Understanding Cross-Channel Abuse with SMS-Spam Support Infrastructure Attribution

Bharat Srinivasan[1], Payas Gupta[2], Manos Antonakakis[1], and Mustaque Ahamad[1,2]

[1] Georgia Institute of Technology, Atlanta, USA
bharat.srini@gatech.edu, manos@gatech.edu, mustaq@cc.gatech.edu
[2] New York University, Abu Dhabi
payasgupta@nyu.edu

Abstract. Recent convergence of telephony with the Internet offers malicious actors the ability to craft cross-channel attacks that leverage both telephony and Internet resources. Bulk messaging services can be used to send unsolicited SMS messages to phone numbers. While the long-term properties of email spam tactics have been extensively studied, such behavior for SMS spam is not well understood. In this paper, we discuss a novel SMS abuse attribution system called CHURN. The proposed system is able to collect data about large SMS abuse campaigns and analyze their passive DNS records and supporting website properties. We used CHURN to systematically conduct attribution around the domain names and IP addresses used in such SMS spam operations over a five year time period. Using CHURN, we were able to make the following observations about SMS spam campaigns: (1) only 1% of SMS abuse domains ever appeared in public domain blacklists and more than 94% of the blacklisted domain names did not appear in such public blacklists for several weeks or even months after they were first reported in abuse complaints, (2) more than 40% of the SMS spam domains were active for over 100 days, and (3) the infrastructure that supports the abuse is surprisingly stable. That is, the same SMS spam domain names were used for several weeks and the IP infrastructure that supports these campaigns can be identified in a few networks and a small number of IPs, for several months of abusive activities. Through this study, we aim to increase the situational awareness around SMS spam abuse, by studying this phenomenon over a period of five years.

1 Introduction

The telephony channel has undergone radical changes in the recent past, including its *convergence with the Internet* via technologies such as smartphones and Voice over IP (VoIP). Although this convergence offers many benefits, it also provides malicious actors the ability to design new attack vectors that combine resources from both the telephony and Internet channels. For instance, text messages containing web links can be sent to phone numbers to direct unsuspecting users to malicious websites [19]. Attacks that exploit the telephony channel can potentially be more effective than traditional attacks over the Internet, as they can abuse the trust that has traditionally been associated with telephony. Similar to traditional email messaging, SMS [18] has become a popular abuse target, as past research efforts have shown [38, 36, 35, 30].

While traditional email spamming activities have been extensively studied, long-term properties of SMS spam operations are not well understood by the community. SMS abuse data and long-term network traffic observation of such abuse are necessary to study

the behavior of SMS spam operations. By using data that spans a period of close to five years, in this study we aim to present such a long-term analysis of SMS spam abuse. Our hope is that such analysis will provide better understanding of the network properties of SMS spam abuse which can be used to build more effective defenses against it.

We call SMS spam *cross-channel abuse* because it relies on and can be observed in both the telephony and Internet channels. In other words, such attacks involve both a telephony resource (e.g., a phone number) and a traditional Internet resource (i.e., a domain name and/or an IP address). To study cross channel abuse, we explore how SMS spam campaigns utilize the domain name system (DNS) and other Internet infrastructure. We build a SMS spam attribution system called CHURN, which is used to analyze abuse data from a period of five years. CHURN analyzes SMS-spam datasets from two different abuse reporting sources: passive DNS datasets from a large Internet Service Provider (ISP), and application layer web information around these SMS spam campaigns. CHURN's ultimate goal is the attribution of SMS spam campaigns with respect to the domain name infrastructure they employ in their abuse activities.

Our SMS spam attribution analysis reveals that cross channel abuse is highly effective and long lived. We found that the Internet IP infrastructure used by the spammers to support SMS spam campaigns is surprisingly stable. For example, abuse campaigns tend to use a handful of IPs in a few networks over several years to continue their activities. This shows current defenses are either unaware of the abuse infrastructure utilized by SMS spam campaigns or they are not effectively using such information to combat cross-channel abuse. We hope that our paper will demonstrate the value of situational awareness around this problem, which could be used to reduce the potential for social engineering and other attacks facilitated through such cross channel abuse. Summarizing, our paper makes the following contributions:

- We build and present a cross-channel attribution system to automate the collection and analysis of SMS spam abuse. Our system, namely CHURN, uses a hierarchical clustering technique that employs network level, application level, and popularity-based statistical features to cluster related SMS spam domain names into campaigns over time.
- Using CHURN, we conduct a five year study that yields attribution results for a plethora of real world SMS spam campaigns. We use (1) 8.32 million SMS abuse reports that consist of messages that directed users to scam websites, (2) more than 56 thousand DNS resource records related to the SMS abuse reports since 2011, and (3) more than 67 thousand web pages reflecting the application layers of the SMS spam campaign. Our experiment helps us conclude the following:
  - We show that a mere 1% of SMS abuse domains appear on public Internet domain blacklists. Among the blacklisted domain names, 94% appeared on blacklists weeks or even months after they were first seen in abuse reports.
  - We show that the domains are long lived during the period of abuse with over 40% of the SMS spam domains being active for over 100 days.
  - We dive deep into the three largest and most long-lived case studies of SMS spam campaigns identified by CHURN. We show that (1) spammers were able to operate these campaigns for more than three years, (2) they consistently used a handful of IPs in a few abuse friendly networks, and (3) the average SMS spam domain name lifetime was in the order of two months, further emphasizing the lack of situational awareness around such cross-channel threats.

2 Background

Spammers have been evolving their operations for more than a decade. It comes as no surprise that as Internet defenses are bolstered, the telephony channel has become an attractive spam target. To better understand this, we aim to study the properties of unsolicited bulk SMS messaging (a.k.a. SMS spam) containing URLs with respect to the Internet infrastructure that supports this abuse. Since the attack relies on both telephony and Internet infrastructure (e.g., domains included in SMS spam URLs and associated IPs), we refer to this problem as "cross-channel abuse". In this section, we provide a high-level overview of the ecosystem that facilitates this cross channel abuse.

Delivering SMS Spam at Scale: To successfully "trick" users into scam operations, spammers need a way to reach potential victims. Because phone numbers come from a limited name space with a defined format, they can be auto-generated randomly or picked selectively. Armed with phone numbers, fraudsters can accomplish large scale distribution of SMS spam in several ways.

1. **Disposable SIMs:** Spammers can purchase disposable subscriber identification module (SIM) cards with gateways having slots to hold hundreds of them or use stolen cell phones and USB modems/Aircards [38] as an entry point into the cellular networks. They can then program these devices using off the shelf bulk SMS software or even Arduino [23] micro-controllers to send well crafted bulk SMS spam.
2. **Exploiting Cloud Telephony Services:** Legitimate cloud telephony Infrastructure as a Service (IaaS) providers such as Twilio [21] and Tropo [20], or even cellular ISPs [38], can be abused by spammers to deliver bulk SMS messages. This is achieved in one of three ways: (1) creating fraudulent accounts on these platforms, (2) hijacking existing (legitimate) accounts, or (3) exploiting unprotected SMS application programming interfaces (APIs) that allow users to transmit a large volume of SMS messages in an automated fashion[3].
3. **Bulk SMS Services:** Spammers can exploit or collude with existing bulk SMS services to deliver messages. Sometimes, services offered by legitimate service providers enable bridging of the email and SMS mediums by allowing email to be sent as an SMS (or vice versa). This can be abused by spammers.

Monetization: After delivering the spam SMS messages, in order for monetization spammers lure victims into responding to, or interacting with, the message. Specially crafted messages with easy-to-click URLs provide an effective way to automate such response. On smartphone-like devices, victims can simply click these URLs and visit a traditional web site that will lure them into the scam. The key point here is that, while the attack vector clearly started as a telephony based communication (vis-à-vis, the SMS spam), these spammers will often try to social engineer the user into a scam using traditional Internet resources. There are multiple reasons to do this, from minimizing the forensic trail in the telephony network to re-utilizing already provisioned Internet infrastructure for abusive actions. Often the content of such illicit webpages can be tailored to the specific scam.

Observing Cross-Channel Abuse: Cross-channel abuse can be observed in both the telephony and Internet channels. Prior work in combating telephony abuse mainly relied on call detail records (CDRs) to identify and block phone numbers that originate spam

[3] Although Twilio and others have a policy against such abuse [22], spammers often find ways to violate it [14].

SMS messages [38, 35]. Cross-channel abuse also requires traditional Internet resources to direct victims to scam websites. This provides an opportunity to observe such communications by passively monitoring network traffic (i.e., the DNS resolutions). For example, when the recipient of an SMS message clicks an embedded link, it typically initiates a DNS resolution process. The end result of this resolution process is the mapping between the requested domain and the IP address hosting it. The client device typically requests the web page associated with the clicked link from the resolved IP address. The DNS visibility at the ISP (cellular or otherwise) recursive resolver level can serve as a great vantage point to study the SMS spam cross-channel abuse with respect to the Internet channel.

3 Cross-Channel Attribution Engine

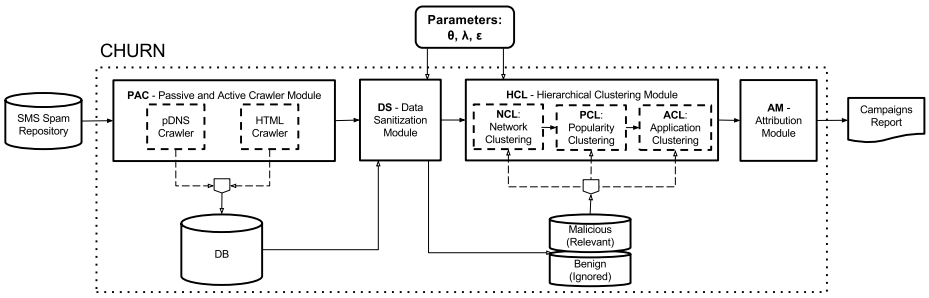


Fig. 1: The cross-channel attribution engine.

In this section, we discuss the details of our Cross Channel Attribution Engine called CHURN. The goal of CHURN is to help understand SMS abuse by attributing domain names in SMS-spam campaigns. CHURN achieves this by clustering network (i.e., domain names and IPs) and application (i.e., HTML content) layer signals that facilitate a given spam campaign. CHURN starts with crowd sourced abuse complaints and produces attributed campaigns with associated network resources. To accomplish this, it performs four tasks serially, as can be seen in Figure 1. Next, we describe in detail each of these four tasks.

3.1 Data Collection Module

Our data collection module takes as input external data source(s) of known SMS-spam. In our case, this dataset comes from two sources: (i) SMS-spam complaint reports filed by consumers to the Federal Trade Commission (FTC) [3], which were made available to participants in the Robocall Challenge [6], and (ii) publicly available SMS complaint reports from the online portal SMS watchdog [15] [4]. While reports from SMS watchdog were crawled between Jan 2011-Aug 2015, the FTC complaint records were limited to the period Jan 2011-Dec 2012 consisting of reports with anonymized destination numbers. Using SMS messages from user complaints as input, we extract the source (e.g., phone number), timestamp t_d, and URL from each SMS-spam report. Using the URLs, we actively crawl different public and private data sources, which provides information about both the website and the network hosting infrastructure facilitating the scam.

[4] smswatchdog.com was down when we last checked as on 02/18/2016 but snapshots of it can be found on the Wayback Machine [9].

Passive DNS Crawler (Network Intelligence): Cross channel attacks, like users responding to SMS-spam messages, can be observed in the Internet when the recipient of the message clicks on the URL of a spam message. In this case, a DNS resolution request will be observable at the local recursive DNS servers. This forensic signal cannot be used to estimate the global abuse properties of a particular SMS-spam campaign, as it is non-trivial to obtain global visibility in the DNS recursive plane. However, given a large enough recursive DNS visibility, it could provide forensic evidence and lower bounds on the following three questions: (i) how long was the campaign active, (ii) what was the average lookup volume and a lower bound on the victims that were targeted by each SMS-spam message, and (iii) what was the domain name and IP network infrastructure that supported this cross channel abuse?

By gaining access to a large private passive DNS repository, we were able to "crawl" and collect datasets that could answer these three questions for every domain name contained in our SMS-spam abuse dataset. As we will discuss in subsection 3.3, the passive DNS (pDNS) dataset plays an important role in our effort to statistically describe the network properties of SMS-based abuse.

HTML Crawler (Application Intelligence): We implement both an active and a passive method to collect datasets that capture application layer properties of the SMS-spam websites. We download and store the full HTML source from the web page pointed to by each URL seen in SMS-spam reports. In many cases, however, the websites of interest were taken down before we could recover any useful intelligence. For such cases, we relied on the Wayback machine [9].

3.2 DS: Data Sanitization Module

The lifecycle of a spam domain involves multiple phases. In the first phase, when the threat is active, the domain will point to IP infrastructure that facilitates the spam operation. Once the spam operation is over, or the domain simply ceases to be used by the spammers, it will enter a phase when it is "parked" or is taken down by network defenders or eventually expires. From the threat analysis and attack attribution point of view, we care to analyze the network infrastructure when the domain is actively used by a spam campaign. The goal of the sanitization module is to weed out the benign infrastructure (in the form of legitimate IP addresses) and HTML sources (related to parked domains) while retaining the network and application information that can be used to analyze the campaigns. Next, we discuss in detail how we can achieve this sanitization of the datasets.

Filtering the pDNS Datasets: Among the domains included in the URLs received in the complaints, we first remove any records containing domains historically appearing in the Alexa [2] top 1 million ranks since 2011. We were able to remove 715 domains using this filter. Next, we use two heuristics to remove DNS information that is related to legitimate IP infrastructure from our datasets. The first heuristic aims to capture the *popularity* of the infrastructure supporting a domain. Parking IP address space is often used to host a relatively large number of domains, at least that is how "domaineers" operate. The number of resource records per IP is a good measure of this as it encapsulates both the diversity in the domains and the popularity in DNS lookup value to domains hosted on certain IPs. The second heuristic aims toward the characterization of the *name server* list supporting a domain. Some name servers (NS) are well known to be

associated with parking activities, as they do not try to hide. We create a hard curated list of such name servers using publicly available information and prior work [44, 8].

More precisely, given a set of pDNS resource records denoted by RR, the sanitization module uses a filter method that uses parking IP threshold θ_p and a name server list, denoted by NS, to create a filtered set, $RR^{\theta_p,NS}$, which consists of all $rr \in RR$ s.t. (i) IP in rr is pointed to by $< \theta_p$ resource records, and (ii) the name server for the domain name d in the $rr \notin \{\text{NS}\}$. Figure 2 shows the cumulative distribution function (eCDF) of the number of resource records hosted per IP in our dataset and the cut-off threshold θ_p. In total we were able to identify ∼1% (232 out of 23,269) IPs as parking and ignore records associated with them for the shown value of θ_p.

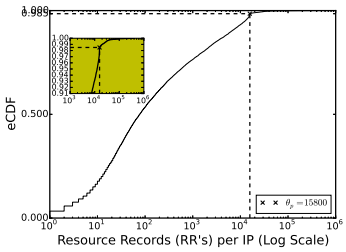


Fig. 2: CDF of Resource Records per IP with cut-off threshold θ_p.

Filtering Application-level Data: To identify the full HTML sources relating to parked domains, we built a supervised binary classifier to identify if an HTML source file was related to a parked domain or not. To train our classifier, we used 20 features extracted from HTML sources. These features included number of links in the source, number of unique domains in the links, minimum, maximum and average link length, number of external links, ratio of internal to external links, website directory presence, source length, text to html ratio based on the number of characters, presence of Javascript redirect and meta refresh redirection mechanisms, boolean value for if the meta domain was external, number of frames and iframes and respective number of distinct frame and iframe domains and boolean values to indicate if any of the iframe or frame domains were external. We also counted the number of images present in the HTML source. Intuition behind these features can be found in the work by Vissers et al. [44].

We trained the SVM model [31] using the 10-fold cross validation technique on a set of 200 parking and 200 non-parking feature vectors extracted from webpages in our dataset. With a threshold of 0.5 we were able to achieve a reasonable TPR of 99.5% and FPR of 1.5%. Table 1 shows the confusion matrix using 10-fold cross validation related to this experiment, where NP denotes non-parking webpages and P denotes parking webpages. In total, the classifier was able to identify ≈10% (7510/75,085) webpages as parking. These were discarded from further processing.

	Predicted: NP	Predicted: P	Total
Actual: NP	197	3	200
Actual: P	1	199	200
Total	198	202	

Table 1: Confusion matrix for the parking classifier.

3.3 HCL: Hierarchical Clustering Module

To find clusters of related domain names associated with cross-channel abuse in a given epoch (time period, t), we follow a hierarchical clustering process. This process can be separated into three different levels. In the first level (NCL), we cluster together domain names based on the network infrastructure properties. In the second level (PCL), first level (NCL) clusters that satisfy a cardinality constraint (based on threshold λ) get further clustered according to the DNS volumetric popularity of the domains within it. In the third and final clustering step (ACL), second level (PCL) clusters that satisfy an entropy (flux) constraint (based on threshold ϵ) get further separated based on the web

content of each domain within it. This way, the entire process produces clusters of high quality at different levels which are then labeled by the attribution module (Section 3.4).

In order to execute these three different clustering steps, we employ the most common statistical features from the areas of DNS [25, 26, 27] and HTML [42] modeling. To be clear, we do not claim novelty about the use of these features. Rather, our goal is to show that already discussed features combined in this novel hierarchical clustering method can provide an efficient and effective attribution system for SMS-spam abuse. Next, we briefly discuss how we used these established statistical features in the context of the three modules of our system.

Network-based Clustering (NCL): To compute network layer features in a given time epoch t, for each domain d in the domain set $\mathcal{D}$ under consideration, we compute two sets: (i) RHIP(d) which is a set of all IPs that have historically mapped to domain d, and (ii) RHDN(IP) which is the set of domains that have historically been linked with the IP in the RHIP set. This could also include domains that are not in $\mathcal{D}$. Using the collection of all domains $\mathcal{D}$, the $pDNS$ dataset and a specified epoch t, the network feature-based clustering submodule generates a matrix $A_{m\times n}$ where $m = |\mathcal{D}|$ represents the total number of domains and $n = |\bigcup_i RHIP(d_i)|$ represents the total number of IPs historically associated with all domains in $\mathcal{D}$ during an epoch t. The matrix A is computed as follows,

$$A_{i,j} = \begin{cases} \frac{H(d_i)}{|RHDN(ip_j)|} & \text{if } ip_j \in RHIP(d_i) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $i \in \{0, 1, ... |\mathcal{D}| - 1\}$ and $j \in \{0, 1, ... | \bigcup_i RHIP(d_i)| - 1\}$. Also, $H(d) = -\sum_{k \in C(d)} p_k * log_2(p_k)$, where $C(d)$ represents the unique set of characters in domain name d and p_k represents the probability of the occurrence of a given character in the domain name. Thus, $H(d)$ gives us the entropy of the name of domain d based on relative character frequencies. The inclusion of the entropy factor in the numerator increases the confidence of producing high quality clusters given the frequent use of DGAs [28, 46] by adversaries.

Finally, we use Singular Value Decomposition (SVD) [45] to reduce the dimensionality of the sparse matrix $A_{m\times n}$ to $A_{m\times \widetilde{n}}$ where $\widetilde{n} < n$. The network clustering module then uses the X-Means clustering algorithm [40] to cluster domains having similar network-level properties.

Popularity-based Clustering (PCL): Sometimes, network level properties may be insufficient to distinguish between unrelated domains, leading to the formation of large clusters. We will see this in Section 4.2. Popularity based clustering uses features extracted from observing the popularity of domain names as measured by the number of the successful DNS resolutions to it within the epoch t. This in turn gives us a lower bound on the number of visits potentially made to the domain name via clicking on a URL embedded in an SMS message. It is computed using the information gathered in the passive DNS dataset. Let Lookup(d, dt) be a function that returns the number of lookups (or in other words, successful DNS resolutions) for domain d on a given date dt. And let C be the set of clusters produced by NCL. Using the pDNS data collection and a specified epoch t, the popularity cluster submodule builds matrices $B_{p\times q}(c_r)$ $\forall c_r \in C$ s.t. $|c_r| \geq \lambda$, $r \in \{0,1,...|C|-1\}$ where λ is a provided threshold and $|C|$ is the number of clusters produced by NCL. Here, $p = |c_r|$, the number of domains in a cluster from NCL and q are the total dates in a given epoch. The matrix B is

computed as follows, $B_{i,j}(c_r) = Lookup(d_i, dt_j)$ where d_i is a domain name and dt_j is a date in epoch t and c_r is a NCL cluster. The intuition behind this matrix follows from the work by Antonakakis et al. [26] which aims to measure the volumetric DNS request patterns to domain names over time, within a NCL cluster (in our case).

Similar to the NCL module, each matrix is dimensionally reduced using SVD followed by X-Means clustering algorithm to cluster domains having similar popularity levels. Therefore, at the end of PCL, we have: (i) smaller clusters from NCL that had sufficient network level information ($|c_r| < \lambda$), and (ii) PCL (sub)-clusters from the larger NCL clusters that required the additional popularity information for further refinement.

Application-based Clustering (ACL): To further refine and resolve any remaining confusion between domain names after PCL, we proceed to a final clustering step that aims to group together domain names with similar domain structure and web content. To cluster similar domains based on their structure, we compute the standard deviation σ of the entropy of domain names in a cluster produced after the PCL module. Let T represent the set of domains in a PCL cluster and $H(T)$ be the set of entropies associated with domain names in T. If $\sigma(H(T)) \geq \epsilon$, i.e., the standard deviation in the entropy of the domain names in the cluster is greater than the threshold ϵ, we apply application based clustering to a PCL cluster. Again, the motivation behind using entropy as a metric to assess the quality of clusters is similar to its purpose during NCL.

Once the clusters requiring application based clustering are identified, we use features extracted from the full HTML source of the web pages associated with domains. Note that there could be multiple and different sources of web pages associated with a certain domain. We use the timestamp of the complaint associated with domains to identify relevant HTML sources in a given epoch. Once we have the domains and their corresponding HTML content, we compute TF-IDF statistical vector on the bag of words on each cluster c [42]. Since the matrix is expected to be quite sparse, the application cluster submodule performs dimensionality reduction using SVD. Once we have the reduced application based feature vectors representing corresponding domains, this module uses the X-Means clustering algorithm to cluster domains hosting similar content.

3.4 AM: Cluster Attribution Module

The cluster attribution module is used to label clusters with keywords that are representative of a campaign's theme. To do this, we leverage the observation that a majority of the domain names involved with cross-channel abuse, despite being auto-generated using domain generation algorithms (DGAs) [28, 46] , have certain keywords in the domain name itself that are relevant to the theme of a campaign. In other words, the domain names are not completely random. The aim is to lure the victim into visiting these domains via their smartphones and a well designed domain name increases the odds of clicking the URL. For example, domain names `yourfastcashsystem[dot]com`, `24hrpaysite[dot]com`, `target.com.ctarg[dot]com`, have keywords cash, pay and target respectively that give us useful clues to what the domain might pertain to.

Using this observation, we use the Viterbi algorithm [33] to filter the domain names in a given cluster to a sequence of words such as [`your`, `fast`, `cash`, `system`] in the case of `yourfastcashsystem[dot]com` and [`24`, `hr`, `pay`, `site`] in the case of `24hrpaysite[dot]com`. More formally, let C be a cluster produced after the entire clustering process and let D be the set of domains in the cluster. For each domain $d \in D$, we create a set $U(d)$ that consists of all the parts of the domain name d except the effective top level domain (eTLD) (e.g. U('abc.example.com') = {abc, example}).

Next, we compute the set of words $W(U(d))$ using the Viterbi algorithm. Therefore, W(U('abc.example.com')) = {example} since 'abc' is not a valid English word. Using W, we increment the frequency counter for the word 'example' in a cluster specific dictionary. In this manner, after iterating over all domains in the cluster, we get a keyword to frequency mapping from which we pick the top most frequent word(s) to attribute the cluster.

4 Results

In this section, we begin by describing the data collected and used in CHURN for SMS-spam attribution. We then dive deeper into both CHURN's clustering results and the attribution accuracy of the system.

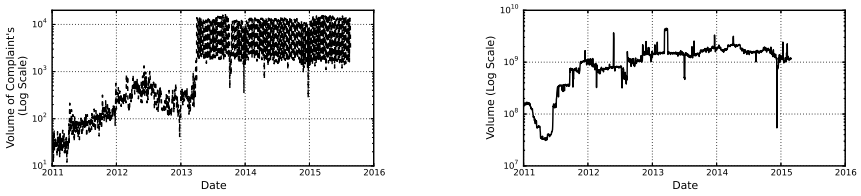
Epoch	RRs (Domain, IP) tuples	Domains (FQDN)	IPs (Hosts)	HTML sources	Complaints
Jan - Dec 2011	17,291	6,159	10,537	16,492	30,973
Jan - Dec 2012	17,316	7,846	8,218	16,321	125,960
Jan - Dec 2013	18,374	7,682	8,793	15,553	2,504,836
Jan - Dec 2014	22,426	7,438	8,858	15,334	3,286,988
Jan - Aug 2015	10,165	5,067	5,627	3,875	2,371,417
Total:	56,940	17,528	23,037	67,575	8,320,174

Table 2: Summary of collected datasets.

4.1 Datasets

CHURN starts with an SMS-spam repository we developed from the sources mentioned in Section 3. It had ≈8.32 million SMS-spam reports. The data collection module used the domain names found in these reports to collect surrounding pDNS, HTML and domain blacklist information using passive and active crawling methods. All these datasets were continuously gathered over a period of four years and eight months, starting in January 2011 and ending in August 2015, ensuring an overlapping time period.

The pDNS crawler was able to observe and record DNS Resource Records (RRs), which gives us a temporal mark between a domain name and an IP address when the SMS-spam was active. We collected 17,528 unique fully qualified domain names, 23,037 distinct IP addresses and 56,940 unique RRs related to the cross-channel abuse. Regarding the HTML datasets around this SMS spam abuse, we were able to download 67,575 distinct pages with the corresponding HTML source code. We summarize all this information across different epochs in Table 2.



(a) Number of daily complaints from both smswatchdog.com and FTC complaints.

(b) Daily aggregated passive DNS lookup volume trend for cross-channel spam domains.

Fig. 3: Temporal characteristics of collected datasets.

Temporal Characteristics of Cross-Channel SMS-Spam Figure 3(a) shows the number of daily SMS complaint reports retrieved and analyzed by our system. Although there are fluctuations in the number of daily complaints, the overall volume of such complaints steadily increased over time. We suspect that the sudden surge in the number of complaints received in early 2013 is due to both a proactive effort by both FTC (and other regulatory parties) to encourage people to report such spam and also an increase in the awareness among consumers of the available reporting tools. The period between mid-2013 to mid-2015 shows a relatively steady volume of SMS-spam reports with only marginal increase in the number of daily complaints. This signals that the more dominant spam campaigns had stablized during this time period. In addition, it is also possible that the number of consumers willing to report such spam had reached a saturation point. Finally, Figure 3(b) shows the daily aggregated DNS lookup volume to SMS-spam domains based on data collected from a large passive DNS repository. We clearly see an uptake and a steady DNS lookup volume over time, showing that the cross-channel SMS based abuse is a persisting phenomenon.

Lifetime of SMS-Spam Domains Figure 4 shows the empirical cumulative distribution function (eCDF) of the lifetime of all domains seen in the campaigns. The lifetime of a domain is derived by using the timestamp of the first and last seen DNS resolution to a particular domain. We observe that ≈30% of the domains had a lifetime of less than 10 days, close to ≈30% of domains had a lifetime between 10 and 100 days and the remaining ≈40% had a lifetime between 100 and 480 days. This indicates that cross-channel spam domains are alive for much longer periods compared to traditional spam abuse, and even certain type of agile botnet abuse such as fast-flux networks [39]. To better study the evolution of SMS-spam abuse, in the remainder of the paper we break and analyze the datasets into yearly epochs.

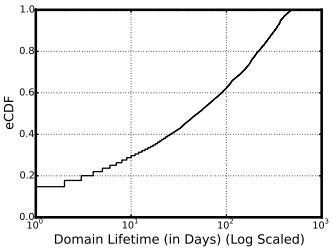


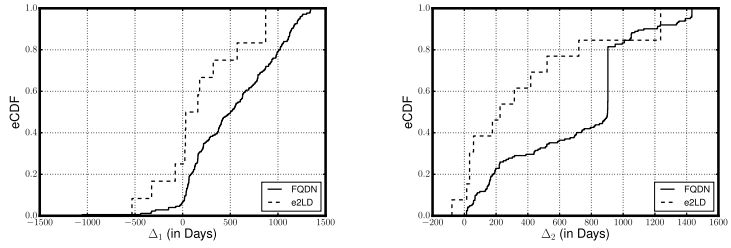
Fig. 4: The eCDF of the lifetime of all domains showing long-lived SMS-spam domains.

Reputation Properties of SMS-Spam Infrastructure Using domains from public blacklists (PBL), namely 'Malware Domains List' [12], 'sans' [17], 'Spamhaus Blacklist' [16], 'itmate' [10], 'sagadc' [13], 'hphosts' [7], 'abuse.ch' [1] and 'Malc0de' Database [11], we verify if and when an SMS-spam domain appeared in any of the PBLs. These PBLs typically include phishing domains, botnet domains, malware sites and other unsafe domains serving malicious content. Given that the cross-channel domains are alive for a long time and the cross-channel spamming is relatively newer, it was not clear whether the traditional blacklists are keeping pace with SMS-spam domains. Indeed, our finding shows that SMS-spam abuse is practically unknown to the PBLs. In total, we had only 177 out of the 17,528, a mere 1%, fully qualified domain names (FQDNs) listed in PBLs. Out of this, 170 domains were listed in a single list while seven domains were listed in two different lists. Moreover, when we checked all the effective second level domain names (e2LD) against the same lists, we only found 15 out of 17,502 (a minuscule 0.08%) e2LDs listed in one or more of the lists — with 11 e2TLDs being listed in a single list while four eTLDs were listed in two different lists. This provides clear evidence that traditional reputation feeds are failing to identify the cross-channel domains even in a postmortem way.

Diving a bit deeper in the blacklisted domains, we wanted to measure the timeliness of the blacklist updates. To achieve this, we computed two metrics Δ_1 and Δ_2. For a blacklisted SMS-spam domain d, $\Delta_1(d)$ measures the difference in days between the earliest date the SMS-spam domain was seen on a blacklist and the earliest date the domain was seen in an SMS-spam message in our complaint repository. $\Delta_2(d)$ measures the difference in days between the earliest date the domain was seen in a blacklist and the earliest date it was looked up, according to the passive DNS visibility we obtained.

Figure 5(a) shows the empirical cumulative distribution (eCDF) of Δ_1 over all blacklisted domains. We show two plots, one for the FQDNs and the other for the e2LDs. A positive value for Δ_1 means that the blacklisting happened after the earliest complaint was received, whereas a negative value implies that the blacklisting happened before the earliest complaint was received. From the eCDF of FQDNs, it is clear that around 94% of the blacklisted FQDNs were blacklisted after the complaint was received ranging from zero to 1,393 days. It is clear that the blacklists are rather slow in incorporating the domains. In some cases, about 6% FQDNs were blacklisted even before a complaint was received, indicating that sometimes either the SMS-spam is not reported on time or existing abuse domains related to traditional spam are being reused to cater to cross-channel spam. We observed a similar pattern in the case of e2LD.

Figure 5(b) shows the eCDF for Δ_2 for FQDNs and e2LDs. A positive value for Δ_2 means that the blacklisting happened after the earliest pDNS lookup as seen by our sensors, whereas a negative value implies that the blacklisting happened before the earliest pDNS lookup as seen in the pDNS database. In majority of the cases we observed a huge lag in the timeliness of the blacklist update. The lag ranged from 13 to 1433 days in the case of FQDNs and from -78 to 1506 days (only one negative value was seen) in the case of e2LDs. Although these findings are for a relatively small number of domains (those that ever appeared in a blacklist), it is clear that the blacklists appear to be lagging in discovering SMS-spam domains.



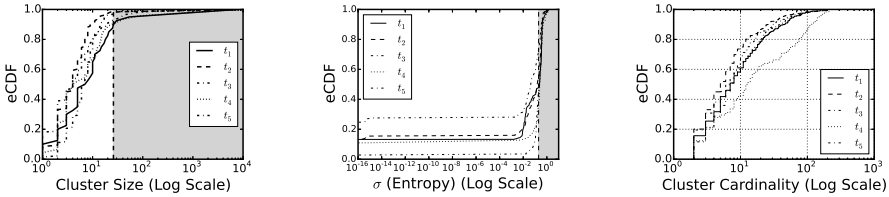
(a) eCDF of Δ_1 (b) eCDF of Δ_2

Fig. 5: Timeliness of blacklists

4.2 Clustering Results

Given a time period or an epoch and a set of domains, CHURN processes them in the hierarchical way as described in Section 3.3. We discuss the clustering results at various levels next.

Clustering Network & Application Level Information Figure 6(a) shows the empirical cumulative distribution of the cardinality (size) of the clusters produced after



(a) eCDF of the cardinality of the clusters produced in the NCL module. Clusters with cardinality $\geqslant \lambda = 25$ (shown as vertical line $x = 25$) are processed further.

(b) eCDF of the standard deviation (σ) of entropy of domain names for clusters after the PCL module. Clusters with $\sigma \geqslant \epsilon = 0.2$ (shown as vertical line $x = 0.2$) are processed further.

(c) eCDF of the cardinality of all the clusters produced after all modules (NCL, PCL and ACL) for five different epochs.

Fig. 6: HCL Thresholds

the network based clustering (NCL) step. Most of the clusters at this level contain few domains, but there exist some clusters that are quite large. We observed that up to 10% of the clusters produced during network level clustering had a cardinality ≥ 25, with one cluster being as large as almost half the number of domains under consideration. For these large clusters we leverage the domain popularity information to further break them down during the popularity based clustering (PCL) phase. By setting $\lambda = 25$, we were able to identify clusters to be processed by the popularity clustering submodule.

Once we have clusters from the NCL and PCL phases, the resulting clusters with disparate domain names are further refined using application level clustering (ACL). This is necessitated for some large clusters produced in the PCL module. Figure 6(b) shows the eCDF of the standard deviation (σ) in entropy of domain names for all clusters thus produced, differentiated based on epoch. Selecting as threshold $\epsilon = 0.2$, we were able to mark up to 60% of the clusters for further processing by the ACL module. Note that both the parameters λ (used in PCL) and ϵ (used in ACL) could be set according to the operator's needs. The application level clustering module gave us fine-grained clusters of very good quality with the largest cluster consisting of 201 domains across all epochs. Figure 6(c) shows the eCDF from the distribution of final cardinalities of all the clusters produced after all modules (NCL, PCL and ACL).

Cluster Level	Domain-(FQDN)	Label(s)	Epoch	Sample Domains
3	8	wire, deposit	2011	`wire600.com`, `deposit1500.com`
1	23	buy, best	2012	`bestbuy.com.bexy.biz`, `bestbuy.com.bwty.biz`
2	20	phone	2012	`mobiletestandkeep.com`, `iphone5tryout.com`
3	58	cash	2013	`startcreatingcash.com`, `trackingyoursuccess.com`
1	4	news	2014	`cnbcnews29.com`, `cnbcnews34.com`
3	129	loans, day, pay	2015	`instanteasyloans.co.uk`, `checkonlinepaydayloans.com`

Table 3: Representative sample of attributed clusters at various levels of the clustering hierarchy. Apart from the above and the case studies, we discovered campaigns related to selling drugs, adult content, free cruises, fake deals and many more.

AM Results The attribution module (AM) is used to label the clusters with keywords based on the domain name patterns. For illustration, Table 3 shows a sample output from this module. It can be seen that domains from certain campaigns can be attributed immediately after the NCL module. Some, however, are attributed after the PCL

module and others after the ACL module. This indicates that some campaigns can be identified just by using network features, while others require a combination of network, popularity and application features.

Evaluation To evaluate the output of CHURN and validate our results, we created ground truth data by labeling domains with group labels. Each group label represents a campaign. We made the judgement of assigning a specific group label to a domain based on looking at the domain names and loading up their associated webpages in a browser. Our experiment consisted of six group labels corresponding to the Bestbuy, Target, Walmart, Financial Freedom, Payday and News campaigns depicted as Group 1-6 in that order. We were able to label 653 (3.7%) domains in total to help us validate our results.

		Group1	Group2	Group3	Group4	Group5	Group6	Total	Parameter Setting
1.	✓	77	65	14	277	205	12	650	$\lambda=25$ & $\epsilon=0.2$
	✗	0	0	0	0	2	1	3	
2.	✓	76	57	14	257	192	12	609	$\lambda=2$ & $\epsilon=10^{-12}$
	✗	1	8	0	20	15	1	44	
3.	✓	67	54	10	208	155	10	504	$\lambda=2$ & $\epsilon=2$
	✗	10	11	4	69	52	3	149	
4.	✓	64	35	8	188	125	7	427	$\lambda=10000$ & $\epsilon=N/A$
	✗	13	30	6	89	82	6	226	
Total		77	65	14	277	207	13	653	

Table 4: CHURN evaluation based on ground truth with different system parameter settings across all epochs.

Table 4 shows how the results from CHURN measured up against the labeled data. System parameters λ and ϵ are varied to show the different cases. When λ is set to a relatively large value (i.e., 10,000), the output from the HCL module of CHURN is reduced to just the output of the NCL module since condition for PCL processing is never satisfied. The fourth threshold configuration shows that 427 out of the 653 domains were correctly attributed by CHURN using this setting. In the case when λ is set to a relatively small value (i.e., 2) and ϵ is set to a relatively large value (i.e., 2), the output from the HCL module of CHURN is reduced to output produced from applying the NCL and PCL modules sequentially but skipping the ACL module altogether. The third configuration shows that we attributed 504 out of 653 domains correctly using this setting.

Next is the case where λ and ϵ both are relatively small (i.e., 2 and 10^{-12} respectively). Such a setting results in all the modules NCL, PCL and ACL being serially applied to all clusters and domains without exception. This second configuration run shows that the number of correctly attributed domains increases from 609 to 653 domains. Finally, when λ and ϵ are set to 25 and 0.2 respectively, based on the justification presented in Section 4.2, NCL, PCL and ACL are applied to domains and clusters depending on the condition(s) being satisfied. This resulted in a marked improvement with 650 out of 653 domains being correctly attributed. The first configuration shows the results using this setting.

5 Case Studies

After CHURN's attribution module generates labels for clusters, these clusters and their associated labels are used to identify and group domains that are part of the same scam campaign. We present case studies for three of the most prominent campaigns that are known SMS scams. As a general takeaway across all three case studies, we observed that the domains supporting the scams were hosted in diverse but few IP locations

and for a long period of time. While the distributed infrastructure ensures reliability, the long term activity behind the domain names suggests the relative ineffectiveness of defenses against these social engineering cross-channel attacks compared to similar attacks via the internet channel.

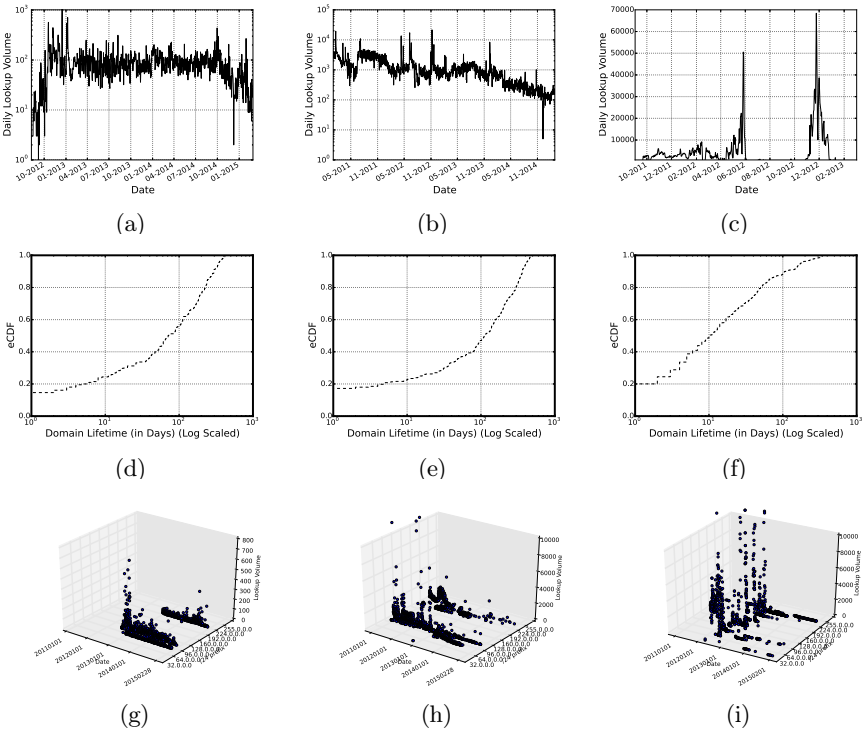


Fig. 7: Three Campaigns: Financial Freedom, Payday and Gift Card. For each we show (7(a)–7(c)) daily lookup volumes according to our pDNS database, (7(d)–7(f)) eCDF of the lifetime of the domains seen and (7(g)–7(i)) 3D view of campaigns based on time, popularity and network infrastructure (IPs binned by /24 prefix).

Financial Freedom: Upon landing on the Financial Freedom web page an embedded video explains the purported benefits of enrolling into the program. The victim is asked to provide her personal information for 'Free Instant Access' to the program. The scam targeted consumers who are financially weak and looking for a solution to credit card debt problems. In our dataset, this scam consisted of 277 FQDNs (e.g. `morefreedomforall[dot]com`) and 187 IPs belonging to 49 distinct /24 subnetworks. None of the domains in this scam were seen in domain blacklists and the domains ended up being clustered in the ACL module. Figure 7(g) shows that the campaign used dedicated infrastructure to operate in a stealthy mode thus surviving for a long time, as can be seen in Figure 7(a), 7(d). Legal proceedings of a law suit initiated against the perpetrators of this scam can be found here [4].

Payday: Payday loan is a short term, high interest cash advance that has been banned in many states in the United States, and the Federal Trade Commission (FTC) has issued warnings regarding it [5]. For example, in one instance the defendants' online contract stated that a $300 loan would cost $390 to repay, but the defendants then charged consumers $975 to repay the loan. This is a case of obscuring the 'Terms of service' specified on the site, which make it hard for the victim to realize they are being scammed. The scam works by sending a victim a SMS message with a URL. Upon clicking the URL, the victim is asked to enter personal information, phone number, and loan amount to proceed further.

A particular online payday loan campaign was clustered in our SMS spam dataset consisting of 207 unique domains; hosted in 212 unique IP addresses; belonging to 142 distinct /24 subnetworks. 68 out of 207 such domains were part of the `.co.uk` TLD. Eight domains in this scam were seen in PBL and they were mainly clustered by the ACL module. Figure 7(b), 7(e) shows that despite the warnings by consumer protection authorities (especially in the USA), this scam has survived and continues to victimize consumers. In addition to this, Figure 7(h) shows the stability behind the network infrastructure used to support the scam domains.

Giftcard: In this case study, the scam works by sending the victim a SMS message with a URL and a code. Upon clicking the URL, the victim is asked to enter his/her personal details including phone number followed by entering the code in order to receive a fake free gift card from the associated brand (e.g., Target, Bestbuy, Walmart etc.). Thereafter, victims were told to sign up for more than a dozen risky trial offers, none of which were free, to qualify for the promised 'free' gift card. In many cases, the correct code confirmed to the gift card scam operators that the mobile number is indeed active and they use this entry as a pretense to falsely subscribe the victim's mobile number to premium rate services.

The giftcard campaign consisted of 207 domains and 215 IPs belonging to 85 distinct /24 subnetworks. Four domains under this scam were seen in PBL and the domains were mostly clustered in the NCL module. This campaign was mostly active during two distinct time periods in 2012 and 2013, as can be seen in Figure 7(c). The resurgence of the campaign the second time coincides with the shopping/holiday season between November 2012 and January 2013 where a lucrative deal for a gift card is more likely to catch the victim's attention. Figure 7(f) shows that ≈45% of the domains had a lifetime of less than 10 days, ≈45% were active between 10-100 days and the remaining ≈10% of the domains were relatively long lived. We found that out of 207 domains, many of them were well crafted 4LDs (4th level domains), named after specific brands such as BestBuy (114), Target (77) or Walmart (16) e.g. `target.com.tthg[dot]biz`. We also noticed that the domains hosting these web pages have very similar layout, structure and content. The majority of the Giftcard scam domains had a relatively shorter lifetime and were more agile in using their network resources.

The FTC pressed charges against the perpetrators of the Gift Card campaign for illegally sending ≈42.5 million text messages to consumers containing bogus offers for 'free' Gift Cards. These charges were publicly reported to be settled in September 2013 [19]. This is reflected in Figure 7(c), where we see very few to no lookups during the second half of 2013.

6 Related Work

Although there has been work in both SMS spam detection [35, 38] and discovering SMS spam campaigns [30], our focus on and characterization of the network infrastructure used by SMS spam campaigns provides new insights that are not available from past research. Jiang et al. [35] use the concept of 'grey' phone numbers, which are phone numbers associated with data-only devices such as laptop data cards and electricity meters, as honeypot end points to capture SMS-spam. They then apply statistical models on the collected data to identify the source phone numbers generating spam. Murynets et al. [38] conducted an empirical analysis of SMS-spam collected from fraudulent accounts in a large cellular provider to uncover spamming sources and their strategies. Our work differs from these works because of our focus on characterization of the network infrastructure rather than source phone numbers of spam. Moreover, while their analysis is based on call detail records (CDR's) generated on the telephony channel, we explore the cross-channel nature of such abuse by attributing the Internet infrastructure that facilitates SMS abuse by using crowd-sourced complaint and passive DNS and application datasets.

Boggs et al. [30] propose a method to discover emergent malicious campaigns in cellular networks by using graph clustering methods with mutual contact graphs that capture interactions between nodes which represent phone numbers or domain names. In addition to discovering SMS spam campaigns, we explore the properties of the infrastructure that supports such campaigns using both passive DNS data and the application level information available from webpages to which users are directed when they click on URLs contained in SMS messages. Our results show that some of the assumptions made in earlier work do not actually hold. For example, [30] assumes that Internet public blacklists can be helpful in detecting and stopping malicious SMS messages but we show that little overlap exists between domains in SMS messages and these public blacklists.

There have been numerous studies that cluster spam infrastructure and campaigns based on URL [43], IP infrastructure [25, 29] and content [24, 32]. Although we do not claim novelty around the individual features used in clustering SMS-spam infrastructure, our contribution lies in observing that it is most effective to use features from different layers of the network stack in a hierarchical manner so as to capture the diverse types of SMS-spam campaigns. Prior work has shown the ineffectiveness of traditional blacklists in protecting services such as instant messaging (IM) [41], and social media [34, 43]. Our demonstration of the poor blacklist coverage of SMS-spam domains is similar. The significant gap in blacklist coverage and longevity of SMS-spam domains shows the limits of using email and malware abuse intelligence to fight cross channel abuse. Lever et al. [37] analyzed malicious cellular DNS traffic generated by mobile applications to conclude that mobile app-level protection (eg. app-market security) suffices to curtail mobile attacks. Our work shows that the emergent cross-channel abuse strategy bypasses this and is a more serious threat to mobile users.

Key Differences: In summary, much of the past work in SMS abuse has focused on the analysis of call detail records to identify spam source phone numbers rather than on the characterization of the network infrastructure that facilitates the abuse. Such network characterization has helped us demonstrate that current publicly available Internet threat intelligence largely fails to identify this infrastructure to stop long-lived SMS spam campaigns. Our work differs in both the long-term analysis of the problem, but also the new methods we propose to cluster and attribute SMS spam messages over time.

7 Limitations

Data collected and analyzed by CHURN, which includes consumer complaints and passive DNS data, is primarily US-centric, making it difficult to generalize the findings to other parts of the world. Indeed, cross-channel spam trends could be different in Europe or Asia as compared to the US. However, our attribution system, CHURN is designed to be easily deployable elsewhere, without much change. In future work, we hope to be able to use CHURN with data from other countries and provide insights on cross-channel abuse from around the world. CHURN's evaluation is based on a limited set of labeled data/ground truth. Although, we consciouly made an effort to label data that is representative of all the spam domains under consideration, by randomizing the selection process for manually inspecting the domains, we recognize the need to scale this experiment and plan to do it in the future while adding more capabilities to our system.

8 Conclusion

In cross-channel abuse, SMS-spammers are able to exploit the ubiquity of mobile devices and trust in the telephony channel to craft attacks that could be more successful than spam on the Internet channel alone. Such illicit activities have become a serious problem, with several reported scams that have lasted for several years. Using data from multiple sources, we seek to attribute cross-channel abuse to the Internet infrastructure that facilitates it. Our research results confirm that SMS-spam is not well defended against, as such campaigns are able to run for long periods of time. Although there is some agility in the network resources used by them, very few of the domains used, appear on traditional domain blacklists.

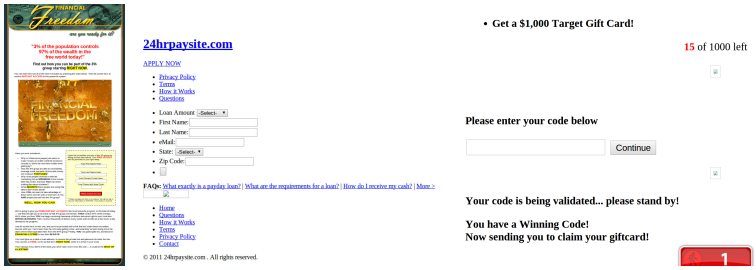
Acknowledgements

This work was supported in part by National Science Foundation grants CNS-1318167 and CNS-1514035. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

9 Appendix

9.1 Prominent Campaigns Snapshots

Figure 8 shows the snapshots of the three campaigns discussed in this paper.



(a) (b) (c)

Fig. 8: Three Campaigns: Financial Freedom, Payday and Gift Card. For each we show (8(a)–8(c)) web pages rendered on a mobile browser.

9.2 Hierarchical clustering module Dendogram

Figure 9 graphically depicts all the attributed clusters in our study at different levels for epoch t_2 (2012) as a radial dendogram plot. The center represents all the domains under consideration and the concentric circles represent the cluster labels at each level starting from NCL (level 1), to PCL (level 2) and ACL (level 3), as we move outward radially.

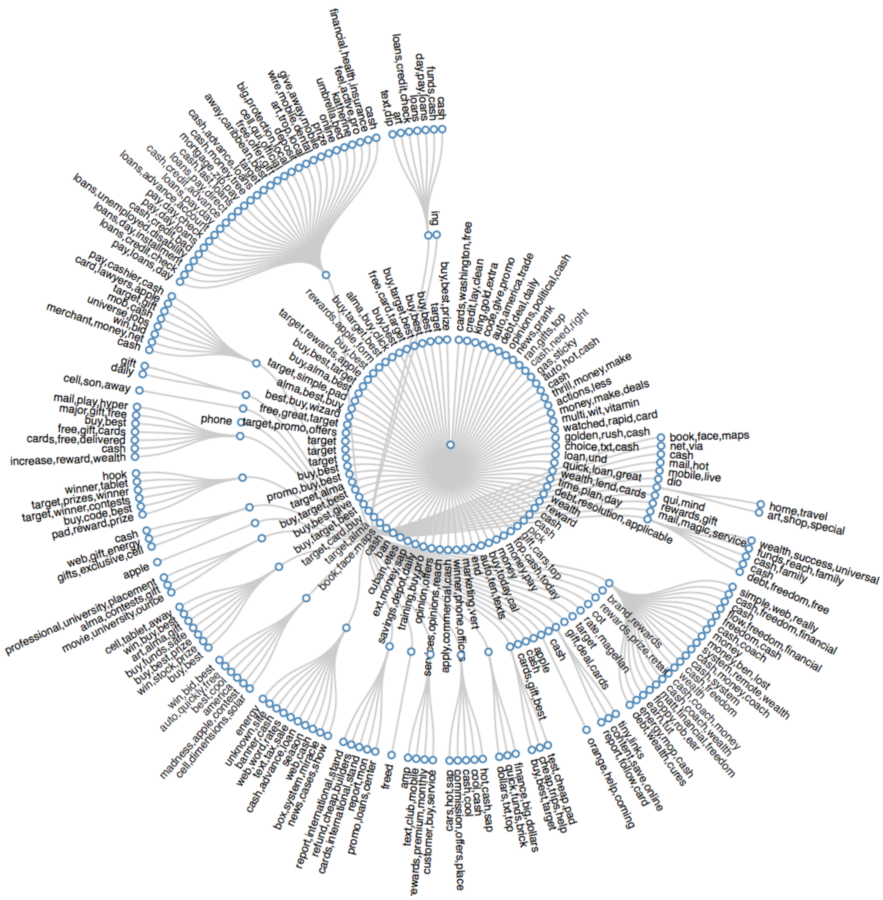


Fig. 9: A radial dendrogram plot illustrating the output from the hierarchical clustering module for a single epoch.

References

[1] abuse.ch - the swiss security blog. `https://www.abuse.ch/`.

[2] Alexa Top Sites. `http://www.alexa.com/topsites`.

[3] Federal Trade Commission FTC Complaint Assistant. `https://www.ftccomplaintassistant.gov/#crnt&panel1-1`.

[4] FTC on Financial Freedom. `https://www.ftc.gov/enforcement/cases-proceedings/092-3056/financial-freedom-processing-inc-formerly-known-financial`.

[5] FTC on Payday Lending. https://www.ftc.gov/news-events/media-resources/consumer-finance/payday-lending.
[6] FTC Robocall Challenge. https://robocall.devpost.com/.
[7] hphosts. http://www.hosts-file.net/.
[8] Identifying parking ip infrastructure: Understanding malware evolution and the implications on data modeling. https://www.damballa.com/identifying-parking-ip-infrastructure-understanding-malware-evolution-and-the-implications-on-data-modeling/.
[9] Internet archive: Wayback machine. https://archive.org/web/.
[10] I.T. Mate Product Support. http://support.it-mate.co.uk/.
[11] Malc0de database. http://malc0de.com/database/.
[12] Malware Domain List. http://www.malwaredomainlist.com/.
[13] sagadc summary. http://dns-bh.sagadc.org/.
[14] SMS Phishers Exploit Twilio and ow.ly to Steal Mobile Account Logins. http://blog.cloudmark.com/2014/02/13/sms-phishers-exploit-twilio-and-owly-to-steal-mobile-account-logins/.
[15] SMSWatchDog. http://www.smswatchdog.com.
[16] SPAMHaus Blocklist. https://www.spamhaus.org/lookup/.
[17] Suspicious domains - sans internet storm center. https://isc.sans.edu/suspicious_domains.html.
[18] Technical realization of the Short Message Service (SMS), 3GPP TS 23.040, v13.0.0. http://www.3gpp.org/dynareport/23040.htm.
[19] Text Spammers Settle FTC Charges They Illegally Sent Consumers Bogus Offers for 'Free' Gift Cards. https://www.ftc.gov/news-events/press-releases/2013/09/text-spammers-settle-ftc-charges-they-illegally-sent-consumers.
[20] Tropo. https://www.tropo.com.
[21] Twilio. http://www.twilio.com.
[22] What kind of SMS messages are not allowed to be sent using Twilio? https://www.twilio.com/help/faq/sms/what-kind-of-sms-messages-are-not-allowed-to-be-sent-using-twilio.
[23] Your very own SMS Internet gateway with Arduino. http://x-ian.net/2012/10/09/your-very-own-sms-internet-gateway-with-arduino/.
[24] D. S. Anderson, C. Fleizach, S. Savage, and G. M. Voelker. *Spamscatter: Characterizing internet scam hosting infrastructure*. PhD thesis, University of California, San Diego, 2007.
[25] M. Antonakakis, R. Perdisci, D. Dagon, W. Lee, and N. Feamster. Building a dynamic reputation system for DNS. In *19th USENIX Security Symposium, Washington, DC, USA, August 11-13, 2010, Proceedings*, pages 273–290. USENIX Association, 2010.
[26] M. Antonakakis, R. Perdisci, W. Lee, N. Vasiloglou, and D. Dagon. Detecting malware domains in the upper DNS hierarchy. In *the Proceedings of 20th USENIX Security Symposium (USENIX Security '11)*, 2011.
[27] M. Antonakakis, R. Perdisci, Y. Nadji, N. Vasiloglou, S. Abu-Nimeh, W. Lee, and D. Dagon. From Throw-Away Traffic to Bots: Detecting the Rise of DGA-Based Malware. In *the Proceedings of 21th USENIX Security Symposium (USENIX Security '12)*, 2012.
[28] M. Antonakakis, R. Perdisci, Y. Nadji, N. Vasiloglou, S. Abu-Nimeh, W. Lee, and D. Dagon. From throw-away traffic to bots: Detecting the rise of dga-based malware. In *Presented as part of the 21st USENIX Security Symposium (USENIX Security 12)*, pages 491–506, Bellevue, WA, 2012. USENIX.
[29] L. Bilge, E. Kirda, C. Kruegel, and M. Balduzzi. EXPOSURE: finding malicious domains using passive DNS analysis. In *Proceedings of the Network and Distributed System Security Symposium, NDSS 2011, San Diego, California, USA, 6th February - 9th February 2011*. The Internet Society, 2011.
[30] N. Boggs, W. Wang, S. Mathur, B. Coskun, and C. Pincock. Discovery of emergent malicious campaigns in cellular networks. In *Proceedings of the 29th Annual Computer Security Applications Conference*, ACSAC '13, pages 29–38, New York, NY, USA, 2013. ACM.

[31] C. J. Burges. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2):121–167, 1998.

[32] M. F. Der, L. K. Saul, S. Savage, and G. M. Voelker. Knock it off: profiling the online storefronts of counterfeit merchandise. In *The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '14, New York, NY, USA - August 24 - 27, 2014*, pages 1759–1768, 2014.

[33] J. Forney, G.D. The viterbi algorithm. *Proceedings of the IEEE*, 61(3):268–278, March 1973.

[34] C. Grier, K. Thomas, V. Paxson, and M. Zhang. @spam: The underground on 140 characters or less. In *Proceedings of the 17th ACM Conference on Computer and Communications Security*, CCS '10, pages 27–37, New York, NY, USA, 2010. ACM.

[35] N. Jiang, Y. Jin, A. Skudlark, and Z.-L. Zhang. Greystar: Fast and accurate detection of sms spam numbers in large cellular networks using grey phone space. In *Proceedings of the 22Nd USENIX Conference on Security*, SEC'13, pages 1–16, Berkeley, CA, USA, 2013. USENIX Association.

[36] N. Jiang, Y. Jin, A. Skudlark, and Z.-L. Zhang. Understanding sms spam in a large cellular network: Characteristics, strategies and defenses. In S. Stolfo, A. Stavrou, and C. Wright, editors, *Research in Attacks, Intrusions, and Defenses*, volume 8145 of *Lecture Notes in Computer Science*, pages 328–347. Springer Berlin Heidelberg, 2013.

[37] C. Lever, M. Antonakakis, B. Reaves, P. Traynor, and W. Lee. The core of the matter: Analyzing malicious traffic in cellular carriers. In *20th Annual Network and Distributed System Security Symposium, NDSS 2013, San Diego, California, USA, February 24-27, 2013*. The Internet Society, 2013.

[38] I. Murynets and R. P. Jover. Crime scene investigation: SMS spam data analysis. In J. W. Byers, J. Kurose, R. Mahajan, and A. C. Snoeren, editors, *Proceedings of the 12th ACM SIGCOMM Conference on Internet Measurement, IMC '12, Boston, MA, USA, November 14-16, 2012*, pages 441–452. ACM, 2012.

[39] J. Nazario and T. Holz. As the net churns: Fast-flux botnet observations. In *3rd International Conference on Malicious and Unwanted Software, MALWARE 2008, Alexandria, Virginia, USA, October 7-8, 2008*, pages 24–31, 2008.

[40] D. Pelleg, A. W. Moore, et al. X-means: Extending k-means with efficient estimation of the number of clusters. In *ICML*, pages 727–734, 2000.

[41] I. Polakis, T. Petsas, E. P. Markatos, and S. Antonatos. A systematic characterization of IM threats using honeypots. In *Proceedings of the Network and Distributed System Security Symposium, NDSS 2010, San Diego, California, USA, 28th February - 3rd March 2010*, 2010.

[42] G. Salton and M. J. McGill. *Introduction to Modern Information Retrieval*. McGraw-Hill, Inc., New York, NY, USA, 1986.

[43] K. Thomas, C. Grier, J. Ma, V. Paxson, and D. Song. Design and evaluation of a real-time URL spam filtering service. In *32nd IEEE Symposium on Security and Privacy, S&P 2011, 22-25 May 2011, Berkeley, California, USA*, pages 447–462. IEEE Computer Society, 2011.

[44] T. Vissers, W. Joosen, and N. Nikiforakis. Parking sensors: Analyzing and detecting parked domains. In *22nd Annual Network and Distributed System Security Symposium, NDSS 2015, San Diego, California, USA, February 8-11, 2014*, 2015.

[45] M. E. Wall, A. Rechtsteiner, and L. M. Rocha. Singular value decomposition and principal component analysis. In *A practical approach to microarray data analysis*, pages 91–109. Springer, 2003.

[46] S. Yadav, A. K. K. Reddy, A. Reddy, and S. Ranjan. Detecting algorithmically generated malicious domain names. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, pages 48–61. ACM, 2010.