International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 7 Issue 03, March-2018

Conceptual Architecture for Identifying and Avoiding Phishing Attacks

Karunambika R
PG Scholor
Department of Computer Science and Engineering
PSG College of Technology
Coimbatore, India

Dr. G.R. Karpagam
Professor,
Department of Computer Science and Technology
PSG College of Technology
Coimbatore, India

***Abstract*— Phishing is a form of social engineering attack often used to steal user data, including login credentials and credit card numbers. It occurs when an attacker, acts as a trusted entity and ask to open an email, instant message, or text message. The recipient is then tricked into clicking a malicious link, which can lead to the revealing of sensitive information. An attack can have devastating results. For individuals, this includes unauthorized purchases, the stealing of funds, or identifies theft. An organization succumbing to such an attack typically sustains severe financial losses in addition to declining reputation, and consumer trust. So, from the wide literature survey we proposed a novel solution for this attack. This work includes the conceptual architecture, components, D-P-P-F (Detection prevention phishing Framework).**

1. MOTIVATIONAL SCENARIO

Phishing is an act of acquiring private and sensitive data from internet users for use in fraudulent activities. It is usually done by sending emails that seem to appear to come from original sources (for example banking website). Some of the criminals behind phishing attack have even gone so far that they create fake websites that appears to be operated by government agencies. Phishing attack steals credential information of the internet user like bank account, credit card information, and user's full identity like personal details, email address, username and password. Phishers simply send an email with a link to the fake site which appears like original site. Email says "your password is going to expired so, please update your password within 24 hours" this, forces user to get into a link and update password. Phishers can steal money or identity of the user and they can use that information for illegal purpose. So, preventing or avoiding phishing is a crucial thing. This section outlines the types of phishing.

1.1 Deceptive Phishing

Deceptive phishing is the most common type of phishing attack which emails you with a link to a bank website with some attractive features like 'if you transfer the amount through this link of our bank, your transaction fee is very less or you have no transaction fee. When you click the link, it redirects you to a phishing site of a bank. Phishing site requires to re-enter username, password, and account information. That bogus site collects all secret information and many other malicious sites are sent to many recipients.

1.2 DNS-Based Phishing

Domain Name System (DNS)-based phishing is also called 'Pharming' where host file is modified. In this pharming scheme, hackers interfere with a company's hosts files or domain name system so that requests for URLs or name service return a bogus address for a website and subsequent communications are re-directed to a fake site which is controlled by hackers and is probably not even in the same country as the legitimate website.

1.3 Content-Injection Phishing

This type of attack modifies the part the original content with fake content in the original site. That fake content misleads the user into giving up their confidential information to the hacker.

1.4 Man-in-the-Middle

In this type of attacks, hackers sits in between the user and the legitimate website or system. They silently record the information which are communicated by those two without interrupting them. Later when the user is not active on the system they can collect those information for fraudulent activities.

1.5 Search Engine Phishing

Here phishers create fake banking websites with too attractive offers like better interest rate than other banks and have them indexed legitimately with search engines. Users can find the sites in the normal searching process for products or services and are fooled into giving up their information.

The remaining section of the paper is organized as follows. Section II of this paper gives the related work about phishing attacks and anti-phishing techniques and prevailing scenario. Section III gives the proposed work D-P-P-F. Section VI concludes the paper.

2. RELATED WORKS

Protecting Users Against Phishing Attacks paper describes an application called 'Antiphish' was developed by Engin Kirda and Christopher Kruegel. (2005). Antiphish integrated with a browser Mozilla Firefox. When user enters sensitive information into the website, it prevents those information from not captured by fake site by encrypting the sensitive information using DES algorithm. Drawback is this application can't be used with other browser.

An Intelligent phishing detection system for e-banking using fuzzy data mining was developed by Aburrous et al. (2010). This paper proposed a model to characterize the e-banking phishing website factors. For that this model uses fuzzy logic with data mining algorithms and also defines six e-banking phishing website attack criteria's with a layer structure.

An Intelligent phishing detection and protection scheme for online transactions was developed by P.A. Barraclough et al. (2013). Their approach was based on utilize a Neuro-Fuzzy Scheme with 5 inputs to detect phishing sites with high accuracy in real-time. The objective of this idea is to extract the phishing features based on 5 inputs, build a Neuro-Fuzzy model and train and validate the Fuzzy Inference model in real-time environment. The advantage is to make users more secure and build their confidence in online transactions.

AuntieTuna:Personalized Content-based Phishing Detection model was developed by Calvin Ardi et al. (2016). AuntieTuna automatically builds personalised list of target sites and test those sites as user browses them. It uses cryptographic hashing of the browser's Document Object Model (DOM) to detect phishing site. Advantages are, AuntieTuna does not slow web browsing time and displays alerts on phishing pages before users give information.

2.1 Prevailing Scenario

In normal case, user is given a requested webpage by web server but, in phishing attacker hosts bogus sites in the server and send email links for that sites to the user. If any user opens the link, then phisher can collects credentials of user from the data collection point. Figure 2 shows different attacks possible in different situation. In that user may access different websites for different purpose. While they open for login (authentication) purpose, phisher may steal username and password so deceptive attack is possible there, if that is for key exchanging purpose then there is possibility for man-in-middle attack, or if they open for credential update purpose then search engine phishing is possible there.

2.1.1 Phishing detection and avoiding

Web crawling setup can be manually programmed for detecting and sending warning message to the user about the phishing site when they are trying to access bogus site instead of original site. For detecting, it compares the entered URL(Uniform Resource Locator) with the sites in blacklist. It collects phishing sites based on some keywords like image, content or certificate and maintains those sites in the blacklist for each original site. If there is a mirror site for original site then it keeps that mirror site with original site in the whitelist.

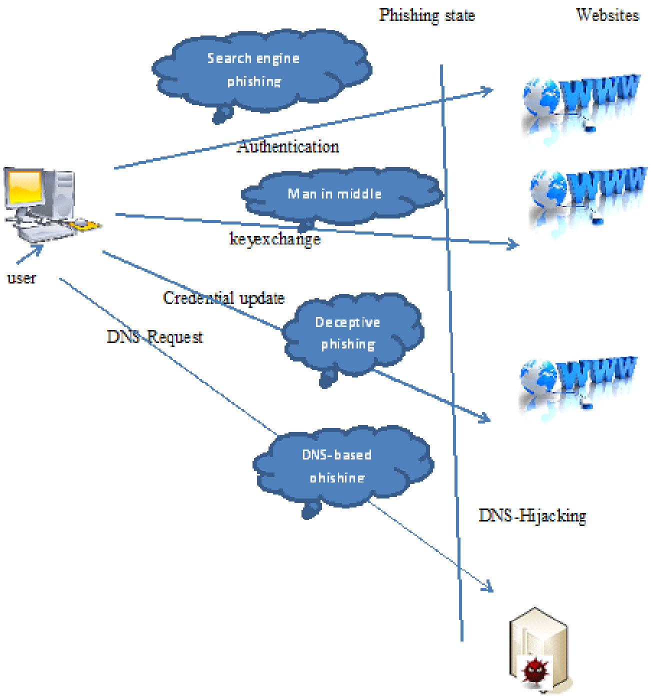

Fig.1. Phishing Attack

2.1.2 Phishing sites identification

Phishers can create fake websites with less amount of work. They use brand name, logo, images and design of the original website to create a fake website. So, identifying these fake websites for a particular brand is a difficult process in the real time, some brand has no fake sites and some has more number of fake sites.

Server logs or referrer logs provided by the visitors can be used by the crawl algorithm to collect fake links or sites of an original site. When someone visits a particular website, server software counts and tracks (logs) that visit. It also keeps a record of it for a certain period of time. Part of the saved information is called a referrer log. Referrer log consists the following data,

- Which engines have sent you traffic
- What keywords were used to find your site
- Which pages were accessed the most or the least
- Who are the visiting spiders
- User profile by region
- Average length of time someone remains on your site
- Average number of user sessions or page views per day
- Top entry and exit pages
- Top referring sites
- Summary of activity by day
- Server errors
- Bandwidth, which is the measure (in kilobytes of data transferred) of the traffic on the site and,
- Type of technology used by your visitors.

Based on the lessons learnt from related work the objective of this work is to design and develop a novel solution for identifying and avoiding phishing attack.

3. PROPOSED WORK

This section focus on the design and development of phishing detection and prevention framework, its components and workflow with an example scenario. The following section list the various components and its purpose.

3.1 COMPONENTS

1. Crawler – Compares entered URL with blacklist websites. For example the original site and mirror site will be in whitelist (www.psgtech.edu) and bogus (fake) sites of this will be in blacklist(www.psgtechh.edu, www.psg-tech.edu, www.psgtechhnology.edu).
2. DB Manager – Maintains those two list (whitelist and black list).
3. Monitoring Agent – It watches users activity(URL) and keep searching for sites with keywords and if it finds it sends to update agent.
4. Update Agent – Update agent get the input sites from monitoring agent and update those sites into the database.
5. Alert Agent – If crawler finds that the user entered site is bogus(fake) one, then it sends alert message to the user.

3.2 CONCEPTUAL ARCHITECTURE

This section gives the work flow of D-P-P-F,(shown in fig.2). Internet user triggers the D-P-P-F. Monitoring agent in D-P-P-F get the URL and send to updating agent which updates database likewise web crawler also gets the URL and compares with two lists in database. If it finds that this is a phishing site then alert agent send alert message to the user.

Figure 3 shows the sequence diagram of the working flow of each component in crawling for phishing site.

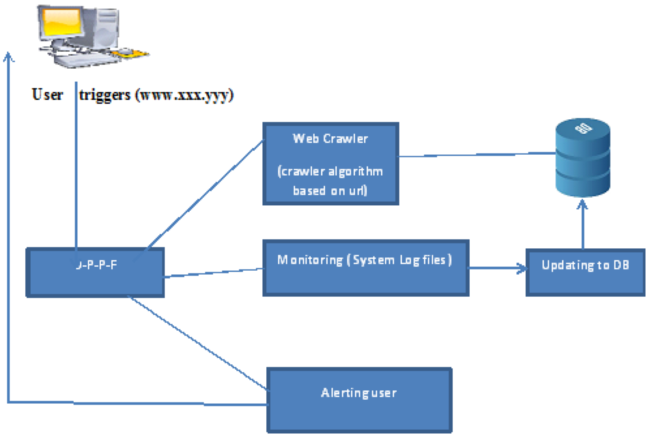

Fig.2. Working flow of D-P-P-F

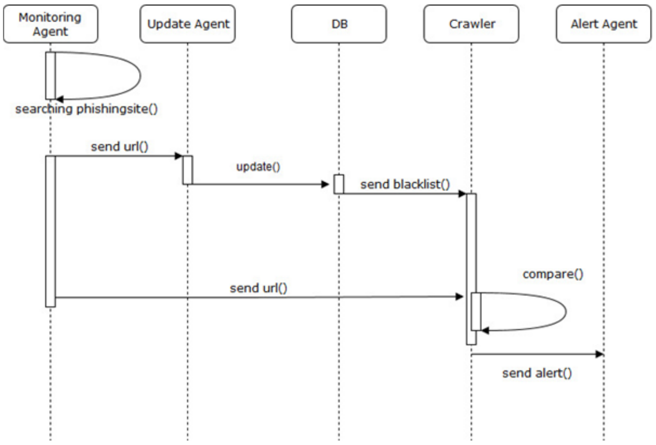

Fig.3. Sequence Diagram

3.3 Banking Scenario

Figure 4 shows D-P-P-F for banking scenario. In that while user enters into a banking website monitoring agent saves this URL into database. At the same time crawler compares entered URL with the whitelist and blacklist in the database. If crawler identifies that user enters into a phishing site then it collects the information from the user, encrypt it and send a warning message to the user. Because of this encryption, attacker can get only the encrypted information so nothing to worry. In case if that entered URL is not matching with both list means it sends to an update agent which adds that URL into blacklist as new phishing site.

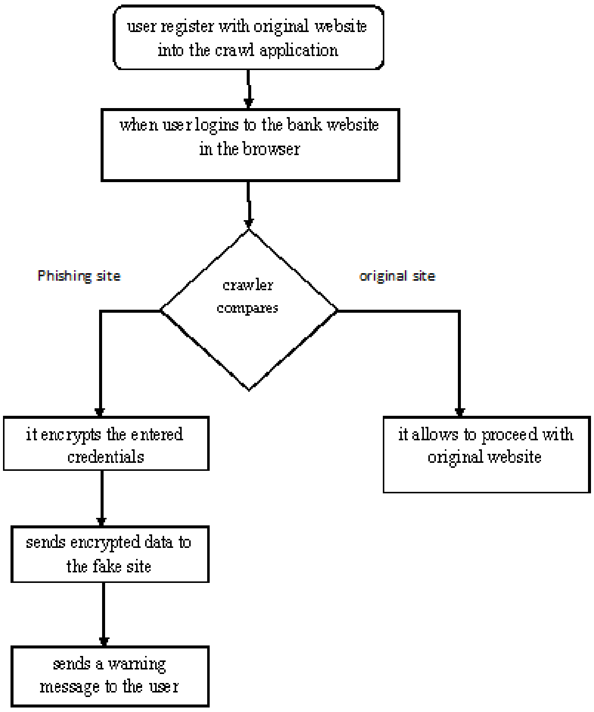
user register with original website into the crawl application
when user logins to the bank website in the browser
crawler compares
Phishing site
original site
it encrypts the entered credentials
it allows to proceed with original website
sends encrypted data to the fake site
sends a warning message to the user

Fig.4. Applying D-P-P-F For Banking Scenario.

4. CONCLUSION

The phishing problem does not have a single solution. This paper gives one solution(concept) for solving phishing attack using web crawling, but identifying phishing sites for particular brand in real time is a difficult thing. If we can find phishing sites dynamically in real time means that gives a better way to detecting avoiding phishing attack.

REFERENCES

[1] Engin Kirda and Christopher Kruegel 2005 ," Protecting Users Against Phishing Attacks with AntiPhish". Computer Software and Applications Conference, COMPSAC 2005. 29th Annual International (Volume: 1).

[2] Maher Aburrous, M.A. Hossain, Keshav Dahal and Fadi Thabtab, "Intelligent Phishing Detection System for e-banking using fuzzy data mining," in *Expert Systems with Applications 37 (2010) 7913–7921.*

[3] P.A. Barraclough, M.A. Hossain, M.A. Tahir, G.Sexton and N. Aslam, "Intelligent Phishing Detection and Protection Scheme for online transactions," in *Expert Systems with Applications 40 (2013) 4697–4706.*

[4] Calvin Ardi, Jhon Heidimann, "Auntie Tuna: Personalized content based Phishing Detection," in *USEC '16, 21 February 2016, San Diego, CA, USA Copyright 2016 Internet Society, ISBN 1-891562—42-8.*

[5] Seoung Yeop Na, Hyun Kim and Dong Hoon Lee 2014," Prevention Schemes Against Phishing Attacks on Internet Banking Systems" International Journal of Advance Soft Computing Application, Vol. 6, No. 1, March 2014 ISSN 2074-8523.

[6] V. Suganya, "A Review on Phishing Attacks and Various Anti Phishing Techniques," Int. J. Computer. Appl., vol. 139, no. 1, pp. 975