Literature Survey On The Selected Project & Information Gathering

In this activity you are expected to gather/collect the relevant information on project use case, refer the existing solutions, technical papers, research publications etc.

H. Huang et al., (2009) proposed the frameworks that distinguish the phishing utilizing page section similitude that breaks down universal resource locator tokens to create forecast preciseness phishing pages normally keep its CSS vogue like their objective pages.

S. Marchal et al., (2017) proposed this technique to differentiate Phishing website depends on the examination of authentic site server log knowledge. An application Off-the- Hook application or identification of phishing website. Free, displays a couple of outstanding properties together with high preciseness, whole autonomy, and nice language-freedom, speed of selection, flexibility to dynamic phish and flexibility to advancement in phishing ways.

Mustafa Aydin et al. proposed a classification algorithm for phishing website detection by extracting websites' URL features and analyzing subset based feature selection methods. It implements feature extraction and selection methods for the detection of phishing websites. The extracted features about the URL of the pages and composed feature matrix are categorized into five different analyses as Alpha- numeric Character Analysis, Keyword Analysis, Security Analysis, Domain Identity Analysis and Rank Based Analysis. Most of these features are the textual properties of the URL itself and others based on third parties services.

Samuel Marchal et al. presents PhishStorm, an automated phishing detection system that can analyze in real time any URL in order to identify potential phishing sites. Phish storm is proposed as an automated real-time URL phishingness rating system to protect users against phishing content. PhishStorm provides phishingness score for URL and can act as a Website reputation rating system.

Fadi Thabtah et al. experimentally compared large numbers of ML techniques on real phishing datasets and with respect to different metrics. The purpose of the comparison is to reveal the advantages and disadvantages of ML predictive models and to show their actual performance when it comes to phishing attacks. The experimental results show that Covering approach models are more appropriate as anti- phishing solutions.

Muhemmet Baykara et al. proposed an application which is known as Anti Phishing Simulator, it gives information about the detection problem of phishing and how to detect phishing emails. Spam emails are added to the database by Bayesian algorithm.

Wang et al., Jain and Gupta and Han et al.use white list-based method for the detection of suspected URL. Blacklist-based methods are widely used in openly available anti-phishing toolbars, such as Google safe browsing, which maintains a blacklist of URLs and provides warnings to users once a URL is considered as phishing. Prakash et al. proposed a technique to predict phishing URLs called Phishnet. In this technique, phishing URLs are identified from the existing blacklisted URLs using the directory structure, equivalent IP address, and brand name.

Felegyhazi et al. developed a method that compares the domain name and name server information of new suspicious URLs to the information of blacklisted URLs for the classification process. Sheng et al. demonstrated that a forged domain was added to the blacklist after a considerable amount of time, and approximately 50–80% of the forged domains were appended after the attack was carried out.