Disclosure E-Mail of Phishing Website

Ahmed Hashim Mohammed

Abstract:

Phishing (pronounced "Fishing") is an online fraud technique used by criminals to entice client to disclose his personal information and its emergence in online communications media allows scammers to reach more people than ever before and at a lower cost, whether through spam, e-mail , instant message scams , faked Web pages or other online avenues for stealing personal finances. To disclose an e-mail of phishing website, three levels have been proposed. The first level where a user assigns static password that embed with all legitimate emailing in the title of message . The second level is to summarize the last visited history to websites and in the third level its server requests a login color from the user without any other requests of confidential information.

In contrast to other proposals, this scheme places a very low burden on the user in terms of effort, memory and time. It also places a high burden of effort on an attacker to spoof personalized security indicators, antiphish is implemented as a client-server by html, php, mysql database, and apachserver.

الكشف عن الرسائل الالكترونية للمواقع الالكترونية المزيفة

م.م أحمد هاشم محمد
قسم الحاسوب/كلية التربية/الجامعة المستنصرية

المستخلص:

Phishing هي تقنية احتيال على الإنترنتِ مستعملة من قبل محتالين لإغْراء الزبون إلى كشف معلوماته الشخصية، وفضحها في أجهزة إعلام الاتصالات على الإنترنتِ والتي تسمحِ للمخترقين الوُصُول الى المزيد مِنْ الناس أكثر من أي وقت مضى، وفي كلفة اقل، سواء من خلال: رسالة الدعاية، بريد إلكتروني، رسالة فورية ، تزييف صفحات الويب أَو إي خدع أخرى على الإنترنتَ الآخرَ لسَرِقَة المعلومات الشخصية للدِفَاع ضدّ هجمات phishing ذات ثلاثة مستويات تم اقتراحها ،المستوى الأول كلمة مرور الزبون تضاف مع عنوان كل الرسائل الحقيقية، والمستوى الثاني: إضافة ملخص أخر زيارة للموقع، والمستوى الثالث: هو إن يطلب الخادم من الزبون فقط اللون السري وليست أية معلومات شخصية أخرى.بالمقارنة مع الاقتراحات الأخرى، هذا المخطط المقترح: يتطلب جهداً ووقتاً اقل من قبل الزبون، وبنفس

الوقت يتطلب النموذج المقترح جهداً ووقتاً أكثر على المهاجم لتزييف الرسائل، تم تطبيق ال antiphish باستخدام لغات ، html, php وقاعدة بيانات mysql والخادم apachserver.

1.Introduction

As people increasingly rely on Internet to do business, Internet fraud becomes a greater and greater threat to people's internet life. Internet fraud uses misleading messages online to deceive human users into forming a wrong belief and then to force them to take dangerous actions to compromise their or other people's welfare. The main type of internet fraud is phishing [1]. Phishing, from "password" and "fishing", is an attempt to get sensitive information, like user names, passwords, credit card information or complete identities from the recipients. This is based on social engineering by masquerading a trustworthy entity in electronic communication and fooling users to submit their data by the use of special techniques, mostly in emails with hyperlinks to forged websites which look nearly like the original ones. Common targets are banks, auction portals or payment services [2].

In this paper, the proposed system authenticates the legitimate website , in order to prevent phishing website and to satisfy that, the user only needs to perform a visual matching operation by recognizing only one static password that is set previously as legitimate for emailing and remember his headline history of last visited websites such as (account, time enter website, time exit website, number of books bought ,types of books), the server generates a unique dynamic password for each user and each emailing to create a trusted outcoming link that provides automatic connection to a website so that the legitimate website never requests creditional information such as (username , password and credit card etc...) from the user but at the same time, the server requests a login color from the client to avoid hacking threats.

The next section reviews related work. And section 3 describes type methods for detection of website phishing while section 4 describes a proposed approach and provides details about the implementation of system , section 5 presents the experimental results that show approach is feasible in practice , and concludes the paper while section 6 offers a suggestion of future of work.

2. Related Work

Several proposals have been made to help internet surfers detect phishing websites and prevent the submission of sensitive information to untrustworthy sites.

In 2006 Robert Ma [3] exploits search results made with unique keywords extracted from a suspicious website on a reputable search engine, Google. Then, these search results are analyzed to determine whether a website is genuine or counterfeit.

In 2007 , Ian Fette Norman Sadeh Anthony Tomasic [4] they present application of machine learning on a feature set designed to highlight user-targeted deception in electronic communication. This method is applicable, with slight modification , for detection of phishing websites, or the emails used to direct victims to these sites. They evaluated this method on a set of approximately 860 such phishing emails, and 6950 non-phishing emails, and correctly identify over 96% of the phishing emails while only misclassifying on the order of 0.1% of the legitimate emails.

C. Yue and H. Wang, 2008 [5] proposed a new approach to protect against phishing attacks with "bogus bites", a unique client-side anti phishing tool, which transparently feeds a relatively large number of bogus credentials into a suspected phishing site. BogusBiter conceals a victim's real credential among bogus credentials, and moreover, it enables a legitimate website to identify stolen credentials in a timely manner. Leveraging the power of client-side automatic phishing detection techniques, BogusBiter is complementary to existing preventive anti-phishing approaches.

In 2008 Andr´e Bergholz, Gerhard Paaß,Frank Reichartz, Siehyun Strobel [6] proposed advanced email features generated by adaptively trained Dynamic Markov Chains and by novel latent Class-Topic Models. On a publicly available test corpus classifiers, using these features is able to reduce the number of misclassified emails.

In 2008 Guang Xiang, Bryan A. Pendleton, and Jason Hong [7] presented the design and evaluation of two blacklist-enhanced content-based algorithms. The key insight behind their algorithms is to leverage existing human-verified whitelists and blacklists, and relax them via probabilistic methods to attain high true positive rates while maintaining extremely low false positive rates.

In 2008, Anirudh Ramachandran, Nick Feamster, Balachander Krishnamurthy,Oliver Spatscheck, Jacobus Van der Merwe [8]

performed a preliminary study to determine the feasibility of detecting phishing attacks in real-time, from the network traffic stream itself. They developed a model to identify the stages wherein network phishing detection is feasible and the data sources can be analyzed to provide relevant information at each stage. Based on this model, they developed and evaluated a detection method based on features that exist in the network traffic itself and are correlated with confirmed phishing attacks.

In 2008, Ram Dantu Srikanth Palla and Joao Cangussu[9] describe techniques for detecting phishers based on their traffic paths, traffic patterns, and on the receivers' social network. Considering such issues, their solution based on the trustworthiness of the relays participating in routing the emails.

In 2009 Michael Blasi [10] improves upon techniques used by popular anti-phishing software and introduces a new faster recognition method of detecting zero day phishing sites which are new phishing sites that have not yet been discovered using cascading style sheets (CSS). This phishing detection techniques are evaluated against hundreds of known phishing sites.

In 2009, Paul Knickerbocker, Dongting Yu, and Jun Li [11] designed Humboldt which is a distributed system that submits poisonous fake data to phishing websites that is indistinguishable from the input of actual phishing victims. The poisonous data collected by a phisher produces detectable behaviors when the phisher attempts to use it and provides a mechanism for tracking activities associated with identity theft.

In 2009, Pawan Prakash, Manish Kumar, Ramana Rao Kompella, Minaxi Gupta [12] proposed five heuristics to enumerate simple combinations of known phishing sites to discover new phishing URLs. The second component consists of an approximate matching algorithm that dissects a URL into multiple components that are matched individually against entries in the blacklist. In evaluation with real-time blacklist feeds, we discovered around 18,000 new phishing URLs from a set of 6,000 new blacklist entries.

3.Methods for Detection Phishing Website

Several methods are being used to create phishing website, in this section list most widespread techniques .

1) E-mail Based Approach

Typically, in an e-mail based phishing attack, phishers dispatch large volume of spoofed e-mails containing embedded URLs to redirect potential victims into fake Websites to trick them into disclosing their confidential information[13]. Some of the approaches attempt to eliminate the phishing problem at the e-mail level by trying to prevent phishing e-mails from reaching the potential victims. E-mail-based approaches typically use filters and content analysis and are, hence, closely related to anti-spam research. Anti-spam solutions are not perfect and some phishing e-mails may still reach potential victims. Microsoft and Yahoo have also defined e-mail authentication protocols (i.e., Sender ID and Domain Keys) that can be used to verify if a received e-mail is authentic [14].

2) Blacklist Based Approach

A blacklist in the context of phishing is a list of untrusted URLs or more simply a list of banned websites that are known to have malicious intentions[10]. Every time a browser opens a new page, the program checks to see if the page is on the list of known phishing sites. If the site is on the list, the user is alerted and prevented from submitting data to the website, thus reducing the risk of identity theft [15].The most popular and widely-deployed antiphishing techniques are based on the use of blacklists of phishing domains. For example, Microsoft has recently integrated a blacklist-based anti-phishing solution into its Internet Explorer (IE) 7 browser[14]. While this solution is effective for users fortunate enough to go to phishing sites on the black list, those who visit phishing sites too new for the blacklist may still fall victim to identity theft. The black-list approach is limited due to the short lived nature of phishing sites which are usually only available for a few hours or days[15].

3) Information Flow Based Approach

Anti-Phish takes a different approach and keeps track of where sensitive information is being submitted [14]. is concerned, every page that contains a form is a potential phishing page. HTML form elements that can be used by the attacker to phish information from the user are text field elements of type text and

password and the HTML text area element. Anti-Phish checks the list of previously captured values (i.e., the "watch list"). For each value in this list that is identical to the one just entered by the user, the corresponding domain is determined. If the current site is not among these domains, a phishing attempt is assumed. The reason is that sensitive information is about to be transmitted to a site that is not explicitly listed as trusted[16].The main disadvantage of Anti-Phish is that it requires user interaction to specify which sensitive information should be captured and monitored. [14].

4) IP-Based URLs

phishing attacks are becoming more sophisticated, IP-based links are becoming less prevalent, with attackers purchasing domain names to point to the attack website instead. However, there are still a significant number of IP-based attacks, and therefore this is still a useful feature [17] A URL containing the '@' symbol will ignore everything to the left of the '@'symbol , Additional checks include excessive dots in the URL, redirection servers, misspelling, non-standard port numbers, and keywords. Current web browsers such as Microsoft Internet Explorer 7 detect some of these suspicious URL techniques and automatically prevent the user from visiting the suspicious site [10].

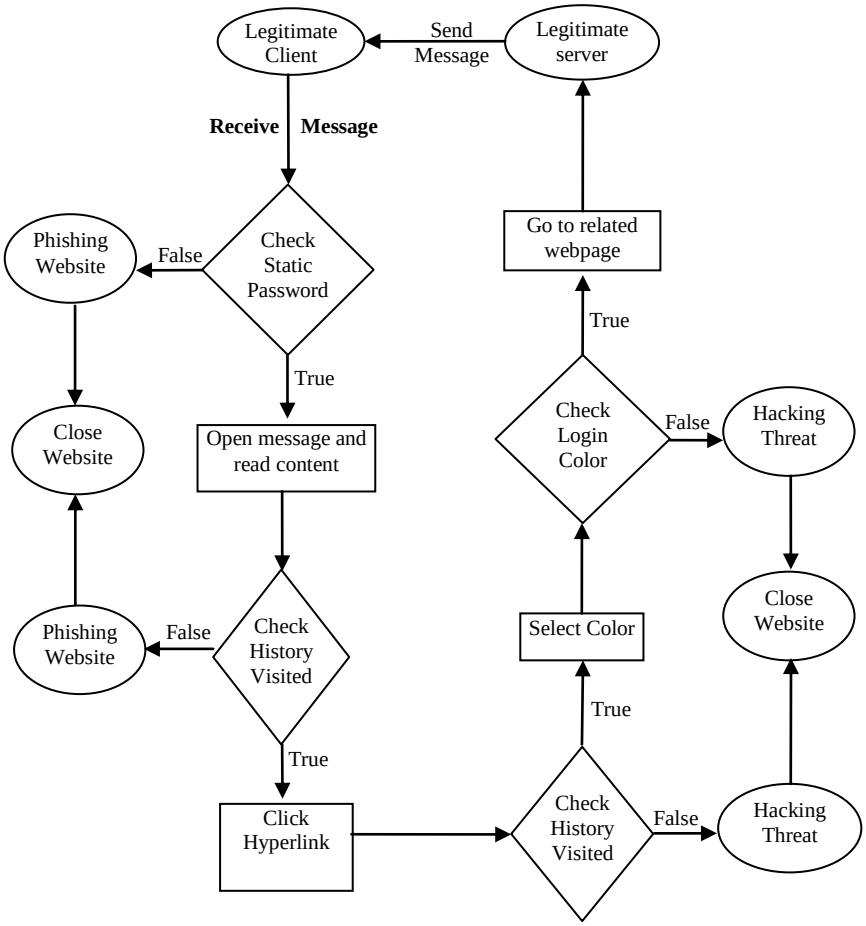
4.1 Proposed System

This paper aims at prevent users for falling in phishing website, through level of authentication process, is well known that users have difficulty in remembering secure passwords. So in the proposed system, users setting only one static password that are meaningful and memorable for legitimate website authentication and favorite color legitimate user authentication.

When a user receives a message he does not know whether it is from a legitimate website or phishing website so to indicate that, first level user read the static password that embed with the title of message if there is matching with his assigned previously password so that message considers its from a legitimate website.

Second level is to emphasize authentication process by user must foind the true heading history of last visited website such as (account, time enter website, time exit website, number of books bought, type of books),third level based on the rule that legitimate website never requests from user any crediational information

except the login color that is assigned previously to prove the legitimate client. Authentication process between server and client can be shown in the figure(1) .



Figure(1) Flowchart of proposed Antiphish system

The algorithm of the proposed Phishing Website is described in the algorithm (1).

Algorithm (1) of Proposed System
Step 1:Client set static password and login color into legitimate website
Step 2:Server store static password and generated dynamic password
Step3:Server sent message to client with two passwords
Step4:Client receives message
Step5:Client checks validity of static password
If false go to(step11)
Else go to next step
Step6:Client checks history of last visited in the content of message
If false go to(step11)
Else go to next step
Step7:Client clicks the hyperlink of message
Step8:Server requests only login color
Step9:Server compare between static and dynamic password and color
If false go to(step11)
Else go to next step
Step10:Server redirects client to related webpage
Step11:Close web browser

4.2 System work

The proposed system is classified into two authentication proces to understand how its system works.

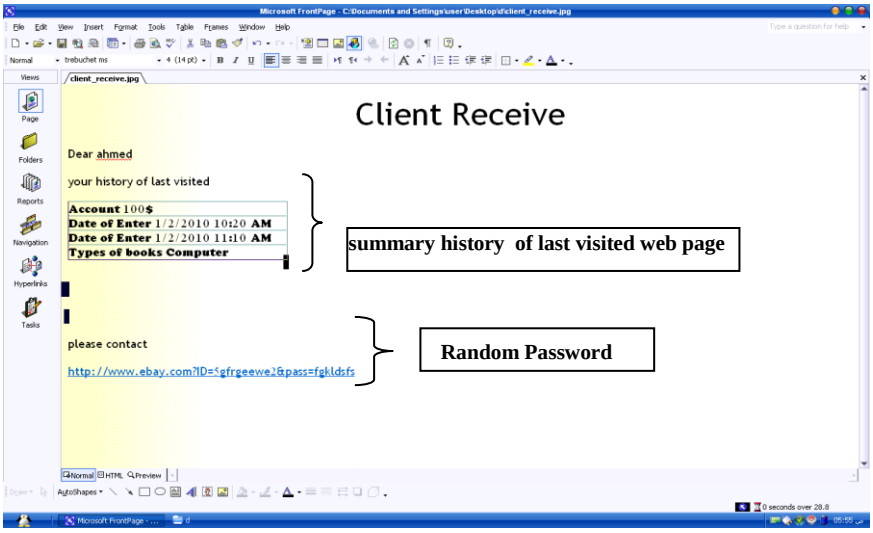
4.2.1 Client-Server Authentication

To authenticate content from a server, the user only needs to read static password that embed with title of message and shortcut history of last visited website its correct .

Table(1) show examples of client data authentication.

Table (1) Client Authentication Details				
User_ID	**Username**	**User_Password**	**Bill**	**account**
1	**Ahmed**	**Hello**	**12$**	**100$**
2	**Ali**	**Computer**	**20$**	**100$**
3	**Mohammed**	**Football**	**30$**	**150$**

If server wants to interact with client urgently must contain random password in the hyperlink of message and summary history of last visited website in the content of message as is shown in figure (2).



Figure(2) Browsing content of message

4.2.2 Server_Client Authenticate

Now, whenever client opens a message and reads the summary history and detects the message it comes from legitimate website, so click link, Server to authenticate webpage requested from legitimate client extract dynamic password from incoming URL hyperlink if there exists a matching between server database and information of querystring server performs process authentication by asking the user to select a true color between many colors

Then allow the client to go to the related web page and replay the request of message, otherwise, the server will be directed to stop connection because a message has been hacked .

Table (2) shows a sample of data that is used by the server.

Table (2) Server Authentication Details			
User_ID	User_Name	Server_Password	Color
1	Ahmed	4356%#@#	Green
2	Ali	Fd#$%@SF	Blue
3	Mohmmeh	F#$F%653	Gray
7	Jassim	Ht5367#$2	Yello
5	Basil	S34%^#@3	Red

At the end of an interaction, the user authenticates the server by static password and shortcut history and the server authenticates the user by selected color and dynamic password.

4.3 Evaluation

In the proposed system, the user correctly verifies that the site is legitimate significantly not to affect the availability of viewing a webpage. Websites load quickly and with minimal delay. The user does not notice a difference between a website loaded with the legitimate website compared to the website loaded with the phishing detection.

To authenticate content from the server, the user can visually verify that the static password matches and last visited history. In contrast to other proposals, proposed scheme places a very low burden on the user in terms of effort, memory and time, two metrics are taken.

a) Zero False Positive Rate

This means that legitimate sites should not be detected as phishing sites. This is probably the most volatile variable because a false positive highly depends on the type of algorithm used to identify potentially phishing sites. The Proposed system ideally the software has a zero false positive rate because a legitimate website never sends error user password and shortcut history.

b) Zero False Negative Rate

This means that phishing sites should be detected as fraudulent sites. This is probably the most volatile variable because a false negative depends a lot on the type of algorithm used to identify potentially fraudulent sites. ANTI-PHISH solution ideally the software has a zero false negative rate because a phishing website sends a true static user password and shortcut history .

Evaluate its efficacy through real experiments over both phishing and legitimate web sites. Experimental results indicate that "anti phish" is a promising anti-phishing approach.

5.Conclution

The most common type of phishing attack attempts is to steal account numbers and passwords used for online banking therefore protecting a user's username/password credential is the primary focus of proposed anti-phishing research work. In particular, anti-phish leverages the power of phishing detection mechanisms.

To conclude with,it is easy for a user to detect the legitimacy of website and at the same time it's hard for an attacker to spoof the indicators, static user password and personal shortcut history and any website asks the user to enter password and other personal information is considered a phishing website.

6. Future Work

When a user receives a message he does not know whether it is from a legitimate website or phishing website so to detect that, first level user reads the static password that embeds with the title of message if there is matching with his previously assigned password so that the message is consider from a legitimate website ,in the future work this operation can be embedded in an email server to detect the phishing website through comparison between two passwords previously set; one is for the email server itself and another for all web server that exists relation with the user.

References

1. **Min Wu. "Fighting Phishing at the User Interface ". 2010 Available at:- http://groups.csail.mit.edu/uid/projects/phishing/proposal.pdf**
2. **Sven Strickroth."Phishing Detection". 2008 Available at:- http://gymnasium-osterode.de/~sstrickroth/phishing-detection.pdf**
3. **Arel Cordero." Detecting Phishing Attacks From RenderedWebsite Images", 2006 Available at:- http://cs.berkeley.edu/~asimma/294-fall06/projects/cordero.pdf**
4. **Ian Fette. 'Learning to Detect Phishing Emails",Carnegie Mellon University 2007Available at:- http://www2007.org/papers/paper550.pdf**
5. **Chuan Y,Haining W "Anti-Phishing in Offense and Defense" The College of William and Mary , 2008 Available at:- http://www.cs.wm.edu/~hnw/courses/cs780S/papers/yuec-antiphishing.pdf**
6. **Andr´e Bergholz, Gerhard Paaß,Frank Reichartz, Siehyun Strobel "improved Phishing Detection using Model-Based Features" Konan Technology. 2008 Available at:- http://www.ceas.cc/2008/papers/ceas2008-paper-44.pdf**
7. **Guang Xiang, Bryan A. "Modeling Content from Human-Verified Blacklists for Accurate Zero-Hour Phish Detection",Carnegie Mellon University, 2008 Available at:- http://www.lti.cs.cmu.edu/Research/TechReports/CMU-LTI-09-ModelingContentfromHumanVerifiedBlacklistsforAccurateZero-HourPhishDetection.pdf**

8. Anirudh Ramachandran, Nick Feamster "Fishing for Phishing from the Network Stream",Georgia Tech, , 2008 Available at:- http://www.cc.gatech.edu/research/reports/GT-CS-08-08.pdf
9. Ram Dantu,Srikanth Palla,Joao Cangussu "Journal of Homeland Security and Emergency Management" ,The Berkeley Electronic Press,2008 Available at:- http://www.usc.edu/dept/ise/assets/007/64791.pdf
10. Michael Blasi,"Techniques for Detecting Zero Day Phishing Websites", Iowa State University, 2009 Available at:- http://archives.ece.iastate.edu/archive/00000498/01/Thesis_Michael_Blasi.pdf
11. Paul Knickerbocker, Dongting Yu, and Jun Li, "A Distributed Phishing Disruption System", University of Oregon, , 2009 Available at:- http://ix.cs.uoregon.edu/~dongting/humboldt.pdf
12. Pawan Prakash, Manish Kumar, Ramana Rao Kompella, Minaxi Gupta, "Predictive Blacklisting to Detect Phishing Attacks",Purdue University, , 2009 Available at:- http://www.cs.indiana.edu/~minaxi/pubs/infocom10-purdue.pdf
13. Shambhu Upadhyaya and Madhu Chandrasekaran,"Challenges in Mitigating Phishing and Spam e-mails", University at Buffalo, 2009 Available at:- http://www.cse.buffalo.edu/~shambhu/documents/pdf/resum.pdf
14. Angelo P. E. "A Layout-Similarity-Based Approach for Detecting Phishing Pages" Technical University Vienna. 2007 Available at:- http://www.seclab.tuwien.ac.at/papers/antiphishdom.pdf
15. James Presley Henshaw,"ENHANCING CONTENT-Triggered Trust Negotiation To Prevent Phishing Attacks ",Brigham Young University, 2009 Available at:- http://isrl.cs.byu.edu/pubs/PhishingWarden.pdf
16. Engin Kirda, Christopher Kruegel,"Protecting Users Against Phishing Attacks",Technical University of Vienna, 2005 Available at:- www.cs.ucsb.edu/~chris/research/doc/cj06_phish.pdf
17. Ian Fette Anthony Tomasic, "Learning to Detect Phishing Emails",Carnegie Mellon University, , 2006 Available at:- http://www2007.org/papers/paper550.pdf