ISSN No. 0976-5697

Volume 8, No. 3, March – April 2017

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Peer-to-Peer Botnet Detection based on Bot Behaviour

Himanshi Dhayal
Department of Computer Engineering
DCRUST Murthal
Sonipat, India

Jitender Kumar
Department of Computer Engineering
DCRUST Murthal
Sonipat, India

Abstract: Peer-to-Peer (P2P) botnets are a significant threat to network security because of their distributed platform. The detection of these botnets becomes very difficult because of their decentralized nature and the situation worsens if an existing P2P network is exploited for botnetwork creation (parasite botnets). In this paper, we propose a two-tier detection framework to detect parasite P2P botnets. The approach can detect botnets in their waiting stage and without any requirement of bots' signature. For detection of bots, we have considered two features: (i) long-living peers, search requests' (ii) intensity and (iii) temporal correlated behaviour. The approach is able to detect bots from a monitored network with high detection accuracy.

Keywords: peer-to-peer; botnets; botnet detection

I. INTRODUCTION

In the last few years, internet security threat has become the major concern of the time. Botnet has been considered as the major threat to the internet security. The botnet attacks are considered as the root cause of various online security attacks such as Distributed Denial of Service (DDoS) attacks, bitcoin mining, click fraud and so on. Botnet, the word is made of two words robot and network. The word robot signifies all the computers which are recruited by various methods such as backdoors, spam mails, trojans etc. and, network is the interconnection of various such systems over the internet. So, a botnet is defined as a coordinated network of bots which are recruited over the internet for doing malicious activity. In this the botmaster controls the botnet and sends the command and control (C&C) messages to the recruited machines or bots to carry out a malicious activity.

Botnet are categorized as [1]: (i) centralized botnets and (ii) de-centralized botnets. The centralized botnets are further divided into two: (i) IRC (Internet Relay Chat)-based botnets and, (ii) HTTP (Hyper Text Transfer Protocol) - based botnets. IRC is an on-line text-based instant messaging protocol and works on client-server architecture [2]. IRC can connect hundreds of clients via multiple servers, clients can be contacted using one-to-many or one-to-one relationships. This feature makes the IRC very suitable for utilizing it to form and control a botnet. The HTTP-based botnets uses HTTP requests to carry out the malicious activities over the internet. However both these types of botnets suffer from single point of failure and thus can be easily taken down.

The decentralized botnets also known as peer-to-peer (P2P) botnets works on the distributed architecture. Because of this property these botnets are considered as most hazardous to computer security, by the researchers. These botnets are difficult to take down because of absence of any central attacker. The bots in P2P architecture can act as both client and server. Therefore the P2P botnets have been considered as the most promising next generation botnets.

Various poplar P2P botnet are: MIRAI botnet [3], [4], Slapper [5], Nugache [6], Storm [6-8] etc. These botnets have targeted many popular websites as their victims. The P2P botnet led security attacks has led to huge financial loss in last few years [9], [10], [11], [12]. The P2P botnets are categorized as: (i) Bot-only botnet and (ii) Parasite botnet. The bot-only botnet uses private network for the formation of botnet. These botnets do not have any benign peers and design their own customized network. Because of unique packet forms and protocols utilized for private network formation, these are easily detectable. The parasite botnet are considered as the most harmful botnets to the internet security. In parasite botnets, an existing peer-to-peer network is utilized for the botnet formation. The network includes both benign peers as well as peers recruited as bots. In this botnet, a bot can communicate with other bots of the network because of malicious behavior, as well as with benign peers because of the requests from its legitimate user and thus making detection hard. Their resemblance to exploited P2P network's benign traffic makes detection even harder.

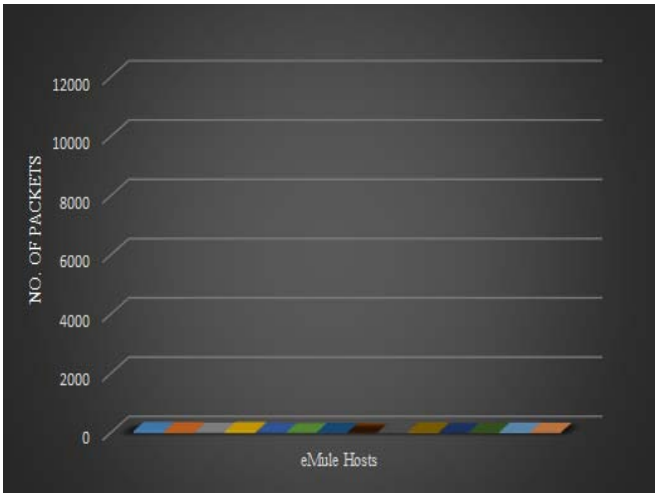
P2P networks hold a major share of internet across the continents. Various popular peer-to-peer networks example are: BitTorrent [13] and eMule [14]. These networks have become popular because of ease of resource sharing. These networks host a large number of files and therefore security has been breached. The attackers thus use the existing P2P networks for the botnet formation. The botmaster have full access to the exploited network's resources and therefore are able to carry out malicious activity at a very large scale.

Various detection techniques have been introduced over the years to detect these botnets, targeting various phases of a botnet cycle. The main problem occurs in the detection of parasite botnets because of their stealthy behavior. In this paper we have proposed a detection framework for the detection of such botnets.

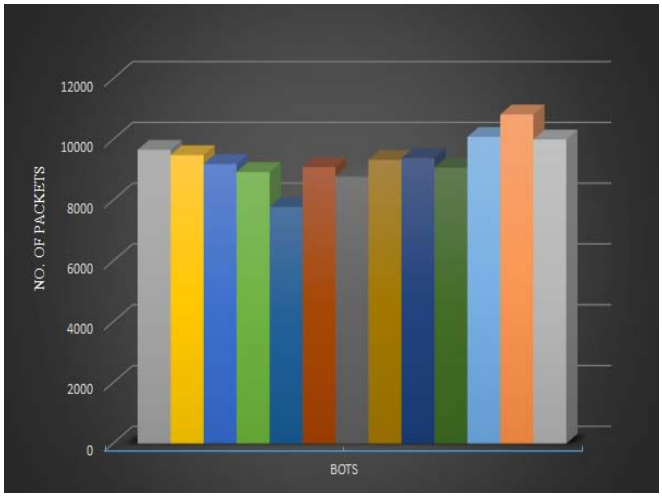
In next section, we discuss related work and their weaknesses. Section III consists of our approach and contribution. Section IV consists of dataset. Implementation details of the approach are discussed in Section V and Section VI presents the evaluation results of our detection approach. Conclusion and future work are discussed in Section VII.

II. RELATED WORK

Various detection techniques have been proposed for detection of P2P botnets. Various techniques are categorized as: (i) honeypot-based, (ii) anomaly-based and, (iii) signature-based detection techniques. Anomaly-based detection is further classified as host-based and network-based detection. All the techniques provide varying accuracy, but none of the technique is able to detect botnets with 100% accuracy.



(a) eMule peers



(b) Storm bots

Figure 1. Comparison of network traces of eMule and Storm

BotMiner [15] is a general botnet detection approach for both IRC and P2P botnets. It is an anomaly based detection approach and utilizes the concept of same communication patterns and malicious activities exhibited by bots. It uses correlation between these two features for identification of bots. Limitation of this technique is that it cannot detect stealthy botnets.

Coskun et al. [16] proposed botnet detection approach based on mutual contacts among the bots, but a seed bot has to be known in advance to start with i.e., to initiate the detection process.

Fan et al. [17] combined two different approaches for botnet detection: offline-based and online-based detection. In offline stage, they check for low connection success rate and similarity of communication traffic. During online stage, suspected machines from stage-I are monitored, if they are initiating any kind of malicious activities. Problem with this technique is, if bots are not in attack stage then cannot execute second stage, resulting in failure.

Entelechia [18] proposed by Hang et al. proposed by Narang et al. considered long-lived conversations and low-volume of data exchanged between bots for their detection. But the approach cannot detect bots if compromised machines along with malicious data also exchange high-volume benign data.

The work presented in [19] provides an overview of advantages and weaknesses of various recent P2P botnet detection techniques. The detection framework presented here is able to deal with almost all of the limitations of various detection approaches listed in [19] and with all the weaknesses of above mentioned approaches in context of P2P botnets and is able to detect bots from a parasite network with high accuracy.

III. PROPOSED DETECTION APPROACH

In P2P botnets, unlike centralized botnets there is no single attacker. Therefore, the detection of these botnets becomes comparatively more difficult. The situation aggravates if the P2P botnet is built on an existing peer-to-peer network. The detection of such botnets require either each bot infected machine be monitored individually to check for any kind of abnormal behavior, or classifying the network traffic either as benign or malicious. The former approach requires individual host analysis, which is difficult if the botnet is composed of large number of computers. The later approach however is more efficient for detection of such botnets.

The P2P botnet lifecycle [20] consists of three main stages:

- Infection stage: in this phase, hosts are recruited to be part of botnet army by infecting them with virus, Trojan horse or worms.
- Waiting stage: in this stage, bots wait for the commands from the botmaster.
- Execution stage: during this stage, coordinated attack is carried out as according to the commands from the botmaster.

The most advantageous way of detecting botnet is before its execution stage, that is, to detect the bots before they can start the attack or could carry any malicious activity. Early detection of bots could help in saving many million dollars and also prevents botnet from further expanding itself. So, in this paper we are proposing the approach which can detect bots before the execution stage. The approach also does not require any seed information about the botnet to initiate the detection. The main strength of the approach is that it can detect bots with high detection accuracy and least false positive rate.

The approach proposed for detection of peer-to-peer botnets in this paper is based on classification of network traffic. It requires the P2P network traffic be analyzed for any kind of abnormality and the nodes associated with the abnormal traffic are finalized as bots. The traffic from a P2P network which is also exploited for the botnet formation consists of both legitimate data from benign peers and malicious data from bots. The abnormality of the traffic lies in the special features that are exhibited by peers that are recruited as part of the botnet. The detection of such features and utilizing them for further classification of traffic finally leads to botnet detection.

For this detection purpose we have utilized some inherent features exhibited by the malicious traffic that are essential for botnet working. The features utilized are: (i) Bots' Lifetime in the P2P network, (ii) Search Request Intensities and, (iii) Time correlated behavior exhibited by the bots.

A. Peers' Lifetime

The first feature utilized is that the bots of a botnet remain part of the network for longer duration of time in anticipation of commands from the botmaster, as compared to the benign peers of the P2P network. Thus the suspected nodes are marked

from among the hosts on the basis of time duration a host is online, i.e., present on the P2P network.

B. Search Request Intensities

To carry attacks against a victim, the bots require commands from the botmaster as soon as possible. Therefore they send out search requests for commands in large number. Thus the second feature utilized is large number of search requests as compared to legitimate peers. The figure 1(a) and 1(b) shows the number of search requests send out by the legitimate peers and the bots of the Storm botnet, correspondingly, at the end of an hour. Thus the nodes are further selected as suspected from among all the peers of the monitored P2P network on the basis of abnormally high number of search requests.

C. Time Corelated Behavior

As bots send out search request in the network for the command files from the botmaster. The search requests sent out by the recruited machines of the botnetwork exhibit similar behavior, as all the bots work on the direction of same bot binaries. For final stage detection we utilize the temporal behavior correlation between the search request packets from the bots

IV. DATASET

We acquired the botnet dataset of Storm bots. The dataset is the same which has been used in the P2P botnet detection related research work of [21]. The data consists of malicious traffic of storm botnet and benign traffic from eMule peers. The data flows contains network traffic of various protocols such as HTTP, SMTP, TCP, UDP etc. and the normal P2P flows that contain legitimate P2P traffic of eMule.

Table I. Packets' Details

Experiment	#Number of Packets
Malicious P2P Data Packets	41941536
Benign Peers (eMule) Packets	25913400

Table 1 shows the number of malicious and benign packets of Storm Botnet and eMule peers, correspondingly, obtained from the authors of the dataset.

V. IMPLEMENTATION DETAILS

The proposed detection framework is shown in Fig. 2. And consists of following modules:

A. Raw Data Processing Module

In this module the raw data obtained is processed and the packets with zero payload are filtered out. The P2P botnet during the waiting stage, works with UDP packets only for searching commands. Therefore, the module selects only the UDP packets and sends them to the next module for further processing. For filtering purpose we have used filtering features available in Wireshark [22].

B. Filtering Module

The packets from the previous module are fed into the filtering module. The module processes packets from each host individually. The module only selects the IP addresses of the hosts that are active for the monitored time and marks them as suspected peers.

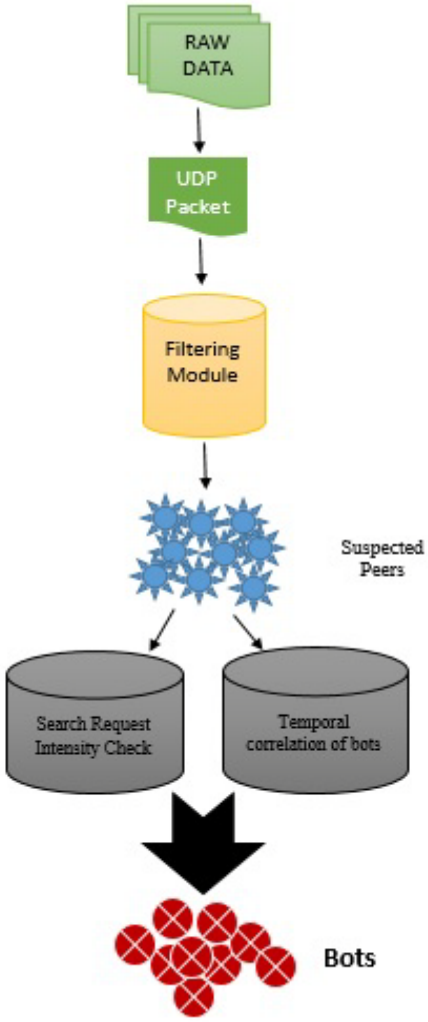


Figure 2. Detection Framework

If the packets from the host are present consistently during the monitored time period i.e., if have at most DUR_GAP time difference between their consecutive packets. The output of this field as depicted in the Fig. 2 is then fed into the search requests' intensity module.

The packets sent to next module consists of following attributes: (i) source IP address, (ii) destination address, (iii) port number, (iv) payload information and, (v) timestamp of the packet.

C. Bot Detection Module

The module consists of two sub-modules which enlists bots on the basis of their packets' behaviors.

Table II: Detection Results

Experiment#	#1	#2	#3	#4
#Known Bots	13	13	13	13
#Benign Peers (eMule)	21	27	30	10
#Bots Detected	13	12	13	13

The first sub-module selects and checks the total number of search requests from the peers enlisted in the previous module. It selects only those peers which show abnormally high number of search requests at the end of the time during which the P2P network is monitored. Such peers are selected and the search requests from such peers along with their temporal behavior are

forwarded to the next detection stage of the module. The second phase of the module identifies and correlates the temporal behavior of the search requests of all the peers. The peers showing higher affinity and correlation in the behavior are finally marked as bots.

VI. RESULTS

The proposed approach has been evaluated for performance of the detection framework. The raw data of botnet and eMule P2P application is extracted and used from the aforementioned dataset. The data is being pre-processed which filtered out unnecessary peers resulting in significantly reducing the number of hosts to be analyzed. The filtered data is then used to detect the malicious peers existing in the dataset collected.

The final detection results are shown in Table II. As the dataset consisted of data from Storm bots and eMule peers. The dataset is divided into sets of fixed duration each and each set is then fed into the detection framework for identification of the bots. As shown in the table, the approach is able to detect bots from all the sets with very high detection accuracy, resulting in detection of almost all the bots with negligible false positive rate.

VII. CONCLUSION

In this paper, we have presented a P2P botnet detection framework, which is able to classify peers from a monitored network as either malicious or not with very high accuracy. Even the classification accuracy of benign peers is also very high, resulting in overall classifying accuracy above 99%. In addition to this, our work is able to detect bots in their waiting stage itself and can even detect bots if they are exchanging benign data.

VIII. REFERENCES

[1] Ping Wang, Baber Aslam, Cliff C. Zou, "Peer to Peer Botnets", Handbook of Information and Communication Security(Springer), 2010, pp. 335-350.

[2] Saman Taghavi Zargar, James Joshi, David Tipper, "A Survey of Defense Mechanisms Against Distributed Denial of Service (DDoS) Flooding Attacks" In IEEE Communications Surveys & Tutorials-2013, Vol. 15, No. 4, pp. 2046-2069.

[3] https://en.wikipedia.org/wiki/Mirai_(malware)

[4] https://www.symantec.com/connect/blogs/mirai-what-you-need-know-about-botnet-behind-recent-major-ddos-attacks

[5] Iván Arce, Elias Levy, "An Analysis of the Slapper Worm", In Security & Privacy, IEEE, 2003, Vol. 1, No. 1, pp. 82 – 87.

[6] Samstover, Davedittrich, Johnhernandez, Andsvendietrich, " Analysis of the Storm and Nugache Trojans: P2P Is Here", In The USENIX Magazine (December 2007), Vol. 32, No. 6, pp. 18-27.

[7] Phillip Porras, Hassen Sa¨ıdi, Vinod Yegneswaran, "A Multi-perspective Analysis of the Storm (Peacomm) Worm", SRI International-2007.

[8] Thorsten Holz, Moritz Steiner, Frederic Dahl, Ernst Biersacky, Felix Freiling, "Measurements and Mitigation of Peer-to-Peer-based Botnets: A Case Study on Storm Worm", SRI report, 2008. Ting-Fang Yen, Michael K. Reiter, "Are your hosts trading or plotting? Telling p2p file-sharing and bots apart", In Distributed Computing Systems (ICDCS), 2010 IEEE 30th International Conference on. IEEE, pp. 241–252.

[9] http://www.symantec.com/connect/blogs/grapplingzeroaccessbotnet.

[10] https://www.incapsula.com/blog/2015-16-ddos-threat-landscape-report.html

[11] https://securelist.com/blog/research/70071/statistics-on-botnet-assisted-ddos-attacks-in-q1-2015/

[12] akamai's [state of the internet] / security Q3 2016 report

[13] BitTorrent, www.bittorrent.com.

[14] eMule, www.emule-project.net.

[15] G. Gu, R. Perdisci, J. Zhang, W. Lee, "BotMiner: Clustering analysis of network traffic for protocol- and structure-independent botnet detection", Proc. 17th USENIX Security Symp. (Security '08), San Jose, CA (2008) pp. 139–154.

[16] Coskun, B., Dietrich, S., Memon, N., "Friends of an enemy: identifying local members of peer-to-peer botnets using mutual contacts, In Proc. of ACSAC-2010, ACM, 2010, pp. 131-140.

[17] Yuhui Fan, Ning Xu, " A P2P Botnet Detection Method used On-line Monitoring and Off-line Detection", International Journal of Security and Its Applications, 2014, Vol.8, No.3, pp. 87-96

[18] Huy Hang, Xuetao Wei, Michalis Faloutsos, Tina Eliassi-Rad, "Entelecheia: Detecting P2P Botnets in their Waiting Stage", IFIP Networking Conference, 2013, pp. 1 – 9.

[19] Priyanka, Mayank Dave, "PeerFox: Detecting Parasite P2P Botnets in their Waiting Stage", In International Conference on Signal Processing, Computing and Control ISPCC, IEEE, 2015, pp. 350-355.

[20] Himanshi Dhayal, Jitender Kumar, "A Review of Botnet and Recent Peer-to-Peer Botnet Detection Techniques", In International Conference on Communication and Signal Processing ICCSP, IEEE, 2017(in press).

[21] Rahbarinia, R. Perdisci, A. Lanzi, and K. Li, "Peerrush: Mining for unwanted P2P Traffic," in Detection of Intrusions and Malware, and Vulnerability Assessment, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2013, vol. 7967, pp. 62–82.

[22] WireShark, https://www.wireshark.org