Heuristic Based Approch for Phishing Site Detection

Project Guide:- Prof.Sunil Sushir Sir

GROUP MEMBERS:-

1]CHIRAG CHAUDHARI (15204030) 2] SWAPNIL KSHETRE (15204003)

3]SWAPNIL GHAWALI (15204009) 4] AAKASH SANE (15204012)

LOGIN
PASSWORD

Index

- Abstract
- Introduction
- Literature Review
- Problem Statement
- Project Scope
- Use Case Diagram
- Architecture
- Project Future Plan
- Summary
- Referance

Abstract

- Damage caused by phishing attacks that target personal user information is increasing.

- Phishing involves sending an email to a user or inducing a phishing page to steal a user’s personal information. This type of attack can be detected by blacklist-based detection techniques.

- However, these methods have some disadvantages and the numbers of victims have therefore continued to increase.

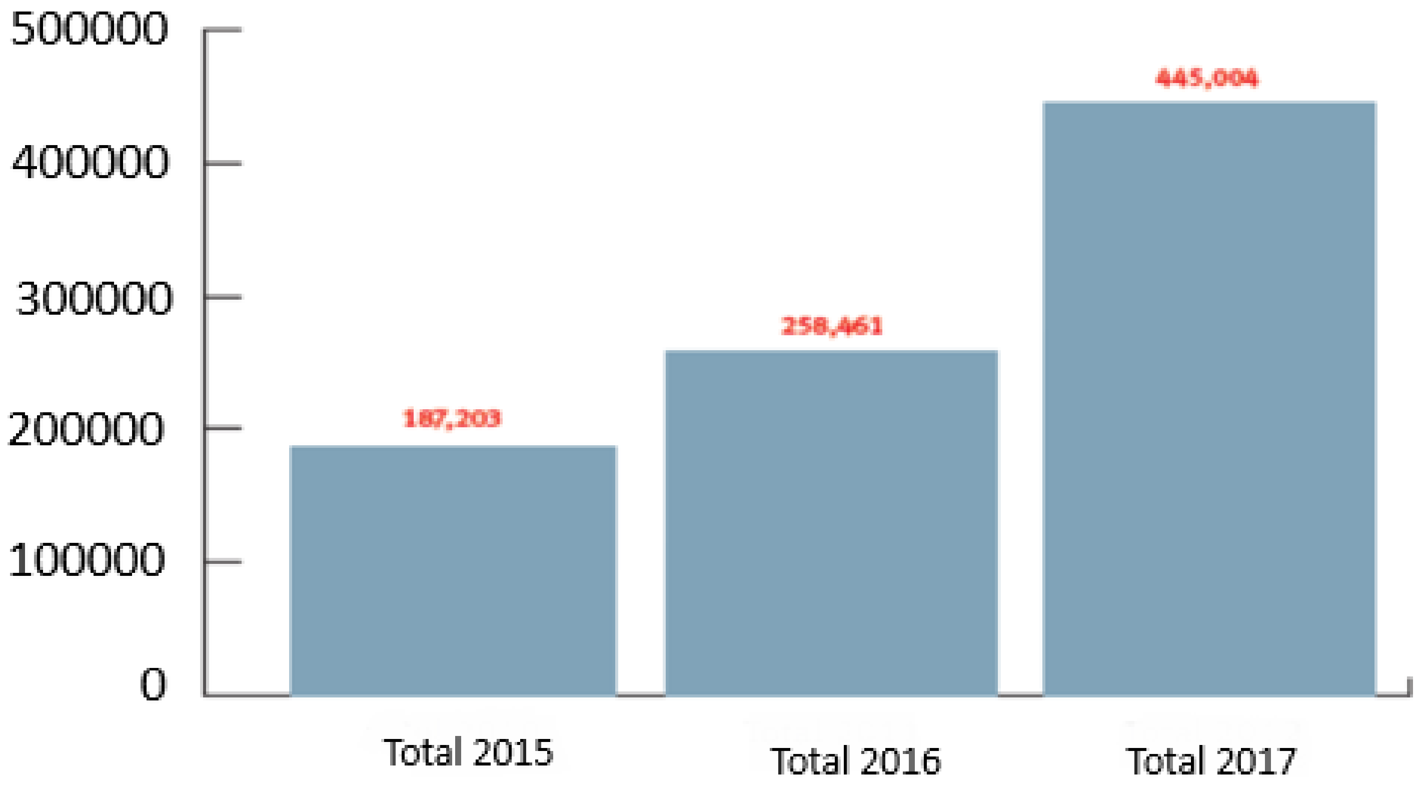
- we propose a heuristic-based phishing detection technique that uses uniform resource locator (URL) features. We identified features that phishing site URLs contain.

Introduction

- Phishing is a malicious use of Internet resources carried out to trick Internet users to reveal personal information, such as usernames, credit card information, and Social Security numbers to the attacker.

- Phishing can appear through a variety of communication forms such as instant messaging, SMS, VOIP, online messenger, and above all the most common form of phishing attack leverages email.

Phishing Attack Per Year



Literature Review

Title	Problem Identified	Methodology	Strength	weakness
The Sceured Anti-Phishing approach Using image based Validation. Y.Yesu Jyothi,D.Srinivas & k.GovindaRaju,2013	To solve the problem of phishing & protect individual personal private information	Visual Cryptography (image based validation)	It prevents attack of phishing websites on financial portal,banking portal & online shopping market	Inability to recover missing or corrupt share
Protecting users Against phishing attacks. Engin kirda & Chistopher Kruegel,2012	Increased email linked to phishing scams	Browser Extension	It protect users against spoofed website – based phishing attacks	It requires that user support to capture & store sensitive information rather than automatically captureing & storing the sensitive information

Problem Statement

- Phishing has been a major security threat in which there is a huge loss for companies as well as customers. These phishing attacks are increasing day by day due to lack of efficient detection techniques and effective preventive measures.

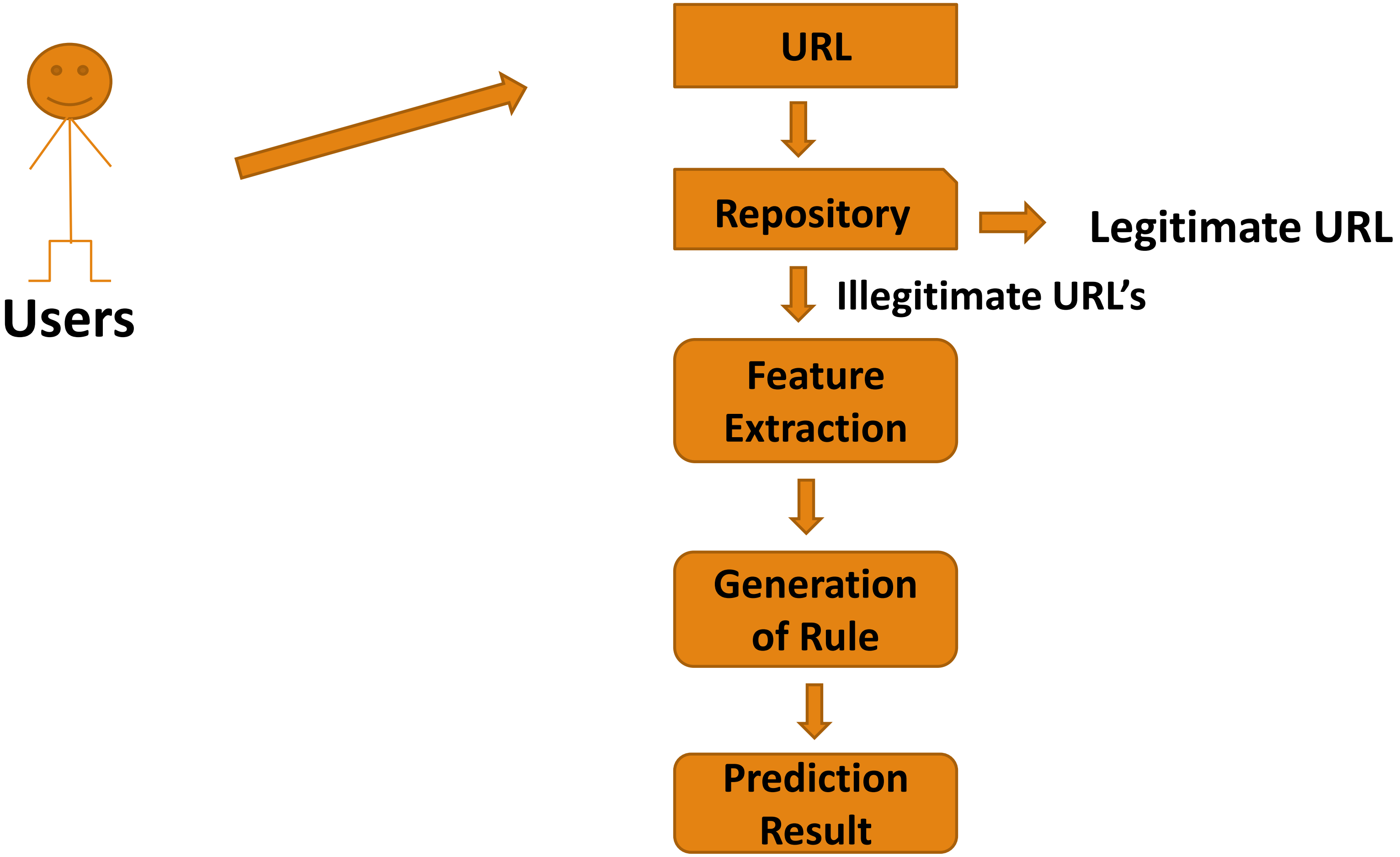
- A comprehensive efficient detection technique should be developed in order to detect and inform the web users about the phishing attacks to make sure that their sensitive data will not be disclosed during these attacks.

- This research project deals with a comprehensive heuristic based method for phishing detection which is based on content of the website through which phishing attacks can be discovered

Project Scope

- Though there are many phishing detection, the scope of the project is limited to feature based phishing detection techniques.
- It extracts the discriminative features from the websites which help in identifying the website class.
- In this process, rules play an important role as they are easily understood by humans.
- The rules are formed in such a way that IF a condition THEN class category where class category represents the category to which a class belongs to.
- This rule induction helps to facilitate the decision making process which ensures reliability and completeness

Usecase Diagram



Users

URL

Repository

Legitimate URL

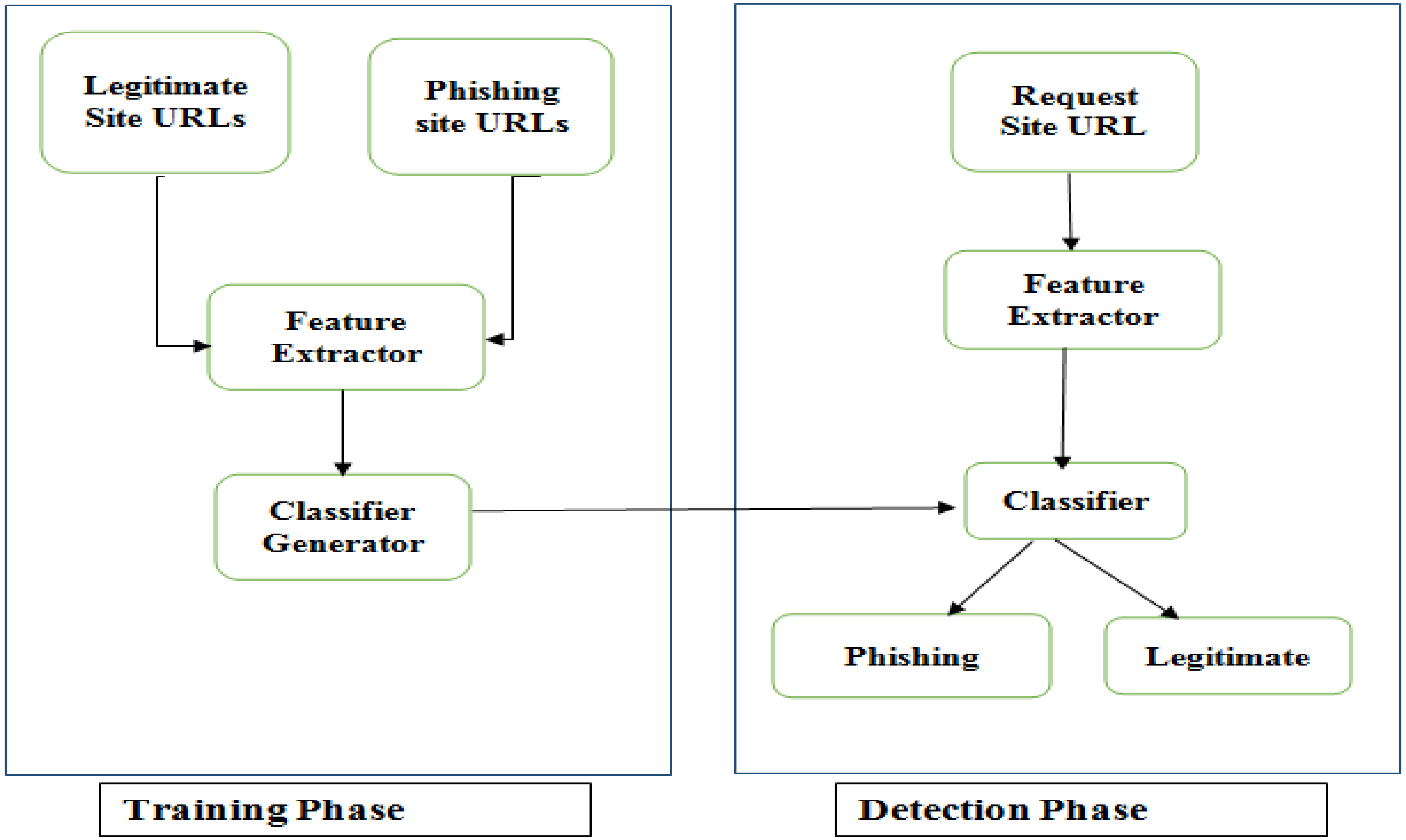
Illegitimate URL’s

Feature Extraction

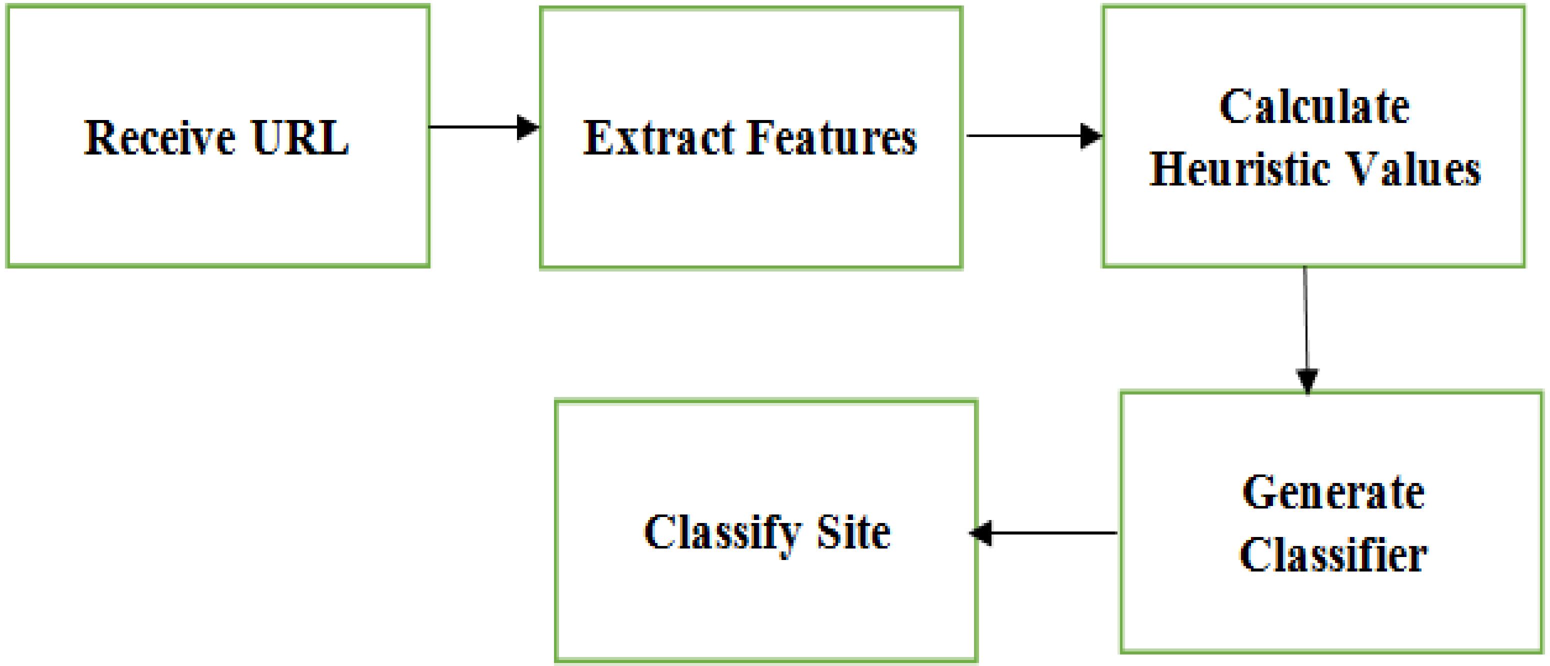
Generation of Rule

Prediction Result

Architecture



Block Diagram



Project Future Plan

- In future work, we intend to address the time-intensive disadvantage of the heuristic-based technique.

- With a large number of features, it is time-consuming for the heuristic-based approach to generate classifiers and perform classification.

- Therefore, we will apply algorithms to reduce the number of features and thereby improve performance.

- In addition, we will examine a new phishing detection technique that uses not only URL-based features, but also HTML and JavaScript features of web pages to improve performance.

Summary

- we proposed a heuristic-based phishing detection technique that employs URL-based features. The method combines URL-based features used in previous studies with new features by analyzing phishing site URLs. Additionally, we generated classifiers through several machine learning algorithms and determined that the best classifier was random forest.

- It showed a high accuracy of 98.23% and a low false-positive rate.

- The proposed technique can provide security for personal information and reduce damage caused by phishing attacks because it can detect new and temporary phishing sites that evade existing phishing detection techniques, such as the blacklist-based technique.

References

- Khonji, Mahmoud, Youssef Iraqi, and Andrew Jones. "Phishing detection: a literature survey." Communications Surveys & Tutorials,IEEE 15.4 (2013): 2091-2121.

- G. Ramesh, I. Krishnamurthi and K. S. S. Kumar, An Efficacious Method for Detecting Phishing Webpages through Target Domain Identification, Decision Support Systems, vol. 61, pp. 12–22, (2014). [Online]. Available at: http://www.sciencedirect.com/science/article/pii/S0167923614000037

- Sunil, A. Naga Venkata, and Anjali Sardana. "A pagerank based detection technique for phishing web sites." Computers & Informatics(ISCI), 2012 IEEE Symposium on. IEEE, 2012.

- Wikipedia. (2015. March) Uniform Resource Loactor Avaliable:http://en.wikipedia.org/wiki/Uniform_resource_locatorWikipedia. (2015. March) Uniform Resource Loactor. Avaliable:http://en.wikipedia.org/wiki/Uniform_resource_locator

Thank
You!