Phishing Detection and Prevention using Chrome Extension

M.Amir Syafiq Rohmat Rose
Software Development Team
Mutiara iTech Sdn Bhd
43300 Seri Kembangan
Selangor, Malaysia
amir.syafiq989@gmail.com

Nurlida Basir
Fakulti Sains dan Teknologi
Universiti Sains Islam Malaysia
71800 Nilai
Negeri Sembilan, Malaysia
nurlida@usim.edu.my

Nur Fatin Nabila Rafie Heng
Fakulti Sains dan Teknologi
Universiti Sains Islam Malaysia
71800 Nilai
Negeri Sembilan, Malaysia
fatin@usim.edu.my

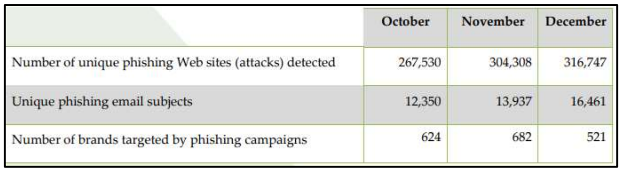
Abstract—During pandemic COVID-19 outbreaks, number of cyber-attacks including phishing activities have increased tremendously. Nowadays many technical solutions on phishing detection were developed, however these approaches were either unsuccessful or unable to identify phishing pages and detect malicious codes efficiently. One of the downside is due to poor detection accuracy and low adaptability to new phishing connections. Another reason behind the unsuccessful anti-phishing solutions is an arbitrary selected URL-based classification features which may produce false results to the detection. Therefore, in this work, an intelligent phishing detection and prevention model is designed. The proposed model employs a self-destruct detection algorithm in which, machine learning, especially supervised learning algorithms such as Support Vector Machine, Random Forest and Artificial Neural Networks were used. All employed rules in algorithm will focus on URL-based web characteristic. URLs are the main component in phishing on which attackers rely upon to redirect the victims to the simulated sites. A datasets from various sources such as Phish Tank and UCI Machine Learning repository were used and the testing was conducted in a controlled lab environment. As a result, a chrome extension phishing detection were developed based on the proposed model to help in preventing phishing attacks with an appropriate countermeasure and keep users aware of phishing while visiting illegitimate websites . It is believed that this smart phishing detection and prevention model able to prevent fraud and spam websites and lessen the cyber-crime and cyber-crisis that arise from year to year.

Keywords—phishing, machine learning, support vector machine, random forest, artificial neural networks

I. INTRODUCTION (*HEADING 1*)

In December 2021, Anti-Phishing Working Group (APWG) reported 316,747 website attacks (see Figure 1), which was the highest monthly total in APWG's reporting history [1]. Based on the observation, between 68,000 and 94,000 attacks increased per month compared to the data recorded in 2020 [2]. In addition, the number of phishing attacks in 2021 has tripled from phishing attacks recorded in early 2020 (see Figure 2 and Figure 3).

In the fourth quarter of 2021 [1], OpSec Security found that phishing attacks against the financial sector became the largest set of attacks, accounting for 23.2 percent of all phishing (see Figure 4). OpSesc observed increasing phishing volume up 20 percent of Q3 in Q4. OpSec also reported smishing and vishing campaigns activities have increased, typically utilizing fake customer support numbers or offering cheap or free products which targeting email and finance organizations. This shows that from day-to-day, phishers keep inventing a new strategy to trick users and thus, bypass the existing anti-phishing technical solutions. It is as well shows that phishers introduced new features or updated URL-



	October	November	December
Number of unique phishing Web sites (attacks) detected	267,530	304,308	316,747
Unique phishing email subjects	12,350	13,937	16,461
Number of brands targeted by phishing campaigns	624	682	521

Figure 1: Phishing Activity 4th Quarter 2021

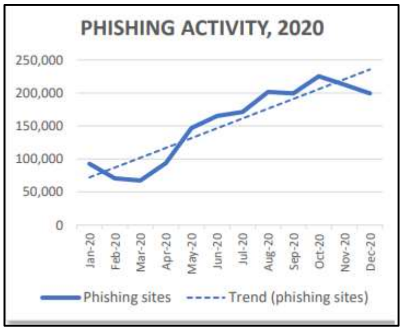


Figure 2: Phishing Attacks 2020

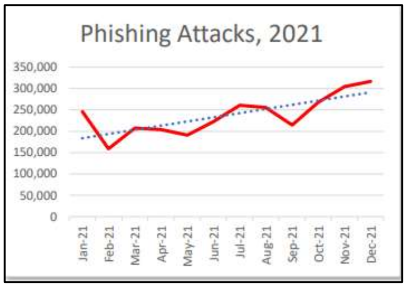


Figure 3: Phishing Attacks 2021

978-1-6654-9796-1/22/$31.00 ©2022 IEEE

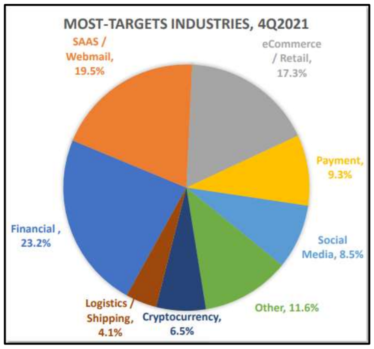


Figure 4: Phishing Most-Targets Industries

based features to continually perform phishing. In many other cases, classification URL-based features are arbitrary selected, thus produce false phishing results to the system [3, 4, 5, 6].

Phishing AttackSpam, cyber-terrorism, theft and phishing are few instances of cyber-attacks. Based on 2020 Annual ReportState of The Phish by proofpoint [7], 88% of organizations worldwide experienced spear phishing attempts in 2019. In addition, based on Verizon in 2020 Data Breach Investigation Report [11], 45% of breaches featured hacking, 17% involved malware and 22% involved phishing were recorded. Phishing attacks are currently targeting not only system end users, but also technical staff at service providers, and may employ advanced techniques such as MitB attacks [11]. One of the advantages that has been utilized by phisher is when personal users cannot differentiate a phishing page from the legitimate ones due to their limited knowledge in phishing. Based on State of The Phish Annual Report 2020 [7], the knowledge on phishing and ransomware among young generations who are the most actively used technologies nowadays are critically low. Therefore there is a need to embed the algorithm into the browser to detect and prevent phishing attacks beforehand. Based on Louis Columbus in his article [8], the one point of failure anti-phishing applications continue to have is lack of adoption. The best solution to closing the gap is by enabling on-device machine learning protection. Therefore, an embedded adaptive machine learning algorithm into the existing web browser is the best way to combat phishing in real environment.

A detailed comparative analysis revealed that machine learning methods are the most frequently used and effective methods to detect a phishing attack [4, 5, 6]. In this work, our main focus will be on URL-based features with a new intelligent algorithm based on a machine learning. This adaptive algorithm will considered an updated classification feature(s) of websites in order to detect and prevent the phishing attacks. The algorithm enables a system to acquire new data patterns as an ideal solution to the phishing detection problem. Though many researches have tried to detect phishing attacks with machine learning in recent years, it shows that most of the researches was ineffective in reducing the false positive rates, even there is a quite high rate of detection. For example, Subasi et. al. 2017 [9] developed an intelligent phishing website detection using random forest classifier. Based on the performance evaluation conducted, 97.36% accuracy was recorded based on the use of random forest classifier. Zamir et.al. 2020 [10] presented a framework to detect phishing websites using stacking model which contains diverse machine learning algorithms in improving the accuracy of all the classifiers. Based on the combination of Random Forest, Neural Networks and Bagging methods, the accuracy is recorded as 97.4%. Based on existing researchers, it shows the accuracy rate still in the range of 97%. It as well showed that an efficient detection system ought to have the option to identify phishing attacks with low false positives. Even though there is a quite high rate of phishing detection, the evolution of phishing attacks and strategies is still a major concern. Thus, many researches [] have shown the significance of integrating the appropriate set of methods in order to obtain successful detection output. Therefore, we believe our method can improve the accuracy rate due to the evolution of phishing attacks in our adaptive algorithm.

II. URL Features Extraction

Features extracted from a URL are the basis to determine if the malicious URL [3, 14, 15, 16]. Here, 16 features have been selected out of 30 available features for test data. Even only 16 been selected it does not much loss in the accuracy on the test data. More number of features increases the accuracy but on the other hand, it also increases the feature extraction time. Thus a subset of features is chosen in a way that the trade-off is balanced. Table 1 shows the list of URL features been selected.

Table 1: URL Features Extraction

IP address	Degree of subdomain	Empty server form handler
URL length	Favicon domain	Use of mailto
URL shortener	HTTPS in domain name	Use of iFrame
'@' in URL	Image cross domain	Status Bar
Redirection with '//'	Anchor tag href domains	
'_' in domain	Script & link tag domains	

The details used to identify the phishing portals based on the selected features are as follows:

- isIPInURL(): Identify IP address presence in the URL
- isLongURL(): Check if the URL is more than 75 characters
- isTinyURL(): Check if the URL is less than 20 characters
- isAlphaNumericURL(): Search in URL for alphanumeric '@'
- isRedirectingURL(): Check if "//" exists more than once within the URL
- isHypenURL(): Search in URL for "-" adjacent to the domain name
- isMultiDomainURL(): The domain name need to be restricted to the top level domain, the country code and the second-level domain

- isFaviconDomainUnidentical(): Check if the links on the web page are from another domain
- isIllegalHttpsURL(): Identify multiple 'https' in a URL string
- isImgFromDifferentDomain(): Verify whether photos are loaded from other domains on the specified web page
- isAnchorFromDifferentDomain(): Check if the links on the web page are from another domain
- isScLnkFromDifferentDomain(): Verify if scripts are on the web-page are from another domain
- isFormActionInvalid(): Search for submissions of invalid / blank form
- isMailToAvailable(): Search for anchor tag that incorporate mailto
- isStatusBarTampered(): Check if onmouseover controls the status bar display
- isIframePresent(): Identify pages that display iframes in the DOM

All 16 features needs to be extracted and encoded for each webpage in real-time while the page is being loaded. A content script is used so that it can access the DOM of the webpage. The 'content.js' script is automatically injected into each page while it loads. The 'content.js' script is responsible to collect the features and then send them to the plugin. The main objective of this work is not to use any external web service and the features needs to be independent of network latency and thus increase the extraction time. Once a feature is extracted it is encoded into values {-1, 0, 1} based on the following notation.

- -1 : Legitimate
- 0 : Suspicious
- 1 : Phishing

The feature vector containing 16 encoded values is passed on to the plugin from the 'content.js' script. The evaluated feature vector, further, passed to predict (data) function reckons the prediction for the website. The output of the binary classification is based on the prediction values and if value return to 1, the user is warned if the webpage is classified as phishing.

III. Machine Learning Algorithm

Based on the identified URL-based features, an intelligent phishing detection and prevention model is designed. This model employs a self-destruct detection algorithm in which, machine learning algorithm been implemented [17]. Based on the performance on classification problems, three supervised machine learning algorithms namely Random Forests, Artificial Neural Networks and Support Vector Machines were used in the model.

- Random Forests

Random forest is a supervised learning algorithm. The random forests are combining many tree predictors, each tree being independently sampled by the values of a random vector. The "forest" it builds, is an ensemble of multiple decision trees and merges them together to get a more accurate and stable prediction. In order to create a tree, assume that n represents the number of training observations and p represents the number of variables in a training set. We choose k « p as the set of variables to be chosen to decide the decision node at a tree. In the training set, we select a bootstrap sample from the n observations and use the rest of the observations to estimate the tree's error during the test phase. Thus, at a given node in the tree, we randomly pick k variables as a decision and determine the best split dependent on the k variables in the training set.

- Artificial Neural Networks

A neural network is formed by a number of identical units (neuron) connected to one another. Signals are transmitted from one neuron to another through interconnections. The connections also have weights to improve neuronal supply. The neurons are not powerful on their own, but can perform complex computations when connected to each other.

- Support Vector Machines

Support Vector Machine (SVM) is a supervised machine learning discriminative model, aligned with the principle of drawing separating hyperplane with maximal protection space, known as margin, to reduce the probability of incorrect predictions. The purpose of the support vector machine algorithm is to locate a hyperplane in an n-dimensional (the number of features) space that classifies data points distinctly. Here, the model accumulates initial responses to a specific phishing attacks and the appropriate countermeasures and responses will be provided. All employed rules in algorithm will focus on URL-based web characteristic to detect phishing. In general, the essence of these methods of machine learning detection is to map all the features of the phishing website and then use the machine learning to detect the phishing sites.

The embedded algorithm predicts the validity of the websites visited by the users and provides warnings of the browsed illegitimate URL on the loading page. The machine learning model been embedded into Chrome browser as an easy-to-use application extension to help end users to deal with phishing attempts. Here, the JavaScript is used to implement the machine learning algorithm and build the Chrome extension. The extension solution integrates implementation of Python-based with JavaScript-based test module. Python was used to improve lagging on performance due to the complex numerical computation libraries with the weight on the web content and features extraction activities. The extension has a minimal web dependence because it gathers several files as a one-time work process into a single file that the user can download. The extension implementation uses a SVM-trained persistent model to identify maliciously sites. Further, the extracted features passed through the SVM model identifies hostile webURLs. The adaptive model process flows is shown in Figure 5.

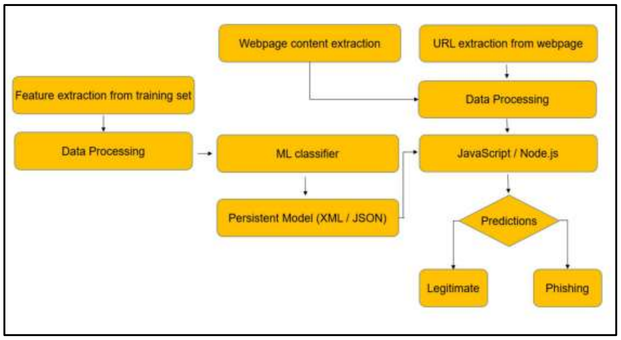

Webpage content extraction
URL extraction from webpage
Feature extraction from training set
Data Processing
Data Processing
ML classifier
JavaScript / Node.js
Persistent Model (XML / JSON)
Predictions
Legitimate
Phishing

Figure 5: Phishing Detection Chrome Extension Implementation

A. *Dataset*

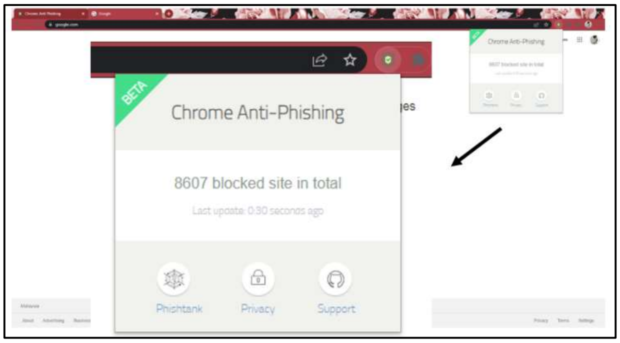
In our work, the datasets were split into phishing and non-phishing dataset. The phishing dataset was obtained from PhishTank [20] while the non-phishing dataset was gathered manually. Here, we also used the 'Phishing Website Dataset' from UCI Machine Learning repository [19] to evaluate our machine learning algorithm. The program includes 11,055 URLs with 6157 instances of phishing and 4898 legitimate instances. There are 30 features in each instance. The features have three distinct values. '1' if the rule is met, '0' if the rule is partial, '-1' if the rule is not met.

B. *Model and Algorithm Evaluation*

The experimental results are obtained by analyzing legitimate and phishing websites from different datasets. The features are evaluated and algorithms are formulated to consider the identified classification features. The algorithms are evaluated based on the Effectiveness Metric (EM) values especially in term of the performance. The accuracy, sensitivity and specificity are analyzed. The experiment is conducted in a controlled lab environment. A browser extension (e.g. Chrome) is embedded with the proposed algorithm. This browser extension is used to test the algorithm in detecting phishing websites.

IV. CHROME EXTENSION

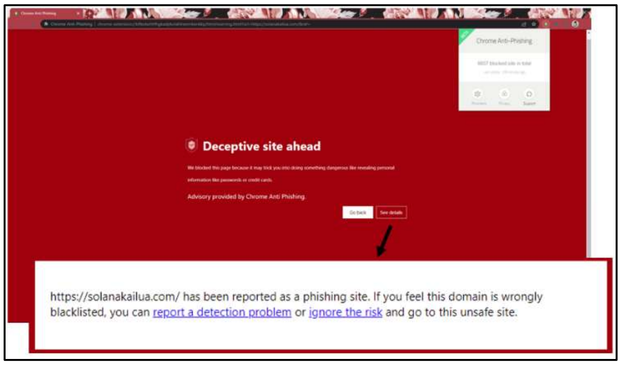
A simple and easy to use User Interface has been designed for the plugin using HTML and CSS [18]. The UI contains the application title, total blocked site, last update runtime and three functional circle buttons. Here, a Chrome extension allows users to easily deploy the learned model on the Chrome Web Store, from where users can download and utilize the extension as phishing detection tool [13]. In case of phishing, the website will be blocked by the Chrome Anti-Phishing warning page and give users an option to Go Back or See Details. If the user chooses to click on "See Details", a report related to the phishing information were given based on the data in the phishing databases. If the users want to continue and ignore the risk of the unsafe site, the users is allowed to do so. The plugin also displays an alert warning based on the machine learning algorithm analysis. The warning page and alert message were set to prevent the user from entering any sensitive information on the website. The UI of the Chrome extension is shown in Figure 6 and the test output of both approaches are shown in Figure 7 and 8.



8607 blocked site in total

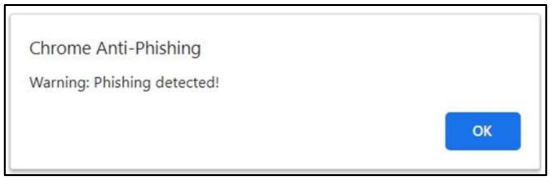
Last update: 0.30 seconds ago

Figure 6: Chrome Extension



https://solanakailua.com/ has been reported as a phishing site. If you feel this domain is wrongly blacklisted, you can report a detection problem or ignore the risk and go to this unsafe site.

Figure 7: Phishing Warning Page



Chrome Anti-Phishing

Warning: Phishing detected!

Figure 8: Alert Warning Message

V. CONCLUSION

This paper describes a Chrome Anti-Phishing extension plugin that has been developed based on a machine learning algorithm and phishing datasets to detect and prevent users from any phishing webpages. The extension provides a warning message and prevent webpage from being displayed. This feature gives users an option to see the blacklist information from the phishing databases as guidelines. The plugin also displays a warning alert based on the machine learning algorithm analysis. The alert message were set to prevent the user from entering any sensitive information of the website. It is believed this plugin able to detect and prevent users from phishing webpages and thus reduce number of phishing activities.

ACKNOWLEDGMENT *(Heading 5)*

This work has been supported by Intelligent Spam Detection Model Based on Immunological Memory of Adaptive Immune System (USIM/FRGS/FST/50116-50).

REFERENCES

[1] [1] Anti-Phishing Working Group (APWG) (2021) “Phishing Activity Trends Report, 4th Quarter 2021” Date accessed: 26th April 2022. Available at:https://docs.apwg.org/reports/apwg_trends_report_q4_2021.pdf?_ga=2.200306155.416943267.1650941114-325193504.1649309719&_gl=1*1b6hsh*_ga*MzI1MTkzNTA0LjE2NDkzMDk3MTk.*_ga_55RF0RHXSR*MTY1MDk0MTExMy4zLjEuMTY1MDk0MTE1OS4w

[2] [2] Anti-Phishing Working Group (APWG) (2020) "Phishing Activity Trends Report 4th Quarter 2020". Date accessed: 26th April 2022. Available at:https://docs.apwg.org/reports/apwg_trends_report_q4_2020.pdf?_ga=2.29581946.416943267.1650941114-325193504.1649309719&_gl=1*4ez4an*_ga*MzI1MTkzNTA0LjE2NDkzMDk3MTk.*_ga_55RF0RHXSR*MTY1MDk0MTExMy4zLjEuMTY1MDk0MTQwOS4w

[3] [4] F. Toolan and J. Carthy (2010) "Feature selection for Spam and Phishing detection," 2010 eCrime Researchers Summit, Dallas, TX, 2010, pp. 1-12, doi: 10.1109/ecrime.2010.5706696.

[4] [5] Basit, A., Zafar, M., Liu, X. et al. (2021) “A comprehensive survey of AI-enabled phishing attacks detection techniques.” Telecommun Syst 76, pp. 139–154. https://doi.org/10.1007/s11235-020-00733-2

[5] [6] Varshney, Gaurav & Misra, Manoj & Atrey, Pradeep. (2016). A survey and classification of web phishing detection schemes. Security and Communication Networks. 9. 10.1002/sec.1674.

[6] [7] Yi, Ping & Guan, Yuxiang & Zou, Futai & Yao, Yao & Wang, Wei & Zhu, Ting. (2018). Web Phishing DetectionUsing a Deep Learning Framework. Wireless Communications and Mobile Computing. 2018. 1-9. 10.1155/2018/4678746.

[7] [8] Proofpoint, Inc. (2020) “Annual Report State of The Phish 2020”. Date accessed: 8th February 2021. Available at: https://www.proofpoint.com/sites/default/files/gtd-pfpt-uk-tr-state-of-the-phish-2020-a4_final.pdf

[8] [9] Louis Columbus [2020] “5 Ways Machine Learning Can Thwart Phishing Attacks” Former Contributor Enterprise Tech. Date accessed: 16th march 2021. Available at: https://www.forbes.com/sites/louiscolumbus/2020/08/12/5-ways-machine-learning-can-thwart-phishing-attacks/?sh=334e51601035

[9] [10] Subasi, E. Molah, F. Almkallawi and T. J. Chaudhery (2017) "Intelligent phishing website detection using random forest classifier," 2017 International Conference on Electrical and Computing Technologies and Applications(ICECTA), Ras Al Khaimah, 2017, pp. 1-5, doi: 10.1109/ICECTA.2017.8252051.

[10] [11] Zamir, A., Khan, H.U., Iqbal, T., Yousaf, N., Aslam, F., Anjum, A. and Hamdani, M. (2020), "Phishing web sitedetection using diverse machine learning algorithms", The Electronic Library, Vol. 38 No. 1, pp. 65-80. https://doi.org/10.1108/EL-05-2019-0118

[11] [12] Verizon (2020) “2020 Data Breach Investigations Report, Executive Summary” Date accessed: 8th February2021. Available at: https://enterprise.verizon.com/resources/executivebriefs/2020-dbir-executive-brief.pdf

[12] [13] M. Khonji, Y. Iraqi and A. Jones (2013) "Phishing Detection: A Literature Survey," in IEEE CommunicationsSurveys & Tutorials, vol. 15, no. 4, pp. 2091-2121, Fourth Quarter 2013, doi: 10.1109/SURV.2013.032213.00009.

[13] [18] Netcraft Ltd (2021) ”Netcraft browser extension for Chrome, Firefox, Edge and Opera”. Date accessed: 8thFebruary 2021. Available at:https://toolbar.netcraft.com

[14] [19] Al-Daeef, Melad Mohamed; Basir, Nurlida; Saudi, Madihah Mohd (2017) “Anti-phishing immune system” Advanced Science Letters, Volume 23, Number 5, May 2017, pp. 4745-4749(5). American Scientific Publishers. https://doi.org/10.1166/asl.2017.8882

[15] [26] Parekh, S., Parikh, D., Kotak, S., & Sankhe, S. (2018). “A New Method for Detection Of Phishing Websites:URL Detection. In 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT) (pp. 949–952). IEEE.

[16] [30] Adebowale, M. A., Lwin, K. T., Sanchez, E., & Hossain, M. A. (2019). “Intelligent Web-Phishing Detection and Protection Scheme Using Integrated Features of Images, Frames And Text.” Expert Systems with Applications,vol. 115, pp. 300–313.

[17] [31] Zamir, A., Khan, H.U., Iqbal, T., Yousaf, N., Aslam, F., Anjum, A. and Hamdani, M. (2020), "Phishing web site detection using diverse machine learning algorithms", The Electronic Library, Vol. 38 No. 1, pp. 65-80. https://doi.org/10.1108/EL-05-2019-0118

[18] [34] Developers G. (2014) “Safe browsing API-Developer Guide V3” Date accessed: 8th February 2021. Available at: https://developers.google.com/safe-browsing/developers_guide_v3

[19] [37] UCI (2012) “UCI Machine Learning Repository: Phishing Websites Data Set,” Date accessed: 8th February2021. Available at:https://archive.ics.uci.edu/ml/datasets/phishing%20websites

[20] [38] PhishTank (2006). “PhishTank” Date accessed: 8th February 2021. Available at:https://www.phishtank.com/s