Abstract

In these days there are two famous internet attacks these are malware and phishing. Malware stands for malicious software. It is designed to damage computer system without knowledge of the user. Phishing website is comparatively new internet crime to malware attack. Phishing is a form of online fraud such as social engineering schemes by sending e-mails, sudden message or online advertising attract users to phishing website that pretend to be trustworthy website in order to trick individuals sensitive information for illustration- financial accounts, password and personal identification numbers, which is used for profit. Malware and Phishing website is share same properties, firstly increasing at a rate of thousands per day and secondly phishing webpage represented by the term frequencies of the website content share comparable characteristic of malware samples represented through instruction frequencies of the program executable code. Past few years many techniques have been develop to detect malware and phishing website. In these techniques firstly extract feature from phishing website or malware and then categorize them into group. In this paper, we proposed Kernel k-means clustering to categorize malware and phishing website. Kernel k-means is advance version of the k-means algorithm. In which vectors are mapped from vector space to a higher dimensional feature space through kernel function and then k-means is applied in feature space. Thus kernel k-means avoids the separable clusters in vector space and improves the accuracy of phishing
website and malware categorization.

References

- Neda Abdelhamid, Aladdin Ayesh, Fadi Thabtah 2014 "Phishing detection based on Associative Classification data mining," 0957-4174/ Elsevier Ltd. All rights reserved.
- "Phishload" Available: http://www.Phishload-Tablesexplained. html Copyright (c) 2012Max-Emanuel Maurer (University of Munich).
- Han Jiawai and Kamber Micheline 2006 Data Mining concept and technique second ed. USA by Elsevier Inc.

**Index Terms**

Computer Science  
Web Service

**Keywords**  
Malware  
Phishing website  
Kernel k-means clustering algorithm.