Abstract

Web Phishing is a major attack nowadays. Web phishing is phisher tries to get users sensitive information like bank details, ATM pin or any personal information. After extracting users information attacker may misuse this information. Nowadays most of the people prefer online shopping or online payment and user has to share his personal or sensitive information on web page. User may not think about website security. So website security is very important. Before doing any transaction and sharing any personal information on web page, user must ensure the security of that website. The best solution for this problem is to protect from phishing is to identify a phish. Phishing emails usually seem to come from well-known organization and ask
your personal information such as credit card number, security number, account number or passwords. What actually attacker does? The attacker creates the no of replicas of authentic sites, and users are forced to direct to that websites by attracting them with offers. As standard mentioned in W3C (World Wide Web Consortium), I am proposing a system which can easily recognize the difference between authentic site and phishing site. There are certain standards which are given by W3C (World Wide Web Consortium), based on these standards I am choosing some features which can easily describe the difference between legit site and phish site. To protect you from phishing, I am proposing a model to determine the fraud sites. To determine the phishing attack, URL features and HTML features of web page are considered. Clustering algorithm such as K-Means clustering is applied on the database and prediction techniques such as Naive Bayes Classifier is applied. By applying this, probability of the web site as valid Phish or Invalid Phish. To check the validity of URL, if still user is not able decide the validity of web page then Naive Bayes Classifier is applied.

References

- RSA Online Fraud Surveyor, "The phishing kit – the same wolf, just different sheep’s clothing", RSA Surveys, vol-1, February-2013.
- Angelo P. E. Rosiello, EnginKirda, Christopher Kruegel, FabrizioFerrandi, and Politecnico di Milano, "A Layout-Similarity-Based Approach for Detecting Phishing Pages", unpublished
- PHISHTANK. COM- The Online Valid Phish Sites Repository, http://data.phishtank.com/data/online valid.csv

Index Terms

Computer Science

Information Sciences
Keywords
Anti Phishing Technique  Bayesian Approach  Data Mining  Database Clustering  And Phishing Attacks.