Web Phishing Detection: A Literature Survey

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 research base systems used for phishing website detection. The systems in this survey range from different detection techniques and tools used by many researchers.

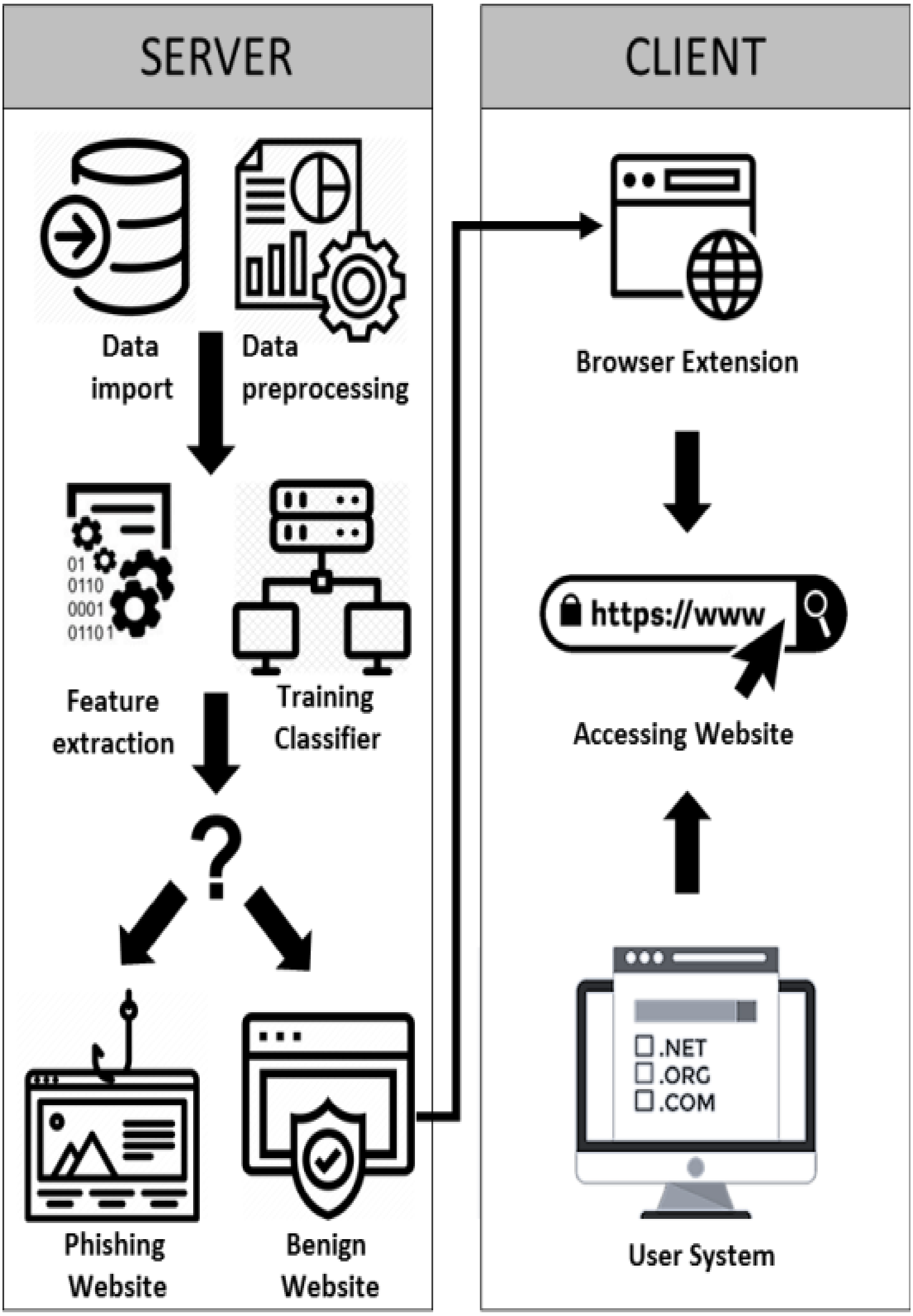
INTRODUCTION:

The advancement of internet is resulting in attracting more and more users into this huge Internet Sea. There are a lot of perks of using internet, one can buy stuff online, way of learning and gaining knowledge has improved, etc. On the contrary, possible threats comes hand in hand. One of them is Phishing Attack. Phishing is an attack where a legitimate user is deceived to disclose sensitive information and assets with economic value. Loss of such sensitive information might cause potential economic or reputational harm an organization. Phishing basically uses social engineering techniques to trick users such as creating fake websites which clones with same attributes and design of the existing legitimate one. In a classic phishing attack a phisher send a link enclosed in a message to the user. The link redirects the user to the cloned malicious page which looks similar to the original webpage but is not and is intended to steal user's sensitive data. Such phishing attacks have proven to cause a lot of financial loss to various organizations. Thus, phishing attacks can be prevented by exterminating such harmful *Website URLs are categorized into the following three classes:*

*a. **Benign**: These are Safe websites that provide normal services to people.*

*b**. Malware**: These websites which are created by attackers look like normal websites can make use of sensitive contents of people.*

*c. **Spam**: These websites flood the user's system with advertisements, fake surveys, etc.*



SYSTEM ARCHITECTURE

APPROACHES:

1.Detection Approaches:

•User training approaches — end-users can be educated to better understand the nature of phishing attacks, which ultimately leads them into correctly identifying phishing and non-phishing messages.

•Software classification approaches — these mitigation approaches aim at classifying phishing and legitimate messages on behalf of the user in an attempt to bridge the gap that is left due to the human error or ignorance.

2.Offensive Defensive Approaches:

Offensive defensive solutions aim to render phishing campaigns useless for the attackers by disrupting the phishing campaigns. This is often achieved by flooding phishing web-sites with fake credentials so that the attacker would have a difficult time to find the real credentials.

3.Correction Approaches:

Once a phishing campaign is detected, the correction pro-cess can begin. In the case of phishing attacks, correction is the act of taking the phishing resources down. This is often achieved by reporting attacks to Service Providers. Phishing campaigns often rely on resources, such as:

*•**Websites** — could be a shared web host owned by the phisher, a legitimate website with phishing content uploaded to it, or a number of infected end-user work-stations in a botnet.*

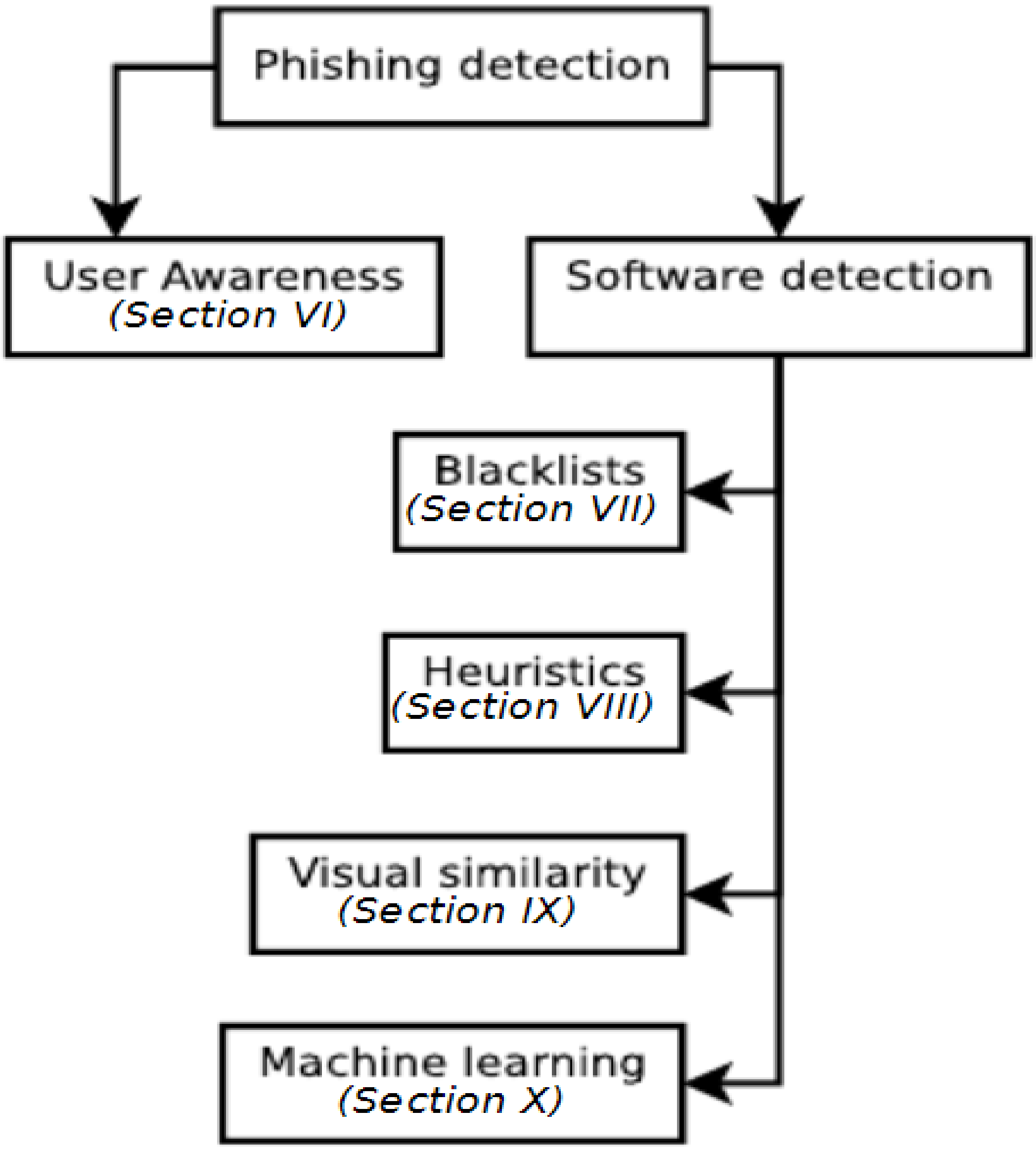
*•**E-mail messages** — could be sent from a variety of sources, such as: free E-mail Service Provider opens Simple Mail Transfer Protocol (SMTP) relays or infected end-user machines that are part of a botnet.*

Prevention Approaches:

The "prevention" of phishing attacks can be confusing, as it can mean different things depending on its context:

•Prevention of users from falling victim — in this case phishing detection techniques will also be considered prevention techniques. However, this is not the context we refer to when "prevention" is mentioned in this survey.

•Prevention of attackers from starting phishing campaigns— in this case, law suits and penalties against attackers by Law Enforcement Agencies (LEAs) are considered as prevention techniques



An overview of phishing detection approaches

CONCLUSION:

Phishing URL detection plays a pivotal role for many cybersecurity software and applications. In this paper, we researched and reviewed works based on the advanced machine learning techniques and approaches that promise a fresh approach in this domain. This article includes summary of the reviewed works after a systematic and comprehensive study on Phishing Website Detection systems. We believe that the presented survey would help researchers and developers with the insight of the progress achieved in the past years. Despite the tremendous progress in the field of cybersecurity, phishing website detection still pose a challenging problem with every evolving technology and techniques.s