International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 6 (2015) pp. 15017-15021

http://www.ripublication.com

An Effective Method of Phishing Detection Using IBC

Mr. N.V.Rajeesh Kumar[#1], S.Vinoth Kumar [*2]
[# 1] *Assitant Professor, Computer Science Department, Sathyabama University*
Chennai, India.
rajeesh555@gmail.com
[#2] *Student, Computer Science Department, Sathyabama University*
Chennai, India
vinothkumarbe16@gmail.com

Abstract

Phishing is one of the online thefts increasing nowadays due to the millions of internet user increased day to day. Previously several methods are used to cheat the victims to steal personal data, some of them are DNS Spoofing, Chat rooms and such. It's a serious threat for our economy for that we have to provide a solution. Some of the alternative methods are Blacklist, Whitelist and Hybrid methods. Even though phishing attack is continued increasingly, to rectify this issue we proposed a new mechanism called Email detection using classification with intelligent approach. In order to find out a phishing mails we are using an Intelligent Based Classification method (IBC). It also differentiates the phishing and non-phishing mails using some of the data set and features of the mails. The results show that the proposed method gives better results when compared to the previous methods. It has to be improved by given the strong rules for IBC. In future we can integrate with the machine learning algorithm for better results.

Keywords: Phishing, Classification, Internet Security, Emails, Anti-Phishing.

Introduction

Nowadays internet is an important factor for each and every one in the world. Through the internet we can communicate with people around the world with the social media networks, emails and such. Phishing is one of the crucial problems facing by the internet users nowadays and it will leads to global security issues for both the user and ISP. Phishing can be done by the following ways SMS, Forums, Calls and Emails. It will cause more financial damages, loss of personal information and identity theft will leads many problems. A statement from UK Cards Association's Press Release states that £21.6 million losses in 6 months in the year

2012 due to phishing attack and it will increase 28% compared to the previous year 2011 [1]. Also RSA statement estimates $687,000,000 loses due to phishing attack in the same period [2].

Normally, an intruder may send emails to the victims it seems to be original email from the organization. These mails look like the authentic one to update your information in the below URL. According to APWG report shows phishing attack aims only 5% targets for individual users, 30% targets for regional banks and 65% targets for the national banks [3]. Identity theft can happen due to the social engineering sites where phishing attack done and information can be stolen.

In this paper, Intelligent Based Classification (IBC) method is presented to reduce the phishing mails in real time analysis. In this method filtering and classification algorithm is used to detect phishing mails. First, basic types of filtering methods are used to classify the mails then Classification algorithm is used to detect the phishing mails by the features of mails.

Related Works

Phishing life cycle gives the complete steps of how the intruder can steal the victim personal information through phishing. First the intruder select the list of person to whom he want to send an emails which seems to be from the authentic organization then he create a phishing emails and send to the list of person. If the person click link in the emails it ask for the user information.

Black List Method

here a list of phishing URL is added in the database, every time when the user click the URL it compared to the database if it is not present then it allow to load the page otherwise it block the URL. Disadvantages of Black List are it will take more time to add a newly created phishing webpage in the database.

Heuristic Method

here a combination of both the whitelist and blacklist approach is used to detect the phishing mails in real time. This method will detect the newly created phishing web sites.

Visual similarity approach

it will compare the website features if it matches with the other websites then it report to the users.

Non-Technical Approach

a) **Education**: Appropriate organization has to educate the consumers regarding the phishing attacks then it will reduce the phishing [4].
b) **Legal Awareness**: where the enact laws punish the phishers for the theft of personal information and financial loses for both the users and ISP [5].

System Overview

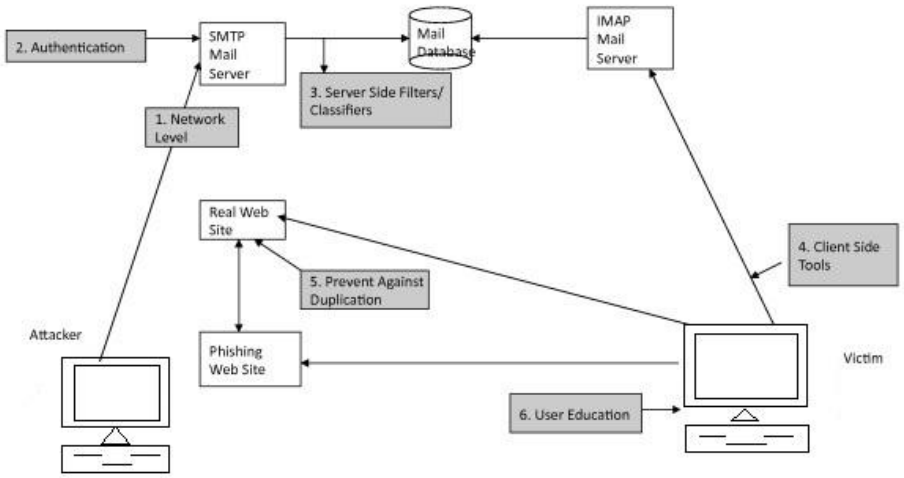

Figure 1: Architecture diagram of Phishing attack

Fig 1. Shows overall architecture of phishing attack and how attacker stolen the information from the victim through phishing emails. IBC algorithm is used to detect the phishing emails then it separate the phishing and non-phishing mails using filtering approach.

Methods

Data Sets

Dataset for this study include number of user's accounts with an e-mail or a message from each account in the form of URLs. Two kinds of data sets are used, one is "phishtank" and other one is "Phishing Archive" Since larger data sets may give inaccurate results, here we have chosen smaller data sets to facilitate more and thorough analysis for a stand alone application operated in offline mode. To have more conclusive results, we elected to use smaller data sets that could be manually created. Content based features were used in the previous studies such as hash tags that uses "@" and "#" respectively.

Criteria	N	Phishing Indicators
URL & Domain Identity	1	Using IP address
	2	Abnormal request URL
	3	Abnormal URL of anchor
	4	Abnormal DNS record
	5	Abnormal URL
Social Human Factor	1	Emphasis on security
	2	Public generic salutation
	3	Buying time to access accounts
Web Address Bar	1	Long URL address
	2	Replacing similar char for URL
	3	Adding a prefix or suffix
	4	Using the @ Symbol to confuse
	5	Using hexadecimal char codes
Page Style & Contents	1	Spelling errors
	2	Copying website
	3	Using forms with *Submit* button
	4	Using pop-ups windows

Figure 2: Based upon the criteria phishing will be classified

Feature Extraction and Classification

(Horng et al., 2011) concluded that these steps will solve the phishing problems. They are Identification of the required data, Training set information, Determination of the input feature, Applying the classification algorithm and Classifier evaluation [6]. (James, 2005) discuss the general non technical methods for phishing activities. They are Legal solutions, Education [7].

In these feature extraction process first it extract features from the phishing e-mails. Depending up on the output, the new selection step is performed to select the subset of applicable features. Using the learning model the most informative features are selected. In classification the data set is divided into two. They are training and testing ratios. Profiling stage is used to train classification algorithm and the training data is used to create profiles.

Intelligent Based Classification

IBC works according to 3 rules, first is discover and generating rules, second is classifier built and last is prediction. In rule one frequently used rule are discovered and generate the complete rules as IBC. According to threshold values and confidence, the rules are ranked as per values. Then the rules are filtered and pruned to remove the unwanted rules, after that remaining rules are predicted as final classifier. The training data set is given to the training test cases for predicting the phishing emails.

As per the data sets collected from the various sources are taken as training data and test cases are done to find the phishing mails, where filtering approach will separate the phishing and non-phishing mails.

Conclusion

Nowadays detecting the phishing emails is one of the very big challenges faced by internet community. In this paper, we proposed an Intelligent Based Classification (IBC) algorithm to detect the phishing emails by using various features present. Here the phishing mails can be detected by the combination of filtering and classification approach. IBC algorithm is based on classification in addition it uses filtering approach. IBC algorithm can detect up to 96% of phishing attacks in real time. Classification algorithm detects phishing attacks and it protects the users from the unreliable links, Instant messages and web pages. Our future scope is to integrate with the machine learning algorithm for better results.

References

[1] Financial Fraud Action UK, Cheque & Credit clearing Company, UKCARDS Association. (2012). Deception crimes drive small increase in cad fraud and online banking fraud losses. Press Release, (pp. 2). [online] <www.financialfraudaction.org.uk>, <www.chequeandcredit.co.uk>, <www.banksafeonline.org.uk>. Accessed 10.11.2012.

[2] RSA Anti-Fraud Command Center. RSA monthly online fraud report. http://www.rsa.com/solutions/consumer_authentication/intelreport/11752_ Online_Fraud_report 0712. pdf?M¼50f5edbd-cabe-4502-bd6c-eab4c4952 bc3; July, 2012.

[3] Anti-Phishing Working Group. Global phishing survey, trends and domain name use in 2H2011. http://www.antiphishing.org/reports/ APWG_GlobalPhishingSurvey_2H2011.pdf; April 2012.

[4] Kumaraguru Ponnurangam, Sheng Steve, Acquisti Alessandro, Cranor Lorrie Faith, Hong Jason. Teaching Johnny not to fall for phish. Article 7 ACM Transactions on Internet Technology May 2010;10(2):31. http://dx.doi.org/10.1145/1754393.1754396, http://doi.acm.org/10.1145/1754393.1754396.

[5] Symantec Global Intelligence Network. State of spam and phishing report. http://eval.symantec.com/mktginfo/enterprise/other_resources/b-state_of_spam_and_phishing_report_02-2010.en-us.pdf; February, 2010. visited on June 2012.

[6] Horng, S. J., Fan, P., Khan, M. K., Run, R. S., Lai, J. L., & Chen, R. J. (2011). An efficient phishing webpage detector. Expert Systems with Applications: An International Journal, 38(10), 12018–12027.

[7] James, L. (2005). Phishing Exposed. s.l.: Syngress Publishing.