International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 4 Issue: 1

ISSN: 2454-4248
187 – 192

A Survey of Website Phishing Detection Techniques

Zinal Shukla
Research Scholar
Marwadi Education Foundation
Rajkot, India
zinalshukla@gmail.com

Kirtirajsinh Zala
Assistant Professor
Marwadi Education Foundation
Rajkot, India
kirtirajsinh.zala@marwadieducation.edu.in

Riddhi Kotak
Assistant Professor
Marwadi Education Foundation
Rajkot, India
riddhi.kotak@marwadieducation.edu.in

***Abstract*—**This article surveys the literature on website phishing detection. Web Phishing lures the user to interact with the fake website. The main objective of this attack is to steal the sensitive information from the user. The attacker creates similar website that looks like original website. It allows attacker to obtain sensitive information such as username, password, credit card details etc. This paper aims to survey many of the recently proposed website phishing detection techniques. A high-level overview of various types of phishing detection techniques is also presented.

Keywords*-Phishing Attack; URL; Phishing Detection; Phishing Website; Anti-Phishing; Social Engineering*

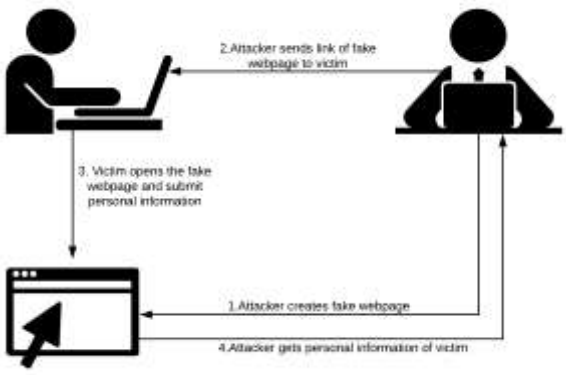
*****

I. INTRODUCTION

Phishing attack is one of social engineering attack which is common security threat used to obtain private and confidential information by simply tricking the users without being detected. The main purpose of attacker is to gain personal information such as username, passwords, credit card details, account numbers etc from the users and to use this information for illegal financial gain. Website phishing is one of the Internet crimes that are growing very frequently.

Various anti-phishing techniques have been developed to combat the problem but these approaches suffer high false positive rates. Studies which use URL address, domain name information, website ranking etc as the features of the webpage always lead to lower recognition rates; Heuristic and machine learning methods which use features that contain the text and the image of the WebPages have been introduced to phishing detection but most of them have high complexity and high false positive rates [4].Moreover, no comprehensive feature which are total representative of phishing strategies have been proposed.

Typically phishing attack exploits the social engineering to lure the victim through sending a spoofed link by redirecting the victim to a fake web page [1]. The spoofed link is kept on popular webpages or sent via Email to the victim. Thus, rather than directing the victim request to the real web server, it will be directed to the attacker server [1]. Figure 1 show steps which are performed in website phishing attack.



Computer security attacks can be classified into three types such as; physical attack, syntactic attack, and semantic attack. Attack against physical infrastructure such as computer parts or memory devices and wires are example of physical attacks. The syntactic attacks targets the operating logic of computers and networks such as software vulnerabilities, cryptographic algorithm vulnerabilities etc and in semantic attacks, targeted at people vulnerabilities such as how to interpret computer messages. Phishing is one of the semantic attacks. People are most likely to believe information they read, without trying to validate the credibility of the information [9], makes false information easier to spread and reach a large number of people with very low cost.

This survey begins by defining the phishing problem in section II, in section III this survey will focus on overview of phishing detection techniques after that in section IV, different methodologies for phishing website detection is discussed and in section V, summaries for different measures used in phishing detection and classification is provided and in last section VI, conclusion for current limitation and future work is discussed.

II. BACKGROUND AND OVERVIEW OF PHISHING ATTACK

Website phishing can be explained as impersonating original website to obtain personal information of user. Internet is accessed for entertainment, shopping, business, financial transaction almost for everything. Due to excessive use of internet, the internet crimes are growing rapidly. Attacker obtains sensitive information of user such as username, password, banking information etc.

Most of the users depend on appearance of the website for its identification. Therefore an attacker creates website that mimic the original website to deceive the user. These fake websites have very much similar visuality to the legitimate website in order to defraud innocent internet user. Typically phishing attack exploits the social engineering to lure the victim through sending a spoofed link by redirecting the victim to a fake web page and the spoofed link is placed on the popular webpages or sent via email to victim [1].

A. Phishing motives

According to weider D. et al. [11], the basic motives of attacker behind phishing attacks are:

 ISSN: 2454-4248
187 – 192

- *Financial gain:* Attacker can steal banking credentials for their financial gain.
- *Identity hiding:* Instead of using stole identities directly, attacker may sell identities to others whom might be criminal seeking ways to hide their identities and activity.
- *Fame and notoriety:* Attacker can attack or the sake of peer recognition.

B. Various types of phishing attack

- *Website phishing attack:* Website phishing is a type of social engineering attack; it is often used to steal user's data including login credentials and credit card numbers etc. Website phishing typically begins by creating website that imitates legitimate website due to the internet users believes on appearance of the website for its identification. This website looks similar to original website which is capable enough to deceive the users so with the use of this attack, attacker aims to obtain sensitive information of user.
- *Email phishing attack:* Email phishing attackers use spam, fake websites which are created to look similar to original websites and emails to trick you into divulging personal information such as bank account passwords and credit card numbers. Attacker sends email to users about the need to verify account information, system failure requiring users to re-enter their information, fictitious account charges, undesirable account changes, new free services requiring quick action and many other scams are sent to large number of people with the hope that the unwary will respond by clicking link to or signing on to bogus websites, Once you enter your information, they can use it to create fake accounts in your name, ruin your credit, and steal your money or identity.
- *Web trojans:* This attacks pop ups invisibly when users are attempt to login in for trusted website. Attacker collects the user's credentials locally and transmits them to the phishers.
- *Content-injection phishing:* In this attack, attacker replaces part of content of a legitimate site with fake content designed to misdirect to user into giving up their personal information to attacker. For example, attacker may insert malicious code to log user's credentials or an overlay which can secretly collect information and deliver it to the attacker's phishing server.

C. Phases of website phishing attack

Website phishing basically works on four phases by which attacker attempts to steal user's personal information. These phases are based on survey of website phishing attack in real time. Phases for website phishing attack are as per mentioned below:

- *Phase 1:* Attacker creates a fake website that looks like original website.
- *Phase 2:* Attacker sends link of fake website to the victim via Email.
- *Phase 3:* While clicking on link victim will be redirected to fake website and submits sensitive information.
- *Phase 4:* Attacker gets sensitive information of user via fake website.

D. Different forms of phishing attack

Due to the broad nature of the phishing problem, it is important to determine forms of website phishing attacks. Phishing attack can take place by several forms which are mentioned below:

- *Creating fake URL*: Attackers usually try to make the Internet address (URL) of phishing sites look similar to legitimate sites to misguide internet users [10]. They cannot use the original URL of the legitimate website so they make fake URL and send it to users via email so user get redirected to fake website.
- *Misspelled URLs*: In this type of form attackers make more spelling mistakes. For example, the URL www.applle.com looks similar to well known website www.apple.com, or http://www.apple.attack.com if user are not careful, they will think that they are on "apple" site [10].
- *Creating anchor text*: This type of form is Similar to the URL feature, but here the links within the webpage may point to a domain different from the domain which is typed in the URL address bar [3].
- *Fake SSL lock*: Now a day it is cheap and easy enough for the attacker to obtain SSL certificates for their malicious sites, therefore users lose one of the methods for identifying trusted sites from phishing targets.
- *URL manipulating using java script*: The most commonly abused is URL manipulating using java script [9]. The attacker will insert a string to be used in the webpage and treated by the user's browser as code and when the browser loads the page, the malicious script executes without the user even knowing that such an attack has taken place.

E. Challenges

There are number of challenges come across when we deal with phishing website detection, some of them can be negligible but some of them are really harsh to ignore so they are taken into consideration as:

- *Link manipulation:* This is one of the most common of all website phishing attacks which works as its name suggests directs user's browser to website which is different from the legitimate website. Link manipulation often comes in the form of an email message for which users thinks its trustworthy website.
- *Website Forgery:* In this attackers imitate the address bar logos of original website and put them beside URLs of their deceiving website.
- *Automatic integrated anti-phishing tool*: This means that user does not need to take additional efforts to integrate and customize tools as per new requirement or updated features to detect fake website.
- *Determining an appropriate threshold*: The threshold is the matching score between original and fake website. If fake website is partially copied from the legitimate website, then none of the visual similarity based approach can detect fake website.

 ISSN: 2454-4248
187 – 192

- *Embedded objects:* The embedded objects available in the website as attacker use images, java script etc., to overcome the anti phishing system.
- I*nsufficient methods using blacklist*: Data set for fake website is provided by number of online dataset sites which can be access through internet but there is no sufficient method which is applicable to detect phished website using online data set.
- *Global page rank*: Numbers of page rank algorithms are available in different sites and tools in internet, however instead of multiple page rank algorithm there must be one global page algorithm.

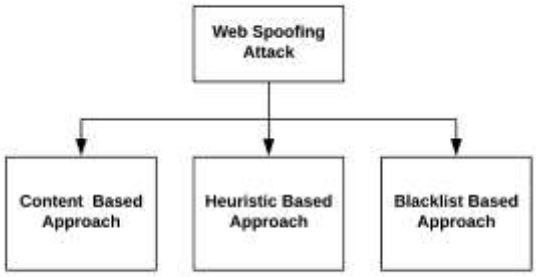
III. Detection of Phishing Attack : An Overview

This section reviews the most related work of website phishing detection. Numbers of researches on website phishing detection have been done for the past few years. Numbers of methods have been created to study web phishing. In this, we consider any anti-phishing solution which aims to classify or identify phishing attack as detection which includes:

A. Web phishing attack detection approaches

Anti-phishing methods of website phishing are classified in various approaches: content based, heuristic based and blacklist based approaches [1][2][3], as shown in figure 2.

- *Content based approach:* This approach detects fake website by inspecting the similarity between legitimate and fake website. The similarity between two website is calculated based on the similarity of web page content. Moreover this approach has more accuracy and low false alarm in detecting fake websites. CANTINA and Goldphish are one of the content based solutions. One research which belongs to this approach is conducted by CANTINA [12]. This research detects fake website by using Term Frequency/Inverse Document Frequency (TF-IDF) with the help of this technique false positive rate is reduced. In another research Goldphish is used which uses Google as a search engine. Goldphish algorithm relies on capturing an image for the current website in the user's web browser. Then the captured image is converted into computer readable text using an optical character recognition technique. The converted text in this technique is used as an input into a search engine for analyzing the page rank and identifying the possible website phishing attack.
- *Heuristic based approach:* Heuristic based approach uses HTML or URL signature to identify fake website. There are different researches conducted based on this approach. SpoofGuard is one of the heuristic based solutions [14]. It is an anti-phishing browser plug-ins. This solution uses a combination of stateless page evaluation and state full evaluation also examination of outgoing post data to compute spoof index value. If the computed spoof index is greater than pre-defined threshold value, the page is classified as fake website and the user is notified about this page. If spoof index is less than threshold value, the page is classified as legitimate website.
- *Blacklist based approach*: This approach has an updated blacklist contains all website that are denied access. Therefore user is prevented from accessing website that appears in blacklist. Most important part of this approach is retrieving URLs from phishing pages in order to create and maintain blacklist [1]. The URLs can be obtained from the users phishing emails, spam or from the organization which serves anti-phishing such as Anti Phishing Working Group (APWG) and Phishtank. Netcraft tool bar is an anti-phishing solution which belongs to blacklist based approach. It detects phishing attack based on few criteria such as time of sitting the Netcraft web server survey, times of visiting the website, country that hosted the website, name of organization that hosting the current website and risk rating.



B. Survey of website phishing detection

There are many researches going on to detect phishing website and distinguish between original and phished website. This survey will compare number of detection techniques. Abdulghani et al. has proposed detection technique of phishing website based on checking Uniform Resources Locators (URLs) of web pages. The proposed solution is able to distinguish between the original website and fake website by checking URLs of suspected web pages [1]. S. kaur et al. has provided approach which utilizes ten features of the webpage and proposes a fitness function for categorizing the webpage. Genetic algorithm is utilized for assigning weights to features [2]. Neda Abdelhamid et al. has investigated the problem of website phishing using a developed AC method called Multi label Classifier based Associative Classification (MCAC) to seek its applicability to the phishing problem[3]. Weiwei et al. have given a principled ensemble classification algorithm to combine the predicted results from different phishing detection classifier. Hierarchical clustering technique has been employed for automatic phishing categorization [4]. Yang et al. has provided a heuristic anti-phishing technology against this increasingly prominent security issues. It will perform six checks in two rounds based on the key elements of the page, including domain name URL, the password field, pictures, links, etc. [5]. P.A. Barraclough et al. has introduces idea to utilize a Neuro-Fuzzy scheme with 5 inputs to detect phishing websites with high accuracy in real time. In this research 2-Fold-cross-validation is applied for training and testing the proposed model [6]. V.Shreeram et al. have proposed an approach to detect phishing hyperlinks using rule based system formed by genetic algorithm. This can be utilized as a part of an enterprise solution to anti-phishing. In this approach, genetic algorithm is used to evolve rules that are used to differentiate phishing link from original link [7]. Ee Hung et al. have proposed an anti-phishing method to protect internet users from the website phishing attack. The proposed method will render a screenshot of the webpage and segment the region of interest which consists of website logo after that Google image

 ISSN: 2454-4248
187 – 192

database is utilized to identify the website identity based on segmented website logo. During the identification process, they employed the content based image retrieval mechanism which is provided in Google search image search engine to locate the most similar logo from Google image database. The results will differentiate original website and phishing website [8]. Mingxing He et al. have presented heuristic method to determine whether a website is legitimate or phishing website. This study converts a webpage into 12 features which are selected based on the existing fake and original WebPages. A training set of WebPages including real and fake phishing pages are then input for a support vector machine to do training. After that testing set is fed into the trained model to do testing [9]. Luong Anh et al. propose a phishing detection approach based on the features of URL. The proposed method focuses on the similarity of phishing site's URL and legitimate site's URL [10].

IV. TECHNIQUES FOR WEBSITE PHISHING DETECTION

Since phishing attacks attempt to take advantage of the unaware users, an obvious solution is educating the users, which would in turn reduce risk of phishing attack. Some features of website can be used to differentiate between legitimate website and spoofed website. These features are many such as URLs, domain identity, security and encryption, source code, page style and contents, web address bar and social human factor [1]. This section will present and discuss various techniques of websites phishing detection as follows:

A. Methodologies for website phishing detection

1) Based on genetic algorithm

In this study [2], Phishing detection consists of four phases which are given as per below:

- *Phase 1 Feature Extraction*: In this phase 10 features are extracted in this paper.
- *Phase 2 Pre-processing*: During this phase the value of each feature is classified into Phishing, legitimate or suspicious class.
- *Phase 3 Weight adjustment*: The aim of the weight adjustment is to find the best weights that can classify the website accurately and genetic algorithm is used for weight adjustment.
- *Phase 4 Results*: In this phase best weights derived in third phase are used to calculate the fitness of the URLs in data set and classified into legitimate or phishy by comparing the fitness with the threshold value.

Information gain of each feature is taken as its initial weight. The information gain is derived from the training dataset. The expected information which are important to classify training data of s samples, where the class attributes has n values { V_1, $V_2 \ldots V_n$} and S_i is the number of samples which belong to class label Vi, is given by,

$$I(S_1, S_2 \ldots . S_n) = \sum_{i=1}^{n} p_i \log_2 p_i$$

Where $P_i = s_i/s$.

If A is an attribute with values $\{v_1, v_2 \ldots . v_m\}$, then A partitions the samples into subset S_1, $S_2, \ldots S_m$ where samples in each S_j have a value of V_j for attribute A. The entropy associated with A is:

$$E(A) = \sum_{j=1}^{m} \frac{S_{1j} + S_{2j} + S_{nj}}{S} * I(S_{ij} \ldots \ldots, S_{nj})$$

Where S is the number of samples in Sj which belongs to class i. I ($S_{1j} \ldots \ldots, S_{nj}$) can be defined using the formulation for I($S_1, S_2 \ldots .., S_n$) with p_i being replaced by pit where $p_{i\,j} = S_{ij}/S_j$.

This technique consist of population of size five where each chromosome shows weights for ten features and performs selection, crossover and mutation until better results are obtained.

- *Selection:* In this one of the chromosomes is selected for performing crossover and mutation. Roulette wheel method is used for selection where the fitness of each chromosome is 1/error and error is the no. of websites misclassified by using the chromosome as the weight of features.
- *Crossover:* Crossover is performed for merging the genetic information of two individuals. This operation is performed on the selected chromosomes and the previous weights to produce new weights.
- *Mutation:* This performs the changes in the new weights randomly. So, value m is added to each element of the new weights and m is generated randomly.

Here, mutation is performed with mutation probability (Mp) 0.2. A random number is generated where, if random number is less than Mp, then mutation is performed.

This process of selection, crossover and mutation continues until the better weights are obtained and genetic algorithm is executed multiple times and then the best weights are selected.

2) Based on associative classification data mining

In associative classification the training phase is about inducing hidden rules using association rules and then classifier is generated after pruning useless and redundant rules. Generally an AC algorithm is performed in three phases. In phase 1, it searches for hidden correlations among the attribute in the training data set and generates them as "Class Association Rule" in "If-Then" format. In phase 2, ranking and pruning procedures start operating thresholds like confidence and support. The output of phase 2 is the set of Class Association Rule which represents the classifier. Finally in phase 3, the classifier derived gets evaluated on the test data to measure its effectiveness in forecasting the class of test data and the output of the last phase is the accuracy or error-rate of the classifier.

In this study [3], the phishing detection process using associative classification from the user side can be explained in the following steps:

- 1. When the end user clicks on a link within an email or browses the internet.
- 2. User will be directed to a website that could be original or fake. Therefore this website is basically the test data.
- 3. A script which is written in PHP is embedded within the browser and starts processing to extract the

 ISSN: 2454-4248
187 – 192

features of the test data and saves them in a data structure.

- 4. Then the intelligent model will be active within the browser to predict the type of the website based on rules learnt from previous websites and the rules of the classifier are utilized to predict the type of the test data based on features similarity.
- 5. When the browsed website is identified as original, no action will be taken. But, when the website turned to be fake, then user will be warned by the intelligent method that he is under risk.

3) Intelligent Phishing website detection and categorization model

In this study [4], the page feature representation method is studied and heterogeneous classifier is built based on these different features and ensemble classification algorithm is proposed to combine all predicted results from heterogeneous classifier, then the hierarchical clustering technique has been adopted for automatic phishing categorization.

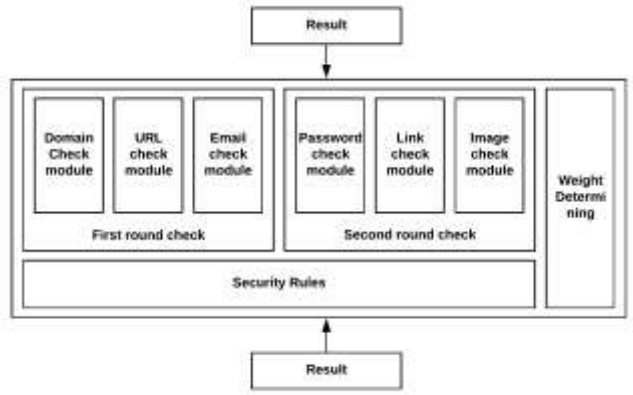
Main component of IPDCM (intelligent phishing website detection and categorization model) is described as per given below:

- *Feature extractor*: feature extractor is used to extract the terms from the webpages and then converts the terms to a group of 32-bit global IDs as the feature of the data collection after that, for training samples these integer vectors are transformed into term frequency features and collected in the database.
- *Classifier training module*: In this, ten heterogeneous classifiers are built based on the characteristic of different features; improved NBC (Naive Bayes Classifier) and SVM (Support Vector Machine) algorithm are employed for the training.
- *Ensemble classification module*: Ensemble classification method is used to combine all the prediction results from heterogeneous classifiers. This has better detection performance than individual classifier.
- *Cluster training module*: Hierarchical clustering algorithm is applied on the term frequency vectors with the TF-IDF weighting scheme.

4) Based on heuristics anti-phishing detection

In this study [5] a method is based on the study of page elements which are possible to be spoofed or can be imitated by an attacker. Which includes, domain name check module, URL checking module, Email checking module, link checking module, the password checking module, picture check module a total 6 modules are checked. Figure 4 shows heuristic based anti-phishing model.

- *Domain check module:* This module will compare the domain name which user is trying to navigate by, with the domain names that are stored by user's browser. If they have certain similarity then system will give warning to user.
- *URL check module:* This module includes three steps,1 Check whether the URL that will be navigated to contains suspicious username; 2. Check whether the host name or domain name in the URL didn't been hidden. For example, the URL contains no suspicious fields like 'www.' or '.com'. 3. Check the page which will being navigated to requested from a standard port.



- *Email check module:* This module checks whether the directed links link to email address, whether the current email domain name is empty also whether an email domain name is from a known website.
- *Password check module:* This module checks whether current page contains fields such as 'password' or 'pass' or 'pwd', if the page contains these fields and the fields hasn't been encrypted the system will give a warning to user.
- *Link check module:* This module checks whether current pages contains suspicious links and suspicious link refers to a link which triggered a warning when it was passing domain check and URL check.
- *Image check module:* This module will compare images in current page with images from pages which are accessed before and compute their hash values. If an image in current page is having same value from one image accessed before, the system will give a warning.

At the first step, the results of the first round of check will compared with the threshold which is set at first, if the result is more than the threshold value then it is considered as fake webpage, navigation canceled, enter the second round of check if it is less than the threshold.

At the second step, the sum of the all above weights is calculated. If the result is more than the alarm threshold value then considered to be a fake webpage, navigation canceled, considered to be a secure page and navigation is continued if the sum is less than the threshold value.

5) Based on Neuro-Fuzzy algorithm

Neuro-Fuzzy is a combination of a fuzzy logic and a neural network with ability of reasoning and learning. Neuro-Fuzzy has abilities of data learning from neural network view point, and forms linguistic rules from fuzzy inference of view which allowing the power of intelligent systems to be used.

In this study [6], five inputs which are five tables where features are extracted and stored for reference and these includes: Legitimate site rules: legitimate site rule is a summary of law which covers phishing laws, User behavior profile: It is list of people's behavior when interacting with legitimate and phishing website and Phishtank: it is a free community website operated by open domain names where suspicious websites are verified and voted as phished by the community experts or user specific side. These five inputs are taken into consideration because they are wholly representative of phishing attack

 ISSN: 2454-4248
187 – 192

techniques and strategies. Among the five inputs, 288 features are extracted which are used as training and testing input data into the Neuro-Fuzzy system for generating Fuzzy IF…THEN rules, and for discriminate between phishing, suspicious and legitimate websites.

6) Based on SVM classifier

Support vector machine is supervised machine learning algorithm which can be used for classification and regression. In this classifier each data item is plot as a point in n-dimension space with the value of each feature being the value of a particular coordinate then it performs classification with finding the hyper plane that differentiate the two classics very well.

In this study [9], first a given webpage is parsed into a DOM (Document Object Model) tree to allow easier processing for further step. DOM which is a World Wide Web Consortium (W3C) standard is a platform and language neutral interface that will allow programs and scripts to dynamically access and update the content, structure and style of document. Therefore after the DOM tree is constructed they will check whether a page contains any text inputs, since a fake page always requires users to input credentials. If a page has at least one text input then it will proceed to next step otherwise the detection is not required since users do not have a way to enter their secret information. After parsing a webpage and conforming that a webpage has at least 1 text input, the identity extraction process would extract the identity set of the parsed DOM tree. After that the webpage DOM tree, its HTTP transaction, and its identity set are then input into the feature generation step to generate features in SVM format. Finally SVM classifier would determine whether the site is phishing or legitimate.

V. SUMMARIES

TABLE I. MEASURES USED FOR DETECTION AND CLASSIFICATION

No	Used Measures		
	Measures	***Formula***	***Meaning***
1.	Accuracy	$=\left(\frac{TF+TN}{TP+TN+FP+FN}\right)\times 100$	Used to calculate accuracy of attack detection [1][8].
2.	False negative	$=\left(\frac{FN}{FN+TN}\right)\times 100\%$	Used to calculate false alarm rate[1]
3.	Fitness	$=x_1*\sum_{i=1}^{n} w_i v_i + x_2*\sum_{i=n+1}^{10} w_i v_i$	Used to calculate the fitness of the website [2].
4.	Precision	$=\frac{TP}{TP+TF}$	Used to evaluate predicted results [4][8].
5.	Recall	$=\frac{TP}{TP+FN}$	Used to evaluate recall rate[4].
6.	Root mean square error	$=\sqrt{\frac{\sum(A_i-D_i)^2}{N}}$	It is a measurement of accuracy [10].
7.	Term frequency	$=\sqrt{\frac{n_{ij}}{\sum_k n_{kj}}}$	The term frequency tf_{ij} value of term t_i in document d_j [9].
8.	Inverse document frequency	$=\ln\left(\frac{\lvert D\rvert}{\lvert(d_j : t_i \in d_j\rvert}\right)+1$	The inverse document frequency is calculated [9].

Figure 1. Example of a ONE-COLUMN figure caption.

VI. CONCLUSION

Phishing attack is currently among the most problematic of trends in cyber crimes. It is a means of obtaining confidential information through fake website which appear to be legitimate.

In this paper, a survey of the protection against these phishing attacks is presented. This survey improves the understanding of the website phishing attack, the current scope of solution and the future scope to detect phishing attack. Approaches given in literature survey still have much limitations regarding accuracy and performance. Many algorithm have been used but still there is no standard technique with high accuracy. In future there is scope of using machine learning algorithm for achieving higher accuracy and efficiency.

REFERENCES

[1] A. Ahmed, N. Abdullah, " Real time detection of phishing website", IEEE 7th Annual Communication Conference (IEMCON), Vancouver, BC, Canada, 2016.
[2] S. Kaur, A. Kaur, "Detection of phishing webpages using weights computed through genetic algorithm", IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE), Amritsar, India, 2015.
[3] N. Abdelhamid, A. Ayesh, F. Thabtah, "Phishing detection based associative classification data mining", Expert Systems with Applications, vol. 41, no. 13, pp. 5948-5959, 2014.
[4] W. Zhuang, Q. Jiang, T. Xiong, "An intelligent anti-phishing strategy model for phishing website detection", 32nd International Conference on Distributed Computing Systems Workshops, Macau, China, 2012.
[5] Y. Liu, M. Zhang, "Financial websites oriented heuristic anti-phishing research", IEEE 2nd International Conference on Cloud Computing and Intelligence Systems, Hangzhou, China, 2012
[6] P. Barraclough, M. Hossain, M. Tahir, G. Sexton, N. Aslam, "Intelligent phishing detection and protection scheme for online transactions", Expert Systems with Applications, vol. 40, no. 11, pp. 4697-4706, 2013
[7] V. Shreeram, M. Suban, P. Shanthi and K. Manjula, "Anti-phishing detection of phishing attacks using genetic algorithm", International Conference On Communication Control And Computing Technologies, Ramanathapuram, India,2010
[8] E. Chang, K. Chiew, S. Sze, W. Tiong, " Phishing detection via identification of website identity", International Conference on IT Convergence and Security (ICITCS), Macau, 2013.
[9] M. He, S. Horng, P. Fan, M. Khan, R. Run, J. Lai, R. Chen and A. Sutanto, "An efficient phishing webpage detector", Expert Systems with Applications, vol. 38, no. 10, pp. 12018-12027, 2011
[10] L. nguyen, B. To and H. Nguyen, " A novel approach for phishing detection using URL-based heuristic", International Conference on Compute, Management and Telecommunications (ComManTel), Da Nang, Vietnam, 2014
[11] W. Yu, S. Nargundkar, N. Tiruthani, " A phishing vulnerability analysis of web based system", 2008 IEEE Symposium on Computers and Communications, Marrakech, Morocco, 2008
[12] Y. Zhang, J. Hong and L. Cranor, "Cantina", Proceedings of the 16th international conference on World Wide Web – WWW '07, Alberta, Canada, 2007
[13] Priyanka.M, Prasanna Kumar K.R, "Real Time and Continuous Detection of Phishing Websites", International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), ISSN: 2321-8169,Vol. 5, No. 5, PP: 923 – 928,2007
[14] Chou, Neil & Ledesama, Robert & Teraguchi, Yuka & C. Mitchell, John, "Client-Side defense against web-based identity theft", Conference of the Network and Distributed System Security Symposium(NDSS) San Diego, California, USA, 2004