Automated Simulation P2P Botnets Signature Detection by Rule-based Approach

Raihana Syahirah Abdullah
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka (UTeM)
Hang Tuah Jaya, 76100 Durian Tunggal, Melaka

Zul Azri Muhamad Noh
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka (UTeM)
Hang Tuah Jaya, 76100 Durian Tunggal, Melaka

Faizal M.A.
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka (UTeM)
Hang Tuah Jaya, 76100 Durian Tunggal, Melaka

Nurulhuda Ahmad
Faculty of Engineering and Built Environment Universiti
Kebangsaan Malaysia (UKM)
43600 Bangi, Selangor

***Abstract*—Internet is a most salient services in communication. Thus, companies take this opportunity by putting critical resources online for effective business organization. This has given rise to activities of cyber criminals actuated by botnets. P2P networks had gained popularity through distributed applications such as file-sharing, web caching and network storage whereby it is not easy to guarantee that the file exchanged not the malicious in non-centralized authority of P2P networks. For this reason, these networks become the suitable venue for malicious software to spread. It is straightforward for attackers to target the vulnerable hosts in existing P2P networks as bot candidates and build their zombie army. They can be used to compromise a host and make it become a P2P bot. In order to detect these botnets, a complete flow analysis is necessary. In this paper, we proposed an automated P2P botnets through rule-based detection approach which currently focuses on P2P signature illumination. We consider both of synchronisation within a botnets and the malicious behaviour each bot exhibits at the host or network level to recognize the signature and activities in P2P botnets traffic. The rule-based approach have high detection accuracy and low false positive.**

***Keywords*—Botnets; P2P Botnets; Signature; Rule-based**

I. Introduction

The botnets population are rapidly growing and they become a huge threat on the Internet. Botnets has been declared as Advanced Malware (AM) and Advanced Persistent Threat (APT) listed attacks which able to manipulate advanced technology where the intricacy of threats need for continuous detection and protection as stated by [1] [2] [3]. These attack will be almost exclusively for financial gain as claimed by [4]. The chronology of botnets attack disclosed the evidence on the seriousness and sophisticated nature in the recent cyber-attack. The growth of Internet itself had reflected to the growth of incidents in security environment. The numerous of security incidents that occurred will cause major loss for the organizations. From a survey conducted by American Society for Industrial Security and Pricewaterhouse-Cooper, there was a loss of USD45 million for 1000 companies due to the security breach [5]. According to marketing networking group Chief Marketing Officer (CMO) Council, a data breach could cost an average of US Dollar14 million on a recovery cost [6]. A survey made by [7], Computer Security Institute (CSI), Computer Crime and Security reported that the lost due to security breach was US Dollar 288 by 618 per respondent. This shows a multiple costs also needed to clean up the botnets infections. Significantly, this research bridged important relationship with botnets technology as depicted in Figure 1 where early emergence of P2P botnets existence in year 2002 and rapidly growth until now with more robust, complicated and flexible P2P botnets. This fact has motivated research communities to do further research on P2P botnets issue.

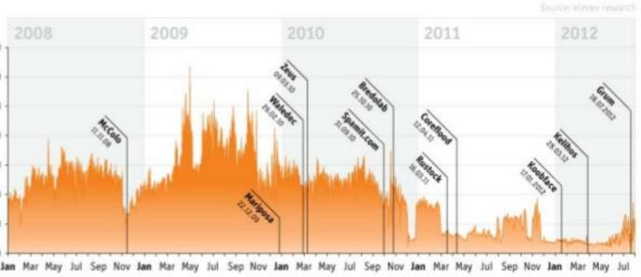

Fig. 1. Evolution of P2P Botnets Technology [8]

So, it was important to build a security mechanism which designed to prevent intrusion from hacker so that it can take action and improves the system security [9]. This research field is known as intrusion detection. Intrusion Detection System (IDS) is a system that continually monitors the dynamic behaviour of the computer system to warn against actions that compromise the integrity, security, and availability [10]. In the real environment practice, IDS focused on detecting known threats or detecting the volume of the traffic generated by bot host after it has been activated [11]. An important problem in the field of intrusion detection is the management alerts [12] as IDS tends to produce high number of false positive alerts as claimed by [13] [14] [15]. Most of the botnets has generating low-volume periodic communication to botmaster which increased false alarm rate and make it harder to be detected as mentioned by [11] [16].

The increment of the traffic volume can cause the IDS to produce large number of alarms as discussed by [17]. Reducing false alarms is a serious problem needs an attention in ensuring the IDS efficiency as mentioned by [14] [18].

As the botnets and advance malware evolve rapidly, the approaches and techniques for detecting botnets need to be improvised. Previous researchers have contributed lots of efforts and works to address the botnets issues. Most of the approaches have developed for detecting Internet Relay Chat (IRC)-based or Hypertext Transfer Protocol (HTTP)-based botnets. These approaches only operate at the network level that focused for traffic signatures on flow patterns [19]. Unfortunately, the detection approaches designed for IRC-based and HTTP-based botnets may become inefficient against the new P2P based botnets as it focused only on the specific protocol [1]. Basically, P2P botnets detector tool is also identified as intrusion detection tool. The exponential growth of P2P botnets and the new distribution channels available to cyber criminals identify that the need for good protection is crucial. So that, it would be feasible to detect P2P botnets through P2P traffic signatures and behaviors [20].

The rest of paper is organized as follows. In Section II, we provides details background on the signature-based detection concept. Section III will describes the methodology of overall process. Next, Section IV will elaborates details on detection module of our proposed P2P botnets automated rule-based detection. While, the results and discussion are also discuss in this section. At last, our paper is concluded in Section V.

II. Background

Generally, signature-based is a supervised learning method where it develop the capability on detecting malicious behaviours on the basis of previously seen malicious events [21]. The signature-based detection modelled the known attack behaviour that learned from attack pattern [22]. By recognizing known attack methods, signature detection is able to recognize when the intrusive patterns occurred. As claimed by [23] during the process of inspection, when the suspicious behaviours partially corresponded with the records of intrusive knowledge base, it automatically can be judged as the intrusion.

The main selection of Java also because it can provides the IF-ELSE statement that referring to the basic of rule-based declaration in terms of generating the signature. All of these benefits are needed in developing the P2P botnets signature detection. Due to the P2P botnets detection have their own signatures and behaviours, therefore a NetBeans IDE version 7.4 are the best selection and option in order to implement the rule-based signature in this research. Generally, NetBeans is an open source and free software that address the needs of developers, users and businesses particularly to enable them to develop the products quickly, efficiently and easily by leveraging the strengths of Java platform [24]. This statement is fully supported by [25] where they were stated that the NetBeans platform provides a reliable and flexible modular architecture to application developers. Conveniently, it also helpful for develop Java desktop, mobile and web application. The strength of Netbeans's offered the best support for latest technologies, fast and smart coding, rapid user interface development and rich set of community provided plugins. Previously, [25] declared the latest version of NetBeans IDE and platform certainly keeping the rhythm, introducing new high-impact features and revamping traditional functionality at full throttle.

Another concern of this paper is how to determine the P2P botnets. This paper highlights the need of analyzing the behaviors and parameters of botnets to determine the anomalous in P2P botnets. Through the prediction of P2P network traffic will define several behaviors and parameters that helped to distinguish between P2P normal and P2P botnets. Otherwise, proposed detection module for P2P botnets will help to solve the issue of limited detection technique in order to detect anomalous P2P botnets. The propose detection module for P2P botnets will improvise the network security. So, the details process flow for P2P botnets detection module will be discovered in next section.

III. Methodology

This section discusses in detail the process involved for the P2P botnets detection as depicted in Figure 2. The process started with input the data from raw of P2P botnet file. Then, the system will be do automated process by show the analysis result in short time. The unknown behavior or signature that has been detected will send to anomaly stage for further action.

Generally, signature-based is a supervised learning method in detecting the malicious behaviour on the basis of previously seen malicious events. The second stage of P2P Botnets Detection Module is the signature-based detection that make detection on known attack. The detection signature has been composed through analysis part. The P2P botnets detection technique has been developed as signature module. This technique has the capability to analyse the malicious activities described as variant behaviours, make classification on P2P variant types and sub-attack types and generate the conclusion either the submitted files is a P2P normal or P2P botnets indeed. This technique also able to produce the report on activities performed by P2P botnets.

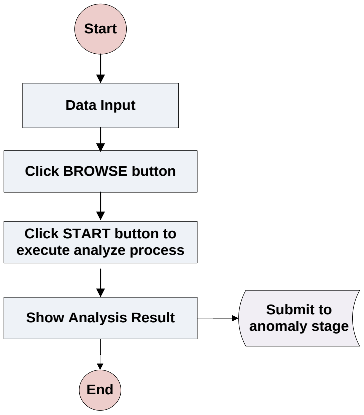

Fig. 2. Process Flow of Signature Module

The signature-based in this detection technique is involved three layer detections as indicated in Figure 3. The three layer detections in signature-based technique consist of:

- Detection only scenario - Mainly decide whether the logs or packets are normal or abnormal
- Detection and classification - Detect the logs or packets by classify them into six main attack types/variants
- Detection and detailed classification - Make classification of sub-attack types based on their attributes and behaviours for every variant

```
The rule-based method role as a signature-based technique to detect
known P2P botnets
    Load captured packet and host log (Input= S1, S2)
    Initialize S with $S_{i}$=normal/abnormal, $S_{ii}$=classification
    attack types, $S_{iii}$=classification sub attack types
    If $S_{i}$=abnormal then $S_{i}$=abnormal;
    Else $S_{i}$=normal
    End
    If $S_{ii}$=classification attack types then $S_{ii}$=classification attack
    types;
    Else $S_{ii}$=none;
    End
    If $S_{iii}$=classification sub attack types then $S_{iii}$=classification
    sub attack types
    Else $S_{iii}$=none;
    End
    If $S_{i}$=abnormal or If $S_{ii}$=classification attack types or
    $S_{iii}$=classification sub attack
    types then Detected, d=1
    Else if $S_{i}$=normal  then Detected, d=0 (proceed with
    statistical test on anomaly-
    based section )
    End
```

Fig. 3. Rule-based detection

The principal step in signature-based has the ability to immediate detection and impossibility of false positives. But signature-based is only capable to be used for detection of well-known botnets. More important, very similar bots with slightly different signature may be missed-out to be detected. However, the anomaly-based technique faced with the problem of detecting unknown botnets through show existence of bots in the network. Anomaly-based technique also has the extra capabilities in terms of reducing false negative alert and detecting multistep attack. Nevertheless, it cannot reduce the false positive alert which can only be reduced by using signature-based technique. Hence, this has given an implication that there are complement each other weaknesses. The fully results are briefly discusses in the next section.

IV. Result and Discussion

Generally, NetBeans has been chosen as its supports java as the main language in its application. As claimed by software programmer and designer that java is simple, free, easy to design, easy to write, easy to compile, debug, and able to learn rather than others programming languages. Besides that, Java is a platform-independent, portable and flexible in nature where a program easily to run from one computer system to another. The most significant feature of Java that appropriate for this research because its support host log and network packet in Packet Capture (PCAP) format and Comma-Separated Values (CSV) format. It also supported by Graphical User Interface (GUI) with integration of network libraries too. Reasonably, the selection of Java also because it can provides the IF-ELSE statement that referring to the basic of rule-based declaration in terms of developing the P2P botnets signature detection. Thus, the NetBeans IDE version 7.4 are the best option in order to implement the rule-based signature. In the signature-based stage, these research requires a PCs to develop and test the system.

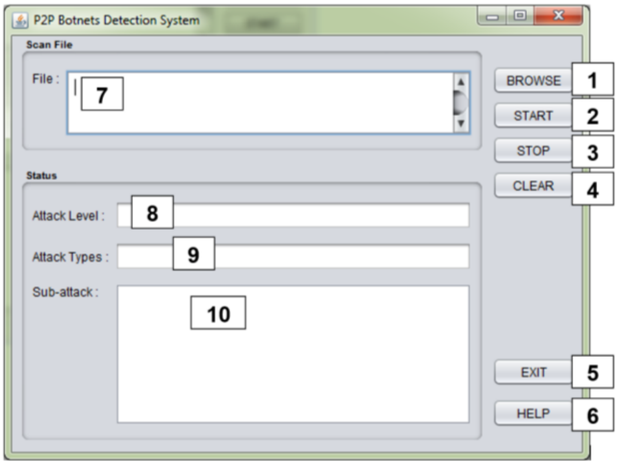

Fig. 4. Design of Signature-based Module

The signature-based system that will be developed here will have an interface that possible to be used by users towards the convenience of designing usage. By designing the interface, it offers the opportunity for the users to perform the detection process, tailored to their needs. The interface is illustrates in Figure 4. Then, the description of each control which marked by number will be discussed in the following table. The explanation about each button and text area are describes in Table 1.

TABLE I. Description of Caption in Signature-based Module

No.	Control Caption	Processing	Description
1.	BROWSE	User: Click the BROWSE button. System: Go to BROWSE file window.	This button is used to open the Browse File window, to choose the file that want to be submitted.
2.	START	User: Click the START button. System: Begin the analysis process.	This button is clicked if user wants to start the process.
3.	STOP	User: Click the STOP button. System: Stop the analysis process.	This button is clicked if user wants to stop the process.
4.	CLEAR	User: Click the CLEAR button. System: Clear all the current process.	This button is clicked if user wants to clear the text area.
5.	EXIT	User: Click the EXIT button. System: Exit the whole system.	This button is clicked if user wants to exit the system.
6.	HELP	User: Click the HELP	This button is

No.	Control Caption	Processing	Description
		button. System: Go to HELP file window.	clicked if user wants to get the related info about this system.
7.	File Text Area	User: Choose a file by clicking BROWSE button first. System: Show the path of chosen file.	Showing the file path of the selected or browsed file.
8.	Attack Level Text Area	User: Click the START button. System: Automatically generate the level of analysis.	Showing the attack level whether the host-level or network-level.
9.	Attack Types Text Area	User: Click the START button. System: Automatically show the types of variants.	Showing the types of variants.
10.	Sub-attack Text Area	User: Click the START button. System: Show the signature for detected variant.	Showing the list of attributes and behaviours for each of detected variant.

The result from Table 2 showed that the signature-based detection has the capabilities to predict 100% correctly for the overall detection rate with 0% of false alarm rate of the P2P network traffic. The improvement of overall detections in the signature-based module from classification table in data mining module are indicated that this signature-based system technically effective for outcome attack detection. Therefore, it can be summarized that this signature-based detection has better prediction and capabilities to distinguish between the normal and attack events reached for thousands of dataset for each variant.

TABLE II. SIGNATURE-BASED MODULE DETECTION RESULT

Variant	False Negative (FN)	Accuracy (A)	Detection Rate (DR)	False Alarm Rate (FAR)
Invalid Hash	0	100%	100%	0%
Allaple.L	2	99.99%	100%	0%
RBot	6	99.98%	100%	0%
Palevo	3	100.00%	100%	0%
Srvcp	2	99.99%	100%	0%
Tnnbtib	0	100%	100%	0%

Inclusively, this signature-based detection system in this research promises the better enhancement in P2P botnets detection technique. The entire six variants have fully detected as the P2P botnets based on the detection result. But, Table 2 shows the detail of the result where the False Negative (FN) emphasize some of the undetectable attributes or undetectable P2P botnets values as the attack declares as normal. Alternately, this problem can be tackled by conducting the anomaly-based detection. In the next of detection stage, the chi-square statistical test with multivariate process has been perform. The tabulated of false negative that indicates undetectable P2P botnets has been proves can be successfully detected through the statistical approach. The detail steps on detecting P2P botnets through anomaly statistical test has been explaining in the next sub-section.

V. CONCLUSION

Currently, the signature that has been analyse by most of researchers are not updated enough. This study presents a new signature and behaviours in detecting P2P botnets. The proposed detection module is based on rule-based approach. The result show that the proposed detection module have high detection accuracy with ability to detect known P2P botnets and produce a high detection rate with low false alarm rate. Hence, the developing detection module based on automated signature-based approach with updated dataset has been the most promising approach to fight against botnets threat in real P2P botnets files. The further work will be done on the developed of automated P2P botnets signature in different attack type and platform.

ACKNOWLEDGMENT

The authors would like to express the appreciation to Inforsnet Group Research of Universiti Teknikal Malaysia Melaka (UTeM) and MyBrain15 Programme by Ministry of Higher Education Malaysia (MoHE) in encouraging the authors to publish this paper. This work was supported by the MoE of Malaysia under Grants FRGS/1/2014/ICT04/FTMK/02/ F00212.

REFERENCES

[1] Zeidanloo, H. R. a. A., A.B. (2010), "Botnet Detection by Monitoring Similar Communication Patterns", *(IJCSIS) International Journal of Computer Science and Information Security Vol. 7(No. 3): 36-45*

[2] Deerman, J. 2012, Advanced Malware Detection through Attack Life Cycle Analysis: The Evolution of Malware, *ISC8 Secure*.

[3] Patrick, T., 2013. Advanced Malware Detection Through Attack Lifecycle Analysis, *ISC8 Secure.*

[4] Husin, J. 2009. ICT, Youth and Terrorism. International Conference on Youth and Terrorism Kuala Lumpur.

[5] Conovan, Q., Ye, N., Emran, S.M., Li, X. and Chen, Q. 2001. Statistical Process Control for Computer Intrusion Detection. *Proceedings of DARPA Information Survivability Conference and Amp; Exposition II DISCEX, IEEE* .

[6] Boonbox, 2009. Types of IT Security Threats and Their Consequences: White Paper Report. *Pacific Coast Information System PCIS.*

[7] Robert Richardson 2007. CSI Computer Crime and Security Survey, The 12th Annual Computer Crime and Security Survey.

[8] Charles, L. 2013. *Malware Threats in our Cyber Infrastructure*, Yogyakarta: Swiss German University.

[9] Sundaram, A. 1996. An Introduction to Intrusion Detection. *ACM Digital Library.*

[10] Razak, S., Zhou, M. and Lang, S. D. 2002. Network Intrusion Simulation Using OPNET. *Proceedings of the OPNETWORK2002*, pp. 1-5.

[11] Chandrashekar, J., Orrin, S., et al. (2009), "The Dark Cloud: Understanding and Defending against Botnets and Stealthy Malware", Intel Technology Journal Vol. 13 (Issues 2).

[12] Moon, S. S. and Kyeong, J. J., 2006. Alert Correlation Analysis in Intrusion Detection. *Proceedings of the 2nd International Conference Advanced Data Mining and Applications ADMA 2006,* pp. 1049–1056.

[13] Emmanuel, H., 2006. Experimental Validation and Analysis of an Intelligent Detection and Response Strategy to False Positives and Network Attacks. *Proceedings of the IEEE Intelligence and Security Informatics Conference*, pp. 711-714.

[14] Tjhai, G. C., Furnell, S. M., Papadaki, M. and Clarke, N. L., 2010. A Preliminary Two-Stage Alarm Correlation and Filtering System Using

SOM Neural Network and K-Means Algorithm. *Journal of Computers and Security*, Vol. 29, pp. 712-723.

[15] Subbulakshmi, T., Mathew, G. and Shalinie, D. S. M. 2010. Real Time Classification and Clustering of IDS Alerts Using Machine Learning Algorithms. *International Journal of Artificial Intelligence & Applications IJAIA,* Vol. 1, No.1, pp. 1-9.

[16] Song, J., Takakura, H., Okabe, Y. and Kwon, Y., 2011. Correlation Analysis Between Honeypot Data and IDS Alerts Using One-class SVM. *In Intrusion Detection Systems, pp. 173-192. InTech Open Access Publisher.*

[17] Cheung, S., Fong W. M. and Lindqvist, U. 2010. Modelling Multistep Cyber Attacks for Scenario Recognition. *Proceedings of the 3rd DARPA Information Survivability Conference and Exposition DISCEX III*, Vol. 1, pp. 284-292.

[18] Lin, H. H., Mao, C. H. and Lee, H. M. 2009. False Alarm Reduction by Weighted Score-Based Rule Adaptation Through Expert Feedback. *Proceedings of the 2nd International Conference on Computer Science and its Applications 2009 CSA 2009*, pp. 1-8.

[19] Yuanyuan, Z., H. Xin, et al. (2010), "Detection of Botnet using Combined Host-and Network-Level Information". *IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*

[20] He, J., Yang, Y., Wang X., and Zeng, Y., 2014. Peer Sorter: Classifying Generic P2P Traffic in Real-Time. *IEEE International Conference on Computational Science and Engineering.* pp. 605-613

[21] Ferragut, E.,M., Laska, J. and Bridges, R.A. 2012. A New, Principled Approach to Anomaly Detection, *11th International Conference on Machine Learning and Applications, IEEE.*

[22] Ngadi, M., A., Yazid, M. I., and Hanan, A., 2005. A Study on Advanced Statistical Analysis for Network Anomaly Detection. *Project Report, Faculty of Computer Science and Information System, Skudai, Johor.*

[23] Bao, X., Xu, T. and Hou, H. (2009), "Network Intrusion Detection Based on Support Vector Machine," International Conference on Management and Service Science (MASS).

[24] NetBeans Community, 2013. Welcome to the NetBeans Community. Available at: https://netbeans.org/about/ [Accessed on 14 March 2014]

[25] Leonardo, G. 2007, Reach Out with the IDE and Platform. *NetBeans Magazines.*

[26] He, J., Yang, Y., Wang X., and Zeng, Y., 2014. Peer Sorter: Classifying Generic P2P Traffic in Real-Time. *IEEE International Conference on Computational Science and Engineering.* pp. 605-613

[27] Karuppayah S., Fischer M., Rossow, C., and Max, M. 2014. On Advanced Monitoring in Resilient and Unstructured P2P Botnets. *IEEE ICC - Communication and Information Systems Security Symposium.* pp. 871-877