Prevention of Phishing Attacks Based on Discriminative Key Point Features of WebPages

Mallikka Rajalingam *mallikka2002@yahoo.com*
School of Computer Sciences
Universiti Sains Malaysia
Pulau Penang,11800, Malaysia

Saleh Ali Alomari *salehalomari2005@yahoo.com*
School of Computer Sciences
Universiti Sains Malaysia
Pulau Penang,11800, Malaysia

Putra Sumari *putras@cs.usm.my*
School of Computer Sciences
Universiti Sains Malaysia
Pulau Penang,11800, Malaysia

Abstract

Phishing is the combination of social engineering and technical exploits designed to convince a victim to provide personal information, usually for the monetary gain of the attacker (Phisher). Attempts to stop phishing by preventing a user from interacting with a malicious web site have shown to be ineffective. In this paper, present an effective image-based anti-phishing scheme based on discriminative key point features in WebPages. We use an invariant content descriptor, the Contrast Context Histogram (CCH), to compute the similarity degree between suspicious pages and authentic pages. To determine whether two images are similar, a common approach involves extracting a vector of salient features from each image, and computing the distance between the vectors, which is taken as the degree of visual difference between the two images. The results show that the proposed scheme achieves high accuracy and low error rates.

Keywords: Image Clustering and Retrieval, Anti-Phishing Mechanism, Digital Image Processing, Security

1. INTRODUCTION

Phishing is also known as "brand spoofing". It is pronounced as fishing. The word has its origin from two words "Password harvesting" or "fishing for passwords". Phishing is a type of deception designed to steal your valuable personal data, such as credit card numbers, passwords, account data, or other information. Phishing is a form of online identity theft associated with both social engineering and technical subterfuge. Attackers might send millions of fraudulent e-mail messages that appear to come from Web sites you trust, like your bank or credit card Company, and request that you provide personal information [25]. As scam artists become more sophisticated, so do their phishing e-mail messages and pop-up windows, they often include official-looking logos from real organizations and other identifying information taken directly from legitimate Web sites. When users unwittingly browse phishing pages and enter their personal information like user name and password their password will get stored in the attackers database and then users are redirected to original sites directly by way of phisher-controlled proxies. Phishing has thus become a serious threat to information security and Internet privacy [16]. To deceive users into thinking phishing sites are legitimate, fake pages are often designed to look almost the same as the official ones in both layout and content. Phishers might insert an arbitrary advertisement banner that redirects users to another malicious Web site if they click on it. So, phishing attacks have become a serious threat. To reduce phishing attacks there is a group called APWG (Anti Phishing Working Group)

which will have a list of phishing pages. If a user finds that the page he visits is a phishing page then he can report it to the APWG [10]. They will add that page in the list of phishing page. First to find whether the page is a phishing page or real page, we've developed a color based image comparison method. Color plays a vital role in an image. Even a small difference can be found by comparing images based on color.

Nowadays, the phishing attack has become a bigger problem. It results in stealing One's personal information like Gmail account and bank password. To avoid phishing attacks many methods are developed, but none of the method is more efficient enough to solve such this kind of problems and still the phishing attacks takes place. The method that we developed here is purely image-based [4]. Snapshot of the requested page is taken. The page is stored as an image then the next step is to get the original page snapshot which is also saved as an image. Then select the source image as well as the select the target directory which contains the images to be compared. The images are compared by using color ratio. The difference is noted and reported to the user. When the difference is zero then the page is not a phishing page. This anti phishing tool is very efficient because it compares the phishing and authentic pages based on the visual appearance level, instead of rather than using text-based analysis [7].

1.1 A Growing Problems in Phishing

The phish attack volume increased 33% in April to 36,557 attacks, continuing the growth trend from March. Phish attacks had been in general decline from August 2009 to February 2010, but now look set to return to the seasonal growth trend that has historically peaked in late Summer/early Fall [9]. In August 2009, for example, the high point of fast-flux phish attacks Produced 60,678 incidents. As shown in Figure. 1, the monthly attacks from April 2009 to April 2010 averaged 45,605. Phish attack volume has not returned to the level seen in April 2009, but note that this chart does not include branded malware attacks, which cybercriminals are likely to have launched during periods of lower phish volumes.

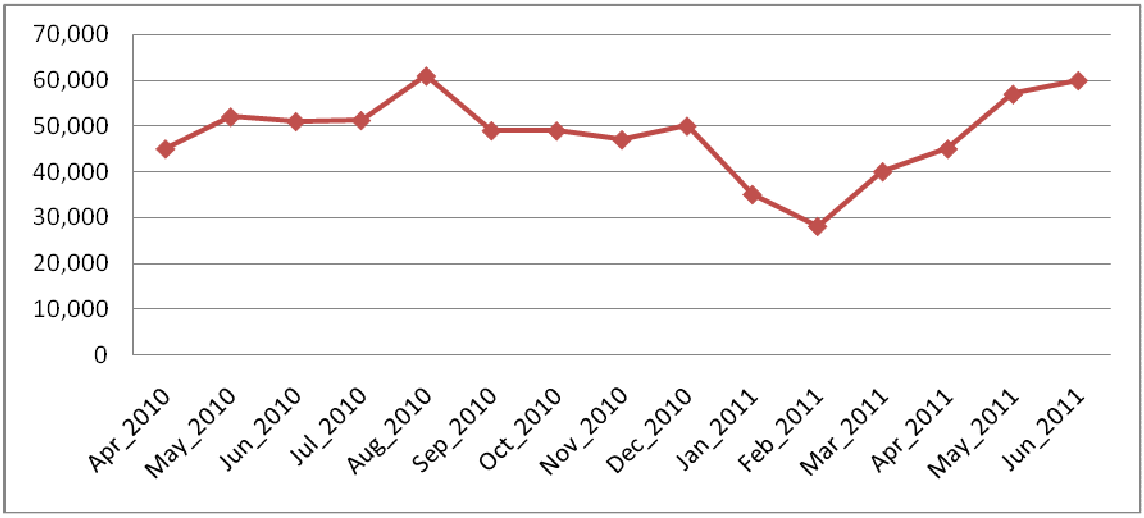

FIGURE 1: Monthly Phishing Attack

As shown in the Figure.2, the payment services sector was the primary sector favored by phishers accounting for 41% of phish attacks in April. The financial sector, historically the most popular phishing sector, accounted for 33% of phish attacks. The auction sector was targeted in 7% of attacks [8]. The "Other" category, which includes social networks, online gaming, online media, various Internet companies, as well as other organizations, accounted for 14% of attacks.

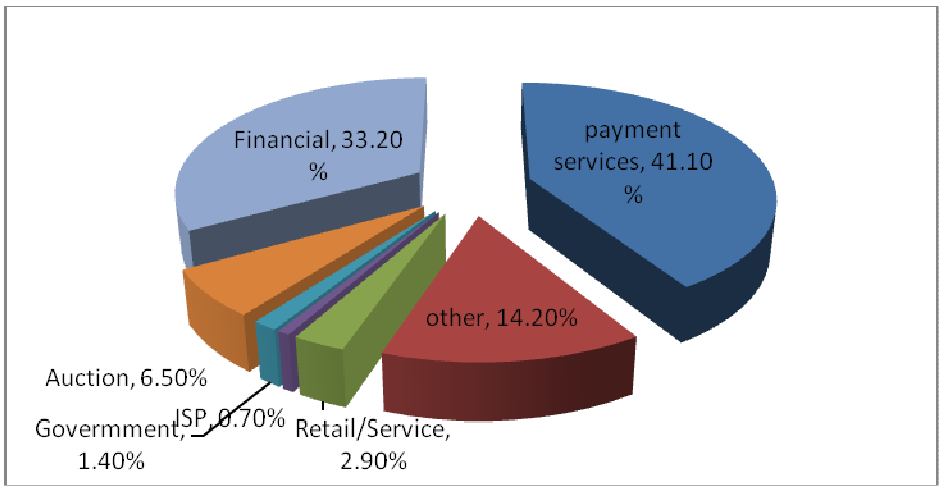

FIGURE 2: Phishing by Industry

1.2 Phishing Attack

Employ visual elements from target site. Most methods of phishing use some form of technical deception designed to make a link in an e-mail (and the spoofed website it leads to) appear to belong to the spoofed organization. Misspelled URLs or the use of sub domains are common tricks used by phishers [11]. Example: www.gmail.com – original link, www.gmai1.com – Fake link. Here are a few phrases that a phishing page may contain verify your account, businesses should not ask you to send passwords, login names, social security numbers, or other personal information through e-mail.

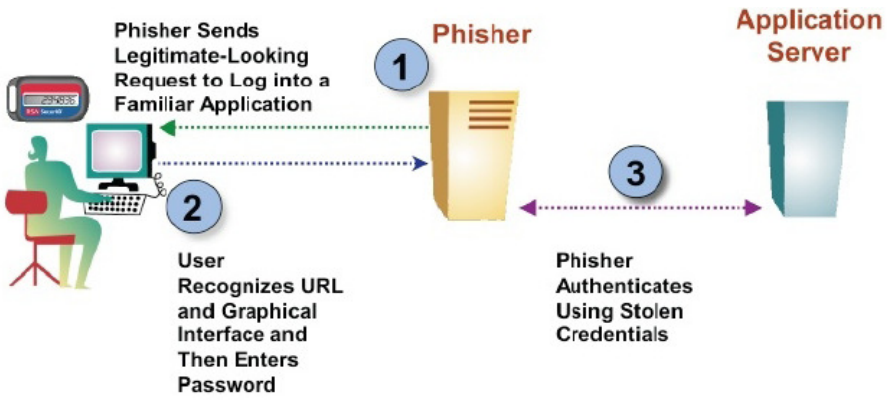

FIGURE 3: Phishing Attack

From the Figure.3, if you receive an e-mail from anyone asking you to update your credit card information, do not respond, this is a phishing scam. If you don't respond within 48 hours, your account will be closed. These messages convey a sense of urgency so that you will respond immediately without thinking. In the Figure.4, the phishing e-mail might even claim that your response is required because your account might have been compromised [3].

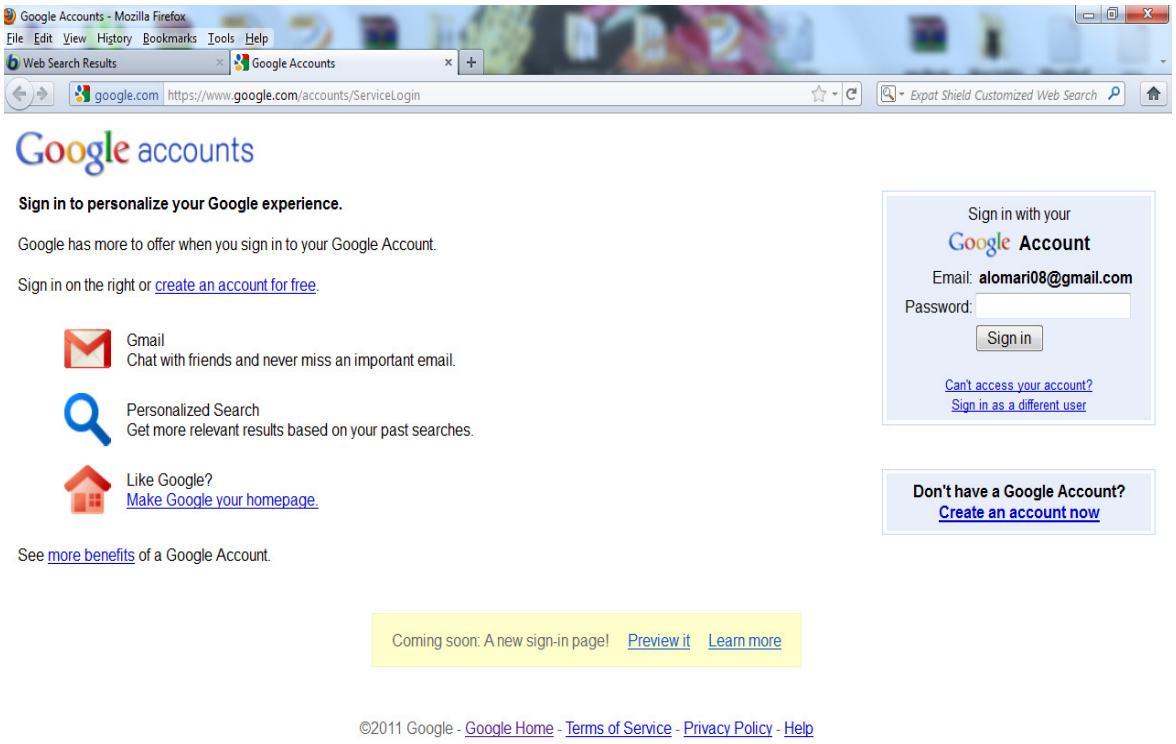

FIGURE 4: Phishing page

The above page looks like an original Gmail page. But it is a phishing page. Whenever this type of page appears before a user then the user enters the user name and password which gets stored in the attackers database. This type of attack will be a serious one if the attacker steals the user's bank user name and password and misuse it.

The remaining of this paper is organized as follows. Basic characteristics of phishing techniques and related works are described in section 2. The methodology and approaches of phishing attacks are discussed in section 3. Various performance testing techniques are in section 4. Results and outputs are in section 5. Then we summarize the whole procedure and draw conclusion in section 6.

2. RELATED WORK

In a SOPHOS white paper-2005, Phish Guru is an embedded training system that teaches users to avoid falling for phishing attacks by sending them simulated phishing emails. People access these training emails in their inbox when they check their regular email. The training emails look just like phishing emails, urging people to go to some website and login. If people fall for the training email that is, if they click on a link in that email. We provide an intervention message that explains that they are at risk for phishing attacks and offers tips they can follow to protect themselves. The training materials present the user with a comic strip that defines phishing, offers steps the user can follow to avoid falling for phishing attacks, and illustrates how easy it is for criminals to perpetrate such attacks [2].

Defending the weakest link, phishing websites detection by analyzing user behaviors, we have used a novel paradigm analysis of the users' behaviors to detect phishing websites. We have shown that it is an accurate method, discussed how it has been designed and implemented to be hard to circumvent, and have discussed its unique strength in protecting users from phishing threats. UBPD is not designed to replace existing techniques. Rather it should be used to

complement other techniques, to provide better overall protection. We believe our approach fills a significant gap in current anti-phishing technology capability.

The image matching is a fundamental problem in computer vision. The existing prevention and detection schemes to combat Phishing in multiple ways. The endless competition between computer virus writers and antivirus software developers, phishers will certainly strive to develop countermeasures against antiphishing solutions. In the existing content-based approach, this analyzes the HTML code and text on a webpage, proved effective in detecting phishing pages. However, phishers responded by compiling phishing pages with non-HTML components, such as images, flash objects, and Java applets. The existing system phisher may design a fake page which is composed entirely of images, even if the original page only contains text information. In this case, the suspect page becomes unanalyzable by content based anti-phishing tools as its HTML code contains nothing but HTML <img/> elements [6]. The drawback of existing systems are ineffective to stop phishing attacks, low degree of accuracy, high error rates, not susceptible to changes in webpage aspect ratio and colors used.

Juan Chen and Chuanxiong Guo proposed a new end-host based anti-phishing algorithm, which we call LinkGuard, by utilizing the generic characteristics of the hyperlinks in phishing attacks. These characteristics are derived by analyzing the phishing data archive provided by the Anti-Phishing Working Group (APWG). Because it is based on the generic characteristics of phishing attacks, LinkGuard can detect not only known but also unknown phishing attacks [21]. Bryan Parno, Cynthia Kuo, and Adrian Perrig proposed using a trusted device to perform mutual authentication that eliminates reliance on perfect user behavior, towards Man-in-the-Middle attacks after setup, and protects a user's account even in the presence of keyloggers and most forms of spyware. We demonstrate the practicality of our system with a prototype implementation [22]. A *spammer* is a person who creates spam messages. *Fraudsters* are people involved in Internet fraud, a practice indulged in by individuals who spam potential victims. It has been reported that in 2003 alone, personal losses amounting to more than 200 million dollars resulted from fraudulent intrusions [23]. Tom Jagatic, Nathaniel Johnson, Markus Jakobsson, and Filippo Menczer proposed context aware phishing, an attacker would gain the trust of victims by obtaining information about their bidding history or shopping preferences (freely available from eBay), their banking institutions (discoverable through their Web browser history, made available via cascading style sheets), or their mothers' maiden names (which can be inferred from data required by law to be public) [24].

In this paper we proposed detecting phishing pages based on the similarity between the phishing and authentic pages at the visual appearance level, instead of rather than using text-based analysis. We first take a snapshot of a suspect webpage and treat it as an image in the remainder of the detection process. We also propose image-based phishing detection scheme that uses the Color ratio and color modes such as RGB to compare images, finally after compared the result will show whether authentic webpage is phishing webpage or not. Our scheme can detect phishing pages with a high degree of accuracy.

2.1 Proposed Solution

Phishing attack has become a serious threat to internet users. It results in stealing one's personal information like Gmail password, bank password, etc. It is an illegal way. To reduce phishing attacks there are many methods developed to avoid phishing attack. The method proposed here is very different. It's purely based on image comparison rather dealing with old text based analysis anti-phishing mechanism. It compares original and authentic webpage images and produces the result [4] [5]. First snapshot of the suspected web page is taken and compared with original web page and the result of the comparison helps the user to identify the phishing page. The main objective of the project is to prevent phishing attacks. To make Online Banking and transaction of money more secured. To prevent the users of gmail, rapidshare, paypal, ebay, etc. getting hacked. To prevent the users loss of data in Internet.

3. METHODOLOGY AND APPROACH

There are four phases 1) Phishing attack demo 2) Web page snapshot 3) Image wizard 4) Comparison of web pages.

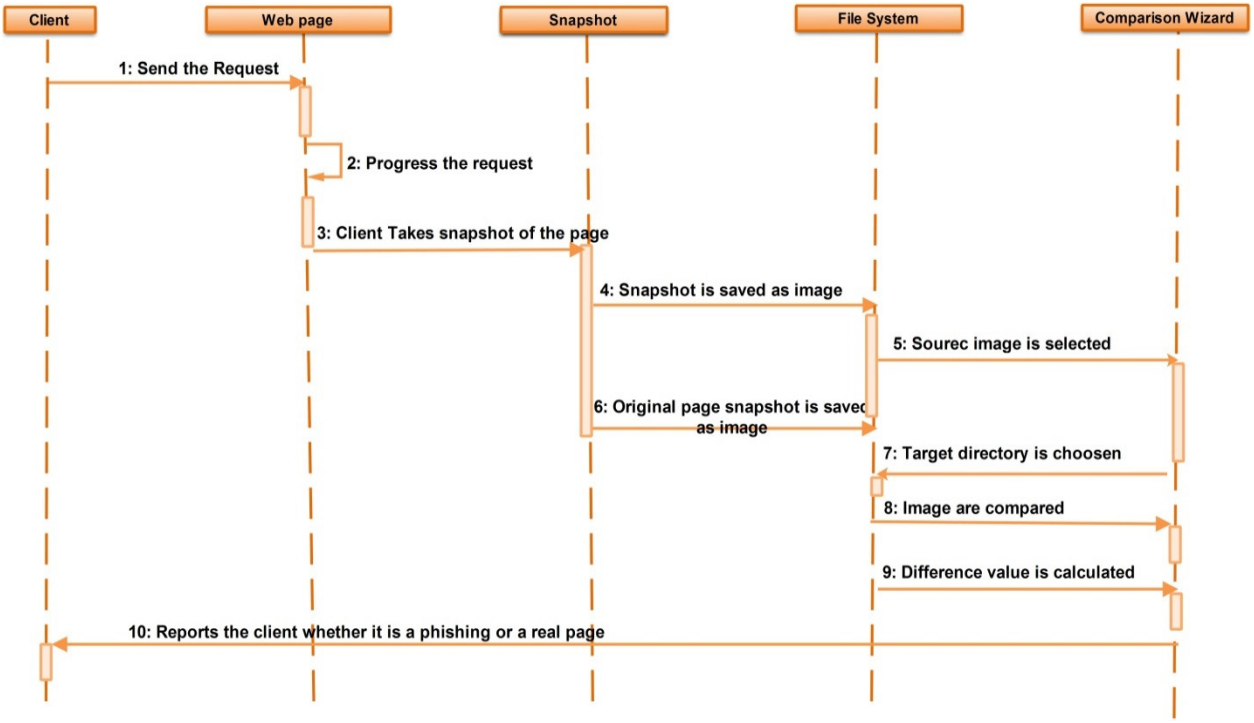

FIGURE 5: Sequence Diagram for Prevention of Phishing Attack

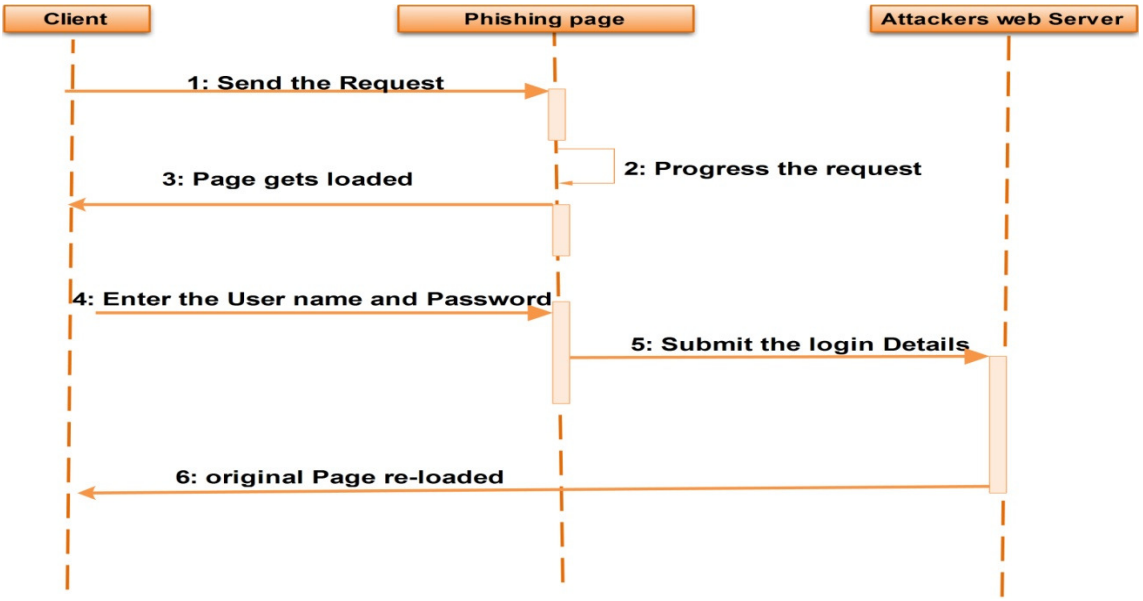

FIGURE 6: Sequence Diagram for Attack

3.1 Phishing Attack Demo

Phishing attack is performed to steal one's personal information. When a user requests a web page the phishers will send their web page which they have developed [13]. They will develop a

phishing page which will be same as that of the original page but there will be some slight differences. Attackers will send it to the user. For example, in the Figure.7 to steal the bank information of a particular user the attackers will send that" the account will expire within today if you fail to fill in the details in the given site" and a link will be provided below. The link which is given will not be the original bank's website. It will contain some official logos which will look similar to the original web site of a bank. Sometimes there will be change in the name of the website in just one character which will be hard for a user to find out. For example, www.ebay.com will be given as www.e6ay.com. Suddenly when the user sees a page with the above specified link they will believe it as an original page. Unknowingly the user enters their password or some personal information which will be taken directly to the phishers server and will get stored in their database. Later the attackers misuse the information given by the user. By using HTML and PHP script, this attack is carried out.

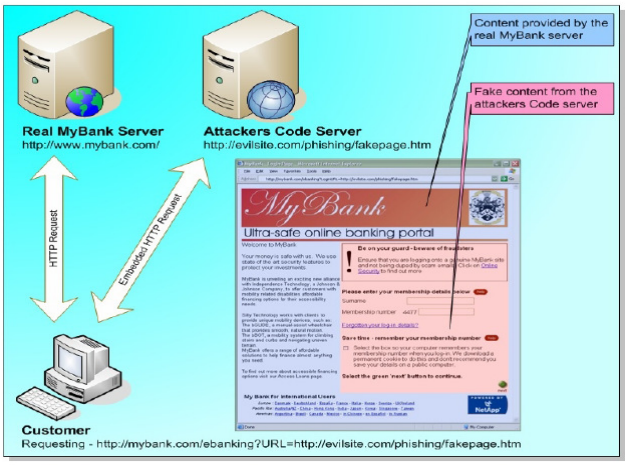

FIGURE 7: E-Banking Fake Page [26]

3.2 Web Page Snapshot

Next step is to take the snapshot of the authentic page. In Figure.9 shows the snapshot of the authentic page we use a tool which will take snapshot of webpage and save in the file system in the required image format. It should be saved only in any of the image format because the web page has to be compared with the original web page. Here specialized tool is needed because we can't take the snapshot normally using print screen key in keyboard. By using print screen key the user can take the whole windows environment and not the required web page alone. This will not give a correct result because if the snapshot of the original webpage image is taken in windows, and when the snapshot of the authentic page is taken in Linux operating system the environment differs and the result will be error prone. So, a tool is necessary to take the snapshot of the web page.

3.3 Image Wizard

Next is to design a wizard to compare the images. In Figure.10 shows the Wizard is designed in such a way that everything appears in the wizard is clear and systematic. Separators are used to clearly distinguish each one. The upcoming window gives instruction to proceed to next window. First the search image is selected and the target directory should be specified in the next step. The panels are designed user friendly and also image panels are used in this wizard to preview the images that are chosen already. If there are no images in the specified target directory the wizard is designed in such a way that an error will be displayed. So, that the user doesn't waste their time by going to the next step of comparing images. Other than that the wizard can get the settings from the user so that sometimes users can give the level of accuracy they needed while comparing images. If out of bound values are given then the wizard takes the default value that is specified. First the original image is taken from the directory which is already stored in that directory. Then we have to specify the target directory which contains similar images. In Figure.11

shows the wizard is designed in such a way that default setting can also be set in preferences window.

3.4 Comparison of Images

Next step is to compare the images. As said earlier in image wizard module user can give the ratio of the accuracy they need while comparing images [1]. First the user is allowed to enter the number of sections by which the image should be divided. Then the image is divided into blocks as given by the user. The target image is also divided into same number of blocks like the original image. The image is divided into blocks by using k-means algorithm. If the user gives "n" number of sections then both the images will be divided into "n*n" number of blocks. For example, if the user gives number of section=3 then both images will be divided into 9 blocks. Get the height and width of the both the images. In this wizard we can also give the number of overlapping value. Instead of taking and comparing each and every whole block, we can also compare blocks that are overlapping. So that can obtain a clear and error prone result. The overlapping value given by the user is taken and it is multiplied with the width and height of the image. By this way we can calculate for overlapping blocks also. Next step is to give the color ratio for the image. First the RGB values of both the images are obtained. Then the average value of the RGB color is obtained. In next step the standard deviation is obtained. Standard deviation determines the range of the colors.

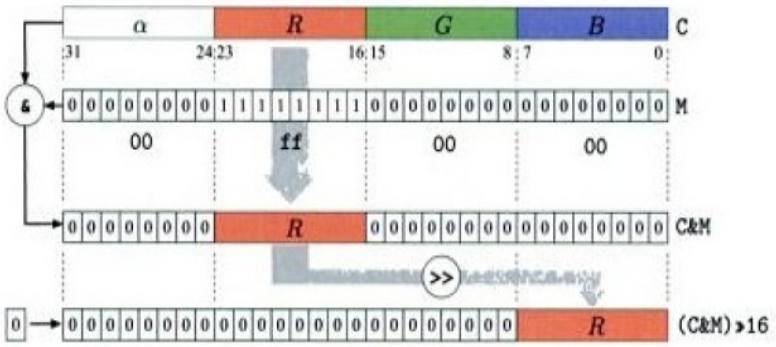

FIGURE 8: RGB Color Images

The Figure.8 shows how to find the RGB value, decomposition of a 32-bit RGB color pixel using bit operations. The R component (bits 16-23) of the RGB pixels C(above in the fig) is isolated using a bitwise AND operation(&) together with a bit mask M=0xff0000. All bits except the R component are set to the value 0, while the bit pattern within R component remains unchanged. This bit pattern is subsequently shifted 16 positions to the right(>), so that the R component is moved into the lowest 8 bits and its value lies in the range of 0 to 255. During this shift operation, zeros are filled in from the left. The construction of an RGB pixel from the individual R,G and B values is done in the opposite direction using the bitwise OR operator(|) and shifting the bits left(<): ((red & 0Xff)<<16) | ((green & 0Xff)<<8) |blue &0Xff Masking the component values with 0Xff works in this case because except for the bits in positions 0 to 7(values in the range 0 to 255), all the other bits are already set to zero. Thus the RGB value is obtained which is then converted into HSV mode. The average color ratios of both the images are obtained by using:

- Red average= sum of all the Red Pixels in the image R (P)/No. Of pixels in the image P
- Green average= sum of all the Green Pixels in the image G (P)/No. Of pixels in the image P
- B average= sum of all the Blue Pixels in the image B (P)/No. Of pixels in the image P

Where,

R (P) = RED component pixels,
G (P) = GREEN component pixels,
B (P) = BLUE component pixels,
P =No. of pixels in the image

After finding the average value, each and every block of the source image is compared with the target image. By using standard deviation, it finds the amount the image is deviated from the average value. The difference between the average values is calculated. If the difference value is zero then the particular page is a real page. If the difference value is more than zero then the page is a phishing page. From the Figure.14, If it is a phishing page then the person can directly report it to the Anti-Phishing Work Group (APWG) using this tool by clicking the button “report phishing”. By clicking this button the user is redirected to APWG web site where the user can give the name of the link which is a phishing page. In this case, the other users will not be fooled by the same link.

4. PERFORMANCE ANALYSIS ON VARIOUS TESTING

There are two general categories of testing. Pre implementation and post implementation. The software testing for the process planning system has been done during the pre-implementation stage using various software testing strategies.

4.1 Unit Testing

The individual modules are tested for proper functioning and are found to be satisfactory as regard to the expected output from the module. The whole work is divided into modules and every module is tested independent of other modules and their functionalities. If the testing of the module requires sub divisions for accurate output they are permissible. The testing is carried out during programming stage itself. There are some validation checks for verifying the data input given by the user for the authentication purpose. The errors could be identified and debugged.

4.2 Interface Testing

After the modules are individually tested we confine the testing process to each and every interface which has been developed in the application since every interface is a master screen. During the interface testing, the GUI interfaces are tested accordingly as per their functionality prescribed. This testing would ensure the proper functioning of the interfaces as per the requirements demanded. Interface testing would improve the performance of the system

4.3 Black Box Testing

This testing focuses on the functional requirements of the software and also it enables the software engineer to drive the sets of input conditions that will fully exercise functional requirements for a program. It attempts to find error such as incorrect missing functions, interface errors, errors in data structures or external database, access, performance errors, initialization and termination errors. The software has been tested to drive a set of cases that satisfy the user requirements

4.4 Integrated Testing

The need for the integrated testing is to find the overall system performance, while testing the whole application there are chances of reoccurrence of errors because, previously all the testing techniques were used to test some individual modules. Now we would integrate all of them and would test for their compatibility as a whole for all the interfaces and the charting process because they are all interdependent on each other. The application has been tested for various kinds of inputs and has successfully passed.

4.5 Validation Testing

At the culmination of Black Box testing, software is completely assembled as a package and tested as a whole unit. Validation testing is where the requirements established as part of the software requirements analysis are validated against the software that has been constructed. It ensures that the software meets all the functional, behavioral and performance requirements. The application was tested on various inputs which authenticates the user as specified by the organization.

5. RESULTS AND OUTPUT

The result shows that the test of our methods is more efficient when compared with existing work. The test case shows the performance analysis with different parameters as shown in Table.1. This makes well to get appropriate output. An example for our test is Gmail original and Gmail fake pages as shown in Figure.15 and Figure.16 respectively.

Test Cases	Input	Expected Input	Output	Expected Output	Result
File name	Text file	Image file	Invalid	Image from the specified directory is loaded	Pass
Directory name	No files in the directory	Image files	No images in the directory	Select the required files from the directory	Pass
Color value	Color value>10&&color value<0.1	Color value<10&&color value>0.1	Invalid values	Values accepted	Pass
RGB or HSV	Nothing is selected	One radio button is chosen	Invalid value	Values accepted	Pass
Number of sections	Number of sections>100 and <1	Number of sections<100 and >1	Invalid values	Values accepted	Pass
Overlapping regions	Overlapping regions>100 and <0	Overlapping regions<100 and >0	Invalid values	Values accepted	Pass

TABLE 1: Performance based on difference parameter

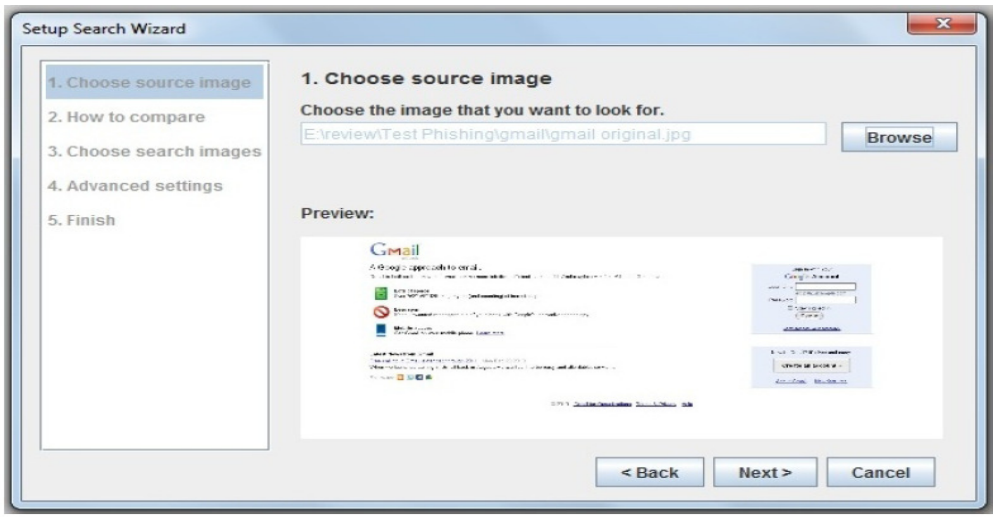

FIGURE 9: Choosing Source Image

In Figure.9, Choose the source image by clique the browse option to display the path. From the root directory required source can be identified.

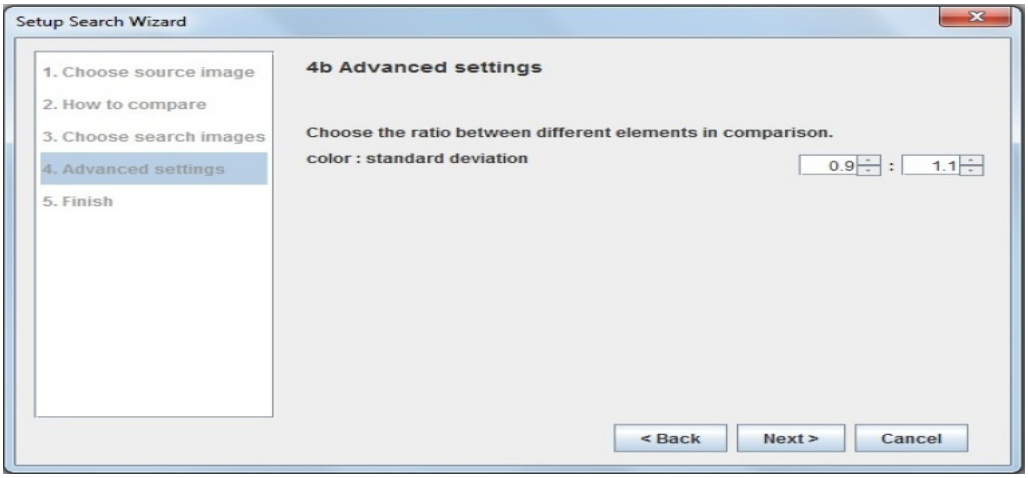

FIGURE 10: Choosing the Parameters

In Figure.10, shows the ratio between different elements in comparison for identification. The standard color deviation gives the appropriate ratio.

FIGURE 11: No Images in the Directory

In Figure.11, shows after selecting the root directory, the appropriate image has to be selected otherwise the pop-up menu will be displayed

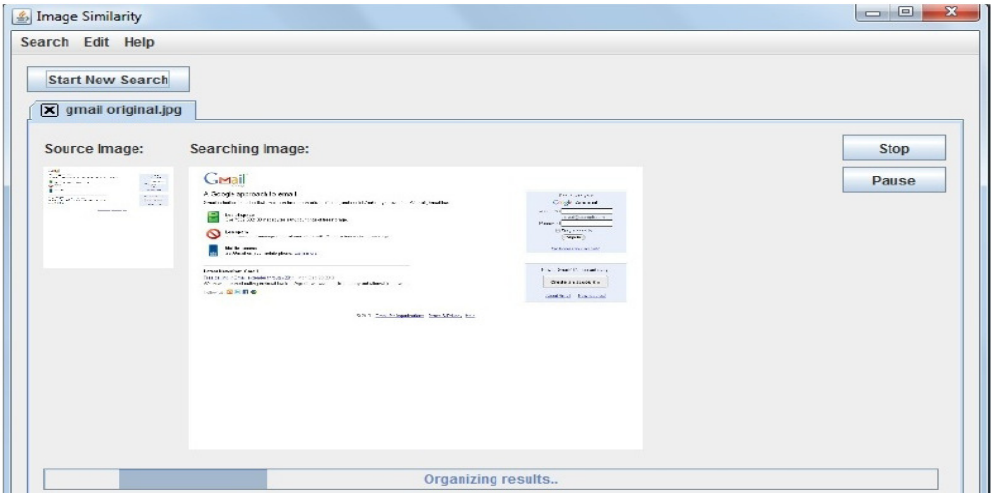

FIGURE 12: Image Similarity

In figure.12, Gives the similarity of search image and existing image. This method is the easy way of find the phishing page. This will improve the efficiency of the web page.

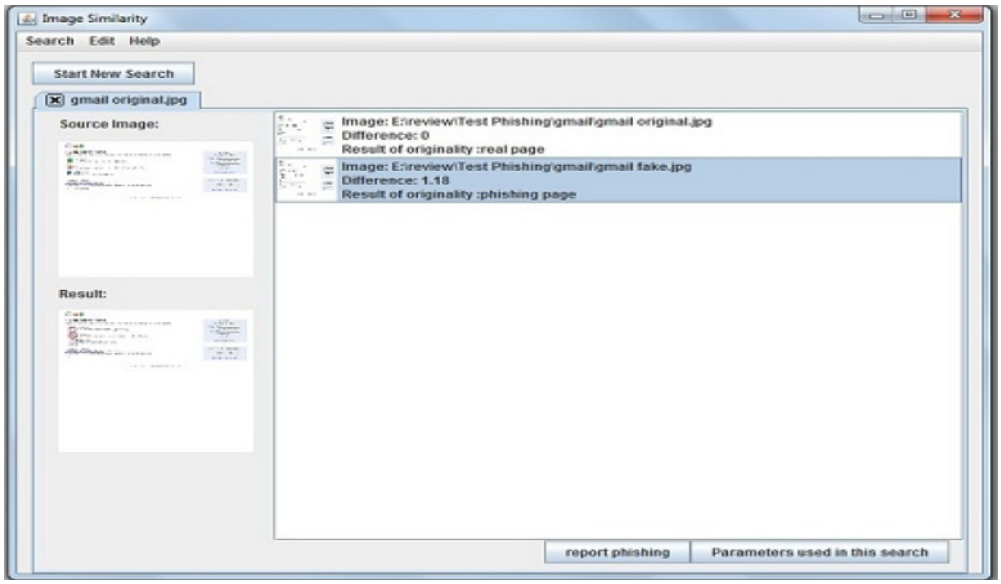

FIGURE 13: Displaying the Difference Value

In Figure.13, based on the difference values can identify the original and fake pages. If the difference value is 0, then real page will be displayed. The difference value is other than 0 will produce phishing page.

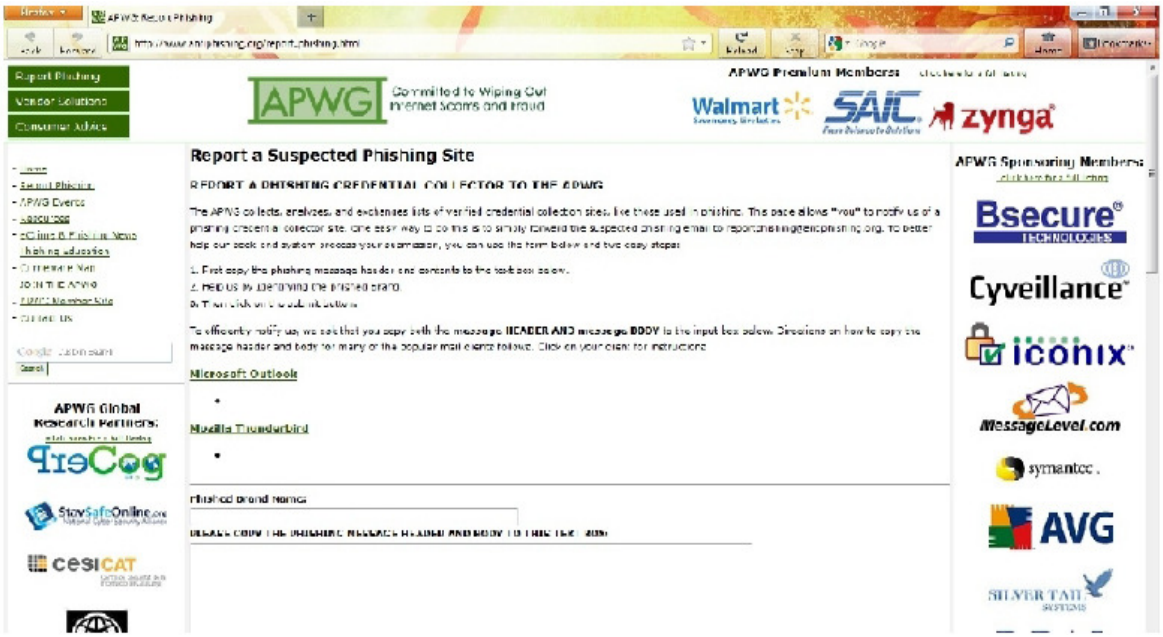

FIGURE 14: Reporting to the APWG site

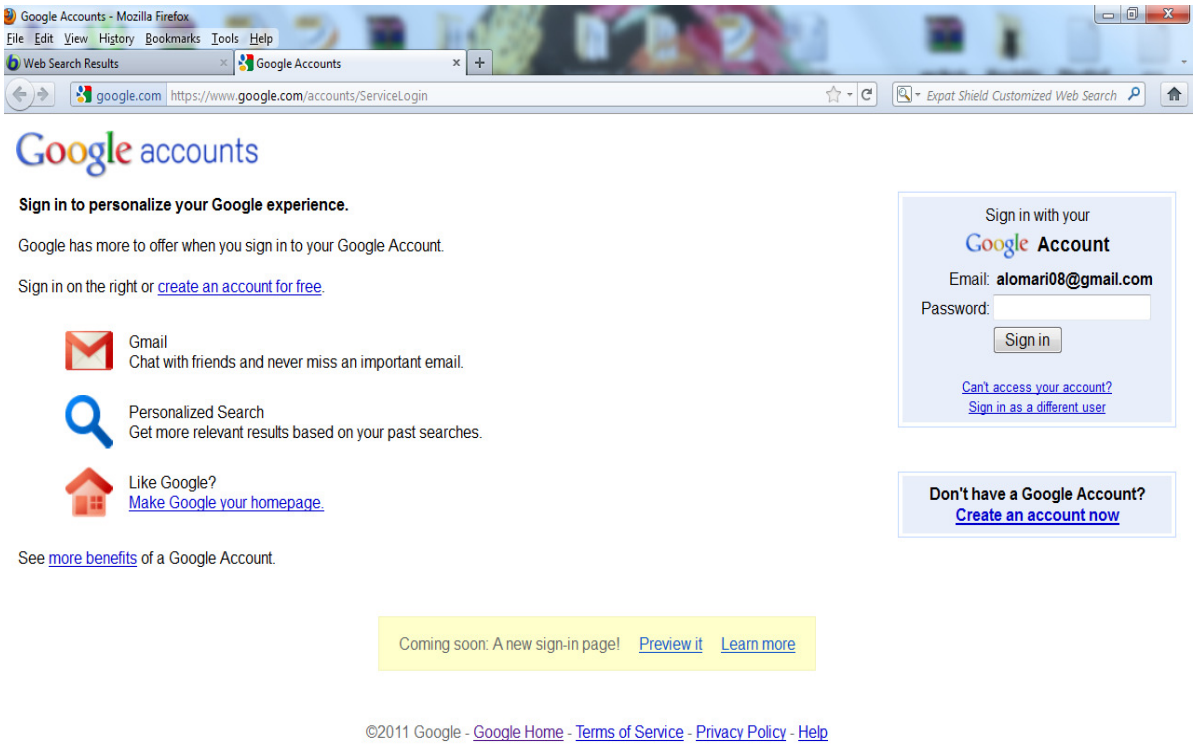

FIGURE 15: Gmail Original Page [www.google.com]

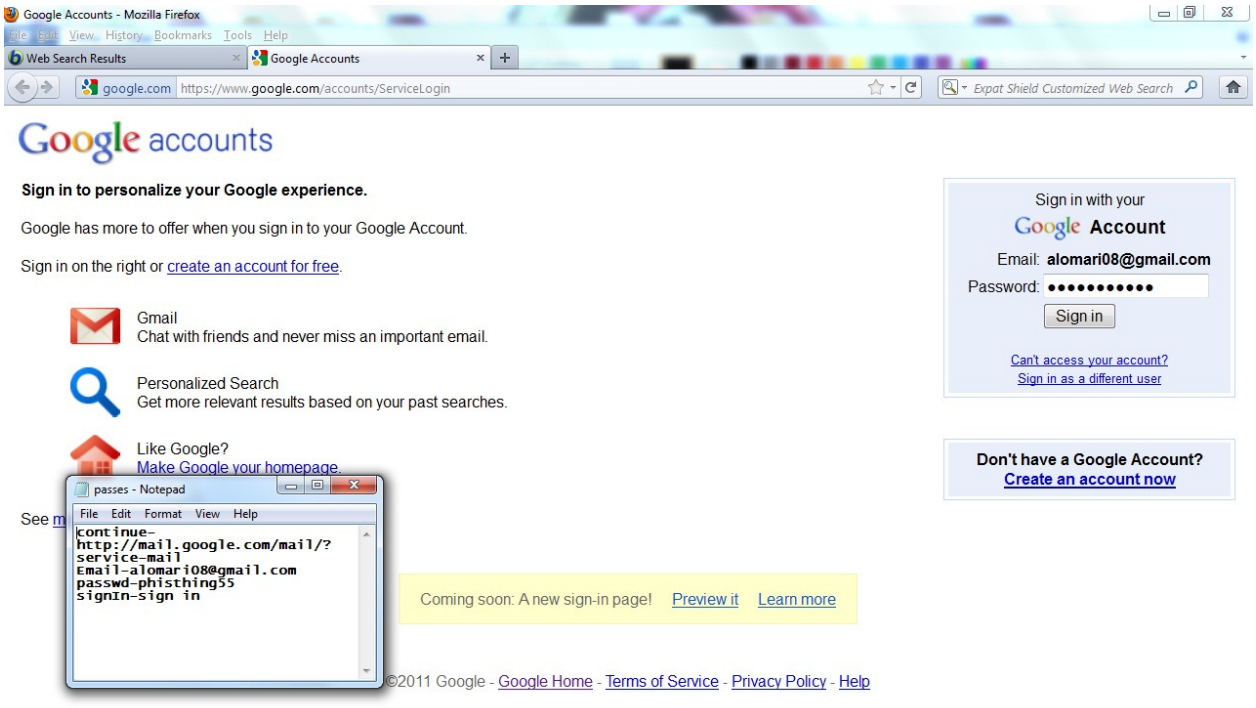

FIGURE 16: Gmail Fake Page

To prevent the phishing attack, the snapshot of the web page that appears before the user in taken. Then the web page is stored as an image in a directory. The anti-phishing tool takes the

snapshot of the original web page and stores it. The original page is chosen from the directory. Then both the images are compared. If the difference value is zero then the page is a “real page” else if the value is a non-zero then the page is a “phishing page”. If the page is reported as phishing page then select report to phishing button which redirects to APWG website where the user can report the page as phishing page and that page will be added to the list of phishing pages.

Techniques	Year	Proposed Work	Result
Spear Phishing	2005	Spear phishing is an e-mail spoofing fraud attempt that targets a specific organization, seeking unauthorized access to confidential data.	92.56%
CCH	2006	A contrast value is defined as the difference in intensity between a point and the salient corner.	98.19%
Whaling	2008	The cybercrime practice of phishing — masquerading online as a trustworthy source to try to steal people's sensitive information — is coming up against some serious competition in the form of "whalers".	96.37%
Spoofing	2004	Spoofing attack is a situation in which one person or program successfully masquerades as another by falsifying data and thereby gaining an illegitimate advantage.	98%
Tabnabbing	2009	Tabnabbing is a computer exploit and phishing attack, which persuades users to submit their login details and passwords to popular web sites by impersonating those sites and convincing the user that the site is genuine.	97.98%

TABLE 2: Phishing techniques compared with other related works

From the above table 2, The CCH performance is high when compared with other techniques like spear phishing, whaling, spoofing and tabnabbing. Tabnabbing is one of the recent techniques introduced in phishing attacks. Whaling is using for cybercrime of phishing, online information can try to steal by the people. Spoofing is well know phishing attacks used to identify false data, is used to improve the efficiency of the web page.

5.1 Non Functional Requirements

- ***Portability:*** This tool is platform independent; it can run in any operating system.
- ***Efficiency:*** It is very speed in nature because this tool does not contain any databases. So, CPU cycles will not be wasted in retrieving data from database.
- ***Time:*** Time is a main constraint in a work. The comparison is done only as per the requirements of the user.

- ***Usability:*** This tool is very easy to use because of its user friendly comparison wizard. Even a normal person can use it without any difficulty.
- ***Scalability:*** In future we can update the tool by adding some extra features. This tool will function properly irrespective of any update.
- ***Performance:*** It can perform high even if there are more images to compare.
- ***Error Handling:*** When there is no image in the specified directory then the wizard will tell the user as "no image". It is robust in nature.
- ***Accessibility:*** It is easily accessible because it is in the universal language English.
- ***Accuracy:*** This tool can detect even a small difference between the images because it matches the color ratio of the images. So, it is highly accurate.
- ***Capacity:*** It can hold many numbers of images to compare.
- ***Visibility:*** The visibility is good. The font used is bigger in size. Each panel contains instructions which will lead the user to the next step.

6. CONCLUSIONS

Nowadays, all activities like banking, shopping, etc. are carried out only using internet. There are more chances for the phishers to steal the information from the user. So security plays a major role. This project is developed to prevent attacks like phishing attack. By this attack the attackers steal the personal information of a user and misuse it. To avoid phishing attack, here we proposed a color based image comparison method is developed. To prevent phishing attacks there are methods which are inefficient. All methods uses only text based comparison which is not error free because the attackers has started to insert images which looks similar to that of the original image. So, by text based comparison the difference between the real and the fake page cannot be found. Color is the most important feature in an image. So, in this project we have developed an image based comparison method which compares the images based on the color values. Only the company which created that website knows about the color range of the images present in the web page. None can design a fake web page similar to the original page with that same color range. So, by comparing images using color values will give an accurate result. Thus, this anti-phishing tool is highly efficient and error free. This anti-phishing tool can be used in online banking, online shopping and to maintain the mail accounts. Even when there is a small variation in the web page this tool can find it and report to the user and another main advantage is that the user need not waste time by searching internet to report the page to APWG. Instead a button is embedded in this tool which will redirect the user to the APWG web site. In Future work, can develop a fully automated crawling framework by using attribute-based phishing attacks that developed for testing, along with main experimental results.

7. ACKNOWLEDGMENT

Sincere thank and recognition goes to my advisor, Associate Professor, Dr. Putra Sumari, who guided me through this research, inspired and motivated me. We also thank the Universiti Sains Malaysia USM for supporting this research.

8. REFERENCES

[1] A. Kannan, V. Mohan and N. Anbazhagan. "Image Clustering and Retrieval using Image Mining Techniques". *IEEE International Conference on Computational Intelligence and Computing Research*, *vol.2, 2010*

[2] SOPHOS 2005, http://*www.sophos.com/whitepapers/sophos-phishing-wpuk.pdf,* accessed *April 2011*

[3] M. Jakobsson, and S. Myers: 'Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft' *Wiley*, *2007*

[4] W. Burger and M. Burge. "Digital image processing: an algorithmic introduction using Java". *Springer, Pages: 240-250, 2008*

[5] S.R. Kodituwakku et al. "Comparison of Color Features for Image Retrieval". *Indian Journal of Computer Science and Engineering, vol.1, no.3, pp.207-211, 2004*

[6] APWG, http://www.antiphishing.org/index.html, accessed March 2011

[7] Wikipedia, http://en.wikipedia.org/wiki/Phishing, accessed April 2011

[8] Webopedia, http://www.webopedia.com/TERM/P/phishing.html, accessed April 2011

[9] M. Aburrous, M.A.Hossain, Keshav Dahal and Fadi Thabtah. "Experimental Case Studies for Investigating E-Business Phishing Techniques and Attack Strategies". *Springer Science, Cong Comput 2010, vol.2, No.242-253, April 2010*

[10] APWG. http://www.apwg.org/reports/APWG_GlobalPhishingSurvey_1H2009.pdf , accessed 8 August 2009

[11] M. Chandrasekaran, K Narayanan and S Upadhaya,"PHONEY:Mimicking User Response to Detect Phishing Attacks", *To appear at TSPUC Workshop, affiliated with IEEE WoWMoM, 2005*

[12] K. Chen, C. Huang and C. Chen. "Fighting Fishing With Discriminative Keypoint Features". *IEEE INTERNET COMPUTING, 2009*

[13] K. Plossl, H. Federrath and T. Nowey. "Protection Mechanisms Against Phishing Attacks". *Proc, 2nd Intl.Conf. on TrusBus 05, LNCS 3592, Springer-Verlag, 2005*

[14] M. Wu, R.C.Miller, S.L.Garfinkel, "Do security toolbars actually prevents phishing attacks?", *in CHI (to appear), 2006*. [online]. Available: http://www.simson.net/ref/2006/CHI-security-toolbar-final.pdf

[15] S. Kierkegaard, "Swallowing the bait, hook, line and sinker: Phishing and Pharming and now rat-ting!", in Managing Information Services in Financial Services H.R. Roa, M. Gupta, S. J. Upadhaya, Eds.USA:IGI publishing, 2008, pp.241-253.

[16] N.P. Singh. "Online Frauds in Banks with Phishing". *Journal of Internet Banking and Commerce, vol.12, 2007*

[17] Phishtank. 2008 http://www.phishtank.com/phish_archive.php, accessed 14 November 2008

[18] A. Abbasi and H. Chen. "A comparison of fraud cues and classification methods for fake escrow website detection". *Springer, Inf Technol March, 2009*

[19] R. Kanthety and S. Saradhi. "Prevention of Phishing Attacks using Link-Guard Algorithm". *International Journal of Computer Science Issues (IJCSI). vol. 7, no. 2, suppl.4, 31p.March 2010*

[20] A. Martin, Na.Ba.Anutthamaa, M. Sathyavathy, Marie Manjari Saint Francois and Dr. Prasanna Venkatesan. "A Framework for Predicting Phishing Websites Using Neural Networks". *International Journal of Computer Science Issues (IJCSI). vol. 8, Issue 2, March 2011*

[21] Juan Chen and Chuanxiong. "Online Detection and Prevention of Phishing Attacks". IEEE Communications and Networking, NSFC, 2005

[22] Bryan Parno, Cynthia Kuo, and Adrian Perrig. "Phoolproof of Phishing Prevention". Financial Cryptography and Data Security, Springer, 2006

[23] Total Number of Fraud Complaints & amount paid. 2003, http://www.consumer.gov/sentinel/states03/fraud_complaint_trends.pdf.

[24] Tom Jagatic, Natheniel Johnson, Markus Jakobsson, and Filippo Menczer. “Social Phishing”. Communications of ACM, 2005

[25] Thomas J. Holt and Danielle C. Graves. “A Qualitative Analysis of Advance Fee Fraud E-mail Schemes”. International journal of Cyber Criminology, vol.1, issue.1, 2006

[26] http://mybank.com/ebanking]