SOCIAL ENGINEERING PHISHING DETECTION USING MACHINE LEARNING- A SURVEY

Devika C.J Nair[1], **Teslin Jacob**[2]

[1]*Computer Science and Engineering Department, Goa College of Engineering, Goa, India*
[2]*Computer Science and Engineering Department, Goa College of Engineering, Goa, India*

---***---

Abstract -Social engineering is the act of breaking security by manipulating the users into divulging confidential information. Social engineering uses psychological tricks to gain trust of the humans to achieve sensitive information for various purposes. Phishing is a method of computer based social engineering attack. Phishing is a criminal act of acquiring personal information by sending out forged emails with fake websites and fraudulent weblinks in web pages. The aim of this research is to develop a system that will detect phishing URLs from the webpage and classify the URL whether it is legitimate or illegitimate URL using machine learning algorithm.

Key Words:Phishing URL, Legitimate URL, Machine Learning, Logistic Regression, Prediction.

1.INTRODUCTION

Social engineering is a way of exploiting people into performing actions like getting into their confidential information. A social engineer uses human psychology to exploit people for his or her own use. The term social engineering applies to deception for the purpose of information gathering, fraud, identity theft or computer system access. Social engineering attacks are more challenging to manage since they depend on human behaviour and involve taking advantage of vulnerable employees.

Phishing is a social engineering attack that aims at exploiting the weakness found in the system at the user's end. A phishing attack is when a criminal sends an email or the URL pretending to be someone or something he's not, in order to get sensitive information from the victim. The victim in regard to his/her curiosity may enter the details like username, password or credit card number and they are likely to get cheated.

Phishing becomes a threat to many individuals, particularly those who are not aware of the threats in the internet. Commonly, users do not observe the URL of a website. Sometimes, phishing scams engaged through phishing websites can be easily identified by observing whether a URL belongs to a phishing or legitimate website.

The problems of phishing implies that computer-based solutions are needed for guarding against these attacks along with user education. Having such a solution might enable the computer to have the ability to identify malicious websites in order to prevent users from interacting with them. A general approach to recognize illegitimate phishing websites is through the Uniform Resource Locator (URL). Even there are cases where the contents of websites are duplicated but still using URLs can distinguish real sites from phished websites.

A common solution is to have a blacklist of malicious URLs developed by anti-virus groups. The drawback of this approach is that the blacklist cannot be exhaustive because new malicious URLs keep cropping up continuously. Therefore, approaches are needed that can automatically classify a new, previously unseen URL as either a phishing site or a legitimate one. For such type of solutions, machine learning based approaches are used where a system can categorize new phishing sites through a model developed using training sets of known attacks. One of the main problems with developing machine learning based approaches is that very few training data sets containing phishing URLs are available in the public domain. As a result, studies are needed that evaluate the effectiveness of machine learning approaches based on data sets that exists.

2. LITERATURE STUDY

The problem of detection of social engineering attacks depend on training potential victims to be resistant to manipulation. The authors Yuki Sawa, Ram Bhakta, Ian G Harris and Christopher Hadnagy [1] proposed an approach that analyzes the transcripts of several attack dialogs. Social engineering attacks involve either questions which request private information or commands which which request the listener to perform tasks which the speaker is not authorized to perform. The paper proposes approach using natural language processing techniques to detect questions and commands and extract their likely topics. Each extracted topic is compared with a topic blacklist to determine if the command or question is malicious or not. The authors have tested on three social engineering attacks and have been successful in detecting malicious sentences in these attacks. The use of NLP techniques to detect sentence type and extract topics is novel. The performance was good enough to provide attack warnings in real time during conversations.

Social media platforms like Facebook have become an indispensable part of every individuals life in today's world. A large number of reported facebook attacks have been observed in past few years. The authors [2] have done a study that proposes a novel model to detect and prevent the social engineering based phishing attacks on facebook. To validate the proposed model, four scenarios have been devised. Finite state machine was used to implement the proposed model since the proposed model is a conceptual model. The model works only for unidirectional attacks. The model detects that the user requested link is safe or deceptive. The proposed model also imposes immediate action if the suspected request is traced even if the attacker's achievement avoids the verification model. The finite state machine described in this paper is an abstract of the model and aims to provide a structured, flexible and deterministic high-level model of the steps taken to mitigate a social engineering attack. The results suggested that the proposed model also prefigure the threatening situation to users with different colors during the validation process. The results showed that in validation the model is feasible for detection and mitigation against phishing attacks on Facebook.

The authors WeinaNiu, Xiaosong Zhang, Guowu Yang [3] have proposed a detection method for phishing attacks using machine learning algorithm. The authors have proposed a model called Cuckoo Search-SVM which extracts 23 features which are used to construct a hybrid classifier. Cuckoo search is integrated with SVM to optimize parameter selection. The paper uses SVM algorithm as a base algorithm to generate hyperplane which minimizes training errors and maximizes the margin with correctly classified data points. Then with respect to the current hyperplane, the algorithm calculates classification error about normal emails and phishing emails. The value is modified according to the Cuckoo search algorithm until classification error remains unchanged or the maximal number of CS iterations are reached. Experimental results show that CS-SVM has a higher phishing email detection accuracy at different training set.

The paper titled "Social engineering detection using neural networks" [4] proposed method that aims to introduce a new technique to extract features that can be used for neural network testing and training. The authors have used benchmark data and developed a new technique to extract features that can be used for NN training and testing. Benchmark data was artificially created which relies on computer systems to analyze phone conversations in real time and determine if the caller is deceiving the receiver. 20 conversation scenarios with 9 social engineering attacks were chosen for this research. The feasibility of using neural networks to detect social engineering attacks in call centres by identifying certain attributes/features of the phone call or caller that can help the system identify whether this is an SE attack or not was investigated. The results indicated that there is a lot of potential for using this technology in SE.

The author Waleed Ali [5] in his research has suggested methods to cope with the problem of growing web phishing attacks. This paper presents a methodology for phishing detection based on machine learning classifiers with a wrapper features selection method. The author has proposed some common supervised machine learning techniques with significant features selected using the wrapper features selection approach to accurately detect phishing websites. In the wrapper based evaluation, a search algorithm is used to search through the space of possible features and evaluate each subset by running a mdel on the subset. Supervised machine learning algorithms such as back propagation neural network (BPNN), radial basis function network (RBFN), support vector machine (SVM), naïve bayes (NB), decision tree (C4.5), random forest (RF) and k-nearest neighbor (KNN) were implemented. The experimental results showed that BPNN, KNN and RF achieved the best CCR while RBFN and NB achieved the worst CCR for detecting the phishing websites. The machine learning classifiers based on wrapper based features selection accomplished the best performance while these classifiers with PCA features selection method achieved the worst performance in terms of CCR, TPR, TNR and GM.

In the paper "Phishing websites detection using Machine Learning" [6], the authors have carried out a method to develop a method of defense utilizing various approaches to categorize websites. The paper uses four classifiers: decision tress, Naïve Bayesian classifier, support vector machine (SVM) and neural network. The features are extracted from the dataset and run on the four classifiers. Using decision trees for classification, the problem of overfitting occurred which was not feasible for the project. Neural network did not perform well for the selected dataset because of less units in the hidden layer and feature values were discrete. SVM performed the best with respect to feature selection and dataset.

Another paper [7] for detection of phishing websites using machine learning proposed two methods for detection that is classification and association which will optimize the system. The proposed model focuses on identifying the phishing attack based on checking phishing website features, Blacklist and WHOIS database. Selected features can be used to differentiate between legitimate and spoofed web pages. Features of URLs and domain names are checked using several criteria. These features are inspected using a set of rules in order to distinguish URLs of phishing webpages from the URLs of legitimate websites.

In a survey paper [8], the authors in their survey study defines social engineering and explains how an attacker can read human mind to capture useful information. The author also provides recommendations on how to protect system against attackers using social engineering techniques. After conducting a survey, the authors have concluded that even after using the best and even the most expensive security technologies, an organization or a company or an individual is completely vulnerable. The authors also say that a key mechanism for combating social engineering must be their education of potential victim in order to raise their awareness of the techniques and how to spot them.

In the paper [9], the authors have proposed method for detecting suspicious URLs in real time system that is for twitter. The proposed system investigates correlations of URL redirect chains extracted from several tweets. Since attackers have limited resources and usually reuse them, their URL redirect chains frequently share the same URLs. The authors develop methods to discover correlated URL redirect chains using the frequently shared URLs and to determine their suspiciousness. The system consists of four components: data collection, feature extraction, training and classification. For classification supervised learning algorithm is used. The classification component executes classifier using input feature vectors, this component flags the corresponding URLs and their tweet information as suspicious. The suspicious will be delivered to security experts or more sophisticated dynamic analysis environments for an in-depth investigation. The proposed system easily fabricates syntactical features of spam messages and some simple modifications can also be applied to other services that can monitor a continuous URL stream.

In the paper “Phishing Website Detection using Machine Learning: A Review” [10], the authors performed a detailed literature survey and proposed a new approach to detect phishing website by features extraction and machine learning algorithm. Phishing is a way to obtain user’s private information via email or website. The study done by authors conclude that as there is lot of research work done, there is not any single technique which is enough to detect all types of phishing attack. As technology increases, phishing attackers uses new methods day by day which enables to find a effective classifier to detect phishing attack. After a detailed research the authors have concluded that tree-based classifiers in machine learning approach is best suitable than any other methods.

Another approach by authors [11] proposed three approaches for detecting phishing websites. The authors first analyzed various features of url and then by checking legitimacy of website by knowing where the website is being hosted and who are managing it. The third approach uses visual appearance based analysis for checking genuineness of website. The authors make use of machine learning techniques like Logistic Regression, Decision trees and Random Forest classifiers for evaluation of different features of URL and websites. The authors concluded that using different approaches altogether will enhance the accuracy of the system. The authors also pointed out a drawback of the system that is detecting of some minimal false positive and false negative results. Random Forest obtained a higher accuracy of 96.58% among the three classifiers used.

In the paper [12], authors have done a study on feature selection for the detection of phishing websites. This paper was aimed to identify the important features required for detection of phishing websites. There are different types of features that could be used for machine learning algorithms. The features that the authors have selected are URL-based features, Domain-based features, Page-based features and Content-based features. Features in URL-based are obtained when the URL is processed. In Domain-based, certain queries are provided that will classify the website as phishing or not. Page-based features check into the pages used that are calculated with respect to the calculation of reputation ranking of pages. Content-based features requires scan to target domain page. The target domain is checked whether it is used for phishing or not by processing the page contents. Therefore, the authors have done a study on these features that could be used for machine learning algorithms where lots of labelled data is required.

In this paper [13], the authors have proposed a phishing detection approach known as PhishZoo that uses profiles of trusted websites’ appearances for detection of phishing. This approach is a combination of having the ability of whitelisting approaches that will detect new or targeted phishing attacks also with the ability of blacklisting and heuristic approaches to warn users about bad sites. The authors have detected these sites by using content similarity between real sites and malicious sites. The method used here is whenever a site is loaded it is matched with the stored profiles. When the SSL and URL of the loaded site matches with the SSLs and URLs of any of the profiles then PhishZoo determines the site to be legitimate. Otherwise, the site’s contents will be matched against appearance profiles. Image matching step is used to reduce false positive rate. The authors have concluded that their method gives an accuracy of 96.10% against current phishing attacks and identifies new and targeted phishing sites. The method shows best accuracy against sites that look most like the real sites.

The proposed paper [14] focuses on improving the accuracy of phishing website detection. The authors have proposed a feature selection algorithm and integrated with ensemble learning methodology which is based on majority voting. The research demonstrates that current phishing detectors have accuracy between 70% and 92 %. The proposed model filters 30 features of initial dataset and the algorithm selects those that are critical in influencing the outcome of the prediction. The prediction model is trained through ensemble learning where multiple learning models are used. The authors have demonstrated that the results from all the models are used and counted to determine the majority of votes. The results state that by using feature selection and ensemble learning models are combined to obtain prediction. Therefore, by having multiple models the prediction is not bias towards one model and is based on majority of predictions such that all predictions from each model influences the final ensemble prediction.

The paper on “Detection and Prevention of Phishing websites using Machine Learning approach” [15], speaks about how the phishing problem is huge and having more than one solution to minimize all the problems effectively. The authors have come up with three approaches for phishing detection. First by analyzing various features of URL, second by checking legitimacy of website and third approach uses visual appearance based analysis for checking genuineness of website. The machine learning algorithms used in this paper are logistic regression, decision trees and random forest. The linear regression plot of expected output verses predicted output was observed for random forest algorithm. The plot had a slight deviation from expected output. The authors concluded by saying the efficiency can be achieved by using hybrid solution of heuristic patterns, visual features and blacklist and whitelist approach to feed them to machine learning algorithms. The new system can be designed in this way to avail more accuracy.

The authors Meenu and Sunil Godara [16] uses several machine learning methods for predicting phishing emails. Their study mainly compares the predictive accuracy f1 score, precision and recall of machine learning algorithms like logistic regression, support vector machine, decision tree and neural networks. The paper also shows the improvement in logistic regression by using additional methods to logistic regression like using feature selection methods. The authors concluded by finally having a comparison with the four classifiers. Among all, the improved logistic regression gave best results with respect to evaluation metrics.

The authors [17] have developed methods for detecting phishing websites with respect to analyzing features of benign and phishing urls. They have used machine learning techniques like Naïve bayes, decision trees, svm and kNN. The authors have developed methods based on lexical features, host properties and page importance properties. Fine-tuned parameters are used in selecting apt machine learning algorithm for separating phishing sites from benign sites. The first step of the proposed model is collecting phishing and benign URLs. The feature value database is formed when host based, popularity based and lexical based features are applied. The database is run using different machine learning methods and the classifiers are evaluated and later implemented in MATLAB. The results have shown that efficiency can be better achieved using lexical features.

The research paper [18] is based on automated real-time phishing detection using machine learning process. The phishing urls have connections between the part of url and by using it the features of phishing urls are extracted. The extracted feature helps real-time detection of phishing webstes using machine learning. Extreme learning machine is used for classification. ELM is a feed-forward artificial neural network and it has single hidden layer. The ANN renews its parameters as gradient-based, input weights are randomly selected while output weights are analytically obtained. To obtain cells in hidden layer of ELM, linear function as well as non-linear, non-derivable or discrete activation functions can be used. The authors have concluded that the feature extraction rules were defined of phishing extraction for obtaining features. They have also said that even users should be trained so as to not to follow the links to websites blindly where their personal information is requested.

The proposed model [19] predicts URL based phishing attacks based on features which gives maximum accuracy. The method uses features from URL. The method uses only those features that are used for phishing detection. The dataset is given to the system which is pre-processed so that unwanted data is removed and is in useable format for analysis. The characteristics of URL are extracted and valid ranges of inputs are identified. These values are assigned to each phishing website risk. After the data is trained, a machine learning algorithm is applied to the dataset. The model uses Naïve Bayes, Random Forest, Support Vector Machine and XGBoost for classification. From this, two classifiers are combined to obtain a hybrid classification using Naïve Bayes nad Random Forest to predict the accuracy of the detection of URL. This hybrid approach is used for testing the data and evaluating the prediction accuracy. The conclusion states that SVM is found to be the best classifier for predicting phishing URLs for having the highest accuracy among the classifiers. The proposed hybrid technique proved to be more secured than previous and existing phishing detection sites.

The paper based on phishing detection [20] makes use of Extreme Learning Machine (ELM) based classification is for 30 features from UCI Machine Learning Repository database. A model was proposed based on machine learning technique like Naïve Bayes to detect phishing web pages. The purpose of this project was to make a classification for the determination of a cyber threat known as phishing. In the dataset used, input attributes are determined in 30 and output attributes in 1. K-fold validation test was implemented where k=10 is for measuring the performance of generated system in the research. The average classification accuracy was measured at 95.05% and highest accuracy was to be measured as 95.93%.

3. CONCLUSIONS

On conducting a survey on various strategies used in phishing detection, the research can be concluded that even after having advanced technologies for detecting fraud URLs they are still not efficient enough. A detailed literature survey was performed in this paper about phishing detection systems. The papers [15] and [16] focusses on using Logistic Regression for detection of phishing URLs. Logistic regression showed to be giving better accuracy than other machine learning algorithms. Datasets used are PhishTank and UCI machine learning repository in most of the papers to obtain an efficient accuracy for classification. The other algorithms used are Support Vector Machine, Neural Network, Decision trees and Random Forest that classifies the URLs efficiently. As technology increases, phishing attackers have also started new methods for acquiring information. This enables us to find effective classifier for phishing detection.

REFERENCES

1. Yuki Sawa, Ram Bhakta, Ian G. Harris and Christopher Hadnagy, “Detection of Social Engineering Attacks through Natural Language Processing of Conversations”, 2016 IEEE Tenth International Conference on Semantic Computing.
2. Abid Jamil, Syed Mudassar Alam, Muhammad Kashif Nazir, Zikra Ghulam, “MPMPA: A Mitigation and Prevention Model for Social Engineering Based Phishing attacks on Facebook”, 978-1-5386-5035-6/18/$31.00 ©2018 IEEE, 2018 IEEE International Conference on Big Data (Big Data).
3. WeinaNiu, Xiaosong Zhang, Guowu Yang, Zhiyuan Ma, ZhongliuZhuo, “Phishing Emails Detection Using CS-SVM”, 0-7695-6329-5/17/31.00 ©2017 IEEE, 2017 IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC).
4. Hanan Sandouka , Dr. Andrea Cullen and Ian Mann, “Social Engineering Detection using Neural Networks”, 978-0-7695-3791-7/09 $26.00 © 2009 IEEE, 2009 International Conference on CyberWorlds.
5. Waleed Ali, “Phishing Website Detection based on Supervised Machine Learning with Wrapper Features Selection”, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 8, No. 9, 2017.
6. Arun Kulkarni, Leonard L. Brown, “Phishing Websites Detection using Machine Learning”, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 7, 2019
7. HemaliSampat, Manisha Saharkar, Ajay Pandey, Hezal Lopes, “Detection of Phishing Website Using Machine Learning”, International Research Journal of Engineering and Technology (IRJET) , e-ISSN: 2395-0056, Volume: 05 Issue: 03 | Mar-2018, p-ISSN: 2395-0072.
8. Anshul Kumar, Mansi Chaudhary and Nagresh Kumar, “Social Engineering Threats and Awareness: A Survey”, European Journal of Advances in Engineering and Technology, 2015, 2(11): 15-19

9.S.Umamaheswari, S.K Srivatsa,"Detection of Suspicious URLs Using Real Time System on Social Networks", June 2014 International Journal of Scientific and Engineering Research, Volume 5, Issue 6, ISSN 2229-5518.

10.Purvi Pujara, M.B. Chaudhari, "Phishing Website Detection using Machine Learning: A review", International Journel of Scientific Research in Computer Science Engineering and Information Technology ,2018, Volume 3, Issue 7, ISSN: 2456-3307

11.Vaibhav Patil, Pritesh Thakkar, Chirag Shah, Tushar Bhat, Prof. S. P. Godse, "Detection and Prevention of Phishing Websites using Machine Learning Approach", 978-1-5386-5257-2/18/$31.00 ©2018 IEEE.

12.Ebubekir Buber, Önder Demir, OzgurKoraySahingoz, "Feature Selections for the Machine Learning based Detection of Phishing Websites", 978-1-5386-1880-6/17/$31.00 ©2017 IEEE.

13. Sadia Afroz, Rachel Greenstadt, "PhishZoo: Detecting Phishing Websites By Looking at Them", Department of Computer Science Drexel University Philadelphia.

14.Alyssa Anne Ubing, SyukrinaKamilia Binti Jasmi, Azween Abdullah, NZ Jhanjhi, Mahadevan Supramaniam, "Phishing Website Detection: An Improved Accuracy through Feature Selection and Ensemble Learning", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 1, 2019.

15.Vaibhav Patil, Pritesh Thakkar, Chirag Shah, Tushar Bhat, Prof. S.P Godse, "Detection and Prevention of Phishing Websites using Machine Learning Approach", IEEE, 978-1-5386-5257-2/18, 2018.

16.Meenu, Sunil Godara, "Phishing detection using Machine Learning techniques", International Journal of Engineering and Advanced Technology, Volume 9, Issue 2, ISSN: 2249-8958, December, 2019.

17.Joby James, Sandhya L, Ciza Thomas, "Detection of Phishing URLs using Machine Learning Techniques", 2013 International Conference on Control Communication and Computing (ICCC), 978-1-4799-0575-1/13,2013 IEEE

18.Sneha Mande, Prof. D.S. Thosar, "Detection of Phishing Website based on Extreme Machine Learning", IJARIIE, Vol-4 Issue 6, 2018.

19.Sophia Shikalgar, Dr.S.D. Sawarkar, Swati Narwane, "Detection of URL based phishing attacks using machine learning", International Journal of Engineering Research and Technology (IJERT), Vol 8 Issue 11, November 2019.

20.Sandeep Kumar Satapathy, Shruti Mishra, Pradeep Kumar Mallick, Lavanya Badiginchala, Ravali Reddy Gudur, Siri Chandana Guttha, " Classification of features for detecting phishing websites based on machine learning techniques", International Journal of Innovative Technology and Exploring Engineering (IJITEE), Volume-8 Issue-8S2, June 2019.