How to Achieve Early Botnet Detection at the Provider Level?

Christian Dietz, Anna Sperotto, Gabi Dreo, Aiko Pras

How to Achieve Early Botnet Detection at the Provider Level?

Christian Dietz[1,2], Anna Sperotto[2], Gabi Dreo[1], and Aiko Pras[2]

[1] Universität der Bundeswehr München,
(christian.dietz, gabi.dreo)@unibw.de
[2] University of Twente,
(c.dietz, a.sperotto, a.pras)@utwente.nl

Abstract. Botnets are an enabler for many cyber-criminal activities and often responsible for DDoS attacks, banking fraud, cyber-espionage and extortion. Botnets are controlled by a botmaster that uses various advanced techniques to create, maintain and hide their complex and distributed C&C infrastructures. First, they use P2P techniques and domain fast-flux to increase the resilience against take-down actions. Second, botnets encrypt their communication payload to prevent signature based detection. Both, the actions to increase the resilience and the prevention of signature based detection are counteractions against detection techniques. In contrast to existing approaches, our novel approach includes DNS registration behaviour, which we currently analyse for the .com, .net and .org domains, representing half of registered domains on the Internet. Hence, the goal of this PhD research is to enable early detection of the deployment and operation of botnets to facilitate proactive mitigation strategies, whereas current approaches usually detect botnets while these are already in active use. Consequently, this proactive approach prevents botnets to fully evolve their size and attack power. Moreover, as many end users are unable to detect and clean infected machines, our approach tackles the botnet phenomenon without requiring any end user involvement, by incorporating ISPs and domain name registrars. In addition, this will enable the discovery of similar behaviour of different connected systems, which allows detection in cases where bots are registered under domains that are not willing to cooperate.

Keywords: Botnet, Early Detection, Provider Network, DNS, IP Flow Monitoring, Coordinated Cyber Threats, Domain Registration Behaviour

1 Introduction

Botnets are an enabler for many cyber-criminal activities and often responsible for DDoS attacks, banking fraud and cyber-espionage. As reported in [9,10] such criminal activities cause substantial economic damage. Recent estimations [15] expect that cyber attacks could cost global economy $3 trillion by 2020. Botmasters use various techniques to create, maintain and hide their complex C&C

infrastructures. First, they use P2P techniques [1] and domain fast-flux to increase the resilience against take-down actions. Second, botnets encrypt their communication payload to prevent signature based detection [12].

However, botnets often use the domain name system (DNS) [2,11,7], e.g., to find peers and register malicious domains. Since, botmasters manage a large distributed overlay network, but have limited personal resources, they tend to automate domain registration, e.g. using domain name generation algorithms (DGAs) [17]. Such automatically generated domains share similarities and possibly appear to be registered in close temporal distance. Such characteristics will be used for bot detection, while their deployment is still in preparation.

Hence, the goal of this PhD research is early detection of botnets to facilitate proactive mitigation strategies. Using such a proactive approach prevents botnets from evolving their full size and attack power. As many end users are unable to detect and clean infected machines, we favour a provider-based approach, involving ISPs and DNS registrars. This approach benefits from its overview of the network that allows to discover behavioural similarities of different connected systems. The benefit of tackling distributed large-scale attacks at provider level has been discussed in [13] and demonstrated by [4]. Further, initiatives to incentivize ISPs to mitigate botnets are already ongoing [8]. In addition, several studies discuss and high-light the role of ISPs in detection and mitigation of various cyber threats, e.g. DDoS, Botnets or SPAM [14,3,16].

The work done in [6] addresses the domain registration behaviour of spammers and [5] demonstrated DGA based malware detection by using flow-based techniques. In contrast, our approach includes the detection of malicious DNS registration behaviour, which we currently analyse for the .com, .net and .org domains. These domains represent half of the registered Internet domains. By combining DNS registration behaviour analysis with passive monitoring of DNS requests and IP flows, we are able to tackle botnets throughout their whole life-cycle. This research is still in its initial state and will result in a PhD thesis.

The remaining parts are structured as follows. Section 2, describes the research problem and questions. Section 3, describes our approach. Next, Section 4 provides early results and the current state of research. Finally, the paper is concluded in Section 5.

2 Research Problem & Questions

The goal of this research is to enable early botnet detection in provider environments. To achieve this goal, our approach is based on large-scale DNS registration behaviour analysis, as this will allow to discover botnet activity in the (pre-)deployment phase of its life-cycle (see Fig. 1). Thus, our novel approach could possibly prevent the botnet from becoming deployed and actively used. Furthermore, the proposed approach takes into account the dynamics of botnet malware and the Internet infrastructure, high data rates, incompleteness of data and encrypted bot communication. In order to tackle the early botnet detection problem, we ask the following questions:

RQ 1: How do botnets interact with the domain name system?
RQ 2: Can domain registration characteristics be used for botnet detection, and if yes, how?
RQ 3: (How) Does early detection work, if some registrars do not cooperate?

The approach used for answering these research questions will be described in the next Section. Fig. 1 shows the bot life-cycle and relating research questions.

3 Approach

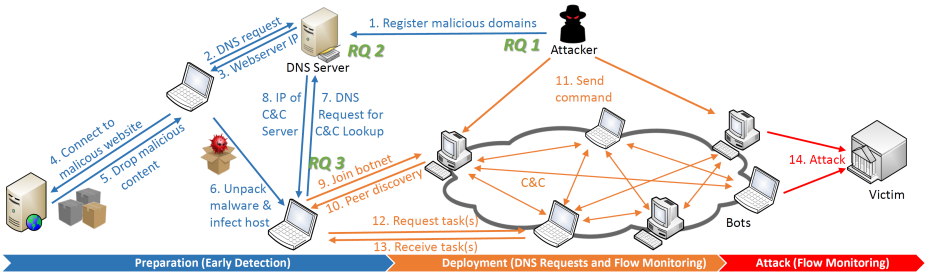

Fig. 1. Detailed overview of botnet operations with mapped research questions.

The goal of this research is to allow faster botnet detection and mitigation. Current approaches are usually limited to detect bots after they already became active or while they are used in attacks. Our approach targets botnet detection in the pre-deployment phase. Therefore, our approach is based on two components: (1) passive monitoring of communication characteristics and (2) DNS registration behaviour analysis. DNS registration analysis allows to detect the preparatory actions of deployment of the C&C infrastructure and the bots. Therefore, our approach allows botnet early detection and consequently facilitates proactive botnet mitigation. In addition, our approach allows botnet detection in the subsequent phases of the bot life-cycle (preparation, infection, peer discovery, malware update, command propagation and attack) by using passive DNS and flow monitoring solutions. This is important, since bots might also be registered at domain providers that are not sharing data.

Research question *1* aims to get insight into the deployment and management of botnets. Therefore, we collect DNS registration data on a daily basis for the *.com*, *.net* and *.net* domains, representing half of the domains registered on the Internet. Second, we query different botnet tracking services and use DGAs to find botnet related records in the domain registration dataset.

Research question *2* aims to extract characteristics of botnets in their deployment phase. Which might allow an early detection and mitigation. To answer this question, we use registration databases of top level domain registrars. Currently, our study involves the *.com, .net, and .org* top level domains.

Research question *3* extends our novel approach to make it applicable in case bots are registered under domains that do not share data. In such cases, our approach might derive flow-based behaviour characteristics based on the knowledge gained in RQ1 and RQ2 for flow-based detection of bots. Flow monitoring solutions provide an overview of large parts of the Internet, in which we expect to find similarities that can be used for detection of bot behaviour.

We will validate our novel approach based on simulations and real-live environments. Further, we compile different datasets. First, we crawl the registration database of multiple top level domains, different botnet domain and IP blocklists with time stamps. This allows us to measure the temporal difference between botnet deployment and detection. Second, we passively capture IP flow data and DNS requests in multiple provider networks to evaluate (a) how accurate our approach can detect the large-scale similarities between distributed bots and (b) determine the temporal delay between malicious domain registration and the first activity. This evaluation also uses IP and DNS blocklists.

4 Early Results

In a first step, we used data captured from Kelihos sinkholing operation, that allowed us to observe real bots in two different states of their life-cycle, peer discovery and job requests. We successfully used our insights gained to developed a concept for flow-based detection. Further, we use multiple DGAs and C&C domain lists to extract the botnet domains (e.g., Zeus, Kelihos.B, Palevo, Drye). Early results show that botnet domains are registered in close temporal distance (bulk registration) and often have structural similarities. Thus, we assume that our approach will be able to accurately detect malicious DNS registration activities and host behaviour of bots.

5 Final Considerations

When provider based solutions are used for bot detection, it is important that data should be accurate and be derived characteristics should be independent of the capture infrastructure. However, as botnets are globally spread, usually one provider can only detect a fraction of a botnet. Therefore, the detection system should run and cooperate across multiple provider networks, by means of providing infrastructure independent detection information and being able to use such data from different networks. ISPs often apply sampling to their flow monitoring to reduce memory consumption, which might be an additional challenge to our approach. Moreover, anti-detection techniques of malware become more sophisticated and often involve encryption and anonymisation techniques. Our approach will be resistant against many of these techniques, due to its high-level overview and independence of packet payloads. The main goal of this approach, should be achieved within a period of four years as part of a PhD thesis.

Acknowledgments This work is partially funded by EU FP7 Flamingo Network of Excellence Project (ICT-318488).

References

1. Andriesse, D., Rossow, C., Stone-Gross, B., Plohmann, D., Bos, H.: Highly resilient peer-to-peer botnets are here: An Analysis of Gameover Zeus. In: 8th IEEE International Conference on Malicious and Unwanted Software (MALWARE) (2013)
2. Antonakakis, M., Perdisci, R., Nadji, Y., Vasiloglou II, N., Abu-Nimeh, S., Lee, W., Dagon, D.: From throw-away traffic to bots: Detecting the rise of DGA-based malware. In: USENIX Security Symposium (2012)
3. Asghari, H., van Eeten, M.J., Bauer, J.M.: Economics of Fighting Botnets: Lessons from a Decade of Mitigation. IEEE Security & Privacy (2015)
4. François, J., Aib, I., Boutaba, R.: FireCol: A Collaborative Protection Network for the Detection of Flooding DDoS Attacks. IEEE/ACM Transactions on Networking (TON) 20(6) (2012)
5. Grill, M., Nikolaev, I., Valeros, V., Rehak, M.: Detecting DGA Malware using NetFlow. In: IFIP/IEEE International Symposium on Integrated Network Management (IM) (2015)
6. Hao, S., Thomas, M., Paxson, V., Feamster, N., Kreibich, C., Grier, C., Hollenbeck, S.: Understanding the Domain Registration Behavior of Spammers. In: Proceedings of the 2013 conference on Internet measurement. ACM (2013)
7. Kwon, J., Lee, J., Lee, H., Perrig, A.: PsyBoG: A Scalable Botnet Detection Method for Large-scale DNS Traffic. Computer Networks (2016)
8. Lone, Q., Moura, G., Van Eeten, M.: Towards Incentivizing ISPs to Mitigate Botnets. In: Monitoring and Securing Virtualized Networks and Services. Springer (2014)
9. McAfee: The Economic Impact of Cyber-crime, `http://www.mcafee.com/mx/resources/reports/rp-economic-impact-cybercrime.pdf`, visited on 05.01.2016
10. Mossburg, E.: A Deeper Look at the Financial Impact of Cyber Attacks, `http://daily.financialexecutives.org/a-deeper-look-at-the-financial-impact-of-cyber-attacks`, visited on 05.01.2016
11. Nguyen, T.D., CAO, T.D., Nguyen, L.G.: DGA Botnet Detection using Collaborative Filtering and Density-based Clustering. In: Proceedings of the Sixth International Symposium on Information and Communication Technology. ACM (2015)
12. Rossow, C., Dietrich, C.J.: Provex: Detecting Botnets with Encrypted Command and Control Channels. In: Detection of Intrusions and Malware, and Vulnerability Assessment. Springer (2013)
13. Steinberger, J., Schehlmann, L., Abt, S., Baier, H.: Anomaly Detection and mitigation at Internet scale: A survey. In: Emerging Management Mechanisms for the Future Internet. Springer (2013)
14. Steinberger, J., Sperotto, A., Baier, H., Pras, A.: Collaborative Attack Mitigation and Response: A Survey. In: IFIP/IEEE International Symposium on Integrated Network Management (IM) (2015)
15. Taylor, B.: Cyber Attacks Fallout Could Cost the Global Economy 3 Trillion Dollar by 2020, `http://www.techrepublic.com/article/cyberattacks-fallout-could-cost-the-global-economy-3-trillion-by-2020/`, visited on 05.01.2016
16. Van Eeten, M., Bauer, J.M., Asghari, H., Tabatabaie, S., Rand, D.: The role of Internet Service Providers in Botnet Mitigation an Empirical Analysis based on Spam Data. TPRC (2010)

17. Yadav, S., Reddy, A.K.K., Ranjan, S., et al.: Detecting Algorithmically Generated Domain-flux Attacks with DNS Traffic Analysis. IEEE/ACM Transactions on Networking 20(5) (2012)