A Novel Customised Anti Phishing Framework for Mobile Environment

Shaik. Shahid¹, Dasari Rajesh²

¹,²Computer Science and Engineering, Rise Group of Institutions, Ongole, India

Abstract: Now a day’s many phishy websites are created by the attackers to extract all the personal information from the users. Many people use mobile internet through mobile phones. The proposed anti phishing tool to detect the phishing websites is very helpful to save the users from many fraud websites. The WIFI internet connection is used to deliver the internet service to the mobile device. We extract some main criteria and relates with the protected websites. Thus the similarities between the protected websites and the fraud websites are tested visually by various algorithms. Finally the phishing report is shown how far the obtain website is genuine. The alarm is fixed and produces an alert sound if the given website is phishy website.

Keywords: Phishing, Websites, Visualization.

1. Introduction

Phishing is a new trend to find the fraud websites form million website pages used now a day’s. As the technology improves our security towards it also should be improve accordingly. A common user can access any kind of websites which they need to. But the concept is no one see to it whether we use a genuine website which we access to transact the money and etc. So these ways of attacks mainly happens while transactions when the user gives their personal details, pin numbers and so on. These phishing websites can be found in many ways. The idea to use data mining concept is very useful than other applications, because it says mining a data or information from huge database. Data mining can provide a way how to find these phishing websites with the help of many applications and algorithms

In [1] the author detects the phishing detection system for the e-banking. The paper states that the phishing websites mostly get the e-banking sites and attack their passwords, credit card number, bank account and personal details of the user. He says it’s a “New Internet Crime”. Comparing with the forma like virus and hacking the phishing is mostly popular now days. In this they introduce a risk assessment model with the help of the fuzzy rule and classification algorithm. This model had six basic categories such as URL& Domain Identity, Security & Encryption, Source Code and so on. In [2] its learning how to detect the phishing emails. This was helpful to find the emails which are sent by the attackers. When a user is sent a mail the first thing he should check the richness of the vocabulary, the structure of the page and so on. Email filtering does all these features. The method followed by the author is PILFER. The paper states the learning of overall approach, features and empirical evaluations. Based on this we can find the differences in IP-based URL’s, Age of the linked-to domain names, Number of links, HTML emails, Number of dots, Number of domains, contains JavaScript, spam-filter output and etc. The conclusion is based on these they are possible to detect the phishing emails by finding the high accuracy by using spam and specialized filters. These are helpful in learning how to detect the phishing emails.

In [3] the paper explains about the behavioral response to phishing risk. This study of paper reports that the pilot survey of 232 computer users to reveal the legitimate emails. It provides the deeper understanding of the web environment. The main behavior of the phishing sites can be found out by the long URL’s and it respond to the user through a intermediate page as linking the main page to the sub pages. The colors will be not standard colors in which its default. The validity of the phishing website may be a short period so we can find the years of the website started and its usage. Many of the phishing websites use PayPal and eBay for the transactions and shopping’s. The table listed below is the features of five emails and corresponding web sites from email and web role play.

<table>
<thead>
<tr>
<th>Email</th>
<th>Legitimacy</th>
<th>Relevant features of email and sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognix</td>
<td>real</td>
<td>Regarding work details Link in email: <a href="http://www.cognix.com">www.cognix.com</a> URL in status bar: <a href="http://www.cognix.com">http://www.cognix.com</a></td>
</tr>
<tr>
<td>NASA</td>
<td>real</td>
<td>Sender is known person Addressed to user Link in email: “this” URL: antwp.gsfc.nasa.gov/apod/astropix.html</td>
</tr>
<tr>
<td>eBay</td>
<td>real</td>
<td>Registered name “Pat Jones” displayed Link in email: “PAY [Click to confirm...]” URL:<a href="http://payments.ebay.com/ws/">http://payments.ebay.com/ws/</a> eBayISAPI.dll?item=6600378513</td>
</tr>
<tr>
<td>PayPal</td>
<td>phishing</td>
<td>Urgent request Lock image in body of webpage Link: “Click here to activate your account” URL:<a href="http://payaccount.me.uk/cgi-bin/webscr.htm?cmd=">http://payaccount.me.uk/cgi-bin/webscr.htm?cmd=</a> login-run</td>
</tr>
<tr>
<td>Laptop</td>
<td>Spear phishing</td>
<td>Generic message about eBay item Link: <a href="http://www.set-ltd.net">www.set-ltd.net</a> URL: <a href="http://www.set-ltd.net">www.set-ltd.net</a></td>
</tr>
</tbody>
</table>

In [4] the paper describes about the social phishing. This social phishing is fully based on the social network database. It has the public data where it finds the social networks and can collect all the emails to check and to authenticate the web logs. Thus it says it is very easy and very effectively done. It takes the note of each authentication part and gives the report of it in graphs. As countermeasures to combat phishing improves, its said that the phishers in their attacks
to appear more convincing. As the social phishing attacks underscore the dangers of the public it takes all the personal information’s and need to adequate countermeasures.

In [5] the antiphishing strategy has been combined with visual similarity to find the phishing sites and create a report based on it. The algorithm proposed in this method is to get all the sufficient blocks from the given URL and then finding the similarities between the protected website and the given website address. Finally it proposed using visual similarity assessment and finds the report by testing websites like eBay, PayPal and so on. In many proposed algorithms they fail to find the phishing websites, but they tried it to a mark upto 50% still they can’t succeed. Many proposed algorithms used association rule and fuzzy algorithms. They are the new mining algorithms in this fast moving world. The association rules for mining the large database is very useful for business websites and marketing.

**Architectural Diagram**

**Proposed Work**

The module description of the phishing tool is classified based on visualization. The modules are based on the data mining algorithms in which the tool can be faster and protective. Each module has different similarities and it can...
be found out when the modules are being match with each other.

**Obtain URL**

The URL is the website address which the user enters in the address bar. When the mobile handset is connected to internet, it extracts the URL from address bar and gives that URL to next modules.

**Web page Extractor**

In this module it extracts the web page with the help of given URL. The URL opens the corresponding website and enters into it. It consists of all styles, images, fonts and the corresponding details of that website. This web page is used for the visualization and divides the web page into categories and sub categories.

**Web page Visualization**

![Block Diagram of Webpage Visualization](image)

Visualization technique is the most powerful devices for identifying patterns hidden in data. We extract three main extraction modules which is necessary to find the phishing sites by visualization. The classifications of visualization are

1. Text extraction
2. Color extraction
3. Image extraction

The text extraction will extract all the text contained in the webpage blocks. The text extraction can be further classified into fonts, border text, style and alignment. We get respected values for this classification module. The color extraction will extract all the colors contained in the webpage.

The color extraction can also be further classified into background colour, foreground colour and content color. We get the respected values for this module.

The image extraction will extract the entire image contained in the webpage. The image extraction can be classified as image color, image type and image size. We get the respected valued for this module.

**Similarities**

We collect the module values for both the obtained website and the protected website from the above modules. Now compare both the values and find the similarities between the websites.

**Report**

With the help of all the similarities we generate a report and if the report is negative we alert a sound alarm and if the report is positive we give the web page to the user.

**4. Conclusion**

The main aim of this paper is to make user safe and a secure access to the mobile internet. These kind of tools had been implemented in early days by WAP servers, but now a day’s our technology had been improved. Thus we use wifi, it’s now a fast internet connection and can access in all places without any device. So its easy for the mobile internet and the tool and secure user in all ways and will be helpful also. In future ways can test this phishi websites through many ways according to our technology development. The future works can be to fix the anti virus also into the tool in which the user will be comfort to access all pages and be secure.

**References**


[5] Phishing Attack Detection by Using a Reputable Search Engine by Robert Ma, Electrical and Computer Engineering Department University of Toronto


