three machine learning classifiers: Artificial Neural Network (ANN), K-Nearest Neighbors (KNN), and Decision Tree (C4.5), with Random Forest Classifier (RFC). This ensemble method significantly improves detection accuracy. Experimental results show that the combination of KNN and RFC achieves an impressive 97.33% accuracy in identifying phishing attacks, outperforming existing studies.

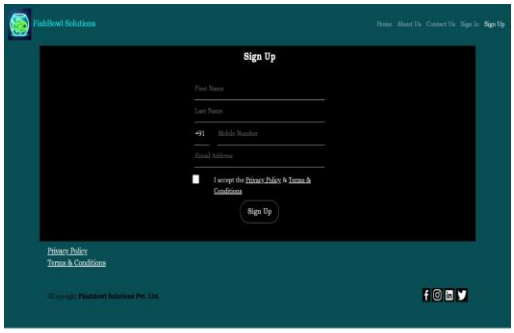
9.Phishing url detection with lexical features and blacklisted domains, Hong, Jiwon, Taeri Kim, Jing Liu, Noseong Park, and Sang-Wook Kim 2020: Many cyberattacks start with phishing to lure victims into malicious web pages where malware codes are hidden. Victim machines are infected by malware and the attacker can intrude the enterprise network, evading firewalls. Therefore, it is of fundamental importance to detect phishing URLs and prevent employees from visiting them. Many machine learning methods were proposed so far. In this work, we collect many lexical features after literature survey and combine them with blacklisted domains to improve the detection performance. We collect many recent phishing URLs because most of open datasets are outdated. Our method shows the F-1 of 0.84.

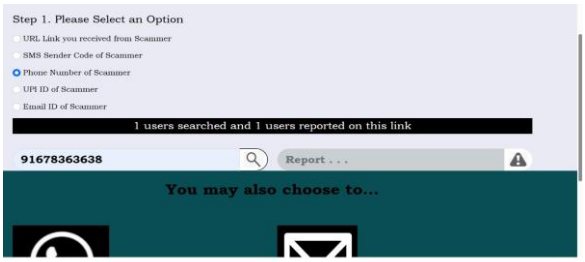
10.Phishing Detection Using Artificial Intelligence, Charan, Kuldeep, and Veeresh G. Kasabegoudar, 2021: The objective of this undertaking is to apply neural systems to phishing email recognition and assess the adequacy of this methodology. We structure the list of capabilities, process the phishing dataset, and execute the Neural Network frameworks. we analyze its exhibition against that of other real Artificial Intelligence Techniques – DT, K-nearest, NB and SVM machine. The equivalent dataset and list of capabilities are utilized in the correlation. From the factual examination, we infer that Neural Networks with a proper number of concealed units can accomplish acceptable precision notwithstanding when the preparation models are rare. Additionally, our element determination is compelling in catching the qualities of phishing messages, as most AI calculations can yield sensible outcomes with it.

Proposed System: The proposed system for "Prevention of phishing attack on various applications" is a novel ensemble model that combines multiple machine learning classifiers (ANN, KNN, C4.5, RFC) to accurately detect phishing attacks. The system leverages feature selection and clustering techniques to enhance its ability to detect unknown attacks. It is designed to provide real-time protection against phishing threats and can be deployed across different platforms, such as email, social media, and websites.

RESULTS







CONCLUSION

In conclusion, phishing attacks pose a significant threat to individuals and organizations, leading to substantial financial losses and compromised personal information. Despite the advancements in detection techniques, traditional methods often fall short in addressing the evolving and sophisticated nature of phishing scams. This project aims to bridge these gaps by leveraging advanced technologies such as machine learning, deep learning, and natural language processing to enhance the accuracy and efficiency of phishing detection and prevention. By developing a comprehensive solution, this project seeks to provide users with effective tools and guidance to recognize and avoid online scams, thereby safeguarding their financial and personal information from cyber fraud.

References

Safi, Asadullah, and Satwinder Singh. "A systematic literature review on phishing website detection techniques." *Journal of King Saud University-Computer and Information Sciences* 35.2 (2023): 590-611.

Razaque, Abdul, Mohamed Ben Haj Frej, Dauren Sabyrov, Aidana Shaikhyn, Fathi Amsaad, and Ahmed Oun. "Detection of phishing websites using machine learning." In *2020 IEEE Cloud Summit*, pp. 103-107. IEEE, 2020.

Zieni, Rasha, Luisa Massari, and Maria Carla Calzarossa. "Phishing or not phishing? A survey on the detection of phishing websites." *IEEE Access* 11 (2023): 18499-18519.

Helali, Rasha Gaffer M. "Phishing Detection Using Hybrid Machine learning Techniques." *Zhongguo Kuangye Daxue Xuebao* 29.2 (2024): 4552.

Linh, Dam Minh, Ha Duy Hung, Han Minh Chau, Quang Sy Vu, and ThanhNam Tran. "Real-time phishing detection using deep learning methods by extensions." *International Journal of Electrical and Computer Engineering (IJECE)* 14, no. 3 (2024): 3021-3035.

Mittal A, Engels DD, Kommanapalli H, Sivaraman R, Chowdhury T. Phishing detection using natural language processing and machine learning. SMU Data Science Review. 2022;6(2):14.

Bhowmick, Alexy, and Shyamanta M. Hazarika. "Machine learning for e-mail spam filtering:

review, techniques and trends." *arXiv preprint arXiv:1606.01042* (2016).

Basit, Abdul, Maham Zafar, Abdul Rehman Javed, and Zunera Jalil. "A novel ensemble machine learning method to detect phishing attack." In *2020 IEEE 23rd International Multitopic Conference (INMIC)*, pp. 1-5.
IEEE, 2020.

Hong, Jiwon, Taeri Kim, Jing Liu, Noseong Park, and Sang-Wook Kim. "Phishing url detection with lexical features and blacklisted domains." *Adaptive autonomous secure cyber systems* (2020): 253-267.

Charan, Kuldeep, and Veeresh G. Kasabegoudar. "Phishing Detection Using Artificial Intelligence." *Design Engineering* (2021): 8424-8436.

Nalawade, V. S., Jagtap, T. G., Jamdar, P. B., Kadam, S. I., & Kenjale, R. S. (2023). Voice-Enabled Traffic Sign Recognition and Alert System using ML: A Review.

Nalawade, V. S., Aoute, Y. P., Dharurkar, A. S., & Gunavare, R. D. (2023). A Survey on Revolutionizing Document Security: A Comprehensive Deep Learning Approach For Signature Detection and Verification.