WEB PHISHING DETECTION

LITRATURE SURVEY

TITLE: Phishing Detection from URLs Using Deep Learning Approach

AUTHOR: Shweta Singh, M.P. Singh, Ramprakash Pandey

ABSTRACT:

Today, the Internet covers worldwide. All over the world, people prefer an E-commerce platform to buy or sell their products. Therefore, cybercrime has become the centre of attraction for cyber attackers in cyberspace. Phishing is one such technique where the unidentified structure of the Internet has been used by attackers/criminals that intend to deceive users with the use of the illusory website and emails for obtaining their credentials (like account numbers, passwords, and PINs). Consequently, the identification of a phishing or legitimate web page is a challenging issue due to its semantic structure. In this paper, a phishing detection system is implemented using deep learning techniques to prevent such attacks. The system works on URLs by applying a convolutional neural network (CNN) to detect the phishing webpage. In paper [19] the proposed model has achieved 97.98% accuracy whereas our proposed system achieved accuracy of 98.00% which is better than earlier model. This system doesn't require any feature engineering as the CNN extract features from the URLs automatically through its hidden layers. This is other advantage of the proposed system over earlier reported in [19] as the feature engineering is a very time-consuming task.

TITLE: WC-PAD: Web Crawling based Phishing Attack Detection

AUTHOR: Nathezhtha, Sangeetha ,Vaidehi.

ABSTRACT:

Abstract- Phishing is a criminal offense which involves theft of user's sensitive data. The phishing websites target individuals, organizations, the cloud storage hosting sites and government websites. Currently, hardware-based approaches for Anti phishing are widely used but due to the cost and operational factors software-based approaches are preferred. The existing phishing detection approaches fails to provide solution to problem like zero-day phishing website attacks. To overcome these issues and precisely detect phishing occurrence a three-phase attack detection named as Web Crawler based Phishing Attack Detector (WC-PAD) has been proposed. It takes the web traffics, web content and Uniform Resource Locator (URL) as input features, based on these features classification of phishing and non-phishing websites are done. The experimental analysis of the proposed WC-PAD is done with datasets collected from real phishing cases. From the experimental results, it is found that the proposed WC-PAD gives 98.9% accuracy in both phishing and zero-day phishing attack detection

TITLE: A Deep Learning-Based Framework for Phishing Website Detection

AUTHOR: LIZHEN TANG ,QUSAY H. MAHMOUD

ABSTRACT:

Phishing attackers spread phishing links through e-mail, text messages, and social media platforms. They use social engineering skills to trick users into visiting phishing websites and entering crucial personal information. In the end, the stolen personal information is used to defraud the trust of regular websites or financial institutions to obtain illegal benefits. With the development and applications of machine learning technology, many machine learning-based solutions for detecting phishing have been proposed. Some solutions are based on the features extracted by rules, and some of the features need to rely on third-party services, which will cause instability and time-consuming issues in the prediction service. In this paper, we propose a deep learning-based framework for detecting phishing websites. We have implemented the framework as a browser plug-in capable of determining whether there is a phishing risk in real-time when the user visits a web page and gives a warning message. The real-time prediction service combines multiple strategies to improve accuracy, reduce false alarm rates, and reduce calculation time, including whitelist filtering, blacklist interception, and machine learning (ML) prediction. In the ML prediction module, we compared multiple machine learning models using several datasets. From the experimental results, the RNN-GRU model obtained the highest accuracy of 99.18%, demonstrating the feasibility of the proposed solution.

TITLE: Phishing Detection in Websites using Parse Tree Validation

AUTHOR: C. Emilin Shyni, Anesh D Sundar, G.S.Edwin Ebby.

ABSTRACT:

Phishing is a technique of tricking people into giving sensitive information like usernames and passwords, credit card details, sensitive bank information, etc., by way of email spoofing, instant messaging, or using fake web sites whose look and feel gives the appearance of a legitimate website. In this work, a technique named parse tree validation is proposed to determine whether a webpage is legitimate or phishing. It is a novel approach to detect the phishing web sites by intercepting all the hyperlinks of a current page through Google API, and constructing a parse tree with the intercepted hyperlinks. This technique is implemented and tested with 1000 phishing pages and 1000 legitimate pages. The false negative rate achieved was 7.3% and the false positive rate achieved was 5.2%.

TITLE: Phishing Web Page Detection Methods: URL and HTML Features Detection

AUTHOR: Humam, Faris, Setiadi, Yazid.

ABSTRACT:

Phishing is a type of fraud on the Internet in the form of fake web pages that mimic the original web pages to trick users into sending sensitive information to phisher. The statistics presented by APWG and Phistank show that the number of phishing websites from 2015 to 2020 tends to increase continuously. To overcome this problem, several studies have been carried out including detecting phishing web pages using various features of web pages with various methods. Unfortunately, the use of several methods is not really effective because the design and evaluation are only too focused on the achievement of detection accuracy in research, but evaluation does not represent application in the real world. Whereas a security detection device should require effectiveness, good performance, and deployable. In this study the authors evaluated several methods and proposed rules-based applications that can detect phishing more efficiently

TITLE: A Machine Learning Approach for URL Based Web Phishing Using Fuzzy Logic as Classifier

AUTHOR: Happy Chapla, Riddhi Kotak, Mittal Joiser.

ABSTRACT:

Phishing is the major problem of the internet era. In this era of internet the security of our data in web is gaining an increasing importance. Phishing is one of the most harmful ways to unknowingly access the credential information like username, password or account number from the users. Users are not aware of this type of attack and later they will also become a part of the phishing attacks. It may be the losses of financial found, personal information, reputation of brand name or trust of brand. So the detection of phishing site is necessary. In this paper we design a framework of phishing detection using URL.