Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 2, Issue. 4, April 2013, pg.523 – 527

RESEARCH ARTICLE

Security Based Phishing Website Detection

Yalavarthi Ravi Theja[1], R. Krishnaveni[2]
[1]Department of Computer Science, Hindustan University, Chennai, India
[2]Department of Computer Science, Hindustan University, Chennai, India

[1] *ravi.theja535@gmail.com;* [2] *rskichu10@gmail.com*

Abstract— Now-a-days most of the people were familiar with the internet and its applications. Along with the internet usage, the attacks also increased. Phishing is the one the most possible attacks in internet and through this the Phisher will get the confidential information like passwords. In this model, MD5 algorithm is used to hash the password. A session key will be send to the authorized person's mobile to perform further transactions. These techniques increase the security levels. The proposed system provides the protection against normal phishing and the In-session phishing using URL checking and session key respectively. According to the report on phishing attacks-2012, India stud in 3rd position all over the world. US was in 1st position and UK was in 2 nd position. In India most of the phishing attacks concentrate only on banking sectors. The proposed system was mainly concentrate on banking sector only. The proposed system gives the best results over the existing system.

Key Terms: - Hashing Algorithm; Phishing Detection; Java Short Messages; GSM; Web Security

I. INTRODUCTION

A. Short Message Service

Short Message Service is a text messaging service component of phone, web, or mobile communication systems, using standardized communications protocols that allow the exchange of short text messages between fixed line or mobile phone devices [4]. SMS text messaging is the most widely used data application in the world, with 3.6 billion active users, or 78% of all mobile phone subscribers. The term SMS is used as a synonym for all types of short text messaging as well as the user activity itself in many parts of the world. SMS is also being used as a form of direct marketing known as SMS marketing. SMS has used on modern handsets originated from radio telegraphy in radio memo pagers using standardized phone protocols and later defined as part of the Global System for Mobile Communications series of standards in 1985 as a means of sending messages of up to 160 characters, to and from GSM mobile handsets [4]. Since then, support for the service has expanded to include other mobile technologies such as ANSI, CDMA networks and Digital AMPS, as well as satellite and Landline networks. Most SMS messages are mobile-to-mobile text messages though the standard supports other types of broadcast messaging as well.

B. Java Short Messages

JSMS is a Java API for sending and receiving Short Messages (SMS) and Multimedia Messages (MMS). The API supports a wide range of communication protocols. Some short message forms are UCS2, Binary, and EMS. Some multimedia messages are MM1, and MM7. Messages may be sent and received by using any GSM device capable of sending SMS messages and also by using the most common SMSC communication protocols like SMPP, TAP/IXO, UCP and CMID2.

The JSMS API [6] has been designed with a modular architecture in focus. This allows an easy integration of other SMS transport facilities. Besides sending and receiving Short Messages, the API also contains a small footprint SMTP (Simple Mail Transfer Protocol) Client which enables your applications to send internet emails according to RFC822. Windowing for applications enables JSMS to initiate more than one operation before receiving responses from the SMSC (Short Message Service Centre).

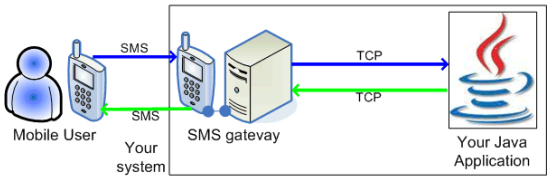

Fig1. Working with jsms

Figure1 explains the sending a SMS to the users mobile. The system contains the java application and the SMS gateway, which was connected with the java application using data cable. Through this a SMS will be send to and received from the users mobile. But the proposed system uses the jsms only to send SMS to users mobile.

C. Global System for Mobile Communications

The Short Message Service – Point to Point was originally defined in GSM recommendation 03.40, which is now maintained in 3GPP as TS 23.040. GSM 03.41 defines the short message service cell broadcast, which allows messages to be broadcast to all mobile users in a specified geographical area [3]. Messages are sent to a short message service centre which provides a "store and forward" mechanism. It attempts to send messages to the SMSC's recipients. If a recipient is not reachable, the SMSC queues the message for later retry. Some SMSCs also provide a "forward and forget" option where transmission is tried only once. Both mobile terminated operations are supported. Message delivery is "best effort", so there are no guarantees that a message will actually be delivered to its recipient, but delay or complete loss of a message is uncommon, typically affecting less than 5% of messages. Some providers allow users to request delivery reports, either via the SMS settings of most modern phones, or by prefixing each message with *0# or *N#. However, the exact meaning of confirmations varies from reaching the network, to being queued for sending, to being sent, to receiving a confirmation of receipt from the target device, and users are often not informed of the specific type of success being reported. Transmission of short messages between the SMSC and the handset is done whenever using the mobile application part of the SS7 protocol. Messages are sent with the MAP MO- and MT-Forward SM operations, whose payload length is limited by the constraints of the signalling protocol to precisely 140 octets. Short messages can be encoded using a variety of alphabets, the default GSM 7-bit alphabet, the 8-bit data alphabet, and the 16-bit UCS-2 alphabet. Depending on which alphabet the subscriber has configured in the handset, this leads to the maximum individual short message sizes of 160 7-bit characters, 140 8-bit characters, or 70 16-bit characters. . Routing data and other metadata is additional to the payload size. Larger content (concatenated SMS, multipart or segmented SMS, or "long SMS") can be sent using multiple messages, in which case each message will start with a user data header (UDH) containing segmentation information. Since UDH is part of the payload, the number of available characters per segment is lower: 153 for 7-bit encoding, 134 for 8-bit encoding and 67 for 16-bit encoding. The receiving handset is then responsible for reassembling the message and presenting it to the user as one long message. While the standard theoretically permits up to 255 segments, 6 to 8 segment messages are the practical maximum, and long messages are often billed as equivalent to multiple SMS messages. Some providers have offered length-oriented pricing schemes for messages; however, the phenomenon is disappearing.

II. EXISTING SYSTEM

The existing system uses the Classifiers, Fusion Algorithm, and Bayesian Model to detect the phishing sites. The classifiers can classify the text content and image content. Text classifier is to classify the text content and Image classifier is to classify the image content. Bayesian model estimates the threshold value. Fusion Algorithm combines the both classifier results and decides whether the site is phishing or not. The performance of different classifiers based on correct classification ratio, F-score, Matthews's correlation coefficient, False negative ratio, and False alarm ratio. The threshold value will be decided by the developer only. This leads to the problems like false positive and false negative. False positive means, the probability of being a phishing webpage is greater than the threshold value but that webpage is not a phishing webpage. False negative means, the probability of being a phishing webpage is less than the threshold value but that webpage is a phishing webpage. This results the reduction in security levels. The existing system handles the only one kind of phishing attacks. If that was a phishing site then the existing system only warns the user. The active and passive warnings

[5] alone were not enough to control the phishing sites. The active warning gives the user options to close the window or displaying the website. The passive warning displays the popup dialog box.

III. MOTIVATION OF THE PROPOSED SYSTEM

Online transactions are nowadays become very common and there are various attacks present behind this. In these types of various attacks, phishing is identified as a major security threat and new innovative ideas are arising with this in each second so preventive mechanism should also be so effective. Thus the security in these cases be very high and should not be easily tractable with implementation easiness. Today, most applications are only as secure as their underlying system. Since the design and technology of middleware has improved steadily, their detection is a difficult problem. As a result, it is nearly impossible to be sure whether a computer that is connected to the internet can be considered trustworthy and secure or not. Phishing scams are also becoming a problem for online banking and e-commerce users. These all things motivate me to do this work.

IV. SYSTEM DESIGN

The proposed system provides the secured online banking transactions. It can handle the two kinds of phishing like Normal and In-session phishing.

A. In-session Phishing

This is one kind of phishing attacks. The user will be diverted by getting alert message like, "your session timeout and please login again". Then that user redirected to phishing site and phisher will get users account number and password. Using them, the phisher can transfer the funds from authorized users account that to without that user's knowledge.

B. Providing Secure Online Banking Transactions

According to IDC report, India has a little than a million active online banking users that might be just 0.096 percentage of the total India population. India has 15 percentage of Internet users in its total population. That was second highest in the world and next to China. But India is lagging behind, when compared to USA, China, UK, and Japan.

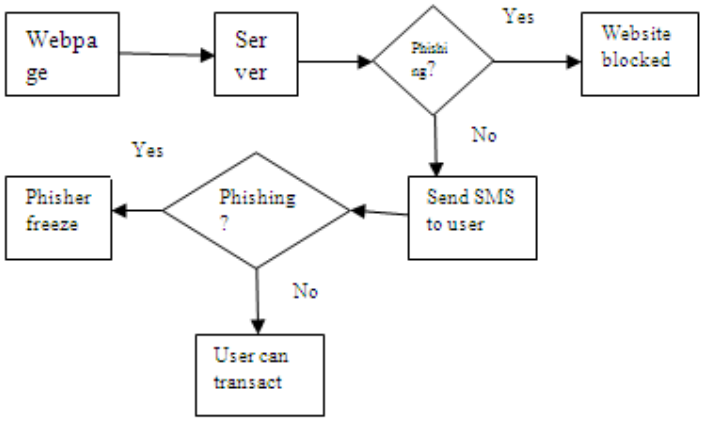

Fig2. Architecture of proposed system

In the case of the online banking transactions, security plays a crucial role. According Figure2, the proposed system provides the secure online banking transactions effectively. It can deal the In-session phishing efficiently. By using username and password, the user can login to his account. Then a session key [6] will be send to the authorized users mobile. Using that session key only the user can perform the further transactions. It was possible for the phisher to get the username and password of a particular user, but without that session key the phisher can't perform any unauthorized transactions.

V. RELATED WORK

Chris Karlof et. al [1] proposed two Locked Same-origin Policies. This system describes a new attack against web authentication, which was called as dynamic pharming. Dynamic pharming works by hijacking DNS and sending the victim's browser malicious JavaScript, which then exploits DNS rebinding vulnerabilities and the name-based same origin policy to hijack a legitimate session after authentication has taken place. As a result, the attack works regardless of the authentication scheme used. Dynamic pharming enables the adversary to eavesdrop on sensitive content, forge transactions, sniff secondary passwords, etc. To counter dynamic pharming attacks, This system proposes two locked same-origin policies for web browsers. In contrast to the legacy same-origin policy, which regulates cross-object access control in browsers using domain names, the locked same-origin policies enforce access using servers' X.509 certificates and public keys. The results suggest one of those policies can be deployed today and interoperate seamlessly with the vast majority of legacy web servers. For other policy, this system presents a simple incrementally deployable opt-in mechanism for legacy servers using policy files, and shows how web sites can use policy files to support self-signed and entrusted certificates, shared sub domain objects, and key updates. Drawback is it's hard to find and get protection against dynamic attacks.

Serge Egelman et. al [5] proposed Active and Passive Phishing warnings. Many popular web browsers now include active phishing warnings since research has shown that passive warnings are often ignored. In this laboratory study we examine the effectiveness of these warnings and examine if, how, and why they fail users. According to a survey, 97% of sixty participants fell for at least one of the phishing messages and 79% of participants ignore the active warnings, which were not the case for the passive warnings where only one participant heeded the warnings. Using a model from the warning sciences we analyzed how users perceive warning messages and offer suggestions for creating more effective phishing warnings. Drawback is Users don't believe those warnings and lack of knowledge about warnings.

Eric Medvet et. al [2] proposed Visual-Similarity based system. Phishing is a form of online fraud that aims to steal a user's sensitive information, such as online banking passwords or credit card numbers. The victim is tricked into entering such information on a web page that is crafted by the attacker so that it mimics a legitimate page. Recent statistics about the increasing number of phishing attacks suggest that this security problem still deserves significant attention. This system presents a novel technique to visually compare a suspected phishing page with the legitimate one. The goal is to determine whether the two pages are suspiciously similar. We identify and consider three page features that play a key role in making a phishing page look similar to a legitimate one. These features are text pieces and their style, images embedded in the page, and the overall visual appearance of the page as rendered by the browser. To verify the feasibility of this approach, need to perform an experimental evaluation using a dataset composed of 41 real-world phishing pages, along with their corresponding legitimate targets. The experimental results are satisfactory in terms of false positives and false negatives. Drawback is it's hardly possible to find phishing sites only by checking home page.

VI. IMPLEMENTATION AND RESEARCH

In this paper password hashing with MD5 algorithm has proposed. If the password is hashed with addition of salt by applying a cryptographic hash function, then Phishing attack can be removed. The salt value will prevent attackers from building a list of hash values [6] for common passwords. It is also shown that the attack on hashed passwords is unsuccessful as getting original password from hashed form is not an easy task due to addition of salt value.

MD5 stands for Message-Digest 5 Algorithm. This was also called as Fingerprint Algorithm. Message-Digest algorithm is a special function which transforms input of arbitrary length into output of constant length. These transformation functions must fulfil these requirements:

- No one should be able to produce two different inputs for which the transformation function returns the same output.
- No one should be able to produce input for given pre specified output.

Message-Digest algorithms serve in digital signature applications for guaranteeing consistency of data. Commonly used model as follows:

- Sender creates input message (M) and computes its message digest (sMD). Then he uses his private key and encrypts message digest (esMD).
- Encrypted message digest (esMD) is attached to the input message (M) and whole message (M-esMD) is sent to receiver.

- Receiver gets the message (M-esMD) and extracts the encrypted message digest (esMD). Then he computes his own message digest (rMD) of the received message (M). He also decodes the received message digest (esMD) with sender's public key and gets decoded message digest (desMD). Then he compares the both message digests (rMD? = desMD). When both message digests are equal, the message was not modified during the data transmission.

In cryptography, MD5 is a widely used cryptographic hash function with a 128-bit hash value. As an Internet standard (RFC 1321), MD5 has been employed in a wide variety of security applications, and is also commonly used to check the integrity of files.

The MD5 algorithm will consist of 5 steps.
Step1: Append Padding Bits
Step2: Append Length
Step3: Initialize MD5 Buffer
Step4: Process Message in 16-Word Blocks
Step5: Output

VII. CONCLUSION

Currently phishing attacks are so common because it can attack globally and capture and store the users' confidential information. This information is used by the attackers which are indirectly involved in the phishing process. The main objective of this paper is to provide the protection from two kinds of phishing attacks. Hence by using the Password hashing technique by MD5 algorithm, the proposed system fulfilled its needs.

The proposed system only confined to the single bank that is Indian bank. The phisher can't do anything without the session key, but it is not so hard to get the session key from the authorized users mobile with the help of hacker. As a future work, the proposed system will be expanding to other banks also and the system has to provide the security not only against the phishing but also the hacking.

REFERENCES

[1] Chris Karlof, J.D. Tygar, David Wagner , Dynamic Pharming Attacks and Locked Same-origin Policies for Web Browsers in IEEE conference, 2011, pp. 14-19.
[2] Eric Medvet, Engin Kirda, Christopher Kruegel, Visual-Similarity-Based Phishing Detection, in conference on secure communications, 2008, pp. 67-78.
[3] M. Markou and C. Panayiotou, Dynamic control and optimization of buffer size for short message transfer in GPRS/UMTS networks, in Proc. of IEEE International Conference on Information & Communication Technologies: From Theory to Applications, 2004, pp. 34-38.
[4] Petros Zerfos , Xiaoqiao Meng, Starsky H.Y Wong A Study of the short message service of the nationwide cellular network , Acm Conference , Oct 2006, pp. 58-65.
[5] Serge Egelman , Lorrie Faith Cranor , Jason Hong , You've been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings in International conference on IT, 2008, pp. 89-96.
[6] Shamir, How to Share a Secret, Communication ACM, 1979, pp. 612-613.
[7] Sid Stamm, Zulfikar Ramzan, Markus Jakobsson, Drive-By Pharming, in IEEE conference, 2006, pp. 24-67.
[8] W. Liu, G. Huang, X. Liu, M. Zhang, and X. Deng, Phishing web page detection, in Proc. 8th Int. Conf. Documents Anal. Recognit., Seoul, Korea, Aug. 2010, pp. 560–564.
[9] Y. Zhang, S. Egelman, L. Cranor, and J. Hong, Phinding phish: Evaluating anti-phishing tools, in Proc. 14th Annu. Netw. Distribut. Syst. Secur. Symp., San Diego, CA, Feb. 2007, pp. 1–16.

