Lies in the Air: Characterizing Fake-base-station Spam Ecosystem in China

Yiming Zhang[1,2], Baojun Liu[1,2,*], Chaoyi Lu[1,2], Zhou Li[3], Haixin Duan[1,2,5,*], Shuang Hao[4], Mingxuan Liu[1,2], Ying Liu[1,2,*], Dong Wang[6] and Qiang Li[6]
[1] Beijing National Research Center for Information Science and Technology [2] Tsinghua University
[3] University of California, Irvine [4] University of Texas at Dallas [5] QI-ANXIN Group [6] 360 Mobile Safe

ABSTRACT

Fake base station (FBS) has been exploited by criminals to attack mobile users by spamming fraudulent messages for over a decade. Despite that prior work has proposed several techniques to mitigate this issue, FBS spam is still a long-standing challenging issue in some countries, such as China, and causes billions of dollars of financial loss every year. Therefore, understanding and exploring the thematic strategies in the FBS spam ecosystem at a large scale would improve the defense mechanisms.

In this paper, we present the first large-scale characterization of FBS spam ecosystem by collecting three-month real-world FBS detection results. First, at "macro-level", we uncover the characteristics of FBS spammers, including their business categories, temporal patterns and spatial patterns. Second, at "micro-level", we investigate how FBS ecosystem is organized and how fraudulent messages are constructed by campaigns to trap users and evade detection. Collectively, the results expand our understanding of the FBS spam ecosystem and provide new insights into improved mitigation mechanisms for the security community.

CCS CONCEPTS

• **Information systems** → *Spam detection*; • **Security and privacy** → *Mobile and wireless security*.

KEYWORDS

Fake Base Station; Spam Ecosystem; Spam campaigns

ACM Reference Format:

Yiming Zhang, Baojun Liu, Chaoyi Lu, Zhou Li, Haixin Duan, Shuang Hao, Mingxuan Liu, Ying Liu, Dong Wang, and Qiang Li. 2020. Lies in the Air: Characterizing Fake-base-station Spam Ecosystem in China. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security (CCS'20), Nov. 9–13, 2020, Virtual Event, USA.* ACM, NY, NY, USA, 14 pages. https://doi.org/10.1145/3372297.3417257

1 INTRODUCTION

For backward-compatibility, so far still many cellular providers and mobile device vendors are supporting the hybrid of 2G/3G/4G protocols. Among those protocols, GSM (2G) is known to be insecure as the authentication vulnerabilities allows an adversary to send spam messages under arbitrary phone numbers [15, 50]. While a user can choose to stay in a safer 3G/4G environment, by using a device called *Fake Base Station (FBS)*, an adversary can still force the user's device to downgrade their active communication mode to 2G and exploit the authentication vulnerabilities [13]. Due to its sender spoofing ability, FBS is extensively used for short text message spamming and advertising illegal and fraudulent businesses, which poses serious threats to mobile users.

Around the world, FBS has become a prominent problem in countries like the US, the UK, China and India [7, 8, 10, 12]. Public reports show that FBS spammers can operate their business at small costs while causing considerable losses to victims. As an example, an FBS device can be purchased at only $700 [11], which is able to generate 60K to 150K text messages and bring in $1,400 revenue per day for a spammer [14]. Such big profit margin stimulates a thriving underground ecosystem filled by FBS spammers. In 2015, Chinese mobile users received over *5.7 billion* unsolicited messages from FBS devices [15], and lost a staggering *$3.13 billion* in 590K reported incidents [18].

To counter FBS spams, prominent efforts have been made by the research community, cellular industries and governments. Government agencies track the positions of FBS devices [9] and arrest their operators. The research community has been focusing on detecting FBS spams based on their unique cellular characteristics [5, 24, 30, 63, 64]. Some approaches have been deployed by security companies [50] to protect vulnerable mobile users. However, despite those long-lasting efforts, FBS still does not die out. According to reports from leading security companies, in Mar. 2016 alone, Chinese users received more than 110 million FBS messages [16], and over 150 million end users still victimized at least one FBS message in the first half of 2017 [19].

Motivation. While the techniques of FBS are well studied and an array of technical approaches are available as defense [24, 26, 30, 32, 38, 44, 56, 63], we still lack deep insights into the *ecosystem* powered by FBS. Similar to online advertising, the FBS ecosystem has hierarchies: an FBS spammer takes orders from a business owner (e.g., fake ID-card seller) to distribute the promotional messages, and a website (or social media) has to be operated to interact with users who follow the contact information embedded in the messages. All the prior works focus on the front-end, i.e., FBS spammer, leaving the other layers unexplored. Even just for the front-end, our knowledge is incomplete and many questions remain open, e.g., *is there a preferred time or location for FBS spammers to send messages? how do spammers entice users to follow the contact information? how do they bypass mitigation?* Therefore, the primary goal of this research work is to comprehensively explore the FBS

CCS '20, November 9–13, 2020, Virtual Event, USA

ACM ISBN 978-1-4503-7089-9/20/11…$15.00
https://doi.org/10.1145/3372297.3417257

ecosystem, in hopes that weaknesses within their hierarchies can be identified. In fact, we are strongly encouraged by prior works characterizing the ecosystems of email [45–47, 55] and social network spamming [36, 39, 59, 60] as those works have inspired more effective mitigation.

More specifically, we take a data-driven approach to analyze the FBS ecosystem. To begin with, we collect *279K real-world FBS messages spanning three months* from a security APP widely used by mobile users in China. On top of this large-scale FBS message dataset, we are able to "recover the crime scenes" and infer the actors and purposes of FBS. Specifically, to learn the business behind, we develop a machine-learning classifier that can accurately label an FBS message into 14 categories. To attribute an FBS message to a campaign (or spammer group), we build another classifier by exploiting the contact information embedded in the message content. Several auxiliary datasets, such as passive DNS, WHOIS and domain blacklists are used to augment the FBS dataset and characterize the supporting web infrastructure of FBS.

Main Findings. Putting together, we uncover the "macro-level" and "micro-level" characteristics of FBS spam ecosystem across a country, including spam categories, temporal patterns, spatial patterns, and evasion strategies. Here we highlight a few major findings. First, FBS spammers prefer to serve different businesses at different cities or places, and most of the messages (over 75%) are associated with illegal or fraud businesses (Section 5). Besides, we find FBS spammers are largely active near main roads and highly-populated regions (e.g., residential areas) to increase their chances of reaching victims. Even after being detected and moved to "Spam Inbox", domains embedded in FBS messages still receive considerable traffic from victims (e.g., 12.6% domains queried by over 5K times). Furthermore, we identify 7,884 FBS campaigns via the contact information of spammers. We find interesting resource sharing (e.g., message templates and contacts) among different groups, even across business categories, suggesting the FBS ecosystem has become hierarchical, where tasks are separated and undertaken by different actors (Section 6). As FBS spammers can spoof sender numbers (Section 2), a few reputable parties are spoofed at high frequencies: e.g., 8.4% FBS messages impersonate the Industrial and Commercial Bank of China (ICBC). To receive money from victims, bank card numbers are embedded in many messages, and we find they are registered at banks with hidden policy flaws that prohibit effective provenance.

Contributions. The main contributions of the paper are outlined as follows.

- Through the first large-scale characterization of FBS ecosystem, we identified FBS spam campaigns, uncovered new insights about how FBS ecosystem is organized and how spammers behave.
- We released our labeled ground-truth dataset (14K FBS messages and their anonymized meta-data) at [22] to help the research community develop better solutions against FBS.

Scope of our study. This work focuses on spam messages distributed by FBS only. Distributing FBS messages requires the actors physically move into the victims' neighborhood, which is different from other channels like public SMS gateways [57]. As such, FBS has a unique ecosystem, which is worth research on its own. We focus on FBS in China where the problem is much more severe than other countries [50]. This research also sheds lights into how underground business is operated in China.

2 BACKGROUND

In this section, we first describe how a FBS device sends spam text messages to mobile devices around and present some examples. Then, we elaborate state-of-the-art approaches to detect FBS and the detection method of our industrial partner which builds our real-world spam message dataset.

2.1 Sending Spam Text Messages with FBS

Base stations (BS) are the basic infrastructure of cellular networks which mobile devices need to connect to reach telephone network. However, a vulnerability of the GSM (2G) protocol allows the creation of FBSes which are not authorized. Attackers could launch attacks such as user identity theft [33] and service hijacking [37] towards victims. In this work, we focus on the ability of FBSes to send spam messages to their connected end-user devices, *from arbitrary phone numbers*, which is a popular approach for spammers to send spam text messages.

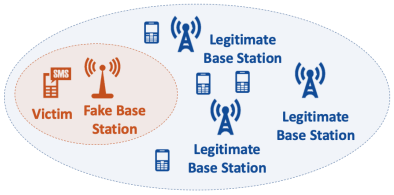


Figure 1: FBS Under GSM (2G) Protocol.

Figure 1 represents the scenario of FBS under 2G. To send spam messages to nearby end devices, an FBS first starts broadcasting its system information (e.g., its BS-ID). As the signal strength of the FBS is stronger than other BSes, nearby mobile devices start connecting to the FBS, by sending Location Updating Requests and their system information (e.g., IMSI and IMEI). On receiving the requests, the FBS sends accept messages and could then send spam messages from any spoofed phone number to the connected devices. Finally, when an FBS no longer wants to be connected, it simply lowers or shuts down the signal, switching its connected devices to other BSes in service.

It should be noted that FBS is not a problem for 2G only. According to previous studies [13], the FBS attackers are also able to force nearby 3G/4G-compatible mobile devices to downgrade to the GSM (2G) mode by sending jamming signals to legitimate base stations. Besides, although SMS Gateway is another channel that enables bulk SMS services, the sending operations via Gateways are subject to many restrictions, e.g., identity verification and content legality check. Therefore, FBSes are still considered as important channels for (especially malicious and illegal) spamming.

Although FBS is not a new threat, limited insights of the corresponding ecosystem have beed discovered. Previous works focus more on algorithms to *detect* spam messages, typically based on content analysis or senders' behaviors [26, 32, 38, 44, 56]. Only a few studies attempt to understand the security and privacy problems in SMS channels, such as malicious behaviors via SMS gateway [57] and communication patterns of SMS spam traffic [54]. However, FBS is a completely different spamming channel. Limited

by previous collected datasets, the *FBS spam ecosystem* still lacks a comprehensive study.

In this research work, we define *spam text messages* as unsolicited or undesired text messages sent from FBS devices. We present two examples of spam messages below.

Advertisement Messages. "*Provide invoices to help you pay less taxes. Contact cellphone 135****2508 or WeChat a135***710.*" – Sent from 10****8989.

Fraud Messages. "*Your phone number has been selected by Satellite TV as a lucky audience. Visit www.xxoosp.com and get your prize!*" – Sent from +8529***7281.

2.2 Detecting FBS

Previous study have uncovered two prerequisites for an FBS to interrupt connections between end devices and legitimate BSes [50]. First, an FBS should broadcast its information in a *significantly higher signal strength* than nearby legitimate BSes to force the nearby end-devices to switch from their already-connected BSes. Second, while the BS-ID of an FBS is correct in syntax, it should be *significantly different* from those of nearby legitimate BSes. Although end devices are not sensitive to small BS-ID changes for energy-saving reasons, a significant change of BS-ID indicates that the device has entered a new cellular coverage area, and will trigger a switch of base station. These two features have been used in a previous study to detect FBS at scale [50], with a promising result of 98% precision. The proposed detection system also inspires tracking techniques of FBS devices recently [64].

This method is also implemented by our industrial partner, which runs a popular mobile security application in China. Users of this application would be prompted to gather the required information for FBS detection and the function can only be activated after the agreement. When a short message is received, the connected BS ID and strength signal would be examined for FBS detection module. This data collection process is done by our industrial partner and we do not consider it as our contribution in this work. The limitations of our datasets, such as the potential false positives and the coverage, will be discussed later (see Section 3.3).

3 DATA COLLECTION

So far, only a few research projects provided datasets of spam messages [56, 57] but none of them contains FBS data. To fill this gap, we collect three-month FBS messages sent by real-world spammers in China, which are detected by a mobile security application. In addition, we also collected several large-scale auxiliary datasets (e.g., passive DNS and blacklists) for in-depth measurement of the FBS ecosystem. Below we elaborate our datasets in detail.

3.1 FBS Detection Logs

Our primary data comes from the detection logs generated by a mobile security application, 360 Mobile Guard [23]. Below we describe its detection workflow and the ethics of data collection.

Workflow of the FBS detector. The mobile security app monitors the messages when permissions (e.g., reading messages, connecting network) are granted by users. On receiving a short text message, the app detects whether it is sent from FBSes, based on the techniques illustrated in Section 2. If one message is classified as being sent from FBSes, the app would move it into `Spam Inbox` (instead of deleting it), and prompt the user with a notification. The message contents in Spam Inbox together with its *anonymized meta-data* are uploaded to cloud servers for further analysis. Each uploaded record contains timestamp, sender's phone number (which is often spoofed), hashed IMEI (International Mobile Equipment Identity) and IMSI (International Mobile Subscriber Identification), message content, and the recent history of BS-IDs that the client has connected to. Besides, to correct false alarms, the software offers an option for mobile users to manually recover a message from Spam Inbox to the normal Inbox. In total, our origin dataset contains 279,017 FBS detection logs from Dec. 1, 2018 to Mar. 7, 2019 (spanning 97 days).

Ethics. The major ethical concern would be about the data collection mechanism of the mobile security app. In fact, users' privacy is taken careful treatment by the app's design. Firstly, the software provides a consent explaining the corresponding functions, data collection details and privacy strategies to users [40]. With the user's authorization, the data collecting process strictly follows that agreement and is supervised by the privacy and legal committee of the industrial partner. Secondly, a user may disable the FBS detection module after installation. All functions of the module relevant to user data, like uploading spam messages, can be enabled/disabled at prominent places in the app UI. Users receive a visible notification when a spam message is detected. Besides, the meta-data of uploaded messages is also carefully anonymized according to the best practices (e.g., the device IDs are hashed and user phone numbers are removed), so that no personal information is revealed. The servers encrypted the anonymized detection logs, and only accessible to the developers and researchers involved in this project. We released our ground-truth dataset at [22] to help research in FBS detection. To avoid privacy risks, the dataset only contains the 14K messages manually labeled by us, and all personally identifiable information (PII) in the contents have been anonymized.

3.2 Auxiliary Datasets

Passive DNS (PDNS). PDNS data contains DNS transactions logged by resolvers, which are extensively used by security researches [51–53]. We use this data to investigate the query volume and lifetime of suspicious domains embedded in FBS messages. While popular PDNS like DNSDB of Farsight Security [35] has good coverage of DNS resolvers, we found its coverage in China mainland is quite limited. Instead, we leverage a PDNS dataset provided by 360 Passive DNS Project [28], whose resolvers are mainly located in China. Each PDNS record provided by [28] contains *daily* aggregate statistics of the DNS queries for a given domain since 2014.

Domain Blacklist. To determine whether a suspicious domain is truly malicious, we leverage three URL blacklists aggregators: VirusTotal, Qihoo 360 and Baidu. If a suspicious domain is alarmed by at least one blacklist under any aggregator, we consider it malicious. Finally, we detect 3,197 malicious domain names from the FBS messages.

WHOIS Database. WHOIS helps us understand the registration behaviors of suspicious domains. We attempt to retrieve WHOIS records for all suspicious domains from a WHOIS database built by our industrial partner, which has already been used in published works [27, 51, 52]. Considering its haphazard format, our industrial partner has done extra parsing of WHOIS data besides processing

it by python-whois [34] to make it more credible. Overall, 94.2% suspicious domains have associated records in the database.

Bankcard Information. We find a lot of FBS messages contain bankcard numbers, which are used by fraudsters to receive money from the victims. In this study, we use a commercial online tool provided by ShowAPI [3] to collect bankcard information of the fraudsters for the purpose of attribution. The collected information includes bank name, bank location and account types (credit/debit card). Note that only fraudsters' bankcard information are retrieved.

BS geo-location Database. Our industrial partners also provide us with a geographic database of BSes in China. By querying a valid BS-ID code, the database returns the province, city, street, and the detailed latitude and longitude information of the BS. It helps us analyze the geographical distribution of FBS spammers.

3.3 Limitation

Although we try to make this study as comprehensive as possible, there are still some limitations in this study and we would discuss them before presenting our measurement results.

Geographic Bias. As other data-driven studies, the collected dataset may have a geographical bias due to user distribution. However, our industry partner has a large user base in China and the security app we utilized has over one hundred million monthly active users, as reported by public reports of market share. The collected FBS messages dataset covers 33 provinces and 332 cities (almost all) in China so it could be considered representative enough for a country-level study.

False Positive. We have little ground-truth about FBS spam campaigns so our results cannot be well-evaluated. To eliminate the impact of False Positives, we build a content-based multi-classifier and only limited the scope of analysis to messages classified as Illegal Promotion, Fraud and Advertisement (see Section4). The selected categories, accounting for 99.3% of all messages reported, are highly involved in FBS spam business.

To the best of our knowledge, our dataset is the largest one of real-world FBS spam text messages so far. Previous works [49,50] collect messages from SMS Gateway, which is a different spamming channel; [42] manually labels 200K suspicious messages while only 0.16% of which are identified as FBS spam. Collectively, we believe that we have taken the first step towards understanding the fraudulent activities of the FBS spam ecosystem.

4 CATEGORIZING SPAM MESSAGES

While the mobile security app can detect FBS message with good accuracy, the detection result does not give much detail like the business and fraud campaign behind each message. The information is crucial for understanding the ecosystem of FBS spam. In addition, categorizing FBS messages can improve the effectiveness of the FBS detector in protecting mobile users. For example, we find users prefer to restore a certain category of FBS messages, like phishing, from "Spam Inbox" which is actually a fraud, however. By learning the FBS category, FBS detector can provide more specific warning to prevent the user from falling in trap (Section 7).

As far as we know, there is no existed public spam message dataset with the labels of category and campaign. For example, the dataset from [56] has 400K messages from SMS gateway but each message is only labeled as spam (8.2% of messages) or not. Therefore, we fill this gap by developing a novel spam message classifer trained on a self-labeled ground-truth dataset. We also cluster the messages into spam campaigns for better observations of FBS spamming ecosystem. The data processing workflow is presented in Figure 2 and the details are explained below.

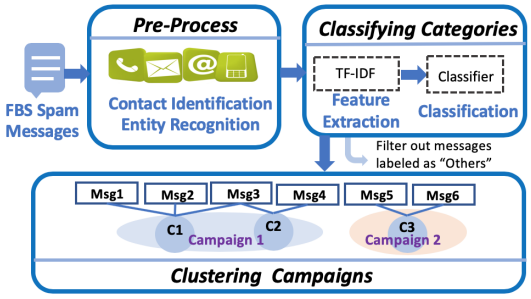


Figure 2: Overview of the data processing flow.

4.1 Classifying FBS Business Categories

Ground truth collection. Our first goal is to classify FBS messages into business categories. To this end, a ground-truth dataset is needed and we address this issue by manually labeling a subset of the FBS messages (5% random samples) provided by our industrial partner. In particular, we ask three security researchers to label this dataset independently, and a message is included in our ground-truth dataset only if at least two of the researchers give the same label. Table 1 presents the categorical labels of spam text messages and their corresponding distribution in the labeled dataset. Note that, the activities in Illegal Promotion are also a subset of Advertisements, but illegal according to Chinese law. In the end, our ground-truth dataset has 14,077 messages under 4 main categories and 14 sub-categories. We ensure that each sub-category contains at least 100 short messages except "Others" to reduce the bias when training the classifier. "Others" type in the last row of Table 1 refers to spam messages that could not fit any of the above categories.

Table 1: Categories of FBS messages

Category	Sub-category	# Labeled	% Labeled
Illegal Promotion (IL)	Gambling	1,681	11.94%
	Fake ID and invoice	495	3.52%
	Political propaganda	290	2.06%
	Escort service	111	0.79%
Fraud (FR)	Phishing (Bank)	1,900	13.50%
	Phishing (Others)	297	2.11%
	Financial fraud	164	1.17%
	Others	711	5.05%
Advertisement (AD)	Retail	4,669	33.17%
	Loan service	283	2.01%
	Real estate	322	2.29%
	Network service	2,127	15.11%
	Others	926	6.58%
Others	-	101	0.72%

* Items in Illegal Promotion and Fraud are illegal according to the Chinese law.

Pre-processing. In the process of content-based classification, the first challenge we encounter is that the length of FBS text messages is usually short (e.g., 40 Chinese characters on average), which lacks sufficient language context for semantic analysis. Worse, the messages usually contain lengthy and variant contact information (e.g.,

Table 2: Eight common types of contact information in spam text messages and pattern examples.

Type		# Contact	# Msg	Pattern Examples
Phone	Cellphone	2,633	84,669	11 digits, starting with ISP code
	Landline	553	22,723	Region code + 7-8 digits
	Toll-free phone	60	328	10 digits, starting with "400"
Web	Domain	3,834	90,558	{[}SLD{]}.{[}TLD{]}
	URL	575	1,696	http{[}s{]}?:// {[}domain{]}/{[}parameters{]}
Social media	Wechat	550	37,364	"Wechat" + name
	QQ	71	425	"qq" + no more than 10 digits
Bank account		40	127	16-20 digits, starting with bank code
Total		**8,316**	**218,543**	-

cellphone number and bank account), which provides virtually no help to analyzing the message semantics but has a negative impact on the classification results [38, 56]. In fact, we find that 78.3% FBS messages contain at least one contact identifier, suggesting leaving contact information in the messages is important to engage victims for the fraudulent activities afterward.

In the end, we address these challenges by identifying the contacts and replacing it with a normalized entity before classifying the message. While existing tools, like Hanlp [43] (an open-source natural-language processing tool) can identify entities, e.g., names and places, they do not work well on contacts that are a combination of digits and letters. For example, it recognizes a phone number as a sequence of "digit". By looking into the samples, we find all contacts can be categorized under eight types, and all of them have unique patterns, as shown in Table 2. Therefore, we build regular expressions according the 8 patterns to find the contacts from messages and then replace them with normalized entities (e.g., each cellphone number is replaced with a constant string "PHONE"). After that, we use Hanlp to segment words and label other identities. If a message is blank or in wrong format, it will be removed before the next step. In the end, 244,240 (87.5%) valid messages and their associated logs are used for the follow-up analysis.

Evaluation. For each normalized spam message, we choose unigram and bi-gram as word terms and calculate their TF-IDF values [25] to model their frequencies, which are used as features of the classifier. Here we do not choose word-embedding feature generation methods [48] because they need high-quality and large-scale datasets for pre-training. Leveraging our ground-truth dataset (14K messages), we train four multi-classifiers with scikit-learn [29]: Support Vector Machine (SVM), Naive Bayes (NB), Logistic Regression (LR) and Random Forest (RF).

Table 3 shows the performance of classifiers under 5-fold cross validation. We find that SVM performs the best (average F1-score 96.87%) so it is used to categorize all other messages (244K) .

Table 3: Five-fold cross validation results

Classifier	Precision	Recall	F1-score
Support Vector Machine	96.90%	96.96%	96.87%
Naive Bayes	95.23%	95.16%	95.06%
Logistic Regression	94.90%	94.64%	94.32%
Random Forest	75.63%	71.88%	72.89%

To further understand the performance of SVM classifer, e.g., which category is more likely to be mis-classified, we generate a confusion matrix from the classification results, as shown in Figure 11 in Appendix A. Each row of the matrix represents the spam messages in one ground-truth category and columns represent their predicted results. It turns out that the classification is accurate for most categories (i.e., all above 85% except for "Others" messages). The main reason for mis-classification is the ambiguous meaning of the message content. For example, some messages only contain few categorical keywords (e.g., "Emergency, please contact me.") besides the contact information of the spammer. It is difficult to learn whether they are about legal advertisement, financial fraud, other new types of spam or just the false positive of FBS detection on top of such limited information. However, based on validation results, those ambiguous cases only have remarkable impact on "Other" type, which accounts for less than 0.7% of the total data set (101 of 14,077 labeled messages). Subsequent measurement and analysis of this work were conducted based on the other 13 sub-categories of spam under "Illegal Promotion", "Fraud" and "Advertisement".

4.2 Clustering FBS Campaigns

In addition to classifying the business category, we develop another classifier to identify the *campaigns* (the groups behind the spam activities) of FBS messages. Attributing the groups can help us get a better understanding of their strategies.

Previous studies on email and twitter spam cluster the spams into campaigns based on URLs [36, 39, 55], text similarity [46, 56], and spamming tools [59]. However, most of those methods are not suitable for our FBS dataset. For instance, URL-based detection only covers less than 1% (1,696 messages, as shown in Table 2) of the whole dataset and text similarity is not an accurate indicator due to the short length of text messages.

Our method of spam campaign identification mainly leverages spam's *contact information*, which is broadly presented in the FBS messages of our ground-truth dataset. Similar core ideas were also used in [31] to group one-click-fraud miscreants. During preprocessing, we extract and normalize eight types of contacts from FBS messages (see Table 2). Then, we identify spam campaigns based on the two assumptions below:

- **Assumption #1: Spam messages sharing the same contact information belong to the same spam campaign.** Normally, different spammers have no incentive to share contact information (e.g., the same phone number or bank account) because they are also competitors.
- **Assumption #2: All contacts in a spam message belong to the same spam campaign.** Spammers typically operate multiple contacts, in case some contacts become unreachable (e.g., being blacklisted). We find multiple contacts can be included in a single message (e.g., the first message in Section 2.1).

The campaign classifier follows a procedure similar to hierarchical clustering. It starts by creating subsets C_i for all messages with contact i. Then it iteratively compares each pair of subsets and merges two subsets C_i and C_j if they have any overlap. The process is repeated until no more subsets can be merged, and each subset is regarded as one spam campaign. For instance, according

to assumption 1, message 1 – 3 in Figure 2 share contact c_1 belong to the same cluster C_1. And it would be merged with C_2 according to assumption 2 (contact c_1 and c_2 appear in the same message 3 belong to the same campaign). In the end, we discover 7,884 spam campaigns and we present the analysis on them in Section 6.

Discussion. Due to the lack of ground truth of spammers' real identities, our clustering result could be inaccurate to some extent. But still, we believe this is a best-effort approach given that getting contact accounts involves manual interaction with the account providers. Among the techniques proposed by previous works, clustering based on text similarity would have less accuracy, as there are many *spam message templates* available and actively leveraged by different spammers. We present the text similarity analysis among our campaigns in Section 6, to compare with current studies. Meanwhile, compared to studies focusing on a single type of contact (e.g., URLs) to identify spam campaigns, our method is more robust because eight types of contact are considered.

5 MEASURING THE PATTERNS OF FBS SPAMMERS

Sending FBS messages requires heavy manual efforts from spammers, e.g., moving to victims' regions and bootstrapping the FBS, different from other spamming channels (Public Gateway SMS, Twitter and Email) that distribute messages in bulk automatically. As such, different spamming strategies are expected. Also FBS spamming in China is expected to run in different business models from other western countries where prior studies focused on [56, 57].

We carry out the first measurement study towards understanding the FBS ecosystem in China and report our findings in this and next Sections. This section, from "macro-perspective", presents an empirical analysis of FBS spammer behaviors based on 243,998 categorized messages (under "Illegal Promotion", "Fraud" and "Advertisement"). To highlight, we find that FBS spam is mainly utilized for illegal business, operated daily with long active hours and targeting crowded regions. We also confirm the severely adverse impact of FBS spamming in China.

5.1 Business with FBS Spam

Finding 1.1: FBS messages are mostly used to advertise illegal businesses. We use the category classifier (Section 4.1) to label the 243,998 FBS messages into the 13 sub-categories and find illegal businesses ("Illegal Promotion" and "Fraud") are behind over 75% FBS messages. As shown in Figure 4(a), on average 38.6% and 38.2% FBS messages are associated with "Illegal Promotion" and "Fraud", while less than one quarter are with advertisements. For the subcategories, **Fake ID (31.4%)** account for the highest proportion, followed by **Bank Phishing (28.8%)**. Top 5 sub-categories are either illegal or fraudulent which pose high threats to victims.

Compared with previous study of spam from SMS Gateway [56], in which Payday Loan (41%) and Job Advertisements (10%) are the top 2 categories that account for more than half of all spam messages, the FBS spam categories are obviously different. The unique social-economic characteristics of China and profit margin of each business category can be the major reasons for this difference.

5.2 Temporal Characteristics of FBS Spam

We measured the timing patterns of FBS spam activities from the timestamp of each spam text message and assess their similarities.

Finding 1.2: FBS spammers are active for long hours each day except during holidays. Different from other spams, they are active on both weekdays and weekends. Figure 3 shows the timing distribution heatmap of spam messages, each week separated by vertical lines. We first notice a significant decrease in spamming activities during Feb 3 and Feb 10, 2019. In fact, this week overlaps with Spring Festival (the most significant holiday in China with 7 days off), and we speculate most spammers also take a break. It also reflects that FBS spam heavily depends on the manual efforts from spammers , which aligns with previous reports [16, 17].

Different from other spam types including domain squatting [62] and spam calls [49], we find that FBS spammers are active on both weekdays and weekends, as we do not see a significant difference on the number of spam messages. Meanwhile, they are also active for long hours in each day to increase their chances of reaching more victims, typically starting from 7 a.m. to midnight. As shown by Figure 5, for special businesses such as Escort Service and Gambling, spammers are active even after midnight.

Some interesting phenomena are also observed from the timing characteristics. Figure 4(b) shows the fluctuation of the sent messages associated with the top 4 sub-categories, all of which are illegal businesses. A spike occurs for Bank Phishing around Jan 1 (New Year's Day) and we find those messages are generated from several spam text templates. Given that a template can be attributed to a FBS campaign, it suggests several FBS campaigns start to get active near that holiday. For days around Feb 5 (Spring Festival), Gambling-related campaigns remain active while other categories all experience activity decrease. Besides, the Gambling-related campaigns add festival-specific contents (e.g., special Spring-Festival offers) into messages, to better target mobile users.

5.3 Spatial Characteristics of FBS Spam

In addition to the temporal analysis above, we also carry out spatial analysis to measure the prefer locations of FBS activities. Tracking spammer's location is challenging due to the lack of exact location information. While the location of legitimate BS can be learnt from the BS-ID dataset (see Section 3), FBS does not have a valid BS-ID and the FBS equipment is a moving target. In this work, we infer the location of FBS based on the information of the last legitimate BS before the user connects to the FBS. When FBS hijacks user's cellular communications, the user should be not too far away from the legitimate BS, therefore our inference method can derive a rough estimation of FBS locations. To obtain more meaningful information about the FBS location, like the types of places, we query the Geocoding API of Tencent Map [6] to retrieve its nearby Places of Interest (POI) and analyze them. For reference, we also provide the average distribution of non-FBS spam victims observed by our data collection software in Figure 12 in Appendix B.

Similar with previous study [50, 54], as shown by Figure 6, we find more FBS messages in east China than the west, which aligns with the distribution of population. We also find quite active FBS spamming activities in several large provincial capitals, such as Chengdu (in southwest China) and Guangzhou (in south China). More efforts should be devoted in these cities to mitigate FBS spam.

Finding 1.3: FBS spammers are largely active near main roads and highly-populated regions (e.g., residential areas and institutions) to increase their influence. In Figure 6, we zoom

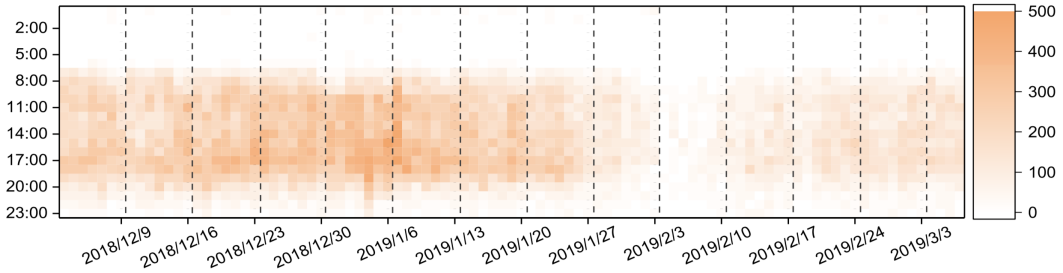
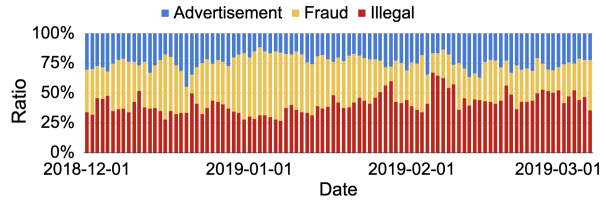
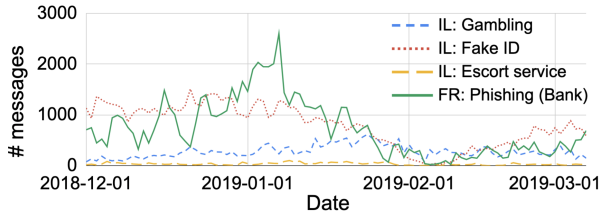


Figure 3: Heatmap of spam activities. Each week is divided by dashed vertical lines.

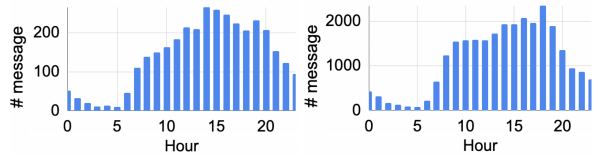


(a) Ratio of three major spam categories



(b) Count of messages associated with the 4 top spam sub-categories

Figure 4: Distribution of spam message categories



(a) Escort service (b) Gambling

Figure 5: Hourly activities of FBS spam.

into Chengdu, where a large number of FBS receivers are observed and find they are concentrated around high-speed ring roads. We speculate that instead of being stable or moving slowly, the spammers prefer to move on expressways to avoid being discovered and increase the spam coverage per time period. For those located elsewhere, we query their POIs to investigate the specific places nearby. When multiple POIs are found near one place, we choose the most well-known one according to the popularity attribute in queried results as a representative. As shown in Figure 7, FBS spammers are found largely active in highly-populated regions to reach more mobile users, such as residential areas and large institutions.

Finding 1.4: FBS spammers prefer serving different businesses at different places. From Figure 7 we can see that the preferred spam businesses vary based on spammers' locations. As an example, Gambling and Escort Service messages show up more often near hotels and transport stations. Also shown in Figure 8, the dominate spam businesses also differ between cities, possibly due to their different cultural background and economic status. For instance, Gambling and Escort Service messages are mostly distributed in Macau (where the two services are legal) and Zhuhai (a city next to Macau). Northeastern China cities (e.g., Dalian and Shenyang) see more financial fraud and bank phishing messages.

5.4 Impact of FBS spam

We also confirm the real-world impact of FBS spam in China. The hashed IMSI and IMEI of each receiver could help us to estimate the affected client population. Besides, for messages containing domains, we use PDNS data to learn the number of follow-up web visits, so how popular each business is can be estimated. To notice, the detected FBS messages in Spam Inbox can still be visible to mobile users and the domain links are clickable.

Finding 1.5: Over 100,000 mobile devices still receive FBS messages within the study period and some device owners are heavily harassed. FBS spamming is not a new problem and methods combating it have already been deployed by a large number of user devices in China (e.g., our security mobile app and Baidu's app[50]). However, we found spammers are relentlessly sending FBS messages despite the efforts on detection. 103,191 IMEIs and 110,185 IMSIs (the difference is caused by dual-card devices) are found receiving FBS messages within the study period. Figure 9 shows the ECDF of spam messages received per IMEI/IMSI. Some users are found heavily harassed, e.g., 38 IMEIs and 34 IMSIs receive more than 100 spam messages during 97-day data collection period (i.e., over 1 message per day on average). Their geo-location indicates that they often appear near regions where FBS spamming is active. Meanwhile, we also observed that these heavily harassed victims usually receive more than one spam messages after connecting to an FBS, or connect to multiple FBS devices consecutively and received messages from all of them in a short period of time.

Finding 1.6: Even after marked as spam, domains in the messages still receive considerable number of visits. We extract 3,834 suspicious domains from all spam text messages and found 3,197 (83.4%) are labeled as malicious by at least one blacklist. We then query for their DNS lookup volume and active time through PDNS. The visit numbers are not evenly distributed, and several domains have long active time and large query volume. For instance, over 70% suspicious domains receive more than 100 queries, 403 (12.6%) domains are queried for over 5,000 times, and 34 Gambling domains even have been visited for more than 100K times.

Admittedly, our visitor estimation can be inaccurate. We are not able to distinguish whether the visits are from FBS victims or other users. Due to DNS cache of resolvers, we may underestimate the actual query volume of malicious domains. Yet, our estimation suggests FBS does have considerable impact on people in China.

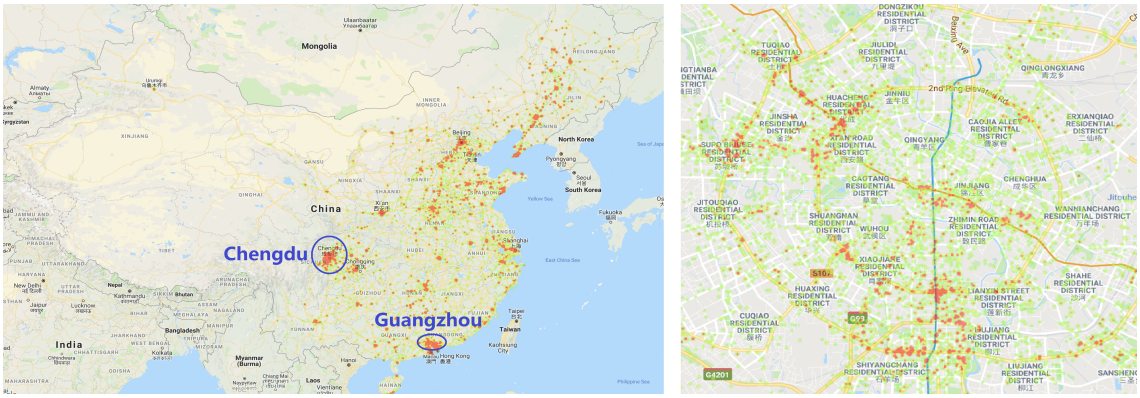


Figure 6: Geo-distribution of FBS spam victims: China-wide (Left), Chengdu city (Right)

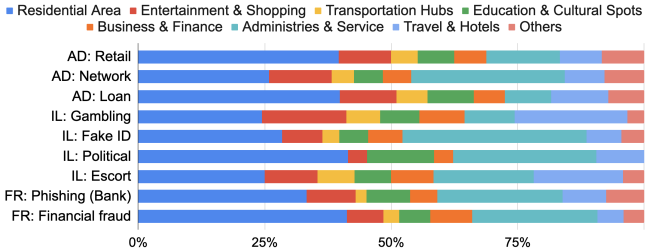


Figure 7: Spam activities near highly-populated regions.

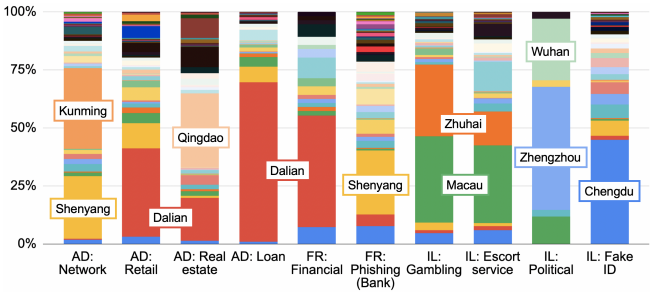


Figure 8: City-level analysis of FBS spam activities.

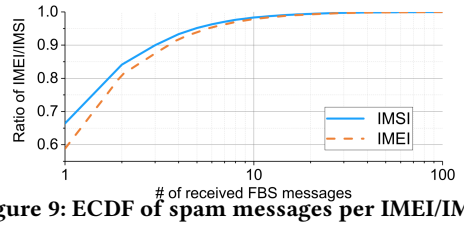


Figure 9: ECDF of spam messages per IMEI/IMSI.

6 MEASURING THE STRATEGIES OF FBS MESSAGE CREATION

The prior section describes the "macro-level" characteristics of FBS spam. In this section, we present its "micro-level" properties, focusing on behaviors of different campaign operators.

Generally, a FBS message is customized under two goals: *trapping more users* and *evading content-based detection and actor attribution.* We organize our findings as follows: Section 6.1 gives an overview on FBS spam campaigns and their characteristics. Sections 6.2 and 6.3 describe the strategies of user trapping. Sections 6.4, 6.5 and 6.6 show how evasion is performed by spammers detection. Lastly, Section 6.7 analyzes the scenario that the detected FBS spam messages are recovered by mobile users.

6.1 FBS Campaign Characteristics

In total, we identify **7,884 spam campaigns** associated with 8,316 unique contacts (see Table 2) by applying the campaign classifier (Section 4.2). We identify new evasion strategies and outsourcing models in FBS spamming, which are never mentioned by prior works [50], suggesting FBS campaigns have evolved their strategies to be more complex and elusive.

First, looking into the distribution of FBS messages over campaigns, the first 100 campaigns (1.3%) account for 34% of all FBS messages, and the first 1,000 (13%) campaigns 80%. In particular,the largest campaign alone is associated with over 11,120 messages (4.55%) in 97 days. On the other hand, contacts are more evenly distributed among campaigns, as over 96% campaigns choose a single contact for all messages. While spammers can spoof sender and evade sender blacklists, their contact in the messages leave them vulnerable to contact-based blocking.

Finding 2.1: Most spam campaigns are short-lived, especially the ones associated with illegal businesses. We count the lifetime of a spam campaign as the number of days its spam messages are observed. We find that most campaigns are short-lived: 92% are only active for less than 10 days. For the remaining, 177 (2.2%) campaigns are active for more than 30 days, and they account for 82,190 (33.7%) spam messages in our dataset.

By manually examining the top 20 long-lived campaigns, we find they are mostly associated with Fake ID and invoice (12 campaigns, 60%), which is illegal in China while only considered as "light" crime. We also examine the 50 least-active campaigns with at most one day lifetime, and find they prefer Phishing messages (25 campaigns, 50%), including 16 campaigns impersonating banks. It might be a reasonable strategy to protect the fraudsters as those campaigns are considered more detrimental (e.g., big financial loss to the victim if she transfers money to fraudsters) and the penalty is much heavier.

Finding 2.2: Multiple spam campaigns can be undertaken by the same FBS operator at the same time, which might be related to delegation. Table 4 presents the timing and locality characteristics of the top 10 spam campaigns, which in total account for 27,702 (11.3 %) messages. One interesting observation we find is that, the timing characteristics of Campaigns 2&3, Campaigns 6&8

Table 4: Top 10 spam campaigns sending most messages

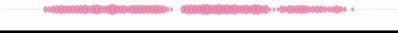
No.	Category	# Message	# IMEI	# Contact	Active Days	Active time (Dec 1, 2018 - Mar 7, 2019)	Hourly Distribution (01 - 24h)	Locality
1	Loan	11,120	1,646	24	95			Dalian
2	Gambling	3,623	2,080	1	97			Macau
3	Gambling	2,971	1,904	8	97			Macau
4	Loan	2,327	687	7	88			Dalian
5	Gambling	1,416	580	1	77			Macau & Zhuhai
6	Fake ID and invoice	1,380	940	1	71			Chengdu
7	Gambling & Loan & Escort service	1,283	460	7	60			Macau & Zhuhai
8	Advertisement (Others)	1,249	889	1	72			Chengdu
9	Phishing (Bank)	1,206	903	1	35			(Several cities of Sichuan Province)
10	Gambling	1,127	486	1	76			Macau & Zhuhai

Table 5: Contacts used for multiple services with overlapped active time

Message Content	Category	Active Days
[Good news] We offer loans to pals who need money. Please contact Manager Chen at WeChat **132****1290.**	Advertisement: Loan	Dec 30, 2018
Freshly squeezed peanut oil from the countryside. Natural without any additives. Contact WeChat **132****1290**.	AD: Other	**Jan 23, 2019**
Celebrating the 10th anniversary of the Crown online casino! Amazing recharge rates and cash rewards! Contact WeChat **132****1290** for more surprise!	IL: Gambling	**Jan 23, 2019** - Jan 27, 2019

and Campaigns 5&10 are similar and the active locations of each pair, as shown in the last column, are also consistent.

These facts indicate the spam messages of each pair of campaigns might be sent from the same set of FBS devices. In other words, the FBS operators (people carrying FBS devices) could work with multiple campaigns and distribute their spam messages simultaneously. As another proof, we check the affected IMEI sets of these campaigns and find at least 54% overlap between each similar campaign pair. Besides, it is worth noting that spam campaigns undertaken by the same FBS operator might belong to different business, such as Fake ID and Other advertisements (Campaigns 6/8).

Interestingly, from their geo-location presented in Table 4, we find that 9 of the top 10 campaigns stay in one same city or nearby cities in our 97-day time span. By contrast, Campaign 9 (associated with Phishing) frequently migrates among several cities in Sichuan Province. Starting from Chengdu in Dec 21, 2018, it travels to Guangan (Dec 24), Dazhou (Dec 25), Suining (Dec 27), Nanchong (Dec 29) and back to Chengdu (Dec 31). As such, while tracing a spammer within a city can lead to successful capture, cooperation of government departments across cities is necessary to combat campaign migration.

Finding 2.3: The same contact can be shared among different spam categories, indicating the task of victim interaction might be delegated as well. While one contact is supposed to interact with victims looking for the same business, we find there are contacts associated with different businesses, which is quite counter-intuitive. Campaign 7, as presented in Table 4, is such a case, in which a WeChat account is reused.

To systematically uncover those shared accounts, we compute *entropy value* of each contact and report the ones with abnormal values. *Entropy value* is defined as $H = -\sum_i p_i \log p_i$, where i indicates one kind of spam category and p_i is the proportion of that category i in all messages associated with one spam contact. Entropy is widely used to assess the closeness of a sample to others: the higher the entropy, the more likely this sample relates to other samples. Similarly in our task, *the larger the entropy value is, the higher the probability that the account is used for multiple services.* After manual verification, we set 0.6, which can detect most such contacts with no false positives, as the experienced threshold of entropy, and find 185 contacts (2.22%) that are shared among multiple categories in all campaigns. Note that, we use *entropy* instead of directly counting the number of message categories to mitigate the impact of mis-classified messages (3.1%, as shown in Table 3).

Specifically for WeChat, 12 of the top 20 accounts (60%) span several services. One example is shown in Table 5. The same contact is used for several unrelated kinds of spam messages and *the active time of different categories overlaps.* We speculate that the contact belongs to an intermediate party, which takes orders from several upstream spammers.

Combined with **Finding 2.2** and **Finding 2.3**, we speculate the FBS ecosystem might be evolved to a hierarchical structure (at least 3 layers), in which the downstream operators act as delegate for distributing messages and the intermediate contact operators take care of victim engagement for the upstream business owners.

Finding 2.4: Message templates are frequently shared among spam campaigns, therefore attributing spam campaign with templates might be erroneous. Spammers can create messages based on existing templates to reduce manual efforts. We find the "templated" messages always share the same descriptive text while only the part of contacts are customized. The same observation has been found in other types of spams like email and twitter and previous works attribute messages to spam campaigns based on their templates [46, 47].

Table 6: Message templates and their categories

Category	# Template	# Msg	% Templated Msg
Phishing (Bank)	168	22,737	32.35%
Fake ID and invoice	19	4,427	5.77%
Gambling	20	10,808	39.62%
Loan service	18	11,135	49.48%
Retail	11	3,385	28.11%
Advertisement (Other)	10	2,314	17.19%
Network Service	5	929	6.46%
Escort Service	6	504	14.97%
Financial Fraud	3	57	13.23%
Real Estate	2	32	1.35%

We aim to study how templates are used for FBS messages. To identify the templates, we calculate the pairwise content similarity (we use edit distance) between spam messages and cluster the similar ones. If the edit distance between two messages is smaller than 10 bytes (note that a Chinese character uses 3 bytes in UTF-8 encoding, so it means only less than 4 characters different), we regard that they belong to the same cluster. A cluster is considered to be generated by a message template, only if it contains more than 5 distinct messages.

In the end, we identify 262 templates from the entire FBS dataset, which are used by 994 spam campaigns. As shown in Table 6, the templates are largely associated with Fraud and Illegal Promotion, particularly Phishing spammers which are impersonating banks (with 90% of all templates). Businesses like Loan, Gambling, and Phishing (Bank) have the highest rate of distributing templated messages. In addition, we found that 83 (31.68%, of 262) message templates are used by more than one spam campaigns, and 858 (86%, of 994) spam campaigns share their templates with others. The most popular template impersonates well-known banks and is shared among 311 spam campaigns.

Admittedly, due to the lack of spammers' real identity information (which is controlled by law enforcement departments), we acknowledge that we cannot associate each campaign with their real spammers behind, thus campaigns using the same templates could belong to the same big group of spammers. However, we believe that this is not the case for at least 83 templates (31.68%), since they are shared among as many as 858 spam campaigns. Consequently, template-based campaign detection methods (i.e., messages using the same template belong to the same spam campaign) could be inaccurate to some extent.

6.2 Tricking Strategy: Sender Spoofing

Finding 2.5: FBS spammers use spoofed sender numbers of well-known companies to make their messages more deceptive. A recent work has uncovered that Sender ID (Caller ID) Spoofing is very effective in telephone scams [61]. Regarding spam messages, spammers can send messages from arbitrary spoofed numbers using FBS (see Section 2), and we found the numbers of well-known companies are impersonated.

As shown in Table 7, large banks, online payment platforms, ISPs and insurance companies are among the top 15 spoofed sender numbers. Particularly, the Industrial Bank of China (ICBC) is impersonated the most by 23,444 (8.4%) FBS messages in our dataset.

To combat such impersonation attack, we suggest user education and extra security measures be performed by the targeted brands.

Table 7: Top 15 spoofed senders

Type	Sender	Description	# Msg
Bank	95588	Industrial and Commercial Bank of China	23,444
	95533	China Construction Bank	13,388
	95599	Agricultural Bank of China	2,963
	95595	China Everbright Bank	308
	95566	Bank of China	213
	95559	Bank of Communications	157
	95558	China Minsheng Bank	137
ISP	10086	China Mobile Communications Corp.	12,161
	1008611	China Mobile Communications Corp.	308
	10010	China United Network Communications	224
Payment	95107	WeChat Pay	5,039
	95188	Alipay	353
Insurance	95518	Peoples Insurance Company of China	149
	95511	Ping An Insurance	136

For instance, enterprises could provide their official phone numbers and domain names to mobile applications. Upon receiving a message that claims to be sent from a well-known company with unmatched domains, the user would be prompted with the potential phishing risk.

6.3 Tricking Strategy: Message Wording

Finding 2.6: The language of FBS messages is usually captivating (e.g., with scares and monetary lures) to engage users. After manually inspecting the 262 message templates, we find 114 (43.5%) templates are designed to scare users, including frozen credit cards (76 templates), blocked online accounts (16 templates) and stolen electronic devices (14 templates). For the remaining, 104 (39.7%) templates attract users by money lures, such as credit card limit increasing (58 templates) and ISP discounts (15 templates). The use of highly captivating language and content is supposed to make FBS messages more effective at trapping users. As a countermeasure, user education would be necessary to discern the language used by the official parties and the impersonators.

6.4 Evasion Strategy: Domain Infrastructure

Finding 2.7: FBS spammers usually use newly-registered domains or domains from domain-squatting services. URL shorteners are extensively used by malicious domains to avoid blacklisting. From our message dataset we find 3,834 domain names. According to their WHOIS data, 2,285 (75.9%) are registered after 2018, and 1,155 (38.4%) are registered after 2019 (less than two months). This suggests that most domains used in the spam messages are newly-registered, which is similar to patterns in other spam businesses [42, 47, 52].

However, we also find that 278 (7.3%) malicious domain names are over 3 years old. After manual analysis of 20 samples based on their WHOIS and passive DNS, we find they are related to one kind of special registration behaviors. First, the spammer registered a large number of (similar) domains at an early time, and then leveraged the domain names for illegal services later in different batches. Spammers can switch to a new batch of domains after the previous domains were blocked by security vendors. As we observe from PDNS, for a domain, there is often a prominent gap between being registered and being activated for spam campaigns.

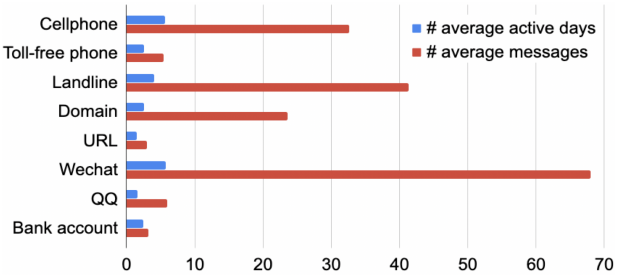


Figure 10: Usage of eight types of contact information

Regarding the 575 URLs in FBS messages, we find 69% use *URL shorteners* to hide their real domain names and paths, which makes detection and blacklisting more difficult. For example, 289 URLs are hosted under `t.cn`, and 27 are under `dwz.cn` (both are URL shortening service providers [1, 4]).

In fact, URL shortening has also been reported in studies of other spams [21, 41, 58]. In previous studies, security researchers usually leverage the APIs of URL shorteners to query domain statistics including number of clicks to evaluate the impact of spam [39, 59].

However, most URLs in our FBS messages use shortening services in China, which do not provide such statistics, making detection and analysis of such domains difficult. At the time of writing all shortened URLs have been either suspended or inactive, and we have not found a channel to request the visit of statistics. Such channel should be established to facilitate spam detection.

6.5 Evasion Strategy: Bank Account

Finding 2.8: FBS spammers are abusing flawed bank card policies to reduce the risks of being blocked. Spammers associated with financial fraud often use bank accounts to receive transactions from victims. As such, bank accounts should directly point to the spammers behind. In total, we find 29 valid bankcard numbers in Fraud messages, and query their account types (credit/debit card), bank names and bank locations by a commercial online tool [3]. All 29 bank accounts are debit card numbers and **all** of them are registered under banks in mid-west China (e.g., Guizhou Province). According to a previous report [20], security procedures for bank account creation in the areas could be flawed (e.g., no strict ID verification) such that the real identities of account owners are untraceable. Therefore spammers prefer to choose these regions in order to lower their risks.

Interestingly, we also find that 16 accounts belong to two special card types (namely "Jinsui Tongbao Card" and "Peony Card"), which allow the creation of *multiple secondary cards* under the same name *for free*. Consequently, if one of them is blacklisted, spammers can switch to other secondary accounts, eliminating their costs of creating new bank accounts.

In general, the creation of secondary cards typically only applies for credit cards, thus we recommend that these banks carefully justify their card policies to avoid being abused by FBS spammers.

6.6 Evasion Strategy: Spammer Contacts

Finding 2.9: Social media (e.g., WeChat) ID has become the major type of spammer contact. Among the top 10 campaigns, 32 (of 52, 61%) contacts are WeChat accounts. Among all campaigns, Figure 10 shows the active time and messages associated with each type of contacts. The results are similar, that WeChat accounts are prominently used by the spam campaigns. In addition, on average, each WeChat account is embedded in 68 FBS spam messages.

Compared with other types of contact, we find that the WeChat accounts are oftentimes long-lived, with a 5.8-day lifetime (embedded in FBS messages) on average, and the longest lifetime of 97 days during our 3-month time span. Further, we also randomly sample 50 WeChat accounts two months after our dataset time span, and find that 26 (52%) are still active. Interestingly, all accounts have nicknames or avatars clearly indicating their relation to promotion FBS businesses. On the other hand, among the 24 inactive accounts, only 7 have been detected and locked by WeChat official platforms (i.e., showing an abnormal account status), while others have already been closed by the users. The low blocking rate and long lifetime of spam accounts suggest more works need to be done on the social media platforms.

6.7 Trapped Mobile Users

As we mentioned in Section 5, we find that malicious domain names embedded in FBS messages received considerable visits, suggesting FBS spammers are effective at trapping end-users, even when their messages are classified as spam. While our mobile security application moves the detected FBS messages into a Spam Inbox (see Section 4), users are allowed to move the detected message back to the normal Inbox. Here we present a further investigation on the issue, by inspecting messages that are marked as FBS spam first but later recovered by mobile users.

Finding 2.10: For security applications, accurate detection and blocking of FBS messages is not enough to combat FBS, as a large number of detected FBS messages are recovered by users. We obtain a sample of 3,775 recovered FBS messages from our industry partner to assess the scale of this issue. We find the messages are largely associated with *Gambling* (3,181, 82.81%) and *Phishing (Bank)* (343, 9.1%). As such, we believe beyond FBS detection, user education and new UI design are also necessary. For example, our fine-grained classifiers would help the security applications to make more detailed and targeted user notifications or warning tips for the detected FBS messages, such as emphasizing the bad consequences of interacting with a specific illegal business and explicating the dangers of Fraud messages.

6.8 Summary

We discussed the characteristics of spam campaigns and the strategies they operate in this section. Although some of the observations, such as shortened URLs and spoofed senders, are similar with previous studies on other types of spam activities [21, 41, 58, 61]. However, this is the first time these phenomena are confirmed to exist in FBS spam. Some founded strategies are novel of FBS spamming, such as flexible locality changes, template sharing, bank card abuse and the high adoption of social media accounts.

From the findings in Section 5 and 6, we found FBS spam is still active and evolving, indicating that more efforts should be taken at multiple layers of the FBS business hierarchy and more parties should be involved to counter FBS. A longitudinal study in this area is necessary and this is planned as one of our future work.

7 DISCUSSION

In this paper, we perform an in-depth measurement study towards the FBS spam ecosystem and investigate how the FBS business is operated. Besides detecting FBS devices, our findings also point to recommendations for different parties evolved in FBS ecosystem to mitigate the security issues together, including the following tips:

• **Mobile Carriers.** Mobile carriers need to upgrade their cell towers and abandon the vulnerable GSM protocol. Even though it will take years, the effort is worth because it will fundamentally eliminate the security threats from FBS spam.

• **Government Agencies.** Based on our analysis of the geographic distribution of FBS spammers, government agencies should make more efforts in some seriously affected cities, like Guangzhou and Chengdu, especially pay more attention to most FBS active places such as express ways and residential areas.

• **Bank.** Banks can also play an important role in the mitigation of FBS spam. We recommend banks to re-evaluate their bankcard policies immediately to avoid being abused, which effectively raise the difficulty for spammers to receive transactions from victims.

• **Social Media Platform.** For the social media platforms, e.g. WeChat, checking whether an account is associated with fraudulent activities would be greatly helpful for fighting against spam.

• **Security Software.** Our multi-classifiers would help to improve the new UI system of security applications, which give more detailed warning up to end-users to avoid being trapped. Furthermore, we provide the extracted popular templates as signatures for detection and our industry partner found almost double amount of FBS messages could be confirmed in the same period, indicating that the templates we found could be effective supplementary to optimize the detection of FBS.

• **Enterprises.**For a number of well-known brand companies, such as banks (especially ICBC) and Chinese ISPs, educating their customers to understand the harm of FBS is necessary. The scenarios of deceptive messages summarized in our work would be valuable materials for user education.

8 RELATED WORK

Fake Base Station Analysis. FBS is a known attack device to arbitrarily propagate malicious content to vulnerable end users, like phishing or fraud messages. Previous studies focus on how to detect the existence of FBS based on network signal characteristics, such as irregularities in mobile network [2] and received cellular signals [5, 24, 30, 63]. Especially, Zhenhua Li *et al.* presented a real-world deployable FBS detection system called FBS-Radar in 2017, which is based on crowd sourced data from millions end users. They found that FBS can be precisely identified by the signal strength and BS-ID difference [50], without analyzing text message contents.

To the best of our knowledge, due to the limitation of collected detection results, there have been no prior studies on in-deep analysis of FBS spammer behaviors. In our study, we leverage the three-month detection results from off-the-shelf approaches as ground truth, and perform a comprehensive measurement study on the strategies of spam campaigns. Our study not only complements existing researches regarding this threat, but also contributes to fundamentally mitigating the security risks of FBS.

Mobile Spamming Ecosystem Analysis. In addition to FBS, fraudsters can also rely on other channels, like SMS gateway, to generate spam messages. Approaches have been proposed to detect spam messages based on their content information [26, 32, 38, 56] or senders' traffic behaviors [44]. In general, obtaining large-scale dataset of spam messages is challenging, inducing a limitation of scale. For example, Almeida *et al.* only obtained 747 spam messages from victim reports [26]. Similarly, Reaves *et al.* collected around 400K short text messages from a public SMS gateway [57] but only around 8.2% of them are classified as spam.

Besides, some malicious call prevention techniques, including white/black-listing, call's ID reputation, and machine learning approaches [49] are also developed to identify the telephony spams and scams. In a more recent work, the authors executed an ethical telephone phishing scam and find that the spoofed Caller ID, had a significant effect in tricking the victims [61].

Spamming Activities Analysis in other fields. In addition to spam messages [54], previous analysis also measured the spam activities on various social platforms, such as on Twitter [39, 59, 60], Facebook [36], emails [45–47, 55] or directly on the websites [31]. In this work, we develop an approach to classify spam messages into spam campaigns. Similar analysis has been done by previous works as well for spam content classification. The proposed methods include template-based clustering, topic analysis, and clustering based on the sending behaviors of spammer accounts. Those methods lead to discoveries about geographic distribution, evasion measures, and attack strategies of spam campaigns [36, 39, 47, 55, 57].

Compared to previous studies, our study reveals unique insights regarding FBS powered spam message distribution based on a real-world and large-scale dataset, which complements the understanding about fraudulent activities of FBS spam ecosystem.

9 CONCLUSION

Currently, despite the increasing abuse of FBS spam in the real world, we still lack an in-depth understanding about FBS ecosystem. In this paper, leveraging three months real-world FBS messages collected from China, we present by far the first comprehensive measurement study on FBS spam ecosystem from a country-level perspective. Through classifying business types of FBS messages and identifying spam campaigns by self-designed algorithms, we are not only able to explore the in-depth characteristics of spammer behaviors, such as business categories, user trapping strategies and evasion techniques, but also gain many valuable insights from the evolving FBS spam campaigns, such as hierarchical organization architecture and resource sharing. To summarize, our findings demonstrate how FBS ecosystem is organized and how spammers behave. To help the security community improve their mechanisms against the FBS, we release our labeled ground-truth dataset and provide recommendations to different parties.

ACKNOWLEDGEMENT

We thank all anonymous reviewers for their valuable comments to improve the paper. This work was supported in part by the National Natural Science Foundation of China (U1836213, U1636204, Grant 61772307), the BNRist Network and Software Security Research Program (Grant No.BNR2019TD01004) and National Key R&D Program of China 2018YFB1800405.

REFERENCES

[1] [n.d.]. Baidu short URL. https://dwz.cn.
[2] [n.d.]. CatcherCatcher - Mobile Network Assessment Tools - SRLabs Open Source Projects. https://opensource.srlabs.de/projects/mobile-network-assessment-tools/wiki/CatcherCatcher.
[3] [n.d.]. Show API. https://www.showapi.com/apiGateway/view?apiCode=30.
[4] [n.d.]. Sina short URL. https://open.weibo.com/wiki.
[5] [n.d.]. SnoopSnitch SRLabs Open Source Projects. https://opensource.srlabs.de/projects/snoopsnitch.
[6] [n.d.]. WebService API. https://lbs.qq.com/webservice_v1/guide-gcoder.html.
[7] 2014. 19 Fake Mobile Base Stations Found Across US. Are They For Spying or Crime? http://ibtimes.co.uk/19-fake-mobile-base-stationsfound-across-us-are-they-spying-crime-1464008..
[8] 2014. Are your calls being intercepted? 17 fake cell towers discovered in one month. http://computerworld.com/article/2600348/mobile-security/areyour-calls-being-intercepted-17-fake-cell-towers-discovered-in-onemonth.html.
[9] 2014. Chinese cops nab 1,530 mobile SMS spammers in raid on fake base. https://nakedsecurity.sophos.com/2014/03/26/chinese-cops-nab-1530-mobile-sms-spammers-in-raid-on-fake-base-stations/.
[10] 2014. Phony cell towers are the next big security risk. http://www.theverge.com/2014/9/18/6394391/phony-cell-towersare-the-next-big-security-risk.
[11] 2014. Qihoo 360: Research reports of Fake Base Sations. http://www.ceocio.com.cn/e/action/ShowInfo.php?classid=69&id=145193.
[12] 2015. Fake Stingray mobile base stations discovered spying on millions of Londoners. http://www.ibtimes.co.uk/fake-stingray-mobile-basestations-discovered-spying-millions-londoners-1505368.
[13] 2016. Demystifying Fake Base Stations. http://business.sohu.com/20160507/n448197405.shtml.
[14] 2016. Demystifying the Industrial Chain of Fake Base Stations. http://m.sohu.com/n/444726367/.
[15] 2016. Mobile Security Reports by Qihoo 360. http://zt.360.cn/2015/reportlist.html?list=1.
[16] 2016. Research Reports: 2016 Fake Base Station of China. http://zt.360.cn/1101061855.php?dtid=1101061451&did=1101741409.
[17] 2016. Research Reports: 2016 Fake Base Station of China. https://m.qq.com/security_lab/news_detail_361.html.
[18] 2016. Undergound economy of Fake Base Station. http://www.ceweekly.cn/2016/0919/164561_4.shtml.
[19] 2017. Chinese Internet Security Report for the first half of 2017. https://s.tencent.com/research/report/242.html.
[20] 2017. National Internet Finance Association of China. http://www.nifa.org.cn/nifa/2955675/2955759/2967869/index.html.
[21] 2018. How Spammers Conduct Mass Spam URL Attacks. https://www.datavisor.com/blog/how-spammers-conduct-mass-spam-url-attacks.
[22] 2020. FBS SMS Dataset. https://github.com/Cypher-Z/FBS_SMS_Dataset.
[23] Qihoo 360. 2019. 360 Mobile Guard Official Website. https://shouji.360.cn.
[24] Dare Abodunrin et al. 2015. Detection and Mitigation methodology for Fake Base Stations Detection on 3G/2G Cellular Networks. (2015).
[25] Akiko Aizawa. 2003. An information-theoretic perspective of TF-IDF measures. *Information Processing & Management* 39, 1 (2003), 45–65.
[26] Tiago A Almeida, José María G Hidalgo, and Akebo Yamakami. 2011. Contributions to the study of SMS spam filtering: new collection and results. In *Proceedings of the 11th ACM symposium on Document engineering*. ACM, 259–262.
[27] Eihal Alowaisheq, Peng Wang, Sumayah A Alrwais, Xiaojing Liao, XiaoFeng Wang, Tasneem Alowaisheq, Xianghang Mi, Siyuan Tang, and Baojun Liu. 2019. Cracking the Wall of Confinement: Understanding and Analyzing Malicious Domain Take-downs.. In *NDSS*.
[28] Netlab at Qihoo 360. [n.d.]. Passive DNS System. http://www.passivedns.cn.
[29] Lars Buitinck, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake VanderPlas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. 2013. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*. 108–122.
[30] Alaaedine Chouchane, Slim Rekhis, and Noureddine Boudriga. 2009. Defending against rogue base station attacks using wavelet based fingerprinting. In *2009 IEEE/ACS International Conference on Computer Systems and Applications*. IEEE, 523–530.
[31] Nicolas Christin, Sally S Yanagihara, and Keisuke Kamataki. 2010. Dissecting one click frauds. In *Proceedings of the 17th ACM conference on Computer and communications security*. 15–26.
[32] Gordon V Cormack, José María Gómez Hidalgo, and Enrique Puertas Sánz. 2007. Spam filtering for short messages. In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*. ACM, 313–320.
[33] Adrian Dabrowski, Nicola Pianta, Thomas Klepp, Martin Mulazzani, and Edgar Weippl. 2014. IMSI-catch me if you can: IMSI-catcher-catchers. In *Proceedings of the 30th annual computer security applications Conference*. 246–255.
[34] DDarko. [n.d.]. Python module/library for retrieving WHOIS information of domains. https://github.com/nri-pl/python-whois.
[35] FarSight-Security. [n.d.]. DNSDB data. https://www.farsightsecurity.com/solutions/dnsdb.
[36] Hongyu Gao, Jun Hu, Christo Wilson, Zhichun Li, Yan Chen, and Ben Y Zhao. 2010. Detecting and characterizing social spam campaigns. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*. ACM, 35–47.
[37] Nico Golde, Kévin Redon, and Jean-Pierre Seifert. 2013. Let me answer that for you: Exploiting broadcast information in cellular networks. In *22nd {USENIX} Security Symposium ({USENIX} Security 13)*. 33–48.
[38] José María Gómez Hidalgo, Guillermo Cajigas Bringas, Enrique Puertas Sánz, and Francisco Carrero García. 2006. Content based SMS spam filtering. In *Proceedings of the 2006 ACM symposium on Document engineering*. ACM, 107–114.
[39] Chris Grier, Kurt Thomas, Vern Paxson, and Michael Zhang. 2010. @Spam: the underground on 140 characters or less. In *Proceedings of the 17th ACM conference on Computer and communications security*. ACM, 27–37.
[40] 360 Mobile Guard. [n.d.]. Pravicy Policy of 360 Mobile Guard. http://shouji.360.cn/about/privacy/index_2.0.html.
[41] Neha Gupta, Anupama Aggarwal, and Ponnurangam Kumaraguru. 2014. bit.ly/malicious: Deep dive into short url based e-crime detection. In *2014 APWG Symposium on Electronic Crime Research (eCrime)*. IEEE, 14–24.
[42] Shuang Hao, Matthew Thomas, Vern Paxson, Nick Feamster, Christian Kreibich, Chris Grier, and Scott Hollenbeck. 2013. Understanding the domain registration behavior of spammers. In *Proceedings of the 2013 conference on Internet measurement conference*. ACM, 63–76.
[43] Han He. 2020. *HanLP: Han Language Processing*. https://github.com/hankcs/HanLP
[44] Nan Jiang, Yu Jin, Ann Skudlark, and Zhi-Li Zhang. 2013. Greystar: Fast and Accurate Detection of SMS Spam Numbers in Large Cellular Networks Using Gray Phone Space. In *Presented as part of the 22nd USENIX Security Symposium (USENIX Security 13)*. 1–16.
[45] Chris Kanich, Christian Kreibich, Kirill Levchenko, Brandon Enright, Geoffrey M Voelker, Vern Paxson, and Stefan Savage. 2008. Spamalytics: An empirical analysis of spam marketing conversion. In *Proceedings of the 15th ACM conference on Computer and communications security*. ACM, 3–14.
[46] Christian Kreibich, Chris Kanich, Kirill Levchenko, Brandon Enright, Geoffrey M Voelker, Vern Paxson, and Stefan Savage. 2008. On the Spam Campaign Trail. *LEET* 8, 2008 (2008), 1–9.
[47] Christian Kreibich, Chris Kanich, Kirill Levchenko, Brandon Enright, Geoffrey M Voelker, Vern Paxson, and Stefan Savage. 2009. Spamcraft: An Inside Look At Spam Campaign Orchestration.. In *LEET*.
[48] Omer Levy and Yoav Goldberg. 2014. Neural word embedding as implicit matrix factorization. In *Advances in neural information processing systems*. 2177–2185.
[49] Huichen Li, Xiaojun Xu, Chang Liu, Teng Ren, Kun Wu, Xuezhi Cao, Weinan Zhang, Yong Yu, and Dawn Song. 2018. A Machine Learning Approach To Prevent Malicious Calls Over Telephony Networks. In *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, 53–69.
[50] Zhenhua Li, Weiwei Wang, Christo Wilson, Jian Chen, Chen Qian, Taeho Jung, Lan Zhang, Kebin Liu, Xiangyang Li, and Yunhao Liu. 2017. FBS-Radar: Uncovering Fake Base Stations at Scale in the Wild.. In *NDSS*.
[51] Baojun Liu, Zhou Liu, Peiyuan Zong, Chaoyi Lu, Haixin Duan, Ying Liu, Sumayah Alrwais, Xiaofeng Wang, Shuang Hao, Yaoqi Jia, et al. 2019. TraffickStop: Detecting and Measuring Illicit Traffic Monetization Through Large-scale DNS Analysis. In *2019 IEEE European Symposium on Security and Privacy (EuroS&P)*. IEEE, 560–575.
[52] Baojun Liu, Chaoyi Lu, Zhou Li, Ying Liu, Haixin Duan, Shuang Hao, and Zaifeng Zhang. 2018. A reexamination of internationalized domain names: the good, the bad and the ugly. In *48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*. 654–665.
[53] Daiping Liu, Zhou Li, Kun Du, Haining Wang, Baojun Liu, and Haixin Duan. 2017. Don't Let One Rotten Apple Spoil the Whole Barrel: Towards Automated Detection of Shadowed Domains. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 537–552.
[54] Ilona Murynets and Roger Piqueras Jover. 2012. Crime scene investigation: SMS spam data analysis. In *Proceedings of the 2012 Internet Measurement Conference*. ACM, 441–452.
[55] Abhinav Pathak, Feng Qian, Y Charlie Hu, Z Morley Mao, and Supranamaya Ranjan. 2009. Botnet spam campaigns can be long lasting: evidence, implications, and analysis. In *ACM SIGMETRICS Performance Evaluation Review*, Vol. 37. ACM, 13–24.
[56] Bradley Reaves, Logan Blue, Dave Tian, Patrick Traynor, and Kevin RB Butler. 2016. 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, 165–170.
[57] Bradley Reaves, Nolen Scaife, Dave Tian, Logan Blue, Patrick Traynor, and Kevin RB Butler. 2016. Sending out an SMS: Characterizing the Security of the SMS Ecosystem with Public Gateways. In *2016 IEEE Symposium on Security and Privacy (SP)*. IEEE, 339–356.

[58] Kurt Thomas, Chris Grier, Justin Ma, Vern Paxson, and Dawn Song. 2011. Design and evaluation of a real-time url spam filtering service. In *2011 IEEE symposium on security and privacy*. IEEE, 447–462.
[59] Kurt Thomas, Chris Grier, Dawn Song, and Vern Paxson. 2011. Suspended accounts in retrospect: an analysis of twitter spam. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*. ACM, 243–258.
[60] Kurt Thomas, Damon McCoy, Chris Grier, Alek Kolcz, and Vern Paxson. 2013. Trafficking Fraudulent Accounts: The Role of the Underground Market in Twitter Spam and Abuse. In *Presented as part of the 22nd USENIX Security Symposium (USENIX Security 13)*. 195–210.
[61] Huahong Tu, Adam Doupé, Ziming Zhao, and Gail-Joon Ahn. 2019. Users Really Do Answer Telephone Scams. In *28th USENIX Security Symposium (USENIX Security 19)*. 1327–1340.
[62] Thomas Vissers, Jan Spooren, Pieter Agten, Dirk Jumpertz, Peter Janssen, Marc Van Wesemael, Frank Piessens, Wouter Joosen, and Lieven Desmet. 2017. Exploring the ecosystem of malicious domain registrations in the. eu TLD. In *International Symposium on Research in Attacks, Intrusions, and Defenses*. Springer, 472–493.
[63] Chen Zhang. 2014. Malicious base station and detecting malicious base station signal. *China Communications* 11, 8 (2014), 59–64.
[64] Zhou Zhuang, Xiaoyu Ji, Taimin Zhang, Juchuan Zhang, Wenyuan Xu, Zhenhua Li, and Yunhao Liu. 2018. FBSleuth: Fake Base Station Forensics via Radio Frequency Fingerprinting. In *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*. ACM, 261–272.

A CONFUSION MATRIX OF SVM MODEL

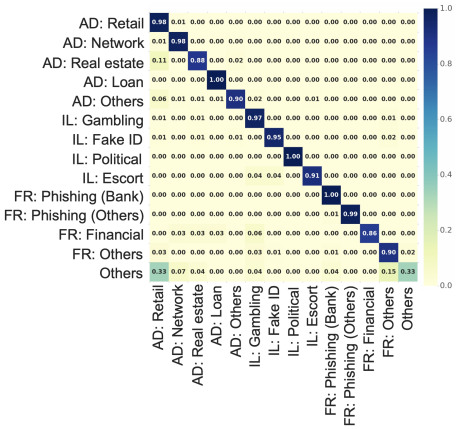


Figure 11: Confusion Matrix of SVM on the Ground-truth Dataset.

B DISTRIBUTION OF SPAM VICTIMS

Besides FBS spam messages, the security application maintained by our industrial partner also could be used to detect common spam messages generated from other channels.

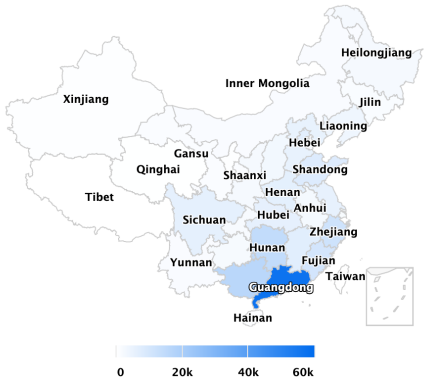


Figure 12: Daily Distribution of Non-FBS Victims Observed by Our Data Collection Software.

We provide the average daily distribution of common spam victims for each province in Figure 12. The data collection period is consistent with our FBS data collection time. Compared with Figure 6, it could be concluded that the FBS spam has unique spatial characteristics, as discussed in Sec 5.