Journal of Computer Science

Original Research Paper

A Novel Botnet Detection System for P2P Networks

[1]Atef Ahmed Obeidat, [1]Majd Mahmoud Al-Kofahi,
[1]Mohammad Jazi Bawaneh and [2]Essam Said Hanandeh

[1]*Department of Information Technology, Al-Huson University College, Al-Balqa Applied University, Salt, Jordan*
[2]*Computer Information System, Zarqa University, Zarqa, Jordan*

Article history
Received: 5-7-2017
Revised: 27-7-2017
Accepted: 3-08-2017

Corresponding Author:
Atef Ahmed Obeidat
Department of Information Technology, Al-Huson University College, Al-Balqa Applied University, Salt, Jordan
Email: atefob@gmail.com

Abstract: Botnets remain an active security problem on the Internet and various computer networks. They are continuously developing with regard to protocols, structure and quality of attacks. Many botnet detection programs are currently available, but only few can detect bots in real-time. The sooner bots are detected the lesser damage they can cause. In this paper, a novel botnet detection system, is proposed to detect peer-to-peer bots. The system consists of three-phases filtering, P2P detection and P2P botnet detection phases. For the third phase, P2P network behavior analysis is performed to detect P2P bots. Experimental results showed that the system exhibits high average true positive rate and extremely low average false positive rate during botnet detection.

Keywords: P2P Networks, Bot Detection, Traffic Analysis, Real-Time, Temporal Groups

Introduction

P2P botnets are among the most common types of P2P malwares (Obeidat, 2016). Botnets are composed of many computers with high bandwidth and computing capabilities, which increase with time. The bot master node controls the other bots by initiating various activities such as, email spamming, distributed denial of service attacks, key-logging, Bitcoin mining, click-fraud scamming and password cracking.

Command-and-Control (C&C) communications in P2P botnets are executed through the exchange of files (resources) shared by nodes in a network. For example, the master node of a P2P botnet can create a file of commands and share it with the bots. Subsequently, the bot master periodically shares the file of the C&C with the bots. Notably, C&C communications are similar to the file download traffic for benign nodes. Thus, constructing a detection system capable of distinguishing the difference between benign and malicious nodes based on network traffic analysis is of great importance.

Numerous methods can provide metrics for the inference or differentiation between benign and malicious networks (Strayer *et al*., 2008; Zhao and Traore, 2012; Dillon, 2014). In this paper, a new metric is proposed based on the behaviors of P2P networks, where members exchange data repeatedly over different time intervals. In benign P2P networks, the repetition of uploaded or downloaded data is minimal. In malicious P2P networks, malicious peers share data several times. A set of characteristics is extracted from the network flow and then used to derive the new metric for the detection phase. These characteristics include timestamp, source and destination IP addresses, protocol and packet size. This metric is based on forming a group for each and every peer in the network, where each group contains all the peers that communicate with this peer either by sending or receiving packets from it. The flow behavior between members of each group are studied and analyzed separately in consecutive and short time intervals according to the following criteria: the rate of change in the size of the group through successive time windows, rate of change in the members forming the group through successive time windows and rate of change in the size of the data transferred between members of the groups through successive time windows. The contribution of this research involves the use of the rate of change in the size (RCS) of the group to distinguish benign peers from malicious ones.

In addition, the proposed system can be characterized by the following features:

- The system uses the behavioral features of the network traffic without the use of the payload in individual packets. Thus, it is not affected by encrypted traffic
- It doesn't require any training to give accurate results. Thus, it can detect a botnet in real time through an efficient approach that works along with a short detection time window

Science Publications

- In real time, the system accurately detects the presence of a bot activity during a significant part of its life during the C&C or attack phase

This paper is organized as follows: the related work section to classify study the related methods. The section of the proposed system discusses proposed method and presents its mechanism in detail. Then the experimental results are discussed and illustrated in following section. The conclusion is presented in the following section. Finally, the future work is discussed.

Review of Related Literature

Botnet detection remains an active research topic. Although many methods were suggested in literature, most of them cannot efficiently detect botnets. P2P botnet detection techniques can be broadly classified according to the type of detection method (Obeidat and Bawaneh, 2016). One such method is botnet detection based on flow analysis (Barthakur *et al*., 2013; Zhang *et al*., 2014). In this method, network flow between the nodes are studied. However, flow-based approaches have two key limitations. First, most of the flows between nodes belong to benign network processes. Second, the flow features must be calculated at runtime and flow analysis requires a high computational overhead at runtime in the absence of an efficient filter. Meanwhile, detection methods based on resource-sharing behavior monitoring (Rodríguez-Gómez *et al*., 2014) model the evolution of the number of peers sharing a resource in a P2P network. The limitation of these methods is its requirement to build a normality model of legitimate resources during the training phase. These resources do not necessarily contain all cases. Node-based detection (He *et al*., 2014; Yin, 2014) examines input and output flow for every node where the approaches aggregate behavioral metrics for each P2P node seen in network communications and use them to distinguish benign P2P hosts from those infected by P2P botnets. The key limitation of this solution is in the use of machine learning, which relies on learning a set of extracted features from real P2P botnets. Conversation-based detection (Dillon, 2014; Fan and Xu, 2014; Narang *et al*., 2014) does not rely on deep packet inspection or signature-based mechanisms. This approach requires a training phase to detect botnets. Thus, the use of new or unknown P2P applications cannot be detected because they do not belong to known classes.

Botnet detection methods based on flow analysis can be classified into two sets. The first set is based on payload inspection. In this set, the methods are usually resource intensive and slow because they require the analysis of big packet data. New bots also frequently utilize encryption and other methods to conceal communication and packet inspection. The second set is based on flow analysis. In this set, encrypted C&C channels are used.

The proposed method belongs to the second set and the following literature reviews the most recent works closely related to this method. In these studies, P2P botnets are detected by analyzing the behavioral characteristics of the network traffic (Saad *et al*., 2011; Zhao *et al*., 2012; Kheir and Wolley, 2013; Dillon, 2014; He *et al*., 2014; Almutairi *et al*., 2016).

PeerDigger (He *et al*., 2014) is a real-time system capable of detecting stealthy P2P bots. At the end of each time window, the system finds the set of destination IP addresses generated by each detected P2P host in Aggregation Flow (AF). The bot detection process is based on the Reconnection Number (RCN) of the AF. The RCN represents the number of repeated elements in the AF. The Reconnection Ratio (RCR) of Host (H) is defined as the maximum RCN of each AF and is used to determine whether the host is a bot. The problem with this approach is that it uses the maximum value of RCR for each AF at the end of each time window and neglects the relationship between consecutive time windows as a metric to identify the botnet network. In addition, the RCN is calculated by counting the number of destination IP addresses for each P2P host ignoring the received packets from other IP addresses to that host with a probability that this behavior is similar to that of normal networks. Thus, this metric cannot measure the temporal behavior of networks accurately. By contrast, the proposed method determines the P2P botnets by analyzing their network behaviors based on the RCS values between consecutive interval windows.

In 2014, the study (Dillon, 2014) on P2P bot detection within a local network was presented on the basis of the communications with the P2P overlay network of the P2P bots. The work used the NetFlow protocol to gain insight in all traffic within the network. The study analyzed and tested the behavior of Zeus as a P2P malware. Detecting this malware is based on either packet ratio (i.e., the sum of up packets divided by the sum of down packets) or traffic pattern. The experiment had limited access to the external network and with the limited data set, predictions cannot be made for results with real data.

The authors in (Kheir and Wolley, 2013) propose a system that detects active P2P bots through network analysis. Through the use of 1,317 distinct malware samples from eight malware families that communicate via P2P, a malware classifier is developed as part of the botnet detection system. P2P botnet traffic can be distinguished by three characteristics, namely, time, space and flow size. Using these characteristics, the authors used machine learning to differentiate P2P botnet traffic from benign P2P traffic with low FPRs. Their approach uses different characteristics with machine learning for botnet detection and thus greatly differs from our approach.

The Proposed system overcomes the previous limitations by analyzing traffic in real time without studying individual packets. The system analyzes network traffic in each phase, filtering out the unlikely flow along each step, so that the most computationally intensive analysis is done on a dramatically reduced traffic set. First, individual flows are subjected to a series of filters and classifiers to filter out as much traffic as possible. In this process, botnet traffic is cautiously prevented from being eliminated. The flows are then correlated with one another to determine the groups of flows that may be related and those that are parts of the same botnet. Finally, the detector module is examined for the presence of malicious networks based on the measurement of temporal node groups.

The Proposed System

The proposed system monitors the traffic in the network to analyze the flow in real time in order to reveal P2P botnets. The process of revealing malicious networks faces a major problem in the small differences between the behaviors of bots and benign networks. The process undergoes three phases, namely, filtering, P2P detection and botnet detection phases, as shown in Fig. 1.

In the first phase, the packets are filtered according to the transport layer protocol used. TCP and UDP traffic flows are extracted from the overall network traffic. The extraction filters out network flows that are unlikely to be generated by P2P network activities (Perényi *et al.*, 2006). Then the flow extraction phase translates the real-time packet stream into several flow streams. In the grouping phase, the flow stream for every host H is partitioned into several time windows of constant size T and a group is created for each H that contains all of the nodes that have communicated with this host. Using the P2P identification phase, the system detects whether H is involved in a P2P communication by checking the number of nodes in each group, which represents a P2P host when it has enough members. In the botnet detection phase, the system detects P2P botnets by analyzing their network behavior based on the RCS.

Filtering Phase

The goal of this phase is to filter out network flows that are unlikely to be generated by P2P network activities (He *et al.*, 2014). It consists of two stages. The first stage filters out only the TCP and UDP packets discarding data from other protocols because these protocols are mainly used as transport layer protocols to communicate and transfer data (Karagiannis *et al.*, 2004; Perényi *et al.*, 2006). The filter module keeps only TCP and UDP flows,because P2P application use them to exchange data. In addition, in the filter module eliminates flows that follow a successful DNS resolution, considering the data flow of non-P2P applications.

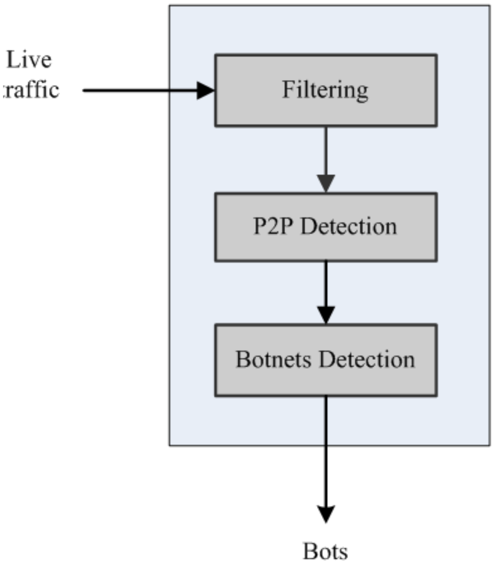


Fig. 1. System architecture

Most non-P2P applications typically need to resolve domain names before beginning flows. By contrast, members of P2P applications frequently join and leave the network and often contact one another directly by looking up IP addresses from a routing table without need to DNS requests. P2P members communicate directly by using IP addresses in the overlay network (Aberer and Hauswirth, 2002). Concluding these stages, a large portion of non-P2P network streams can be filtered, while retaining P2P network flows (see algorithm 1).

P2P Detection Phase

In this phase, all P2P types are detected before identifying P2P bots. The stream of packets represents a set of IP packets exchanged between two nodes. It is uniquely identified by the five-tuple set that contains the following information: protocol, source IP address, destination IP address, source port number and destination port number. These packets are generated by various P2P network activities, such as continuation of communication between network members, peer discovery, content request and data transmission.

All the nodes sending packets to a specific node Pj and all nodes receiving packets from that node as a group g(j) are considered. The peer Pj is considered as the master peer in g(j).

For every time windowi, a set of groups are captured and stored in vector $g_i(j)$=<T_l, P_i, R_p, S_p>, where T_i is the timestamp associated with the packet that belongs to a specific time window, P_j denotes the master node, which is the source or the destination for the packets within this time window, R_p represents the distinct addresses of the set of source nodes for the packets received by the master node P_j and S_p is the distinct addresses of the set of destination nodes for the packets sent by P_j. g_i as a proposed group composed of sender and receiver nodes that communicate with the master node. In this approach, the real-time packet stream generated by every host H can be translated into a set of groups $G(H) = \{g_i(j)\}$(see Algorithm 1).

P2P botnets communicate with each other without a C&C server. That is, P2P bots have a network behavior that is similar to those of benign P2P applications.

The system detects all P2P hosts by identifying the groups that present P2P network behaviors. To detect the group in real-time, the flow stream for every H is divided into time windows of constant size T according to the timestamp *Ti*. For every time window, a set of groups is extracted by H. These groups are denoted as $G(H) = \{g_i(j)\}$. At the end of each time window, the size of each group inG(H) is calculated. For each group $g_i(j)$, only distinct members are considered and the size of group j are denoted by $\delta_j = \delta_i(j)$ and $\delta(\mathrm{H}) = \{\delta_i(j)\}$. Groups with δ_j smaller than the threshold θ_δ are discarded and the remaining groups are considered as P2P groups that may represent a botnet or benign network (see Fig. 2). Thus, for each H, a set of groups G(H) can be extracted from a segment of the flow stream at the end of the time window, that is, $\mathrm{G(H)} = g_i = \{g_1,\ldots,g_m\}$. An H is considered as a P2P host when it generates at least one group (see lines 3-12 in Algorithm 1 in Fig. 3).

P2P Botnet Detection Phase

The goal of this phase is to identify malicious P2P or benign P2P networks (groups) resulting from the previous phase. Both types of networks share similar network behavior patterns. However, little differences exist between the two types because their goals in using the P2P protocol vary. The members of the Botnet groups must periodically recommunicate with the botmaster. That is, the group is constructed for a reasonably long time. The reasons for this situation are the following: First, P2P bots are likely to experience less peer churn than benign P2P members (Stutzbach and Rejaie, 2006). Second, most P2P bots store a list of known peers for bootstrapping itself into the botnet (Holz *et al*., 2008; Obeidat, 2016) and determining the number of peers communicating with them.

By contrast, benign P2P systems communicate with the master node, such as file-sharing systems and IPTV platforms, which are extremely dynamic because of the availability of the desired files and their short lifetimes (Aberer and Hauswirth, 2002). These features are expected of P2P bots that build groups containing bots that tend to terminate communication with the same botmasters.

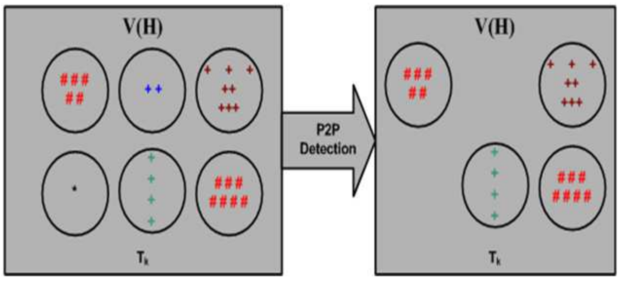


Fig. 2. P2P Detection Model. All groups with θδ≥2 are discarded

To exploit these features for the identification of P2P bots in real time, the members in each group generated by P2P hosts are counted at the end of a time window. For every group in G(H) extracted from the P2P H, the size of all groups of H $\delta(H)=\{\delta_{i,j}\}$ are calculated, where j=1,...,m and $\delta_{i,j}$ represent the size of the group G_j at time slicei. The extent of changes in the sizes of these groups $\Delta\delta(H)$ are calculated at time slice i for each P2P H, as shown in the following equation:

$$\Delta\delta(H)_i = \{\Delta\delta_{\delta i,1},\ldots,\Delta\delta_{i,m}\}$$

Where:

$$\Delta\delta_{i,j} = \delta_{i,j} - \delta_{i-1,j}$$

Then, we define the rate of change in the size of the group in time window i (RCS)i as:

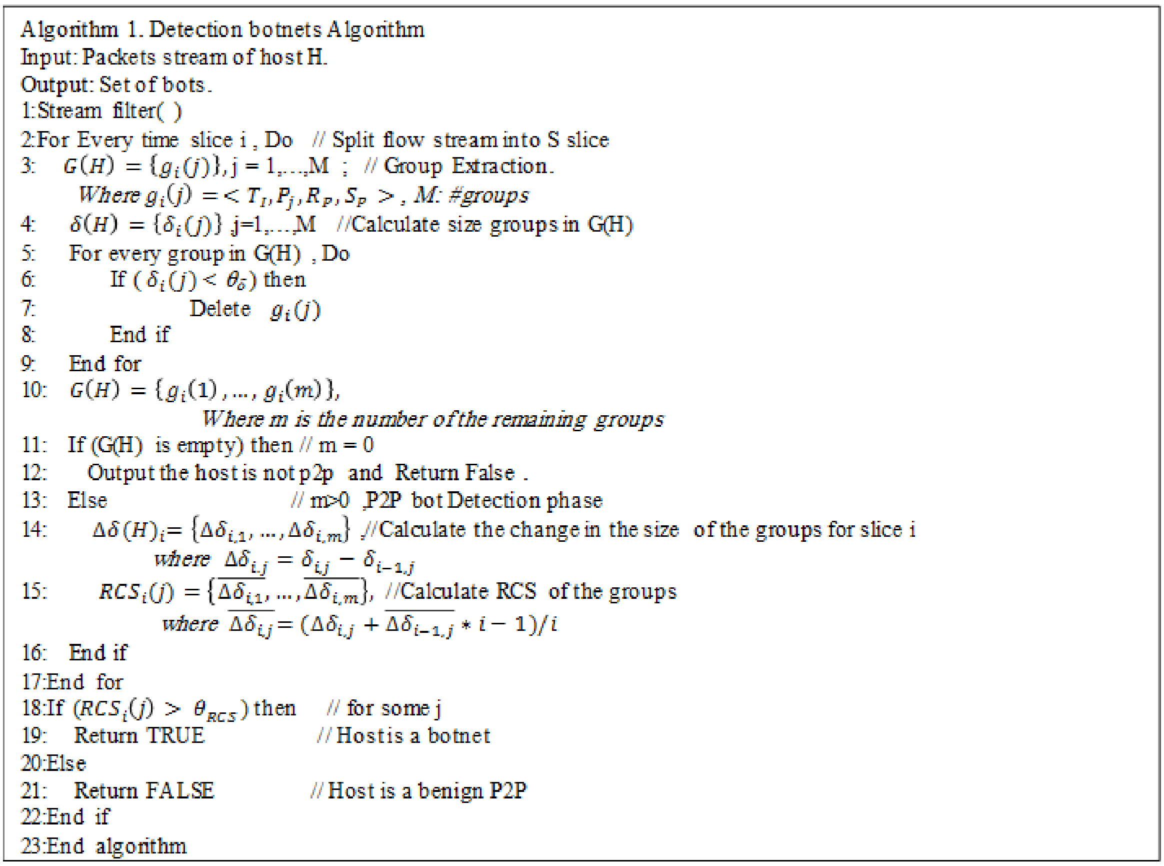
$$(\mathrm{RCS})_i = RCS_i = \{\overline{\Delta\delta_{l,1}},\ldots,\overline{\Delta\delta_{l,m}}\}$$

Where:

$$\overline{\Delta\delta_{l,j}} = (\overline{\Delta\delta_{i,j}} + \overline{\Delta\delta_{l-1,j}} * i - 1) / i$$

The system can provide an early decision at the end of the second time window according to the behavior of P2P traffic. The communication between members and botmasters are repeated. The members and botmasters then produce a positive rate of change in the number of group members. After computing the RCS for every detected P2P host, H is labelled as a P2P bot when the RCS is greater than or equal to a threshold θ_{RCS}.

This method is simple yet successful and does not require additional tools for detection, such as machine learning. The Proposed system provides a real-time bot detection mechanism that works well in high-traffic networks and its efficiency is due to the constant filtering of the data flow through all stages apart from the fact that it does not require data storage for more than two consecutive time windows.



```
Algorithm 1. Detection botnets Algorithm
Input: Packets stream of host H.
Output: Set of bots.
Stream filter( )
For Every time slice i , Do   // Split flow stream into S slice
    $G(H) = \{g_i(j)\}$, j = 1,…,M ;  // Group Extraction.
      $Where\ g_i(j) = < T_I, P_j, R_P, S_P >$ , $M$: #groups
    $\delta(H) = \{\delta_i(j)\}$ j=1,…,M   //Calculate size groups in G(H)
    For every group in G(H) , Do
        If ( $\delta_i(j) < \theta_\delta$ ) then
                Delete  $g_i(j)$
        End if
    End for
    $G(H) = \{g_i(1), \dots, g_i(m)\}$,
                  $Where\ m\ is\ the\ number\ of the\ remaining\ groups$
    If (G(H) is empty) then // m = 0
      Output the host is not p2p  and  Return False .
    Else                              // m>0 ,P2P bot Detection phase
        $\Delta\delta(H)_i = \{\Delta\delta_{i,1}, \dots, \Delta\delta_{i,m}\}$ ,//Calculate the change in the size  of the groups for slice i
              $where\ \Delta\delta_{i,j} = \delta_{i,j} - \delta_{i-1,j}$
        $RCS_i(j) = \{\overline{\Delta\delta_{i,1}}, \dots, \overline{\Delta\delta_{i,m}}\}$, //Calculate RCS  of the groups
              $where\ \overline{\Delta\delta_{i,j}} = (\Delta\delta_{i,j} + \overline{\Delta\delta_{i-1,j}} * i - 1)/i$
    End if
End for
If ($RCS_i(j) > \theta_{RCS}$) then   // for some j
    Return TRUE                 // Host is a botnet
Else
    Return FALSE                // Host is a benign P2P
End if
End algorithm
```

Fig. 3. Botnet detection algorithm

Experimental Results and Analysis

Dataset Collection

The experiment of the proposed system used a dataset of non-P2P traffic, dataset of P2P traffic generated by a variety of popular P2P applications and dataset of traffic from three famous P2P botnets. Table 1 summarizes the details of all datasets with respect to the duration of data capture, number of hosts involved and the size of the data collected.

Dataset of Non-P2P Traffic

Non-P2P traffic dataset collection involves the following processes: monitoring of the traffic crossing the campus network over the period of 1 day and collecting all packets from hosts not running P2P applications. The stream of packets contain a large number of general traffic from a variety of applications, such as web-browsing and email.

Dataset of P2P Traffic

The P2P traffic dataset is collected in a fully controlled network. Three of the common P2P applications are selected, namely, BitTorrent, eMule and Ares. An experimental local network is built in the campus such that it consists of four hosts capturing the network traffic generated by these hosts into the dataset.

Dataset of P2P Botnet Traffic

The dataset of P2P botnet traffic is obtained from a third party (Rahbarinia *et al*., 2013). This dataset includes a five-hour trace of Waledac, which contains three bots; a 24-hour trace of Zeus, which contains one bot; and a 6.15-hour trace of Neris, which also contains one bot. Table 1 summarizes these traffic datasets.

The three dataset types were merged together into a single dataset to construct a strong experimental dataset. The proposed system is tested using different lengths of time windows and the performance is discussed for every case in the subsequent sections.

Evaluation of P2P Host Detection

Flows from the P2P network can lead to relatively large groups, while unrelated flows can form smaller groups. So, the threshold value has a very important role in detecting P2P hosts. To achieve a high TPR while keeping the FPR low, the value of should be selected carefully.

Table 1. Traffic datasets

Category	Application	Duration	Number of Hosts	Size
Non-P2P dataset	Web, emails,...etc	5 hours	18	502 MB
Benign P2P Dataset	BitTorrent	5 hours	4	1.1GB
	eMule	5 hours	4	
	Ares	5 hours	4	
P2P	Waledac,	5 Hours	3	43 MB
Botnet	Zeus,	24 hours	1	3 MB
dataset	Neris	6.15 hours	1	56 MB

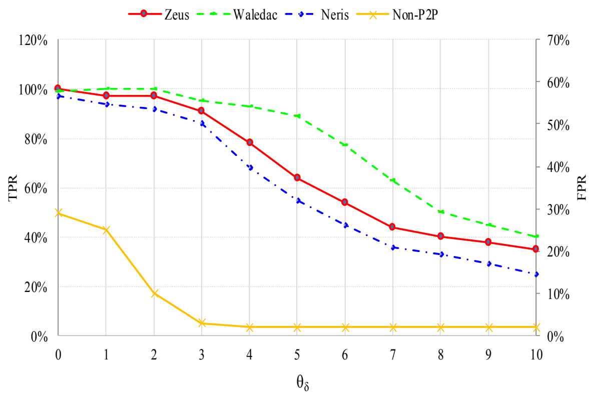


Fig. 4. TPR and FPR of P2P host detection for different values of θ_δ

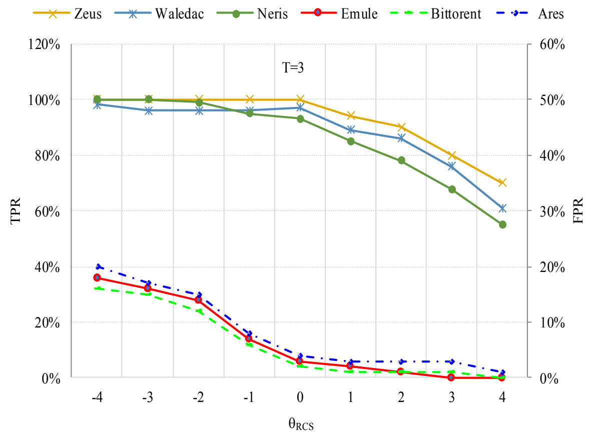


Fig. 5. TPR of P2P bot detection and FPR of benign P2P detection for different values of θ_{RCS} =0

Separately assigning different values to, ranging from 0 to 10. To determine the best value of threshold, the P2P host detection is applied for these values. The results with respect to the TPR and FPR are explained in Fig. 4. In this phase, the hosts within the experimental dataset are classified into two categories, namely, the positive category, which represents a P2P host (either benign or malicious) and the negative category, which represent a non-P2P host. As shown in Fig. 4, small values of results in high TPR values, but provide also worst FPR value.

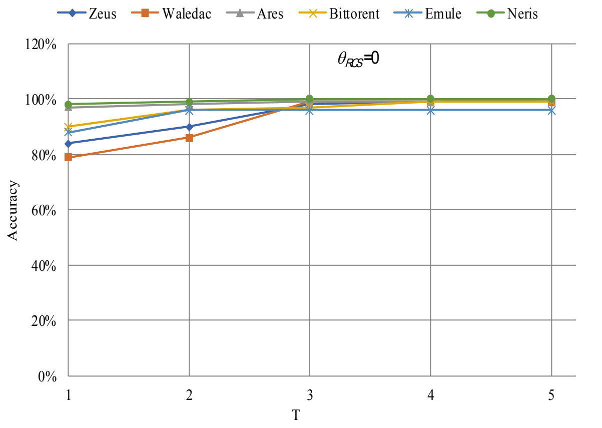


Fig. 6. Accuracy of P2P detection for different time window lengths

By contrast, with a high value leads to low TPR and FPR values. The best results are obtained when = 3 and the average TPR is 91% and FPR is 3% for T of 3 min.

Evaluation of P2P Botnet Detection

To evaluate the effectiveness of the differences between benign P2P hosts and bots, different values of θ_{RCS} ranging from -4 to 4 are investigated and the results are shown in Fig. 5. In this phase, the positive category consists of one Neris bot, one Zeus bot and three Waledac bots, whereas the negative category consists of four hosts that only run benign P2P applications. As seen from the curves, the system has a high TPR value when θ_{RCS} is small. However, the FPR values for BitTorrent, Emule and Ares are extremely low at the same θ_{RCS} values. When θ_{RCS} =0, the average TPR is 97.0% for botnet networks and average FPR is 3.0% for benign networks at T of 3 min.

The curves in Fig. 6 represent the accuracy in detecting P2P hosts running benign P2P and botnet applications for different time windows where T= 1 to 5 min. The accuracy of detection is proportional to T reaching 100% for some applications and starts giving the best results for all applications at T=3.

Comparing the New Method and other Similar Works

The experimental results show that the proposed method in this paper has satisfied good results that are better than many other methods found in the same field.

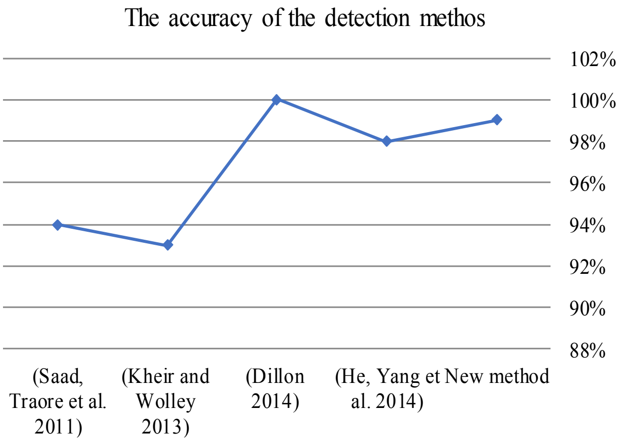


Fig. 7. A comparison between the new method and related work for the accuracy of detection of P2P bots

Figure 7 shows a comparison between the new method and some previous works in terms of bot setection accuracy (Saad *et al*., 2011; Kheir and Wolley, 2013; Dillon, 2014; He *et al*., 2014).

Conclusion

In this paper, a novel system to detect P2P bots within a monitored network through traffic analysis is proposed. It first detects all hosts engaged in P2P communication based on the size of the groups which present P2P applications then identifies P2P bots among the detected P2P hosts based on the rate of change in the size of the groups. The strength of the system lies in the following features: The system is not affected by encrypted traffic because it does not rely on payload data. Second, it is simple such that it does not involve the use of complicated statistical features or sophisticated algorithms. Third, the system does not undergo any training phase and thus detect bots in real time. Fourth, it can detect bots during the C&C or attack phase. Finally, its results are superior to those of other similar work.

The evaluation results demonstrated that the proposed system can detect P2P hosts with an average TPR of 99% and average FPR of 1.45%. While, P2P bots can be identified with an average TPR in the range of 62-99% and an average FPR in the range of 18-0.003% for different values of θ_{RCS}.

Future Work

The current approach has several limitations, which we intend to resolve in our future work. Given that the results obtained are based on the availability of existing malicious data, the experiments must be developed to include more types of P2P applications to produce more realistic results. Strengthening the bot detection model based on other factors is possible with respect to the rate of change between members of the groups and RCS of the data transferred between these members. These concerns may be addressed by developing a hybrid botnet detection system that utilizes two factors in addition to the current factor used in the bot detection model.

Acknowledgment

This research is supported by Dept. of Information Technology, Al-huson University College, Al-Balqa Applied University.

Author’s Contributions

Dr. Atef Ahmed Obeidat: Proposed the main idea of Botnet detection. Compared between ideas of related works and the suggested one. Found out the similiarites and differences between ideas and the suggested idea. Participated in carrying out the modeling stage (analysis and design). Participated in writing introduction and the propsed system sections.

Dr. Majd Mahmoud Al-Kofahi: Participated in palnning for the required cost, effort and time for building the detection system. Participated in carrying out the modeling stage (analysis and design). Evaluated and depolyed the constructed system. Wrote absatrct and results with analysis sections.

Dr. Mohammad Jazi Bawaneh: Constructed and converted the desgin of idea to real system by using visual C# langauge and SQL server. Tested the validation and verification of system. Participated in writing introduction and the propsed system sections.

Dr. Essam Said Hanandeh: Collected the studies in this area and summerized them in simple form. Participated in palnning for the required cost, effort and time for building the detection system. Wrote related work and references section.

Ethics

This article is original and contains unpublished Material, the corresponding author confirms that no ethical issues involved.

References

Aberer, K. and M. Hauswirth, 2002. An overview of peer-to-peer information systems. WDAS.

Almutairi, S., S. Mahfoudh and Jalal S. Alowibdi, 2016. Peer to peer botnet detection based on network traffic analysis. Proceedings of the 8th IFIP International Conference on New Technologies, Mobility and Security, Nov. 21-23, IEEE Xplore press, Larnaca. DOI: 10.1109/NTMS.2016.7792467

Barthakur, P., M. Dahal and M.K. Ghose, 2013. An efficient machine learning based classification scheme for detecting distributed command and control traffic of P2P botnets. Int. J. Modern Education Comput. Sci., 5: 9-18. DOI: 10.5815/ijmecs.2013.10.02

Dillon, C., 2014. Peer-to-peer botnet detection using netflow.

Fan, Y. and N. Xu, 2014. A P2P botnet detection method used on-line monitoring and off-line detection. Int. J. Security Its Applic., 8: 87-96. DOI: 10.14257/ijsia.2014.8.3.10

He, J., Y. Yang, X. Wang, C. Tang and Y. Zeng, 2014. PeerDigger: Digging stealthy P2P hosts through traffic analysis in real-time. Proceedings of the IEEE 17th International Conference on Computational Science and Engineering, Dec. 19-21, IEEE Xplore press, China. DOI: 10.1109/CSE.2014.283.

Holz, T., M. Steiner, F. Dahl, E. Biersack, F. Freiling *et al.*, 2008. Measurements and mitigation of peer-to-peer-based botnets: A case study on storm worm. LEET 8: 1-9.

Karagiannis, T., A. Broido, M. Faloutsos and K. Claffy, 2004. Transport layer identification of P2P traffic. Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement, Oct. 25-27, ACM, Italy, pp: 121-134. DOI: 10.1145/1028788.1028804

Kheir, N. and C. Wolley, 2013. BotSuer: Suing Stealthy P2P bots in Network Traffic through Netflow Analysis. Cryptology and Network Security, Springer, pp: 162-178.

Narang, P., C. Hota and V.N. Venkatakrishnan, 2014. PeerShark: Flow-clustering and conversation-generation for malicious peer-to-peer traffic identification. EURASIP J. Inform. Security, 2014: 1-12. DOI: 10.1186/s13635-014-0015-3

Obeidat, A.A. and M.J. Bawaneh, 2016. Survey of the P2P botnet detection methods. Int. J. Emerging Trends Technology Computer Sci., 5: 12-23.

Obeidat, A.A., 2016. Analysis the P2P botnet detection methods. IPASJ Int. J. Comput. Sci., 4: 1-11.

Perényi, M., T.D. Dang, A. Gefferth and S. Molnár, 2006. Identification and analysis of peer-to-peer traffic. J. Communications, 1: 36-46.

Rodríguez-Gómez, R.A. and G. Maciá-Fernández, R.A. Rodríguez-Gómez, M. Steiner and D. Balzarotti, 2014. Resource monitoring for the detection of parasite P2P botnets. Computer Netw. 70: 302-311. DOI: 10.1016/j.comnet.2014.05.016

Saad, S., I. Traore, A. Ghorbani, B. Sayed and D. Zhao *et al.*, 2011. Detecting P2P botnets through network behavior analysis and machine learning. Proceedings of the 9th Annual International Conference on Privacy, Security and Trust, July, 19-21, IEEE Xplore press, Canada. DOI: 10.1109/PST.2011.5971980.

Strayer, W.T., D. Lapsely, R. Walsh and C. Livadas, 2008. Botnet Detection based on Network Behavior. Botnet Detection, Springer, pp: 1-24.

Stutzbach, D. and R. Rejaie, 2006. Understanding churn in peer-to-peer networks. Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement (SCIM'06), ACM.

Yin, C., 2014. Towards accurate node-based detection of P2P botnets. Scientific World J., 2014: 1-10. DOI: 10.1155/2014/425491

Zhang, J., R. Perdisci, W. Lee, X. Luo and U. Sarfraz, 2014. Building a scalable system for stealthy p2p-botnet detection. IEEE Trans. Information Forensics Security, 9: 27-38. DOI: 10.1109/TIFS.2013.2290197

Zhao D., I. Traore, A. Ghorbani, B. Sayed and S. Saad et al., 2012. Peer to Peer Botnet Detection based on Flow Intervals. Information Security and Privacy Research, Springer, pp: 87-102.

Zhao, D. and I. Traore, 2012. P2P Botnet detection through malicious fast flux network identification. Proceedings of the 7th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, Nov. 12-14, IEEE Xplore press, Canada. DOI: 10.1109/3PGCIC.2012.48

Rahbarinia, B., R. Perdisci, A. Lanzi and K. Li, 2013. Peerrush: Mining for unwanted p2p traffic. Proceedings of the International Conference on Detection of Intrusions and Malware and Vulnerability Assessment, Jul. 18-19, Germany, pp: 62-82.