Département de mathématiques et Informatique

Travail personnel de l'Etudiant (TPE) *Système et logiciel en environnement Distribué (SLED)*

19 Mars 2017

Phishing Detection

NGOTCHIE CHRISTIAN JOEL

13A031FS

VANDI TIZE

10A665FS

Supervisé par : Dr-Ing. **FRANKLIN TCHAKOUNTE**

Sécurité des systèmes et des réseaux M1-SLED 2016-2017

1. INTRODUCTION

1.1- Context

Phishing is a social engineering attack that aims at exploiting the weakness found in system processes as caused by system users, it is a method used by cybercriminals to steal personnl credentials and information through the Internet in oder to use that for financial fraud. According to AntiPhishing Working Group (APWG) [1], Phishing is a crimminalmechanism employing both social engineering And technical subterfuge to steal consumers, personnal identify data and financial account credentials. Over the years phishing has becom a very drastic criminal act on the Internet. Phishing attacks are directed to reveal and access confidential information for instance credentials, credit card, regarding information, in oder to make use of these data maliciously. Since phishing attacks aim at exploiting weaknesses found humans (system end-users), it is difficult to mitigate them.

Though much research on antiphishing techniques has been done and new techniques and methodologies are being proposed regularly, online scammers manage to come up with innovative schemes to improve existing detection technologies and lure potential victims to their phishing campaigns.

For example, as evaluated many case, end-users failed to detect 29%of phishing attacks even when trained with the best performing user awareness program. On the other hand, software phoshing detection technique are evaluated against bulk phishing attacks, which makes their performance practically unknown with regards to targeted forms of phishing attacks.

1.2- Background

a) History

According to APWG, the term phishing was coined in 1996 due to social engineering attacks against America On-line (AOL) accounts by online scammers. The term *phishing* comes from *fishing* in a sense that fishers (attackers) use a bait (socially-engeneering messages) to fish (steal personal information of victims). The origins of the ***ph*** replacement of the character ***f*** in fishing is due to the fact that one of the earliest forms of hacking was against telephone networks, which was named ***phone phreaking***

As a result, ph became a common hacking character replacement of ***f***.

According to APWG stolen account via phishing attacks were alsoused a currency between hackers by 1997 to trade hacking software in exchange of the stolen accounts.

Phishing attacks were historically started by stealing AOL accounts, and over the years moved into attacking more profitable targets, such as on-line banking and e-commerce services.

b) Phishing motives

Primary motivation behing phishing attacts , from an attacker's perspective, are :

• ***Financial gain:*** phishers can use stolen banking credentials to their financial benefits.

• ***Identity hiding:*** instead of using stolen identities directly, phishers might sell the identities to others whom might be crimmals seeking ways to hide their identities and activities.

• ***Fame and notoriety:*** phishers might attack victims for the sake of peer recognition.

As ather phishing motives we can also enumerate : ***identity theft, Identity Traffcking, industrial espionage, malware distribution, exploit secure Hole, Password Harvesting.***

1.3- Problem and research questions

Phishing attack is a dangerous cyber-threat that target computer users and leverage the human vulnerability rather than exploring technical errors and and weaknss. Many reseacgers repeatedly mentioned that, most important factor in computer system sis human focused ; so that means the users must be trained and educated to ensure that are well protected and safe when using online services. Some researchers have also mentioned that well designed end-users training and education can be effective to mitigate such threats. This can be webbased educational materials or embedded training system to protect user against phising attacks. Howerver, that it is amazing that the increasing development of antiphishng services and technologies, phishing education, the number of phishing email messages and other phishing technique continues to increase rapidly.

The main research is to give some answer for this question : **Can Phishing Education Enable Users To Recognize Phishing Attacks?**

2. Related works

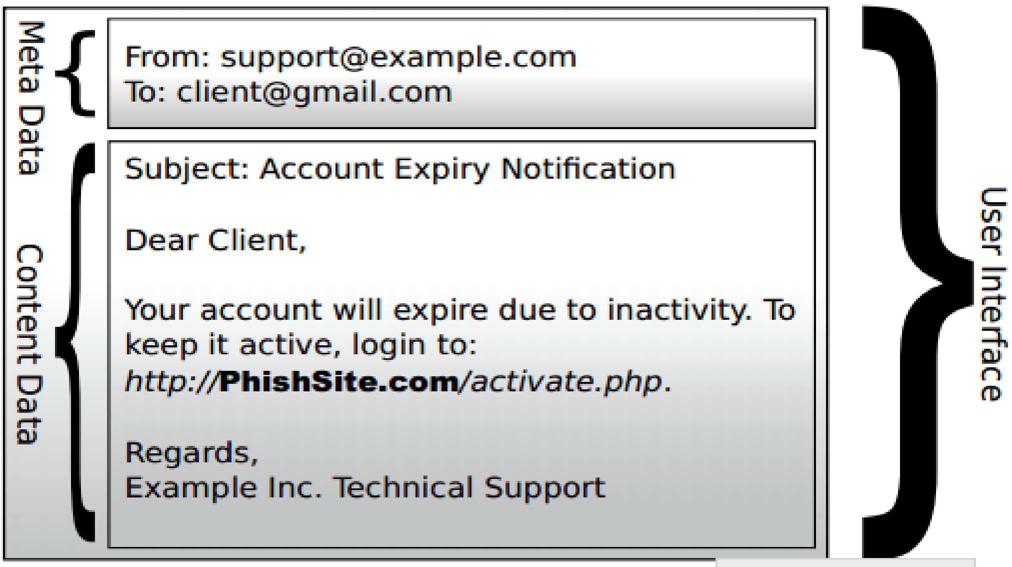
Work	***Phishing email or SMS***	***Phishing URLs***	***Phishing webpages***	***Phishing VoIP***
[1]	✓	✓	✓	
[2]	✓		✓	
[3]	✓		✓	✓
[4]	✓	✓	✓	✓
[5]		✓		
[6]		✓	✓	
[7]		✓	✓	✓
[8]	✓			
[9]	✓	✓	✓	✓
[10]	✓			✓
[11]	✓	✓	✓	✓
[12]	✓	✓	✓	✓
[13]	✓		✓	
[14]	✓			
[15]	✓			✓
[16]				✓
Phishing attack detection techniques based on the machine learning methodology				
URL blacklist technique	✓	✓	✓	
A Web Mining Approach to Detecting Phishing URLs		✓		
Learning to Detect Malicious Web Sites from Suspicious URLs	✓	✓	✓	
Google Safe Browsing API		✓		
SpoofGuard		✓	✓	
Public Switched Telephone Network (PSTN) and Voice over IP (VoIP)				✓
Visual Similarity-based Detection without Victim Site Information	✓	✓	✓	

Automatic Detection of Phishing Target from Phishing Webpage		✓	✓	
Detecting phishing emails using hybrid features	✓		✓	

3. Medology employed

Phishing is a semantic attack that uses electronic communication channels to deliver content with natural laguages(English, Aarabic, French, …) to persuade victims to perform certain actions. Lately phishers have been making use of various channels, these channels range from instant messages to e-mails, facility of phone (text message and calls), and Social Network. Several form of phishing attacks are employed day-to-day by the ciber-attackers.

- **Emails Phishing Attack**
  Today, the importance of making use of e-mail is extremely significant, and almost all the people throughout the world need to use it for either personal or business use. The difference betwen clean e-mail and a phishing e-mail is not evident ; Following important signs have been identified which can help anyone to identify a phishing email. These signs are**:**
  - ***Using urgent words in messages***, Attackers usually use words that show urgency and make the victims get results very quickly.



From: support@example.com
To: client@gmail.com

Subject: Account Expiry Notification

Dear Client,

Your account will expire due to inactivity. To keep it active, login to: *http://***PhishSite.com***/activate.php*.

Regards,
Example Inc. Technical Support

- ***Nesting a request of sensitive and personal information*** : This is the aim of most attackers to deceive their targeted victims into trusting them to gain their personal information in details.
- ***Unknown Sender:*** This kind of phishing attack is more concentrated and what the attacker does in that within the same corporate domain they impersonate a co-worker.
- ***The contents of the mail are structured using poor language and it usually contains spelling and grammatical errors.***
- ***Contains images and logos to visually represent impersonation.***

- **Vishing**
  The stated term has been derived from two different words ***voice and phishing***. It is the term used to illustrate the practice of using voice messages based on IP (VoIP : Voice over Internet Protocol) in order to socially plot a targeted user and make them reveal their financial ,personalor any other kind of confidential to gain some financial benefit from it. The fact that Vishing is likely to have an even greater success rate than those of other phishing methodologies can be the fact that : The number of people that can be accessed via the phone is greater than that via email ,Telephones are a great way of personalizing the message of social engineering, Message timing can be made use of in order to increase the chances of success .

- **SMishing**
  This phenomena which is termed as Smishing, the word has been derived from the combination of SMS and phishing.

- **Social Media Phishing**
  Phishing attackers can refine and improve their tactics through the use of readily and publicity available informatio from the greatestsource (social media).

4. Results found and Perspectives

Often information that transit throught Internet being transmitted is valuable and sensitive such that effective protection mechanisms are desirable in order to prevent information from being manipulated or to protect confidential information from being revealled by unauthorized parties. From a high-level perspective, there are generally two commonly suggested solutions to mitigate phishing attacks : ***Implementation of technical procedure,***

User education, commonly designed by ***technical*** and ***non-technical phishing detection solution.***

- ***Implementation of technical procedure.***

  There are number of different technical solutions available, all require some level of technical implementation and more importantly awareness of the issue of phishing in the first place. Different technical Efforts are use to Reduce Phishing Attacks, most importantly ***Antiphish in web browser toolbars, strong authentication and autorisation, Virus, Spyware, and spyware prevention.***

  - ***Antiphish in web browser toolbars*** ***:*** AntiPhish is a browser extension which is used to protect inexperienced users against spoofed web site-based phishing attacks . it is a plug-in tool which keeps track of user's sensitive information and prevents this information from being passed to a web site that is considered untrustworthy or unsafe.

  - ***Strong authentication and authorization*** ***:*** Essentially, users ought to be authenticated with the help of information that may not be revealed readily by them to parties which may use it for hazardous use. This can be achieved if security is not relied on users and at the same time the process of authentification is enhanced to ensure maximun security and frauds are elimined. Malware is developing to attain specific information that is sensitive and which has particularly been created to authorize online transaction of commerce. Another kind of authentication that can be utilized to increasesecurity is the two factor authentification, which requires multiple authentificators. This can have information like password or PIN or something that can be possessed for instance hardware token or a credit card.

  - ***Virus, Spyware, and spyware prevention*** ***:*** A common definition for spam does not exist, but majority of the sources approve the core idea of the phenomenon that such message are not solicited. The most famous and well developed technique to anti-spamming is a filter thatis based on learning. Currently, filters use machine learning, and artificial intelligence's algorithms. Solutions that aim at protecting users from Trojans, viruses, spam and spyware play a significant role when it comes to phishing threats.

- ***User education***

  The human is educated in an attempt to ehnace his/hre classification acccuracy to correctly identify phishing message, and then apply proper actions on the correctly classified phishing message, such as reporting attacks to system administrators. Anti – phishing researchers have developed different methods to help internet users recognize and avoid phishing attacks. Much of their work focused on developing technique towarm users about phishing website. However, less effort has been concentrad on training users and educate them how to identify such threats.

Users should be trained so that they do not blindly follow links to websites, particularly, where they are used to provide their credentials is the most rffective remedy of phishing threats. But, thing that the entire group of users will perceive and understand the threats of phishing browse accordingly is not realistic.

Always to strunggle against phishing attacks another solution are found :

- **Phishing detection by BLACKLIST and Visual similarity :** Blacklists are frequently updated lists of previously detected phishing URLs, Internet Protocol addresses or Keywords. The detect phishing attacks based on their visual appearance, as oposed to analyzing the the underlying source code or network-level information.
- **Phishing detection by data mining**
- **Phishing detection by neural network.**

Phishing attack is a kind of social engineeringwhichbecoming increasingly sophisticate, all of the proposed solutions attempt to minimize the impact of phishing attacks. Phishing has spread to not only include e-mails but also SMSs, VoIP and social media. Despite of the number of mechanisms an organazationuse to avoid phishing attacks such as firewalls, encryption software, certificates, and two factor authentications, i twill not be useful if the person behind the computer falls for phish.

For the future work the aim of developing phishing quiz application is to create a better educational environment that suites regular web users and provide feedbacks which enhance them to learn from anti-phishing materials and get progress on their scores. Therefore, future research can be focused on developing a highly sophisticated anti-phishing mobile application that can simulate a real life phishing threats with attractive visual objects and graphics.

5. Most important publications for this work

Phishing detections are an current events in internet universe and many researchs and publications to concentrate on this subject. Among several publications for this work we can to list some important :

[1] Aanchal Jain, Prof. Vineet Richariya, ***Implementing a web Browser with Phishing Detection Technique,*** *(WCSIT) ISSN 2221-0741 vol 1, no 7,289-292, 2011*

[2] Jean-François PILLOU,fondateur de CommentCaMarche.net, ***Le phishing (hameçonnage)***, septembre 2015

[3] *Gérard Peliks, Président de l'atelier sécurité Forum ATENA,* ***Le business model du Phishing****, Janvier 2008*

[4] Youssef Iraqi, Andy Jones, ***Phishing Detection: A Literature Survey*** *, IEEE Communications Surveys & Tutorials · April 2013*

[5] Ram B. Basnet, Andrew H. Sung, Quingzhong Liu, ***Rule-Based Phishing Attack Detection***

[6] Joshua S. White, Jeanna N. Matthews, John L. Stacy, ***A Method For The Automated Detection Of Phishing Websites Through Both Site Characteristics And Image Analysis,***

[7] Ningxia Zhang, Yongqing Yuan, ***Phishing Detection Using Neural Network,***

[8] OUCH, security the Human, ***Attaques par phishing*** Février 2013

REFERENCES

[9] The Anti-phishing Working Group, http://www.Antiphishing.org

[1] Aanchal Jain, Prof. Vineet Richariya, ***Implementing a web Browser with Phishing Detection Technique,*** *(WCSIT) ISSN 2221-0741 vol 1, no 7,289-292, 2011*

[2] Jean-François PILLOU,fondateur de CommentCaMarche.net, ***Le phishing (hameçonnage)***, septembre 2015

[3] *Gérard Peliks, Président de l'atelier sécurité Forum ATENA,* ***Le business model du Phishing****, Janvier 2008*

[4] Youssef Iraqi, Andy Jones, ***Phishing Detection: A Literature Survey*** *, IEEE Communications Surveys & Tutorials · April 2013*

[5] Ram B. Basnet, Andrew H. Sung, Quingzhong Liu, ***Rule-Based Phishing Attack Detection***

[6] Joshua S. White, Jeanna N. Matthews, John L. Stacy, ***A Method For The Automated Detection Of Phishing Websites Through Both Site Characteristics And Image Analysis,***

[7] Ningxia Zhang, Yongqing Yuan, ***Phishing Detection Using Neural Network,***

[8] OUCH, security the Human, ***Attaques par phishing*** Février 2013

[10] Sudhir Shashidhar, Research Fellow, IIT Kanpur , ***Spear Phishing - The New Face of Phishing,*** *April 26,2016 from* ***http://www.trendmicro.com/***

[11] Mahmoud Khonji, Youssef Iraqi, *Senior Member, IEEE,* and Andrew Jones, ***Phishing Detection: A Literature Survey,*** *IEEE COMMUNICATIONS SURVEYS & TUTORIALS, VOL. 15, NO. 4, FOURTH QUARTER 2013*

[12] Hanaa Alghamdi, ***Can Phishing Education Enable Users To Recognize Phishing Attacks ?*** 03 January *2017*

[13] L. James, ***Phishing Exposed****. Syngress Publishing, 2005*

[14] P. Kumaraguru, Y. Rhee, A. Acquisti, L. F. Cranor, J. Hong, and E.Nunge, ***"Protection people from phishing :the disign and evaluation of an embedded training email system"****,New York, NY, USA: ACM, 2007, pp. 905–914.*

[15] A. Stone, ***"Natural-language processing for intrusion detection,"*** *Computer, vol. 40, no. 12, pp. 103 –105, dec. 2007.*

*[16] G. Liu, B. Qiu, and L. Wenyin****, "Automatic detection of phishing target from phishig webpage "*** *in patter Recognition(ICPR), 2010 20th internationnal Conference on aug.2010 pp4153-4156*