“DETECTION OF PHISHING WEBSITES USING MACHINE LEARNING”

Qureshi Abdul Hai[1], Mohammed Fazl Ur Rehman[2], Omar Mohammed Ali[3], Mr. Syed Mushtaq Ali[4]

[1,2,3]UG Students, Department of CSE – Data Science, LIET, India

[4]Assistant Professor, Department of CSE – Data Science, LIET, India

ABSTRACT: *Phishing attacks are a rapidly expanding threat in the cyber world, costing internet users billions of dollars each year. It is a criminal crime that involves the use of a variety of social engineering tactics to obtain sensitive information from users. Phishing techniques can be detected using a variety of types of communication, including email, instant chats, pop-up messages, and web pages. This study develops and creates a model that can predict whether a URL link is legitimate or phishing.*

The data set used for the classification was sourced from an opensource service called ‘Phish Tank’ which contain phishing URLs in multiple formats such as CSV, JSON, etc. and also from the University of New Brunswick dataset bank which has a collection of benign, spam, phishing, malware & defacement URLs. Over six (6) machine learning models and deep neural network algorithms all together are used to detect phishing URLs.

This study aims to develop a web application software that detects phishing URLs from the collection of over 5,000 URLs which are randomly picked respectively and are fragmented into 80,000 training samples & 20,000 testing samples, which are equally divided between phishing and legitimate URLs. The URL dataset is trained and tested base on some feature selection such as address bar-based features, domain-based features, and HTML & JavaScript-based features to identify legitimate and phishing URLs.

In conclusion, the study provided a model for URL classification into phishing and legitimate URLs. This would be very valuable in assisting individuals and companies in identifying phishing attacks by authenticating any link supplied to them to prove its validity.

INTRODUCTION

The Internet has become an important part of our lives for gathering and disseminating information, particularly through social media. Internet is a network of computers containing valuable data, so there are many security mechanisms in place to protect that data, but there is a weak link: the human. When a user freely gives away their data or access to their computer, security mechanisms have a much more difficult time protecting their data and devices. Therefore, social engineering (a type of attack used to steal user data, including login credentials and credit card numbers) as a type of attack that is one of the most common social engineering attacks. The attack happens when an attacker fools a victim into opening an email, instant message, or text message as if it were from a trusted source. Upon clicking the link, the recipient is fooled into believing that they've received a gift and unsuspectingly clicks a malicious link, resulting in the installation of malware, the freezing of the system as part of a ransomware attack, or the disclosure of sensitive information.

DOI: https://doi-ds.org/doilink/09.2024-62585588

Computer security threats have increased substantially in recent years, owing to the rapid adoption of technology improvements, while simultaneously increasing the vulnerability of human exploitation. Users should know how the phishers do it, and they should also be aware of techniques to help protect themselves from becoming phished.

The strategies employed by cybercriminals are becoming more complex as technology advances. Other than phishing, there are a variety of methods for obtaining personal information from users. Following are some types of phishing:

- Vishing (Voice Phishing): This kind of phishing includes the phisher calling the victim to get personal information about the bank account. The most common method of phone phishing is to use a phone caller ID.
- Smishing (SMS Phishing): Phishing via Short Message Service (SMS) is known as Smishing. It is a method of luring a target through the SMS text message service by sending a link to a phishing website.
- Ransomware: A ransomware attack is a type of attack that prevents users from accessing a device or data unless they pay up.
- Malvertising: Malvertising is malicious advertising that uses active scripts to download malware or push undesirable information onto your computer. The most prevalent techniques used in malvertisements are exploits in Adobe PDF and Flash.

Hence, this is a rapidly evolving threat to individuals as well as big and small corporations. Criminals now have access to industrial-strength services on the dark web, resulting in an increase in the amount of these phishing links and emails, and, more frighteningly, they are increasing in ‘quality,' making them tougher to detect.

LITERATURE SURVEY

The Phishing Landscape: Scope, Impact, and Industry Trends

Phishing attacks remain a pervasive threat in today's digital landscape, posing significant risks to both individuals and organizations. Verizon's 2023 Data Breach Report highlights the severity of this issue, indicating that phishing accounted for 36% of all data breaches in the US that year. Further emphasizing this concern, a Proofpoint study revealed that a staggering 71% of companies fell victim to a successful phishing attack in 2023. The magnitude of the problem is further underscored by data from the Anti-Phishing Working Group, showing that the number of unique phishing sites reached a record-breaking 5 million in 2023. Phishing attacks are not only widespread but also financially damaging, ranking as the second costliest source of compromised credentials according to the IBM Cost of a Data Breach Report.

Attackers employ a variety of techniques to deceive victims. The most prevalent method, as described in the aforementioned studies, is email phishing, where deceptive emails seemingly from legitimate sources are sent to steal sensitive information. Spear phishing tactics involve increased personalization, with attackers researching individuals beforehand to craft more believable messages and increase success rates. High-profile individuals are targeted in whaling attacks, aiming to gain access to crucial information or initiate fraudulent transfers. Another technique, pharming, redirects users to fraudulent websites that mimic legitimate ones to steal personal information.

The impact of phishing scams extends beyond just financial losses. Individuals targeted by these attacks can suffer from identity theft, damaged credit scores, and emotional distress. Businesses, on the other hand, face significant financial repercussions due to data breaches, lost

DOI: https://doi-ds.org/doilink/09.2024-62585588

productivity, potential regulatory fines, and reputational harm. The true cost of phishing attacks can be even greater, encompassing the intangible costs associated with a damaged business reputation, the erosion of consumer trust, and the ongoing expenses associated with implementing preventative measures.

Phishing Trends and Techniques: A Look at the 2023 Cloudflare Report

This section summarizes key findings from the 2023 Cloudflare Phishing Threats Report, offering valuable insights into evolving phishing tactics.

Deceptive Links Dominate: The report highlights deceptive links as the primary phishing tactic. Attackers continuously refine methods to entice users into clicking. This emphasizes the importance of user awareness and training to recognize suspicious links, regardless of their displayed text.

Identity Deception Remains Prevalent: The report underlines the continued threat of identity deception, encompassing tactics like BEC and brand impersonation. These techniques can bypass email authentication, making them a significant concern. Phishing attempts often impersonate trusted entities users interact with regularly, increasing the likelihood of success.

Targeted Impersonation for Maximum Impact: The report reveals that attackers target a broad range of organizations but primarily impersonate entities users trust and interact with for work purposes. This targeted approach personalizes the attack and increases the chance of victims falling prey.

Advanced Detection Techniques for Elusive Threats: The report acknowledges the complex nature of phishing emails, often combining social engineering with technical obfuscation. To combat this, Cloudflare utilizes advanced detection methods like sentiment analysis, structural analysis, and machine learning models to identify "fuzzy" signals beyond just visible content. Trust graphs are also employed to evaluate sender legitimacy and potential impersonations.

Threat Intelligence for Enhanced Protection: The report emphasizes the importance of threat intelligence in email security. Cloudflare integrates threat intelligence from its global network, blocking billions of cyber threats daily. This intelligence feeds into email classifications (malicious, BEC, spoof, or spam) alongside specific threat indicators for each disposition.

Top Email Threat Indicators: The report details the most common threat indicators observed, including deceptive links, domain age (newly registered domains), identity deception, credential harvesting, and brand impersonation. Understanding these indicators helps identify suspicious emails and develop appropriate security measures.

By incorporating these insights from the Cloudflare report into your project's literature survey, you gain valuable knowledge about current phishing trends and techniques. This knowledge empowers you to develop a more robust system for detecting and mitigating phishing attacks.

The Phishing Landscape: Scope, Impact, and Industry Trends

Phishing attacks remain a pervasive threat in today's digital landscape, posing significant risks to both individuals and organizations. Verizon's 2023 Data Breach Report highlights the severity of this issue, indicating that phishing accounted for 36% of all data breaches in the US that year. Further emphasizing this concern, a Proofpoint study revealed that a staggering 71% of companies fell victim to a successful phishing attack in 2023. The magnitude of the problem is further underscored by data from the Anti-Phishing Working Group, showing that the number of unique phishing sites reached a record-breaking 5 million in 2023. Phishing attacks are not only widespread but also financially damaging, ranking as the second costliest

DOI: https://doi-ds.org/doilink/09.2024-62585588

source of compromised credentials according to the IBM Cost of a Data Breach Report.

SYSTEM DESIGN

Architectural design is concerned with understanding how a system should be organized and designing the overall structure of that system, it shows how different components of the system work together to achieve its main objectives. It is the process for identifying sub-systems making up a system and the framework for sub-system control and communication. The diagram below represents a graphical overview of the architectural design of the proposed system.

Figure 4.1 shows the architecture view of the proposed phishing detection system such that a user enters a URL link and the link moves through different trained machine learning and deep neural network models and the best model with the highest accuracy is selected. Thus, the selected model is deployed as an API (Application Programming Interface) which is then integrated into a web application. Hence, a user interacts with the web application which is accessible across different display devices such as computers, tablets, and mobile devices.

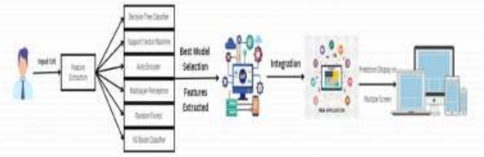

Figure 1 Architectural Design of the Proposed System

FLOWCHART DIAGRAM

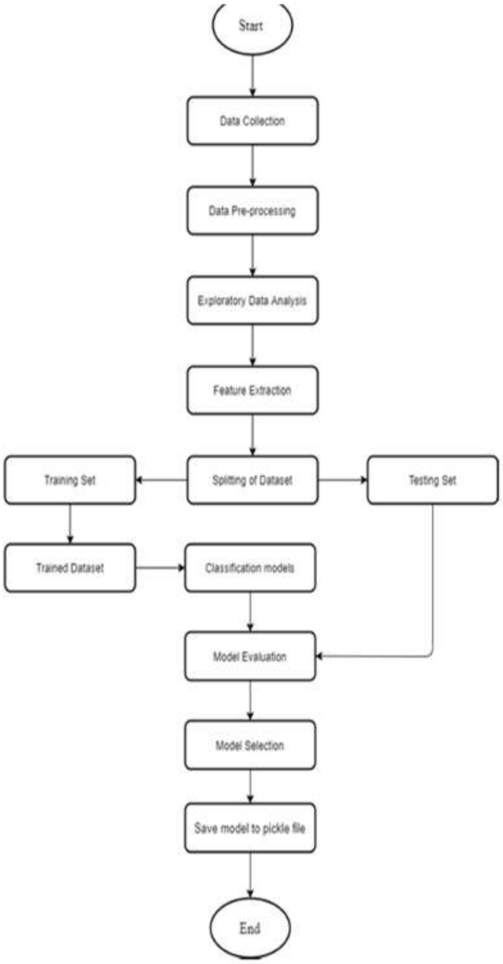

Figure 2 Flowchart Diagram

Data Analysis & Visualization

The image as shown in figure 3 shows the distribution plot of how legitimate and phishing datasets are distributed base on the features selected and how they are related to each other.

In figure 4 shows the plot of a correlation heat-map of the dataset. The plot shows correlation between different variables in the dataset.

DOI: https://doi-ds.org/doilink/09.2024-62585588

In figure 5 and figure 6, it shows the feature importance in the model for Decision tree classifier and Random Forest classifier respectively.

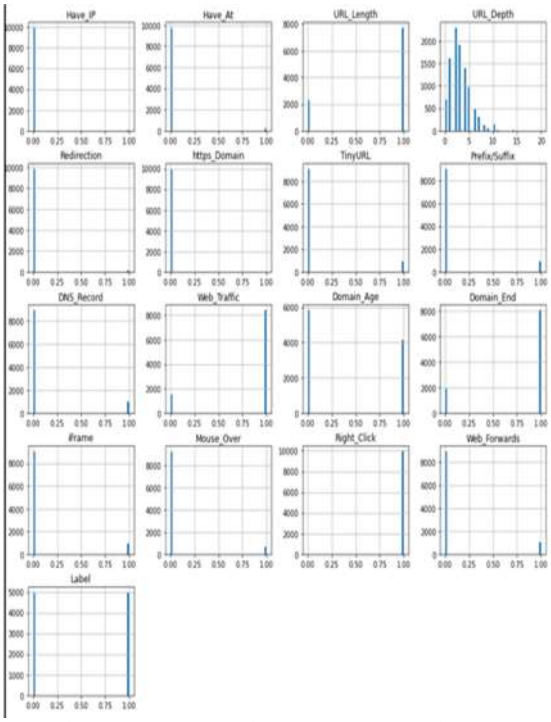

Figure 3 Distribution Plot of Dataset based on the features selected

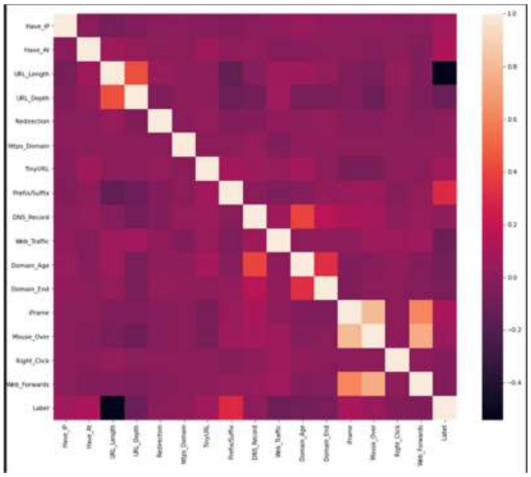

Figure 4 Correlation Heat map of the dataset

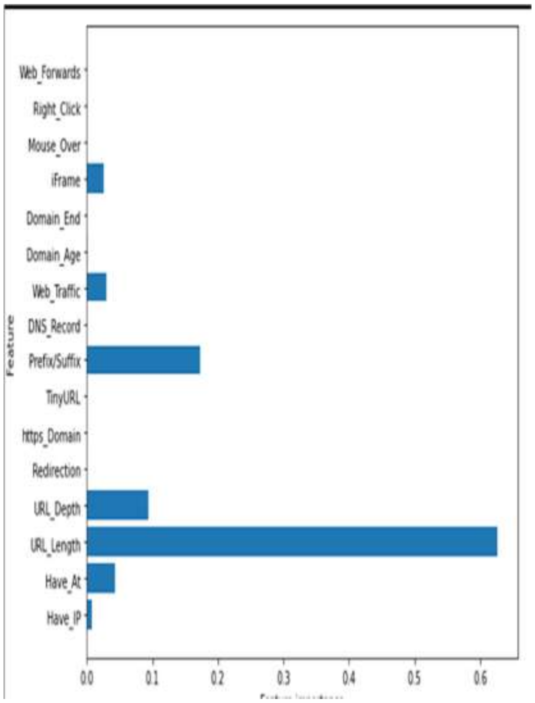

Figure 5 Feature Importance for Decision Tree Classifier

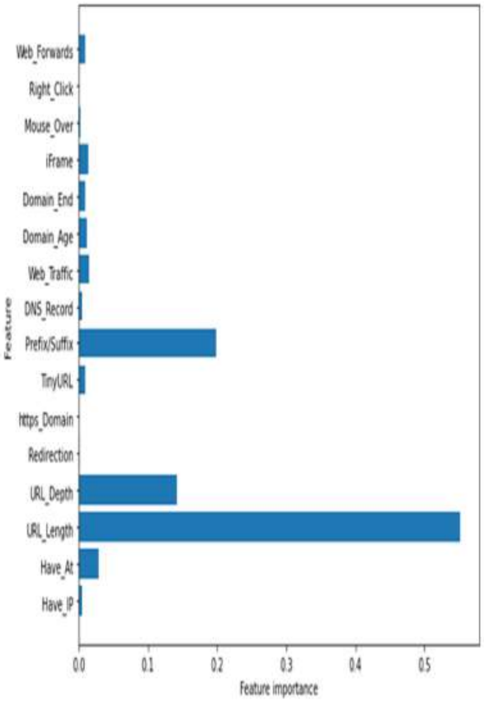

Figure 6 Feature Importance for Random Forest Classifier

IMPLEMENTATION

This chapter deals with the implementation of multiple machine learning models for the detection of phishing websites.

The implementation is concerned with all the activities that took place to put up the newly developed system into operation (using the approach that was stated in the methodology) to achieve the objectives of the project to convert the theoretical design into a working system.

Black Box Testing

Black box testing involves testing the functionality of the models without looking into the internal code or logic. Here are the test cases for black box testing:

Data Input Validation

Test Case 1: Verify the model handles valid input data correctly.

Input: A valid dataset with all required features.

Expected Output: Model processes the data without errors and proceeds to training.

Test Case 2: Verify the model handles missing values in the dataset.

Input: A dataset with some missing feature values.

Expected Output: The model either handles missing values gracefully or raises an appropriate error.

Model Training

Test Case 3: Verify the model training process completes successfully.

Input: Valid training dataset.

Expected Output: Model trains without errors and produces a trained model.

Test Case 4: Verify model training with imbalanced data.

Input: Imbalanced training dataset.

Expected Output: Model identifies the imbalance and handles it (e.g., through weighting).

Model Output Validation

Test Case 5: Verify the model's predictions on the test dataset.

Input: Valid test dataset.

Expected Output: Model produces predictions within an acceptable range of accuracy.

Test Case 6: Verify the model's ability to detect phishing sites.

Input: Test dataset containing known phishing and legitimate websites.

Expected Output: Model correctly identifies phishing sites with high accuracy

6.1.2 White Box Testing

White box testing involves testing the internal logic, code, and structure of the model. Here are the test cases for white box testing:

Code Logic Verification

Test Case 1: Verify the feature selection process.

Method: Check the code that selects features for correctness.

Expected Output: Code correctly identifies and selects relevant features.

Test Case 2: Verify the data preprocessing steps.

Method: Check the code for handling missing values, normalization, and encoding.

Expected Output: Data preprocessing is performed correctly without errors.

Model Architecture Validation

DOI: https://doi-ds.org/doilink/09.2024-62585588

Test Case 3: Verify the neural network architecture.

Method: Check the defined layers, activation functions, and connections.

Expected Output: Neural network architecture is correctly implemented as per the design.

Test Case 4: Verify the parameters of the XGBoost model.

Method: Check the parameter settings for XGBoost.

CONCLUSION

Phishing attacks are a rapidly expanding threat in the cyber world, costing internet users billions of dollars each year. It involves the use of a variety of social engineering tactics to obtain sensitive information from users. Hence, Phishing techniques can be detected using a variety of types of communication, including email, instant chats, pop-up messages, and web pages.

This project was able to categorize and recognize how phishers carry out phishing attacks and the different ways in which researchers have helped to solve phishing detection. Hence, the proposed system of this project worked with different feature selection and machine learning and deep neural networks such as Decision Tree, Support Vector Machine, XGBooster, Multilayer Perceptions, Auto Encoder Neural Network, and Random Forest to identify patterns in which URL links can be detected easily.

The Model with the highest accuracy based on the feature extraction algorithm used to identify phishing URL from legitimate URL links was integrated to a web application where users can input website URL links to detect if it is legitimate or phishing.

REFERENCES

1. AO Kaspersky lab. (2017). The Dangers of Phishing: Help employees avoid the lure of cybercrime. [Online] Available: https://go.kaspersky.com/Dangers-Phishing-Landing-Page- Soc.html [Oct 30, 2017].
2. " Financial threats in 2016: Every Second Phishing Attack Aims to Steal Your Money" Internet: https://www.kaspersky.com/about/press-releases/2017 financial-threats-in-2016. Feb 22, 2017 [Oct 30, 2017].
3. Y. Zhang, J. I. Hong, and L. F. Cranor," Cantina: A Content-based Approach to Detecting Phishing Web Sites," New York, NY, USA, 2007, pp. 639-648.
4. M. Blasi," Techniques for detecting zero-day phishing websites." M.A. thesis, Iowa State University, USA, 2009.
5. R. S. Rao and S. T. Ali," PhishShield: A Desktop Application to Detect Phishing Webpages through Heuristic Approach," Procedia Computer Science, vol. 54, no. Supplement C, pp. 147-156, 2015.
6. E. Jakobsson, and E. Myers, Phishing and Counter-Measures: Under- standing the Increasing Problem of Electronic Identity Theft. Wiley, 2006, pp.2–3.
7. L. A. T. Nguyen, B. L. To, H. K. Nguyen, and M. H. Nguyen, "Detecting phishing web sites: A heuristic URL-based approach," in 2013 Inter-national Conference on Advanced Technologies for Communications (ATC 2013), 2013, pp. 597-602.
8. Z. Zhang, Q. He, and B. Wang," A Novel Multi-Layer Heuristic Model for Anti-Phishing," New York, NY, USA, 2017, p. 21:1-21:6.
9. N. Sanglerdsinlapachai and A. Rungsawang," Web Phishing Detection Using Classifier Ensemble," New York, NY, USA, 2010, pp. 210-215.

DOI: https://doi-ds.org/doilink/09.2024-62585588

10. [10] G. Xiang, J. Hong, C. P. Rose, and L. Cranor," CANTINA+: A Feature- Rich Machine Learning Framework for Detecting Phishing Web Sites," ACM Trans. Inf. Syst. Secur., vol. 14, no. 2, pp. 21:1-21:28, Sep. 2011.

11. R. M. Mohammad, F. Thabtah, and L. McCluskey," Predicting phishing websites based on self-structuring neural network," Neural Comput & Applic, vol. 25, no. 2, pp. 443-458, Aug. 2014.

12. Pradeepthi K V and Kannan A," Performance study of classification techniques for phishing URL detection," in 2014 Sixth International Conference on Advanced Computing (ICoAC), 2014, pp. 135-139.

13. "PhishTank — Join the fight against phishing." [Online]. Available: https://www.phishtank.com/. [Accessed: 29-Nov-2017].

14. J. VanderPlas, Python data science handbook, 1st ed. 1005 Gravenstein Highway North, Sebastopol, CA 95472.: OReilly Media, Inc., 2016, pp. 331–515.

15. M. S. I. Mamun, M. A. Rathore, A. H. Lashkari, N. Stakhanova, and A. A. Ghorbani," Detecting Malicious URLs Using Lexical Analysis," in Network and System Security: 10th International Conference, NSS 2016, Taipei, Taiwan, September 28-30, 2016, Proceedings, J. Chen,V. Piuri, C. Su, and M. Yung, Eds. Cham: Springer International Publishing, 2016, pp. 467- 482.

DOI: https://doi-ds.org/doilink/09.2024-62585588