Analyzing the Aftermath of the McColo Shutdown

Steve DiBenedetto, Dan Massey, Christos Papadopoulos
Department of Computer Science
Colorado State University
Fort Collins, Colorado

Patrick J. Walsh
eSoft, Inc.
Broomfield, Colorado

Abstract

This paper examines how spam behavior was impacted by the shutdown of McColo, a service provider known for its lax security enforcement. Since the shutdown, a variety of sources have reported significant changes to global spam patterns. In an effort to clarify how spam has changed, we examine reputation data provided by a leading security vendor and present an analysis of spam before and after the McColo shutdown. We show that the actual number of spammers has decreased. We also examine the distribution of spammers both geographically and across the IP space. Our results show that 87% spam sending regions suffered some reductions. Despite this however, the number of sources identified as spammers is still monotonically increasing and the spam volume has recovered to its pre-shutdown levels.

1 Introduction

Spam is a pervasive problem in the Internet and finding new ways to combat spam has been the goal of much research. Increasingly, spam is generated by botnets[11]. While tracking botnets command and control servers may be difficult, spam can be a useful tool in the hunt for the bots themselves. Previous work[3] has shown that botnet operators recognize the risk of their bots becoming discovered by spam sending patterns. As a result, most bots exhibit behavioral patterns which help them to stay unnoticed. Ideally, these command and control (C&C) servers are hosted in sites who are unaware that of their existence and/or have lax enforcement of security policies.

Given the goal of one day stopping spam, it is helpful to gain insight in the way botnets react to disruptions in their infrastructure. In doing so, we hope to develop a strategy for systematic disruption of botnets. In this work, we analyze the effect of the McColo shutdown on spam in terms of both volume and distribution.

On November 11, 2008, the McColo service provider, which housed the command and control servers for a number of major botnets, was shutdown[4]. In the following weeks, many news sites[7][5] reported a sizable reduction in spam volume. However, our results show some of the effects of this shutdown were short lived as many of the bots located alternative control servers and others picked up the remaining slack.

2 Related Work

Ramachandran and Feamster[3] attempted to determine differences between spam and normal email at the network level. Intuitively, network-level traits are considered to be more difficult to forge and would therefore allow for more accuarate countermeasures. However, they show that differences can be difficult to uncover. For example, the overall distribution of spam is very similar to that of legitimate email, but there are a number of hotspots at granularity of /8 routing prefixes. Similar to this work, we show the distribution of spam and how it changes over time. However, we are specifically interested in the McColo shutdown event and the time period surrounding it. As such, we hope to learn more about spammers based on their reactions to a large disruption in their infrastructures.

Collins *et al*[8] investigate the current bot address to determine where they will be in the future. Their work shows that networks with existing bot activity are very likely to continue to have bot activity in the future. Additionally, there is a tendency for bots to cluster together in *unclean* networks. Our work also looks at how spammer activity changes with respect to space and time, but rather than looking for general trends we are focused on a particular event. Overall, focusing on the McColo shutdown provides us with new insights on spammers react to infrastructure changes.

3 Methodology

In order to assess the impact of the McColo shutdown, we examined data provided by eSoft, Inc., a network security whose customers are businesses of many different sizes from around the globe. Our primary source of data is from updates to an IP address reputation list provided to us by eSoft, Inc.

We receive an updated copy of the reputation list every 30 minutes. Potential spammers (IP addresses) are scored by a proprietary algorithm and stale addresses are automatically removed from the list after several days. The list itself is compiled from a number of geographically diverse collectors deployed by the security company's customers. This distributed nature gives us a view of spam as seen by a wide collection of sites.

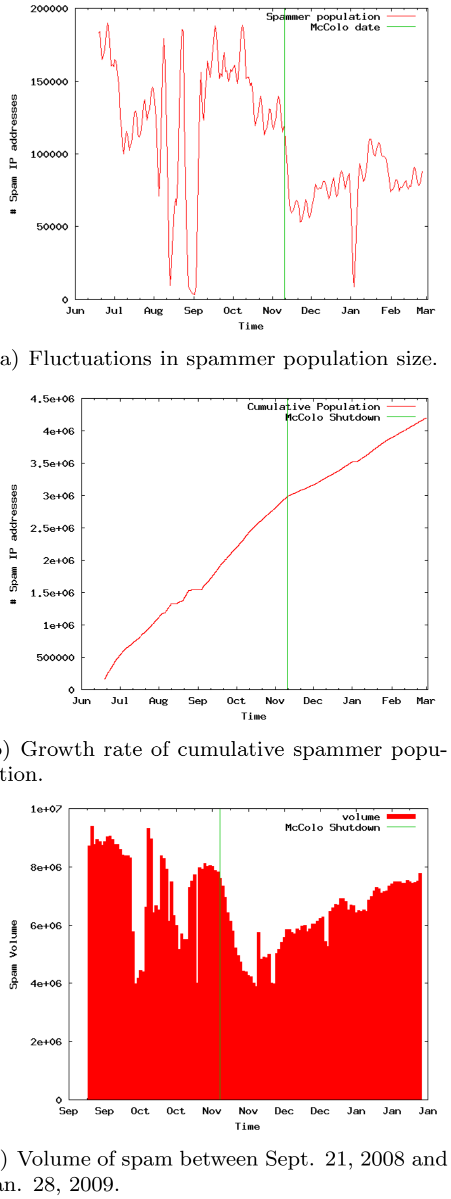
Due to the nature of the updates, we are unable to tell precisely which customer received a particular piece of spam or the exact volume. To account for this, eSoft, Inc. has also provided us with a report detailing the number of spam emails which were blocked each day since June 1, 2008.

3.1 IP Space and Geography Distribution

In order to investigate the reaction of spammers to the McColo shutdown, we examined how their distribution changes in regards to both space and time. As such, we are interested in both IPv4 space distribution as well as geography.

Our primary point of comparison is the number of spammers in a particular region of the IPv4 space. To achieve this, we first divide the IPv4 space into regions equivalent in size to a /16 prefix. It is important to note these regions are in no way tied to the actual space divisions as would appear in allocation records or the default routing table. As shown later in the paper, this choice of address distribution allows us to follow the approach of [6] and display the spam sources in a Hilbert graph.

To examine the geographic distribution of spammers, we make use of the IP::Country tool[9] to determine country of origin. This freely available Perl module makes use of the IP registrars databases. It should be noted that there can be some inaccuracies in the registrar database. However, our study uses a vast set of addresses and we are only interested in the country granularity.



(a) Fluctuations in spammer population size.

(b) Growth rate of cumulative spammer population.

(c) Volume of spam between Sept. 21, 2008 and Jan. 28, 2009.

Figure 1: Changes in the spammer population and spam volume over time.

4 Analysis

4.1 How does spammer population vary?

Figure 1(a) shows the daily fluctuations in the size of the spammer population. The vertical line at 11/11 shows the McColo shutdown date. The results show a notable drop in the size of spammer population immediately following the McColo shutdown. There are three periods where the graph bottoms out due to data corruption.

Before the McColo shutdown, each update listed between 100,000 and 200,000 distinct IP addresses as spammers. With the exception of spikes in activity near the beginning and end of November, the number of spammers listed is relatively consistent.

Immediately after the McColo shutdown, the spammer population underwent a drastic reduction. In the short term, the daily peaks in the spammer population approach the 70,000 mark. The long term effects of the shutdown have thus far limited the spammer population to below 100,000.

4.2 How does the cumulative spammer population grow?

Figure 1(b) illustrates the growth rate of the cumulative spammer population since the start of our monitoring in June. The plateaus shown here are the result of technical problems and do not correspond to a lack of new IPs appearing. Again, the vertical line represents the McColo shutdown.

To date, we have seen over 4 million distinct address and there is a strong growth trend. Interestingly, the growth trend noticeably changes slope after the McColo shutdown and we can clearly see a new trend afterwards. Approximately 409 new spammers are added every half hour prior to the shutdown. However, the post shutdown period sees new spammers added at a rate of 242 per half hour (41% reduction).

4.3 How does the volume of spam change?

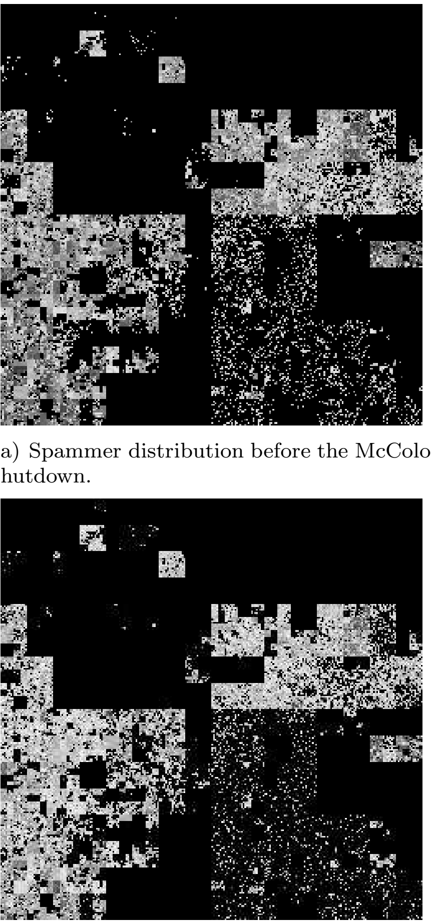
Figure 1(c) shows the volume of spam sent by the addresses in our data feed. Surprisingly, this graph shows multiple drop offs in the volume which are nearly equal to that caused by McColo. We suspect these earlier drop offs signify the ends of various spam campaigns. If so, what separates the shutdown from a campaign ending is the *duration* of the reduced spam levels.

While the duration of the reductions in spam volume differs, the shutdown ultimately had little effect in the long term. Within a few months, the spam volume once again returned to its previous level. However, the spammer population during this time period (figure 1(a)) is still low (less than 100,000). This suggests that while some spammers were taken offline by the shutdown, others must have increased their output to fill the void.

4.4 How does the distribution of spammers change across the IPv4 space?

Figures 2(a) and 2(b) are heat maps of spammer densities mapped on a Hilbert curve[1]. The Hilbert curve allows for neighboring address blocks to be grouped closer together. This is a somewhat popular technique that has been utilized by others in the past for similar purposes [10][6] Due to the drastic range between the extremes of spammer activity, 2(a) and 2(b) are shown in log scale.

The heat maps show spammer concentration levels within regions equivalent in size to a /16 routing prefix. In gray scale, the darker shaded regions are the most active. Black here denotes space which does not have any spam originating from it. Many of the most active regions have experienced a serious reduction in number of spammers.[1]



(a) Spammer distribution before the McColo shutdown.

(b) Spammer distribution after the McColo shutdown.

Figure 2: Changes in the spammer distribution over time.

In the months leading up to the McColo shutdown, 17,532 regions contained at least one spammer. After the shutdown, 1,572 of these regions ceased to have any activity and 975 others began originating spam. The majority of spam sending regions suffered reductions rather than outright elimination. Approximately 87% experienced a decrease in activity. However, 19% of the original regions gained more spammers.

4.5 How does the distribution of spammers change geographically?

Table 1 lists the the top ten origins of spam before the McColo shutdown, ordered by change percentage. The majority of spammers are located in the United States. This observation is consistent with previous work in [3].

[1] For a better visualization of the change in activity, a color version of the map and a movie chronicling the changes for November are available on our website[2].

Country	Before	After	% Change
China	169,367	133,307	-21.29%
Brazil	174,090	126,647	-27.25%
South Korea	109,208	76,522	-29.93%
India	85,393	55,068	-35.51%
United States	465,048	295,591	-36.44%
Spain	92,188	43,487	-52.83%
United Kingdom	116,873	47,580	-59.29%
Argentina	88,705	36,032	-59.38%
Russia	215,563	79,491	-63.12%
Turkey	281,571	74,832	-73.42%

Table 1: Top 10 spam origins

Surprisingly, the actual ordering of the top ten changes and shows a disproportionate results based on origin. While the population of spammers in the United States was dealt a substantial blow (down 169,457 or -36.44%), it was not the largest proportional change. Spammers originating from Turkey suffered the largest set back (206,739 or -73.42%) followed closely by their Russian counterparts.

5 Discussion and Conclusion

Our work analyzed the impact of the McColo shutdown on spammers. The shutdown had very different impacts on spam volume and spammer population. Before the McColo shutdown, spammers were sending relatively high volume campaigns. Additionally, there were approximately 150,000 active spammers on average at any one time. While the spammer population suffered significant casualties after the shutdown, spam itself is only temporarily set back. Within a month's time, the volume of spam once again returned to its normal state. However, the spammer population has yet to return to its previous levels.

Our results provide a unique perspective on how a major infrastructure disruption impacted spammers. We believe the data here not only shows the actual impact of a major security event, but more importantly suggests direction for additional studies.

One possible approach for reducing spam is to target botnet C&C servers. This study provides the first view of what happens when a large set of C&C servers are disabled. The shutdown of McColo is widely believed to have eliminated a large number of C&C servers. Our data shows this resulted in a substantial reduction in the spammer population. We conjecture that this drop in spammers does not suggest that a large number of boxes were patched. Instead, the loss of the C&C servers resulted in a loss of control over these compromised boxes.

This data is encouraging for approaches that target botnet C&C servers. Our data suports the claim that disabling C&C servers can reduce the number of bots. However, these effects are both short-lived as the number of spammers continued to grow. More importantly, users care most about the overall volume of spam email and this showed a short-lived gain at best. We believe that this shows targeting C&C servers is not sufficient and a more comprehensive defense strategy is needed.

References

[1] Hilbert curve. http://en.wikipedia.org/wiki/Hilbert_curve.

[2] Spam Monitoring Project. http://www.netsec.colostate.edu/~dibenede/smp/main.php.

[3] A. Ramachandran and Nick Feamster. Understanding the Network-Level Behavior of Spammers. *ACM SIGCOMM*, 2006.

[4] Brian Krebs. Host of Internet Spam Groups is Cut Off. http://www.washingtonpost.com/wp-dyn/content/article/2008/11/12/AR2008111200658.html.

[5] Brian Krebs. Major Source of Online Scams and Spams Knocked Offline. http://voices.washingtonpost.com/securityfix/2008/11/major_source_of_online_scams_a.html.

[6] J. Heidemann, Y. Pradkin, R. Govindan, C. Papadopoulos. Census and Survey of the Visible Internet (extended). Technical report.

[7] Joel Hruska. Spam sees big nosedive as rogue ISP McColo knocked offline. http://arstechnica.com/security/news/2008/11/spam-sees-big-nosedive-as-rogue-isp-mccolo-knocked-of
ars.

[8] M. P. Collins, T. Shimeall, S. Faber, J. Janies, R. Weaver, M. De Shon. Using uncleanliness to predict future botnet addresses. *IMC*, 2007.

[9] Nigel Wetters Gourlay. IP::Country. http://search.cpan.org/~nwetters/IP-Country-2.26/lib/IP/Country.pm.

[10] Team Cymru. Internet Malicious Activity Map. http://www.team-cymru.org/Monitoring/Malevolence/hilbert.html, 2009.

[11] Y. Xie, F. Yu, K. Achan, R. Panigrahy, G. Hulten, I. Osipkov. Spamming Botnets: Signatures and Characteristics. *ACM SIGCOMM*, 2008.