TuijinJishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 4 (2023)

SmishSMS- The Latest Detection of SMS Phishing Trends

[1]Anisha Asirvatham,[2]Dr.C.Meenakshi

[1]Research Scholar Department of Computer Science,Vels Institute of Science Technology and Advanced Studies, Pallavaram, Chennai, India

[2]Department of Computer Applications, Vels Institute of Science Technology and Advanced Studies Pallavaram, Chennai, India

***Abstract*—** Banking frauds is an emergency issue of concern in today's busy schedule of the people around the world which has left many banking organizations bankrupt, and also has led many customers to a deadly torment. It is a cyber-security attack, where each day the fraudsters are inventing newer techniques which uses Short Message Service (SMS) to steal the credentials of mobile users. This paper has recklessly examined the e-banking frauds, the different methods for detection of e-banking frauds, control of banking frauds and the challenges associated with the detection and control of e-banking frauds. This paper focuses on Comparative analysis with different machine-learning techniques and the methods of the features extracted and classified. The paper also shows the further scope of study to boom the efficiency of the system.

Keywords*—SMS, ham, legitimate messages, spam messages,SmishSMS*

Introduction

This research review is done with few research papers to detect the Smishing among the SMS received by the users. Text message is the most used form of communication around the world as it is cheap and can be used as a medium without the use of internet. Text messages are not safe due to improper message filtering methods in recent times. This gives way to Spammers or Hackers to intrude the users and steal the confidential information of the users.The SMS received can be of different types as verbal, non-verbal and written. Among the given type of messages, the messages can be ham or spam messages. Spam messages are unwanted or unsolicited messages which are sent out in bulk. Ham messages are non-spam messages.

Literature Review

This content proposes a new concept for categorizing the SMS messages. Here, useful features are inspected first using two different features of selection techniques based on chi-square method and information gain method, and then SMS messages are classified along with these selected feature sets using Bayesian based classification algorithms. Based on the previously described Android application, a real-time SMS spam filtering mobile application is built that employs the suggested categorization method. As feature size rises, spam accuracy improves but genuine accuracy decreases. This is true for all feature selection and classification techniques used in this investigation. When all trial findings are taken into account, the binary classification model employing the top-10 CHI-SQUARE based features achieves the greatest overall accuracy of 90.17%. The effectiveness of the suggested plan is evaluated using a sizable sample of SMS messages, both valid and spam. The experimental analysis's findings plainlyshow that the suggested approach is quite good at identifying both spam and legal SMS messages. [1]

SNORT is a Network Intrusion Detection System that detects harmful network traffic activities. It is a popular Intrusion Detection System. Snort may operate in four modes: packet sniffer, packet logger, network-based intrusion detection system, and inline mode. Its key benefit is ability to recognize a variety of threats, including viruses, DoS attacks, malwares, ransom wares, spywares etc. The biggest disadvantage of this system is that it uses signature-based approaches to identify intrusions, which means that if aberrant behavior takes place, it

won't be picked up by SNORT. The use of SNORT IDS for signature detection will more effectively identify and trigger the right alert for signature-based attacks. Simultaneously, for Anomaly Detection approach, Anomaly detection based IDS would identify novel assaults accurately and produce proper alerts. [2]

Phishing is a big issue in the cyber-world, causing financial losses to the businesses and people. Visual similarities based approaches are advantageous for efficiently detecting phishing websites. Consumers are tricked on viewing the proper website where a phishing website appears strikingly similar to its matching real website. Visual similarity based phishing detection systems show appearances such as textual content, text format data, HTML tags, Cascading Style Sheet (CSS), image and so on to help them make a decision. These techniques use several attributes to compare the suspect website to the equivalent authentic website, and if the resemblance is exceedingly predetermined with a threshold value, it is termed as phishing. Users are exposed to malware through these phishing attacks that sends them to fake websites or proxy servers. When the user clicks on the fake hotlink, destructive software is installed on the user's system, gathering personal information from the system, and transferring it to the attacker. Attackers embed malware or hazardous links in the fake emails and sms. Potentially, attackers could gain remote access to the victim's computer and gather data at any time. This study focuses on social engineering methods since they are the most common way for phishing victims' information to be stolen. [3]

The accuracy of the SmiDCA model was tested using experiments on datasets in both English and non-English, and the results of both tests showed 96.40 percent accuracy in the English data and 90.33 percent accuracy in the non-English data. Additionally, even after nearly half of the characteristics were pruned, the model still had an accuracy of 96.16%. The experiment's findings indicate that the Random Forest classifier helped the model in BFSA to evaluate accuracy at 96.40%. The model reduced the features' of English-dataset dimension by 40.71 %, and AFSA assessed its correctness at 96.16%. Additionally, the effectiveness of both algorithms demonstrates that AFSA was more effective than BFSA. [4]

Text messages, personally identifiable information, banking information, credit card information, and passwords are all targets of the cybercrime known as phishing. Online identity theft includes scams like phishing. The victim's personal details and account information are obtained by the phisher through social engineering. This article gives a general overview of phishing attacks, the many kinds of phishing attempts that are made, and strategies for detection and defense. Deceptive phishing, spear phishing, clone phishing, whaling, link manipulation, and voice phishing are different forms of phishing. Avoiding spam, sharing sensitive information only over the phone or secure websites, not opening attachments or clicking on links in emails from unknown senders, following good security procedures, and receiving security awareness training can all assist to avoid phishing attempts. Phishing attacks are detected using custom DNS services, browser phishing lists, examining links in websites, utilizing personal ninja skills, seeking secure connections, verifying domain of URL, and inspecting the site itself. [5]

The 12 distinct unique qualities of hyperlinks in the authors' method proposed are utilized to train the machine learning algorithms. This phishing detection method works solely on the client side and doesn't need any assistance from a third party. It is able to recognize websites written in any textual language since it is language neutral. The proposed methodology has a reasonably high accuracy in identifying attacked websites when compared to previous approaches since it achieved more than 98.4% accuracy using logistic regression classifier. [6]

The goal of this paper is to design a method that could effectively locate junk mail messages with excessive accuracy. It works by identifying capabilities that may be used to categorize messages as junk mail or ham message. The study is accomplished with decision tree, Support vector machine, ANN with Back propagation and K-Nearest Neighbours in which the accuracy with ANN with Back propagation is 95.81% in contrast with different techniques considering the 14 functions with all of the algorithms mentioned. [7]

The search engine-based technical method uses a simple and reliable query to determine whether the suspected URL is legitimate. As some freshly formed legitimate web sites won't show up in the search engine, five criteria to the search engine-based technique is furnished to improve the detection accuracy. The suggested method can

also successfully categorize newly established legitimate websites that aren't classified using existing search engine-based techniques. Results of evaluations reveal that our method outperforms competing search-based techniques and achieves 98.15 % TPR and 0.05 % FPR. [8]

Demand Response (DR) is a key element to a sustainable and reliable grid. The success of DR initiatives is significantly influenced by consumer behavior. This research concentrates on how short messaging service (SMS) message responses from customers might interfere with demand response. The best notification method for Demand response programs is SMS. Mobile is the next frontier for cybercrime, according to assessments. Mobile devices are the target of 48 percent of all phishing assaults, and every year there are twice as many mobile phishing attacks. Additionally, 42% of users click on malicious URLs from a mobile device, according to Stanford University. [9]

This survey report generated at a large number of transactions from 2016 to 2019 that the Deep Neural Network learned and rated as valid or fraudulent. Deep learning is classified into several categories based on the data and methods utilized, including Deep Neural Network, Auto encoders, Convolution Neural Network, Generative Adversarial Network and Recurrent Neural Network. It is an unbalanced report. As a result, deep learning algorithms will be widely employed to detect fraudulent transactions in order to minimise financial losses in financial institutions. [10]

In the financial industry, temporal networks are employed to discover fraud detection systems. The cycle detection approach is utilised, which detects the transaction pattern using DFS. This approach is a valuable tool for detecting different sorts of transaction cycles in the network, and the parameterization may be customised. [11]

Smishing messages are extra dangerous as compared to unsolicited mail messages. The impact of smishing assault is economic loss and identification theft. There are numerous approaches to hit upon the unsolicited mail messages. There are no filters that separates the smishing message from the unsolicited mail message. This method indicates the technique to clear out smishing message from unsolicited mail message. There are levels of this method. The undesired mail and ham communications are filtered in the first part. The key element separates spam mail from smishing texts. The effectiveness of the suggested technique is assessed using a variety of expert system based classifiers and a dataset of spam and undesired messages. According to the simulation results, the suggested technique can identify unwanted mail with a precision value of 94.9 percent and can filter spam messages with a precision of 96 percent using a neural community classification algorithms. Naive Bayes algorithm, neural networks, decision trees, J48, random forests, and logistic regression were applied in this suggested approach to filter the message type. Analysis of the different URL in the message is done to see whether it takes the user to a malignant causing login page or prompts them to download the relative malicious software. Analysis of the URL present in the message can redirects the person to a malicious login web page or ends in the down load of a few malicious application. [12]

Smishing (SMS Phishing) is a text-messaging-based assault that targets smartphone users. The authenticity of the URL in the message is checked to distinguish spam messages from smishing communications. In this research, the short messaging service detection model is introduced with two phases: domain checking and SMS classification. The dataset is analyzed using three machine learning algorithms that is Decision Tree algorithm, Random Forest algorithm and Naive Bayes algorithm. The Random Forest Algorithm had a reasonable forecast accuracy of 97.85 %, while the Back propagation Algorithm provided the greatest prediction accuracy of 97.93 %. The Naive Bayes algorithm (NB) performed well as well, with an accuracy of 97.76 %. The Back Propagation Approach can detect smishing messages with a 97.93% accuracy. It is also found that the phone number and email id included in a message are not checked for their maliciousness in this study. [13]

This paper presents an efficient smishing detection system using ANNs where smishing can be very well illustrated using a problem with a binary classification. Text messages are used by attackers to target smartphone users via text messages and URLs included in the messages. Users who visit links attached to SMS typically redirect users to malicious websites and ultimately lead to downloading rogue applications that attempt to steal users' sensitive records. Smishing detectors are implemented using Naive Bayes classifiers, neural back

propagation networks, and decision tree classifiers. An accuracy of 97.40% is achieved by a neural network with a true positive rate value of 92.37% and a true negative rate value of 97.91%. [14]

This paper considers four distinct strategies, focusing on deep learning, ensemble machine learning techniques and aberration detection in particular. The trials showed that the mentioned strategies still have a strong ability in showing the results with supreme precision of 99.95% for artificial neural network, anomaly detection having accuracy as 99.78%, XGBoost algorithm with 99.62% and 99.69% accuracy using random forest algorithm. The main contributions of this study are tools for mobile network operators to improve operational security with respect to money transaction fraud by adopting machine learning strategies to secure the varied users from mobile money transactions. The drawback of this study is that just one deep learning strategy that is the artificial neural network was taken into account. To get a broader context and more information deduced, the dataset will be resampled in subsequent work using both under sampling and oversampling techniques. [15]

Proposed Work

We have done a small scale research after pandemic to know the frequency of the unknown SMS received from the mobile users, where most of the message are related to updation of PAN details,Update your KYC,etc. which shows to the users as the genuine messages from the banks. And once the link is clicked the user's confidential financial bank details are asked like the credit card or debit card details, expiry date, name of the card holder, and CVV (Card Verification Value), number. It just replicates the normal banking processes where once the users card details are added, the OTP (onetime password) is sent to the registered mobile number of the user and the whole transaction appears so genuine, that the mobile users share the confidential OTP also to the malicious contact and the consecutive transaction happens in few minutes, leading to the bankrupt status of the account. These group of the malicious hackers, they loot the common people with these kind of messages sent to the users and the savings of the users are being depleted in few minutes and few transactions, where they take the control of your mobile devices and the consecutive transactions are carried out.

This leads to a novel idea to the users where, the users can be made aware of these contact details, once the SMS is delivered, so that the users can be protected from such scam.

The proposed work SmishSMSshows the different steps of evaluation of the SMS received with different steps as illustrated in Fig1.

1. The SMS is first scrutinized with the text pre-processing methods like count vectorizer and Tfidf vectorizer to convert the text into numbers.
2. The URL pattern is checked for the genuineness of the domain name in the link provided with the SMS.
3. If, the URL is not present, then the source is checked for the phone number.
4. When the source is known, the received message is legitimate message, otherwise the message is checked for the misspelled words, leet words, symbols and special characters.
5. Furthermore, the URL pattern and the type of SMS is checked with the classification models and predict the actual source of the message, so as to know the origin of spam messages.

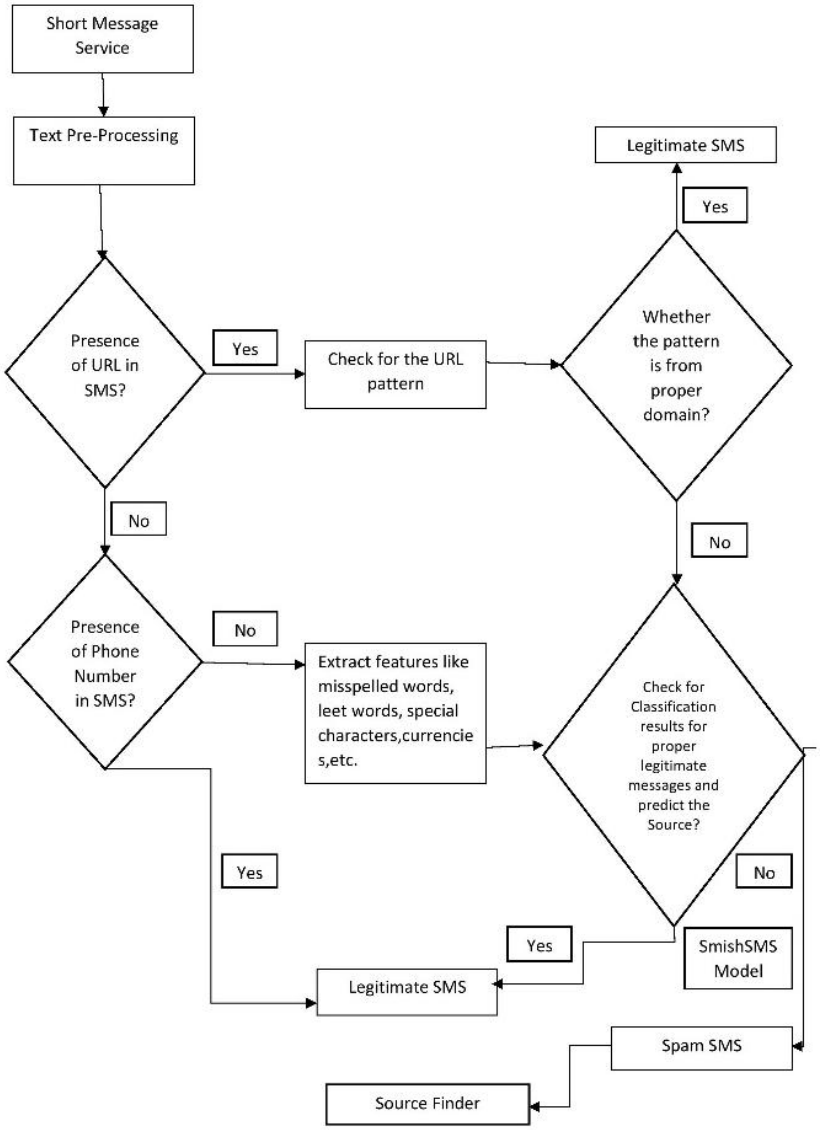

Figure.1 – Flowchart of the Proposed System

In fig.2, the SMS regarding the account being blocked is received from the unknown malicious phone number. A common mam, when after reading the message will get panic, and will tend to click on the abnormal link sent in the SMS. This will then lead to the revealing of personal information of the user to the spam message sender and will bankrupt the users account. This is a serious problem which needs to be made aware to the common man.

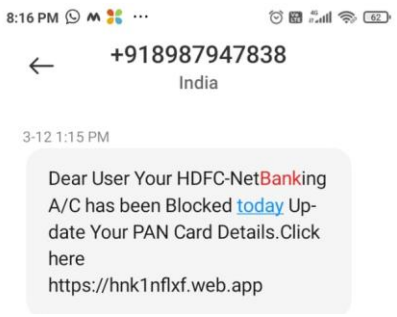

Figure.2- SmishSMS message showing bank scam

Fig.1 shows the flow of the analysis and the processes carried out to get the optimal results. The different ways performed are elaborated as follows:

Text Pre-Processing Phase: Text pre-processing is the phase where the users input is tokenized in the machine readable format (integers). The countVectoriser which helps to tokenize the collection of documents, creates its own vocabulary and then helps to split the sentences in words format. TfidfVectorizer is another vectoriser used to tokenize the documents with new set of vocabularies and also allows to encode the said vocabularies into a new document. The different pre-processing techniques,also includes, converting the text to lowercase, stop word removal means the common words in the text provided are eliminated.

Checking for URL pattern: Many URLs which are received in SMS message type are normally mis-spelled, like www.google.com can be written as www.go0gle.com. Both the URL given must open the official page of Google.com. But most of the time, the mobile users and elderly people in the family, tend to go through the URL once again. In the given example, one alphabet "O" is mistyped as "0(Zero)".

*Checking for the domain pattern:*In this process, it extracts the domain name, which follows the proper format with the web address followed by com,net,in,ac.in,org,etc. The anatomy, of the web address can be illustrated with the given example-https://www.example.com. The sent SMS will be can be checked for the proper anatomy of the web address. In the given example in fig 2, the web address given as https://hnk1.nflxf.web.app. In the given example, the web address follows the anatomy but the names given are not meaningful, where it shows irrelevant address with the combination of alphabets and numbers. This model finds the anatomy of the web addresses authenticity and responds for further investigations.

Check for the authenticate contact number: The next step checks for the valid contact number, from where the spam message is delivered to the user. The phone number is checked for the contact among the contact list trying to fool the user and also the other directory numbers. The proposed model tries to find the pattern of the contact number as genuine or not. If, the pattern is genuine and among the contact list, then the process id terminated. Otherwise, the model checks for the type of the contact number and informs the user about the genuinity of the contact.

Leet words: Leet words are the words which are spelled incorrectly. For example, Sorry can be written as Sry,Srry. The words can also be entered with a combination of alphabets and numbers. For example, hac0R which can be pronounced as hacker.

Special Character Check: Furthermore, in the SMS, there can be currency denominations, special characters like $, €, £, etc. along with the numeric values. These features are extracted and checked for the authentic messages through the model. If the source is reliable, the process stops, otherwise the further scrutiny of the message takes place and passed on to the proposed model to check for the spam messages. With the dataset collected, 28.62% of the SMS consists of special characters.

Keywords: There are few keywords like Congratulations, Winner,Offer, etc. These words are included in the SMS,so that it prompts the users to click on the message to read and follow the steps. In banking messages also,promotional discounts and offers for loans are used to make the users.

CRC (Cyclic Redundancy Check): CRC is the method to show the error check in the communication channel between the nodes with the graph G, having vertices V and Edges E, denoted as G (V.E). [16]

Source Finder: The path is found using Depth First Search (DFS) with the CRC algorithm and Modulo2 algorithm, where the error is calculated within the network generated using the remainder generated. When the remainder is 1, then the MSB is eliminated and the modulo is evaluated till the remainder is zero. [17]

Discussion And Result Evaluation

This paper has proposed a new idea to reduce the SMS phishing model as SmishSMS to detect the spam message. The foremost objective of this study is to segregate the messages as spam and ham messages and then to also know the source of the message with phone number and the IP address of the source. The proposed

model examined the SMS messages received, where 20 features are extracted. The different machine learning models were also deployed to get the maximum optimal results.

4.1 Data Collection: As per the Table 1, there are total 1001 SMS analyzed for Spam and Legitimate messages. There were 152 spam messages and 849 legitimate messages. With reference to the table data, the data was collected from different age group of mobile users.

Table 1: Text messages (Spam and Ham)

Total SMS	Spam Message	Legitimate SMS
1001	152	849

*Evaluation of Metrics:*The proposed methodology and the algorithm used for comparative study are evaluated using Accuracy Score, Recall Score, Precision Score and F1 Score. These metrics and evaluated in percentage values.

- True Positive (TP): True Positive denotes the total number of SMS messages identified as spam messages by the system.
- False Positive (FP): False Positive denotes the total number of SMS messages identified as legitimate or ham messages by the system.
- False Negative (FN): False Negative denotes the total number of legitimate or ham messages identified as spam by the system.
- True Negative (TN): True Negative denotes the total number of legitimate messages identified as legitimate or ham messages by the system.
- Accuracy: Accuracy is values calculated as the proportion of True Positive and True Negative over the total number of messages as given by the formula below:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision: Precision is calculated as follows:

$$\text{Precision} = \frac{TP}{TP + FP}$$

- Recall: Recall is calculated as follows:

$$\text{Recall} = \frac{TP}{TP + FN}$$

- F1-Score: F1-Score is the Harmonic mean of Precision and Recall.

$$\text{F1-Score} = 2 * \frac{Precision * Recall}{Precision + Recall}$$

- Use of BFS and DFS methodology to detect the path or source of the SMS sender.

- Cyclic Redundancy Check (CRC) is used to detect error in data and also information transmitted through the network,the uses checksum value and Modulo Division method helps to find the error in the bit,when the remainder is non zero. When the remainder is zero,then there is no error in the bit generated. It uses the redundancy factor type CRC-32.[16]

4.2 Results:

The feature extraction with text preprocessing is performed with countvectorizer and TFIFD vectorizerto extract the strings to numbers from the dataset to predict the SmishSMS accurately. Table 2, shows the parameters calculated using the different Machine Learning Algorithms.

Table 2: Performance of different Algorithm

Classifier	Precision	Recall	F1-scores	Accuracy
SmishSMS	0.94	0.98	0.96	98.97
Support Vector Machine Classifier	0.92	0.98	0.94	98.23
Naïve Bayes Classifier	0.90	0.97	0.93	98.00
Back Propagation Algorithm	0.92	0.97	0.93	97.48
Random Forest Classifier	0.91	0.97	0.95	96.10

To establish the significance of features employed in our system, the frequency of each characteristic in our dataset is determined. The frequency is generated based on the features in 152 Spam messages present in our dataset. The feature is tested by using python code to extract the features like mis-spelled words, Leet words, currency denominations, keywords and other features on the given figure 3.From the results generated,themisspelled words feature is most used for detecting spam SMS and it exists in 36.18% of spam SMS, Currency denominations which shows 28.29%, key words which shows 22.37%, leet words which is 8.55% and special characters which shows 4.61% of spam SMS. Special characters fetches the least features among all the others.

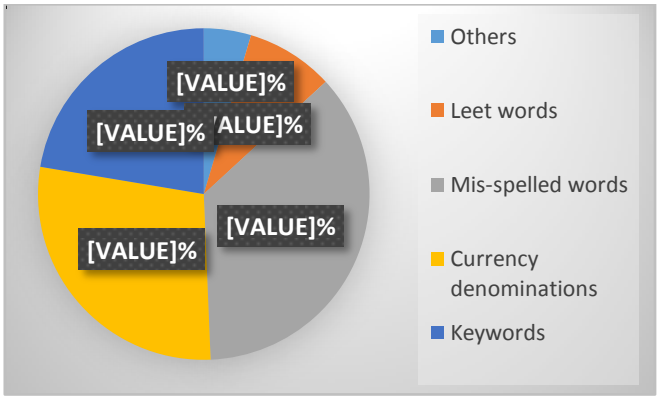

Figure 3: Frequency of the features in Percentage

The above given features defined by our system shows best accuracy while classifying the messages using proposed SmishSMS Algorithm. The values found by the above heuristics have been passed on to machine learning algorithms. The same dataset is being evaluated using other machine learning algorithms also, namely Support Vector Machine, Naïve Bayes, Random Forest and Back Propagation algorithm. The prediction results of each algorithm is based on the feature values which have been recorded. The prediction result of each algorithm is shown in Table 2. The evaluation shows that the Support Vector Machine Algorithm (SVM) gave

an accuracy of 98.23%, while SmishSMS gavethe good accuracy of 98.97% with the dataset. The Naive Bayes algorithm (NB) also evaluated a decent performance with an accuracy of 98%. The recall value of SmishSMS and SVM are almost the same with a value of 98%, respectively. The precision value using SVM is 92% and SmishSMS is 94%. Also, F1-Scores for the SVM is 94% and for SmishSMS is 96%, respectively. Hence, the SmishSMS gave the best outcome than SVM in the given experiment. The graph is shown below in Fig.4.

The proposed system not only finds the SMS received in mobile user as spam SMS or legitimate SMS,but it also checks the contents in the SMS like the URL,contact numbers as the spam or not. Furthermore, the source of the message received is checked with the Depth first search (DFS) method with the time to capture the network protocols, so that the mobile users will be aware of the spammers and will not attempt to read or click the irrelevant contents in the SMS. Currently, with the dataset provided to the system, the accuracy is only upto 35%, which needs to be researched on more. Also, there needs to be more accurate path to be found wherethe setback is.

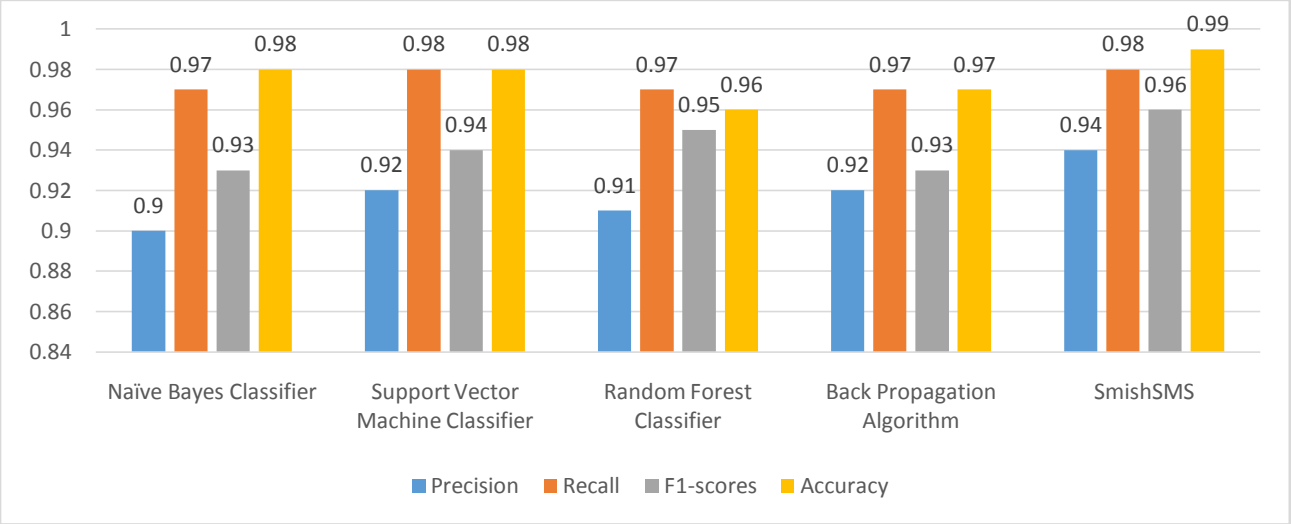

Figure 4: Frequency of the parameters in ML algorithm

Conclusion

Phishing is the more prominent fraud which is on-going in the current world leading to the financial losses to individual people and businesses too. The findings reveal that the SMS is classified into 2 different types like, legitimate and spam messages.

This paper totally focused on the development of SmishSMS system. The evaluation and prediction is done with the 1001 dataset collected in 3 months duration from different sources like Google forms,authentic dataset sources,etc.

This system had three phases of evaluation. The first phase is to check the type of the SMS received, where the contents of the SMS were checked for features around 20 in numbers, wherein few were URL, phone number, leet words, currency denominations, offers, etc. The second phase is the checking of domain from the URL received, to check for the contact numbers as spam or ham. The third phase was to get the source details with the internet protocol or the path using the data structures algorithm.

There has been a comparison of our technology with other smishing detecting techniques. It demonstrated how various cutting-edge methods were applied in this study project to find SMS phishing. There are many blacklisted numbers which are detected as spam, but these blacklisted numbers are not available in public. So, the database of the blacklisted numbers are listed in the system and the further evaluation of the system is carried out.

The SMS dataset is scrutinizedwith algorithms like Naive Bayes algorithm, Random forest classifier, Support Vector Machine algorithm, backpropagation algorithm, cycle detection approach, etc. Among the different algorithms evaluated,the SVM based SmishSMS model gave the accuracy of 98.97% of spam messages received in the mobile users and 35 % accuracy to the source finder, which the next aspect of research in challenge. Reinforcement learning techniques with deep learning can also be used to get the source found for the end users.

References

[1] S. S. G. A.K.U Serkan Gunal, "A novel framework for SMS Spam Filtering," *IEEE*, 2012.

[2] A. V.H.Modi, "Intrusion Detection System: A Review," *International Journal of Engineering Research & Technology(IJERT)*, no. 2013.

[3] Jain, Ankit Kumar; Gupta, B.B, "Phishing Detection: Analysis of Visual Similarity Based Approaches," *Hindawi Security and Communication Networks*, 2017.

[4] G. Sonowal and K. S. Kuppusamy, "SmiDCA An Anti-Smishing Model with Machine Learning Approach," *Security in computer systems and networks the computer journal*, 2018.

[5] Sharma, Shabnam et.al, "Study on Phishing Attacks," *International Journal of Computer Applications*, 2018.

[6] Jain, Ankit Kumar; Gupta, Brij B, "A machine learning based approach for phishing detection using hyperlinks information," *Springer Nature*, 2018.

[7] Ankit Kumar Jain et.al, "Predicting Spam Messages Using Back Propagation Neural Network," *Springer Nature*, 2019.

[8] B. B. Gupta and A. K. Jain, "Phishing Attack Detection using a Search Engine and Heuristics-based Technique," *Journal of Information Technology Research*, 2020.

[9] E. U. Soykan and M. Bagriyanik, "The Effect of SMiShing Attack on Security of Demand Response Programs," *energies 2020*, 2020.

[10] K. and D. Singla, "A Survey of Deep Learning based Online Transactions Fraud Detection Systems," in *International Conference on Intelligent Engineering and Management (ICIEM)*, 2020.

[11] L. Hajdu and M. Kresz, "Temporal Network Analytics for Fraud Detection in the Banking Sector," *Springer Nature Switzerland*, 2020.

[12] Jain Ankit Kumar et.al, "A Novel Approach to Detect Spam and Smishing SMS using Machine Learning Techniques," *International Journal of E-Services and Mobile Applications*, 2020.

[13] Mishra, Sandhya; Soni, Devpriya, "DSmishSMS-A System to Detect Smishing SMS," *Neural Computing and Applications*, 2021.

[14] Mishra, Sandhya; Soni, Devpriya, "Implementation of 'Smishing Detector': An Efficient Model for Smishing Detection Using Neural Network," *SN Computer Science*, 2022.

[15] Frances Effirim Botchey et.al, "An Evaluation of Machine Learning Methods to Predict Fraud in Mobile Money Transactions," *International Journal of Engineering Research & Technology (IJERT)*, 2022.

[16] "Towards Data Science," [Online]. Available: https://towardsdatascience.com/shortest-path-distance-with-

deep-learning-311e19d97569.

[17] "Quick Bird Studios," [Online]. Available: https://quickbirdstudios.com/blog/validate-data-with-crc/.

[18] "University of Massachusetts Amherst," [Online]. Available: https://www.umass.edu/it/security/phishing-fraudulent-emails-text-messages-phone-calls.