The Fuzzy Experiment Approach for Detection and Prevention of Phishing attacks in online Domain

Article · December 2020

DOI: 10.36349/easjecs.2020.v03i10.001

4 authors:

Sezuo BASHIR Tenuche
Kogi State University
6 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Agbata Celestine
Kogi State University
14 PUBLICATIONS **1** CITATION
SEE PROFILE

Emmanuel Ogala
Federal University of Agriculture
14 PUBLICATIONS **4** CITATIONS
SEE PROFILE

William Obeng-Denteh
Kwame Nkrumah University Of Science and Technology
75 PUBLICATIONS **133** CITATIONS
SEE PROFILE

East African Scholars Journal of Engineering and Computer Sciences
Abbreviated Key Title: East African Scholars J Eng Comput Sci
ISSN: 2617-4480 (Print) & ISSN: 2663-0346 (Online)
Published By East African Scholars Publisher, Kenya
Volume-3 | Issue-10 | December-2020 | DOI: 10.36349/easjecs.2020.v03i10.001

OPEN ACCESS

Research Article

The Fuzzy Experiment Approach for Detection and Prevention of Phishing attacks in online Domain

Bashir[1], Tenuche[2*], Agbata, B.C[2], Emmanuel Ogala[3], William Obeng-Denteh[4]

[1]Department of Mathematical Sciences, Computer Option, Kogi State University, Anyigba, Nigeria
[2]Department of Mathematics, University of Nigeria, Nsukka Nigeria
[3]Department of Maths/Stat/Computer Science, Federal University of Agriculture, Makurdi Nigeria
[4]Department of Mathematics, College of Science, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

Article History
Received: 21.11.2020
Accepted: 03.12.2020
Published: 11.12.2020

Journal homepage:
https://www.easpublisher.com

Quick Response Code

Abstract: Phishing is an aspect of identity theft that uses engineering and social paradigms to steal personal information from unsuspecting users for their selfish gain. Most attacks are in the form of luring the user into clicking a link that directs the user to a rogue page. The major target of phishing attacks are online customers of e-banking and payment service providers, and these groups suffer huge financial loss. Phishing is not all about sending spoofed mails to users as most people assume, rather it is a multifaceted techno-social issue for which there is no particular solution to end its reign. This has given rise to a series of research in this field as scholars are working towards creating more efficient anti-phish solutions by quantifying risks and degree of vulnerability of users. Most approaches to combat phishing are not able to make dynamic decisions to determine the risk rate of the website and this allows for a large number of false positive. The use of blacklists and whitelists has their limitations due to poor scalability and time constraint. Anti-phishing solutions are methods put in place to protect internet users from attacks aimed at defrauding them of their finances. Browser plug-ins have been the most recent methods adopted, though a lot of questions have been raised to question the effectiveness of these plug-ins. In this research we aim at developing an intelligent anti-phishing plug-in for e-banking, capable of detecting phishing attacks based on existing knowledge about features and patterns of phishing websites. The proposed system is developed to protect users from deceptive tricks used by phishers by giving them the ability to identify phishy or fraudulent websites. The proposed model uses fuzzy logic to define rules and assign linguistic indicators in the form of if-then rules to each phishing criteria. The approach adopted is a combination of fuzzy reasoning in quantifying dynamic and unclear phishing characteristics, with the proficiency to categorize the phishing rules.

Keywords: Internet, e-banking, phishing attack, phishing detection, fuzzy based model, website detection, software, system design.

INTRODUCTION

The Internet and Internet Banking

The internet is a network of networks, all freely exchanging information. The networks range from large corporate networks to just normal or ordinary ones [1]. Advances in today's world of technology are eroding the luxury of privacy which we have somehow learnt to take for granted... daily activities like credit cards, cell phones, and key cards allow daily purchase and movement to be tracked. Per Hammoud *et al.* [2], these advances in technology have succeeded in making various aspects of life easier for the societies of today [3]. More importantly, it has become a fundamental component in improving the quality of services in general and E-Banking services in particular [4]. E-Banking service is said to rely on the exchange of information between customers and providers using technological methods devoid of face-to-face interaction [5]. Shah and Clarke [6] explain CDed that internet banking can mean the provision of information about a bank and its services via a home page on the World Wide Web or as a mobile phone application. Internet banking is the easiest way transactions can be carried out in the hectic schedule of individual users. Per Ming-Chi Lee [17] e-banking has emerged as one of the most profitable e-commerce applications developed; it is widely used and enjoyed by the masses because the benefits it offers are numerous. The online banking facility offered by banks to all customers is an advantage; it has opened doors to all customers to operate beyond boundaries [7]. It is obvious that while financial institutions and their customers are gaining immensely from the convenience and automation of online banking, there is a new risk to

***Corresponding Author:** Tenuche

be faced which is associated with delivering services over a public network. Criminals are aware of the security risks associated with this situation; they are taking advantage of this situation by attacking security weaknesses, posing as valid customers and stealing money from online accounts [8]. Phishing is one of the methods used to accomplish this.

Phishing and Trust in E-Banking

Service quality remains a dominant factor in keeping a competitive edge and sustaining satisfying relationships with customers [9]. Service quality is one of the aspects contributing to customer satisfaction judgement. However, Phishing websites are extremely dangerous to internet business, as users tend to lose their trust in internet transactions out of fear of becoming victims of fraud. It's only natural for users to believe that using online banking increases the chances of becoming victims of phishing websites and identity theft despite the protection provided by online transactions as compared to paper and mail based systems. The worst effect is the trust crises. The trust the customers have will gradually be eroded without any effective countermeasures to deal with the fraud and in the end all the parties involved in network transactions will be harmed. Trust is an essential determinant of e-banking [10] as it is an essential aspect of understanding interpersonal behaviour in e-banking. Internet technology is very pervasive today, for example, from online social networking to online banking; it has improved human to human interaction [11].

Phishing Attack

Phishing is a form of cybercrime aimed at deceiving users into providing personal and/or financial information or to send money directly to the attacker [12]. A phishing attack is instigated via some form of message such as a link to a deceptive domain name which may seem legitimate initially but is actually controlled by an attacker. The term "Phishing" was first used in 1996 and has since continued to grow and evolve. Phishing is no longer restricted to email as was the case in the past, it may also be carried out through voice-messaging, SMS, instant messaging, social networking sites and even multi-player games [13]. The idea is to deceive the victim into visiting the spoofed site, which appears to be a lot like the original one, hence making the user comfortable enough to enter a username and password or other personal information [12]. Phishing sites are commonly created to acquire personal information such as credit card numbers, personal identification numbers (PINs), social security numbers, banking numbers, passwords etc. or to install some form of malware on the victim's computer. According to the anti-phishing work group (APWG), Phishing is a crime that employs both social engineering and technical subterfuge to steal an individual's personal identity data and financial account credentials. The APWG indicated about 50000 unique reports by the end of 2019, 45072 unique phishing e-mail reports (campaigns) received by APWG from consumers, 341 brands targeted by phishing campaigns, also, Almost three-quarters of all phishing sites now use SSL protection, highest recorded since early 2015, and an indicator that users can't rely on SSL alone to understand whether a site is safe or not. Per Bhadane & Mane [14] Business E-mail Compromise (BEC) reported a loss of 3 Billion USD (Internet Crime Complaint Center (IC3), 2016) and the numbers of attacks are continually increasing (Phishing Activity Trends Report). More than 80% of organizations have faced phishing attacks [12]. Phishing attacks take advantage of recent events such as Equifax hack to hook users into traps Equifax or Equiphish? - Krebs on Security, 2017; Kennedy, 2017. An attack impersonating Google Docs affected almost a million users [15], and the DNC impersonation hack led to leak of political data [16].

Motivation

Although there are several applications in place for phishing website detection, only a few solutions use machine learning techniques in detecting phishing websites. More so, most of the techniques used to combat phishing attacks are impractical or rather inaccurate and suffer from high levels of false positives and misdetection [17]. The motivation behind this study is to be able to determine whether phishing activities are ongoing or not on an e-banking website (FirstBank Nigeria) by creating an effective and resilient model capable of detecting phishing websites, with an aim to prevent innocent users from being deceived or hacked. The methodology implemented in this research is quantitative; it investigates detecting phishing websites intelligently based on neuro-fuzzy techniques. The technique adopted uses fuzzy logic to process the phishing features, for extracting classification rules to be implemented inside the fuzzy inference engine. The fuzzy rules allow for the construction of if-then rules, which creates a bigger picture of the relations between different phishing features and their association with one another, to be used to determine the final phishing website detection rate. As a constant user of e-banking sites and e-commerce, I realized the trouble and hassles phishing sites have posed over the years; moreover I have been a victim of these attacks myself. This pushed me to investigate and do some research towards finding a solution to overcome this issue, especially for some naive banking clients and customers to several e-commerce sites.

RESEARCH AIMS AND OBJECTIVES

The proposed research is aimed at validating the fuzzy experiments proposed in a previous study by Aburrous [18] by building an intelligent plug-in using the proposed specifications i.e. rules and layers to be able to determine how smart the system can be in detecting phishy sites. The following objectives must be achieved to complete this task;

- Explore existing literature from recent articles in this field, understand existing tools and techniques for phishing website detection and protection.
- Fuzzification of values i.e. assigning linguistic descriptors to a range of values
- Defuzzification: converting the input data into crisp data
- To develop an intelligent tool bar using values from fuzzy tests, this will be smart enough to determine if a certain URL is a phishy website or not.

Phishing Detection

Detecting phishing attacks is a difficult problem and is still so far away from being solved. The prevalence of phishing websites and emails confirms the success phishers are having with their attacks. When a website or email matches a known and legitimate website or email, it is very difficult for a user to spot the difference and hence he gets fooled easily. One popular solution used by many is to add additional features into an internet browser to warn users when they wander into a phishing website. This browser security is provided by a mechanism called blacklisting: a process that matches any given URL with a list of URLs belonging to a blacklist. But phishers today have a lot of technical know-how to be able to evade blacklists [19]. Some previously proposed models to detect these attacks are discussed below:

Wu *et al*. [29] proposed a method that requires web page creators to follow some rules to create web pages, like adding sensitive information location attributes to HTML code. However, all web page creators can't be persuaded to follow the same rules. Liu *et al*. [30] compared legitimate and actual phishing web pages to define metrics that can be used to detect phishing pages on visual similarity (i.e. block level similarity, layout similarity and overall similarity). The DOM-based visual similarity [20] of web pages is oriented and the idea behind visual approach to phishing detection was first introduced. With this approach, a phishing web page can be detected and hence reported in an automatic way instead of involving too much human effort. This method firstly decomposes the web pages (in HTML) into block regions that can easily be distinguished. With metrics, the visual similarity can then be evaluated: block level similarity, layout similarity and overall style similarity which are all based on matching the salient block regions. If the visual similarity of a web page is above a predefined threshold, the web page is classified as a phishing page. Joshi *et al*. [21] proposed a mechanism that detects phishing attacks by submitting wrong credentials in login processes. The idea is a novel algorithm aimed at identifying a forged website by submitting random credentials before the actual credentials in a login process of a given website. Alongside this they proposed a mechanism for analysing the responses received from the server against the submissions of all those credentials to determine if the website is legitimate of phishy. The issue with the prototype developed is – it is developed for sites supporting HTTP Digest authentication and accepting user id and password pair as credential. A similar idea used by C. Yue and Wang [22] called the bogus-biter which is a unique client-side anti-phishing tool, transparently feeds a large amount of bogus credentials into a suspected phishing site. It conceals the victim's actual credentials among bogus credentials and also enables a legitimate website to immediately identify stolen credentials. Jain & Gupta [11] proposed a novel approach to protect against phishing attacks using auto-updated white-list of legitimate sites accessed by the individual user. The approach has both fast access time and high detection rate. When users try to open a website which is not available in the white-list, the browser warns users not to disclose their sensitive information. The approach checks the legitimacy of a webpage using hyperlink features. For this, hyperlinks from the source code of a webpage are extracted and apply to the proposed phishing detection algorithm. The experimental results show that the proposed approach is very effective for protecting against phishing attacks as it has 86.02 % true positive rate while less than 1.48 % false negative rate. Moreover, the system is efficient to detect various other types of phishing attacks (i.e., Domain Name System (DNS) poisoning, embedded objects, zero-hour attack). Zhang *et al*. [23] proposed a content-based phishing detection technique called CANTINA, which takes feature set from various fields of a webpage. The proposed technique calculates TF-IDF (term frequency-inverse document frequency) of the content of a website and creates a lexical signature. Then, the top five terms with highest TF-IDF values are submitted to the search engine. The top "n" results are used to check the legitimacy of a website, though the performance of CANTINA is affected by the language used in the website. Xiang *et al*. [24] present CANTINA+, an effective, rich feature-based machine learning approach to detect phishing webpages. The rich features are taken from the various field of a webpage like Document Object Model (DOM) tree and the URL of a website. They filtered the website without login forms in the first step to decrease false positive rate. CANTINA+ achieved a true positive rate of 92 % and a false positive rate of 0.4 %. Reddy *et al*. [35] presented an anti-phishing technique which protects user at client side against phishing attacks. The proposed technique provides facility for the user to select specific image corresponding to every website he/she visits. Next time, when a user visits the same website and if the images do not match, then the system will alert the user. However, maintaining the image database required a lot of memory, and matching the images of suspicious sites with the stored images required a lot of time.

Practical Approach

The visual similarity of a phishing site when compared to a real site is one of the major deceptive concept used by phishers. Bursztein [36] conducted a test using Amazon's mechanical Turks to measure the exact similarity level of some of these sites and it came as a surprise that some of the phishing sites are exactly similar to the original sites and this is one of the issues that tends to deceive several users who still try to be as careful as possible. To be able to identify the factors that actually affect users' judgement about phishing websites and the reason for their vulnerability we decided to interact with some users using questionnaires. The questions asked were aimed at assessing and evaluating the accuracy and precision of phishing sites and understanding why so many users fall for their tricks. Another reason for conducting this survey was to identify what attack strategies are most convenient for attackers to deceive users and why.

A second approach was a supposed mail from the team of Facebook asking the user to follow a certain link to make changes to his account. It seemed official with a logo to deceive the user into actually believing it's true.

Dear valued customer,
You recently changed your Facebook password on the August 10th, 2019. As a precaution, this notification is sent to all email addresses associated with your account. If you did not change your password, your account may have been the victim of a phishing scam. Please follow the link below to regain full control over your account.
Note that as the owner of this account you are required to log into our system securely via the link provided below and adjust as few of your security settings, while you are at that, we believe it is in your best interest to set up a new password and restrict the audience that had initial access to your account.
http://www.fbworldrecovery.uhostall.com/
Please ensure that all fields are completed so we can adjust your settings as soon as possible. We apologize for any inconveniences.

The Facebook team.

Another survey to confirm this was by selecting a group of users and showing them a few websites and asking them to confirm if the websites shown to them were legitimate or fake based on characteristics that have been displayed and shown to them about phishing websites. A total of 60 users were tested. 20 of them were well aware of the characteristics of phishing websites while the other 40 had basically no idea what phishing websites were all about, they were rather naive on the topic of phishing. One point most users still fail to understand is that most phishing e-mail will appear similar in both appearance and structure. Virtually all spam mails sent will have one common feature: a clickable link [37]. Also, Based on previous studies and research conducted by Aburrous [18], 27 phishing characteristic indicators were considered for the research and development of the Antiphishing plug in to detect online phishing activities and protect users from online scams.

Table-1

CRITERIA	PHISHING INDICATORS
URL AND DOMAIN IDENTITY	Using IP address
	Suspicious URL
	Abnormal URL of Anchor
	Abnormal DNS record
	Suspicious URL in address bar
SECURITY AND ENCRYPTION	Using SSL certificate
	Certificate authority
	Abnormal cookie
	Absence of lock key on status bar
SOURCE CODE AND JAVASCRIPT	Redirect pages
	Pharming attack
	on mouse over to hide link
	Server form handler
PAGE STYLE AND CONTENTS	Spelling errors
	Copying websites
	forms with submit buttons
	pop up windows
WEB ADDRESS BAR	Long URL
	Replacing similar characters for URL
	Adding a prefix or suffix to the url
	Confusing users with the @symbol
	Using hexadecimal character codes
SOCIAL HUMAN FACTOR	Emphasis on security
	Public generation salutation
	Buying time to access accounts

The results inferred from the survey still shows that humans are the weakest link in any given attack as most times users are so vulnerable they tend to fall for almost anything. The survey also showed that though most users try to be careful but they tend to feel anything official is definitely real and so they go with it. The results displayed above turned out to be as expected. Users with no prior knowledge of computing and phishing attacks dominated the chart as most of that group gave positive responses by clicking on the link provided in the mail before realizing it was probably an attack.

Fuzzy Based Model for Phishing Website Detection

From research and investigation in phishing attacks several characteristics of phishing websites and emails were discovered. Aburrous *et al*. [18] mentioned 27 main that can be used to differentiate a phishing website from a legitimate one. The following are the characteristics majorly used and considered for this work.

i. IP Address: most attackers tend to conceal the destination of the website by making the URL unclear or ambiguous and the favorite method to accomplish this is to use the IP address of the website instead of using the host name.
ii. Redirect pages: attackers can use programming bugs in a mail message or website to redirect users to other untrusted pages. This is same as the test conducted to deceive users about their Facebook accounts where they are asked to follow a certain link into a phishy site. http://www.fbworldrecovery.uhostall.com/
iii. Suspicious URL: In this case the host name in the URL may contain the legitimate website's URL only as a substring but it won't match its claimed identity.
iv. Abnormal URL of anchor: Aburrous *et al*. [18] explained that a web page is considered suspicious when the domains of most of its URLs are different from the page's domain and the anchors do not link to any page. An example is <a href=http://www.ebay.com/> in the eBay website. Generally webpage name should be short, all letters in lowercase with no space and note that hyphens are used not underscores.
v. Long URL address: Websites with short URL addresses are considered to be more reliable than those with long URL addresses.
vi. URL prefix/suffix: this is a popular method used to deceive users for example www.firstbankonline.com the suffix online is added to deceive the users as it is not on the domain of the site.
vii. SSL certificates: Websites with secure certificates (https ://) are more reliable to deal with than the unsecured website (http ://) as most phishy websites don't register for this third party security feature. Users are advised to check for the padlock usually situated at the bottom of the browser frame.
viii. Using the @ symbol to confuse: Using the @ symbol in a URL is a way of deceiving the browser because all the text before the symbol will be ignored by the browser and only the text after the symbol will be considered.
ix. Using Hexadecimal Character Codes: Fraudsters are sometimes known to hide their URLs by using hexadecimal character codes to represent numbers in the IP address. Each hexadecimal code begins with - %
x. Using Pop-Ups windows: Fraudulent web pages open as pop-ups which redirect the browser window to the real company site [18]. The phishers use pop ups to gain a user's personal information. These pop-ups usually ask the user to update, validate or confirm account information.
xi. Spelling errors: a common feature in phishing mails and websites misspellings. A lot of phishing websites are characterized by grammatical errors, sloppiness in the use of English and several inconsistencies.
xii. Keywords: phishers can send links within an email posing as messages from a genuine organization e.g. first bank Nigeria plc and this links lead to a page which is actually a phishing form and it reacts to incomplete forms and forms that contain certain keywords. Users may unintentionally go ahead to fill the form and give away their credentials.
xiii. Salutations and security: methods of salutation in a phishing mail is most usually the same as the attackers are sending this to as many people as possible, there is no personalization in the mail sent. Personalized mails are more likely to be legitimate than mails with just the usual greeting e.g. Dear customer, or our esteemed customer. In some attacks there is a sort of emphasis laid on security issues which seems rather unnecessary.

The Figure below shows an example of how phishing characteristic indicators can be represented by linguistic descriptors. The URL address is used here and a plot of the fuzzy membership functions displayed. On the x-axis are the ranges of possible values for the corresponding key phishing indicators and on the y-axis is a representation of these values to linguistic descriptor

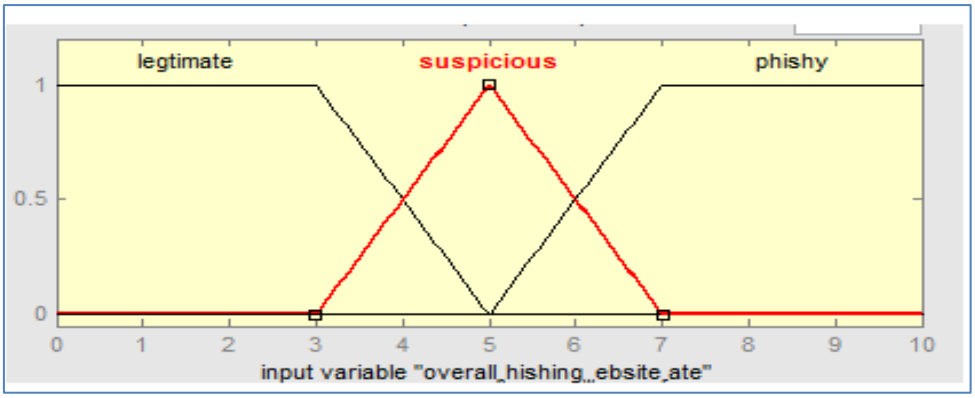

Fig-1

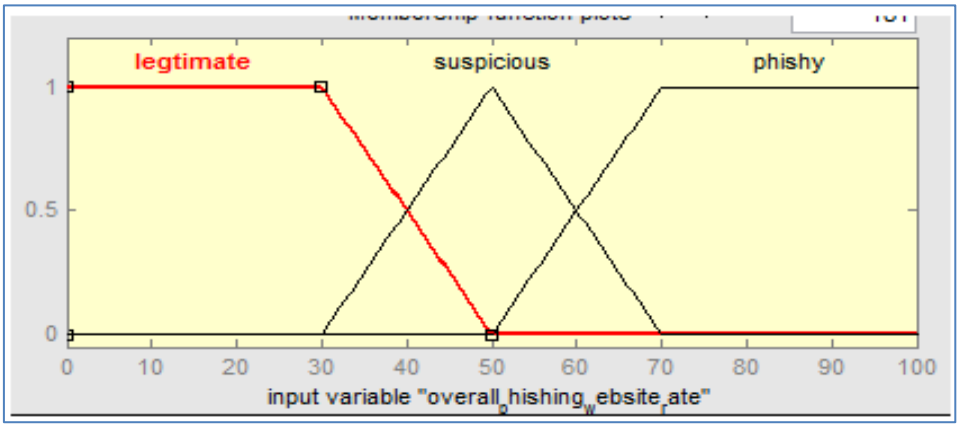

Fig-2

The plot above i.e. Figure 1and 2 show the overall phishing website risk rate displaying the linguistic indicators of the values generated in the order of legitimate, suspicious and phishy.

	Linguistic value	Numerical range
xiv.	Legitimate	[0, 0, 30, 50]
xv.	Suspicious	[30, 50, 70]
xvi.	Phishy	[50, 70, 100]

Phishy: there is a high possibility that the website is phishy and the user is at the risk of losing all his personal information.

Suspicious: the website cannot be considered as totally legitimate and there should be some kind of caution when visiting this website as it could have some serious consequences.

Legitimate: the website is considered safe for use.

System Design

In [18] phishing website risk rate is performed based on six criteria: URL and domain identity, Security & encryption, Source code & JavaScript, Page style and contents, web address bar and social human factor. A total of 27 in all. Each criteria has a number of components. In the proposed phishing website model, the criteria were all divided into 3 layers. In the first layer is the URL and domain identity criterion with a total weight of 0.3 (weighed according to its effect), the second layer is the security and encryption criteria; and source code and JavaScript with a total weight of 0.2 and the third layer page contents, web address bar and social human factor criteria. The final phishing fuzzy criteria is calculated as

0.3 * URL & Domain Identity crisp [First layer] + ((0.2 * Security & Encryption crisp) + (0.2 * Source Code & Java script crisp)) [Second layer] + ((0.1 * Page Style & Contents crisp) + (0.1 * Web Address Bar crisp) + (0.1 * Social Human Factor crisp)) [Third layer]

Rule Base 1

Rule base 1 for layer 1

There are 5 input parameters in total and a single output. It contains the if-then rules of the entire system. Each component of the rule base is assumed to be one of three given values based on linguistic descriptors and each criterion is made of 5 components, hence rule base 1 contains = 243 entries. The output of this rule base represents one of the phishing rate fuzzy sets i.e. genuine, suspicious and fraud representing URL identity. Table 4.4 shows a sample of the structure and entries for rule base 1, layer 1. The system structure for URL identity involves joining the six components which in turn produces the URL identity criteria.

Rule #	IP Address	Abnormal Request URL	Abnormal URL of Anchor	Abnormal DNS record	Abnormal URL	Phishing Criteria
1	Low	Low	Low	Low	Low	Genuine
2	Low	Low	Low	Low	Moderate	Genuine
3	Low	Low	Low	Moderate	Moderate	Suspicious
4	Low	Low	Low	Moderate	High	Suspicious
5	Low	Low	Moderate	Moderate	High	Fraud
6	Low	Moderate	Moderate	Low	High	Fraud
7	Moderate	Low	High	Moderate	High	Fraud
8	High	Moderate	Low	Low	Low	Fraud
9	Low	High	Low	Low	Moderate	Suspicious
10	High	Moderate	High	High	High	Fraud

There are two inputs in layer two, Security and Encryption and source code and JavaScript, and a single output. This layer has four components; they are (Using SSL certificate, certification authority, Abnormal cookie and Distinguished Names certificate. this is the rule base 2-1 and the second input which is the source code and JavaScript has 5 components and they are (Redirect pages, Straddling attack, Pharming attack, Using the mouse to hide links, server form handler.

Rule base 2 for layer 1

Rule No	SSL Certificates	Abnormal cookies	Certification authority	DNS	Phishing website risk rate
1	Low	Low	Low	Low	Genuine
2	Low	Moderate	Low	Low	Genuine
3	Moderate	Low	Low	Moderate	Suspicious
4	Low	Moderate	Low	Moderate	Doubtful
5	Low	Low	Moderate	Moderate	Fraud
6	Low	Moderate	Moderate	Low	Suspicious
7	Moderate	Low	High	Moderate	Fraud
8	High	Moderate	Low	Low	Suspicious
9	Low	High	Low	Low	Fraud
10	High	Moderate	High	High	Fraud

The structure of the rule base layer 2 is shown below. The system structure for this layer is the combination of two criteria i.e. Security and Encryption / Source code and JavaScript and the final rule base for layer two is derived based on this [18]. Table 7 shows the final rule base for layer two. The rule base contains 9 entries.

Rule No	Security and Encryption	Source code and JavaScript	Final phishing risk rate
1	Genuine	Genuine	Legal
2	Genuine	Suspicious	Legal
3	Genuine	Fraud	Uncertain
4	Suspicious	Genuine	Legal
5	Suspicious	Suspicious	Uncertain
6	Suspicious	Fraud	Uncertain
7	Fraud	Genuine	Uncertain
8	Fraud	Suspicious	Uncertain
9	Fraud	Fraud	Fake

In the final phase of the phishing website rule base, all input layers are to be considered and a final output identifying the risk rate of the fraudulent website. The structure of the system for this fuzzy model is a combination of all three layers to get the final phishing website rule base. The rule base is made up of (3^3) = 27 entries and the output is one of the classes (Legitimate, Suspicious and phishy) which represent the final rate of the phishing website risk.

Rule #	Layer one	Layer 2	Layer 3	Final website Risk rate
1	Valid/Legitimate	Legitimate	Legitimate	Legitimate
2	Valid	Legitimate	Suspicious	Legitimate
3	Valid	Legitimate	Fraud	Suspicious
4	Valid	Suspicious	Legitimate	Suspicious
5	Valid	Suspicious	Suspicious	Suspicious
6	Valid	Suspicious	Fraud	Phishy
7	Valid	Fraud	Legitimate	Suspicious
8	Valid	Fraud	Suspicious	Phishy
9	Valid	Fraud	Fraud	Phishy
10	Suspicious	Legitimate	Legitimate	Legitimate
11	Suspicious	Legitimate	Suspicious	Suspicious
12	Suspicious	Legitimate	Fraud	Phishy
13	Suspicious	Suspicious	Legitimate	Suspicious
14	Suspicious	Suspicious	Suspicious	Suspicious
15	Suspicious	Suspicious	Fraud	Phishy
16	Suspicious	Fraud	Legitimate	Phishy
17	Suspicious	Fraud	Suspicious	Phishy
18	Suspicious	Fraud	Fraud	Phishy
19	Fraud	Legitimate	Legitimate	Suspicious
20	Fraud	Legitimate	Suspicious	Suspicious
21	Fraud	Legitimate	Fraud	Phishy
22	Fraud	Suspicious	Legitimate	Suspicious
23	Fraud	Suspicious	Suspicious	Suspicious
24	Fraud	Suspicious	Fraud	Phishy
25	Fraud	Fraud	Legitimate	Phishy
26	Fraud	Fraud	Suspicious	Phishy
27	Fraud	Fraud	Fraud	Phishy

IMPLEMENTATION

Software requirements

Client: Mozilla Firefox

Programming language: JavaScript and python 2.7.3: General purpose open source, high-level language and the mode of design emphasizes code readability. The proposed plug-in was developed using JavaScript, a compatible version of Firefox and the Mozilla Firefox Addon sdk (software development kits) which is a development tool will aid in the development of this application. The SDK requires the need for Python to be running on the system.

For the implementation of the proposed fuzzy based phishing detection, the system proposed is a tool bar as a plug in to be installed on the client's computer for Mozilla Firefox. The function of this toolbar is to dynamically protect users online by alerting them when they wander into a phishing website. The major work to be carried out was to validate the values derived and understand the rules these values have been able to generate. In the initial state, phishing features and patterns were extracted and assigned to different classes known as fuzzy sets. The fuzzy sets are used as a guide to generate rules that will determine the risk rate of different web pages.

Using Firefox addon sdk along with python 2.7.3 we developed a system called phish addon which functions by JavaScript and the rules defined earlier in the literature to determine the status of the site in question. If a suspicious or fraudulent site is detected the user will be automatically notified using a friendly alertbox. The Add-on SDK is a set of APIs bundled with command line utility than allows for easy development of add-ons with your own tool chain. The Add-on SDK is set of APIs bundled with a command line utility that allows you to develop add-ons with your own tool chain. In its usage it allows free choice of code-editors, you can locally access development files and get control over the building steps of the sdk [38].

User Manual

The system is very easy to use and does not require any expertise at all to operate. At the moment this research has not been published so the plug in developed cannot be directly installed from the internet and used. Installing the add-on at the moment will require any intending user to open complete the following steps:

- Open Mozilla Firefox on your personal computer.
- Key in ctrl + o. This action will open the file browser using which, you can
  Browse to the location containing the plug-in installation file. As shown below

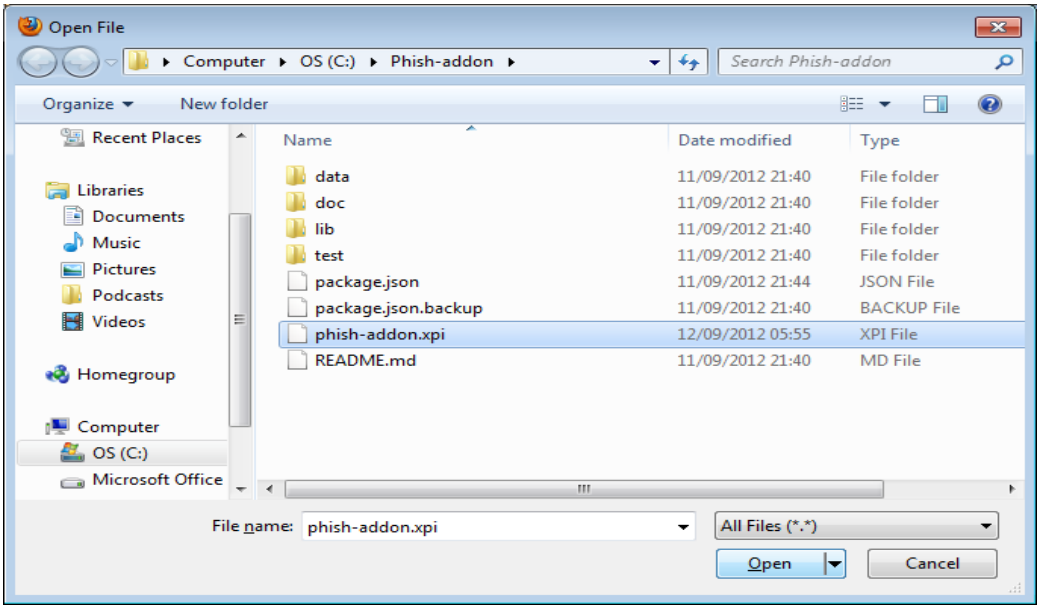

Fig-3

- Selecting the file will install the plug-in in Firefox as shown in the screenshot above.
- After installation browse the web normally.
- You will be notified when a suspicious or fraudulent site is detected as shown below.

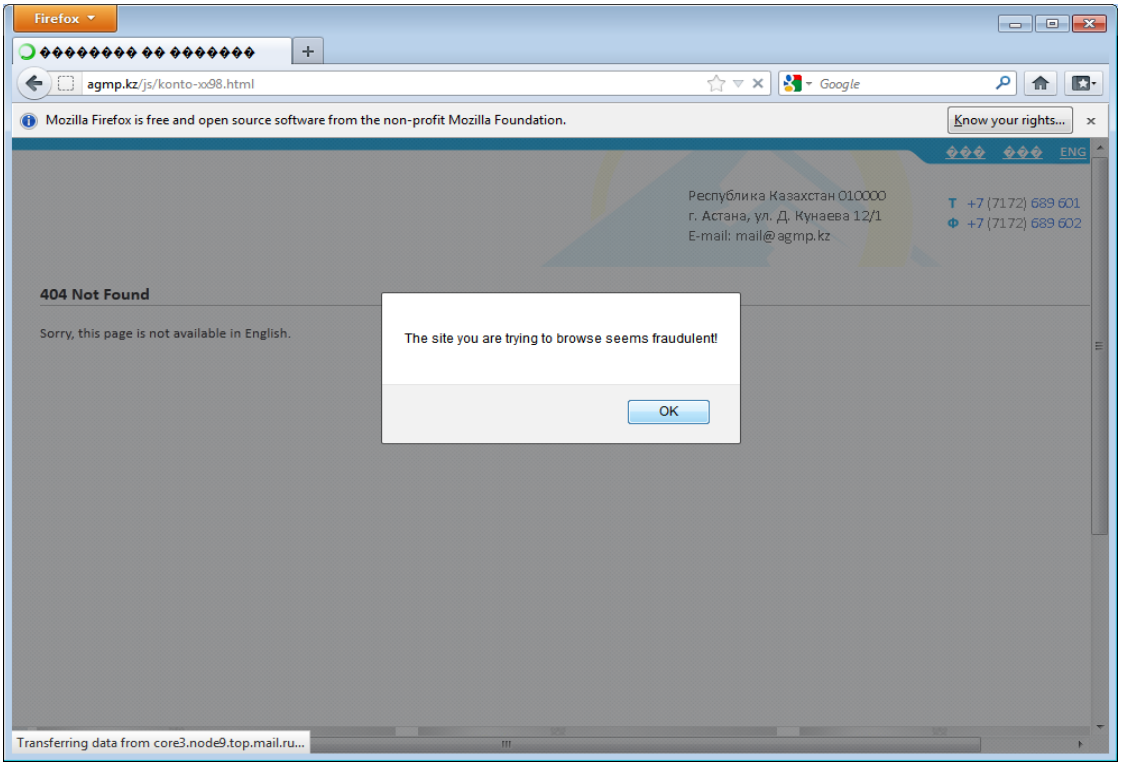

Fig-4

CONCLUSION

The initial objective of this research is to develop an intelligent plug-in capable of detecting phishing websites and also alert customers of the first bank e-banking website when they are in danger of attacks. Several approaches to counter phishing activities have proved unsuccessful because some Antiphishing tools examine the content of web pages [39] to be able to tell if they are legitimate or otherwise, in this event it is necessary to wait till the examination is complete while the page you requested is still loading. The alert indicator in this case has to change to actually show that a phishing website has been detected. A delay in this change may mean the user has already fed his personal information before realizing.

Blacklists can't be considered as very safe either as each URL is matched against a list of blacklists. Some tools automatically receive updates while others don't get often updates and in this case identifying phishing websites will not be effective enough.

This led to the question behind this dissertation – how to develop an Antiphishing solution that had the ability to automatically detect phishing websites based what they look like, their patterns and the common features they possess. This model does not need to receive updates and maintain a database of previous phishing sites. Our model is an intelligent fuzzy based system that uses fuzzy logic to access a set of rules. Based on the weights of each layer of the rules we can determine the phishing risk rate of any webpage and rank them as legitimate, suspicious or phishy. The structure is made of three main phases (Fuzzification, inference and defuzzification). Fuzzy rules are derived from previous expert knowledge; they are processed by fuzzy sets operations in the inference engine for final calculation of the phishing website detection rate. These results were generated by Aburrous *et al*. [18] as the 27 phishing characteristic indicators and a careful selection of some of these features were implemented in the design of an intelligent browser plug in. From the results it can be seen that the criteria URL and Domain identity is very significant and it appeared in all the phishing tests that were carried out. The browser Antiphishing toolbar developed was able to detect some phishing features and determine which pages are legitimate and which are fake.

As part of this research two experiments were carried out to determine website attack techniques, phishing detection and user involvement. A set of questions were drafted for users to determine the vulnerability of users and understand why users are quick to fall for attacks and a group users were tested to determine the level of awareness individuals have gotten to in determining phishing attacks.

For phishing activities to be tackled properly it is necessary to educate more users on the dangers of giving out information on the web, naivety of many users still makes them fall for attacks despite the sophisticated methods used to fight against phishing attacks.

RECOMMENDATIONS AND FUTUREWORK

The proposed fuzzy based phishing detection system can be implemented not only in e-banking websites but in all activities carried out on the internet (e-commerce sites too). It only requires for the user agent to be installed on the client's computer. In future approaches, predicting phishing websites is essential and can be done by the use of neural networks. Depending on data mining classification algorithms alone has a high error rate as stated by A.Martin [24] and Aborrous [24]. In an event of failure of neural networks, the system still continues to function properly without any issues because of its parallel nature, therefore in this method performance is made better and the error rate will be drastically reduced. Furthermore, the 27 phishing characteristics can be used as inputs to neural networks first layer and then the output generated are used as the inputs for the second layer and the same logic will be applied to the third layer and the output of this final layer will give the final detection rate as legitimate, suspicious or phishy.

Another point to consider is the validation of grammatical errors, spelling mistakes and sloppy use of English. There are so many words to consider especially when checking for key words. There are some nouns that are not listed as proper or correct words and can be considered as spelling errors whereas they are not. In the future there will be a need to use a sort of keyword extraction algorithm to solve this issue of spelling errors.

REFERENCES

1. Gardner, E. M., McLees, M. P., Steiner, J. F., del Rio, C., & Burman, W. J. (2011). The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. *Clinical infectious diseases*, *52*(6), 793-800.
2. Benhard, M. Hammerli, R. S. (2007). Detection of Intrusions and Malware and vulnerability assessment. Berlin Heidelberg, *New York*, 22-27
3. Rust, R. T., & Oliver, R. W. (1994). The death of advertising. *Journal of Advertising*, *23*(4), 71-77.
4. Joseph, M., & Stone, G. (2003). An empirical evaluation of US bank customer perceptions of the impact of technology on service delivery in the banking sector. *International Journal of Retail & Distribution Management*.
5. Darwish, A., & Lakhtaria, K. I. (2011). The impact of the new Web 2.0 technologies in communication, development, and revolutions of societies. *Journal of advances in information technology*, *2*(4), 204-216.
6. Soman, C., Pathak, H., Shah, V., Padhye, A., & Inamdar, A. (2008). An intelligent system for phish detection, using dynamic analysis and template matching. *World Academy of Science, Engineering and Technology*, *42*, 321-327.
7. Rampur, S. (2011). Video Surveillance in the Workplace.
8. Afroz, S., & Greenstadt, R. (2011, September). Phishzoo: Detecting phishing websites by looking at them. In *2011 IEEE fifth international conference on semantic computing* (pp. 368-375). IEEE.
9. Firdous, S., & Farooqi, R. (2017). Impact of internet banking service quality on customer

satisfaction. *The Journal of Internet Banking and Commerce*, *22*(1), 1-17.
10. Suh, B., & Han, I. (2002). Effect of trust on customer acceptance of Internet banking. *Electronic Commerce research and applications*, *1*(3-4), 247-263.
11. Gupta, B.B., Arachchilage, N.A.G., & Psannis, K.E. Defending against phishing attacks: taxonomy of methods, current issues and future directions. *Telecommun Syst*, 67, 247–267 (2018). https://doi.org/10.1007/s11235-017-0334-z
12. Li, Y., Yao, J., Han, C., Yang, J., Chaudhry, M. T., Wang, S., & Yin, Y. (2016). Quercetin, inflammation and immunity. *Nutrients*, *8*(3), 167.
13. Huang, H., Zhong, S., & Tan, J. (2009, August). Browser-side countermeasures for deceptive phishing attack. In *2009 Fifth International Conference on Information Assurance and Security* (Vol. 1, pp. 352-355). IEEE.
14. Bhadane, A., & Mane, S. B. (2018). Detecting lateral spear phishing attacks in organisations. *IET Information Security*, *13*(2), 133-140.
15. Robertson, J. L., Jones, M. M., Olguin, E., & Alberts, B. (2017). *Bioassays with arthropods*. CRC press.
16. Auffray, C., Balling, R., Barroso, I., Bencze, L., Benson, M., Bergeron, J., & Del Signore, S. (2016). Making sense of big data in health research: towards an EU action plan. *Genome medicine*, *8*(1), 1-13.
17. Wu, M., Miller, R., & Garfinkel, S. (2006). Do Security Toolbars Actually Prevent Phishing Attacks?, CHI.
18. Aburrous, M., Hossain, M. A., Dahal, K., & Thabtah, F. (2010). Intelligent phishing detection system for e-banking using fuzzy data mining. *Expert systems with applications*, *37*(12), 7913-7921.
19. Crain, J., Opyrchal, L., & Prakash, A. (2010, February). Fighting phishing with trusted email. In *2010 International Conference on Availability, Reliability and Security* (pp. 462-467). IEEE.
20. Wood, W., Tam, L., & Witt, M. G. (2005). Changing circumstances, disrupting habits. *Journal of personality and social psychology*, *88*(6), 918.
21. Joshi, Y., Saklikar, S., Das, D., & Saha, S. (2008, December). PhishGuard: a browser plug-in for protection from phishing. In *2008 2nd International Conference on Internet Multimedia Services Architecture and Applications* (pp. 1-6). IEEE.
22. Yue, C., & Wang, H. (2008, December). Anti-phishing in offense and defense. In *2008 Annual Computer Security Applications Conference (ACSAC)* (pp. 345-354). IEEE.
23. Chen, Y., Ma, W. Y., & Zhang, H. J. (2003, May). Detecting web page structure for adaptive viewing on small form factor devices. In *Proceedings of the 12th international conference on World Wide Web* (pp. 225-233).
24. Martin, A., Anutthamaa, N., Sathyavathy, M., Francois, M. M. S., & Venkatesan, D. V. P. (2011). A framework for predicting phishing websites using neural networks. *arXiv preprint arXiv:1109.1074*.