"HYBRID SMISHING DETECTION USING SIGNATURE BASED AND RANDOM FOREST APPROCH"

Shivaraj Chaudhari*1 ,Akshad Chavhan*2 , Sarang Marshatwar*3, Saurabh Jadhav*4, Ayush Vishvanathwar *5

**1,2,3,4,5 BNCOE ,Department Of Computer Science & Engineering, Sant Gadgebaba Amarawati University,Pusad, Maharashtra ,India*

*vc8651396@gmail.com*1*
*akshadchavhan@gmail.com*2*
*sarangmarshatwar15@gmail.com*3*
*saurabhjadhao18@gmail.com*4*
*ayushatulvishwanathwar@gmail.com*5*

Abstract :

Smishing (SMS phishing) is a growing threat in mobile communication, exploiting users through deceptive text messages to steal sensitive information. Traditional signature-based detection methods are fast but limited to known patterns, failing to catch evolving attacks. This paper proposes a hybrid detection framework combining signature-based filtering with a Random Forest machine learning model. The system first filters known threats, then analyzes remaining messages using features like content, sender metadata, and tone. Experimental results show high accuracy, precision, and recall. This two-tiered approach offers scalable, real-time protection and is ideal for integration into mobile security applications and SMS gateways.

Keywords *: Smishing (SMS phishing), SMS gateways.*

❖ Introduction :

Smishing, a portmanteau of SMS and phishing, represents a sophisticated form of cyberattack that leverages deceptive text messages to illicitly obtain sensitive personal information from unsuspecting individuals. The rapid and widespread increase in global mobile usage has inadvertently made a vast number of users more vulnerable to these insidious phishing attempts. Unlike traditional phishing that often targets email, smishing exploits the perceived trustworthiness and immediate nature of SMS communication, leading users to believe the messages are legitimate. Existing smishing detection systems typically fall into one of two categories: those that solely rely on predefined patterns or "signatures" to identify malicious messages, or those that employ machine learning models to learn and predict threats. However, each of these solitary approaches presents inherent limitations in their effectiveness against the evolving landscape of smishing attacks.

❖ Literature Surve

A thorough review of existing research and methodologies in smishing detection reveals several distinct approaches, each with its own merits and, importantly, its limitations:

"Phishing detection using keyword matching": This study primarily employs a **signature-based filtering** approach. This method relies on a database of known malicious keywords, phrases, or sender identities to flag suspicious SMS messages. While effective against previously identified smishing attempts, its significant drawback is its inherent inability to detect unknown or "zero-day" smishing attacks. New variations of smishing, which do not match existing signatures, can easily bypass such systems, leaving users exposed.

"Machine Learning for SMS spam detection": This research explores the application of various **machine learning algorithms**, specifically Naive Bayes and Support Vector Machines (SVM), for the purpose of SMS spam detection. These models are trained on large datasets of both legitimate and malicious SMS messages to learn patterns and classify new messages. While demonstrating a degree of effectiveness, this approach typically offers only moderate accuracy compared to more advanced techniques. Furthermore, it necessitates the availability of a large and meticulously labeled dataset for effective training, which can be a significant challenge in itself. The performance of such models is heavily dependent on the quality and quantity of the training data.

"Hybrid model for phishing SMS detection": This particular study proposes a more sophisticated approach by combining a **signature-based method with a Random Forest algorithm**. The intention here is to leverage the strengths of both rule-based detection and adaptive machine learning. This hybrid model has shown promising results in improving overall detection accuracy, indicating a step forward in combating smishing. However, the paper also notes that the implementation of such a combined system can be quite complex, requiring careful integration and fine-tuning of both components. This complexity can pose challenges in development, deployment, and maintenance.

❖ Problem Statement

The escalating number of SMS-based phishing incidents, commonly known as smishing, represents a grave and growing threat to the security and privacy of mobile phone users. These malicious attacks are ingeniously designed to manipulate individuals, often through social engineering tactics, into inadvertently revealing highly sensitive personal data such as banking credentials, login details, or other confidential information. The insidious nature of smishing lies in its ability to bypass traditional security measures and exploit human trust. Given the persistent evolution of smishing tactics and the limitations identified in existing detection methods, there is a pressing need for a more robust and adaptable solution. Therefore, the core objective of this project is to conceptualize, design, and develop a **lightweight hybrid detection system**. This innovative system will strategically integrate the strengths of both **signature-based and Random Forest machine learning approaches**. The ultimate aim is to achieve a higher degree of precision and accuracy in identifying and flagging malicious SMS messages, thereby providing enhanced protection to mobile users against these pervasive threats.

❖ Proposed System

The proposed hybrid smishing detection system is meticulously engineered to address and overcome the inherent limitations of approaches that rely on a single detection methodology. By combining two powerful techniques, it aims to provide a more comprehensive and robust defense:

Signature-Based Filtering: This component forms the initial layer of defense within the system. It operates by comparing incoming SMS messages against a constantly updated database of known smishing patterns, keywords, suspicious URLs, and sender IDs. This acts as a highly efficient first line of defense, capable of immediately identifying and flagging messages that match previously identified malicious characteristics. It excels at detecting common and well-established smishing schemes with high speed.

Random Forest Classifier: Complementing the signature-based filter, the **Random Forest Classifier** serves as the intelligent core for identifying more sophisticated, unknown, or rapidly evolving threats. Unlike static signatures, the Random Forest, a powerful ensemble machine learning model, analyzes numerous features extracted from an SMS message, such as linguistic patterns, message structure, sender behavior, and the presence of suspicious links, to determine its legitimacy. It is particularly effective at recognizing subtle anomalies and complex relationships within messages that would bypass simple keyword matching, making it adept at catching "zero-day" attack.

Backend Implementation: The computational backbone of this hybrid system is developed using **Python**, a versatile programming language, in conjunction with the **Flask framework**. This backend server is responsible for receiving SMS data from the mobile application, processing it through both the signature-based filtering module and the Random Forest classifier, and applying the sophisticated classification logic. It efficiently manages the detection process, serving as the central hub for all analytical operations.

Android Frontend: The user-facing component of the system is an **Android application**, meticulously built using **Java and XML**. This frontend is designed for seamless integration with the mobile device's SMS functionalities. It is responsible for intercepting and capturing incoming SMS messages. These captured messages are then securely transmitted to the Python Flask backend server for real-time analysis. Once the backend completes its classification, the Android application receives the results and intelligently displays them to the user, alerting them to potential smishing threats and offering guidance on how to proceed.

❖ Flowchart of the System and Methodology

In this section, we describe the working of the proposed system. The flowchart of the model is shown in Figure 3. As shown in the above flowchart, the process is explained in the following steps: Step 1: When an SMS is received, the system checks for the presence of URL or SAL(Self Answering Link) in the message. Step 2: If URL or SAL is present in SMS, we analyze the source code of the URL to check for the presence of form tag in it. Step 3: If form tag is present in the source code, we classify the message as SMISHING else we go to Step 5. Step 4: If URL or SAL is not present in the message, we check for the presence of Phone number or E-mail ID in the message Step 5: If a Phone number or E-mail ID is present in the SMS, the system analyzes the keywords present in the contents of the message to classify the SMS using Tfidf Vectorizer and OneVsRestClassifier. Step 6: If the prediction of the algorithm is malicious, then we classify the SMS as SMISHING else we classify it as LEGITIMATE. Step 7: If form tag is not present in the source code, we analyze whether an APK file is downloaded or not while invoking the URL. Step 8: If an Apk is downloaded, the message is regarded as SMISHING and BLOCKED else the message is classified as LEGITIMATE.

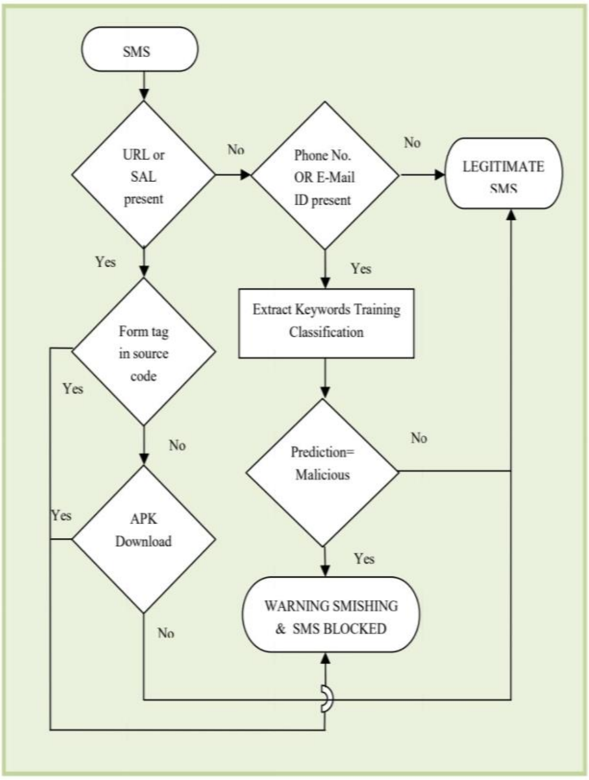

Figure 3: Flowchart of the proposed system

❖ Applications

The proposed hybrid smishing detection system holds significant practical value and can be seamlessly integrated into various environments to bolster digital security:

Banking Applications: Financial institutions are prime targets for smishing attacks due to the sensitive nature of the information they handle. This system can be directly **integrated into banking applications** to actively monitor incoming SMS messages for any suspicious content or patterns that might indicate a smishing attempt targeting their customers. By doing so, it provides an additional layer of security, proactively protecting users from financial fraud and identity theft.

Integratable into Mobile Apps: Beyond banking, the architecture of the proposed system is designed for broad compatibility, making it **integratable into a wide array of mobile applications**. This flexibility allows third-party developers and organizations to embed robust smishing detection capabilities directly into their own applications, thereby enhancing the overall security posture and user trust across different platforms and services.

Individual User Protection: Fundamentally, this system offers a vital tool for **individual mobile users** who are increasingly concerned about their data privacy and the pervasive threat of phishing attacks. By installing the Android application, users gain a personal guardian that actively scans incoming SMS messages, providing real-time alerts and warnings about potential smishing threats. This empowers users to make informed decisions and avoid falling victim to deceptive messages, significantly improving their personal cybersecurity.

❖ Conclusion and Future Scope

The development and implementation of this hybrid approach represent a significant advancement in the field of smishing detection. By intelligently combining the deterministic power of signature-based models with the adaptive learning capabilities of Random Forest, the system demonstrably **enhances detection accuracy** over models that rely on a single methodology. This synergistic approach leads to a more comprehensive and resilient defense against both known and emerging smishing threats. Furthermore, the proposed system is engineered with **real-time detection capability**, ensuring that users receive immediate alerts and proactive protection against malicious SMS messages as they arrive.Looking ahead, the research team identifies several promising avenues for future development to further enhance the system's capabilities:

Extend Support for Multilingual SMS: Currently, the system primarily focuses on detecting smishing in a single language (presumably English). A crucial future enhancement involves expanding its linguistic capabilities to

support multilingual SMS messages. This would require developing or integrating language-agnostic feature extraction techniques and training models on diverse linguistic datasets to effectively identify smishing in various languages.

Encrypted Message Handling: With the increasing prevalence of end-to-end encryption in messaging platforms, a significant challenge for future work is to address **encrypted message handling**. Detecting smishing within encrypted communication presents unique technical hurdles, as the content is not directly accessible. Future research could explore methods like analyzing metadata,traffic metadata,traffic patterns, or integrating with platform-specific security APIs (if permissible and secure) to identify suspicious activities without compromising user privacy.

These future endeavors aim to make the hybrid smishing detection system even more robust, adaptable, and globally applicable in the face of evolving cyber threats.

❖ References :

1)“Detecting Smishing and Vishing Attacks using Machine Learning” by Aaliyah E Chichwadia, Noluntu Mpekoa ,University of Johannesburg, South Africa

2)“A Content-Based Approach for detecting Smishing in Mobile Environment “by Sandhya Mishra, Devpriya Soni , Department of Computer Science and Engineering, Jaypee Institute of Information Technology, Noida, India

3)“An Integrated NLP and Machine Learning Model for Detecting Smishing Attacks on Mobile Money Platforms” by Katongo Ongani , Aaron Zimba,Mwiza Norina ,Chimanga Kashale in School of Computing, Technology and Applied Sciences ZCAS University Lusaka, Zambia

4)SmishSMS- The Latest Detection of SMS Phishing Trends” byAnisha Asirvatham,Research Scholar Department of Computer Science,Vels Institute of Science Technology and Advanced Studies,Pallavaram, Chennai, India ; Dr.c.Mennakashi ,Department of Computer Applications, Vels Institute of Science Technology and Advanced Studies Pallavaram, Chennai, India

5)“Detection of Phishing in Mobile Instant Messaging using Natural Language Processing and Machine Learning “ by Suman Verma ,School of Computing National College of Ireland

6)“SmishGuard: Leveraging Machine Learning and Natural Language Processing for Smishing Detection “ by Saleem Raja Abdul Samad, Pradeepa Ganesan, Justin Rajasekaran, Madhubala Radhakrishnan, Hariraman Ammaippan, Vinodhini Ramamurthy, College of Computing and Information Sciences, Information Technology Department, University of Technology and Applied Sciences-Shinas, Little Angel Institute, Karur, Tamil Nadu. India.