A Generic Process Model for Botnet Forensic Analysis

Meenakshi Thapliyal
Department of Computer Science and Engineering, Graphic Era University
Dehradun, Uttarakhand, India
itsmemeenu123@gmail.com

Anchit Bijalwan
Department of Computer Science and Engineering, Uttarakhand Technical University
Dehradun, Uttarakhand, India
anchit_bijalwan2000@yahoo.com

Neha Garg
Department of Computer Science and Engineering, Graphic Era University
Dehradun, Uttarakhand, India
nehagarg.february@gmail.com

Emmanuel Shubhakar Pilli
Department of Computer Science and Engineering, Graphic Era University
Dehradun, Uttarakhand, India
emmshub@gmail.com

Abstract

Botnets are becoming more hazardous in cyber crime when compared to other malicious activities. Security against botnets is a major concern. Botnet forensics is young science which can answer questions about how, what and where of damage done by bots. The forensic system deals with capturing, recording, and analysis of botnet traffic. This paper outlines the process of Botnet forensic analysis and its implementation. A generic process for botnet forensics is proposed based on previous digital forensics models. The specific research gaps existing in implementation are identified and presented as challenges. The contribution of this work is that it presents an overview on botnet forensics analysis and implementation which will be more valuable for security.

Keywords: Bot, Botnet, Analysis, Forensic.

1. Introduction

Botnet is an army of infected computers that take instructions from a botherder. A botherder is corrupt hacker who uses the botnet for financial gain or as a destructive behavior of civilization and the Internet community without ethics. The bad guys have been using the latest killer web application, is the advanced security Web technology. Many of the security professionals who pioneered the fight against botnets are demoralized by the realization that taking out the C&C [1] does not helps. Botnets a call to Action (Command and Control) server is less effective now days. Botnet infection is adaptive means for chain to chain system. Single virus/worm spread the infection another module through malicious code that prevents previous antivirus action [2].

Social networking sites like orkut, facebook, skype and Google blogger were infected by distributed denial of service attacks, spam etc. Computer forensics was introduced by law enforcement with proper guidelines of judicial system [3]. Computer forensics involves protection, detection, mining, records, and analysis of computer data. Botnet forensics analysis is a natural extension of computer forensics. Botnet forensics analysis is a permanent monitoring process that deals with capturing, recording, and analysis of botnet traffic in packets format. It can also indicate alerts when thresholds (maximum limit) are exceeded. If the present attack could not be prevented, the fundamental information is used to defend against similar attack in future event. Botnet forensics can be used to evaluate how and where the attack occurred, who was culprit, duration of the exploit, and the line of attack. Botnet forensics can be used as a device for monitoring the

© 2013. The authors - Published by Atlantis Press

activity of botmaster and hacker, business transaction analysis and investigation of irregular performance issues sources. Botnet forensics involves postmortem research of the attack means notitia criminis (after crime announcement) which takes particular duration of action and way of tackle for particular cases. Botnet forensics and the attacker both are at the identical proficiency altitude. The hacker uses a set of tools to launch the attack and the botnet forensic authority applies similar tools to explore the attack. The hacker has all the time at his disposal and will regularly get better his skills, provoked by the millions of dollars in resolving risk .
The paper is structured as follows: in Section 2 introduces the botnet forensic methodology and motivation, Section 3 brief summary of previous botnet forensic analysis. We propose a generic process model for botnet forensics analysis. Key Challenges are presented in Section 4, conclusion is given in Section 5.

2. Background

Internet Relay Chat (IRC) is a text-based chat-system that communicates bot in channels. They were capable to interpret simple commands, provide administration support, suggest simple games and regain information about operating systems, logins, email addresses, aliases, etc [4]. Denial of service, then distributed denial of service attacks were implemented in these IRC bots. They could proliferate like worms, remain hidden as viruses and could initiate huge, corresponding attacks like AgoBot and SDBot. The recent generation of bots can multiply through file-sharing networks, peer-to-peer networks, email attachments and infected websites. Bots communication can be accomplished by several protocols, such as IRC, HTTP and P2P. Latest sophisticated botnet is CoolBot, which could explain the problems that how to recover shut down and accept delay problem of C&C and contradict against insecurity (routing table poisoning). CoolBot could control C&C model automatically and repair the broken C&C, which means C&C reconstruction [5].

2.1. *Botnet forensic methodology*

Botnet forensic methodology consists of three steps:

- Malware collection

A 'Catch-it-as-you-can' system follows where all bots packets passing through an exacting traffic point are captured. Analysis is consequently done which require huge storage. A piece of bots packet is individually analyzed and stored in definite memory for future analysis [6].

- Malware analysis

The botnet forensic system is a security device with hardware and pre-installed software. It is defined behavior based analysis.

- Botnet tracking

The extracted information from malware collection interferes into the control channel of the botnet network which tracks that what kinds of attack has occurred.

2.2. Motivation for botnet forensics Analysis

The defensive approaches of botnet forensics follows that prevention is better than attack, detection and response perspectives. Botnet forensics ensures that attacker spends more time and energy to cover his tracks for making the attack. Hackers/ criminals will be more careful to avoid prosecution for their prohibited events. This restriction reduces network crime rate and thus security improves. Internet Service Providers (ISPs) are also being made answerable for what's going on their network. Now a days companies doing business on Internet can not hide a security breach and are expected to prove the state of their security as a compliance measure for regulatory purposes.

3. Botnet Forensic Analysis

Various advanced botnet forensic concepts, designs and analysis approaches were planned to handle the botnet network environments. In recent years, a behavior-based bots detection tool has been developing fast, which gives a serious malware drive.

3.1. *IRC traffic analysis*

Mazzariello et al. [7] focused that how an mass of probably strong hosts can be control from being infected. IRC user behavior model organized in a channel to make difference between normal and botnet-related activity. They will concentrate on the problem of detecting botnets, by introducing network traffic analysis architecture, and describing a behavioral model, for a specific class of network users, able to identify botnet-related activities. Time analysis patterns of botnet activity and taxonomy both phenomenons are established through Botnet traffic analysis. This IRC traffic analysis is a versatile approach which based on systematic monitoring, for botnets detection which based on the network behavior. Initially analysis has been supervised then next testing of pure invalid anomaly was detected. Kugisaki et al. [8] studied on IRC behavior which directly related to an IRC server. This novel approach can be used to make judgments between of present bots more than existing viruses by computerization. More objects (problem of verification

and generality of objecs), duration of detection by real visible and Measures against bots that does not use IRC.

3.2. *Asprox botnet analysis*

Borgaonkar et al. [9] studied the design and structure of the Asprox botnet, the communication protocols which drive-by downloading of spreading malicious substance and the advanced fast-flux service network. The main features of the Asprox botnet are the use of centralized command and control structure(C & C structure), HTTP based communication, advanced double fast-flux service networks, SQL injection attacks for recruiting new bots and social engineering tricks to spread malware binaries. Hydra fast-flux network, SQL injection attack tool is advanced features of Asprox**.** Asprox botnet does not suitable strong for Cryptography. In the botnet architecture, authenticity and integrity of the bot commands is important.

3.3. *Cross-analysis of botnet victims*

Balzarotti et al. [10] provide an in depth analysis passive (depend on human action or other) and active(bot frequently use network scanning techniques to get susceptible hosts for spreading infection) measurement study that how data get infected through major botnets like Conficker, MegaD, and Srizbi. They observed commonly-malware infected networks. IP address space and physical location of Conficker botnet can be observed by CAIDA. A Cross-botnet prediction technique is proposed to predict unknown victims of one botnet from the information of the other botnet if they have similar infection vectors. They will further provide new approaches to explain relationships between geopolitical locations and malware infection.

3.4. *Financial botnet analysis*

Financial botnets particularly aimed at carrying out financial fraud, popular threat for banking institutions. This is authenticated by one of the biggest savings banks in Spain which helps to fight against financial cybercrime. A financial botnet is identifying, analyzing, and mitigating through Financial Botnet Analysis. Banking Trojan is a representative malware which evaluated in financial institution. This analysis automatically emphasized on the detection, imagination and sharing intelligence about financial botnets [11].

3.5. *Graph-based analysis*

Nagaraja et al. [12] proposed Graph-based analysis, a powerful approach because it lacks protocol semantics or packet statistics dependence. It depends on being able to accurately model valid network growth. A botnet can be detected based on the observation that an attacker will increase the number of related graph components due to a rapid growth of edges between suspect neighboring nodes. BotGrep is a graph theory based approach to botnet detection and analysis. Nodes of communication graph represent Internet hosts and edges represent communication between them.

4. Generic Process model

We use a diagrammatic approach (Fig.1) for Botnet forensic analysis which is based on previous accessible forensics analysis.

4.1. *Preparation of security tool*

Botnet forensics is relevant only to environments where bot security tools (sensors) like intrusion detection systems, packet analyzers, firewalls, traffic flow measurement software are installed. The required authorizations and legal warrants are obtained so that privacy is not dishonored. The preparation phase ensures the monitoring tools are in well place.

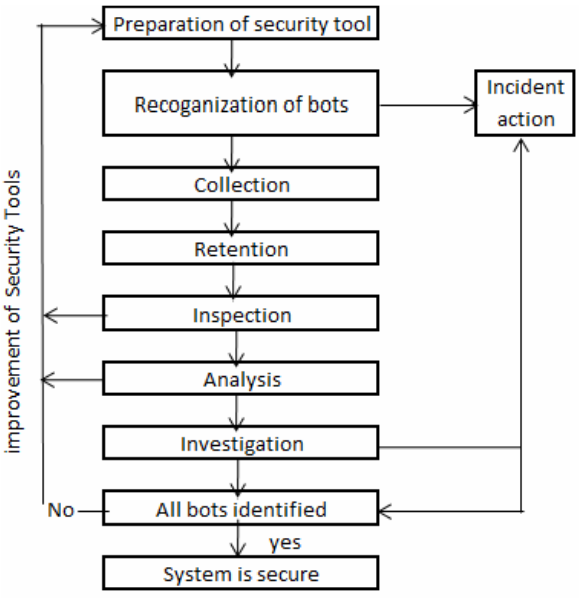

Fig 1. Generic Process model of Botnet Forensics Analysis

4.2. *Recoganizations of bots*

Various security tools produced a specific security infringe to indicate alert during unauthorized actions and inconsistency. Different parameters are used for the determination of presence and nature of the attack. An immediate validation is done to assess and confirm the suspected attack and ignore the false alert. TCPDump, Wireshark, PADS, Nepenthes, Snort etc. devices are used to confirm accurate alarm. Alert and collection of responses during attack is accomplished, bots is identified.

4.3. *Incident Response*

Crime response is convenient on the collected information which validates and assesses the event by organization strategy, authorized and business constraint. Securities against future attacks and recovers from the existing damage, preplanning is initiated. At the same time, the decision whether to continue the investigation and gather more information is also taken. A similar response is to be initiated after the investigation phase where the information obtained may requires certain actions to control and reduce the attack.

4.4. *Collection*

The traffic data is assembled from the botnet sensors and most of the facts causing minimum contact to the infected machine. Traffic data rapidly change and it is difficult to create the same trace later. TCPDump, Wireshark, Snort tools assist in collections of traffic data.

4.5. *Retention*

Traces of data and logs are stored on a backup device. The original security traffic data are unaffected for legal requirements. Single copy of the data will be analyzed. TCPDump, Wireshark, Snort tools is applied for retention phases.

4.6. *Inspection*

The whole traces data and specific evidences of the attack are composed of combined data format which can be analyzed. TCPDump, Wireshark, Flow-tools, NfDump, Bro, Snort tools provide a proper way to check each and every facts of the attack which improves the security tools.

4.7. *Analysis*

Previous botnet attack's pattern helps to classify the particular infection sign and reformed to understand the intention and methodology of the attacker and are classified and correlated. The data searching and matching attack patterns can be done by Statistical, soft computing and data mining. The tools which support in analysis of botnet attack are TCPDump, Wireshark, TCPFlow, TCPTrace, Olly Dbg, IDA Pro, NetFlow, TCPXtract, Snort etc.

4.8. *Investigation*

Botnet forensics is targeted to define communication pathway between an infected machine and back to the origin point of attack. Incident response and prosecution of the attacker is used to identify the attacker, the tough measurement of the botnet. IP spoofing and stepping stone attack is still prevalent technique of the attacker to hide himself.

4.9. *Results (Max. Possibility of bot's identified)*

The legal systematic format is designed to arrive at the conclusion. The botnet forensic analysis provides visual conclusion of the attacker methodology and feedback for future investigations to conduct the deployment and improvement of security products.

This generic process model of botnet forensics analysis is used in both real-time and post attack scenarios [13]. The first five phases (including incident action) handle concurrent botnet traffic. The next four phases are common for real time and post attack scenarios. The post attack investigation begins at the inspection phase, where a copy of the packet capture (lib pcap format) file is given for investigation. The inspection phase fuses inputs from various sources and identifies attack indicators. The analysis phase classifies attack patterns using data mining and statistical approaches. The investigation phase involves trace back and attribution. Finally bot is captured, if results are not satisfactory for further improvement of the security tools.

5. Research Challenges

5.1 *Collection and Detection*

The main problem is that different kinds of bots having their different characterization makes it very tough to detect these bots, to define the logical relationship between different bots, arises the new challenges to the accuracy of population counting techniques. Botnet Overlapping between botnet counting and detection creates potential hidden relationships among botnets.

5.2. *Botnet Size*

Each and every minute new advanced bots (TDL-4, Grum) are generated in the internet field and its very tough to guess the kind of next bot's attack and how powerful it will be. Botnet forensic analysis involves capturing of bots and tells, what is the position of attack with its maximum possible botnet size. Waledac Botnet size estimation is difficult because botnet itself was capable of sending about 1.5 billion spam messages a day, or about 1% of the total global spam volume. It will be difficult to operate on a new botnet especially on the encrypted botnet.

5.3. *Botnet traffic filter*

It does not contribute to any information between Failover pairs. Failovers or Reboots requires re-download of the Dynamic Database. Currently there is no support for IPV6.

5.4. *Investigation*

It is biggest challenge to investigate a robust botnet unwanted traffic detection algorithms and how to filter botnet command and control (C&C) traffic early.

5.5. *Analysis*

Active analysis is required because honeypot setup is difficult on large scale network. Passive anomaly analysis usually independent of the traffic content and has the potential to find different types of botnets (e.g., HTTP, IRC and P2P).

5.6. *Temporal correlation technique*

This technique utilize between DNS queries and entropy based correlation between domain names, for speedier detection. It is difficult to applied a more system level logs such as Process/service executions, memory/CPU utilization, disk reads/writes. It is a biggest challenges server failure based DNS failures, or failures related to the name servers, as a means for detecting botnets which exhibit double fast flux.

6. Conclusion

Botnet forensics analysis is helps to capture, detect and trace particular bot among piles of bots. It also defines the possible complicated relationship between different kind of bots and tracing the source of attack with hackers. It provides a systematic pathway to predict future attacks by using intrusion data of previous botnet attack strategy.

References

1. M. Bailey, E. Cooke, F. Jahanian, X. Yunjing, and M. Karir, "A Survey of Botnet Technology and Defenses," in *Conf. For Homeland Security, (CATCH '09. Cybersecurity Applications & Technology*, 2009) pp. 299-304.

2. J. Govil, "Examining the criminology of bot zoo," in *2007 6th Int. Conf. on Information,(Communications & Signal Processing*, 2007) pp. 1-6.

3. E. S. Pilli, R. C. Joshi, and R. Niyogi, "Network forensic frameworks: Survey and research challenges,"(*Digital Investigation,* 2010) vol. 7, pp. 14-27.

4. S. r. S. C. Silva, R. M. P. Silva, R. C. G. Pinto, and R. M. Salles, "Botnets: A survey," *Computer Networks,* 2012.

5. L. Chaoge, L. Weiqing, Z. Zhiqi, L. Peng, and C. Xiang, "A recoverable hybrid C&C botnet," in *Malicious and Unwanted Software(MALWARE), 6th Int. Conf. on*, 2011, pp. 110-118.

6. M. A. R. J. Z. Fabian and M. A. Terzis, "A multifaceted approach to understanding the botnet phenomenon," in *Proc. of the ACM SIGCOMM Internet Measurement Conference (IMC)*, 2006.

7. C. Mazzariello, "IRC traffic analysis for botnet detection," in *Information Assurance and Security, ISIAS'08. Fourth Int. Conf. on*, 2008, pp. 318-323.

8. Y. Kugisaki, Y. Kasahara, Y. Hori, and K. Sakurai, "Bot Detection Based on Traffic Analysis," in *Intelligent Pervasive Computing, IPC, Int Conf. on*, 2007, pp. 303-306.

9. R. Borgaonkar, "An Analysis of the Asprox Botnet," in *Emerging Security Information Systems and Technologies (SECURWARE), Fourth Int. Conf*, 2010, pp. 148-153.

10. R. Sommer, D. Balzarotti, G. Maier, S. Shin, R. Lin, and G. Gu, "Cross-Analysis of Botnet Victims: New Insights and Implications," in *Recent Advances in Intrusion Detection*, vol. 6961(Springer Berlin Heidelberg, 2011) pp. 242-261.

11. M. Riccardi, D. Oro, J. Luna, M. Cremonini, and M. Vilanova, "A framework for financial botnet analysis," in *Crime Researchers Summit (eCrime),* 2010, pp. 1-7.

12. S. Nagaraja, P. Mittal, C.-Y. Hong, M. Caesar, and N. Borisov, "BotGrep: finding P2P bots with structured graph analysis," in *Proc. of the 19th USENIX conf. on Security*, 2010, pp. 7-7.

13. S. Krasser, G. Conti, J. Grizzard, J. Gribschaw, and H. Owen, "Real-time and forensic network data analysis using animated and coordinated visualization," in *Information Assurance Workshop, IAW '05. Proc. from the Sixth Annual IEEE SMC*, 2005, pp. 42-49.