Can We Trust Social Media Data?
Social Network Manipulation by an IoT Botnet

Masarah Paquet-Clouston
GoSecure Research
800 Boulevard René-Lévesque Ouest
#1860, Montréal, QC H3B 1X9
mcpc@gosecure.ca

Olivier Bilodeau
GoSecure Research
800 Boulevard René-Lévesque Ouest
#1860, Montréal, QC H3B 1X9
obilodeau@gosecure.ca

David Décary-Hétu
Université de Montréal
2900 Boulevard Edouard-Montpetit,
Montréal, QC H3T 1J4
david.decary-hetu@umontreal.ca

ABSTRACT

The size of a social media account's audience – in terms of followers or friends count – is believed to be a good measure of its influence and popularity. To gain quick artificial popularity on online social networks (OSN), one can buy likes, follows and views, from social media fraud (SMF) services. SMF is the generation of likes, follows and views on OSN such as Facebook, Twitter, YouTube, and Instagram. Using a research method that combines computer sciences and social sciences, this paper provides a deeper understanding of the illicit market for SMF. It conducts a market price analysis for SMF, describes the operations of a supplier – an Internet of things (IoT) botnet performing SMF – and provides a profile of the potential customers of such fraud. The paper explains how an IoT botnet conducts social network manipulation and illustrates that the fraud is driven by OSN users, mainly entertainers, small online shops and private users. It also illustrates that OSN strategy to suspend fake accounts only cleans the networks a posteriori of the fraud and does not deter the crime – the botnet – or the fraud – SMF – from happening. Several solutions to deter the fraud are provided.

CCS Concepts

Human-centered computing → Collaborative and social computing → Collaborative and social computing theory, concepts and paradigms → Social media • Security and privacy → Intrusion/anomaly detection and malware mitigation → Malware and its mitigation • → Human and societal aspects of security and privacy

Keywords

Social media fraud (SMF); Online social networks (OSN); IoT botnets; Market analysis.

#SMSociety'17, July 28-30, 2017, Toronto, ON, Canada © 2017 Copyright is held by the owner/author(s). ACM ISBN 978-1-4503-4847-8/17/07.
http://dx.doi.org/10.1145/3097286.3097301

1. INTRODUCTION

Online social networks (OSN) are primary outlets for many activities such as advertising, personal communications, news broadcasts, political announcements and advocating social causes. They now engage a large portion of the world's population, making it possible for individuals, companies and governments to reach a large audience through the acquisition of a fan base, also known as 'followers' and/or 'friends'. In most cases, attracting new followers and friends is done by publishing interesting content online. In some cases, however, actors elect to buy their fan base, a strategy that is part of social media fraud (SMF). SMF is the process of creating likes, follows, views or any other online actions on OSN like Facebook, Twitter, YouTube and Instagram to artificially increase an account's fan base. This method falsifies social media data and creates disinformation that could lead to a decrease in users' trust in OSN. This paper studies the illicit market where SMF services are bought and sold to better understand the potential impact of that market on OSN. With a research method that combines computer sciences and social sciences, this paper evaluates the supply and demand for SMF services. The supply analysis is two-fold: a market price analysis for SMF services and an in-depth evaluation of the operations of an IoT botnet acting as a supplier in the market. The demand analysis contains a profiling of 'potential customers' of SMF, retrieved from accounts found in the IoT botnet communications. The results provide an in-depth understanding of the extent to which social media data can be trusted and who contributes to falsifying it.

The following text presents a literature review of what is known about SMF and social network manipulation by botnets. Then the research methodology is developed, followed by the result and the discussion section.

2. LITERATURE REVIEW

OSN have been the target of various malicious activities such as identity theft [3], spam campaigns [22] and political online manipulation [12]. Those behind the malicious activities exploit the trust relationship between OSN users to manipulate, distort, utilize or steal OSN data [19]. One malicious activity that contributes to online disinformation, through its manipulation and distortion OSN data, is SMF. SMF is the process of creating likes, follows, views, or any other online actions on OSN, to artificially grow an account's fan base. This fraud aims to increase the digital influence of an OSN account [22].

Indeed, OSN are rapidly becoming the preferred vehicle for marketing campaigns allowing individuals, companies and governments to increase their visibility online. To increase brand recognition, marketers continuously develop interactive techniques that seek to enhance the number of likes, followers or comments of their clients' accounts [10]. The size of an account's audience – in terms of followers count – is believed to be a good measure of an account's influence and popularity [14,19]. A substantial amount of time and resources need to be invested to develop a large fan base on OSN and this time-consuming process is not necessarily accessible to everyone.

To create an illusion of popularity at relatively low cost, some account owners have resorted to buying likes, follows, and views. Of course, having thousands of followers on an OSN does not directly translate into influence over the followers [8], but it does create an illusion of popularity that an account's owner can leverage. Search engines can point account owners to many SMF service providers that will happily provide 10,000 followers on Twitter for a price ranging from $40 to $216 [19]. This price seems relatively inexpensive when compared to the money and time required to organically grow a fan base of 10,000 followers.

SMF services rely on one of two strategies to gain control over the accounts they use to artificially increase the follower count of their clients: 1) compromising real existing accounts or 2) creating new and fake accounts. The former is achieved by luring users to click on links with the promise of free followers but instead compromising their account credentials and using them to like, follow or view other accounts [19]. The latter consists of creating new fake accounts and using them to generate follows, likes or views [9]. Large numbers of fake accounts for SMF can be created by using either click farms or botnets. Click farms consist of large groups of low-paid laborers, usually settled in developing countries, hired to conduct SMF on demand [16]. Botnets, on the other hand, are groups of compromised computer systems remotely controlled by a third party, a bot master. Using the compromised systems, fake accounts are created automatically and used to perform likes, follows and views. This automated process increases the potential scope of SMF as tens if not hundreds of thousands of compromised systems can like, follow, or view other accounts as well as post on OSN, all with very little cost since the systems used to conduct SMF are compromised and not owned by the fraudsters.

Such network of fake OSN accounts are also known as social bots, which are "software-controlled OSN accounts that mimic human users with malicious intentions" [23] (p.46). Social bots on Twitter for either spam distribution or digital influence manipulation -including SMF- have been known to be very efficient [23]. Yet, to prevent the creation of thousands of fake accounts, OSN have developed several techniques to avoid large, automated account creations by botnets. Such techniques include the use of CAPTCHAS, phone number verification and IP blacklisting [20] [21]. These strategies are not foolproof, however, since offenders can rely on "CAPTCHA solving services; fraudulent email credentials from Hotmail, Yahoo, and mail.ru; and tens of thousands of hosts located around the globe to provide a diverse pool of IP addresses to evade blacklisting and throttling" [21]. To bypass the phone verification required by OSN during an account's creation, offenders have used Voice Over Internet Protocol (VoIP) services, which reduce the costs associated with buying and maintaining multiple SIM cards [21]. While automated registration barriers raised by OSN are routinely circumvented by offenders, they are believed to represent one of the biggest challenge that offenders face when conducting SMF [21]. Indeed, once fake accounts are created, automating the browsing and actions on OSN appears to be relatively safe as demonstrated by researchers who simulated a social bot and showed that they could successfully log in and interact with the Twitter network [11]. In 2011, other researchers [6] built an automated social botnet on Facebook and found that the OSN could be infiltrated successfully at a rate of up to 80%.

To prevent malicious activities on OSN, researchers have also developed classifiers, tools, and applications with the help of machine learning techniques and network connectivity information in order to detect patterns of malicious activity on OSN [7]. An example of this is the "CopyCatch" software tool, which studies how people like pages made by spammers on Facebook to detect SMF based on the time the likes were created [2]. Another example is the "Synchrotap" tool, which seeks to identify accounts that perform loosely synchronized actions such as uploading many spam photos with a small set of IP addresses or quickly inflating the number of followers by targeting a set of users in batches [7]. Researchers have also created a system that uses supervised machine learning techniques to dynamically detect fake Twitter accounts based on their characteristics. Fake accounts usually have no face in their profile picture, write fewer words when posting links and have a lower number of followers to following ratio. Recently, researchers developed a strategy to retroactively detect fake profiles out of the 62 million public accounts available on Twitter "using a pattern-matching algorithm on screen-name with an analysis of tweet update times" [14]. They found that the profile creation time and promoted URLs of fake accounts were significantly different from those of legitimate users. Finally, a study directly compared the behavior of fake accounts on Facebook with legitimate users and found that fake accounts "tend to often re-share content, use fewer words and poorer vocabulary, and more often generate duplicate comments and likes compared to normal users" [15]. Its authors developed a classifier based on the behavior of fake accounts and achieved 99% precision and 93% recall. Taken as a whole, these studies demonstrate a real effort by the research community to detect and block botnets' operations on OSN. They also demonstrate the ability of SMF service providers to circumvent the defensive mechanisms put in place by OSN.

SMF, no matter how innocuous it may look, deserves further attention by the research community for three reasons. First, SMF incurs important costs to OSN companies, who need to develop algorithms and tools to detect fraudulent activities. Second, it creates biased data and artificially boosts the popularity of certain accounts. This generates distortions and drives disinformation. As such, a company falsely endorsed by 100,000 followers may look legitimate when, in fact, it is not. When people are fooled and misinformed, their trust in the signals established by OSN decreases, lessening the overall value of the entire OSN ecosystem. Third, the SMF supply is driven in part by botnets that pollute the Internet infrastructure. Overlooking the infection of thousands of systems by a botnet because its end-activity seems harmless is a risky proposition in the long term. Indeed, the botnet may end up being used for other illicit activities such as launching distributed denial of service (DDoS) attacks or stealing credentials.

To better understand the illicit market for SMF, this paper evaluates the supply for such fraud. A market analysis of the price for SMF on OSN is first presented. Then the operations of a supplier, an IoT botnet conducting such fraud, are described. To do so, real-time intelligence was gathered on a botnet conducting such fraud by launching an active man-in-the-middle attack against the botnet's secure traffic. This paper is the first attempt by researchers at gathering information on the inside of a botnet conducting SMF, providing an exclusive view of how SMF is achieved. The data extracted from the traffic also provides information on the potential customers of SMF. Profiles of such consumers are thus developed to better understand those that buy fraud – the market's demand side. Understanding this illicit market through a global perspective – using both the supply and the demand side – is the first step in effectively addressing SMF.

3. METHODOLOGY

The methodology section is divided into three subsections. The strategy behind the open-source data collection of the prices for SMF services is presented. The IoT botnet, monitored over several months in 2016, is then introduced. Next, how the honeypots were built and the man-in-the-middle attack conducted to decrypt the traffic is explained. This traffic data exposes the botnet's operations and the information about its potential customers. The last subsection presents the strategy for profiling the accounts belonging to potential customers.

3.1 Data Collection on the Price of SMF

In January and February of 2016, 83 websites advertising SMF and one freelancer platform – a website where many self-employed individuals post ads for SMF services – were identified. A search on Google.com for the keywords "buy likes" and "buy followers" was used. The website data was collected manually given the small size and the difficulty associated with building an automated tool that could understand the different interfaces of each website. In the case of the freelancer platform, an automated web-crawler was used to gather the information. The platform had a social marketing section grouping all its ads for SMF services. The data collected with the web-crawler was validated manually.

Most websites offered a large variety of SMF services targeting a variety of OSN. A comprehensive list of the SMF services and their prices was built and is presented below. The information gathered included the type of SMF service offered, its price and the number of followers or likes, included for that price. A total of 5,687 ads were collected for various services offered on popular OSN. This dataset is used to analyze the market for SMF services, comparing the price for 1,000 follows or likes on Facebook, Instagram, Twitter and YouTube.

3.2. Catching Linux/Moose: The IoT Botnet that Conducts SMF

The IoT botnet studied in this research is called *Linux/Moose* and was first investigated by the ESET research team in 2015 [5]. Linux/Moose is still active and targets routers and IoT devices, such as smart TVs. It uses them as proxies to connect to OSN (for technical information on the structure of the compromised network, see [5,11,17]). As Linux/Moose is a botnet primarily used for its proxy service, intelligence about its operations can be gathered by intercepting the communications between the command and control (C&C) servers and the infected systems.

3.2.1 Building Honeypots

Rather than monitor infected systems in the wild, we created and infected our own 'honeypots', which are virtual environments specifically designed to mimic a system vulnerable to Linux/Moose. These honeypots were indifferentiable from real consumer routers from the botnet's point of view and designed to securely log all their activities. The honeypot creation process was facilitated using the Cowrie honeypot project which already includes most of the functionality necessary to emulate Linux/Moose's targets. This architecture enabled system monitoring from outside of the compromised operating system. It also protected the infrastructure from tampering attempts by the botnet operators and enabled the deployment of honeypots in different geographic regions of the world.

Since the botnet uses the Telnet protocol to spread – by guessing usernames and passwords – the honeypot needed to support that protocol to parse and interpret the command-line interactions between the C&C and the infected system. Telnet protocol support was developed for Cowrie as part of the research project [4]. An additional difficulty was faced regarding honeypot infection. For a honeypot system to become part of the Linux/Moose botnet, its infecting system must indicate to the C&C servers that the infection was successful. This happens during the infection process where a contact is made to the C&C servers with a custom protocol. This was a new behavior that was not witnessed in 2015 by the ESET research team where infection simply consisted of executing the malicious binary. The infection mechanism was therefore adapted (for more technical details on the mechanism, see [17]) and 12 systems were successfully infected. These honeypots relayed the botnet's traffic between January and December 2016 at different time periods and with varying levels of activity. The proxy traffic consisted in majority of encrypted HTTP traffic commonly referred to as HTTPS.

3.2.2 Man-in-the-Middle Attack

The HTTPS traffic our honeypots relayed gave us no visibility into its content due to its encryption. Therefore, to gather further information on the botnet's activity, Linux/Moose's traffic needed to be decrypted.

We launched a man-in-the-middle attack against the connection between the C&C servers and our honeypots. A 'man-in-the-middle attack' is an attack where a third-party intercepts, but still relays, connections that were intended to be direct between two endpoints. HTTPS is designed to prevent this type of attack through the use of certificates for authentication and the use of verification chains that are validated by browsers. To circumvent this control, we created our own certificate authorities and used them in our man-in-the-middle attack tool. Such an attack raises a "Certificate Error" message informing the user that the connection certificate is not valid. It appears that the botnet operators were not logging or paying attention to these warnings since our untrusted certificates were accepted to encrypt the traffic. This is not particularly surprising, as certificate error warnings are common and often ignored whenever a Web task is automated. Running a large-scale botnet requires this type of automation. At this point, given that we are providing the encryption keys (via the certificates) we can decrypt the traffic going through our honeypots.

The attack provided valuable information on the botnet's activities on OSN, such as the name of the fake accounts used by the botnet,

its modus operandi in conducting SMF and the identities of potential consumers. Overall, the 12 infected honeypots relayed 273,776 requests and 86% of them were directed toward the Instagram social network. Given its prevalence, only the traffic going to and from Instagram is studied in this research.

3.2.3 Analyzing the Botnet's Operations

To analyze the decrypted traffic and identify the botnet patterns, a number of Python libraries including SciPy, NumPy, Panda and mitmproxy were used. Each pattern is a series of requests sent to Instagram to perform an action, such as creating an account or following an account. Requests patterns used by the botnet were identified by searching through the network traffic logs. The findings are presented in Section 4.2 and are complemented with a discussion of the techniques employed by the botnet to avoid detection by OSN. Patterns also provide information on the botnet's fake accounts, as the bots need to be logged in to a fake account to perform SMF. A generic profile of the botnet's fake accounts was developed by looking at the fake accounts' username, name, description, number of posts, number of followers, number of followings, and photos.

3.3 Profiling Potential Customers

Some of the requests relayed by the infected systems sought to follow an account on Instagram. They provided the Instagram ID of the account followed, which could belong to a potential customer of the botnet.

To discriminate between the botnet's customers and the accounts that were followed to legitimize the fake accounts, three criteria were used. First, several celebrities followed by the botnet could not be considered "potential customers" because the botnet operators may like or follow celebrities to avoid being detected by Instagram, acting as a "normal user" would. Thus, any accounts that had a "verified badge" given by Instagram or more than one million followers were excluded from the list of potential customers. Second, to avoid adding to the list of customers any accounts on which a bot randomly performed a follow, more than five follows had to be performed on an account for it to be considered. Third, the ratio between the average number of likes of posts and the number of followers was assessed. An account with 150,000 followers who posted a picture and had 50 reactions was a signal that the followers were likely false and that the owner of the account was a buyer of SMF services. This third criterion was more of a final check, to ensure that no false positives were added to the list.

A list of all accounts on which Linux/Moose performed more than five follows was retrieved from the decrypted traffic. Then, public information on these accounts was gathered. Overall, Linux/Moose traffic identified 765 accounts on which more than five follows were performed by the end of December 2016. From these 765 accounts, 15 were private, and 55 had been deleted or were no longer available. Moreover, ambiguous profiles, with random pictures and incomprehensible descriptions, were ruled out, forcing the elimination of 134 more profiles. Finally, 34 profiles had a verified badge and more than one million followers and were therefore eliminated from the list. The list of potential customers included a total of 522 accounts. These 522 accounts all had few reactions to their posts compared to their number of followers.

To develop a profile of the potential customers of the botnet, a content analysis of the account description was conducted. The qualitative analysis was the only option because the account descriptions were in various languages and translation was required. First, the language used was coded and the profile description was translated, if needed. Second, categories were created according to the profile description. To avoid any bias in the data analysis, the first category was created based on the result of the first account analyzed and more categories were created as the analyses of new accounts were completed. The categories were created throughout the analysis until the last potential customer was analyzed, yielding the creation of a total of 75 categories.

3.4 Ethical Considerations

Three ethical considerations related to this research need to be mentioned. First, while building the honeypots, many measures were undertaken to avoid infecting other systems and preventing the botnet from spreading further. To do so, outbound Telnet connections were blocked on the honeypot environment outside of the QEMU virtual machine that was infected. This was a reliable way to prevent the Linux/Moose worm from spreading as it uses Telnet for propagation. Second, the honeypots conducted SMF during a specific period. However, since one of the aims of this research is to understand the operations of a supplier conducting the fraud, such externalities were inevitable. The hope is that the outcome of this research will surpass the harm to OSN. Third, even though all of the information gathered on potential customers was publicly available on OSN, all data presented in the results section was anonymized. The names and profiles of potential customers are therefore never disclosed.

4. RESULTS

This section begins with a presentation of the distribution of the prices for SMF services, followed by a description of the Linux/Moose botnet operations. It ends with a review of the profiles of SMF customers found in the botnet's traffic.

4.1 The Price of Social Media Fraud

SMF services really are only a few clicks away. Indeed, hundreds of websites specializing in offering these fraudulent services are accessible through a search for "buying likes," or "buying followers" on popular search engines. Also, several freelancer platforms, where self-employed individuals can advertise their services, also display specific sections where SMF services are advertised.

Most of the websites advertising SMF services offer diversified sets of services for various OSN including tweets on Twitter or likes on Instagram. On the other hand, ads on freelancer platforms are more specific since they generally target only one OSN: "I will Add 3000 Permanent FACEBOOK Likes" (Seller 48). Also, some services are more expensive than others; for example, buying a comment is more expensive than buying a like. Some OSN and the size of the service – number of likes – also impact prices. Generally, the price per endorsement, such as the price per like, decreases as the quantity offered increases.

Table 1 presents the average price for 1,000 follows and likes for four popular social networks: Facebook, Instagram, Twitter and YouTube.

Table 1. Price of follows and likes on the illicit market for SMF

	$ USD /1,000 *follows*		**$ USD / 1,000 *likes***	
	Mean (Std.)	Median	Mean (Std.)	Median
Facebook	$34 ($21)	$29	$28 ($21)	$20
Instagram	$16 ($13)	$13	$21 ($18)	$14
Twitter	$15 ($12)	$12	$24 ($26)	$15
YouTube[a]	$49 ($15)	$51	$52 ($31)	$50

[a] *Followers are subscribers in the case of YouTube*

Table 1 shows just how dispersed the price distribution for 1,000 follows or likes is. On Facebook, the mean price for 1,000 follows is $34, but the standard deviation of $21 shows that the price for the service varies greatly. The large standard deviation is found in almost all mean prices for likes and follows and for all OSN. The median is thus a better indicator of the price for SMF services. Competitive markets are usually markets with similar prices, as each seller tries to catch a share of the market. The price variation in this market may indicate that it is still young and immature. Sellers may not know the worth of the services they offer; some of them therefore overprice while others underprice. Moreover, the fact that the service is fraudulent and buyers have no legal recourse when deceived may attract scammers. Scammers may price SMF services without considering the real costs related to providing them, creating false signals in the market.

Table 1 also displays important differences in pricing for each OSN. The median prices for 1,000 follows and likes on Instagram are respectively $13 and $14 whereas the same median prices for YouTube are $51 and $50. SMF services targeting Instagram therefore seem to be advertised as cheaper than SMF services targeting YouTube. To verify the hypothesis that there are significant differences in the mean price between each OSN, an analysis of variance (ANOVA) was conducted on the mean price of follows. At 95% confidence level, the results indicate that there are significant differences in the mean prices between each OSN. The mean price of follows on YouTube is significantly different from – and higher than – the mean price of Facebook followers ($p<0.05$), Instagram followers ($p<0.05$) and Twitter followers ($p<0.000$). The mean price of Facebook followers is also significantly different from – and higher than – the mean price of Instagram followers ($p<0.05$) and Twitter followers ($p<0.05$). However, the mean price for 1,000 Instagram followers is not significantly different from the mean price for 1,000 Twitter followers.

The fact that the prices for similar services are different depending on the social network on which the fraudulent services are offered indicates that suppliers are practicing price discrimination. This price discrimination may be related to the challenges faced by suppliers when trying to conduct the fraud without detection. For example, since more subscribers, likes or views can be translated into a pay check for YouTube users, YouTube may invest more time and effort in detecting fraudulent actions on its OSN, increasing costs for SMF suppliers.

4.2 The Operations of a Supplier Conducting SMF

The analysis of Linux/Moose operations on Instagram yielded information on the bots' approach to registering accounts, on the profiles of the fake accounts, on the techniques the bots used to perform SMF and on their overall rate of success at avoiding detection.

4.2.1 Fake Account Creations

Registering new accounts is an essential part of SMF. Indeed, a customer ordering 6,000 follows expects 6,000 different accounts to follow his/her own account. Registration barriers are believed to be the highest challenges for offenders to circumvent [21]. The decrypted traffic showed that the bots registered accounts either through a web browser using the Instagram web application or through the Instagram mobile application. In both cases, the requests relayed to register accounts are the same as those sent from a normal user. Linux/Moose does not use any hacking technique and does not take advantage of security vulnerabilities to create accounts since all it needs to provide in order to register an account is an email address or a phone number. This email address is not verified at registration since no verified links are sent to the user. The botnet can thus randomly generate as many email addresses as it needs. Of the email addresses used for registration attempts, 67% came from large email providers such as Yahoo, Gmail, Outlook, Hotmail, AOL, Yandex.ru, and Mail.ru. The other 33% of email addresses used domains that were most likely generated on the fly. Below is a sample of emails that were used by the botnet:

Al***aspn*236[at]ozly[.]com
***kbo*m12700b[at]zvjmfl[.]jp
Chyn***th1*471p[at]ogxvf[.]org
Groe****elwub*nhwt[at]wzgvf[.]org
557*6dod***mb[at]coolsite[.]net
T*a***39[at]knpt[.]org
Da***rto*uvg[at]salesperson[.]net

Emails created on the fly can be used to register accounts on Instagram because of the lack of identity verification by the OSN. Since there is no need to confirm that the email address is valid, anything can be written up in the registration form.

4.2.2 Profiles of the Botnet's Fake Accounts

At the end of August 2016, a list of 1,732 fake accounts was built. However, 72% of these fake accounts had already been flagged and suspended by Instagram. From August to December 2016, 51 new fake accounts were found in the traffic and 92% were flagged and suspended by January 2017. These results suggest that Instagram manages to suspend most fake accounts associated to Linux/Moose and that its fake accounts have a very short life span of less than six months.

Information on the remaining 493 fake accounts that were not flagged by Instagram was retrieved online. Also, with the network traffic, information from 611 other fake accounts that had been suspended was found. The profile of the botnet's fake accounts is therefore based on an analysis of 1,104 fake accounts.

The profile pictures of the Linux/Moose fake accounts display mostly animals, plants, landscapes and buildings. The names of the fake accounts look legitimate and are often composed of a first and last name. They are used to create the fake accounts' username, which needs to be unique on Instagram. Linux/Moose's most common pattern for creating a username is simple and is shown in Table 2. It uses the fake accounts' first and last name and either adds random characters at the beginning of the full

name (Example A), in between the first and the last name (Example B) or at the end of the full name (Example C).

Table 2. Linux/Moose strategy for creating Instagram unique usernames

Example	Name	Unique Surname
A	Cirilli Jose	173epcirillijose
B	Bielak Koob	bielak09koob
C	Nunn Mizia	nunnmiziahp

About 47% of fake accounts had a short profile description. The description sometimes consisted of sentences such as "like honey you need to chill" or "I cry Don't disturb me." Looking at languages used, 41% of profile descriptions were written in English, 26% in Dutch, 6% in Spanish, 1% in Malay and fewer than one percent in Portuguese and German. Some profile descriptions did not make any sense, such as "iiiiii" or "fo", or consisted of numbers like "52314." The great majority of the fake accounts also had no posts at all.

The number of accounts followed by the fake accounts ranged between 0 and 822. The number of accounts following the fake account – their followers – was low: 72% of the fake accounts had no followers and 97% had fewer than three followers. The profile of the suspended accounts was not different from that of accounts that were never flagged. It is therefore not possible at this time to understand how or why some fake accounts were flagged while others were not.

This analysis shows that flagging Linux/Moose fake accounts should be possible by targeting accounts with a short profile description, no posts, a large number of followings but no follower. Linux/Moose appears to put very little effort into keeping its fake accounts alive on Instagram. This is perhaps due to the ability for the botnet to create new fake accounts at will.

4.2.3 Performing SMF

Once the fake accounts are created, Linux/Moose conducts SMF on Instagram. The decrypted traffic showed that Linux/Moose performs mostly likes and follows. No requests were made to post comments and very few requests were made to view videos. Further investigations showed that many requests not directly related to fake endorsements were needed to reduce the odds of detection by the OSN. Indeed, of the 184,952 requests sent to Instagram, only 4,199 (2%) performed a like and 18,720 (10%) performed a follow. This means that 88% of Linux/Moose's requests on Instagram were either meant to register an account or to browse through profiles, probably to reduce the odds of detection.

The decrypted traffic displayed different modus operandi used by the botnet to successfully conduct follows. These predetermined sets of requests were probably aimed at reducing detection by making the bots act more like humans. Two approaches were used to conduct follows: mobile application or web browser, which were characterized by the User-Agent field of the HTTP headers.

Via a mobile application, two different sets of requests were observed. Bots can take either a short path toward successfully following or a long path toward successfully following. The short path consists of six predetermined requests:

1- Searching for the targeted account to be followed;
2- Requesting basic information about the account;
3- Looking at a friendship status with the account;
4- Loading the feed of the account;
5- Requesting a list of other accounts the botnet could be interested in according the profile loaded;
6- Following the targeted account.

The long path consists of 15 predetermined requests:

1- Visiting own inbox
2- Looking at potential recipients to send a message
3- Visiting own personal timeline
4- Visiting own inbox
5- Looking at potential recipients to send a message
6- Visiting again own inbox
7- Loading the feed of the account of the targeted account to be followed;
8- Requesting basic information about the account;
9- Searching for the targeted account to be followed;
10- Searching again for the targeted account;
11- Requesting basic information about the account;
12- Looking at the friendship status with the account;
13- Loading the feed of the account;
14- Requesting a list of other accounts according to the profile loaded;
15- Following the targeted account.

The long path contains twice as many requests as the short one and they are used alternatively. More specifically, bots use the short path until the attempted follow is flagged and denied by Instagram. In this case, bots turn to the longer path until a follow is successfully conducted, and then revert to the short path until its requests are flagged again. The fact that Linux/Moose takes the short path whenever possible indicates that the operators want their botnet to be as efficient as possible by maximizing successful follows with the minimum number of requests required. When Linux/Moose connects to Instagram using a web browser, its modus operandi for creating a follow is much shorter: it consists of visiting the account once before trying to follow the account. The modus operandi for creating likes, both with web browser and mobile application, also consists of visiting the picture once before performing the like. All patterns are used alternatively, which shows that bots need to use these different paths when conducting SMF to prevent detection by Instagram.

4.2.4 SMF Success Rate on Instagram

Assessing Linux/Moose's success rate for the requests it makes on Instagram is tricky. Instagram spam detection strategies are private, limiting one's ability to test hypotheses. Honeypots were also hosted in cloud servers around the world, except for one, which was set up in the home router of one of the researchers. Instagram may therefore have been flagging the botnet's requests more often because they originated from cloud services.

Still, the results suggest that overall only 18% of the botnet's requests to create an account were successful. When considering only the traffic from the home router, the success rate jumped to 80%. Account creation attempts are therefore more successful when they originate from IP addresses that belong to a consumer Internet Service Provider (ISP). Requests to like a post or follow an account were successful 88% of the time overall. Those that originated from the home router indicate that the likes and follows were successful 99.9% of the time.

4.3 Profiling the Demand for SMF Services

As mentioned earlier, the profiles of 522 potential customers were evaluated and categorized. To the best of our knowledge, this is the first assessment of the demand for SMF services based on raw traffic from a supplier. It provides a first glance at those that buy from SMF services to artificially increase their popularity on OSN.

Among the pool of customers, 38% had a description in English, 35% in Arabic, 9% in Russian, 8% in Portuguese, and 6% in Spanish, Turkish and Indonesian. This suggests that SMF service customers have a very diverse origin.

Individuals involved in the entertainment industry (actor/actresses, singers, TV presenters), the modeling industry, and the photography industry made up 20% of the customer sample. Another 21% of the sample consisted of shops selling products. Within this group, more than one-third were online shops selling jewels, watches, shoes and clothes, and two-thirds were related to the sale of various products such as electronics, books, dental products and makeup. The range of products sold was quite broad, and included anything from bikes, to protein for hair, to animals. Accounts that belonged to individuals not involved in an industry made up 26% of the sample. These individuals posted pictures of their everyday life, their friends and their activities, while having a fan base of thousands of Instagram followers. Thus, more than one-quarter of the demand in the sample consisted of individuals wanting to boost their popularity on Instagram mainly for what seemed to be ego purposes. This shows that online popularity is important not only for business purposes but also for private users.

Among other categories found, 5% of the sample consisted of designers, 4% of bloggers or magazines, 4% of social media services ads and 2% of decoration or construction services. The other 18% consists of profiles, such as magicians, hypnotists, accounts for religious purposes, politicians and export consultants that could not be classified in larger categories.

Potential customers of SMF service seem to be entities – individuals, companies and entrepreneurs, that may not have the time and the resources required to gain a large fan base through sophisticated marketing strategies. Yet, since they value popularity on OSN, they turn to the illicit market for SFM to gain a large fan base at a low cost.

5. DISCUSSION

The results of this research illustrate that an IoT botnet is able to perform SMF on Instagram easily, monetizing a fraudulent service provided through a network of infected devices. Linux/Moose is capable of registering fake accounts on Instagram and this is made even easier by the lack of identity validation. Registration barriers are assumed to be an important challenge faced by botnet operators [21], but Linux/Moose focuses on a social network on which there are none, facilitating its operations. The botnet operators do not even need to rely on CAPTCHA solving services or Voice Over Internet Protocol (VoIP) as others do [20,21], as they can simply generate random email addresses on the fly. Adding email verification loops at the time of registration could significantly increase the challenges the Linux/Moose operators face and reduce their ability to massively create fake accounts. This would indeed force the botnet operators to create real email accounts and confirm the registration of each account created on Instagram by clicking a link sent to the account's mailbox. Although not insurmountable, this would harden the process of conducting SMF and likely increase the price of the fraudulent service [21].

Linux/Moose is quite successful at performing follows on Instagram. Two elements may help its bots to conduct SMF without detection. First, Linux/Moose is a network of infected IoT devices – mainly routers – that are used as proxies to connect to OSN. It leverages its bots' IP addresses, which are from credible consumer ISPs, to make OSN believe that its requests are coming from normal users [5,11,17]. Second, the bots use patterns to try to act like humans to avoid detection. They will therefore visit account profiles (either theirs or others) and view posts. Combined, these two strategies help the bots to successfully conduct SMF. It appears that few roadblocks are put in place to prevent these bots from operating.

However, most of the botnet fake accounts are flagged within six months of their creation. This may be due to the multiple classifiers developed by the community in the past year to flag bots' accounts and suspend them [1,2,7]. SMF service customers must therefore see their follower count decrease as the fake accounts are flagged. This does not seem to be a problem for the Linux/Moose operators, who do not appear to be putting any effort in making their fake accounts look more legitimate. Linux/Moose's fake accounts look like many of the bot accounts described in the literature [14,15]. It is possible that the botnet's customers complain to the supplier when they lose their fake followers, but they have no legal recourse since they bought a fraudulent service. Flagging fake accounts does not prevent SMF and only serves to clean the OSN a posteriori. As such, it is unlikely to deter SMF service providers although it could impact the faith that customers have in them, pushing them toward more legitimate means to acquire new followers.

Linux/Moose is a successful SMF service that targets Instagram. Looking at the market price analysis, the significant price differentiation among different OSN led to the hypothesis that this differentiation is related to the cost suppliers face when conducting SMF. Considering how easily Linux/Moose performs SMF on Instagram could explain why the price for fraudulent services on this OSN is much lower than on Facebook and YouTube. More research on the difficulties associated with conducting SMF on these OSN should be conducted to confirm this hypothesis. By setting higher registration barriers, it is likely that the price of SMF on Instagram will increase. The market price analysis also showed that SMF services are easily accessible through online searches. Customers can thus easily find the service they are looking for. The large price variation in the sale of SMF also shows that the same service can be priced very cheaply or very expensively, making it accessible to any type of consumer. A buyer unsure about the quality of the service can always go for the lowest price. However, the large price variation may also indicate that there are scammers in the market, pricing without considering the real costs of providing the fraudulent service. Buyers can be fooled and ripped off by scammers. A first step at preventing SMF from happening would be to bring together OSN, hosting services and law enforcement agencies to shut down websites offering SMF services, a strategy used in the global effort in 2015 to shut down 37,000 websites selling counterfeit goods [18]. This would reduce the availability of SMF services and likely decrease the overall number of purchases. The

illicit market for SMF is so easily accessible that it even seems to be a legitimate practice undertaken by many.

Finally, the demand analysis showed that those that contribute to the falsification of social network data are not only the suppliers of the fraud, but also those that buy the service. Entertainers, online shops and private users are part of the illicit market for SMF and contribute to disinformation on OSN. The people that use the network are also those that try to manipulate it. SMF is therefore not just an activity undertaken by serious online offenders but one that includes a wide range of actors who may not see their activity as illegal. It is therefore somewhat reminiscent of the peer-to-peer exchange of intellectual property like movies and music in the 2000s, when millions of people illegally shared their library with others. Although many knew it was illegal, it was not deemed to be a crime in and of itself and led to widespread social acceptance for the practice. SMF service customers see value in their online popularity but may not have the resources or patience needed to reach a large fan base legitimately, through time-intensive marketing strategies. A first step to combating the fraud could be to target these users and offer them strategies to gain a larger legitimate fan base at low cost. A more drastic approach for OSN could be to target potential buyers and warn them that buying such service is fraudulent and contributes to online crime when provided by botnets. Raising awareness that such fraud is facilitated by criminal means and is prohibited, is essential to counter the feeling of impunity that buyers may feel while purchasing SMF.

Looking at the supply and the demand for SMF illustrated that many classes of actors are part of the problem, from SMF service providers to customers and to OSN themselves. Social network data manipulation creates disinformation, which in turn may be hurtful for the online ecosystem. There is a need to take actions to decrease this phenomenon from all angles: take down fraudulent websites, develop more sophisticated detection techniques, set up strong registration barriers and contribute to potential buyers' awareness while giving them the tools they need to gain online credibility and popularity. Taking measures against SMF facilitated by botnets is essential to ensure that trust is preserved among members of the social network community. With the rise in social network manipulation, the reliability of the information provided online may be questioned. As OSN are primary outlets for many social, economic and political activities and engage a large portion of the world's population, part of their responsibility lies in preventing such data manipulation on their platforms.

6. ACKNOWLEDGMENTS

This research was partially financed by the MITACS Accelerate Program, under Project N. IT06861. We thank them for financing the project and our partners, Université de Montréal and ESET, for helping us throughout the year.

7. REFERENCES

[1] Alsaleh, M., Alarifi A., Al-Salman, A. M., Alfayez, M. and Almuhaysin, A. 2014. TSD: Detecting Sybil Accounts in Twitter. In *Proceedings of the 13th International Conference on Machine Learning and Applications*. IEEE, Detroit, MI, 463–469. DOI=http://dx.doi.org/10.1109/ICMLA.2014.81.

[2] Beutel, A., Xu, W., Guruswami, V., Palow, C. and Faloutsos, C. 2013. CopyCatch: stopping group attacks by spotting lockstep behavior in social networks. In *Proceedings of the 22nd international conference on World Wide Web*. ACM, New York, NY, 119–130. DOI=http://dx.doi.org/10.1145/2488388.2488400.

[3] Bilge, L., Strufe T., Balzarotti, D. and Kirda E. 2009. All your contacts are belong to us: Automated identity theft attacks on social networks. In *Proceedings of the 18th international conference on World wide web*. WWW, Madrid, Spain, 551–560. DOI= 10.1145/1526709.1526784.

[4] Bilodeau, O. 2015. *Telnet Support for Cowrie*. Cowrie GitHub Repository, available at: https://github.com/micheloosterhof/cowrie/pull/222.

[5] Bilodeau, O. and Dupuy, T. 2015. *Dissecting Linux/Moose: The Analysis of a Linux Router-Based Worm Hungry for Social Networks.* ESET Research Technical Report, Montreal, QC, available at: http://www.welivesecurity.com/2016/11/02/linuxmoose-still-breathing/.

[6] Boshmaf, Y., Muslukhov, I., Beznosov, K., and Ripeanu, M. 2011. The socialbot network: when bots socialize for fame and money. In *Proceedings of the 27th annual computer security applications conference,* ACM, Florida, USA, 93-102. DOI= 10.1145/2076732.2076746

[7] Cao, Q., Yang, X., Yu, J. and Palow, C. 2014. Uncovering large groups of active malicious accounts in online social networks. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security.* ACM, Scottsdale, Arizona, 477–488. DOI= http://dx.doi.org/10.1145/2660267.2660269.

[8] Cha, M., Haddadi, H., Benevenuto, F. and Gummadi, K. P. 2013. Measuring user influence in Twitter: The million follower fallacy. In *Proceedings of the Fourth International Conference on Weblogs and Social Media.* AAAI, Washington, DC, 30. Available at: http://www.aaai.org/Library/ICWSM/icwsm10contents.php.

[9] De Cristofaro, E., Friedman, A., Jourjon, G., Kaafar, M. A. and Shafiq, M. Z. 2014. Paying for likes?: Understanding Facebook like fraud using honeypots. In *Proceedings of the 2014 Conference on Internet Measurement Conference*. ACM, Vancouver, BC, 129–136. DOI= http://dx.doi.org/10.1145/2663716.2663729.

[10] De Vries, L., Gensler, S. and Leeflang, P. S. 2012. Popularity of brand posts on brand fan pages: an investigation of the effects of social media marketing. *Journal of Interactive Marketing,* 26 (01-04-2012), 83–91. DOI= http://dx.doi.org/10.1016/j.intmar.2012.01.003.

[11] ESET Research. 2016. *Linux/Moose: Still Breathing*. Blog We Live Security, ESET, Montreal, QC, available at: http://www.welivesecurity.com/2016/11/02/linuxmoose-still-breathing/

[12] Forelle, M. C., Howard, P. N., Monroy-Hernández, A., and Savage, S. 2015. *Political bots and the manipulation of public opinion in Venezuela*. Available at SSRN: https://ssrn.com/abstract=2635800

[13] Freitas, C., Benevenuto, F., Ghosh, S. and Veloso, A. 2015. Reverse engineering socialbot infiltration strategies in Twitter. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and*

Mining 2015. ACM, Paris, France, 25–32. DOI= http://dx.doi.org/10.1145/2808797.2809292.

[14] Gurajala, S., White, J. S., Hudson, B. and Matthews, J. N., 2015. Fake Twitter accounts: profile characteristics obtained using an activity-based pattern detection approach. In *Proceedings of the 2015 International Conference on Social Media & Society.* ACM, Toronto, ON, 9. DOI= http://dx.doi.org/10.1145/2789187.2789206.

[15] Ikram, M., Onwuzurike, L., Farooqi, S., De Cristofaro, E., Friedman, A., Jourjon, G., Kaafar, M. A. and Shafiq, M. Z. 2015. *Combating Fraud in Online Social Networks: Detecting Stealthy Facebook Like Farms*. Available at: https://arxiv.org/abs/1506.00506v3.

[16] Nguyen, C. 2016. *What Life for a Bangladeshi Click Farmer Looks Like*. Motherboard–Vice. Published 30 March 2016 at: http://motherboard.vice.com/read/what-life-for-a-bangladeshi-click-farmer-looks-like.

[17] Paquet-Clouston M., Bilodeau, O., Décary-Hétu, D. and Dupuy, T. 2016. *EGO MARKET: When Greed for Fame Benefits Large-Scale Botnet*s. GoSecure Technical Report, Montreal, QC, available at: http://gosecure.net/wp-content/uploads/2016/11/Ego-Market_When-Greed-for-Fame-Benefits-Large-Scale-Botnets.pdf.

[18] SC Magazines. 2015. Global efforts take down 37,000 websites selling counterfeit goods, Published 4 December at: https://www.scmagazine.com/global-efforts-take-down-37000-websites-selling-counterfeit-goods/article/533253/

[19] Stringhini, G., Egele, M., Kruegel, C. and Vigna, G. 2012. Poultry markets: on the underground economy of Twitter followers. In *Proceedings of the 2012 ACM Workshop on Online Social Networks*. ACM, Helsinki, Finland, 1–6. DOI=http://dx.doi.org/10.1145/2342549.2342551.

[20] Thomas, K., Iatskiv, D., Bursztein, E., Pietraszek, T., Grier, C. and McCoy, D. 2014, November. Dialing back abuse on phone verified accounts. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. ACM, Scottsdale, AZ, 465–476. DOI=http://dx.doi.org/10.1145/2660267.2660321.

[21] Thomas, K., McCoy, D., Grier, C., Kolcz, A. and Paxson, V. 2013. Trafficking fraudulent accounts: the role of the underground market in Twitter spam and abuse. In *Proceedings of the 22nd USENIX Security Symposium. USENIX Security*, Washington, DC, 195–210.

[22] Zhang, J., Zhang, R., Zhang, Y., and Yan, G. 2016. The rise of social botnets: Attacks and countermeasures. *IEEE Transactions on Dependable and Secure Computing*, 99 (06-04-2016), 1-14.

[23] Zhang, J., Zhang, R., Zhang, Y., and Yan, G. 2013. On the impact of social botnets for spam distribution and digital-influence manipulation. In *Proceedings of the 2013 Communications and Network Security (CNS)*, Chicago, USA, 46-54. DOI=10.1109/CNS.2013.6682691.