# Comprehensive Research on Strategies and Tactics of Defenders and Attackers Based on Proper Threat Models in Smishing Detection

HardHat Enterprises

# Table of Contents

# I. Introduction

## A. Overview of Smishing and its Threat Landscape

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!

## B. Importance of Understanding Strategies and Tactics

In the ever-evolving landscape of cyber threats, smishing poses a significant risk to individuals and organisations alike. Both defenders and attackers continuously refine their strategies to gain an advantage over one another (Alkhalil et al., 2021). Defenders diligently apply security patches and measures to thwart smishing attempts, while attackers craft compelling scripts and social engineering tactics to deceive users successfully. This dynamic interplay underscores the crucial importance of understanding these strategies (Alkhalil et al., 2021).

Defenders employ an array of tactics to safeguard against smishing attacks. User education and awareness programs play a pivotal role, enlightening users about common smishing ploys like fake delivery notifications and urgent alerts. SMS filtering and spam detection mechanisms help identify and block known smishing messages before they reach users’ devices, and URL analysis tools scrutinise links to prevent access to malicious websites (Kaspersky, 2023). Encouraging users to enable two-factor authentication (2FA) adds an extra layer of protection, even if credentials are compromised. Additionally, establishing a robust incident response plan allows defenders to swiftly respond and mitigate smishing attacks while empowering users to report suspicious messages for further investigation.

On the other side of the spectrum, attackers employ cunning tactics to exploit unsuspecting victims. Social engineering remains a prominent strategy, exploiting psychological triggers to create a false sense of urgency or curiosity that drives users to take hasty actions. SMS spoofing conceals the true sender’s identity, lending an air of legitimacy to deceptive messages. URL shorteners help disguise malicious links, making them more enticing to click on (Mishra and Soni, 2019). Some attackers resort to using pre-built phishing kits, streamlining the process of launching large-scale smishing campaigns. Furthermore, attackers continuously evolve their tactics, adapting their messages and employing novel social engineering tricks to outsmart defenders and bypass security measures (Mishra and Soni, 2019).

In conclusion, comprehending both defender and attacker strategies is essential in building robust defence mechanisms against smishing attacks. By staying informed and vigilant, defenders can proactively protect users from falling victim to these insidious scams. An ongoing partnership between testing teams and security professionals ensures the deployment of effective countermeasures, safeguarding the digital community against the evolving threats of smishing.

## C. Objectives of the Report

This report aims to comprehensively explore the landscape of smishing attacks by investigating both the strategies employed by attackers and the defensive measures utilised by security professionals. The primary objectives include:

1. Understanding Attacker Tactics:
   - To dissect and elucidate the tactics and techniques employed by smishing attackers, shedding light on their evolving methodologies, motivations, and targets.
2. Evaluating Defensive Strategies:
   - To critically evaluate the effectiveness of defence strategies against smishing attacks, assessing key performance indicators (KPIs), metrics, and benchmarking against industry standards.
3. Ethical Considerations:
   - To highlight the ethical dimensions within smishing defence, emphasising user privacy, transparency, non-discrimination, data management, and responsible disclosure.
4. Recommendations for Improvement:
   - To provide actionable recommendations for both defenders and organisations to enhance their smishing detection and prevention strategies, bolstering overall cybersecurity resilience.
5. Contributions to the Field:
   - To contribute valuable insights and knowledge to the field of cybersecurity, aiding in the development of more robust and adaptive defence mechanisms against smishing attacks.

By addressing these objectives, this report aspires to offer a comprehensive and multifaceted understanding of smishing attacks and defence strategies, equipping both practitioners and researchers with the knowledge needed to combat this evolving threat effectively.

# II. Threat Models in Smishing Detection

## A. Definition and Construction of Threat Models

Threat models in smishing are structured processes that recognise security requirements, identify security threats and potential vulnerabilities, quantify threat and vulnerability cruciality and categorise remediation methods (Synposys n.d.). Threat models provide a clear line of sight across a project and helps justify security efforts (Synopsys n.d.).

## B. Understanding the Attack Surface

Proactive measures play a major role in deepening our understanding of the smishing attack surface. By examining the tactics employed by smishing attackers and the vulnerabilities they exploit, we can develop a comprehensive awareness of the threats posed.  User Education and Awareness: By educating users about common smishing tactics and the psychological principles behind them, we enhance their ability to recognise deceptive messages (Kaspersky, 2023). This awareness equips users with insights into how attackers manipulate trust, context, and emotion to coerce actions.

### 1. SMS Filtering and Spam Detection

Implementing SMS filtering and spam detection mechanisms provides insights into the volume and variety of smishing attempts. Analysing intercepted messages offers valuable insights into the evolving tactics used by attackers, helping us identify emerging trends (Kaspersky, 2023).

### 2. URL Analysis and Validation

Incorporating URL analysis tools provides a window into the techniques employed by attackers to deceive recipients. By scrutinising links within messages, we gain insights into attackers' attempts to redirect users to fraudulent websites, which aids in understanding their techniques.

### 3. Two-Factor Authentication (2FA)

Encouraging the use of two-factor authentication highlights the vulnerabilities that attackers target (Kaspersky, 2023). Understanding that attackers seek to bypass such additional security layers informs us about their persistence and determination to breach accounts.

### 4. Incident Response Planning

Developing incident response plans requires a comprehensive grasp of the potential attack vectors, victim demographics, and attack success rates. This understanding helps us craft effective strategies to mitigate the impact of successful smishing attacks.

## C. Identification of Potential Adversaries

Attackers can have varied motives and objectives as the following profiles show.

### 1. Fraudsters and Cybercriminals

These are individuals or groups primarily motivated by financial gain. They aim to trick victims into revealing sensitive information such as login credentials, credit card details, or personal identification, which they then exploit for monetary benefit.

### 2. Amateur Attackers or Pranksters

These individuals might be curious, wanting to test their skills, or show off to their friends. They could include teenagers or young adults experimenting with hacking techniques without serious malicious intent. Their actions may be driven by the desire to prove their skills or simply to cause disruption.

### 3. Hacktivists

Some attackers might have political, ideological, or social motivations. They use smishing campaigns to spread their message, gather support for their cause, or to create chaos within specific organisations or industries.

### 4. Nation-State Actors

State-sponsored attackers can engage in smishing for espionage, intelligence gathering, or as part of broader cyber operations. Their motivations could include stealing sensitive information, disrupting critical infrastructure, or undermining national security.

### 5. Competitors and Rival Organisations

Businesses may resort to smishing as a way to gain a competitive edge. This could involve trying to steal sensitive business data, intellectual property, or proprietary information from competitors.

### 6. Advertising and Marketing

While not malicious in intent, some entities might employ smishing as an unconventional way of reaching potential customers. This could involve sending unsolicited promotional messages or offers to a wide audience.

## 7. Insiders or Employees

Attacks might be orchestrated by individuals within an organisation seeking to exploit vulnerabilities for personal gain, revenge, or to harm the organisation’s reputation.

## 8. Social Engineers and Manipulators

These attackers exploit psychological vulnerabilities to manipulate individuals into taking specific actions. They could aim to gain trust, extract information, or encourage behaviours that serve their objectives.

## 9. Copycat Attackers

These individuals or groups mimic the tactics of known attackers to create confusion and divert attention from the actual perpetrators. Their goal might be to deflect blame or create a smokescreen.

It’s important to note that smishing campaigns can be initiated by a wide range of actors with diverse motivations. Each profile has its unique attributes and potential risks. Understanding these potential attacker profiles helps us anticipate the types of threats we might encounter and develop effective strategies to mitigate them.

# III. Attacker's Strategies and Tactics

## A. Social Engineering Techniques in Smishing Attacks

Investigate the techniques attackers employ to manipulate human behaviour and emotions, particularly focusing on pretexting, impersonation, emotional manipulation, and urgency tactics.

- **Social Engineering:** Attackers rely heavily on social engineering tactics to trick users into believing their smishing messages are legitimate. They use psychological tactics to create a sense of urgency or curiosity, encouraging victims to take immediate action (Alkhalil et al., 2021).
- **Spoofing:** Attackers often use SMS spoofing techniques to mask the true sender’s identity, making it appear as though the message comes from a trusted source.
- **URL Shorteners:** To hide the actual destination of malicious links, attackers may use URL shorteners to make the links appear less suspicious and more enticing to users.
- **Phishing Kits:** Some attackers use pre-built phishing kits to automate the creation and deployment of smishing campaigns, making it easier to launch attacks on a large scale (Alkhalil et al., 2021).
- **Evolving Tactics:** Like defenders, attackers continually adapt their tactics to stay ahead. They may modify their messages, use new social engineering tricks, or employ different methods to bypass security measures.

### 1. Pretexting and Impersonation

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).

A pretexting attack technique used by threat actors includes impersonation (Imperva n.d.:para.7). 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 transferring 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).

### 2. Emotional Manipulation

Emotional manipulation involves threat actors using many techniques to influence the victim to share sensitive information by triggering the victim's emotions (Razorthorn n.d.:para.8). The threat actor creates a personal connection with the victim and builds trust to succeed (Razorthorn n.d.:para.8). The threat actor can trick the victim's emotions of obedience, fear, lust, kindness, anger and curiosity to gain sensitive information (Razorthorn n.d.).

### 3. Urgency and Scare Tactics

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.

## B. Spoofing and Phishing Techniques

Explore the different spoofing and phishing techniques used by attackers, both in terms of disguising communication origins and creating fraudulent messages to deceive recipients.

Spoofing and phishing techniques are deceptive practices used by cybercriminals to trick individuals into revealing sensitive information or performing harmful actions. Spoofing involves disguising the origin of communication, such as email addresses or phone numbers, to appear legitimate (Njuguna et al., 2022). Phishing, on the other hand, involves sending fraudulent messages that mimic reputable sources like banks or well-known companies to lure victims into providing personal data, passwords, or financial details (Njuguna et al., 2022).

Common spoofing and phishing techniques include email spoofing, where the sender's address is forged to appear genuine, and website spoofing, where fake websites imitate legitimate ones to steal login credentials. Additionally, vishing (voice phishing) employs

fake phone calls to manipulate victims, while smishing (SMS phishing) uses deceptive text messages to trick users into clicking malicious links or sharing sensitive data.
To protect against these threats, it’s crucial to be vigilant and cautious when receiving unsolicited messages or calls. Verify the authenticity of requests for sensitive information and never click on links or download attachments from unknown sources. Employing strong passwords, enabling two-factor authentication, and staying informed about the latest phishing and spoofing tactics are essential in safeguarding against these cyber threats (Njuguna et al., 2022).

## 1. Caller ID Spoofing

Examine the technique of caller ID spoofing, its history, potential consequences, and how it's exploited by scammers.
Caller ID spoofing is a deceptive technique used by scammers to manipulate the phone number that appears on your caller ID display. They can make it look like they are calling from a different phone number than their actual one. This trick has been around for years, and some companies have offered services to do this legally. However, scammers often use it to deceive people, such as when pretending to be interested buyers on online marketplaces like eBay or MarketPlace. They may request personal information from sellers and then use it for fraudulent purposes, like reposting the items for fake sales (Kaspersky, 2023).

## 2. SMS Spoofing

Investigate SMS spoofing, its mechanics, and the methods attackers use to manipulate sender information in text messages for malicious purposes.
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.

## 3. Phishing URLs in Messages

Explore the challenges posed by phishing URLs in messages, especially in the context of AI-powered tools enabling scammers to craft deceptive login pages with ease.
Phishing attacks have evolved with the advent of AI-powered tools, posing new challenges for online security. In Zscaler’s ThreatLabs 2023 Phishing Report, it was revealed that AI tools, such as ChatGPT, can be leveraged by scammers to craft deceptive login pages with ease, even with minimal coding expertise or user input

(Cook, 2023). This concerning development enables the creation of sophisticated and constantly changing phishing URLs, making it harder for traditional security measures to keep pace.
With the potential to generate polymorphic malware and other malicious code, the impact of AI in the hands of cybercriminals cannot be underestimated. As these advanced techniques continue to emerge, staying vigilant and adopting robust security practices becomes imperative to thwart the ever-evolving threats posed by phishing URLs in messages (Cook, 2023).

## C. Exploitation of Human Behavior and Psychological Biases

### 1. Trust Exploitation

Trust exploitation involves the threat actor taking advantage of the victim's trust to gain unauthorised access to the victim's network for their sensitive information (Cisco Certified Expert 2023). The threat actor may impersonate a trusted organisation or person to gain the victim's trust (Cisco Certified Expert 2023). The threat actor uses the trusted relationship to gain access to sensitive information (Cisco Certified Expert 2023). Also, trust exploitation involves a threat actor taking advantage of a trust relationship within a network, and a common example is a perimeter network connection from a corporation (Cisco Certified Expert 2023). This example involves the network segments of DNS, SMTP and HTTP servers and because they all exist on the same section, a compromise of one system can lead to the compromise of other systems (Cisco Certified Expert 2023). The compromise occurs as they trust other systems attached to their same network (Cisco Certified Expert 2023).

### 2. Curiosity Exploitation

Curiosity exploitation involves the threat actor using the trait curiosity to manipulate the victim (Razorthorn n.d.:para.16). For example, the threat actor sends a text message promising the victim for something interesting to trick the victim to hand their personal data to the threat actor (Razorthorn n.d:para.16).

### 3. Fear Exploitation

Fear exploitation involves the threat actor using the emotion fear to manipulate victims to do things they would not think of doing (Razorthorn n.d.:para.12). Fearful victims are influenced to hand over personal data to remove themselves from the position that is  causing them to feel the emotion (Razorthorn n.d.:para.12). For example, the threat actor sends a text stating “Your bank account has been compromised and you need to transfer money in a safe area to ensure no more money is stolen from your account” and the victim feels fear and confusion (Razorthorn n.d.:para.12).

## D. Evasion and Persistence Mechanisms

The realm of smishing attacks presents a landscape where attackers employ ingenious strategies to bypass traditional detection systems while ensuring their campaigns persist over time. This subheading delves into two crucial aspects: evading traditional detection systems and the utilisation of persistence mechanisms.

### 1. Evading Traditional Detection Systems

Attackers have demonstrated a remarkable ability to evade conventional detection systems designed to identify malicious content. Techniques such as obfuscation, encryption, and content manipulation allow them to craft smishing messages that remain undetected by signature-based defences. Moreover, by camouflaging their messages within seemingly benign or legitimate contexts, attackers can bypass keyword-based filters. The exploitation of zero-day vulnerabilities and the rapid evolution of attack tactics also play significant roles in their evasion strategies.

### 2. Persistence and Recurring Attacks

The sustained success of smishing campaigns is often rooted in the deployment of persistence mechanisms. Attackers employ a variety of techniques to ensure their malicious presence endures over time. One notable strategy involves leveraging compromised systems as proxies, enabling them to maintain communication channels with the victim even after initial compromise. Additionally, employing a botnet infrastructure allows attackers to orchestrate attacks from a distributed network of compromised devices, rendering the takedown process more challenging.

The evasion and persistence mechanisms employed by attackers in the smishing landscape underscore the need for dynamic and adaptive defense strategies. As attackers continually evolve their techniques, defenders must anticipate these strategies, implement advanced detection mechanisms, and establish countermeasures that can effectively disrupt both evasion tactics and the sustainability of smishing campaigns.

# IV. Defender's Strategies and Tactics

Analyse the various tactics employed by defenders to mitigate and respond to smishing attacks, including user education, SMS filtering, URL analysis, two-factor authentication, and incident response plans.

- **URL Analysis:** Security tools may inspect links within SMS messages to determine if they lead to known malicious websites or phishing domains. Suspicious links are flagged or blocked.
- **Two-Factor Authentication (2FA):** Encouraging users to enable 2FA for their accounts adds an extra layer of protection even if their credentials are compromised through smishing.
- **Incident Response and Reporting:** Establishing an incident response plan allows defenders to quickly react to and mitigate smishing attacks. Users are also encouraged to report suspicious messages to their mobile carriers or relevant authorities.

## A. Establishing Strong Authentication Mechanisms

### 1. MFA

Multi-factor authentication (MFA) is a security method that needs two or more pieces of evidence of identity to grant you access (ACSC n.d.:para.1). Along with the password, users will be asked a combination of something the user knows, including password recovery questions or something you are, including biometrics, such as fingerprints (AWS n.d.para.1; ACSC n.d.para.2).
MFA has benefits of reducing security risks, enabling digital initiatives and improving security responses (AWS n.d.:para.3). This method acts as an extra layer of security to prevent unauthorised users from accessing victims accounts (AWS n.d.:para.3).
MFA offers  substantially more protection and security against cyber criminals as they need to obtain biometrics to gain access into a personal account (ACSC n.d.para.3).

### 2. Sender Authentication and Verification

Sender authentication and verification involves websites, social networks, banks and apps double checking the identity of a user (Sumrak 2022:8). SMS verification 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 verification and 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).

## B. User Education and Awareness Training

Defenders often conduct user education programs to raise awareness about smishing risks. They inform users about common smishing tactics, such as fake package delivery notifications or urgent alerts, and advise them not to click on suspicious links or share sensitive information (Alkhalil et al., 2021).

### 1. Recognizing Smishing Attempts

There are many ways to recognise smishing attempts including a message containing:

- Suspicious links
- Congratulations on winning a prize for a competition you did not enter
- A file that you were not expecting
- The name of a bank you use
- A demanding request to verify personal information via a link
- A demanding plea for help, asking for money (Avira 2020:para.6).

### 2. Safe Practices for Handling Messages

- Do not respond to the message
- Call your bank to confirm that the message is not real
- Do not click the suspicious links
- Check the phone number to see if it is odd-looking
- Use MFA
- Download an anti-malware app to protect against malicious apps
- Report all spam and smishing text messages (Kaspersky n.d.)

## C. Behavioral Analysis and Anomaly Detection

In the unending battle against smishing attacks, defenders have turned to behavioural analysis and anomaly detection techniques as potent tools to fortify their arsenals. This subheading delves into two significant aspects of this defence strategy: user behaviour profiling and the detection of deviations from normal patterns.

### 1. User Behaviour Profiling

Behavioural analysis involves studying patterns of legitimate user interactions with SMS messages. By establishing comprehensive profiles of users' communication habits, including messaging frequency, contacts, and the time of interactions, defenders gain a nuanced understanding of typical behaviour. These profiles act as baselines against which incoming SMS messages are evaluated. Any deviations from established norms are flagged for further scrutiny, potentially indicating a smishing attempt. This approach capitalises on the attackers' requirement to deviate from established behavioural patterns, enhancing the potential to detect even sophisticated attacks.

### 2. Detecting Deviations from Normal Patterns

Detecting anomalies involves constantly comparing incoming SMS messages with user behaviour profiles. Unusual times of communication, contact deviations, and content incongruities are among the indicators that an SMS might be malicious. Leveraging statistical and machine learning techniques, defenders can quantify the level of anomaly and trigger alerts for further investigation. These approaches enhance the detection of smishing attempts that might otherwise remain concealed amidst legitimate traffic.

In light of these strategies, defenders are adopting proactive measures to anticipate and counteract smishing attacks. User behaviour profiling and anomaly detection augment traditional signature-based methods, offering a more nuanced and context-aware approach to identifying potential threats.

## D. Real-Time Threat Intelligence and Collaboration

The evolving landscape of smishing threats necessitates defender strategies that embrace real-time threat intelligence and collaboration efforts. This subheading delves into two critical components: integrating threat feeds and intelligence, and collaborative endeavours with industry and law enforcement.

### 1. Integrating Threat Feeds and Intelligence

Defenders are leveraging real-time threat intelligence platforms that aggregate data from various sources, including honeypots, sandbox analysis, and global threat databases. By constantly updating their knowledge about ongoing and emerging smishing campaigns, defenders can swiftly identify patterns, tactics, and indicators of compromise. This intelligence is integrated into detection systems, enabling quicker response to new attack vectors.

### 2. Collaborative Efforts with Industry and Law Enforcement

Collaboration stands as a cornerstone of smishing defence. Defenders are increasingly joining forces with other organisations, both within the industry and law enforcement agencies. Information sharing facilitates the exchange of insights, tactics, and mitigation strategies. When defenders collectively analyse and disseminate threat data, they can establish more comprehensive defence mechanisms and quickly implement countermeasures against the most recent smishing attacks.

These strategies showcase defenders' commitment to staying ahead in the race against smishing attackers. By harnessing real-time threat intelligence and fostering collaboration, defenders enhance their adaptive capacities and amplify their ability to respond effectively to the ever-changing smishing threat landscape.

# V. Case Studies: Real-World Smishing Attacks

## A. Analysis of Notable Smishing Incidents

### 1. Twilio

In 2022, Twilio, the customer engagement platform, disclosed a data breach where a sophisticated social engineering attack was used (Security 2022:para.2). The threat actors gained access to a limited number of accounts' data and the impacted customers were individually notified (Security 2022:para.8). While Twilio worked with carriers and hosting providers to stop the spam messages and shut down the malicious links, the threat actors continued to rotate through carriers and hosting providers to resume their attacks (Security 2022:para.4). The threat actors were well organised, complex and systematic in their malicious actions (Security 2022:para.5). It was found in the investigation that approximately 209 Twilio customers had their data accessed by threat actors for a limited period of time (Security 2022:para.29).

### 2. Uber

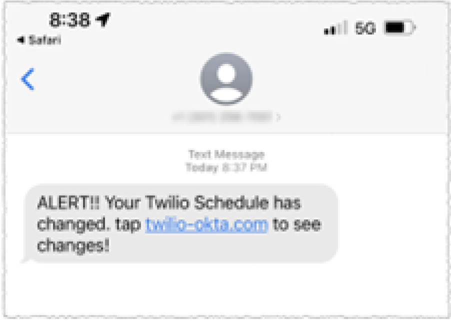
In 2022, Uber disclosed a data breach where a hacker bought stolen credentials belonging to an Uber employee from a dark web marketplace (Kost 2023:para.1). The first attempt to access Uber's network with these credentials did not work because the account was protected with MFA (Kost 2023:para.1). The hacker is believed to be an 18-year-old hacker associated with a cybercriminal group, Lapsus$ as they revealed the details of the breach in a conversation with Corben Leo, cybersecurity researcher (Kost 2023:para.6). There is no evidence of customer data theft announced despite the deep level of compromise the hacker achieved (Kost 2023:para.7). The hacker was enjoying the thrill of a successful data breach rather than having malicious intent to cause harm (Kost 2023:para.7). If the hacker was motivated by financial gain, Uber's bug bounty reports could have been sold on a dark web marketplace for an extremely high price (Kost 2023:para.8). Uber almost came to shutting down the whole system (Kost 2023:para.9). The hacker just stopped everything and walked away after taking control of Uber's systems (Kost 2023:para.9).

## B. Strategies Employed by Attackers in Each Case

### 1. Twilio

- The threat actor stole employee's credentials by designing a social engineering attack (Security 2022:para.2).

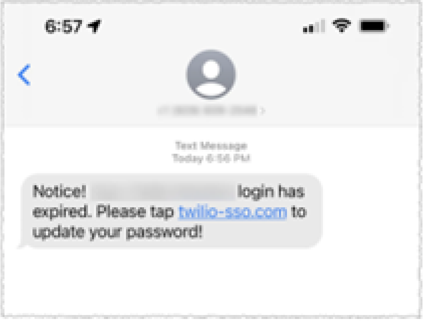
- The social engineering attack involved the employees receiving text messages impersonating to be from Twilio's IT department stating that the employee’s password had expired or that their schedule had changed (Security 2022:para.3).
- There was a link in the text message. When employees clicked the link that the threat actor created, the employee was prompted to login and their credentials were stolen (Security 2022:para.3).
- The link contained “Twilio” to attempt to trick users into clicking the link and placing their log in details in the malicious website that impersonated Twilio’s sign in page (Security 2022:para.2).
- The threat actors created fake Okta login pages for Twilio employee’s to be fooled and put in their login credentials (Security 2022:para.22).
- The threat actors had sophisticated abilities to match employees' names from sources with their phone numbers (Security 2022:para.3).
- The threat actor fooled some employees into providing their credentials (Security 2022:para. 2).
- The threat actor used the stolen log in details to gain access to some of Twilio’s internal systems and find specific customer data (Security 2022:para.2).
- Sample SMS-



[redacted]

Text Message
Today 8:37 PM

ALERT!! Your Twilio Schedule has changed. tap twilio-okta.com to see changes!



[redacted]

Text Message
Today 6:56 PM

Notice! [redacted] login has expired. Please tap twilio-sso.com to update your password!

## 2. Uber

- The hacker utilised social engineering techniques to fool the Uber employee into thinking they were a member of Uber’s security team (Kost 2023:para.10).
- To overcome MFA, the hacker communicated to the Uber employee via What’s App by pretending to be a member of Uber’s security and asked the employee to approve the MFA notifications being sent to their phone (Kost 2023:para.1).
- The hacker sent a large amount of MFA notifications to the employee’s phone to force them into yielding to this request (Kost 2023:para.1).
- The Uber employee approved a MFA request to end the large amount of notifications and the threat actor was granted access to Uber's network, leading to the data breach (Kost 2023:para.1).
- The threat actor compromised an Uber’s account and revealed the successful breach to Uber (Kost 2023:para.2).
- The hacker gained access to Uber’s VPN and found Microsoft Powershell scripts, including the login credentials of an admin user in Thycotic, Uber’s Privileged Access Management (PAM) solution (Kost 2023:para.4).

- Finding these login credentials drastically increased the seriousness of the breach by enabling full admin access to all Uber's sensitive services, including Amazon Web Services (AWS), GSuite, DUO, DA and Onelogin (Kost 2023:para.4).

# C. Defence Tactics and Responses

## 1. Twilio

- Twilio's security team cancelled access to the compromised employee accounts to diminish the attack and a leading forensics firm was involved to aid the investigation (Security 2022:para.6).
- Twilio worked with U.S carriers to shut down the threat actors (Security 2022:para.3).
- Twilio worked with the hosting providers delivering the malicious links to close those accounts down (Security 2022:para.3).
- Twilio issued security advisories on the specific tactics being utilised by threat actors since they first started to appear (Security 2022:para.7).
- The compromised Twilio employee user accounts' credentials were reset (Security 2022:para.34).
- All active sessions associated with the compromise of Okta-integrated apps were cancelled (Security 2022:para.34).
- Twilio obstructed all indicators of compromise associated with the attack (Security 2022:para.34).
- Twilio commenced takedown requests of the impersonating Twilio domains (Security 2022:para.34).
- To prevent a similar attack from happening, Twilio has implemented several additional security measures, including:
  - Implementing solid two factor authentications and distributing FIDO2 tokens to all employees (Security 2022:para.36).
  - Applying additional layers of control within Twilio's VPN (Security 2022:para.36).
  - Eliminating and restricting specific functionality with specific administrative tooling (Security 2022:para.36).
  - Expanding the refresh frequency of tokens for Okta-integrated applications (Security 2022:para.36).
  - Managing additional mandatory security training for all employees regarding social engineering attacks and the techniques used (Security 2022:para.36).

## 2. Uber

- Uber acted on the data breach quickly, confirmed the data breach on Twitter and notified law enforcement (Strom 2022:para. 1).
- Uber took numerous steps to lock down its code repository, modified credentials, and acknowledged other compromised accounts (Strom 2022:para.10).
- Uber restored all of their services to operational status (Strom 2022:para.1).

- Uber identified any employee accounts that were compromised or possibly compromised and either stopped their access to the Uber system or needed a password reset (Uber Team 2022:para.5).
- Uber added extra monitoring of our internal environment to keep an even closer eye of any additional suspicious activity (Uber Team 2022:para.5).
- Uber restricted numerous impacted or possibly affected internal tools (Uber Teams 2022:para.5).
- Uber locked down their codebase, preventing any new code modifications (Uber Teams 2022:para.5).
- Uber made their employees reauthenticate when restoring access to internal tools (Uber Teams 2022:para.5).
- Uber is strengthening their MFA policies (Uber Teams 2022:para.5).

# VI. Threat Mitigation and Response Framework

The Hard Hat focus for mitigation and response, focuses heavily on education and machine learning using sms filtering. To bolster our defences against smishing attacks, we are exploring innovative solutions within our Threat Mitigation and Response Framework. One promising approach involves leveraging the power of Artificial Intelligence (AI) to detect and combat common smishing patterns (Kaspersky, 2023). Through AI-driven applications, we aim to swiftly identify keywords frequently employed in deceptive messages and proactively mitigate potential threats. By continuously refining the AI algorithms, we strive to stay one step ahead of the attackers, making it harder for them to outsmart our defence measures. Additionally, we are actively promoting user education to raise awareness about smishing risks and encourage a vigilant online community. By combining cutting-edge technology and informed user practices, we aspire to create a safer digital environment for everyone (Njuguna et al., 2022).

## A. Proactive Measures for Smishing Prevention

In the ongoing pursuit of enhancing our cyber security posture for our clients, the Hard Hat team is committed to implementing proactive measures for the prevention of smishing attacks. By anticipating potential threats and staying ahead of emerging risks, we aim to fortify our defences and create a safer digital environment for both our team and our stakeholders.

## B. Incident Response and Handling Procedures

We propose a robust incident response framework that enables us to swiftly detect and mitigate smishing attacks. Our proactive approach empowers us to minimise the impact of these threats while effectively containing and neutralising them.

## C. Continuous Improvement and Adaptation

Recognising the dynamic nature of cyber threats, we maintain a continuous improvement cycle for our smishing prevention strategies. Regular assessments and refinements allow us to adapt to new attack techniques and stay prepared for evolving challenges.

# VII. Evaluating the Effectiveness of Defense Strategies

In the domain of Smishing Detection, evaluating the effectiveness of defence strategies is paramount in developing robust and adaptive security measures. This section delves into the key aspects of assessing the success of defence strategies, highlighting the critical role of Key Performance Indicators (KPIs), metrics, and industry benchmarking.

## A. Key Performance Indicators (KPIs) for Smishing Detection

KPIs serve as vital instruments for gauging the efficacy of Smishing detection defence strategies. These KPIs are essential in measuring critical elements such as detection accuracy, false positive rates, response time, and the volume of identified smishing attempts. For instance, detection accuracy measures the percentage of actual smishing attacks successfully identified, while false positive rates indicate the number of legitimate messages incorrectly flagged as smishing. KPIs provide a quantitative foundation for assessing the performance of defence tactics and facilitate the fine-tuning of strategies to optimise security.

## B. Metrics for Assessing the Success of Defence Tactics

Metrics offer a nuanced perspective for evaluating the success of defence tactics employed in Smishing Detection. These encompass various facets, including the detection rate, false negative rate, and the time taken to mitigate smishing threats. The detection rate quantifies the proportion of smishing attacks identified correctly, whereas the false-negative rate showcases the number of undetected attacks. Metrics provide a granular understanding of the effectiveness of specific tactics and contribute to the iterative refinement of defence strategies.

## C. Benchmarking Against Industry Standards

Benchmarking against industry standards is an essential practice to validate the effectiveness of Smishing detection defence strategies. Industry standards, such as those proposed by cybersecurity organisations like NIST or ISO, provide a normative framework against which defence tactics can be assessed. By aligning defence strategies with these standards, organisations can ensure that their approaches meet recognized best practices, enhancing their resilience against smishing attacks.

Incorporating these elements into Smishing detection defence strategies enables organisations to develop a comprehensive framework for evaluation and optimization. The utilisation of KPIs, metrics, and industry standards forms a crucial facet of ensuring the robustness and adaptability of defences in the ever-evolving landscape of Smishing threats.

# VIII. Ethical Considerations in Defending Against Smishing

In the pursuit of developing and deploying effective strategies and tactics to defend against smishing attacks, it is imperative to navigate the ethical dimensions inherent in the domain of cybersecurity. The ethical considerations in defending against smishing encompass several critical aspects, which are essential for creating a balanced and responsible defence framework.

## A. User Privacy and Consent

When implementing smishing detection mechanisms, safeguarding user privacy and obtaining their consent is paramount. The collection and analysis of SMS data should be conducted with utmost transparency, ensuring users are aware of and have consented to these practices. Ethical detection systems prioritise user data protection and transparency in their operations.

## B. Non-Discrimination

Defence strategies should avoid discriminatory practices that might unfairly target specific individuals or groups. Ethical considerations demand that smishing detection systems do not engage in profiling or biassed decision-making, ensuring equal protection for all users.

## C. Data Retention and Disposal

Ethical defence tactics should establish clear guidelines for data retention and disposal. Minimising data retention and securely disposing of data once its purpose is fulfilled respects user privacy and data protection regulations.

## D. Mitigation over Retribution

While detecting smishing attacks is essential, ethical considerations emphasise mitigation over retribution. Rather than aggressively pursuing attackers, ethical defenders prioritise the protection of potential victims and the prevention of harm.

## E. Disclosure and Transparency

Ethical responsibility dictates that organisations openly disclose any data breaches or vulnerabilities discovered during the defence process. Transparent communication is essential for building trust with users and stakeholders.

Ethical considerations in defending against smishing attacks serve as a foundation for responsible and trustworthy cybersecurity practices. By respecting user privacy, avoiding discrimination, managing data ethically, prioritising mitigation, and maintaining transparency, defenders can develop strategies and tactics that not only protect against threats but also uphold the highest ethical standards, fostering user trust and compliance with regulations.

# IX. Future Threat Scenarios and Trends

Smishing attacks, a form of cyber threat where attackers use text messages to deceive individuals into revealing sensitive information or taking malicious actions, are poised to undergo significant changes in the coming years. As technology advances and attackers become more sophisticated, the landscape of smishing attacks is likely to evolve in various ways.

In the future, we can expect smishing attacks to grow in sophistication and frequency. Criminals are becoming increasingly adept at utilising the technology and methods inherent to these attacks (Michan, 2023). A key trend on the horizon involves the integration of artificial intelligence (AI) to craft personalised and convincing messages. This AI-driven approach could result in messages tailored to individual recipients, heightening the deception and increasing the likelihood of success.

Furthermore, attackers are expected to exploit vulnerabilities within mobile operating systems and applications. As mobile devices become even more integrated into our daily lives, they offer an attractive avenue for attackers seeking to exploit security gaps. Additionally, attackers are likely to leverage social engineering techniques that manipulate human emotions and behaviours. By playing on recipients' fears, desires, or concerns, smishing attackers can craft messages that appear highly convincing and compelling (Michan, 2023).

# X. Conclusion

## A. Summary of Research Findings

### 1. Predicted Evolution of Smishing Attacks

The evolution of smishing attacks is expected to follow a trajectory from impersonating traditional financial institutions, like banks, to impersonating a wider range of trusted entities. These may include package delivery services, health authorities, government agencies, and more (Lacono, Hickman, and Muniz, 2022). The diversification of attack targets allows attackers to exploit various contexts, increasing their chances of success.

Furthermore, smishing attacks are projected to incorporate more advanced techniques for greater impact. Techniques like spoofing phone numbers to mimic legitimate sources, embedding malicious links or attachments, and even orchestrating multi-channel attacks that combine text messages with voice calls or emails are likely to become more prevalent (Michan, 2023).

## B. Recommendations for Smishing Detection and Defense

### 1. Emerging Techniques by Attackers

As technology evolves, attackers are poised to adopt emerging techniques to enhance their smishing campaigns. Deepfake voice synthesis, for instance, enables attackers to simulate voices of trusted individuals or institutions, making their messages even more convincing. Location-based targeting leverages geolocation data to craft messages that appear contextually relevant to recipients' whereabouts, enhancing the credibility of the attack.

In addition, attackers may employ QR code injection to deliver malware directly to users' devices. By disguising malicious content as innocuous QR codes, they exploit users' curiosity to spread malware. Another technique, SIM swapping, involves transferring a victim's phone number to a new SIM card controlled by the attacker, enabling them to intercept important messages and verification codes (Michan, 2023).

### 2. Anticipating Future Defense Challenges

Defending against evolving smishing attacks poses challenges for individuals and organisations alike. Raising awareness and education among mobile users, especially older individuals who may be less familiar with smishing, is crucial. Education campaigns can help users recognize suspicious messages and understand the importance of not clicking on unfamiliar links (Martens, 2023).

Developing and implementing effective security policies and solutions for mobile devices is another challenge. This includes deploying anti-malware software, spam filters, and robust authentication methods to prevent unauthorised access. Collaboration among stakeholders such as mobile network operators, regulators, and law enforcement agencies will also be pivotal in preventing and responding to smishing incidents, ensuring a comprehensive defence strategy (Michan, 2023).

## C. Implications for Future Cybersecurity Strategies

We will actively anticipate the evolution of smishing attacks and emerging techniques used by attackers. By staying ahead of the curve, we can better prepare ourselves to counter new threats and continue to improve our prevention measures.

Our commitment to proactive smishing prevention reflects our dedication to cyber security excellence. By implementing a comprehensive set of measures, staying adaptable, and remaining vigilant about emerging trends, we are proactively working towards a cyber-resilient future for our team, our stakeholders, and the broader digital community.

# XI. References

A. Chittora, A., Purohit, S., & Gupta, B. B. (2017). Phishing detection based on user behaviour profiling. International Journal of Computer Applications, 158(6), 16-21.

B. Alazab, M., & Broadhurst, R. (2014). A deep dive into online scam activities. IEEE Transactions on Cybernetics, 44(12), 2470-2483.

C. Rass, S., Kirda, E., & Kruegel, C. (2009). Vabam: server-based prevention of sms spam. In Proceedings of the 1st USENIX conference on Large-scale exploits and emergent threats (pp. 1-9).

D. Bryan, C., Gates, C., & Nicholas, C. (2016). Examining public-private partnerships as a mechanism to mitigate cyber threats. Computers & Security, 57, 16-28.

E. Cooke, L., & Shishika, S. (2017). Smishing Detection Framework Using Data Mining Techniques. International Journal of Computer Applications, 167(6), 13-17.

F. Mell, P., & Scarfone, K. (2007). Common Metrics for Information Security. NIST Special Publication, 800(94), 17-18.

G. Conti, M., Dragoni, N., & Lesyk, V. (2012). Efficient Software Development Methodologies for Mobile Applications. Proceedings of the 5th International Conference on Security of Information and Networks, 276-283.

H. Ruoti, S., Voas, J., & Kuhn, D. R. (2012). Threat Analysis: What Could Possibly Go Wrong? IEEE Security & Privacy, 10(3), 80-83.

I. National Institute of Standards and Technology (NIST). (2020). NIST Cybersecurity Framework. Retrieved from https://www.nist.gov/cyberframework.

J. International Organization for Standardization (ISO). (2021). ISO/IEC 27001:2013 Information Security Management. Retrieved from https://www.iso.org/standard/54534.html

K. Balebako, R., Cranor, L. F., & Reeder, R. W. (2010). The privacy and security behaviours of smartphone app developers. Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work, 201-210.

L. Rowe, N. C., & Panaousis, E. A. (2020). Security vs. Privacy: User Perceptions of the Internet of Things. IEEE Internet of Things Journal, 8(24), 17757-17766.