A Survey of Phishing Attack Techniques

Article *in* International Journal of Computer Applications · May 2014

DOI: 10.5120/16197-5460

Siddharth Singh Chouhan
Shri Mata Vaishno Devi University

A Survey of Phishing Attack Techniques

Minal Chawla
Department of Information Technology
Laxmi Narayan Collage of Technology, Bhopal

Siddarth Singh Chouhan
Department of Information Technology
Laxmi Narayan Collage of Technology, Bhopal

ABSTRACT

Now in a day’s phishing is a special type of network attack where the attacker creates a replica of an existing web page to fool users in to submitting personal, financial, transaction or password data to what they think is their service provider’s website. Phishing has two techniques, deceptive phishing and malware – based phishing. Here we focus on deceptive phishing using social engineering schemes. To protect users against phishing, various anti-phishing techniques have been proposed. In this paper we have reviewed various phishing and anti-phishing methods for detecting and preventing phishing attack.

1. INTRODUCTION

The word 'Phishing' initially emerged in 1990s. The early hackers often use 'ph' to replace 'f' to generate new words in the hacker's community, since they usually hack by phones. Phishing is a new word produced from 'fishing', it refers to the act that the attacker appeal users to visit a faked Web site by sending them faked e-mails (or instant messages), and silently get victim's personal information such as user name, password, national security ID, etc. This information then can be used for future target advertisements or even identity theft attacks (e.g., transfer money from victims' bank account).

One of the primary aims of phishing is to dishonestly carry out fraudulent financial transactions on behalf of users using a forged email that contains a URL pointing to a fake web site masquerading as an online bank or a government entity. A phisher may

Tempt a victim into giving his/her Social Security Number, full name, & address, which can then be used to apply for a credit card on the victim’s behalf.

The act of sending an e-mail to a user falsely claiming to be an established legitimate enterprise in an attempt to scam the user into conceding personal information that will be used for identity theft. The e-mail directs the user to visit a Web site where they are asked to renew personal information, such as passwords and credit card, social security, and bank account numbers, that the legitimate organization already has. The Web site, on the other hand, is bogus and set up only to steal the user's information.

Types of phishing attack- Phishing is a particular type of spam that employs two techniques-

1. Deceptive phishing
2. Malware-based phishing

The first technique is associated to social engineering schemes, which depend on forged email claims that emerge to originate from a legitimate company or bank. Subsequently, through an embedded link within the email, the phisher attempts to redirect users to fake Websites. These fake Web sites are designed to fraudulently achieve financial data (usernames, passwords, credit) from victims.

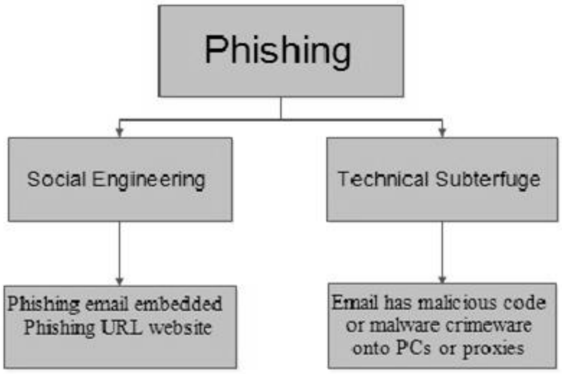
Phishing
Social Engineering
Technical Subterfuge
Phishing email embedded Phishing URL website
Email has malicious code or malware crimeware onto PCs or proxies

Fig 1- Types of phishing

The second technique involves technical subterfuge schemes that rely on malicious code or malware after users click on a link rooted in the email, or by detecting and using security holes in the user’s computer to achieve the victim’s online account information directly. Sometimes, phisher attempts to misdirect the user to a fake Web site or to a legitimate one monitored by proxies.

1.1 Procedure

In general, phishing attacks are performed with the following four steps:

1) Phishers set up a counterfeited Web site which looks precisely like the legitimate Web site, including setting up the web server, applying the DNS server name, and creating the web pages similar to the destination Web site, etc.

2) Send large amount of spoofed e-mails to target users in the name of those legitimate companies and organizations, trying to convince the budding victims to visit their Web sites.

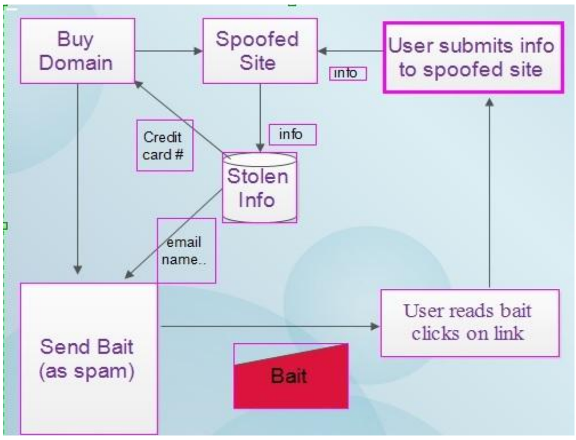

Fig 2 – Procedure of phishing attack

3) Receivers receive the e-mail, open it, click the spoofed hyperlink in the e-mail, and enter the required information.

4) Phishers steal the personal information and act upon their scam such as transferring money from the victims 'account.

1.2 Life cycle of phishing email

The procedure of phishing email transfer in a computer network. It contains three components: a Message Transfer Agent (MTA), Message Delivery Agents (MDA), and a Mail User Agent (MUA).

Message Transfer Agent (MTA) is responsible for sending and receiving mail between systems using SMTP.

Message Delivery Agents (MDA) are responsible for receiving a message from an MTA and arranging for it to be received by the local system (e.g. delivered to a mailbox).

Mail User Agent (MUA) is the program that an end user uses to read and process mail. Typical examples include Microsoft outlook, Pegasus, exmh, mutt, Eudora, etc.

2. LITERATURE SURVEY

In [1] they present an allegation of the various techniques currently used to detect phishing email, at the different stages of attack, mostly focusing on machine-learning techniques. A comparative study and evaluation of filtering methods is carried out. This provides an understanding of the problem, its recent solution space, and the future research directions anticipated. Classifiers used to identify phishing email are based on: supervised learning, i.e. they must learn before they can be used to detect a new attack; unsupervised learning, which is faster, but has a low level of accuracy; or a hybrid (supervised and unsupervised) learning, which is time consuming and costly.

In [2] they explain various approaches to detection for replication of web site layout and structure through source code (and optionally image) fingerprinting. This Anti phishing technique based on URL and Domain Identity, and Image Based Webpage Matching .It first identifies the related authorized URL in which approximate string matching algorithm is used. The image based matching mechanism uses key point's detection and feature extraction methods.

In [3] they present a novel technique to visually compare an alleged phishing page with the legitimate one. The target is to determine whether the two pages are warily similar. Signature based algorithm is used, the proposed approach is inspired by two previous open source anti-phishing solutions: the Anti Phish browser plug in and its DOM Anti Phish extension. This results in impressive speed-ups. A negative comparison between two pages is produced in a few milliseconds.

In [4] they present that participant using the high exactness anti-phishing tool considerably outperformed those using the less accurate tool in their ability to: (1) differentiate legitimate websites from phish; (2) avoid visiting phishing websites; and (3) avoid transacting with phishing websites. URL and DNS Based spoofing technique are used. The results indicate that using more accurate anti-phishing tools can significantly improve users' ability to identify phishing websites and to better avoid visiting and transacting with phish. It is also imperative to improve methods for conveying tool warnings.

In [5] they implement a model i.e. – IPDCM, can handle about 50 pages per second, which make it feasible for real internet security production applications. This intelligent model for detecting phishing websites, extract 10 different types of features such as title, keyword and link text information to represent the website. Heterogeneous classifiers are then built based on these different features Hierarchical clustering technique has been employed for automatic phishing categorization. And also SVM is used in this technique.

In [6] this technique is purely based on image comparison using discriminative key point features in WebPages. They used an invariant content descriptor, the Contrast Context Histogram (CCH), to compute the similarity degree between suspicious pages and authentic pages. This anti phishing tool is highly efficient and error free. It can be used in online banking, online shopping and to maintain the mail accounts.

In [7] they proposed a preventive and upbeat technique, which detects phishing activity even without opening a phishing web page .Using an intelligent hybrid technique, Adaptive Neuro - Fuzzy Inference System (ANFIS). Neural networks and Fuzzy logic have been used to successfully counter the phishing attacks. The detection rate is 98% .There is no false positives present, which may lift up a false alarm and classify a genuine e-mail as a phishing e-mail.

In [8] Code Generation Technique is used. This technique ensures users about the legitimacy of the webpage, he visits where he is already registered and makes him conscious about phishing. It is not browser dependent; fairly it is related with user's own saved information, so he can log on from any computer and from anywhere. This technique needs initial registration of the user to the correct website which has this facility. There is no chance of any false positive or negative as it is prevention based and not detection based.

In [9] they proposed Document authorship techniques. Anti-Spear phishing Content-based Authorship Identification (ASCAI) is used for mismatches between the writing styles of a received email body, ID-based authentication techniques which can be spoofed or mistreated. As a proof, SCAP is implemented. A classification accuracy of 87% can be considered up to standard compared to the simplicity of the implementation and the noisy dataset, higher classification accuracy is desirable to reduce false positives

In [10] The proposed model is based on FL operators which is used to illustrate the website phishing factors and indicators as fuzzy variables and produces six measures and criteria's of website phishing attack dimensions with a layer structure. fuzzy logic techniques is the use of linguistic variables to represent Key Phishing Characteristic Indicators and relating website phishing possibility .This experimental results showed the significance and importance of the phishing website criteria (URL & Domain Identity) represented by layer one, and the variety influence of the phishing characteristic layers on the final phishing website rate.

In [11] they identify a rule that can be explained as: if there exists an IP address of the URL in e-mail and it does not match the defined Rule Set for White List then the received mail is a phishing mail. It provides the feature of malicious status notification before the user reads the mail. Genetic algorithm is proposed, this algorithm is used to evolve rules that are used to differentiate phishing link from legitimate link. The parameters like evaluation function, crossover and mutation are evaluated.

In [12] Phish zoo technique is implemented , it is a blend of 5 features that is Profile making, Profile matching, Image matching using SIFT Running , Phish Zoo in Bulk and Online and offline profile matching. This approach provides similar accuracy to blacklisting approaches (96%), with the advantage that it can categorize zero-day phishing attacks and targeted attacks against smaller sites (such as corporate intranets). A key contribution of this paper is that it includes a recital analysis and a structure for making use of computer vision techniques in a practical way.

In [13] it has four features HTML Crosslink check, false info feeder check, SSL handshake and Certificate Suspicious check. This implies that using attribute-based anti-phishing checks can provide a viable defense against phishing, complementing signature and block list based defenses. This technique has been implemented in Phish Bouncer tool. The preface experimentation results show a 65% rejection rate for zero-day attacks using only 4 types of checks.

In [14] they give the research of why people fall for phishing attack. Phishers use the personal loaded webpage from the real Web site to make the phishing webpage appears exactly the same as the real one does. Digital watermarking is one of the most extensively used measures to protect digital information from copyright infringements. Counter measure techniques that are used for classifier.

In [15] Content-based phishing detection is greater to detection using white and black lists because it does not require the maintenance of lists. The keyword extraction method (if- idf) presently used for content-based detection is insufficient and causes a high rate of false positives. A system using these methods was implemented and evaluation with 172 legitimate Web sites demonstrated a reduction in the false positive rate from 14.0% to 7.6%.

3. FINDINGS

Phishing is the act of sending an e-mail to a user falsely claiming to be an established legitimate enterprise in an attempt to scam the user into surrendering private information that will be used for identity theft. Phishing is being combated through user education, legislation, and integrated anti-phishing measures in modern Web browsers. We focus on deceptive phishing using social engineering schemes. To protect users against phishing, various anti-phishing techniques have been proposed. To detect phishing web site mostly filtering methods, classifiers based on machine learning algorithm i.e. supervised and unsupervised learning and Visual Similarity Assessment based technique is to be applied.

4. FUTURE WORK

Here we will use Ant colony algorithm to detection of phishing attack, while processing of algorithm it generates multiple rules for the phishing data and suspicious links within short of time. This policy will help to protect client and server side attacks. URL and Domain Identity mechanism is used in this technique. ACO algorithm is used to illustrate and recognize all the factors and rules in order to classify the phishing website and the relationship that correlate them with each other. This algorithm is implemented in PHP and advanced java.

5. REFERENCES

[1] Ammar Almomani, B. B. Gupta, Samer Atawneh, Meulenberg, and Eman Almomani "A Survey of Phishing Email Filtering Techniques" IEEE Communications Surveys & Tutorials, Vol. 15, No. 4, 2013.

[2] T.Balamuralikrishna, N.raghavendrasai, M.Satya Sukumar "Mitigating Online Fraud by Ant phishing Model with URL & Image based Webpage Matching" International Journal of Scientific & Engineering Research Volume 3, Issue 3, March -2012.

[3] A.V.R.Mayuri "Phishing Detection based on Visual-Similarity" International Journal of Scientific & Engineering Research Volume 3, Issue 3, March -2012.

[4] Ahmed Abbasi, Fatemeh "Mariam" Zahedi and Yan Chen "Impact of Anti-Phishing Tool Performance on Attack Success Rates" 978-1-4673-2104-4/12/©2012 IEEE

[5] Weiwei Zhuang1,2, Qingshan Jiang2,3*, Tengke Xiong2 "An Intelligent Anti-phishing Strategy Model for Phishing Website Detection" 1545-0678/12 © 2012 IEEE.

[6] Mallikka Rajalingam, Saleh Ali Alomari, Putra Sumari "Prevention of Phishing Attacks Based on Discriminative Key Point Features of WebPages" International Journal of Computer Science and Security (IJCSS), Volume (6): 2012.

[7] Shivender Singh, Anil K. Sarje, Manoj Misra "Client-Side Counter Phishing Application using Adaptive Neuro-Fuzzy Inference System" 978-0-7695-4850-0/12 © 2012 IEEE.

[8] Madhuresh Mishra, Gaurav, Anurag Jain "A Preventive Anti-Phishing Technique using Code word" International Journal of Computer Science and Information Technologies, Vol. 3 (3) , 2012,4248 – 4250.

[9] Mahmoud Khonji, Youssef Iraqi, Andrew Jones "Mitigation of Spear Phishing Attacks: A Content-Based Authorship Identification Framework" 978-1-908320-00-1/11© 2011 IEEE.

[10] Maher Aburrous, M. A. Hossain, Fadi Thabatah, Keshav Dahal "Intelligent Phishing Website Detection System using Fuzzy Techniques".

[11] V.Shreeram, M.Suban, P.Shanthi, K.Manjula, "Anti-phishing detection of phishing attacks using Genetic Algorithm" 978-1-4244-7770-8/10/ ©2010 IEEE.

[12] Sadia Afroz, Rachel Greenstadt "PhishZoo: Detecting Phishing Websites By Looking at Them".

[13] Michael Atighetchi, Partha Pal "Attribute-based Prevention of Phishing Attacks" 978-1-4673-2104-4/12/ ©2012 IEEE.

[14] Huajun Huang, Junshan Tan, Lingxi Liu "Countermeasure Techniques for Deceptive Phishing Attack" 978-0-7695-3687-3/09 © 2009 IEEE.

[15] Shinta Nakayama and Hiroshi Yoshiura, Isao Echizen "Preventing False Positives in Content-Based Phishing Detection" 978-0-7695-3762-7/09 $26.00 © 2009 IEEE.