The Underground Economy of Spam: A Botmaster's Perspective of Coordinating Large-Scale Spam Campaigns

Brett Stone-Gross[§,⋆], Thorsten Holz[‡,⋆], Gianluca Stringhini[§], and Giovanni Vigna[§,⋆]

[§]University of California, Santa Barbara
{bstone,gianluca,vigna}@cs.ucsb.edu

[‡] Ruhr-University Bochum
thorsten.holz@rub.de

[⋆] LastLine, Inc., Santa Barbara CA 93111, USA
{brett,tho,vigna}@tllod.com

Abstract

Spam accounts for a large portion of the email exchange on the Internet. In addition to being a nuisance and a waste of costly resources, spam is used as a delivery mechanism for many criminal scams and large-scale compromises. Most of this spam is sent using botnets, which are often rented for a fee to criminal organizations. Even though there has been a considerable corpus of research focused on combating spam and analyzing spam-related botnets, most of these efforts have had a limited view of the entire spamming process.

In this paper, we present a comprehensive analysis of a large-scale botnet *from the botmaster's perspective*, that highlights the intricacies involved in orchestrating spam campaigns such as the quality of email address lists, the effectiveness of IP-based blacklisting, and the reliability of bots. This is made possible by having access to a number of command-and-control servers used by the Pushdo/Cutwail botnet. In addition, we study Spamdot.biz, a private forum used by some of the most notorious spam gangs, to provide novel insights into the underground economy of large-scale spam operations.

1 Introduction

In the Internet's vast underground economy, unsolicited bulk email (or *spam*) serves an important role. There are a number of spam forums that specifically cater to illicit businesses from advertising cheap pharmaceutical drugs, to distributing malware, to performing a variety of scams [10]. Purveyors of spam offer a plethora of services such as custom software to manage spam operations, email address lists (commonly referred to as *bases*), and CAPTCHA solvers [13]. In addition, there are criminal organizations that provide fee-based services to send spam on behalf of third-party customers.

In order to send large amounts of bulk email most efficiently, many of these ventures employ the use of *botnets*, a collection of compromised computers (i.e., bots) under their control. According to a recent study by Symantec, more than 89% of all email messages on the Internet were attributed to spam in the year 2010. Furthermore, about 88% of these spam messages were sent with the help of botnets [12]. This huge percentage of botnet-related spam is due to several advantages that a botnet can provide with respect to other kinds of spam delivery mechanisms. First, since a botnet operates as a distributed system where each infected machine receives a subset of the overall tasks, the amount of resources required by the spam operator is greatly reduced. This increases the effective throughput, as the bots perform the majority of the work on their own. Second, most botnets contain a degree of geographic diversity that makes spam filtering techniques such as IP-based blacklisting more difficult. That is, identifying the sources of bot-related spam is challenging due to constant changes in the sources of spam messages. As a result, blacklists must be continuously updated to remain valuable.

In this paper, we provide an in-depth analysis of *spam campaigns* orchestrated by the *Pushdo*/*Cutwail* botnet. A spam campaign is a coordinated effort to deliver a particular spam message to a target population. In August 2010, we obtained access to 13 Command & Control (C&C) servers and 3 development servers (16 servers in total) used by botnet operators of the Cutwail spam engine. This software has been used by some of the most prolific spammers over the last few years, and is frequently installed by a separate Trojan component known as Pushdo. Cutwail utilizes an encrypted communication protocol and an automated template-based spamming system to dynamically generate unique emails with the goal of evading existing spam filters. Interestingly, each Cutwail bot maintains highly detailed statistics about its own spam activities, which are reported back to the C&C server. The data we obtained from these C&C servers provides us with a novel, deeper insight into the *modus operandi* of cyber criminals and the dynamics behind some of the most sophisticated spam operations to-date.

In addition to the data retrieved from the Cutwail servers, we were also able to obtain a copy of a popular web-based forum known as *Spamdot.biz*. This forum is dedicated to spam activities, and used by several operators of Cutwail and the *Bredolab* botnet. From this vantage point, we can observe communications, transactions, and the exchange of ideas between some of the largest players in the underground spam economy, which enables us to get a better understanding of their techniques. We apply this intelligence to the data collected from the Cutwail C&C servers to approximate the cost of running a spam campaign, which provides us with an estimate of the economic aspects of spam operations.

What makes our research novel is the unique perspective and the depth of our analysis. As a result of having gained access to a large number of C&C servers, we are able to observe an almost complete view of how modern spam operations work. In particular, we can identify the problems that make spam operations difficult, and the value of spam from a financial point of view. In addition, we believe that our findings will improve the understanding of the amount of spam delivered by botnets.

In summary, we make the following three contributions:

- We provide an in-depth analysis of the Cutwail spam operation and present detailed statistics based on the analysis of 16 servers belonging to this botnet. This is the most comprehensive, large-scale study of spam botnets, highlighting different aspects of such operations, based on information *directly* collected from the botmaster's hosts.
- We examine how modern spam campaigns are managed and delivered, from the botmaster's point of view. We discuss the software infrastructure that is used, the functionality it provides, and its role in the underground economy.
- We analyze the *Spamdot.biz* forum and study the communication and transactions of spammers in order to understand the economics of spam campaigns and their role in the underground economy.

2 Technical Background

In this section, we present an overview of the key components of the Cutwail botnet and the process through which an infected computer becomes a spam bot. A detailed technical analysis is beyond the scope of this paper, and was presented by Decker et al. [4].

The original Cutwail botnet emerged back in 2007, and has evolved in sophistication from using simple HTTP requests to a proprietary, encrypted protocol. A typical Cutwail infection occurs when a compromised machine executes a so called *loader* called Pushdo, that behaves as an installation framework for downloading and executing various malware components. Depending on the victim's system configuration, the Pushdo malware will contact a C&C server, and request additional malware components (as shown in Figure 1, Step 1).

After Pushdo contacts the C&C, several malware modules are typically downloaded and installed. This commonly includes a rootkit component to hide the presence of malware on the infected system, the Cutwail spam engine, and a list of IP addresses of Cutwail C&C servers (Step 2). At this point, the infected machine executes the Cutwail spam engine and becomes a spam bot (Step 3). Next, the Cutwail bot will contact one of the IP addresses from the list provided through the Pushdo bootstrap process, and wait for instructions (Step 4). The Cutwail C&C server provides several critical pieces of information to begin a spam campaign (Step 5). More specifically, the C&C server provides the actual spam content delivered through the use of *spam templates*, a target list of email addresses where spam will be delivered, a dictionary consisting of 71,377 entries for generating random sender/recipient names, and a configuration file containing details that control the spam engine's behavior, such as timing intervals and error handling. Optionally, a list of compromised SMTP credentials can be distributed to bots for "high-quality" spam campaigns [15]. These techniques are used by similar botnets to perform template-based spamming [16].

There have been previous attempts at disrupting Cutwail's activities, such as the shutdown of the bulletproof hosting provider McColo in 2008 [5], and of the Triple Fiber Network (3FN) in 2009 [9]. The security firm FireEye attempted a second takedown effort in 2010 [14], which diminished spam levels for several weeks. Despite the attention that Cutwail has received, it is still able to function and delivered billions of spam email messages

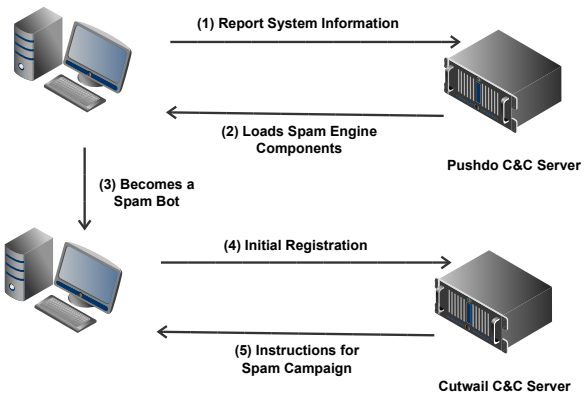

Figure 1: Overview of the Cutwail installation and infection process.

per month, as we will show later. At its peak, in May 2009, Cutwail was estimated to be responsible for 46.5% of all the spam on the Internet [21].

The Cutwail spam engine is known in spam forums by the name *0bulk Psyche Evolution*, where it is rented to a community of spam affiliates. These affiliates pay a fee to Cutwail botmasters in order to use their botnet infrastructure. In return, clients are provided with access to a web interface (available in Russian or English language) that simplifies the process of creating and managing spam campaigns (referred to by Cutwail as *bulks*). The interface includes features to fine-tune nearly every part of an email message and spam campaign. For instance, a user can choose to customize the email's headers to impersonate legitimate mail clients (e.g., Microsoft Outlook, Windows Mail, and TheBat), or may opt to define their own headers. After defining the headers, the user may define other fields, including the sender address, email subject line, and body. All of these fields can make use of *macros* that will instruct each individual bot to dynamically generate (and fill-in) unique content for each email sent in order to evade detection by spam filters, similar to other spam botnets [16]. In order to increase the spam campaign's effectiveness, each Cutwail C&C runs a local instance of SpamAssassin, a free open-source mail filter, that uses a set of heuristics to classify spam. Once the email template has been created, it is automatically passed through SpamAssassin and can be tweaked until it successfully evades detection. After creating the spam message, the user must specify several parameters such as a target email address list, a configuration file that controls a number of bot parameters (e.g., sending rates, timeouts, retries, etc.), and the time when the spam campaign will commence. If a Cutwail user requires assistance, they can refer to an instruction manual that is included, or contact Cutwail's support team.

3 Data Collection

In the following section, we describe the process that facilitated our efforts in disrupting the majority of Cutwail botnet C&C servers, the results of our takedown efforts, and the data we collected. The primary tool that we utilized was ANUBIS [6], a framework for dynamic, runtime analysis of binary programs. ANUBIS runs a Windows executable and records during runtime the program's behavior such as file system modifications and network activity. At the moment, the system processes tens of thousands of malware samples per day and offers us an insight into the latest malware trends [1].

By searching through the ANUBIS database, we were able to identify 30 distinct Cutwail C&C servers based on their unique communication signatures. We then contacted the hosting providers whose servers were being used for controlling the botnet. We provided them with evidence that specific servers within their network were used for malicious purposes and requested the take down of these servers. Note that we had previously established relationships with some of these hosting providers through our network reputation service called FIRE [20]. This service tracks the network locations of servers used for malicious purposes and enables ISPs to proactively clean their networks from compromised/malicious hosts.

Data Sets. As a result of our notification and mitigation steps, more than 20 servers were shut down and we were able to obtain access to 16 servers used by Cutwail controllers from some of the hosting providers. These servers contained a wealth of information, including:

- More than 2.35 TB of data.
- 24 databases that contain detailed statistics about the infected machines and overall spam operations.
- Spam templates and billions of target email addresses for spam campaigns.
- The botnet's source code and a highly detailed instruction manual for botnet operators.

An analysis of this data enables us to obtain unique insights into modern spam operations. It is evident from the content on these servers that there are several different crews renting these botnets. As we will see later in Section 4, these various groups also carry out different spam campaigns. Interestingly, there was only one server that was running both Pushdo and Cutwail services, which we believe indicates that most current Cutwail botmasters recruit the services of other criminal organizations (primarily those of Pushdo, but also those of other botnets, such as *Butterfly*) to assist with malware installations (or *loads*), on a *pay-per-install* (PPI) basis.

The short-term effect on the overall spam levels associated with the Cutwail botnet was significant. In addition to the reduction in spam, the takedown greatly disrupted the distribution of the Bredolab malware loader [21]. This happened because, as we will discuss in Section 4.1, Cutwail is commonly used to spread malware.

Cutwail Botnet Size. When a Cutwail bot initially connects to a C&C server, it will receive a bot identifier. By analyzing the Cutwail source code, we found that the identifier is simply an automatically incremented field in their botnet database. We found that while these values are unique, they do not appear to be an accurate indicator of the total number of bots. First, a Cutwail bot may connect to multiple C&C servers over its lifetime, and, thus, several C&Cs may have their own identifier for a single bot. In addition, we observed some bots that appeared to

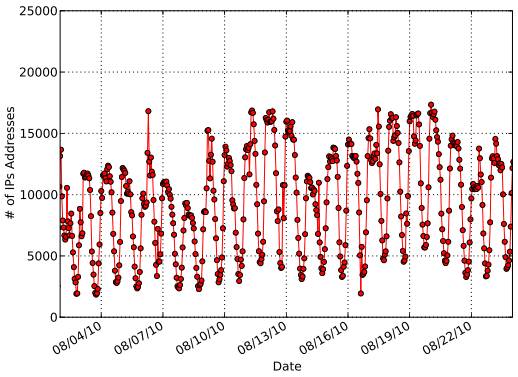

Figure 2: Unique online bot IP addresses per hour.

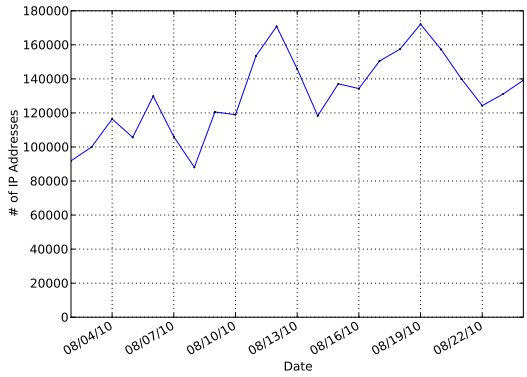

Figure 3: Unique online bot IP addresses per day.

receive multiple identifiers each time they connected to the C&C server, possibly due to a bug in the malware. Since Cutwail bots do not have a unique identifier, we can only estimate the total number of infected machines based on IP addresses. That is, we counted the number of unique IP addresses on an hourly basis, which was shown in [19] to be a reasonable approximation of a botnet's live population. On average, there were 121,336 unique IPs online per day, and 2,536,934 total IPs observed over the whole analysis time frame.

Figures 2 and 3 show the number of bots that came online on an hourly and daily basis, respectively. Interestingly, the highest concentration of bots by far was in India (38%). One possible explanation is that the Cutwail controllers may specifically target Indian machines because the cost per bot is cheaper than those in other geographic regions, as we will see in Section 5. Australia (9%), Russia (4%), Brazil (3%), and Turkey (3%) account for the next largest number of spam bots.

4 Spam Campaign Dynamics

The most interesting information retrieved from the C&C servers was stored in the databases containing meticulous records for each spam bot. More specifically, the botnet controllers maintain detailed statistics per infected machine (identified via a unique IP address) in order to measure the effectiveness of their spam campaigns. We found that a spammer's job is complicated by a number of factors including invalid email addresses, SMTP errors, and blacklisting. As a result, the amount of spam that was actually delivered (i.e., accepted by mail servers) was only around 30.3%, and the actual volume was likely much less after client-side spam filters are taken into account. This delivery rate is slightly higher than the 25% delivery rate of the *Storm* botnet [8].

The largest cause of failure was invalid email addresses accounting for 53.3% of errors, followed by SMTP blacklisting (16.9%), miscellaneous STMP errors (11.8%), and connection timeouts (11.3%). Interestingly, 3.5% of mail servers notified the sender (in this case, the bots), that the content of the email was flagged as spam. Despite these complications, the amount of spam that is sent by Cutwail bots is immense. During one period from July 30, 2010 and August 25, 2010 the database records show 87.7 billion emails were successfully sent.

Overall, records contained on these Cutwail servers dated as far back as June 2009 and reported 516,852,678,718 messages were accepted for delivery out of a total of 1,708,054,952,020 attempts. Note that we obtained roughly one-half to two-thirds of the active Cutwail C&C servers, so the overall numbers are likely higher. Figure 4 displays the number of spam emails sent per campaign hourly and Figure 5 shows the cumulative number of successfully delivered spam messages. The rate of spam followed a highly linear pattern and there were virtually no periods of inactivity.

4.1 Spam Content

The content of the email messages sent by Cutwail included pornography, online pharmacies, phishing, money mule recruitment, and malware. The malware (e.g., the ZeuS banking Trojan) is typically distributed by enticing a user to open an attachment in the form of a greeting card, resume, invitation, mail delivery failure, or a receipt for a recent purchase. In addition, many of the emails contained links to malicious websites that attempted to covertly install malware on a victim's system through drive-by-download attacks. Cutwail operators also advertised content to Russian speakers such as real estate and ski resorts.

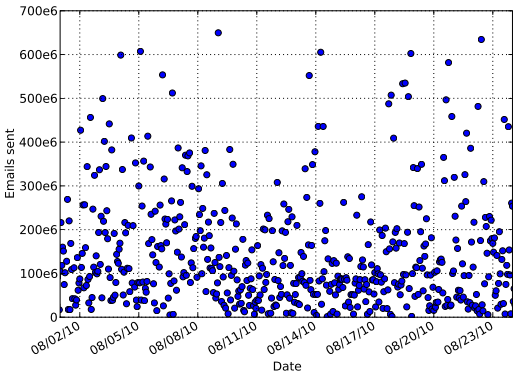

Figure 4: Spam delivered per campaign per hour.

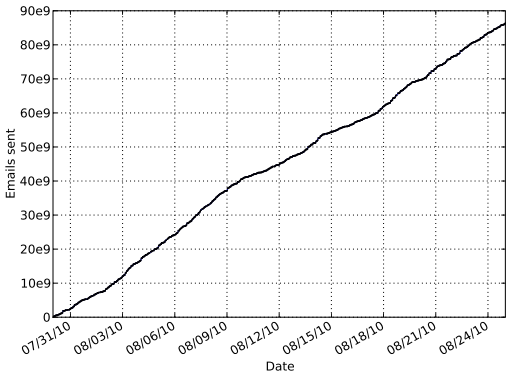

Figure 5: Aggregate spam sent per day.

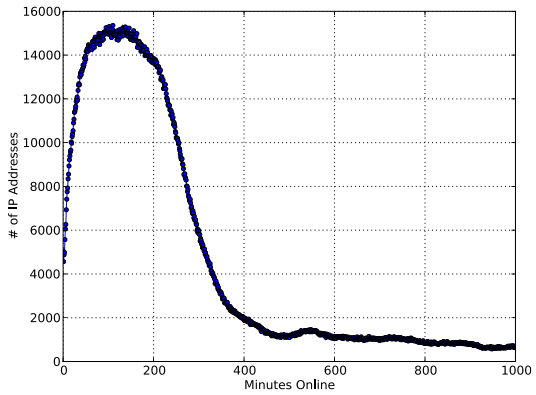

Figure 6: Spam bot blacklisting time.

4.2 Blacklisting

One of the most important aspects of a spam campaign is the ability to pass through both IP-based blacklists and content-based filters. Bots that have not been added to a spam blacklist are the most valuable, since there is a higher chance that their mail will be delivered to the recipient. Each Cutwail bot periodically queries several blacklists (i.e., SORBS, SpamCop, DNSBL), in order to determine its reputation. This information is reported back to the C&C server and recorded. Figure 6 shows the amount of time it takes for bots to appear on at least one of these blacklists. Interestingly, only about 12.8% of bots are blacklisted after an hour when they come online. At two hours about 29.6% of bots are blacklisted, and 46.4% are blacklisted after three hours. By six hours, roughly 75.3% are blacklisted. The rate reaches 90% after a period of about 18 hours.

4.3 Spam Organizations

During our takedown, we got access to 13 Cutwail C&C servers. An interesting aspect that we studied was how these servers were controlled, and whether they were rented to different organizations or individuals, running different spam operations. By analyzing the usernames in each database, we found evidence that the servers are set up by one group of people, perhaps the creators of Cutwail. In particular, although most servers had unique usernames, every database contained a common set of accounts (usernames/passwords). Thus, we believe each server was likely controlled by this core group that rented the servers to different clients. Another interesting observation was that nine of the C&C servers (out of the 13) were running two distinct versions of the *Psyche Evolution* software. Since each version utilized its own database, one server could have been rented to two different groups at the same time. By correlating the user accounts on the different databases, we tracked the servers rented to each client to carry out their activities. We discovered that one client rented at least eight instances, while three other clients rented at least two instances. All the other instances were rented by individual clients.

Table 1 shows spam statistics for each instance controlled by a Cutwail client. Note that there were four instances where we could not definitively identify a client. Thus, the aggregate mail sent in Table 1 is less than the 500 billion discussed earlier. In general, the content of the spam campaigns varied by client. For example, Client-1 coordinated phishing campaigns (e.g., Google Mail, Friendster, etc), while Client-9 was advertising only Russian real estate. Overall, phishing was the most popular campaign with 16 instances, followed by six instances of malware campaigns (i.e., the mail included a malicious link or attachment), and two instances of pharmaceuticals and online education campaigns. Although in some cases the types of campaigns were similar, the content of the campaign was unique to each client.

The instances controlled by each client displayed varying levels of performance. For example, the group controlling the largest number of instances did not send the

Client (ID)	Instances (#)	Unique Bot IPs (#)	Avg. Lifespan (Days)	Mails Sent (#)	Average Mails/ Active Bot (Per Day)	Campaign Type
1	8	2,251,156	17	98,401,907,545	2,571	Phishing, Malware
2	2	40,924	168	45,555,535,375	6,626	Phishing
3	2	56,733	54	155,098,090,946	50,626	Diplomas
4	2	34,742	22	17,941,545,204	23,473	Phishing, Pharm.
5	1	21,993	8	60,169,427,197	341,980	Money Mule
6	1	29,471	13	4,309,066,448	11,247	Pharmaceuticals
7	1	27,658	55	9,408,910,232	6,185	Phishing
8	1	30,503	135	12,485,832,067	3,032	Phishing
9	1	29,415	18	2,365,652,828	4,467	Real Estate

Table 1: Statistics for individual spam operations run by Cutwail.

most spam, having 5.5 million messages sent per day. On the other hand, Client-5, who controlled a single instance, was able to send out more than 7 billion messages per day. These results may have been caused by a larger number of unreliable bots, or due to bad lists of email addresses. Thus, we attempted to verify these assumptions by examining the country distribution of the bots per instance to isolate other factors such as email address lists and other configuration parameters. We found that in general the reliability and quality of a bot was not tied to its geographic location, and that bots from nearly all regions exhibited similar spam capabilities. Another important component of spam campaigns that we examined was the quality of email address lists used by each client. The size of the lists that we found on these instances differed considerably. In particular, Client-8 possessed lists containing more than two billion addresses, whereas Client-9 used one list of 169 million. By comparing the email addresses, file names and the vendor who provided the list, we were able to assess their relative effectiveness. More specifically, we found that the most successful clients had custom lists, that contained unique addresses that were not shared by any other clients, and most importantly, the addresses were valid. On the other hand, Client-1 appeared to be using default lists (possibly included for free), on every server. Thus, these lists were lower in quality, and therefore contributed to less effective spam campaigns. Finally, we analyzed the partitioning of jobs to bots during a spam campaign. We found that the email list for a campaign is usually partitioned depending on the number of bots. However, there is a tradeoff on the size of the partitions and the number of bots. That is, orchestrating a campaign with small partitions and a large number of bots has the consequence of keeping most of them idle (e.g., Client-1). On the other hand, choosing larger partitions allows more bots to be active, which may result in substantially more spam mails being sent at a faster rate (e.g., Client-5).

5 Analysis of the Spamdot.biz Web Forum

In this section, we will analyze *Spamdot.biz*, an underground web forum, devoted to spam operations. This forum is a sister site of a well-known spam affiliate program known as *SpamIt*, that promoted cheap prescription drugs carrying the infamous "Canadian Pharmacy" brand name. *SpamIt* officially closed on October 1, 2010 due to increased public scrutiny [10]. Shortly before *SpamIt*'s closure, we were able to obtain a complete copy of the *Spamdot.biz* forum. Based on the information obtained from the forum, we will take a closer look of how the underground spam economy operates and, in particular, how it relates to our discussion of Cutwail.

We begin by reviewing the format of *Spamdot.biz*, which is built on the open source forum software *phpBB*. In order to join the forum, there is a strict vetting process that requires at least three referrals from existing members that have a minimum number of posts, or a recommendation from at least two *trusted* members, who are part of the top echelon of the spam community. The forum's members consist almost entirely of Russian speakers (91.3%) from Eastern Europe (the remaining 8.7% selected English as their primary language). In total, there are 1,929 registered users who posted 35,423 messages in the forums, and sent 11,638 private messages. Members typically advertise services through the forum and conduct transactions through private messages via the forum software or ICQ instant messages.

The forum is divided into two primary categories: *spam community* and *vendor services*. These categories are further subdivided into sections for proxies, hosting providers, CAPTCHA and webmail spam, email address databases, spam products and services, and botnets. Note that the forum operates based on a system of trust (as described previously), where members routinely vouch for the quality of goods and services provided by others. Members deemed to be grifters were banned by the

forum moderators, and, therefore, we believe that most of the advertised goods, services, and prices on the site were legitimate and reflected current market values.

An area that we are interested in is the economics of the various spam-related goods and services. One of the hottest commodities are lists of email addresses. There are a number of factors that make an email list valuable, including whether the addresses are valid (the recipient exists) and if another spam group has been targeting the addresses recently. In addition, some email address lists may be localized to a particular gTLD (e.g., .us, .uk, .ru), or regionalized through the use of IP-based geolocation services. The cost of these lists range in price, depending on the targeted region. Furthermore, the value of an email address is contingent upon whether it belongs to a free email service, such as Gmail, Hotmail, or Yahoo. Email addresses that belong to one of these free services sell at a rate of nearly one-half the value of a standard email address, likely due to the fact that many free email providers use more sophisticated spam filters, and, hence, spam is less likely to reach its intended target. Rates for one million email addresses range from $25 to $50, with discounted prices for bulk purchases.

Those interested in building a botnet or installing their malware on a large number of systems often seek the services of groups who provide so called *loads*, i.e., the ability to install malware on compromised machines. The loads come from a variety of sources such as drive-by-download attacks using HTML iframes and other malware. We observed several individuals offering 10,000 malware installations for approximately $300–$800. However, the market price per load is highly dependent on its geographic location, and whether it was the sole malware executable running on the victim's system. For example, infected computers in the U.S. are more valuable than those in Asia, probably because they have a faster and more reliable Internet connection. To put this into perspective, loads sold per thousand in Asia went for around $13, Europe at $35, and $125 for the U.S. Similar to the sale of email address lists, quantity discounts are given when larger amounts of loads are purchased. Bots that have not been blacklisted (referred to on the forum as "clean") sell at higher prices, since they are especially valuable for spam campaigns. Another problem that controllers of spam botnets struggle with is maintaining a sufficient number of bots that are online, and there were reports that bot populations sometimes drop by 50% per day, and, thus, botmasters may have to frequently replenish their supply.

After a group has acquired the resources necessary to form a botnet, they may launch their own spam campaigns or rent out parts of the botnet to other spam organizations. Some of these organizations share a percentage of sales (e.g., *SpamIt* paid affiliates a 40% commission), with botnet controllers who drive traffic to their sites. Alternatively, spam-as-a-service can be purchased for approximately $100–$500 per million emails sent. Botnets can also be rented out to groups interested in sending out larger campaigns that are capable of sending 100 million emails per day for $10,000 per month. Before renting a botnet, potential buyers are traditionally offered a free trial to evaluate the performance of the spam botnet.

Based on the value of the products and services that we previously described, we can estimate on a high level the cost of operating Cutwail's spam campaigns, and approximate the transaction volume related to such an operation. As we discussed in Section 3, there were an average of 121,336 unique IPs online per day. Thus, the Cutwail operators may have paid between $1,500 and $15,000 on a recurring basis to grow and maintain their botnet (assuming they did not develop their own loads system). If we estimate the value of the largest email address list (containing over 1,596,093,833 unique records) from advertised prices, it is worth approximately $10,000–$20,000. Finally, we estimate the Cutwail gang's profit for providing spam services at roughly $1.7 million to $4.2 million since June 2009 (contingent on whether bulk discounts were provided to customers).

6 Related Work

In the past few years, there have been several studies of spam botnets similar to the analysis presented in this paper. We now briefly discuss how our study relates to previous work in this area and what novel insights this paper provides. Cho et al. infiltrated *MegaD* by automatically reverse-engineering the communication protocol [2], impersonating a bot, and "milking" the C&C servers for spam templates [3]. This technique was also used to analyze the Storm botnet [11], and provided insights into the types of spam messages that are sent out via a given botnet. We provided a similar analysis for Cutwail, providing further insights into spam campaigns.

Nunnery et al. [15] cooperated with third-party hosting providers to get access to two C&C servers of the *Waledac* botnet and studied the file system found on these servers. As a result, they were able to uncover detailed information about the mechanisms used to run this botnet, an analysis that complemented work in that area [17, 18]. We also cooperated with hosting providers to get access to C&C servers, but focused our analysis on the details of spam operations, for example by studying the numerous complexities in sending spam.

Systems such as BOTLAB [7] or AUTORE [22], and proprietary analysis performed by antivirus companies were used to study the spam volume and relative size of spam botnets [12], while Zhuang et al. introduced a

technique to cluster spam campaigns based on collected spam messages [23]. In contrast, we focus on one particular spam operation and provide actual statistics for the absolute size and success rates of such operations.

7 Conclusions

In this paper, we have presented a study on how large-scale spam campaigns are carried out and managed using botnets, from a botmaster's perspective. Our vantage point on this type of malware infrastructure was made possible by gaining access to a number of the actual command-and-control servers that were part of the Pushdo/Cutwail botnet. This provided us with a novel view into the statistics maintained by the botmasters, and the software they use to manage both the bots and the clients, to whom they offer their services. In addition, we highlighted the role of these types of botnets in the underground economy by accessing a private forum used by well-known criminal organizations.

We believe that these insights will improve the security community's understanding of the underground economy. In addition, the data that we provide can be used to validate or refute results built on simulation, or by speculations based on the observations of subsets of botnet components.

Acknowledgments

This work has been supported by the Office of Naval Research (Grant N000140911042), the Ministry of Economic Affairs and Energy of the State of North Rhine-Westphalia (Grant 315-43-02/2-005-WFBO-009) and the Federal Ministry of Education and Research (Grant 01BY1020 – MobWorm). We also thank the anonymous reviewers for their valuable insights and comments.

References

[1] U. Bayer, I. Habibi, D. Balzarotti, E. Kirda, and C. Kruegel. A View on Current Malware Behaviors. In *USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET)*, 2009.

[2] J. Caballero, P. Poosankam, C. Kreibich, and D. Song. Dispatcher: Enabling Active Botnet Infiltration Using Automatic Protocol Reverse-engineering. In *ACM Conference on Computer and Communications Security (CCS)*, 2009.

[3] C. Cho, J. Caballero, C. Grier, V. Paxson, and D. Song. Insights from the Inside: A View of Botnet Management from Infiltration. In *USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET)*, 2010.

[4] A. Decker, D. Sancho, L. Kharouni, M. Goncharov, and R. McArdle. A Study of the Pushdo / Cutwail Botnet. `http://us.trendmicro.com/imperia/md/content/us/pdf/threats/securitylibrary/study_of_pushdo.pdf`, 2009.

[5] S. DiBenedetto, D. Massey, C. Papadopoulos, and P. J. Walsh. Analyzing the Aftermath of the McColo Shutdown. In *Annual International Symposium on Applications and the Internet*, 2009.

[6] International Secure Systems Lab. Anubis: Analyzing Unknown Binaries. `http://anubis.iseclab.org`, 2010.

[7] J. P. John, A. Moshchuk, S. D. Gribble, and A. Krishnamurthy. Studying Spamming Botnets Using Botlab. In *USENIX Symposium on Networked Systems Design and Implementation*, 2009.

[8] C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. Voelker, V. Paxson, and S. Savage. Spamalytics: An Empirical Analysis of Spam Marketing Conversion. In *ACM Conference on Computer and Communications Security (CCS)*, 2008.

[9] B. Krebs. FTC Sues, Shuts Down N. Calif. Web Hosting Firm. `http://voices.washingtonpost.com/securityfix/2009/06/ftc_sues_shuts_down_n_calif_we.html`, 2009.

[10] B. Krebs. Spam Affiliate Program Spamit.com to Close. `http://krebsonsecurity.com/2010/09/spam-affialite-program-spamit-com-to-close/`, 2010.

[11] C. Kreibich, C. Kanich, K. Levchenko, B. Enright, G. M. Voelker, V. Paxson, and S. Savage. Spamcraft: An Inside Look at Spam Campaign Orchestration. In *USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET)*, 2009.

[12] MessageLabs. MessageLabs Intelligence: 2010 Annual Security Report. `http://www.messagelabs.com/mlireport/MessageLabsIntelligence_2010_Annual_Report_FINAL.pdf`, 2010.

[13] M. Motoyama, K. Levchenko, C. Kanich, D. Mccoy, G. M. Voelker, and S. Savage. Re:CAPTCHAs – Understanding CAPTCHA-Solving Services in an Economic Context. In *USENIX Security Symposium*, 2010.

[14] A. Mushtaq. Infiltrating Pushdo – Part 1. `http://blog.fireeye.com/research/2010/01/infiltrating-pushdo-part-1.html`, 2010.

[15] C. Nunnery, G. Sinclair, and B. B. Kang. Tumbling Down the Rabbit Hole: Exploring the Idiosyncrasies of Botmaster Systems in a Multi-Tier Botnet Infrastructure. In *USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET)*, 2010.

[16] A. Pitsillidis, K. Levchenko, C. Kreibich, C. Kanich, G. Voelker, V. Paxson, N. Weaver, and S. Savage. Botnet Judo: Fighting Spam with Itself. In *Symposium on Network and Distributed System Security (NDSS)*, 2010.

[17] G. Sinclair, C. Nunnery, and B. Kang. The Waledac Protocol: The How and Why. In *Annual Conference on Malicious and Unwanted Software*, 2010.

[18] B. Stock, J. Gobel, M. Engelberth, F. Freiling, and T. Holz. Walowdac Analysis of a Peer-to-Peer Botnet. In *European Conference on Computer Network Defense*, 2009.

[19] B. Stone-Gross, M. Cova, L. Cavallaro, R. Gilbert, M. Szydlowski, R. Kemmerer, C. Kruegel, and G. Vigna. Your Botnet is My Botnet: Analysis of a Botnet Takeover. In *ACM Conference on Computer and Communications Security (CCS)*, 2009.

[20] B. Stone-Gross, A. Moser, C. Kruegel, E. Kirda, and K. Almeroth. FIRE: FInding Rogue nEtworks. In *Annual Computer Security Applications Conference (ACSAC)*, 2009.

[21] M. Vicario. Cutwail Takedown Cripples Bredolab Trojan; No Effect on Spam Levels. `http://www.symantec.com/connect/blogs/cutwail-takedown-cripples-bredolab-trojan-no-effect-spam-levels`, 2010.

[22] Y. Xie, F. Yu, K. Achan, R. Panigrahy, G. Hulten, and I. Osipkov. Spamming Botnets: Signatures and Characteristics. *SIGCOMM Comput. Commun. Rev.*, 38, August 2008.

[23] L. Zhuang, J. Dunagan, D. R. Simon, H. J. Wang, and J. D. Tygar. Characterizing Botnets From Email Spam Records. In *USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET)*, 2008.