ANTI-PHISHING MECHANISM USING IMAGE CAPTCHA.

Nirbhay Shah[1], Shyam Shah[2], Jaikit Jilka, [3] Yash Tare[4], Amruta Pokhare[5]

Information Technology Department, Mumbai University
Atharva College of Engineering, Malad, Mumbai, India

nirbhay.r.shah@gmail.com,
shyamshah2210@gmail.com,
jaikit54@gmail.com,
yash_tare@yahoo.in,
amrutapokhare09@gmail.com

Abstract- **With the emergence of a new era where a large number of transactions take place over the internet, there arouse new threats and security issues for the information being transferred over the network. One such threat is 'phishing' where in a mischievous or dangerous entity might acquire confidential and sensitive information passed over the network, by masquerading as a trustworthy system (a trustworthy website in particular). The user might end up presenting all his/her confidential information like user-id, password, and bank-account details, etc. without doubting the authenticity of the system. This is a concern for both the user as well as the rightful organisation as their reputation is at stake. We thereby put forth a technique called "Image CAPTCHA Authentication, Based on Visual Cryptography" which not only ensures the candid nature of the user but also tests the website for its genuineness. Here an original CAPTCHA asked from the user, after presenting it to him/her prior to the registration, is decomposed into 2 shares, of which one share is kept with the user and the other with the server. Neither share of the CAPTCHA reveals the original one until the user logs in into the server and both present their share to prove their truthful nature. Thus we can mitigate phishing effectively.**

Keywords— **CAPTCHA, Visual Cryptography, Security, Phishing, Anti-Phishing, Authenticity, Genuine, User Share, E-commerce, Banking, Website, Theft, Registration, Login, Database**

I. Introduction

Online Commerce or E-Commerce is the new term in the market. People are getting their hands on this tool, but at the same time falling prey to various malicious activities that are affecting the privacy of the system. Phishing is one such malicious activity that can harm the user. So various mechanisms have to be developed to safeguard the user from the online thieves or also called as attacker.

II. ABOUT PHISHING

Basically in Phishing the phisher tries to get the information of the user by posing himself as a genuine online entity, to which the user is bound to give the data. So, sometimes it becomes difficult for the user to differentiate between the genuine entity and fake entity. This gives rise to Anti-phishing, there are various tools to fill the anti-phishing arsenal and one of them is CAPTCHA.

Phishing web pages are forged web pages that are created by mischievous people to mimic Web pages of real web sites. These web pages tend to be pretty much like an exact replica of the original web pages. In fact intelligent phishers use automated tools which download the real web pages instantaneously at the moment when the user unintentionally visits the phished page and then those phishers present phished pages with the latest of the contents of the real site to the users. Thus, all the updates, recent announcements, news etc. of the real site are reflected on the phished pages, leaving no place for doubt in the user's mind. The victims may end up presenting their bank account details, password, credit card number, or other important information to the phishing web page owners. Phishers normally use e-mails and spam messages that claim to offer you a false jackpot or ask you to update your bank-account details. Techniques such as man in the middle attacks, installation of key loggers and screen captures are common. So a better system is required to counter phishing attacks effectively and efficiently.

III. WHAT IS CAPTCHA

CAPTHCHA stands for Completely Automated Public Turing Test to Tell Computer and Humans Apart. It is basically a string of characters that is turned and twisted into an image that cannot be filled or solved by computers. So, computerized theft or robotic intrusion cannot surpass it. Nowadays it is an integral part of all forms for online registration for a particular web site.

IV. Visual Cryptography

One of the best known techniques to protect data is cryptography. It is the science of transferring encrypted messages that can be decrypted only by the rightful receiver. Encryption and decryption of the data sent over the network are accomplished by using mathematical algorithms in such a way that no one but the intended recipient can decrypt and read the message. Naor and Shamir [2] introduced the visual cryptography scheme (VCS) as a simple and secure way to allow the secret sharing of images without any cryptographic computations.

V. Existing Views

In the current scenario the user has very less chances to differentiate between the genuine website and the phishing i.e. fake website. Whenever the user logs into his account he provides his username and password, so the data is given just by the user and there is no information provided by the server, just the user authenticity is checked. This vulnerability makes a host-point for the phisher to attack. So phisher can easily develop a similar page but with a different URL and snatch the users crucial credentials. But if the user is smart enough he can check the URL and safeguard his or her credentials. The existing scenario I depicted in the diagram 4.1.

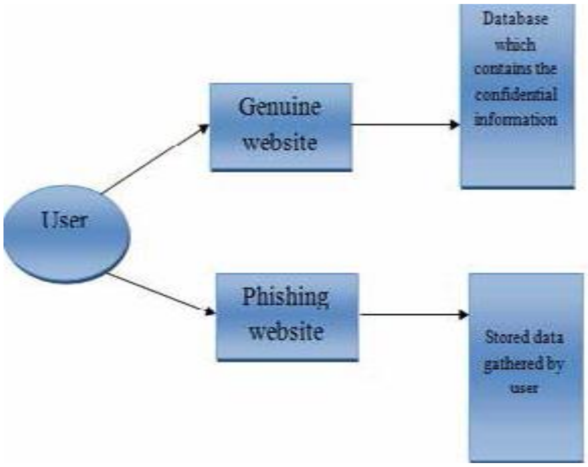
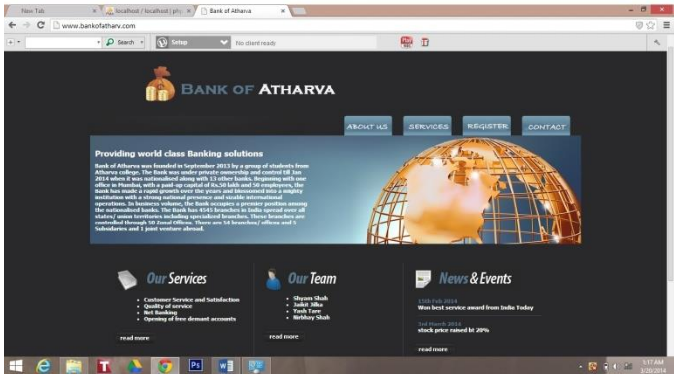

Database which contains the confidential information
Genuine website
User
Phishing website
Stored data gathered by user

Fig. 4.1 Existing Scenario

The phisher may send a hoax mail to the user asking him to fill in the user credential details for security purposes. Then as shown in the figure instead of the genuine it he will be redirected to the fake website which has a similar design and layout just the URL will be different. Figure 4.2 shows an example of the fake website which is based on the genuine website which is shown in figure 4.3.



Providing world class Banking solutions

Bank of Atharva was founded in September 2013 by a group of students from Atharva college. The Bank was under private ownership and control till Jan 2014 when it was nationalised along with 13 other banks. Beginning with one office in Mumbai, with a paid-up capital of Rs.50 lakh and 50 employees, the Bank has made a rapid growth over the years and blossomed into a mighty institution with a strong national presence and sizable international operations. In business volume, the Bank occupies a premier position among the nationalised banks. The Bank has 4545 branches in India spread over all states/ union territories including specialized branches. These branches are controlled through 50 Zonal Offices. There are 54 branches/ offices and 5 Subsidaries and 1 joint venture abroad.

Our Services

- Customer Service and Satisfaction
- Quality of service
- Net Banking
- Opening of free demant accounts

Our Team

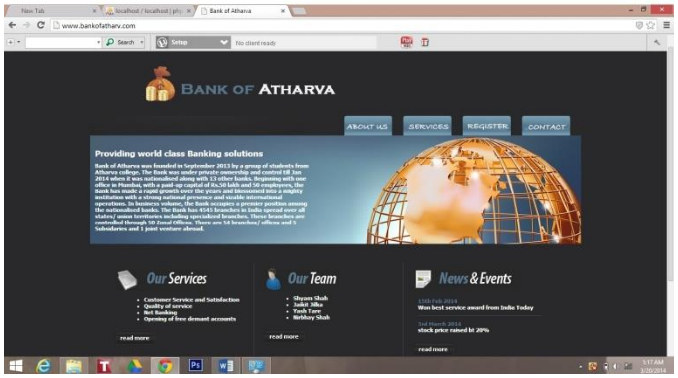
- Shyam Shah
- Jaikit Jilka
- Yash Tare
- Nirbhay Shah

News & Events

15th Feb 2014
Won best service award from India Today

3rd March 2014
stock price raised bt 20%

Fig. 4.2 Fake website



Providing world class Banking solutions

Bank of Atharva was founded in September 2013 by a group of students from Atharva college. The Bank was under private ownership and control till Jan 2014 when it was nationalised along with 13 other banks. Beginning with one office in Mumbai, with a paid-up capital of Rs.50 lakh and 50 employees, the Bank has made a rapid growth over the years and blossomed into a mighty institution with a strong national presence and sizable international operations. In business volume, the Bank occupies a premier position among the nationalised banks. The Bank has 4545 branches in India spread over all states/ union territories including specialized branches. These branches are controlled through 50 Zonal Offices. There are 54 branches/ offices and 5 Subsidaries and 1 joint venture abroad.

Our Services

- Customer Service and Satisfaction
- Quality of service
- Net Banking
- Opening of free demant accounts

Our Team

- Shyam Shah
- Jaikit Jilka
- Yash Tare
- Nirbhay Shah

News & Events

15th Feb 2014
Won best service award from India Today

3rd March 2014
stock price raised bt 20%

Fig. 4.3 Genuine Website

This is how the user may get influenced by the phisher and may end giving him the essential credentials to him. This is the one of the main vulnerability which needs to be addressed immediately to safeguard the security of the user.

Certain techniques have been developed in this regard. However they are not very effective. These techniques are:

VI. Existing flaws

- Only data is taken from the user during authentication, no crucial information is provided by the genuine website server to prove or take make the user confident about the authenticity of the website.
- User awareness regarding the phishing activities and understanding the need to secure their crucial credentials.
- Lack of efforts to reduce the phishing attacks.
- High cost of the existing anti-phishing techniques.

VII. Existing Approaches

A. Detect and Block the phishing Websites in Time

The first step to prevent phishing is to realize that we are being a victim of the same, at the right time. It is because the phisher might be sitting in any part of the world. If we end up giving our confidential information without knowing the nature of the website we are dealing with, by the time we know that we have become a victim of phishing, the phisher would have taken away our money or misused the information in some other way and also would have fled beyond our reach in any way. Since its' the matter of reputation of the genuine websites, they too are required to keep a check on the possibilities of phishing attacks to their users. For this their server needs to scan the root DNS from time to time and see if there is any kind of domain name similar to theirs. If so they can alert the user preventing him/her to mistake that website as theirs. Detecting Phishing manually is also an option with the users. Smart phishers however use automated tools to download the genuine website's content at that moment to deceive the user. The website entirely resembles the original one with all latest updates.

B. Enhance the Security of the Websites

Some large banks use different techniques to prevent phishing of their websites. Hardware devices are used by certain companies. For example, the Barclays bank provides its' users with a credit card with a PIN. Prior to any online shopping, the users have to insert their credit card into the card reader, and input their PIN code. The reader generates a onetime password and sends it to the user. After putting in the password the user is free to shop. Biometric techniques can be thought of as an option. Various unique characteristics of humans like finger prints, voice, and iris scanning etc. can be taken into consideration. PayPal uses voice recognition for security. However it must be understood that the user might operate from any environment (home, cybercafé etc.). It might not be feasible for the user to install the hardware required for these systems. Also it is a bad practice to ask the user to take efforts especially on client PC.

C. Taking help of Spam Filters in our mail box, to distinguish a potentially phished mail from normal ones

Amongst the various ways to make a potential victim fall in the trap of phishing, 'email' is the most common and effective one. An e-mail that seems to come from a trusted source or from a site where in we hold an account, requesting to update our information at a link provided or an email the allures us with a jackpot are some kind, are some of the traps. Phishers are seen to have a tendency to be able to easily counterfeit their identity as a trusted sender identity. Amongst the large number of potential victims who receive these bulk emails, very few seem to realize or even suspect the genuineness of the sender. The Simple Mail Transfer protocol (SMTP) gives the phishers a liberty to do so. Since it is not secure enough and does not possess any counter measure to block e-mails, possibly coming from malicious sources. One solution to this is enhancing the ability of our spam filters to determine whether the e-mails that have come to us are from truthful senders. i.e. (They are actually the ones whom the claim to be). This would eliminate most of the risks of becoming a phishing victim. Such attacks would then reduce considerably.

Certain techniques that are developed in this respect are Microsoft's SIDF. It is like the sender's ID helping the filters to distinguish various senders. SIDF is a combination of Microsoft's Caller ID for E-mail while the SPF (Sender Policy Framework) is another tool developed by Meng Weng Wong. Both these tools check the domain name of the e-mail's sender to verify if the e-mail is sent from an authorized server and in turn to determine whether that e-mail has a spoofed address of the sender. If so, the ISP (Internet Service Provider) can then declare that particular e-mail as a spam e-mail. Phishing can thus be greatly reduced. The spoofed e-mails used by phishers are one type of spam e-mails. Thus, spam filters help us greatly to curb phishing attacks. Some more techniques that are developed are, blacklist, whitelist, keyword filters, Bayesian filters (possessing self-learning abilities), E-Mail Stamp, etc. These techniques can all be used at both server level and client level systems. Generally such techniques perform filtering at the client side by scanning the contents and the address of the received e-mails. They are further discussed below in detail. Blacklist-whitelist techniques fail if the names of the spammers are not known in advance while Keyword filter and Bayesian filters can detect spam based on content and hence can detect even unknown spasm. However even these techniques can mislead resulting in false negatives and false positives.

D. Installation of Software on Client PC to Check Phishing

To further ensure the protection against the phished e-mails that still manage to penetrate in our PCs, certain online tools must be installed. These tools can be divided into 2 types: blacklist-whitelist based tools and rule-based tools.

1) Type 1: The developers of these tools maintain a database called blacklist, consisting of those URLs which have a bad record in history and might lead to a phished page. Against this they also maintain a whitelist which holds the list of secure URLs which the user can trust. Whenever the user visits a particular website, the URL in the address bar is checked and accordingly the user is notified if that URL is blacklisted. Thus the user would refrain from staying on the website or providing any information to it.

2) Type 2: These techniques use certain rules in their software to decide whether a particular website is phished or not. Such types of tools are Spoof Guard developed by Stanford,

TrustWatch of the Entrust, etc. Spoof Guard scans the domain name and the URL, including the port number of Web site. It also verifies if the browser is directed to the current URL via the links in the contents of e-mails. On noticing that the domain name of the website appears somewhat similar to that of a well-known website which could have been phished by this particular domain name or that a standard port is not used, the Spoof-Guard immediately notifies the user about it. TrustWatch completely depends upon some third party. A trusted third party organization evaluates the website on the grounds of trustworthiness and determines whether the website is phished or not. In both the above tools a separate toolbar is provided to notify the analysis of the web site.

VIII. Proposed Methodology

Based on the various existing techniques and with an attempt of overcoming their glitches, we propose a new methodology to know if a particular website is phished or is a genuine one.
The process is split into two phases:

A. Registration Phase

This is a phase wherein the website and the user would develop an understanding of each other, which would, in future, guarantee each other’s authenticity. Let us say that a user is creating an account on a website that could be providing any type of service, such as e-banking, e-commerce, etc. While registering, apart from all the other information that the user trusts the website with, an additional random sequence of characters, displayed in the form of a CAPTCHA is asked from the user. At the server level, this CAPTCHA is broken into 2 shares and stored in the database along with its shares, against that particular user. Accordingly the user is notified about his/her CAPTCHA share along with the entire CAPTCHA through an e-mail and an SMS. The user needs to remember both of these to prove his/her authenticity.

B. Login Phase

In future when the user wishes to enter his/her account, to access some confidential information, make monetary transactions or to update his account etc., he/she needs to enter the username along with the user’s CAPTCHA share.

Correspondingly the entire CAPTCHA of the user is displayed by that website guaranteeing its truthful nature. This would suggest the user to proceed further by entering the password and other confidential information and to carry out the desired tasks. However a failure to display the correct CAPTCHA image would mean that the website is a phished one and that the user should immediately stop to stay on that page. We strongly suggest that the process of displaying the CAPTCHA image should be triggered on pressing the key in the text space of the user’ CAPTCHA share (key listener) rather than having to press the submit button. This is because a phisher might even get the CAPTCHA image after a certain attempts.

Thus with this technique one can verify whether the website is genuine/secure website or a phishing website and the website can also verify whether the user is a human or an automated system. This phase is shown in Fig.4.

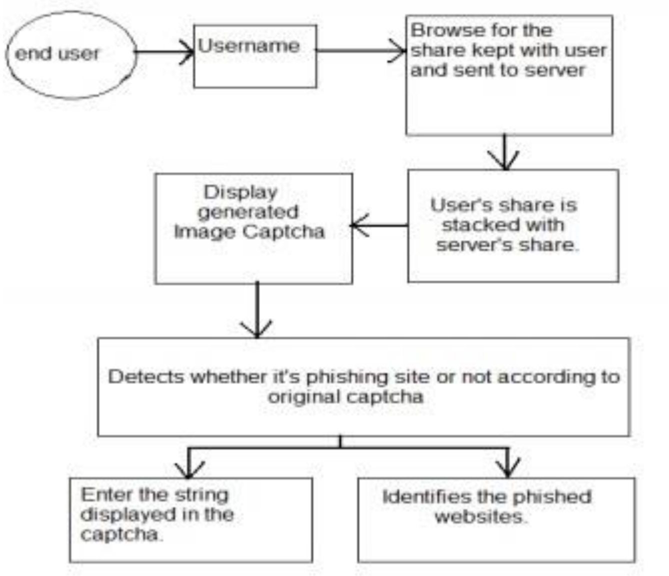

end user
Username
Browse for the share kept with user and sent to server
User's share is stacked with server's share.
Display generated Image Captcha
Detects whether it's phishing site or not according to original captcha
Enter the string displayed in the captcha.
Identifies the phished websites.

Fig. 4 Login Phase

IX. IMPLEMENTATION AND ANALYSIS

This project is implemented on a dummy banking website developed by us, to prove the effectiveness and the extent of this authentication system.

We have developed three websites one is a genuine website, similar to the other websites and like their authentication methods we had developed the registration and login phase. Then we had developed phishing website that influences the user to provide their phishing website users the crucial credentials.

And then we developed a banking website with our methodology to reflect the changes and improved efficiency in the field of security and user service.

The entire project was tested on the WAMP server and later it was implemented on the actual server to check for any performance or last minute error that may come up.

X. Conclusion

Over these years phishing has emerged as great threat to all those who carry out monetary transactions or pass sensitive information on the World Wide Web. Mitigation of phishing attacks remains to be a major concern for all the large business entities active on the internet. Thus, considering the scale of the attacks, it can be inferred that phishing is undoubtedly the gravest threat to the users, all over the globe, as compared to other traditional ones. In a research submitted at the Conference on Human Factors in Computing System, the authors suggest that even in the best of environment of knowledgeable individuals, a perfectly phished site could deceive more than 90% of the users. Very few users seem to care about the indicating factors of phishing which were inbuilt in the browsers that were used. There arises a need to educate the users all over the world regarding such issues. 'The Anti Phishing Working Group' is an organization attempting to create awareness among people. The organization maintains an archive of all the phishing attacks in history and any phishing victim can report about the attack over there. As far as our proposed methodology is concerned, it checks the credibility of the website giving user a confidence in the transaction that he/she performs. It is a simple framework that could be plugged in into any web site. The image CAPTCHA that is displayed to the user by the website serves as the authenticity of that site. It is only then, will the user be able to enter further details like password etc. A phished website however would not be able to provide the actual CAPTCHA in any case, failing to win the user's trust for that transaction. The CAPTCHA which is then validated at the server proves whether the user is a human or an automated system. Thus this way both the parties gain confidence in each other and the transaction is completed successfully. In future, we look forward to curb phishing and having more awareness and alertness among all the active entities on the internet.

REFERENCES

[1] Ollmann G., the Phishing Guide Understanding & Preventing Phishing Attacks, NGS Software Insight Security Research.

[2] M. Naor and A. Shamir, "Visual cryptography," in Proc. EUROCRYPT, 1994, pp. 1–12.

[3] A. Shamir, .How to Share a Secret,. Communication ACM, vol. 22, 1979, pp. 612-613.

[4] G R. Blakley, .Safeguarding Cryptographic Keys,. Proceedings of AFIPS Conference, vol. 48, 1970, pp. 313-317.

[5] A. Menezes, P. Van Oorschot and S. Vanstone, .Handbook of Applied Cryptography, CRC Press, Boca Raton, FL, 1997.

[6] B. Borchert, .Segment Based Visual Cryptography,. WSI Press, Germany, 2007.

[7] W-Q Yan, D. Jin and M. S. Kanakanahalli, .Visual Cryptography for Print and Scan Applications, IEEE Transactions, ISCAS-2004, pp.572-575.

[8] T. Monoth and A. P. Babu, .Recursive Visual Cryptography Using Random Basis Column Pixel Expansion,. in Proceedings of IEEEInternational Conference on Information Technology, 2007, pp. 41-43.

[9] Antiphishing Framework theglobaljournals.com/paripex

[10] H. J. Kim, V. Sachnev, S. J. Choi and S. Xiang, .An Innocuous Visual Cryptography Scheme, in Proceedings of IEEE-8th International Workshop on Image Analysis for Multimedia Interactive Services, 2007.

[11] C. Blundo and A. De Santis, .On the contrast in Visual Cryptography Schemes, in Journal on Cryptography, vol. 12, 1999, pp. 261-289.

[12] P. A. Eisen and D. R. Stinson, .Threshold Visual Cryptography with specifed Whiteness Levels of Reconstructed Pixels,. Designs, Codes, Cryptography, vol. 25, no. 1, 2002, pp. 15-61.

[13] E. R. Verheul and H. C. A. Van Tilborg, .Constructions and Properties out of n Visual Secret Sharing Schemes, Designs, Codes, Cryptography, vol. 11, no. 2, 1997, pp. 179-196.

[14] H. Yan, Z. Gan and K. Chen, .A Cheater Detectable Visual Cryptography Scheme, Journal of Shanghai Jiaotong University, vol. 38, no. 1,2004.

[15] G. B. Horng, T. G. Chen and D. S. Tsai, .Cheating in Visual Cryptography, Designs, Codes, Cryptography, vol. 38, no. 2, 2006, pp. 219-236.

[16] C. M. Hu and W. G. Tzeng, .Cheating Prevention in Visual