PROFESSIONAL READINESS FOR INNOVATION, EMPLOYABILITY AND ENTREPRENEURSHIP

LITERATURE SURVEY

MENTOR NAME : DR.S.RAJA KUMAR

TEAM LEADER NAME: K.HARISH SHORKKO

TEAM MEMBERS NAME :

1**.** S.AKASH-2019PECEC364

2. K.HARISH SHORKKO-2019PECEC292

3. R.KISHORE KUMAR-2019PECEC307

4. R.LOKESH-2019PECEC314

TEAM LEADER MAIL ID:harishshorkko02@gmail.com

DOMAIN: APPLIED DATA SCIENCE (ADS)

PROJECT: WEB PHISHING DETECTION

PHASE 2 DESCRIPTION: IDEATION PHASE (LITERATURE SURVEY,EMPATHIZE,DEFINING PROBLEM STATEMENT,IDEATION)

LITERATURE SURVEY

PROJECT TITLE: Detection And Prevention On Web Phishing Using Machine Learning

AUTHOR: Malaika Rastogi ,Anmol Chhetri,Divyanshu Kumar Singh,Gokul Rajan V

YEAR OF PUBLISH: 2021

ABSTRACT: In this paper discusses on the phishing websites prevention and detection. A phishing website is a common social engineering method that mimics trustful uniform resource locators(URLs) and webpages. phishing is the most commonly used social engineering and cyber-attack.

Through such attacks,the phisher targets naïve online users by tricking them into revealing confidential information ,with the purpose of using it fraudulently.

In our paper we will be discussing and listening a few of the artificial intelligence models, that will help us to detect these phishing websites so that in the future these data and technique can be used in machine learning to make our system better and efficient .the problem of phishing is widespread and there is no particular single solution available to effectively reduce all vulnerabilities,so many techniques are often used to reduce certain attacks.Machine learning is a useful tool used to reduce phishing attacks.

As innovation keeps on developing ,phishing strategies began to advance quickly several anti-phishing tools are available and have their own disadvantages . The paper concentrates on basic machine learning supervised classification to seek out an answer to phishing attacks.

The four classification models are KNN,Kernel-SVM,Rando Forest Classifier and Decision tree. The supervised classification contains a labeled dataset that is used to train the models.All the four algorithms used:KNN,Kernel-SVM,Rando Forest Classifier and Decision tree are classification models.With machine learning ,cybersecurity systems can analyze patterns and learn from them to assist prevent similar attacks and answer changing behavior. It can help users to be more active in preventing threats .

PROJECT TITLE: Phishing Website Detection Based On Effective Machine Learning Approach,Journal Of Cyber Security Technology

AUTHOR: HARINAHALLI LOKESH.G,BOREGOWDA.G

YEAR OF PUBLISH:2020

ABSTRACT : Phishing a form of cyber-attack, which has an adverse effect on people where the user is directed to fake websites and duped to reveal their sensitive and personal information which includes passwords of accounts, bank details, atm pin-card details etc. Hence protecting sensitive information from malwares or web phishing is difficult. Machine learning is a study of data analysis and scientific study of algorithms, which has shown results in recent times in opposing phishing pages when distinguished with visualization, legal solutions, including awareness workshops and classic anti-phishing approaches. This paper examines the applicability of ML techniques in identifying phishing attacks and report their positives and negatives. In specific, there are many ML algorithms that have been explored to declare the appropriate choice that serve as anti-phishing tools. We have designed a Phishing Classification system which extracts features that are meant to defeat common phishing detection approaches. We also make use of numeric representation along with the comparative study of classical machine learning techniques like Random Forest, K nearest neighbours, Decision Tree, Linear SVC classifier, One class SVM classifier and wrapper-based features selection which contains the metadata of URLs and use the information to determine if a website is legitimate or not

PROJECT TITLE: Review on Phishing and Anti-Phishing Techniques.

AUTHOR : Ayesha Arshad , Attique Ur Rehman , Sabeen Javaid, Tahir Muhammad Ali , Javed Anjum Sheikh , Muhammad Azeem

YEAR OF PUBLISH: 2021

ABSTRACT- 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.

PROJECT TITLE: Web phishing detection techniques: a survey on the state-of-the-art, taxonomy and future directions

AUTHOR: M. Vijayalakshmi , S. Mercy Shalinie, Ming Hour Yang, Raja Meenakshi U

YEAR OF PUBLISH:2020

ABSTRACT: Internet dragged more than half of the world's population into the cyber world. Unfortunately, with the increase in internet transactions, cybercrimes also increase rapidly. With the anonymous structure of the internet, attackers attempt to deceive the end-users through different forms namely phishing, malware, SQL injection, man-in-the-middle, domain name system tunnelling, ransomware, web trojan, and so on. Amongst them, phishing is the most deceiving attack, which exploits the vulnerabilities in the end-users. Phishing is often done through emails and malicious websites to lure the user by posing themselves as a trusted entity. Security experts have been proposing many anti-phishing techniques. Till today there is no single solution that is capable of mitigating all the vulnerabilities. A systematic review of current trends in web phishing detection techniques is carried out and a taxonomy of automated web phishing detection is presented. The objective of this study is to acknowledge the status of current research in automated web phishing detection and evaluate their performance. This study also discusses the research avenues for future investigation.

PROJECT TITLE: A Survey of Phishing Website Detection Systems.

AUTHOR: Prachit Raut, Harshal Vengurlekar, Rishikesh Shete.

YEAR OF PUBLISH: 2020

ABSTRACT : Phishing URL is a widely used and common technique for cybersecurity attacks. Phishing is a cybercrime that tries to trick the targeted users into exposing their private and sensitive information to the attacker. The motive of the attacker is to gain access to personal information such as usernames, login credentials, passwords, financial account details, social networking data, and personal addresses. These private credentials are then often used for malicious activities such as identity theft, notoriety, financial gain, reputation damage, and many more illegal activities. This paper aims to provide a comprehensive and comparative study of various existing free service systems and researchbased systems used for phishing website detection. The systems in this survey range from different detection techniques and tools used by many researchers. The approach included in these researched papers ranges from Blacklist and Heuristic features to visual and content-based features. The studies presented here use advanced machine learning and deep learning algorithms to achieve better precision and higher accuracy while categorizing websites as phishing or benign. This article would provide a better understanding of the current trends and existing systems in the phishing detection domain.

PROJECT TITLE： A survey of phishing attack techniques, defence mechanisms and open research challenges

AUTHOR: A.Jain, B.Gupta

YEAR OF PUBLISH: 2021

ABSTRACT: Phishing is an identity theft, which deceives Internet users into revealing their sensitive data, e.g., login information, credit/debit card details, and so on. Researchers have developed various anti-phishing methods in recent years. However, the problem still exists. Therefore, this paper presents a detailed analysis of phishing attack methods and defense techniques. This survey is presented in five folds. First, we discuss in detail the lifecycle of phishing attack, its history, and motivation behind this attack. Second, we present various distribution methods that are used to spread phishing attacks. Third, we provide taxonomy of various phishing-attacking techniques in desktop and mobile environments. Fourth, we provide numerous phishing protection mechanisms and their comparisons. Finally, the article presents various performance challenges faced by developers while dealing with this crucial attack. This paper also provides the consequences of phishing attacks in emerging domains like mobile and online social networks. This paper will help the different users in avoiding phishing attacks while using Internet for their day-to-day activities, and will guide business administrators in designing new effective solutions for their enterprise against various types of phishing threats.

PROJECT TITLE: Detection of Phishing Websites using Machine Learning

AUTHOR: Atharva Deshpande,Omkar Pedamkar,Nachiket Chaudhary,Dr. Swapna Borde.

YEAR OF PUBLISH: 2021

ABSTRACT: Phishing is a common attack on credulous people bymaking them to disclose their unique information usingcounterfeit websites. The objective of phishing website URLs isto purloin the personal information like user name, passwordsand online banking transactions. Phishers use the websiteswhich are visually and semantically similar to those realwebsites. As technology continues to grow, phishing techniquesstarted to progress rapidly and this needs to be prevented byusing anti-phishing mechanisms to detect phishing. Machinelearning is a powerful tool used to strive against phishingattacks. This paper surveys the features used for detection anddetection techniques using machine learning.Phishing is popular among attackers, since it is easier to tricksomeone into clicking a malicious link which seems legitimatethan trying to break through a computer's defense systems. Themalicious links within the body of the message are designed tomake it appear that they go to the spoofed organization usingthat organization's logos and other legitimate contents.Here, we explain phishing domain (or Fraudulent Domain)characteristics, the features that distinguish them fromlegitimate domains, why it is important to detect these domains,and how they can be detected using machine learning andnatural language processing techniques.

PROJECT TITLE: Intelligent phishing URL detection using association rule mining.

AUTHOR: S.Carolin Jeeva,Elijah Blessing Rajsingh

YEAR OF PUBLISH:2016

ABSTRACT: Phishing is an online criminal act that occurs when a malicious webpage impersonates as legitimate webpage so as to acquire sensitive information from the user. Phishing attack continues to pose a serious risk for web users and annoying threat within the field of electronic commerce. This paper focuses on discerning the significant features that discriminate between legitimate and phishing URLs. These features are then subjected to associative rule mining--apriori and predictive apriori. The rules obtained are interpreted to emphasize the features that are more prevalent in phishing URLs.

Analyzing the knowledge accessible on phishing URL and considering confidence as an indicator, the features like transport layer security, unavailability of the top level domain in the URL and keyword within the path portion of the URL were found to be sensible indicators for phishing URL. In addition to this number of slashes in the URL, dot in the host portion of the URL and length of the URL are also the key factors for phishing URL.

PROJECT TITLE: Detecting Phishing-Sites using Hybrid Model

AUTHOR : Poonam Kumari, Apoorva H R Gowda, Bhandhavya K, Bhavya M U, Spurthi M N

YEAR OF PUBLISH:2020

ABSTRACT: There is a drastic increase in number of online users due to highly enhanced online technologies. This led to increase in security threats as people now a days are using services from chatting to banking transactions through online. There are many security issues that people are facing from hackers/attackers. There are many types of attacks like keylogger, waterhole attacks, eavesdropping, phishing and many more. The attacker here is called the phisher, where phisher tricks the online users to reveal their confidential/sensitive information like bank account number, password, social network password etc., using phishing websites. There already exists many approaches and techniques to detect and filter out the phishing websites but still researches are going on to find a solution that provides best accuracy. Phishing website has a certain features/patterns through which it can be identified using data mining techniques and the phenomenon called as classification. An hybrid model for classification is presented in this paper to overcome phishing-sites problem. Through this approach we obtain higher accuracy as result.

REFERENCES:

1.Malaika Rastogi,Anmol Chhetri,Divyanshu Kumar Singh,Gokul Rajan V "Detection And Prevention On Web Phishing Using Machine Learning" Applied Data Science pp.45-47, 2021.

2.Harinahalli Lokesh, G. and BoreGowda, G.,"Phishing website detection based on effective machine learning approach,"Journal of Cyber Security Technology,pp.1-14,2020

3. Ayesha Arshad , Attique Ur Rehman , Sabeen Javaid, Tahir Muhammad Ali , Javed Anjum Sheikh , Muhammad Azeem "Review on Phishing and Anti-Phishing Techniques" Information Science pp.0-14,2021

4. M. Vijayalakshmi , S. Mercy Shalinie, Ming Hour Yang, Raja Meenakshi U "Web phishing detection techniques: a survey on the state-of-the-art, taxonomy and future directions",IET Networks pp.07-11,2020

5. Prachit Raut, Harshal Vengurlekar, Rishikesh Shete "A Survey of Phishing Website Detection Systems",Applied Data Science ,2020

6. **:** A.Jain, B.Gupta "A survey of phishing attack techniques, defence mechanisms and open research challenges",Applied Data Science,pp.0-14,2021

7. Atharva Deshpande,Omkar Pedamkar,Nachiket Chaudhary,Dr. Swapna Borde. "Detection of Phishing Websites using Machine Learning",Cyber Security,pp.2-25,2021

8. S.Carolin Jeeva,Elijah Blessing Rajsingh "Intelligent phishing URL detection using association rule mining", Applied Data Science,pp.35-37,2016

9. poonam kumari, apoorva h r gowda, bhandhavya k, bhavya m u, spurthi m n, 2020, detecting phishing-sites using hybrid model, international journal of engineering research & technology (ijert) ncetesft – 2020 (volume 8 – issue 14)