International Journal of Grid and Distributed Computing
Vol. 14, No. 1, (2020), pp. 513-529

Integrated Machine Learning Model for an URL Phishing Detection

Gouse Baig Mohammad[a*], S. Shitharth[b], Puranam Revanth Kumar[c]
[a,b]Department of Computer Science Engineering, Vardhaman College of Engineering, Shamshabad, Hyderabad – 501218, India.
[c]Department of Electronics and Communication Engineering, IcfaiTech (Faculty of Science and Technology), Hyderabad - 501203, India.

Abstract:
The problem of phishing attacks in enterprise is next issue rising in wide scale and complexity, as phishers use email phishing via obfuscated, malicious or phished URLs and continuously adapt or innovate their strategies to lure victims. To gain trust and confidence of victim's phishers have started using visceral factors and Familiarity cues. Although in most cases a phisher's clear motive is to commit identity theft in order to benefit from it financially; it is wrong to assume that phishing is always money centric. A phisher can also rob an internet user of his goodwill and character. There are no limits to what a phisher can do in such a scenario. Earning a bad name for oneself in a professional or academic arena can prove much more traumatic than being embarrassed at a social networking site. It is a challenging task to address this issue. It is evident through extensive literature review that single phishing detection filter approaches are insufficient to detect different categories of phishing attempts in enterprise environ. Therefore, a novel anti-phishing model for enterprise using artificial neural network is proposed. In addition, this model effectively identifies whether the phishing email is known phishing or unknown phishing to reduce the trust and familiarity-based email phishing enterprise environ. The Feed-Forward Backpropagation and Levenberg- Marquart methods of Artificial Neural Network (ANN) are adopted to enhance the URL classification process and with Fuzzy Inference System to get result with imprecise data of social features. The proposed model can accurately classify the known and unknown email phishing via URLs.

Keywords: E-mail Phishing, Enterprise cyber threat, Social engineering , Artificial Neural Network, Fuzzy Inference System

1. Introduction

In digital era, email phishing is a critical issue causing loss of finance during online transactions. At present, different anti-phishing approaches are being proposed in order to detect email phishing attacks. Despite the various developed anti-phishing approaches are profoundly incompetent to tackle the real-time hassles. In literature review, a group of researchers argued upon about URLs blacklisting, which are mostly used in industry or organizations but not well off to detect email phishing attacks with accuracy. The main reason, why still phishing attacks prevails is due to lack of computer operational knowledge and unawareness of cyber threats among internet users. Many internet users still remain unacquainted of recognizing phishing emails (Krombholz et al., 2015) (Stolfo et al., 2003).

For Instance, users or employees do not understand the syntax or the semantic of the URLs and it is actually a hard nut to crack to differentiate the between fake webpages and legitimate ones. Since the internet user are less aware and alert, the phishers gets the scope of crafting phishing emails via spoofed websites / obfuscated/ malicious URLs/links and forward them towards targeted internet user or spread bulk emails randomly. Although there are many different ways to launch phishing attack via URLs, emails, instant messages, forum post and comments and social media etc. Email Phishing in an enterprise is an emerging problem now-a-days and solving these problems is a challenging task (Vandermeer, 2006), (Lee et al., 2015).

Therefore, this paper is focused on email phishing detection in an enterprise. There are a number of possible research findings present with email phishing detection systems or models with sophisticated machine learning techniques to meet out the challenges. Currently the security mechanism broadly classifies into two categories (Tyler et al., 2005). One is list based approach and other is heuristic based approach. List-based approach assesses the existence of the legitimate website and accordingly stores in the predefined list which includes blacklist or white list or both (McGrath & Gupta, 2008). On the other hand, heuristic based approach is based on some distinctive features of phishing website or malicious URLs to facilitate the detection and identification of phishing website (Cialdini, 2001). Heuristic based model is designed to detect email phishing

ISSN: 2005-4262 IJGDC

via URLs (obfuscated, malicious or phished URLs) using Naïve Bayes and Support Vector Machine classifiers. This model includes a URL detection algorithm which efficiently detects phishing and legitimate URLs (Social Network Analysis, 2017).

Research evidences suggest that the use of rule base phishing detection method has application in industry for identifying phishing attacks to analyze website information. As of now, many different phishing filtering approaches are existing with ML techniques such as logistic regression, Support Vector Machine, Markov Model, decision tree, and random forest for detecting phished URLs/links (Sankhwar et al., 2019) (Medvet et al., 2008). In this chapter, a novel approach for email phishing detection in enterprise environ is proposed with ANN.

2. Phishing – An Enterprise Threat

Since the last decade, social networking has come into lime light and has drawn users' remarkable standard of attention towards it. Social Networking refers to web-based services which allow users to connect with other users within their specific arena. Creation of profiles which could be public or protected are both meant to socially link with others in acquaintance or randomly. Social networking has emerged as a significant mode of online conferencing and sharing and exchanging contents or personal data, approaches, sentiments, opinion expression and feelings etc. The reliance and inclination of more and more users towards e-communication and social networking for information, news, opinions and other diverse matters has caught the eyes of phishers and has lured them towards the same (Wang et al., 2012).

2.1. Phishing Attack through Social Engineering

Current adversaries have understood the ease, effectiveness, and efficiency of social engineering. It includes persuasion, data gathering, foot-printing to launch phishing attack using email as a vector has proven particularly successful. Phishing attacks have also been quite extensive in the email sector and social networking. To plan and execute these attacks it becomes mandatory for phishers to do deep social engineering to make there snares foolproof and result oriented (Kaivanto, 2014). The attacks using social engineering have become very common in Enterprise Environ. The specific attacks based on social engineering comprise psychological manipulation, fooling and impersonation. So that the unsuspecting users or employees may handover there confidential and sensitive data.

Usually, email communication or others means are used as machete to bring the theory in framework. The emails or other communication methods are structured in a form which could invoke fear, urgency, reply inducement or other similar feeling in the victim leading him to promptly trust or reveal his/her sensitive information. For instance, clicking on malicious link or file which would redirect the user to an illegitimate destination. In this era, when social engineering is an elephantine element to secede from these attacks, becomes very much impossible for the enterprise employees (Stringhini & Thonnard, 2015).

Phishers are well verse that mostly users do not want to encounter or challenge an authority or create an uncomfortable social interaction. Keeping this in mind, phishers do social engineering and take the advantage of users' behavior and starts information gathering and foot-printing such as employment history, hobbies and likes, family background or any other favoritism of the potential victim, using these observations the attacker create the email or profile which would potentially motivate the user to build a sense of authenticity. Then next the attackers try and built a remote relationship with the user using a cloak of his preferences, generally luring them towards there area of interest like sending online invitations for specific seminar, tender, projects, assignments or unsolicited promotions, job offers or a call for job interview (Adedoyin-Olowe et al., 2013).

2.2. Phishing Strategies in Enterprise Environ

In enterprises environ, phishers gather internal information about the enterprise, names of collogues, relevant post to craft tailor-made emails which would contain a phished/malicious URLs/links. Looking at the reference/ source of this email the victim automatically believes the authenticity and falls prey. This type of attack crafted by phishers and called as in insider attack in the enterprise environ. The attackers often disguise as a trustworthy and genuine entity and build contacts with their targets (naive user/ employee) through email & social media. Often, they tend to attack a random individual or enterprise itself (Keshtkar & Inkpen, 2009). The individual attacks may pose specific spurious emails or post of an individual's concern or interest like using the posted official details of travel plans or reimbursement sanction, awards & promotion related with enterprise protocols and others, asking confirmation of debit/credit card details, username, password, desk number, date of birth etc(Kim et al., 2013).

When it comes to the phishing attack in an enterprise, the boomerang thrown by attackers is usually the business email or familiarity cues. The targeted emails in this concern are sent to employees appearing to have sent from enterprise executives or IT staff ordering them to make wire transfers. The sly attackers already do their homework about gathering user data and acquiring information details of the working structure of enterprise with foreign associates or firms with which their

instruction and actions are unquestioned or undoubted by the enterprise personnel (Wakker, 2010). This states that the phishing attacks executed through social engineering comprises four steps:

- Collating information and fake pretenses.
- Developing relationship or nexus.
- Exploitation of identified vulnerabilities.
- Execution of phishing attack.

3. Illustrations

The term social facets define the social human factors which lead to cyber-crime. Here, major social facets responsible for phishing are explained as given below:

3.1. Familiarity Exploit

Familiarity exploitation refers to the phishing attacks targeted to prey a known user and sometimes unknown user as well targeted, phishers turn him as an amiable by generating a touch of friendliness or amity. One of the most likely auras in phishing attacks and also corner stone of social engineering is the exploitation done through the foundation of familiarity (Keshtkar & Inkpen, 2009). In a nutshell, making the appearance look perfectly normal to everyone that it is genuine to proceed and be a part of concerned webpage/website. Different people react to different situations or people differently but a person with acquaintance or familiarity will never leave any mark in the eye. The presence of a familiar person turns off the alarm bells of scepticism in one's head. Hence, an insider in enterprise itself is usually a snake in the grass (Wang et al., 2012).

3.2. Pretexted/ Imposter Account

Imposter account are crafted to target the users, customers or employees of an enterprise by tricking them and eventually provoking them to hand over their confidential information or login credentials i.e. personal data or information relating to the enterprise usually which is done through sending a hoax hyperlink to the target. Clicking on the link present in email may end up landing on a website designed to steal the sensitive information or credentials (Kaivanto, 2014).

3.3. Insider Threat

An insider threat is a former or current employee/ freelancer or any one in an enterprise who is unauthorized to access the sensitive data, technology, systems or other areas of interest which means people among and amidst us (Adedoyin-Olowe et al., 2013). It is the main cause of more than three fourths of the total cyber security breaches in enterprise environ attributed to a trusted one or an insider. The increasing pace of accessing email, social networking via techno gadgets has made it very accessible for the adversaries or phishers to pose an insider threat or launch a phishing attack (Kim et al., 2013). The negligent and unwitting users fall into the snares of the insider fraudsters who contribute the largest and biggest pool of potential phishers in enterprise environ (Vandermeer, 2006).

3.4. Human Error

The phishers attack the users or employees using the psychology of deception through the means of persuasion, manipulation of thoughts inflicting the feelings of trust, pity, fear, remorse, greed, anxiety, pleasure and other such overwhelming visceral factors ultimately making them fall prey in the snares (Kim et al., 2013). So far, the phishers have attacked the employee or users in enterprise leading them vulnerable enough to fall victims (Page et al., 1999). The ultimate flaw and loophole are the unawareness and negligence of users themselves (Baeza-Yates & Davis, 2004).

4. Proposed Approach for E-mail Phishing Detection

An Anti-Phishing in Enterprise Environ (Anti-PhiEE) is integrated Approach to Detect Email-Phishing through malicious URLs in Enterprise Environ. Anti-PhiEE is an email phishing detection model as shown in Figure 1. Anti-Phishing Multi-Filter (APMF) is developed with 25 heuristics that work as multi-layer filters. APMF consists of five layers which are used to discriminate between the legitimate and phished URLs as well as Social Facet filter (social human factor scanner) to identify the known and unknown phishing as shown in Figure 2.

ISSN: 2005-4262 IJGDC

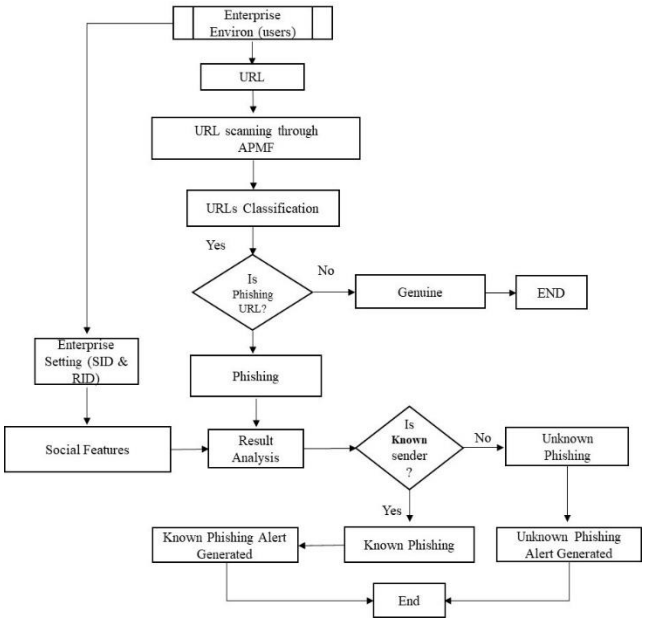

Fig. 1. Architecture of Anti-PhiEE Model

Pertinent 25 heuristics is identified through the exhaustive literature review, statistical investigations and analysis on phished and legitimate URLs/websites. In this study, a neural network toolbox of MATLAB9 is used; a backpropagation learning algorithm has been used here. TRAINLIM (that is, Levenberg-Marquardt back-propagation) algorithm is also used in this approach. The Artificial neural network methods Feed-Forward Backpropagation and Levenberg-Marquardt Neural Network used for URLs classification, FRBS is used for Social facet filter with Mamdani FIS method to determine the known and unknown phishing.

4.1. Architecture of Proposed Approach

An Anti-Phishing in Enterprise Environ (Anti-PhiEE) Model is integrated Approach The foremost aim of the architecture is to detect the email phishing in an enterprise with the help of APMF and another objective is to identify whether the phishing attack is known or unknown as shown in Figure 1. Anti-PhiEE works basically in three Phases as explained with steps below:

Phase I:

Step 1: URLs are Fetched from the e-mail.
Step 2: Anti-Phishing Multi-Filter (APMF) is applied on fetched URLs to check whether the URLs are legitimate or phished.
Step 3: ANN is applied on URLs dataset for classification i.e. Feed-Forward Backpropagation Neural Network and Levenberg- Marquardt Neural Network.
Step 4: Classifier classifies the URLs nature whether it is legitimate or phished
Step 5: The result/output produced by classifier is accumulated or stored for further analyses.

Phase II

Step 1: In this step, to identify Known Phishing or Unknown Phishing the Social Facets as an input (imprecise data) is taken which is divided into four Social Features i.e.X1, X2, X3, X4 (see in section 4.4.3).
Step 2: The Social Feature Algorithm is applied to check whether the sender is known or unknown.
Step 3: The Mamdani FIS used is for analysis of input values (social features).
Step 4: The output of phase I and the output value gained from step 3 of phase II are merged together to get the output- Known Phishing or Unknown Phishing. At last, the phishing alert message generated. Alert message for known phishing is quiet emphasized according to sender and receiver 'known' intensity as per fuzzy linguistic rule output.

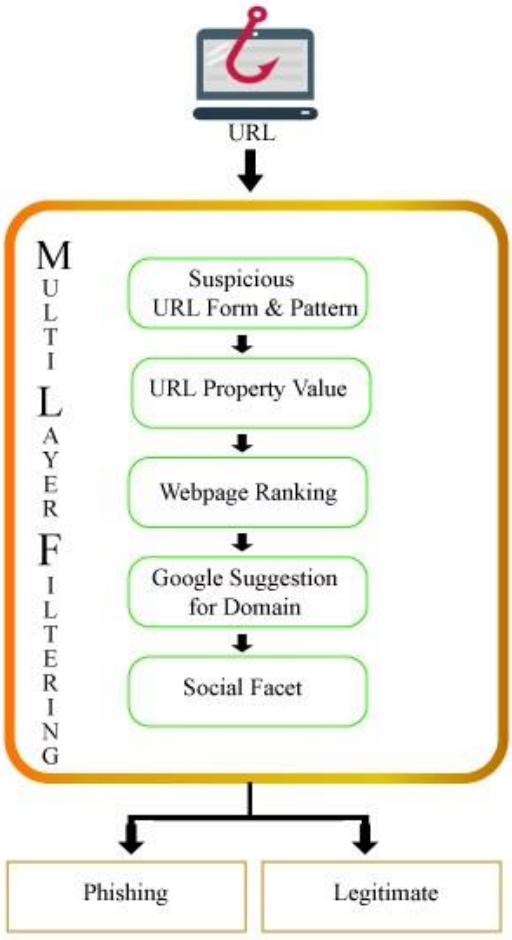

Fig.2. Design of Anti-Phishing Multi-Filter (APMF)

4.2. URL Feature Set

Heuristics approach is used to discriminate phishing URL/websites and legitimate URLs/webpage/links. Based on Advanced pertinent threats the phishing heuristics is identified; through the exhaustive literature review, statistical investigations and analysis on phishing and legitimate websites or URLs (Madhusudhanan Chandrasekaran et al., 2006),(Kotsiantis et al., 2007), (Fette et al., 2007), (Medvet et al., 2008), (Shah et al., 2009), (Xiang et al., 2011), (Zhang et al., 2008), (Greg Aaron & Rod Rasmussen, n.d.), (Center, 2012), (McGrath & Gupta, 2008), (Sankhwar & Pandey, 2017). Total 25 Heuristics are defined here to effectively determine the legitimate and phished URL as listed below:

Table 1 - Suspicious URL Forms or Patterns

Heuristics	Description
IP Address, Hexadecimal or ASCII code in URL	If URL in the form of IP Address, If URL in the form hexadecimal or Unicode.
Abnormal URL	URL- phishing Page Redirection
No. of Sub Domain	Length of sub domain
No of Dot '.' In URL	More than 5 Dots in URL
URL of length	Length of URL
Special Characters	Whether URL has '- ', '@' symbol or '//'
Phishing Keyword	Phishing words as a hyperlink like- verify, click here, submit, login, sign-in etc.
Age of Domain (in Days)	Domain is less than 43 days

Port number matching	Whether explicit port number and protocol port no. are equal.
Number of TLDs	More than one TLD in a URL
Primary Domain Spelling Mistakes	Whether primary domain is
Number of Slash '/'	Number of '/' slash
Login Form	Login Form in Fake webpage

4.2.1. Suspicious URL Forms or Patterns

These Heuristic are associated with Suspicious URL forms or patterns and symbols, The Characters such as '@' and more than one time '//' rarely appear in a URL (Yearwood et al., 2012) (Suriya et al., 2009). The legitimate sites have one TLD so if URL containing more than one considered as phishing site. Phishing sites have very less life-time as get block listed. Fake Login form in a phishing page is a dangerous sign of loss money or sensitive information as listed in Table.1(Madhusudhanan Chandrasekaran et al., 2006), (Netcraft Anti-Phishing Toolbar, 2004), (Fette et al., 2007), (Suriya et al., 2009), (Gansterer & Pölz, 2009), (Firake et al., 2011), (Al-Momani et al., 2011),(Yearwood et al., 2012), (Almomani et al., 2012), (Almomani et al., 2013), (Malaysia, n.d.), (Nguyen et al., 2014), (Smadi et al., 2015),(Gupta et al., 2017), (Jayakanthan et al., 2017),(M. Chandrasekaran et al., 2006), (Herzberg, 2009).

4.2.2. URL Property Values

These heuristics are based on URL Property Values for identification of phishing URL/website (Almomani et al., 2013), (McGrath & Gupta, 2008). The fake or temporary phishing a site does not contains required properties as listed in Table 2. (Bedingfield Sr & Gehl, 2012), (Berners-Lee et al., 1994).

Table 2 - URL Property Values

Heuristics	Description
Country matching	TLD country and domain country-code are equal.
HTTPS protocol	Whether URL use HTTPS or not.
DNS record	Whether URL has DNS record or not.
Reverse DNS look-up	Query of DNS to determine the domain associated with an IP Address.
WHOIS Record	WHOIS record (Domain name, Registration, Expiry details etc.)
Value of TTL	TTL value of domain.
PTR record	Whether domain has PTR record or not.

4.2.3. Page Ranking

These heuristics are based on page ranking, it is a numerical value calculated by the number of visitors and degree of popularity (Page et al., 1999), (Baeza-Yates & Davis, 2004).

Table 3 - Page Ranking

Heuristics	Description
Google page rank (Indexing)	Domain's PageRank value
Alexa rank	Alexa Rank value of domain
Alexa reputation	Alexa reputation value of domain

It is seen that the phishing site has very low page rank value as rarely visited by bulk users and these sites are exist for less time.

Therefore, domain page rank value is very low as mentioned in Table 3.

4.2.4. Google suggestion for URL Authenticity

These heuristics are made on the idea of Google suggestion with existing algorithms to match or compare the String (Mohammad et al., 2013). String matching algorithms are used to detect phishing URLs as listed in Table 4. When a user enters a single term in google suggested word is returned. Using this idea of entering the URLs of phishing sites and legitimates sites for google suggestion the result analysis done (Damerau, 1964) (Phelps & Wilensky, 2000).

Table 4 - Google suggestion for URL Authenticity

Heuristics	Description
Similarity of primary Domain and Google suggestion	Levenshtein distance between primary domain and Google Suggestion
Similarity of Subdomain and Google suggestion (subdomain)	Levenshtein distance between primary domain and Google Suggestion
Safety of Google Suggestion for Primary domain	To check Google Suggestion result of Primary domain is present in the whitelist or not
Safety of Google Suggestion for Subdomain	Google Suggestion result of Subdomain is present in the whitelist or not

If URL of phishing sites is searched and is similar to suggested result, the input URL is considered as suspicious as the site could be emulating an existing site. The URL as a string is considered in APMF and two well-known string-matching algorithms (Damerau-Levenshtein edit distance and Longest Common Subsequence) are employed to assess the similarity between two strings on the ground of different acts(Damerau, 1964)(Levenshtein, 1966).

4.3. Social Feature Set

It is featured and demonstrated that the behavioral aspects are taken from the information of email header. The proposed heuristichas been used in this regard to identify and discriminate between the known and unknown phishing attacks in Enterprise Environ as given below:

4.3.1. Social Media Contacts (X1)

In this heuristic, friendship on social media account of Sender (S_id) and Receiver (R_id) with help of their email's IDs can be identified. It has been observed through the trend of fraudulent activities and the brainstorming discussion with experts and professionals that mostly Primary ID is used for both Email Service Provider as well as on Social Media. There is always a possibility that someone who is a rightful worker associated of the relevant is an enterprise insider threat, fraudster or phisher and, in this case, they can easily send the phishing email from within the arena.

4.3.2. Social Media Common Contacts (X2)

In this heuristic, targeting the common friend that means friends of friend's social media account is identified. Adversary or insider threat do social engineering, for instance- peeing into the social media friend list of targets. Adversary tries to gain trust from targets or naive user. They gather information from friends of the target via social media through scraps, post, pages to launch the attack on email.

Illustration- if the Boss is friend of naïve user and Boss's Personal Secretory is common friend. He does social engineering and launch the email phishing attack with the content and URL showing his/her Boss instruction.

4.3.3. Social Media Common Activities (X3)

In order to catch the eyes of user the phisher involve and caste themselves into the same swing of user which may grab user's attention. For instance, liking the same post, pages and stuff on social media which user likes or declaring to visiting an event on social media site in which the user has shown interest. Further, volunteering or participating in specific professional project or a social task or cause like becoming a blood donor in which the user has also applied for. So, it simply means the phishers try to make an analysis of target's behavior in depth and grabbing their attention by literally chasing them on social media but not approaching them straight away however, at the same time being visible in the sight of users

constantly by their parallel activities in order to make the user feel that they both are on the same boat or likely to be in the same swing.

4.3.4. Frequency of Email Communication (X4)

Suppose, an employee in enterprise has not registered himself on social media so it becomes difficult to identify known and unknown phishing. Thereby, this heuristic is coined to check the Frequency of email communication so as to identify known phishing or unknown phishing through the means of Email Service Provider (ESP Contact list) as well as frequency of email communication between sender and receiver through the threshold value 1 (Tt).

4.4. Anti-Phishing Multi-Filter

Anti-Phishing Multi-Filter (APMF) Algorithm is developed to detect known and unknown Phishing in enterprise environ. APMF algorithm consist of five layers i.e. Page Ranking, URL Property Values, Suspicious URL Forms or Patterns, Google suggestion for URL Authenticity and Social Feature Set as shown in Figure 3.

```
Phishing URL Detection (PUD) input= URL;
output= Phishing [1] /legitimate [0]
Social Facets Analysis (SFA)   input= S_id, R_id;
output=known phishing [1] / unknown phishing [0]

PUD
{
Filter 1: execute EMUD algorithm
If Phishing
Return 1 else Return 0
Filter 2: Find the login Page
If (login page is phishing)
Return 1 else Return 0
Filter 3: Page ranking
If (page ranking is less)
Return 1 else Return 0
Filter 4: Google suggestion
Execute string matching algorithm
If (similarity score of domain is less)
Return 1 else Return 0
}

SFA
{
Filter 5: Social facets
Function Social Media Contacts (X1)
If (Sender and receiver are in direct contact in social media)
Return 1 else Return 0
Function Social Media Common Contacts (X2)
If (sender and receiver have common contacts in social media)
Return 1 else Return 0
Function Common Activity in Social Media (X3)
if (common activities exist between sender and receiver in social media)
Return 1 else Return 0
Function Times of Email Communication (X4)
Return 1 else Return 0
}

Start
If PUD returns 0 then display legitimate
else
KP=process result (PUD, SFA)
If KP = 1 then display known Phishing
else display unknown Phishing
End
```

Fig.3. Anti-Phishing Multi-Filter (APMF) Algorithm

These five layers contains total 30 heuristics. Fourteen Heuristics are already explained and discussed in (Sankhwar et al., 2019), and included in Enhanced Malicious URL Detection (EMUD) algorithm. EMUD algorithm is modified by adding few more relevant heuristics which is used to detect phishing email via Malicious or phished URL (Sankhwar et al., 2019).

In this research paper, rest sixteen heuristics are identified and discussed in above section (see in section 4.2). By merging both the algorithm that means Modified EMUD algorithm and four social feature Set together APMF is developed through which known and unknown phishing is determined. The APMF Algorithm as shown in Figure 3.

5. Basic Principle of Artificial Neural Network

An Artificial Neural Network (ANN) is based on mimicry of human brain having basically three parts receiving the input from the surrounding called input layers and number of nodes (processing element) connected each other for transforming

signal as a data is also called hidden, intermediate layer. In this research, we have implemented multi-layer feed forward neural network (Mohammad et al., 2013) (Kotsiantis et al., 2007). We have used Feed-Forward Back propagation Neural Network for analysis.

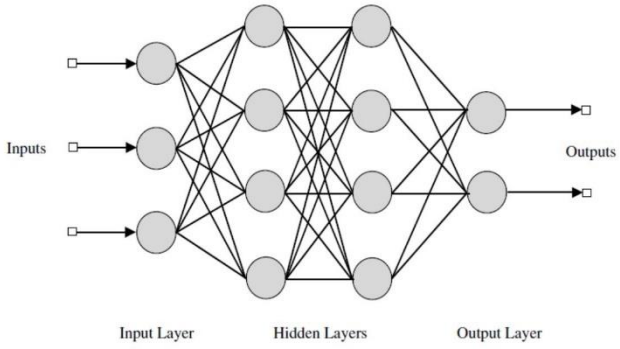

Fig.4. 25x2 Input Output Neural Network Architecture

Feed-forward Back propagation Neural Network (FFNN) consists of one and more hidden neural layer. In feed forward learning algorithm, we have used ten sigmoid algorithms for training process as shown in Figure 4. The number of hidden layers and their respective number of neurons depends upon the nature and complexity of the problem being mapped by neural network. The amount of output signal will always be précised with number of neurons from their respective layers (Kotsiantis et al., 2007), (Mitchell et al., 2013). We have also used Levenberg-Marquardt Neural Network for data analysis. The Levenberg Marquardt (LM) algorithm is an approximation to the Newton method used for training ANNs.

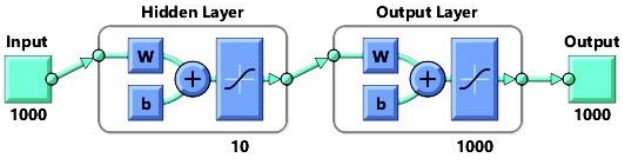

Fig.5. Artificial Neural Network

6. Implementation of Proposed Approach Phase-I

The APMF test the collected data of the 2000 phishing URLs and legitimate as input and thereafter, employed to machine learning for evaluation of the performance of APMF. The ANN methods, Feed forward back propagation and Levenberg-Marquardt Neural Network is used. In both ANN methods, same dataset is segregated in the form of training, testing 50%, 50% respectively. The neural network toolbox of MATLAB version R2016a is used for ANN applications and validation using statistical analysis is done (Negnevitsky & Intelligence, 2005). The result of this approach has been depicted in the Table 5. and graphically Figure 6 and Figure 7.

6.1. Dataset

The collected dataset is online data of the 2000 phishing URLs and legitimate. Specifically, the distribution ratio of phishing and legitimate data is in 60:40 ratio respectively. Phishing URLs data source is Phishing tank and legitimate URLs data source is DMOZ and Alexa (Phishing Tank, n.d.), (Alexa Topsites Datasource, n.d.).

In both ANN methods, same dataset is segregated in the form of training, testing 50%, 50% respectively. The result of this approach has been depicted in the Table 5.and graphically Figure 6 and Figure 7.

6.2. Performance Evaluation

In order to demonstrate the overall performance of the study experiments performed on two aspects: The performance of both the methods where measured using Root Mean Square Error (RMSE) and R2 value. The result of both the methods are shown in Table 5. For getting the accuracy of classifier with small dataset, k-fold Cross-Validation is adapted. k-fold cross-validation is used for validation of proposed phishing detection techniques. The total 2000 dataset divided into 10 smaller

datasets with equal size. This dataset is used for training, testing, validation. 10-fold cross-validation is applied in both Feed-Forward Back propagation and Levenberg-Marquardt Neural Network.

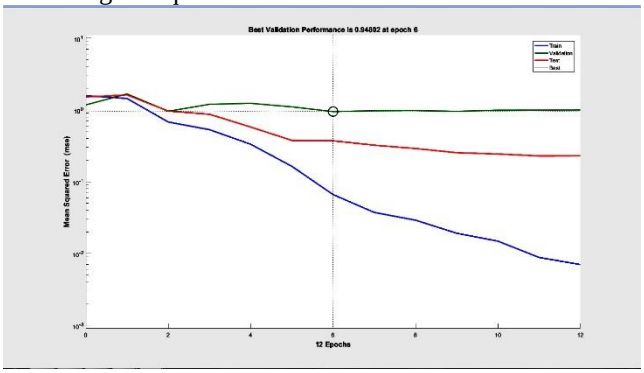

Fig.6. Performance of Feed-Forward Back propagation Neural Network

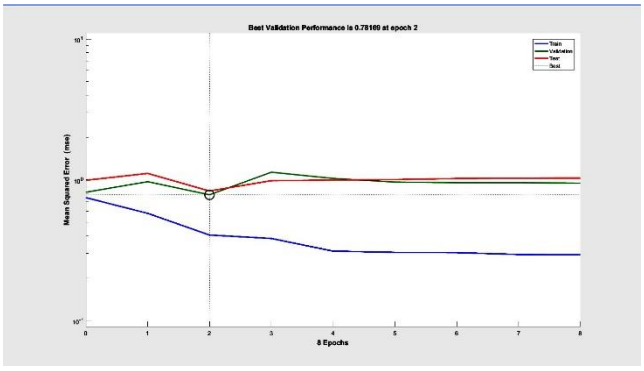

Fig.7. Performance of Levenberg-Marquardt Neural Network

6.3. Results and discussion

In both the approaches we have considered 50% of input data as a training set and 50% as a testing set. The method FFNN result reveals that the RMSE is 0.28 in case of training with R2 Value 0.92, whereas the testing error was less than 20% with R2 value 0.82. 10-fold cross-validation method used to validate the model. Similarly, the result of Levenberg- Marquardt Neural Network RMSE having 0.3% with R2 value 0.91. In comparison of both the methods FFNN yield quiet precise result are shown in Table 5. This optimization technique is more powerful than standard Backpropagation Neural Network (BPNN). LM algorithm is very efficient and fast, having also a quite good global convergence property. For these reasons, LM algorithm is used in this research as shown in Figure 5 (Mitchell et al., 2013), (Jang et al., 1997).

Table 5 - Comparison between FFNN and LM Neural Network

Different NN Architecture	Process	Sample size	RMSE	R^2
Feed-Forward Backpropagation	Training Set	1000	0.28	0.92
	Testing Sets (10 set each size 100)	1000	0.42	0.82
	10-Fold Cross Validation	100	0.24	0.94
Levenberg-Marquardt Neural Network	Training Set	1000	0.30	0.91
	Testing Sets (10 set each size 100)	1000	0.55	0.69
	10-Fold Cross Validation	100	0.46	0.78

7. Implementation of the Proposed Approach Phase-II

As we already executed the phase I of proposed approach and achieved 94% performance in terms of accuracy in phishing detection. In this phase II, the known and unknown phishing are determined. Therefore, same approach is applied in real world data of and an enterprise which has combined data of email and social media contacts.

The neural network toolbox of MATLAB version R2016a is used for ANN applications and validation by statistical analysis [109]. In this study, Mamdani FIS method is used with four linguist variables such Low, Medium, High, Very High each having input variable i.e. X1, X2, X3, X4 (see in section 4.3).

7.1. Dataset

Data set has been collected from real world data to check whether the sender is the known and unknown for known and unknown phishing attacks. Implementation is done on total 2000 emails with corresponding 2000 social media contacts of an enterprise. The source of all the data are originated from a real enterprise of Software Development and Training, Private Limited(Askysoftware Development Pvt. Online Available Https//:Www.Askysofware.Com., n.d.).

8. Fuzzy Rule-Based Systems (FRBS)

A Knowledge Base (KB) and Inference Engine (IE) are two main components of FRBS. There are various ways to represent knowledge. Perhaps the most common way to represent human knowledge is to form it into natural language expression (Jang et al., 1997). KB generally represents the knowledge about the problem being solved in the form of fuzzy linguistic IF-THEN rules, and the Inference Engineering (IE), which puts into effect the fuzzy inference process, is needed to obtain an output from the FRBS, when an input is specified. This form in expression is commonly referred to as the IF-THEN rule-based form like IF premise (antecedent), THEN conclusion (consequent) parameters. The schematic view of an FRBS is shown in Figure 8.

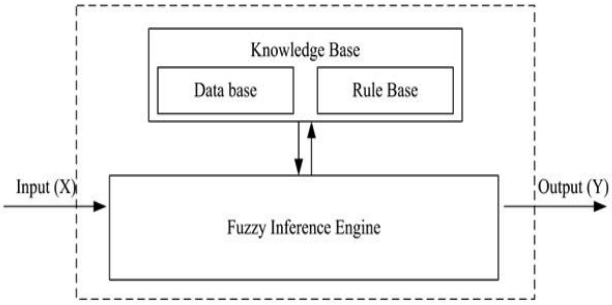

Fig.8. A schematic view of an FRBS

An FRBS consists of three modules, namely fuzzification, inference and defuzzification. Fuzzification is the process, in which the input parameters are converted into appropriate fuzzy sets to express measurement uncertainty. The fuzzified measurements are then used by inference engine to evaluate the control rules stored in the fuzzy rule base and a fuzzified output is determined. The fuzzified output is then converted into a single crisp value. This conversion is called de-fuzzification (Nauck & Kruse, 1996).

9. Fuzzy Linguistic variable and Membership functions

Fuzzy linguistic approach provides a systematic way to represent linguistic variables in a natural evaluation procedure (Buckley & Hayashi, 1994). A fuzzy linguistic label can be represented by a fuzzy number, which is represented by a fuzzy set (Nomura et al., 1992). Fuzzy sets capture the ability to handle uncertainty by approximation methods (Mamdani & Assilian, 1999). A fuzzy set α is represented by a pair of two things – the first one is the element x and the second one is its membership value μα (x) (varying in the range of [0, 1]), as given below:

$$\alpha = \{(x, \mu_\alpha (x)): x \in X\}.$$

For the inputs and output, triangular membership functions were used in order to keep the design of the FLCs simple. A degree of overlapping of two was used, as shown in Figure 8. Furthermore, a universe of discourse normalized to the range

of [0.0, 1.0] was utilized. This value, called membership value or degree of membership (as given below), quantifies the grade of membership of the element in X to the fuzzy set A.

$$\mu_A(x) = \begin{cases} 0 & x \le a \\ \frac{x-a}{m-a} & a < x \le m \\ \frac{b-x}{b-m} & m < x < b \\ 0 & x \ge b \end{cases} \quad (1)$$

Here, a, b, m are real numbers. In this formula, b and a are the upper and lower values of the support of A, respectively, and m is the median value of A(Negnevitsky & Intelligence, 2005).

9.1. Description of Fuzzy Input Variables

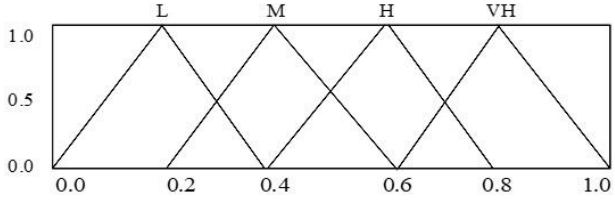

Fig.9. Membership Function Distributions for the Variables

The input fuzzy variables are X1 = {Social Media Contacts}, X2= {Social Media Common Contacts}, X3 = {Common Activity in Social Media}, X4 = {Times of Email Communication} and each of them was represented using four linguistic term Low (L), Medium (M), High(H), Very High (VH). The linguistic term and their ranges are shown in the Figure 9.

Table 6 - Linguistic term and their range

Linguistic terms	Membership Function	Range of parameter
Low (L)	Trimf	[0.0,0.4]
Medium (M)	Trimf	[0.2,0.6]
High (H)	Trimf	[0.4,0.8]
Very High (VH)	Trimf	[0.6, 1.0]

9.2. Description Fuzzy output variables

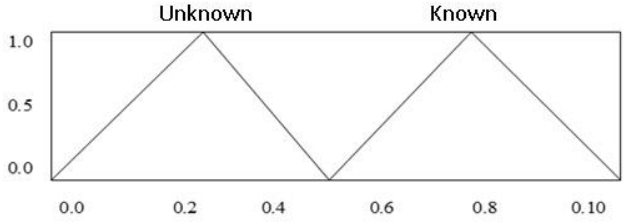

Figure 10: Membership Function Distributions for Output Fuzzy Variable

Linguistic term and their range for the variables are X_1 = {Social Media Contacts}, X_2= {Social Media Common Contacts}, X_3 = {Common Activity in Social Media}, X_4 = {Times of Email Communication} mentioned in Table 6.

Membership function distributions for output fuzzy variable areX_5 = {Known Phishing (1) / Unknown phishing (0)} which is crisp in nature as shown Figure 10.

9.3. Determining Fuzzy Rule Base from input and output variables

Rules are the core of the FRBS (Fuzzy Rule Base System) which represent the relationship between inputs and output. In the present problems four input variable were considered and each of them was represent using four linguistic terms. Thus, there could be maximum of rules in the FRBS. In this study, 256 fuzzy rules (4x4x4x4=256) are generated. For instance, the first and last rule as follow:

R1: If X1 is L AND X2 is L AND X3 is L AND X4 is L THEN output is Known Phishing

Similarly,

.

.

R256: If X1 is VH AND X2 is VH AND X3 is VH AND X4 is VH THEN output is Unknown Phishing.

9.4. Fuzzy Rules and Coding

Four input variables are in this section each have four linguistic terms which constitute 64 rules.

Table 7 - Description of fuzzy linguistic term.

Abbreviation	Expression	Index representation
L	Low	0.25
M	Medium	0.35
H	High	0.55
VH	Very High	0.85

Linguistic terms are denoted with their index value as listed in the Table 7.

10. Working Principle of Traditional FLC (Mamdani Approach)

An FLC consists of a set of rules presented in the form of IF (a set of conditions are satisfied) THEN (a set of consequences can be prepared). Here, antecedent is a condition in its application domain and the consequent is a control action for the system under control. Both the antecedents and consequents of the IF-THEN rules are represented using some linguistic terms. The inputs of fuzzy rule-based systems should be given by fuzzy sets, and therefore, we have to fuzzify the crisp inputs. Moreover, the output of an FLC is always a fuzzy set, and therefore, to get the corresponding crisp value, a method of defuzzification is to be used. The fuzzification of input variables involves the following steps:

1. Measure all the input variables.
2. Perform a scale mapping that transfers the range of values of inputs variables into corresponding universes of discourse.
3. Perform the function of fuzzification that converts input data to suitable linguistic values, which may be viewed as label of fuzzy sets.

The rule base comprises of knowledge of the application domain by using the information of data base. Thus, the data base provides necessary data to design the control rules involving linguistic terms. The rule base characterizes the control goals and policy of the domain experts by means of a set of linguistic control rules [120]. The Inference Engine of an FLC has the capability of simulating human decision-making based on fuzzy concepts and of inferring fuzzy control actions by employing fuzzy implication and the rules. A method of defuzzification is used to obtain the crisp value corresponding to the fuzzified output. In this study, Center of Sums (COS) method of defuzzification was utilized, which is given below:

$$U'_{f'} = \frac{\sum_{j=1}^{P} A(\alpha_j) \times f_j}{\sum_{j=1}^{P} A(\alpha_j)} \quad (2)$$

Where $U'_{f'}$ is the output of the controller, $A(\alpha_j)$ represents the firing area of j-th rule, p is the total number of fired rules and f_j represents the center of the area. Like fuzzy logic control based on Mamdani approach. We have also applied other method called Fuzzy Takagi-Sugeno Inference Engine (Back & LaPrade, 2019), (Negnevitsky & Intelligence, 2005). In this approach the outcomes of the variables depend upon a function of several input variables. A function may be linear or non-linear depend upon problem specification. We assumed all the features of URLs are dependent each other which is a nature

of complex task (non-linear) for getting output for function(Sahingoz et al., 2019). During experiment it is observed that the Mamdani FIS results were quite appropriate than Fuzzy Takagi-Sugeno Inference Engine.

11. Results and Discussion

The traditional fuzzy reasoning tools was developed using four inputs namely Social Media Contacts, Social Media Common Contacts, Common Activity in Social Media, Times of Email Communication and each having four different responses (i.e. Low, Medium, High, Very High) a set of 256 rules designed manually for analysis of the result. The result of this approach suggested that Social Media Contacts (X1) and Social Media Common Contacts (X2) are more influenced the output variable known and unknown person (S_id & R_id) refer Figure. 11 whereas other input variables are Common Activity in Social Media (X3), and Social Media Common Contacts (X2) are that much more influenced output variables as shown in Figure 12.

The outcome variables depend detection of phishing URL and Enterprise Environ. This shows that the Social media common contact and social medial common activity are most used Social facets used to launch phishing attack.

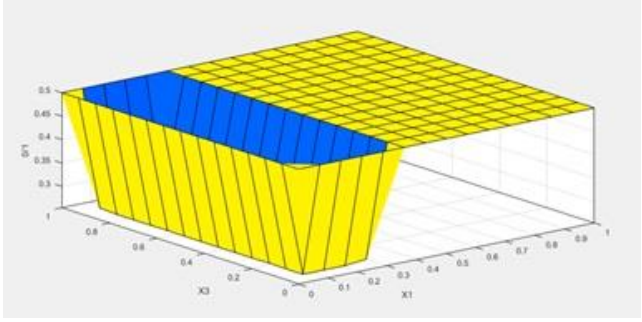

Fig.11. Social Media Contacts (X_1) vs Social Media Common Contacts (X_2)

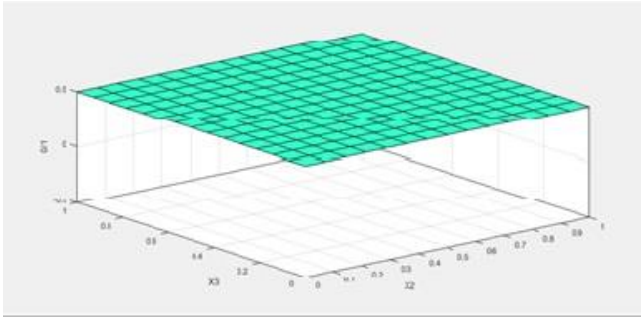

Fig.12. Social Media Common Contacts (X_2) vs Common Activity in Social Media (X_3)

12. Conclusion

Phishing is a catastrophic enterprise risk which relies on decentralized decision making of enterprise's employees. For example, enterprise environ security risk depends on quite an extent over the response and reversion to phishing attacks laid. The most alluring target for the phishers here is the enterprise employees or personnel. Employees Behavioral towards internet usage shows lay decision making, that is the reason, victims typically deviate from systematic measure of rationality and get into trap of phishers. As the phishers plot successful deceptions over employee peripheral route permission and shun direction logical argumentation. Further, manipulating splanchnic emotions i.e. fear, lust, greed, pity, anxiety urgency and acquainted contextual cues play a major role in persuasion or undermining the employee's confidence as well as trust to proceed towards the snares laid by the hoodwinkers. In this research, a novel anti-phish model is proposed and implementation of the proposed model is done in two phases, in first phase, performance of the model is measured using two different methods of ANN i.e. FBNN and LM neural network by training and testing the dataset. The results show the advantages of FBNN over LM neural network in terms of RMSE and R^2 value. Feed-Forward Backpropagation approach was able to yield better results as compared to other one. It might happen to the reason of each processing element of neural network properly trained with optimized weight. In the second phase, adoption of Mamdani FIS with four social features (Social Media Contacts, Social Media Common Contacts, Common Activity in Social Media, Frequency of email

communication) are used for detection of known and unknown e-mail sender in enterprise environ. Finally, in both the output of two phase, combined used as input to method to determine unknown or known phishing e-mail sender. The result of this approach suggested that Social Media Contacts (X_1) and Social Media Common Contacts (X_2) are more influenced the output variable known phishing attack whereas other input variables are Common Activity in Social Media (X_3), and Social Media Common Contacts (X_2) are that much more influenced output variables. The outcome variables depend detection of phishing URL and Enterprise Environ. This shows that the Social media common contact and social medial common activity are most used Social facets (features) used to launch email phishing

REFERENCES

[1] Adedoyin-Olowe, M., Gaber, M. M., & Stahl, F. (2013). A survey of data mining techniques for social media analysis. ArXiv Preprint ArXiv:1312.4617.

[2] Alexa Topsites datasource. (n.d.). https://www.alexa.com/topsites

[3] Almomani, A., Gupta, B., Atawneh, S., Meulenberg, A., & Almomani, E. (2013). A survey of phishing email filtering techniques. IEEE Communications Surveys & Tutorials, 15(4), 2070–2090.

[4] Al-Momani, A., Wan, T.-C., Al-Saedi, K., Altaher, A., Ramadass, S., Manasrah, A., Melhim, L., & Anbar, M. (2011). An online model on evolving phishing e-mail detection and classification method. Journal of Applied Science, 11(18), 3301–3307.

[5] Almomani, A., Wan, T.-C., Altaher, A., Manasrah, A., ALmomani, E., Anbar, M., ALomari, E., & Ramadass, S. (2012). Evolving fuzzy neural network for phishing emails detection. Journal of Computer Science, 8(7), 1099.

[6] Askysoftware Development Pvt. Online available https//:www.askysofware.com. (n.d.). https//:www.askysofware.com.

[7] Back, S., & LaPrade, J. (2019). The Future of Cybercrime Prevention Strategies: Human Factors and A Holistic Approach to Cyber Intelligence. International Journal of Cybersecurity Intelligence & Cybercrime, 2(2), 1–4.

[8] Baeza-Yates, R., & Davis, E. (2004). Web page ranking using link attributes. Proceedings of the 13th International World Wide Web Conference on Alternate Track Papers & Posters, 328–329.

[9] Bedingfield Sr, J. C., & Gehl, J. M. (2012). Substitute uniform resource locator (URL) generation. Google Patents.

[10] Berners-Lee, T., Masinter, L., McCahill, M., & others. (1994). Uniform resource locators (URL).

[11] Buckley, J. J., & Hayashi, Y. (1994). Fuzzy neural networks: A survey. Fuzzy Sets and Systems, 66(1), 1–13.

[12] Center, R. A.-F. C. (2012). RSA monthly online fraud report, May 2012.

[13] Chandrasekaran, M., Chinchani, R., & Upadhyaya, S. (2006). PHONEY: Mimicking User Response to Detect Phishing Attacks. 2006 International Symposium on a World of Wireless, Mobile and Multimedia Networks(WoWMoM'06), 668–672. https://doi.org/10.1109/WOWMOM.2006.87

[14] Chandrasekaran, Madhusudhanan, Narayanan, K., & Upadhyaya, S. (2006). Phishing email detection based on structural properties. NYS Cyber Security Conference, 3.

[15] Cialdini, R. B. (2001). The science of persuasion. Scientific American, 284(2), 76–81.

[16] Damerau, F. J. (1964). A technique for computer detection and correction of spelling errors. Communications of the ACM, 7(3), 171–176.

[17] Fette, I., Sadeh, N., & Tomasic, A. (2007). Learning to detect phishing emails. Proceedings of the 16th International Conference on World Wide Web, 649–656.

[18] Firake, S. M., Soni, P., & Meshram, B. (2011). Tool for Prevention and Detection of Phishing E-Mail Attacks. International Conference on Network Security and Applications, 78–88.

[19] Gansterer, W. N., & Pölz, D. (2009). E-mail classification for phishing defense. European Conference on Information Retrieval, 449–460.

[20] Greg Aaron, & Rod Rasmussen. (n.d.). Anti-Phishing Working Group. Global phishing survey, trends and domain name use in 2H2011. (Global Phishing Survey, Trends and Domain Name Use in 2H2011.) [Research, Analysis Support, and Graphics]. https://docs.apwg.org/reports/APWG_GlobalPhishingSurvey_2H2011.pdf

[21] Gupta, B. B., Tewari, A., Jain, A. K., & Agrawal, D. P. (2017). Fighting against phishing attacks: State of the art and future challenges. Neural Computing and Applications, 28(12), 3629–3654.

[22] Herzberg, A. (2009). DNS-based email sender authentication mechanisms: A critical review. Computers & Security, 28(8), 731–742.

[23] Jang, J.-S. R., Sun, C.-T., & Mizutani, E. (1997). Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence [Book Review]. IEEE Transactions on Automatic Control, 42(10), 1482–1484.

[24] Jayakanthan, N., Ramani, A., & Ravichandran, M. (2017). Two phase classification model to detect malicious URLs. International Journal of Applied Engineering Research, 12(9), 1893–1898.

[25] Kaivanto, K. (2014). The effect of decentralized behavioral decision making on system-level risk. Risk Analysis, 34(12), 2121–2142.

[26] Keshtkar, F., & Inkpen, D. (2009). Using sentiment orientation features for mood classification in blogs. 2009 International Conference on Natural Language Processing and Knowledge Engineering, 1–6.

[27] Kim, Y., Hsu, S.-H., & de Zúñiga, H. G. (2013). Influence of social media use on discussion network heterogeneity and civic engagement: The moderating role of personality traits. Journal of Communication, 63(3), 498–516.

[28] Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. Emerging Artificial Intelligence Applications in Computer Engineering, 160, 3–24.

[29] Krombholz, K., Hobel, H., Huber, M., & Weippl, E. (2015). Advanced social engineering attacks. Journal of Information Security and Applications, 22, 113–122. https://doi.org/10.1016/j.jisa.2014.09.005

[30] Lee, J.-L., Kim, D., Changhoon, & Lee. (2015). Heuristic-based Approach for Phishing Site Detection Using URL Features.

[31] Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. Soviet Physics Doklady, 10(8), 707–710.

[32] Malaysia, U. Ы. (n.d.). An Enhanced Online Phishing E-Mail Detection Framework Based On Evolving Connectionist System.

[33] Mamdani, E., & Assilian, S. (1999). An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Human-Computer Studies, 51(2), 135–147.

[34] McGrath, D. K., & Gupta, M. (2008). Behind Phishing: An Examination of Phisher Modi Operandi. LEET, 8, 4.

[35] Medvet, E., Kirda, E., & Kruegel, C. (2008). Visual-similarity-based phishing detection. Proceedings of the 4th International Conference on Security and Privacy in Communication Netowrks - SecureComm '08, 1. https://doi.org/10.1145/1460877.1460905

[36] Mitchell, R., Michalski, J., & Carbonell, T. (2013). An artificial intelligence approach. Springer.

[37] Mohammad, R., McCluskey, T., & Thabtah, F. A. (2013). Predicting phishing websites using neural network trained with back-propagation.

[38] Nauck, D., & Kruse, R. (1996). Neuro-fuzzy systems research and applications outside of Japan. Fuzzy-Neural Networks, Soft Computing Series, 108–134.

[39] Negnevitsky, M., & Intelligence, A. (2005). A guide to intelligent systems. Artificial Intelligence, 2nd Edition, Pearson Education.

[40] Netcraft Anti-Phishing Toolbar,. (2004). [Security, Netcraft Services]. Netcraft. https://news.netcraft.com/archives/2004/12/28/netcraft_antiphishing_toolbar_available_for_download.html

[41] Nguyen, L. A. T., To, B. L., Nguyen, H. K., & Nguyen, M. H. (2014). A novel approach for phishing detection using URL-based heuristic. 2014 International Conference on Computing, Management and Telecommunications (ComManTel), 298–303.

[42] Nomura, H., Hayashi, I., & Wakami, N. (1992). A learning method of fuzzy inference rules by descent method. [1992 Proceedings] IEEE International Conference on Fuzzy Systems, 203–210.

[43] Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The pagerank citation ranking: Bringing order to the web. Stanford InfoLab.

[44] Phelps, T. A., & Wilensky, R. (2000). Robust hyperlinks and locations. D-Lib Magazine, 6(7/8), 1082–9873.

[45] Phishing Tank. (n.d.). https://www.phishingtank.com

[46] Sahingoz, O. K., Buber, E., Demir, O., & Diri, B. (2019). Machine learning based phishing detection from URLs. Expert Systems with Applications, 117, 345–357.

[47] Sankhwar, S., & Pandey, D. (2017). A Comparative Analysis of Anti-Phishing Mechanisms: Email Phishing. International Journal of Advanced Research in Computer Science, 8(3).

[48] Sankhwar, S., Pandey, D., & Khan, R. (2019). Email Phishing: An Enhanced Classification Model to Detect Malicious URLs. EAI Endorsed Transactions on Scalable Information Systems, 6(21).

[49] Shah, R., Trevathan, J., Read, W., & Ghodosi, H. (2009). A proactive approach to preventing phishing attacks using Pshark. 2009 Sixth International Conference on Information Technology: New Generations, 915–921.

[50] Smadi, S., Aslam, N., Zhang, L., Alasem, R., & Hossain, M. (2015). Detection of phishing emails using data mining algorithms. 2015 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), 1–8.

[51] Social network analysis (4th edition). (2017). SAGE Publications.

[52] Stolfo, S., Hu, C.-W., Li, W.-J., Hershkop, S., Wang, K., & Nimeskern, O. (2003). Combining Behavior Models to Secure Email Systems. https://doi.org/10.7916/D8F47VVN

[53] Stringhini, G., & Thonnard, O. (2015). That ain't you: Blocking spearphishing through behavioral modelling. International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment, 78–97.

[54] Suriya, R., Saravanan, K., & Thangavelu, A. (2009). An integrated approach to detect phishing mail attacks: A case study. Proceedings of the 2nd International Conference on Security of Information and Networks, 193–199.

[55] Tyler, J. R., Wilkinson, D. M., & Huberman, B. A. (2005). E-mail as spectroscopy: Automated discovery of community structure within organizations. The Information Society, 21(2), 143–153.

[56] Vandermeer, J. (2006). Seven Highly Successful Habits of Enterprise Email Managers: Ensuring that your employees' email usage is not putting your company at risk. Information Systems Security, 15(6), 64–75. https://doi.org/10.1080/10658980601051359

[57] Wakker, P. P. (2010). Prospect theory: For risk and ambiguity. Cambridge university press.

[58] Wang, M.-F., Tsai, M.-F., Jheng, S.-L., & Tang, C.-H. (2012). Social feature-based enterprise email classification without examining email contents. Journal of Network and Computer Applications, 35(2), 770–777.

[59] Xiang, G., Hong, J., Rose, C. P., & Cranor, L. (2011). CANTINA+: A Feature-Rich Machine Learning Framework for Detecting Phishing Web Sites. ACM Transactions on Information and System Security, 14(2), 1–28. https://doi.org/10.1145/2019599.2019606

[60] Yearwood, J., Mammadov, M., & Webb, D. (2012). Profiling phishing activity based on hyperlinks extracted from phishing emails. Social Network Analysis and Mining, 2(1), 5–16.

[61] Zhang, J., Porras, P. A., & Ullrich, J. (2008). Highly Predictive Blacklisting. USENIX Security Symposium, 107–122.

ISSN: 2005-4262 IJGDC