On Botnet Detection in Networks, based on Traffic Monitoring

[1]Shamsul Haq, [2]Yashwant Singh,
[1]M.Tech. Student, [2]Associate Professor
Department of Computer Science and Information Technology, Central University of Jammu, Jammu and Kashmir, 181143
[1]s.haq266@gmail.com, [2]yash22222k1@gmail.com

ABSTRACT

One of the serious and widespread attacks in cyber security is Botnet. Using command and control infrastructure or peer-to-peer communication between bots, botmasters can perform a variety of attacks on internet system-users. To mitigate this, multiple techniques have been developed for botnet detection over the past two decades. In this paper we have discussed various botnet structures and the different techniques of botnet detection proposed in literature. We evaluated these techniques based on their distinctive features and presented their detailed comparative analysis. We also proposed a method for botnet detection using network traffic monitoring. Our approach is based on combining signature and anomaly detection systems that complement each other. Our proposed hybrid detection system may decrease false positive rate in anomaly detection by finding the well-known bots using signature detection and thereby may increase overall detection efficiency.
Keywords: Botnet; malicious activities; P2P; anomaly detection

INTRODUCTION

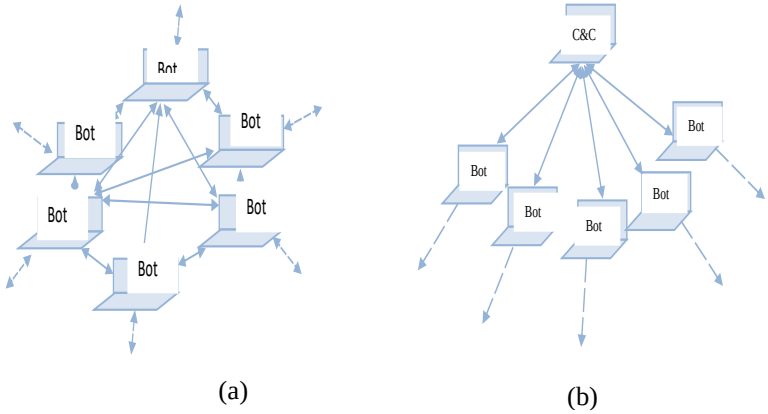
A bot is a malicious program that runs and works as an operative for hackers in stealth to computer users. Hackers installing bots on user systems are known as bot masters or bot headers. Bot masters work remotely and forward bots to the vulnerable computers to turn them into their slave zombies (infected nodes). These zombies can be PCs, mobile devices, servers, internet of things that are infected with bots. The bots then wait for the commands of master, so to perform any type of service/attack without users' knowledge. A hacker can infect multiple interconnected computers with bots which interact to form Botnets. Botnets can perform any type of attack on the Internet like a distributed denial of service (DDoS), spamming mail, phishing and key-logging attack for the Bot master etc.. They can also be used for malicious activities like Trojans, viruses and worm implementations as well as for identity theft such as to steal personal data of the users (e.g. bank details). Malicious use of botnets include secondary local infection, trading of bandwidth or computational resources to run some distributed computing tasks for the master, back doors and host illegal trade. These impending threats makes Botnets an major challenge to cyber security.

LITERATURE SURVEY

A. Botnet Structures

A key aspect to a botnet's viability is the underlying communication architecture it employs. From this standpoint, Botnets are usually of two typical structures: centralized and decentralized Botnets. Centralized Botnets traditionally worked by connecting to an Internet Relay Chat (IRC) server under the command and control (C&C) of Bot master. However, such architectures are prone to disruption if the servers are shut down by law enforcement agencies. Furthermore, network traffic monitoring of such Botnets is easy and well-studied in literature. Over the recent years, command and control structures have switched to the use of HTTP protocol for server-client messaging. HTTP based Botnets utilize data packet encryption and control of authenticated communication channels to circumvent a systems' Firewall security. Botmasters can also change the IP addresses and domain names of the control and command servers to obstruct any closure attempts by the law enforcement.

To avoid the weaknesses of centralized structures, Decentralized Botnets have evolved recently and rely on peer-to-peer (P2P) communication networking. Each bot can work as a server or a client as per the botmaster's needs to transfer master's commands and updates throughout the network. Such networks are robust to breakdown as the network can continue to operate even if a few bots are down. However, P2P botnets do suffer from delays in message communication across the network as well as resulting synchronization problems. Figure 1 shows representations of the two types of botnet structures[1][2][3].



(a) (b)

Figure - 1. (a) Centralized Botnet and (b) Decentralized Botnet (p2p) [2]

B. Bot Attacks

To make attacks successful, botmasters focus on a) a variety of means to install bots on victim's computers, b) methods to create, update and control large botnets and c) ways are stay stealth on a victims' machine. While traditionally, hackers have been writing the bot programs themselves, novice hackers can use existing available programs or some modifications of the existing code. Hackers often avoid using use their own computers to forward bots. Instead they rely on using already infected machines, which act as proxy servers for further attacks. These proxy servers help protect the hackers' identities from the security investigators [4].

C. Botnet Detection Techniques

As most of attackers concentrate on command and control protocols (e.g. HTTP, IRC) using centralized structure, research literature has mostly focused on the detection techniques for better detection of centralized Botnet structures. botmasters have therefore more recently focused on using P2P architectures for illegal attacks. Due to these un-centralized structures, most of the current detection techniques have been rendered ineffective [5][6][7][8]. In this section we give an overview of the different detection techniques and their classification based on various features (see Figure 2).

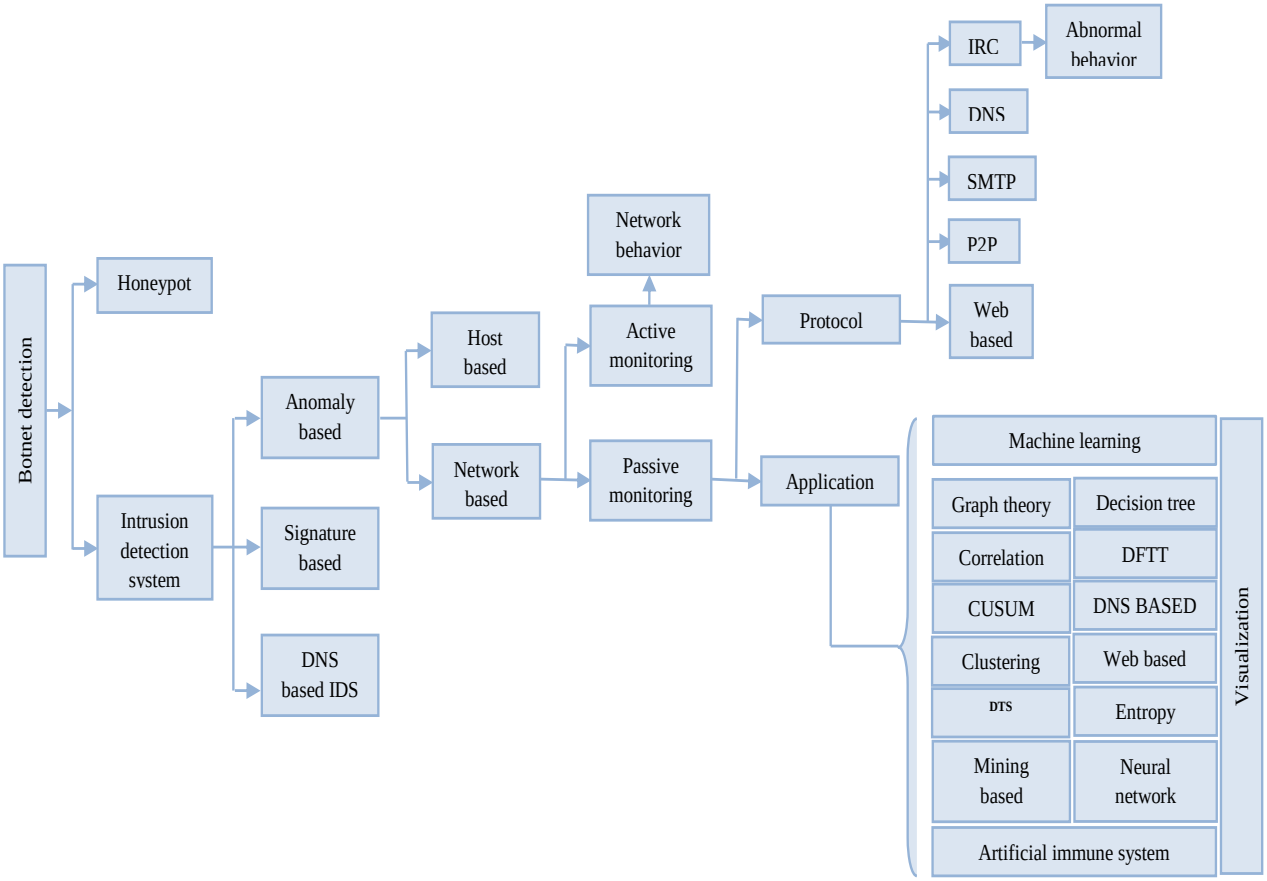


Figure - 2. Botnet detection techniques [7]

Broadly, detection techniques can be divided into two camps: Honeypots and Intrusion-Detection Systems (IDS) as discussed below. These detection mechanisms can focus on either C&C detection, bot detection or Bot-master detection. Detection mechanisms employing traffic monitoring can work in active or passive modes [9]. Active monitors send enquires into the network to acquire data on packet flow and bandwidth functioning etc. This is not a preferable method as packets interfere with ongoing traffic data. Passive monitors employ different measures such as packet inspection, analysis of flow records etc., to directly observe traffic flow for any suspicious bot activity.

(1) Honeypot – It is generally a trap that diverts the attention of attackers to attack a bait computer system and avoid attention to the main system, which needs to be prevented from

illegal activities. This detection technique is effective for tracking Botnets. These bait systems can help with providing full examination of Bot behaviors. Various such useful data that honeypots gather include information about Bots C&C mechanisms, the motivations of the attacker, signatures of bots, the unknown security holes in the network that enable bots to penetrate in [10].

(2) Intrusion Detection System – An IDS may be a software application or hardware machine that detects and reports any illegal or unauthorized bot activity on computer systems in a network. There are several types of such detection systems: for example, signature based, anomaly based and DNS based intrusion detection systems. Signature based detection compares data packets with predefined rules or patterns (known as signatures) extracted from previous intrusion activity. One of the examples used as signature based detection is SNORT [11]. Although this form of detection has high response time for known attacks, there are minimum chances of detecting unknown, zero-day botnet attacks and signature databases need to be updated frequently.

Anomaly based detection is based on analyzing several networks traffic irregularities from an established baseline behavior in a specific network environment. Irregularities may include for example, traffic passing through unusual ports, increased traffic volume etc. Anomaly based detection can be host-based and network-based according to the baseline profile and the location where the incoming traffic is analyzed. Network based detection can operate in active or passive mode for Botnet detection. Although such systems can detect zero-day attacks, they often report false positives at a high rate. These systems are slow to work when placed in new environments, as they need to customize the baseline to a different network traffic. An added limitation in cases where active monitoring is employed, is that additional packets introduced in the network increases traffic payload [8].

Another type of botnet detection technique is based on Domain Name System (DNS) detection. Bots use DNS to communicate with C&C servers to receive commands and to pass harvested information. To establish a malicious network and communicate instructions to bots, Botnet creators or hackers generally use fixed IP addresses through code level establishment within the malicious software which can be tracked. However, with the use of domain names that are generated via DGA algorithms, attackers can change IP addresses of C&C servers with no need to modify bot codes. Consequently, bots perform DNS queries to connect to C&C servers or to download its update from the botnet server. These DNS queries that are sent simultaneously by distributed bots can be detected for example by DNS anomaly detection. Recent works use data mining techniques as key approaches for DNS detection [12].

D .Traffic Monitoring

In case of monitoring network traffics, the bot detection may include various phases, which are network traffics, filtering, application classifier (IRC and HTTP based centralized part and P2P based classification), malicious activity detector, traffic monitoring and report[7]. This phased process is useful to identify the typical communication and malicious activity patterns within the same botnet. The process go as follows. Network flows are captured, and special information is recorded on each flow. Information that is generally recorded includes protocol, duration, source port, destination port, source IP addresses, destination IP addresses, number of packets and number of bytes transferred in both direction. Following this, the average number of bytes per

packet (number of bytes/number of packets) and average number of bytes per second (number of bytes/duration) are calculated. The similarities among databases can be drawn or represented in different ways. For example, one way to represent is draw a graph in x-y axis. The x-axis can be considered as number of bytes per packet and y-axis as the number of bytes per second. The curves that are formed by joining the points are clustered into the same category of similar characteristics of network flow.

OUR PROPOSAL OF BOTNET DETECTION

The proposed model for traffic monitoring consists of network traffic collection, filtering, application classifier, Hybrid detection (signature and anomaly detection) as shown in figure 3.

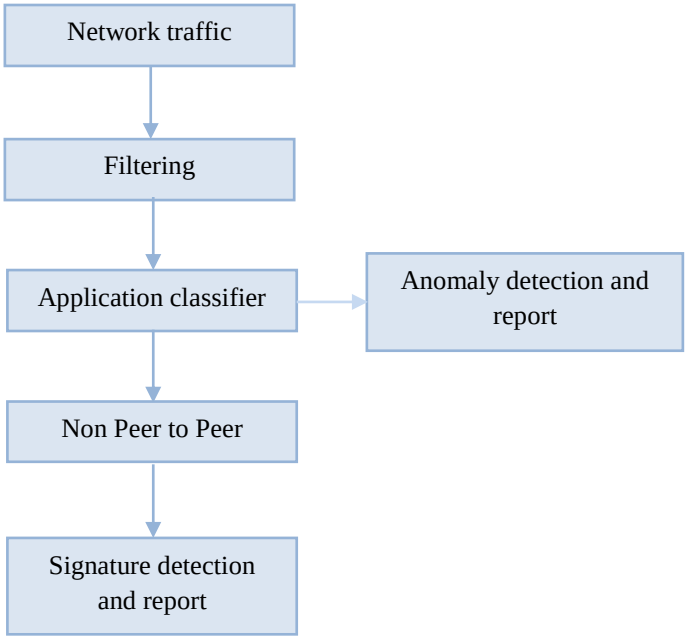


Figure - 3. Proposed model for Botnet detection

There may be similar communication pattern traffics and malicious activity of hosts. The filtering is responsible for minimizing the traffic by filtering in anomalous traffics. Application classifier separates the traffic of IRC and HTTP based activity from the rest of traffics, i.e. the separation p2p and non-p2p traffics. Hybrid detection consists of detecting sequentially or parallelly with signature-based (for known signatures) and anomaly detection (based on k-means clustering or self-organizing map) for the remaining traffic.

COMPARATIVE ANALYSIS OF DETECTION TECHNIQUES

Most of the times the statistical protocol features for detection methods are not a poor practice. However, it is rigid to depend on statistically generated features too much. Table 1 helps to realize the design of proposals and how to avert being detected by such methods. Table 2 shows

the Botnet detection comparison based on independence of signature as well as protocols, approaches considering encrypted C&C Botnets, Detection in real-time, and accuracy.

PAPERS	IRC	P2P	DNS	UDP	TCP	SMTP	HTTP
FLUXOR [13]	-	-	-	-	-	-	-
BOTHUNTER[14]	✓	-	-	-	-	-	-
BOTMINER[15]	-	-	✓	-	-	-	-
MARKOV[16]	-	-	-	-	✓	-	-
BOTSNIFFER[6]	-	-	✓	-	✓	✓	✓
N-GRAM[5]	✓	✓	✓	✓	✓	✓	✓

Table – (1) – Protocol Dependent Comparison

The signature-based detection systems are unable to detect without the prior knowledge (only known Botnets are detected). The other types of detection approaches can detect unknown Bots. The advantages of signature detection are that the response time is high for known Botnets and false negative rate is low, whereas its disadvantage is limited ability to notice zero-day-attacks. A few Botnet detection techniques are effective when there is a change in the C&C protocol or the structure, but these methods are less effective in case of encrypted channels. Anomaly based approaches have the advantage to notice or observe zero-day attack attempts and false negative rate is low, whereas it has high false positive rate. Furthermore, when placed in new environment, it works very slow. The various techniques of DNS and data mining provide the tradeoff. They are also effective for command and control communication.

Approaches of detection		Structure and protocol independent	Real- time detection	Bot detection (unknown)	Bot detection (encrypted)
Anomaly detection	[17]	-	-	✓	✓
	[6]	-	-	✓	✓
Mining detection	[18]	-	-	✓	-
	[2]	✓	-	✓	✓
	[19]	-	-	✓	-
Signature detection	[20]	-	-	-	-
DNS detection	[1]	✓	-	✓	✓
	[21]	-	-	✓	✓
	[22]	-	✓	✓	✓

Table – (2) – Botnet Detection Comparison

CONCLUSION

Botnets are a major serious threat for cyber security. They can affect many services of information and security like data confidentiality, integrity and availability. Botnet can also enable illegal and unauthorized system attacks like spam, Trojan, viruses, worm, key-loggers, DDOS. Though different techniques have been developed for bot detection, an analytical survey of the literature reveals that it is hard to find a universal solution. In this paper, our analysis suggests that hybrid detection techniques combining anomaly and signature approaches can significantly increase overall bot detection accuracy. As the signature detection is for known fingerprints, so this type of detection checks the known bots very quickly and properly. The priority or classification can generally separate traffics signatures for signature detection and the remaining traffic for anomaly approach. This way, signature detection can reduce the traffic for anomaly detection, thereby reducing its false positive rates.

REFERENCES

[1] H. Choi, H. Lee, H. Lee, and H. Kim, "Botnet detection by monitoring group activities in DNS traffic," *CIT 2007 7th IEEE Int. Conf. Comput. Inf. Technol.*, pp. 715–720, 2007.

[2] "64 BotMiner_ Clustering Analysis of Network Traffic for Protocol- and Structure-Independent Botnet Detection." .

[3] A. K. Tyagi and G. Aghila, "A Wide Scale Survey on Botnet," *Int. J. Comput. Appl.*, vol. 34, no. 9, pp. 975–8887, 2011.

[4] D. Geer, "Malicious bots threaten network security," *Computer (Long. Beach. Calif).*, vol. 38, no. 1, pp. 18–20, 2005.

[5] W. Lu, G. Rammidi, and A. A. Ghorbani, "Clustering botnet communication traffic based on n-gram feature selection," *Comput. Commun.*, vol. 34, no. 3, pp. 502–514, 2011.

[6] G. Gu, J. Zhang, and W. Lee, "BotSniffer : Detecting Botnet Command and Control Channels in Network Traffic Georgia Institute of Technology Roadmap • BotSniffer Experimental Evaluation," pp. 1–27, 2008.

[7] I. Technology, H. R. Zeidanloo, A. B. Manaf, P. Vahdani, F. Tabatabaei, and M. Zamani, "Botnet Detection Based on Traffic Monitoring," pp. 97–101, 2010.

[8] A. Karim, R. Bin Salleh, M. Shiraz, S. A. A. Shah, I. Awan, and N. B. Anuar, "Botnet detection techniques: review, future trends, and issues," *J. Zhejiang Univ. Sci. C*, vol. 15, no. 11, pp. 943–983, 2014.

[9] S. Khattak, N. R. Ramay, K. R. Khan, A. A. Syed, and S. A. Khayam, "A Taxonomy of botnet behavior, detection, and defense," *IEEE Commun. Surv. Tutorials*, vol. 16, no. 2, pp. 898–924, 2014.

[10] P. V. Amoli, "A Taxonomy of Botnet Detection Techniques Hossein Rouhani Zeidanloo , Moh amm ad Jorjor Zadeh M . Safari , Mazdak Zamani B . Intrusion Detection System (IDS)," *Ind. Eng.*, pp. 158–162, 2010.

[11] A. R. Baker *et al.*, *Snort 2.1 Intrusion Detection*. 2004.

[12] V. Krmicek, "Inspecting DNS Flow Traffic for Purposes of Botnet Detection," *GEANT3*

JRA2 T4 Intern. Deliv., pp. 1–9, 2011.

[13] D. Hutchison and J. C. Mitchell, *Detection of Intrusions and Malware , and Vulnerability Assessment*. 1973.

[14] “62 29BotHunter Detecting Malware Infection Through IDS-Driven Dialog Correlation.” .

[15] “62 28BotMiner Clustering Analysis of Network Traffic for.” .

[16] “77 37Adaptive pattern mining model for early detection of botnet-propagation scale.” .

[17] A. Karasaridis, B. Rexroad, and D. Hoeflin, “Wide-scale Botnet Detection and Characterization.”

[18] J. Goebel and T. Holz, “Rishi : Identify Bot Contaminated Hosts by IRC Nickname Evaluation.”

[19] T. Strayer and R. Walsh, “Botnet Detection,” vol. 36, no. June 2014, pp. 0–29, 2008.

[20] U. Snort and M. Tcp, “n s t i t u t e u t h o r r e t a i n s f u l l r i g h t s.”

[21] D. Dagon, “Botnet Detection and Response The Network is the Infection,” 2005.

[22] “Revealing Botnet Membership Using DNSBL Counter-Intelligence.” .