Literature Survey :

1. Project Title : Detecting phishing websites using machine learning technique

Author Name **:** Ashit Kumar Dutta

Year : 2021

Explanation : In recent years, advancements in Internet and cloud technologies have led to a significant increase in electronic trading in which consumers make online purchases and transactions. This growth leads to unauthorized access to users' sensitive information and damages the resources of an enterprise. Phishing is one of the familiar attacks that trick users to access malicious content and gain their information. In terms of website interface and uniform resource locator (URL), most phishing webpages look identical to the actual webpages. Various strategies for detecting phishing websites, such as blacklist, heuristic, Etc., have been suggested. However, due to inefficient security technologies, there is an exponential increase in the number of victims. The anonymous and uncontrollable framework of the Internet is more vulnerable to phishing attacks. Existing research works show that the performance of the phishing detection system is limited. There is a demand for an intelligent technique to protect users from the cyber-attacks. In this study, the author proposed a URL detection technique based on machine learning approaches. A recurrent neural network method is employed to detect phishing URL. Researcher evaluated the proposed method with 7900 malicious and 5800 legitimate sites, respectively. The experiments' outcome shows that the proposed method's performance is better than the recent approaches in malicious URL detection.

2. Project Title : Phishing Websites Detection using Machine Learning

Author Name **:** Arun Kulkarni1 , Leonard L. Brown, III2

Year : 2019

Explanation : Tremendous resources are spent by organizations guarding against and recovering from cybersecurity attacks by online hackers who gain access to sensitive and valuable user data. Many cyber infiltrations are accomplished through phishing attacks where users are tricked into interacting with web pages that appear to be legitimate. In order to successfully fool a human user, these pages are designed to look like legitimate ones. Since humans are so susceptible to being tricked, automated methods of differentiating between phishing websites and their authentic counterparts are needed as an extra line of defense. The aim of this research is to develop these methods of defense utilizing various approaches to categorize websites. Specifically, we have developed a system that uses machine learning techniques to classify websites based on their URL. We used four classifiers: the decision tree, Naïve Bayesian classifier, support vector machine (SVM), and neural network. The classifiers were tested with a data set containing 1,353 real world URLs where each could be categorized as a legitimate site, suspicious site, or phishing site. The results the experiments show that the classifiers were successful in distinguishing real websites from fake ones over 90% of the time.

3. Project Title : Detecting Phishing Websites Using Machine Learning

Author Name : Aniket Garje1, Namrata Tanwani1 , Sammed Kandale1 , Twinkle Zope1 , Prof. Sandeep Gore2

Year : r 2021 |

Explanation : Phishing is a type of cybersecurity attack that involves stealing personal information such as passwords, credit card numbers, etc. To avoid phishing scams, we have used Machine learning techniques to detect Phishing Websites. Therefore, in this paper, we are trying to find the total number of ways to find Machine Learning techniques and algorithms that will be used to detect these phishing websites. We are using different Machine Learning algorithms such as KNN, Naive Bayes, Gradient boosting, and Decision Tree to detect these malicious websites. The research is divided into the following parts. The introduction represents the focused zone, techniques, and tools used. The Preliminaries section has details of the preparation of the information that is required to move further. Later the paper emphasizes the detailed discussion of the sources of information.

4. Project Title : Phishing Detection using Machine Learning based URL Analysis: A Survey

Author Name : Arathi Krishna V, Anusree A, Blessy Jose, Karthika Anilkumar, Ojus Thomas Lee

Year : 2021

Exlpanation : As we have moved most of our financial, work related and other daily activities to the internet, we are exposed to greater risks in the form of cybercrimes. URL based phishing attacks are one of the most common threats to the internet users. In this type of attack, the attacker exploits the human vulnerability rather than software flaws. It targets both individuals and organizations, induces them to click on URLs that look secure, and steal confidential information or inject malware on our system. Different machine learning algorithms are being used for the detection of phishing URLs, that is, to classify a URL as phishing or legitimate. Researchers are constantly trying to improve the performance of existing models and increase their accuracy. In this work we aim to review various machine learning methods used for this purpose, along with datasets and URL features used to train the machine learning models. The performance of different machine learning algorithms and the methods used to increase their accuracy measures are discussed and analysed. The goal is to create a survey resource for researchers to learn the current developments in the field and contribute in making phishing detection models that yield more accurate results.

5. Project Name : Applications of deep learning for phishing detection: a systematic literature review

Author Name : Cagatay Catal1 , Görkem Giray2 , Bedir Tekinerdogan3 , Sandeep Kumar4 , Suyash Shukla4

Year : 2022

Explanation : Phishing attacks aim to steal confidential information using sophisticated methods, techniques, and tools such as phishing through content injection, social engineering, online social networks, and mobile applications. To avoid and mitigate the risks of these attacks, several phishing detection approaches were developed, among which deep learning algorithms provided promising results. However, the results and the corresponding lessons learned are fragmented over many different studies and there is a lack of a systematic overview of the use of deep learning algorithms in phishing detection. Hence, we performed a systematic literature review (SLR) to identify, assess, and synthesize the results on deep learning approaches for phishing detection as reported by the selected scientific publications. We address nine research questions and provide an overview of how deep learning algorithms have been used for phishing detection from several aspects. In total, 43 journal articles were selected from electronic databases to derive the answers for the defined research questions. Our SLR study shows that except for one study, all the provided models applied supervised deep learning algorithms. The widely used data sources were URL-related data, third party

6. Project Name : A Systematic Literature Review on Phishing and Anti-Phishing Techniques

Author Name : Ayesha Arshad1 , Attique Ur Rehman1 , Sabeen Javaid1 , Tahir Muhammad Ali2 , Javed Anjum Sheikh1 , Muhammad Azeem

Year : : 2021

Explanation : Phishing is the number one threat in the world of internet. Phishing attacks are from decades and with each passing year it is becoming a major problem for internet users as attackers are coming with unique and creative ideas to breach the security. In this paper, different types of phishing and anti-phishing techniques are presented. For this purpose, the Systematic Literature Review(SLR) approach is followed to critically define the proposed research questions. At first 80 articles were extracted from different repositories. These articles were then filtered out using Tollgate Approach to find out different types of phishing and anti-phishing techniques. Research study evaluated that spear phishing, Email Spoofing, Email Manipulation and phone phishing are the most commonly used phishing techniques. On the other hand, according to the SLR, machine learning approaches have the highest accuracy of preventing and detecting phishing attacks among all other anti-phishing approaches.

7. Project Name : Phishing Detection: A Literature Survey

Author Name : Mahmoud Khonji, Youssef Iraqi, Senior Member, IEEE, and Andrew Jones

Year : 2013

Explanation : This article surveys the literature on the detection of phishing attacks. Phishing attacks target vulnerabilities that exist in systems due to the human factor. Many cyber attacks are spread via mechanisms that exploit weaknesses found in endusers, which makes users the weakest element in the security chain. The phishing problem is broad and no single silver-bullet solution exists to mitigate all the vulnerabilities effectively, thus multiple techniques are often implemented to mitigate specific attacks. This paper aims at surveying many of the recently proposed phishing mitigation techniques. A high-level overview of various categories of phishing mitigation techniques is also presented, such as: detection, offensive defense, correction, and prevention, which we belief is critical to present where the phishing detection techniques fit in the overall mitigation process

8. Project Name : Web Phishing Detection Using a Deep Learning Framework

Author Name : Ping Yi , 1 Yuxiang Guan,1 Futai Zou,1 Yao Yao,2 Wei Wang,2 and Ting Zhu 2

Year : 2018

Explanation : Web service is one of the key communications sofware services for the Internet. Web phishing is one of many security threats to web services on the Internet. Web phishing aims to steal private information, such as usernames, passwords, and credit card details, by way of impersonating a legitimate entity. It will lead to information disclosure and property damage. Tis paper mainly focuses on applying a deep learning framework to detect phishing websites. Tis paper frst designs two types of features for web phishing: original features and interaction features. A detection model based on Deep Belief Networks (DBN) is then presented. Te test using real IP fows from ISP (Internet Service Provider) shows that the detecting model based on DBN can achieve an approximately 90% true positive rate and 0.6% false positive rate.

9. Project Name : Survey on Phishing Websites Detection using Machine Learning

Author Name : Mr. B. Ravi Raju1 , S. Sai Likhitha2 , N. Deepa3 , S. Sushma4

Year : 2022

Explanation : Phishing is a widespread method of tricking unsuspecting people into disclosing personal information by using fake websites. Phishing website URLs are designed to steal personal information such as user names, passwords, and online banking activities. Phishers employ webpages that are visually and semantically identical to legitimate websites. As technology advances, phishing strategies have become more sophisticated, necessitating the use of anti-phishing measures to identify phishing. Machine learning is an effective method for combating phishing assaults. This study examines the features utilised in detection as well as machine learning-based detection approaches. Phishing is popular among attackers because it is easier to persuade someone to click on a malicious link that appears to be legitimate than it is to break through a computer's protection measures. The malicious links in the message body are made to look like they go to the faked organisation by utilising the spoofed organization's logos and other valid material. We'll go through the characteristics of phishing domains (also known as fraudulent domains), the qualities that distinguish them from real domains, why it's crucial to detect them, and how they can be discovered using machine learning and natural language processing techniques.

10. Project Name : Phishing Webpage Classification via Deep Learning-Based Algorithms: An Empirical Study

Author Name : Nguyet Quang Do 1 , Ali Selamat 1,2,3,4,* , Ondrej Krejcar 4 , Takeru Yokoi 5 and Hamido Fujita 6,7,8

Year : 2021

Explanation : Phishing detection with high-performance accuracy and low computational complexity has always been a topic of great interest. New technologies have been developed to improve the phishing detection rate and reduce computational constraints in recent years. However, one solution is insufficient to address all problems caused by attackers in cyberspace. Therefore, the primary objective of this paper is to analyze the performance of various deep learning algorithms in detecting phishing activities. This analysis will help organizations or individuals select and adopt the proper solution according to their technological needs and specific applications' requirements to fight against phishing attacks. In this regard, an empirical study was conducted using four different deep learning algorithms, including deep neural network (DNN), convolutional neural network (CNN), Long Short-Term Memory (LSTM), and gated recurrent unit (GRU). To analyze the behaviors of these deep learning architectures, extensive experiments were carried out to examine the impact of parameter tuning on the performance accuracy of the deep learning models. In addition, various performance metrics were measured to evaluate the effectiveness and feasibility of DL models in detecting phishing activities. The results obtained from the experiments showed that no single DL algorithm achieved the best measures across all performance metrics. The empirical findings from this paper also manifest several issues and suggest future research directions related to deep learning in the phishing detection domain.