INTERNATIONAL JOURNAL OF
RECENT TRENDS IN MULTIDISCIPLINARY RESEARCH
November-December 2024, Vol 4(06), 01-03. ISSN No: 2583-0368

Ontology-Based Domain Framework for Enhanced Anti-Phishing Measures

Sana Khan[1], Mani Priya Mishra[2], Rukaiya Khatoon[3]
[1,2,3]B. Tech Student, Department of Computer Science, Institute of technology and Management (ITM), Gida, Gorakhpur, Uttar Pradesh, India.

OPEN ACCESS
Article Citation:
Sana Khan[1], Mani Priya Mishra[2], Rukaiya Khatoon[3] "Ontology-Based Domain Framework for Enhanced Anti-Phishing Measures", International Journal of Recent Trends in Multidisciplinary Research, November-December 2024, Vol 4(06), 01-03.

Publishedby5thDimension Research Publication

Abstract: Phishing in the mobile network is a recent threat that extracts sensitive data from the users especially in online shopping, social networking etc. The hackers can target the mobile users using phishing web pages when they are browsing. The various reasons behind the mobile phishing attacks are small screen size, lack of identity index, preferences and usage of the mobile users etc. In recent times, the mobile browsers have become capable enough to support every kind of web browsing such as online banking, online shopping, online socializing, etc. Finally, it is ended to share sensitive data to the phishing websites. Users deal with certain massive experiences in mobile platform. The mobile phones have certain limitations when it comes to the hardware part of the device. In order to enhance the user experience, various well-known companies have developed mobile applications that have in turn brought the phishing in mobiles into the spotlight. The phishing application succeeds in deducing the user 's credentials from the unsuspecting users that are stored in the phishing servers.

Key Words: Anti-Phishing, index, Account Phish Defense,

1.System Architecture

System architecture provides a domain ontology based system. OMPD is proposed as an automated method for tackling the mobile phishing attacks. OMPD has two key components, namely, Ontology based Webpage Phish Defense (OWPD) and Ontology based Account Phish Defense (OAPD) that are capable of protecting the mobile web pages and persistent accounts, respectively.

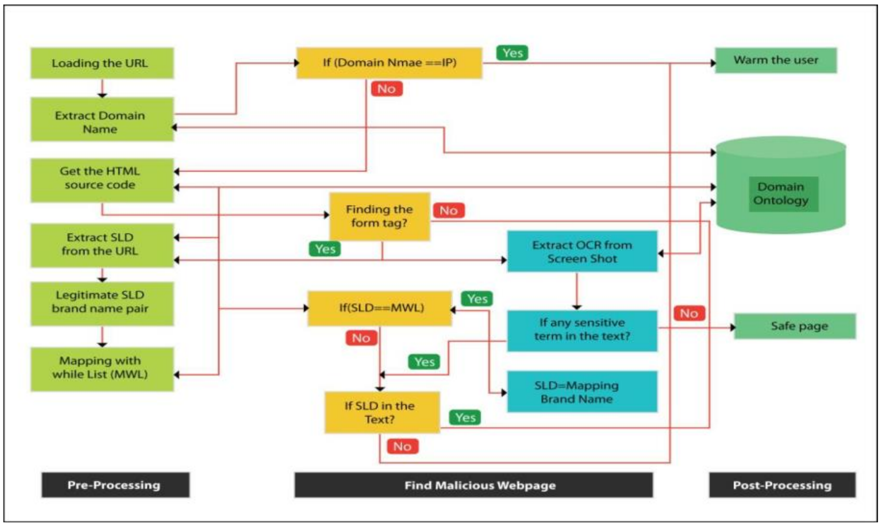

Figure 1 Architecture of OWPD

2. Working Methodology

This section focuses the techniques that implement the OMPD technique. This includes both account phishing and web page phishing in mobile phones. We have used 1000 random phishing URLs to identify the phishing target.
The OWPD scheme starts with the loading of the URL. The OWPD scans the URL when the browser loads the webpage and then it is sent to the ODB (Ontology Database) to investigate if the URL is valid or trusted.

The contrast between an original site and a phishing site is that domain names are utilized as confirmations in the event of the trusted web pages, whereas IP addresses are used by the phishing site. The user is intimated about a fake web page in case an IP address is used. The OWPD gets the HTML source code of the loading page to find the form of the page.

Structures are required by the phishing pages with the goal that the clients can enter the information for submission like a trusted site. This strategy does not require that every one of the pages are examined as the OWPD, checks the presence of structures. In any case, the centre module of identity extraction does not depend on any piece of HTML source code. The HTML source goes to the ODB for the checking of the form tag. It will move to a protected page if there is no form. If there should be an occurrence of finding of the form, the further identity extraction and verification is started by the OWPD. The genuine character of the site that is signified by the second level domain name (SLD) from the URL is extracted. At that point the SLD is ordered in the Mapping White-List (MWL). If there is any match in the MWL

3. OAPD Method

All the applications in the mobile are powerless against account registry phishing attacks as the clients may not recollect all the applications introduced in the gadget. In light of the three sorts of phishing attacks, specifically, sort A, B and C, the domain ontology is grouped. In the sort A attack, the objective account gets an unexpected application in comparison to the original application (e.g. an amusement application enrolls a Face book account). In the sort B attack, the vindictive application does not show up inthe main menu by any stretch of the imagination. In the sort C attack, the malicious application specifically appears as the objective application (e.g. repackaged application), which implies they will have a similar application name appeared in the main menu as the account name that shows up in the account list. The possibility of the detection mechanism for the sort A and sort B account phishing attacks is to think about the application name in main menu and the account name in the account list.

ODB returns sort B attack if the application name is NULL; else the APP name is checked in the Account mapping white list (AMWL). The malicious application can powerfully enlist and change the account data. The OAPD ought to be skilled such that it can check the enrollment of accounts in runtime that should be possible just by the adjusting source code. Each time an account gets included, the account label is separated and contrasted and the host application name.

4. Domain Ontology

The domain ontology is created on the basis of the phishing keywords taken from the accounts and the web pages. In order to find the domains, all the extracted keywords are mapped to the ontology by using the OBIE (Ontology Based Information Extraction). In the first step of the domain identification, Triplet Extraction Algorithm is employed and the subject, predicate and object are extracted from the keywords. By the usage of NLP method, the subject is identified and mapped to the semantic class. It uses the predicate and object as name and value of the attribute respectively. Theme identification forms the next step which is the nouns (extracted tokens) and they form the concept set. The identified subjects are included in the subject list. Max Occur Concepts are those concepts that occur for the maximum number of times. The theme identification is done by intersecting the three sets obtained.

Domain selection Process:

For URLs, there can be found many variations between phishing URLs and legitimate URLs which can be used to identify very easily depending on URL characteristics. In particular, there are 3 features:
Primary Domain, Subdomain and Path Domain of the URL. Primary Domain: Phishers are unable to utilize the original Primary domain because it has already been authorized by the authentic organization. Therefore, phishers register misspellings or perhaps identical Primary domain of phishing websites in order to trick end users. For instance, URL www.amazonn.com seems identical to the well-known website www.amazon.com.

- **Sub Domain:** Phishers frequently prepend the domain associated with phishing sites to their website. As an example, phishers prepend the Sub Domain ―amazon.com‖ to some different domain (e.g., ―.io‖, ―.biz‖) that could mislead users into the phishing URLs.
- Path Domain: It is a sub-folder in the URL. Phishers may also utilize the Path domain to deceive users. As an illustration, phishers may possibly navigate users to the URL www.attack.com/amazon, in which a phishing web page interface resembles the original one. Thoughtlessly, the users will certainly believe that this specific URL is actually from the ―amazon.com‖ web site. Especially, utilizing mobile devices along with small graphical interfaces, it might be very hard to identify these kinds of phishing URLs.

5. Conclusion

The previously mentioned experiments approve the productivity of the two sub plots that we have created. The execution of the OWPD is evaluated by the estimation of the execution time of this plan (Figure 6.7).
There are three key techniques that are utilized in the OWPD – making the domain ontology, looking of the SLD from the separated content of the screenshot and blocking of the SMS and socket connections for some interval of time. OCR method is utilized for the SLD looking through that expends much time. In this manner the evaluation of the delay overhead of OCR based

strategies and other non-OCR systems are done independently. Table 6.5 shows the outcomes. There are three stages in the OCR-based methods - taking a screenshot, extracting the content from the screenshot, and looking if the SLD exists in the content.

References

1. *Abu-Salih, B.; Qudah, D.A.; Al-Hassan, M.; Ghafari, S.M.; Issa, T.; Aljarah, I.; Alqahtani, S. An intelligent system for multi-topic social spam detection in microblogging. J. Inf. Sci.* ***2022****.*
2. *Zantal-Wiener, A. 47% of Social Media Users Report Seeing More Spam in Their Feeds, even as Networks Fight to Stop It. 2019.*
3. *Barati, R. Security Threats and Dealing with Social Networks. SN Comput. Sci.* ***2022****, 4, 9.*
4. *Rodrigues, A.P.; Fernandes, R.; Shetty, A.; Lakshmanna, K.; Shafi, R.M. Real-time twitter spam detection and sentiment analysis using machine learning and deep learning techniques. Comput. Intell. Neurosci.* ***2022****, 2022, 5211949.*