ISSN No. 0976-5697

Volume 4, No. 9, July-August 2013

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Detecting Botnets in View of an Efficient Method

Narges Arastouie, Naeemeh Arastooie
Computer Engineering and Information Technology
Amirkabir University of Technology
Tehran, Iran
Arastouie@aut.ac.ir, arastooie@aut.ac.ir

Elham Salimi, Saeed Soltanali
Computer Engineering
Azad University of Khorasgan , Tehran university
Esfahan, Tehran, Iran
e.salimi@khuisf.ac.ir, soltanali@ut.ac.ir

Abstract: With the increasing expansion of Botnets, techniques for identifying and analyzing the behavior of bots extensively can be seen. There are a variety of Tools and techniques to identify Botnets classified into two categories, anomaly-based and signature-based methods. Botnets cannot be detected by signature-based methods through their rapid changes thus, signature-based systems are not suitable for detection. Therefore, we are about to provide a solution in addition to not requiring a specific architecture in order to be able to help us detecting infected client Bots. In this regard, a list of network addresses that were not assigned to any host was allocated to a system in the network by examining the network traffic using Pcab, collecting and analyzing network packets and finally the analysis of network packets the behavior of an infected host is been scrutinized.

Keywords: Bot; Botnet; Bot's behaviour; network traffic; command-control channel

I. INTRODUCTION

Today we see Bots growing as an example of computer worms in the real world [1]. Bots run series of attacks and malicious activities automatically; they get into the user's computer as an unauthorized user and get the command of it. These captive systems are called Zombie. In contrast to the Bot there is a concept called Botnet. Botnet is a network of Bots that do malicious activities through a command and control channel under the command of an intruder known as Bot Master [2]. Figure 1 shows an example of Botnets that are considered as the biggest threat on the internet network and security. Botnets can be generally defined as follows:

a. A Botnet comprises a minimum of a server Bot or controller and one or more client Bot systems.
b. The center of each client Bot is the command interpreter that can retrieve orders independently.
c. Botnets are a series of malwares such as viruses, worms and other infections.
d. Bots are controlled remotely by a hacker.

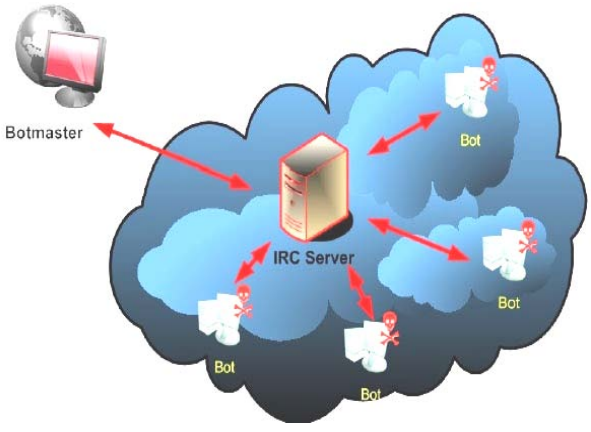

Figure 1: a General View of Botnet's Structure.

Bots sometimes use techniques existing in computer worms, malicious software key loggers (in order to register users entries like usernames and passwords) and Rootkit techniques (in order to stay hidden in the victim system). Like worms they can propagate themselves rapidly on the internet and expand the network's contamination extent. [3] They can be disseminated through sending e-mails, quick messages and using shared folders. In some cases BotMaster does the dissemination through captive computer's web servers, visiting different websites and downloading items. Dispersion and contamination mechanisms vary for different kind of Bots according to Bot Master's discretion [4].

In general words Bots use a range of activities such as DDOs, sending spam emails with infected attachments, the use of weak security, key loggers and the latest vulnerabilities for distribution. Unlike other kind of attacks, Bot nets can captivate and exploit hundreds and thousands of computers for their personal purposes and as for their motives can gain the highest achievement by sending a set of commands to infected computers [5].

So the BotMaster uses different ways inorder to contaminate new systems. The main difference between Botnet and malicious code is their ability to take the command via command-control channel and also self-updating in the course of this channel.

Another difference is that Bot nets look for predetermined goals and they have considered other menaces mentioned before like viruses and malicious code aiming to infect destination computer's files, using bandwidth, CPU, memory and etc. consequently they can damage the system in such way that the user may not realize. As an example disabling the antivirus can be declared, thus host-based methods cannot be used for identifying and defeating Bot's attacks. Since antivirus tools are based on signature-based methods Bots can easily remain hidden from being identified by antivirus signatures via faster self-updating, in other words Bots can hide themselves from antivirus tools.

Therefore instead of host-based methods our focus can be kept on network-base methods in order to detect Bots [6][7].

According to the mentioned reasons using network based method and anomaly based methods is suggested to detect attacks. Follows that the normal traffic is defined for each network and any kind of deflection from this normal traffic is considered as anomaly. The only problem with this method is that it causes false positive alerts. False positive alerts are in fact the packet rates that have been identified as malicious packets so the lower rate the better detection method. For this purpose using snort which is based on signature is usually suggested for existing detection methods in order to balance the rate of false positive alerts. Snort is used for different purposes that will be discussed further. Anomaly-based methods have been much considered as one of the network-based methods in Botnet detecting [8].

In this methodology normal traffic is been defined for each network and any kind of deviation from the normal traffic is been reflected on as an anomaly. The worst thing with these methods is the number of false positive alerts. These alerts are actually those packet rates that have been detected as vulnerable packets. Thus batter detection way results the lower alert rate percentage. So as to balance false positive alert rates usually Snorts are being used [9].

II. BOT'S DIFFUSIONAL BEHAVIORS CLASSIFICATION

Bots can do any devastating operation on computers and networks that have under control. In other words Bot's classification depends on their operational characteristics and behaviors specially their dissemination methods and services that they provide for the Bot Master[10]. Here some of Bots devastative and illegal actions are listed:

a. Client-Bots exploitation:

Indeed in order to obtain sighted passwords or searching for system's vulnerabilities use infected clients.

b. Directing DDoS attacks :

To begin an attack the attacker should command a group of Bots to start working. This kind of attack is presented only for an internet service like a web site and it can cause abuse of system's resources. One of these resources is the band width used by sending a stream of HTTP, UDP, TCP SYNC and this operation persists until disabling the whole system or the time that it can't provide any service to the user. Figure 2 is a view of DDoS attack.

c. Acquiring personal and financial information:

Bots may use this technique to thief significant information; in reality it often directs users to enter details like username, ID's of credit cards and etc.at a fake website whose look and feel is almost identical to the legitimate one [11].

d. Directing spam companies to send spam email:

Using Botnets systems can change into a spam mail distributing net. Some Botnets send the spam mail over to the spam email's proxy for further diffusion. Due to the spam proxy, spammers can prevent themselves to be identified instead of forwarding spam mails autonomously from each Bot. figure 3 shows spam mail sender Bots using proxy [12].

e. Directing defrauding companies:

Click fraud defines the act of directing users to click on company's infected ad wares. A BotMaster can easily capture a computer so as to add it to its Botnet. This is simply by clicking on an advertisement banner on the internet. These attacks follow personal and commercial purposes and encompass 27.8% clicks on daily ads.

f. Ad wares installation without user permission (installing unauthorized ad wares) :

Botnets usually have an appropriate format for worm dispensation since they can legitimize all their illegal activities [13].

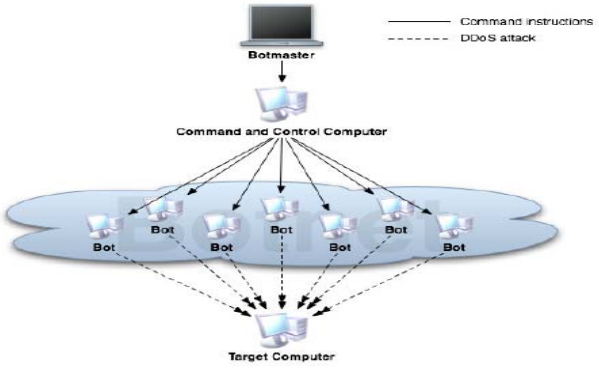

Figure 2: a View of DDoS Attack.

Figure 3: Spam Mail Sender Bots Using Proxy.

Botnet's dissemination mechanism is investigated through comparing their design complexity, recognizing ability, dispensation rate and size and the result is shown in table I. Table II also figures different categories for attacks. [14]

Table: 1 Botnet's Dissemination Mechanism

PopulationSize	Propagation Speed	Detectability	Design Complexity	Propagation Methodology
High	Low	High	Medium	Exploit: Operating System
Medium	Medium	Medium	Medium	Services
Low	High	Low	High	Applications
High	Low	Medium	Low	Social Engineering

Table: 2 Attacks Categories.

Topology	Detectability	Design Complexity	Attack Value
Single Host DDOS	High	Low	Low
Multi Host DDOS	Medium	Medium	Medium
Identity Theft	Low	High	Medium
Spam	Medium	Medium	High
Phishing	Medium	High	Medium

III. DIFFERENT COMMUNICATION PROTOCOLS AND COMMAND-CONTROL PROTOCOLS

Botnets in general don't create a new communication protocol but they use the protocol that is compatible with the network. The following protocols are commonly used: [23][3]

A. IRC protocol

Almost all today Bots use IRC protocol to communicate and it has the most use. This is basically a protocol designed for group communication (multipoint) and if necessary it can be converted to a point to point communication by sending private messages. Consequently this type of flexible connection has been used mostly for BotMasters so their commands can be posted to the group or one of the Bots [15].

B. HTTP protocol

This protocol is been used due to IRC protocol's problems on firewall configuration and not passing IRC packets through.

C. Peer to peer protocol

In this protocol Bots have point-to-point (mutual) communication. Accordingly it is intricate to identify them because finding a clientBot cannot cause the whole Bot net destruction And hence the BotMster should be connected to at least one of the infected computers to control the entire botnet [11].

Future progresses go toward using these kinds of Bots [16].

One of the key parts of the Bots' important features is the command-control service. The mentioned services consider kinds of architectures for their configuration. These architectures can be divided into three different categories[17]:

D. Centralized model :

In this type of Botnet all computers are connected to a command-control server. This server waits for being connected to new Bots in order to record them in its database, follow their status and send them the issued instructions. The BotMaster ought to be connected to this sever so as to send its commands [18].

E. Peer-to-peer model :

In a decentralized Botnet bots connect to a large number of infected machines instead of being connected to the control center. Commands are transmitted from Bot to Bot and the same they will be transmitted to others. In this case, the Bot should be connected to at least one of the infected computers to control the entire Botnet [19].

F. Stochastic model :

The communicational system of this model is based on the principle that a Bot is not informed of other Bots.

In this topology a Bot or controller randomly scans the internet and sends its encrypted messages to the discovered Bots [20].

These three models cover almost all cases of today Bots. Table III shows an investigation on types of command-control topologies from the view of design complexity, detection, the delay of commands, durability and stability [21].

Table: 3 Different Topologies for Command-Control.

Topology	Design Complexity	Detectability	Message Latency	Survivability
Centralized	Low	Medium	Low	Low
Peer-to-peer	Medium	Low	Medium	Medium
Random	Low	High	High	High

For sending commands to Client Bots, controlling and better communication with them there are two ways of sending commands from the command-control server's side which includes:

a. Push style:

In this methodology commands are forwarded to the Client Bots and will be replied in real time Like the IRC based command-control, figure 4 shows a view of it [15].

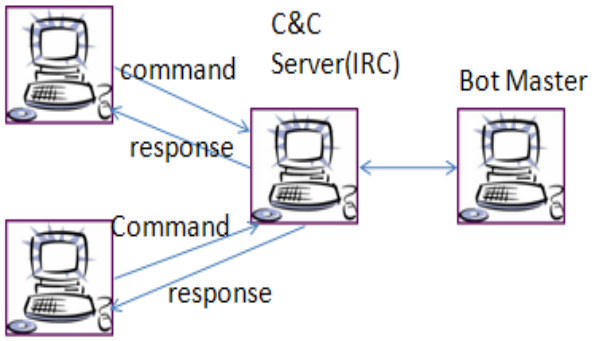

Figure 4: a View of Push Style Command.

b. Pull style:

Unlike the previous method here Bots first have to download commands in order to send them thus there is no real time mode. Figure 5 is a view of it [22].

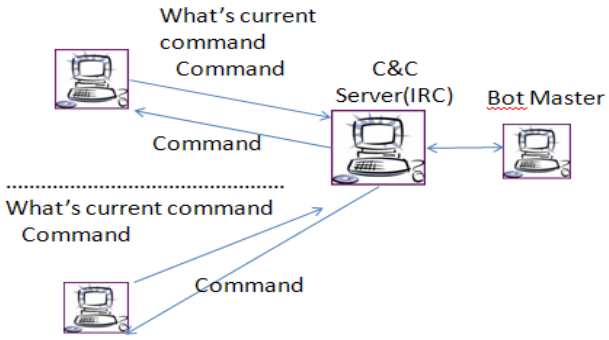

Figure 5: a View of Pull Style Command.

IV. TYPES OF ATTACK DETECTION SYSTEMS AND MECHANISMS

One of the attack detection system types is the signature-based detection model. Intrusion detection system (IDS) is based on looking for those activities that are compatible with known signatures of attacks or vulnerabilities as a matter of fact it detects infections based on a set of predefined defaults [10].

They download current traffic then compare it with the existing traffic in the database so they can take the necessary measures. This method is used less for detecting Bots since Bots can be updated faster than known signatures.

a. Advantages of signature-based models:

(a). The real time detection.
(b). Users are being special by applying any rule.
(c). There is no false positive alert.

b. Disadvantages of signature-based models:

(a). Zero-day attacks are not detectable. (For the reason that signatures are predefined).
(b). Large databases of new/old signatures along with their effects are kept.
(c). Malicious code's similar body and small differences between their lost signatures would create difficulties.
(d). Overloading in encryption causes incompatibility with known signatures.

In contrast to the signature-based models there are anomaly-based detection models. In this approach the identification process of intrusion detection system is done by tracking any kind of anomaly in the network traffic. A fully dynamic method which already has no information about the malicious behavior as statistical technique, evolutionary and etc[12]. anomaly-based detection techniques are more effective than the signature-based ones because:

(a). They don't need to have a basic knowledge of command-control server and signature's contents.
(b). Are able to analyze the encrypted packets and identify them if necessary.
(c). To monitor the network not only they don't need a lot of Bots in the network but also they are able to continue their activities with a few of Bots. It means that they don't require a lot of communication with the command-control server.
(d). They have less false positive alerts and false negative alerts rates.
(e). Having a single algorithm they can support a large number of malicious codes.

c. Disadvantages of anomaly-based models:

(a). The learning process may take a long time (have to watch the network traffic at least for a short period of time).
(b). The user must be more involved in the algorithm and its implications.

Today the new mechanisms have tried to use both methods together since the use of signature-based methods is necessary for balancing the malicious false positive alerts rates [15]. Based on different models, mechanisms to detect infections of intrusion detection systems takes place in two categories:

a) Host-Based intrusion detection systems (Host-Based IDS):

Studies attacks on a host system or several host systems like the Operating System's contents, files and system applications. There are specific tools for this task on the server for the network administrator.

b) Network-Based intrusion detection systems (Network-Based IDS):

Analyzes and identifies the existing attacks based on captured packets from the scattered sensors over the network. For example Snort can be used to capture and analyze the network packets. Figure 6 is a view of network traffic patterns used in anomaly-based detection models in a network-based ID.

Malicious activities of network traffic patterns may express different network characteristics in different levels. Pattern types are as follows: pattern of the package level (package's header and content), patterns of the network stream (statistical and evolutionary pattern), pattern of the host level (statistical dialog patterns based on a person or group).

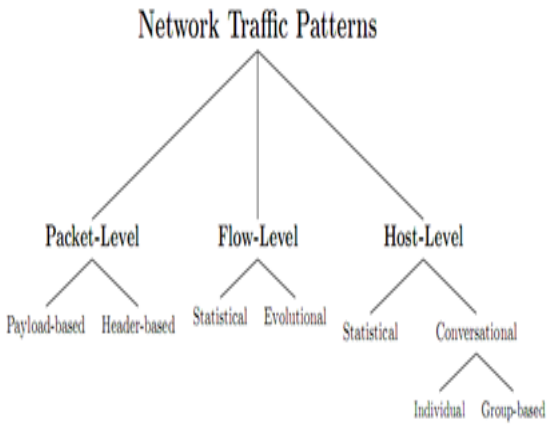

Figure 6: Network Traffic Pattern View

V. THE PROPOSED METHOD FOR THE DETECTION

Monitoring the performance of similar systems, analyzing and reviewing the information of a statistical society of Botnets the following results were obtained:

a. Intrusion to the victim's machine without the user's knowledge along with other applications.
b. Sitting in the system's startup.
c. Settle down in the most important system's directories.
d. Settle down in the system's temporary directories.
e. The ability to update in a hidden manner.
f. Using the Rootkit technology.

The presented method presents a new mechanism for analyzing Bot's behavior and detecting them using infected clients over the local network (LAN). This mechanism tries to distinguish the network normal traffic and the produced traffic from the malicious code by differentiating traffics and data analysis. The main idea of this approach is inspired from the way Bots propagate. Once a system is infected with Bots the contamination tries to be dispersed over the network with different ways. In other words it tries to find vulnerable clients and send them an exploit according to the existing vulnerability in thoes systems. In a real network usually all the physical addresses are not been used in that network. For instance from the 100 percent received IP addresses in an organization 90 percent are left for development and the other 10 percent are unused. Hence if a network traffic is been sent to those 10 percent IPs it is not a normal traffic because while these IP addresses are not going to provide any services having request to these addresses is unusual so it shows that somebody is trying to attack and identifying the server. Attackers forward their requests to communicate through all existing addresses in the network. If the request is been responded they seek to server. ARP requests can do this job. By allotting those unused IP addresses in the local network (LAN) to a host computer in that LAN incoming packets could be captured with libpcap software and by analyzing these packets the infected Client Bot can be diagnosed.

If there were packets sent to those addresses we could be suspicious to those packets and their sender as an infected Client Bot. There for sending any packets to these addresses expresses an aggressive attack on the entire network which can be counted as a DDoS attack that is so effective for disseminating Bots over the network. Due to the internet packet sender protocols the host can be considered as a host infected with Bots. Bots at first recognize all the systems in a network by forwarding ARP messages and if they received any response send the following packets to the system:

a. Scanning system vulnerabilities.
b. The exploit packets that a vulnerability uses.
c. Backscatter [Reflection of DoS attack].
d. Flaws in configuration.

To implement this approach using ARP protocol unallocated addresses should be identified and an address is allotted to a victim system. then run the pcap on the victim system to monitor the network traffic and capture it (in Linux versions libpcap library and in windows winpcap is used).

The first step in communicating with pcap is to identify network connections on the operating system and providing a list of them. This is done by two functions on the pcap that their general form is as follows:

```
intpcap_findalldevs (
pcap_if_t ** alldevsp,
char * errbuf
)
intpcap_findalldevs_ex (
char * source,
structpcap_rmtauth * auth,
```

*pcap_if_t ** alldevs,*
*char *errbuf*
)

Since data in IP layer has a 1500 Bytes limitation on length if there were more data in a packet the data will be divided into 1500 Bytes packets.

Most of malicious data encompassing exploitation use the TCP protocol to ensure the accuracy of their connection also Bots which work with central command-control server use the TCP protocol in order to send their request after communication phase. On the other hand peer to peer Bots use the UDP protocol to send information which provides unreliable services. The majority of malicious messages are sent via TCP packets to guarantee the packets have reached their destination hence in this method more focus is kept on TCP packets.

VI. THE EVALUATION OF METHOD FOR IN REAL WORLD

To test the proposed method a scenario is been designed. This scenario is formed by Metasploit tool and a statistical society of existing Botnets. The testing scenario is done through a Hub Switch in a LAN environment.

The network range is 192.168.0.X and entails two phases to evaluate the proposed method's efficiency:
First phase: recognizing suspect origin.

For one thing a list of unallocated addresses is been prepared. Then an address is allotted to a system and waits for a packet using pcap. For example in the network mentioned above the unallocated address was 192.168.0.101. As soon as a packet is received in this destination the source address considers it as suspected address which tries to contaminate this address so as to contaminate the whole network range. Henceforth the destination address considers 192.168.0.101 as a suspected address out of the network that aims to contaminate the whole network range and all of its interactions and behaviors go under control in order to prove its infection if it is

This review is done based on captured packets by pcap for that source. All transferred packets related to 192.168.0.101 are sent from the port number 56302 using TCP protocols to all IP addresses over the network range. Live IP addresses on the network do the essential hand-shakings. After hand-shaking between 192.168.0.101 and 192.168.0.102 the transferred information can be seen which states the communication with an IRC server since the source port is 194 that is specifically used for IRC communications. Figure 7 shows the communications. It has to be mentioned that for ensuring the accuracy of the test the system has been infected with some Bots in order to analyze its behavior and performance. To ensure that the suspect system is already infected the exchanged data have been examined of course the raw data are in Hexadecimal that are been changed to String manually and the information in the table IV will be resulted.

In those figures it can be clearly seen that the system is infected with a kind of Bot which uses IRC channel to communicate out the network.

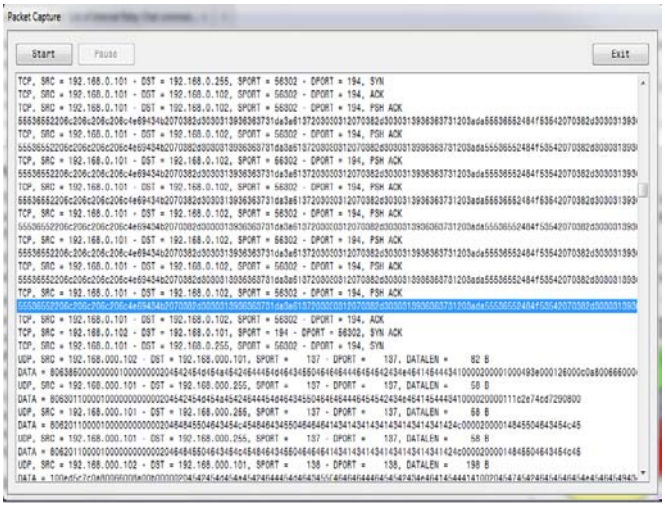

Figure 7: a View of Captured Packets by Pcap

Table: 4 Information of infected system

USeR l l l l	55536552206c206c206c206c
NiCK p8-00196671	4e69434b2070382d303031393636373 1da
a7 001 p8-00196671 : :	3a6137203030312070382d303031393 6363731203ada
USeRHOST p8-00196671	55536552484f53542070382d30303139 3636373120
:a7 302 p8-00196671 :p8-00196671=+l@192.168.0.101	3a6137203330322070382d303031393 6363731203a70382d303031393636373 313d2b6c403031302e3132392e323131 2e313320202020
JOiN #p8 ihodc9hi	3a4f694e202370382069686f64633968 692020
:a7 332 p8-00196671 #p8 :!Q gfcagihehehadkcpcpgigpgngfheg phhgocogbgpgmcogdgpgncphihih igmgpgmhhhegggjgigbhihihihicp hdgpgdglhddjgbcogkhagh	6137203333332207038 2d3030313936 3637312023703 8203a2151206766636 1676968656865 6861646b6370637067 696770676e676668656770686867 6f636f67626770676d636f67646770676e63 706869686968696 76d6770676d68686 8656767676a6769676268696869 6869637068646770676476c6864646a 6762636f676b68616768
:a7 333 p8-00196671 #p8 a 1134159047	3a6137203333332070382d303031393 6363731202370382061203131333431 3539303437
:a7 366 p8-00196671 #p8:	3a6137203336362070382d303031393 6363731202370382 03a

VII. CONCLUSION

Generally there are a variety of tools used to detect Botnets which are classified into two categories: anomaly-based and signature-based.

Realizing Botnet's behavior is so useful but it is not sufficient for detection since Botnet's rapid changes prevent their identification by signature-based systems so we tried to present a solution that can detect infected ClientBots without any need to a special architecture or topology. In this regard a list of unallocated IP addresses was allotted to a system in the network. Therefore by exploring the network traffic using pcap network packets can be obtained and after all the behavior of an infected host has been analyzed.

VIII. REFERENCES

[1] Y. AI-Hammadi, "Behavioural correlation for Malicious Bot Detection", Ph.D thesis submitted to The University of Nottingham, 2010.

[2] G. Gu, "Correlation-based Botnet Detection In enterprise Networks", Ph.D thesis submitted to College of Computing, Georgia Institute of Technology, Pp. 1-6, 2008.

[3] W. Lu and A. A. Ghorbani, "Botnet Detection Based on IRC-Community", published in international journal of computer science and information security , 2008.

[4] A. K. Seewald and W.N. Gansterer, "On the detection and identification of botnets", Submitted to Computers and Security, 2009.

[5] D. Tran, "Propagating Mailicious Codes: Theory and exprements", Ph.D thesis submitted to State University of New York at Buffalo, 2009.

[6] H. R. Zeidanloo, A. A. Manaf, R. Ahmad, and M. Zamani," A Proposed Framework for P2P Botnet Detection", IACSIT International Journal of Engineering and Technology, Vol.2, 2010.

[7] J. Liu, Y. Xiao, K. Ghaboosi, H. Deng, and J. Zhang, "Botnet: Classification, Attacks, Detection, Tracing,and PreventiveMeasures" Published in EURASIP Journal on Wireless Communications and Networking, 2009.

[8] C. A. Schiller, J. Binkley, D. Harley, G. Evron, T. Bradley, C. Willems, and M. Cross, "the killer web application."published in syngress, 1st Edition,2008.

[9] Ping Wang, S. Sparks, C.C.Zou, "An Advanced Hybrid Peer-to-Peer Botnet."published in IEEE Transactions. Dependable and Secure Computing, vol. 7, issue. 2, April-June 2010.

[10] DAI Wei, "The Comparative Study of the Botnet Detection Algorithm". Published in Electronic Engineering Journal, Issue. 3, Page 9-11, 2011.

[11] Raihana Syahirah Abdullah, Mohd Faizal Abdollah, Zul Azri Muhamad Noh, Mohd Zaki Mas'ud, Siti Rahayu Selamat, Robiah Yusof, "Revealing the Criterion on Botnet Detection Technique". Published in IJCSI International Journal of Computer Science Issues, Vol. 10, Issue. 2, No. 3, March 2013.

[12] M. Bailey, E. Cooke, F. Jahanian, Y. XU, and M. Karir, "A Survey of Botnet Technology and Defenses", In Proceeding of Cyber Security Applications and Technology Conference, Pages229-304, 2009.

[13] W. Lu, M. Tavallaee, G. Rammidi, and A. A. Ghorbani, "BotCop: An Online Botnet Traffic Classifier", In proceedings of 7th Communication Networks and Services Research Conference, 2009.

[14] R. F. Erbacher, A. Cutler, P. Banerjee, and J. Marshalla, "A Multi-Layered Approche to Botnet Detection", In Proceeding of Cyber Security Applications and Technology Conference, 2008.

[15] P. Wurzinger, L. Bilge, T. Holz, J. Goebel, C. Krugel, and E. Kirda, "Automatically Generating Models for Botnet Detection", In proceedings 14th European Symposium on Research in Computer Security(ESORICS) , 2009.

[16] L. Liu, S. Chen, G. Yan, and Z. Zhang, "BotTracer: Execution-based Bot-like Malware Detection", In proceedings of 11th Information Security Conference (ISC), 2008.

[17] J. R. Binkley and S. Singh, "An algorithm for anomaly-based botnet detection", In Proceedings of USENIX Steps to Reducing Unwanted Traffic on the Internet Workshop (SRUTI), 2006.

[18] Jun'ichi Takeuchi, "On Botnet Detection using sparse Structure Learning", In proceedings of 7th Communication Networks and Services Research Conference ,2009.

[19] M. A. Rajab, J. Zarfoss, F. Monrose, and A. Terzis, "A multifaceted approach to understanding the botnet phenomenon," in Internet Measurements Conference (IMC '06), Oct 2006.

[20] E. Cooke, M. Bailey, F. Jahanian, and R. Mortier, "The Dark Oracle: Perspective-Aware Unused and Unreachable Address Discovery", In proceedings of 3rd Symposium on Networked Systems Design & Implementation (NSDI), 2007.

[21] Lima,peru,"Guide on Policy and Technical,Approaches against Botnet", Submitted to Security & Prosperity Steering Group (SPSG) , 2006.

[22] G. Gu, J. Zhang, and W. Lee."BotSniffer: Detecting botnet command and control channels in network traffic,"In Proceedings of the 15th Annual Network and Distributed System Security Symposium (NDSS'08), 2008.

[23] P. Barford and V. Yegneswaran, "An inside look at botnets", Special Workshop on Malware Detection, Advances in Information Security, Springer Verlag, 2006.