On SMS Phishing Tactics and Infrastructure

Aleksandr Nahapetyan*, Sathvik Prasad*, Kevin Childs*,
Adam Oest†, Yeganeh Ladwig†, Alexandros Kapravelos*, Bradley Reaves*
*North Carolina State University
†PayPal, Inc.

{anahape,snprasad,krchilds,akaprav,bgreaves}@ncsu.edu
{aoest,yesafaei}@paypal.com

***Abstract*—In 2022, the Anti-Phishing Working Group reported a 70% increase in SMS and voice phishing attacks. Hard data on SMS phishing is hard to come by, as are insights into how SMS phishers operate. Lack of visibility prevents law enforcement, regulators, providers, and researchers from understanding and confronting this growing problem. In this paper, we present the results of extracting phishing messages from over 200 million SMS messages posted over several years on 11 public SMS gateways on the web. From this dataset we identify 67,991 phishing messages, link them together into 35,128 campaigns based on sharing near-identical content, then identify related campaigns that share infrastructure to identify over 600 distinct SMS phishing operations. This expansive vantage point enables us to determine that SMS phishers use commodity cloud and web infrastructure in addition to self-hosted URL shorteners, their infrastructure is often visible days or weeks on certificate transparency logs earlier than their messages, and they reuse existing phishing kits from other phishing modalities. We are also the first to examine in-place network defenses and identify the *public* forums where abuse facilitators advertise openly. These methods and findings provide industry and researchers new directions to explore to combat the growing problem of SMS phishing.**

1. Introduction

Phishing refers to sending legitimate-looking communication to steal a user's sensitive information [1]. It remains an effective attack vector that harms the end-user and the companies the attackers impersonate. Despite this, the state of the art is always stuck in a game of cat and mouse between phishing detection and detection evasion [2]. In this game, mobile users are more likely to arrive first at a phishing website and are more likely to submit their credentials [3]. While dating back to 1996, phishing has seen an increase in traffic as of late [4], [5] with the FCC reporting a 70% increase in complaints between 2015 and 2022.

Between 2014 and 2016, the NIH identified at least ten hospitals where the hackers gained unauthorized access to hospital systems through phishing in the United States [6]. With COVID-19, we saw an increase in SMS Phishing campaigns [7], and in March of 2023, the FCC adopted its first rules focused on scam texting. The regulation required wireless carriers to block texts from illegitimate numbers, citing a rise of robotext complaints from 3,300 to 18,900 per year. [5]

Prior work has studied phishing by dynamically examining the web page contents[2] for cloaking behavior, looking at features within the URLs[8], or by collecting and examining phishing kits [9]. However, all these works used a set of URLs and did not examine them in the context of their delivery mechanism. On the other hand, Tang *et al.* crawled Twitter using Optical character recognition (OCR) and clustering to isolate and examine phishing and spam SMS reports [10], yielding a dataset of 21,918 phishing messages, and concluding that modern A2P providers are bad at distinguishing the messages that they observed as spam.

To expand our understanding of SMS Phishing, we use Public SMS gateways: websites that allow users to view messages sent to numbers owned by the gateway via a web interface. They often facilitate fraud or bypass proof-of-humanity verifications [11], [12]. Reaves *et al.* were the first to identify malicious traffic on public SMS gateways, but at the time, they only identified 1,393 URLs with one detection on VirusTotal for two years. In 2018, they reported that VirusTotal marked only 64 URLs as a "phishing site" [11].

We use public SMS gateways and capture 67,991 phishing messages over 396 days. With a more rigorous definition of what constitutes a phishing URL, we capture URLs due to the gateways we choose and their overall traffic than prior work[12], [11]. Over a year, we captured 2,866 URLs with three or more detections in VirusTotal and at least one "phishing site" detection. While prior work has emphasized using A2P providers as the cause of the increase in phishing traffic, we present semi-corporate bulk SMS advertisements on LinkedIn and Telegram and show that A2P providers are not the only way to send bulk SMS messages.

In this paper, we present the following contributions:

C1. We characterize modern SMS Phishing, highlighting the size, duration, and targeting pattern in individual messages, larger clusters of SMS campaigns, and even larger clusters of operations. Using APWG as a feed

of general phishing URLs, we describe the differences between SMS and general phishing URLs. We isolate 35,128 SMS Phishing campaigns by grouping identical messages when excluding URLs, phone numbers, email addresses, and One-Time codes (OTCs) in the message body. After using various threat intelligence sources to collect metadata about their web infrastructure, we conclude that most campaigns last 13.02 days, target English-speaking audiences and phone numbers in the US and UK, and have a trimodal distribution of the ratio of the number of messages to the number of destinations. We define what campaigns and operations mean in Section 4.

C2. We deterministically cluster phishing operations based on their infrastructure and discuss different insights that can be gained from this approach. We isolate 2,106 operations, showing that most operations are short-lived and are hosted on popular cloud providers. We identify 652 multi-campaign operations with an average of 52 campaigns, a lifespan of 275 hours, and 1.6 urls (1.097 hostnames) per operation. Most operations keep the same hostname (96%), and 84% of operations do not mention a public organization detected by SpaCy. The top organizations mentioned by the remaining 334 operations were Apple, Apple Pay, CommBank, DHL, and MyGov.

C3. A dataset of phishing messages, campaigns, and operations.[1]

Overall, we expand upon the status quo of phishing research by examining a large dataset of SMS phishing messages from an atomic scale of messages and URLs to a larger cluster of operations. We capture more messages, to our knowledge, than past research has and examine patterns that only telco providers could. The rest of the paper is structured as follows: Section 2 will bring up relevant background in both Phishing and SMS, Section 3 will talk about prior work in the field in more detail, Section 4 will talk about our crawling setup, and analysis rational, Section 5 will present our findings and discuss their implications and Section 6 will examine case studies and findings through manual examination.

2. Background

In this section, we provide helpful background on SMS and phishing concepts.

2.1. The Short Message Service

The Short Message Service (SMS) originated in early versions of second-generation cellular networks. SMS eventually came to be a popular, if not critical, function in mobile networks. What was originally a minor add-on feature came to rival if not actually surpass voice as a predominant communications medium.

1. https://github.com/wspr-ncsu/sms-phishing

SMS is technically limited. Famously, messages have a 160-character limit (using 7-bit characters, leading to 140 octets in length) with no formatting or other annotation or metadata. [2] These limitations are an artifact of the design process: SMS was effectively "shoe-horned" into extra space in control channels of the voice network.

Each message is transferred through mobile networks in a store-and-forward model analogous to email. Messages are sent and received on behalf of subscribers by entities called SMS Centers (SMS-Cs) in each mobile network. In addition to the official SMS-Cs, there are a large number of External Short Message Entities (ESMEs). This category describes network elements that are operated by third parties in order to provide SMS service to non-mobile customers. Virtually all systems that provide the ability to send an SMS without an actual phone and mobile subscription use an ESME.

SMS-Cs and ESMEs communicate to their host and external networks over telco signaling protocols. The most common and prominent of these is Signalling System 7 (SS7). Ideally, ESMEs must still engage with network gatekeepers, pay required fees, and conduct themselves according to industry standards of behavior. These standards include preventing and disconnecting abusive actors and going through official processes and channels by mobile carriers to ensure bulk messaging is acceptable to their subscribers and is not a nuisance. For example, an SMS provider for a package delivery service may be asked to register each messaging "campaign" with mobile providers, including message contents and send rates. Sending messages outside of those parameters — even if there is a legitimate reason like a spike in delivery demand — can lead to messages being blocked or future access being denied.

This state of affairs can be frustrating for legitimate senders, but it is a massive problem for bulk message abusers (as intended). Because carriers are proactive in limiting bulk access — often called "application to person" messaging or "A2P" — phishers and other malicious senders have to find an alternative means for message distribution.

The answer in practice seems to be to exploit the fact that while bulk, A2P messaging is heavily monitored for abuse, blocking and censorship of messages between individual subscribers (person-to-person or "P2P" messages) is virtually unheard of. As such, providers to malicious or otherwise undesirable traffic purchase individual phone subscriptions (often pre-paid, largely anonymous SIM cards) and multiplex many cards in a VoIP-to-SMS gateway, colloquially termed a "simbox" [13].

2.1.1. Public SMS Gateways. While there are a wide variety of use cases for ESMEs, one that seems to be perennially popular is web interfaces that provide public phone numbers to receive text messages. Prior work termed these services "public SMS gateways." While these services advertise as meeting an unmet need on behalf of individuals

2. There is also a rare form of data-carrying SMS message primarily used by networks to facilitate over-the-air updates, but these are not visible or available to end-users.

who cannot otherwise receive SMS, in practice, these services largely facilitate the creation of phone-verified accounts. Web services use phone verification to limit the number of accounts or identities an individual can create or to geographically limit users to certain countries through testing for access to a phone number in that area. Public gateways are the simplest among several techniques to evade this restriction.

Public SMS gateways also provide one of the few sources of real, large-scale text message data for research. While a service that openly and publishes messages sent to a number that will likely not demonstrate "typical" personal SMS traffic, it still provides a window into certain kinds of messages, including malicious and phishing SMS messages.

2.2. Why bother with SMS in the '20s?

As SMS grew in popularity, users came to expect more functionality. This led mobile devices to support large messages that exceeded the original 160-character limit through message concatenation. Later came network support for the Multimedia Message Service (MMS) in third-generation networks to support images and short videos. These features rely on both the networks' and devices' support. Only the original SMS is guaranteed to work.

The concurrent rise of smartphones and widespread and affordable data service led to a tectonic shift in mobile messaging. First, Apple introduced iMessage with the iPhone; while the app supported SMS, it preferred its proprietary protocol transmitted over the phone's data channel. Many third parties, such as WhatsApp, Telegram, and Signal, created analogous products not tied to a device and sent messages "over the top" (OTT) using the data channel.

Frustrated by the fact that iMessage is only available on Apple products, Google Android has been championing the deployment of a successor to SMS and MMS called Rich Communication Services (RCS), and it is now available to millions of subscribers worldwide. It is not, however, supported by Apple devices. Further, while WhatsApp and others became the defacto messaging platforms in many countries, iMessage and SMS/RCS remain the most popular forms of mobile messaging.

The upshot of this history is that **SMS remains the only available messaging service that can reach all US.** This fact is equally true for individuals messaging friends and associates, legitimate businesses communicating with their customers, and malicious actors who seek a communication channel to precipitate fraud and other abuses.

2.3. Phishing

Phishing is a type of social manipulation wherein a malicious individual pretends to be a reliable coworker, acquaintance, or reputable entity to trick the target into divulging confidential data or granting access to their network. These deceptive tactics can take various forms, such as emails, text messages, or phone calls. 2022 was a record year for phishing, with the APWG logging more than 4.7 million attacks. Since the beginning of 2019, the number of phishing attacks has grown by more than 150% per year. [14] Phishing websites face an active adversary of threat intelligence companies and law enforcement. Many phishing pages employ cloaking to prolong their lifetime and avoid detection and classification.

Cloaking has been used for search engine optimization purposes [15] and to hide malicious pages from security research and automated detection tools. There are two forms of cloaking on the web: **Server-Side** and **Client-side** [2]. Server-side cloaking refers to deciding whether to server a website based on the HTTP request's IP addresses, user agents, and other markers. These decisions are made on the server and can not be examined without looking at the website's source code. [15]. Client-side cloaking is more sophisticated and usually requires the mechanism (mainly in JavaScript) to be sent over to the user [2]. Thus, they employ heavy obfuscation to hide their internal logic. Client-side cloaking can employ different fingerprinting techniques and require interaction, such as dismissing alerts of captchas.

With increased phishing attacks came ready-to-use software to set up a web page, **Phishing-Kits**. These kits are either used by the authors or sold to malicious actors [9]. During the set-up process, inexperienced actors might leave the zip file for the phishing kit on the server, which can be discovered using URL-fuzzing tools like KitPhisher [16]. These phishing kits can provide insights into cloaking behavior, such as block-listed IPs/User-Agents, data exfiltration techniques, and setup functionality.

3. Related Work

3.1. Phishing and Internet Abuse

Phishing has emerged as one of the most frequent and effective attack vectors. Correspondingly, it has received an overwhelming amount of attention from security researchers. Here, we briefly highlight some recent work most relevant to this project. For a more thorough treatment, we recommend readers review any one of several research surveys on the topic [17], [18].

Phishing Ecosystem: Previous research has focused on understanding the phenomenon; In [4] and [19] Varshney *et al.* and Banu *et al.* both present a comprehensive overview of the lifecycle, styles, and taxonomy of phishing. However, they focused primarily on email and study URLs gathered from threat intel sources and not in the context of their delivery method.

Cloaking: Cloaking and the different medium of delivering a phishing URL poses a hard problem for detection tools. Verma *et al.* [8] studies different features of malicious URLs and proposes a set of online and batch learners for classification purposes. Oest *et al.* [9] demonstrated that phishing kits are a viable way to study the phishing ecosystem, and that proposes a generic 4-type breakdown for phishing URLs. They highlight the targeting of anti-phishing infrastructure in server-side *.htaccess* filtering.

Oest *et al.* in [20] demonstrated the effectiveness of cloaking against modern detection systems. Past studies into Phishing infrastructure have shown that this kind of adversary employs Server-side and Client-side cloaking; however, to avoid server-side cloaking, they rotated only user agents/IPs. Zhang *et al.* examines both APWG data and a public dataset of phishing websites and taxonomizes them into eight different types of cloaking. Broadly, they identify and study cloaking behavior that requires user interactions, attempts to fingerprint the target, and attempts to evade automated bots. [2] Alghamdi *et al.* [21] shows that users tend to fall for phishing due to not knowing how to make good decisions online. SMS Phishing has been shown to have an immense impact, even going as far as disrupting EV charging grid, as shown by Soykan *et al.*. [22]

Threat intelligence: While helping guide our research efforts, past research has found OpenPhish and Phishtank to be limiting as Phishing BlackList due to their auto-removal policy [23] and Phishtank, in particular, is less complete than proprietary sources of reports [24].

3.2. SMS Abuse

SMS Spam: The bulk of research on SMS messaging abuse has focused on unwanted messages, and interested readers would be well served by reading surveys on the topic [25], [26]. Detecting SMS Spam has been at the heart of much of this work, with notable works including Murynets and Jover [27] and papers from Jiang et al. [28] characterizing the state of SMS spam in 2013. That same year, Jiang et al. also released a paper on Graystar, where they examined messages sent to unused telephone numbers to detect and characterize spam activity [29]. Later efforts described detection in the context of large-scale bulk messaging collected through a public SMS gateway [30] and using more modern language embeddings for detection [31].

SMS Data Collection and Characterization: Virtually all work on SMS abuse has been data-driven, yet few datasets are available to the public, leading to creativity on the part of prior work. We follow in the footsteps of Reaves et al. [30], [32] in using public SMS gateways as a data source, as have several others [33], [34]. Srinivasan et al. leveraged datasets of SMS spam complaints collected by the FCC and third parties and combined them with historical DNS and blacklist datasets to characterize SPAM URLs [35]. Balduzzi et al. described directly deploying SIM cards to collect unsolicited calls and SMS [36], while Tang et al. combined optical character recognition and Twitter feed monitoring to identify tweets and posts about SMS spam and collect spam messages and commentary [37]. Finally, Li *et al.* [38] use messages identified as coming from fake base stations in China as a large-scale dataset to study SMS Spam in that country.

SMS Phishing Compared with SMS spam, SMS phishing has a far less substantial body of research. Earlier work on SMS abuse found SMS-based phishing to be rare, even compared to SMS spam as a whole [32]. Our findings in this paper confirm this phenomenon. The most notable work on SMS phishing is quite recent. Rahman et al. also conducted controlled simulated SMS phishing campaigns, demonstrating a high hit rate among victims, which provides empirical evidence that these campaigns remain effective and are thus likely to continue [39]. Timko and Rahman evaluated several applications on a dataset of 20 confirmed SMS phishing messages, finding most apps that they tested failed to filter them [40]. Mambina et al. work with a dataset of 297 phishing campaigns (described by the authors as "unique Smishing SMS") identified by a mobile provider in Tanzania. While the total number of campaigns was relatively low, the total volume of messages collected over two days was 1.8 million, indicating that the number of individuals affected by a single campaign can be quite high [41]. Finally, Liu et al. work with a dataset of 11,475 "spearphishing" SMS campaigns that they derive from labeled spam SMS provided by a major mobile security vendor in China; they focus primarily on the semantic properties of these messages, including customization with personal information (such as name or other PII). [42]

Our work differs from all prior work in having a broad, longitudinal lens (1 year) over multiple countries to focus solely on SMS phishing, and it happens to be the largest dataset of SMS phishing campaigns to date. Moreover, our study is entirely replicable because our oracles and data sources are publicly available. Beyond scale, we are distinct in our deep analysis of phisher operational habits and infrastructure, covering everything from campaign lifetime to resource registration to shared infrastructure. To the best of our knowledge, we are also the first to explore the forums and platforms that serve as the black markets for services that enable SMS phishing.

4. Methodology

This section describes our data collection process, how we aggregate individual messages into campaigns, and how we identify SMS phishing operations. We then describe our SMS defenses measurement experiment and crawling setup to uncover bulk SMS service providers.

4.1. Data Collection

We collect SMS message data by crawling eleven SMS gateways. These SMS gateways are accessible through a web browser. Messages received by SMS gateways are published on a website hosted on a specific domain listed in Table 1. We developed web crawlers to scrape these websites using Scrapy. These webpages publish the raw message body along with any identifiers (URLs, One-Time Passwords or OTPs, etc.), and describe useful metadata information about the message. The metadata for each message contains the source phone number, the destination phone number, and a timestamp. Although various gateways implement these fields differently, our robust crawlers extract the metadata and the message body from each gateway, storing them in a local database. We run our crawlers every hour, except for sms24, receive-sms-online, and receive sms, which we run every two hours.

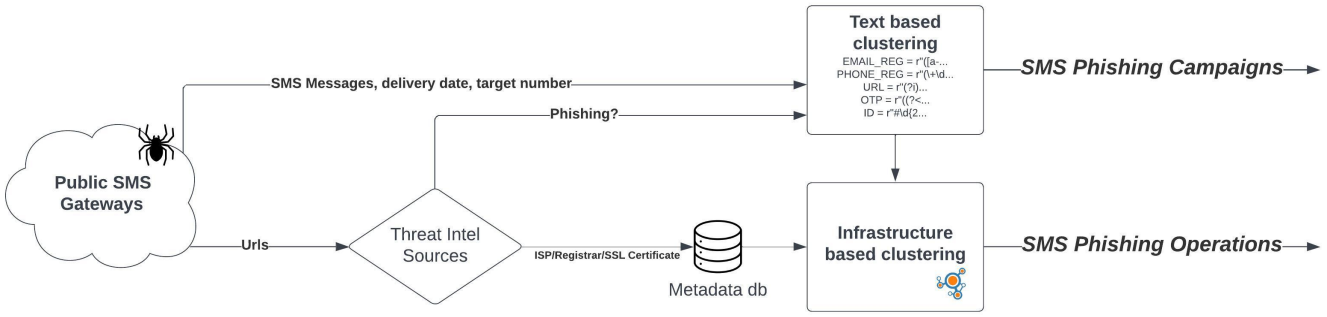

Figure 1: Overview of the crawling and clustering pipeline

We stopped receiving data from receive-sms-online because their reverse proxy provider banned our IP address.

Computing Message Timestamp: Although the metadata contains the message's timestamp, most SMS gateways do not display a precise timestamp. Instead, they display a relative timestamp (e.g. *Two hours ago, 4 days ago, etc.*). When a precise timestamp is unavailable, we use the unit of the displayed timestamp to compute a time window for each message.

Deduplication across gateways: Although we collected data from eleven different gateways, these gateways occasionally contain the same destination phone number and, therefore, the same message. We combine the target number, the message body, and the time window to discard any duplicate messages collected by the crawlers.

Extracting and enriching identifiers: We enrich the raw data collected from the gateways by post-processing the message body and the metadata. Using the country code and numbering format, we identify the country to which the destination phone number belongs using *libphonenumber* [3]. Post-processing also involves extracting URLs, email addresses, phone numbers, and One-Time Codes embedded within the message body.

Using a combination of VirusTotal[43], APWG[44], and Google Safe Browsing[45], we detect if a URL extracted from the message body is malicious or benign. A URL is flagged as **malicious** if it meets at least one of the three criteria : (i) at least three vendors on VirusTotal flagged the URL as malicious, (ii) APWG flagged the URL as malicious, (iii) Google Safe browsing marks the URL as phishing. We mark the message containing the URL as a **malicious SMS Message**. Among all the malicious messages, we identify a subset of messages and flag them as **phishing SMS Messages**. These messages contain a URL flagged as a **phishing URL** by either VirtusTotal or APWG. We do this to remove further scam-like behavior that may not strictly be phishing, like crypto-casinos, loan applications, and malicious file downloads.

Collecting domain and registrar information: For each phishing URL, we collect Internet Service Provider information and domain registrar information using WHOIS. We query TLS transparency logs[46] to capture the issue time, start date, and expiration date for TLS certificates assigned to that domain. Since there can be multiple certificates issued with different start dates and expiration dates for the same domain, we use a combination of the certificate issue date, expiration date, and the message timestamp to identify the certificate that was active when the campaign was active. This active certificate is the one that was issued before the timestamp of the earliest message containing the URL with a certificate expiration date that is beyond the timestamp of the most recent message containing the URL. If none of the certificates meet these criteria, we flag the certificate issued closest to the timestamp of the first message with the URL as the active certificate. We then compute the *time of deployment* for each URL using the TLS entry timestamp of the active certificate.

Cloaking and redirection: We perform two analyses involving the server the phishing page is deployed on, measuring the time of deployment using SSL certificates and querying WHOIS to identify the cloud service provider (CSP). For this, we are not using the URLs in the messages but rather the domain names, which may or may not be hosted on the same server as the phishing page (or be a URL shorter, cloaking endpoint, etc). After discovering the initial link, mobility and cloaking can change the chain's final endpoint. For example, 272 of the URLs' final destination, according to VirusTotal, was *google.com*. Our TLS certificate issuing time analysis and looking at cloud hosting providers (CSPs) for the servers behind these websites would have incredibly unreliable results in such cases, potentially leading to incorrect conclusions since the final destination might not be a phishing page. In Section 5, we find that the majority of the e2LDs in these messages are public or private URL shorteners, and the deployment strategies for those services might vary from that of phishing sites; thus, we exclude these domains only for the TLS certificate and CSP analysis.

Collecting phishing kits: Phishing kits are low-effort resources that enable bad actors to set up functional phishing domains. Occasionally, threat actors deploying these phishing kits unintentionally leave the archive of the phishing kit source code on the phishing domain [47]. After identifying phishing domains, we use KitPhisher [48] to collect any phishing kits hosted on these domains.

3. https://github.com/google/libphonenumber

TABLE 1: Total number of messages, the total number of phishing messages, and dates crawled for each gateway[4]

Gateway	Number of messages	Shows past recent messages	Dates Crawled	Number of phishing messages
sms24[.]me	199,561,540	Yes	2019-08 - 2023-07	33,240
receivesms[.]org	1,037,620	No	2021-12 - 2023-07	2,207
freephonenum[.]com	594,523	No	2021-12 - 2023-07	939
7sim[.]org	17,383,181	Yes	2021-08 - 2023-07	82,172
temp-number[.]com	154,240	Yes	2022-09 - 2023-03	31
receivesms[.]cc	4,868,011	No	2023-03 - 2023-07	9,182
sms-online[.]co	256,847	No	2021-01 - 2023-07	2,379
freeonlinephone[.]org	826,625	Yes	2021-12 - 2023-07	1,073
sms-online[.]co	427,932	No	2021-01 - 2023-07	2,362
receive-sms-online[.]com	1,050	No	2019-12 - 2021-07	0
receivesms[.]co	8,093,113	Yes	2021-12 - 2023-07	13,220

4.2. SMS Phishing Campaign

SMS phishing campaigns have been defined in numerous ways. In this work, we define an SMS phishing campaign as a collection of identical phishing messages which may have different identifiers. These identifiers include URLs, phone numbers, email addresses, and One-Time Codes. We uncover SMS phishing campaigns by aggregating individual phishing messages that use identical language while being agnostic towards any changes in the identifiers present in the message. We operationalize our phishing message aggregation technique by first normalizing individual messages by replacing phone numbers, email addresses, One-Time Codes, and URLs with an appropriate custom token. For example, *"Your OTP is 1234"* is converted to *"Your OTP is **#OTP**"*, and *"Check this offer on xyz[.]com"* becomes *"Check this offer on **#URL**"*. After normalizing the messages, we group identical phishing messages into phishing campaigns. Aggregating individual phishing messages into SMS phishing campaigns enables us to study the variations across different campaigns and study their evolution, as shown in Section 5. We present the top 4 campaigns (by destination) in Table 2.

4.3. SMS Phishing Operations

We project phishing messages and the web infrastructure used by these messages onto a non-directional bipartite graph. By extracting the connected components of the graph, we uncover SMS phishing operations. A bipartite graph, by definition, has two types of nodes. In our construction, the two node categories are (i) content nodes and (ii) infrastructure nodes. The content nodes represent the body of the phishing messages, and the infrastructure nodes represent the phishing URLs identified previously. The construction of the bipartite graph captures the relationship between the content nodes and their corresponding infrastructure node. For each content node, we draw an edge to connect the content node to an appropriate infrastructure node representing the URL embedded in the message content. We then extract the connected components from the graph to uncover SMS operations. A graph-based approach to uncovering SMS phishing operation ensures that the connections between messages and web infrastructure is non-probabilistic, thereby eliminating challenges related to selecting an appropriate clustering algorithm or tuning hyper-parameters.

We uncovered numerous campaigns where the complete message body was the URL, with no other content. Some other campaigns contained a few characters of additional text along with a URL. These campaigns (listed in Table 6 in the appendix) were removed from the bipartite graph due to the lack of an accompanying message body along with the URL.

Phishing campaigns often impersonate well-known brands and consumer companies [49]. We used SpaCy's Named Entity Recognition pipeline [5] to extract references to brands and organizations within the message body. Using polyglot [50], we identify the language used in the message body.

4.4. Measuring SMS Delivery Rates

We measured the SMS delivery rate of Application-to-Peer (A2P) services to wireless phones. These messages were generated in a controlled fashion using a bulk messaging provider and were sent to eleven different wireless phone numbers across various carriers, as shown in Table 5 in the appendix. We used a well-known type of SMS phishing scam (package delivery scam) to craft the message body. The message body also contained a benign URL that we controlled. The crafted message was *"ASAP! Your parcel is waiting to be shipped. Please confirm the shipping information #URL."* Although the URLs in each of these messages used the same top-level domain (*redacted-research-domain[.]com*), we used different subdomains for every message. We recorded the requests made to these subdomains. This allowed us to study any requests made to the specific subdomain while the SMS traversed across the network until it was delivered to the destination.

4.5. Bulk SMS Service Marketplace

Anecdotal evidence from industry professionals and telecom fraud experts has indicated that malicious bulk SMS

4. receive-sms-online blocked our IP after 24 hours, however, the messages from the gateway were still included in our analysis pipeline

5. Model Name: en_core_web_trf

TABLE 2: Top 4 SMS Phishing campaigns by the number of destinations

Credit Agricole: Vous avez cree un nouveau destinataire le 18/08 a 14:30:41. Si ce n'est pas vous, veuillez annuler ici.: #URL
ASSURANCE MALADIE : Votre nouvelle carte vitale est disponible. Remplissez ce formulaire afin de rester couvert : #URL
Apple Customer, Your Lost iPhone 13 Switched ON. Check Turn ON last location: #URL Apple Support.
USPS has dispatched your package. However, there is a typo in your delivery address. Your address must be corrected: #URL

delivery services use social media and messaging platforms to advertise their services. To study this ecosystem, we crawled LinkedIn and Telegram phishing groups for posts related to bulk SMS services. We successfully uncovered 49,417 posts advertising 3rd party bulk SMS services. Some posts on LinkedIn were re-advertised in Telegram phishing channels.

5. Results

In this section, we will start by examining overall trends in phishing traffic, look at the phishing URLs and messages at an individual scale, and cluster them into campaigns and then further into operations.

Finding 1: *On average, these gateways received 172 phishing messages per day ($\sigma = 431$), making them excellent honeypots for phishing messages*. Between *01-05-2022* and *01-05-2023*, we isolated 67,991 phishing messages. The average percentage of phishing messages for a day was 0.037% with a median of 0.010% (of that day's traffic). While these might sound low, we were processing, on average 536,814 messages ($\sigma = 276,664$) per day. We should note that we encountered 28 days that had no phishing messages [6]. As these gateway numbers are fast changing [11], there is no central repository of disposable phone numbers malicious actors can use to ensure that their traffic does not end up here. This means that many bad actors do not block these gateway numbers or that these gateways are being used for testing purposes, a hypothesis we expand upon in Finding 3 and Section 6. Through the year, the highest number of phishing messages we recorded was 2,758 on 2023-05-15.

Finding 2: *We do not observe periodicity in phishing message volume*. Prior work studying robocall abuse [51] found that robocall volume had a strong periodicity with a heavy bias towards weekdays and US working hours. We may see something similar in phishing messages. We examine the autocorrelation of daily phishing message volume to answer this question, as shown in Figure 2.

Autocorrelation computes the correlation of a time series with a time-shifted version of itself by several data points, called "lags." For example, the value at lag 3 indicates the average correlation of all data points with a data point 3 days later. This analysis shows that the strongest predictor of phishing message volume is the message volume of the immediately preceding days. Such a result is consistent with a time series with a strong trend component.

We see in Figure 3 that the typical message volume varies wildly over time, indicating that it is indeed factors other than periodicity dictating message volume. Nevertheless, it

6. 26 days were between 2023-01-07 and 2023-03-26, during which we were experiencing infrastructure issues.

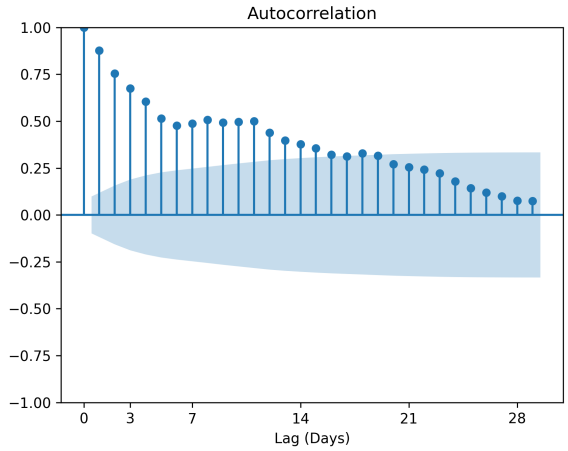

Figure 2: Daily SMS phishing volume in public gateways is not periodic and largely dominated by outside factors.

is still possible that there may be a short-term (e.g., weekly) periodicity, even with a varying trend. We explored this possibility using the standard practice of computing the autocorrelation of the daily difference in phishing message volume. We further did similar analyses of the total message volume and the ratio of phishing to messages and found no meaningful indicators of periodicity, as judged by a low autocorrelation at calendrically significant lags (e.g., seven days, 14 days, 30 days).

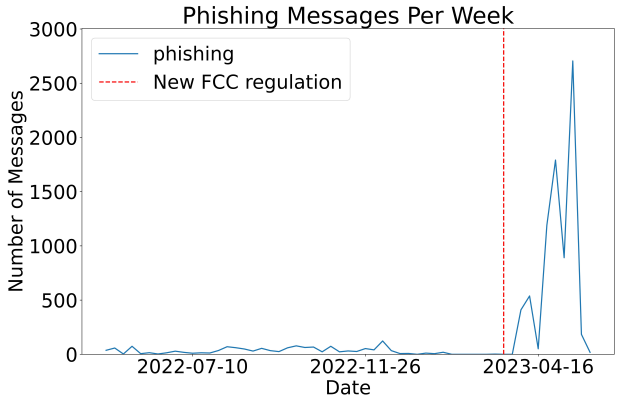

Figure 3: Line plot of daily received phishing message volume. We see a notable increase in volume between the FCC publishing an order to block robotexts, and it's implementation deadline

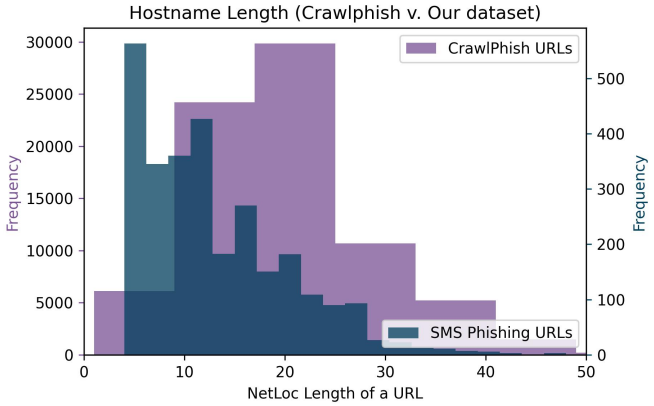

Figure 4: Naturally, a histogram of phishing domain lengths would show that URLs used in SMS phishing tend to be shorter than less-limited channels[2]. However, this plot more interestingly shows the upperbound of domain length that phishers will tolerate – more than 30 characters in some cases.

5.1. Phishing URLs, Domains, and Hosting

Finding 3: *We find that 70.1% of phishing URLs redirect to a different hostname by looking at metadata provided by VirusTotal, and despite their many redirection, hostnames and URLs tend to be short-lived.* Between *01-05-2022* and *01-05-2023*, we gathered 2,866 distinct phishing URLs. Most URLs redirect to a different hostname, sometimes more than once. On average, the URLs redirect 2.3 times, with a standard deviation of 1.7, a median of 2.0, and a max of 9. The average observed lifetime of a hostname, determined by subtracting the first time we captured a phishing message containing that hostname from the first time we observed it, is 12.241 days ($\sigma = 46.436$), with the median being 0.53 hours long, indicating URLs are observed for a short amount of time. Out of 1,696 URLs contained in at least two messages, the average observed duration of a URL, that is, the number of days between the first and last message observed, is 4 days with a median of 0.19 hours. The top 75th percentile URLs had an observed duration of 0.34 hours. It should be noted that VirusTotal may be deceived by cloaking URLs; we discuss this further in Section 6.4 and Section 7.

While cloaking is hard to measure using our oracles (discussed in Section 7), we found a case with selective cloaking based on the mobile device's operating system in **strongdry[.]com/iccu**. The page would redirect non-desired targets (based on browser configuration) to *google.com* if you did not have a mobile User-Agent and would switch assets based on an Android and IOS User-agent. Additionally, 272 URLs report final redirection to *google.com/* in VirusTotal, suggesting that they successfully blocked VirusTotal from crawling their page (though we can not be sure if this was an IP block or a User-agent one).

Finding 4: *Some URLs had their TLS certificates issued after we observed the first message with none active when we did.*

TABLE 3: Top 5 e2LD from our dataset. We see that the top 5 e2LDs are all URL shorteners.

Second-Level Domain	Number of URLs	Public shortener?
tx[.]vc	263	No
shrtlink[.]net	173	No
qi[.]lv	83	Yes
shor[.]td	70	Yes
kvo6[.]io	50	No

This could be an indication of a phishing campaign that is using these gateways as a testing ground for their delivery routes. In Figure 5, we show the distribution of the time between the TLS certificate being issued and the first message being sent in 24 hours. We isolated 592 hostnames that do not redirect to a different hostname and have a TLS record. 16 (2.703%) hostnames have their TLS certificate issued 12 or more hours after we first see them on the gateway; we speculate that these could be instances of malicious actors testing their delivery rate and if some keywords stop their messages from being delivered.

Finding 5: *Majority of TLS certificates are issued to these phishing domains more than a week before we first see them in phishing messages*. We look at hostnames for URLs that do not appear to be redirected to a different hostname and are themselves the final destination. We do this because, otherwise, intelligence regarding URL shorteners would affect the results. Redirects, especially from public shorteners, do not have a meaningful delay between deployment and the message observed. Most certificates are issued more than six days before the first message appears. There are 239 (40.372%) hostnames that have their TLS certificates issued a week before we first see a message appear on the gateway. This could indicate an actor trying to avoid detection via someone watching TLS certificate logs.

Finding 6: *Phishing infrastructure observed relies on popular hosting providers, the most common being CloudFlare, followed by Amazon and Google-Cloud-Platform*. There are many tricks cybercriminals use to build resilient infrastructure for their phishing operations. However, using VirusTotal, we found that most of these domains were hosted using common cloud providers. There were 551 hostnames with an A or AAAA record that did not redirect to a different hostname. The top service provider is AWS, with CloudFlare being a close second. The top two locations for these servers are in the US, with the 3rd being Moscow. In Table 4, we show the top 5 service providers hosting by the number of phishing hostnames we've captured.

As a comparison, we used VirusTotal to identify the autonomous system (AS) owners of the 3471 phishing hostnames collected from APWG on July 28th, 2023. We find that the top 4 service providers were CloudFlare (39%), Hostinger International Limited (6%), AWS (5%), and Tencent Cloud (2%). Illicit market hosting and multiple proxies require complicated setups and resources; our finding suggests that these actors find a low barrier to entry by abusing resources from these cloud providers.

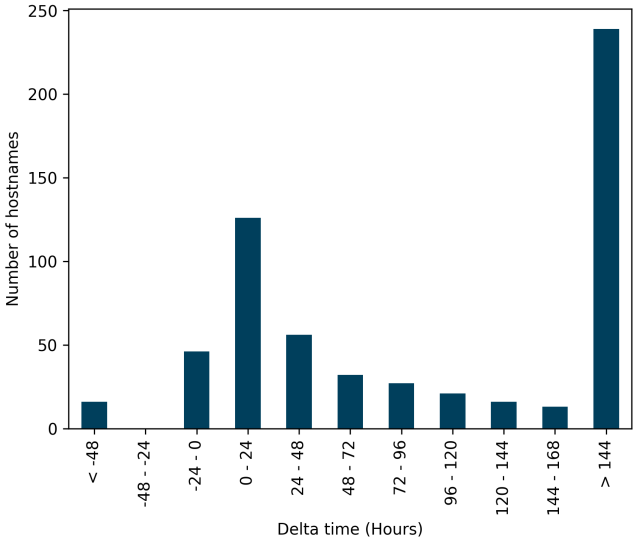

Figure 5: This histogram shows that the time between when a phishing domain TLS certificate is issued and when we see the first message using it is most often more than 6 days.

TABLE 4: The top 5 ISP hosting phishing websites (as reported by VirusTotal)

CSP	Hostname count
Cloudflare	44
Amazon	37
Google Cloud Platform	30
Delis LLC	23
Digital Ocean	19
Other	298

5.2. Phishing Kits

Kitphisher is a tool that finds phishing kits left on the server after setup by fuzzing the URL. Phishing kits are ready-to-use software, and sometimes they are left behind after setup. In total, we studied three kits that were identified from phishing URLs.

Phishing kit 1 was for a website that arrived with an SMS of "CITIZENS -We noticed an unusual activity due to security update; kindly visit." However, it contained a simple HTML file redirected to a page flagged as Malicious/phishing by three vendors on VirusTotal. Navigating to the domain shows a loading screen with no content.

Phishing kit 2 used the SMS message "informed invalid delivery address update here." It contained a list of blocklisted IPs and loaded a favicon from the original USPS page while linking to it; it loaded local CSS/JS and image assets. The page made sure to add any X-Forwarded-IPs, an HTTP header identifying the originating IP address of a client connecting to a web server through a proxy server, to the block list. Moreover, it contained hardcoded CNC credentials and disallowed access from one specific user-agent (Chrome 104 on Windows 10) while permitting other desktop user agents. This kit was in PHP and had no deployment script.

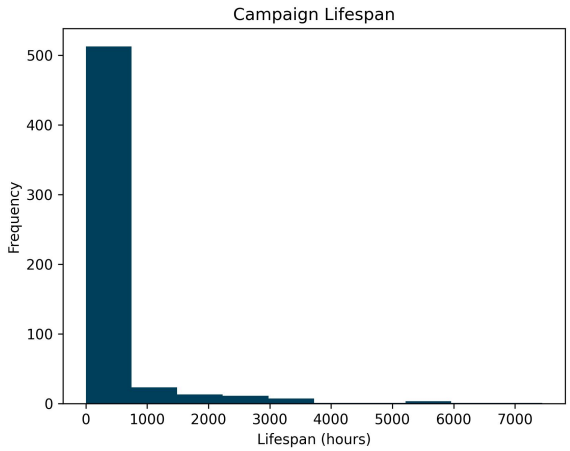

Figure 6: Histogram of time between first- and last-seen phishing messages in a campaign shows a concise operation window in most cases.

Lastly, phishing kit three was designed to resemble a UPS clone. It delivered the message "Shipped. / Your parcel is arriving today," with the link leading to *ups[.]com[.]track[.]perfhost[.]net* instead of the legitimate UPS website. The kit hosted all its assets locally and employed HTML/CSS/JS without server-side logic. Notably, it used Telegram for data exfiltration. Although it reported users' IP addresses, no robots.txt, .htaccess, or any other logic in the JavaScript attempted to cloak the contents. Similar to the previous kits, phishing kit 3 lacked a deployment script.

5.3. SMS Phishing Campaigns

We observe a total of 35,128 SMS phishing campaigns during the 396 day observation period. We find that average phishing campaigns last 13.02 days with a median lifetime of 0.82 days. The longest-lived campaign, *"Begin your PocketWin adventure today with an up to £10 FREE BONUS! #URL 40xWR. 14Days. MaxWD£50. T&C's Apply. Stop Msg? Text PW"*, lasted 265 days, and the shortest-lived campaign lasted 2.0 seconds. In Figure 6, we show the distribution of the lifetime of the SMS phishing campaigns. Due to the longest campaign being so general, we manually removed the campaigns deemed too general before clustering operations; we discussed this methodology in Section 4.

Finding 7: *SMS phishing campaigns tend to have a message-to-destination ratio of either 3, 2, or 1.* We find a trimodal distribution of the ratio of destinations to the number of messages in a campaign and the number of destinations it targeted. Some campaigns target either once, twice, or 3-4 times the same numbers. In Figure 7, we present the distribution of the ratio of the number of messages to the number of destinations (1 meaning for every number a campaign targetted it sent out one message, while .2 meaning for every one destination, the campaign sent out five messages). Malicious actors may be spamming the same

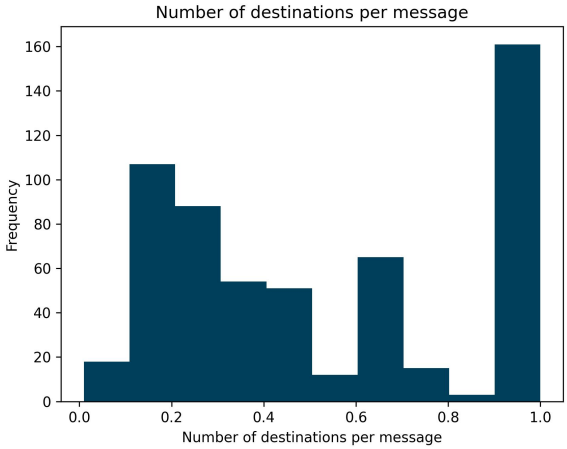

Figure 7: The distribution of the ratio of the number of messages to the number of destinations.

numbers multiple times to compensate for poor delivery rates; no prior work has measured the delivery rates for bulk messaging providers we identify in Section 5.5.

Finding 8: *We find the tolerable upper bound for domain length in SMS phishing to be 30+ characters*. We observed an abundance of short domains that redirect to a secondary page. The top e2LD was **tx[.]vc**, which accounted for 263 of the URLs across 2,492 messages. We manually go through the top 5 e2LDs and check if they are publicly available URL shorteners while the URLs share similar characteristics to some, notably, *shrtlink[.]com* being a valid URL shorter, we fail to find a public interface to register URLs for 3 out of the five. We find that the top domains used in SMS phishing messages are relatively short. With all of the top second-level domains in Table 3.*VirusTotal* reports that most of the URLs would redirect to a different hostname (sometimes more than once). Compared to the phishing URLs analyzed in Crawlphish, the URLs we observed in the SMS phishing messages are relatively short; we compare the hostname length in Figure 4. The character limits of SMS are a limitation for the text of the phishing campaigns and the URL length, and on top of that, due to a lack of hyperlinks, longer messages might also instill confusion in victims. The average campaign (with all PII removed) has 90 characters, which leaves 70 characters for URLs, phone numbers, and email addresses.

Finding 9: *Phishing SMS campaigns in our dataset target the English-speaking audience primarily in the US and UK*. Using polyglot to extract the language of these campaigns, we find that the top language (in numbers from across countries) is English. The top 4 languages sent to US numbers outside English were Russian, Danish, and Norwegian. The highest number of phishing campaigns were sent to numbers in the US, UK, Canada, Portugal, and Ukraine. In Figure 7 in the appendix, we show the top 10 languages used in phishing campaigns. Polyglot reports that most campaigns do not have more than one language, with only 1.22% having multiple languages. This finding is a drastic shift from findings in prior work in Reaves *et al.* [11]. There are many reasons behind this pattern; in 2022, the FCC reported an increase in package delivery phishing messages [52]; to mimic USPS and UPS messages, malicious actors might be using English. Moreover, bypassing the numbers through *libphonenumber*, we see that the top countries that received phishing traffic in our dataset are the United States, France, and the United Kingdom.

5.4. SMS Phishing Operations

We clustered 35,128 into 2,106 operations. Viewing our SMS phishing campaign dataset as a graph of phishing Campaigns and URLs, we isolate connected components of this graph as operations. Out of 2,106 operations, we find that, on average, most operations are short-lived (less than an hour).

Finding 10: *Phishing operations, even multi-campaign and multi-URL operations, tend to share hostnames but may vary service providers*. We find that most operations (96.01%) share only a single hostname, with only 50 operations having more than one hostname. Out of those 50 operations, 17 had the same service provider behind them, while 25 had two different service providers. The most hostnames used in a single operation is 13, with four operations using more than three hostnames. On average, an operation has used 1.347 URLs ($\sigma = 1.63$), with 114 (5.4%) operations having more than two URLs We separate the campaigns into 1606 short-lived (observed for less than two hours) and 500 long-lived operations (observed for more than two hours). Short-lived operations, on average, have 1.055 destinations, mentioned 0.2 organizations, and had 1.4 campaigns (3.3 messages). Long-lived operations, on average, have 6.9 destinations, mentioned 0.214 organizations, and had 65.7 campaigns (124.7 messages). We find 652 multi-campaign operations with an average of 51.629 campaigns ($\sigma = 350.587$), a lifespan of 274.6 hours, and 1.6 URLs (1.1 hostnames).

Finding 11: *Despite clustering around shared infrastructure, our definitions of phishing operations yield groups of closely related text*. We used Levenshtein distance to quantify how similar campaigns within an operation were. By using the python library *TheFuzz* [53], we can get a similarity score between 0 and 100 (100 being two identified strings). These operations have an average score between campaigns (not including the URLs) of **88.4** ($\sigma = 14.7$) with a median **93.0** and an average of **21.8** campaigns per operation. We present a sample operation in Figure 8.

Finding 12: *Most operations (84.141%) do not mention a detectable organization using SpaCy, with only 334 operations mentioning an organization, and the top organizations being Apple (22 operations), Apple Pay (12), CommBank (10), DHL (9), and MyGov (9)*. We find that NLP pipelines have a hard time identifying named entities in our dataset as SpaCy finds only six operations of USPS, while a single regex search yields 23 operations. Formal grammar, regular expressions, or pipelines trained specifically on SMS might be better suited to identify brand impersonation, as in some cases,

these names are used as labels in the SMS, for example, *USPS: Your package needs your attention (3.00$ unpaid fee). and confirm the delivery address here: #URL* was not detected via spacy and was with a simple regex.

These operation clusters open an avenue for enforcement agencies and researchers to study the phishing ecosystem by identifying its most prominent actors, as they link large volumes of traffic into small clusters that share web infrastructure. In our case, starting with 65K messages, we reduced it to clusters of 2,106 operations, which are not just textually similar, but also share web infrastructure. This approach can also be used to build a campaign reputation engine.

5.5. Bulk SMS Services

Finding 13: *We confirm that URLs sent in P2P traffic are not monitored by any US carrier we tested.* As discussed in Section 4, we attempted to measure the delivery rate of phishing campaigns against SMS Firewalls. We used ten numbers from 6 wireless providers and sent a text message from a personal phone line and an A2P provider. In Table 5 in the appendix, we break down the carriers and the delivery status for both sent from a personal number and sent from A2P, but we observe a 100% delivery rate of the message. We also find no crawls on the URL used in the message, suggesting no active fingerprinting of a website happening in SMS firewalls. Only four requests are sent to the URL; all visits have Google-page-render or Googlebot in the user agent with IPs within Google's AS. We assume these are messenger-dependent renderers that proxy through Google's infrastructure to avoid privacy leaks. We conclude that the primary deterrents of SMS phishing are the monitoring and incident response of the A2P provider with no SMS Firewalls in place that engage in active URL crawling.

Finding 14: *There is an illicit market of individuals advertising bulk SMS services. Bulk SMS services posts openly advertise on public platforms their willingness to transit nuisance or illegal traffic*. To understand the ecosystem of SMS abuse better, we crawl LinkedIn for posts advertising bulk SMS services. We find and analyze 49,417 posts that advertise bulk SMS services. We show an example of such a post in Figure 9 in the appendix. Most of these actors use messenger apps like WhatsApp, Telegram, Skype, or Viber. 20717 posts link at least to one of these apps in their posts, with Whatsapp being the most popular accounting for 72% of the usernames. Only 17% of the emails used in these posts were Gmail, while the rest were custom domains or other email providers. We discuss sample posts in Section 6.

6. Discussion

6.1. Why SMS Abuse Continues

Our work shows that SMS phishing is rampant and, given its continued use, likely effective. This raises the question: why does SMS abuse continue despite regulatory attention and measures taken by providers? The answer lies in how the measures taken fail to address the fundamental problems: weak notions of identity in telephony and distinguishing legitimate and fraudulent messages and services.

Rather than address these incredibly difficult problems, providers have treated SMS abuse as an *access control* problem. Carriers have added increasingly deployed onerous vetting processes before an entity can send application-to-peer (A2P) traffic. A2P includes messaging used to provide desired, urgent communications like delivery updates as well as appointment reminders, not just bulk marketing. Each carrier's process is different, but is time-consuming and difficult to comply with on the part of legitimate messaging enterprises. Our work provides circumstantial evidence that these measures have succeeded in closing the A2P path to malicious operations, but in practice, it has likely served to simply move abuse onto P2P channels.

In the US, the FCC is increasingly concerned with SMS abuse, and their approach to-date has been to propose or adopt measures from voice networks that were unsuccessful in that domain. These include requiring providers to block messages from invalid sources [5], expanding the Do-Not-Call list to SMS, and even proposing STIR/SHAKEN style attestation in SMS [54]. The first such order was issued in March 2023, with a September 2023 implementation deadline. In the interim, we observed a large spike in message volume in the SMS gateways we monitor, indicating that the problem is certainly not solving itself. Of course, all of these solutions assume that either the adversaries will suddenly begin respecting the law or that a trustworthy claimed source telephone number will actually help subscribers know whether a message's content is legitimate. The former is implausible, and the latter overlooks the fact that legitimate messages from enterprises rarely have a consistent SMS source number.

6.2. Advertisements for Bulk SMS Service

Looking through Telegram Groupchats labeled phishing general; we found that the author of the post in Figure 9 in the appendix cross-posts advertisements in Telegram as shown in Figure 10 in the appendix; we also observed cross-posting between LinkedIn and Fiverr. This indicates that these actors use multiple platforms for advertising their services. Within the posts we've collected, we have seen offers of free trials and content delivery. It should be noted that in Figure 10 the user is openly listing "SPAM" as a traffic category, but not in Figure 9; we theorize that this is to avoid being flagged by the platform.

6.3. SMS Gateways as a Source of Threat Intelligence

Finding 15: *These SMS gateways are being used to test delivery routes*. With a free resource like the public SMS gateways, phishing campaign operators might be tempted to use it as a testing ground. Figure 8 is an example of an operation that displays characteristics of testing a delivery

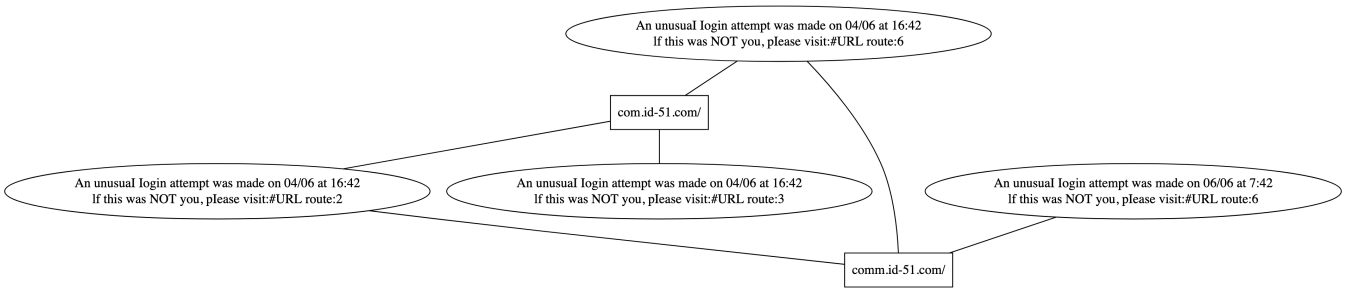

Figure 8: Sample operation with four different campaigns and two URLs. Ovals are the texts of the campaigns, and rectangles are the URLs. An edge between an oval and a rectangle indicates that a variation of that campaign uses that URL.

network. Establishing when the message might have been seen in the wild is tricky as we are most likely the first to submit it to *VirusTotal*. However, this is not the only operation we have seen with this text pattern. Using a simple regex of *route:[0-9]*, we see that 161 campaigns across 45 different operations employ this tactic. Out of 12 hostnames that these campaigns used that did not redirect to a different one, 7 TLS certificates were issued before we first saw the message, while 4 were issued after.

Considering the near-constant traffic of phishing messages, these gateways can be a valuable tool for brands to detect ongoing attacks or, in some cases, pre-meditated attacks (while actors are testing their delivery routes) against their customers by constantly crawling these gateways.

7. Limitations

All data-driven research on messaging abuse faces methodological limits. Prime among them are ground truth and accurately detecting a phenomenon that, by definition and design, does not want to be detected. In our work, we leverage the decades of research and experience in detecting phishing embodied in VirusTotal and Google SafeBrowsing. These tools are arguably the best available but are not perfect. They may produce false positives or negatives or they may face delays in detecting malicious activity.

VirusTotal aggregates dozens of other detection engines, each of which has its own biases and limitations. Our choice to require multiple detections before labeling an SMS "phishing" is consistent with prior work, but the detection threshold used by all researchers is generally arbitrary. We aimed to increase precision at the cost of recall without threshold, though if we had chosen a different value, we may have seen more or less SMS phishing. VirusTotal's phishing is also known to have minimal resilience against client- and server-side cloaking by malicious websites, which limited our ability to make claims about the endpoints of redirection chains.

Our data source, public SMS Gateways, is unique and may present a biased view of SMS phishing in a few ways. The gateways drop any MMS traffic, so phishing/malicious MMS messages are not captured. The gateways we monitor focus almost exclusively on Western countries, and the numbers they have and use are well outside the norm of typical use by an individual. Gateways also limit our visibility in some ways. Analyzing the message's sender is often unreliable, if not impossible. Gateways vary in how they present sender identity, and they may not provide the full sending phone number or any other caller ID information. Finally, we may see more or less SMS phishing than is typical for an individual subscriber. On one hand, we have a relatively small window into the totality of the SMS network, and public gateways tend to keep numbers live for only a short time. On the other hand, the possibility that phishers use gateways as test infrastructure may give us above-average visibility. Only further research with better, currently unavailable datasets would be able to address this question.

8. Conclusion

This paper presented an in-depth analysis of modern SMS Phishing campaigns and operations, utilizing data from public SMS Gateways as honeypots. Our findings shed light on the SMS Phishing activities' scale, duration, and targeting patterns. From the analysis of individual SMS campaigns, we observed a trimodal distribution in the ratio of messages to destinations, indicating that many campaigns target the same numbers multiple times. We also found that SMS Phishing URLs tend to be shorter, often utilizing URL shorteners, due to the limitations imposed by the character limits of text messages. Analyzing larger clusters of SMS campaigns, we identified operations that share common infrastructure and URLs. The presence of multi-campaign operations points to sophisticated and organized phishing operations that attempt to maximize their reach and impact. Our approach of clustering operations based on infrastructure provided valuable insights into the organization and characteristics of SMS Phishing operations.

This study highlights the urgent need for robust defenses against SMS Phishing attacks. Traditional SMS Firewalls and A2P providers should actively engage in URL crawling and incident response to protect users from falling victim to these scams. Additionally, regulatory measures like those implemented by the FCC can act as deterrents, but they must be complemented with active monitoring and response mechanisms. In conclusion, the insights gained from our analysis provide valuable information for cybersecurity professionals,

policymakers, and researchers in the ongoing fight against SMS Phishing and similar cyber threats. By understanding the tactics and infrastructure used by malicious actors, we can develop more effective countermeasures and protect users from falling prey to these deceptive attacks.

Acknowledgement

We thank the anonymous reviewers and the shepherd for their feedback and suggestions. We extend our thanks to APWG and VirusTotal for allowing us to use their APIs for this research. This material is based upon work supported by the National Science Foundation under Grant Numbers CNS-2142930 and CNS-2047260, the North Carolina Partnership for Cybersecurity Excellence, the Office of Naval Research (ONR) under grant N00014-21-1-2159, funds from the 2020 Internet Defense Prize, and Paypal.

Competing Interests Disclosure: Dr. Reaves has received in-kind support from Bandwidth, ZipDX, VirusTotal, Anti-phishing Working Group, and Google Cloud. He is a member of the ACM, IEEE, USENIX, the Communications Fraud Control Organization (CFCA), SIPForum, and the Alliance for Telecommunications Industry Solutions (ATIS). No third party directly influenced the conduct or publication of this work.

References

[1] What Is Phishing? Examples and Phishing Quiz. Cisco. [Online]. Available: https://www.cisco.com/c/en/us/products/security/email-security/what-is-phishing.html

[2] P. Zhang, A. Oest, H. Cho, Z. Sun, R. Johnson, B. Wardman, S. Sarker, A. Kapravelos, T. Bao, R. Wang, Y. Shoshitaishvili, A. Doupé, and G.-J. Ahn, "Crawlphish: Large-scale analysis of client-side cloaking techniques in phishing," in *2021 IEEE Symposium on Security and Privacy (SP)*, 2021, pp. 1109–1124.

[3] IMB, "Mobile Users 3 Times More Vulnerable to Phishing Attacks," Security Intelligence. [Online]. Available: https://securityintelligence.com/mobile-users-3-times-more-vulnerable-to-phishing-attacks/

[4] G. Varshney, M. Misra, and P. K. Atrey, "A survey and classification of web phishing detection schemes," *Security and Communication Networks*, vol. 9, no. 18, pp. 6266–6284, 2016. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/sec.1674

[5] FCC, "FCC Adopts Its First Rules Focused on Scam Texting," 2023. [Online]. Available: https://www.fcc.gov/document/fcc-adopts-its-first-rules-focused-scam-texting

[6] A. Wright, S. Aaron, and D. Bates, "The big phish: Cyberattacks against u.s. healthcare systems," *Journal of General Internal Medicine*, vol. 31, pp. 1115–1118, 2016. [Online]. Available: https://api.semanticscholar.org/CorpusID:39941124

[7] Ivan-Righi. Clickbait to Checkmate. [Online]. Available: https://www.digitalshadows.com/blog-and-research/sms-based-scam-targets-us-smartphones-and-accesses-victim-locations/

[8] R. Verma and A. Das, "What's in a url: Fast feature extraction and malicious url detection," in *Proceedings of the 3rd ACM on International Workshop on Security And Privacy Analytics*, ser. IWSPA '17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 55–63. [Online]. Available: https://doi.org/10.1145/3041008.3041016

[9] A. Oest, Y. Safei, A. Doupé, G.-J. Ahn, B. Wardman, and G. Warner, "Inside a phisher's mind: Understanding the anti-phishing ecosystem through phishing kit analysis," in *2018 APWG Symposium on Electronic Crime Research (eCrime)*, 2018, pp. 1–12.

[10] S. Tang, X. Mi, Y. Li, X. Wang, and K. Chen, "Clues in Tweets: Twitter-Guided Discovery and Analysis of SMS Spam," 2023-04-09. [Online]. Available: http://arxiv.org/abs/2204.01233

[11] B. Reaves, L. Vargas, N. Scaife, D. Tian, L. Blue, P. Traynor, and K. R. B. Butler, "Characterizing the security of the sms ecosystem with public gateways," *ACM Trans. Priv. Secur.*, vol. 22, no. 1, dec 2018. [Online]. Available: https://doi.org/10.1145/3268932

[12] J. M. Moreno, S. Matic, N. Vallina-Rodriguez, and J. Tapiador. Your Code is 0000: An Analysis of the Disposable Phone Numbers Ecosystem. [Online]. Available: http://arxiv.org/abs/2306.14497

[13] B. Reaves, E. Shernan, A. Bates, H. Carter, and P. Traynor, "Boxed out: Blocking cellular interconnect bypass fraud at the network edge," in *USENIX Security Symposium*, Washington, D.C., Aug. 2015.

[14] T. A.-P. W. Group, "2022 4th quarter report," 2022. [Online]. Available: https://docs.apwg.org/reports/apwg_trends_report_q4_2022.pdf

[15] R. Duan, W. Wang, and W. Lee, "Cloaker Catcher: A Client-based Cloaking Detection System," [Accessed 29-07-2023]. [Online]. Available: http://arxiv.org/abs/1710.01387

[16] cybercdh, "Kitphishr: A tool designed to hunt for phishing kit source code," 2023. [Online]. Available: https://github.com/cybercdh/kitphishr

[17] Z. Alkhalil, C. Hewage, L. Nawaf, and I. Khan, "Phishing attacks: A recent comprehensive study and a new anatomy," *Frontiers in Computer Science*, vol. 3, 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fcomp.2021.563060

[18] S. Salloum, T. Gaber, S. Vadera, and K. Shaalan, "A systematic literature review on phishing email detection using natural language processing techniques," *IEEE Access*, vol. 10, pp. 65 703–65 727, 2022.

[19] M. N. Banu and S. M. Banu, "A comprehensive study of phishing attacks," *International Journal of Computer Science and Information Technologies*, vol. 4, no. 6, pp. 783–786, 2013.

[20] A. Oest, Y. Safaei, A. Doupe, G.-J. Ahn, B. Wardman, and K. Tyers, "PhishFarm: A Scalable Framework for Measuring the Effectiveness of Evasion Techniques against Browser Phishing Blacklists," in *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2019, pp. 1344–1361. [Online]. Available: https://ieeexplore.ieee.org/document/8835369/

[21] H. Alghamdi, "Can Phishing Education Enable Users To Recognize Phishing Attacks?." [Online]. Available: http://arrow.dit.ie/scschcomdis/99/

[22] E. U. Soykan, M. Bagriyanik, and G. Soykan, "Disrupting the power grid via ev charging: The impact of the sms phishing attacks," *Sustainable Energy, Grids and Networks*, vol. 26, p. 100477, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352467721000485

[23] S. Bell and P. Komisarczuk, "An Analysis of Phishing Blacklists: Google Safe Browsing, OpenPhish, and PhishTank," in *Proceedings of the Australasian Computer Science Week Multiconference*, ser. ACSW '20. Association for Computing Machinery, 2020, pp. 1–11. [Online]. Available: https://doi.org/10.1145/3373017.3373020

[24] T. Moore and R. Clayton, "Evaluating the Wisdom of Crowds in Assessing Phishing Websites," in *Financial Cryptography and Data Security*, G. Tsudik, Ed. Springer, 2008, pp. 16–30.

[25] L. N. Lota and B. M. Hossain, "A Systematic Literature Review on SMS Spam Detection Techniques," *International Journal of Information Technology and Computer Science (IJITCS)*, vol. 9, no. 7, pp. 42–50, 2017.

[26] S. M. Abdulhamid, M. S. Abd Latiff, H. Chiroma, O. Osho, G. Abdul-Salaam, A. I. Abubakar, and T. Herawan, "A review on mobile sms spam filtering techniques," *IEEE Access*, vol. 5, pp. 15 650–15 666, 2017.

[27] I. Murynets and R. Jover, "Crime scene investigation: SMS spam data analysis," *Internet Measurement Conference*, 2012.

[28] N. Jiang, Y. Jin, A. Skudlark, and Z.-L. Zhang, "Greystar: Fast and accurate detection of SMS spam numbers in large cellular networks using gray phone space," in *22nd USENIX Security Symposium (USENIX Security 13)*. Washington, D.C.: USENIX Association, Aug. 2013, pp. 1–16. [Online]. Available: https://www.usenix.org/conference/usenixsecurity13/technical-sessions/presentation/jiang

[29] ——, "Understanding sms spam in a large cellular network," in *Proceedings of the ACM SIGMETRICS/International Conference on Measurement and Modeling of Computer Systems*, ser. SIGMETRICS '13. New York, NY, USA: Association for Computing Machinery, 2013, pp. 381–382. [Online]. Available: https://doi.org/10.1145/2465529.2465530

[30] B. Reaves, L. Blue, D. Tian, P. Traynor, and K. R. Butler, "Detecting SMS Spam in the Age of Legitimate Bulk Messaging," in *Proceedings of the 9th ACM Conference on Security & Privacy in Wireless and Mobile Networks*. ACM, 2016, pp. 165–170.

[31] C. Oswald, S. E. Simon, and A. Bhattacharya, "SpotSpam: Intention Analysis-driven SMS Spam Detection Using BERT Embeddings," *ACM Transactions on the Web*, 2022.

[32] B. Reaves, N. Scaife, D. Tian, L. Blue, P. Traynor, and K. R. Butler, "Sending out an SMS: Characterizing the Security of the SMS Ecosystem with Public Gateways," in *2016 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2016, pp. 339–356.

[33] J. M. Moreno, S. Matic, N. Vallina-Rodriguez, and J. Tapiador, "Your Code is 0000: An Analysis of the Disposable Phone Numbers Ecosystem," *arXiv.org*, 2023.

[34] Y. Cheng, H. Wang, Z. Zhang, and N. Li, "Characterizing the security threats of disposable phone numbers," in *Frontiers in Cyber Security*, G. Xu, K. Liang, and C. Su, Eds. Singapore: Springer Singapore, 2020, pp. 491–507.

[35] B. Srinivasan, P. Gupta, M. Antonakakis, and M. Ahamad, "Understanding Cross-Channel Abuse with SMS-Spam Support Infrastructure Attribution," *LecEuropean Symposium on Research in Computer Security (ESORICS)*, 2016.

[36] M. Balduzzi, P. Gupta, L. Gu, D. Gao, and M. Ahamad, "Mobipot: Understanding mobile telephony threats with honeycards," in *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security*, ser. ASIA CCS '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 723–734. [Online]. Available: https://doi.org/10.1145/2897845.2897890

[37] S. Tang, X. Mi, Y. Li, X. Wang, and K. Chen, "Clues in Tweets: Twitter-Guided Discovery and Analysis of SMS Spam," *Conference on Computer and Communications Security*, 2016.

[38] Z. Li, W. Wang, C. Wilson, J. Chen, C. Qian, T. Jung, L. Zhang, K. Liu, X. Li, and Y. Liu, "FBS-Radar: Uncovering Fake Base Stations at Scale in the Wild," in *Proceedings of the Network and Distributed System Security Symposium (NDSS 2017)*, San Diego, CA, February 2017.

[39] M. L. Rahman, D. Timko, H. Wali, and A. Neupane, "Users Really Do Respond To Smishing," in *Proceedings of the Thirteenth ACM Conference on Data and Application Security and Privacy*, ser. CODASPY '23. Association for Computing Machinery, 2016, pp. 49–60. [Online]. Available: https://doi.org/10.1145/3577923.3583640

[40] D. Timko and M. L. Rahman, "Commercial anti-smishing tools and their comparative effectiveness against modern threats," in *Proceedings of the 16th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, ser. WiSec '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 1–12. [Online]. Available: https://doi.org/10.1145/3558482.3590173

[41] I. S. Mambina, J. D. Ndibwile, and K. F. Michael, "Classifying Swahili Smishing Attacks for Mobile Money Users: A Machine-Learning Approach," *IEEE Access*, vol. 10, pp. 83 061–83 074, 2022.

[42] M. Liu, Y. Zhang, B. Liu, Z. Li, H. Duan, and D. Sun, "Detecting and Characterizing SMS Spearphishing Attacks," in *Annual Computer Security Applications Conference*. ACM, 2021, pp. 930–943. [Online]. Available: https://dl.acm.org/doi/10.1145/3485832.3488012

[43] Virustotal - virustotal.com. [Online]. Available: https://www.virustotal.com/gui/home/upload

[44] APWG | Unifying The Global Response To Cybercrime. [Online]. Available: https://apwg.org/

[45] Google Safe Browsing. [Online]. Available: https://safebrowsing.google.com

[46] Crt.sh | Certificate Search. [Online]. Available: https://crt.sh/

[47] B. Tejaswi, N. Samarasinghe, S. Pourali, M. Mannan, and A. Youssef, "Leaky kits: The increased risk of data exposure from phishing kits," in *2022 APWG Symposium on Electronic Crime Research (eCrime)*, 2022, pp. 1–13.

[48] cybercdh, "GitHub - cybercdh/kitphishr: A tool designed to hunt for Phishing Kit source code — github.com," https://github.com/cybercdh/kitphishr, 2023, [Accessed 29-07-2023].

[49] K. Tian, S. T. K. Jan, H. Hu, D. Yao, and G. Wang, "Needle in a haystack: Tracking down elite phishing domains in the wild," in *Proceedings of the Internet Measurement Conference 2018*, ser. IMC '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 429–442. [Online]. Available: https://doi.org/10.1145/3278532.3278569

[50] Cybercdh, "Polyglot: Multilingual text (nlp) processing toolkit," 2023. [Online]. Available: https://github.com/aboSamoor/polyglot

[51] S. Prasad, E. Bouma-Sims, A. K. Mylappan, and B. Reaves, "Who's calling? characterizing robocalls through audio and metadata analysis," in *29th USENIX Security Symposium (USENIX Security 20)*. USENIX Association, Aug. 2020, pp. 397–414. [Online]. Available: https://www.usenix.org/conference/usenixsecurity20/presentation/prasad

[52] FCC. (2022) How to identify and avoid package delivery scams. [Online]. Available: https://www.fcc.gov/how-identify-and-avoid-package-delivery-scams

[53] seatgeek, "Thefuzz: Fuzzy string matching in python," 2023. [Online]. Available: https://github.com/seatgeek/thefuzz

[54] Somos, Inc., "2023 telecom fraud symposium – robotext prevention, stir/shaken and more take center stage," 2023, accessed: 2023-12-01. [Online]. Available: https://www.somos.com/insights/2023TelecomFraudSymposium

Appendix A. Tables and Screenshots

TABLE 5: Bulk SMS delivery rate

Carrier	Sent from personal	Sent from A2P
T-Mobile	Received	Received
At&T	Received	Received
At&T	Received	Received
H2O	Received	Received
T-Mobile	Received	Received
Verizon	Received	Received
GoogleFi	Received	Received
T-Mobile	Received	Received
Visible	Received	Received
Visible	Received	Received

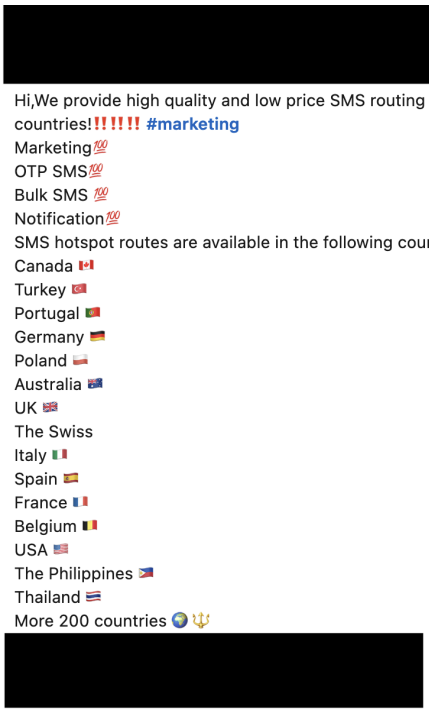

Figure 9: A sample post in LinkedIn advertises bulk SMS services.

TABLE 6: Campaigns we deemed too general to include in the operations graph as they would cluster URLs together that may not be related.

#URL
#URL
#URL: #URL
dear #URL
Hi! #URL
#URL plan
#URL #OTP 2
: #URL
fyavyayvayva #URL
Confirm: #URL
HII: #URL
W0W #URL

TABLE 7: Top 3 languages used in SMS Phishing campaigns.

Language	Campaigns
English	29433
Unknown	3697
Danish	802

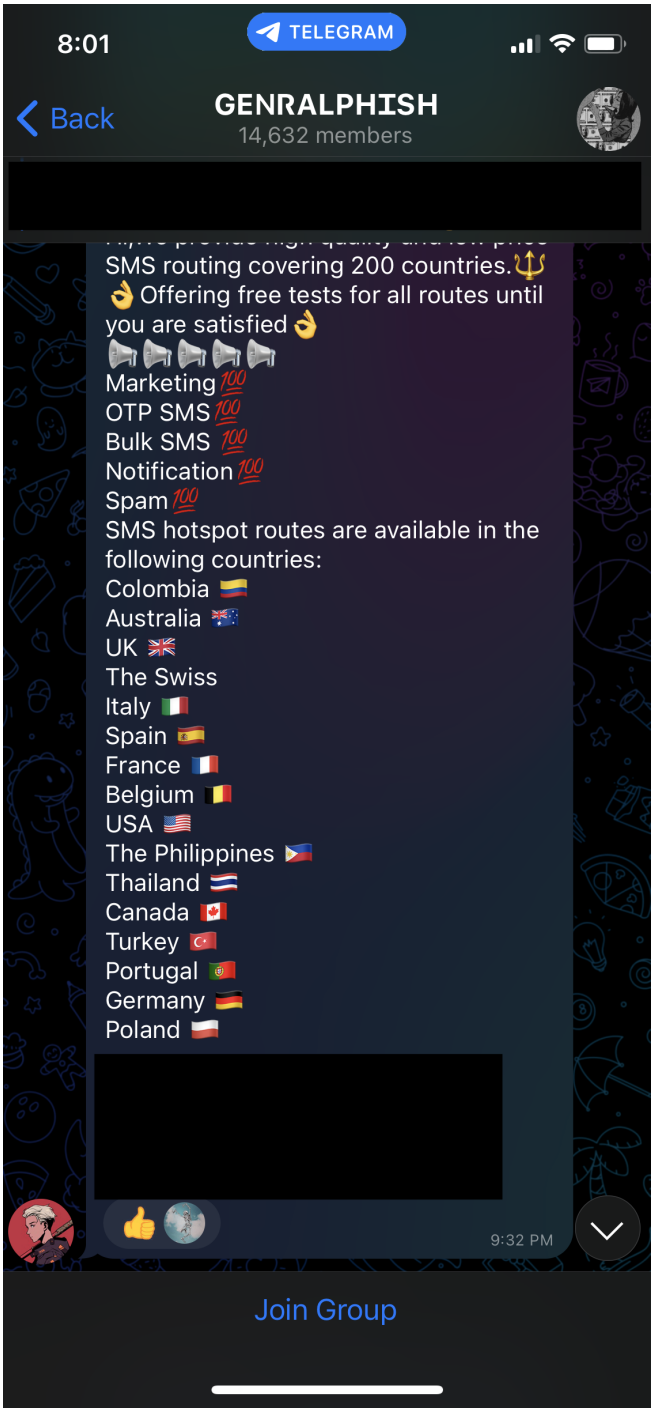

Figure 10: A message sent on a Telegram group chat titled "PHISHING GENERAL"

Appendix B.
Meta-Review

The following meta-review was prepared by the program committee for the 2024 IEEE Symposium on Security and Privacy (S&P) as part of the review process as detailed in the call for papers.

B.1. Summary

This paper presents a measurements study where large-scale SMS messages are collected over a long period, looking for malicious messages within and extracting particularly phishing links from these messages. The data was obtained by crawling SMS gateways. This approach enabled authors to make conclusions on the state of phishing campaigns.

B.2. Scientific Contributions

- Provides a New Data Set For Public Use
- Independent Confirmation of Important Results with Limited Prior Research
- Addresses a Long-Known Issue

B.3. Reasons for Acceptance

1) The paper provides a larger SMS messages dataset, collected over a longer time period and scattered over wider geographic regions. The dataset is rich, and the authors' analysis methodology appears sound and valid.
2) The paper sufficiently justifies that this remains a problem in dire need of a solution. The methodology is clear, and the limitations are clearly spelled out and justified. The findings are presented clearly and concisely, which is useful for those looking at this work when trying to address the problems of SMS phishing.
3) The paper addressed a long-known issue regarding SMS phishing attacks by highlighting several observations within the SMS phishing ecosystem.

B.4. Noteworthy Concerns

1) While the findings are informative, the paper provides limited discussion on practical implications or recommendations for mitigating SMS phishing.