Malaysian Journal of Computing, 5(1): 362-373, 2020

eISSN: 2600-8238 online

ANTI-PHISHING WITH GOOGLE EXTENSION 3H1M USING BLACKLIST ALGORITHM

Muhammad Hafiz Ramli[1], Ahmad Izatul Hisham Faugi[1], Nur Mawaddah Mohd Faizal[1], NorHaziqah Mohd Khadri[1] and Jasni Mohamad Zain[2]

[2]Advanced Analytics and Engineering Center
[1]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA
40450 Shah Alam, Selangor, Malaysia
2017734685@isiswa.uitm.edu.my, 2017346465@isiswa.uitm.edu.my,
2019754965@isiswa.uitm.edu.my, 2019577059@isiswa.uitm.edu.my,
jasni@tmsk.uitm.edu.my

ABSTRACT

The advancement of technologies has rapidly risen for the past few years, which brings a lot of benefits especially to the users. However, the developments have also contributed many security issues, where security attacks have also become more advanced. Phishing is one of the security attacks that describe spoof emails or websites to trick users into exposing their personal or any confidential information. Hence, this project develops a solution, 3H1M extension, to help in mitigating the phishing issues when browsing the Internet. This extension will compare the visited website URL with the blacklisted website lists in the database to identify the validity of the website. If the user is visiting a malicious website, a pop-up message will be displayed on the screen and the user will be directed to the search engine page. A comparison is made between a web browser implemented with and without the 3H1M extension. It can be observed that the extension is able to help the users to distinguish between the real and malicious websites.

Keywords*: Anti-phishing, 3H1M, Blacklist Algorithm and Google Extension.*

Received for review: 27-06-2019; Published: 06-03-2020

1. Introduction

Phishing is one of the top cybercrimes around the globe that affects customers and companies. It's the Internet's most popular scams. The term "phishing" occurred in the mid-1990s, when hackers started using fraudulent emails to "fish" unsuspecting users for information. Since these early hackers have often been referred to as "phreaks," the term has become known as "phishing", with a "ph." Phishing is an example of social engineering in which an attacker attempts by impersonating a trustworthy third party to fraudulently acquire sensitive information from a victim (Issac, Chiong & Jacob, 2014).

Phishing is when cybercriminals send malicious by use email, mobile, or social channels designed to trick people into falling for a scam. Phishing emails are trying to lure victim in and get victim to take the bait. The intent is often to get users to reveal financial information, system credentials, or other sensitive data. The techniques of social engineering include forgery, misdirection and lying, all of which can contribute to phishing attacks. Phishing use social engineering at a basic level to encourage victims to act without thinking

through things. According to Khonji, Iraqi & Jones (2013), a system may be technically sufficiently secure against password theft, however unaware end users may leak their passwords if they are asked by an attacker to update their passwords via a given Hypertext Transfer Protocol (HTTP) link, which ultimately threatens the overall system security.

Recently, the phishing concept has expanded to include a wider variety of electronic financial crimes. In relation to the extensive use of these false email messages and websites to attract victim to disclose their private data. These programs, once installed on a victim's desktop, it uses a variety of methods to spy on communications with web sites and gather account data. This technique varies from the general technical subterfuge connected with phishing scams and can also be included in the spyware definition. In conclusion, for criminals phishing is an extremely profitable activity.

1.1 How the Attacker Gain Information?

Phishing is a cyberattack used as a weapon by disguised email (Alam & El-Khatib, 2016). The goal is to trick the email recipient into believing that the message, for example, is something they want or need from their bank, or a note from someone in their company, and to click a link or download an attachment. These messages look authentic and try to get victims to disclose their personal information. The attackers masquerade as some kind of trusted entity, often a real or plausibly real person, or a business with which the victim could do business.

Phishers frequently use true business logos and duplicate lawful e-mail messages to replace the connections with those that direct the victim to fraudulent site (Routhu Srinivasa Rao, 2017). It uses spoofed or false e-mail addresses in the fields of the message "From:" and "Reply-to," and blur connections to create them appear legitimate. But recreating an official message's presence is just part of the method.

Most phishing messages offer a reason for instant action to the victim, causing the victim to act first and think subsequently (Jagatic *et al.*, 2005). If the victim does not respond quickly, messages often threaten the victim with account cancellation. Some thank the victim that they have never created a buy. Because the victim doesn't want to lose money that they didn't really lose, they follow the link of the message and ends up offering the phishers precisely the sort of information that they were scared they had first (Saiful *et al.*, 2017).

Moreover, many individuals believe that automatic procedures are safe from human error. This is why many emails argue that a computerized inspection or other automated method found that there is something wrong with the account of the victim. The person is more probable to believe that someone has tried to break into his account than to think that an error has been created by the machine doing the inspection.

Address spoofing is the most common trick. Many email programs enable customers to access the "From" and "Reply-to" areas of their required data. While this makes it convenient for individuals using various email addresses, it makes it simple for phishers to generate emails that sound like they originated from a lawful origin. Some email servers also enable machines to connect without using a password to the simple mail transfer protocol (SMTP) gateway. This enables phishers to immediately connect to the e-mail server and deliver emails to victims.

Phishers can use proxy to track operations of victims and track operation of site victims. They can also benefit from poor security on the web site of a company and insert malicious software into particular websites (Banday & Qadri, 2007). Phishers using these techniques do not have to hide their links because the victim is on a lawful website when the victim data is stolen.

1.2 Effect Attack of Phishing

It is now commonly widely known the dangers of being phished, but the extent of the harm is often overlooked. Successful phishing includes the scammer having unauthorized access to the private data of an organization and then using it for personal profit. Phishing's effect is far more

insidious than a privacy intrusion. Phishing is used by social engineering to compromise software safety. It can be used for stealing data, disrupting software activities, stealing cash, ruining reputations, destroying significant data or feeding an attacker's ego.

The most evident damage induced to legitimate companies and organisations is the financial damage induced by phishing. Phishing can cause the business to lose huge sums of money, sometimes even millions of dollars (Vayansky & Kumar, 2018). In 2003, it was projected that phishing caused US banks and credit card companies around $1.2 billion in immediate financial losses. Indirect business losses are much greater because they include customer service expenses, account replacement costs, and greater online service expenses owing to a reduction in use caused by absence of trust in data security. This absence of trust in the organizations ' online services is understandable.

In addition to the financial loss incurred through phishing, companies may also experience reputational harm in relation to the economic loss caused through phishing. They may be seen as incompetent and untrustworthy if a business has fallen victim to a scam. If a business is a third-party supplier, an event of violation may cause their customers to terminate their agreements instantly. It requires time and commitment to build up a brand reputation, all of which can be almost instantly wiped out if a phisher attack. The harm to the reputation of a company arises not only from being phished, but also from being spoofed. If a hacker gets the customer list of the organization and sends them spoof emails, the reputation of the organization will take a shot (Ragucci & Robila, 2006).

Recently, it has been reported that 46 million of personal data belongings to Malaysia data plan users were leaked (ictmalaysiablog, 2019). So, when it goes to individuals and community, the internet is really damaged by phishing scams. In their junk mail directory or advertisements on Facebook and twitter they can always discover some scams that attempt to connect to a false website. With the rapidly increasing phishing technology and increasing social networking, when sharing personal information online, individuals are becoming more at risk. This could result in confidential details losses for users and even prevent them from accessing their own accounts.

1.3 Prevention Techniques from Phishing Attack

Phishing attack are one of the most prevalent safety problems facing both people and businesses in maintaining safe data. Whether accessing passwords, credit cards, or other delicate information, hackers use email, social media, phone calls, and any type of interaction to steal precious information. Of course, businesses are a worthwhile target. Many companies, such as Google and IBM, have involved in the campaign of against phishing. For example, Google has created a list of phishing sites to avoid web users from falling into the scheme (Cui *et al.*, 2017). There are many techniques of preventing the phishing attacks.

Phishing can be prevented by blacklisting or preventing phishing sites or wiping out phishing emails before it hits the user. The first technique is to look at the URLs and the locations they pretend to be manually or automatically using machine teaching. The second technique can be considered more efficient because if it is carried out effectively it will prevent the user from ever being subjected to the phishing sites link. There are many effective spam filters been used by email servers, but few phishing filters due to its more complicated nature.

Some recognized patterns are available that can be noted to avoid phishing. These include the suspect grammar and punctuation (Qin *et al.*, 2011). To generate messages with well-tested material, topic line and call-to-action, professional copywriters go to excellent lengths. Any email containing bad grammar, punctuation or displaying an illogical flow of content is probable to be written by inexperienced scammers and fraudulent.

There are a number of other best methods that users can use to avoid phishing irrespective. These include paying attention to shorter links. Short links do not indicate the true name of a website and can therefore be used more readily used to trick recipient by clicking.

Hackers can use reduced links to redirect user and catch delicate data to false websites look alike. Place cursor on the abbreviated line before clicking on it to see the target location.

Other pattern that can be recognize is hackers send an email about some pending date. For instance, a hacker may write a renewal letter about an expiring insurance policy, or a restricted discount on some deal that may be of concern to the target. Emails that typically direct people to data collection sites that wind up stealing valuable private or financial information.

The best way to solve phishing is design strict user education programs that not only assist users recognize fraudulent email, but also provide particular advice on how to manage suspicious communication (Rajab, 2018). In the parts below, it will concentrate on safe handling of email that violate the software layer's safety. This involves rules to identify suspicious messages based on frequently known historical pattern, as well as a range of best methods to prevent victimizing email that fail to get through.

2. 3H1M Anti-Phishing Google Extension

3H1M was an acronym from team member name which is Hafiz, Hisham, Haziqah and Mawar. 3H1M was Google Extension develop for Google Chrome, helping to detect phishing website using blacklist algorithm.

2.1 Introduction of 3H1M Google Extension

After understanding the concept and overall structure of a phishing attack, an anti-phishing technique can be taken into consideration to prevent users from being a victim. An anti-phishing system has been instigated to mitigate this issue by exploiting the function of the google extension since attackers mainly will create a phishing website by imitating the legitimate page content. The phishing URLs created are mostly similar to the valid URLs, where mostly users will become unaware of the ploys and causes them to accidently browse the malicious website. For example, the domain name www.maybank2u2.com might seems legit, when in fact the real URL for the website is actually www.maybank2u.com. Hence, this system can help users from visiting any malicious web pages unconsciously and falling into the trap.

2.2 System Objectives of 3H1M Google Extension

There are several objectives that are considered during the development of this system that help in describing the desired results, which is to prevent the phishing attack when browsing the websites. The objectives of this project are as stated below:

a) To develop an anti-phishing system in a form of a Google extension, named as 3H1M extension.
b) To detect any fraud web pages by looking into the database containing the blacklisted pages.
c) To notify and alert users upon accessing a blacklisted website.

2.3 System Implementation of 3H1M Google Extension

This section will explain on the components used in this system. It is important to understand the components included for this project and its functions for the phishing websites detection. Next, this section also includes the explanation about the flow of the system in detailed forms to show how it works and to understand the functions of the 3H1M extension.

2.3.1 System Components

Before going through the explanation on how the system works, it is important to know the components that have been taken into consideration to develop this anti-phishing mechanism. Table 1 shows the system components and its functions respectively.

Table 1. System Components

Components	**Details**
3H1M extension	A plug-in feature installed on Google Chrome that allows the detection of any phishing websites.
Databases	Consist of the blacklisted URLs of malicious websites.
Web services	A software service allowing the interoperability between applications on World Wide Web.

2.3.2 How The 3H1M Google Extension Work

The development of the system focuses on the phishing attack on web browsers, hence, google extension is one of the best methods for the implementation. Figure 1 shows a web browser, Google Chrome, implemented with the extension.

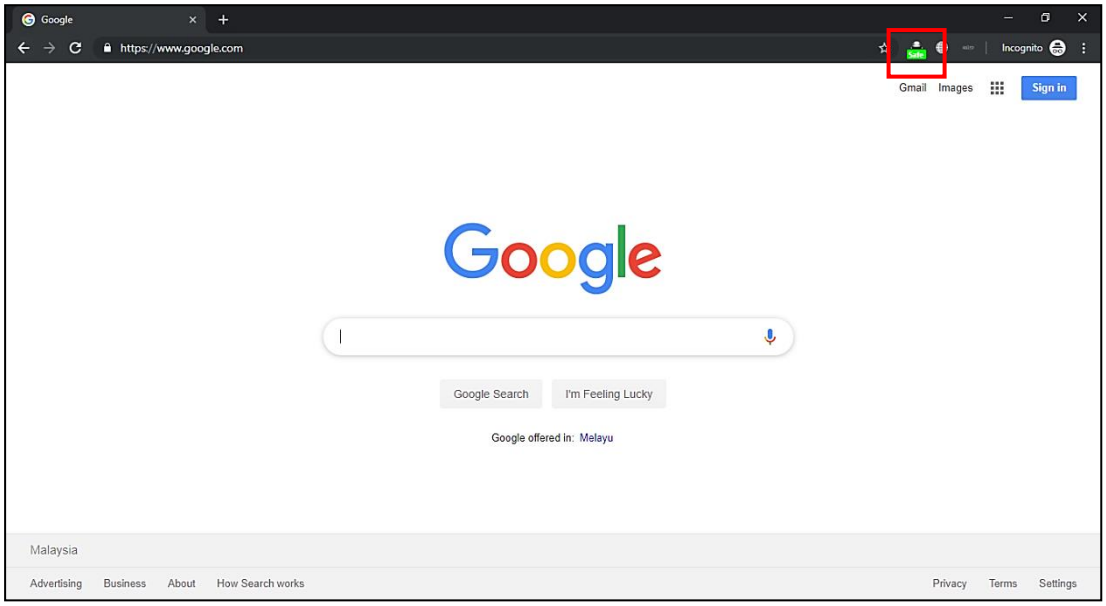

Figure 1. Implementation of the 3H1M Extension

As can be seen from the above figure, the 3H1M extension is located at the most right of the toolbar with a green box indicating that the current web is safe for browsing and it is not listed as a phishing website. Figure 2 shows the close-up of the extension.

Figure 2. 3H1M Extension

When a user inserts a website URL at the search engine, the extension will look into the databases to identify whether the website searched is legit or malicious. If the URL of the website matches with the one in the database, then the website is malicious. Otherwise, the website is safe to be visited. For example, a user is trying to visit a website by typing its URL at the search engine, which is amanz.my as shown in Figure 3.

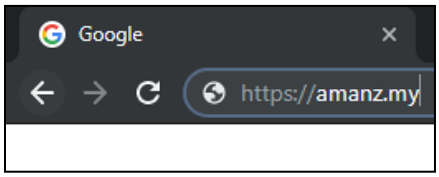

Figure 3. Visiting a Website

When the website is being accessed after clicking on the URL, the extension will compare the current URL with its database lists for blacklisted URLs to check the validity of the website. Based on the example, the 3H1M extension has found a similar URL in its database, which indicates that the user is accessing a phishing website as shown in Figure 4 Stimulate by adding URL website into blacklisted database, used https://amanz.my/ as example purpose.

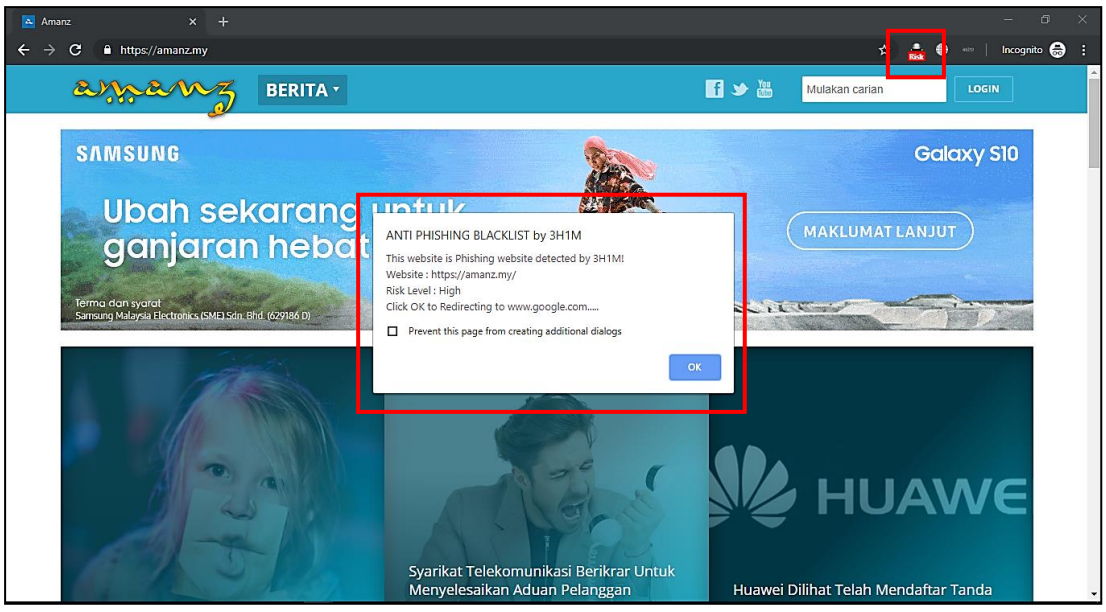

Figure 4. Malicious Website

Based on the above figure, when a user accidentally accessed a malicious website, a pop-up message will be displayed on the screen saying "This website is phishing website detected by 3H1M" to notify the user about the phishing page. The message will also instruct the user to click the OK button to be directed back to the Google homepage. A warning page will be displayed on the screen once user has clicked on the OK button as shown in Figure 5.

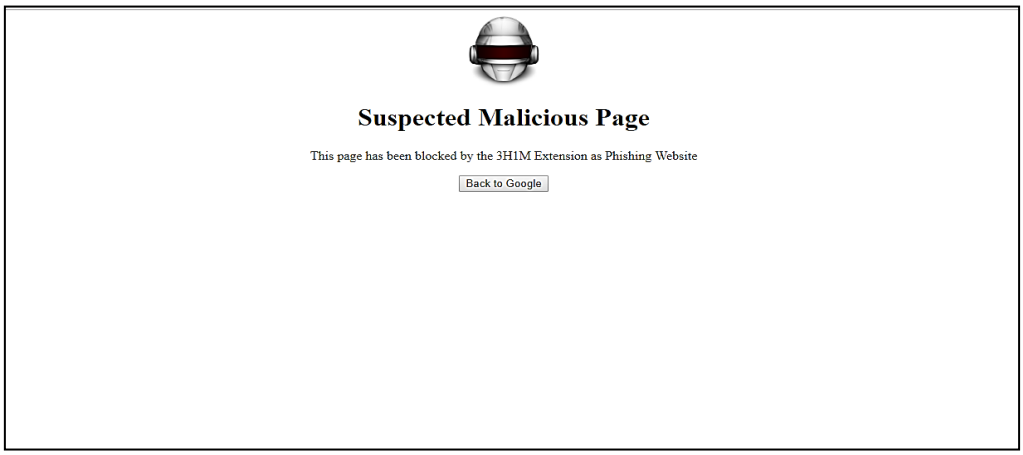

Figure 5. Warning Website

Based on the above figure, when a user accidentally accessed a malicious website, a pop-up message will be displayed on the screen saying "This website is phishing website detected by 3H1M" to notify the user about the phishing page. The message will also instruct the user to click the OK button to be directed back to the Google homepage. A warning page will be displayed on the screen once user has clicked on the OK button as shown in Figure 5. Figure 6 shown 3H1M Google extension system architecture where shows five main steps and process.

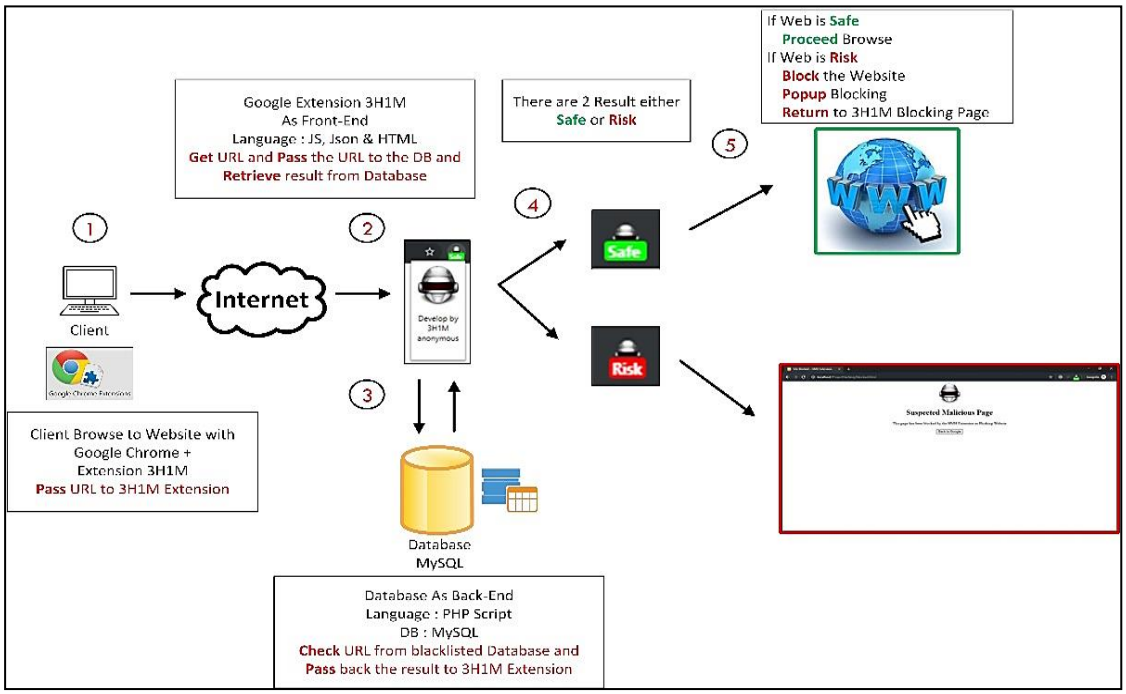

Figure 6. 3H1M Google Extension System Architecture

2.3.3 System Coding

The 3H1M extension uses JSON language, which is a text written with JavaScript object notation. The code is developed with a name Manifest.json as shown below.

```
{
"name": "ANTI PHISHING BLACKLIST by 3H1M",
"description": "Anti Phishing Blacklist Method",
"version": "3.0",
"browser_action": {"default_icon": "icon2.png",
"default_popup": "popup.html"
}
```

Based on the coding above, the name, description, and version of the plugin are assigned for the users' guide. The browser_action is included to display the icon bar and the file execution to plugin bar when it is being clicked. The next part of the coding is as shown next:

a) To permit the tabs and all HTTP/HTTPS web requests to the server to access the extension API `"permissions": [ "tabs","http://localhost/*", "http://*/*", https://*/*]`.
b) To allow the API to gain access to the chrome background `"background": {"scripts": ["background.js"], "persistent": false }, "manifest_version":2`.
c) To display the extension icon in the list of the extensions at the `toolbar "icons": {"128": "icon2.png"}`.
d) Content_scripts syntax matches all the HTTP/HTTPS web requests `"content_scripts":[{"matches":["http://*/*","https://*/*" ],"js":["popup.js","jquery.js"],"run_at":"document_end"}] }`.

Next, the implementation code for the website access using js file extension, which is a text file that contains JavaScript code to execute JavaSript instructions in websites.

Popup.js

```
chrome.runtime.sendMessage(document.getElementsByTagName('ti
tle')[0].innerText).
```

All the information about a website, such as file name, can be collected with this code when accessing a website. The information gathered will then be sent to the receiver API code. The next coding is implemented to identify the validity of the websites visited which is also written using the js file extension.

Background.js

```
chrome.runtime.onMessage.addListener(function(response,
sender,sendResponse){
 var web = sender.tab.url;
 var xhr = new XMLHttpRequest();
      xhr.open("GET",
 "http://localhost/ProjectHacking/trace.php?webs="+web,
 false);
      xhr.send();
```

```
        var result = xhr.responseText;

        if(result == "phishing"){
              alert("This website is Phishing website detected
        by 3H1M!");
              chrome.tabs.update({ url:
        "http://localhost/ProjectHacking/blocked.html" });
        }
        else{
              //alert("This is not phishing website");
        }
  });
```

All the information sent from the sendMessage API code will be received with onMessage API code. The URL of the website can be retrieved for comparison with the implementation of send.tab.url. XMLHttpRequest is important to establish the connection to the local or live server. The server will reply with a message to indicate validity of the websites based on the information retrieved. If the server replies with a message “phishing”, the access to the website will be blocked and user will be redirected to a non-phishing website through a pop-up message. However, to allow the comparison of the URLs, a connection to the database need to be established through the next implementation code.

Trace.php

```
<?php
include "db.php";
$web = mysqli_real_escape_string($con,$_GET['webs']);
$sel_web = "select * from website where websitename='$web'";
$run_web = mysqli_query($con, $sel_web);
$check_web = mysqli_num_rows($run_web);
   if($check_web==1)
   {
        echo "phishing";
   }
   else
   {
        echo "notphishing";
   }
?>
```

This file will receive a website link and it will find and query to the table(website)/column(websitename). If the website is listed on the table, it will echo message “phishing”, otherwise, it will echo message “notphishing”. The connection of the database and blacklistwebsite can be established through the following implementation code.

Db.php

```
<?php
header("Access-Control-Allow-Origin: *");
$con =
   mysqli_connect("localhost","root","","blacklistwebsite")
   or die ("could not connect database");
?> //database connection to blacklistwebsite
```

A complete MHM extension can be created through all the implementation codes which have been discussed in this section from the placement on the Google Chrome toolbar to the list of the blacklisted websites.

3. Simulation of 3H1M Anti-Phishing Google Extension

A simulation has been conducted to verify the functionalities of the solution proposed. The simulation involved the testing using a real phishing website. A comparison between a browser with and without the extension was analysed.

Website Link (Phishing Website):
http://tasktg.com/wpcontent/plugins/ubh/agrement/39cc04fcda9d7d7b638969a90fc11563/customer_center/customer-IDPP00C731/myaccount/signin/

Scenario 1: Browser implemented without Google Extension 3H1M

Figure 7 displays the phishing page which is similar to the legitimate PayPal web page. It shows that users are unable to detect whether it is a malicious or the original website due to the similarities between the two pages. This indicates that, without the use of Google Extension 3H1M, users are prone to the phishing attack and can be deceived by this act.

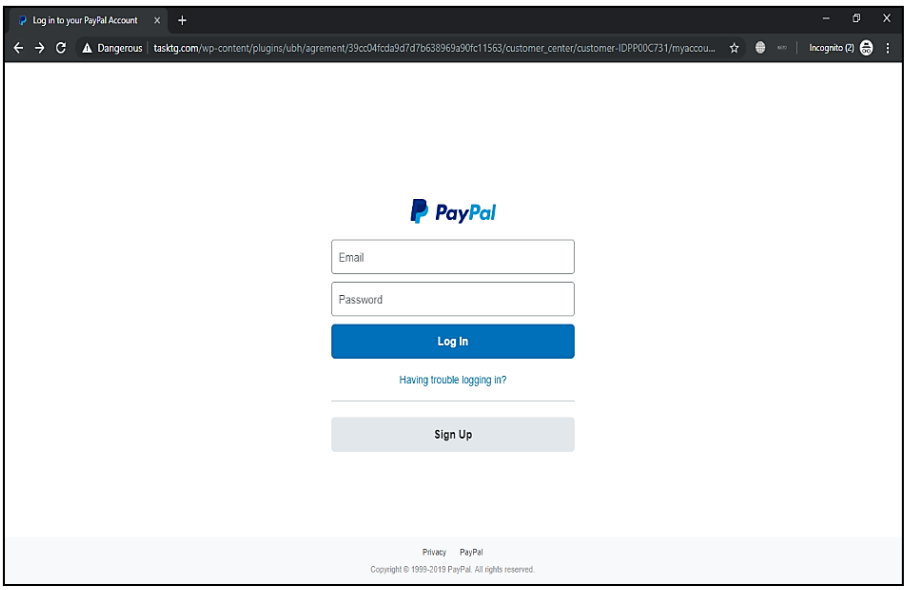

Figure 7. Phishing Page without using Google Extension 3H1M

Scenario 2: Browser implemented with Google Extension 3H1M

This scenario involves the implementation of the solution, which is Google Extension 3H1M. Figure 8 shows that when a user browses the phishing PayPal page, the extension helps preventing the phishing attack by giving alert through the pop-up message. Hence, users are able to identify the validity of the web pages that they are currently browsing.

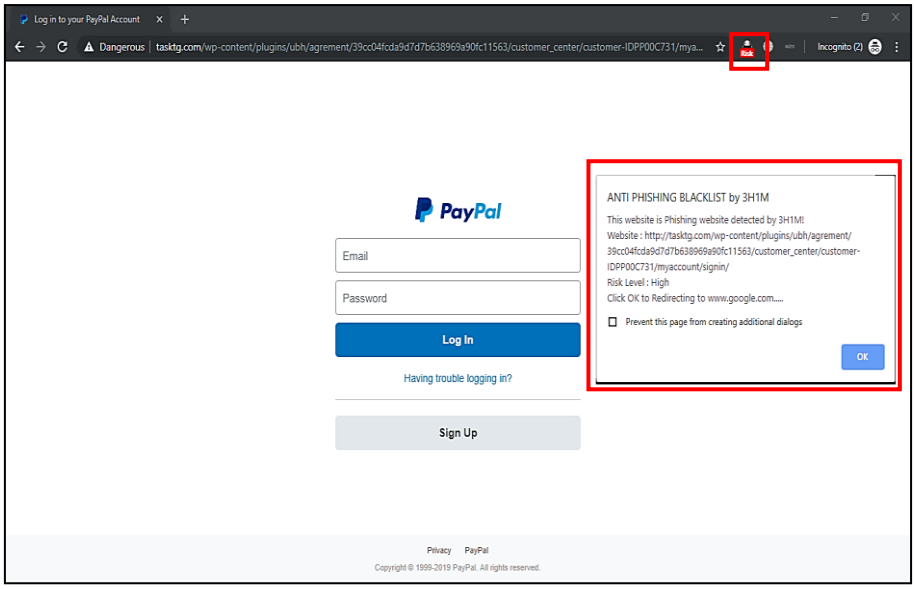

Figure 8. Phishing Page with using Google Extension 3H1M

Based on both scenarios in this simulation, the probability of the users to fall for a phishing attack is higher when the browser is not implemented with the Google Extension 3H1M. Hence, it can be concluded that the Google Extension 3H1M is able to identify any malicious web pages and help the users from being a phishing victim.

4. Conclusion

This project helps in mitigating the issues of accessing phishing websites through the use of google extension. It shows that by using the existing lists of blacklisted websites that are stored in the database, any suspicious URLs that refer to progressing phishing threats in real time can be discovered. The technique that has been selected in this project is list-based detection where an extension is created for an anti-phishing detection to be used by the end users. A database is used to store the phishing websites which consumes fewer computational resources to filter the malicious accessing. However, the dynamic features of the changing web require the database lists to be updated frequently for reliable blocking performance. The browser can detect whether the URLs that are being accessed are legitimate or fraud by referring to the list. A popup message will be displayed to alert the user about the phishing website. Experimental results show that this blacklist anti-phishing extension method is an effective and efficient approach to prevent the users from being a phishing victim when browsing the Internet.

References

Alam, S., & El-Khatib, K. (2016). Phishing Susceptibility Detection through Social Media Analytics. Association for Computing Machinery, (July), 61–64.

Banday, M. T., & Qadri, J. a. (2007). Phishing – A Growing Threat to E-Commerce. The Business Review, 12(2), 76–83.

Cui, Q., Jourdan, G., Bochmann, G. V, Couturier, R., & Onut, I. (2017). Tracking Phishing Attacks Over Time. In International World Wide Web Conference Committee (pp. 667–676).

Khonji, M., Iraqi, Y., and Jones, A. (2013). Phishing Detection: A Literature Survey, in IEEE Communications Surveys & Tutorials, vol. 15, no. 4, pp. 2091-2121, Fourth Quarter 2013.

ictmalaysiablog. (2019). Serious Cyberattack of Phishing in Malaysia. Retrieved from https://ictmalaysiablog.wordpress.com/2017/12/13/serious-cyber-attack-of-phishing-in-malaysia/

Issac, B., Chiong, R., & Jacob, S. M. (2014). Analysis of Phishing Attacks and Countermeasures. Proceedings of the 6th International Business Information Management Association (IBIMA) Conference, 339–346.

Jagatic, T., Johnson, N., Jakobsson, M., & Menczer, F. (2005). Social Phishing. Association for Computing Machinery, 2005, 1–10.

Ragucci, J. W., & Robila, S. A. (2006). Societal aspects of phishing. International Symposium on Technology and Society, Proceedings, (June). https://doi.org/10.1109/ISTAS.2006.4375893

Rajab, M. (2018). An Anti-Phishing Method based on Feature Analysis. Association for Computing Machinery.

Routhu Srinivasa Rao, A. R. P. (2017). Detecting Phishing Websites using Automation of Human Behavior, 33–42.

Saiful Azad, Musfiq Rahman, M. S. A. Noman Ranak, B. M. F. Kamal Ruhee, N. Nourin Nisa, Nazrul Kabir, Arafatur Rahman, Jasni Mohamad Zain (2017) VAP code: A secure graphical password for smart devices, Computers & Electrical Engineering, 59, pages 99-109.

Qin H., Ma X., Herawan T., Zain J.M. (2011). An Adjustable Approach to Interval-Valued Intuitionistic Fuzzy Soft Sets Based Decision Making. In: Nguyen N.T., Kim CG., Janiak A. (eds) Intelligent Information and Database Systems. ACIIDS 2011. Lecture Notes in Computer Science, vol 6592. Springer, Berlin, Heidelberg

Vayansky, I., & Kumar, S. (2018). Phishing – challenges and solutions. Computer Fraud and Security, 2018(1), 15–20. https://doi.org/10.1016/S1361-3723(18)30007-1