# A MALICIOUS URL DETECTION FRAMEWORK USING PRIORITY COEFFICIENT AND FEATURE EVALUATION

AHMAD SAHBAN RAFSANJANI

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy

Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia

JANUARY 2023

## DEDICATION

This thesis is dedicated to my family for their unlimited support, encouragement, and love. This thesis is specially dedicated to my son, Ayeen who was born during this journey and has given me smiles, hope, and endless joy.

**ACKNOWLEDGEMENT**

First and foremost, I thank the Almighty God for the grace he bestowed upon me, without which this work would not have been possible. I would like to thank University Technology Malaysia for the unlimited support. I would like to express my gratitude to my supervisor, Dr. Norshaliza Binti Kamaruddin, for her patience, continuous support, assistance, encouragement, and priceless advice from a humanitarian and scientific point throughout this research work. Furthermore, I also acknowledge my prior supervisor, Prof. Dr. Mohd Shahidan bin Abdullah for his invaluable advice, corrections, and guidance for the betterment of this research.

I am very grateful to my parents for their support, encouragement, and love, which enabled me to reach this point. The words cannot explain my sincere thanks to my father, Mohammad, through his teaching and philosophy of life, he taught me how to face real life with a brave heart and to appreciate the worth of hard work and the acquisition of knowledge. I express all my gratitude to my beloved mother, Nooshin, asking god to give her all the best with long life. Last but most importantly, I am expressing all grateful to my adored wife, Fatemeh, for her patience, understanding, support, love, and many sacrifices to keep me on track and help me survive during this adventure. Finally, I dedicate my work to my son, Ayeen, who was born during this journey and has given me smiles, hope, and endless joy.

## ABSTRACT

Malicious Uniform Resource Locators (URLs) are one of the major threats in cybersecurity. Cyber attackers spread malicious URLs to carry out attacks such as phishing and malware, which lead unsuspecting visitors into scams, resulting in monetary loss, information theft, and other threats to website users. At present, malicious URLs are detected using blacklist and heuristic methods, but these methods lack the ability to detect new and obfuscated URLs. Machine learning and deep learning methods have been seen as popular methods for improving the previous method to detect malicious URLs. However, these methods are entirely data-dependent, and a large, updated dataset is necessary for the training to create an effective detection method. Besides, accuracy and detection mostly depend on the quality of training data. This research developed a framework to detect malicious URL based on predefined static feature classification by allocating priority coefficients and feature evaluation methods. The feature classification employed 39 classes of blacklist, lexical, host- based, and content-based features. A dataset containing 2000 real-world URLs was gathered from two popular phishing and malware websites, URLhaus and PhishTank. In the experiment, the proposed framework was evaluated with three supervised machine learning methods: Support Vector Machine (SVM), Random Forest (RF), and Bayesian Network (BN). The result showed that the proposed framework outperformed these methods. In addition, the proposed framework was benchmarked with three comprehensive malicious URL detection methods, which were Precise Phishing Detection with Recurrent Convolutional Neural Networks, Li, and URLNet in terms of accuracy and precision. The results showed that the proposed framework achieved a detection accuracy of 98.95% and a precision value of 98.60%. In sum, the developed malicious URL framework significantly improves the detection in terms of accuracy.

# ABSTRAK

*Malicious Uniform Resource Locators* (URL) adalah salah satu ancaman utama dalam keselamatan siber. Penyerang siber menyebarkan URL perosak untuk melakukan serangan seperti pancingan data dan perisian perosak yang mampu menarik perhatian pengguna internet kepada penipuan siber, mengakibatkan kerugian kewangan, rompakan maklumat dan ancaman lain terhadap pengguna internet. Pada masa ini, URL perosak boleh dikesan menggunakan kaedah senarai hitam dan *trial and error*, tetapi kaedah ini tidak mempunyai keupayaan untuk mengesan URL baharu dan yang tidak jelas. Pembelajaran menggunakan mesin dan kaedah pembelajaran yang mendalam telah dilihat sebagai kaedah yang disukai untuk menambah baik kaedah sebelumnya untuk mengesan URL perosak ini. Walau bagaimanapun, kaedah ini bergantung sepenuhnya kepada data, dan set data yang besar dan dikemas kini diperlukan untuk latihan bagi mewujudkan kaedah pengesanan yang berkesan. Selain itu, ketepatan dan pengesanan sangat bergantung pada kualiti dan latihan. Penyelidikan ini membangunkan pengesanan rangka kerja URL perosak berdasarkan pengelasan ciri statik yang telah ditetapkan dengan memberikan keutamaan kepada pekali keutamaan dan kaedah penilaian ciri. Pengelasan ciri menggunakan 39 senarai hitam, leksikal, *host-based*, dan ciri yang berasaskan kandungan. Set data yang mengandungi 2000 *real-word* URL telah dikumpulkan daripada dua laman web terkenal untuk pancingan data dan perisian perosak, iaitu URLhaus dan PhishTank. Dalam kajian ini, rangka kerja yang dicadangkan telah dinilai dengan tiga kaedah pembelajaran yang diselia, iaitu: *Support Vector Machine* (SVM), *Random Forest* (RF) dan *Bayesian Network* (BN). Keputusan yang diperoleh mendapati bahawa rangka kerja yang dicadangkan mengatasi prestasi kaedah ini. Selain itu, rangka kerja yang dicadangkan telah ditanda aras dengan tiga kaedah pengesanan URL perosak secara menyeluruh, iaitu *Precise Phishing Detection with Recurrent Convolutional Neural Networks* (PDRCNN), kaedah Li dan URLNet dari segi ketepatan dan kejituan. Keputusan menunjukkan bahawa rangka kerja yang telah dicadangkan mencapai ketepatan pengesanan sebanyak 98.95% serta nilai ketepatan sebanyak 98.60%. Secara ringkas, rangka kerja URL perosak yang dibangunkan dengan ketara meningkatkan pengesanan dari segi ketepatan.

**TABLE OF CONTENTS**

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | | |
|---|---|---|
| ANN | - | Artificial Neural Network |
| AdaBoost | - | Adaptive Boosting |
| API | - | Application Programming Interface |
| BN | - | Bayesian Network |
| BoW | - | Bag of Words |
| CA | - | Certificate Authority |
| CatBoost | - | Categorical Boosting |
| CNN | - | Convolutional Neural Network |
| CPT | - | Conditional Probability Table |
| CSRF | - | Cross Site Request Forgery |
| DAG | - | Directed Acyclic Graph |
| DL | - | Deep Learning |
| DML-NYS | - | Distance Metric Learning- Nyström |
| DOM | - | Document Object Model |
| DOM | - | Document Object Model |
| FM | - | Factorization Machines |
| GBDT | - | Gradient Boosting Decision Tree |
| GRU | - | Gated Recurrent Unit |
| HTML | - | Hyper Text Markup Language |
| HTTP | - | Hypertext Transfer Protocol |
| HTTPS | - | Hypertext Transfer Protocol Secure |
| IP | - | Internet Protocol |
| JS | - | JavaScript |
| K-NN | - | K-Nearest Neighbour |
| KL | - | Kullback Leibler |
| L-SVM | - | Linear - Support Vector Machine |
| LDA | - | Linear Discriminant Analysis |
| LightGBM | - | Light Gradient Boosting |
| LR | - | Logistic Regression |

| | | |
|---|---|---|
| LSTM | - | Long Short Term Memory |
| ML | - | Machine Learning |
| MVVM | - | Model-View-ViewModel |
| NB | - | Naive Bayes |
| RF | - | Random Forest |
| RNN | - | Recurrent Neural Network |
| ROC | - | Receiver Operating Characteristic |
| SQL | - | Structured Query Language |
| SSL | - | Secure Sockets Layer |
| SVD | - | Singular Value Decomposition |
| SVM | - | Support Vector Machine |
| TCN | - | Temporal Convolutional Network |
| TF-IDF | - | Term Frequency - Inverse Document Frequency |
| TLD | - | Top Level Domain |
| URL | - | Uniform Resource Locator |
| WWW | - | World Wide Web |
| XGBoost | - | Extreme Gradient Boosting |
| XSS | - | Cross Site Scripting |

## LIST OF APPENDICES

# CHAPTER 1

# INTRODUCTION

## 1.1 Introduction

With the development of Internet technology, the number of world wide web services has increased, like online banking and electronic commerce, which deal with important information such as credit card numbers, bank accounts, and personal information. The crux of this growth on the internet is the rise of various types of cyberattacks on unsuspecting users, and securing this information during any transactions has become a necessity (ALfouzan, 2022). The field of computer security is conducted with a variety of threats and vulnerabilities that can lead to the lack of accessibility of network resources and compromise the system's confidentiality, integrity, and availability (Butt, 2020). Cybersecurity refers to the practice of keeping computer systems and networks safe against unauthorized access, theft, or destruction (Rakotoasimbahoaka, 2019).

Malicious (Harmful) URLs have become one of the most significant cybersecurity threats on the Internet today (Naveen, 2019). The vast majority of cyberattacks are launched as a result by clicking on malicious websites (Manyumwa, 2020). Malicious URLs host unwanted material, which leads unsuspecting visitors into scams, resulting in monetary loss and information theft (Chatterjee, 2019). The field of cybersecurity has developed into a dominating focus for researchers as they attempt to eliminate all threats and dangers provided by malicious URLs (Park, 2022). The fast growth of harmful URL trends makes detection extremely difficult since attackers are sufficiently skilled to utilize numerous obfuscated methods to keep the URLs ambiguous (Alshehri, 2022).

Phishing and malware are the most popular malicious URL attacks, which occur daily and harm millions of people, and can target various operating systems.

(Catak, 2021; Patil, 2018). Phishing is a type of online fraud attack in which attackers use multiple fraudulent social engineering tactics to trick unsuspecting victims into disclosing their personal data (Benavides, 2020). Malware, short for malicious software, is a type of code created by cyberattackers with the goal of causing extensive damage to data and systems or gaining unauthorised access to a network (Krombholz, 2014). The number of phishing and malware websites has increased rapidly from a few thousand in January 2007 to more than 2 million in January 2021(Google-Safe-Browsing, 2022). According to (Johnson, 2022) there were less than 145,000 different phishing websites detected around the world in 2013. But by 2021, this number had jumped to over 630,000, which is a huge increase.

In this research, an efficient framework for detecting malicious URLs is proposed. This framework is based on predefined static feature classifications to overcome the multiple limitations of current malicious URL detection methods. The priority coefficient and feature evaluation methods are allocated to enhance detection accuracy. The priority coefficient is allocated to the classes and features according to the level of importance in order to lend greater weight to the essential classes and features that are effective in detecting malicious URLs. The feature evaluation method assesses the value delivered from feature classification. It determines if all of the features deliver a value, and if any of them fails to deliver, the method will decide to use the other feature's coefficient value instead, which leads to increasing the detection accuracy of the framework. The proposed framework is being evaluated and benchmarked with several methods, and the result shows the proposed framework achieved 98.95% accuracy, which represents a significant improvement over the previous attempts.

## 1.2 Background of the Problem

There have been a lot of scientific research that have demonstrated a variety of methods for detecting malicious URLs. Existing solutions are generally classified into four categories: blacklist, heuristic, machine learning, and deep learning-based (Liang, 2021; Xuan, 2020). In past years, the blacklist feature was the most preferred approach

for identifying harmful websites. URLs that have already been identified as potentially dangerous (phishing, malware) are located in blacklist databases and have gathered over time. (Patil, 2018; Yuan, 2021).

The primary objectives of detecting harmful websites are to defend users against cyber-crime attacks and detect newly generated websites in real time (Rupa, 2021). The malicious URLs used to be discovered mostly by blacklist-based methods previously due to some limitations. The main issue is that this approach is unable to detect newly created malicious URLs since it is impossible to have a complete database of all old and newly generated websites (Sahoo, 2017; Xiao, 2020). The other issue is obfuscation techniques, which can easily pass through databases by simply converting a harmful URL to a safe one and redirecting users to a fake site rather than their intended destination. This approach is currently only used as a supplemental method for detection (Afzal, 2021; Mourtaji, 2021).

The heuristic-based detection methods rely on the statistical similarities between phishing and malware sites, extracted features and gathered critical information regarding a website, and expert knowledge to detect malicious URLs. Malicious URL detection is performed on the basis of these features, which are derived from many observations of known harmful webpages and generalized into a specific set of heuristic rules (Wang, 2019 ; Yuan, 2021).

The researchers utilized heuristics to reduce a large collection of online sites to a more manageable collection of suspect web pages. Although this method overcomes the blacklist method and does not require a large database of malicious URLs, most of the suggested methods are still incapable of detecting harmful websites due to the fact that rules are created based on the behavior of already existing malicious URLs (Almeida, 2020). In addition, analyzing harmful websites involves a great deal of subjective expertise. Currently, the behaviour of phishing and malware websites is diverse, and utilizing rule-based techniques is ineffective .

Due to the weakness of the blacklist and heuristic-based methods to predict new malicious URLs and to overcome this issue, researchers have applied Machine

Learning (ML) techniques in the last decade and achieved significant results (Al-Janabi, 2017; Rakotoasimbahoaka, 2019). The most extensively used method for identifying malicious URLs is machine learning. The goal of machine learning is to create a method from a sample dataset, referred to as training data, and to create a pattern. These patterns may then be used to predict, classify, and cluster URLs (Patil, 2018).

Despite the fact that machine learning algorithms have made significant improvements in detecting malicious URLs over the past decade, there are still numerous critical limitations that remain. Supervised machine learning is one of the most important approaches for identifying malicious URLs with an accuracy rate of more than 90%, and the majority of research has been conducted using this method (Al-Janabi, 2017; Ramesh, 2021; Xuan, 2020). However, this strategy required a massive, tedious, and time-consuming process to collect labelling training data. Another issue with the ML methods is that it needs massive quantities of data for training in order to develop an appropriate detection method with acceptable levels of accuracy.

The most important weakness of machine learning is data-dependent. The detection method's reliability and level of accuracy are entirely dependent on the dataset's quality and a large and updated dataset is necessary for the training. Also, the method that was built based on a training dataset with a high level of accuracy is ineffective for detecting URLs in another dataset (Janet, 2021; Kumar, 2021). The other limitation is retraining time and cost due to short-lived URLs, which refers to the difficulty and complexity of retraining the method with an updated dataset to enable it to detect and interact with new characteristics of websites (Khonji, 2013).

Deep learning approaches have made great progress in detecting malicious URLs over the last few years. Deep learning is a subset of machine learning approaches based on artificial neural networks, and uses current URL datasets to develop a pattern for predicting future URLs. Some machine learning problems have been overcome, such as tedious feature classification, which leads to a training model with minimal effort and results in an appropriate pattern for detecting malicious URLs (Bu, 2021).

This method is currently used widely to develop a pattern for predicting future URLs (Benavides, 2020; Le, 2018; Saxe, 2017).

However there are still a number of major issues remaining (Benavides, 2020; Le, 2018; Saxe, 2017). The main limitation of DL is interpretability. However, deep learning overcomes feature classification, which is one of machine learning's primary issues, but also leads to a new difficulty (Liang, 2021). Deep learning algorithms provide automated feature classification according to the training URLs, which usually leads to fairly good results. But the algorithms do not disclose the details and specifics of the method's prediction and feature classification, which behave like black boxes (Park, 2022).

It is necessary to adhere to specific criteria to develop an efficient malicious URL detection method. In reality, detecting malicious URLs faces a variety of challenges which are (Sadique, 2020; Huang, 2014)

**Realtime Detection.** The malicious URL detection method must be able to notify users about a harmful website before visiting the website to properly protect them.

**Detection of New URLs.** The primary objective of detecting malicious URL methods is to defend users against cyber-crime attacks and related threats in the real world (Rupa 2021).

**Effective Detection.** The malicious URL detection method ought to be efficient and demonstrated by evaluation metrics such as accuracy.

There have been lots of scientific studies that have demonstrated a variety of methods for detecting malicious URLs. However, some drawbacks can be found in methods that must be addressed.

Data dependency is the main challenge to the current effective malicious URL detection methods, which is common to all big data learning methods. This refers to the fact that the reliability and accuracy of the detection method are entirely dependent

on the quality of the dataset (Afzal, 2021). Furthermore, these methods need huge datasets for training in order to create an adequate method, and accuracy is totally dependent on the quality of the dataset. Finally, they necessitate large retraining costs, which require a significant amount of time and software resources (ALfouzan, 2022). Figure 1.1 shows the scenario leading to the problem addressed by this study.

Lack of effective feature classification and evaluation is the other limitation to be addressed related to the current detection methods (Ghaleb, 2022). These methods do not prioritize the classes of features according to level of importance, and all have the equal detection priority level, which may cause a low level of accuracy. Also, feature classification is extremely hard and needs expertise, and may lead to a decreased detection rate due to ineffective feature selection (Sahoo, 2017; Alshehri, 2022). The feature evaluation assesses the feature's value, which is delivered from feature classification, and in the case that the value is not delivered, it needs to identify the URL with other available information. The majority of malicious URL detection methods do not contain feature evaluation, resulting in decreased detection accuracy and inefficiency. Furthermore, it allocates dynamic values to the classes according to the training dataset. These values change over time in the learning methods based on their training dataset.

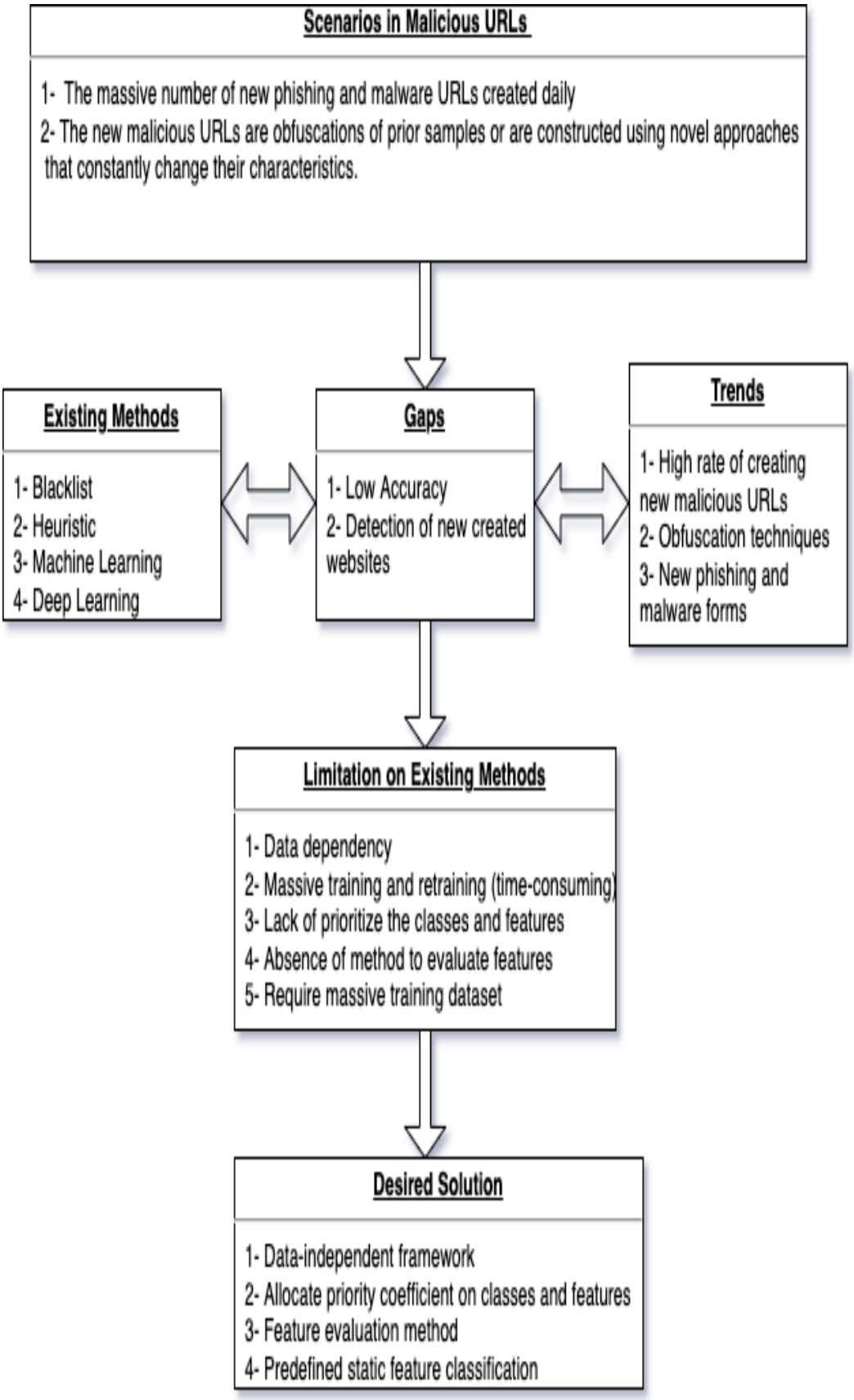


Figure 1.1 Scenarios leading to the problems

## 1.3 Problem Statements

Although several studies have been done on detecting malicious URLs, there are still several fundamental issues that need to be addressed. Data dependency is the primary concern regarding the current malicious URL detection methods that have attracted more attention. The method's reliability and level of accuracy are entirely dependent on the quality of training dataset. In order to develop an effective framework for detecting malicious URLs, a massive and updated dataset is necessary for the training, which takes a lot of time and resources and it gets even worse when retraining of new data is required. However, the method that was built based on a training dataset with a high level of accuracy is ineffective for detecting URLs in another dataset. The second challenging issue to be addressed is the current detection methods do not prioritize the classes and features according to their level of importance, which leads to a low level of accuracy. Each class and feature has a different level of importance in detecting malicious URLs, and facing them should provide distinct results. However, current methods do not distinguish between them. The third challenging issue, which focuses less but is very effective for detecting malicious URLs is feature evaluation. It evaluate the feature's value which deliver from feature classification and determining whether a URL is benign or malicious in the case that the value do not delivered. It may occur for various reasons, but the method must be able to detect URL according to the values of other available features.

This research proposed a framework to enhance the detection accuracy of malicious URLs in real time by allocating predefined static feature classifications and implementing priority coefficients and feature evaluation methods.

## 1.4 Research Questions

1) How to design an accurate framework to detect newly generated malicious URLs in real time?
2) What are the most effective methods to enhance the detection accuracy of malicious URL framework?

3) How to overcome the data dependency limitation of current malicious URL detection methods?
4) How to evaluate and benchmark the proposed framework with other malicious URL detection methods?

## 1.5 Research Objectives

The primary objective of this research is to design a malicious URL detection framework that is capable of detecting newly generated malicious websites in real time with a high level of accuracy. Figure 1.2 illustrates the mapping of research objectives, questions, and research methods. The objectives of the study are as follows:

1) To propose a priority coefficient and feature evaluation methods to provide a more efficient framework.
2) To design predefined static feature classifications by allocating a range of values for the classes.
3) To test and evaluate the proposed framework by benchmarking with supervised machine learning and other malicious URL detection methods.

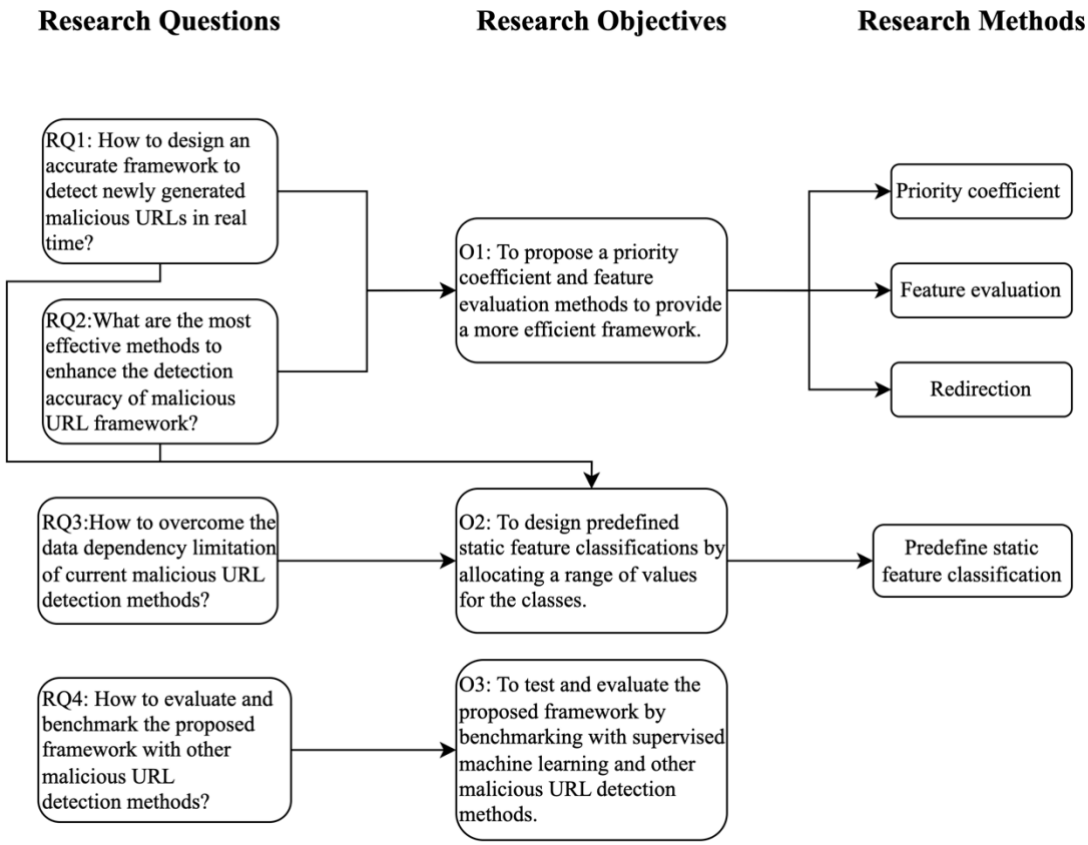


Figure 1.2 Mapping research objectives, questions and research methods

## 1.6 Research Scope

1) This research only focuses on phishing and malware attacks, which are the most popular types of URLs attacks (Patil, 2018; Rupa, 2021).

2) The application is developed based on the MVVM architecture on Kotlin programming languages in Android.

## 1.7 Significant of the Study

The research is important and significant from theoretical and practical perspectives. The rationale and motivation for this research were:

1) The number of phishing and malware websites has increased rapidly from a few thousand in January 2007 to more than 2 million in January 2021 (Google-Safe-Browsing, 2022). The new generated and obfuscated malicious URLs require developing a new framework to overcome the shortage of available methods.

2) An efficient framework is required to overcome the limitations of available methods by detecting newly generated malicious URLs in real time in order to protect computer systems and data.

3) Phishing and malware websites are becoming more complicated and widespread as a result of obfuscation methods (short-URLs), which make users more vulnerable.

4) The malicious URL detection framework requires efficient feature classification and evaluation.

## 1.8 Research Contributions

This research designs an efficient malicious URL detection framework which is able to identify new generated and obfuscated malicious websites with a high level of accuracy. The framework is based on a predefined static feature classification. It utilizes the priority coefficient and feature evaluation methods to provide an accurate malicious URL detection framework. The priority coefficient gives more weight to the important classes and features that are effective at detecting malicious URLs and feature evaluation, assessing the feature's value, which is delivered from feature classification, and determining whether a URL is benign or malicious in the case that the value is not delivered.

The research findings provide a significant improvement in malicious detection accuracy compared with the existing methods. This framework is based on predefine feature classification and does not require a massive dataset for training. The outcome is an Android application that has the possibility to detect newly generated malicious

URLs with a high level of accuracy, and the results can be used both in research and industry.

This application is able to protect users and computers from various types of online attacks such as phishing and malware, which lead unsuspecting visitors into scams and fraud, resulting in monetary loss and information theft. The application analyses the website's status and returns the scanning result, whether it is malicious or benign, before accessing the website. Additionally, it alerts users to any potential threats. Furthermore, it can operate under a variety of conditions and deliver an accurate result.

## 1.9 Definition of Terms

In this section, the main definition and terms adapted in the thesis are presented. Table 1.1 presents the source of all definitions and terms adapted in the thesis.

Table 1.1 Definitions and terms

| **Benign URL** | **Refer to the safe websites.** |
|---|---|
| **Class** | The classes are a subset of features and refer to the sort of information that is extracted from the URL. Each feature includes several classes, and in this research 39 classes are employed. |
| **Feature** | Refers to a category of information that retrieves information regarding a URL. In total, four features are utilized in this research that are blacklist, lexical, host-based, and content-based. |
| **Feature Classification** | The feature classification refers to features and their classes that are utilized to find out the characteristics of a URL to determine whether it is malicious or benign. It includes the processes of creation, extraction, and selection. |
| **Feature Evaluation** | Refer to evaluate the value of features delivered from feature classification for the final calculation. If any of them fails to deliver a value, it uses the other feature priority coefficient’s value instead. |

| | |
|---|---|
| **Malicious URL** | Refers to a website that was created with the purpose of promoting scams, attacks, and frauds that result in monetary loss and information theft. |
| **Predefined feature classification** | This method is part of feature classification and is implemented in feature creation. The predefined values are generated and configured based on the static characteristics of URLs and assign a range of values for the classes in analysis and comparison. |
| **Priority Coefficient** | Refer to lending greater weight to the important classes and features. |
| **Redirect URL** | Refers to types of URLs that are shortened or redirected. |

## 1.10 Thesis Organizations

This chapter presents the introduction of the study. The chapter has provided an introduction to malicious URL detection issues and problems and the motivation for the research by reviewing the background to the problem, as well as outlining the problem statement and the objectives of the research. In addition, the potential contribution of the proposed research has also been highlighted.

Chapter 2 appraises the state of the art of the previous research related to malicious URL detection. This chapter reviews issues and existing techniques in the development of malicious URL detection systems in order to find the problems in these methods. The limitations of their work are discussed to guide the research's direction.

Chapter 3 presents the methodology of the proposed framework to detect malicious URLs. The proposed framework contains three phases: identification, feature classification, and feature evaluation. The specific purpose of the identification phase is that if a URL is redirected (using a short URL), it should be transmitted to the original website. The feature classification is based on a predefined static feature classification and allocating a priority coefficient to selected classes. The feature evaluation assesses the value delivered from feature classification. It determines if all

of the features deliver a value for the final calculation (except lexical features that always return a value), and if any of them fails to deliver a value, it uses the other feature priority coefficient's value instead.

Chapter 4 presents the design and implementation of the framework. This chapter discusses the methods and pseudocodes utilized to design an efficient malicious URL detection framework.

Chapter 5 highlights the experimental results of the framework and evaluates the methods utilized. Also, discuss and benchmark the outcome with other methods.

Finally, Chapter 6 summarizes the findings and the overall results, as well as the conclusion of this study. Also, present suggestions for future work are at the end of the chapter.

# REFERENCES


Acharya, J., Chuadhary, A., Chhabria, A., & Jangale, S. (2021). Detecting Malware, Malicious URLs and Virus Using Machine Learning and Signature Matching. Paper presented at the 2021 2nd International Conference for Emerging Technology (INCET).

Afzal, S., Asim, M., Javed, A. R., Beg, M. O., & Baker, T. (2021). Urldeepdetect: A deep learning approach for detecting malicious urls using semantic vector models. Journal of Network and Systems Management, 29(3), 1-27.

Akiyama, M., Yagi, T., & Itoh, M. (2011). Searching structural neighborhood of malicious urls to improve blacklisting. Paper presented at the 2011 IEEE/IPSJ International Symposium on Applications and the Internet.

Akiyama, M., Yagi, T., & Hariu, T. (2012). Improved blacklisting: inspecting the structural neighborhood of malicious URLs. IT Professional, 15(4), 50-56.

Al-Janabi, M., Quincey, E. D., & Andras, P. (2017). Using supervised machine learning algorithms to detect suspicious URLs in online social networks. Paper presented at the Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017.

Al-Milli, N., & Hammo, B. H. . (2020). A convolutional neural network model to detect illegitimate URLs. Paper presented at the 2020 11th International Conference on Information and Communication Systems (ICICS).

Al-Zahrani, M. S., Wahsheh, H. A., & Alsaade, F. W. (2021). Secure Real-Time Artificial Intelligence System against Malicious QR Code Links. *Security and Communication Networks, 2021*.

Alexa, I. (2022). Alexa. Retrieved from https://www.alexa.com/siteinfo

ALfouzan, N. A., & Narmatha, C. (2022). A Systematic Approach for Malware URL Recognition. Paper presented at the 2022 2nd International Conference on Computing and Information Technology (ICCIT).

Aljofey, A., Jiang, Q., Qu, Q., Huang, M., & Niyigena, J.-P. (2020). An effective phishing detection model based on character level convolutional neural network from URL. *Electronics, 9*(9), 1514.

Aljofey, A., Jiang, Q., Qu, Q., Huang, M., & Niyigena, J.-P. (2020). An effective phishing detection model based on character level convolutional neural network from URL. *Electronics, 9*(9), 1514.

Alkhudair, F., Alassaf, M., Khan, R. U., & Alfarraj, S. . (2020). *Detecting malicious url.* Paper presented at the 2020 International Conference on Computing and Information Technology (ICCIT-1441).

Almeida, R., & Westphall, C. (2020). *Heuristic phishing detection and URL checking methodology based on scraping and web crawling.* Paper presented at the 2020 IEEE International Conference on Intelligence and Security Informatics (ISI).

Alshehri, M., Abugabah, A., Algarni, A., & Almotairi, S. (2022). Character-level word encoding deep learning model for combating cyber threats in phishing URL detection. Computers & Electrical Engineering, 100, 107868.

Althobaiti, K., Rummani, G., & Vaniea, K. (2019). *A review of human-and computer-facing URL phishing features.* Paper presented at the 2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW).

Arceri, V., & Mastroeni, I. . (2021). Analyzing dynamic code: a sound abstract interpreter for evil eval. *ACM Transactions on Privacy and Security (TOPS), 24*(2), 1-38.

AV-TEST. (2020). *SECURITY REPORT 2019/2020*. Retrieved from https://www.av-test.org/fileadmin/pdf/security_report/AV-TEST_Security_Report_2019-2020.pdf

Bahnsen, A. C., Bohorquez, E. C., Villegas, S., Vargas, J., & González, F. A. (2017). *Classifying phishing URLs using recurrent neural networks.* Paper presented at the 2017 APWG symposium on electronic crime research (eCrime).

Begum, A., & Badugu, S. (2020). A study of malicious url detection using machine learning and heuristic approaches. In *Advances in Decision Sciences, Image Processing, Security and Computer Vision* (pp. 587-597). Springer, Cham.

Benavides, E., Fuertes, W., Sanchez, S., & Sanchez, M. (2020). Classification of phishing attack solutions by employing deep learning techniques: A systematic literature review. Developments and advances in defense and security, 51-64.

Bharadwaj, R., Bhatia, A., Chhibbar, L. D., Tiwari, K., & Agrawal, A. (2022). *Is this URL Safe: Detection of Malicious URLs Using Global Vector for Word Representation.* Paper presented at the 2022 International Conference on Information Networking (ICOIN).

Bitly. (2022). Short links, big results. Retrieved from https://bitly.com/

Bu, S.-J., & Cho, S.-B. (2021). Deep character-level anomaly detection based on a convolutional autoencoder for zero-day phishing URL detection. *Electronics, 10*(12), 1492.

Butt, U. A., Mehmood, M., Shah, S. B. H., Amin, R., Shaukat, M. W., Raza, S. M., . . . Piran, M. J. (2020). A Review of Machine Learning Algorithms for Cloud Computing Security. Electronics, 9(9), 1379. Retrieved from https://www.mdpi.com/2079-9292/9/9/1379

Catak, F. O., Sahinbas, K., & Dörtkardeş, V. (2021). Malicious URL detection using machine learning. In Artificial intelligence paradigms for smart cyber-physical systems (pp. 160-180): IGI Global.

Chatterjee, M., & Namin, A. S. (2019). Detecting phishing websites through deep reinforcement learning. Paper presented at the 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC).

Chatterjee, M., & Namin, A. S. (2019). Deep reinforcement learning for detecting malicious websites. *arXiv preprint arXiv:1905.09207*.

Chen, J., Yuan, J., Li, Y., Zhang, Y., Yang, Y., & Feng, R. (2020). A Malicious Web Page Detection Model based on SVM Algorithm: Research on the Enhancement of SVM Efficiency by Multiple Machine Learning Algorithms. Paper presented at the 2020 3rd International Conference on Algorithms, Computing and Artificial Intelligence.

Cisco. (2021). Cybersecurity threat trends: phishing, crypto top the list. Retrieved from https://umbrella.cisco.com/info/2021-cyber-security-threat-trends-phishing-crypto-top-the-list

CTI. (2021). RANSOMWARE ATTACK STATISTICS 2021 – GROWTH & ANALYSIS. Retrieved from https://www.cognyte.com/blog/ransomware_2021/

Da Silva, C. M. R., Feitosa, E. L., & Garcia, V. C. (2020). Heuristic-based strategy for Phishing prediction: A survey of URL-based approach. *Computers & Security, 88*, 101613.

Darling, M., Heileman, G., Gressel, G., Ashok, A., & Poornachandran, P. (2015). A lexical approach for classifying malicious URLs.

Darling, M., Heileman, G., Gressel, G., Ashok, A., & Poornachandran, P. (2015). *A lexical approach for classifying malicious URLs.* Paper presented at the 2015

international conference on high performance computing & simulation (HPCS).

Desai, A., Jatakia, J., Naik, R., & Raul, N. (2017). *Malicious web content detection using machine leaning.* Paper presented at the 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT).

Ding, C. (2020). *Automatic detection of malicious urls using fine-tuned classification model.* Paper presented at the 2020 5th International Conference on Information Science, Computer Technology and Transportation (ISCTT).

Ding, C. (2020). *Automatic detection of malicious urls using fine-tuned classification model.* Paper presented at the 2020 5th International Conference on Information Science, Computer Technology and Transportation (ISCTT).

DomCop. (2022). OpenPageRank. Retrieved from https://www.domcop.com/openpagerank/

Dudheria, R. (2017). *Evaluating Features and Effectiveness of Secure QR Code Scanners.* Paper presented at the 2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC).

Eshete, B., Villafiorita, A., & Weldemariam, K. (2012). *Binspect: Holistic analysis and detection of malicious web pages.* Paper presented at the International conference on security and privacy in communication systems.

Fukushima, Y., Hori, Y., & Sakurai, K. (2011). *Proactive blacklisting for malicious web sites by reputation evaluation based on domain and IP address registration.* Paper presented at the 2011IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications.

Ghaleb, F. A., Alsaedi, M., Saeed, F., Ahmad, J., & Alasli, M. (2022). Cyber Threat Intelligence-Based Malicious URL Detection Model Using Ensemble Learning. *Sensors, 22*(9), 3373.

Google-Safe-Browsing. (2022). Making the world's information safely accessible. Retrieved from https://safebrowsing.google.com/

Gupta, N., Aggarwal, A., & Kumaraguru, P. (2014). *bit. ly/malicious: Deep dive into short url based e-crime detection.* Paper presented at the 2014 APWG Symposium on Electronic Crime Research (eCrime).

Hemavathi, P. (2018). Anti-Malware Phishing QR Scanner. International Journal of Innovative Science and Research Technology, 3(5).

Hoi, S. C., Sahoo, D., Lu, J., & Zhao, P. (2021). Online learning: A comprehensive survey. *Neurocomputing, 459*, 249-289.

Huang, D., Xu, K., & Pei, J. (2014). Malicious URL detection by dynamically mining patterns without pre-defined elements. *World Wide Web, 17*(6), 1375-1394.

Janet, B., & Kumar, R. J. A. (2021). *Malicious URL Detection: A Comparative Study.* Paper presented at the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS).

Johnson, J. (2020). Annual number of malware attacks worldwide from 2015 to 2020. Retrieved from https://www.statista.com/statistics/873097/malware-attacks-per-year-worldwide/

Johnson, J. (2021). Online industries most targeted by phishing attacks as of 1st Quarter 2021. Retrieved from https://www.statista.com/statistics/266161/websites-most-affected-by-phishing/

Johnson, J. (2022). Number of unique phishing sites detected worldwide from 3rd quarter 2013 to 1st Quarter 2021. Retrieved from https://www.statista.com/statistics/266155/number-of-phishing-domain-names-worldwide/

Joshi, A., Lloyd, L., Westin, P., & Seethapathy, S. (2019). Using lexical features for malicious URL detection--a machine learning approach. *arXiv preprint arXiv:1910.06277*.

Kaur, G., & Lashkari, A. H. (2021). Understanding Android Malware Families (UAMF) – The Foundations (Article 1). Retrieved from https://www.itworldcanada.com/blog/understanding-android-malware-families-uamf-the-foundations-article-1/441562

Khonji, M., Iraqi, Y., & Jones, A. . (2013). Phishing detection: a literature survey. *IEEE communications surveys & tutorials, 15*(4), 2091-2121.

Kim, S., Kim, J., Nam, S., & Kim, D. (2018). WebMon: ML-and YARA-based malicious webpage detection. *Computer Networks, 137*, 119-131.

Krombholz, K., Frühwirt, P., Kieseberg, P., Kapsalis, I., Huber, M., & Weippl, E. (2014). QR code security: A survey of attacks and challenges for usable security. Paper presented at the International Conference on Human Aspects of Information Security, Privacy, and Trust.

Kumar, Y., & Subba, B. (2021). *A lightweight machine learning based security framework for detecting phishing attacks.* Paper presented at the 2021 International Conference on COMmunication Systems & NETworkS (COMSNETS).

Kumi, S., Lim, C., & Lee, S. G. (2021). Malicious url detection based on associative classification. *Entropy, 23*(2), 182.

Kunhare, N., Tiwari, R., & Dhar, J. (2020). Particle swarm optimization and feature selection for intrusion detection system. Sādhanā, 45(1), 1-14.

Le, H., Pham, Q., Sahoo, D., & Hoi, S. C. (2018). URLNet: Learning a URL representation with deep learning for malicious URL detection. *arXiv preprint arXiv:1802.03162*.

Lemay, A., & Leblanc, S. P. (2018). *Is eval () evil: A study of JavaScript in PDF malware.* Paper presented at the 2018 13th International Conference on Malicious and Unwanted Software (MALWARE).

Li, T., Kou, G., & Peng, Y. . (2020). Improving malicious URLs detection via feature engineering: Linear and nonlinear space transformation methods. *Information Systems, 91*, 101494.

Liang, Y., Wang, Q., Xiong, K., Zheng, X., Yu, Z., & Zeng, D. (2021). Robust Detection of Malicious URLs with Self-Paced Wide & Deep Learning. IEEE Transactions on Dependable and Secure Computing.

Liu, C., Wang, L., Lang, B., & Zhou, Y. (2018). *Finding effective classifier for malicious URL detection.* Paper presented at the Proceedings of the 2018 2nd International Conference on Management Engineering, Software Engineering and Service Sciences.

Ma, J., Saul, L. K., Savage, S., & Voelker, G. M. (2009). *Identifying suspicious URLs: an application of large-scale online learning.* Paper presented at the Proceedings of the 26th annual international conference on machine learning.

Madhubala, R., Rajesh, N., Shaheetha, L., & Arulkumar, N. (2022). *Survey on Malicious URL Detection Techniques.* Paper presented at the 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI).

Maggi, F., Frossi, A., Zanero, S., Stringhini, G., Stone-Gross, B., Kruegel, C., & Vigna, G. (2013). *Two years of short urls internet measurement: security threats and countermeasures.* Paper presented at the proceedings of the 22nd international conference on World Wide Web.

Mamun, M. S. I., Rathore, M. A., Lashkari, A. H., Stakhanova, N., & Ghorbani, A. A. (2016). *Detecting malicious urls using lexical analysis.* Paper presented at the International Conference on Network and System Security.

Manyumwa, T., Chapita, P. F., Wu, H., & Ji, S. (2020). Towards Fighting Cybercrime: Malicious URL Attack Type Detection using Multiclass Classification. Paper presented at the 2020 IEEE International Conference on Big Data (Big Data).

Marchal, S., François, J., State, R., & Engel, T. (2014). *PhishScore: Hacking phishers' minds.* Paper presented at the 10th International Conference on Network and Service Management (CNSM) and Workshop.

McGahagan, J., Bhansali, D., Pinto-Coelho, C., & Cukier, M. (2019). *A comprehensive evaluation of webpage content features for detecting malicious websites.* Paper presented at the 2019 9th Latin-American Symposium on Dependable Computing (LADC).

Mondal, D. K., Singh, B. C., Hu, H., Biswas, S., Alom, Z., & Azim, M. A. (2021). SeizeMaliciousURL: A novel learning approach to detect malicious URLs. *Journal of Information Security and Applications, 62*, 102967.

Morishige, S., Haruta, S., Asahina, H., & Sasase, I. . (2017). *Obfuscated malicious javascript detection scheme using the feature based on divided url.* Paper presented at the 2017 23rd Asia-Pacific Conference on Communications (APCC).

Mourtaji, Y., Bouhorma, M., Alghazzawi, D., Aldabbagh, G., & Alghamdi, A. (2021). Hybrid rule-based solution for phishing URL detection using convolutional neural network. *Wireless Communications and Mobile Computing, 2021*.

Mutchler, P., Doupé, A., Mitchell, J., Kruegel, C., & Vigna, G. (2015). *A large-scale study of mobile web app security.* Paper presented at the Proceedings of the Mobile Security Technologies Workshop (MoST).

Nadar, V. K., Patel, B., Devmane, V., & Bhave, U. (2021). *Detection of Phishing Websites Using Machine Learning Approach.* Paper presented at the 2021 2nd Global Conference for Advancement in Technology (GCAT).

Naresh, R., Gupta, A., & Giri, S. (2020). Malicious url detection system using combined sym and logistic regression model. *International Journal of Advanced Research in Engineering and Technology (IJARET), 11*(4).

Naveen, I. N. V. D., Manamohana, K., & Verma, R. (2019). Detection of malicious URLs using machine learning techniques. International Journal of Innovative Technology and Exploring Engineering, 8(4S2), 389-393.

Ndichu, S., Kim, S., & Ozawa, S. . (2020). Deobfuscation, unpacking, and decoding of obfuscated malicious JavaScript for machine learning models detection performance improvement. *CAAI Transactions on Intelligence Technology, 5*(3), 184-192.

Palaniappan, G., Sangeetha, S., Rajendran, B., Goyal, S., & Bindhumadhava, B. S. . (2020). Malicious domain detection using machine learning on domain name features, host-based features and web-based features. *Procedia computer science, 171*, 654-661.

Park, K. H., Song, H. M., Do Yoo, J., Hong, S.-Y., Cho, B., Kim, K., & Kim, H. K. (2022). Unsupervised Malicious Domain Detection with Less Labeling Effort. Computers & Security, 102662.

Patgiri, R., Katari, H., Kumar, R., & Sharma, D. . (2019). *Empirical study on malicious URL detection using machine learning.* Paper presented at the International Conference on Distributed Computing and Internet Technology.

Patil, D., & Patil, J. (2018). Feature-based malicious url and attack type detection using multi-class classification. The ISC International Journal of Information Security, 10(2), 141-162.

Patil, D. R., & Patil, J. B. . (2018). Malicious URLs detection using decision tree classifiers and majority voting technique. Cybernetics and Information Technologies, 18(1), 11-29.

Pham, T. T. T., Hoang, V. N., & Ha, T. N. . (2018). *Exploring efficiency of character-level convolution neuron network and long short term memory on malicious url detection.* Paper presented at the Proceedings of the 2018 VII International Conference on Network, Communication and Computing.

Prakash, P., Kumar, M., Kompella, R. R., & Gupta, M. (2010). *Phishnet: predictive blacklisting to detect phishing attacks.* Paper presented at the 2010 Proceedings IEEE INFOCOM.

Prasad, S. D. V., & Rao, K. R. (2021). A novel framework for malicious url detection using hybrid model. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12*(7), 68-76.

Qamar, A., Karim, A., & Chang, V. . (2019). Mobile malware attacks: Review, taxonomy & future directions. *Future Generation Computer Systems, 97*, 887-909.

Rakotoasimbahoaka, A., Randria, I., & Razafindrakoto, N. R. (2019). Malicious URL Detection by Combining Machine Learning and Deep Learning Models. Artificial Intelligence for Internet of Things, 1.

Ramesh, K., BENNET, M. A., VEERAPPAN, J., & RENJITH, P. . (2021). *Performance Metric System for Malicious URL Data using Revised Random Forest Algorithm.* Paper presented at the 2021 5th International Conference on Computing Methodologies and Communication (ICCMC).

Ranveer, S., & Hiray, S. (2015). SVM based effective malware detection system. International Journal of Computer Science and Information Technologies, 6(4), 3361-3365.

Rebrandly. (2022). Your Brand on Your Links. Retrieved from https://www.rebrandly.com/

Ren, F., Jiang, Z., & Liu, J. (2019). A bi-directional lstm model with attention for malicious url detection. In *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)* (Vol. 1, pp. 300-305). IEEE.

Rupa, C., Gautam Srivastava, Sweta Bhattacharya, A Praveen Reddy, A Thippa Reddy Gadekallu. (2021). A Machine Learning Driven Threat Intelligence System for Malicious URL Detection. Paper presented at the The 16th International Conference on Availability, Reliability and Security, Vienna, Austria. https://doi.org/10.1145/3465481.3470029

Sadique, F., Kaul, R., Badsha, S., & Sengupta, S. . (2020). *An automated framework for real-time phishing url detection.* Paper presented at the 2020 10th Annual Computing and Communication Workshop and Conference (CCWC).

Sahoo, D., Liu, C., & Hoi, S. C. (2017). Malicious URL detection using machine learning: A survey. arXiv preprint arXiv:1701.07179.

Sameen, M., Han, K., & Hwang, S. O. . (2020). PhishHaven—an efficient real-time ai phishing URLs detection system. *IEEE Access, 8*, 83425-83443.

Saxe, J., & Berlin, K. . (2017). eXpose: A character-level convolutional neural network with embeddings for detecting malicious URLs, file paths and registry keys. *arXiv preprint arXiv:1702.08568*.

Seshagiri, P., Vazhayil, A., & Sriram, P. (2016). AMA: static code analysis of web page for the detection of malicious scripts. *Procedia computer science, 93*, 768-773.

Shibahara, T., Yamanishi, K., Takata, Y., Chiba, D., Akiyama, M., Yagi, T., . . . Murata, M. (2017). *Malicious URL sequence detection using event de-noising convolutional neural network.* Paper presented at the 2017 IEEE International Conference on Communications (ICC).

Shibahara, T., Takata, Y., Akiyama, M., Yagi, T., & Yada, T. (2017). *Detecting malicious websites by integrating malicious, benign, and compromised redirection subgraph similarities.* Paper presented at the 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC).

Similarweb-LTD. (2022). Similarweb. Retrieved from https://www.similarweb.com/

Singhal, S., Chawla, U., & Shorey, R. . (2020). *Machine learning & concept drift based approach for malicious website detection.* Paper presented at the 2020 International Conference on COMmunication Systems & NETworkS (COMSNETS).

Sistemas, H. (2022). VIRUSTOTAL. Retrieved from https://www.virustotal.com/gui/home/url

Siteefy. (2022). How Many Websites Are There in the World? Retrieved from https://siteefy.com/how-many-websites-are-there/#:~:text=Currently%2C%20there%20are%20around%201.18,active%2C%2083%25%20are%20inactive.

Song, J., Gao, K., Shen, X., Qi, X., Liu, R., & Choo, K. K. R. (2018). QRFence: A flexible and scalable QR link security detection framework for Android devices. *Future Generation Computer Systems, 88*, 663-674.

Sonowal, G., & Kuppusamy, K. (2020). PhiDMA–A phishing detection model with multi-filter approach. *Journal of King Saud University-Computer and Information Sciences, 32*(1), 99-112.

Talal, M., Zaidan, A. A., Zaidan, B. B., Albahri, O. S., Alsalem, M. A., Albahri, A. S., ... & Alaa, M. (2019). Comprehensive review and analysis of anti-malware apps for smartphones. *Telecommunication Systems, 72*(2), 285-337.

Tzacheva, A., Ranganathan, J., & Mylavarapu, S. Y. (2019). *Actionable pattern discovery for tweet emotions.* Paper presented at the International Conference on Applied Human Factors and Ergonomics.

Ulevitch, D. (2022). PhishTank. Retrieved from https://www.phishtank.com/

Vanhoenshoven, F., Nápoles, G., Falcon, R., Vanhoof, K., & Köppen, M. (2016). *Detecting malicious URLs using machine learning techniques.* Paper presented at the 2016 IEEE Symposium Series on Computational Intelligence (SSCI).

Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2018). Evaluating deep learning approaches to characterize and classify malicious URL's. *Journal of Intelligent & Fuzzy Systems, 34*(3), 1333-1343.

Vu, S. N. T., Stege, M., El-Habr, P. I., Bang, J., & Dragoni, N. (2021). A Survey on Botnets: Incentives, Evolution, Detection and Current Trends. *Future Internet, 13*(8), 198.

Vundavalli, V., Barsha , F., Masum, M., Shahriar, H., & Haddad, H. . (2020). *Malicious URL Detection Using Supervised Machine Learning Techniques.* Paper presented at the 13th International Conference on Security of Information and Networks.

Wang, W., Zhang, F., Luo, X., & Zhang, S. . (2019). Pdrcnn: precise phishing detection with recurrent convolutional neural networks. *Security and Communication Networks, 2019*.

Wejinya, G., & Bhatia, S. (2021). Machine learning for malicious url detection. In *ICT Systems and Sustainability* (pp. 463-472): Springer.

WhoXy. (2022). WhoXy. Retrieved from https://www.whoxy.com/

Wu, C.-m., Min, L., Li, Y., Zou, X., & Qiang, B. (2018). *Malicious website detection based on urls static features.* Paper presented at the Proceeding of International Conference on Modeling, Simulation and Optimization.

Xiao, X., Zhang, D., Hu, G., Jiang, Y., & Xia, S. (2020). CNN–MHSA: A Convolutional Neural Network and multi-head self-attention combined approach for detecting phishing websites. Neural Networks, 125, 303-312.

Xu, L., Zhan, Z., Xu, S., & Ye, K. (2013). *Cross-layer detection of malicious websites.* Paper presented at the Proceedings of the third ACM conference on Data and application security and privacy.

Xuan, D., C., Nguyen, H. D., & Nikolaevich, T. V. (2020). Malicious url detection based on machine learning.

Yang, W., Zuo, W., & Cui, B. (2019). Detecting malicious URLs via a keyword-based convolutional gated-recurrent-unit neural network. *IEEE Access, 7*, 29891-29900.

Yao, H., & Shin, D. (2013). *Towards preventing qr code based attacks on android phone using security warnings.* Paper presented at the Proceedings of the 8th ACM SIGSAC symposium on Information, computer and communications security.

Yerima, S. Y., & Alzaylaee, M. K. (2020). *High accuracy phishing detection based on convolutional neural networks.* Paper presented at the 2020 3rd International Conference on Computer Applications & Information Security (ICCAIS).

Yuan, J., Liu, Y., & Yu, L. (2021). A Novel Approach for Malicious URL Detection Based on the Joint Model. Security and Communication Networks, 2021.

Zeltser, L. (2021). Free Online Tools for Looking up Potentially Malicious Websites. Retrieved from https://zeltser.com/lookup-malicious-websites/

Zhang, L., Liu, D., Wang, Z., & Wang, X. (2022). Bayesian Network Structure Learning and Application. *Mobile Information Systems, 2022*.

Zhang, Y., Hong, J. I., & Cranor, L. F. (2007). *Cantina: a content-based approach to detecting phishing web sites.* Paper presented at the Proceedings of the 16th international conference on World Wide Web.

Zhao , P., & Hoi, S. C. (2013). *Cost-sensitive online active learning with application to malicious URL detection.* Paper presented at the Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining.

## LIST OF PUBLICATIONS

### Indexed Journal

1. **Rafsanjani, A. S.,** Kamaruddin, N. (2022). A Evaluating Security and Privacy Features of Quick Response Code Scanners: A Comparative Study. *Open International Journal of Informatics*, 10 (2).

### Non-Indexed Conference Proceedings

1. **Rafsanjani, A. S.** (2018). Comparison Cover Image of Digital Watermarking Based on Discrete Cosine Transform by Using Quick Response Code. *International Conference on Emerging Trends in Engineering, Technologies and Social Sciences (ICETS-2018)*