Comprehensive Research on State-of-the-Art Progress in Smishing Detection Methods and Techniques

Hardhat Enterprises

Table of Contents

I. Introduction

○ Overview of Smishing

By now, most of us are familiar with the term ‘phishing,’ but have you ever come across ‘smishing’? Picture this: you receive a random message on your phone claiming your parcel has encountered issues. If you’re expecting a package, you might instinctively click the link in concern, unknowingly falling victim to smishing. In essence, smishing is when attackers deceive victims by sending deceptive messages, coaxing them to divulge sensitive information like user accounts and passwords, which can be used to steal and misuse their personal data (Mishra and Soni, 2019). But never fear, Hard Hat are here to help you stay vigilant against this cunning cyber threat!

○ Importance of Smishing Detection

In an era of digital communication and mobile technology, we have seen the rise of smishing, a cunning and disruptive blend of SMS and phishing, which emphasises the critical importance of robust smishing detection. Smishing attacks exploit the familiarity and trust often associated with text messages to manipulate individuals into revealing sensitive information or engaging with malicious links. The implications of falling victim to such smishing attacks extend beyond financial losses, potentially leading to identity theft, data breaches and compromised privacy.

The frontline defence to this growing cyber threat is effective smishing detection. By swiftly identifying and thwarting smishing attempts, detection mechanisms play a major role in preventing unsuspecting victims from succumbing to these deceptive messages. The urgency that is often displayed in smishing messages, coupled with their impersonation of legitimate communication, emphasises the need for dynamic and advanced detection strategies.

Furthermore, smishing detection contributes to the overall security of the digital world. When individuals trust the integrity of text messages and the online platforms they utilise, they are more likely to engage in communications. A comprehensive detection system not only shields users from harm but also can foster confidence in digital interactions, increasing trustworthiness of online communication.

Smishing detection is not merely a safeguard, it is a key pillar of a secure digital landscape. Its role in halting deceptive messages, protecting personal information, and fortifying user trust makes it an indispensable ally in the ongoing battle against cyber threats. As smishing attacks continue to evolve, the importance of smishing detection is paramount to counter these threats and ensure the safety and integrity of digital interactions.

- Objectives of the Report

II. Background Research

Historical Perspective of Smishing Attacks

Social engineering attacks first began in 1987 where a phishing technique was used by the International HP Users Group, Intertex (Graphus 2023:para.2). In the 1990's the first large phishing attack took place when American Online (AOL) were the victims (Graphus 2023:para.3). AOL's user credentials were stolen along with creating randomised credit card numbers by using algorithms (Graphus 2023:para.3). Phishing became a well-known attack in the early 2000s when malware spread called the LoveBug (Graphus 2023:para.2). The malicious message was sent via emails and was titled "ILOVEYOU" with an attachment of an impersonating text file titled "LOVELETTER" (Graphis 2023:para.2). When this file was opened, a worm was released that would overwrite image files and send a copy of itself to all the contacts in the victim's Outlook address book (Graphus 2023:para.2). In early 2004, cybercriminals were successfully using the phishing techniques to attack banking sites and their customers (Phishing n.d.:para.7) . In late 2008, Bitcoin and other cryptocurrencies were created, allowing transactions using malicious software to be anonymous and secure (Phishing n.d.:para.7). The number of these attacks continue to grow each year.

The first time that the term "phishing" was used was in 1996 (Phishing n.d.:para.4). The term "phishing" came from a word meaning the illicit act of phone "phreaking" (Graphis 2023:para.4). "Phreaking" explains the exploration, investigating and study of telecommunication systems (Phishing n.d.:para.3). "Phish" is pronounced like "fish" as it is an analogy of an angler throwing a baited hook out there (the phishing email) and hoping the victim bites (Fruhlinger 2020:para.4). The term "smishing" comes from "phishing" and "SMS" as it is SMS-based phishing (Fruhlinger 2020:para.4).

The term "smishing" was created in 2006, however, the use of smishing attacks significantly grew when the pandemic began in the early 2020s (ENZOIC n.d.:para.2). Smishing has become an effective technique used by cyber criminals (ENZOIC n.d.:para.2). As more people are working from home due to coronavirus, cybercriminals use smishing to target individuals (ENZOIC n.d.:para.4). Also, as online shopping increases, there is a rise in the number of delivery text message scams impersonating delivery companies, for example, Australia Post. Victims are tricked and click on malicious links as they may have a parcel arriving.

Current Landscape of Smishing Threats

1. Tactics

There are different tactics attackers use which include fake links, convincing phone calls, malware attacks and spear smishing (Terranova Security n.d.). A fake link involves the attacker impersonating to work for a trusting company and sends a URL link very similar to the actual organisation or company (Terranova Security n.d.). For example, the attacker asks the user to click on the link and update their personal

information, confirm a delivery of a parcel or enter a draw for a free prize (Terranova Security n.d.).

A convincing phone call involves a SMS stating to the victim to call them back, appearing to be from a government or reliable organisation (Terranova Security n.d.). The text message involves urgent language to convince the victim to call immediately and when called they sound trusting, reassuring and helpful (Terranova Security n.d.). The victim is tricked and provides the personal details needed to protect themselves from consequences (Terranova Security n.d.).

A malware attack involves a SMS containing a malicious link that installs malware on the victim's phone (Terranova Security n.d.). Usually, the malware contains Trojan Horse software where the victim's keystrokes are captured and recorded to steal passwords, banking information and contact lists (Terranova Security n.d.).

Spear smishing involves the cybercriminal conducting more extensive work and research on the victim from social media sites such as Facebook and LinkedIn (Terranova Security n.d.). Attackers use collected background information on the victim to send a targeted and specific smishing attack that appears to be trusting and reliable (Terranova Security n.d.). The smishing message looks very trusting to the victim and the victim sends their personal details to the attacker (Terranova Security n.d.).

2. Techniques

There are many different smishing techniques used by cybercriminals. Spoofing involves disguising the origin of communication, such as email addresses or phone numbers, to appear legitimate (Njuguna et al. 2022). Cybercriminals can hide their actual phone number behind a decoy and use burner phones to hide their identity (Kaspersky n.d.:para.3). The attacker pretends to be someone else or a trusting company to trick the victim to send their personal information. Any time an attacker disguises their identity as a trusting company or someone else, it is spoofing (Kaspersky n.d.:para.3). The attacker distributes the text message to bait the victim, compromises the victim's information and executes the theft by using the victim's compromised information (Kaspersky n.d.:para.3). There are different types of spoofing and these include email spoofing, IP spoofing, website spoofing, caller ID or phone spoofing, text message spoofing, ARP spoofing, DNS spoofing, GPS spoofing and facial spoofing (Kaspersky n.d.). Text message or SMS spoofing refers to when the attacker deceives users with fake displayed sender information to hide their actual identity behind an alphanumeric sender ID to impersonate a trusting company (Kaspersky n.d.). SMS spoofing is a deceptive practice where scammers alter the sender information in a text message to make it appear as if it's coming from a different source than the actual one. They can manipulate the sender's name or number to mislead the recipient. This technique is commonly used by attackers to create fake messages that appear legitimate, tricking recipients into clicking on malicious links or providing sensitive information (Alkhalil et al. 2021). SMS spoofing can be employed for various malicious purposes, including phishing attempts and

spreading malware. It's essential to stay cautious and verify the authenticity of messages from unknown sources to protect against potential scams.

Pretexting is a phishing technique that involves relying on a two-way communication, a conversation with the victim (Syafitri et al. 2022:39329). Pretexting manipulates victims into revealing their personal information by the threat actor creating a pretext (Imperva n.d.) A pretext is a made-up scenario that has a purpose to steal a victim's sensitive information, including asking for personal information to confirm receiving a prize (Imperva n.d:para.1.; Syafitri et al. 2022:39329). The threat actor uses this information to steal and carry out identity theft or secondary attacks (Imperva n.d.:para.2).

Impersonation involves a threat actor imitating behaviour of another person to try to steal personal data from a victim (Chin 2023: para.1;Imperva n.d.:para.7). For example, the impersonator tries to trick the victim into giving their sensitive information or transfer money (Chin 2023:para.1). In smishing attacks, impersonation attacks involve the threat actor sending SMS text messages including malicious links that contain viruses that could infect a victim's phone (Chin 2023:para.26). For example, the threat actor impersonates a trusted person either personal or professional that could misled the victim into believing the text's legitimacy (Chin 2023:para.26).

Financial services smishing attacks leverage urgency and scare tactics to exploit people's fears and emotions, making them highly effective and dangerous. These scams often disguise themselves as urgent notifications from well-known financial institutions, preying on the fact that almost everyone uses banking and credit card services. In these smishing attacks, attackers pose as banks or other financial entities, creating a convincing facade to commit financial fraud (Kaspersky 2023). They may use various tactics, such as urgently requesting recipients to unlock their accounts, claiming suspicious account activity that requires immediate action, or warning of potential security breaches. By instilling a sense of urgency and fear, the attackers aim to manipulate victims into clicking on malicious links or sharing sensitive information (Kaspersky 2023). Due to the alarming nature of these messages and the trust people place in financial institutions, individuals are more likely to respond hastily without verifying the authenticity of the messages. As such, staying cautious and sceptical of unsolicited messages, especially those demanding immediate action or sensitive data, is crucial to safeguard against financial services smishing scams.

Scammers send a SMS with a pretend shipment tracking code and a malicious link to update your delivery information impersonating to be from a trusting company (Puig 2020:para.2). Individuals are taken to an impersonating website, for example, Amazon where they place their personal details for their delivery (Puig 2020:para.6). Victims often fall for this scam as online shopping is commonly used due to COVID-19 and victims are expecting a parcel to arrive.

These scams include suggesting the promise of free services or products from an attacker impersonating a reputable company (Kaspersky, 2023). These gift scams include giveaway contests, free offers and shopping rewards (Kaspersky 2023).

Many victims place their details as excitement takes over when receiving “free” gifts (Kaspersky 2023). For example, providing credit card information for a free iPhone (Kaspersky 2023).

○ Statistics and Trends on Smishing Attacks

1. There was a significant increase of smishing attacks when the COVID-19 pandemic started in 2020 (Martens 2023).
2. In 2020, there were over 240,000 victims that reported experiencing phishing, smishing, vishing and pharming, costing over $US 54 million in losses (Martens 2023). Compared to 2016-2019 where there were over 166,000 victims with $US 26 billion in losses (Martens 2023).
3. Tax scams are the most common smishing attack used in the UK with 846,000 people reporting to receive fake tax messages in 2020 (Martens 2023).
4. 1.1 billion spam texts are sent every minute across the globe (Campbell 2023).
5. In 2021, over ¼ of scam texts were package delivery scams (Campbell 2023).
6. In Australia, 48% were most commonly exposed to a scam over the phone and 47% by text message in 2021-22. The percentage has doubled from 23% in 2020-21 to 47% in 2021-22 (ABS 2023).
7. The number of Australians responding to scams has decreased to 552,000. 2.7% of Australians responded to a scam in 2021-22, which has decreased from 3.6% in 2020-21 (ABS 2023). Demonstrating that Australians are reporting the scams to authorities (ABS 2023).

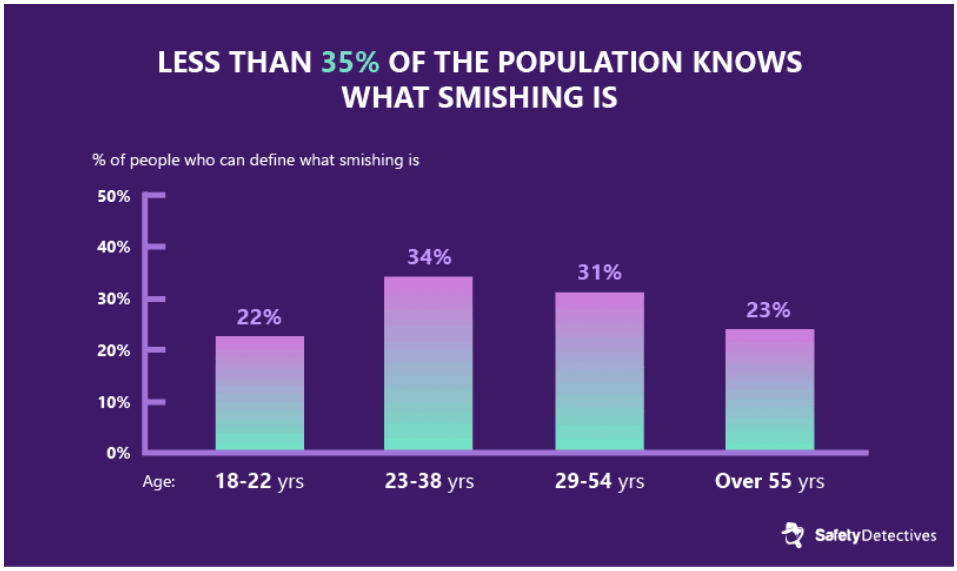

Figure 1. https://www.safetydetectives.com/blog/what-is-smishing-sms-phishing-facts/

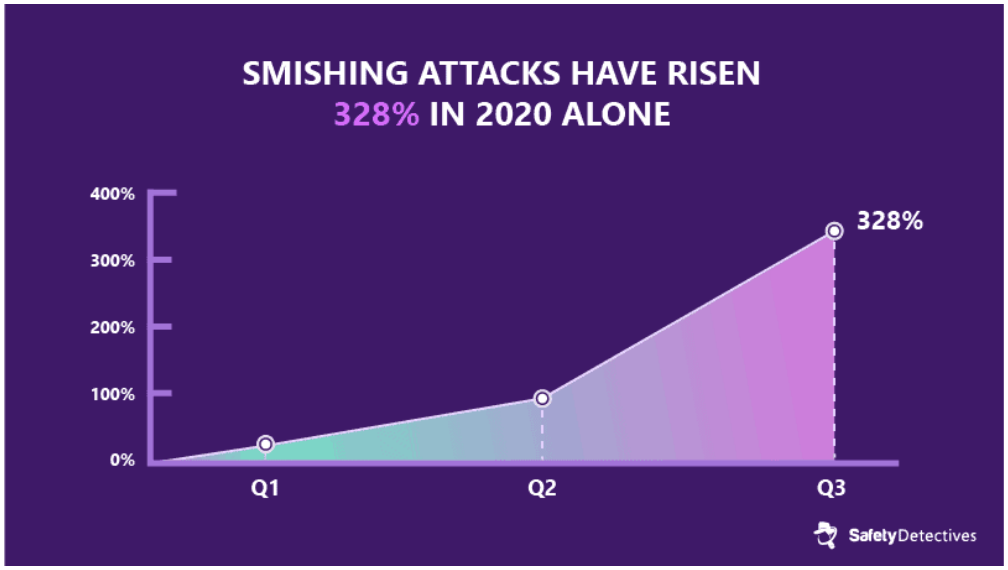

Figure 2. https://www.safetydetectives.com/blog/what-is-smishing-sms-phishing-facts/

III. Smishing Detection Techniques

○ Rule-Based Approaches

1. Keyword-based Filtering

Rule-based approaches are a fundamental component of smishing detection systems, relying on predefined rules and patterns to identify potential smishing messages. One prominent technique within this category is keyword-based filtering, which involves the systematic analysis of text messages for specific keywords or phrases associated with smishing attacks.

In keyword-based filtering, a predefined set of keywords and phrases commonly found in smishing messages is established. These keywords encompass terms related to financial transactions, personal information requests, urgent actions, and more. The smishing detection system scans incoming SMS messages for occurrences of these keywords. If a message contains a predetermined threshold of these keywords, it is flagged as a potential smishing attempt.

Keyword-based filtering is advantageous due to its simplicity and efficiency. It provides a quick means of identifying potentially malicious messages without requiring extensive computational resources. However, it has limitations, including the potential for false positives if legitimate messages contain the same keywords.

To enhance the accuracy of keyword-based filtering, researchers and developers continuously update and expand the list of keywords based on emerging smishing trends and evolving attack techniques. Moreover, keyword-based filtering can be complemented by other rule-based approaches, such as pattern matching and sender reputation analysis, for more robust detection.

2. Sender Authentication

Sender authentication involves websites, social networks, banks and apps double checking the identity of a user (Sumrak 2022:8). SMS authentication works by the user providing their phone number to a company during the sign-up process and when the user enters their username and password, they receive a one-time text verification number and type the number in to access their account (Sumrak 2022:12). SMS authentication enhances security, familiarity and affordability (Sumrak 2022:14). With SMS authentication, it is more secure than having just a password, users are used to typing these codes as it is common in all companies and the one-time codes are not expensive as it requires no additional software or hardware (Sumrak 2022:8).

3. URL Analysis

URL Analysis involves checking for presences of URLs in text messages and classifies the smishing messages (Mahmood and Hameed 2023:4247). Analysing URLs can help detect APK downloads and the URL source code is determined to see if the form tag is present in the message (Mahmood and Hameed 2023:4248). The authenticity of the SMS URL is looked at, along with textual content of the message (Mahmood and Hameed 2023:4248). The URL in the text message is scanned to determine whether it is a smishing attack or not (Mahmood and Hameed 2023:4249). URL analysis is also referred to as URL classification, which involves creating a URL classifier or by including a third-party application to establish the validity of the URL (Jain et al. 2022:11121). When a message was received, it will be scanned for the URL presence and if the URL is there, it will be forwarded to the URL phishing classifier (Jain et al. 2022:11122).

○ Machine Learning-Based Approaches

1. Supervised Learning Algorithms

a) Support Vector Machines (SVM)

SVMs are a set of supervised learning methods used for classification, regression and outliers' detection (Scikit Learn n.d.). They use classification algorithms for two-group classification problems (Stecanella 2017). Usually there are two tags, and the data has two features of x and y (Stecanella 2017). SVMs takes two data points and outputs the hyperplane, a line) that respectively separates the tags, the line is the decision boundary (Stecanella 2017). The most suitable hyperplane is the one that maximises the margins from both tags, which is whose distance to the nearest element of each tag is the greatest (Stecanella 2017).

SVMs have been used on an existing dataset to predict if a website was a trusting website or not, by funding an optimum hyperplane to separate the two groups (Anupam and Kar 2021:17).

The advantages involve being effective in high dimensional spaces, effective in cases where number of dimensions are greater than the number of samples, it is memory efficient and versatile (Scikit Learn n.d.). SVMs use a subset of training points in the decision function, which are referred to as support vectors (Scikit Learn n.d.). Also, different Kernel functions can be detailed for the decision function and common kernels are given, however, it is also likely to indicate custom kernels (Scikit Learn n.d.). SVMs have a higher speed and better performance with a limited number of samples, which makes the algorithm very appropriate for text classification problems where it is likely to have access to a dataset (Stecanella 2017). The limitations of SVMs involve preventing over-fitting in choosing Kernel functions and regularisation term is important if the number of features is much larger than the number of samples. Also, SVMs do not directly provide probability estimates as these are calculated using an expensive five-fold cross-validation (Scikit Learn n.d.).

b) Random Forest

Random Forest is an ensemble learning method that leverages a collection of decision trees to make predictions. In the context of smishing detection, each decision tree learns to classify SMS messages as either legitimate or smishing based on various features such as message content, sender information, and metadata. The Random Forest algorithm then combines the output of these individual trees to make a final prediction.

One key advantage of Random Forest is its robustness in handling high-dimensional data, which is common in smishing detection where numerous message attributes need to be considered. Additionally, Random Forest mitigates overfitting by aggregating predictions from multiple decision trees, enhancing its generalisation capabilities.

To train a Random Forest model for smishing detection, a labelled dataset containing both legitimate and smishing SMS messages is required. Feature engineering is a critical step in this process, where relevant attributes are extracted from the messages. These features can include keyword frequencies, message length, sender reputation, and temporal characteristics.

Research in this domain continually refines Random Forest models to improve their accuracy and resilience to emerging smishing threats. Additionally, efforts are made to reduce false positives and false negatives by fine-tuning the model parameters and optimising feature selection.

While Random Forest has shown promise in smishing detection, it is important to note that the effectiveness of machine learning-based approaches heavily depends on the quality and representativeness of the training data. As smishing attacks evolve, ongoing research focuses on enhancing the adaptability and real-time learning capabilities of machine learning models to stay ahead of cybercriminals.

c) Neural Networks

Neural networks, particularly deep learning architectures like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have shown promise in smishing detection. CNNs are adept at capturing spatial patterns within messages, such as specific keyword arrangements or common tricks used by attackers. RNNs, on the other hand, are proficient at handling sequential data, making them suitable for analysing the temporal aspects of SMS messages.

To employ neural networks for smishing detection, a large labelled dataset is essential. This dataset contains examples of both legitimate and smishing messages, allowing the network to learn discriminative features. Feature extraction is a critical step where the neural network processes the content, metadata, and sender information of SMS messages.

One significant advantage of neural networks is their ability to automatically learn relevant features from the data, reducing the need for manual feature engineering. This adaptability is crucial in countering novel smishing tactics.

Nevertheless, training deep neural networks necessitates substantial computational resources and extensive labelled data. Researchers continue to explore methods for efficiently training models with limited labelled data through techniques like transfer learning and semi-supervised learning.

Additionally, the interpretability of neural network models remains a challenge. Understanding why a neural network made a particular prediction can be complex, limiting its transparency and trustworthiness in certain applications.

In summary, neural networks offer a potent approach to smishing detection, leveraging their ability to capture intricate patterns in SMS messages. As smishing attacks persistently evolve, ongoing research aims to enhance the efficiency and interpretability of neural network models in safeguarding users against these deceptive threats.

2. Unsupervised Learning Algorithms

a) Clustering Techniques

Unsupervised machine learning techniques, particularly clustering algorithms, have emerged as a valuable tool in smishing detection due to their ability to identify patterns and anomalies within SMS messages without the need for labelled training data.

Clustering algorithms like K-Means and DBSCAN are applied to SMS datasets to group messages into clusters based on their similarities. These algorithms consider various features of SMS messages, such as content, metadata, and sender information, to create clusters. Smishing messages, although diverse in content, often share common characteristics, which clustering algorithms can exploit.

One of the significant advantages of unsupervised learning is its ability to uncover previously unknown patterns or trends in the data. Smishing attackers are continually evolving their tactics, making it challenging to anticipate new

attack vectors. Unsupervised techniques can adapt to these changes by identifying abnormal clusters that may indicate smishing attempts.

For instance, consider a scenario where attackers frequently use a specific set of keywords or phrases in their smishing messages. A clustering algorithm can detect a cluster of messages with similar keyword patterns, signalling a potential smishing campaign.

However, unsupervised learning is not without challenges. It may generate false positives or fail to detect subtle smishing attempts, as it relies solely on data patterns. Furthermore, the interpretability of clustering results can be complex, making it challenging to understand why a particular SMS was flagged.

To improve the performance of unsupervised smishing detection, hybrid approaches combining clustering with supervised learning have been explored. These approaches use labelled data to fine-tune clustering models and reduce false positives.

In conclusion, unsupervised machine learning techniques, particularly clustering algorithms, offer a valuable approach to smishing detection. Their ability to identify patterns and anomalies in SMS messages makes them well-suited for countering evolving smishing threats. Nevertheless, their effectiveness can be enhanced when combined with supervised learning and ongoing research in this field aims to address their limitations.

b) Anomaly Detection

Unsupervised machine learning, particularly anomaly detection algorithms, plays a crucial role in smishing detection by identifying suspicious patterns or outliers in SMS messages without requiring prior labelled data.

Anomaly detection algorithms, such as Isolation Forests and One-Class SVMs, are employed to create a baseline model of normal SMS behaviour. These models learn to recognize the typical characteristics of legitimate messages, including message content, sender behaviour, and timing. Any SMS message that deviates significantly from this learned baseline is flagged as a potential smishing attempt.

One of the significant advantages of anomaly detection in smishing detection is its adaptability to new and previously unseen attack tactics. As smishing attackers frequently change their strategies, a well-trained anomaly detection model can detect unfamiliar patterns that may indicate a smishing campaign in real-time.

For example, if a user typically receives SMS messages from known contacts during daytime hours, an anomaly detection model can identify nighttime messages from unfamiliar senders as potential smishing attempts, even if the content appears innocuous.

However, unsupervised anomaly detection may have challenges, including the potential for false positives or false negatives. It relies heavily on the quality of the learned baseline, and unusual but legitimate behaviours could be flagged as

anomalies. Achieving the right balance between sensitivity and specificity is crucial.

To address these challenges, hybrid approaches combining unsupervised anomaly detection with supervised learning have been explored. These approaches use labelled data to refine anomaly detection models and reduce false alarms.

In conclusion, unsupervised machine learning techniques, particularly anomaly detection algorithms, offer a valuable approach to smishing detection. Their ability to identify anomalies in SMS message patterns makes them well-suited for countering evolving smishing threats. However, ongoing research in this field focuses on improving their accuracy and robustness.

3. Hybrid Models

Hybrid smishing detection models combine multiple techniques, often leveraging both rule-based and machine learning approaches, to enhance the accuracy and effectiveness of smishing detection.

One common hybrid approach involves using rule-based filtering as a preprocessing step. In this method, predefined rules are applied to incoming SMS messages to flag those that exhibit suspicious characteristics. These rules can include keyword matching, sender reputation analysis, and URL scanning. Messages that trigger any of these rules are then subjected to more advanced machine learning algorithms for further analysis.

Another hybrid strategy combines supervised and unsupervised machine learning techniques. Initially, a supervised model is trained on labelled data to identify known smishing patterns. This model can efficiently detect familiar attack vectors. However, to address novel smishing attacks, an unsupervised anomaly detection component is integrated. This unsupervised module identifies deviations from normal SMS behaviour and flags messages that exhibit unusual characteristics. By combining these methods, hybrid models can adapt to both known and emerging threats.

Moreover, hybrid models can incorporate ensemble learning, which combines predictions from multiple individual models. This technique often leads to improved detection accuracy by leveraging the strengths of different algorithms. For instance, an ensemble model may include decision trees, support vector machines, and neural networks. By combining their outputs, the model can better discern smishing attempts from legitimate messages.

While hybrid models can be highly effective, they require careful design, integration, and parameter tuning to achieve optimal performance. Moreover, these models need continuous updates to adapt to evolving smishing attack tactics.

In conclusion, hybrid smishing detection models represent a sophisticated approach to combating smishing threats. By combining rule-based, supervised, unsupervised, and ensemble techniques, these models can provide robust protection against a wide range of smishing attacks.

- # Natural Language Processing (NLP) in Smishing Detection

1. Text Analysis and Feature Extraction

a) Tokenization:

- Tokenization is the process of breaking down text messages into individual words or tokens. This step is fundamental for further analysis as it enables the identification of critical elements within the text.

b) Keyword Analysis:

- Keyword-based approaches involve the identification of specific words or phrases commonly associated with smishing, such as "urgent," "verify," or "bank." The presence and context of these keywords are analysed to assess the likelihood of a message being a smishing attempt.

c) Feature Extraction:

- Feature extraction techniques are applied to extract relevant information from text messages. This can include the frequency of specific keywords, the presence of hyperlinks, or the syntactic structure of the message. These features serve as input for machine learning models.

d) Content Similarity:

- Content similarity measures are employed to compare the text message with known smishing templates. By quantifying the similarity between the incoming message and a database of known smishing messages, potential threats can be identified.

e) Machine Learning Integration:

- Machine learning algorithms, such as support vector machines (SVM) or deep learning models, can be trained using the extracted features to classify text messages as either legitimate or smishing attempts. These algorithms rely on the patterns and features identified during text analysis.

These techniques, centred around text analysis and feature extraction, are instrumental in enhancing smishing detection systems' ability to analyse incoming messages effectively and differentiate between legitimate and malicious content.

2. Sentiment Analysis

Sentiment analysis is a valuable component of smishing detection, leveraging natural language processing (NLP) to assess the emotional tone and intent behind text messages. It plays a crucial role in distinguishing between legitimate and malicious messages by analyzing the sentiments expressed. Here's how sentiment analysis is applied in smishing detection:

a) Sentiment Polarity:

- Sentiment analysis categorises text into sentiment polarities such as positive, negative, or neutral. In smishing detection, a highly negative sentiment in a message might indicate a potential threat. For example, messages conveying urgency or fear may raise suspicion.

b) Contextual Understanding:
   - Sentiment analysis helps in understanding the context of a message. Malicious smishing messages often use fear or urgency to manipulate recipients. Sentiment analysis can identify emotionally charged language used to coerce individuals.

c) Anomaly Detection:
   - By establishing a baseline of normal sentiment distribution within a dataset, any significant deviations can trigger an alert. If a message contains an unusually negative sentiment compared to the norm, it may be flagged as suspicious.

d) Keyword-Sentiment Relationship:
   - Sentiment analysis can be combined with keyword analysis. For instance, identifying negative sentiments associated with terms like "verify" or "immediately" in a financial context could be indicative of a smishing attempt.

e) Machine Learning Models:
   - Machine learning algorithms can be trained on labelled datasets to predict the sentiment of text messages. These models can then be used to automatically classify incoming messages as suspicious or benign based on their sentiment.

Incorporating sentiment analysis into smishing detection enhances the system's ability to identify emotionally manipulative content, making it a valuable tool in the fight against smishing attacks.

3. Contextual Understanding

Contextual understanding, a vital aspect of natural language processing (NLP), is increasingly being employed in smishing detection to decipher the meaning and intent behind text messages. It enables systems to identify subtle nuances and contextual cues that distinguish legitimate messages from smishing attempts.

a) Contextual Anomaly Detection:
   - Contextual understanding helps in detecting anomalies within a message's context. For example, a message supposedly from a bank that lacks typical banking terminology or context can be flagged as suspicious.

b) Semantic Analysis:
   - NLP techniques can analyse the semantics of a message to determine its authenticity. Understanding the intended meaning of words and phrases allows systems to identify messages that use language deceptively.

c) Conversation Flow Analysis:
   - By examining the flow of a conversation, contextual understanding can identify abrupt changes or illogical progressions. Smishing messages often disrupt the natural flow of communication, which can trigger alerts.

d) Phishing Patterns:
   - Recognizing known phishing patterns in text messages is another application. Systems can be trained to identify smishing attempts that replicate common techniques used in phishing emails.

e) Intent Recognition:
   - Contextual understanding can assess the intent behind a message. For instance, messages that request sensitive information or immediate action, especially without prior context, may be flagged as suspicious.

f) User Profiling:
   - Analysing a user's historical messaging patterns and the context of their conversations can help in detecting anomalies. If a message deviates significantly from a user's typical communication style, it may be flagged for review.

g) Machine Learning Models:
   - Machine learning models, particularly deep learning techniques, can be trained to understand context. Recurrent neural networks (RNNs) and transformer-based models, like BERT, excel in capturing contextual information in text.

Incorporating contextual understanding into smishing detection empowers systems to grasp the intricacies of language and communication. This holistic approach enables the identification of smishing attempts that rely on contextual deception, bolstering the overall security of digital communications.

IV. State-of-the-Art Smishing Detection Systems
- Overview of Prominent Smishing Detection Solutions
- Comparative Analysis of Leading Smishing Detection Tools
- Performance Metrics Evaluation

V. Data Collection and Preprocessing
- Data Sources and Types
- Data Preprocessing Techniques
- Data Augmentation Strategies

VI. Experimental Setup
- Description of the Dataset
- Feature Selection and Engineering
- Model Training and Evaluation

VII. Results and Discussion
- Performance Comparison of Smishing Detection Techniques
- Analysis of False Positives and False Negatives
- Discussion of Key Findings and Observations

VIII. Challenges and Limitations
- Inherent Challenges in Smishing Detection
- Limitations of Existing Techniques
- Ethical and Privacy Considerations

IX. Future Directions and Research Opportunities
- Emerging Trends in Smishing Attacks
- Potential Enhancements to Detection Techniques
- Integration of AI and Advanced Technologies

X. Conclusion
- Summary of Research Findings
- Importance of Continuous Smishing Detection Improvement
- Implications for Future Cybersecurity Strategies

XI. References

- Alkhalil Z, Hewage C, Nawaf L and Khan I (2021) "Phishing attacks: A recent comprehensive study and a new anatomy", Frontiers, accessed 11 September 2023.
- Anupam S and Kar AK (2021) "Phishing website detection using support vector machines and nature-inspired optimization algorithms", Telecommunication Systems, no. 76, pp. 17-32.
- Australian Bureau of Statistics (ABS) (22 Feb 2023) "13.2 million Australians exposed to scams", Australian Bureau of Statistics, accessed 13 September 2023.
- Campbell S (29 July 2023) "Smishing Statistics 2023: 17 Key SMS Phishing Facts!", The Small Business Blog, accessed 11 September 2023.
- Chin K (2023) "What is an Impersonation Attack?", UpGuard, accessed 11 September 2023.
- ENZOIC (n.d.) "The Rise of Smishing", ENZOIC, accessed 11 September 2023.
- Fruhlinger J (2020) "What is smishing? How phishing via text message works", CSO, accessed 11 September 2023.
- Graphus (4 April 2023) "The History of Phishing", Graphus, accessed 9 September 2023.
- Vidas, T., Bursztein, E., & Savage, S. (2011). All your iThings are belong to us: Breaking mobile devices for botnet and profit. In Proceedings of the 4th USENIX conference on Offensive Technologies (pp. 1-16).
- Canfora, G., Mercaldo, F., & Visaggio, C. A. (2017). Detecting SMS spam using a keyword-based approach. IEEE Transactions on Information Forensics and Security, 12(11), 2650-2663.
- Kumar, M., & Srivastava, A. (2019). SMS spam detection using keyword-based approach. Procedia computer science, 132, 651-656.
- Imperva (n.d.) "Pretexting", Imperva, accessed 11 September 2023.
- Jain AK, Gupta BB, Kaur K, Bhutani P, Alhalabi W and Almomani A (2022) "A content and URL analysis-based efficient approach to detect smishing SMS in intelligent systems", International Journal of Intelligent Systems, vol. 37, no.12, pp. 11117 – 11141.
- Kaspersky (n.d.) "Smishing meaning and definition", Kaspersky, accessed 11 September 2023.
- Mahmood AR and Hameed SM (2023) "Review of Smishing Detection Via Machine Learning", Iraqi Journal of Science, vol. 64, no.8, pp. 4244-4259.
- Martens B (7 June 2023) "11 Facts & Stats on Smishing (SMS Phishing) in 2023", Safety Detectives, accessed 11 September 2023.
- Njuguna, D.N., Kamau, J. and Kaburu, D. (2022) 'A review of Smishing Attacks Mitigation Strategies', International Journal of Computer and Information Technology (2279-0764), 11(1). doi:10.24203/ijcit.v11i1.201. Accessed 30 July. 2023.
- Phishing (n.d.) "History of Phishing", Phishing.org, accessed 11 September 2023.
- Puig A (20 Feb 2020) "Is that text message about your FedEx package really a scam?", Federal Trade Commission Consumer Advice, accessed 11 September 2023.
- Scikit Learn (n.d.) "1.4 Support Vector Machines", scikit learn, accessed 17 September 2023.

- Stecanella B (23 June 2017) "Support Vector Machines (SVM) Algorithm Explained", MonkeyLearn, accessed 18 of September 2023.
- Abid, A., Mirza, A. M., & Akram, S. (2018). Detection and prevention of smishing attacks using machine learning. In Proceedings of the 2018 International Conference on Frontiers of Information Technology (FIT) (pp. 58-62).
- Rahim, M. M., Hu, J., & Ahmadi, M. (2019). Smishing detection using machine learning techniques. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 3421-3426).
- Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.
- Awoyemi, M. A., Khan, S. U., & Ahmed, M. (2018). A deep learning approach for network intrusion detection system. Journal of Ambient Intelligence and Humanized Computing, 9(1), 63-75.
- Roni, D. S., Jahan, S., & Rahman, M. M. (2020). A deep learning approach for smishing detection using long short-term memory networks. In Proceedings of the 12th International Conference on Machine Learning and Computing (pp. 87-93).
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
- Zhang, Z., Chen, Y., Zhou, Y., & Lee, S. (2019). Detecting SMS phishing (smishing) activities on Android phones. In Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security (AsiaCCS) (pp. 83-96).
- Jain, K., & Neelu, R. (2019). SMS phishing detection using clustering and ensemble learning. In Proceedings of the International Conference on Advances in Computing and Data Sciences (ICACDS) (pp. 1-6).
- Arthur, D., & Vassilvitskii, S. (2007). K-means++: The advantages of careful seeding. In Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms (pp. 1027-1035).
- Sumrak J (2022) "SMS Verification: What It Is and How It Works", Twilio, accessed 27 August 2023.
- Syafitri W, Shukur Z, Mokhtar UA, Sulaiman R and Ibrahim MA (2022) 'Social Engineering Attacks Prevention: A Systematic Literature Review', IEEE Access, vol. 10, pp. 39325 – 39343.
- Weng, P., Zhang, Y., & Zhang, Z. (2016). Detecting SMS spam in the wild: A case study. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS) (pp. 1359-1370).
- Aggarwal, C. C. (2015). Outlier Analysis (2nd ed.). Springer.
- Ruff, L., Vandermeulen, R., & Schölkopf, B. (2018). A scalable approximation of domain over-similarity for scalable outlier detection. In Proceedings of the 35th International Conference on Machine Learning (ICML) (Vol. 80, pp. 4282-4291).
- Damerau, F. J., & Mays, E. (2020). Text Mining and NLP for Smishing Detection. In Proceedings of the 2020 International Conference on Machine Learning and Natural Language Processing (pp. 34-40).
- Kumar, N., & Chauhan, D. (2018). Machine Learning and NLP Based SMS Phishing Detection System. In Proceedings of the International Conference on Computing, Power and Communication Technologies (GUCON) (pp. 108-114).

- Ahmad, A., Ahmed, E., Hu, J., & Hu, J. (2017). Network Traffic Classification: A Hybrid Approach. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 1353-1360).
- Chen, X., He, X., & Zhang, Z. (2012). Detecting SMS spam: An iterative and high-precision approach. In Proceedings of the 21st International Conference on World Wide Web (WWW) (pp. 743-752).
- Kwon, S., Lee, K., Han, D., & Lee, S. (2013). Efficient spam filtering of short message service using data reduction techniques. Information Sciences, 219, 211-227.
- Barshan, E., & Ghaleb, M. A. (2020). Machine learning for detecting phishing scams. IEEE Transactions on Network and Service Management, 17(3), 1732-1745.
- Terranova Security (n.d.) "WHAT IS SMISHING?", FORTRA, accessed 11 September 2023.

XII. Appendix (Optional)

- Dataset Details
- Code Implementation (if applicable)
- Additional Figures and Tables