International Journal of Technical Research and Applications e-ISSN: 2320-8163,
www.ijtra.com Volume 3, Issue 2 (Mar-Apr 2015), PP. 45-51

A STUDY OF MAJOR PHISHING TARGETS AND THEIR ANTI-PHISHING SOLUTIONS

[1]Rani S. K, [2]D. Kavitha M. E.,
[1]PG Student, [2]Assistant Professor
Department of Computer Science and Engineering,
Valliammai Engineering College,
Chennai, India
[1]rsk1411@gmail.com, [2]kavi_pranav1@yahoo.co.in

***Abstract*— Motivation of all the types of Phishing attack is to acquire personal information from users without the user knowledge. Attacker uses different medium to perform phishing on user environment. These types of phishing attack increasing day by day in different forms. Several losses faced by people in all over the world. One needs to put extra care to overcome this fraudulent activity in the initial stage itself. This survey focuses on different types of phishing environments and the major targets of phishing such as Emails, Websites and Mobile Phones. The anti-phishing solution for such phishing targets protect your business from the earlier stages of phishing attack and the over-all study of different anti-phishing techniques will help the user to choose appropriate anti-phishing technique with respect to the nature of the environment the phishing takes place.**

***Keywords*— Authentication Based Anti-Phishing, Email Anti-Phishing, Mobile Anti-Phishing, Phishing Target, Website Anti-Phishing.**

I. INTRODUCTION

Phishing will direct the user to suspicious link to obtain personal information such as username, password, Bank account number, and credit card details and it pretends as a trustworthy entity. Phishing also tends to spread on Voice over IP, Multiplayer online games, and Social media [1]. However, the effectiveness of the attack increasing day by day, several anti-phishing techniques was introduced by researchers to overcome this attack. According to APWG second quarter Phishing trends report [2] of 2014, there are 128,378 phishing sites detected in second quarter and it is comparatively greater than first quarter report. The most targeted service is online payment services and crypto-currency sites. Also there is a drastic increase in potentially unwanted programs such as spyware and adware and United States is considered to be one of the top most Country hosting phishing sites.

II. TYPES OF PHISHING ATTACKS

A. Session Hijacking

Session Hijacking is also known as cookie hijacking. It is a security attack on user sessions. IP spoofing and man in the middle attack is the most common method of session hijacking. The http communication needs different TCP connections and the web server uses token method to recognize those connections. Generally those tokens are string of variable like URL or other parts of http header or body.

B. Key loggers

Key loggers is a type of monitoring software that monitors the behavior of people through the actions they perform using keyboard. It records the information typed on the keyboard, stores those information in a specific log file and send it to the higher authorities. For example, In Business environment, it is mainly used to monitor the activities of the employees to ensure that they are using the computers only for business use and not for any other unwanted activities. However, Sometimes the log file information will be passed to the unknown third parties. This may leads to revealing the user personal information to unknown parties.

C. System reconfiguration attacks

The malicious nature of system reconfiguration attack is performing changes on the settings of the user PC. For example changes in the favorites file of user, usually favorites files contains user visited URLs. So the modification to those files may lead to malicious attacks.

D. Malware-based phishing

In Malware-Based phishing, users are tricked by an alert or update message. The alert or update notification supplied by the malware companies to the user computer. These update message contains information like “your computer has a virus and it needs protection from virus”. So these kinds of messages may insist the user to perform some unwanted actions which leads to the malicious attackers to get their personal information like credit card details.

E. Data Theft

Data theft is known as stealing company confidential data and misusing it. The person who is stealing the data might be working in the same company or he/she might have already quit the company.

F. Search engine phishing

Search engine phishing occurs through the fake website created by attackers. Here the attackers wait until the unaware user to place their personal information in the places like conforming order for purchase or any signup process.

G. Pharming

Pharming is one of the types of phishing attack in which malicious practice is done on the DNS server by creating false information and redirecting the user into fraudulent websites, without their consent.

H. Web Trojans

The motive of Web Trojans is to collect user credentials by creating fake popup on login screens.

I. Deceptive Phishing

The term "phishing" in general refers to stealing of personal information using instant messaging. But nowadays, a deceptive email messaging is used as a common method for phishing. In other words deceptive phishing refers to creating an impression different from the original, making unwary to respond to the fake link and getting personal information through it.

J. Host File Poisoning

When a user tries to navigate to a website, the IP address of the user computer can be determined either by using DNS

server or by using a local file called host file. By poisoning that host file, the hacker can redirect the user from legitimate website to fake site, where they can steal the confidential information of users.

III. Targets Of Phishing Attacks

Phishing Target mainly focused on Emails, Marketing Websites and Mobile phones.

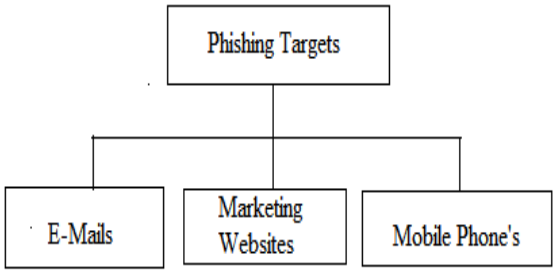

Fig 1: Major Target of Phishing Attacks

A. Characteristics of Phishing Emails

- Phishing email appears as an important notice, urgent update or alert with an enticed subject and it makes the user to believe that the email has come from a trusted source. In order to bypass spamming filters it arrives with numeric characters or other letters.
- Sometimes the email arrives like promising messages such as prize or reward.
- It normally forged the identity of the organization and making the email appear as it's comes from the trusted organization it claimed to be.
- The attacker creates fake websites like legitimate website and the suspicious link is attached to the user email by the attacker. The fake websites contains the similar contents such as texts, logos, images and styles to obtain the recipient confidence.
- It also contains illegal forms for the recipient to fill in personal/financial information and let recipient submit it.

B. Characteristics of Phishing websites

- Phishing websites uses genuine looking content such as images, texts, logos to entice the user and making them to enter their accounts or financial information.
- It may use a similar domain name or sub-domain name as that of the legitimate website.
- It may use forms to collect visitors' information where these forms are similar to that of in the legitimate website.
- Creating fake address bar in place of the original address.

C. Characteristics of Mobile Phishing

- SMS phishing is one of the types of phishing in mobile phones in which the user is tricked by the attacker using text messages to perform criminal activity.
- Sometimes fake phone calls from someone who pretends to be a legitimate person for acquiring personal information from the user.
- The user interface of mobile phones uses simplified version. This will lead to attacker easily implement phishing on mobile phones.

IV. Authentication Based Anti-Phishing Techniques

Authentication is the process of giving individuals access to system objects based on their identity. Market segments such as E-commerce, Online Banking which is the major target of Phishing attackers. In [3], [4] and [5] which provides authentication based anti-phishing solution in different environment.

TABLE I - Anti-Phishing Techniques Based on Authentication

Year	Author	Concept of Anti-phishing Approach
2008	HwanJin Lee, InKyung Jeun, Kilsoo Chun, Junghwan Song	Two factors authentication is used to secure Open ID.
2008	Antonio San Martino, Xavier Perramon	Multifactor mutual authentication is used to secure E-Banking environment.
2010	K.Nirmal, K.Geetha, S.E Vinodh Ewards	3-factor authentication is used to maximize online security.

Open ID is a user centric ID management system. The work in [3] is based on anti-phishing technique to prevent phishing attack in Open ID. Open ID is an open standard and decentralized-protocol that authenticates the user and allows them to get the relying party services. The users of Open ID can able to access those web service which recognize and having the facility to use Open ID without any signup process. Thus Open ID simplifies the signing process. This means websites that takes the advantage of Open ID which would not need the personal information of user over and over again. However, there is a several means of phishing attack in Open ID is possible. The author uses portable tokens and authentication emails as a second medium to prevent phishing in Open ID.

The anti-phishing solution for E-Banking environment [4] uses multi-factor mutual authentication. Multi-factor authentication is a security system that requires more than one form of authentication for the verification of any transaction process. The multi-factor authentication system proposed in this paper [4] is listed as, 1. Password like authentication that is called 'something you know authentication'. 2. Smart card like authentication that is called 'something you have authentication' and 3. Fingerprint like authentication that is called 'something you are' authentication. The mutual authentication between the user and the Bank server is secured by HTTP and SSL/TLS protocol. The authentication process consists of basic model, factor protection model and mixed factor protection model. All these models perform multi-factor authentication in E-Banking based on different hypothesis.

Similar to the work presented in [3] and [4], paper [5] also uses authentication based anti-phishing method. This authentication based anti-phishing system uses 3-factors for authentication. So it is called 3-factor authentication. In this paper [5] the 3-factor authentication is used to prevent counter-attack phishing and it is based on phish-secure phishing algorithm which uses image similarity detection method.The phish-secure algorithm verifies the page the user tends to visit and detects the phishing using the factors such as URL verification, Black listing, and Layer 3 Destination Address Verification. The system maintains Database which consists of the calculated values such as actual mean RGB of various websites, actual mean RGB of particular portion of the

websites, actual IP address, actual URL in database, and Blacklisted IP.

The entire workflow of this prevention system starts by capturing the image of web page (i.e., visual page) and extracts the corresponding elements with respect to the database elements. Finally the elements of visual web page and the elements of database are compared to find the page similarity between them. The result shows the website is phishing or not.

Thus the work in [3], [4] and [5] focuses authentication based anti-phishing technique. In Open ID system [3] the authentication is based on two factors such as portable tokens and authentication email. The advantage of portable token is easily portable in USB memory. The Open ID provider performs encryption on Open ID provider authentication key by using Open ID provider encryption key and symmetric algorithm and saves the Open ID provider authentication key on user computer. In case the user loses the authentication key, no one can able to perform decryption why because the encryption key is known only to the Open ID provider. The authentication email can be used at anytime and anywhere.

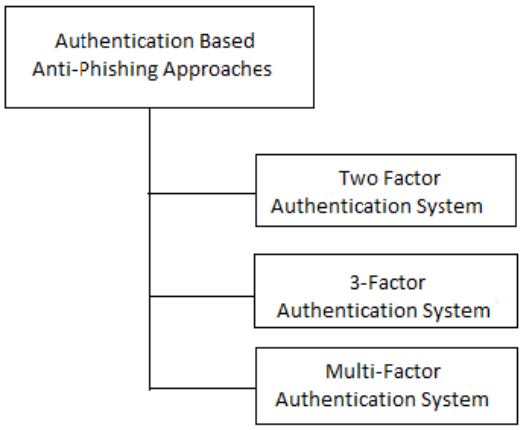

Fig 2: Types of Authentication System

Antonio San Martino, Xavier Perramon uses multi-factor mutual authentication based anti-phishing in [4]. In this type of authentication, In case anyone of the authentication factor is hacked by the attacker in the sense there are still atleast one more authentication factor the attacker has to break to reach into the target. Compared to two factor authentication system [3], the multi-factor authentication [4] is highly immune to phishing attack.

One of the advantages of 3-factor authentication system [5] is URL verification. URL verification finds out 79% of the phishing attack. The second factor blacklisting verification reduces the detection time of phishing attack. The third factor i.e the verification of destination server IP address reduces the probability of phishing attack and the image similarity detection method in [5] uses RGB mean value so we can get accurate input for exact detection of phishing attack.

V. Anti-Phishing Approaches In Phishing Emails

The work in [6] is based on hybrid method for phishing mail detection. Blacklisting, White listing and heuristic techniques are presented in hybrid method. Blacklisting [7] is a list of register of entities that stores fake DNS. White listing [8] is a list identifying entities that are accepted and recognized. It contains registered DNS. Heuristic technique is based on the analysis of some specific features. Some of the features like text, content, DNS and URL features from email messages are used for analysis of phishing email. The modules in [6] are 1.DNS analyzer and Pattern Matching 2.Classifier System and 3.Lookup System. DNS Analyzer compares the visual DNS and the actual DNS and verifies it is blacklisting or white listing. If the result is null, that is if it is not presented in blacklisting or white listing then the pattern matching module verifies sender DNS and actual DNS if the actual DNS and sender DNS varies then it returns phishing. The classifier module performs analysis based on the heuristic features such as URL and textual features. Lookup System maintains blacklisting and white listing. The DNS analyzer and pattern matching is implemented by using analyze DNS algorithm. The analyze text algorithm performs analysis on text and maintains blacklisted tokens, if these tokens exceeds the threshold it is considered as phishing. The analyze URL algorithm converts the URL into tokens and find out the blacklisted features, if these features exceeds the threshold then it is considered as phishing mail.

Table II - Anti-Phishing Approaches for Phishing Emails

Year	Author	Concept of Anti-phishing Approach
2009	Liping Ma, Rosemary Torney, Paul Watters, Simon Brown	Phishing Email Detection using automatically generating classifier.
2012	Jayshree Hajgude, Lata Ragha	Hybrid method for Phishing Mail Detection
2012	Taiwo Ayodele, Charles A. Shoniregun,Galyna Akmayeva	Machine Learning Framework for detection and prevention of Phishing Email.

The work in [9] is based on detection and prevention of phishing email. It is a MLAPT (Machine Learning Anti-Phishing Technique) framework. The change of interface, color-miss match, domain verification, re-direction verification and data request verification are some of the behavior of phishing email is discussed in [9] and MLAPT is designed to handle detection of such kind of phishing behavior effectively.

The MLAPT learns and analyze those behaviors and then compare the analysis with the database of human judged phishing model. This model consists of thousands of sample of real life phishing attack, false contents, false hostname, and IP address patterns. After the comparison, the MLAPT classifies the phishing email into 3 components. 1. Number of phishing attack, 2.Number of phishing attempt, and 3. Number of no phishing attack based on the decision and the learning information gained by it.

Another type of anti-phishing technique for phishing email detection is automatically generating classifier [10].Normally classifier approach produces accurate result on any system. The author proposed generation of decision parser for the translation of classifier into an implementable code. The overall Process of the detection system as follows. The input consists of collection of emails which is fed into the feature generator and the resultant output is presented in the form of feature matrix. The feature matrix is inputted to the machine learning method which uses C4.5 algorithm [11] to train and test the classifiers. The resultant information gained by the machine learning i.e., information about features is given to the inductor which runs the machine learning algorithm with minimum

features. Finally the classifier classifies the email by using decision tree learning algorithm.

In hybrid method [6] blacklisting and white listing verification greatly reduce the detection time and achieves low false positive. Although there is some advantages in blacklisted technique, some drawbacks are also available. This technique detects only 20% of zero-day phishing attack, because many web pages are short lived. The implementation of the heuristic technique decreases the false positive rate.

In MLAPT [9] anti-phishing technique, MLAPT acts as a decision maker. Machine learning technique performs automatic learning based on the input data. The acquired knowledge from the learning process is used for making future decision. The advantage of machine learning is more accurate than user crafted rules. It won't need a human expert or programmer. It is fast, flexible, customizable and scalable.

The classifier based email detection [10] uses C4.5 algorithm for the implementation. It is one of the types of machine learning algorithm.C4.5 algorithm is used to generate a decision tree and it can be used for classification. The C4.5 can handle continuous and discrete attributes, reduces misclassification errors, handle missing attribute values of training data, and produces good classification accuracy.

Both [9] and [10] paper implements machine learning approach, The major drawback of machine learning is it needs lot of labeled data and in c4.5 algorithm [11] the machine learning implementation leads to error prone i.e small variation in the data sets can lead to different decision tree.

VI. Anti-Phishing Approaches in Phishing Websites

The anti-phishing technique for phishing website detection through the logo is done in this paper [12].The main function of this system is logo segmentation and website identity identification. In logo segmentation the webpage screenshot is captured by using Google Chrome plugins and then segmentation is done on the screenshot page to get the best cropped logo image.

In website identity identification the segmented region is loaded into Google image search engine. The search engine locates the most similar logo from Google Image Database using content based image retrieval mechanism. The search result of the Google image search is given as input to Google text search. The domain of the given input webpage is compared with the Google text search. If match is not found then the webpage is considered as a phishing website.

Authentication is the process of giving individuals access to system objects based on their identity. Market segments such as E-commerce, Online Banking which is the major target of Phishing attackers. In [3][4] and [5] which provides authentication based anti-phishing solution in different environment.

Guang-Gang Geng, Xiao-Dong Lee, Wei Wang, and Shian-Shyong Tseng proposed anti-phishing solution based on the characteristics of favicon [13]. It is a file containing one or small icon, website icon or bookmark icon. It is usually placed in favicon.ico root directory. Attacker often perform trick on favicon and make the user to believe that he/she is connected to the proper website. By analyzing the characteristics of favicon, an alternative grayscale information based phishing sites detection is proposed by Guang-Gang Geng, Xiao-Dong Lee, Wei Wang, and Shian-Shyong Tseng. It is implemented by using favicon image recognition algorithm in which the recognition between two image sets are done. The main idea is comparing the gray scale values in two levels such as global and core pixels. The overall process consists of fetching and extraction of all the images on two favicons then transform them into grayscale and re-size all the grayscale images into 16×16 pixels. Finally comparison between the image pairs is done using favicon image recognition algorithm. If match found the suspicious URL is classified as phish.

TABLE III - Anti-Phishing Approaches for Website Based Phishing

Year	Author	Concept of Anti-phishing Approach
2010	Gang Liu, Bite Qiu, Liu Wenyin	Automatic phishing target detection by using phishing webpage
2010	Jordan Crain, Lukasz Opyrchal, Atul Prakash	Prevention of phishing by combining automatic and transparent email signing with email an email client plugin
2011	Sadia Afroz, Rachel Greenstadt	Phishing Website detection by using profiles of trusted websites
2013	Ee Hung Chang, Kang Leng Chiew, San Nah Sze and Wei King Tiong	Phishing detection by using website identity
2013	Guang-Gang Geng, Xiao-Dong Lee, Wei Wang, and Shian-Shyong Tseng	Phishing Websites detection by analyzing the characteristic of favicon
2013	Hossain Kordestani, Mehdi Shajari	Automatic Phishing Detection by entice resistant
2013	Weibo Chu, Bin B. Zhu, Feng Xue, Xiaohong Guan, Zhongmin Cai	Phishing websites protection by using lexical and domain features.

In Phishing website detection through the logo [12] Google image search engine is used to get the search results faster. In favicon based detection [13] Google page rank and DNS information is used to filter the non-phishing sites and reduce the false positive rate. To evaluate the performance of Logo based detection System [12], the phishing websites are collected from PhishTank and legitimate website are collected from Alexa. The detection results are separated into four datasets (dataset1, dataset2, dataset3, dataset4). The dataset1 performance is based on 1×3 segmentation, the dataset2 is based on 2×2segementation, the dataset3 is based on 3×3 segmentation and the dataset4 is performance is based on Best Fit Logo Segmentation.

According to the performance result, the dataset4(i.e the Best Fit logo segmentation) achieves a true positive of 92.5%, true negative of 100%, false positive of 0%, and false negative of 7.5% and which is considered as the best performance compared to other datasets.In favicon based detection [13] the dataset are collected from phish Tank, APWG, DMOZ and Google. The performance evaluation of [13] results true positive of 99.5% and false positive of rate of 0.15%. The favcion detection [13] can lead to reduce the page load time due to the need of checking it in a fixed location. Many of the websites do not use favicon. In this case The favicon based detection method is ineffective. Logo Based detection system [12] also has the same type of drawback like favicon based detection [13].

In phishing website through the logo [12] considers only the image segmentation of the website (i.e., the logo on the top left side of the webpage). Although this method produces

accurate result, The resultant is entirely based on the given right input. This means many of the websites not at all placed its logo on the top left side of the webpage so there is a possibility of wrong detection and it leads to false result. There are also some other situation like some logo will look similar to other logo. This will cause Google Image Search return an undesired result and decrease the detection accuracy.

Sadia Afroz and Rachel Greensand introduces PhisZoo [14] a phishing detection which uses profiles of trusted websites, contents and images displayed. A profile of a site is a combination of different metrics that uniquely identifies the particular site. In a profile PhishZoo having SSL certificate, URL , and contents of a site such as HTML files, extracted features of the logo. Phiszoo stores these profiles in a local database.

Whenever a site is loaded it is matched against the stored profile at the time of loading. If the SSL and URL of the loaded matches with the stored profile, it is considered as legitimate site, if match is not found then it will be checked against appearance of the profiles. Profile Matching is done by extracting the tokens in the hostname, the path URL and HTML files. Then these tokens are searched for specific keywords. TF-IDF Technique is used to select keywords from the domain name of the protected URL and HTML files. The selected keywords also used to select the relevant profile whose logo will be matched with the images of the newly loaded site. Match score is computed as follows,

Match Score = Number of keypoints matched/ total keypoints in the original logo.

Scale Invariant Feature Transform (SIFT) is used to extract the features from the site logo. PhishZoo [14] uses computer vision algorithm for the detection of phishing attack. This method detects 90% of current phishing sites with a false positive of only 0.5%. This approach uses lexical features of URL along with the site contents and image analysis to improve performance and reduce false positive rate. PhishZoo can be used for both online and offline phishing detection. Profile matching approach is based on current contents, so that the phishing can be detected as soon as the page is loaded. SIFT can detect 84% logos even after applying various transformations, Gaussian noise rotations (Upto 30 degrees and scaling.) SIFT fails to detect images that are rotated more than 30 degrees. Favicon based detection [13] and logo based detection [12] uses only particular features but Phiszoo [14] uses entire profile of a websites for detection.

The work in [15] discussed about the effectiveness of machine learning based phishing detection method and it uses only lexical and domain features. The proposed system consists of two stages model training learning stage and detection stage. In both stages the re-direction parse module is used to convert the original URL into true URL if the original URL and true URL exists the feature extraction module extract the lexical and domain features.

In learning stage these features are fed into the support vector machine. In machine learning support vector machine is a supervised learning model with associated learning algorithm that analyze data and recognize pattern used for classification and regression analysis. Support Vector Machine performs linear and non linear classification. The resultant data from Support Vector Machine and the resultant data from feature extraction of detection stage is inputted to the detection model. Finally the detection model produces the detection result.

The author named Hossain Kordestani, Mehdi Shajari introduced an entice resistant automatic phishing detection [16]. It is a offline machine learning method. It consists of learning phase and detection phase. In learning phase the system gets the input from the user and learns about the difference between the phishing and legitimate website, based on the properties of each class of input. Here the Machine Learning engine i.e ML engine is used for learning purpose.

In this phase the collection of websites and vector of properties from third party services is given as input to the feature extractor. The feature extraction was implemented by jsoup library and weka implementation of the classifier. The feature extractor extracts the online features, content based features and URL based features. These features will be inputted to the ML engine based on the input information from feature extractor. ML engine learns about each class properties (i.e legitimate and phishing properties) and generate a classifier. In detection system the input URL of a given webpage is fed to feature extractor and then using the classifier the detection is performed. The evaluation metric F1 is calculated from false positive and false negative.

The work in [15] uses 18discriminative features, among those the domain brand-name distance, path brand-name distance, domain age and domain confidence level are newly introduced by the author. Plus 2 minus 1 algorithm is used to evaluate the performance of the features.

Function of plus m minus r: Begin with zero function chosen, sequentially append m features to chosen ones and pop r features from them. Select the features set that yields the Best classification result. Successively performing the above algorithmic steps, finally it yields best classification accuracy.

The main advantage of this paper [15] is it uses lexical and domain features which are available even when the phishing web pages are inaccessible. The detection system achieves 98% accuracy with a false positive rate of 0.64% or less. The detector is effective even when the phishing URL changes. The lexical features have the accuracy rate of 95.88%, false positive of 0.82%, and false negative of 9.38%. The domain features has the accuracy of 98.14%, false positive is 2.56%, and false negative is 1.37%. While considering the performance, the lexical features which has little worse performance than domain features. The Support Vector Machine is simultaneously minimizing the empirical classification and maximizing the geometric margin classification space. These properties reduce the structural risk of over learning with limited samples. The Gaussian Radial Basis Function (RBF) and plus m minus r algorithm gives the best classification accuracy.

The disadvantage of Support Vector Machine is it requires large number of labeled training samples. We know Support Vector Machine is a Binary Classifier. In order to perform multi-class classification in Support Vector Machine pair-wise

classification can be made. But it is computationally expensive and runs slow.

In [16] Random based evaluation is used to evaluate overall performance of the system and for evaluating the parameters, cross validation is used. There are 3 classifier used in this system. Each classifier produces different result. The F1 Measure of Bayesian Network classifier is 0.991, support vector machine is 0.981 and Random Forest is 0.993. The advantage of entice resistant method produces true positive of 99.9%, false positive of 0% , precision of 100%, recall of 99.9%, and F1-Measure of 99.9% based on the comparison of performance of entice resistant with other method like Cantina and Cantina+. The majority of information is provided from trusted third party services like alexa, google etc. Therefore it is not easy for the attacker to mislead the information.

Reference to the paper "[17]" the author uses DBSCAN algorithm to detect phishing website. DBSCAN (density based spatial clustering of application with noise) is a data density algorithm which finds the clustering between the nodes. The main idea behind this paper is it finds the cluster between the given webpage and its associated web pages. There are two kinds of associated given web pages, one is directly associated pages and indirectly associated web pages. The directly associated pages can be found by extracting the HTML source hyperlink and indirectly associated pages can be found by similarity in text or visual information.

The attributes of the webpage such as Link, Rank, Text Similarity and Layout similarity are used to find the relationship between the given webpage and the associated webpage. The calculation of these attributes is briefly explained in [17]. Then the clustering is performed by using DBSCAN algorithm. In DBSCAN algorithm the specification of the number of clusters in the data a priori is not required. It can find arbitrarily shaped clusters and it can select any points as start point for clustering. In this paper [17] author uses Endpoints, Min-Points, Core points, directly density reachable and density-reachable are used for implementation. The End points and Min-Points are used to evaluate the accuracy rate and false alarm rate.

The disadvantage of DBSCAN is not entirely deterministic and it cannot perform clustering on those data sets having large differences in densities. The advantage of this method [17] is it produces phishing detection accuracy of 91.44% with the dataset of 8745 pages and false alarm rate of 3.4% with the dataset of 1000 pages.

In [18] the author introduces trusted email for prevention of phishing by combining automatic and transparent email signing with an email client plugin. This method will help the user to identify and distinguish legitimate websites from spam and alert before the user is exposed to phishing. The Trusted email system requires the company or institution (i.e Bank, credit card companies etc) to either obtain a certificate for their key pairs or to generate their own key pairs. In this system the communication between the user and the institution takes place by using RSA public key encryption algorithm.

In this implementation [18] the author used two system components, one is proxy server and the other one is email client plugin. The entire process of communication is described between the user and the institution (for example: Bank)

1. User Wishes to establish a trust relationship with the bank. They send an initialization equest to bank for secure communication.
2. Bank send response email to user through bank proxy. The bank proxy establish encryption(i.e proxy sign) and send email to user proxy.
3. User Proxy delivers authorization request email by verifies proxy signs and add verification header before delivering email.
4. Email plugin parses headers and displays trust information.
5. User authorized the delivered email (i.e the trust information) The authorized response will be processed by the user proxy and send to the bank proxy. Successfully the user proxy completes establishment of keys with Bank proxy.
6. Bank proxy sends confirmation back to user proxy that is bank confirmation message to user.

VII. ADVANTAGES

1. The proxy generates an RSA key-pair or use a certificate generated by a certificate authority and majority of the functions is done by proxy server.
2. Email Plugin is used for user interface.
3. All communication between the proxy and the plugin take place in the headers of email messages passed between the two. This ensures that the plugin and the proxy will be able to communicate in any situation where sending and receiving email is possible.
4. RSA is secure and convenience.

VIII. DISADVANTAGES

1. The trusted email method of email verification is not designed to provide protection over already compromised communications channels.
2. Many users do not use their custom message of security during Initial request communication between the user and the bank.
3. Even the custom message is added sometimes there is a possibility of eavesdrop attack.

IX. ANTI-PHISHING APPROACHES IN MOBILE PHISHING

The work in [19] is based on anti-phishing technique against mobile phishing attack. In this technique, web-to-native phishing attack is developed on iOS mobile platform. The real demonstration of the developed web based attack proved that it has the capability of easily fakes the real apps in iOS. The defense implementation against this web based attack uses keyloggers, alert system and policy whitelist. This defense implementation is effective on iOS built-in keyboard. The major drawback is if any system uses its own implementation of keyboard. The proposed system is ineffective.

TABLE IV - Anti-Phishing Approaches for Mobile Phishing

Year	Author	Concept of Anti-phishing Approach
2012	Jie Hou, Qi Yang	Keyloggers, alert system and policy whitelist is used to defend against mobile phishing
2014	Longfei Wu, Xiaojiang Du and Jie Wu	MobiFish defense targets mobile web pages and Apps

The author introduces mobifish [20] antiphishing scheme for mobile phones. The mobifish has two components. One of the component is Webfish which has the capability to handle phishing attacks on mobile webpages and the another component is Appfish which has the capability to handle phishing attacks on mobile applications. Here the Webfish component which gets the actual identity from the URL of a webpage and the Appfish which gets the actual identity from the URL of a remote server. The claimed identity of the webpages and the apps is extracted from the screenshot of a given page.

Optical character recognition is used to convert the screenshot from the text. Here the Tesseract tool is used to extract the text from screenshot. Comparison is made between the actual identity and the claimed identity if any variation between the identities this tool sends warning to the user. The performance evaluation of mobifish is done by using 100 phishing URL and the corresponding legitimate URL. The results shows the detection and prevention of mobile phishing attack is efficiently handled by mobifish.

X. Conclusion

In this paper different types of phishing targets and their anti-phishing solution is surveyed. However, there are several medium are available to attempt phishing, the major cause behind all these types of phishing is user unawareness. This study will help the user to get full-fledged knowledge about phishing, because it thoroughly discussed about each types of phishing attack, its nature and major target of attackers. Advantages and drawbacks of different anti-phishing technique corresponding to phishing targets also discussed.

XI. Future work

The future work is increasing the authentication factor and implementing it with real life phishing attack using several types of real datasets.

References

[1] J. Hong, “The state of phishing attacks,” Communications of the ACM, vol. 55, no. 1, pp. 74-81, 2012.

[2] Anti-Phishing Working Group (APWG), “Phishing activity trends report second quarter 2014,” http://docs.apwg.org/reports/apwg_trends_report_q2_2014.pdf, 2014, accessed Oct 2014.

[3] H.J. Lee, I.K. Jeun, K. Chun and J. Song. “A new anti-phishing method in OpenID,” in Second International Conference on Emerging Security Information, Systems and Technologies, 2008, IEEE, pp. 243-247.

[4] A. San Martino and X. Perramon. “A model for securing e-banking authentication process: anti-phishing approach,” in IEEE Congress on Services - Part I, 2008, IEEE, pp. 251-254.

[5] K. Nirmal, S.E.V. Ewards and K. Geetha. “Maximizing online security by providing a 3-factor authentication system to counter-attack 'phishing',” International Conference on Emerging Trends in Robotics and Communication Technologies, 2010, IEEE, pp. 388-392.

[6] J. Hajgude and L. Ragha. “Phish mail guard :phishing mail detection technique by using textual and URL analysis”, World Congress on Information and Communication Technologies, 2012, IEEE, pp. 297-302.

[7] M. Sharifi and S.H. Siadati, "A phishing sites blacklist generator," in IEEE/ACS International Conference, 2008, pp. 840-843.

[8] Y. Cao, W. Han, and Y. Le, "Anti-phishing based on automated individual white-list," Proceedings of the 4th ACM workshop on Digital identity management, 2008, pp. 51-60.

[9] T. Ayodele, C.A. Shoniregun and G. Akmayeva. ”Anti-phishing prevention measure for email systems,” World Congress on Internet Security, 2012, IEEE, pp. 208-211.

[10] L. Ma, R. Torney, P. Watters and S. Brown. “Automatically generating classifier for phishing email prediction,” in 10th International Symposium on Pervasive Systems, Algorithms, and Networks, 2009, IEEE, pp. 779-783.

[11] J. R. Quinlan. “C4.5: programs for machine learning,” 1993.

[12] E.H. Chang, K.L. Chiew, S.N. Sze and W.K. Tiong.”Phishing detection via identification of website identity,” International Conference on IT Convergence and Security, 2013, IEEE, pp.1-4.

[13] G.G Geng, X.D Lee, W. Wang and S.S Tseng. ”Favicon - a clue to phishing sites detection,” in eCrime Researchers Summit, 2013, IEEE, pp.1-10.

[14] S. Afroz and R. Greenstadt. “PhishZoo: detecting phishing websites by looking at them,” in Fifth International Conference on Semantic Computing, 2011, IEEE, pp. 368-375.

[15] W. Chu, B.B. Zhu, F. Xue, X. Guan and Z. Cai, “Protect sensitive sites from phishing attacks using features extractable from inaccessible phishing URLs,” International Conference on Communications, 2013, IEEE, pp. 1990-1994.

[16] H. Kordestani and M. Shajari. “An entice resistant automatic phishing detection,” in 5th Conference on Information and Knowledge Technology, 2013, IEEE, pp. 134-139.

[17] G. Liu, B. Qiu and L. Wenyin. “Automatic detection of phishing target from phishing webpage,” in 20th International Conference on Pattern Recognition, 2010, IEEE, pp. 4153-4156.

[18] J. Crain, L. Opyrchal and A. Prakash. “Fighting phishing with trusted email,” in 10th International Conference on Availability, Reliability and Security, 2010, IEEE, pp. 462-467.

[19] J. Hiou and Q. Yang. “Defense against mobile phishing attack,” in EECS 588 Project, 2012.

[20] L. Wu, X. Du and J. Wu. “MobiFish: a lightweight anti-phishing scheme for mobile phones”, in 23rd International Conference on Computer Communication and Networks, 2014, IEEE, pp. 1-8.