PROPOSED SOLUTION

WEB PHISING DETECTION

PROBLEM STATEMENT:

Internet has dominated the world by dragging half of the world's population exponentially into the cyber world. With the booming of internet transactions, cybercrimes rapidly increased and with anonymity presented by the internet, Hackers attempt to trap the end-users through various forms such as phishing, SQL injection, malware, man-in-the-middle, domain name system tunnelling, ransomware, web trojan, and so on. Among all these attacks, phishing reports to be the most deceiving attack. Our main aim of this paper is classification of a phishing website with the aid of various machine learning techniques to achieve maximum accuracy and concise model.

IDEA/SOLUTION DESCRIPTION:

There are two approaches that are typically used in detecting phishing websites. The first approach is typically based on a blacklist, where in the given URL is compared with the URLs present in the blacklist. The other part of this approach is that the blacklist usually cannot identify all phishing sites, hence a new fraudulent website is launched. The alternate or the second approach is referred to as heuristic based methods, where few of the features are collected from the sites to distinguish it as either phishing or legitimate.

NOVELTY/UNIQUENESS:

In terms of accuracy, it was primarily due to the capability of the proposed PSO based feature weighting to successfully weight the website features used for enhancing phising website detection. In addition to the classification accuracy, we can have the TPR, TNR, FPR, FNR of machine learning classifiers before and after applying the proposed PSO based feature weighting.

SOCIAL IMPACT/CUSTOMER SATISFACTION:

An exhaustive systematic search was performed on all the indexing databases. The state-of-the-art research related to the web phishing detections was collected. The papers were classified based on the methodologies. A taxonomy was derived by performing a deep scan on the classified papers. The contributions listed in this survey are exhaustive and lists all the state-of-the-art development in this area.

BUSINESS MODEL (FINANCIAL BENEFIT):

Phishing attacks are categorized according to Phisher's mechanism for trapping alleged users. Several forms of these attacks are keyloggers, DNS toxicity, Etc., [2]. The initiation processes in social engineering include online blogs, short message services (SMS), social media platforms that use web 2.0 services, such as Facebook and Twitter, file-sharing services for peers, Voice over IP (VoIP) systems where the attackers use caller spoofing IDs [3, 4]. Each form of phishing has a little difference in how the process is carried out in order to defraud the unsuspecting consumer. E-mail phishing attacks occur when an attacker sends an e-mail with a link to potential users to direct them to phishing websites.

SCALABILITY OF SOLUTION:

The methods are evaluated in terms of learning rate, accuracy, and precision. It presents the learning rate of the methods during the training phase. The performance of three detectors during the training phase are similar. It is evident that the learning ability of methods are same. Authors maintained

similar parameters for all detectors. The learning rate of LURL is reasonable comparing to other two methods. It indicates that ML based methods able to scan an average of 84% of dataset to learn the environment at the rate of 1.0.