Mazephishing: The COVID-19 pandemic as credible social context for social engineering attacks

Kristjan Kikerpill\(^1,1\) and Andra Siibak\(^2\)

\(^1\)University of Tartu
\(^2\)Affiliation not available

May 18, 2021

Abstract

The first months of the COVID-19 pandemic witnessed a surge of social engineering attacks. Although the pandemic is certainly not the first occurrence of socially disruptive circumstances that drive cybercriminals to action, relevant academic scholarship has remained scarce. To fill this gap in literature, and propose the analytical framework of mazephishing that places particular emphasis on the importance of credible social context in the functioning of the online scam ecosystem, we carried out a content analysis of (\(N=563\)) international news stories reporting on social engineering attacks. Our results indicate that criminals make heavy use of social context and impersonation to make scams seem more credible. Major themes used in the scam messages include health information, personal protective equipment, cures, financial relief and donations. Additionally, scammers diversify their use of mediums depending on the type of scam being perpetrated. Our analysis also shows a significant presence of principles of persuasion in the circulated scam attempts.

(Preprint for https://doi.org/10.3176/tr.2021.4.01)

Mazephishing: The COVID-19 pandemic as credible social context for social engineering attacks

Kristjan Kikerpill, University of Tartu
Andra Siibak, University of Tartu

Abstract

The first months of the COVID-19 pandemic witnessed a surge of social engineering attacks. Although the pandemic is certainly not the first occurrence of socially disruptive circumstances that drive cybercrime, relevant academic scholarship has remained scarce. To fill this gap in literature, and propose the analytical framework of mazephishing that places particular emphasis on the importance of credible social context in the online scam ecosystem, we carried out a content analysis of (\(N=563\)) international news stories reporting on social engineering attacks. Our results indicate that criminals make heavy use of social context and impersonation to make scams seem more credible. Major themes used in the scam messages include health information, personal protective equipment, cures, financial relief and donations. Additionally, scammers diversify their use of mediums depending on the type of scam being perpetrated. Our analysis also shows a significant presence of principles of persuasion in the circulated scam attempts.

Keywords: social engineering, phishing, message context, salient events, social context, cybercrime, content analysis

Introduction
In *The Science of Human Hacking*, Hadnagy (2018: 7) defines social engineering as “any act that influences a person to take an action that may or may not be in his or her best interests”. According to Hadnagy, the definition is broad and general because the use of social engineering is not always negative. For instance, children persuade their parents to play games, parents convince their children to visit the dentist and spouses are coaxed into attending social events. However, the use of influencing techniques and the application of psychological principles also manifests on the dark side of our ubiquitously connected society – in the form of social engineering attacks and, in particular, phishing. Phishing attacks are cyber-attacks “that communicate socially engineered messages to humans via electronic communication channels in order to persuade them to perform certain actions for the attacker’s benefit” (Khonji et al. 2013: 2092). Provided that the basic tenet of phishing is to deliver messages that elicit action, various mediums such as e-mails, voice calls (vishing) and text messages (smishing) are employed in carrying out the attacks (Chiew et al. 2018). Alongside the mediums used, phishing attacks are also categorized based on how the attacks target the potential victims, e.g. more elaborate attacks are dubbed ‘spear-phishing’ or even as ‘whaling’ when perpetrated against high-level targets such as CEOs (Hong 2012).

However, between phishing messages replete with grammatical errors (Chiluwa 2019) and meticulously tailored spear-phishing attacks targeting specific individuals exists a swathe of territory populated by ‘context aware phishing’ (Jakobsson and Myers 2007: 176), i.e. attacks “mounted using messages that somehow – from their context – are expected or even welcomed by the victim”. The importance of context in phishing attacks was predicted to increase due to improvements in countermeasures more than a decade ago (Jagatic et al. 2007). Nevertheless, literature pertaining to the analysis of contexts (Greene et al. 2018, Steinmetz et al. 2021) and salient current events that can impact susceptibility to phishing (Verma et al. 2018, Williams and Polage 2018, Kikerpill and Siibak 2021) is currently scarce. In part, this could be because the focus on human-centric solutions to phishing attacks have started to significantly increase only relatively recently (Ferreira and Vieira-Marques 2018). To fill this gap - and propose a specific term for social engineering attacks with a heavy reliance and emphasis on context - our article presents a study of mazephishing. In the tradition of using ‘fishing’ references when naming social engineering attack types, mazephishing is inspired by the age-old fishing technique of ‘almadraba’, a term of Arabic origin meaning ‘a place to smite’ (see Richardson 2007: 56), where fishermen set up complex underwater mazes of nets to catch tunas during their seasonal migration journeys through the Strait of Gibraltar. Thus, a successful catch depends on (1) proper timing, i.e. understanding the reason why fish are on the move in large numbers at certain times, (2) place, i.e. interrupting the tunas’ movement at a location and in a manner suitable for the fishermen, and (3) trap-setting technique.

In our study, we focus on the social context created by the COVID-19 pandemic because no other interpretive backdrop in recent history compares to the disruption in social circumstances created by this disease. The virus’ spreading forced an increase in people’s reliance on online resources and digital technologies (De et al. 2020, Vargo et al. 2021). From the perspective of cybercriminals crafting social-engineering attacks, a larger number of people using means of online communication more frequently constitutes a larger pool of readily available potential victims. In fact, researchers noted a substantial spike of 667% in COVID-19 phishing attacks in the first months of the pandemic (Shein 2020). Hence, the COVID-19 pandemic is operationalized as the credible social context, i.e. the ‘timing’ aspect of mazephishing, in our study. Context has both an interpretive and a constitutive dimension (Rigotti and Rocci 2006), which means that it helps us interpret received messages but also influences how messages are crafted in specific contexts.

Having fixed the mobilizing social context on the COVID-19 pandemic, our focus in this article is on the ‘place’ and trap-setting technique aspects of mazephishing, i.e. how cybercriminals attempted to spring their social-engineering traps. For this, we carried out a content analysis of international news media articles (N=563) from January – April 2020 that reported on and warned about relevant online scams. More specifically, within the overarching salient circumstances created by the COVID-19 pandemic, we analyse (1) what kind of communicative strategies and topics cybercriminals covered, (2) who scammers impersonated for the purposes of perceived source credibility, (3) what types of communication mediums were employed and (4) to what extent can the six principles of persuasion suggested by Cialdini (2009) be used to explain
the message content of the sample social engineering attacks.

**Theoretical background**

**Salient current events and phishing**

Health crises add a huge burden on the media to keep the public constantly informed (Ogbodo et al. 2020). In fact, previous scholarship (Liu 2020) indicates that the nature of media framing of health information not only help to form people’s understanding of the health crisis, but also shape people’s responses, i.e. influence public behaviour. In short, the media has the power to accentuate or mitigate the crisis depending on the frames adopted in their coverage. Although rumours and questionable information have oftentimes been associated with pandemics and crises (Eysenbach 2011), the dramatic increase in the dissemination of bogus information during the COVID-19 initiated an infodemic that enabled to create a “fertile information ecosystem for cybercriminals to exploit” (Naidoo 2020: 317).

Even though the COVID-19 pandemic is unique in its reach and social impact, it is certainly not the first salient current event to be featured as credible social context in fraud campaigns. Examples from recent history include the aftermaths of forest fires in Australia and Portugal, a hurricane in Puerto Rico and an earthquake in Japan (Grad 2020). Health-related social circumstances have been the credible social context in cyberattacks during the Ebola outbreak in 2014, the Zika virus in 2016 and influenza in 2019 (RiskIQ 2020). However, academic scholarship on the connection between salient current events and social engineering attacks has been scarce (Holt and Graves 2007, Greene et al. 2018, Steinmetz et al. 2021), leaving technology news stories, - blog posts or - reports as the relevant available sources. To some extent, the COVID-19 pandemic has been the exception. In addition to general COVID-19 themed cybersecurity overviews (see Pranggono and Arabo 2021), researchers have drawn connections between earlier disease outbreaks, the COVID-19 pandemic and changes in the cyberthreat landscape (Mouton and de Coning 2020), constructed event and cyberattack timelines (Lallie et al. 2020), proposed approaches on how the COVID-19 pandemic influences cybercrime (Naidoo 2020) and analysed the general communicative approaches employed in pandemic-themed social engineering attacks (Kikerpill and Siibak 2021).

Although social engineering and phishing have received considerable attention in research literature (see Montanez et al. 2020, Nguyen et al. 2020: Appendix A), our current understanding of the specificities of psychological mechanisms or demographics at work in online fraud victimisation remain limited (Button and Cross 2017, Norris et al. 2019). For instance, the relationship between an individual’s personality and phishing susceptibility has been considered weak (Sommestad and Karlzén 2019) or inconclusive (Montanez et al. 2020). Additionally, no anti-phishing training tools that actually use adjustments based on people’s personality traits have been identified (Jampen et al. 2020). Furthermore, no one demographic is necessarily more or less vulnerable to online fraudulent activity (Button and Cross 2017, Norris et al. 2019).

Steps to limit the impact of fraud ought to clearly recognise the universal nature of compliance (Norris et al. 2019: 242), including an acknowledgement of the ease with which people can come into contact with cybercriminals (see Kikerpill 2021). Since the basic tenet of social engineering attacks is to elicit compliance and action (Khonji et al. 2013), the observable tool in such compliance-gaining efforts is the transmitted message. Thus, the background of cybercriminals, which is rarely known even to the law enforcement community (Button et al. 2009: 13), or the specific demographic of the victims (see Button and Cross 2017) notwithstanding, it is paramount to further study how influencing and compliance-gaining efforts appear in the messages that connect perpetrators and potential victims. Given that context impacts both the creation and interpretation of such messages (Rigotti and Rocci 2006), its inclusion in any such analysis is crucial.

**‘Timing’, ‘place’ and principles of persuasion as ‘technique’**

The proposed construct of *maze phishing* emphasises the combination and interplay of two important aspects, i.e. timing and place (context and medium) and technique (action-eliciting message). Scam messages for and in which context is mechanically manufactured, e.g. ail-and-wail stories of fictional widows wishing to depart with large sums of money (Kikerpill and Siibak 2019: 57-58), are evidently opportunistic. In that sense,
transmitting such messages is akin to setting up elaborate mazes of nets to catch fish that may or may not be swimming towards these traps. Nevertheless, the success of such less elaborate social engineering attacks is not entirely negated even if the context is manufactured (Chiluwa 2019: 904). In part, this may be due to the sheer volume of messages that perpetrators are able to transmit, i.e. the year 2020 was predicted to witness the daily exchange of 306 billion emails (Radicati Group 2020) and approximately one in every 2000 emails is a phishing attack (Palmer 2020). To continue the simile, if there is no virtual limit on how many traps can be set, some fish at some point in time will end up in those traps even when their placement was entirely arbitrary.

The situation changes when perpetrators craft messages that reference existing companies or events (Chiluwa 2019: 898), i.e. credible social context, or ‘timing’, is introduced. This credible social context then also starts influencing individuals’ decision-making about the message (Greene et al. 2018, Carter 2021). Since recipients of scam messages tend to pay more attention to the content of the messages rather than the technical aspects of mediums such as email (Alsharnouby et al. 2015), criminals attempt to increase the chances of success by also altering the contextual placement of traps. Thus, the emergence of salient social circumstances like the COVID-19 pandemic results in some scammers adapting their previously used messages to fit the new context (EC News Desk 2020) and choose ‘places’ suitable for the traps such as emails, voice calls, text messages, fake websites and social media. However, the technique that is used in setting the traps is as important as the ‘timing’ (context) and ‘place’ (medium).

Since eliciting compliance is quintessential to any type of phishing (Norris et al. 2019), the trap ‘technique’ used in social engineering attacks is therefore bound to manifest in the content of the messages that perpetrators deliver. Without establishing lines of communication, no mediated convergence between criminals and victims could occur. Once this convergence occurs, however, the perpetrators’ goal is to influence the recipient enough to elicit specific actions (see Khonji et al. 2013). Thus, established persuasion scholarship can provide explanations of such trapping ‘technique’. Cialdini’s (2009) six basic principles of influence - authority, reciprocity, social proof, commitment/consistency, liking/similarity, and scarcity – all of which are “used ubiquitously in human interactions to influence and to persuade people to do, act, and think the way one wants” (Ferreira et al. 2015: 37) is an influencing technique framework often applied in phishing studies (see Zielinska et al. 2016, Lawson et al. 2020: 8).

According to Cialdini (2009), **authority** is the principle that describes people’s tendency to comply with the request of authoritative figures. Empirical research (Algarni et al. 2014) suggests that messages coming from a source that has expertise and is respected or authoritative, e.g. government agencies or reputable companies, have a higher likelihood of gaining end-user compliance. The **scarcity** principle mainly exists in two forms: “the limited number tactic” and “the deadline” tactic (Cialdini 2009). In fact, as posed by Cialdini (2009: 257) “newly experienced scarcity is the more powerful kind” as we tend to want an item more when there is a competition for it, or its availability decreases. The **liking and similarity** principles suggest that individuals are more easily persuaded by someone they know and like, which is common in how people interact socially (Ferreira et al. 2015). However, the usage of these principles does not always guarantee success. For example, findings by Zielinska and colleagues (2016) suggest that the success rate when using the principles of ‘liking/similarity’ and ‘authority’ is quite unpredictable – they could either increase phishing rate if successfully applied but could also decrease compliance if not used correctly. Recent research has also detected that phishing emails where the ‘scarcity’ principle has been employed are most often met with strong scepticism, i.e. individuals are least susceptible to such emails (Lawson et al. 2020).

Still, there are several additional principles of influence that can potentially impact the content of scam messages. The principle of **social proof**, for instance, is an important aspect of online commerce (Talib and Saat 2017), which makes use of community recommendations and product reviews in influencing purchasing decisions. Provided that such reviews can be intentionally manipulated (Zhuang et al. 2018), fake reviews can be considered as an easy-to-implement option when promoting the sale of non-existent products. The principle of **reciprocity**, however, is often employed in the solicitation of donations (Cialdini 2009), and donations to bogus charities are a common scam tactic used in perpetrated cybercrime, frequently in use
during natural or manmade disasters (Stabek et al. 2010). Reliance on reciprocity is also used in advance of fee scams where perpetrators offer a large reward in exchange for a minimal initial contribution from the would-be victim (Holt and Graves 2007, Kikerpill and Siibak 2019). The final principle of influence introduced by Cialdini (2009) is **consistency**, which pertains to humans’ inclination to act in accordance with their previously made decisions and commitments. For instance, scammers can leverage this principle, by sending users of specific services bogus reminders to update login information (Wright et al. 2014) or fraudulently impersonating utility companies and suggesting that the would-be victim’s utility bill is past due (KNEB 2020).

Utilizing the previously described principles of persuasion as the frame of reference in studying scam messages provides a revealing insight into the interplay between the ‘timing’, ‘place’ and ‘technique’ aspects of mazephishing. In other words, it allows us to analyse and explain how criminals employed the credible social context of the COVID-19 pandemic (‘timing’) and various convergence mediums (‘place’) to spring their social engineering traps (‘technique’).

**Data and methods**

Since the mediated convergence of perpetrators and potential victims in social engineering attacks occurs in the form of fraudulent messages, qualitative assertions regarding the interplay of messages and credible social context require ample data from which the specific messages-in-context can be extracted. Thus, we collected news story data that covered a four-month period from January-April 2020 to capture the start and the evolution of the COVID-19 pandemic as credible social context. We opted for news stories as our source of data due to our data requirements, i.e. descriptions of scams reported as such scams occurred, and to minimize any content-related bias that may entail from using blog posts or reports by specific companies.

We obtained the initial sample (N=1924) for our content analysis by carrying out a parameter-restricted Google keyword search. The search results’ language was set to English, the time period included the first and last dates of each month in the period studied, and the search was carried out using the ‘allintitle’ search operator to restrict obtained results solely to matches appearing in the results’ titles. In the keyword search, we used four pairs of two-word phrases, i.e. “covid scam”, “covid phishing”, “coronavirus scam” and “coronavirus phishing”. Using both “covid” and “coronavirus” in the search was necessary due to the fact that the disease was named COVID-19 only in the middle of the period for which data were collected (Nelson 2020). The ‘linkclump’ Google Chrome extension was used to gather the results’ links from the list of search results. The obtained results were divided based on the respective months and search phrases.

Following the initial preparation of the data, numerous restricting parameters were implemented to exclude unsuitable data from the original sample. As our aim was to gather news stories, we excluded blog and social media posts, scam warning notices that private companies, e.g. security vendors, and universities posted on their websites, alerts posted by governmental entities, civil service and law enforcement agencies (e.g. the FBI or local police forces) as well as links to images and news videos, and news stories that were not originally written in English. Based on the aforementioned exclusions, the sample was reduced to N=831 stories.

Since the process required manual verification, we also separated the sample news stories into categories of “traditional media”, e.g. international news outlets and local news that cover a variety of topics, and “specialist media” outlets focusing solely on technology news, e.g. BleepingComputer. Furthermore, we used the initial reading to develop our codebook for the subsequent systematic analysis of news stories’ text (Krippendorff 2004). Thus, the resulting codes covered five major categories: type of media source, (impersonated) transmitter of the scam, the general theme and the communication style employed in the scam, and the medium used to convey the scam.

As mentioned previously, the type of media was divided into traditional and specialist media. Faked, or spoofed, transmitters of scams were coded separately to account for the possibility that fraudulent activities were committed under a person’s real name. With respect to the impersonated scam transmitters, the initial reading resulted in three subcategories of “international”, e.g. the World Health Organization (WHO), “governmental” such as the U.S. Center for Disease Control (CDC) or HMRC (Her Majesty's Revenue and
Customs) in the United Kingdom, and “private”, which included individual transmitters such as fake private companies.

Codes for the general themes used in the communicated scams included “health information”, “personal protective equipment (PPE) offers”, “cures”, “vaccines”, “(home) testing”, “relief pay”, “donations” and “fines”. Additionally, we created a failsafe category “other” to account for scams that possessed all other necessary elements but did not fit under any of the previously mentioned sub-categories. The communication styles were coded as “Good Samaritan” and “Shock and Awe” (Kikerpill and Siibak 2021), where the first sub-category noted the creation of a sense of gain by the scam, e.g. by offering to fulfil immediate needs, and the second sub-category marked a respective sense of loss evoked by the deceptive communication, e.g. threatening to turn off utilities.

We coded mediums of transmitting the scam under sub-categories “phishing” (scams via email), “vishing” (scams by phone), “smishing” (scams via text message) as well as “website” for fake websites used as landing pages, “social media” and “medium unspecified”. The latter sub-category was used as a failsafe for instances in which all other necessary elements of an online scam were reported in the news story, but the specific medium of communication was not listed. Due to the initially implemented search criteria, we also separately coded scams that were carried out offline (N=42), marking their respective medium as “traditional” and excluding these stories from the final sample if the news story only covered one scam that was perpetrated offline. The initial reading also made it clear that news stories, e.g. ‘round-up’ articles, include descriptions of more than a single scam, which necessitated separating the count of news stories from the number of scams reported therein. In the course of coding, further news stories were excluded where such news reports only included ambiguous warnings about scams being circulated without providing the elements required under our coding scheme. The final sample comprised N=563 news stories from traditional and specialist media outlets, which reported on a total of N=1040 online scams (see Table 1).

Table 1: Description of the sample (N=563)

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Sample size</th>
<th>Source of the news story</th>
<th>Source of the news story</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of news stories</td>
<td>Number of scams</td>
<td>Traditional media</td>
<td>Specialist media</td>
</tr>
<tr>
<td>January 6</td>
<td>7</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>February 50</td>
<td>93</td>
<td>29</td>
<td>21</td>
</tr>
<tr>
<td>March 260</td>
<td>517</td>
<td>190</td>
<td>70</td>
</tr>
<tr>
<td>April 247</td>
<td>423</td>
<td>182</td>
<td>65</td>
</tr>
<tr>
<td>Total 563</td>
<td>1040</td>
<td>406</td>
<td>157</td>
</tr>
</tbody>
</table>

Results

Main communicative strategies and topics in scams

Our analysis of news stories published in international media suggests that two main types of communicative strategies were used in the scams – the gain-based “Good Samaritan” and the loss-based “Shock and Awe” strategies (see Table 2).

Table 2: Communicative strategies used in the scams

<table>
<thead>
<tr>
<th>Communication style</th>
<th>Communication style</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good Samaritan</td>
<td>Shock and Awe</td>
</tr>
<tr>
<td>January 4</td>
<td>3</td>
</tr>
<tr>
<td>February 83</td>
<td>10</td>
</tr>
<tr>
<td>March 463</td>
<td>54</td>
</tr>
<tr>
<td>April 350</td>
<td>73</td>
</tr>
</tbody>
</table>
Scammers primarily employed “the Good Samaritan” communicative strategy (used in 86.5% of scams), which allows to view the victim as in need of help (demand), the Good Samaritan as recognizing the importance of the event (salience of credible social context, or ‘timing’) and thus aiding the person, who was struggling to fulfill their needs (supply) (see Kikerpill and Siibak 2021). Our analysis suggests that cybercriminals were quick to initiate scams that purported to offer important health information to people in need (Zorz 2020), i.e. 37.5% of all scams across four months concerned health information but a relatively higher proportion of such scams were circulated in January and February (see Table 3). As the pandemic progressed, more advanced attacks started to be employed aiming to take advantage of people’s desire to stay up-to-date on COVID-19 related information. For instance, malware was embedded into a fake live-tracking map that mimicked the original Johns Hopkins University interactive dashboard and bogus COVID-19 smartphone applications (Austin Daily Herald 2020).

As the first months of the crisis witnessed a severe shortage of personal protective equipment (PPE) (see Table 3), con-artists quickly attempted to act as “Good Samaritans” by unloading their “imaginary” stock of highly sought-after supplies (Campbell 2020). Thus, with coronavirus fears providing salience to the situation, various other demand and non-existent supply scams were employed by scammers, e.g. when creating bogus offers for PPE but also when promising to deliver cures and vaccines.

Table 3: Themes used in the scams

<table>
<thead>
<tr>
<th>Scam themes</th>
<th>Scam themes</th>
<th>Scam themes</th>
<th>Scam themes</th>
<th>Scam themes</th>
<th>Scam themes</th>
<th>Scam themes</th>
<th>Scam themes</th>
<th>Scam themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health information</td>
<td>PPE offer</td>
<td>Cure</td>
<td>Vaccine</td>
<td>(Home) testing</td>
<td>Relief pay</td>
<td>Donation</td>
<td>Fines</td>
<td>Other</td>
</tr>
<tr>
<td>January</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>February</td>
<td>62</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>March</td>
<td>214</td>
<td>33</td>
<td>32</td>
<td>28</td>
<td>36</td>
<td>69</td>
<td>38</td>
<td>6</td>
</tr>
<tr>
<td>April</td>
<td>109</td>
<td>43</td>
<td>13</td>
<td>5</td>
<td>18</td>
<td>119</td>
<td>36</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>390 (37.5%)</td>
<td>82 (7.9%)</td>
<td>50 (4.8%)</td>
<td>36 (3.5%)</td>
<td>58 (5.6%)</td>
<td>188 (18.1%)</td>
<td>76 (7.3%)</td>
<td>16 (1.5%)</td>
</tr>
</tbody>
</table>

For instance, a domain registered in Russia offered to sell willing online shoppers “the best and fastest test” for detecting the coronavirus (Venkat 2020). Our analysis of news stories revealed that scammers oftentimes aimed to increase the credibility of such offers by creating fictitious customer reviews, manipulating the number of ‘likes’ of a social media post or the times such posts are shared, i.e. they made use of the principle of social proof (Cialdini 2009).

Although criminals became bolder and more nuanced in their approach to victimizing recipients, most scams in March and April still relied on the creation of an illusionary sense of gain. For instance, due to the growing financial difficulties, scammers started to take advantage of people by communicating various versions of relief payment schemes (see Table 3), e.g. emergency money for groceries (Capodanno 2020) and bogus unemployment grants (Lourie 2020). Moreover, surging unemployment entailed various scams that included fraudulent job offers, e.g. as a secret shopper (WMC 2020).

Solicitations of donations for non-existent charities also started to be more actively spread in March and April (see Table 3), indicating that, in addition to the continued use of the principle of scarcity, the principle of...
reciprocity (Cialdini 2009) became more prevalent in the perpetrated scams. Provided that public knowledge of ongoing hardship, e.g. financial problems due to the spread of the virus, is important for the legitimacy of donation solicitations, the perceived legitimacy of donation requests increased once circumstances entailing from COVID-19 became more widely acknowledged. Furthermore, given that occurring disasters do not affect all socioeconomic groups equally (Parker et al. 2020), requesting aid on behalf of those suffering the most can be viewed as an act of kindness, which is to be returned in the form of a donation. Thus, people fell victim to bogus emails sent, for example, on behalf of the CDC asking for a donation to „help fund its ‘incident management system’ that is coordinating the response to the coronavirus” (Weisbaum 2020).

Although cybercriminals often acted as “the Good Samaritan” coming to the rescue of individuals in need, in a number of scams, the perpetrators attempted to create a sense of loss in the recipients. In this extortion type of communication, which could be seen as a non-military equivalent of a “shock and awe” approach (Kikerpill and Siibak 2021), cybercriminals exploit the credible social context, and use urgency cues (Norris et al. 2019), to present the startled victims with a bifurcation fallacy, i.e. startling the recipient with a choice between bad and worse.

Our results indicate that criminals responded to the imposition of lockdowns and requests to self-quarantine by making assumptions about people’s behavioural tendencies, i.e. the scammers countered lockdowns by communicating messages about fake fines. In the UK, a scam was spread that involved text message alerts pertaining to fake fines imposed on people for violating lockdown rules (Salisbury 2020). In such cases, recipients who consider the contents of the smishing attack as legitimate, have a choice of either paying the fine or facing other, unknown consequences, which are left implicit in the message but could be worse than accepting the fine. Yet, people receiving these fake fine notifications also had the options of ignoring the communication, notifying law enforcement, or attempting to verify the communication through methods external to the original message. Thus, the initial “shock” is necessary to try and get the recipient off balance, making them stop paying attention to what was omitted from the message (Kikerpill and Siibak 2021).

Utilities scams, which entered the picture in March, attempted to exploit unaware individuals by falsely informing them of unpaid bills. The scams subsequently stated that unless a payment is made, the recipients’ utilities will be shut off (KNEB 2020). In April, however, criminals quickly adapted the utilities scams and instead of threatening immediate shut-off of electricity, which was prohibited under local regulations, started to offer discounts on electrical bills (WHSV 2020).

Both of the above examples indicate that perpetrators were also making use of the consistency principle to assure the victims that they needed to honour their previously made commitments and decisions. For instance, in the case of the utilities scam, not only is electricity a necessity in modern society but contracts concluded with utility companies also constitute “previous commitments”. Thus, the combination of assuming compliance with the consistency principle and catching unsuspecting recipients off-guard by alerting them to unpaid bills can result in the payment of non-existent bills. Text messages notifying people of fake fines for violating lockdown rules also follow from the principle of consistency. Not only was the recipient supposedly caught violating rules, which were temporarily enforced to hinder the virus’ spread, but would additionally have to face the possibility that non-payment of the fine can entail even worse consequences, i.e. messages pertaining to fake fines operated on both the principle of consistency and authority.

Furthermore, our analysis suggests that since March, certain health information related communications became more personal and aggressive instead of merely offering information about the virus’ spread. In a smishing attack, for instance, criminals tried to use a so-called “mandatory online COVID-19 test” in order to get recipients clicking on the links included in the text message (Inveiss 2020). The month of March also saw an increase of reported cases of blackmail. In one such instance, criminals contacted people and claimed that they had gained access to the recipient’s personal information, had knowledge of the person’s whereabouts, and threatened to infect the recipient and their family with COVID-19 unless a sum of money was paid (Shein 2020). Except for the instances of blackmail, which is blatantly criminal, other previously described scams are more persuasive when coming from institutions and individuals of authority. Thus, it
is important to analyse which persons and/or institutions were impersonated in such scams.

**Impersonating authority figures**

Regarding credible social contexts such as the COVID-19 pandemic, the embedding of authority cues, i.e., exercising the principle of authority (Cialdini 2009), in scam messages functionally follows a two-step process: 1) acknowledging the existing circumstances of a salient event and 2) choosing and impersonating a relevant authority figure fit for the context.

In the first two months of the crisis, news articles in our sample reported on a plethora of fake emails apparently originating either from certain international organizations, e.g., the WHO, the CDC, or various national entities, e.g., a Japanese disability welfare service provider (Zorz 2020) (see Table 4). The scammers’ view on influencing and gaining the compliance of unsuspecting victims is therefore premised on the perceived legitimacy and authority of the WHO, the CDC or other relevant healthcare institutions. The choice to impersonate the abovementioned institutions enabled to lend further credibility to communications that were unsolicited, exhibiting how scammers rely on the context of the salient current event as well as the perceived authority figure. In addition to fraudulent emails, people also received phone calls, text messages and even faxes impersonating the WHO (Olenick 2020).

**Table 4**: The apparent identity of the sender of the scam (N=1040)

<table>
<thead>
<tr>
<th>Message Source</th>
<th>Message Source</th>
<th>Message Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>False sender</td>
<td>Fraudulent activity committed under person’s real name</td>
<td>International organization</td>
</tr>
<tr>
<td>January 7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>February 91</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>March 508</td>
<td>9</td>
<td>85</td>
</tr>
<tr>
<td>April 413</td>
<td>10</td>
<td>46</td>
</tr>
<tr>
<td><strong>Total</strong> 1019 (98 %)</td>
<td><strong>21 (2 %)</strong></td>
<td><strong>155 (14.9 %)</strong></td>
</tr>
</tbody>
</table>

In March and April, the impersonation of the WHO and organizations such as the Red Cross, was largely replaced by impersonations of government institutions (182 and 199 scams, respectively), indicating that scammers adjusted their lures regarding who or what carries relevance and authority at any given moment during the pandemic (see Table 4). For example, due to the deepening financial hardships caused by the pandemic, financial relief scams and scams related to the distribution of stimulus checks in the U.S. had grown so out of control in April that the Federal Trade Commission opted to warn people by releasing the “FTC Scam Bingo” to get legitimate recipients’ attention (Wolff-Mann 2020).

Additionally, private individuals and those impersonating private companies (250 scams in March, 176 in April), continued to offer non-existent PPE, cures, vaccines but also came up with new approaches to exploiting people’s hardships. For example, surging unemployment entailed various scams that included bogus job offers, offers for free goods, e.g., Netflix passes (Bisson 2020), or free grocery shopping gift vouchers (Morton 2020).

Our analysis suggests that some scammers were brazen enough to not even hide their real identity whilst engaging in fraudulent activity (see Table 4). For instance, a doctor operating out of San Diego was arrested for offering bogus “COVID-19 Treatment Packs”, which included hydroxychloroquine, i.e., the drug promoted by the former U.S. President Donald J. Trump (Briquelet 2020). By directly referencing former President Trump, the latter case also implies the use of the principle of liking. Hence, scammers widely utilized the option of impersonating well-known health authorities and public figures in lending credibility to their fraudulent attempts.

**The main communication mediums used**

Our analysis suggests that although fraudulent emails i.e., phishing (53.5% of the scams) were the dominant
type of communication mediums used by the cybercriminals during the first four months of the pandemic (see Table 5), the social-engineering scam ecosystem reached a type of equilibrium already in March with a wider implementation of smishing, vishing, social media, and fake websites.

Table 5: Communication medium used for the scams

<table>
<thead>
<tr>
<th>Communication medium</th>
<th>Phishing</th>
<th>Vishing</th>
<th>Smishing</th>
<th>Website</th>
<th>Social Media</th>
<th>Medium unspecified</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>2</td>
<td></td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>February</td>
<td>66</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>March</td>
<td>275</td>
<td>83</td>
<td>52</td>
<td>59</td>
<td>19</td>
<td>29</td>
</tr>
<tr>
<td>April</td>
<td>214</td>
<td>55</td>
<td>71</td>
<td>49</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>557 (53.5%)</td>
<td>142 (13.6%)</td>
<td>131 (12.6%)</td>
<td>117 (11.3%)</td>
<td>32 (3.1%)</td>
<td>61 (5.9%)</td>
</tr>
</tbody>
</table>

Our findings indicate that the choice of a medium was dependent both on the expected audience of the scams, as well as the main purpose of the scam. For example, analysis of news stories suggests that numerous scams crafted to target elderly people were carried out by phone. Vishing gained prevalence in March, when robocalls were employed to offer unaware victims the chance to obtain fake COVID-19 home testing kits (Romm 2020). In April elderly people also became the targets of fake charity calls (Williams 2020) as well as a typical grandparent scam, which involves contacting elderly people by phone later at night, was adapted to reflect COVID-19 conditions (Levine 2020).

Fraudulent websites (used in 11.3% of the reported scams), often with bogus customer reviews, were predominantly created for the sale of non-existent PPE, fake cures, and vaccines (see Table 5). Donation solicitations were most often spread on social media where platform affordances also enable the use of the principles of reciprocity and social proof (Cialdini 2009). Our analysis also suggests that smishing attacks were used to lure people into clicking links sent in text messages, which notified recipients that a fake fine had been imposed or that someone they had come into contact with had tested positive for COVID-19 (Wall 2020).

Overall, perpetrators used a variety of mediums, often giving preference to specific mediums depending on the type of scam.

Discussion

Motivated by the scarcity of current literature on the role and importance of context in social engineering attacks, our study analysed scam messages that were spurred by the COVID-19 pandemic. To frame our study, we proposed and employed the concept of mazephishing, which allows to analyse the interplay between credible social context and the scam messages it spurs. In the following, we discuss the empirical support our results and analysis provided for the concept of mazephishing, including the potential of this concept for future studies, and address certain limitations of our study.

Our analysis suggests that two primary types of communicative strategies - the “Good Samaritan” and the “Shock and Awe” strategies - were used in the scam messages. The observation that the majority of scams in our sample employed the gain-based “Good Samaritan” strategy is important, as influencing derived from fear of loss is usually considered more effective (Williams and Polage 2018). It is possible that the credible social context of the COVID-19 pandemic prompted a stronger focus on using the fulfillment of people’s psychological and material needs as the relevant lure. This type of general psychological approach has also been emphasised in previous studies of social engineering attacks (Norris et al. 2019). In our sample of news stories, initial scams circulated in January and February used offers of “health information” as the lure and were later joined by bogus offers for PPE, cures and vaccines, which were either in short supply or did not exist at all. Similarly, scams that employed the “Shock and Awe” strategy were also adapted to the specifics of the COVID-19 pandemic. For example, without actual lockdown restrictions in place, messages notifying
of the imposition of fake fines would be irrelevant. Furthermore, absent an ongoing health crisis, threatening to infect the families of message recipients if the scammers’ demands are not met, would make little or no sense. Hence, while scams in our sample exhibited an apparent preference for the “Good Samaritan” strategy, the significance of the ‘timing’ or credible social context element of mazephishing was strongly present with the use of both strategies.

Our study also indicated that the choice of impersonation in the scams often directly followed from the credible social context of the COVID-19 pandemic. As mentioned previously, initial scams purporting to offer important health information were made to appear as if originating from well-known health and disease control institutions, e.g. the WHO and the CDC. This reliance on a specific choice for impersonation is significant, as the scam themes were altered, e.g. the WHO either offering health information or soliciting donations, but the perceived credible source was retained in the scams. Similarly, other aspects that surged to accentuated prevalence during the pandemic, e.g. provision of “Good Samaritan” support for grocery shopping (Morton 2020) or home entertainment (Bisson 2020), were covered in the scams by the impersonation of relevant sources (supermarkets and Netflix). The impersonation of health authorities during the COVID-19 pandemic was in line with similar previous occurrences, e.g. during the 2014 Ebola outbreak or the 2016 Zika virus (RiskIQ 2020, Mouton and de Coning 2020). The public’s need for information is increased during health crises (Ogbodo et al. 2020) and the initial information voids are often exploited by scammers. Furthermore, while free offers of goods or services are common in scams (Atkins and Huang 2013), gift vouchers and home entertainment offers can be viewed as having taken on a different meaning in the context of the COVID-19 pandemic. For instance, under non-pandemic circumstances, gift vouchers may be simply viewed as bonus gains that supplement other sources of income or help save money. During the COVID-19 pandemic, however, supermarket gift vouchers can additionally take on the meaning of necessary resources that help mitigate temporary unemployment and Netflix passes provide activities for people that are expected to abide by lockdown rules. Hence, otherwise common scam tactics can take on new meaning depending on the credible social context within which such scams are circulated.

With respect to the mediums used in scam delivery during the first four months of the COVID-19 pandemic, our study provided significant insights into the general ecosystem of scams by including the emails, text messages, phone calls, fake websites as well as scams perpetrated through social media. Previous studies into the presence of principles of persuasion have been limited to email-based scams (Zielinska et al. 2016; Lawson et al. 2020). As other studies have emphasised (Kikerpill and Siibak 2021), the scam ecosystem is far more varied than email-based scams alone. This notion carries importance in our study, because the medium used in scams notes the ‘place’ element in the mazephishing framework. For instance, based on our sample, scams that offered home testing kits or solicited donations from the elderly were predominantly perpetrated by calling the victim on the phone. Additionally, offers for bogus cures and vaccines were often presented on fake websites. Although our data does not allow us to claim that people visited such websites directly, e.g. without prior email or text message solicitation, the medium remains significant because it constitutes the final destination, i.e. the ‘place’ element, of the scam. Furthermore, without the inclusion of other mediums such as text messages, important scam types, e.g. fake fine notifications (Salisbury 2020), would have been completely absent from the study. Important empirical support for the mazephishing concept comes from the general observations that scammers often, in fact, did prefer different mediums for delivering different types of scams. To an extent, this variation in preferences can also be explained by the affordances provided by various mediums. For instance, scams that utilize websites and social media posts can benefit from embedding the principle of social proof in the form of fake positive reviews.

The ‘technique’ element of mazephishing, i.e. how scammers elicit action from and gain the compliance of unsuspecting victims through social engineering (Hadnagy 2018), received support with respect to all six of Cialdini’s principles of persuasion. However, certain principles such as that of authority were more prominent in the scams studied. As mentioned previously, the impersonation of health authorities during the first months of the COVID-19 pandemic indicated strong support for the principle of authority. Offers for difficult-to-obtain PPE, but also for non-existent cures and vaccines, constituted clear use cases of the principle of scarcity, while solicitations of donations employed the principle of reciprocity. In the scams
studied, there were considerably fewer use cases of the principles of social proof and liking. This can partly be explained by the fact that while news stories reported on scams involving the creation of bogus websites for the sale of PPE and/or cures and vaccines, the actual websites were not displayed in the news stories, e.g. as screenshots. In the majority of cases, this effectively prevented us from analysing whether fake reviews or endorsements had been used alongside bogus product offers. However, use of the principle of consistency/commitment was clearly present in the cases of utilities scams, fake fines as well as bogus messages notifying people that they had come into contact with an infected person. The latter two examples drew their significance only from the ongoing pandemic and therefore provided strong support for the *mazephishing* concept.

Overall, our analysis indicated that perpetrators at least perceive salient current events as able to provide credibility to various scams and rush to exploit such circumstances as credible social context for their messages. Thus, *mazephishing* is located between phishing attempts that rely entirely on manufactured social contexts and spear-phishing attacks, which add elements of personal or individual context in addition to relying on relevant but broader social context. Although *mazephishing* attacks are opportunistic in nature, such attempts carry additional importance because these fully utilize the affordances provided by credible social context. All constituent elements of the proposed *mazephishing* concept, i.e. ‘timing’, ‘place’ and ‘technique’, were empirically supported by our study results. The ‘timing’ element was supported by the fact that the COVID-19 pandemic spurred scams that would have carried little relevance (or made little sense) in other contexts. Additionally, our analysis also suggested that otherwise common scams, e.g. free offers, can take on an alternative or a more specific meaning depending on the credible social context within which such scams appear. The ‘place’ element found support in the fact that perpetrators seemingly prefer different scam delivery mediums for different scams, which does not, however, preclude cross-use. The ‘technique’ element was supported by the possibility of detecting Cialdini’s principles of persuasion in the actual scam messages.

Hence, the concept of *mazephishing* allows to explain the core aspects of context-aware phishing under sustained salient circumstances like the COVID-19 pandemic. The *mazephishing* framework may also be utilized in analysing well-known periods of scam circulation, e.g. tax declaration periods or holidays with a strong commercial emphasis such as Christmas. Additionally, the analytical framework can provide support in explaining the proliferation of scams in the immediate aftermath of disruptive events such as natural disasters or disease outbreaks. Thus, the three-element *mazephishing* concept may support media literacy and cybersecurity training initiatives (Kikerpill 2021) by introducing and emphasising the significance of credible social context in the general scam ecosystem.

**Limitations and future research**

This study has some limitations. The collected sample of news stories did not provide us with information on the actual intentions of scammers nor the relative success rates of the analysed scams. However, and importantly, the news stories provided descriptions of circulated scams that allowed us to analyse the contents of the scams in the credible social context these appeared in at the time, similarly to the work by Lallie and colleagues (2020). Previous studies (Williams and Polage 2018) have attempted to construct such credible scam message contexts for study participants, but did not detect particular relevant effects regarding the efficacy of the scam messages. This could potentially be explained by the fact that constructed contexts are understood by participants and/or recipients of scam messages, but are not ‘lived’, i.e. constructed contexts do not provide a similar interpretative backdrop for scam message recipients.

Regarding the actual intentions of perpetrators circulating scams, previous research has suggested that even the law enforcement community is often not aware of the criminals’ backgrounds (Button et al. 2009). Although this type of information would certainly make for an insightful addition, it did not present a limitation on our study. While we are now learning that social engineers may pay specific attention to salient social contexts (Steinmetz et al. 2021), it is not clear whether all scammers knowingly and intentionally embed principles of persuasion into scam messages or if certain criminal actors simply mimic the formats and/or contents of already available scams. Nevertheless, assessing the presence of principles of persuasion in
scam messages still carries importance, for instance, in the future design of cybersecurity training and digital literacy initiatives. The concept of mazephishing is a suitable framework for further investigations into how scams are created on the basis of, and can be explained by, emerging and salient social circumstances. Future studies utilizing the concept of mazephishing could additionally focus on local salient events as well as events or circumstances that, unlike the COVID-19 pandemic, provide a credible social context for phishing scams only for brief periods of time.

**Conclusions**

The uncertainty and fears entailed by the COVID-19 pandemic drove cybercriminals to adapt their main tools, i.e. socially engineered messages circulated on a mass scale, to fit the salient social context. Focussing on the first four months of pandemic circumstances in early 2020, our study showed that scams and scam messages reported in international news media were quickly adapted as real-life events unfolded. When people were confused about credible pandemic information, cybercriminals were there to ‘inform’. Where people looked for personal protective equipment, COVID-19 treatments or even vaccines, cybercriminals were there to ‘supply’. As the financial burdens and uncertainty caused by the pandemic circumstances deepened, scammers were present to ‘offer relief’. Alongside bogus offers to fulfil people’s immediate material and psychological needs, cybercriminals also threatened to make things worse for recipients, e.g. by suggesting that people’s electricity will be shut off at a time where the use of digital means had increased significantly.

Our proposed concept of mazephishing captures the under-emphasised significance of social context in social engineering attacks. The presented analysis showed that criminals follow the changes taking place in people’s lived experiences, choose suitable communication mediums, and employ relevant persuasion tactics in a concerted effort to exploit difficult social circumstances. Given that our content analysis provided clear evidence on the diverse use of the COVID-19 pandemic as social context in social engineering attacks, future studies of online scams should focus more on what connects potential victims rather than what separates them. In other words, instead of demographics or personality traits, the focus should be on the shared lived experiences of people under specific social circumstances, the relevant scam messages being circulated, and the media used to deliver such messages.

**Acknowledgements**

The preparation of this article was supported by a personal scholarship (no. 1.1-17/31-1) awarded to the primary author by the Estonian Police and Border Guard Board.

**Author information**

Kristjan Kikerpill (kristjan.kikerpill@ut.ee); ORCID: 0000-0002-1523-3208

Prof. Andra Siibak (andra.siibak@ut.ee)

**References**


