SMS Spam
A Holistic View

Lamine Aouad[1], Alejandro Mosquera[1], Slawomir Grzonkowski[1] and Dylan Morss[2]
[1]*Symantec Ireland, Ballycoolin Business Park, Dublin 15, Dublin, Republic of Ireland*
[2]*Symantec San Francisco, 303 Second Street, 94523, San Francisco, CA, U.S.A.*

Keywords: Mobile Messaging, Abuse, Spam, Concept Drift, Targeting Strategies, Filtering.

Abstract: Spam has been infesting our emails and Web experience for decades; distributing phishing scams, adult/dating scams, rogue security software, ransomware, money laundering and banking scams… the list goes on. Fortunately, in the last few years, user awareness has increased and email spam filters have become more effective, catching over 99% of spam. The downside is that spammers are constantly changing their techniques as well as looking for new target platforms and means of delivery, and as the world is going mobile so too are the spammers. Indeed, mobile messaging spam has become a real problem and is steadily increasing year-over-year. We have been analyzing SMS spam data from a large US carrier for over six months, and we have observed all these threats, and more, indiscriminately targeting large numbers of subscribers. In this paper, we touch on such questions as what is driving SMS spam, how do the spammers operate, what are their activity patterns and how have they evolved over time. We also discuss what types of challenges SMS spam has created in terms of filtering, as well as security.

1 INTRODUCTION

From the early stages of this research, we realized that the question of how spam differs from legitimate communications was not a trivial one. Typically, spam would refer to unsolicited, undesirable, and mostly commercial communications. However, there is sometimes a thin line between spam and legitimate advertising or bulk communications for other purposes, be it political campaigns, event organization, religious communities, subscriber lists, crowdfunding campaigns and so on. We are all subjected to these types of communications and advertisements on a daily basis while browsing the Internet, on social networking sites, or in our email inboxes. However, most of this is based on services we have used before or opted into and, although still sometimes irrelevant or even irritating, this is somehow accepted by the masses as the price paid to access valuable content and services. On the other hand, we consider spam to be generally imposed, and with no potential value or benefit for the recipient. Moreover, most of spam-advertised products and services are usually deceptive to the user, frequently involving scams and financial fraud.

As the world increasingly turns to mobile devices, spammers do too. In a survey from Tatango, an SMS marketing company, 68% of mobile users in the US reported receiving SMS spam in 2011. Reportedly that equates to 4.5 billion spam messages received that year, with a 45% increase from the previous year (Kharif, 2012). SMS spam is also an emerging issue in many other areas of the world, representing 20 to 30% of the traffic in parts of Asia, including China and India (GSMA Spam Reporting, 2011). In addition to being a nuisance, as any other unsolicited and unwanted communication, SMS spam represents a possible financial loss for the subscribers, with risks of phishing attacks or malware downloads, for instance, leading to subscription to premium rate services. SMS spam can also be highly damaging to the reputation of the mobile carrier brand and cause increased operating costs.

Since its inception, the spam landscape has always been a cat-and-mouse playground between the spammers on the one hand, and anti-spam vendors and email providers on the other. The same applies to SMS spam, as mobile operators are becoming increasingly involved in the anti-spam field. Many cooperative organizations, such as M3AAWG (M3AAWG, 2014) and working groups within GSMA (GSMA, 2014) for instance, are very active in mobile messaging abuse and related security issues.

On the research side, the literature presents a range of studies focusing on SMS spam filtering.

Aouad L., Mosquera A., Grzonkowski S. and Morss D..
SMS Spam - A Holistic View.
DOI: 10.5220/0005023202210228
In *Proceedings of the 11th International Conference on Security and Cryptography* (SECRYPT-2014), pages 221-228
ISBN: 978-989-758-045-1

Some of them derived from the email space, mainly content-based technologies, based on Natural Language Processing (NLP) and machine learning techniques (Charles Lever and Lee, 2013), (Gómez Hidalgo et al., 2006), (M. Zubair Rafique, 2010), among many others. We will mention additional research in this paper. However, we will mainly focus on the multi-faceted aspect of the spam chain describing the full set of resources and entities involved as well as the prospect of implementing effective defenses at each stage of the spam chain.

The reminder of this paper is organized as follows. Section 2 will describe the spam chain and the analysis of spam activities during more than six months and the implications in terms of anti-spam strategies. Section 3 will highlight the challenges in terms of filtering and concept drift, as well as security. It will also provide some background and related research. Concluding remarks are then given in section 4.

2 THE SPAM CHAIN

As opposed to spammers back in the early days, where they use to operate mostly individually and on the whole chain, nowadays spammers are more organized and in some ways more specialized. They mostly organize the content of the campaigns, and send bulk messages. They usually do not author the target content, nor do they process the orders. This section will highlight the multi-faceted nature of the spam chain and will also analyze the spam and its evolution.

2.1 What's behind the Traffic?

A tipping point of the spam evolution was the adoption of affiliate networks taking away different responsibilities in the spamming chain, which increased the army of spammers, and consequently the traffic as well. There are hundreds of legitimate affiliate networks that promote legitimate products in lawful ways. We have, however, looked at those responsible for the spam we observed, in order to identify potential ways of being proactive against new campaigns, mostly based on the type of content they are serving.

The majority of openly spamming networks require an invitation. They are usually run as gangs and have proven hard to get access to. However, we have seen many cases of *shady* networks, that are relatively open and with no clear policy against spamming. It can also be argued that affiliates abusing products and services do not necessarily make the affiliate network spammy. We could, however, identify some of those behind the most virulent campaigns, and in a few cases register and get additional information on the process and the services provided by these affiliate networks.

The most popular kind of affiliate promotion we observed was in the adult/dating space, initially using hundreds of throwaway domains, then increasingly using a range of URL shortening services, all advertising a relatively small number of products and usually offering attractive percentages or commissions per sale. During the observation period, we saw only half a dozen products, most of which were also operated and hosted by the same entities. This pointed to rather a small number of affiliates targeting a limited number of high-commission products. Identifying affiliate IDs behind the links seen in the spam is straightforward as they are usually part of the destination URL parameters, then preserved via parameters or cookies throughout the entire process. For one of the biggest adult dating campaigns, we observed the same affiliate ID for the whole six-month period, switching from SMS spam, to social media, and adult dating scams based on online classified Ads. In this case, it seems content providers including hosting companies, registrars and affiliate networks are not taking enough responsibility, as this spammer, and others, have repeatedly been reported, but are still active.

The process itself is fairly straightforward; once a new domain is registered, the spammer logs in with the affiliate network, select the program, the landing page, and then the campaign and its features. The system will then give you a customized link to the site ready to use in the spam campaign. The traffic generation is very basic; a large number of SMS messages are sent to recipients that are randomly generated or scraped off the Internet. In other cases, spammers might choose to keep a low profile by sending low volumes of targeted spam. The good news is that the network traffic related to those links is easily identifiable, indicating affiliate marketing as opposed to a possibly registered user receiving an SMS from that service.

Another type of affiliate-driven spam we observed included rogue pharmacy and 'work from home' types of campaigns, which mostly used URL shortening services or disposable domains and pointed to fast-flux hosted pages. These types of scam sites were easily traceable using network analysis tools and we were consequently able to pro-actively identify large numbers of additional links and compromised service networks serving this type, and similar types, of scam and spam content. Crossovers at the IP level could also highlight existing correlations between these spamming activities and some systems vulnerabilities. However, this is quite dynamic and needs to be updated frequently.

Although representing a large chunk, the traffic was not all driven by affiliate marketing. Other contributors included a large number of bank scams and phishing campaigns, in addition to some payday loan and other spam and scam campaigns, including social media, junk cars, fake prizes, and so on. The following section will present some activity patterns used by SMS spammers.

2.2 How do Spammers Operate?

We have seen a range of activity patterns related to domain registration, landing pages and target content, recipient lists generation, spam domain naming, and so on. This section introduces the landscape and frequent patterns we have observed in the SMS spamming world.

2.2.1 Domain Registration Patterns

More than 70% of SMS spam uses URL-based call-to-action (CTA). Among these, more than 50% were newly registered domains. While this is a common pattern, we have seen an increase in use of URL shortening services over time, seeing them used in as much as 7% of the traffic in certain weeks and averaging 2% overall. Further analysis of the short links generated by spammers revealed that most of them were generated at the same time that the message itself was sent. The destination URL is usually modified by adding a dummy parameter to generate a completely different short link while still redirecting to the same target website. We also observed spammers making use of hacked websites and public hosting services but to a much lower extent, representing less than 0.1% of the traffic during the six-month period.

During the analyzed period, we noticed how SMS spammers reused keywords in domain names and contact information, such as administrator names and email addresses for their mass domain registration processes. This contact information is usually publicly available using the registrar WHOIS service, which would make it possible to track and identify new potentially spammy domains. We also identified several hosting services and registrars that were popular among SMS spammers. One of these registrars accounted for more than 40% of spam domains, mostly those used in adult and dating spam campaigns. Using this knowledge, it is possible to extract valuable information about the domain registration process in order to detect and predict new domain names that can be potentially used in spam campaigns.

Table 1: Popularity ranking of spam domain naming keywords per category.

Rank	Adult and Dating	Gambling	Finance and Loans	Health	Apps.
1	date	system	loan	cure	quiz
2	sex	betting	pay	diet	mobile
3	be	roulette	cash	los	mob
4	flirt	sport	payday	to	cell
5	mob	click	advance	fat	gana
6	my	bank	credit	secret	fun
7	ero	pick	my	your	win
8	dating	win	the	in	app
9	love	bet	secure	weight	bala
10	vie	lottery	now	body	mo
11	club	the	score	for	game
12	fun	poker	in	fitnes	yep
13	vid	football	online	workout	skill
14	girl	to	debt	free	thrill
15	single	secret	life	health	m
16	the	lot	money	stop	club
17	offer	racing	finder	natural	me
18	black	winning	for	how	blink
19	in	lay	quick	solution	play
20	gay	profit	rate	get	get
21	mobile	pro	usa	life	gold
22	shag	cash	tax	garcinia	zi
23	secure	money	free	training	hoch
24	match	blackjack	relief	treatment	yu
25	game	bingo	offer	guide	wi
26	partner	best	account	day	iu
27	sm	soccer	car	and	the
28	free	horse	holiday	muscle	kazoo
29	friend	online	auto	skin	dorado
30	local	casino	network	now	mundo
31	xxx	vega	first	program	dragonfly
32	video	tipster	pro	plan	sm
33	survey	winner	shop	lose	izz
34	and	tip	get	power	ring
35	just	crusher	of	slim	master
36	hookup	bot	your	healthy	score
37	meet	crap	daily	max	champion

We have seen many cases of algorithmically-generated domain names taking into account variable-length substrings. For spammers, good domain names are those that contain keywords that give semantic information relevant to the campaign (e.g. meetnicegirls or advancepaydaynow). Because SMS messages are short, the URL is an important part of the message and should be enticing to the victim. The strategy for choosing domain names affects the success of the spam campaign. We have compared the domain names for different SMS spam campaigns by splitting them into variable-size n-grams and ranking the top n-grams for each category, as shown in Table 1.

The keywords used in the domain name generation are not very different to high ranking keywords entered into search engines for these niches. Similarly, if we cluster the non-TLD part of the domain name by n-gram we can see high-density clusters for popular spam campaigns such as adult/dating or payday loans as shown in Figure 1. These keywords can be helpful in pre-emptive detection and categorization.

2.2.2 Landing Pages

As mentioned in the previous section, the use of good domain names seems to be critical for the success of a spam campaign in the SMS world. These domains

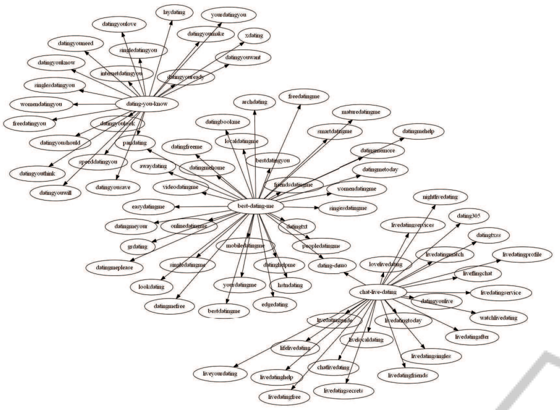
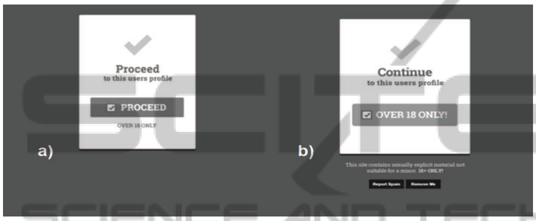


Figure 1: Example of adult/dating domain clusters.


a)

Proceed
to this users profile

☑ PROCEED

OVER 18 ONLY

b)

Continue
to this users profile

☑ OVER 18 ONLY!

This site contains sexually explicit material not suitable for a minor. 18+ ONLY!

Report Spam | Remove Me

Figure 2: Example of changing landing pages in an adult/dating campaign.

would then redirect to the affiliate link, or show an interstitial Web page. The latter option seems to be more popular in adult/dating campaigns in order to hide credit card scams such as advertising fake age verifications that end up in the subscription to several adult services with subsequent credit card charges, or for credibility purposes, e.g. by adding an unsubscribe button.

All landing pages of the spam domains were analyzed in order to detect and track new campaigns. While most of them remained constant during the duration of the campaign, others evolved as the product was changing. We have used clustering techniques and extracted content and structural features from the target pages in order to detect near-duplicates, such as those shown in Figure 2. Additional campaigns were also seen elsewhere, as identified by the network analysis mentioned earlier, and despite being global scams (such as the ’work from home’ campaigns), they were still reusing the same content and very similar components.

2.2.3 Targeting Strategies

Random Generation. One of the most frequent targeting strategies adopted by spammers is the random generation of recipient phone numbers (Murynets and Piqueras Jover, 2012), (Jiang et al., 2013). Generally, these would be generated within a specific area code but could also target different area codes or exchange codes for the same campaign, or even hit the entire phone number space. These campaigns are also quite aggressive and usually generate high rates of spam traffic.

In the US, the sequence of target phone numbers is a concatenation of three components, namely the three digits of the area code, the three digits of the exchange code (usually considered as part of the subscriber number), and the last four digits of the subscriber number. Figure 3 shows an example of an adult/dating campaign using random generation, targeting the same area code in this case, with occasional increments in the exchange code.

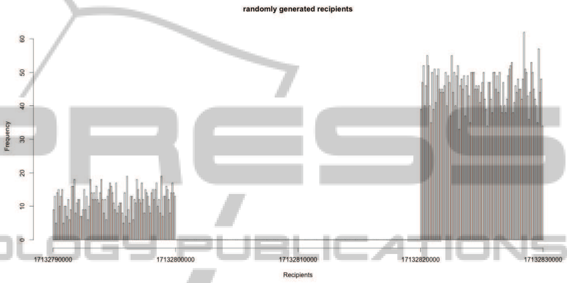


Figure 3: Example of random generation in an adult/dating campaign.

As can be seen in Figure 3, these numbers were uniformly generated. They perfectly fit the uniform distribution using maximum likelihood for instance as shown in Figure 4, with Q-Q and P-P plots showing that the empirical data, representing the first thousand instances of destination phone numbers here, comes from the population with uniform distribution, as the points perfectly fall along the reference lines.

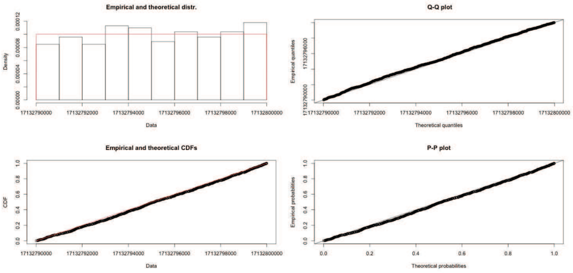


Figure 4: Recipients fitting the uniform distribution.

There are a variety of goodness-of-fit tests in the literature which would summarize the disparity between observed and expected values under a given model. We have tested the most commonly used ones, including Chi-Square, Kolmogorov-Smirnov, Anderson-Darling, and Cramer-Von Mises. Most of the spam traffic we have seen in the spam falls within the uniform distribution as shown earlier. P-values returned by these methods vary depending on the implementation, the random numbers generator, number of bins, and how the data is binned. Presenting p-values associated with each test might not be relevant here. Note however that at the beginning of the observation period, more than 50% of the spam traffic used random generation of the recipient lists. However, we have seen this number fall to about 20% as certain campaigns primarily using it were being blocked and hence losing their momentum.

Social Spam. Spammers are continually adopting new platforms and technologies and using any channel available to drive the traffic. Social media as a channel is seeing an astonishing rise in spam and scam activities. There was a 355% increase in the first half of 2013 according to Nexgate, a company specializing in the field of social Web security and compliance. Social media is also used as a mean to drive traffic to other channels, including the SMS world. For instance, we have seen many targeted campaigns initiated through adult sites and many chat and mobile dating applications. This usually involves chat bots engaging with online users and attempting to get them to register for a given product so spammers will earn referral and affiliate bonuses. In the SMS conversation example transcribed in Table 2, the victim received an initial message from a chat bot through a mobile dating application.

Since the emergence of Web 2.0 technologies, spammers have always spread and collected information through online forums, the blogosphere, compromised accounts and websites, and other targeted sites. We have seen for instance compromised Facebook accounts spreading lottery scam, compromised Twitter accounts spreading “miracle diet” spam, compromised branded short domains used in targeted spam campaigns on popular mobile applications like Snapchat, Kik, and the list goes on (Narang, 2014). We have also seen a sheer number of SMS campaigns collecting phone numbers off classified ads sites and use them in targeted spam campaigns. One particularly nasty scam campaign involved a website claiming to expose online prostitution solicitors. Spammers were creating fake profiles on the target website using phone numbers extracted from online ads and asking for money (200$ to 500$) in order to remove the data from the website. The message reads: **ALERT* You are listed on [WEBSITE]/[PHONE NUMBER] for soliciting a prostitute online for sex. Go to the above link to Delete your profile.*

Table 2: Example of a subscriber engaging in SMS conversation with a chat bot. Initiated on social media.

victim: Hey sexy...this is [NAME]..from [DATING APP]
bot: so i don’t have xrated pics online but i have a couple on my phone... [LINK] ... now send me urs bby
victim: Yep..didn’t want to let u get away...lol
victim: I must say...I think you are just beautiful:P:P:P;)
bot: u like my shirt baby? haha want sum more??
bot: sum private pix babes, i want the dirty stuff
victim: Yes I do.....want all u wanna give...
. . . . . . *Discussion continues with the chat bot sending links of adult content to the victim, leading to a subscription on a scam dating site*
bot: its free to join but it will ask for a card i think.. im gonna get naughty and i cant have kids watching..
victim: I’m in now
bot: ok babe.. talk to you in there.. gonna put my phone to charge.. mwa! xoxo

Low-volume Campaigns. Another common trick used by spammers is to keep a low profile by sending a low volumes of targeted spam. In the email world, this is usually botnet-driven using compromised accounts. We have seen a range of under-the-radar spam and scam campaigns, some of which use SMS gateways or SMS through email. However, during the observation period, we did not see any proof of compromised phones being used in spam campaigns. The most persistent type of campaigns were essentially ’get rich quick’, ’work from home’, and fake lottery campaigns. Table 3 shows what SMS through email messages look like. These grew in volume throughout the observed period, showing the constant evolution of delivery methods used by spammers.

Table 3: Example of SMS through email spam.

FRM:[EMAIL ADDRESS] SUBJ:Hello MSG:You pumped to make a bunch cash online? Then go here: [LINK]
FRM:[EMAIL ADDRESS] MSG:[BANK NAME] NOTICE: Your ACCOUNT has been Locked. Please call [PHONE NUMBER].

When and where? We were also interested in the ’when’, ’where’ and ’how long’ of the observed campaigns. Overall, more than 80% of recurrent campaigns were initiated outside of working hours (late afternoon/early morning), and on weekends. Campaigns taking place during working hours had the shortest lifespan amongst similar campaigns. This is one of the reasons spammers generally work staggered hours. We have also monitored several spammers’ phone numbers (sometimes used in testing) and we have found that some of them do indeed have legitimate daytime jobs besides their illegal activities.

Further analysis of the area and exchange codes of recipient phone numbers shows at least two recurrent geo-targeting strategies that correlate with the previously described patterns; uniformly-generated recipients are usually focused on one or few geographic areas that they sweep trying to reach potential victims. Moreover, spammers that use phone lists leaked from databases or crawled from the Internet show a more uniform geographic distribution where the most populated areas are more affected, as shown if Figure 5.

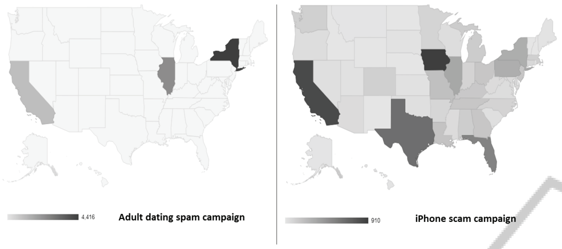


Figure 5: Different geographic targeting for two recurrent campaigns.

We also noticed a number of campaigns initiated in the US and targeting foreign networks. One of the more significant campaigns involved real estate spam, written in Chinese, with a relatively large number of variants and call-to-actions associated with them. We have also seen spam and phishing campaigns, in Spanish, targeting Central and South-American networks and subscribers. Typically, we would see 30 to 50 different languages weekly, which highlights the multi-lingual challenges faced by spam filtering methods and tools.

2.3 How Does it Evolve?

As mentioned in previous sections, spam campaigns change over time by selecting new affiliate products, registering new domains, changing the targeting strategy, evolving the textual content of the messages, or applying evading techniques such as obfuscation in order to avert filters. The following section will briefly describe concept drift and some of the techniques used in some of the long-lasting campaigns, namely lose-weight (rogue pharmacy), adult/dating, and bank scams.

2.3.1 Campaign Drift

It was initially thought that, given their size, SMS messages provided little room to create substantially different content to evade filters. As it turns out, this was not the case. The lack of context in SMS messages also makes it difficult to link campaigns and offers even less information to work with, compared to email spam for instance.

On the other hand, we have observed overlaps with email and social media spam for some campaigns, especially on Twitter and classified ads websites. The latter is abused frequently as people usually post their contact details including phone numbers, which makes it low-hanging fruit for SMS spammers who collect this data and use it to create personalized attacks. It was also quite common to see that after the initial campaign stops generating traffic or gets blocked in one channel, spammers recycle the domain for another one. There were also recurrent probes in the analyzed SMS spam trying to reuse old domains for new campaigns.

Lexical Variation and evading Techniques. Recurrent campaigns that are active for a long time usually exhibit a high level of lexical variation. These message variants are generated by paraphrasing the original message or replacing the call-to-action URL or phone number. An example message reads: *Elizabeth, this is what worked for her [LINK-REMOVED]*. Some of the variants found in this high-volume lose-weight campaign can be seen in Figure 6.

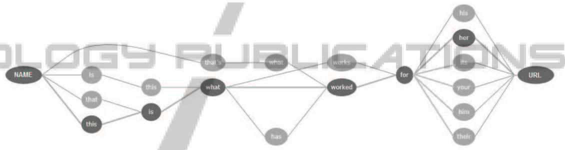


Figure 6: Different message variants in a lose-weight campaign.

These lexical variations can be based on synonyms or semantically related words as shown in Table 4, but it can also involve the inclusion of misspellings, slang, SMS-style contractions or phonetic substitutions. As it has been shown in NLP literature, these are good candidates for text normalization before using other techniques such as LSA or LDA in order to lower the dimensionality of the data (Yvon, 2010).

Table 4: Examples of lexical variation in bank scam campaigns.

CARD Service ALERT: Your DEBIT-CARD has been BLOCK. Please call [PHONE NUMBER]
Metro C.U. Alert: Your DEBIT-CARD has been DE-ACTIVATED. Please call [PHONE NUMBER]
Your CARD starting with 440336 has been temporary FLAG. Please call CREDIT UNION SERVICES at [PHONE NUMBER]
Credit Union Mobile ALERT: Your VISA has been temporarily SUSPENDED. Please call Cardholder Services 24hrs line [PHONE NUMBER]

Another interesting example of variation involves the insertion of common keywords frequently used in text messages (see Table 5). Taking into account n-gram overlaps, these messages would be more difficult to filter by just using word occurrence vectors. So any trained classifier on content-only features would fail or generate a high number of false positives. Because the language used in these texts is very common in non-spam (ham) messages, the only textual fingerprint that can be probably extracted in this case would be the URL. Although quite common in other channels such as email or Web spam, we have not seen advanced obfuscation techniques in SMS spam besides some recurrent car scrapping campaigns. We have also seen rare instances of obfuscation using multipart messages where the URL is split in order to evade the filtering, or encoding obfuscation using injected Chinese, Korean and other characters that cannot be encoded in GSM 7-bit and need UTF-16, which would usually be triggered in newer devices, and should not be an issue. Some common tricks used by spammers include the use of interleaved spaces between all characters and number/letter substitutions of visually similar tokens. We have also seen the use of simple anti-URL detection measures such as interleaving a space before the dot or replacing it with the word *dot* in several messages.

Table 5: Examples of keyword poisoning for adult campaign.

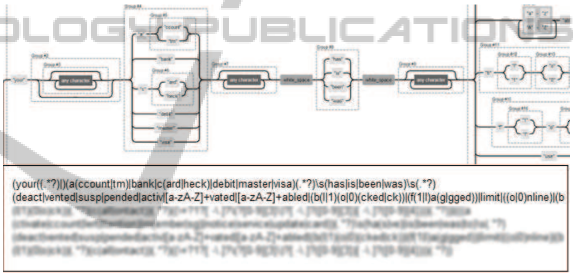
[URL] that is what the website is called
Hi! You looked nice when i sawyou . Its Melanie, respond back to me at [URL]
Hey watsup Its heather. Messsage me at [URL] my pics are up too.
Helllo , you were cute the other day, Textmessage me back at [URL] its Katie
Hi i saw you the other day . i was too nervous to ask then but do u want to talk . Is me on my profile at [URL]
Hey you were lookin good when i saw you, Its jasmine . Hit me back at [URL]
Hi! You looked nice when i sawyou . Its Melanie, respond back to me at [URL]

3 FILTERING AND SECURITY CHALLENGES

SMS spam has proven more challenging than expected; in content-based filtering for instance, where the length of these messages gives little material to work with and makes misclassification more likely. The language used in SMS messages, which contains extra linguistic challenges with abbreviations, phonetic contractions, bad punctuation and so on. There is also far less context compared to email for instance and information found in headers. In addition, concept drift happens more quickly in the mobile world, with campaigns running for a much shorter period of time, and spammers being far more reactive and responsive.

A large number of filtering techniques have been applied to SMS spam including traditional content-based filtering using regular expressions, supervised and non-supervised machine learning techniques, evolutionary algorithms, crowd-sourcing, and many content-less methods based on features of the network, temporal analysis, reputation, and so on (Delany et al., 2012). We have tested the efficacy of many of these methods in dealing with real-world issues and the increased sophistication we have seen in the SMS spam.

As mentioned in the previous section, campaigns that exhibit a high lexical variability proved to be more challenging in terms of filtering. We have seen thousands of variants of the previously mentioned bank scam, but its textual patterns can still be inferred by taking into account common n-grams and subsequences. In this case, using a variant of the Aho-Corasick algorithm, extracted textual patterns were compiled into regular expressions.



```
(your((.*?)|)(a(ccount|tm)|bank|c(ard|heck)|debit|master|visa)(.*?)\s(has|is|been|was)\s(.*?)
(deact|vented|susp|pended|activ|[a-zA-Z]+vated|[a-zA-Z]+abled|(b(l|1)(o|0)(cked|ck))|(f(1|l)a(g|gged))|limit|((o|0)nline)|(b
```

Figure 7: Automatically-evolved regular expression filter for bank scam campaigns.

With an aim to maximize the detection coverage and minimize the chance of false positives, the previously obtained regexes were combined using an evolutionary algorithm, discarding the ones that match ham messages and giving more weight to the most successful ones. The highest ranked regular expression in our experiments obtained a 98% matching coverage with unique bank/card scam messages without generating any false positive (see Figure 7). However, the high complexity of the generated regular expressions would have a strong performance impact, showing that traditional filtering techniques should be adapted in order to deal with these kinds of continuously-evolving threats.

Another interesting challenge in SMS filtering is how to provide additional perspectives on the little information contained in the messages. CTA fingerprinting results in ordinary conversations, which included references to spam URLs, phone numbers, etc., to be filtered as well. The question is then; how to differentiate *semantic territories* of text messages? There are a range of NLP techniques with tentative solutions to this issue and to a more general case of differentiating ’meanings’ attached to text. We have tested LSA (Latent Semantic Analysis) - mentioned before as a pre-processing method to other classifiers, but it can also be solely used as a classifier - and also LDA (Latent Dirichlet Allocation). LSA is a bag-of-words model that represents word co-occurrences, meaning the structure within the documents is not maintained. LDA on the other hand can be seen as a mixture of topics that splits out words with certain probabilities, so if applied to a set of documents and topics, it will output topic representations for each document.

Models have successfully been populated with representatives of important campaigns, including the bank scam mentioned earlier, to be used in blocking. If the cosine distance of any incoming message is higher than a certain threshold, it represents an actual spam, as opposed to a message including the CTA. The higher the threshold the more accurate the model is, which can be tuned to avoid false positives. The case of forwarding however is a lost cause and would still be blocked.

Other important challenges ahead of mobile messaging abuse are bot-driven campaigns mentioned earlier, whether originating from ordinary phone numbers belonging to spammers, or infected mobiles. During the analyzed period, we saw an instance of a campaign distributing malware, which was a Trojan SMS Agent /Opfake, representing a variant of a continually evolving infection typically used to send text messages from infected mobile devices to premium rate numbers. This malicious application creates a mobile botnet by sending malicious links to numbers in the contact list via SMS. Analysis of command-and-control (C&C) activities revealed a wide spread in a short period of time, from the initial infected devices in Egypt reporting to the C&C server, to tens of thousands of SMS messages sent in the US, South Korea, India, and many other countries. This highlights the importance of defenses at the client side, as well as preventing malicious messaging, whether internal to operators or across borders.

4 CONCLUSION

Although we have not seen an increase in volume, we have come across a relatively high level of sophistication in the SMS spam world. The mobile ecosystem is also undergoing major developments, with an increased in the market share of smartphones, and the wider adoption of IP-messaging over text messaging, but this is not necessarily taking the heat off mobile network operators. We have seen evidence that spammers are using multiple delivery channels and are by no means abandoning SMS messages just yet. Unlimited text plans and the trusted nature of text messages will always attract attackers and make it necessary for operators to deploy effective defenses. This paper has described the SMS spam ecosystem, covered some of the most effective counter measures to a wide range of SMS spam, along with the trends and challenges ahead.

REFERENCES

Charles Lever, Manos Antonakakis, B. R. P. T. and Lee, W. (2013). The core of the matter: Analyzing malicious traffic in cellular carriers. In *NDSS 2013*.

Delany, S. J., Buckley, M., and Greene, D. (2012). Review: Sms spam filtering: Methods and data. *Expert Systems with Applications*, 39(10):9899–9908.

Gómez Hidalgo, J. M., Bringas, G. C., Sánz, E. P., and García, F. C. (2006). Content based sms spam filtering. In *Proceedings of the 2006 ACM Symposium on Document Engineering*, DocEng ’06, pages 107–114, New York, NY, USA. ACM.

GSMA (2014). The GSM association. http://www.gsma.com/. [Online; accessed 20-June-2014].

GSMA Spam Reporting (2011). Sms spam and mobile messaging attacks - introduction, trends and examples. Technical report.

Jiang, N., Jin, Y., Skudlark, A., and Zhang, Z.-L. (2013). 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. USENIX Association.

Kharif, O. (2012). Mobile Spam Texts Hit 4.5 billion Raising Consumer Ire. http://www.bloomberg.com/news/2012-04-30/mobile-spam-texts-hit-4-5-billion-raising-consumer-ire.html. [Online; accessed 20-June-2014].

M. Zubair Rafique, M. F. (2010). Sms spam detection by operating on byte-level distributions using hidden markov models. In *Virus Bulletin 2010*.

M3AAWG (2014). Messaging, malware and mobile anti-abuse working group. http://www.maawg.org/. [Online; accessed 20-June-2014].

Murynets, I. and Piqueras Jover, R. (2012). Crime scene investigation: Sms spam data analysis. In *Proceedings of the 2012 ACM Conference on Internet Measurement Conference*, IMC ’12, pages 441–452, New York, NY, USA. ACM.

Narang, S. (2014). Snapchat spam: Sexy photos lead to compromised branded short domains. http://www.symantec.com/connect/blogs/snapchat-spam-sexy-photos-lead-compromised-branded-short-domains. [Online; accessed 16-January-2014].

Yvon, F. (2010). Rewriting the orthography of sms messages. *Natural Language Engineering*, 16:133–159.