PHISHING ATTACKS DETECTION USING MACHINE LEARNING APPROACH

Abstract

Phishing is a type of network assault in which an attacker makes a false website in an effort to convince users to provide personal, financial, or password information to the phony website they think is their service. Build an app using machine learning algorithms to predict genuine or phishing URLs, identifying top threats targeting sensitive personal information such as usernames, passwords and credit card details. This method helps protect individuals and organizations from potential cyber-attacks.

Author

Madhura Eknath Sanap
Assistant Professor
SAE,kondhawa
Pune, Maharastra, India.

I. MOTIVATION

Phishing attacks are a growing problem in the digital world, causing data breaches, privacy issues and money fraud for internet users. Phishers create fake identities to defraud, steal legitimate information, and collect personal information such as passwords, account details, and transaction credentials.

II. GOALS AND OBJECTIVES

The goal of the paper is to learn techniques to avoid phishing using machine learning and improve the prediction rate of phishing. It analyzes phishing patterns, helps businesses avoid scams or hacks, and spreads awareness about phishing attacks. The main goal is to improve the success rate of anti-phishing methods and help business owners avoid phishing using various techniques. The main goal is to gain more trust from customers and handle data responsibly.

III. SCOPE OF PHISHING DETECTION

The online industry is constantly evolving, leading to the development of new phishing techniques. As a result, individuals, websites, and organizations are increasingly vulnerable to phishing risks. Anti-phishing software is essential for preventing these attacks. Cloud service providers have implemented various anti-phishing mechanisms to proactively identify attacks. To stay updated on the latest trends, supervised learning techniques are used to classify and detect cyber-attacks. Two popular machine learning algorithms are used in this research: Naïve Bayes and Naive Bayes.

- Principal Component Analysis
- Decision Tree
- Random Forest

IV. LITERATURE SURVEY

1. **Phishing attack detection using Machine Learning (Year-2020):** Decision tree and random forest analyze datasets for classification and detection, but difficulty in creating confusion matrix is a limitation.

2. **A Novel Approach to Detect Phishing Attacks using Binary Visualization and Machine Learning (Year-2022):** The experiment utilized neural network and binary visualization, but limitations include a 4,000-dot dataset, affecting model efficacy and predication.

3. **User Experiences of Torpedo: tooltip-powered phishing email detection (Year-2017):** Introduces TOROEDO to map out existing email phishing problems. The limitation was the use of IT literate to carry out research.

4. **Phi Boost- A novel phishing detection model Using Adaptive Boosting approach (Year-2021):** Blacklist and cloud-threat inspection approaches have limitations, including

phishing detection methods like decision tree and associative classification, but none have been demonstrated as applicable in this literature publication.

This research highlights challenges in machine learning techniques, such as overfitting, low accuracy, and ineffectiveness due to insufficient training data. It recommends internet users be aware of phishing to prevent cyberattacks and proposes an automated solution to phishing websites. While no single technology can completely eradicate phishing, a combination of good organization, proper technology application, and security improvements can significantly reduce its prevalence and losses.

- There may various type of phishing attacks face by many peoples in day to day life.

  - Deceptive phishing
  - Spear phishing
  - Whaling
  - Algorithm-Based phishing
  - URL phishing
  - Hosts File poisoning
  - Content-Injection technique
  - Clone phishing

- A variety of machine learning methods can be used to prevent these kinds of assaults.

Algorithms SVM (Support Vector Machine)

  - Import the dataset.
  - Explore the data to figure out what they look like.
  - Pre-process the data.
  - Split the data into attributes and labels.
  - Divide the data into training and testing sets.
  - Train the SVM algorithm.
  - Make some predictions.

V. CONCLUSION

User education aims to increase technical awareness to reduce phishing attacks. However, education alone may not guarantee positive behavioral responses. Users are kept safe online by technologies like machine learning, email filtering, and harmful URL detection. When users reply to emails from outside their corporate domains, especially in business settings, some providers even warn users.

REFERENCES

[1] "WC-PAD: Web Crawling based Phishing Attack Detection" Nathezhtha T., Sangeetha D., Vaidehi V.
[2] "Detection of Phishing Attacks with Machine Learning Techniques in Cognitive Security Architecture" Ivan Ortiz-Garces, Roberto O. Andrade, and Maria Cazares
[3] "A Methodical Overview on Phishing Detection along with an Organized Way to Construct an Anti-Phishing Framework" Srushti Patil, Sudhir Dhage.
[4] 4 "A survey of the QR code phishing: the current attacks and countermeasures" Kelvin S. C. Yong, Kang Leng Chiew and Choon Lin Tan.

[5] "Fuzzy Rough Set Feature Selection to Enhance Phishing Attack Detection" Mahdieh Zabihimayvan and Derek Doran.
[6] N. Goel, A. Sharma, and S. Goswami, "A way to secure a QR code: Sqr," in 2017 International Conference on Computing, Communication and Automation (ICCCA). IEEE, 2017, pp. 494–497.
[7] Mavroeidis and M. Nicho, "Quick Response Code Secure: A Cryptographically Secure Anti-Phishing Tool for QR Code Attacks," in International Conference on Mathematical Methods, Models, and Architectures for Computer Network Security. Springer, 2017, pp. 313– 324.
[8] K. Krombholz, P. Fruhwirt, T. Rieder, I. Kapsalis, J. Ullrich, and E. Weippl, "QR Code Security–How Secure and Usable Apps Can Protect Users Against Malicious QR Codes," in 2015 10th International Conference on Availability, Reliability and Security. IEEE, 2015, pp. 230–237
[9] Denso Wave, "QR Code development story," 2019, [Accessed: 28-Mar2019]. [Online]. Available: https://www.denso-wave.com/en/technology/ vol1.html .
[10] Dabrowski, K. Krombholz, J. Ullrich, and E. R. Weippl, "QR inception: Barcode-in-barcode attacks," in Proceedings of the 4th ACM Workshop on Security and Privacy in Smartphones Mobile Devices. ACM, 2014, pp. 3–10

