Phishing Website Classification using Least Square Twin Support Vector Machine

Mayank Arya Chandra, S S Bedi, Shashank Chandra, Suhail Javed Quraishi

Abstract-- Phishing is one among the luring procedures used by phishing attackers in the means to abuse the personal details of clients. Phishing is earnest cyber security issue that includes facsimileing legitimate website to apostatize online users so as to purloin their personal information. Phishing can be viewed as special type of classification problem where the classifier is built from substantial number of website's features. It is required to identify the best features for improving classifiers accuracy. This study, highlights on the important features of websites that are used to classify the phishing website and form the legitimate ones by presenting a scheme Decision Tree Least Square Twin Support Vector Machine (DT-LST-SVM) for the classification of phishing website. UCI public domain benchmark website phishing dataset was used to conduct the experiment on the proposed classifier with different kernel function and calculate the classification accuracy of the classifiers. Computational results show that DT-LST-SVM scheme yield the better classification accuracy with phishing websites classification dataset.

Keywords: Least square, SVM, Twin SVM, Phishing, kernel, Classification, Machine Learning.

I. INTRODUCTION

Phishing is an illegal activity of collecting the sensitive information of people through fake website. In recent years, the analysis of phishing websites has attracted much interest. There is a dramatic increment in the quantity of phishing website since most recent couple of years. However internet users are wary of such phishing attacks, and a large portions of them move toward becoming casualties to these attacks.

A generic phishing site page may imitate trusted third party, such as a bank or e-commerce object etc. that convinces the clients to disclose their confidential data. The main objective of phishing is to collect the credential and personal information like as password, pin, and account detail by impersonating a legitimate entity in the cyber space [1]. There are constantly inventive ways that made consistently by phishing attackers to confuse the anti-phishing schemes. Subsequently, sustainable demands are essential to come up with keenly intellective anti-phishing strategies that are predicated on web mining and machine learning [2].

Phishing recognition is treated as special types of classification problem in data mining context [3].

The main objective is to predict the class of website either it is “legitimate” or “phishy” or “suspicious” on the basis of input data set, which is refer as training data set. The training data set contains the information of different features of websites with target class [4].

It is important that these website features must contain enough information to make accurate prediction. A number of traditional methodologies are utilized to categorize the phishing website, these methodology are not so much efficient to tackle the nascent and emerging patterns. This incentivized numerous researchers into searching for other efficacious strategies that can deal with existing and emerging extortion, inspired to the discovery of machine learning schemes.

Machine learning, a well-known word in researcher community, is a field of artificial intelligence that utilize the strategy for data mining to build and study of systems that can learn from datasets and retrieve the new or existing patterns (or features) from a data which can then be used to resolve the classification area[5]. In this paper, the general way to deal with exhibit novel machine learning strategy is utilized to classification phishing sites. The strategy depends on SVM which is prominent terms in the field of classification [6]. Here the scheme of decision tree least square twin support vector machine (DT-LST-SVM) is presented for classification of phishing website.

TABLE 1. DATA USED FOR TESTING

Total Sample	1353 websites
Total Legitimate Websites	548 websites
Total Phishing Websites	702 websites
Total Suspicious Websites	103 websites

In this work, the UCI public domain benchmark data set of website phishing consisting of a set of 10 renowned phishing features from a dataset consisting of 1353 websites was used. In this paper designed classifier trained by these features is presented in Table 1 and machine learning scheme is described in detailed. Section 2 gives Prelude of phishing and SVM schemes. Section 3 Explore the website phishing data set and respective features. Section 4 presents the proposed classifier details on the machine learning scheme. In Section 5 investigates the performance of proposed classifier and compare with different machine learning schemes. Section 6 gives concluding remarks and future perspective.

II. PRELUDE OF PHISHING AND SVM TECHNIQUES

This section discusses the concept of phishing and SVM and their extended schemes. These schemes have been explained briefly as under.

Revised Manuscript Received on November 30, 2019.
***** Correspondence Author

Mayank Arya Chandra*, MJP Rohilkhand University Bareilly, India
S S Bedi, MJP Rohilkhand University Bareilly, India
Shashank Chandra, Research Fellow, RAC, DRDO Delhi, India
Suhail Javed Quraishi, Invertis University, Bareilly, India

Retrieval Number: A3905119119/2019©BEIESP
DOI: 10.35940/ijitee.A3905.119119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

A. *Phishing*

Phishing is well known word used in social engineering to seduce clients, and exploit weaknesses in current web security [7]. Phishing websites is a semantic intrusion that attacks directly to the client rather computer and deceive users into giving up their own personal information [8].

Gartner is an American leading information technology research and advisory firm, published a report that exposed that phishing attacks continue to escalate in near future [9]. According to Gartner published report "Estimated cost of theft via phishing activities is approximately $2.8 billion annually as per the US banks and credit card issuers. According to the cisco's security group, personalised and focused on assaults that attention on accessing corporate bank accounts and valuable confidential sensitive data, intellectual property are on the embarkation [10].

Media that is mostly targeted by attacker is explained in table 2. Email and website are the main venerable media which is mainly targeted by hackers. Phishing Activity Trends Report 2018 and 2019 published by The Anti-Phishing Working Group (APWG) Inc., describes the statistical highlights of different types of attacks on different media.

TABLE 2. MEDIA THAT IS MOSTLY TARGETED BY ATTACKER

Email	Send a mail for updating information, whenever user clicks over link it leads the phishing website.
Website	Most phishing strategies use a special misleading scheme designed to make hyperlinks in emails (the spoofed sites they suggest) appear to belong to fraudulent organizations.
URL	Phishing attacks are usually accomplished through suspicious URLs
Social Network	Social media Phishing on Instagram, Facebook, Twitter have witnessed a rapid growth of phishing attacks for number of reasons. 1. Easily imitate personal information, 2 User's ready to click
Blog	Forged e-cards, online job scams and donation scams are some examples of phishing attacks against blogs and forums. For example, use an online job scam to collect the credentials of a job seeker.

TABLE 3 (A) APWG PHISHING ACTIVITY TRENDS REPORT STATISTICAL ANALYSIS FOR 2018 QUARTERLY (Q) [11, 12, 13, 14]

Attribute	1st Q	2nd Q	3rd Q	4th Q
Number of Unique Phishing Websites Detected	2,63,538	2,33,040	1,51,014	1,38,328
Number of Unique Phishing E-Mail	84, 444	2,64,483	2,70,557	2,39,910
Number of Brands Targeted By Phishing Campaigns	746	786	777	836

TABLE 3 (B) APWG PHISHING ACTIVITY TRENDS REPORT STATISTICAL ANALYSIS FOR 1ST QUARTER AND 2ND QUARTER 2019 MONTHLY[15,16]

Attribute	Jan	Feb	Mar	1st Q	April	May	June	2nd Q
Number of Unique Phishing Websites Detected	48,663	50,983	81,122	**180,768**	59,756	61,820	60,889	**182,465**
Number of Unique Phishing E-Mail	34,630	35,364	42,399	**112,393**	37,054	40,177	34,932	**112,163**
Number of Brands Targeted By Phishing Campaigns	327	288	330	**945**	341	308	289	**938**

The, APWG, the global industry unifies the global response to cyber-crime through development of data resources. APWG published Phishing Activity Trends Report (PATR) quarterly/Half-yearly which explores the phishing attacks reported by its research partner and business organization. Then APWG analyse the type of attacks, measures the evolution advancement, expansion, and spread of crime-ware by drawing from the examination of our partner industries based on different parameterAs indicated by APWG Phishing Activity Trends Report, 4th Quarter 2018 [14] [See table 3(a)].

- The approximately 138,328 phishing attacks were launched in 2018. 58 % of phishing sites used SSL certificates, a significant increase from the previous quarter in which 46% SSL certificates were used [15].
- In Brazil Attackers launched both conventional phishing scheme and social media attacks to cheat the internet users.

In Brazil hackers used both conventional phishing scheme and social media attacks to cheat the internet customers. They also use specialized traps to make it harder for user to avoid theses scams.

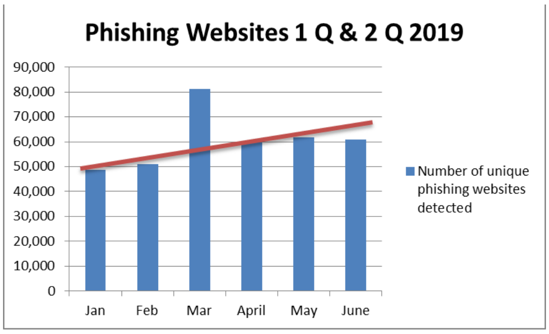

FIG.1 PHISHING WEBSITES 1 Q & 2 Q 2019

According to APWG Phishing Activity Trends Report 1st Quarter 2019 that phishing attack occurred most frequently in the payment service about 27% financial sector about 16% , next to SAAS/webmail (36%) and e-commerce (3%) telecom (3%) [15], As the Table 3(b) indicated total no of phishing sites at 2nd Q is 182,465, up slightly than 1st Q 180,768 as represented in fig.1 .

B. Support Vector Machine

Mid- 1990's Support Vector Machine comes in a picture and introduced by Vapnik et.al. Support Vector Machine is supervised learning scheme [17]. Linear as well as non-linear data can be classified by SVM. SVM transforms an input data into higher dimensional and constructs a hyper-plane which classifies the input set.

For the classifying data into higher dimensional SVM finds the vector point known as a support vector (S V), to represent the decision boundary plane and gives the maximal marginal distance between the different classes [18]. This large marginal separation refers as maximum marginal distance in decision space (shown in fig.2). The main feature of SVM is that it separates different classes with maximum marginal distance [18] [19].

These support vectors play crucial role in training phase of SVM. Basic SVM shows the outstanding performance for binary classification problems. If input data represent by x_i has $m * n$ dimension and their corresponding output classes is $Y_i \in \{-1, +1\}$. According to Vapnik's theory, decision boundary expressed as a linear equation (1). This decision boundary refers as a hyper-plane.

$$w^T \varphi(x) + b = 0 \qquad (1)$$

For binary classification following inequality satisfied by the processed data and built a hyper plane for both classes on the basis of given condition

$$w^T \varphi(x_i) + b \geq +1 \quad Y_i = +1 \qquad (2)$$

$$w^T \varphi(x_i) + b \leq -1 \quad Y_i = -1 \qquad (3)$$

Equation (2) and (3) is equivalent to

$$Y_i(w^T \varphi(x_i) + b) \geq +1 \quad i = 1 \ldots\ldots N \qquad (4)$$

The role of activation function φ (•) is transform input data set into higher dimensional feature space. Some situation may be arise where such hyper-plane cannot be built in higher dimensional space, at that point present another variable which is non-negative known as a slack variable (scalar variable) ξ_i.

$$Y_i(w^T \varphi(x_i) + b) \geq 1 - \xi_i \qquad (5)$$

$$\& \qquad \xi_i \geq 0 \qquad (6)$$

C. Twin SVM

Twin SVM is emerging scheme for classification problem. The main functionality of twin SVM is that it constructs two non-parallel hyper-planes in higher dimensional space as compare to one hyper-plane as conventional SVM. Twin SVM is used the concept of two hyper-plane for classifying the given data set. TW-SVM finds the solution of pair of quadratic programming problems (QPP's) rather than single one [20] [21]. The SVM has $O(n^3)$. Computational complexity for given 'n' size input data. If data size is $n/2$ then running time complexity of TW-SVM is $O(2 * (n/2)^3)$. So the runtime ratio is approximately

$$\frac{n^3}{2\,(^n/_2)^3} = 4 \qquad (7)$$

Time complexity of Twin SVM is $^1/_4$ original SVM as per the equation (7). The objective of Twin SVM is to build two non-parallel hyper-planes for each category by solving a couple of quadratic programming problems (QPP's) so that each hyperplane is closer categories of dataset, while away from another dataset. The working of Twin SVM is, let's separate the two categories (Category -1 and Category 1) with two non-parallel hyperplanes such that each hyperplane is closer to the dataset of one category and further away from the other.

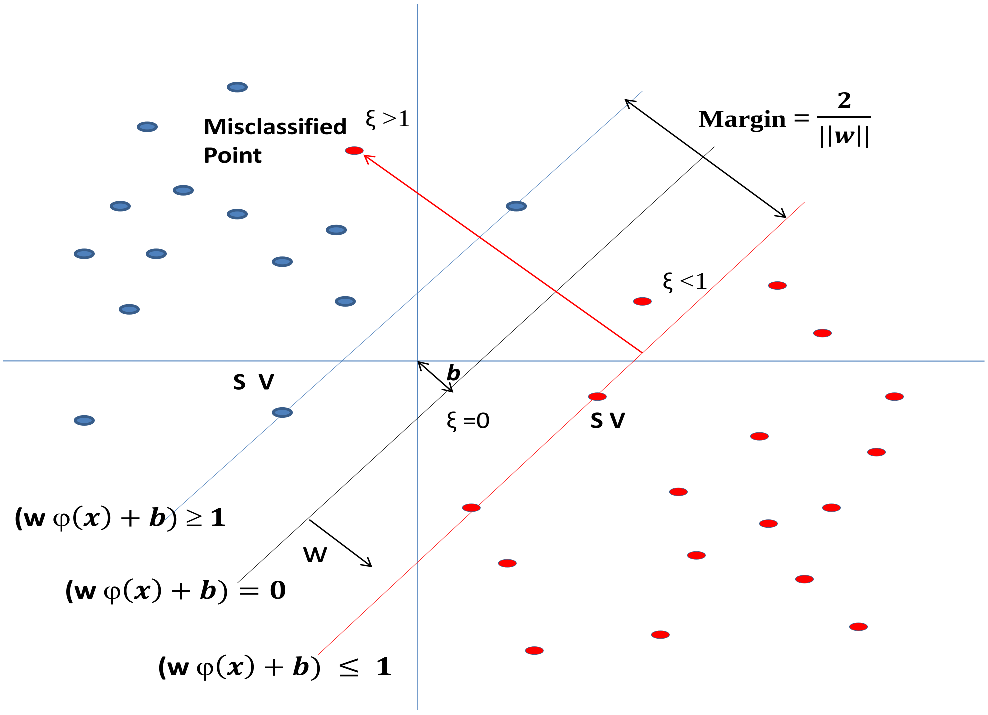
Misclassified Point
ξ >1
Margin = $\frac{2}{||w||}$
ξ <1
b
S V
ξ =0
S V
$(w\ \varphi(x) + b) \geq 1$
W
$(w\ \varphi(x) + b) = 0$
$(w\ \varphi(x) + b) \leq 1$

FIG.2 FUNDAMENTAL OF SVM [6]

Mathematical representation of the Twin SVM is represented by linear equation (8) and (9), which shows the two hyper-planes.

$$x^T w_1 + b_1 = 0 \qquad (8)$$

$$x^T w_2 + b_2 = 0 \qquad (9)$$

The TW-SVM classifier is achieve by resolving the given couple of quadratic programming problems (QPP)

Twin SVM-1

$$\min_{w_1\, b_1\, \xi_1} \frac{1}{2}(Aw_1 + e_1 b_1)^T (Aw_1 + e_1 b_1) + c_1 e_2^T \xi_1$$

$$Subject \;\; -(Bw_1 + e_2 b_1) + \xi_1 \geq e_2, \qquad \xi_1 \geq 0$$

$$(10)$$

Twin SVM-2

$$\min_{w_2\, b_2\, \xi_2} \frac{1}{2}(Bw_2 + e_2 b_2)^T (Bw_2 + e_2 b_2) + c_2 e_1^T \xi_2$$

$$Subject \;\; (Aw_2 + e_1 b_2) + \xi_2 \geq e_1 \qquad \xi_2 \geq 0$$

$$(11)$$

where e_1, e_2 are vectors parameters of arbitrary length, c_1 and c_2, are penalty parameters [22].

The scheme discovers the two hyper-planes for every class to classify data such that hyper-plane nearest to the data sample. The main working procedure of twin SVM is that it calculates the sum of squared distance from the points of one class (class 1) to hyper-plane. Applying optimization concept which tends to that hyper-plane nearer to one class with respect to other class and the calculated distance should be the smallest distance from the points of the other class (class -1) [22]. One error variable introduces for calculating the error. The error must be minimize because it reduce the chances of mis-classification.

D. LST-SVM

Least Square Twin Support Vector Machine (LST-SVM) is extremely fast scheme and exhibits improved generalization performance of classification. LST-SVM construct a two hyper-plane by using two linear equations as compare to solve the pair of complex QPP's [23][24][25][26][27].

$$\min_{w_1\, b_1\, \xi_1} \frac{1}{2}(Aw_1 + eb_1)^T (Aw_1 + eb_1) + \frac{c_1}{2} {\xi_1}^T \xi_1$$

$$Subject \;\; -(Bw_1 + eb_1) + \xi_1 = e, \qquad \xi_1 \geq 0$$

$$(12)$$

$$\min_{w_2 b_2\, \xi_2} \frac{1}{2}(Aw_2 + eb_2)^T (Aw_2 + eb_2) + \frac{c_2}{2} {\xi_2}^T \xi_2$$

$$Subject \;\; -(Bw_1 + eb_1) + \xi_2 = e, \qquad \xi_2 \geq 0$$

$$(13)$$

The equation (12) and (13) can be resolved by

$$\begin{bmatrix} w_1 \\ b_1 \end{bmatrix} = -\left(B^T B + \frac{1}{c_1} A^T A\right)^{-1} B^T e \qquad (14)$$

$$\begin{bmatrix} w_2 \\ b_2 \end{bmatrix} = -\left(A^T A + \frac{1}{c_2} B^T B\right)^{-1} A^T e \qquad (15)$$

Calculated decision function from these equations (14) and (15) is

$$\text{Class N} = \text{argmin}_{(i-1,\ 2)} \frac{|w_i^T x + b_i|}{\|w_i\|} \qquad (16)$$

III. WEBSITE PHISHING DATA SET

"Website Phishing Data Set" is taken from "University of California, Irvine, Machine Learning Repository: Website Phishing Data Set," 2016. That is used in this work [28]. The main resource of this data set is: Phish Tank archive. This paper describes the important features that have been proven to be reliable and effective when classifying the phishing websites.

The dataset have 3 types of instances classes as Legitimate, Phishy & Suspicious. Table 4 represents the numeric value respectively to their instance class values. The legitimate websites data were taken from Yahoo using PHP script plugged into browser and collect 548 legitimate websites from 1353 websites. A total of 702 phishing URL's and 103 suspicious URLs were calculated.

TABLE 4. FEATURES CATEGORY NUMERICAL VALUES

Features Category	Numerical Values
Legitimate	1
Phishy	2
Suspicious	3

Phishing dataset has the following 10 attributes.

1. "SFH"
2. "Pop-Up-Window"
3. "SSL-Final-State"
4. "Request URL"
5. "URL Anchor"
6. "Web Traffic"
7. "URL Length"
8. "Age of Domain"
9. "IP"
10. "Result Class"

IV. PROPOSED SCHEME

In this paper, we employed a scheme of training classifiers that classify phishing dataset. LST-SVM with Decision Tree scheme with different kernel functions is proposed to classify the phishing website. The results show that linear kernel function has highest accuracy. For optimize the given classifier performance we used different penalty parameters and also observe the impact of penalty parameter over classification accuracy. For calculating the best penalty parameters for the given model, we employ thumb rule.

DT-LST-SVM scheme can generate binary classifier of i^{th} & j^{th} classes and construction is started from root node and data set movement is decided by the decision function values. Process repeated until reach the leaf node and this is actual class of given data set. The website phishing data set S, divided into two subsets, namely training and test sets, which have similar sizes (k) and similar category distributions. It generates N binary classifier by solving N linear equations.

1. $for\ i = 1\ to\ N, \quad j = i + 1\ to\ N$

2. $If\ Class\ is\ i^{th}\ Trained\ with\ j^{th}\ Class$
3. $Lets\ Two\ Matrix\ \ A_{ij}\ \&\ \ B_{ij}$
4. $Arrange\ penalty\ parameter\ c_1, c_2 > 0$

$$\min_{w_1\, b_1\, \xi_1} \quad \frac{1}{2}(Aw_1 + eb_1)^T\,(Aw_1 + eb_1) + \frac{c_1}{2}\xi_1{}^T\xi_1$$

$$Subject\ -(Bw_1 + eb_1) + \ \xi_1 = e, \qquad \xi_1 \geq 0$$

$$\min_{w_2 b_2\, \xi_2} \quad \frac{1}{2}(Aw_2 + eb_2)^T\,(Aw_2 + eb_2) + \frac{c_2}{2}\xi_2{}^T\xi_2$$

$$Subject\ -(Bw_1 + eb_1) + \ \xi_2 = e, \qquad \xi_2 \geq 0$$

5. Calculate Hyper-Plane Parameters According to Eqs

$$\begin{bmatrix} w_1 \\ b_1 \end{bmatrix} = -\left(B^T B + \frac{1}{c_1}A^T A\right)^{-1} B^T e$$

$$\begin{bmatrix} w_2 \\ b_2 \end{bmatrix} = -\left(A^T A + \frac{1}{c_2}B^T B\right)^{-1} A^T e$$

and Create Hyper-Planes $h_{ij} and\ h_{ji}$.

6. Further Calculate Distance bd_{ij} based on Euclidean distance for Testing.

V. EXPERIMENTAL RESULT

In this Section, we conducted experiment on Website Phishing Data retrieved from UCI repository [28]. The experiment conducted by partitioning a data set into training and testing. To verify the efficiency of proposed scheme are performed 10 fold cross validation where The data set is divided into 10 different parts, 9 of which are used to train the classifier and 10th part is used for testing, which is done in 10 times. The predicting accuracy of the classifier depends upon the information obtained during the training phase. We tested our scheme on Website Phishing Data Set the result reported in table 5 and table 6. Select the optimal values of the penalty parameter $c_1 = c_2$, from $\{10^{-6}, \ldots 10^{6}\}$ and kernel parameter σ from$\{\ 2^{-8} \ldots\ 2^{8}\}$. The comparison chart of the different classifiers predicted accuracy depicted in fig.3.

TABLE 5. AVERAGE ACCURACY OF DIFFERENT KERNEL.

Data Set	Linear Kernel Accuracy (%)	Polynomial. Kernel Accuracy (%)	RBF Kernel Accuracy (%)
"UCI-Website Phishing Data Set-2016"	87.69	82.81	80.30

The basic SVM Vapnik's classifier is used to decide if the site is phishy or not. The classification accuracy of Vapnik's SVM classifier was 84%, which is considered very low as compare to other schemes [29]. Classification accuracy of DT-LST-SVM on different kernel function shows in table 5. It is seems that proposed scheme with linear kernel function gives the higher predicted accuracy over polynomial kernel and RBF kernel function.

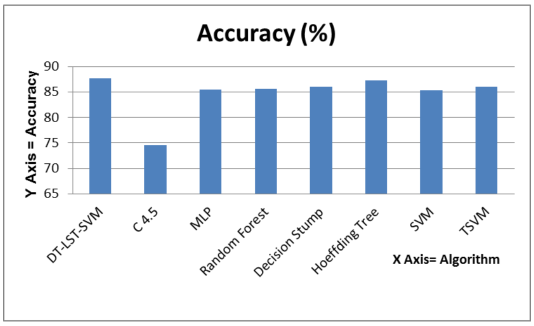

FIG.3 AVERAGE ACCURACY CHART

Performance comparison of different classifiers on Phishing Datasets with proposed scheme and similar work in literature shows in table 6 and result shows that DT-LST-SVM has better accuracy than other schemes.

TABLE 6. PERFORMANCE COMPARISON OF DIFFERENT CLASSIFIERS ON PHISHING DATASETS.

Scheme	Accuracy (%)
DT-LST-SVM	**87.69**
C 4.5	74.6
MLP	85.5
Random Forest	85.7
Decision Stump	86.1
Hoeffding Tree	87.3
SVM	84.1
TSVM	86.3

Table 5 and 6 indicate the performance comparison of our and different other schemes. Our scheme gives the excellent performance with linear kernel function represented in table 5. As per the table 6 our scheme achieved higher prediction accuracy (87.69%) compared to other classifiers.

VI. CONCLUSION

Phishing has turned to be a genuine danger for worldwide security and economy. Rapidly changing increasingly sophisticated social-engineering attack and new phishing websites attacks have made it difficult to keep blacklists up to date. This scheme as the phishing website classification is a machine learning scheme for classification of given data set whether the given website is legitimate, phishy or suspicious.

Therefore, the presented a DT-LST-SVM Scheme produced high predictive classification accuracy of 87.69 %. Applying this scheme will undoubtedly increase the performance of a classifier since classification of website totally dependent on the phishing features identified in the learning phase of the classification itself. We plan to improve this work in the near future by combining this scheme with a naturally inspired scheme. .

REFRENCES

1. Z. Ramzan and C. Wuest, "Phishing Attacks: Analyzing Trends in 2006," *CEAS 2207- 4th Conference on Email and Anti-spam*, Mountain View, California USA, 2007.
2. H. Zuhir, A.Selmat and M. Salleh, "The Effect of Feature Selection on Phish Website Detection an Empirical Study on Robust Feature Subset Selection for Effective Classification," *International Journal of Advanced Computer Science and Applications*, vol.6, no.10, pp. 221-232, 2015.
3. N. Abdelhamid, A. Ayesh and F. Thabtah, "Phishing Detection based Associative Classification Data Mining," *Expert System with Applications*, vol.41(13), pp.5948-5959, 2014.
4. M. Al-diabat, "Detection and Prediction of Phishing Websites using Classification Mining Techniques," *International Journal of Computer Applications*, vol.147, no.5, pp.5-11, 2016.
5. A. A. Akinyelu and A. O. Adewumi, "Classification of Phishing Email Using Random Forest Machine Learning Technique," *Journal of Applied Mathematics*, vol.2014, pp.6, 2014.
6. M. A. Chandra and S. S. Bedi, "Survey on SVM and their application in image classification," *International Journal of Information Technology*, 2018. https://doi.org/10.1007/s41870-017-0080-1.

7. A. Jøsang, B. AlFayyadh, T. Grandison, M. AlZomai and J. McNamara, "Security Usability Principles for Vulnerability Analysis and Risk Assessment," *Twenty-Third Annual Computer Security Applications Conference (ACSAC 2007)*, Miami Beach, FL, pp. 269-278, 2007.
8. A. Hodzic, J. Kevric and A. Karadag, “Comparison of Machine Learning Technique in Phishing Website Classification,” *International Conference of Economic and Social Studies (ICESoS)*, pp.249-256, Sarajevo, 2016.
9. Gartner, Incorporated. Available at, http://www.gartner.com/technology/home.jsp.
10. Lennon, M. Security Week. Available at, http://www.securityweek.com/cisco-targeted-attacks-cost-organizations-129-billion-annually, 2011.
11. "APWG Phishing Attack Trends Reports," 1st Quarter July 31, 2018.
12. "APWG Phishing Attack Trends Reports," 2nd Quarter October 18, 2018.
13. "APWG Phishing Attack Trends Reports," 3rd Quarter December 11, 2018.
14. "APWG Phishing Attack Trends Reports," 4th Quarter Published March 4, 2019
15. "APWG Phishing Attack Trends Reports" 1st Quarter May 15, 2019.
16. "APWG Phishing Attack Trends Reports" 2nd Quarter Sept 12, 2019
17. C. Cortes and V. Vapnik, “Support Vector Network,” *Machine Learning*, vol. 20 (3), pp.273-297, 1995.
18. I. Guyon, J. Weston, S. Barnhill and V. Vapnik, “Gene Selection for Cancer Classification using Support Vector Machines,” *Machine Learning*, vol. 46 (1-3), pp.389-422, 2002.
19. G. L. Prajapti and A. Patle, “On performing Classification using SVM with Radial Basis and Polynomial Kernel Functions,” *3rd International Conference of Emerging Trends in Engineering and Technology*, Goa, India, pp. 512-515, 2010.
20. J. A. K. Suykens and J. Vandewalle, “Least Square Support Vector Machine Classifiers,” *Neural Processing Letters*, vol. 9 (3), pp. 293-300, 1999.
21. J. A. K. Suykens, T. V. Gestel, J. D. Brabanter, B.D. Moor and J .Vandewalle, “ Least Square Support Vector Machines,” *World Scientific publishing Co.*, Singapore, ISBN 981-238-151-1, 2002.
22. G. Fung and O. L. Mangasarian, “Proximal Support Vector Machine Classifiers,” *Proceedings KDD 01, 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.77-86, San Francisco, California, 2001.
23. M. A. Kumar and M. Gopal, “Least Square Twin Support Vector Machines for Pattern Classification,” *Expert Systems with Applications*, vol. 36 (4), pp. 7535-7543, 2009.
24. M. A. Kumar and M. Gopal, “Least Squares Twin Support Vector Machines for Text Categorization,” *2015 39th National Systems Conference (NSC)*, pp.1-5, Noida, 2015.
25. M. A. Chandra and S. S. Bedi, “Survey on Support Vector Machine (SVM) and Their Application in Technical Diagnostics,” *JUET Research Journal of Science and Technology*, vol.4, no.1, 2017.
26. M. A. Chandra and S. S. Bedi, “Benchmarking Tree based Least Squares Twin Support Vector Machine Classifiers” *International Journal of Business Intelligence and Data Mining*, 2018, DOI: 10.1504/IJBIDM.2018.10009883.
27. M. A. Chandra and S. S. Bedi, “A Twin Support Vector Machine Based Approach to Classifying Human Skin,” 2018 *4th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, 2018, pp. 1-5.
28. UCI Machine Learning Repository: Phishing Websites Data Set, from https://archive.ics.uci.edu/ml/datasets/Phishing+Websites, Retrieved May 9, 2019.
29. R. M. Mohammad, F. Thabtah and L. McCluskey, “Intelligence Rule Based Phishing Website Classification,” *IET Information Secur*ity, vol. 8, no. 3, pp. 153-160, 2014.