Skudlark, Ann

Conference Paper

Characterizing SMS spam in a large cellular network via mining victim spam reports

20th Biennial Conference of the International Telecommunications Society (ITS): "The Net and the Internet - Emerging Markets and Policies" , Rio de Janeiro, Brazil, 30th-03rd December, 2014

Provided in Cooperation with:
International Telecommunications Society (ITS)

Suggested Citation: Skudlark, Ann (2014) : Characterizing SMS spam in a large cellular network via mining victim spam reports, 20th Biennial Conference of the International Telecommunications Society (ITS): "The Net and the Internet - Emerging Markets and Policies" , Rio de Janeiro, Brazil, 30th-03rd December, 2014, International Telecommunications Society (ITS), Calgary

This Version is available at:
https://hdl.handle.net/10419/106899

Characterizing SMS spam in a large cellular network via mining victim spam reports

Ann Skudlark, AT&T Labs, skudlark@att.com
2600 Camino Ramon
San Ramon, CA 94583

The author wishes to thank Yu Jin and Nan Jiang who contributed the analytics for this paper.

Abstract

In this paper[1] a study of SMS messages in a large US based cellular carrier utilizing both customer reported SMS spam and network Call Detail Records (CDRs) is conducted to develop a comprehensive understanding of SMS spam in order to develop strategies and approaches to detect and control SMS spam activity. The analysis provides insights into content classification of spam campaigns as well as spam characteristics based on sending patterns, tenure and geolocation.

[1] This paper is concerned with SMS spam sent from mobile devices via a Subscriber Identification Module (SIM) card to a mobile device, not SMS spam generated from email to a mobile device which is typically captured via filters at email gateways.

1. Introduction

The explosion of mobile devices in the past decade has brought with it an onslaught of unwanted SMS (Short Message Service) spam [1]. It was reported that the number of spam messages in the US rose 45% in 2011 to 4.5 billion messages [2]. Furthermore a 2012 Pew Research Center study reported more than 69% of mobile users have received text spam [3]. The sheer volume of spam messages not only inflict an annoying user experience, but also incur significant costs to both cellular carriers and customers alike. Due to the proliferation of unwanted messages, SMS spam may be compared to Email spam. However, in contrast to email spam where the number of possible email addresses is unlimited, SMS spammers can more easily reach victims by, e.g., simply enumerating all numbers from the finite phone number space. This, combined with wide adoption of mobile phones, makes SMS a medium of choice among spammers. Furthermore, the increasingly rich functionality provided by smart mobile devices also enables spammers to carry out more sophisticated attacks via both voice and data channels, e.g., using SMS spam to entice users to visit certain websites for product advertisement or other illicit activities.

Despite the importance and urgency of the SMS spam problem and its wide impact on cellular networks, the scarcity of representative spam datasets makes network-wide SMS spam studies a rather challenging task. The volume of SMS messages makes it difficult to collect SMS messages (with content) inside cellular networks. Meanwhile, what lacks is a reliable automated approach to differentiate spam messages from legitimate ones. For these reasons, existing research focuses primarily on building content-based SMS spam filtering at end user devices, e.g., [4,5], as opposed to studying large-scale SMS spam across the entire network. Though anonymized SMS records can be employed to characterize the network behaviors of individual phone numbers that initiate spamming, e.g., without the spam message content, it is difficult to correlate these spam numbers so as to understand how different phone numbers collaborate to launch large scale spam campaigns.

To circumvent this challenge, in this paper, a novel data source is employed – user (victim) generated spam reports (a.k.a. victim spam reports or spam reports in short) – to study SMS spam in a large cellular network. As a means to combat SMS spam, many cellular network carriers have adopted and deployed an SMS spam reporting mechanism for mobile users. In

particular, once receiving a spam message, a victim can report it via a text message forward. Cellular providers can then investigate and confirm the reported spam and restrict the offending spam phone numbers. Such victim spam reports not only contain the entire spam text, but they represent a reliable and clean source of SMS spam samples, as spam messages contained in spam reports have been vetted and classified by mobile users.

In addition to detecting spammers, the content, as reported by the spam victims, also serves as a valuable asset to understand spammers' approaches and strategies. Taking advantage of this SMS spam reporting mechanism, a year of spam reports was collected from one of the largest cellular carriers in the US which contains approximately 543K spam messages – and an extensive and multi-facet analysis of SMS spamming was carried out using these messages. The research objectives are three-fold: 1) to devise an effective approach for identifying large-scale SMS spam campaigns which are initiated collaboratively by many offending phone numbers; 2) to assess the scale and impact of today's SMS spam campaigns in large cellular networks; 3) to infer the intents and strategies of spammers behind these spam campaigns in order to develop mitigation approaches.

2. Background and datasets

In this section the SMS architecture of the cellular network under study is introduced. Then the dataset collected from this network for the analysis is introduced.

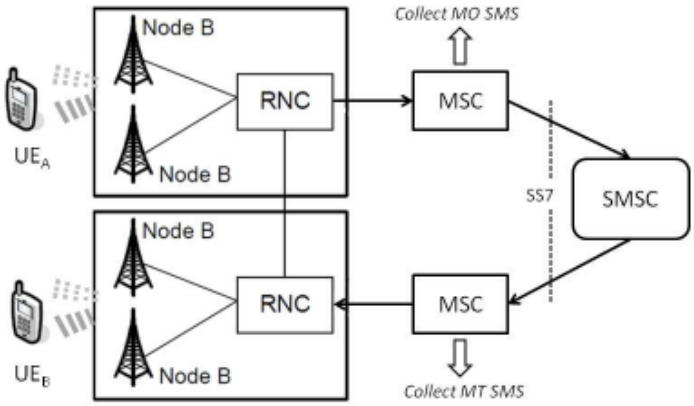

Figure 1: SMS architecture in UMTS networks.

2.1 User spam report dataset

The cellular network under study utilizes primarily UMTS (Universal Mobile Telecommunication System), a popular 3G mobile communication technology adopted by many mobile carriers across the globe. The (high-level) architecture for delivering (text-based) SMS messages inside a UMTS network is depicted in Fig. 1. When sending an SMS message, an end user equipment (UE*A*) directly communicates with a cell tower (or node-B), which forwards the message to a Radio Network Controller (RNC). The RNC then delivers the message to a Mobile Switching Center (MSC) server, where the message enters the Signaling System 7 (SS7) network and is stored temporarily at a Short Message Service Center (SMSC). From the SMSC, the message will be routed to the serving MSC of the recipient (UE*B*), then to the serving RNC and Node-B, and finally reach UE*B*. [2]

The said cellular service provider deploys an SMS spam reporting service for its users: when a user receives an SMS text and deems it as a spam message, s/he can forward the message to a spam report number designated by the cellular service provider. Once the spam is forwarded, an acknowledgment message is returned, which asks the user to reply with the spammer's phone number (referred to as the spam number hereafter). Once the above two-stage process is completed within a predefined time interval, a spam record is created. The dataset used in the study contains spam messages reported by users over a one-year period (from June 2011 to May 2012). The dataset contains approximately 543K complete spam records. Each spam record consists of four features: the spam number, the reporter's phone number, the spam forwarding time and the spam text content.

Spam Number Extraction. During a one year observation period, a phone number can be deactivated, e.g., abandoned by users or shut down by cellular providers, and can be recycled after a predefined time period. In other words, a phone number can be owned by some users for legitimate communication and by some others for launching SMS spam. To address this issue, the service plans of the phone numbers and their service starting times and ending times were identified to help to uniquely identify each phone number. For example, even with the same 10-digit sequence, a phone number which has a service plan that ends in January and is reopened in May will be counted as two different numbers in these two months. Hereafter this definition will

[2] This study of SMS messages are those send through the SS7 – Signaling System 7 network, as opposed to messaging services which deliver content through data channels such as iMessage and WhatsApp.

be used to identify spam numbers. The one-year user-generated spam reports contain a total of 78.8K spam numbers.

2.2 SMS spam call records

To assist the analysis of spamming activities from multiple dimensions, the author also utilizes the SMS (network) records – SMS call records – associated with the reported spam numbers over the same one year time period. These call records are collected at MSCs primarily for billing purposes: depending on the specific vantage point where call records are collected, there are two types of call records (see Fig. 1): whenever an SMS message sent by a user reaches the SS7 network, a Mobile Originating (MO) record is generated at the MSC serving the sender (even when the terminating number is inactive); once the recipient is successfully paged and the message is delivered, a Mobile Terminating (MT) record is generated at the MSC serving the recipient. Unlike the user-generated SMS spam reports, these SMS call records do not contain the content of SMS messages. Instead, they contain only limited network related information such as the SMS sending time, the sender's and receiver's phone numbers, the serving cell tower and the device International Mobile Equipment Identity (IMEI) number for the sender (in MO records) or receiver (in MT records). Using SMS spam numbers from spam reports; all call records associated with these numbers were extracted during the same one-year period, and were used to study the network characteristics of spam numbers and spammers. Since all the spam numbers are inside the cellular network under study, only MO calls were utilized for the studies, which cover the complete spamming history of each spam number.

No customer private information is used for the study, as all customer identities have been anonymized before any analysis is carried out. In particular, for phone numbers, only the area code (i.e., the first 3 digits of the 10 digit North American numbers) is kept and the remaining digits are shifted by adding a large prime number. Similarly, for the IMEI numbers associated with mobile devices, only the first 8-digit Type Allocation Code (TAC) is retained in order to identify device types and hash the remaining 8-digit to preserve customers' privacy. In addition, to adhere to the confidentiality under which the data was provided, in places only normalized views of results are presented while retaining the scientifically relevant magnitudes.

2.3 Related work

In a related study [6], the researchers characterized the network behaviors of individual spam numbers in a large cellular network. In comparison, this work focuses on characterizing large-scale SMS spam campaigns that are launched by multiple spam numbers. Although network-level analysis of SMS spam is conducted, the purpose is to infer the intentions and strategies adopted by SMS spammers, and to identify and explain the correlation among different spam numbers. Other work on SMS spam focused primarily on content-based spam filters at mobile devices, especially using machine learning techniques to filter spam [4, 5].

This work is also related to email spam analysis and detection. [7–10] analyzed various aspects of spam campaigns conducted by botnets. [11] studied the network behaviors of spammers, and the findings are applied to develop Spam-Tracker to identify spammers based on their sending behaviors [12]. A methodology is employed similar to [7, 13] and the analysis reveals unique characteristics of SMS spam campaigns, such as their target selection strategies, etc.

Furthermore, the author has analyzed various aspects of SMS related spam [14, 15] and developed approaches to detect SMS spam [16].

3. Methodology

Similar to email spam campaigns, spammers launch various SMS spam campaigns to achieve specific goals; each SMS spam campaign may comprise of a number of distinct spam activities. Borrowing a similar idea used in email spam campaign studies [7, 13], in this paper the URLs contained in SMS spam messages are used to track and group spam messages into spam campaigns, and perform content-based analysis to identify distinct spam activities in each spam campaign. In this section the motivation and a brief overview of the basic methodology developed for SMS spam analysis is provided. Expanding on this basic methodology, Section 3.2 describes how URLs contained in spam messages are used to track and group individual spam messages into various spam campaigns, and in Section 3.3 presents the content-analysis method that is employed to cluster spam messages within each spam campaign into distinct spam activities.

3.1 Methodology overview

As mentioned earlier, in this paper SMS spam with potentially malicious intent is the key focus. SMS spam messages generated by such malicious spammers often contain a URL, with the goal to entice users to visit a malicious website so as to, e.g., obtain users' personal information. Fig. 2 provides a schematic illustration of the strategies and mechanisms often employed by malicious spammers. A malicious spammer first purchases a number of (often prepaid) mobile phone numbers (and Subscriber Identification Module cards) and spam sending devices (e.g., 3G modems). The spammer may also create one or multiple fraudulent websites with malicious advertising and other content. S/he then launches a spam campaign. Such a campaign may be conducted in multiple acts that occur at similar or different times. In each act, the spammer uses a subset of mobile devices to send out spam messages with similar content and URLs to a group of targeted users to entice users to visit the malicious site. In Fig. 2, the spammer uses three sets of mobile devices to engage four acts of SMS spamming activities with different spam messages and URLs, with the goal to entice users to visit a fraudulent website, xx.xx.xx.xx. Once a user visits the website, the victim is told that s/he has won a $1000 gift card. Upon clicking the link to claim the prize, a new page pops up, asking the victim to submit personal information (e.g., name, address, phone number, birthday, etc.) in order to receive the gift card. Through such trickeries, the spammer is able to harvest victims' private information and use it to target them for further malfeasance.

Similar to prior studies of email spam campaigns [7, 13], a spam campaign is defined as a collection of SMS spam messages which advertise, or are associated within, the same fraudulent website (as identified by its IP address). Note that the URL of the fraudulent site may not be explicitly contained in the SMS spam messages. For example, a spammer may use a URL shortening service [17] to condense or obfuscate the URL of the fraudulent site. S/he may also use multiple (often similar-looking) domain names which resolve to the same IP address, and thus point to the same website so as to circumvent DNS blacklisting. Within each SMS spam campaign, a group of spam messages with similar content are referred to as a spam activity. Fig. 2 shows four spam activities (labeled ACT 1, ..., ACT 4) to advertise the same fraudulent site. Activities 1 and 3 contain the same URL pointing to the fraudulent site, but with different spam content. Activities 2 and 4 contain different spam messages with different URLs; in particular, Activity 4 contains a shortened URL linked to the fraudulent site. The purpose of conducting

different activities is plausibly to avoid content-based spam filtering. In conducting Activities 1 and 2, the spammer employs the same set of spam phone numbers (TN 1, where TN stands for telephone number(s)), whereas for conducting Activities 3 and 4, s/he uses two different sets of spam phone numbers. The existence of multiple TN sets are often caused by cellular carriers restricting certain spam numbers, e.g., based on user spam reports, hence spammers need to invest in a different set of phone numbers to continue the campaign.

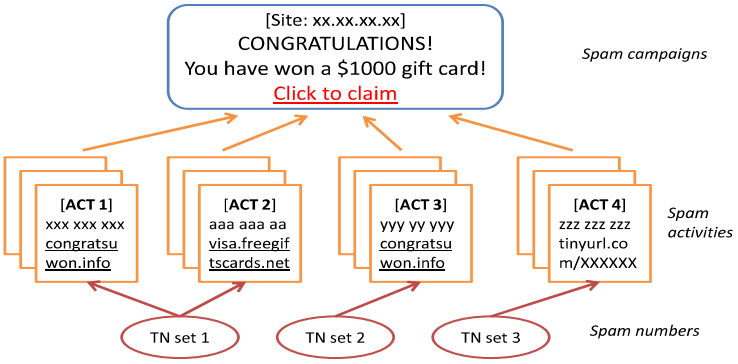

Figure 2: Entities in a typical SMS spam campaign.

Roadmap. Using the user generated SMS spam reports, a two-level SMS spam analysis is performed: first reported spam messages are grouped into spam campaigns by resolving and tracking the URLs contained in the spam messages; and within each SMS campaign, next cluster the spam messages into distinct spam activities based on content similarity. These two steps are presented in more detail in Sections 3.2 and 3.3, respectively. In Section 4 conduct an in-depth analysis of the top 10 SMS spam campaigns uncovered using the method in Section 3.2; the spam activities will also be correlated within these spam campaigns with the spam numbers used. In Section 5 use the spam numbers extracted to characterize the spam sending rates and victim target selection strategies employed by spammers. In Section 6 analyze the collective network characteristics of spam numbers that are used in each SMS spam activity, which further corroborate the correlation between spam activities and spam numbers. In Section 7 summarize the conclusions and future work.

3.2 Detecting SMS spam campaigns

Recall that the study relies on URLs embedded in spam content to identify spam campaigns. However, many URLs do not directly point to the destination site. Techniques like URL redirections and URL shortening services are commonly used by spammers possibly to reduce the message length and to avoid content-based detection. In addition, some URLs point to a survey site where manual input (to fill out a survey) is required in order to proceed to the destination site. Due to these reasons, URLs with different forms can point to the same site. There is a need to develop a technique to identify the real site behind each URL and group the spam messages accordingly. Moreover, there is also a need to address the issue that that some URLs are expired at the time of the analysis.

From the one-year spam reports, 5,249 distinct embedded URLs from the spam reports are identified. 26.7% of these embedded URLs have been shortened through URL shortening services. First these URLs are un-shortened to their original forms. Next a crawler is run to visit each URL in order to identify the destination site. 40.1% of the URLs require redirections (including redirections through URL shortening services) and 7.8% of them will be redirected more than once.

By investigating the web pages downloaded (if any) by the crawler, the crawler fails to reach the destination site under two circumstances. First, 42.3% of the URLs point to a survey site where manual input is required in order to proceed with the URL redirection. Second, for 27.1% of the URLs, the crawler may stop when it fails to resolve an expired domain name. For the former case, the manual input is supplemented in order to reach the target site. For the latter case, a best-effort approach is adopted by querying the DNS history of these expired domain names to find the IP addresses they are associated with when spam reports occurred. These IP addresses are considered to be the destination sites advertised by the URLs.

After identifying all the sites (labeled by their IP addresses) behind the URLs, spam messages are grouped pointing to the same site into a spam campaign. In this way, a total of 820 spam campaigns are found. The results are validated by comparing URLs classified as the same campaign and find that they either look alike, such as best-buy-1k.com, bestbuy-1k.com and bestbuy1k.com, or in a similar form, such as bbiy.biz, bxsy.biz and sbxt.biz. The sizes of spam campaigns in terms of the number of associated spam reports were studied to identify a few

dominant spam campaigns. In particular, the top 10 (1.2%) campaigns account for more than 46% of spam reports and the top 100 campaigns (12.2%) contribute to more than 81% of spam reports. Since small campaigns cannot provide us useful information due to the inadequacy of spam report samples, in Section 4, only the top campaigns were selected to conduct in-depth studies of their behaviors.

3.3 Identifying SMS spam activities

Recall that a spam activity is defined as a set of spam messages with URLs from the same spam campaign and have similar spam text content. Hence a text mining tool–CLUTO [18, 19]–is applied to cluster spam messages within each spam campaign into spam activities. CLUTO incorporates many different algorithms for a variety of text-based clustering problems, which have been widely applied in research domains like analyzing botnet activities [20].

Before applying CLUTO, the analyst first computes a similarity matrix for all the spam messages within each campaign, using the tf-idf term weighting and the cosine similarity function. Operating on each similarity matrix, CLUTO repeatedly selects one of the existing clusters and bi-partitions it in order to maximize a predefined criterion function. The algorithm stops when K clusters are formed. For each campaign, different choices of K are explored and select the largest K such that trivial clusters (i.e., which contain only one message) start to appear after further increasing K.

The analyst manually investigates and validates the activities identified by CLUTO. Interestingly, that spam messages within the same activity are generally similar except for one or two words. Table 1 demonstrates examples of spam messages that belong to two different activities, where the variant text content is highlighted in blue italics.

Table 1: Example messages from the same activities.

Raymond you won ... Go To apple.com.congratsuwon.info/*codeIrkfxxxxxx* *Laurence* you won ... Go To apple.com.congratsuwon.info/*codercryxxxxxx*
You have been chosen ... Goto ipad3tests.com. Enter: *68xx* on 3rd page You have been chosen ... Goto ipad3tests.com. Enter: *16xx* on 3rd page

It is suspected that such variant content is specific to each spam victim. Spammers rely on such content to distinguish and track responses from different victims and possibly get compensated according to the number of unique responses.

In the end, 2,540 spam activities that cover all the spam messages are obtained. Fig. 3 shows the number of messages (y-axis) and spam numbers (x-axis) associated with each activity. For better illustration, the marginal density plots are shown along the axes. Most of the spam activities (92%) contain multiple spam numbers and 48% can cover more than 10 spam numbers. In addition, though there is a general positive correlation between the number of spam numbers employed by the activity and the number of user spam reports, the report rate varies across activities, e.g., from 1 report per spam number to over 10 reports per number on average. As will be demonstrated in Section 5, such divergence is due to the specific spamming strategies adopted by SMS spammers.

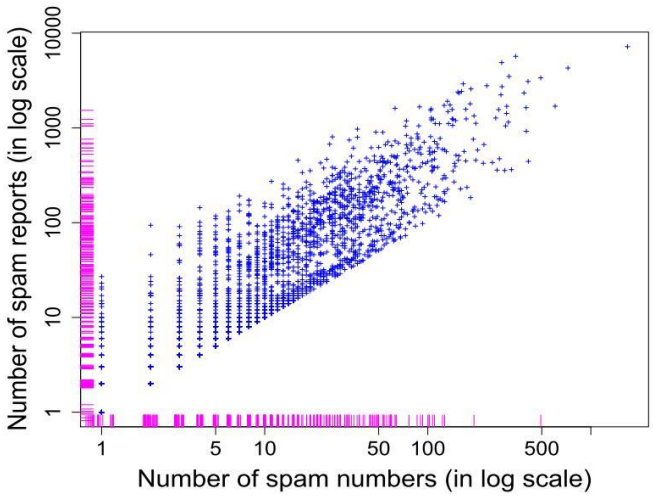

Figure 3: No. of spam reports vs. spam numbers for spam activities.

4. Spam Campaign Analysis

After clustering related spam messages based on their content, in this section, an analysis of SMS spam campaign characteristics is conducted.

4.1 Topics of spam campaigns

First the topics addressed by different spam campaigns are examined. Table 2 summarizes the top 10 topics and the proportion of spam campaigns that are associated with each topic. Note that for campaigns involved in multiple topics, the analyst labels them with the most dominant ones. Various types of phishing spam account for a large majority of all spam activities, where the top three categories cover 86.8% of all campaigns. The URLs associated with these campaigns often lead a user to a site where the user is required to enter certain private information like the phone number and home address in order to claim the gift cards or free smartphones or apply for cash loans. The remaining categories are primarily related to advertising spam, where the spam messages contain advertisement for dating sites, prescription drugs, insurance plans, products, jobs, etc.

Table 2: Spam campaign topics.

Category	Pct. (%)
Offer cash advance services	43.8
Offer a gift card to claim	29.0
Provide popular mobile devices for testing	14.0
Subscribe to services with monthly charges	4.0
Bank related scam, e.g., bank account phishing	1.3
Offer work-from-home job opportunities	1.2
Advertise dating sites	1.1
Ask a trivial question and request for answers	0.7
Auto related scam, e.g., buy junk cars	1.0
Advertise prescribed drugs	0.5
Others (lottery, free electronics, adult sites)	3.3

4.2 Study of dominant spam campaigns

In Table 3, the top 10 SMS spam campaigns in the dataset are summarized (recall that these top 10 campaigns cover nearly half of the spam reports). Approximately 10% of spam numbers are shared by multiple spam campaigns. These spam numbers are temporarily removed before calculating the statistics.

Table 3: Top 10 dominant spam campaigns in terms of the number of spam reports.

ID	Campaign topic	# Reports	# Spam number s	# Report per number	# URLs	# Activities	# Sender locations	# Reporter locations	Duration (days)
1	Walmart gift card/free apple device test	60304	6213	9.71	269	196	38	50	223
2	Free apple product	11778	1840	6.40	35	21	20	42	321
3	Walmart and Bestbuy gift card	11124	736	15.11	28	37	29	50	116
4	Cash advance / cash loans	10941	2207	4.96	22	42	32	50	+365
5	Walmart and Starbucks gift card	9972	816	12.22	30	13	24	38	46
6	Free apple product/Walmart gift card	8251	2257	3.66	29	111	20	47	360
7	Bestbuy gift card	7243	524	13.82	19	19	20	42	66
8	Cash loan services	7168	786	9.12	44	35	29	39	+363
9	Apple device test and keep	5403	303	17.83	7	7	15	36	110
10	Cash loan services	4275	618	6.92	18	16	12	30	+363

The second column in Table 3 shows the topics of the campaigns. In accordance with the ranking in Table 2, these 10 campaigns all utilize the top 3 topics: cash advance, gift card and free device for testing. In addition, half of the campaigns 1,3,4,5 and 6 involve multiple topics to attract victims to visit the same site.

Next the campaign sizes from multiple dimensions are quantified. Column 3 and 4 show the total number of spam reports and spam numbers associated with each campaign, respectively, and column 5 displays the number of reports received by each spam number on average. Each campaign employs quite a few spam numbers, ranging from a few hundred to more than 6K. However, the number of reports per spam number varies greatly from ~3.5 (campaign 6) to almost 18 (campaign 9). As shall be seen in Section 5, such a difference is mainly due to the specific spamming strategies and spamming rates adopted by spammers. In the 6th and 7th columns, the number of URLs and activities associated with each campaign are shown. Each campaign contains many URLs and activities. However, these two quantities are not always equal to each other. Some campaigns use different content to advertise the same URL, i.e., with more activities than URLs (like 3, 4 and 6). Others broadcast the same URLs through different text content, i.e., with more URLs than activities.

The analyst investigates the distributiveness of the spam numbers involved in each campaign based on the area codes in the spam numbers. From column 8, surprisingly even the smallest campaign employs spam numbers from 12 states and the largest one can cover spam numbers from 38 states. Since a spammer needs to physically connected to the network to launch spam, the wide distributiveness of spam numbers implies that these spam campaigns are

launched collaboratively by different spammers across the country. It is conjectured that these spammers are possibly hired by the owner of fraud sites to advertise the site, and are rewarded based on the number of victims attracted by the spam to visit the site. To further validate this hypothesis, the temporal correlation of spam activities and spam numbers is studied.

In column 9 and 10, the analyst measures the impact of spam campaigns in terms of their victim distributiveness and duration, respectively. Column 9 shows the total number of states of the US where the spam reporters reside based on phone number area codes. These campaigns receive complaints from mobile users from at least 30 US states. From column 10, these spam campaigns are long-lasting. The campaign duration as the time interval from the first report to the last one observed in the dataset is calculated. Note also that "+" sign indicates the first report came in the first month of observation (i.e., June 2011). Also, the analyst observes reports for all these 10 spam campaigns in the last month of the observation period (i.e., June 2012). In other words, these spam campaigns are still active and the durations here only serve as underestimates of their real life span. Notwithstanding, the shortest campaign still lasts for more than one month and half of the campaigns have a duration of approximately one year. These long spamming periods result in a wide impact on the cellular network and mobile users.

4.3 Correlation of spam activities and spam numbers

Spam numbers or activities are considered to be correlated if they initiate SMS spamming at similar times. Arrival times of spam reports are used as a proxy to measure the correlations of spam activities and spam numbers. Intuitively, temporally correlated spam activities/numbers will have accompanying user reports at similar times, since most of the reports come within one day after the spam occurs.

The spam report arrival times for different spam activities corresponding to the largest spam campaign (campaign 1 in Table 3) are illustrated in Fig. 4. Similarly, the arrival times for spam numbers from the top 10 largest activities of campaign 1 in Fig. 5. The x-axis represents the days during the one-year observation period, and the y-axis stands for each spam activity or spam number. A point indicates that y were reported by users during day x. Only the largest campaigns are shown due to space limit. Similar observations apply for other campaigns.

In Fig. 4, there is significant divergence among spam activities participating in the same campaign. First of all, the starting times vary greatly across activities, with even month long gaps

between each other. In addition, some activities attract persistent user complaints, and thereby displaying horizontal bars, while others contain noticeable temporal clusters user reports (displayed as intermittent horizontal bars), with noticeable gaps between clusters. These clusters are associated with similar spam text content but different domain names. This is likely because the old domain names are blocked in the middle of the activity, and the spammer switches to a different domain name pointing to the same site. In most cases, there is strong similarity of these domain names, such as ipad3tests.mobi, ipad3tester.info and ipad3winner.co, etc. In comparison, spam numbers within the same activity are in accordance with each other. Their related user reports often arrive at the same time, thereby exhibiting clear vertical bars in Fig. 5. Such vertical bars often appear at beginning of a spam activity and last persistently for a few days.

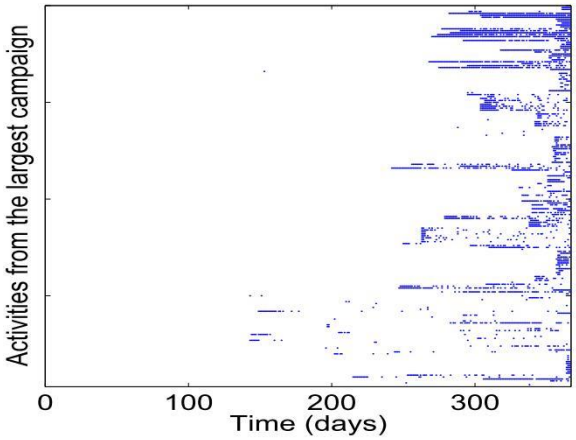

Figure 4: Temporal correlation of spam activities.

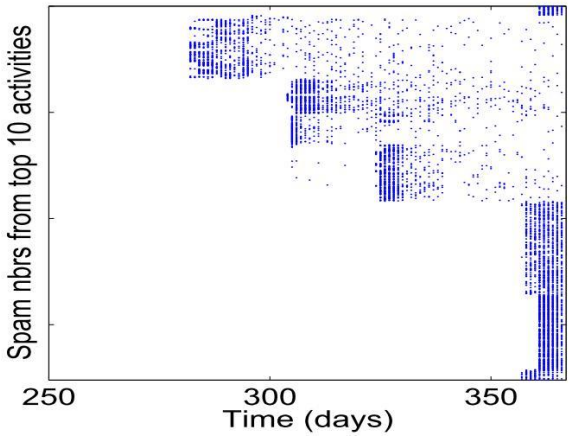

Figure 5: Temporal correlation of spam numbers.

In summary, from user spam reports, interesting properties of large-scale SMS spam campaigns in cellular networks are identified, which also sheds light on the organization of SMS campaigns and spammer behaviors. In the following, the author takes one step further to study the strategies adopted by spammers for launching the spam, e.g., how do they choose targets, and their impact on the cellular network. To achieve these goals, SMS call records for the studies are incorporated, which contain complete SMS sending history of individual spam numbers. Using such SMS network datasets, the author wants to find evidence to corroborate the observations of the correlation of spam numbers and to develop more effective spam number detection methods based on the findings.

5. Characterizing spam numbers

5.1 Spam sending rate

The SMS spamming rate using the average number of SMS messages sent from each number per hour are measured. The variability of spamming rates is assessed using the coefficient of variation, which is defined as cv = σ /μ, where σ and μ represent the standard deviation and mean spamming rate of each spam number. The coefficient of variation shows the extent of variability relative to the mean sending rate. Fig. 6 displays the mean spamming rate and the corresponding coefficient of variation for individual spam numbers. The spamming rate varies from a few spam messages to over 5,000 spam messages per hour. In addition, while the majority of spamming activities are at a constant rate (i.e., with a low cv close to the x-axis), some numbers exhibit more bursty spamming behaviors, i.e., with a cv greater than 3. From these two metrics, three distinct regions are observed, which are referred to as "slow," "moderate," and "fast" spammers (from left to right in Fig. 6). "Moderate" spammers cover 63% of all spam numbers, while "fast" spammers and "slow" spammers account for 20% and 17%, respectively.

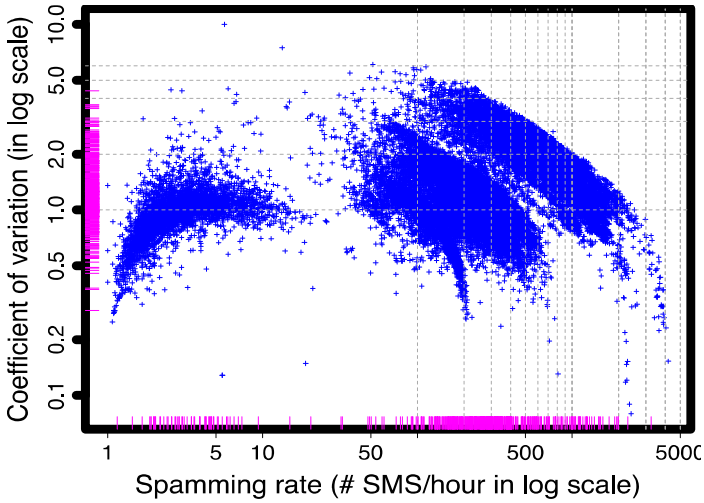

Figure 6: Rate and variability.

Further investigation shows that the spamming rates generally depend on the devices used and the network location of the spammers. For example, for the top 10 campaigns in Table 3, the ones with a large number of reports per spam number are often associated with fast

spamming devices which allow sending spam aggressively at a high rate (e.g., a SIM card bank); in contrast, the other campaigns often employ moderate-rate devices like laptop cards. With the same amount of time, the former spam numbers can reach far more victims, which consequently leads to more user spam reports.

5.2 Target selection strategy

The next section shows how spammers select spamming targets. Given the fact that each phone number is a concatenation of two components: the 3-digit area code, which (except for number portability) is location specific, and for the 7-digit subscriber number the analyst characterizes the target selection strategies at two levels, i.e., how spammers choose area codes and phone numbers within each area code.

Figure 7 plots the area code relative uncertainty (the x-axis) and the random spamming ratio (the y-axis) for individual spam numbers. For ease of visualization, the marginal densities along both axes are illustrated.

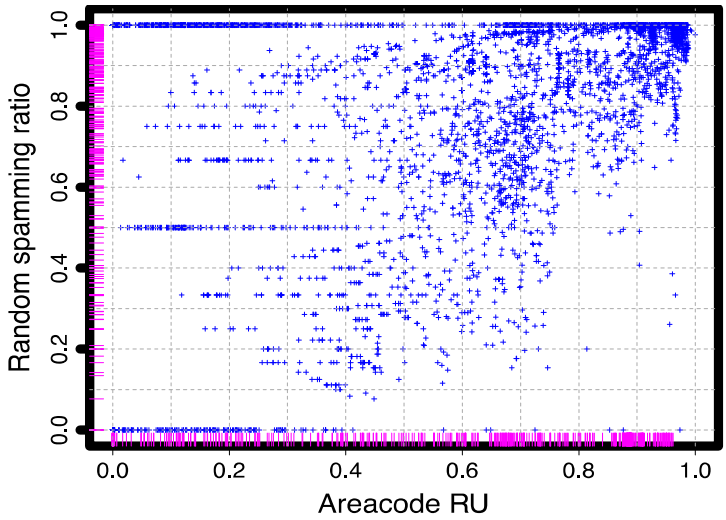

Figure 7: Target selection strategies.

Based on the marginal density of the area code relative uncertainty the majority of spam numbers (78% using a 0.8 as a cut-off threshold) concentrate on phone numbers within certain area codes. This spamming strategy is referred to as target spamming. In comparison, the remaining 22% spam numbers adopt a random spamming strategy. The analyst ranks the area

codes by their popularity among spam numbers, i.e., how many spam numbers select the most target numbers from a particular area code. In fact, the top 20 popular area code among spammers are investigated and find that most of them correspond to large cities and metro areas, e.g., NYC (with 3 area codes), Chicago (2), LA (2), Atlanta, and so on.

Based on the y-axis, no matter how a spam number chooses area codes, a predominant portion of them select targets randomly within each area code. This explains why spammers favor large metro areas, because they are likely to reach more active mobile users by randomly selecting phone numbers within these area codes.

5.3 Spamming locations and impact on the cellular network

This section is ended by an assessment of the sending locations of spam messages and the potential impact of spamming traffic on the cellular network. The location of a spam number is defined as the serving node-B from which a spam message is sent from that spam number. In Fig. 8, for each spam number the total number of node-B's (the x-axis) vs. the proportion of spam messages from the most dominant node-B (the y-axis) is illustrated.

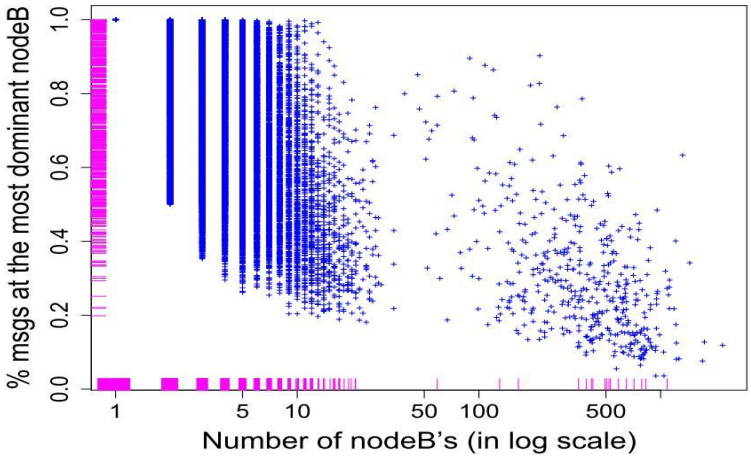

Figure 8: Spamming locations.

Same as the observation from [6], there are a few spam numbers (4.9%) which are highly mobile, i.e., they utilize more than 10 node-B's and distribute their workload among these node-B's (i.e., with the proportion of spam messages from the most dominant node-B less than 40%). However, most spam numbers initiate spam at less than 5 node-B's (78.2% spam numbers) and the most dominant node-B carry more than 60% of the traffic (74.5%). These dominant node-B's

are referred to as the primary spamming locations for spam numbers. The primary spamming node-B's are mapped to their geographical locations. 19 out of the top 20 popular primary spamming locations are located at large metro areas, e.g., 11 from LA, 3 from NYC and 3 from Miami, etc. At these node-B's, the sheer volume of spamming traffic is astonishing. The spamming traffic can exceed normal SMS traffic by more than 10 times. Even at the RNC's, which serve multiple node-B's, the traffic from spamming may account for 80% to 90% of total SMS traffic at times. Such a high traffic volume from spammers can exert excessive loads on the network, affecting the performance of non-spam SMS traffic. Furthermore, since SMS messages are carried over the voice control channel, excessive SMS traffic can deplete the network resource, and thus can potentially cause dropped calls and other network performance degradation.

6. Detecting Spam Numbers

Using the seed list of victim SMS spam reports and studying the characteristics of spam campaigns, it is observed that not all spam activity was uniform, and as a result different methods were needed to detect different types of SMS spam. The objective was to have 0 false positives to avoid incorrectly identify a legitimate SMS sender, thus detection methods were set conservatively. One approach was to combine >1 victim complaints with X messages to Y recipients over T period of time. This approach combines both complaints and volumetrics.

Another detection method was based on the observation that spammers employ multiple spam numbers for the same activity, once a spam number is confirmed, e.g., by user spam reports, other spam numbers that are temporally and spatially correlated with the confirmed number can be detected. A two-step detection algorithm was implemented. First, all SMS senders in the network are monitored and a watch list of phone numbers in different geolocations (node-Bs) that have sent SMS messages to more than Y recipients over T interval is maintained. Second, detection is triggered once a spam report arrives to confirm a particular spam number in the watch list, and look for other numbers from the same watch list whose spamming locations (i.e., node-Bs) are the same as the confirmed one [14].

A third detection method is based on understanding spamming target selections. Since many SMS spammers adopt random target selection strategies, mobile users (within the same

area code) have the same exposure to spam. Using device (IMEI) and mobile plan information SMS recipients are identified who do not normally send or receive SMS – which defined as "Grey Phone Space" [16]. Using an algorithm and a set of parameters candidate SMS spammers based on sending to Grey Space numbers are identified.

The three detection methods described above are focused on domestic SMS. However, there are also malicious Mobile Originating spammers who target international victims through unlimited international county text plans – primarily prepaid plans. Since international victim spam reports are not available, volumetrics (X messages to Y recipients over T period of time) is the principle detection method.

7. Conclusion and future work

In this paper, an extensive analysis of SMS spam was conducted using user reported spam messages and network-level SMS call records collected from a large cellular network. A two-level text clustering method was proposed to identify spam campaigns and activities from spam reports and studied interesting properties of representative SMS spam campaigns, which can last for months and have a wide impact the network. Assisted with call records, spamming strategies adopted by spammers were analyzed and inferred. As a result of understanding SMS spammer behavior spam detection algorithms were developed. Future work for this project involves continuing to study spammer behavior in order to tweak current algorithms based on changes in behavior, and to develop methodologies as spammer behaviors evolve.

References

[1] Wikipedia: mobile phone spam. http://en.wikipedia.org/wiki/Mobile phone spam

[2] Mobile spam texts hit 4.5 billion. http://www.businessweek.com/news/2012-04-30/mobile-spam-texts-hit-4-dot-5-billion-raising-consumer-ire.

[3] 69% of Mobile Phone Users Get Text Spam. http://pewinternet.org/Media-Mentions/2012/69-of-Mobile-Phone-Users-Get-Text-Spam.aspx

[4] K. Yadav, P. Kumaraguru, A. Goyal, A. Gupta, and V. Naik. Sms assassin: Crowd sourcing driven mobile-based system for sms spam filtering. HotMobile '11, 2011.

[5] G. Cormack, J. Hidalgo, and E. S´anz. Feature engineering for mobile (sms) spam filtering. SIGIR ’07, 2007.

[6] I. Murynets and R. Jover. Crime scene investigation: SMS spam data analysis. IMC’12, 2012.

[7] A. Pathak, F. Qian, C. Hu, M. Mao, and S. Ranjan. Botnet spam campaigns can be long lasting: evidence, implications, and analysis. SIGMETRICS ’09, 2009.

[8] C. Kreibich, C. Kanich, K. Levchenko, B. Enright, G. Voelker, V. Paxson, and S. Savage. On the Spam Campaign Trail. In LEET’08, 2008.

[9] C. Kreibich, C. Kanich, K. Levchenko, B. Enright, G. Voelker, V. Paxson, and S. Savage. Spamcraft: An inside look at spam campaign orchestration. In LEET’09, 2009.

[10] C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. Voelker, V. Paxson, and S. Savage. Spamalytics: An empirical analysis of spam marketing conversion. Communications of the ACM, 52(9):99–107, 2009.

[11] A. Ramachandran and N. Feamster. Understanding the network-level behavior of spammers. SIGCOMM ’06, 2006.

[12] A. Ramachandran, N. Feamster, and S. Vempala. Filtering spam with behavioral blacklisting. CCS ’07, 2007.

[13] Y. Xie, F. Yu, K. Achan, R. Panigrahy, G. Hulten, and I. Osipkov. Spamming botnets: signatures and characteristics. SIGCOMM ’08, 2008.

[14] N. Jiang, Y. Jin, A. Skudlark, and Z.-L. Zhang. Understanding SMS spam in a large cellular network. SIGMETRICTRICS ’13, 2013.

[15] N. Jiang, Y. Jin, A. Skudlark, and Z.-L. Zhang. Understanding SMS spam in a large cellular network: Characteristics, strategies and defenses. RAID’13, 2013.

[16]N. Jiang, Y. Jin, A. Skudlark, and Z.-L. Zhang. Greystar: Fast and accurate detection of SMS spam numbers in large cellular networks using gray phone space. USENIX SEC’13, 2013.

[17] Tinyurl.http://tinyurl.com/.

[18] Cluto -software for clustering high-dimensional datasets. http://glaros.dtc.umn.edu/gkhome/views/cluto.

[19] Y. Zhao, G. Karypis, and U. Fayyad. Hierarchical clustering algorithms for document datasets. Data Min. Knowl. Discov., 2005.

[20] G. Jacob, R. Hund, C. Kruegel, and T. Holz. Jackstraws: picking command and control connections from bot traffic. SEC'11, 2011.