Poster: Looking at the landscape of SMS Phishing

Aleksandr Nahapetyan
NC State University
anahape@ncsu.edu

Kevin Childs
NC State University
krchilds@ncsu.edu

Sathvik Prasad
NC State University
snprasad@ncsu.edu

Bradley Reaves
NC State University
bgreaves@ncsu.edu

Alexandros Kapravelos
NC State University
akaprav@ncsu.edu

***Abstract*—SMS Phishing or SMISHING attacks have grown in popularity over the last few years. To launch a web-based SMS Phishing campaign, malicious actors need web infrastructure (hosting, domain, SSL certificates), a phishing page (usually generated by a PhishingKit), and means of delivering this to their targets. We leverage public SMS Gateways to study phishing campaigns, identify their infrastructure and look at temporal trends. Additionally, we scrape LinkedIn posts to highlight avenues a malicious actor can use to deliver phishing messages that do not include conventional A2P providers. By submitting URLs from the public gateway messages, we find 23,000 messages containing a malicious link; we find that Amazon provides the most (45.9%) of the hosting services for these URLs. We isolated 188 SMS campaigns and manually classified 116 of them as phishing. We find over 40,00 posts mentioning bulk sms services on LinkedIn. Overall, we provide a bird's eye view of the SMS Phishing ecosystem, looking at existing trends, commonly used web infrastructure, and less regulated delivery mechanisms.**

I. Extended Abstract

The Federal Communications Commission reports a rise in complaints for "unwanted text messages" from 5,700 in 2019 to 14,000 in 2020 and 15,300 in 2021. [1] The academic study of SMS phishing requires an existing pipeline for users to report messages to researchers or to scrape public forums like Twitter for posted screenshots [3] of spam content. We opt to scrape Public SMS Gateways, external entities within the SMS ecosystem that allow users to receive text messages to a select list of numbers via a web page [2]. Scraping these pages allows you to monitor the traffic to all of these phone numbers, effectively using them as honey pots. To filter the malicious messages, the URL from the messages to VirusTotal. Any message with a URL flagged by at least 5 VirusTotal vendors, GoogleSafebrowsing or was on a phishing blocklist from OpenPhish, gitlab.com/malware-filter, or Ublock origin we tagged as malicious.

From these messages, we isolated campaigns via these steps:

- We remove all URLs, phone numbers, emails, and one-time codes from these messages using regular expressions
- We group these patterns if they have a similarity ratio of greater than 95. We use python's thefuzz library to compute this ratio, which intern computes the Levenshtein distance between two given texts. We note that TheFuzz uses words at the basic tokens, not individual characters.
- We manually go through all 188 campaigns, tagging them with relative themes as we observe them and marking out the false positives based on the message. We mark out any message as not phishing if they are in a language we do not speak, is an advertisement for casinos or loan services, or links to a legitimate service.

We manually reviewed the phishing campaigns and tagged them with the following tags.

- **General**: Messages that lacked intent, things like "Hello user" or "click here." Or campaigns that used multiple scare tactics to try and get the user to click on a link.
- **Unknown**: Non-English messages that we could not classify.
- **Account**: Messages falsifying account login notices or changes.
- **Financial**: Messages that mention bank login attempts or false transactions.
- **Mail Delivery**: Messages that were focused on miss-delivered packages.
- **Casino** Non-Phishing messages were ads for an online Casino.
- **Surveys**: Messages asking for a survey response.
- **Medical**: Messages that mention medical tests or appointments.
- **Apple**: Messages about Apple accounts (iCloud, FindMy iPhone, Apple Card, etc.)
- **Download**: Messages that tried to get you to download apps/files.
- **Sales**: Messages that were trying to encourage buying or selling property.
- **Loans**: Messages that mentioned loan applications.
- **Betting**: Messages that mentioned betting in a non-casino setting.
- **Taxes**: Messages that mention the IRS and Taxes.

We present the breakdown of the campaigns in Figure 1.

In addition, crawl all maliciously tagged URLs and identify their registrars, SSL Certificate issuers, and the Anonymous System (AS) delivering their content. We report a summary of the top registrars, SSL Certificate issuers, and AS Orgs in Table 1.

We look at LinkedIn posts where individuals advertise bulk SMS services and, in some cases, hint at not caring about

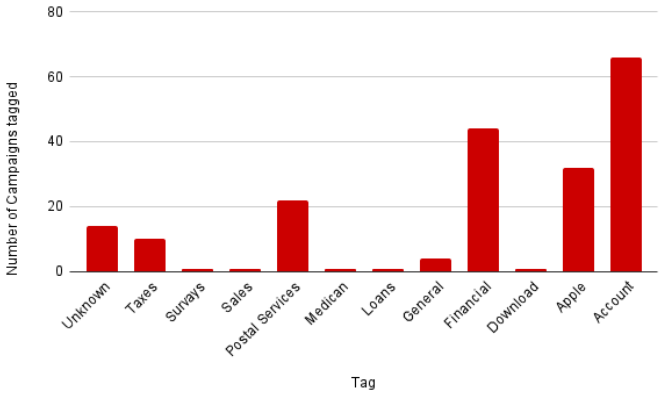


Fig. 1: breakdown of the 188 campaigns by theme

Top SSL Provider	Let's Encrypt	736 (67.4)
Top AS Org	Amazon Inc	953 (45.9%)
Top Registrar[1]	GoDaddy LLC	65 (10.3%)

TABLE I: Figure 1: Summary of the Top SSL provider, Domain registrar, and antonymous system that malicious URLs used.

the content—combined with PhishingKits that we gather by fuzzing the malicious URLs; we look at all three steps of SMS phishing campaign deployment.

REFERENCES

[1] FCC Warns Consumers of Rising Threat of Scam Robotexts.

[2] Bradley Reaves, Logan Blue, Dave Tian, Patrick Traynor, and Kevin R.B. 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*, WiSec '16, pages 165–170. Association for Computing Machinery.

[3] Siyuan Tang, Xianghang Mi, Ying Li, XiaoFeng Wang, and Kai Chen. Clues in Tweets: Twitter-Guided Discovery and Analysis of SMS Spam. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, CCS '22, pages 2751–2764. Association for Computing Machinery.

[1] 17.5% of the URLs had an unknown registrar

We leverage data from **public SMS gateways** and **LinkedIn posts** to study the mobile phishing ecosystem. We identify over **116 phishing campaigns** and **40642 LinkedIn posts** advertising bulk messaging services

Looking at the landscape of SMS Phishing

Campaigns, Infrastructure, and ecosystem

Aleksandr Nahapetyan, Kevin Childs, Sathvik Prasad, Alexandros Kapravelos, Brad Reaves

Introduction

- SMS Phishing or **SMISHING** is still a widespread problem. Despite next-generation SMS firewalls, FCC reported in 2022 "[a rise from complaints from] **5,700 in 2019**, to **14,000 in 2020**, **15,300 in 2021**, and **8,500 through June 30, 2022**."
- **Public SMS Gateways usually come in the form of a webpage full of phone numbers that can bypass "Proof of humanity" verifications** in the form of OTPs. *Reaves at el.* has previously used these gateways to study malicious actors in the SMS landscape as malicious actors will spam these numbers thinking they are used by humans.
- The building blocks of web-based Phishing campaigns are **web-infrastructure, bulk SMS services, and a phishing page** (usually created by a PhishingKit.)

Methodology

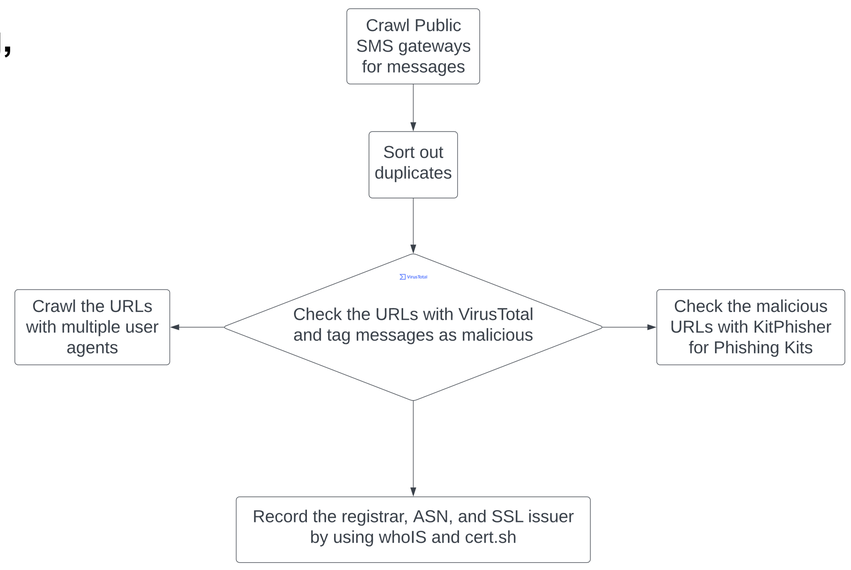
- We use scrapy to scrape multiple public SMS gateways, filtering out duplicate messages and storing them for analysis.
- We rely on *VirusTotal* to identify malicious content.
- We use python's *TheFuzz* library **with a constant threshold (95) for the** similarity ratio to group similar SMS messages.
- We manually tag and classify the identified campaigns. After going through all of the campaigns, we tags them as **General, Unknown, Account, Financial, Postal Services, Casino, Surveys, Medical, Apple, Download, Sales, Loans, Betting, or Taxes**
- We use a *puppeteer-powered* LinkedIn Crawler to gather data on posts that advertise bulk SMS service.

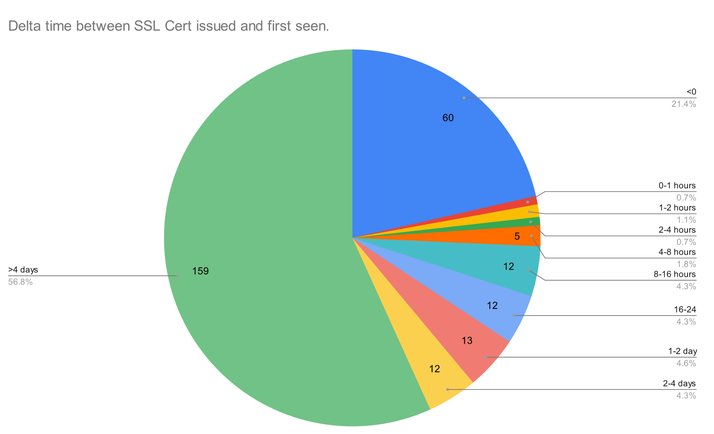
Results

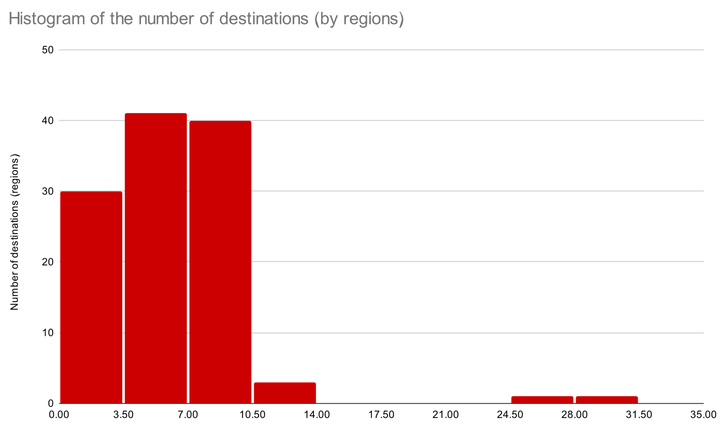
- We classified **116 (61.7%) of the campaigns as phishing.**
- **953 (45.9%) of the malicious URLs** were hosted on an **AS owned by Amazon**, according to its whois data.
- **Most common SLD (33 URLs / 2.2%) was web.app,** showing that these actors do not rely on accessible hosting providers.
- KitPhisher helped us extract **two phishing kits.**
- We **found 40642 LinkedIn posts** that advertised bulk SMS services. **With 26.7% of the posts using WhatsApp** as their points of contact, **7.7% using Skype**, and **4.3% using Telegram**.
- Unsurprisingly most common SSL certificate issuer was **Let's Encrypt at 67.4% (736),** followed by **Greater Manchesterat at 5.2% (57).**

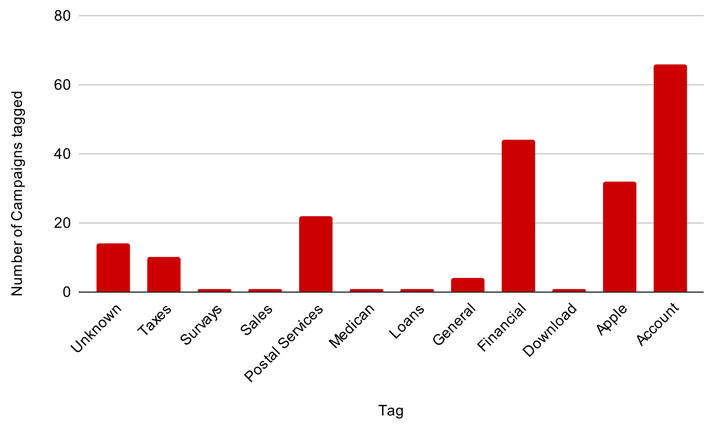
Takeaways

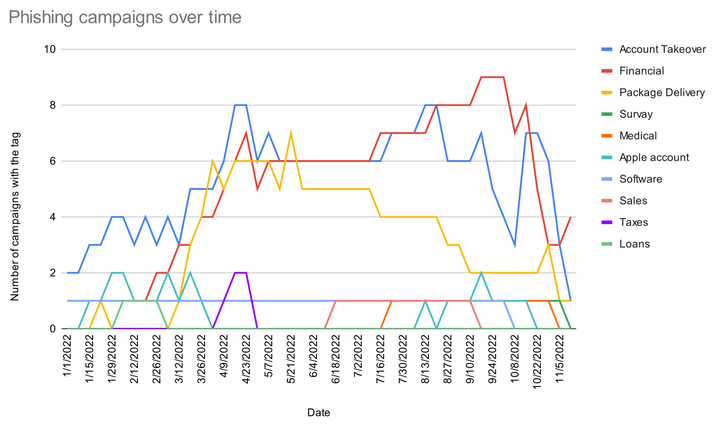
Our findings give us insight into the parts of the SMS ecosystem, common points used in their infrastructure, currently available defenses, and their effectiveness.











NC STATE UNIVERSITY

WSPR