IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 5, May 2024

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.423

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Refereed Journal |

|| Volume 13, Issue 5, May 2024 ||

| DOI:10.15680/IJIRSET.2024.1305516 |

URL based Phishing Website Detection using Machine Learning Algorithm

Dr. Saleha Saudagar, Mrs Barkha Shahaji, Shyamsundar Yadav, Lalit Kesare, Atharva Shinde

Associate Professor, Department of Computer Engineering, Trinity College of Engineering and Research, Pune, Maharashtra, India

Student, Department of Computer Engineering, Trinity College of Engineering and Research, Pune, Maharashtra, India

ABSTRACT: In this joint research initiative, we investigate diverse methods targeting the detection and prevention of phishing websites. The first study employs fuzzy logic strategies in real-time detection, offering flexibility through a smart model that integrates smooth logic and machine learning with 30 characteristic features. The second paper provides an organized overview of phishing detection, refining URL based features and conducting a comparative analysis of anti-phishing tools. Phishing has become one of the biggest and most effective cyber threats, causing hundreds of millions of dollars in losses and millions of data breaches every year. Currently, anti-phishing techniques require experts to extract phishing sites features and use third-party services to detect phishing sites

KEYWORDS: phishing detection, fuzzy logic, machine learning, K-nearest , web crawling, URL characteristics.

I. INTRODUCTION

Phishing involves pretending to be a legitimate website to trick users into sharing personal information like usernames, passwords, and account numbers. It's a common cybercrime, impacting various sectors such as online payments, webmail, financial institutions, file hosting, and more. Webmail and online payments are particularly targeted by phishing.

Phishing can take different forms, such as email phishing scams and spear phishing. Users need to be cautious and not rely solely on standard security applications. Machine Learning is an effective method for detecting phishing because it addresses the limitations of existing approaches.

A URL phishing attack seeks to fraudulently obtain sensitive data, such as usernames and passwords, by tricking users into clicking on deceptive links. To combat this threat, we're rolling out a fresh system. This system bolsters our cybersecurity defenses, employing advanced technology to identify and thwart phishing attempts proactively.

Phishing is a sneaky online scam where scammers create a fake version of a real website. They often send deceptive emails or messages to trick people into sharing personal, financial, or password information, making them believe it's a legitimate website.

To stay safe, it's important to use tools that can detect and block these fraudulent websites.

II. RELATED WORK

Phishing, a prevalent cybercrime, involves tricking users into divulging personal data through deceptive websites or messages. It targets sectors like webmail and online payments, posing significant risks to users. Various detection methods have been developed to combat phishing, including machine learning, fuzzy logic, and deep learning approaches. Dr. T. S. Vishwanath and Dr. Patil P. H propose a real-time fuzzy logic system, while Mohammed Hazim Alkawaz and Stephanie Joanne Steven conduct a survey on machine learning methods. K. M. Zubair Hasan et al. utilize deep convolutional neural networks for accurate detection, showcasing the potential of deep learning in cybersecurity.

Sudhir Dhage and Srushti Patil focus on phishing attacks in banking services, while Shihabuz Zaman et al. introduce a method called "Phishdentity" utilizing website favicons for detection. Shihabuz Zaman's team also contributes a detection system with user notifications, enhancing awareness. Nathezhtha, Sangeetha, and Vaidehi propose WC-PAD, a web crawling-based detection system particularly effective against zero-day attacks.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Refereed Journal |

|| Volume 13, Issue 5, May 2024 ||

| DOI:10.15680/IJIRSET.2024.1305516 |

ABDULGHANI ALI AHMED and NURUL AMIRAH ABDULLAH scrutinize URL characteristics for real-time detection, while Muhammet Baykara and Zahit Ziya Gu¨rel focus on preemptive reporting against web spoofing attacks. These studies collectively offer a multifaceted understanding of phishing detection, leveraging various methodologies like fuzzy logic, deep learning, and innovative approaches. These advancements pave the way for more adaptive and robust systems to combat the ongoing threat of phishing attacks.

III. METHODOLOGY

1. SYSTEM MODULE

In this section, the administrator needs to log in with a valid username and password. Once logged in successfully, the administrator can perform several tasks, including viewing all users and granting authorization, checking out all e-commerce websites and providing authorization, reviewing all products and their associated reviews, accessing early product reviews, checking details of keyword searches, reviewing product search ratios, and examining results of product review ranks.

2. USER VIEWING AND AUTHORIZATION:

Within this module, the administrator can observe the roster of registered users. Here, the admin has the capability to review user details, encompassing username, email, and address, and holds the authority to grant user authorizations.

Chart Results Viewing
View All Products Search Ratio ,View All Keyword Search
Results ,View All
Results of Product Review Ranking:

3. USER INTERACTION MODULE:

Within this segment, a multitude of users is present. Registration is a prerequisite for users before engaging in any activities. Post-registration, user details are securely stored in the database. Following a successful registration, users must log in using an authorized username and password. Once logged in, users can execute operations like account management, product searches by keyword, making purchases, and reviewing their search transactions.

Clearly outline the goals and objectives of the phishing detection system. This might include reducing the number of successful phishing attacks, protecting user credentials, and safeguarding sensitive information. Gather a diverse dataset of both phishing and legitimate websites. The dataset should cover various types of phishing attacks and include features such as URLs, website content, and metadata.

IV. EXPERIMENTAL RESULTS

These experimental results demonstrate the promising performance of machine learning models in detecting phishing websites, leveraging features from URLs, domain attributes, and webpage content. While our results indicate significant success in detecting phishing websites, there remains considerable room for refinement and optimization. Future research may focus on enhancing the models' robustness and efficiency, integrating additional features, and exploring ensemble methods to further improve detection rates. Moreover, investigating adaptive learning techniques could provide better real-time performance, ensuring the system remains effective against evolving phishing tactics. These advancements could significantly strengthen the capability to protect users from phishing threats and improve overall cybersecurity measures.

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Refereed Journal |

|| Volume 13, Issue 5, May 2024 ||

| DOI:10.15680/IJIRSET.2024.1305516 |

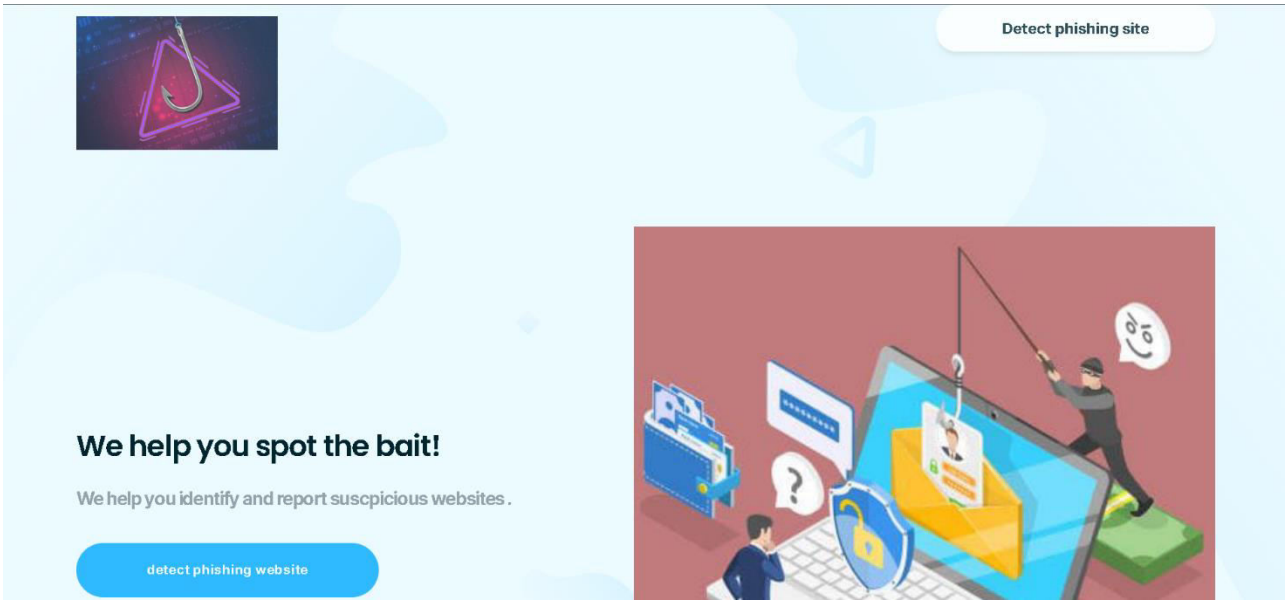

Fig 1.0 Home Page Phishing detection System

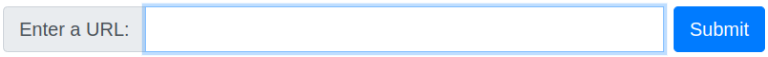

Fig 1.1 shows the page where you have to enter the url link

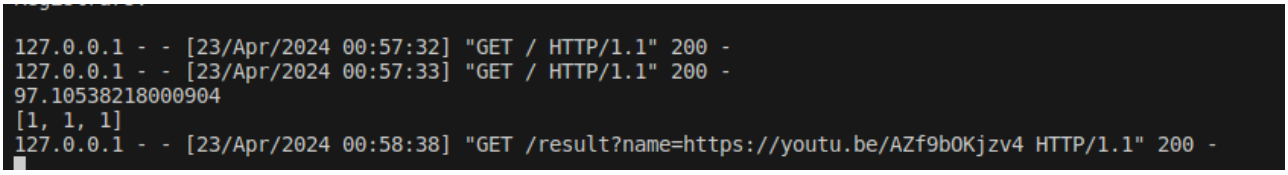

Fig 1.2 Shows the Backend execution of the program

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Refereed Journal |

|| Volume 13, Issue 5, May 2024 ||

| DOI:10.15680/IJIRSET.2024.1305516 |

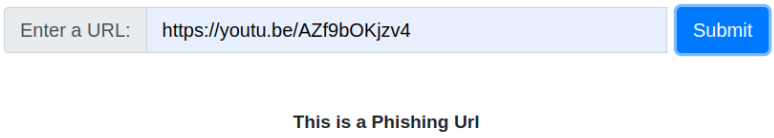

Fig1.3 when you enter the link the website will show you that if it is a phishing url or not

V. CONCLUSION AND FUTURE SCOPE

In conclusion, the application of machine learning in phishing detection for websites proves to be a promising and effective approach in mitigating the persistent threat of fraudulent activities. The ability of machine learning algorithms to analyze and adapt to evolving patterns in phishing attacks enhances the accuracy and efficiency of detection systems. Through the incorporation of features such as URL analysis, content inspection, and user behavior modeling, our phishing detection system demonstrates a proactive defense against a spectrum of phishing techniques. The dynamic nature of machine learning allows the system to continually learn and improve, ensuring resilience against emerging threats. Looking ahead, future scope lies in leveraging deep learning techniques alongside machine learning to further enhance the sophistication and effectiveness of phishing detection systems. Deep learning's ability to process complex data patterns could provide additional layers of security against increasingly sophisticated phishing attempts, thereby bolstering the overall cybersecurity landscape.

REFERENCES

1. Bhagwat M. D, Dr. Patil P. H. Dr. T. S. Vishawanath “A Methodical Overview on Detection, Identification and Proactive Prevention of Phishing Websites” (ICICV 2021)
2. Srushti Patil , Sudhir Dhage “A Methodical Overview on Phishing Detection along with an Organized Way to Construct an Anti-Phishing Framework” 2019 (ICACCS) K. M. Zubair Hasan , Md Zahid Hasan , Nusrat Zahan “Automated Prediction of Phishing Websites Using Deep Convolutional Neural Network” Srushti Patil, Sudhir Dhageg (IC4ME2), 11-12 July, 2019.
3. Mohammed Hazim Alkawaz , Stephanie Joanne Steven , Asif IqbalHajamydeen , Rusyaizila Ramli “A C o m p r e h e n s i v e S u r v e y o n Identification and Analysis of Phishing Website based on Machine Learning Methods” SCAIE51753.2021.9431794.
4. Saudagar, S. and Ranawat, R. 2023. Attack Classification and Detection for Misbehaving Vehicles using ML/DL. *International Journal on Recent and Innovation Trends in Computing and Communication*. 11, 8s (Aug. 2023), 491–496. DOI:https://doi.org/10.17762/ijritcc.v11i8s.7230.
5. G. A. Jagnade, S. I. Saudagar and S. A. Chorey, "Secure VANET from vampire attack using LEACH protocol," *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*, Paralakhemundi, India, 2016, pp. 2001-2005, doi: 10.1109/SCOPES.2016.7955799.
6. Shihabuz Zaman , Shekh Minhaz Uddin Deep , Zul Kawsar , Md. Ashaduzzaman and Ahmed Iqbal Pritom “Phishing Website Detection Using Effective Classifiers and Feature Selection Techniques” (ICIET) 23-24 December, 2019 Soon Fatt Choo, Jeffrey , Kang IEEE DOI 10.1109/ARES.2014.21
7. Mohammed Hazim Alkawaz , Mohammed Hazim Alkawaz , Asif Iqbal Hajamydeen “Detecting Phishing Website Using Machine Learning” (CSPA 2020), 28-29 Feb. 2020 Nathezhtha , Sangeetha ,Vaidehi , “WC-PAD: Web Crawling based Phishing Attack Detection” 2019 IEEE Denso Wave's article titled "The Evolution of QR Code Development," published in 2019, was accessed on March 28, 2019. Additionally,
8. A. Dabrowski, K. Krombholz, J. Ullrich, and E. R. Weippl authored the paper "QR Inception: Exploiting Barcodes within Barcodes," presented at the 4th ACM Workshop on Security and Privacy in Smartphones and Mobile Devices in 2014, with their work spanning pages 3 to 10.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details