International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

AI Based Advanced Real-time Phishing Detection System Using Supervised Machine Learning Model

Dr. K. Sudha[1], S. Agalya[2], R. Akshaya[3], S. Keerthana Devi[4]

[1,2,3,4]B.Tech Computer Science and Business System, RMD Engineering College, Chennai, India

Abstract:
As digital connectivity continues to proliferate, the threat landscape is dominated by the pervasive and insidious nature of phishing attacks. Phishing, a form of social engineering, has become the predominant method for cyber attackers to exploit unsuspecting online users, compelling them to divulge sensitive information with the ultimate aim of financial or data theft. This paper addresses the pressing need for a robust defense mechanism against phishing attacks by introducing an innovative and proactive solution.Our mission is to safeguard users from malicious phishing attempts through an automated system that activates seamlessly when users interact with hyperlinks. Unlike traditional methods, this application transcends platform boundaries, offering comprehensive protection across diverse mobile applications such as WhatsApp, Facebook Messenger, and Instagram. The system's independence from user initiation ensures an added layer of defense, mitigating the risk of falling victim to phishing attacks.
Key features of the proposed system include automatic activation upon link interaction, cross-platform monitoring, and platform independence. By operating seamlessly in the background, our solution provides users with a consistent and proactive defense against phishing attempts, ensuring a secure digital environment for online interactions.

1. INTRODUCTION

Given the developing intricacy of organizations that has turned out to be progressively huge and far-reaching in
the proficient field as well as for individuals, we will be confronted with the challenge of adding to the quest for answers to safeguard against hackers and malware that are increasingly duplicated and numerous through the Web [1]. Antivirus and firewalls are security solutions but are limited even with the rapid development of hacking techniques, thus the need for an intrusion detection and prevention system (IDS/IPS) [2], similar to the Grunt [3]. This is one of the extremely durable monitoring methods to detect any violation of the security policy and to report the attack's information to network security administrators.
The purpose of this article is to work on the most traditional computer security, using "Machine Learning" techniques, to increase intrusion detection capabilities, compared to the new attacks. This new development will be a combination of both approaches, scenario, and behavior with a first proposed model. In a subsequent one, our proposed solution will be improved by the use of new automatic learning techniques. The field of "AI" essentially includes building a model from data using specific

algorithms. It allows, according to the given data in input, to predict unknown ones as well as to better understand existing ones, then to detect the new threats.

This paper is organized as follows: The main area will be dedicated to the introduction of the security problems, subject of our review, and the instruments developed to provide them solutions, particularly security intrusion detection tools on the networks. The subsequent area will be dedicated to an applied work wherein we present some research works in the field of security intrusion detection involving the contributions of recent trends in artificial intelligence, specifically AI [4]. The third area will be devoted to the introduction of our approach based on the use of AI techniques to improve intrusion detection in networks. The fourth area will present some simulation results of the implementation of our approach that clearly demonstrates its effectiveness compared with the traditional procedures. Finally, we present a conclusion and some prospects for our next work in the field of intrusion detection.

Phishing is a web-based deceitful act that uses social engineering and technical deception to mislead Web users and acquire their sensitive data or critical online information [5]. Social engineering techniques aim to obtain unsuspecting users' identity or sensitive private information using spoofed emails, fake websites, questionable online adverts/promotions, fake SMS from service providers or online companies, pin phishing, etc. The common targets in social engineering schemes include large corporations, financial institutions, payment companies, military, and government agencies who usually experienced significant financial and brand credibility damages [6]. For example, Details and Trends, 2017 security reports showed that almost about $5 billion were lost between October 2013 and December 2016 affecting more than 24,000 victims worldwide in a W-2 type of phishing attack. The W-2 phishing emails have been reported to be the most hazardous phishing email scams recently as its aim is to file false tax returns and claim refunds.

On the other hand, technical trick schemes typically involve the use of malicious software or crimeware which is usually installed on a computer or its associated devices without the knowledge of the victim [7]. Several techniques used in technical deception include DNS poisoning, keyloggers, session hijacking, host file poisoning, script injection, etc. In recent times, phishers have developed "ransomware" which executes a malicious code that adversely affects computing resources and demands a ransom payment to restore the resources to the original state. The incidence of these ransomware-based phishing emails as reported by CSO, indicated that 93% of phishing emails are now "ransomware". The report noted that most victims tend to pay quickly due to the critical nature of their resources.

Based on the phishing risk, various countermeasures called Anti-Phishing Systems (APS) were developed. However, phishers keep adopting evolving new sophisticated patterns to overcome current defense systems. Specifically, most existing APS have issues of the possibility of a zero-day attack, unnecessary computational overhead, high false positive and negative rates [8]. While some existing methods extracted one or more features to achieve promising results, others extracted a subset of existing feature corpus to achieve similar results [9]. Nevertheless, some APSs using AI (ML) and data mining techniques achieved a promising accuracy rate peaking at 99.62% [10], the selection and the performance of the feature vector on these algorithms limit the effective detection system.

Inspired by this reason, in this work, we pursue an effective based anti-phishing scheme to overcome the recent challenges facing existing APS by posing the following questions: a. Is it possible to achieve a more significant detection accuracy result by selecting or combining existing feature corpus? b. Might the extracted subset of the feature vector ever achieve similar results on more than one ML technique? c. Should new features be continuously proposed to combat phish despite the large available phishing

feature corpus ranging from visual to text-based?. These questions indicate the need to examine the possibility of using existing phishing feature corpus to build a robust anti-phishing scheme with significant efficiency. Our focus is to propose an efficient phishing "fingerprints" i.e., feature vector with significant detection accuracy across more than one ML algorithms. In particular, we choose Support Vector Machine and Naive Bayes for the evaluation of our feature vector because most existing APSs have used these ML algorithms more than others to benchmark their approach in most surviving literature available. Although other ML, such as KNN have been used in phishing issue, both SVM and NB have been found to be most suitable due to their binary classification nature and simplicity

2. State Of Art

A. Literature Review

Draw attention to the drawbacks of conventional approaches to managing complex and dynamic attack plans. Examine the difficulties presented by the unique idea of digital dangers and the development of interruption discovery frameworks over the long run. Underline how the review "Phishing Recognition Utilizing AI Method" has progressed the region. Explain phishing and its variants, focusing on the social engineering component. Clarify the need of executing crafty strategies for strengthen protections against cyberattacks. Identify the most significant methods, results, and research from the body of recent literature. Give an explanation of the broader effects of phishing detection using machine learning in cybersecurity. Give an outline of prior research in the interruption identification area utilizing imaginative philosophy. Sum up the advancing danger scene in network safety and the requirement for reliable interruption recognition frameworks.

It is necessary to provide a summary of the IEEE publication "Improved Security Intrusion Detection Using Intelligent Techniques." Point out any deficiencies or openings in the current information that the IEEE article plans to address. Point out the constraints of conventional rule-based strategies for distinguishing novel phishing procedures. Depict the elements, dataset, and AI procedures applied in the paper's technique. Think about the viability of the proposed model with other phishing recognition techniques that are as of now being used and have been examined in the writing. Compare the model's advantages and disadvantages to those of other approaches. Talk about any new elements or changes the creators apply to tried and true strategies.

Discuss the prevalence of phishing attacks and the potential negative effects they could have on individuals, organizations, and systems. Make sense of the fundamental job that interruption location plays in recognizing and stopping unlawful or threatening exercises. Numerous enemy of phishing frameworks are currently being created to distinguish phishing content in web correspondence organizations. Despite the widespread availability of anti-phishing tools, phishing still occurs. Summarize the main findings of the review of the literature. The piece of the prescient model that is utilized to create a proficient element vector is known as the Component Choice Module. The incremental component-based method is used to extract these attributes from the URL, webpage properties, and webpage activity. The resulting feature vector is then sent to the prediction model. The proposed framework's Help Vector Machine and Credulous Bayes calculations are prepared on a 15-layered include set. The exploration utilized datasets including 2541 phishing assaults and 2500 harmless examples. With 99.96% accuracy and 0.04% False Positive for both SVM, the experimental results demonstrate outstanding performance with 10-fold cross-validation.

3. METHODOLOGY

Phishing attacks often involve manipulating links to deceive users and make them believe that they are clicking on legitimate URLs. There are several techniques used to detect and prevent from people falling as victims to this manipulation. One method is to represent malicious URLs as hyperlinks with names on websites. This means that the displayed text for a link appears trustworthy, while the actual link goes to the phisher's site. For example, a phishing email may contain a link that appears to be from a legitimate organization, but when clicked, it takes the user to a fraudulent website designed to steal their information. Another technique is to use misspelled URLs or subdomains that resemble legitimate ones. These misspelled URLs can deceive users who may not notice the subtle differences and believe they are accessing a legitimate website. For instance, a phishing URL could be something like "www.yaahoo.com" instead of "www.yahoo.com".

URL shortening services, like Bitly, can also be used to hide the true destination of a link. Attackers can create shortened URLs that appear harmless, but actually lead to malicious websites. Victims have no way of knowing if the shortened URL points to a legitimate or malicious website. Phishing attacks often rely on tricking users into divulging sensitive information, downloading malware, or exposing themselves or their organizations to cybercrime. By manipulating links, attackers can make their phishing attempts appear more legitimate and increase the chances of success.

To protect against phishing attacks, it is important to be cautious when clicking on links, especially in emails or messages from unknown sources. Hovering over a link to check the destination URL can provide some indication of its legitimacy, but it is not foolproof as some phishers may be able to bypass this security measure. It is also recommended to use websites that can check URLs and inspect them for any signs of phishing.

SUPPORT VECTOR MACHINE: Support Vector Machine (SVM) is a machine learning algorithm used for regression tasks. It is known for its ability to produce significant accuracy with less computation power.To train an SVM model, a labeled data set is required, where each example is marked as belonging to one of the two categories . The SVM training algorithm then builds a model based on these examples, which can be used to classify new examples into one category or the other.

When using SVM, there are several parameters that can be tuned to improve model performance, such as the choice of kernel, gamma, and C values. A data set containing both phishing and legitimate URLs is collected. The data set should be diverse and representative of different types of URLs.Various features are extracted from the URLs to represent their characteristics. These features can include the number of hyphens, dots, numeric characters, presence of IP addresses, and different distance metrics like Liechtenstein distance and longest common sub sequence. The goal is to capture patterns and differences between phishing and legitimate URLs.The collected datasets is processed to ensure the data is in a suitable format for training the SVM model. This may involve removing duplicates, handling missing values, and normalizing or scaling the features.The trained SVM model is evaluated using the testing set. Performance metrics such as precision, accuracy, recall, and F1 score are calculated to assess the effectiveness of the model in detecting phishing URLs.The SVM model may be fine-tuned by adjusting hyper parameters, such as the kernel type, regularization parameter, and gamma value, to improve its performance. Techniques like cross-validation can be used to find the optimal hyper parameters. The system can process incoming URLs and classify them as phishing or legitimate based on the trained SVM model.

LOGISTIC REGRESSION: Logistic regression is a machine learning algorithm that is commonly used

for binary classification tasks, including the detection of phishing URLs in a URL phishing detection system. It works by fitting a logistic function to the training data and using input features to predict the probability of a URL being a phishing URL. While logistic regression is just one of the many algorithms that can be used for this task, it has been shown to achieve high accuracy rates in detecting phishing URLs.Phishing URLs are fraudulent URLs that are designed to deceive users and steal their sensitive information.

It involves training a model on a dataset of URLs labeled as phishing or legitimate. The model learns the patterns and features that distinguish between the two classes. Once trained, the model can then be used to predict whether a new URL is phishing or not.The features used in logistic regression for phishing URL detection can vary, but they often include characteristics such as the length of the URL, the presence of certain keywords or symbols, the domain age, and the presence of suspicious or malicious patterns .To train the logistic regression model, a dataset of labeled URLs is required. This dataset should contain a sufficient number of examples of both phishing and legitimate URLs to ensure the model can learn the relevant patterns. The dataset should also be diverse and representative of the types of URLs the model will encounter in real-world scenarios

Once the logistic regression model is trained, it can be evaluated on a separate test dataset to assess its performance. Common evaluation metrics include accuracy, precision, recall, and F1 score. These metrics provide insights into how well the model is able to correctly classify phishing and legitimate URLs.It is worth noting that logistic regression is just one approach among many for detecting phishing URLs. Other machine learning algorithms, as mentioned earlier, can also be used, and the choice of algorithm may depend on factors such as the specific problem domain, the available data, and the desired performance .

RANDOM FOREST MODEL: The Random Forest algorithm is a popular machine learning technique that falls under the category of supervised learning. The algorithm is based on the concept of ensemble learning, which involves combining multiple classifiers to solve complex problems and improve model performance.

As the name suggests, Random Forest is a classifier that consists of a collection of decision trees. Each decision tree is built on a different subset of the given dataset. The algorithm takes the predictions from each tree and, based on the majority votes of these predictions, predicts the final output. By averaging the predictions from multiple trees, Random Forest improves the predictive accuracy of the dataset.The greater the number of trees in the Random Forest, the higher its accuracy and ability to solve problems. This also helps prevent over fitting, which is when a model performs well on the training data but fails to generalize to new data.

Random Forest is a versatile algorithm that can handle various types of data, including binary, continuous, and categorical data. It is known for its efficiency and is widely used in different industries. One of the advantages of Random Forest is its ability to handle missing values. However, it also has some limitations.

In this we are training an ensemble of decision trees on a dataset of labeled URLs. Each decision tree is trained on a random subset of the data and features, and the final prediction is made by aggregating the predictions of all the individual trees.In the context of phishing URL detection, the features used in Random Forest can include various characteristics of the URL, such as its length, the presence of suspicious keywords or symbols, the domain age, and the presence of known phishing patterns or indicators.To train the Random Forest model, a dataset of labeled URLs is required. This dataset should

contain examples of both phishing and legitimate URLs. The dataset should be diverse and representative of the types of URLs the model will encounter in real-world scenarios.
Once the Random Forest model is trained, it can be used to predict whether a new URL is phishing or not. The model takes the features of the new URL as input and outputs a prediction based on the aggregated predictions of the individual decision trees.Evaluation of the Random Forest model is typically done on a separate test dataset. Common evaluation metrics include accuracy, precision, recall, and F1 score. These metrics provide insights into how well the model is able to correctly classify phishing and legitimate URLs.

4. RESULT AND DISCUSSION

When it comes to detecting phishing URLs, several machine learning models can be used. In our research paper we had analyzed the accuracy of every models and we have choosen Support Vector Machine (SVM), Random Forest, and Logistic Regression. These models are chosen for their effectiveness and accuracy in identifying phishing websites. Let's compare these models and understand why they are preferred over others. While there are other machine learning algorithms available for detecting phishing URLs, SVM, Random Forest, and Logistic Regression are often preferred due to their proven effectiveness and accuracy in this specific task. These models have been extensively studied and implemented in our research paper. They have shown good performance in distinguishing between legitimate and malicious URLs, and their ability to handle high-dimensional data and complex relationships makes them suitable for this task. It's important to note that the choice of the machine learning model depends on various factors, including the specific requirements of the task, the nature of the dataset, and the computational resources available.

Figure.1: Result of checking genuine link in the website "Not a Phishing Website!"

Figure.2: Result of checking phishing link in the website "Phishing Website!"

5. CONCLUSION

In this paper, we had a brief idea of how the various supervised machine learning models can be used to detect the phishing website. The proposed system has satisfied all the drawbacks analyzed in the existing systems such as detecting phishing websites not only confined to google chrome browser but all the other platforms. Also the accuracy rate achieved by the randomforest model is 98.9%, support vector machine(svm) is 94.6% and logistic regression is 97.4% approximately. Thus the system developed is highly efficient, user friendly and detects phishing websites/urls precisely.When it comes to detecting phishing URLs, several machine learning models can be used. In our research paper we had analyzed the accuracy of every models and we have chosen Support Vector Machine (SVM), Random Forest, and Logistic Regression. These models are chosen for their effectiveness and accuracy in identifying phishing websites. Let's compare these models and understand why they are preferred over others. While there are other machine learning algorithms available for detecting phishing URLs, SVM, Random Forest, and Logistic Regression are often preferred due to their proven effectiveness and accuracy in this specific task. These models have been extensively studied and implemented in our research paper. They have shown good performance in distinguishing between legitimate and malicious URLs, and their ability to handle high-dimensional data and complex relationships makes them suitable for this task. It's important to note that the choice of the machine learning model depends on various factors, including the specific requirements of the task, the nature of the dataset, and the computational resources available.

6. FUTURE WORK

Our future work would be further enhancing the system developed so far by analyzing and comparatively studying the other aiml models with highest precision and accuracy rate. Also by increasing the amount of dataset collected we aim to achieve a good detection speed(currently 1-2 secs) as well as accuracy of the result predicted by the system. We are currently planning to update adware detection and prevention system which will be very useful for the users to further prevent themselves from falling victims to malware attacks.

REFERENCES

1. (Urvashi Modi1 and Anurag Jain,"AN IMPROVED METHOD TO DETECT INTRUSION USING MACHINE LEARNING ALGORITHMS", Informatics Engineering an International Jorunal(IEIJ),Vol.4,No.2,DOI:10.5.121/ieij.2016.4203 17,June2016;
2. Junaid Rashid and Muhammad Wasif Nisar ,"Phishing Detection Using Machine Learning Technique" December 22,2022.04:37:53
3. H.Bleau,Global Fraud and Cybercrime Forecast,.2017..
4. Valero Leon,Andres,"INsIDES:A new Machine Learning-based Intrusion Detction System",Curs 2016-2017,Spain;
5. EI-Alfy, E.-S.M.,Detection of phishing websites based on probabilistic neutral networks and k-medoids clustering.The Computer Journal,2017. 60(12): p.1745-1759
6. Peng, T., I. Harris, and Y. Sawa. Detecting Phishing Attacks Using Natural Language Processing and Machine Learning. in 2018 IEEE 12th International Conference on Semantic Computing (ICSC). 2018.
7. Sahingoz, O.K., et al., Machine learning based phishing detection from URLs. Expert Systems with Applications, 2019. 117: p. 345-357.

8. Toolan, F. and J. Carthy. Feature selection for Spam and Phishing detection. in 2010 eCrime Researchers Summit. 2010.
9. Aburrous, M., Hossain, M. A., Thabatah, F. and Dahal, K. 2008. Intelligent Phishing Website Detection System using Fuzzy Techniques
10. A.A. Orunsolu , A.S. Sodiya , A.T. Akinwale,"A predictive model for phishing detection",13 December 2019.