**A MODIFIED ALGORITHM FOR INCREASING MACHINE LEARNING PERFORMANCE FOR DETECTING AND CLASSIFYING PHISHING ATTACKS**

**Yertayev Ansar, MIS-812**

**Izmir Bakircay University**

**Annotation**

Phishing websites refer to an attack where cyber criminals spoof official websites to lure people to access to illegally obtain user's identity, password, privacy and even properties. This attack poses a great threat to inexperienced Internet users and is becoming more and more sophisticated. Many proposals for detecting phishing websites have shown to be effective and advantageous methods for detection and classification, unified resource locator (URL). Although several approaches have been proposed to detect and classify phishing attacks, URL-based machine learning and artificial intelligence approaches provide better performance results, but they all depend on the feature set used. To improve the accuracy of phishing website detection, this paper proposes a novel decision tree-based model and feature set for Internet of Things (IoT) devices that have limited capabilities and low power consumption. This study examines how feature set selection from a training dataset, which significantly improves the speed and performance of phishing attack classification in IoT devices. According to the experimental and comparative results of the implemented classification algorithms, the piecewise linear decision tree algorithm based on novel activation features provides the best performance with an accuracy of 97.50% for detecting phishing URLs.

**Keywords:** decision tree, classifier, phishing attack, URL resource, training set, machine learning

**Introduction**

With the advent of the Internet, it is clear that a new technological advancement in cybercrime is taking place. As part of this advancement, many areas of business and information technology have moved from traditional services to online forms. Along with information technology and taking advantage of the ubiquity of online transactions, many offenses have also moved to online forms, i.e. cybercrime. Today, one of the most common variants of cyberattacks is phishing attacks. In 2019, global online banking transaction fraud amounted to 1920 million dollars, of which phishing attacks accounted for 318 million dollars [1], making it one of the most effective and widespread frauds on the Internet [2].

Normal web browser users are asked to enter their personal information during a phishing attack, usually through a Uniform Resource Locator (URL). Typically, the URL in a web browser used in a phishing attack is masked by using long sequences of alphanumeric characters and/or entering characters similar to the original URL (e.g., www.mibank.com instead of www.mybank.com). If the malicious URL is delivered to devices with small screens (e.g., cell phones, tablets, gadgets, etc.), the phishing attack becomes even more effective and efficient for cyber fraudsters [8]. In web browsers, the address bar for entering a website address is usually reduced or sometimes hidden from the Internet user. Such devices make up a set of devices called the Internet of Things (IoT) [3]. Many IoT devices are used to exchange messages, documents, listen to online radio, watch movies, shop online, communicate with friends/colleagues, etc. Considering these facts that the targets of cyber attacks are considered to shift to IoT devices and their users are becoming more and more [11]. In addition, phishing cyberattack, which is expected to grow faster than any other, which is very attractive to cybercriminals due to the physical features and low security level of devices such as IoT.

Given the characteristics of phishing attacks, the research focuses on learning the features in training datasets to improve the performance of machine learning and artificial intelligence algorithms for phishing detection and classification. However, in many works on phishing attack detection and classification, most of them have focused on determining which classifier performs better given pre-defined features derived from third-party services and sources of training datasets found in publicly available repositories [4]. These works also utilize complex data structures and data representations combined with computationally intensive processes, making them unsuitable for use in IoT devices. In addition, some works acquire the characteristics of visiting a suspicious web page, implying that they have been victimized by an attack. IoT devices are characterized by limited computational capabilities and low power consumption, which makes such classification methods and algorithms unsuitable for use in these systems [5].

In such cases, such classification algorithms running in IoT devices should be lightweight, energy efficient and it is recommended to avoid the use of complex data structures, and the training dataset sources used and their features should be as simple as possible. Considering the above requirements, this paper proposes a decision tree based method which is proposed in [6] for detecting phishing URLs in IoT environments, which maximizes the detection rate and classification accuracy of phishing attacks. The choice of feature set is crucial to propose a phishing detection approach applicable in practice. In addition, the proposed method can detect real-time and zero-day attacks and is independent of third-party services.

The main contributions of this paper are:

1. To select features from a training dataset for detection and classification of phishing URLs, exactly suitable for IoT systems;

2. To serve as a starting point for researchers and practitioners in developing solutions to phishing attack classification problems for systems with limited properties like IoT.

**2.Basic concepts and techniques for phishing attacks**

2.1. Phishing Attacks Phishing attacks can be performed via URLs from users' web browser. Generally, URL-based phishing attacks are mainly performed by embedding special words and/or characters into URLs:

(a) generate similar words but with not significant errors;

(b) contain a set of special characters/letters to redirect the web page;

(c) apply shortened and/or unnecessarily too long URLs that are not suitable for understanding;

(e) use attractive keywords that appear to be correct;

(f) in most cases add a malicious file to the link, which after being automatically downloaded goes to the victim-user's IoT device.

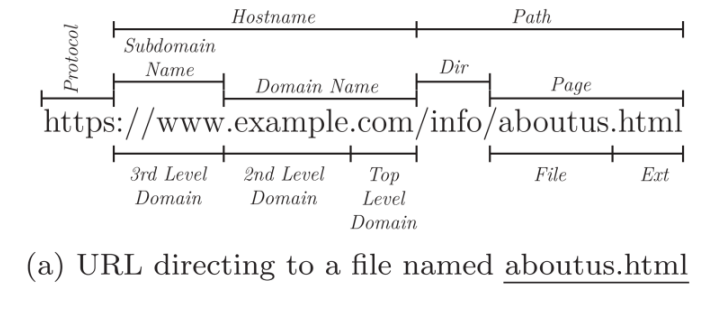
One approach to detecting and classifying phishing URLs is based on blacklists that rely on a repository of already classified websites (https://phishtank.com). This approach is high-speed and efficient, but has some drawbacks. For example, a URL that does not exist in the training dataset will not be correctly classified, especially URLs of zero-day attacks. In such approaches found in blacklist based methods that traditional machine learning algorithms are quite good at solving problems with phishing URL detection problems [7-11]. These requirements are particularly suitable for application to IoT devices due to their relatively low computational power and resources.

2.2. URL This research is related to phishing attacks, so this paper analyzes the data related to an address, which is called a Uniform Resource Locator (URL), which can be found in the RFC1738 standard. The general view of URLs is shown in Figure 1.

2.3. Methods and Algorithms for Classifying Phishing URLs A common technique used for phishing attacks is to create a very large number of URLs with all sorts of different variants. This provides a distraction for regular Internet users, which means that the likelihood of a successful phishing attack increases by several times the order of magnitude. In the URL string, introducing multiple slashes that point to multiple directories in the URL and look correct to inexperienced users of IoT devices. Similarly, introducing a few dots and some characters/digits in the domain name to create multiple sub-domains gives the impression of a correct URL.

Such generated malicious URLs very often replace alphanumeric characters with other characters, i.e. Unicode characters and/or hexadecimal character representation. English text has a relatively low entropy, i.e. it is predictable, and entropy changes more when different characters are introduced. Hence, the use of entropy can lead to the detection and correct classification of malicious URLs.

Considering such aspects of phishing attacks, this paper focuses on proposing a lightweight URL representation and a better performance algorithm for detecting and classifying phishing URLs in IoT systems. Cybercriminals during a phishing attack very often deliver a malicious URL using a common application (email, Telegram, Tweeter, Facebook, etc.). If an inexperienced Internet user accesses the phishing URL, the malicious activity will work in favor of the cybercriminal.



Изображение выглядит как текст, чек, линия, Шрифт

Автоматически созданное описание

*Figure 1. General view of the unified resource index form and its parts*

**3. Feature selection for phishing attack detection**

In [12-17], the authors investigated some approaches for detecting phishing URLs. These approaches utilize several features extracted from URLs in practice. In this study, the proposed set of attributes was built by considering the nature of phishing attacks and projecting them onto URLs, e.g., phishers try to confuse ordinary users of IoT devices by making URLs unreadable and unintelligible. As a result, the generated phishing URLs become larger in number and use different characters/numbers than the correct addresses.

Based on this fact, in this study it is recommended to use such types of features measuring the length of some parts of the URL, the number of characters/digits and features related to HTTP/S. For a better understanding of the structure it is recommended to look at Figure 1. In this study it is recommended to use the following set of features:

F-1. URL length; F-2. Length hostname; F-3. IP; F-4. Number of dots; F-5. Number of question marks; F-6. Number of equals; F-7. Number of slashes; F-8. Number of “www”; F-9. Ratio digits url; F-10. Ratio digits host; F-11. TLD in subdomain; F-12. Prefix suffix; F-13. Shortest word host; F-14. Longest words raw; F-15. Longest word path; F-16. Phishing hints; F-17. Number of hyperlinks; F-18. Ratio int Hyperlinks; F-19. Empty title; F-20. Domain in title; F-21. Domain age; F-22. Google index; F-23. Page rank; F-24. Status(for legitimate 0, for phishing 0);

**4. Computational experiment**

The objective of this project was to build a machine learning model to classify URLs as either phishing or legitimate. This classification helps in identifying malicious websites and preventing phishing attacks. The dataset used for this project contains various features extracted from URLs, which are utilized to train the model.

***Data Preprocessing***

Libraries such as pandas, matplotlib, and seaborn were imported for data manipulation and visualization. The dataset was loaded from a CSV file into a pandas DataFrame. The dataset contained 11,430 rows and 89 columns. Basic statistics were calculated for the dataset to understand the distribution of features. Then a checking was performed for missing values, and it was found that there were no missing values in the dataset.

***Feature Selection***

A correlation matrix was constructed to identify the relationship between the features and the target variable (status). A heatmap was generated to visualize the correlations. Then the Features with a correlation absolute value greater than 0.2 with the target variable were selected. This resulted in selecting 23 features that had significant correlations with the status.

***Model Training***

The selected features are extracted from the original dataset to form the predictor variables and the target variable (status) is mapped to numeric values, where 'legitimate' is mapped to 0 and 'phishing' is mapped to 1. The dataset was split into training (80%) and test (20%) sets using train\_test\_split from sklearn.model\_selection. Then RandomForestClassifier with 350 estimators and a random state of 42 was used for training and the model was trained on the training dataset. The model is trained on the training set using the fit method of the Random Forest classifier.

***Evaluating Model Accuracy***

The accuracy of the model on the training set is evaluated, resulting in perfect accuracy (100%). The model's accuracy on the test set is evaluated, yielding an accuracy of 96.5%. A confusion matrix is generated for both the training and test sets to provide a detailed view of the model's performance in terms of true positives, true negatives, false positives, and false negatives.

Confusion Matrix for Training Data:

Изображение выглядит как текст, Шрифт, белый, снимок экрана

Автоматически созданное описание

True Negatives (TN): 4558; False Positives (FP): 0; False Negatives (FN): 0; True Positives (TP): 4588;

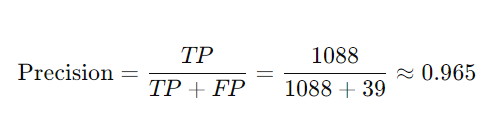
Confusion Matrix for Testing Data:

Изображение выглядит как текст, Шрифт, снимок экрана, белый

Автоматически созданное описание

True Negatives (TN): 1118; False Positives (FP): 39; False Negatives (FN): 41; True Positives (TP): 1088;

**Precision:** Precision is the ratio of correctly predicted positive observations to the total predicted positives.



**Recall (Sensitivity):** Recall is the ratio of correctly predicted positive observations to all observations in the actual class.

Изображение выглядит как текст, Шрифт, линия, белый

Автоматически созданное описание

**F1 Score:** The F1 score is the weighted average of Precision and Recall.

Изображение выглядит как текст, Шрифт, снимок экрана, линия

Автоматически созданное описание

**Conclusion**

In this study was created a novel Random Forest classifier to enhance the recognition and categorization of malicious web addresses. This model was focused on IoT devices. By carefully selecting an optimal feature set from training data sets, it can be improved phishing prevention performance in such cases. Experimental results of our model are remarkable with 96.5 percent accuracy of phishing URLs detection which demonstrates the effectiveness in our approach in properly identifying phishing threats. Proposed method is particularly suitable for deployment on IoT devices due to its emphasis on efficient algorithms. The contribution of this study significantly advances the field of phishing attack detection by providing a practical and effective solution tailored to the constraints and needs of IoT environments. Future research could build upon this work by incorporating additional machine learning techniques and exploring advanced feature selection methods to make more robust and efficient phishing detection systems. Finally, the Random Forest classification model presented in this paper represents a significant leap forward in detecting phishing attacks against cyber threats that target IoT devices at large.

**References**

1. R.J. Anderson, C. Barton, R. Böhme, R. Clayton, M.J.G. van Eeten, M. Levi, T. Moore, S. Savage, in: R. Böhme (Ed.), Measuring the Cost of Cybercrime, The Economics of Information Security and Privacy, Springer, 2013, pp. 265-300.

2. Q. Cui, G.-V. Jourdan, G.V. Bochmann, I.-V. Onut, J. Flood, Phishing Attacks Modifications and Evolutions, in: J. Lopez, J. Zhou, M. Soriano (Eds.), Computer Security, Springer International Publishing, Cham, 2018, pp. 243-262.

3. L. Atzori, A. Iera, G. Morabito, Understanding the Internet of Things: definition, potentials, and societal role of a fast evolving paradigm, Ad Hoc Networks. 56 (2017) 122-140.

4. O.K. Sahingoz, E. Buber, O. Demir, B. Diri, Machine learning based phishing detection from URLs, Expert Systems and Applications. 117 (2019) 345-357.

5. M.T. Suleman, S.M. Awan, Optimization of URL-Based Phishing Websites Detection through Genetic Algorithms, Automatic Control and Computer Sciences. 53 (2019) 333- 341.

6. Marakhimov A.R., Kudaybergenov J.K., Khudaybergenov K.K., Ohundadaev, U.R. A multivariate binary decision tree classifier based on shallow neural network. Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2022, 22(4), pp. 725-733.

7. S. Afroz, R. Greenstadt, Phishzoo: Detecting phishing websites by looking at them, in: Fifth IEEE International Conference on Semantic Computing, 2011.

8. S. Sheng, B. Wardman, G. Warner, L. Cranor, J. Hong, An empirical analysis of phishing blacklists, 2009.

9. S. Haruta, H. Asahina, I. Sasase, Visual similarity-based phishing detection using image and CSS with target website finder, in: IEEE Global Communications Conference, 2017.

10. S. Abdelnabi, K. Krombhoiz, M. Fritz, WhiteNet: Phishing website detection by visual whitelists, in: Cryptography and Security, 2019.

11. Peng, I. Harris, Y. Sawa, Detecting phishing attacks using natural language processing and machine learning, in: IEEE 12th International Conference on Semantic Computing(ICSC), 2018.

12. Yazan Ahmad Alsariera, et al., Ai meta-learners and extra-trees algorithm for the detection of phishing websites, IEEE Access 8 (2020) 142532-142542.

13. A. Tewari, B.B. Gupta, Security, privacy and trust of different layers in Internet-of-Things (IoTs) framework, Future Gener. Comput. (2020).

14. Sahingoz OK, Buber E, Demir O, Diri B. Machine learning based phishing detection from URLs. Expert Systems and Applications. 2019. Vol.117, pp.345-57.

15. Shekokar NM, Shah C, Mahajan M, Rachh S. An ideal approach for detection and prevention of phishing attacks. Procedia Computer Science. 2015. pp.49-82.

16. Nezhad JH, Jahan MV, Tayarani-N M-H, Sadrnezhad Z. Analyzing new features of infected web content in detection of malicious webpages. ISC International Journal of Information Security. 2017. Vol. 9(2), pp. 63-83.

17. Rao RS, Pais AR. Detection of phishing websites using an efficient feature-based machine learning framework. Neural Computing and Applications. 2019. vol. 31, pp.3851-3873.