Friend or Foe: An Investigation into Recipient Identification of SMS-Based Phishing

Max Clasen, Fudong Li, David Williams

Friend or Foe: An investigation into recipient identification of SMS-based phishing

Max, Clasen; Fudong, Li; David Williams

School of Computing, University of Portsmouth, UK
Fudong.Li@port.ac.uk; David.Williams2@port.ac.uk

Abstract.
Short Message Service (SMS) messaging plays a key role in many people's lives, allowing communication between friends, family and businesses through the convenient use of a mobile phone. At the same time, criminals are able to utilise this technology to their own benefit, such as by sending phishing messages that convince their victims into sharing sensitive information or installing dangerous software on their devices. Indeed, Proofpoint's State of the Phish report found 81% of surveyed US organisations had faced smishing attacks – which is a type phishing attack via SMS message in 2020.
Although phishing is well studied, the amount of research in SMS-based phishing is somewhat limited. Therefore, this study addresses the lack of SMS-based phishing insight, investigating which techniques/tactics are used by malicious senders and honest recipients to disguise/identify SMS-based phishing. By using an online questionnaire, a total of 576 participants' options upon 20 text messages (10 genuine and 10 phishing) were gathered. The result shows 73.4% of the SMS messages were categorised correctly; also a number of factors such as shortened URLs, inconsistent metadata/content, urgency cue, and age play a positive role in identifying phishing attacks.

Keywords. short message service (SMS), phishing, text message, mobile phishing

1 Introduction

Smartphone ownership and usage continues to increase; 88% of people in the United Kingdom in 2019 own at least one smartphone, up from 52% in 2012 (Deliotte, 2019). Smartphones are not just used to communicate with individuals they are used to access data sensitive services such as mobile banking, managing healthcare data and appointments and conducting business. This creates both a complex environment for

users to manage and a data rich environment presenting malicious actors additional vectors of attack, including for socially engineering and phishing.

SMS messaging, social media and email are the top three distribution methods for mobile phishing (Wandera, 2018), with 17.3%, 16.4% and 15.4% of the share respectively. Wandrera's Mobile Threat Landscape Report found over half of surveyed organisations to have experienced at least one mobile phishing incident in 2019 (Wandrera, 2020); more alarmingly, Proofpoint's State of the Phish report found 81% of surveyed US organisations had faced smishing attacks – which is a type phishing attack via SMS message (Proofpoint, 2021).

Consequences of phishing attacks can be damaging to both individuals and businesses. According to Verizon's 2020 Data Breach Investigation report, 22% of data breaches in 2020 involved phishing (Verizon, 2020) and phishing was involved in 78% of cyber-espionage incidents 2019 (Verizon, 2019). Cyber incidents, such as these, lead to direct monetary loss, operational costs and brand damage (Wardman, 2016).

In order to detect phishing attacks, a wide range of methods have been investigated. For instance, Ho et al. (2019) investigated characteristics of phishing based upon a dataset of 113 million employee-sent emails from 92 organisations; their results demonstrate latest state of enterprise phishing attacks. Sahingoz et al. (2019) proposed machine learning based phishing detection from URLs with a 97.98% accuracy rate on a dataset containing 36,400 legitimate and 37,175 phishing links. In addition, a number of studies were proposed for investigating the issues for Smishing specifically, such as Balim and Gunal (2019), Mishra and Soni (2019), and Sonowal and Kuppusamy (2018). Nonetheless, those proposed were mainly focused upon the detection of smishing attacks by using machine learning techniques rather than from users' point of view. It is well known that users are the weakest link in the security chain. Indeed, 97% of people around the world cannot identify a sophisticated phishing email (Inspired eLearning, 2017). To this end, this paper presents a survey study which investigates the accuracy with which individuals can differentiate between phishing and legitimate SMS messages as well as the methods used to differentiate them.

The remainder of the paper is structured as: Section 2 reviews related academic literature; Section 3 provides the research methodology; Section 4 presents key results; and Section 5 draws conclusions and highlights future research directions.

2 Literature Review

This literature review explores the techniques and tactics used by attackers when constructing SMS phishing messages and external factors that influence the capability of recipients to detect phishing.

2.1 Phishing Techniques and Tactics

Techniques used to hide the malicious intent of the message and exploit the trust recipients have in known websites include spoofing URLs to appear similar to the genuine website (Patel & Lou, 2007; Kim et al., 2011). By using URL shortening services, attackers can obfuscate the destination of malicious links in **shortened URLs**. This helps attackers to visually deceive the recipient; it also allows them to bypass URL blacklisting software (Le Page et al., 2018; Joo et al., 2017). Other message elements used to deceive include the use of security components like HTTPS to trick victims into believing it is a legitimate website (Dong et al., 2008). According to the APWG (2019, p.10), this has become increasingly common, providing a possible indication to its effectiveness. It follows that the presence of HTTPS in SMS messages with a URL may have some influence on how the recipient determines the legitimacy of the message.

Harrison et al. (2016) explain how phishing messages rely on a sense of urgency to reduce the recipient's ability to make rational decisions by acting quickly. Two types of **urgency cues** should be considered: fear-based and reward-based. Fear-based cues use some form of threat, such as imminent account closure if action is not taken; while reward-based cues attempt to offer something of value, but within a limited time. In an experiment with 194 participants, Harrison et al. (2016) investigated whether fear-based cues were more successful than reward-based cues, but no solid evidence to prove this hypothesis was obtained. It may be that young people (the mean age of the experiment's participants was 20 years) are less influenced by the difference between fear-based and reward-based messages.

From the recipient's perspective, Nicho et al. (2018) believe a sense of urgency in an email presents an easy indicator of a phishing attack; nonetheless, victims still fall for them because vigilance of phishing is often not a priority. Jensen et al. (2017) agree as determining whether a message is a threat is often an "ancillary task". They advocate the use of mindful training techniques to promote recipients into thinking about the message request (including whether the message invokes a feeling of urgency).

Abroshan et al. (2018) suggest that part of the phishing process on the recipient-side involves two steps of decision making: whether to trust the sender and then whether to share information with them. A commonly accepted method for gaining trust is to exploit trust imparted by third-parties by masquerading as them (APWG, 2019; Abroshan et al., 2018; Whittaker et al., 2010). Attackers seek to acquire trust by making their messages mirror legitimate ones, e.g., by spoofing the sender ID and making the implication of being from a legitimate service in the body of the message (Jensen et al., 2017). Dong et al. (2008) state that **metadata/content inconsistencies**, i.e., differences between metadata (e.g., the sender ID) and content data (i.e., the text message), can lead to detection of phishing, so harmonising these two factors is to the attacker's benefit.

Harrison et al. (2016) mention leakage cues, such as grammar and spelling mistakes in the message impact trust, leading to increased attention by the recipient and thereby reduce phishing success. This is intuitive as one might expect genuine messages from services to have limited mistakes. Nonetheless, Jakobsson (2018) argues that it is easier to fake an SMS text message due to its simplicity, as it consists mostly of just a sender ID, plain text message and a timestamp.

2.2 External factors

While evidence suggests **age** may be a factor in how susceptible one is to phishing, there is disagreement over gender being a factor. Siadati et al. (2017) found that older people fall for phishing less frequently than younger people. This is consistent with the demographic study on phishing susceptibility conducted by Sheng et al. (2010); their study suggests that people between 18-25 years were more likely to fall victim than other age groups due to a lack of sufficient technical knowledge and experience. Sheng et al. (2010) also suggest that females are more prone to phishing attacks as they clicked on links more often than males, but Siadati et al. (2017) contradict this having found no significant difference in verification code phishing susceptibility by **gender**.

In their experiment to compare the more common rule-based training techniques against mindfulness techniques, Jensen et al. (2017) found that many participants reported a high level of **confidence and expertise** in identifying phishing messages. They also discovered that those trained to use mindfulness techniques, such as thinking about the request and whether the message felt rushed, were more successful than those who had received rule-based training, where elements like an unusual sender ID or the appearance of embedded links are used. It is generally considered that training users to understand and become more familiar with phishing attacks increases their ability to correctly identify phishing and non-phishing messages (Khonji et al., 2013; Jensen et al., 2017; Jain and Gupta, 2018).

According to Harrison et al. (2016, p. 270), participants who are **study aware** (i.e., participants are aware they are being tested on phishing detection) exhibit increased cognitive processing of messages. However, it is unlikely that users scrutinise every message in real life; as a result, it is more likely that users make a rapid decision based on straightforward cues found in the message. Jackson et al., (2007) found that aware participants are more likely to categorise both real and fake messages as phishing.

2.3 Summary

As demonstrated earlier, visual elements in phishing messages that may be influenced by attackers can be used to trick recipients; also a number of external factors should be considered on whether an individual can detect a phishing message through these elements. Nonetheless, the influence of these elements has not been wholly investigated

in terms of the SMS communication medium. Therefore further work is required to investigate those elements within the domain of phishing via SMS.

3 Methodology

To investigate which factors are used to identify SMS-based phishing, a questionnaire that can be used to test the participant's ability to distinguish genuine SMS messages from phishing messages was designed; also, several factors, including: suspicious requests, urgency cues, leakage cues and inauthentic URLs, were considered when messages were selected. For each chosen SMS message, a screenshot was presented, asking participants to judge whether the message was phishing, not phishing, or if they did not know, and (optionally) to justify their answer (Harrison et al., 2016). In addition to the chosen messages, the following were asked from participants: their age, gender, their experience in phishing, and their ability to identify phishing attempts.

With the aim of harvesting phishing emails, five managed phone numbers were exposed within publicly viewable spaces at 30 of the most visited websites as rated by Alexa.com and Quantcast (Balduzzi et al., 2016). Phishing messages were determined as such if they were unsolicited and attempted to convince the recipient to do something that would likely result in harming them. Other factors within the message could also be examined, such as analysing any provided URLs via sandboxing tools and seeking information on the sender phone number in known "bad number" lists. Logical factors in the message content were also used, such as attempting to request information from an iPhone user despite being received on an Android phone, or supposed banks requesting recipients to share information that they would not ask for through SMS. By using this method, a total of six SMS-based phishing messages were obtained; another four phishing messages that had been shared publicly on social media were collected. Also, ten genuine SMS-based messages were included; most of the genuine messages were received via those five managed phones and they were related to verification code. The total of 20 SMS messages is inline with the work of Siadati et al. (2017) which had a total of 18 messages. Examples of both phishing and genuine SMS messages are presented in Figures 1 and 2 below; also, the rest of the messages are available in the Appendix.

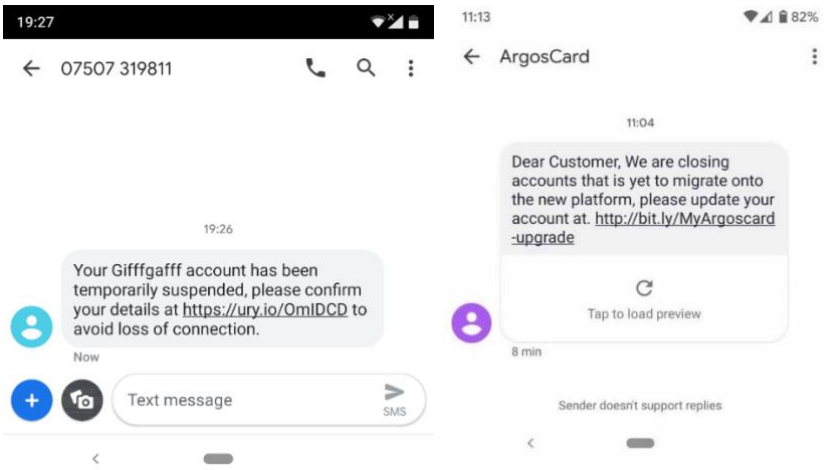


Fig. 1. Example of two phishing messages (Messages 1 and 2)

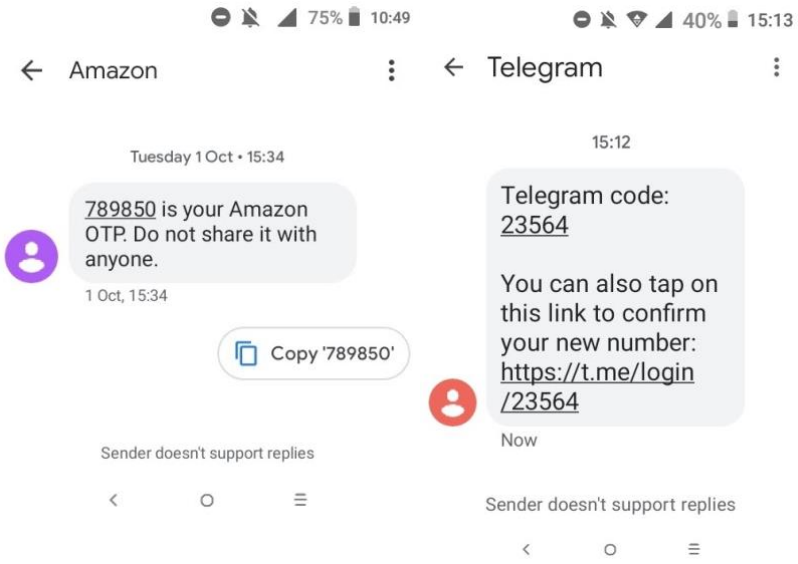


Fig. 2. Example of two genuine messages (Messages 3 and 20)

3.1 Data Collection

The primary research was collected using an online questionnaire via Google Forms due to its popularity, ease of use and accessibility. The questionnaire was distributed via a post on the "r/SampleSize" page of a popular social media website Reddit.com on Friday 24th January 2020. Also, it was shared to all students within the authors' department on Wednesday 29th January 2020. The questionnaire was closed to further

responses on Friday 31st January 2020 after receiving 576 responses (around 500 responses were gathered via Reddit). Amongst those 576 participants, 50% were males, 45% were females and the rest (5%) were classified as others. The majority (i.e. 86.4%) of the participants were younger people with 40.5%, 27.9% and 18% for age groups of 18-21, 22-25 and 26-30 respectively. Also, the proportions for age groups 31-40 and 40+ are 9.2% and 4.3% respectively. 3 participants of the total 576 did not select an age group.

4 Survey Findings

The 576 survey participants correctly **categorised** SMS messages 73.4% of the time. Messages were incorrectly categorised 14.8% of the time and stating they did not know if the message was phishing or not at a rate of 11.8%. The correct categorisation rate for the phishing messages (87.6%) was significantly higher than that for the genuine messages (59.2%). The genuine messages received an overall incorrect categorisation rate of 23.3% (where participants stated the message was phishing when in fact it was not), with the remaining 17.5% accounting for occasions when participants responded that they did not know whether or not it was phishing. Phishing messages were incorrectly categorised 6.3% of the time and were unknown 6.1% of the time (see Figure 3).

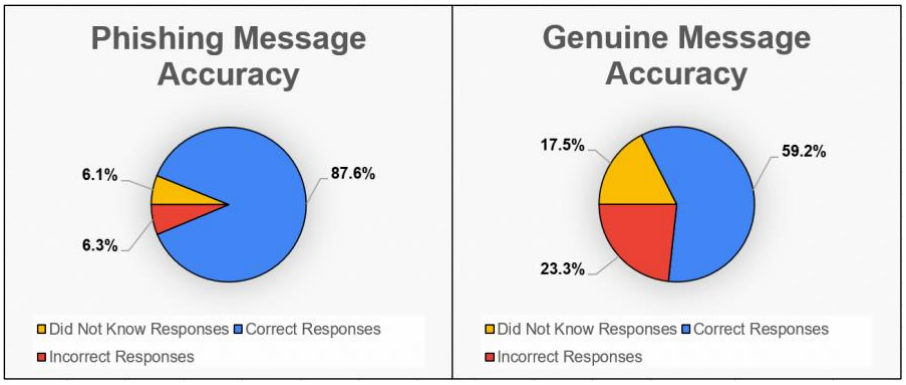


Fig. 3. The overall accuracy rate for the phishing messages (left) and the same for the genuine messages (right).

Messages that appeared later in the questionnaire gradually received fewer responses than those nearer to the beginning. For example, M1 received 351 responses out of the 576 participants (61.8% of participants) while M20 received 151 responses (26.2%). The reduced response rate for later messages may be attributed to response fatigue. Randomising the message order may have equalised the response rate. The total number of responses to the optional open-ended questions was 6531, averaging to 227.8 per message (39.5%).

Table 1 breaks down the accuracy of responses for each message. The accuracy rates are conditionally formatted to display a different background colour. Higher percentages have a green background, while lower percentages have a red background.

Table 1. Accuracies for individual messages

Individual Message Accuracy										
Phishing Messages	**M1**	**M2**	**M4**	**M8**	**M10**	**M12**	**M13**	**M14**	**M16**	**M17**
I don't know	12	44	37	29	5	74	84	15	41	12
No - This is NOT a phishing message	6	42	28	19	5	98	105	7	46	7
Yes - This is a phishing message	558	490	511	528	566	404	387	554	489	557
I don't know rate	2.1%	7.6%	6.4%	5.0%	0.9%	12.8%	14.6%	2.6%	7.1%	2.1%
Incorrect rate ("No" answers)	1.0%	7.3%	4.9%	3.3%	0.9%	17.0%	18.2%	1.2%	8.0%	1.2%
Correct rate ("Yes" answers)	96.9%	85.1%	88.7%	91.7%	98.3%	70.1%	67.2%	96.2%	84.9%	96.7%
Genuine Messages	**M3**	**M5**	**M6**	**M7**	**M9**	**M11**	**M15**	**M18**	**M19**	**M20**
I don't know	42	72	107	102	111	54	164	165	47	144
No - This is NOT a phishing message	514	415	269	376	210	463	191	185	478	310
Yes - This is a phishing message	20	89	200	98	255	59	221	226	51	122
I don't know rate	7.3%	12.5%	18.6%	17.7%	19.3%	9.4%	28.5%	28.6%	8.2%	25.0%
Correct rate ("No" answers)	89.2%	72.0%	46.7%	65.3%	36.5%	80.4%	33.2%	32.1%	83.0%	53.8%
Incorrect rate ("Yes" answers)	3.5%	15.5%	34.7%	17.0%	44.3%	10.2%	38.4%	39.2%	8.9%	21.2%

4.1 Phishing Messages

As shown in Table 1, the phishing messages M1, M10, M14 and M17 were accurately categorised more than any other message, each having a correct identification rate of over 95%. Messages M2, M4, M8 and M16 were also all correctly categorised relatively often, with success rates of between 91.7% and 84.9%. In contrast, M12 (70.1% accurately categorised) and M13 (67.2%) were correctly categorised the least of all the phishing messages.

Shortened URLs are apparently one reason for the high correct response rate for the phishing messages M1, M14 and M17. Shortened URLs within these three messages were cited as cause for suspicion by 13.4%, 16.4% and 32.1% of responses, respectively. Outside of these three messages, only M2 had a comparatively high rate for this point as well, at 15.0%, with the other messages being 4.0% and below.

The messages most frequently described as a suspicious request were M1, M10 and M17 (10.4%, 6.0% and 6.4% of responses respectively stating as such). By contrast, M14 had one of the lowest rates for this, with only 0.3% stating that it seemed unusual. The message itself did not contain a request, instead using a reward-based **urgency cue** notifying the recipient that they had won a sum of money. This was noticed by respondents; 13.9% indicated that the reward was “too good to be true” and 16.4% stated there was a clear attempt to entice the user with a reward.

Urgency cues in M10 were picked up frequently, with the sense of fear or other threat being mentioned in 9.6% of its responses, more than any other message. It was also the message for which leakage cues were mentioned most frequently, with 42% of responses pointing out the spelling and grammar mistakes in the text. The lack of

urgency cues in M12 and M13 may have been behind them being the most miscategorised. Only 0.4% of responses to M13 mentioned any sense of fear or threat from the message, lower than any other except M12, which received no responses that indicated a sense of urgency.

For M12, 28.6% noticed the **metadata/content inconsistency**, mentioning that it appeared that the URL was not for the service provider claimed in the message, higher than any other phishing message with the next closest comparison being M16 at 21.2%. Relating the appearance of the URL to the supposed service the message was often used to help categorise the messages, as witnessed in the responses to M4 (19.7% of responses made this observation), M10 (16.3%) and M12 (19.6%).

M13 was the only message that lacked a URL. Interestingly, this was mentioned as a sign of legitimacy in 3.6% of responses. However, 31.1% successfully recognised that the message was an attempt to obtain a verification code. One respondent noted:

> "This requires you to be a little savvy. When you get a verification code you should never share it with anyone. Sometimes the company tells you that when they send the code, and sometimes they don't. This is the most sophisticated attempt on this survey so far."

4.2 Genuine Messages

The genuine messages most frequently categorised correctly were M3 (correctly categorised in 89.2% of responses) , M11 (80.4%) and M19 (83%); also mostly categorised correctly were M5 (72.0%) and M7 (65.3%), respectively. The remainder of the genuine messages were accurately categorised only between 53.8% and 32.1% of cases. M9 (36.5%), M15 (33.2%) and M18 (32.1%) were correctly categorised as genuine the least frequently.

The most correctly categorised genuine messages M3, M11 and M19, were often simply considered legitimate second-factor authentication code messages, as mentioned in 28.2%, 32.6% and 34.2% of responses for each message, respectively – higher than any other message. Responses for M3 (13.2% of responses), M11 (18.7%) and M19 (24.8%) noted that the message lacked any request for information or a foreseeable way to acquire information.

In contrast, M9, M15 and M18, were correctly categorised the least. None of these messages used an authentication code, instead asking the recipient to respond to an application installation request via a web link. Respondents were sceptical of the request, especially in M9 where 14.4% mentioned that it was inherently suspicious. In both 9.5% of M15 responses and 9.0% of M18 responses, the use of an in-house URL shortening service rather than a public one was considered an indication of legitimacy.

4.3 Analysis of External Factors

Participants were asked to indicate their age; there was a significant imbalance in representation across age ranges, reflective of the demographics of Reddit. Participants aged 18-21 years comprised the overall majority at 40.5%, followed by 27.9% for 22-25, 18% for 26-30, 9.2% for 31-40 and 4.4% for 41+. Table 2 combines data across the age groups of 41-50, 51-60 and 61+ into a 41+ age group to compensate for the small number of participants above the age of 40. The percentage of responses that correctly categorised phishing messages marginally increased with age before dropping off at 41+ (see Table 2). For example, participants aged 18-21 years correctly categorised phishing messages 87.4% of the time, while participants aged 31-40 correctly categorised 90%. At 41-50 this dropped to 79.3%; all those 41+ collectively categorised phishing messages at a rate 79.6%. The categorisation of genuine messages revealed an opposite trend where younger participants correctly categorised them at a higher rate than for older participants.

Table 2. Accuracy by Age Groups

Accuracy by Age Group									
	Phishing			Genuine			Overall		
Age Group	Correct	Incorrect	Did not know	Correct	Incorrect	Did not know	Correct	Incorrect	Did not know
18-21	87.4%	7.7%	4.9%	62.0%	25.3%	12.7%	74.7%	16.5%	8.8%
22-25	87.9%	5.2%	6.9%	59.9%	20.7%	19.4%	73.9%	12.9%	13.2%
26-30	88.0%	6.4%	5.6%	58.3%	20.9%	20.9%	73.1%	13.6%	13.3%
31-40	90.0%	3.8%	6.2%	54.9%	24.5%	20.6%	72.5%	14.2%	13.4%
41+	79.6%	6.0%	14.4%	42.0%	26.8%	31.2%	60.8%	16.4%	22.8%

Each participant was also asked to rate their **confidence** in identifying phishing messages on a scale of 1 to 5, with 1 being lowest and 5 being highest, the vast majority of participants rated themselves highly. 31.9% rated themselves a 5, while 51% rated themselves a 4. This drops significantly at 14.6% for a 3, 2.1% for a 2 and just 0.3% for a 1. Those that rated themselves with a 5 accurately categorised phishing messages 90.8% of the time. This gradually decreased with those rating themselves a 4 (accurately categorising phishing messages 87.2% of the time) and those rating themselves a 3 (82.4%). However, the 2 (resp. 12) participants that rated themselves as a 1 (resp. 2) accurately identified phishing messages 90% (resp. 84.2%) of the time. The extent to which conclusions could be drawn from any apparent correlation between confidence rating provided and accurate categorisation of phishing and genuine SMS messages is limited by the significant under-representation of the lower confidence ratings.

No significant differences based upon **gender** were witnessed. The 288 participants that identified as male accurately categorised messages 74.8% of the time, in comparison to 71.9% for the 259 participants that identified as female and 74.7% for the 29 participants that identified as other.

5 Conclusion

Phishing has been used widely to attack the users, who are the weakest point of the cyber security system. Also, phishing attack can be distributed across many platforms (e.g. SMS message). The result of this research demonstrates that still a significant amount of users cannot differentiate between phishing and genuine SMS messages. Also, a number of factors (e.g. age and urgency cue) that can be used to identify phishing messages are investigated and the outcome is positive. In future, additional factors that may affect user's ability in identifying the legitimacy of a message should be investigated; this would provide a better understanding in training users to spot malicious messages.

References

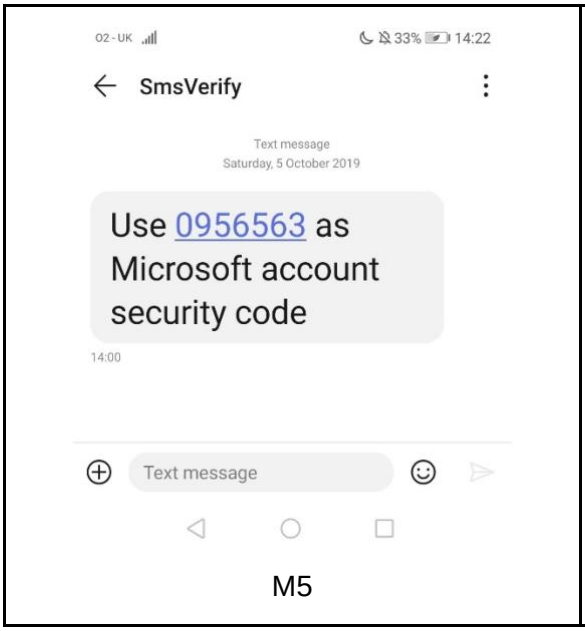
1. Abroshan, H., Devos, J., Poels, G., & Laermans, E. (2018). Phishing attacks root causes (Vol. 10694 LNCS). Springer International Publishing AG, part of Springer Nature 2018N. Cuppens et al. (Eds.): CRiSIS 2017, LNCS 10694, pp. 187–202. https://doi.org/10.1007/978-3-319-76687-4_13
2. Anti-Phishing Working Group [APWG] (2019). APWG Phishing Activity Trends Report 3rd Quarter 2019. Retrieved from https://docs.apwg.org/reports/apwg_trends_report_q3_2019.pdf
3. Balduzzi, M., Gupta, P., Gu, L., Gao, D., & Ahamad, M. (2016). MobiPot: Understanding Mobile Telephony Threats with Honeycards. Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security. 723-734. http://dx.doi.org/10.1145/2897845.2897890
4. Balim, C. and Gunal, E.S. "Automatic Detection of Smishing Attacks by Machine Learning Methods," 2019 1st International Informatics and Software Engineering Conference (UBMYK), Ankara, Turkey, 2019, pp. 1-3, doi: 10.1109/UBMYK48245.2019.8965429.
5. Deloitte (2019) "Smartphone accessories market to ring up revenues of £1.9bn in 2020 as UK reaches 'peak' handset ownership", available at: https://www2.deloitte.com/uk/en/pages/press-releases/articles/smartphone-accessories-market-to-ring-up-revenues-of-1-point-9-billion-pounds-in-2020.html
6. Dong, X., Clark, J. A., and Jacob, J. (2008). Modelling user-phishing interaction. 2008 Conference on Human System Interactions, Human System Interactions, 2008 Conference On, 627–632. https://doi.org/10.1109/HSI.2008.4581513
7. Harrison, B., Svetieva, E., & Vishwanath, A. (2016). Individual processing of phishing emails. Online Information Review, 40(2), 265-281.
8. Ho, G., Cidon, A., , Gavish, L., Schweighauser, M., Paxson, V., Savage, S., Voelker, G.M., and Wagner, D. (2019) "Detecting and Characterizing Lateral Phishing at Scale", 28th {USENIX} Security Symposium ({USENIX} Security 19) isbn: 978-1-939133-06-9, pages 1273-1290
9. Inspired eLearning (2017) "Phishing Statistics – The Rising Threat To Business", available via https://inspiredelearning.com/blog/phishing-statistics-facts/

10. Jain, A., K. & Gupta, B., B., (2018). Rule-Based Framework for Detection of Smishing Messages in Mobile Environment. Procedia Computer Science, 125, 617-623. https://doi.org/10.1016/j.procs.2017.12.079
11. Jakobsson, M. (2018). Two-factor in authentication – the rise in SMS phishing attacks. Computer Fraud & Security, 2018(6), 6–8. https://doi.org/10.1016/S1361-3723(18)30052-6
12. Jensen, M. L., Dinger, M., Wright, R. T., and Thatcher, J. B. (2017). Training to Mitigate Phishing Attacks Using Mindfulness Techniques. Journal of Management Information Systems, 34(2), 597–626. https://doi.org/10.1080/07421222.2017.1334499
13. Joo, J., Moon, S., Singh, S., & Park, J. (2017). S-Detector: an enhanced security model for detecting Smishing attack for mobile computing. Telecommunication Systems, 66(1), 29–38. https://doi.org/10.1007/s11235-016-0269-9
14. Khonji, M., Iraqi, Y. & Jones, A. (2013). Phishing Detection: A Literature Survey. IEEE Communications Surveys & Tutorials, Vol. 15 (No. 4), 2091-2121. https://doi.org/10.1109/SURV.2013.032213.00009
15. Kim, W., Jeong, O., Kim, C. and So, J. "The dark side of the Internet: Attacks, costs and responses", Information Systems, Volume 36, Issue 3, 2011, Pages 675-705, ISSN 0306-4379, https://doi.org/10.1016/j.is.2010.11.003.
16. Le Page, S., Jourdan, G.V., Bochmann, G. V., Flood, J., and Onut, I.V. (2018). Using URL shorteners to compare phishing and malware attacks. ECrime Researchers Summit 2018–May, 1–13. https://doi.org/10.1109/ECRIME.2018.8376215
17. Mishra, S. and Soni, D. (2019) "A Content-Based Approach for Detecting Smishing in Mobile Environment", Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity University Rajasthan, Jaipur - India, February 26-28, 2019
18. Nicho, M., Fakhry, H., and Egbue, U. (2018). When Spear Phishers Craft Contextually Convincing Emails. Proceedings of the IADIS International Conference on WWW/Internet, 313-320.
19. Patel, D. and Luo, X. (2007). Take a Close Look at Phishing. In Proceedings of the 4th Annual Conference on Information Security Curriculum Development, Kennesaw, GA, USA
20. Proofpoint (2021) "2021 State of the Phish - An In-Depth Look at User Awareness, Vulnerability and Resilience", available at: https://www.proofpoint.com/sites/default/files/threat-reports/gtd-pfpt-uk-a4-r-state-of-the-phish-2021.pdf
21. Sahingoz, O.K., Buber, E., Demir, O. and Diri, B. "Machine learning based phishing detection from URLs," Expert Systems with Applications, Volume 117, 2019, Pages 345-357, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2018.09.029.
22. Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L. and Downs, J. (2010) "Who falls for phish? a demographic analysis of phishing susceptibility and effectiveness of interventions", In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10). Association for Computing Machinery, New York, NY, USA, 373–382. DOI:https://doi.org/10.1145/1753326.1753383
23. Siadati, H., Nguyen, T., Gupta, P., Jakobsson, M., & Memon, N. (2017). Mind your SMSes: Mitigating social engineering in second factor authentication. Computers & Security, 65, 14–28. https://doi.org/10.1016/j.cose.2016.09.009

24. Sonowal, G. and Kuppusamy, K.S. (2018) SmiDCA: An Anti-Smishing Model with Machine Learning Approach, The Computer Journal, Volume 61, Issue 8, August 2018, Pages 1143–1157, https://doi.org/10.1093/comjnl/bxy039
25. Verizon (2019) "2019 Data Breach Investigations Report", available via https://enterprise.verizon.com/resources/executivebriefs/2019-dbir-executive-brief.pdf
26. Verizon (2020) "2020 Data Breach Investigations Report", available via https://enterprise.verizon.com/resources/reports/dbir/
27. Wandera (2018) "Mobile Phishing Report", available at http://go.wandera.com/rs/988-EGM-040/images/mobile-phishing-report.pdf
28. Wardman, Brad, "Assessing the Gap: Measure the Impact of Phishing on an Organization" (2016). Annual ADFSL Conference on Digital Forensics, Security and Law. 2. https://commons.erau.edu/adfsl/2016/thursday/2
29. Whittaker, C., Ryner, B. and Nazif, M., 2010. Large-scale automatic classification of phishing pages.

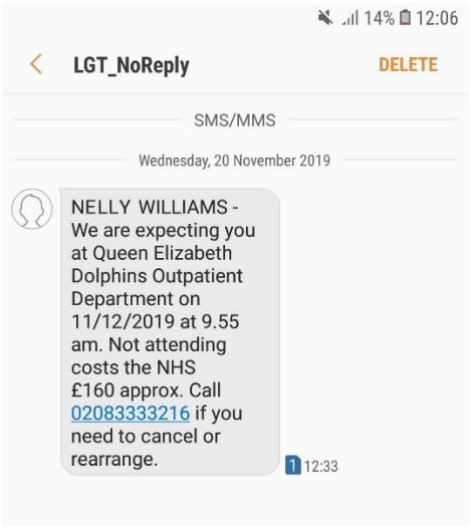
Appendix:

Genuine messages

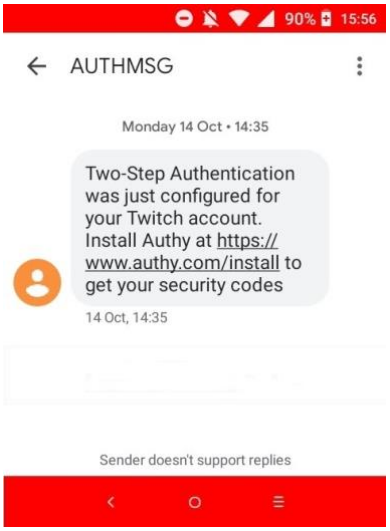


M5

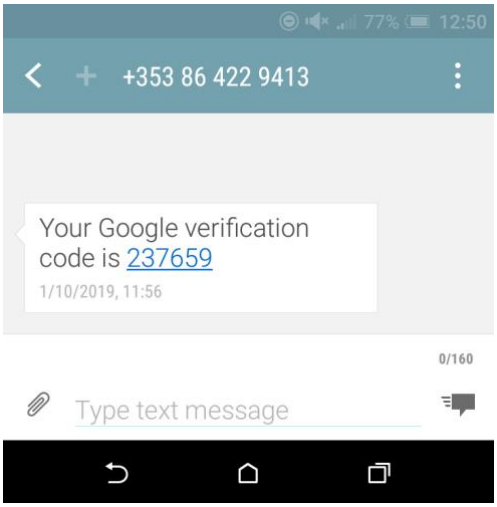
M6


14% 12:06
LGT_NoReply
DELETE
SMS/MMS
Wednesday, 20 November 2019
NELLY WILLIAMS - We are expecting you at Queen Elizabeth Dolphins Outpatient Department on 11/12/2019 at 9.55 am. Not attending costs the NHS £160 approx. Call 02083333216 if you need to cancel or rearrange.
12:33

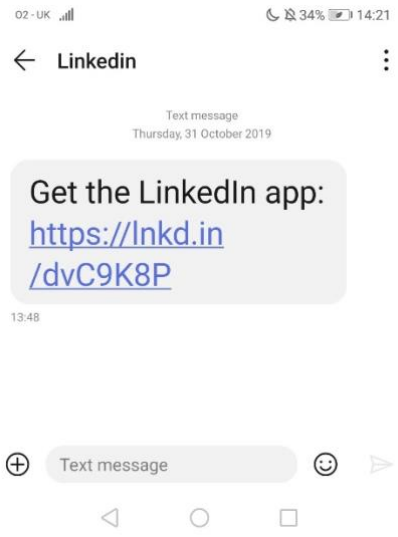
M7


90% 15:56
AUTHMSG
Monday 14 Oct • 14:35
Two-Step Authentication was just configured for your Twitch account. Install Authy at https://www.authy.com/install to get your security codes
14 Oct, 14:35
Sender doesn't support replies

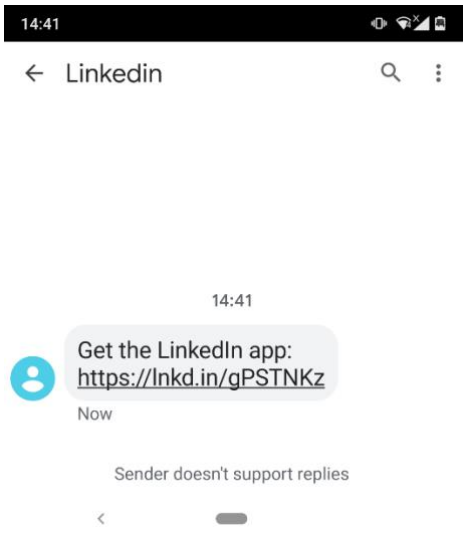
M9


77% 12:50
+353 86 422 9413
Your Google verification code is 237659
1/10/2019, 11:56
0/160
Type text message

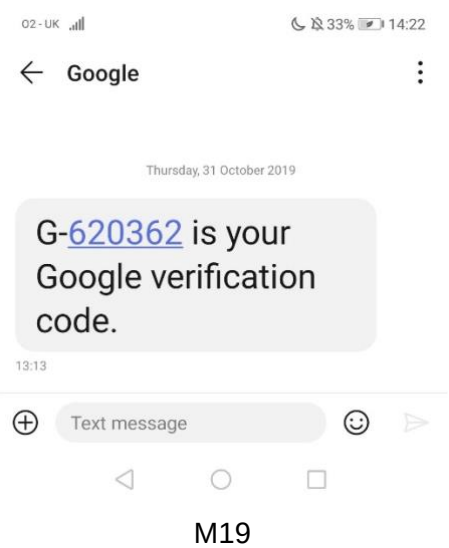
M11


02 - UK
34% 14:21
Linkedin
Text message
Thursday, 31 October 2019
Get the LinkedIn app: https://lnkd.in/dvC9K8P
13:48
Text message

M15

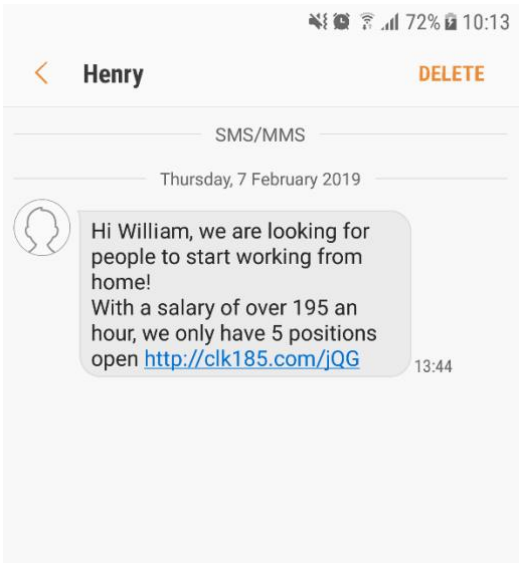


M18

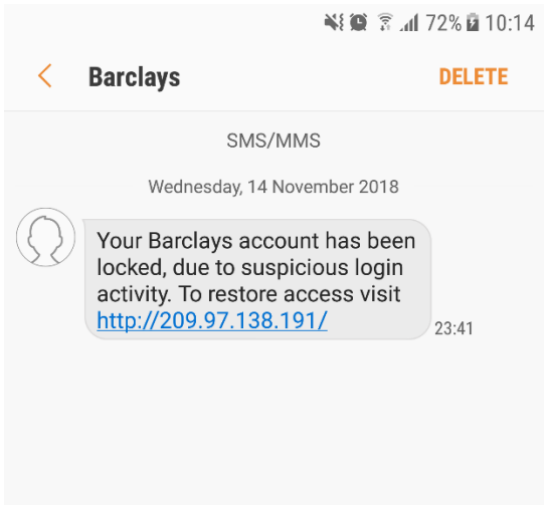


M19

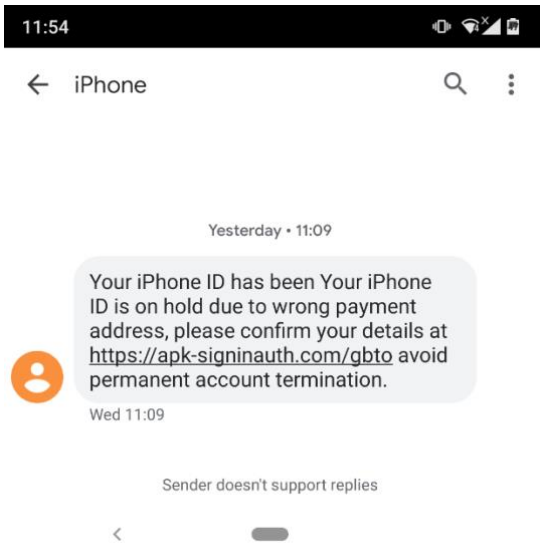
Phishing messages



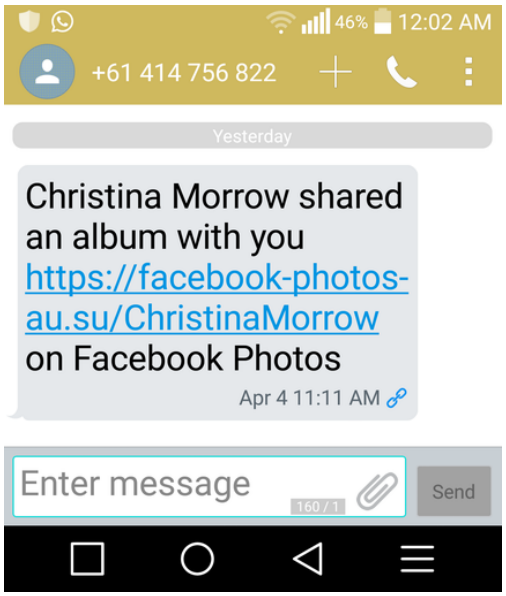
M4



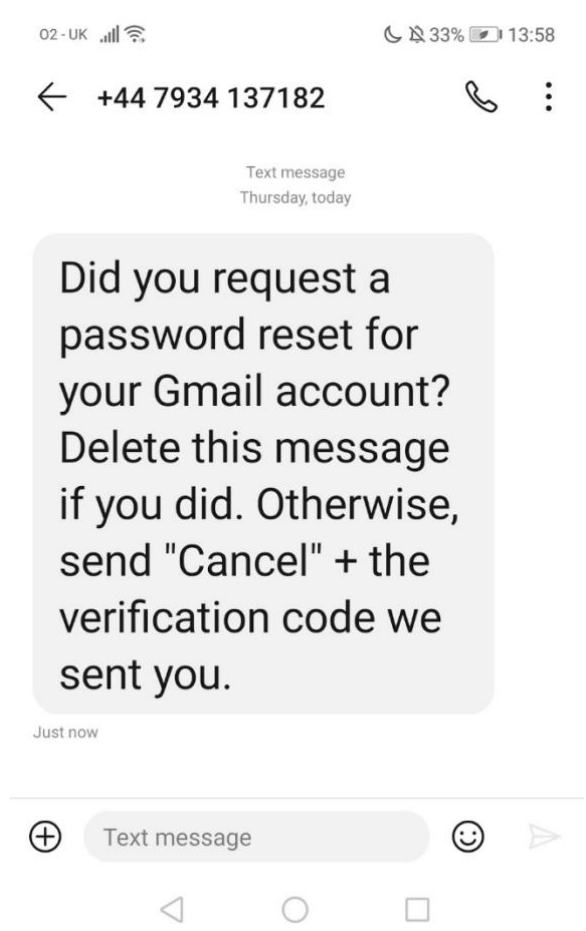
M8


11:54
iPhone
Yesterday • 11:09
Your iPhone ID has been Your iPhone ID is on hold due to wrong payment address, please confirm your details at https://apk-signinauth.com/gbto avoid permanent account termination.
Wed 11:09
Sender doesn't support replies

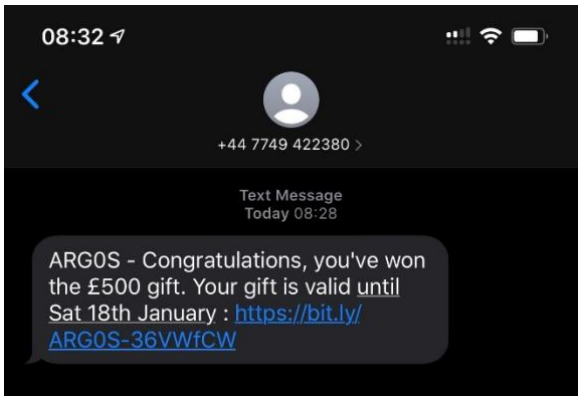
M10


46% 12:02 AM
+61 414 756 822
Yesterday
Christina Morrow shared an album with you https://facebook-photos-au.su/ChristinaMorrow on Facebook Photos
Apr 4 11:11 AM
Enter message
160 / 1
Send

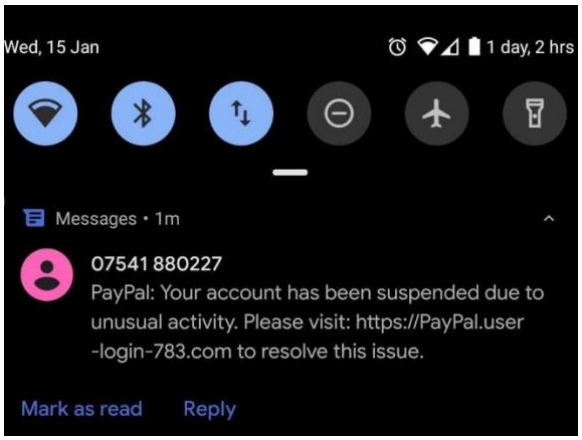
M12


02 - UK
33% 13:58
+44 7934 137182
Text message
Thursday, today
Did you request a password reset for your Gmail account? Delete this message if you did. Otherwise, send "Cancel" + the verification code we sent you.
Just now
Text message

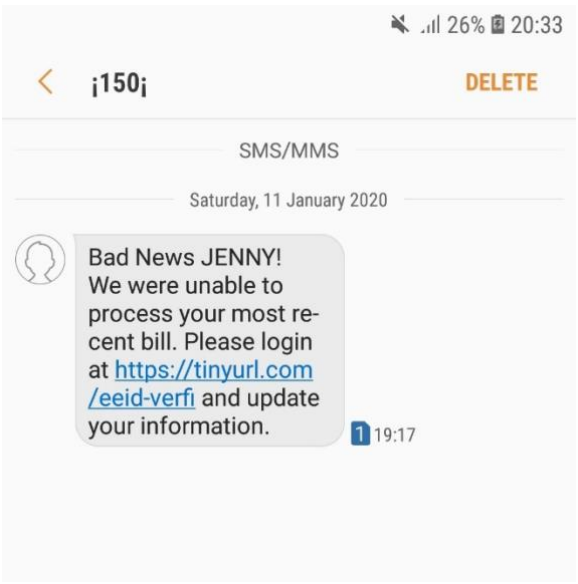
M13


08:32
+44 7749 422380
Text Message
Today 08:28
ARGOS - Congratulations, you've won the £500 gift. Your gift is valid until Sat 18th January : https://bit.ly/ARGOS-36VWfCW

M14


Wed, 15 Jan
1 day, 2 hrs
Messages • 1m
07541 880227
PayPal: Your account has been suspended due to unusual activity. Please visit: https://PayPal.user-login-783.com to resolve this issue.
Mark as read
Reply

M16


26% 20:33
¡150¡
DELETE
SMS/MMS
Saturday, 11 January 2020
Bad News JENNY!
We were unable to process your most recent bill. Please login at https://tinyurl.com/eeid-verfi and update your information.
19:17

M17