Behavioral Patterns of Fast Flux Service Networks

Alper Caglayan
Milcord LLC
acaglayan@milcord.com

Mike Toothaker, Dan Drapaeau
Milcord LLC
mtoothaker@milcord.com

Dustin Burke, Gerry Eaton
Milcord LLC
dburke@milcord.com

Abstract

We present behavioral pattern analysis of fast flux service networks (FFSNs) using our database of FFSNs collected over a period of 12 months with our real-time fast flux network detection algorithm [1]. FFSNs exploit a network of compromised machines (zombies) for illegal activities such as spam campaigns, phishing scams and malware delivery using DNS record manipulation techniques. Our results, which build upon our analysis results [2], show that such networks share common lifecycle characteristics, and form clusters based on size, growth and type of malicious behavior. In particular, we introduce a social network connectivity metric, and show that (Command and Control and phishing), (malware and spam botnets) have similar scores with this metric.

1. Introduction

ICANN describes [4] fast flux as 'rapid and repeated changes to host and/or name server resource records, which result in rapidly changing the IP address to which the domain name of an Internet host or name server resolves'. While fast flux methods do have a legitimate use as a load balancing technique for high availability and high volume Web sites, its malicious use enables concealment of the Command and Control server using compromised machines ('zombies') that are used in DDoS, spam, phishing, malware delivery. There are three main variants of fast flux hosting: (1) basic fast flux hosting where IP addresses of malicious web sites are fluxed, (2) Name Server (NS) fluxing where IP addresses of DNS name servers are fluxed, and (3) double flux, where IP addresses of web sites and name servers are fluxed [4].

Published works on fast flux databases include the ISOC Network and Distributed System Security Symposium (NDSS) paper [5] on measuring and detecting fast flux service networks, the FluXOR paper [6] on detecting and monitoring fast-flux service networks, and our CATCH paper detecting and classifying fast flux service networks in real time. The ISOC paper uses a direct DNS monitoring approach over seven weeks of collected data, and is based on building a linear classifier using a flux score, which is a function of number of unique A records in all lookups, number of NS records in a single lookup, and number of unique ASNs (Autonomous System Number). The FluXOR method collects domains from spam emails in honeypots, monitors their DNS over a period of 3 hours and uses a trained Naïve Bayes classifier to classify as benign or fast-flux.

Our approach detailed in [1] complements current data collection research by focusing on the real time detection and classification of fast flux service networks using both active and passive DNS monitoring. We employ a Bayesian classifier that fuses multiple indicators including fast flux activity index, network footprint index, TTL, guilt by association, and others. In our earlier paper [2], we analyzed the short-term, long-term, organizational and operational fast flux service networks. In this paper, we analyze the structural relationships (domain, nameserver, IP connectivity) of fast flux botnets, identify recurrent structural clusters across different botnet types, and demonstrate the guilt by association knowledge encoded in these structures. For instance, for a new suspicious domain, having the IP it resolves to or having one of its nameservers in our database tagged as fast flux speeds up the detection process. Using a social network connectivity metric, we show that {Command and Control and phishing}, {malware and spam botnets} have similar structural scores with this metric.

Other related research includes the correlation of multiple DNS returns [7] at the University of Melbourne. This research has shown that the correlation of evidence from multiple DNS servers offers substantial speed up in the detection of fast flux botnets. In our approach, the use of different DNS servers corresponding in our active and passive monitoring subsystems offers the same advantage. Research at Indiana University [8] focused on the longevity of phishing botnets lasting less than 10 days to see if fast flux evasion increases the expected lifespan. This research shows that double flux increases the lifespan of phishing botnets. In our research, we focused on the lifespan of spam, CnC, malware and phishing networks. Research at Georgia Tech [9] analyzed the rate of change in DNS records for spam botnets using fast flux evasion techniques. In particular, scam domain change on shorter time intervals than their TTL values.

IEEE computer society

978-0-7695-3869-3/10 $26.00

2. Fast Flux Behavior Patterns

2.1 Data Collection

We collected our fast flux database using our Fast Flux Monitor (FFM); a Web service application designed to detect whether a domain exhibits fast flux (FF) or double flux (DF) behavior. The primary technical components of FFM include: (1) sensors which perform real-time detection of FF service networks using behavioral analysis that examine various indicators, (2) a database of known FF service networks – zombie IPs used for domain names, nameservers, and (3) analytical knowledge harvested from the database, which can include: (i) the fast flux service network's size and growth rate estimates , (ii) the social network of a fast flux service network where IPs are shared by different fast flux service networks, (iii) the footprint of a fast flux service network for a given enterprise, (iv) the footprint of a fast flux service network for a given ISP, and (v) the footprint of a fast flux service network for a given country.

We have employed multiple sensors for our FFM active sensors: (1) FF Activity Index, (2) Footprint Index, and (3) Time To Live (TTL), and (4) Guilt by Association Score. In active monitoring, we perform DNS lookup with dig, and record the A records returned with each query. For nameservers, we perform dig in order to resolve a set of nameservers. For each nameserver, we perform an nslookup in order to resolve the set of IP addresses associated with the nameservers. We then query our database to see if any of the resultant IP addresses have been associated with other domains that we have been monitoring.

Table 1. Fast flux botnet domain coverage

Type	Total (count)	Fast Flux (count)	Fast Flux (%)	Inactive (count)
Spam	207,497	12,927	6.0	10,558
CnC	3,085	55	1.7	5
Phishing	42,052	1,149	2.7	682
Malware	27,405	219	0.1	93
Total	280,039	14,350	5.0	11,338

Table 1 shows the coverage of our Fast Flux Botnet database, which began monitoring in March 2008. This database is the input into the pattern analysis reported in this paper. Our Fast Flux Botnet Database contains over 280,000 domains. About 5% of these domains have been classified as fast flux by our real-time fast flux detection algorithm. Fast flux spam botnets constitute the largest category in our database followed by fast flux phishing, Command and Control (CnC), and malware delivery botnets. There are a significant number of inactive fast flux domains, enabling us to study the long-term behavior of fast flux botnets.

Table 2. Botnet attributes in fast flux database

Entities	Total (count)	Fast Flux (count)	Active (count)
Domains	376,209	16,887	5,807
IPs	424,940	187,785	91,545
Nameservers	162,939	16,954	2,403
Total	964,088	221,626	99,755

Table 2 shows the coverage of fast flux network attributes - domains, domain IPs, and nameservers. In this paper, we find the patterns in the relationships between these attributes of various types of botnets. Our database contains over 5,800 active fast flux domains with over 90,000 IPs and 2,400 nameservers. When inactive ones are taken into account, there are over 280,000 fast flux domain, IP, and nameserver entities in our fast flux botnet database.

In contrast to the database restricted to phishing botnets in [8] and the database restricted to spam botnets, having multiple types of botnets for spam, phishing, malware and Command and Control enables us to compare the lifespan, size, and structural connectivity of different fast flux botnet types. Our analysis answers these questions for fast flux service networks: *Are spam botnets larger in size than phishing botnets? Is the lifespan of malware botnets longer than that of phishing botnets? Do malware botnets use more nameserver fluxing than phishing botnets? Does (the number of domains/the number of nameservers) change across botnet types? Is there sufficient guilt by association knowledge encoded in domain/nameserver/IP graphs that can be exploited for botnet detection?*

2.2 Short Term Behavior

Figure 1 shows the Fast Flux Activity Index, in the form of a sparkline, for a known fast flux domain involved in the

Canadian Pharmacy family of spam. Fast Flux Activity Index for domains is an empirical belief function that weighs the evidence for fast flux (i.e. new A records) against the evidence of not fast flux (i.e. A records previously seen) over a moving window of 10 minutes. The red bar shows the presence of fast flux activity, and green shows the absence of fast flux activity, while the yellow bar shows the borderline behavior. The height of the bar signifies the confidence level associated with the labeled activity.

Each row in Figure 1 represents the last 20 Fast Flux Activity Index computation scores. Subject to the domain priority and server load associated with the number of domains monitored, we compute the Activity Index on the average every 2 minutes by using the past 10 minutes of data. The number of Activity Index computation snapshots shown every 3 minutes changes depending on scheduled domains of different priorities and loading on the server. For instance, between 1:55 and 1:58, the Fast Flux Activity Index has been computed 3 times whereas the Fast Flux Index has been computed only once between 1:58 and 2:01 for this example.

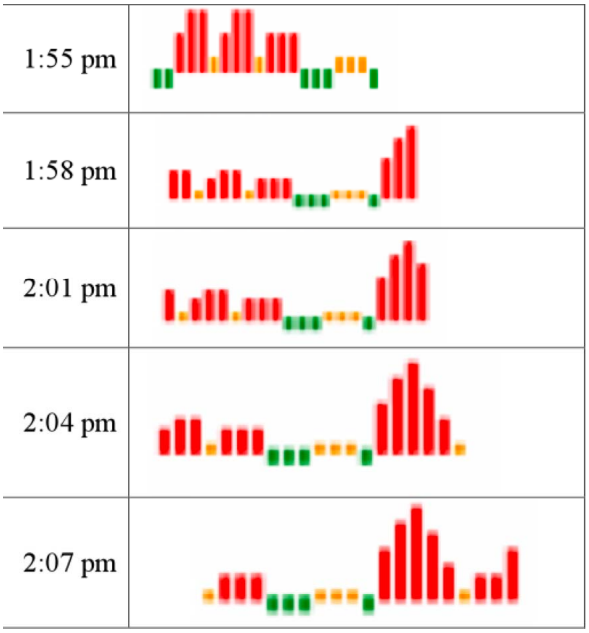

Figure 1. Fast flux activity index

Validating the results reported in [9], this spam botnet fluxes faster than its TTL value. Within a span of 12 minutes starting at 1:55, our Fast Flux Monitor checked this web site 9 times and found significant flux activity 8 times, and border line fast flux activity once.

The use of a sparkline gives us the ability to show large amounts of data in a small space. Figure 1 shows 12 minutes of activity sampled every 2 minutes. As seen, the monitored domain does not consistently exhibit fast flux behavior, but there are intervals where such behavior is dominant. According to our research, fast flux activity occurs much slower at the nameserver level, sometimes as slow as once every 12-24 hours. Known safe domains had no change after weeks of monitoring.

2.3 Long Term Behavior

We analyzed the long-term behavior of fast flux service networks. Figure 2 shows the lifespan distribution of fast flux botnets. Here the x-axis scale denotes the lifespan of an inactive fast flux botnet in number of days. The y-axis shows the number of inactive domains corresponding to specific lifespan measured in days. As shown in the figure, we found fast flux botnet domains have a shelf life of less than 3 months. There are spikes in the number of fast flux botnets for a lifespan of 30 and 70 days.

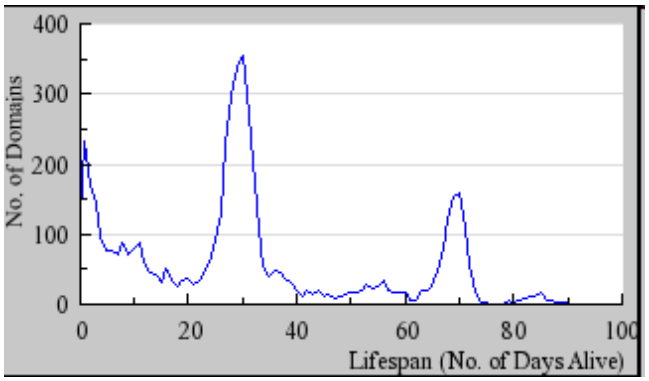

Figure 2. Lifespan distribution of fast flux botnets

Figure 2 shows the lifespan distribution for the general population of botnets. We also analyzed the lifespan distribution of fast flux botnets used for spam campaigns, phishing scams, and malware delivery separately. Figure 3 shows the lifespan of fast flux malware (blue), phishing (green), and spam (red) botnets. In terms of the fast flux botnet lifespan, spam botnets outlive malware botnets, which, in turn, outlive phishing botnets. Phishing fast flux botnets tend to die out very quickly, most living less than a week. In contrast to double flux getting more attention as posited in [8], we believe that phishing botnets targeting well-established brands, receive the attention of brand protection takedown services. In contrast, malware delivery and spam botnets attack the general population, do not have well defined targets for retaliation, and live up to 30 and 90 days, respectively.

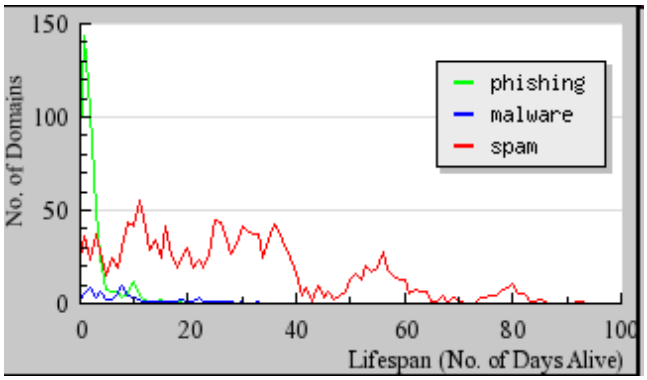

Figure 3. Lifespan: spam vs. malware vs. phishing

Figure 4 shows the semi-log plot of fast flux botnet domain size (defined as the number of IPs that this domain has resolved to) against lifespan in number of days lived. As seen from the figure, the network size is clustered between 100 to 1,000 zombie botnets. It is worth noting that we are referring to the size of the fast flux domain used in the specific malware delivery, phishing spam, and spam campaigns. The botnet sizes are much larger than the domain sizes as multiple campaigns are conducted by the same botnet. The larger botnet domains tend to live longer than their smaller counterparts. The data shows that the network size seems to converge to an asymptote after 1,000 zombies, which implies that lifespan gained per number of additional zombies recruited diminish. Our data also shows the truncation of lifespan after 3 months. This observation agrees with the findings of other researchers in the field. It is not clear whether this behavior is forced as a result of takedowns, or voluntary on the part of criminal organizations to avoid detection.

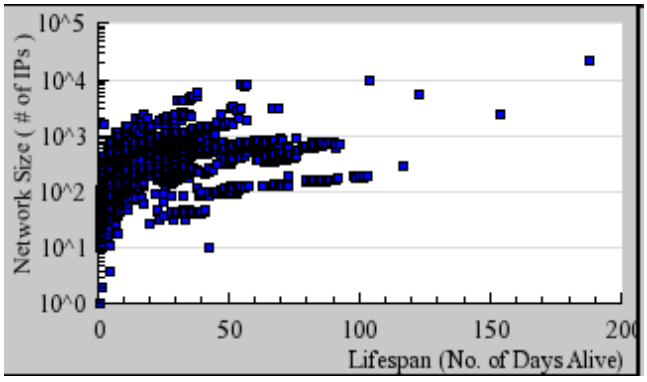

Figure 4. Botnet domains: size vs. lifespan

Figure 4 shows the composite picture of botnet domain size vs. lifespan. If we look at the same distribution across different types of botnets like malware, phishing and spam, some differences emerge. The network size seems to converge to an asymptote after 1,000 zombies, which implies that lifespan gained per number of additional zombies used diminish. Figure 5 shows the same distribution for fast flux botnet domains used for phishing campaigns, showing shorter lifespan rarely reaching the 1,000 zombies limit.

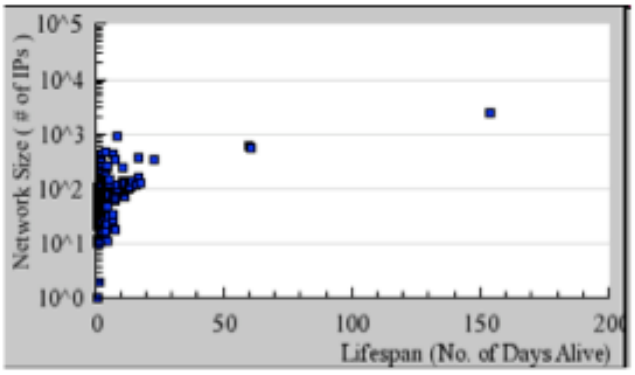

Figure 5. Phishing Domains: Size vs. Lifespan

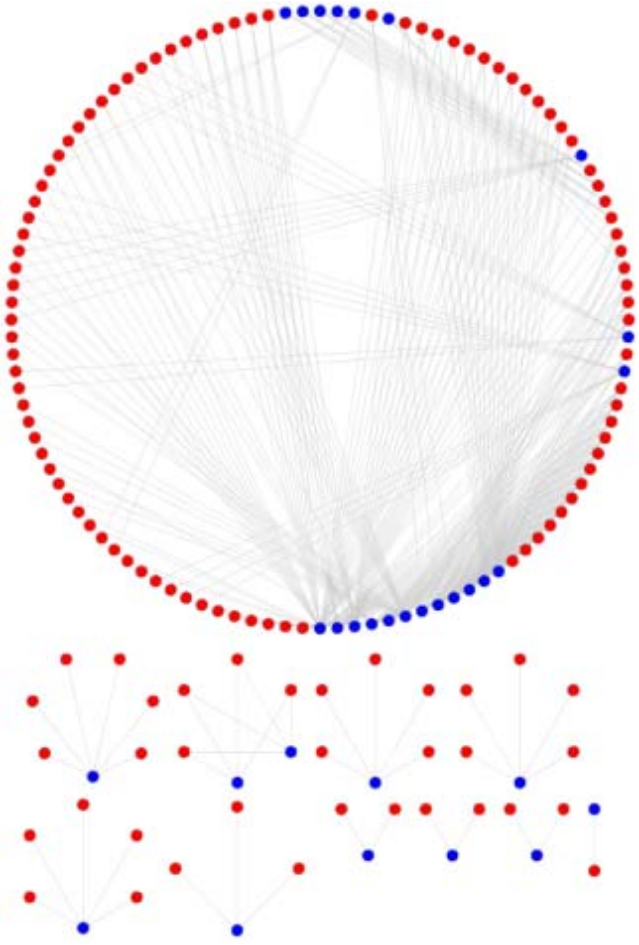

Figure 6. CnC Botnet Social Network

2.4 Organizational Behavior

Analytic sensors are derived from our cumulative collection of observed activities. We have developed a number of 'Guilt by Association' sensors. These analytical sensors include domains sharing a known guilty nameserver, domain names resolving to a known guilty IP, and nameserver domains resolving to a known guilty IP. In addition to helping real-time detection, such guilt by association relationships generate a rich social network view of the fast flux domain networks, which we used to analyze and cluster fast flux botnet domains.

Figure 6 shows the 10 social network clusters for Command & Control (CnC) domains where the red node signifies a nameserver, a blue node denotes a domain of a fast flux service network, and the edges show the relationship between nameservers and domains. Each cluster represents a set of domains and nameservers where any 2 nodes can be linked through the social network of domains and nameservers used for Guilt by Association. For instance, in the top large cluster in Figure 6, any two domains, or nameservers, or domain/nameserver combination is linked though a shared domain or nameserver. The dominant pattern in Figure 6 is that there are several more nameservers than domains in each CnC social network cluster.

Figure 7 shows the 62 social network clusters of fast flux networks used for malware delivery. Similar to the CnC fast flux social networks, on the average, there are more nameservers than domains in each cluster. Counting from the top, except for the 9th, 21st, 36th cluster, each cluster has more nameservers than domains. In contrast, the social network for fast flux networks used by phishing shown in Figure 8 shows that there are on the average more domains than nameservers in the phishing clusters.

The structural connectivity is important in speeding up the detection of fast flux botnets. Referring to the clusters in Figure 7 with multiple domains (blue dots), a new domain can be declared as fast flux if its nameservers have been classified as fast flux. In addition, the new domain can be assigned to an existing cluster, which is typically an ongoing botnet campaign like Waledac, thus improving situation awareness.

Table 3. Botnet Social Network Cluster Statistics

Type	Cluster Size		Domain		Nameserver		Domain to Nameserver Ratio	
	Average	Std Dev	Average Count	Std Dev	Average Count	Std Dev	Average	Std Dev
Malware	13	38	3	8	10	30	0.34	0.23
Spam	71	445	51	341	20	105	0.96	1.14
CnC	15	33	3	6	12	26	0.39	0.25
Phishing	19	77	11	50	9	28	1.20	2.04

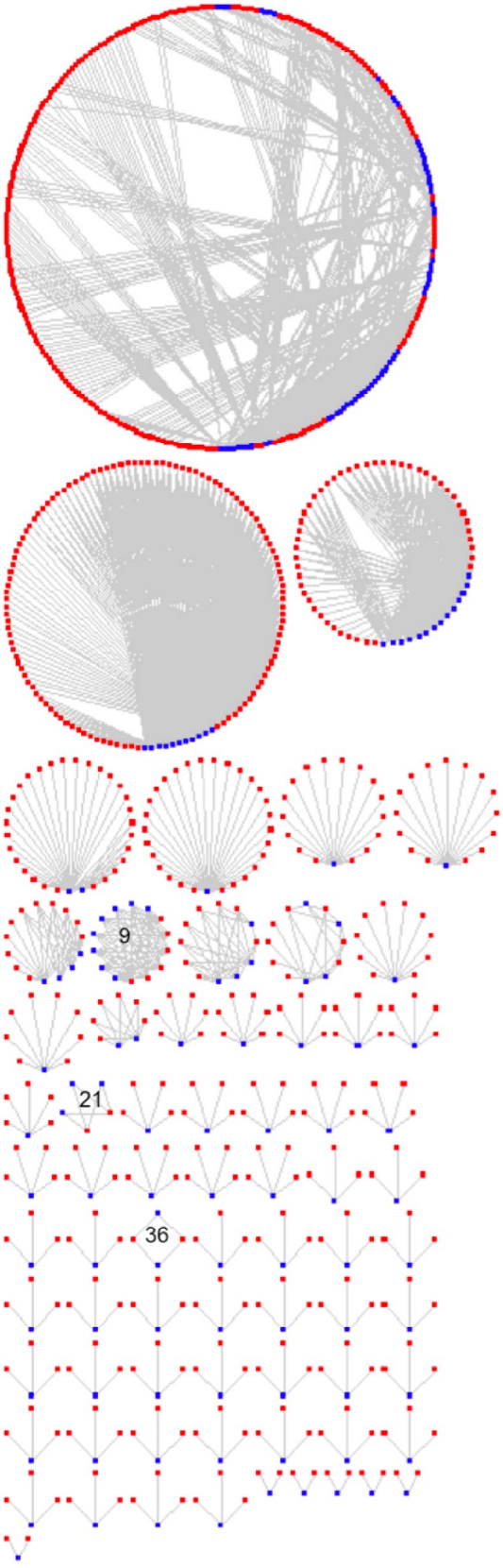

Figure 7. Malware Fast Flux Social Network

Table 3 gives the cluster statistics of fast flux networks used for malware delivery, spam campaign, Command and Control, and phishing scam. The first two columns give the average cluster (counting both domains and nameservers) size and the associated standard deviation. The next two columns show the average number of domains in a cluster and the associated standard deviation. The next two columns show the average number of nameservers in a cluster and the associated standard deviation. The final two columns show the average and standard deviation of a connectivity metric – D/N, which we define as the number of domains divided by the number of nameservers. Based on the cluster attribute statistics, we observe the following trends:

- With respect to size, fast flux spam networks have the largest clusters.
- With respect to D/N connectivity metric, malware and CnC are similar; spam and phishing are similar.

Based on our D/N connectivity metric, Command and Control and malware related domains can be grouped together (average ratio of about 0.3) and spam and phishing domains can be grouped together (average ratio of approx. 1.0). Compared to spam and phishing, D/N standard deviation for both malware and CnC is tighter, thus indicating significance. This grouping is most likely related to the ease of updating domains primarily accessed through websites versus domains primarily accessed via a binary. For a spam or phishing network operator sending out large quantities of email trying to get people to go to websites and buy products or input personal information, it is easy to change domains or nameservers on a campaign-by-campaign basis. On the other hand, it is more difficult to update the botnet software on all zombies of a large botnet in order to hide the command and control system. For this reason, the botnet operators change nameservers as a way to keep from having to shut down the botnet.

Table 4. Small Phishing Cluster

Name	Type
"ns1.mydnsonyouiscool.com"	nameserver
"ns5.mydnsonyouiscool.com"	nameserver
"ns3.mydnsonyouiscool.com"	nameserver
"web.da-us.citibank.com.p6v22qa74.cn"	domain
"ns4.mydnsonyouiscool.com"	nameserver
"ns7.mydnsonyouiscool.com"	nameserver

Let's take a closer look at some of the clusters from Figure 8 to gain insight about the reasons for the specific graph structures. Starting from the top, let's look at the *first* and *second* clusters in the *second* row from the top. Looking at the node data for these two clusters we can see that they are both very specialized and targeted attacks.

The *first* cluster in the *second* row from the top is a phishing attack on eBay. The two domains (blue nodes) are phishing domains for "pages.ebay.com" and "cgi.ebay.com" and take on the form of "<target domain>.<phisher domain>". All of the nameservers (red nodes) are of the form "<ip address>.ip4.<phisherdomain>". Examination of the names of these nodes reveals the following. First, the phisher thought it was more important to keep the domains static, in order to do so they chose to rapidly change the location of the nameservers. The names of the nameservers would further indicate that they are likely zombie nodes of a botnet.

The *second* cluster from the left in the *second* row from the top depicts a different method. Here there are a large number of domains all being handled by only 5 name servers. Here again the examination of the names used for the domains reveals the rationale behind this structure. The domains in this case are set up to look like Google ads session url's and take the form "adwords.google.com.session-<random numbers>.<phisher domain>". Unlike the previous example, the domains are the disposable piece of the puzzle and the nameservers are more important, possibly containing specialized code for creating unique domain names on the fly.

Table 4 shows the domain and nameservers depicted in the *first* cluster in the *fifth* row from the bottom. This network along with the three following appear to be structures that were on their way to growing into a network such as the one for phishing eBay discussed above. Considering the highly visible target (Citibank), it is reasonable to assume that this network was probably taken down early in its life cycle by a Digital Brand Protection company, and, therefore, did not get to grow to its full potential.

Figure 9 shows the number of shared IP addresses across fast flux domains using a log-log scale. The plot shows a linear trend in this scale in that the number of IP addresses shared decreases with increasing number of fast flux domains. For instance, there are 100 botnet domains sharing 100 IP addresses whereas there are only 10 botnet domains sharing 1,000 IP addresses, resulting from having more small botnets than large ones.

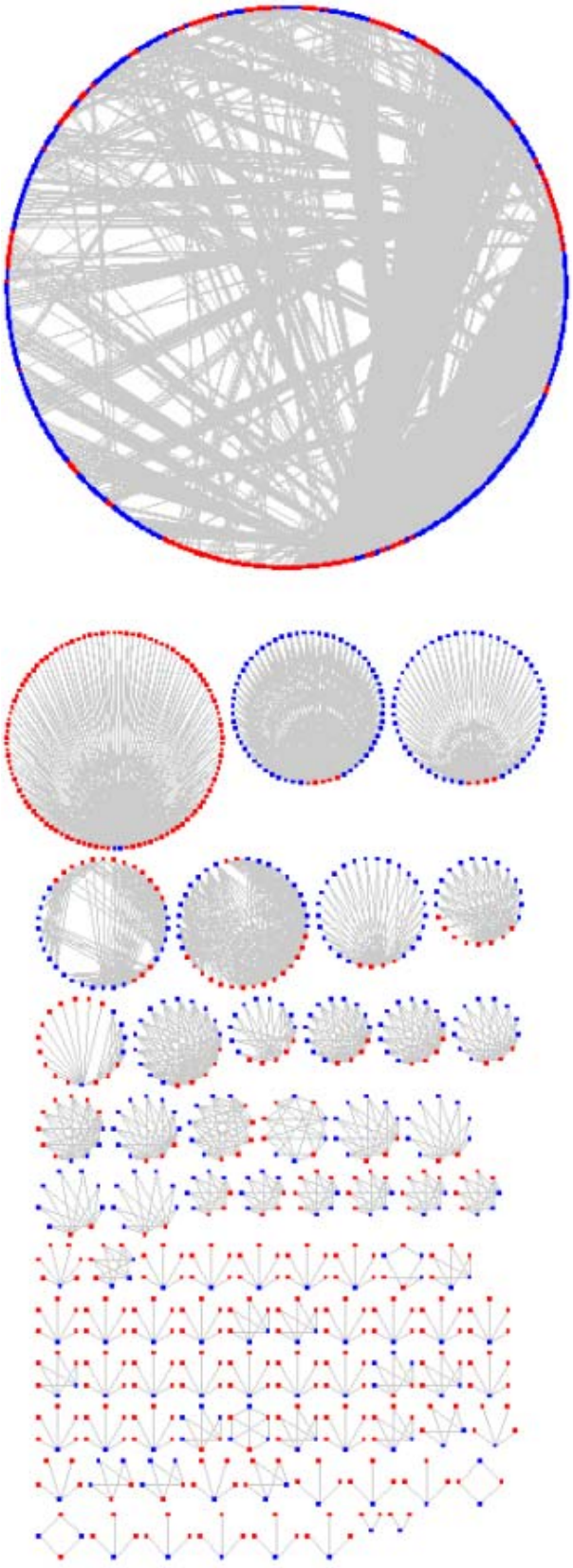

Figure 8. Phishing Botnet Social Network

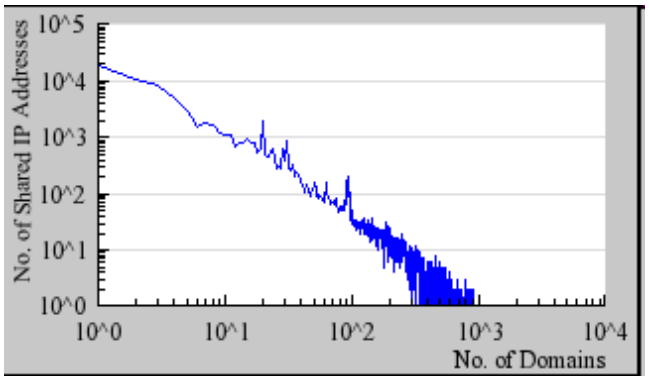

Figure 9. Distributions of IP Shared

The previous social network graphs show the relationships between domains and nameservers. Figure 10 extends the graph by including IP address to domain relationships where the red, blue and green nodes signify nameserver, domain, and domain IP's, respectively. Domain-> Nameserver relationships graph shows 11 networks, and the IP address->Domain relationship graph shows 19 networks. The examination of the two graphs together reveals that these CnC domains all belong to just 9 networks, making it easier to identify new CnC domains, thus increasing situation awareness.

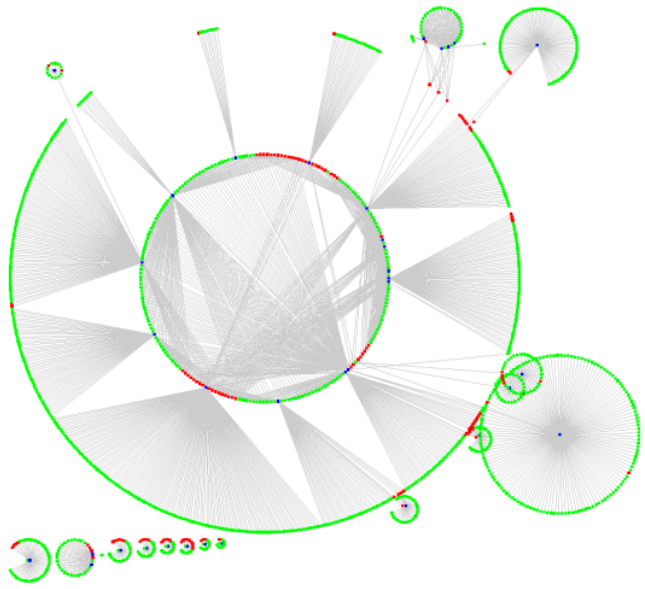

Figure 10. CnC IP/Domain/Nameserver Social Network

2.5 Operational behavior

We analyzed each botnet's distribution across different countries, and Autonomous System Numbers (ASNs). Figure 11 shows the number of countries where fast flux botnets operate. While there are a large number of botnets operating in a few (less than 5) countries, most of the botnets operate in between 20 and 40 countries. The number of botnets operating in more than 40 counties falls off sharply.

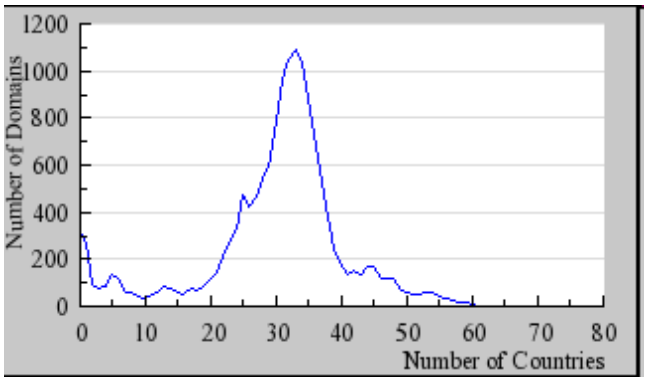

Figure 11. Operations in Multiple Countries

We also analyzed the distribution of fast flux botnet across ASNs. Figure 12 shows the number of ASNs in which fast flux botnets operate. While there are a large number botnets (100) operating in less than 10 ASNs, fast flux botnets operating in 10 – 250 ASNs have a bimodal distribution with sizable mass at modes correspondiung to 100 and 175 ASNs. We believe that the bimodal distribution is due to the presence of two large botnets (e.g. Waledac) in our database. Removal of the large botnets from consideration yields a uniform distribution between 10 and 250 ASNs. The ASN distribution has a sharp fall off after 250 ASNs with practically no botnets operating in more than 375 ASNs. We suspect that his behavior is due to the limited size of the botnet ASN targets, namely, ISPs and universities.

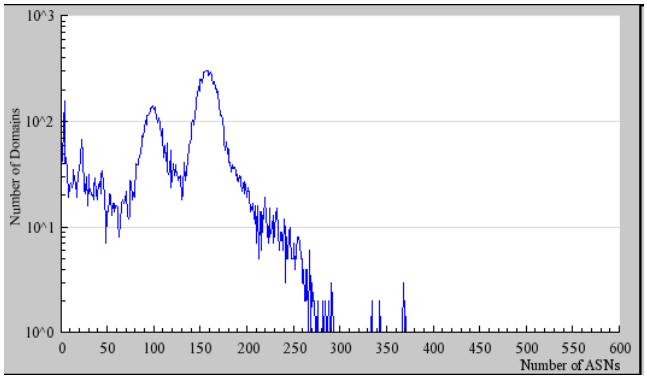

Figure 12. Operations in Multiple ASNs

3. Conclusions

In this paper, we presented a behavioral analysis of fast flux service networks using our botnet database collected over 12 months. Our results show that such networks share common lifecycle characteristics, and form clusters based on size, growth, attack campaigns, and social network structure. Our results show:

- Fast flux spam networks have the longest lifespan compared to fast flux malware and phishing networks.
- Fast flux phishing networks have the shortest lifespan compared to fast flux malware and spam networks.
- The social network of fast flux spam networks have the largest clusters compared to fast flux malware and phishing networks.
- Domain to nameserver ratio serves as a useful metric capturing the connectivity characteristics of the fast flux social network behavior across different types of botnets.
- Fast flux malware and CnC social networks have similar domain/nameserver connectivity metric scores whereas fast flux spam and phishing social networks have similar scores, indicating the use of nameserver fluxing for CnC and malware networks.
- The number of shared IP addresses across fast flux domains shows a log/log linear trend where the number of IP addresses shared decreases with increasing number of fast flux domains.
- The combination of {IP, Domain} and {Domain, Nameserver} social networks increase situation awareness for guilt by association.
- Most botnets operate in between 20 and 40 countries and in less than 250 ASNs.

Understanding the domain, nameserver, and IP connectivity patterns is essential in detecting and shutting down botnet operators. For detection, our results show that there is quite a large body of Guilt by Association knowledge captured in domain/nameserver/ botnet graphs that can be harvested in botnet detection algorithms. For takedown operations, CnC IP/Domain/Nameserver graphs offer the insight necessary for investigations.

4. Acknowledgements

This research was supported by Department of Homeland Security, Science and Technology Directorate Cybersecurity R&D program.

5. References

[1] Caglayan, A., Toothaker, M., Drapeau, D., Burke, D., and Eaton, G., "Real-time Detection and Classification of Fast Flux Service Networks", Cybersecurity Applications and Technology Conference for Homeland Security (CATCH), March 3 - 4, 2009, Washington, DC.

[2] Caglayan, A., Toothaker, M., Drapeau, D., Burke, D., and Eaton, G., "Behavioral Analysis of Fast Flux Service Networks", Cyber Security and Information Intelligence Research Workshop (CSIIRW-09), April 13 - 15, 2009, Oak Ridge, TN.

[3] ICANN. GNSO Issues Report on Fast Flux Hosting, March 2008.

[4] ICANN Security and Stability Advisory Committee. SAC 025: SSA Advisory on Fast Flux Hosting and DNS, March 2008.

[5] Holz, T. Gorecki, C. Rieck, C. Freiling, F. "Measuring and Detecting Fast-Flux Service Networks." Presented at NDSS Symposium (2008).

[6] Passerini, E. Paleari, R. Martignoni, L. Bruschi, D. "FluXOR: detecting and monitoring fast-flux service networks." Detection of Intrusions and Malware, and Vulnerability Assessment (2008), pp. 186-206.

[7] Zhou, C. V., Leckie, C. and Karunasekera, S., "Collaborative Detection of Fast Flux Phishing Domains", Journal of Networks, Vol. 4, No. 1, February 2009.

[8] McGrath, D. K., Kalafut, A., Gupta, M., "Phishing Infrastructure Fluxes All the Way", IEEE Security and Privacy Magazine Special Issue on Securing the Domain Name System, September/October 2009.

[9] Konte, M., Feamster, N. , and Jung, J., "Dynamics of Online Scam Hosting Infrastructure", Proceedings of Passive and Active Measurement Conference (PAM), Seoul, Korea, April 2009.