Web Phishing Detection

LITERATURE SURVEY

Date: 15 October 2022

Project Name: Web Phishing Detection

Abstract:

Phishing is a common attack against Internet users that causes them to reveal their information using fake websites. The goal of the fake website is to steal personal information such as usernames, passwords and online banking transactions. Scammers use websites that are visually and semantically similar to the real ones. As technology continues to advance, phishing techniques begin to advance rapidly, and this should be prevented by using anti-phishing mechanisms such as spoofed URL detection. Machine Learning is a powerful tool used to combat spoofing attacks. This report covers machine learning technology to detect fake URLs by extracting and analyzing different characteristics of legitimate and fake URLs. Random Forest, Logistic Regression and algorithms are used to detect fake websites.

Introduction:

Nowadays, the Internet plays an important role in communication, where people create an online environment to manage business functions, online activities of banks, social networks… However, the Internet also contains hidden things. a lot of risk because when users operate in an online environment, they can be vulnerable to attackers. And their identity is often a fake URL. And spoofed URLs are often placed on popular websites or sent to user emails.

A meta-analysis of field experiments on phishing susceptibility

Jason Hong

Published 2009

Phishing attacks are a significant security threat to users of the Internet, causing tremendous economic loss every year. Past work in academia has not been adopted by industry in part due to concerns about liability over false positives. However, blacklist-based methods heavily used in industry are slow in responding to new phish attacks, and tend to be easily overwhelmed by phishing techniques such as fast-flux and the proliferation of toolkits

Mahmoud khonji, Y Iraqi Andrew Jones

Computer science

IEEE communication surveys tutorial 2013

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 end-users, 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

T. Sommestad, Henrik Karlzén

Published 1 November 2019

2019 APWG Symposium on Electronic Crime Research (eCrime)

Phishing is a serious threat to any organization allowing their employees to use messaging systems and computers connected to the internet. Consequently, researchers have undertaken a large number of studies to identify the variables that determine this threat, i.e., variables that influence users' susceptibility to phishing emails. Both relative risks and association tests showed that technical warning systems, email personalization, training, and the use of established deceptive tactics influence the susceptibility rate. The type of scam as such also appears to be important, with some types of scams being orders of magnitude more successful than other types. Many of the results had limitations in control and sampling, which may explain unexpected and contradictory results.

1. URL-based Phishing Websites Detection via Machine Learning

2021 - IEEE - International Conference on Data Analytics for Business and Industry (ICDABI)

Phishing is a social engineering cybersecurity attack that involves an attacker who provides a counterfeit piece of information that is hand-crafted skillfully to trick a user (human victim usually) to provide sensitive information to the attacker or to install malicious software on the victim’s computing platform. The system developed in this paper for phishing websites is detected using URL addresses and solves binary classification problems where websites are classified into either authentic or phishing websites. . Aimed at improving the computational efficiency of our system by providing an optimized implementation using fast machine learning frameworks built using low-level programming languages

2. A Machine Learning Approach for URL Based Web Phishing Using Fuzzy Logic as Classifier

2019 International Conference on Communication and Electronics Systems (ICCES)

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.

References:

1. URL-based Phishing Websites Detection via Machine Learning | IEEE Conference Publication | IEEE Xplore

2. A Machine Learning Approach for URL Based Web Phishing Using Fuzzy Logic as Classifier | IEEE Conference Publication | IEEE Xplore

Conclusion:

This paper aims to enhance detection method to detect phishing website using machine learning technology. Also, classifiers generated by machine learning algorithms identify legitimate phishing websites. The proposed technique can detect new temporary phishing sites and reduce the damage caused by phishing attacks. The performance of the proposed technique based on machine learning is more effective that previous phishing detection technologies.