A Trend Micro Research Paper

An In-Depth Analysis of Abuse on Twitter

Jonathan Oliver and Paul Pajares
Threat Research Team

Christopher Ke, Chao Chen, and Yang Xiang
Deakin University

CONTENTS

ABSTRACT

In this paper, we examine Twitter in depth, including a study of 500,000,000 tweets from a two-week period to analyze how it is abused. Most Twitter abuse takes the form of tweets with links to malicious and spam websites.

These websites take many forms, including spam websites, scam sites involved in compromising more Twitter accounts, phishing websites, and websites with malware or offering cracked versions of software. Many of the malicious tweets are sent from legitimate accounts that have been compromised, creating a range of problems for their owners.

The scale of the threat is significant. Previous research, notably "@spam: The Underground on 140 Characters or Less," (Grier, 2010), indicates that using URL blacklists is ineffective in detecting threats. Our research shows otherwise—approximately 5% of all tweets with links contained malicious and/or spammy content.

We also applied graph algorithms to the Twitter data and were able to find various clusters of interrelated websites and accounts. We were able to identify specific spam tweet campaigns as well as groups carrying out these campaigns.

The data from this analysis leads us to conclude that blacklisting, in conjunction with other analytical tools, is an effective tool for identifying malicious tweets.

* This work was supported by ARC Linkage Project LP120200266.

INTRODUCTION

Researchers from Trend Micro and Deakin University worked together to investigate the Twitter threat landscape. This paper features a comprehensive study that lasted for two weeks from 25 September to 9 October 2013, which includes further analysis of some of the threats we discovered throughout the given period. The study revealed a significant level of abuse of Twitter, including spamming, phishing, and sharing links that led to malicious and potentially illegal websites. The majority of the malicious messages we observed were sent from abused and compromised accounts, many of which have subsequently been suspended by Twitter.

A 2010 study examined 400 million public tweets and 25 million URLs.[1] The authors identified 2 million URLs (8%) that pointed to spamming, malware-download, scam, and phishing websites, which led them to conclude that:

- Blacklists were ineffective, as these only protected a minority of users
- URL-shortener usage made the task of identifying malicious links very difficult

This research paper begins by giving a brief overview of the types of Twitter abuse we discovered within the study period. It then provides a summary of the data we collected to learn more about the abuse. Given the data, we examined a range of issues, including:

- The use of blacklists to catch Twitter spam
- The coordinated nature of certain Twitter spam outbreaks
- The timing of spam outbreaks
- Details related to particular Twitter scams

In Section IV, we propose an approach for analyzing Twitter spam outbreaks, which is very useful in augmenting blacklists for detecting Twitter spam.

OVERVIEW OF THE ABUSE ON TWITTER

This section provides a brief overview of the Twitter threats we found. It also provides examples of the most active threat types, which revealed a significant level of Twitter abuse, including:

- Traditional spam similar to email spam
- Searchable spam, which differed from email spam
- Phishing messages
- Suspended and compromised accounts

Traditional Spam

The following are some of the features of traditional Twitter spam:

- They typically promoted weight-loss drugs, designer sunglasses and bags, etc. like email spam.
- They used unrelated but often-trending hash tags to increase tweet distribution and to attract more people to click the link.
- They used misspelled words, sometimes substituting numbers for letters, which was typical of email spam 10 years ago.
- Some used URL shorteners to make it more difficult for security analysts to identify which tweets point to spam websites.

Searchable Spam

Figure 1 shows examples of searchable Twitter spam.

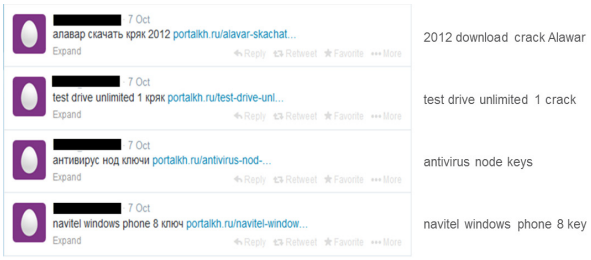

Figure 1: Sample searchable spam with translations on the right

The following are some features of searchable spam:

- They typically promote free access to copyrighted and licensed materials or offer gadget knockoffs such as:
  - Solutions to homework and exams
  - Free movie downloads
  - Cracked versions of software
  - Computer, printer, and mobile device knockoffs
- They did not have or sparingly used hash tags.
- Many of them were written in Russian.
- They used several domains, many of which were hosted in Russia and in the Ukraine.

Our analysis of searchable spam revealed that the probability of Twitter suspending an account involved with an incident was significantly lower than if it was involved in sending out traditional Twitter spam or other malicious messages. In addition, we found that 50% of those who clicked links in searchable spam written in Russian were from non-Russian-speaking countries such as the United States and Japan (see Section VII). This spam type typically stays on Twitter after transmission and can be readily searched for. For example, Group A, described in Section V, consists of over 7.8 million searchable spam. Approximately 90% of these remained accessible on Twitter at the time of writing.

We conclude that searchable spam attempt to avoid irritating users so they would not be reported with the help of the Abuse button Twitter has made available. They cover a wide range of content, which some users might be motivated to use Twitter's Search function to find. They might even be willing to use automated translation tools so to understand the content of such spam.

Twitter Phishing

We examined a long-running phishing scam that exploits certain Twitter features.[2] The scam starts with a compromised user sending messages to friends (using the @ syntax on Twitter) that ask them to click a shortened URL. Clicking the link would start a redirection chain that ends on a phishing page that tells them their session timed out and that they need to log in again. In the course of doing research, we attempted to estimate the scale of this problem.

Suspended and Compromised Accounts

While doing research, we followed accounts that were involved in spamming. We attempted to access them in December 2013 (i.e., two months after our period of analysis). We found that Twitter suspended tens of thousands of accounts involved in spamming and in other malicious activities. Many of these appeared to have been specially created for this purpose. The accounts were created then immediately started sending out spam. In some cases, account owners identified the problem and took some corrective actions to restore their accounts. These were significantly rarer than account suspension. We do not have statistics on this in this paper because it was difficult to establish when compromises occurred. We only had anecdotal evidence of their occurrence.

RESEARCH SCOPE AND METHODOLOGY

We collected as many tweets with embedded URLs as possible within the two-week period from 25 September to 9 October 2013. We restricted the tweets we examined to those with embedded URLs. While it is possible to use Twitter to send out spam and other messages without URLs, the majority of the spam and other malicious messages we found on Twitter had embedded URLs. Among the thousands of spam that humans inspected in the course of research, we only found a handful of tweets without URLs that could be considered abusive or harmful.

We categorize tweets that contain malicious URLs as "malicious tweets." The data we collected is shown in Table 1. We gathered a total of 573.5 million tweets containing URLs and identified 33.3 million malicious tweets, which accounted for approximately 5.8% of all of the tweets with URLs.[1] We used two methods to identify malicious tweets. The first method involved the use of the Trend Micro Web Reputation Technology, which used a blacklist.[3] The second method involved identifying groups of malicious tweets using the clustering algorithm described in Section IV. Note that we experienced a disruption in our data-collection process on 29 and 30 September 2013, which accounted for data loss during the said period.

1 The authors understand that the two-week study period was during a period of spam activity that was significantly higher than the norm.

Table 1: Data Collected

Day/Date	Number of Tweets with URLs	Number of Malicious Tweets	Percentage of Malicious Tweets
Wednesday, 09/25/2013	39,257,353	2,292,488	5.8%
Thursday, 09/26/2013	47,252,411	3,190,600	6.8%
Friday, 09/27/2013	49,465,975	3,947,515	8.0%
Saturday, 09/28/2013	37,806,326	2,018,935	5.3%
Sunday, 09/29/2013	-	-	-
Monday, 09/30/2013	-	-	-

Table 1: Data Collected			
Day/Date	**Number of Tweets with URLs**	**Number of Malicious Tweets**	**Percentage of Malicious Tweets**
Tuesday, 10/1/2013	48,778,630	2,511,489	5.1%
Wednesday, 10/2/2013	51,728,355	3,739,597	7.2%
Thursday, 10/3/2013	51,638,205	3,932,186	7.6%
Friday, 10/4/2013	49,230,861	3,398,526	6.9%
Saturday, 10/5/2013	44165664	2293539	5.2%
Sunday, 10/6/2013	45,089,730	2,006,447	4.4%
Monday, 10/7/2013	50,457,403	2,305,794	4.6%
Tuesday, 10/8/2013	42,031,232	1,152,119	2.7%
Wednesday, 10/9/2013	16,612,318	538,133	3.2%
TOTAL	**573,514,463**	**33,327,368**	**5.8%**

CLUSTERING ALGORITHM TO IDENTIFY MALICIOUS TWEETS

One of our research goals was to obtain a high-level understanding of the various types of spam and scams on Twitter. We determined that one approach to achieve this understanding would be to cluster malicious tweets into groups. Forming clusters of malicious tweets would be successful if we could adequately explain why tweets in a group are considered similar to one another and why they are considered malicious.

Several possible variables could be extracted from tweets, including

- Content
- Embedded URLs
- Hash tags
- Sender data, including frequency

It would prove very useful if it were possible to group Twitter spam into distinct outbreaks rather than try to understand a huge mass of data. Traditional approaches for doing this include grouping Twitter spam that have similar content or applying machine-learning approaches. Applying machine-learning approaches involves extracting numerical or categorical variables from tweets and users (e.g., how often they send messages, dramatic changes in their behavior, etc.) and applying a statistical or machine-learning approach to the data (e.g., SVMs or Nearest Neighbor).

We took another approach. Our proposal for identifying certain classes of high-volume spam is to create a graph consisting of senders and domains in tweet URLs and to identify bipartite cliques in this graph.[4] Such graphical approaches to identifying cliques in data have been previously applied to computer security problems.[5] To do this, we constructed a graph where the Twitter users are nodes on the left-hand side of the graph while the domains in links are nodes on the right-hand side. For each tweet from User U that contains a link with Domain D, we include an arc in the graph from User U to Domain D. Some spammers use applications that employ a round-robin approach for sending spam. Given a number of sending accounts and destinations for URLs in the tweets, the use of a round-robin approach maximizes the number of spam while minimizing the effects of having the accounts they are using suspended and of blacklisting in blocking their spam. When the graphical approach described above is used, a set of users using a round-robin approach will generate a bipartite clique in the graph. Hence, bipartite cliques in such a graph are very suspicious—the probability of real users behaving this way in the normal course of events is extraordinarily small. There are scalable approaches for using map-reduce to identify cliques in large data sets.[6, 7]

Figure 2 provides an example of a bipartite clique found in the data consisting of 727 users who sent tweets containing links to 11 domains; all of the users in the clique sent tweets containing links to all of the domains in the clique.

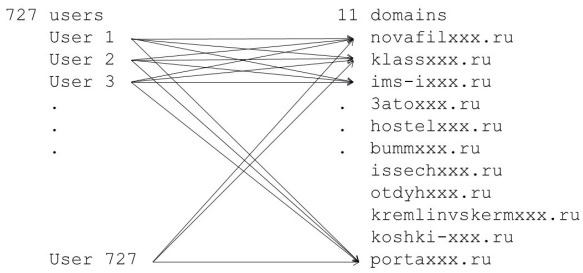

Figure 2: Sample bipartite clique

This approach is suited to understanding certain types of Twitter spamming behaviors but unsuited for others. For example, it is not suited for analyzing the Twitter follower scam described in Section VI since it did not use a round-robin approach for sending scam messages. The Twitter follower scam was confirmed malicious by installing the app and monitoring its behavior.

Other groups of malicious behavior were identified by following the links through to the final website and confirming that the website was malicious.

HIGH-LEVEL PERSPECTIVE

We applied the clique algorithm described in Section IV to the Twitter data we collected. [6] The algorithm identified 16 cliques, each of which accounted for 1% or more of the Twitter spam. Table 2 describes each of the cliques generated. In addition, Group G was a Twitter follower spam group, which accounted for 2.5% of the Twitter spam.

Table 2: High-Level Perspective

Description	Percentage of Malicious Tweets	Number of Senders	Hash Tags	Number of Domains	Percentage of Suspended Accounts
A. Education spam, etc.	27.28%	797	None	24	10.3%
B. Cracked software and game spam	8.11%	578	None	20	31.5%
C. Education spam	6.26%	539	None	20	19.7%
D. Cracked software spam	6.19%	9,509	Limited	21	12.0%
E. Cracked software spam	4.39%	727	None	11	11.6%
F. Printer/ Mobile spam	3.72%	12,275	Low	3	89.1%
G. Twitter follower spam	2.54%	59,205	Yes	1	2.1%
H. Video/ Mobile/ Cracked software/game spam	2.23%	8,987	Low	50	95.2%

Table 2: High-Level Perspective					
Description	**Percentage of Malicious Tweets**	**Number of Senders**	**Hash Tags**	**Number of Domains**	**Percentage of Suspended Accounts**
I. Game and computer spam	2.04%	608	None	19	97.9%
J. Education spam, etc.	1.99%	284	None	14	47.9%
K. Shirt spam	1.91%	1,699	None	5	74.7%
L. Game, mobile, and printer spam	1.81%	1,197	None	18	98.8%
M. Computer/ Printer spam	1.77%	26,603	Low	60	42.3%
N. Game/ Hardware spam	1.53%	2,514	Yes	70	90.0%
O. Computer game/mobile device spam	1.41%	1,491	None	73	94.7%
P. Credit and education spam	1.08%	8,541	None	32	72.5%
Q. Cracked software and game spam	1.02%	9,066	None	4	98.6%
Other spam	24.74%	N/A	N/A	N/A	N/A

The columns in Table 2 are defined as follows:

- The "Description" column describes the content of the tweets.
- The "Percentage of Malicious Tweets" column gives the percentage of tweets out of the total 28 million tweets in the group.

- The "Senders" column shows the number of confirmed senders in a clique. As such, a confirmed sender should have sent tweets to all of the domains in a clique. For example, 797 senders sent at least 24 messages with links going to all of the 24 domains in Group A. The number of senders in Group G is simply the number of senders who sent tweets with URLs that led to a Twitter follower scam website. In this case, there was no convenient confirmation step to separate legitimate users who retweeted spam from those whose accounts were under spammers' control.
- The "Hash Tags" column summarizes the use of hash tags in spam that belong to each group.
- The "Domains" column lists the number of domains. Some groups used multiple hosts from the same domain. For example, Group H in Table 1 had five separate domains and used 10 distinct hosts to each of the domains.
- The "Percentage of Suspended Accounts" column shows the percentage of accounts that have been suspended when we checked their status in December 2013—two months after the study period.

We noted the following from Table 2:

- The 17 groups listed account for 75% of the Twitter spam we identified.
- It is highly likely that there were other types of abuse and spam that we were not able to identify in the study.
- Twitter is very effectively responding to some spam outbreaks. For example, it has identified and suspended over 95% of the accounts in Groups H, I, L, and Q. Other spamming behaviors were not detected. For example, in Group A, which accounted for over 27% of the spam we found, approximately 10% of the accounts were suspended.

DETAILS ON SPECIFIC OUTBREAKS

Russian-138 Spam

Six of the groups described in Section V had a set of features in common. We coined the term "Russian-138 spam" to describe Twitter spam with the following features:

- They were primarily written in Russian.
- Many of the domains in the tweets were *.ru* domains.
- The URLs were followed by a date stamp.

For example, a tweet with the URL, *http://xxxxxx.ru/angliyskiy-fizik-moss-t-1380765135.html,* was sent on 5 October 2013. *1380765135* appears to be a timestamp that translates to "Thursday, October 3, 01:52:15 2013 UTC," two days before the tweet was sent.

The six groups that were characterized as Russian-138 spam were Groups A, B, C, E, I, and J. Figure 3 shows the number of tweets per hour in each of the groups monitored within the study period.

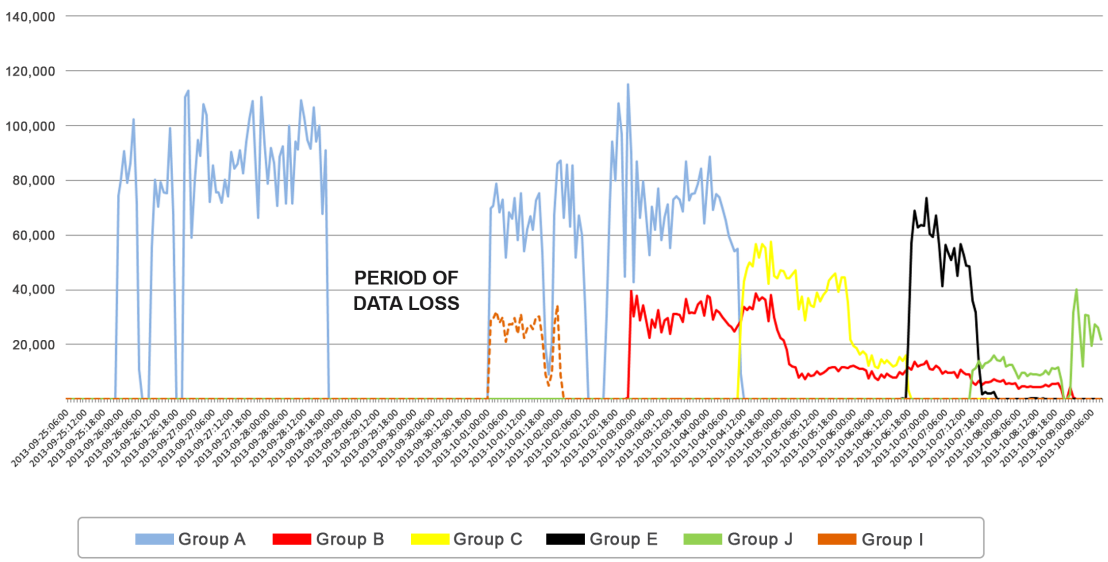

Figure 3: Number of tweets per hour for the six Russian-138 spam groups

Figure 3 highlights the spammy nature of the groups:

- The groups of spamming Twitter users are acting in a coordinated manner. They start and stop spamming at roughly the same time.
- In some situations, one group of users will stop spamming to a set of domains while at the same time another group will start spamming

another set of domains. Examples of this include the following:

- At 2013-10-04 11:00 UTC, Group A (blue) stopped spamming and Group C (yellow) started spamming.
- At 2013-10-06 18:00 UTC, Group C (yellow) stopped spamming and Group E (black) started spamming.

Twitter Follower Scams

In January 2014, we reported about a Twitter follower scam that used spam to entice users to install and authorize an app access to their accounts.[8] Once authorization is granted, the users' accounts would get more followers (i.e., other users of the app), become a follower of other users of the app, and possibly send out Twitter spam that advertise the app. The IP addresses that host the scam are shown in Table 3. The majority of the victims were from the United States and Turkey. The premium service access prices cost 5–10 euros.

Table 3: Summary of Twitter Scam Infrastructure and Domains

Host Details	Number of Domains	Sample Domains
IP address: 68.178.255.65, ns1.dershanelerkapatilmasin.com, ns2.dershanelerkapatilmasin.com **Country:** United States **ASN:** 26496	35	askfollow.com, askfollow.net, bestfollow.info, worldfollowers.info, etc.
IP address: 208.109.108.124, ns1.ip-68-178-255-209.secureserver.net, ns2.ip-68-178-255-209.secureserver.net **Country:** United States **ASN:** 26496	26	bestfollowers.org, biturlx.com, bulkfollowers.co, utf8more.info, etc.
IP address: 69.175.70.173, ns05.domaincontrol.com, ns06.domaincontrol.com **Country:** United States **ASN:** 32475	26	15c.info, azmh.info, cefpua.info, yigm.info, etc.
IP address: 172.70.175.69, 69.175.70.172, ns35.domaincontrol.com, ns36.domaincontrol.com **Country:** Namibia, United States **ASN:** 32475	5	followback.info, hitfollow.info, letgetmorefollowers.info, newfollow.info, plusfollower.info

Table 3: Summary of Twitter Scam Infrastructure and Domains		
Host Details	**Number of Domains**	**Sample Domains**
IP address: 54.225.82.214, ns75.domaincontrol.com, ns76.domaincontrol.com **Country:** United States **ASN:** 14618	7	ferrastudios.com, followmania.co, followmania.com, unfollow.ferrastudios.com, etc.

At the end of January 2014, we saw a spike in the number of users attempting to visit sites involved with scams as shown in Figure 4. Hundreds of users attempted to access domains that contained instructions that, if followed, would cause their Twitter accounts to be compromised. Figure 4 also shows the distribution of users that were targeted by this scam, the majority of whom were from the United States. A significant number also came from Turkey, most likely because of the keyword "takip" in some of the domains, which means "follow up" in Turkish. Most of the Web pages' contents are written in English so U.S. users could be their primary targets.

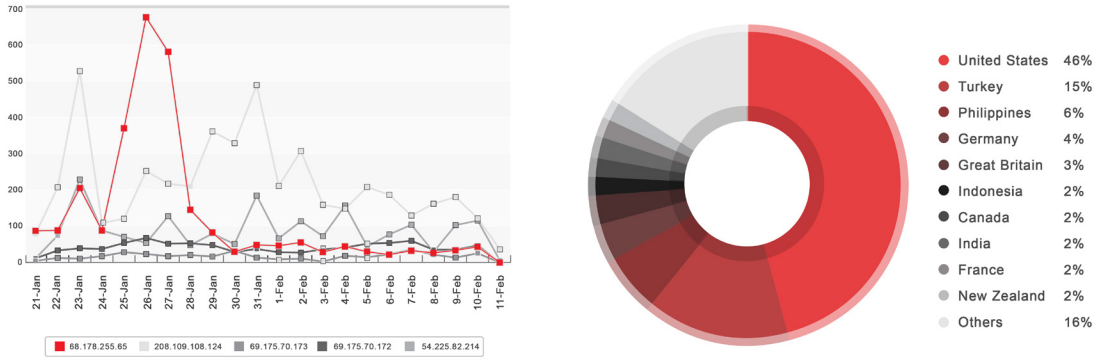

Figure 4: Impact of Twitter scams from January to February 2014

Users must be cautious of providing third-party apps access to their Twitter accounts (see Figure 5). If they have been victimized by the scam above, they should revoke the malicious apps' access rights through their settings.

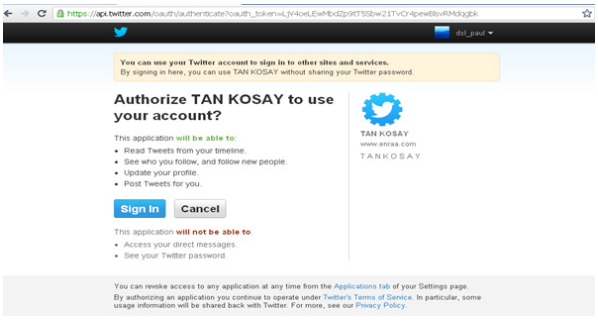

Figure 5: Authorizing Twitter-related apps

IMPACT ANALYSIS OF CLICK-THROUGH DATA

Previous studies on email spam found that click-through and conversion rates considerably varied.[9, 10, 11] The estimate click-through rates (i.e., the number of people who arrive at the website having clicked the link in the email) ranged from 0.003% to 0.02%.[9, 10] The 2010 study on Twitter spam estimated the click-through rate at 0.13%, which suggests that the click-through rate for Twitter spam was two times higher in magnitude higher than for email spam.[1]

The Trend Micro Web Reputation Technology has a component that allows users to obtain malicious anonymized feedback if they wish to.[3] We examined the feedback data to determine which malicious URLs embedded in tweets were clicked. However, without access to the platform's backend infrastructure, it was difficult to determine the absolute Twitter spam click-through rate. However, we were able to sensibly compare the relative effectiveness of malicious campaigns and determined that there was great variability across campaigns.

We classified the groups and domains we analyzed in Section V into the following categories:

- **Malware:** Tweets with embedded links that led to malware-distribution websites.
- **Traditional phishing:** Tweets with embedded links that led to phishing websites.
- **Twitter-specific scam:** Tweets that led to the Twitter follower scam described in Section VI.
- **Spam:** Tweets that were sent by groups or domains involved in spam distribution. We split this category into three subcategories because the different spam flavors had distinct characteristics. The subcategories include the following:
  - Traditional spam
  - Spam with shortened URLs
  - Russian spam, including the most prolific type, Russian-138 spam, described in Section VI
  - Spam related to a viral Japanese campaign

There were enormous variations in the effectiveness of the different approaches to Twitter spamming. For example, the viral Japanese campaign was approximately 5,000 times more effective than the Russian spam campaign.

Table 4: Clicks per Tweet

Abuse Category	Clicks per Tweet
Viral Japanese spam campaign	0.26862

Table 4: Clicks per Tweet	
Abuse Category	**Clicks per Tweet**
Malware	0.03065
Traditional phishing	0.00959
Spam with shortened URLs	0.00388
Spam	0.00239
Twitter-specific scam	0.00112
Russian spam	0.00005

Viral Japanese Spam Campaign

The viral Japanese spam campaign continued until February 2014. The vast majority (99%+) of users that were victimized were Japanese users.

Malware Tweets

While conducting the study, we witnessed an outbreak of Arabic tweets with embedded links that led to malware-laden websites. The majority of the affected users were from Saudi Arabia, Egypt, and Sudan, followed by the United States (see Figure 6).

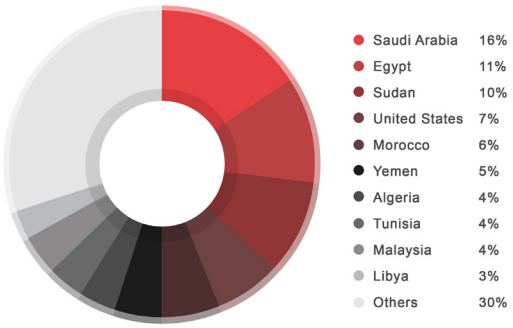

Figure 6: Distribution of clicks that led to malware-laden websites

Traditional Phishing Tweets

The traditional phishing tweets are similar to phishing emails. The tweets attempt to convince users that they came from legitimate users. As shown in Figure 7, the phishing tweets we studied primarily targeted users in the United States.

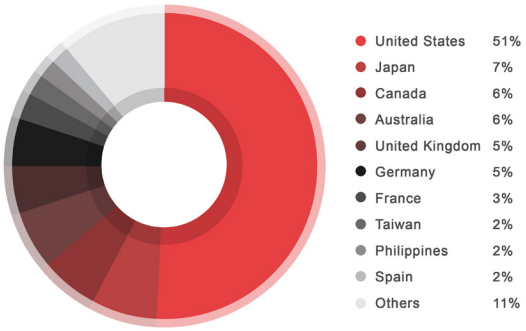

Figure 7: Distribution of clicks that led to phishing websites

Spam with Shortened URLs

A range of URL shorteners and proxy-avoidance domains were also used to further obscure links in tweets. This issue was discussed at length in the 2010 study

on Twitter spam.[1] Within the study period, apart from the commonly abused bit.ly shortener, we also saw URL shorteners such as 17q.org, bitlyjmp.com, kisalink.tk, lima.pp.ua, qwapo.es, redir.ec, shortredirect.us, and shortn.me used in malicious tweets. The distribution in Figure 8 reflects the use of region-specific URL shorteners such as kisalink.tk and qwapo.es in some outbreaks.

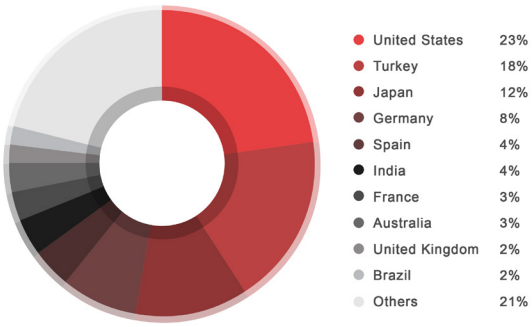

Figure 8: Distribution of clicks for tweets with shortened URLs

Traditional Spam

The distribution of traditional spam attacks (shown in Figure 9) primarily focused on U.S. users. We saw a large-scale health spam outbreak within the study period.

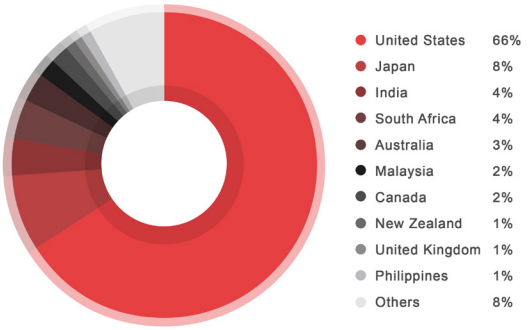

Figure 9: Distribution of clicks for traditional Twitter spam

Twitter-Specific Scams

We discussed the impact of Twitter follower scams in Section VI.

Russian Spam

The majority of users who clicked links embedded in Russian spam (shown in Figure 10) were from Russia (50%). However, many users from non-Russian-speaking countries also clicked links in this kind of spam. We theorize that the contents advertised in this spam type sufficiently appealed to some users (e.g., cracked software and games, free movies, cracks for mobile devices, exam and homework answers, etc.) so they use automated translation tools to access inappropriate content.

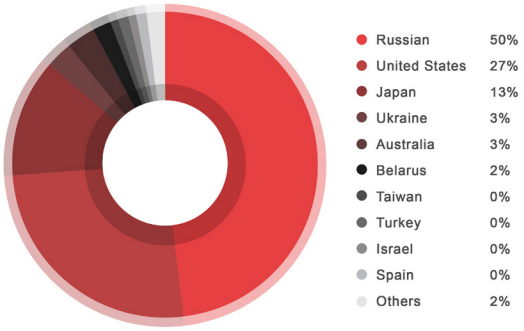

Figure 10: Distribution of clicks related to Russian spam

IMPACT OF TWITTER PHISHING

In Section II, we briefly described a Twitter-specific phishing scheme that has been going on for some years now.[2] We will discuss how such a scheme impacted Twitter and its users. This and similar schemes exploit the following features of Twitter in order to spread:

- They use URL shorteners.
- They have complex infection chains.
- The phishing tweets were sent out via accounts that have been compromised.

Figure 11: Typical infection chain for a Twitter phishing scheme

In Figure 11, we considered the final page in the infection chain as the "phishing landing page."

We approached this scheme from two ends—we determined how many posts on Twitter matched our phishing criteria and how many users attempted to load the phishing landing pages. We studied one particular scheme from March to May 2014.

The largest outbreak we monitored occurred on 15–19 March 2014. On 18 March 2014, we identified 22,282 compromised users who sent out phishing tweets with 13,814 distinct shortened URLs. On 19 March 2014, we identified 23,372 compromised users who sent out phishing tweets with 5,148 distinct shortened URLs. The shortened URLs described here were confirmed to have infection chains that ended with phishing landing pages.

We tracked the number of users who landed on phishing websites and what countries they came from within the study period (see Figures 12 and 13). Throughout the study period, we noticed changes in cybercriminal tactic. In mid-March, we saw an ongoing attack develop into sporadic outbreaks in May. In March and April, the phishing landing pages had literal IP addresses as URLs while the attacks in late May used more socially engineered host names using free Web-hosting services.

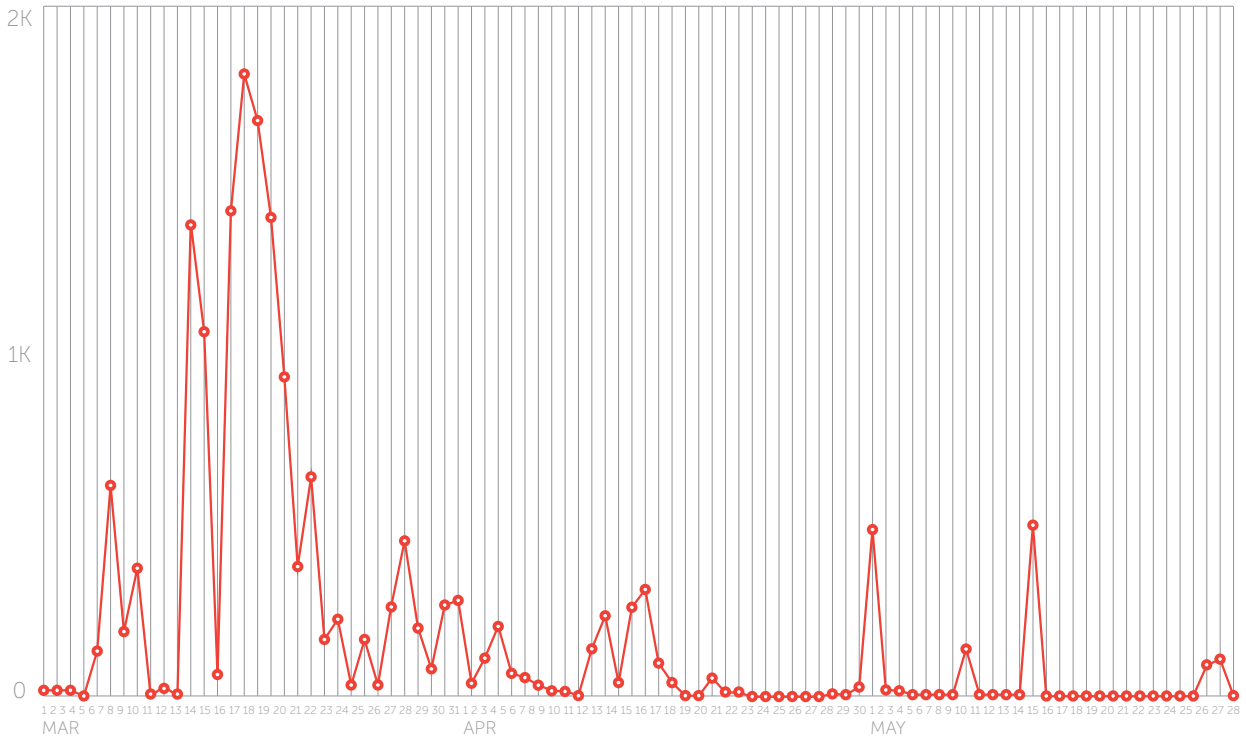

Figure 12: Number of users who attempted to access phishing landing pages

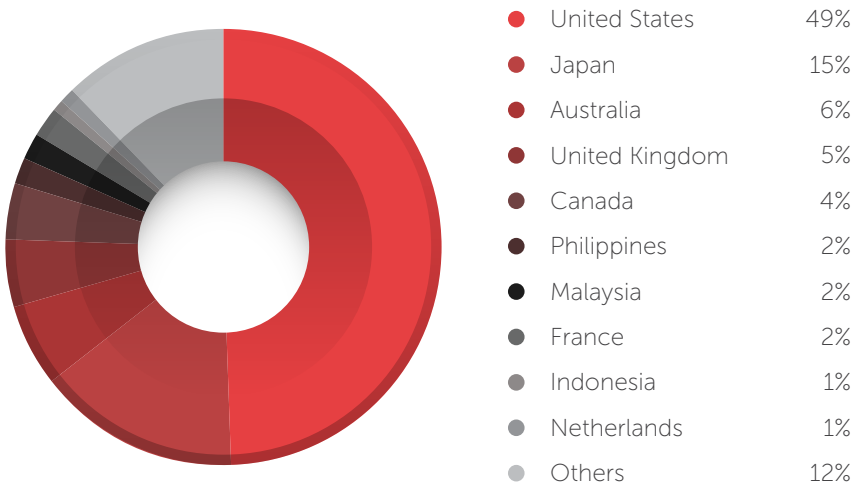

Figure 13: Distribution of users who attempted to access phishing landing pages

CONCLUSION

This research paper presented a study of various types of abuse on Twitter. We analyzed 500 million tweets with embedded URLs and found that, during a period of high spam activity, 5.8% of them were spam or malicious in nature.

We applied a hybrid technique of combining a blacklist augmented by algorithms suited for social networks to the problem of identifying spam and malicious tweets that proved reasonably effective. The blacklist was augmented with a clique-discovery approach, which also very effectively identified large-scale spam outbreaks. We came to a different conclusion—that blacklists when augmented in this way are a useful tool in uncovering Twitter spam.

We examined the response rates for various types of Twitter spam and found that they widely varied, depending on the spam's content and other factors. We therefore conclude that quoting a single response rate for Twitter spam is inadequate; it is important to quote response rates for each type of spam instead.

We also examined the regional response rates for various Twitter outbreaks and found that they greatly differed across countries and regions.

REFERENCES

1. Chris Grier, Kurt Thomas, Vern Paxson, and Michael Zhang. (2010). "@spam: The Underground on 140 Characters or Less." In Proceedings of the 17th ACM Conference on Computer and Communications Security, pages 27–37. Last accessed May 20, 2014, http://www.icir.org/vern/papers/ccs2010-twitter-spam.pdf.

2. Biz Stone. (February 26, 2010). *Twitter Blog.* "Avoid 'Phishing' Scams." Last accessed May 21, 2014, https://blog.twitter.com/2010/avoid-phishing-scams.

3. Trend Micro Incorporated. (2014). *Smart Protection Network—Data Mining Framework.* "Key Components." Last accessed June 2, 2014, http://cloudsecurity.trendmicro.com/us/technology-innovation/our-technology/smart-protection-network/#key-components.

4. Wikimedia Foundation Inc. (May 4, 2014). *Wikipedia.* "Clique Problem." Last accessed May 21, 2014, http://en.wikipedia.org/wiki/Clique_problem.

5. Yun-Chian Cheng. (October 2012). "Hadoop Success Stories in Trend Micro SPN." Presented During the Hadoop in Taiwan Workshop. Last accessed May 21, 2014, http://www.gwms.com.tw/TREND_HadoopinTaiwan2012/1002download/04.pdf.

6. Jingen Xiang, Cong Guo, and Ashraf Aboulnaga. "Scalable Maximum Clique Computation Using MapReduce." Last accessed May 21, 2014, https://cs.uwaterloo.ca/~ashraf/pubs/icde13maxclique.pdf.

7. Michael Steven Svendsen. "Mining Maximal Cliques from Large Graphs Using MapReduce, Masters Thesis." Last accessed June 3, 2014, http://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=3631&context=etd.

8. Paul Pajares. (January 30, 2014). *TrendLabs Security Intelligence Blog.* "Does the Twitter Follower Scam Actually Work?" Last accessed May 21, 2014, http://blog.trendmicro.com/trendlabs-security-intelligence/does-the-twitter-follower-scam-actually-work/.

9. Chris Kanich, Christian Kreibich, Kirill Levchenko, Brandon Enright, Geoffrey M. Voelker, Vern Paxson, and Stefan Savage. (2008). "Spamalytics: An Empirical Analysis of Spam Marketing Conversion." In Proceedings of the 15th ACM Conference on Computer and Communications Security, pages 3–14. Last accessed May 21, 2014, http://www.icsi.berkeley.edu/pubs/networking/2008-ccs-spamalytics.pdf.

10. Alex Mindlin. (July 3, 2006). *The New York Times.* "Seems Somebody Is Clicking on That Spam." Last accessed May 21, 2014, http://www.nytimes.com/2006/07/03/technology/03drill.html?_r=2&.

11. Josh Catone. (November 11, 2008). *SitePoint.* "Spam ROI: Profit on 1 in 12.5m Response Rate." Last accessed May 21, 2014, http://www.sitepoint.com/spam-roi-profit-on-1-in-125m-response-rate/.

12. Eva Zangerle and Günther Specht. (2014). "'Sorry, I Was Hacked': A Classification of Compromised Twitter Accounts." Last accessed May 22, 2014, http://www.evazangerle.at/wp-content/papercite-data/pdf/sac14.pdf.

13. Manuel Egele, Gianluca Stringhini, Christopher Kruegel, and Giovanni Vigna. (2013). "COMPA: Detecting Compromised Accounts on Social Networks." In ISOC Network and Distributed System Security Symposium (NDSS). Last accessed May 22, 2014, http://cs.ucsb.edu/~gianluca/papers/

thjp-ndss13.pdf.

14. Chung-Tsai Su, Wen-Kwang Tsao, Wei-Rong Chu, and Ming-Ray Liao. (2012). "Mining Web Browsing Log by Using Relaxed Biclique Enumeration Algorithm in MapReduce." In Volume 3, pages 54–58, IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology. Last accessed May 22, 2014, http://www.computer.org/csdl/proceedings/wi-iat/2012/4880/03/4880c054-abs.html.

Trend Micro Incorporated, a global leader in security software, strives to make the world safe for exchanging digital information. Our innovative solutions for consumers, businesses and governments provide layered content security to protect information on mobile devices, endpoints, gateways, servers and the cloud. All of our solutions are powered by cloud-based global threat intelligence, the Trend Micro™ Smart Protection Network™, and are supported by over 1,200 threat experts around the globe. For more information, visit www.trendmicro.com.

TREND MICRO™

Securing Your Journey to the Cloud

225 E. John Carpenter Freeway, Suite 1500
Irving, Texas 75062 U.S.A.

Phone: +1.817.569,8900