My Botnet is Bigger than Yours (Maybe, Better than Yours) :
why size estimates remain challenging

Moheeb Abu Rajab Jay Zarfoss Fabian Monrose Andreas Terzis
Computer Science Department
Johns Hopkins University

Abstract

As if fueled by its own fire, curiosity and speculation regarding botnet sizes abounds. Among researchers, in the press, and in the classroom—the questions regarding the widespread effect of botnets seem never-ending: what are they? how many are there? what are they used for? Yet, time and time again, one lingering question remains: how big are today's botnets? We hear widely diverging answers. In fact, some may argue, contradictory. The root cause for this confusion is that the term botnet size is currently poorly defined. We elucidate this issue by presenting different metrics for counting botnet membership and show that they lead to widely different size estimates for a large number of botnets we tracked. In particular, we show how several issues, including cloning, temporary migration, and hidden structures significantly increase the difficulty of determining botnet size with any accuracy. Taken as a whole, this paper calls into question speculations about botnet size, and more so, questions whether size really matters.

1 Introduction

It is widely accepted that botnets pose one of the most significant threats to the Internet. For the most part, this belief has been supported by the conjecture that at any moment in time, there is a large collection of well-connected compromised machines that can be coordinated to partake in malicious activities at the whim of their botmaster(s). Indeed, the potential threat of botnets comprising several hundred thousands bots has recently captured the headlines of the press [11, 18], but the question of size itself, continues to be a point of debate among the research community.

In particular, the question of how we arrive at size estimates, or more importantly, just what they mean, remains unanswered. As a case in point, while earlier studies (*e.g.,* [4, 5, 14]) have proposed a number of techniques to measure the size of botnets, they provide very inconsistent estimates. For example, while Dagon *et al.* [5] established that botnet sizes can reach 350,000 members, the study of Rajab *et al.* [14] seems to indicate that the effective sizes of botnets rarely exceed a few thousand bots. Clearly, something is amiss.

In this paper, we attempt to shed light on the question of botnet membership. Our study primarily focuses on IRC botnets because of their continuing prominence in the Internet today. Specifically, we survey a number of techniques for determining botnet membership and examine the different views they generate. As we show later, the inconsistency among the resulting outcomes is largely tied to the counting techniques being used, and does not necessarily reflect a change in underlying activity during the time that these studies were undertaken. For example, one of the botnets we tracked appeared to consist of 48,500 bots over the entire tracking period. However, if we examine the bots that simultaneously appeared on the bot server in question, the size does not exceed 3,000 bots. At a high level, this suggests that expecting a single definitive answer to the question of botnet membership is unreasonable. Instead, "botnet size" should be a qualified term that includes the specifics of the counting method, its caveats, and the context in which the measurements are relevant.

Additionally, we show that the issue of botnet membership extends beyond single botnet considerations in that potential cross-botnet relationships add another challenge to estimating membership. Specifically, our preliminary insights raise questions about the extent to which we can assert that two or more botnets are different, or more importantly, the degree of overlap among the populations of different botnets.

In summary, this paper makes the following contributions: *(i)* we explore a number of mechanisms (including prior work) for estimating botnet sizes and highlight the challenges associated with each, *(ii)* we present results of applying these techniques to data derived from a large-scale measurement study and show the extent of

the discrepancy between the different size estimates, and *(iii)* we examine potential hidden structures among botnets we tracked and highlight their implication on determining botnet membership.

The remainder of this paper is organized as follows. Section 2 provides a comprehensive list of botnet size estimation techniques and highlights the challenges associated with each. In Section 3 we present the results of applying these techniques to botnet data collected from a wide-scale monitoring experiment. In Section 4 we examine the existence of hidden structures among the botnets we tracked, while Section 5 discusses related work. We conclude in Section 6 with a discussion about the subtleties associated with counting botnet membership.

2 Size Matters, But What Does It Mean?

While the topic of large botnets has certainly captured the attention of academicians and practitioners alike [5, 7, 12, 14, 15, 16], there seems to be little, if any, agreement on what specifically the size of a botnet refers to. Arguably, the only consensus seems to be that a botnet's size is the main determining factor of its perceived impact. However, unlike other classes of malware (*e.g.*, worms), where the size of the infected population determines the impact of the outbreak, botnet size can convey several meanings. Therefore, to clear the fog on this issue, we start by providing different definitions of botnet size and detail the context in which each definition is relevant.

In what follows, we draw the distinction between two main terms. First, we denote a botnet's *footprint* as the overall size of the infected population at any point in its lifetime. While this measure reflects how wide spread a botnet infection is, it fails to capture the actual capacity of the botnet to execute a particular command issued by the botmaster at any given point in time. Second, we consider the botnet's *live population* as the number of live bots simultaneously present in the command and control channel. Therefore, unlike its footprint, the live population of a particular botnet indicates the botnet's capacity to perform a malicious task posted as a command message by the botmaster.

Generally speaking, the estimation techniques we survey belong to two broad categories based on the information used. Next, we elaborate on each category, detailing the estimation techniques, their challenges, and their relevance in light of the aforementioned notions.

2.1 A View From Within

The first category includes techniques that directly count bots connecting to a particular server. Two main variants are relevant here: *infiltration* and *redirection*.

Botnet Infiltration. An obvious way to learn several aspects of a botnet's activity is to infiltrate the botnet by joining the command and control channel. Botnet infiltration provides valuable information about several malicious activities such as DDoS attacks as shown earlier by Freiling *et al.* [7]. In our earlier work [14], we used botnet infiltration to provide in-depth analysis of several facets of botnets, including inferring their membership by directly counting the bots observed on individual command and control channels. To achieve this, we developed a lightweight IRC tracker (see [14] for details). In a nutshell, the tracker intelligently mimics the behavior of actual bots and joins a number of botnets, all the while recording any information observed on the command and control channel. This information may include the identities of all active bots. In this case, the botnet's footprint is simply the total number of unique identities observed on the channel over the entire tracking period. Similarly, the botnet's live population is measured by counting the number of bots simultaneously present on the channel at a particular time. In some cases, this estimate can also be derived from the IRC server's welcome message.

Despite its simplicity, this technique suffers from a number of limitations. First, botmasters may suppress bot identities from being transmitted to the channel and in doing so render this technique useless. Second, even when this information is available, counting can lead to different estimates depending on whether we count the fully qualified unique user IDs or the IP addresses—be it cloaked or plain. As we show later, temporal population variations due to bot cloning and temporary migration of bots complicate this issue even further. What this means is that it is difficult to provide an accurate bot count in these cases, as distinguishing between actual bots and temporary clones or migrants is nontrivial.

DNS Redirection. As an alternative to botnet infiltration, Dagon *et al.* explored a technique for counting infected bots by manipulating the DNS entry associated with a botnet's IRC server and redirecting connections to a local sinkhole [5]. The sinkhole completed the three-way TCP handshake with bots attempting to connect to the (redirected) IRC server and recorded the IP addresses of those victims. Their results suggest the existence of large botnets with populations up to 350,000 bots. Unfortunately, although this approach allows us to observe the IP addresses of different bots, it has a number of limitations. First, this technique can only measure the botnet's footprint. The reason is that although the sinkhole observes bot connection attempts, it is impossible to know how many live bots are simultaneously connected to the actual server channel. Second, as the sinkhole does not host an actual IRC server, there is no

way of knowing if the bots are connecting to the same command and control channel. Finally, as Zou *et al.* [19] suggest, it is conceivable that botmasters can detect DNS redirection and subsequently redirect their bots to another IRC server thus distorting the estimate provided by this technique.

2.2 When The Lights Go Out

When insider information is not available because bot activities are not echoed on the channel (and so can no longer be overheard by an IRC tracker), it is still possible to estimate a botnet's size by exploiting external information. In this case, however, techniques that rely on externally visible information can only provide an estimate of a botnet's footprint.

A natural source of externally visible information about a botnet's prevalence is DNS. In our earlier work [14], we explored the use of DNS cache snooping to uncover a botnet's footprint. In short, our approach takes advantage of the fact that bots normally make a DNS query to resolve the IP address of their IRC server before joining the command and control channel. Our technique estimates a botnet's DNS footprint by probing the caches of a large collection of DNS servers and recording all cache hits. A cache hit implies that at least one bot has queried its nameserver within the time to live (TTL) interval of the DNS entry corresponding to the botnet server. The total number of cache hits provides an indication of the botnet's DNS footprint.

That said, a botnet's DNS footprint provides (at best) only a lower bound of its actual footprint. For one, using DNS to estimate size is only possible for DNS servers that allow probes from arbitrary clients and reply to queries for cached results. To yield an accurate estimate, this technique requires a large list of such servers. Moreover, botnet servers that have DNS names with low TTL furhter complicates this technique because, for such names, the probability of a cache hit from an infected domain is low. Finally, a hit indicates only the existence of at least one infected host associated with that DNS server.

Recently, Ramachandran *et al.* [15] suggested another DNS-based technique. Their approach infers bot counts by observing DNS lookups for hosts listed in DNS-based blackhole lists. The rationale behind this approach is that botmasters tend to query these lists to detect if their bots are blacklisted and thereby unusable for certain tasks (*e.g.,* sending spam e-mails). This approach has the potential to provide an overall estimate of possible bots in DNS-based blackhole lists, but clearly it cannot estimate the footprint or the live population of a specific botnet.

Duration of data collection period (days)	300
Total number of unique malicious binaries	1,400
Total number of IRC botnet binaries	1,058
Number of unique botnet channels tracked	472

Table 1. Summary of collected data.

3 Just How Big is Your Botnet?

To elucidate the discrepancies among different counting techniques, we now provide botnet size estimates using the different approaches discussed in Section 2. Where possible, we outline the factors that contribute to inflating or deflating the botnet population estimates derived by these techniques. For comparison purposes we analyze the traces of a large collection of botnets captured and tracked over a period of more than 9 months using a distributed data collection infrastructure. We established this infrastructure as part of an ongoing effort to study the botnet phenomenon. In short, we use a combination of lightweight responders (based on the `nepenthes` framework [1]) as well as deep interaction honeypots to collect malware binaries. The collected binaries are analyzed in an isolated environment to elicit any IRC related features and then produce configuration templates. These templates are used to create several customized IRC tracker instances that infiltrate the botnets specified in the collected binaries (see [14] for more details). Table 1 summarizes the data we collected, including traffic traces captured at our distributed darknet, IRC logs gathered from 472 botnet channels either visited by our IRC tracker or observed on our honeynet, and DNS cache hits from tracking 100 IRC servers for more than 45 days.

3.1 One Botnet to Rule Them All

Before addressing our main question, let us first begin by analyzing the global statistics that we can infer about the botnet problem in general from the available data. Despite earlier predictions [11], even this seemingly simple task is laden with challenges. For example, a crude count of the number of unique bots (based on user IDs) across all botnets we tracked results in an estimate of 1,153,371 bots, while counting the IP addresses (either cloaked or plain) yields a more moderate figure of 426,279 bots. However, notice that these estimates do not account for two important factors, namely, the overlap across different botnet populations (which may be substantial) and the impact of dynamic addressing (*e.g.,* DHCP and NATting), which is generally difficult to quantify [3] especially when the IP addresses of the bots are cloaked.

From another viewpoint, we note that our cache prob-

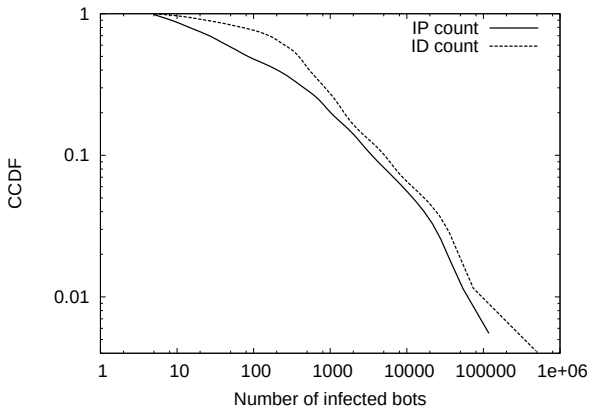

Figure 1. CCDF of the aggregate infected host population counted by unique IDs and unique IP addresses.

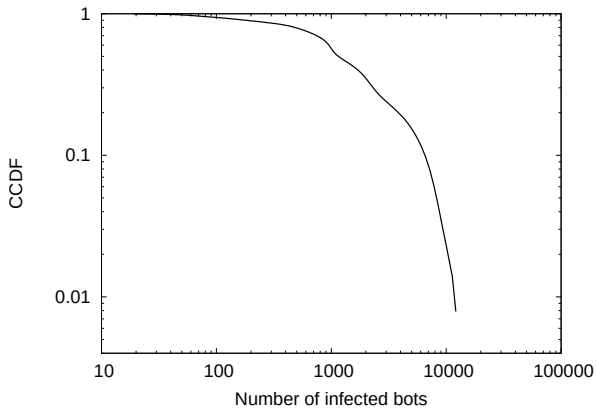

Figure 2. CCDF of the maximum number of simultaneous online bots.

ing results show evidence of at least one botnet infection in 11% of the 800,000 DNS servers we probed. While one could speculate that this figure is in agreement with the conjecture that bots reside in 11% of the overall Internet host population (*e.g.*, [11]), this claim can not be easily justified. In fact, our DNS results can not be directly extrapolated to actual bot counts.

3.2 Large Botnets May Not Be So Big After All

Returning to our outstanding question, we turn our attention to a simple question: how big is a botnet? In short the answer is, it depends. To see why, let us explore the results for estimating the size of the botnets we tracked, based on the strategies given in Section 2. Figure 1 shows the complementary cumulative density function (CCDF) of botnet footprint sizes, counted by user IDs and by IP addresses for the botnets that broadcast that information. Overall, 52% of the botnets we tracked make such data available. Notice that counting bot IP addresses versus IDs already leads to one discrepancy. While botnet sizes, by ID count, can exceed 450,000 bots, counting by IP addresses yields sizes in the range of 100,000 bots.

Figure 2 shows the CCDF of the live botnet population size for the same set of botnets. Clearly, there is an even more substantial discord in this case. While botnet footprint sizes can exceed 100,000 infections [1], their live populations are normally in the range of a few thousand bots—a significant decrease in size which has profound impact on the perceived vivaciousness of these botnets. This discrepancy can be explained by the fact that the live population of a botnet is normally constrained by the capacity of the botnet server and affected by high bot churn rates [14].

Finally, we resorted to DNS cache snooping to estimate the DNS footprints of the remaining 48% of the botnets that do not publish membership data. Figure 3 presents the CCDF of DNS footprint sizes. Because in this case we count domains rather than bots, the discrepancy between DNS footprints and infection footprints (cf. Fig. 1) is wide. Refining the estimate of a botnet's infected population from its DNS footprint is a subject that warrants further work and one we are currently pursuing.

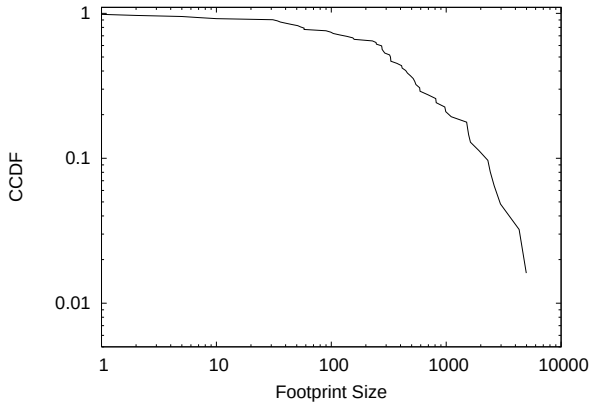

Figure 3. CCDF of the DNS footprint sizes.

3.3 Challenges and Caveats

As we alluded to earlier, there are several additional issues that complicate the task of counting botnet members. Temporary bot migration and bot cloning are major contributors to this effect. In several occasions, we

[1] These results are in accord with the estimates of Dagon *et al.* [5] derived using DNS redirection.

observed botmasters commandeering their bots to temporarily migrate from one botnet to another. In cloning, botmasters command bots to create copies of themselves and join a new channel on the same server, or to connect to a different server altogether [14]. Generally, we observed two types of cloning: *(i)* clone flooding, in which bots create a large number of instances to overwhelm a target IRC server, and *(ii)* normal cloning events in which botmasters command their bots to create a new IRC connection and join another channel on the same server or on a different server.

These observations raise the following important question: when we count botnet members are we really counting *actual* compromised machines? Although direct counting of bots by botnet infiltration seems to be the most direct way of estimating a botnet size, it is, unfortunately, unclear whether or not the resulting estimate is a count of real bots. For one, temporary botnet migration can significantly inflate the membership of a particular botnet. Figure 4, for example, presents an instance of temporary migration observed by our IRC tracker. In this example, if we were to count the population of Botnet II immediately after the migration, we would arrive at an inflated count. While on the surface this may not seem as a big concern, if such migrations occur frequently, then we could be substantially overcounting the cumulative bot population.

To further illustrate the impact of bot cloning on size estimation we extracted all clone commands observed in the IRC traces of the botnets we tracked. In this case, we only consider the events corresponding to the second type of cloning and therefore we exclude all commands corresponding to "clone flood" attacks. Overall, we observed cloning behavior in 20 tracked botnets. Interestingly, our results show that although the total footprint of these botnets was near 130,000 bots, they created a total of 2,383,500 clone instances of which roughly 10% connected to new botnet servers. Figure 5 presents an example of one such cloning event in which bots are asked to join another channel on the same server. The graph shows a sudden surge in the number of online bots reported in the server's welcome message shortly after the botmaster posted a command to her bots to create clones and join a new channel on the same server. Obviously, the population count in this case is not indicative of actual bots. Coupled with the issue of bot migration, this may be one of the underlying reasons for the wide variation in botnet sizes quoted in the literature. Unfortunately, without more qualified discussions of what botnet sizes represent, it is difficult to come to any definitive conclusions.

4 Exposing hidden botnet connections

One of the most challenging facets of the botnet membership problem lies in discerning the relationship among (seemingly) different botnets. To highlight this, we examine the existence of hidden relations among the botnets we tracked. The presence of these relations raises new challenges to the accuracy of botnet population counting techniques. Specifically, for botnets that are related, is the aggregate population count simply the sum of the different botnet populations? Or more importantly, how do we characterize the overlap between different botnet populations? In what follows, we discuss our methodology for finding potential hidden relationships among botnets.

First, we create for each botnet a d-dimensional structural feature vector $\vec{v}_i = x_1, x_2, \ldots, x_d$. We choose the following features to represent a botnet's unique identity:

1. DNS name and/or IP address of IRC Server.
2. IRC server or IRC network name (*e.g.,* `ToXiC.BoTnEt.Net`).
3. Server version (*e.g.,* Unreal3.2.3).
4. IRC channel name.
5. Botmaster ID. All these IDs are extracted from the IRC trace by observing the identity of the user with operator privileges who posts commands to the channel.

To reveal the existence of clusters of related botnets we then create a proximity matrix $\mathbf{M}$ by calculating a pair-wise scores across all botnet vectors, $\mathcal{V} = \vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}$. For a pair of vectors $\vec{v_i}, \vec{v_j}$ the pair-wise score $m_{i,j}$ is a weighted dot product of the two vectors.

$$m_{i,j} = \sum_{k=0}^{d} w_k \left(x_{i,k} \cdot x_{j,k}\right)$$

where w_k is the weight assigned to dimension k and the product of the two vector fields is one if they are identical, or zero otherwise. Considering that similarity in the names of the IRC servers implies strong correlation between two botnets, we assign a weight of 1.5 to the IRC server dimension, while all other dimensions are given equal weights of 0.5.

Given the matrix $\mathbf{M}$, we infer related botnets by extracting botnet groups that have pairwise similarity scores above a threshold δ. We choose $\delta = 1.5$, so that two botnets are related if they have the same IRC server DNS name or match in at least three other dimensions.

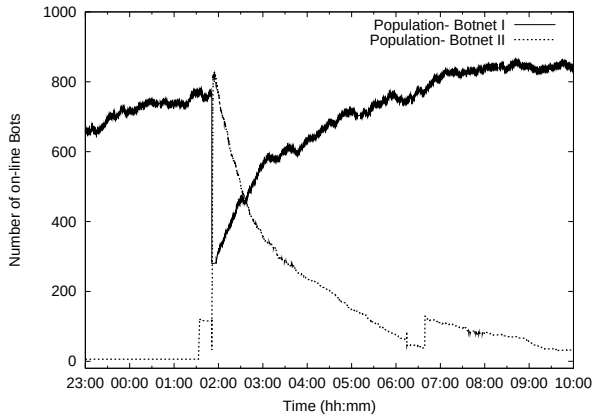

Figure 4. Botnet temporary migration instance.

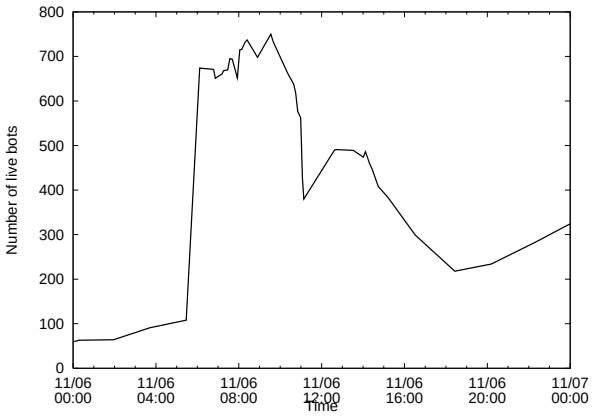

Figure 5. Bot count for a botnet with cloning.

```
 # <DNS name>    <Channel>  <Server ID>   <Botmaster ID>    <Server Version>
[1] hid.shgon.net #!GT!# IRC.Death.TeaM.KW [Lindi_Cracker]-1!HackPimp Unreal3.2.5
[2] bruimi.shgon.net #!GT!# IRC.Death.TeaM.KW ChanServ!Coder Unreal3.2.5
[3] newbot.shgon.net #.rxb0t IRC.Death.TeaM.KW Chan!Coder Unreal3.2.5
[4] bb.shgon.net #.rxbot IRC.Death.TeaM.KW Chan!Coder Unreal3.2.5
```

Figure 6. Example of a botnet cluster.

Preliminary Results. We applied this methodology to the 472 botnets we captured and tracked. Our results revealed 90 groups of related botnets covering 25% of the botnets we tracked. Figure 6 presents the features of the botnets in one of these clusters. As the figure illustrates, while these botnets used different servers, similarities across other dimensions can be used to detect their potential relationship. Notice that in this example (and many similar ones) the names of all IRC servers belong to the same DNS domain which provides additional evidence of the relationship among these botnets.

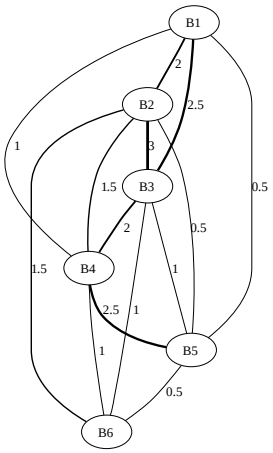

Figure 7. Example of a botnet cluster.

Figure 7 provides a graphical representation of one example cluster, with nodes indicating distinct botnets and edges indicating relationships between different botnets. The label on each edge reflects the pairwise similarity score. It is evident from this graph that botnet relationships can evolve to form rather complex clusters that significantly complicate the task of estimating botnet membership.

Figure 8 plots the CDF of the number of botnets affiliated with botnet cluster we discovered. The graph indicates that botnet clusters can span relatively large collections of botnets. Finally, we note that while code reuse [2] could explain the commonalities across some of the features we chose (*e.g.*, IRC server version), other common features, such as channel names and botmaster IDs, are more likely to indicate intentional botnet relationships. Further research into feature selection and assigning proper weights for each feature is a subject of our ongoing work.

5 Related Work

Of late, articles about botnets with hundreds of thousands of members have captured the headlines on several occasions (*e.g.*, [11, 17, 18]). This attention is warranted, as botnets undoubtedly pose a significant threat to the Internet. Starting from the early work of Freiling *et al.* [7], a number of research efforts have explored

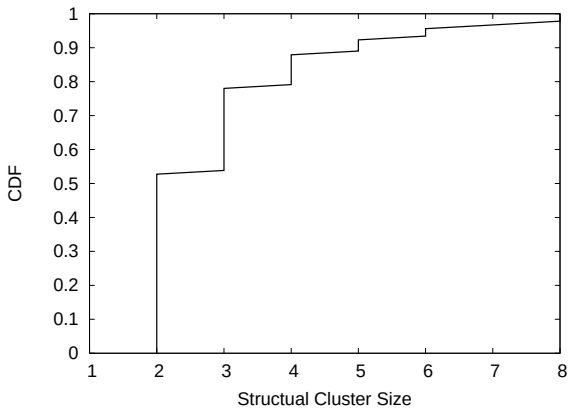

Figure 8. CDF of the number of botnets affiliated to each observed cluster.

the rise of botnets. However, the issue of determining botnet size still remains contentious. In particular, Dagon *et al.* used DNS redirection to study the size and evolution of several botnets and reported botnets with 350,000 members [5]. Similar observations were also reported in [8, 13]. In contrast, the work of Cooke *et al.* [4] and Jahanian [9] seem to point to a trend towards smaller botnets with sizes ranging from several hundreds to a few thousand hosts; many of these botnets emerge and then become defunct after relatively short periods of time [9]. In this paper we examine two techniques for gleaning information about a botnet's size namely, IRC tracking and DNS snooping [6]. Our results show that while the footprints of the botnets we tracked can grow to several tens of thousands of bots, their effective sizes usually are limited to a few thousands at any given point in their lifetime. These discrepancies argue that botnet size should be a qualified term that is relevant only within the context of the counting method used to generate the result.

Equally important to the question of size is that of the overall prevalence of botnets. While the earlier work of Rajab *et al.* [14] provided partial insights about this issue, more recent work has attempted to answer this particular question. Specifically, Ramachandran *et al.* [15] monitored queries sent to servers maintaining the DNS names of blacklisted hosts to infer the overall prevalence of bots in these lists. In this paper, we show that the same discrepancies that plague size measurements of individual botnets apply to total populations counts as well, and we attempt to expose the causes that lead to these inaccurate and conflicting size estimates.

Lastly, Dagon *et al.* [10] presented a taxonomy and analysis of potential botnet structures. In this paper, we sketch a technique for unveiling the existence of hidden clusters among botnets.

6 Summary

From a high-level perspective, this paper underscores the need for better clarity in studies related to botnet dynamics. Specifically, given the variety of botnet size estimation techniques and the diversity of results they provide, it seems only natural that botnet size should be a qualified term reflecting the context in which the resulting estimate should be interpreted.

That said, the results in this paper (and the questions they raise), should not be construed as an indication of our opinion on the prevalence of the botnet problem. Rather, our goal is simply to emphasize the fact that no single metric is sufficient for describing all aspects of a botnet's size. Moreover, given the variable temporal behavior that botnets exhibit and the inherent inaccuracies of existing estimation techniques, a prudent step towards providing more reliable size estimates is to synthesize the results from multiple concurrent and independent views of a botnet's behavior.

Finally, while we focus primarily on IRC botnets, many suggest that a migration to more sophisticated topologies and protocols (*e.g.,* P2P botnets [19]) is inevitable. If (or when) this transition occurs, the adoption of such technologies will pose substantial challenges to existing botnet tracking efforts, and brings its own set of difficulties.

Acknowledgments

This work is supported in part by National Science Foundation grant CNS-0627611. We extend our gratitude to the anonymous reviewers for their insightful comments and their help in improving this paper.

References

[1] Paul Baecher, Markus Koetter, Thorsten Holz, Maximillian Dornseif, and Felix Freiling. The Nepenthes Platform: An Efficient Approach to Collect Malware. In *Proceedings of the* 9^{th} *International Symposium on Recent Advances in Intrusion Detection (RAID)*, September 2006.

[2] Paul Barford and Vinod Yagneswaran. *An Inside Look at Botnets*. Advances in Information Security. Springer, 2007.

[3] Martin Casado and Michael Freedman. Peering through the shroud: The effect of edge opacity on IP-based client authentication. In *Proceedings of* 4^{th} *USENIX Symposium on Networked Systems Design and Implementation (NDSI)*, April 2007.

[4] Evan Cooke, Farnam Jahanian, and Danny McPherson. The Zombie Roundup: Understanding, Detecting, and Disturbing Botnets. In *Proceedings of the first Workshop on Steps to Reducing Unwanted Traffic on the Internet*, July 2005.

[5] David Dagon, Cliff Zou, and Wenke Lee. Modeling Botnet Propagation Using Time Zones. In *Proceedings of the* 13^{th} *Network and Distributed System Security Symposium NDSS*, February 2006.

[6] DNS Cache Snooping or Snooping the Cache for Fun and Profit. Available at: `http://www.sysvalue.com/papers/DNS-Cache-Snooping/files/DNS_Cache_Snooping_1.1.pdf`.

[7] Felix Freiling, Thorsten Holz, and Georg Wicherski. Botnet Tracking: Exploring a root-cause methodology to prevent denial-of-service attaks. In *Proceedings of* 10^{th} *European Symposium on Research in Computer Security, ESORICS*, September 2005.

[8] Allen Householder and Roman Danyliw. Increased Activity Targeting Windows Shares. Technical Report CA-2003-08, CERT, 2003.

[9] Farnam Jahanian. Enter the Botnet: An Introduction to the Post-Worm Era. ARO-DARPA-DHS Special Workshop on Botnets, 2006.

[10] Wenke Lee, Cliff Wang, and David Dagon. *Botnet Detection: Countering the Largest Security Threat*. Springer Verlag, July 2007.

[11] John Markoff. Attack of the Zombie Computers Is a Growing Threat, Experts Say. In *New York Times*, January 2007.

[12] Bill McCarty. Botnets: Big and bigger. *IEEE Security and Privacy Magazine*, 1(4):87–90, 2003.

[13] Laurianne McLaughlin. Bot software spreads, causes new worries. *IEEE Distributed Systems Online*, June 2004.

[14] Moheeb Abu Rajab, Jay Zarfoss, Fabian Monrose, and Andreas Terzis. A Multifaceted Approach to Understanding the Botnet Phenomenon. In *Proceedings of ACM SIGCOMM/USENIX Internet Measurement Conference (IMC)*, pages 41–52, Oct., 2006.

[15] Anirudh Ramachandran, Nick Feamster, and David Dagon. Revealing Botnet Membership using DNSBL Counter-Intelligence. In *Proceedings of the* 2^{nd} *Workshop on Steps to Reducing Unwanted Traffic on the Internet (SRUTI)*, July 2006.

[16] L. Spitzner. The Honeynet Project: Trapping the Hackers. *IEEE Security and Privacy Magazine*, 1(2):15–23, 2003.

[17] T. Sterling. Prosecutors say Dutch suspects hacked 1.5 million computers world wide. Associated Press, October 2005.

[18] CNN Technology. Expert: Botnets No. 1 emerging Internet threat. Online article, see `http://edition.cnn.com/2006/TECH/internet/01/31/furst/index.html`.

[19] Cliff C. Zou and Ryan Cunningham. Honeypot-Aware Advanced Botnet Construction and Maintenance. In *Proceedings of the 2006 International Conference on Dependable Systems and Networks (DSN)*, 2006.