Detecting Smishing and Vishing Attacks using Machine Learning

Aaliyah E Chichwadia, Noluntu Mpekoa
University of Johannesburg
South Africa

Abstract

Mobile devices have become an integral part of daily human activities. The rapid growth and impact of these devices have made them increasingly susceptible to cyber-attacks. Attackers often exploit Short Text Messages (SMS) to target smartphone users. Smishing, also known as SMS phishing, is a type of attack that targets smartphone users through text messages. Smishing, a type of phishing, differs from traditional phishing in several ways, such as the information available in the SMS and the attack strategy. As such, detecting smishing is challenging because attackers share minimal information. The research explores the risks of smishing and vishing attacks targeting mobile device users. To tackle this, machine learning algorithms were developed to detect smishing attacks and a blacklist to identify potential vishing attacks. The models evaluated in this study include the Naïve Bayes, Decision Tree, and Random Forest algorithms. The results obtained indicated that the Random Forest model performed best in the selected environment. When addressing vishing attacks, the prototype solely relied on a blacklist which can be valuable, but its effectiveness is reliant on continuous updates and monitoring.

1. Introduction

In today's world, technology has democratized access to resources, allowing individuals from diverse backgrounds to avail themselves of a wide range of opportunities [1], [4], [19]. Moreover, technology is also being used to assist people who need help, enhancing their quality of life and opening up opportunities that would otherwise be out of reach [3]. Technology has revolutionized the way people live, making lives more convenient and efficient than ever before. The world has seen the benefits of using technology, such as automating tasks, setting reminders, efficient communication, easy bill payments, and simple online shopping for groceries to investing in valuable assets right in the comfort of our homes [2], [14], [18].

With the evolution of technology and the expanding market, mobile devices have undergone significant advancements [5]. These small devices have become integral to daily life and are relied upon for work, personal, financial, and leisure activities [12]. Our daily lives are greatly affected by mobile devices. These devices are used for a wide range of purposes in the military, government, education, and entertainment sectors [8]. Due to the ease of use, individuals keep sensitive data and carry out sensitive operations [18], [21]. Continuous development and growth make these devices more vulnerable to threats [7], [11], [13].

There has been a rise in attacks targeting these devices, which directly threaten personal privacy, data integrity, and network security [10]. In 2023, [20] noted a consistent rise in mobile device attacks, totaling nearly 33.8 million, marking a 50% increase from the previous year. Spam, phishing, malicious apps, and ransomware are just a few of the threats that mobile device users face today [16]. The attack techniques are becoming more sophisticated each year. According to [9], emails comprise only 15% of mobile phishing attacks, placing them behind messaging, social media, and "other" applications. Smishing involves sending fake text messages or misleading links over Short Message Service (SMS) [22], [30]. The user may inadvertently disclose personal information such as credit card numbers and account numbers. Also, the user may unknowingly download malware through website visits, infecting the device [15], [24]. While the consequences of a security breach will vary for each individual or organization, the most common impacts include damage to reputation, loss of data, financial impact, downtime, penalties and fines, loss of additional equipment, and increased insurance costs [10], [26]. Mobile devices store a significant amount of personal and sensitive data, including contact lists, emails, passwords, and financial information [31], [36]. It is crucial that mobile security protects this data from unauthorized access and potential misuse [38].

The use of SMS has expanded to become the most commonly used feature on mobile devices. The use of SMS has not been phased out by other messaging applications, despite technological advancements [27], [33]. However, the SMS feature has been exploited by hackers to carry out smishing attacks [28], [29]. Existing research has focused on detecting and separating spam SMS from legitimate messages but has not effectively prevented smishing. This study

 | DOI: 10.20533/ijicr.2042.4655.2024.0151

focuses on smishing and vishing attacks common on mobile devices.

This article is sectioned as follows: Section 2 discusses the literature on phishing and smishing, as well as on mobile device security, Section 3 presents the methods used to conduct the study, while Section 4 outlines the proposed model. Then Section 5 discusses the implementation and results, and Section 6 concludes the article.

2. Literature review

Smishing is a cybercriminal attack that targets mobile Short Message Service (SMS) devices. This type of attack involves sending a malicious link, phone number, or email via text message with the intent of stealing the recipient's sensitive information, such as passwords, bank account details, and credit card information [13], [23], [39]. One technique of combating smishing is to increase awareness and educate users about the various tactics used by SMS phishers. However, this method has been criticized for becoming inefficient because smishing tactics are continually evolving [25], [34]. This section discusses models and techniques to detect Smishing and Vishing attacks.

[37] in their study state that phishing scams via SMS are increasingly common due to smartphone and mobile internet use. Identifying a phishing SMS by analyzing unstructured short texts is a significant challenge in AI-driven cybersecurity. Their study made use of machine learning-based techniques integrated with natural language processing because they have massive potential in identifying differentiating patterns between phishing and legitimate SMS. Their study experimented with several machine learning algorithms on a benchmark dataset. They also integrated NLP-based feature extraction and feature selection steps to develop an automated phishing detection strategy. Their findings suggest that the support vector machine classifier when applied after feature extraction and feature selection outperforms the tenfold cross-validation score of 98.27%, F1-score of 99.08% for legitimate SMS, and accuracy of 98.39%. The performance of the tested methods were assessed using popular evaluation metrics on a standard dataset.

[32] support [37] by emphasizing that phishing is a serious cyber-security issue spreading through various media to capture critical profile information. The study describes Smishing as a type of phishing attack that uses text messages to trick people into giving away their online credentials. Their study presents an anti-phishing model entitled 'SmiDCA' (SMIshing Detection based on Correlation Algorithm). Their proposed model collected various smishing messages from different sources, and initially, 39 distinct features were extracted. The SmiDCA model encompasses dimensionality reduction, and machine Learning-based experiments were conducted without (BFSA) and with (AFSA) reduction of features. The model was validated with experiments and the results of these experiments indicated that in terms of accuracy, the model achieved: 96.40% for the English dataset and 90.33% for the non-English dataset. After nearly half of the features were pruned, the model achieved an accuracy of 96.16%.

This study differs in their definition of Smishing [26] from the previous authors. They define Smishing as an incarnation of phishing, but they state that Smishing and Phishing differ in the information available and attack strategy, causing detection to be difficult. Their study proposes a machine-learning based model to classify Swahili Smishing text messages targeting mobile money users. Experimental results show a hybrid model of Extratree classifier feature selection and Random Forest using TFIDF (Term Frequency Inverse Document Frequency) vectorization yields the best model with an accuracy score of 99.86%. A Swahili dataset with 32259 messages was used for performance evaluation and their results were measured against a baseline Multinomial Naïve-Bayes model. Their model returns the lowest false positive and false negative of 2 and 4, respectively, with a Log-Loss of 0.04.

Likewise, [6]'s study describes phishing as a type of cyber-attack aiming to steal user's confidential information using bait. They also define smishing as a form of a phishing attack that is carried out via SMS messages sent to mobile phones. In this study, a machine learning-based model is proposed to detect smishing messages. The study used experiments on a moderately large dataset, which contained legitimate and smishing messages. The results indicate that the proposed model offers a promising detection performance. Alam et al. (2020) conducted a study focusing on defending users against phishing attacks. Their study proposed and developed a model that detects phishing attacks using machine learning (ML) algorithms like random forest (RF) and decision tree (DT). A standard legitimate dataset of phishing attacks from Kaggle was used for model evaluation. The proposed model used feature selection algorithms like principal component analysis (PCA) to analyze the attributes of the dataset. The findings show that a maximum accuracy of 97% was achieved through the random forest algorithm.

[17] in their study allude that Smishing attacks have become a severe cyber-security difficulty and trigger incredible monetary losses to the victims. Therefore, they proposed an approach that analyzes text content and uniform resource locator (URL) presented in the SMS. Their solution integrates the URL phishing classifier with the text classifier to improve accuracy. Two data sets for both text as well as for URL phishing classifier were used and also a

 | DOI: 10.20533/ijicr.2042.4655.2024.0151

synthetic minority oversampling technique was used to balance the training data. The voting classifier merges the findings of each classifier passed into it and predicts the output on the basis of voting. The findings suggest that the integration of KNN, RF, and ETC can detect smishing messages with a 99.03% accuracy and 98.94% precision rate.

The study by [35] aims to develop a robust, effective, sophisticated, and reliable solution for phishing detection through the optimal feature vectorization algorithm (OFVA) and supervised machine learning (SML) classifiers. In this study, the OFVA was utilized to extract 41 optimal intra-URL features from a novel large dataset comprising 2,74,446 raw URLs (134,500 phishing and 139,946 legitimate URLs). Then, data cleansing, curation, and dimensionality reduction were performed to remove outliers, handle missing values, and exclude less predictive features. The study then evaluated and compared 15 SML algorithms arising from different machine learning (ML) families, including Bayesian, nearest-neighbors, decision trees, neural networks, quadratic discriminant analysis, logistic regression, bagging, boosting, random forests, and ensembles. The evaluation was performed based on various metrics such as confusion matrix, accuracy, precision, recall, F-1 score, ROC curve, and precision-recall curve analysis. The findings indicate that random forests (RF) outperformed the other classifiers, achieving a greater accuracy rate of 97.52%, followed by 97.50% precision, and an AUC value of 97%.

A plethora of anti-smishing solutions for mobile devices have been proposed (as shown above) to date but still, the battle against phishing continues to be an ongoing challenge. Firstly, the detection of smishing is a challenge not only because of feature constraints but also due to the scarcity of real smishing datasets. Secondly, there is a lack of solutions that attempt to combat both smishing and vishing attacks, lastly, there is a lack of fully-fledged solutions.

3. Methodology

The main aim of this paper was to design and develop a system to detect and defend against smishing and vishing attacks. A literature review was conducted to provide a comprehensive overview of previous research on this topic. The proposed model was implemented in Python, and a simple interface was created to demonstrate the model. The second implementation was developed using Android Studio so that the solution could be tested on a mobile device emulator. A database was set up to store a list of blacklisted numbers. The database also stores all messages and calls that are flagged as smishing or vishing.

The dataset used for training and testing of the model was the “SMS Spam collection dataset”. The dataset contains 5,574 SMS messages in English, tagged as either “HAM” or “SPAM”. The dataset is a combination of SMS messages from various sources, but it contains more legitimate messages than SPAM messages. For data preprocessing, irrelevant columns were removed keeping only “label” and “raw text”. The researcher also made sure that there were no missing values in the dataset. Tokenization, stemming and removal of stop words were techniques that were used to transform raw text. In this research, tokenization was utilized in text preprocessing, sentiment analysis, and language modeling, whereas stemming was used for information retrieval, and text mining.

The models evaluated in this study include the Naïve Bayes, Decision Tree, and Random Forest algorithms for detecting smishing attacks.

4. Proposed Model

The proposed model for detecting and preventing smishing and vishing attacks is an Android studio application that provides real-time analysis of both smishing and vishing attacks. The application uses machine learning algorithms in conjunction with a database of blacklisted numbers to perform analysis on incoming messages and calls and then notifies the user of the outcome.

4.1. Data Collection

The data used to train the model comes from a publicly available dataset that is a set of SMS-tagged messages collected for research. It contains 5574 English messages that are tagged as ham (legitimate messages) or spam. The file containing the messages has one message per line, and each line contains two columns: a label (ham or spam) and the raw text. This has been collected from free research sources on the internet.

4.2. Machine learning algorithm

A machine learning model was developed to analyze text messages for smishing attacks by detecting anomalies or suspicious patterns. The model was trained using labeled data from the dataset to classify text messages. The messages were evaluated and assigned to a likelihood score of being a spam message or not. The model can be regularly retrained to adapt to new patterns and achieve improved accuracy.

4.3. Database

A database has been established to store a list of blacklisted numbers. These numbers are used to identify if an incoming call has been reported as a vishing call by other users. The database also logs all

 | DOI: 10.20533/ijicr.2042.4655.2024.0151

messages and calls that are flagged as smishing or vishing. This information is used to enhance the system and to generate reports related to the system.

4.4. End-User

User notifications - Upon receipt of a message, the message content is automatically analyzed using a machine learning model to detect if it is a spam message. The user is promptly notified of the analysis results and can then decide how to proceed. Similarly, for incoming calls, the caller is immediately analyzed, and the user is notified accordingly. This allows the user to make an informed decision on whether to answer the call or not.

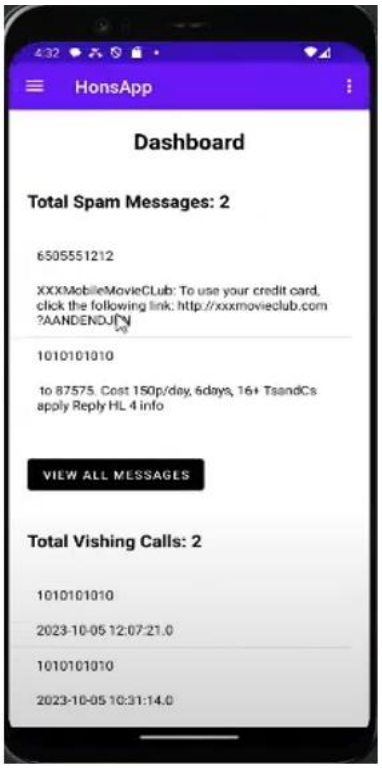

Figure 1. Smishing and Vishing Analysis

User actions - Every system user can take action when an analysis is conducted (Figure 1 above). These actions include reporting a message as smishing, reporting a call as vishing, and managing the application's permissions over received messages and calls.

5. Implementation and results

The proposed prototype for detecting and preventing smishing and vishing attacks involves developing an application using Android Studio. Integrated into this application is a machine learning model trained to identify spam text messages and a blacklist of reported numbers to identify vishing attacks. The core of the solution involves training machine learning models. Once trained, these models are used to analyze text messages for any signs of a smishing attack. In this research, three models have been trained in Python, each using a different algorithm.

The first model utilizes the Naïve Bayes classification technique, which is based on the Bayes Theorem. It assumes that all features predicting the target value are independent of each other. The model calculates the probability for each class and selects the class with the highest probability. However, due to limitations, a second model was developed.

The second model was trained using the Decision Tree algorithm which recursively divides the dataset into subsets according to the most important attribute at each stage. The objective is to build a decision tree composed of decision nodes, with branches for feature values and nodes for features. The predicted value is contained in the leaves. Using this algorithm, the third model was developed.

The third model utilizes the Random Forest algorithm. This algorithm works by combining multiple Decision Trees to form a Random Forest. This approach improves prediction accuracy and reduces overfitting. The algorithm creates a forest of decision trees, with each tree using a random subset of features at each node and trained on a random portion of the dataset. The final prediction is determined by averaging the scores from individual trees. The model was trained using the TensorFlow Lite framework, which enables efficient model execution on mobile devices. This framework was chosen to facilitate proper demonstration on the Android emulator.

Prior to training the model, the data required preparation. The data used was sourced from a publicly available dataset containing tagged SMS messages as either ham (legitimate) or spam. Unnamed columns were dropped, missing values were addressed, and duplicate rows were eliminated. The text underwent preprocessing, including conversion to lowercase, removal of non-alphanumeric characters, elimination of English stopwords and punctuation marks. Subsequently, the text was transformed using TF-IDF (Term Frequency-Inverse Document Frequency) and vectorization to provide the relevant features for the machine learning algorithm to detect and classify spam messages.

After training the model, it was integrated into the Android Studio application. For every incoming message, the content is analyzed, and a prediction is made to determine whether the text received is spam or not (shown in Figure 2 below). Additionally, the solution utilizes a database that maintains a list of blacklisted numbers and tracks smishing and vishing encounters by storing information such as flagged messages, reported incidents, and blacklisted numbers.

 | DOI: 10.20533/ijicr.2042.4655.2024.0151

The blacklisted numbers are used to identify potential vishing attacks from incoming calls. When a message or call is received, the application notifies users in real-time if it has been identified as smishing or vishing. Users can take actions such as reporting a message as smishing or a call as vishing. They can then decide how to handle the message or whether to answer the call.

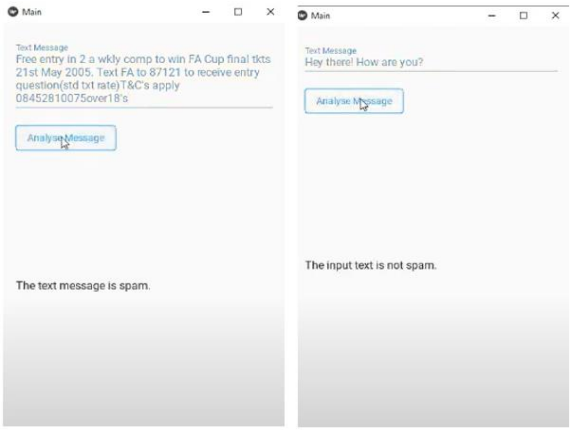

Figure 2. Scanning of incoming messages

5.1. Results

The prototype was tested for its ability to detect smishing attacks. The prototype identifies vishing attacks based on a database of blacklisted numbers. This section presents the results of these evaluations and assesses the effectiveness of the blacklist in detecting and preventing smishing and vishing attacks. The assessment of each machine learning model involved evaluating accuracy, precision, recall, and F1 score. The results of this assessment are shown in Figure 3.

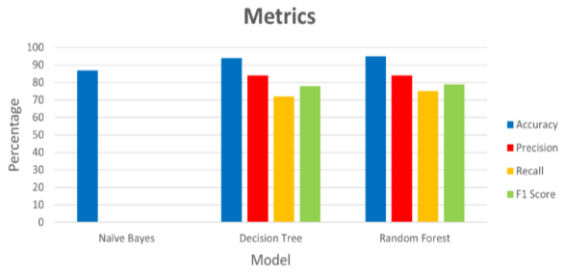

Figure 3. Summary Results on Three Models

The comparison of models shows that the Random Forest model outperforms other models in detecting smishing attacks. The Decision Tree comes in second with scores close to that of the Random Forest, and the Naïve Bayes model is far behind. The model's 0 precision and recall rates are a result of the dataset imbalance. The ability to detect vishing attacks depends entirely on the blacklist status. To maintain effectiveness, the list must be regularly updated with the most recent numbers. Failure to do so reduces the method's ability to detect and prevent vishing attacks.

The Naïve Bayes model had a lower accuracy, which raised several considerations. The precision and recall rates, both being 0%, indicate the model's inability to successfully classify messages. This could be the result of the imbalanced dataset, resulting in a failure to capture smishing attacks. The F1 score reflects the poor trade-off between precision and recall, further pointing out the need for adjustments in this approach.

The Decision Tree model showed improved accuracy, indicating its ability to correctly identify patterns and make more accurate predictions. However, the precision and recall rates suggest that there is room for improvement. The model's precision score of 84% signifies an acceptable rate of true positive predictions, but the recall score of 72% indicates a significant number of false negatives. The F1 score of 78% highlights the need to enhance the model's ability to detect smishing attacks without generating as many false positives.

The Random Forest model achieved the highest accuracy of 95%. This model effectively mitigated overfitting and captured a wide range of patterns related to smishing attacks. With a precision score of 84%, it demonstrated its ability to minimize false positives, while the 75% recall score indicated a higher rate of false negatives, highlighting the need to improve the model's capabilities. The F1 score of 79% demonstrated a balanced trade-off between precision and recall, but there is room for improvement to increase the score.

The performance differences among the models are attributed to their specific characteristics. The Naïve Bayes model relies on independent assumptions, which can make it overly cautious and result in poor precision and recall rates. In contrast, Decision Trees are effective at capturing patterns, while Random Forests take a collaborative approach to balance precision and recall rates.

To combat vishing attacks, the prototype uses a blacklist. This blacklist stores numbers that have been identified as sources of vishing attacks. While using the blacklist helps in spotting potential vishing attacks, it's not entirely reliable. New spoofed vishing numbers may not be detected initially, and there's a risk of legitimate phone numbers being wrongly flagged as potential vishing sources, leading to false positives. Keeping the blacklist up to date will require a lot of resources and ongoing monitoring. Additionally, blacklists are not very effective against targeted vishing attacks, like spear phishing.

6. Future areas of research

Future research could concentrate on utilizing advanced machine learning models, such as deep

 | DOI: 10.20533/ijicr.2042.4655.2024.0151

learning techniques, natural language processing, and behavioral analysis, to enhance the accuracy of detecting smishing and vishing. Additionally, research can explore the integration of hybrid models that combine different algorithms to create a more comprehensive solution. Real-time analysis could also be pivotal in detecting and preventing smishing and vishing attacks.

Improving the dataset is crucial for enhancing the performance of machine learning models. It is important for future research to prioritize data collection by incorporating information from diverse sources, regions, and contexts. Additionally, integrating user behavior data can offer valuable insights into how individuals interact and respond to these attacks. Furthermore, it is essential to consider data privacy and ethical concerns when creating and adding datasets.

7. Conclusion

The research paper focuses on the increasing threat of smishing and vishing attacks targeting mobile device users. As mobile devices have become essential for various activities, they have also become prime targets for cybercriminals. Addressing these challenges is crucial to ensure safety in the digital world. The dynamic and sophisticated nature of phishing attacks, alongside the inadequacy of anti-phishing tools, has made phishing detection a significant challenge. Consequently, new gaps have emerged in phishing detection, including the challenges and drawbacks of existing techniques.

In order to address these issues, this study proposed a solution for detecting and preventing smishing and vishing attacks. The prototype utilizes machine learning models to analyze incoming text messages and employs a blacklist to identify potential vishing attacks. The research evaluated the performance of three machine learning models – Naïve Bayes, Decision Tree, and Random Forest – for detecting smishing attacks. The findings showed that the Random Forest model performed the best in the given context. When it came to addressing vishing attacks, the prototype relied solely on a blacklist, which can be effective, but its success depends on regular updates and monitoring. Without real-time capabilities, the solution may not fully address these threats.

In future research and development, it has been noted that more advanced machine learning models could be developed to effectively detect both smishing and vishing attacks. It is also important to improve the datasets used to train the model. Data diversity and balance between the categories being trained are important to ensure an unbiased model is trained. In the ever-evolving digital world, safeguarding against smishing and vishing attacks is crucial. The prototype and research in this paper provide an understanding of these threats and lead to opportunities to develop a more effective countermeasure to the threat. By addressing the existing challenges, new paths can be explored so that individuals can be better protected.

8. References

[1] O. N. Akande, O. Gbenle, O. C. Abikoye, R. G. Jimoh, H. B. Akande, A. O. Balogun, and A. Fatokun, “SMSPROTECT: An automatic smishing detection mobile application”, ICT Express, 9(2), 2023, pp. 168-176.

[2] M. N. Alam, D. Sarma, F. F. Lima, I. Saha, and S. Hossain, “Phishing attacks detection using machine learning approach”, in 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT) IEEE. 2020, pp. 1173-1179.

[3] Z. Alkhalil, C. Hewage, L. Nawaf, and I. Khan, “Phishing attacks: A recent comprehensive study and a new anatomy”, Frontiers in Computer Science 3, 2021, 563060. doi: 10.3389/fcomp.2021.563060

[4] M. F. Ansari, P. K. Sharma, and B. Dash, “Prevention of phishing attacks using AI-based Cybersecurity Awareness Training”, Prevention, 3(6), 2022, pp. 61-72.

[5] M.E. Armstrong, K.S. Jones, and A.S. Namin, “How perceptions of caller honesty vary during vishing attacks that include highly sensitive or seemingly innocuous requests”, Human Factors 65, no. 2, 2023, pp. 275-287.

[6] C. Balim, and E. S. Gunal, “Automatic detection of smishing attacks by machine learning methods”, in 2019 1st International Informatics and Software Engineering Conference (UBMYK), IEEE, 2019, pp. 1-3.

[7] M.A.S. Bubukayr, and M.A. Almaiah, “Cybersecurity concerns in smart-phones and applications: A survey”, in 2021 International Conference on Information Technology (ICIT), IEEE, 2021, pp. 725-731.

[8] J. Bullee, L. Montoya, M. Junger, and P. Hartel, “Spear phishing in organisations explained”, Information and Computer Security 25, no. 5, 2017, 593-613.

[9] CheckPoint. (2024) Top 6 Mobile Security Threats and How to Prevent Them, https://www.checkpoint.com/cyber-hub/threat-prevention/what-is-mobile-security/top-6-mobile-security-threats-and-how-to-prevent-them/ (Access Date: 12 February 2024).

[10] N. Choudhary, and A. K. Jain, “Comparative analysis of mobile phishing detection and prevention approaches”, In Information and Communication Technology for Intelligent Systems (ICTIS 2017) -Volume 1 2, Springer International Publishing, 2018, pp. 349-356.

[11] W. Dimitrov, “The impact of the advanced technologies over the cyber attacks surface”, In Artificial Intelligence and Bioinspired Computational Methods: Proceedings of the 9th Computer Science On-line Conference 2020, Vol. 2 9, Springer International Publishing, 2020, pp. 509-518.

 | DOI: 10.20533/ijicr.2042.4655.2024.0151

[12] B. Dupont, (2012). The Cyber Security Environment to 2022: Trends, Drivers and Implications. SSRN: DOI: 10.2139/ssrn.2208548,

[13] T. Fadziso, U.R. Thaduri, S. Dekkati, V.K.R. Ballamudi, and H. Desamsetti, "Evolution of the cyber security threat: an overview of the scale of cyber threat", Digitalization and Sustainability Review 3, no. 1, 2023, pp. 1-12.

[14] T. Halevi, N. Memon, and O. Nov. (2015). Spear-phishing in the wild: A real-world study of personality, phishing self-efficacy and vulnerability to spear-phishing attacks. Phishing Self-Efficacy and Vulnerability to Spear-Phishing Attacks. SSRN: DOI: 10.2139/ssrn.2544742,

[15] M. Hijji, and G. Alam. (2021). A multivocal literature review on growing social engineering based cyber-attacks/ threats during the COVID-19 pandemic: challenges and pr ospective solutions. https://us.norton.com/ blog/ emerging -threats/cybersecurity-statistics (12 February 2024)

[16] Internet Crime Complaint Center (IC3). (2021). https:// www.ic3.gov/Media/PDF/Annual-Report/2021IC3Report. pdf. (12 February 2024)

[17] A. K. Jain, B. B. Gupta, K. Kaur, P. Bhutani, W. Alhalabi, and A. Almomani, "A content and URL analysis-based efficient approach to detect smishing SMS in intelligent systems", International Journal of Intelligent Systems, 37(12), 2022, pp 117-141.

[18] Kalaharsha, P., and Mehtre, B. M. (2021). Detecting Phishing Sites: An Overview. arXiv preprint arXiv: 2103.1 2739.

[19] Kaspersky. (2022). What is Smishing and How to Defend Against it, 2022. [online] https://www.kaspersky. co.za/resource-center/threats/what-is-smishing-and-howto-defend-against-it (Access Date: 12 February 2024).

[20] Kaspersky. (2024). Attacks on mobile devices significantly increase in 2023. https://www.kaspersky.com/ about/press-releases/attacks-on-mobile-devices-significantl y-increase-in-2023 (Access Date: 12 February 2024).

[21] P. Kelley, "Evolution of Cyber Attacks and Their Economic Impact", TechRxiv Authorea Preprints. 2023, 10.36227/techrxiv.21670718.v1

[22] A. Kovač, I. Dunđer, and S. Seljan, "An overview of machine learning algorithms for detecting phishing attacks on electronic messaging services", In 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO) IEEE, 2022, pp. 954-961.

[23] K. Krombholz, H. Hobel, M. Huber, and E. Weippl, "Advanced social engineering attacks", Journal of Information Security and Applications 22, 2015, pp 113-122.

[24] M. Lee, and E. Park, "Real-time Korean voice phishing detection based on machine learning approaches", Journal of Ambient Intelligence and Humanized Computing, 14(7), 2023, pp 8173-8184.

[25] A. R. Mahmood, and S. M. Hameed, "Review of Smishing Detection Via Machine Learning", Iraqi Journal of Science, 2023, pp 4244-4259.

[26] I. S. Mambina, J. D. Ndibwile, and K. F. Michael, "Classifying swahili smishing attacks for mobile money users: A machine-learning approach", IEEE Access, 10, 2022, 83061-83074.

[27] G. Martin, (2022). South Africa in top five countries affected by cybercrime in 2022, https://www.defence web.co.za/cyber-defence/south-africa-in-top-five-countries affected-by-cybercrime-in-2022/. (Access Date: 12 Februar -y 2024).

[28] E.M. Maseno, (2017). Vishing attack detection model for mobile users. PhD dissertation, KCA University.

[29] S. Mishra, and D. Soni, "Dsmishsms-a system to detect smishing sms", Neural Computing and Applications, 35(7), 2023, pp. 4975-4992.

[30] D. Njuguna, J. Kamau, and D. Kaburu, "Model for mitigating smishing attacks on mobile platforms", In 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET), IEEE, 2021, pp. 1-6.

[31] P.S. Seemma, S. Nandhini, and M. Sowmiya, "Overview of cyber security", International Journal of Advanced Research in Computer and Communication Engineering 7.11, 2018, pp. 125-128.

[32] G. Sonowal, and K. S. Kuppusamy, "SmiDCA: an anti-smishing model with machine learning approach", The Computer Journal, 61(8), 2018, pp 1143-1157.

[33] Stouffer, C. (2022). 115 cybersecurity statistics + trends to know in 2023, https://us.norton.com/blog /emerging-threats/cybersecurity-statistics (Access Date: 12 February 2024).

[34] Surfshark. (2024). Cybercrime statistics, https://surfs hark.com/research/data-breachimpact/statistics (Access Dat -e: 12 February 2024).

[35] M. A. Tamal, M. K. Islam, T. Bhuiyan, A. Sattar, and N. U. Prince, "Unveiling suspicious phishing attacks: enhancing detection with an optimal feature vectorization algorithm and supervised machine learning", Frontiers in Computer Science, 6, 2024, 1428013. DOI: 10.3389/fcomp.2024.1428013

[36] V, R. Truchero (2021). Efficient smishing detector using Machine Learning techniques. Dissertation, Universitat de Lleida.

[37] R. E. Ulfath, I. H. Sarker, M. J. M. Chowdhury, and M. Hammoudeh, "Detecting smishing attacks using feature extraction and classification techniques", In Proceedings of the International Conference on Big Data, IoT, and Machine Learning: BIM, Springer Singapore, 2022, pp. 677-689.

[38] P. Weichbroth, and Ł. Łysik, "Mobile security: Threats and best practices", Mobile Information Systems 2020, 2020, pp. 1-15.

 | DOI: 10.20533/ijicr.2042.4655.2024.0151

[39] E. O. Yeboah-Boateng, and P. M. Amanor, "Phishing, SMiShing and Vishing: an assessment of threats against mobile devices", Journal of Emerging Trends in Computing and Information Sciences, 5(4), 2014, pp 297-307.

 | DOI: 10.20533/ijicr.2042.4655.2024.0151