Enhanced Peer-to-Peer based Botnet Detection Method with Intrusion Free Network Scheme

Megha Godage
M.E CSE
N B Navale Sinhgad College of Engg. Solapur

A. A. Phatak
Assistant Professor
(HOD CSE)
N B Navale Sinhgad College of Engg. Solapur

R. S. Dayama
Assistant Professor (CSE)
N B Navale Sinhgad College of Engg. Solapur

ABSTRACT
Peer-to-Peer [P2P] networks usually has the ability to perform bidirectional communication efficiently, which means both the sending and receiving end has the same transactional power and ability to communicate with one and another. In the modern world, lots of misuses occurred via network schema; now-a-days, most of the malicious activities are held via this kind of P2P data sharing mechanisms, which are held by Botmasters. The botmasters acquire the usage of P2P network and utilize it for their own purposes and make this network to act like malicious one and took the effortness of others. Lots of existing approaches are available, but all are having certain limitations, so a new botnet detection scheme is required to resolve these issues and save the network from this kind of activities, which is called "Scalable Botnet Detection Mechanism". This Scalable Botnet Detection Mechanism initiates its first activity by means of finding the connected systems into the network and make the summary of it. The next step is to manipulate the profile-handling of P2P traffic-estimations as well as classify the P2P botnet traffic and legitimate P2P traffic. Simultaneously this system efficiently identifies the performance scenario of the P2P network and makes the system more scalable in further processing. The experimental results proves that our proposed approach is producing the result with more accuracy as well as more scalable than past schemes.

Keywords
Botnet Detection, Scalable Botnet Detection Scheme, Peer-to-Peer, P2P, Network Intrusion Avoidance, Botnet Master.

1. INTRODUCTION
The group of affected or conciliation machines are handled by Botnet, which is activated and manipulated by the intruder/attacker from remote-environments via Command User Interface [CUI] methodologies. Botnets function the infrastructures accountable for a range of cyber-crimes, like spamming, distributed denial-of- service [DDoS] attacks, fraud, click fraud, etc. The Command-and-Control channel is an important element of a botnet as a result of botmasters has confidence the Command-and-Control channel to issue commands to their bots and receive data from the compromised machines[1],[3],[4].

Botnets could structure their Command-and-Control channels in numerous ways in which. During a centralized design, all bots during a botnet contact one (or a few) Command-and-Control server(s) in hand by the botmaster. However, a elementary disadvantage of centralized Command-and-Control servers is that they represent one purpose of failure [2]. so as to beat this downside, botmasters have recently began to build botnets with a a lot of resilient Command-and-Control design, employing a peer-to-peer (P2P) structure or hybrid P2P/centralized Command-and-Control structures. Bots happiness to a P2P botnet kind associate degree overlay network within which any of the nodes (i.e., any of the bots) will be utilized by the botmaster to distribute commands to the opposite peers or collect data from them. Notable samples of P2P botnets area unit described by Nugache, Storm, Waledac, and even Confiker, that has been shown to plant P2P capabilities. Storm and Waledac area unit of specific interest as a result of they use P2P Command-and-Control structures because the primary thanks to organize their bots [5].

Whereas a lot of advanced, and maybe a lot of expensive to manage compared to centralized botnets, P2P botnets supply higher resiliency against take-down efforts (e.g., by law enforcement), since albeit a big portion of bots during a P2P botnet area unit non-continuous the remaining bots should still be able to communicate with one another and with the botmaster detection botnets is of nice importance. However, coming up with an efficient P2P-botnet detection system is sweet-faced with many challenges [6], [7].

First, the P2P file-sharing and communication applications, like Bittorrent, emule, and skype, area unit very talked-about and thence Command-and-Control traffic of P2P botnets will simply mix into the background P2P traffic. This challenge is additional combined by the very fact that a bot-compromised host could exhibit mixed patterns of each legitimate and botnet P2P traffic (e.g., attributable to the beingness of a file-sharing P2P application and a P2P larva on a similar host) [5], [8], [10].

Second, fashionable botnets tend to use more and more furtive ways in which to perform malicious activities that area unit extraordinarily laborious to be discovered within the network traffic. For instance, some botnets send spam through giant well-liked webmail services like Hotmail, that is probably going clear to network detectors attributable to coding and overlap with legitimate email use patterns [11], [13], [14].

Third, because the volume of network traffic grows speedily, the deployed detection system is needed to method a large quantity of knowledge with efficiency. To date, a number of approaches capable of detection P2P botnets are projected [12], [15].

However, these approaches cannot address all the said challenges. for instance, BotMiner identifies a gaggle of hosts as bots happiness to a similar botnet if they share similar

communication patterns and meantime perform similar malicious activities, like scanning, spamming, exploiting, etc. Sadly, the malicious activities is also furtive and non-observable, thereby creating BotMiner ineffective. Additionally, BotMiner's quantifiability is considerably strained. Yen et al. projected associate degree algorithmic program that aims to differentiate between hosts that run legitimate P2P file sharing applications and P2P bots. Nonetheless, this algorithmic program doesn't take into consideration the very fact that a larva could exist with a legitimate P2P application on a similar host. As a consequence, the mixed traffic profile of the compromised host could disguise the communication patterns associated with the larva, rendering the algorithmic program ineffective [3].

Bot Grep analyzes network flows collected over multiple giant networks (for instance. ISP networks), and tries to discover P2P botnets by analyzing the communication graph fashioned by overlay networks. Though BotGrep doesn't have confidence malicious activities for detection, it needs a worldwide read of web traffic and a priori detection results from extra systems to bootstrap the detection. However, it's extraordinarily laborious to amass such data in observe [17]. During this paper, we have a tendency to gift a unique ascendible botnet detection system capable of detection furtive P2P botnets. we have a tendency to seek advice from a furtive P2P botnet as a P2P botnet whose malicious activities might not be evident within the network traffic. significantly, our system aims to discover furtive P2P botnet albeit P2P botnet traffic is overlapped with traffic generated by legitimate P2P applications running on a similar compromised host and ii) attain high quantifiability [16], [19].
To the current finish, our system identifies P2P bots among a monitored network by detection the Command-and-Control communication patterns that characterize P2P botnets, notwithstanding however they perform malicious activities in response to the botmaster's commands [18].

Specifically, it derives applied math fingerprints of the P2P communications generated by P2P hosts and leverages them to differentiate between hosts that area unit a part of legitimate P2P networks (for instance., file sharing networks) and P2P bots. The high quantifiability of our system stems from the parallelized computation with delimited machine quality [20].

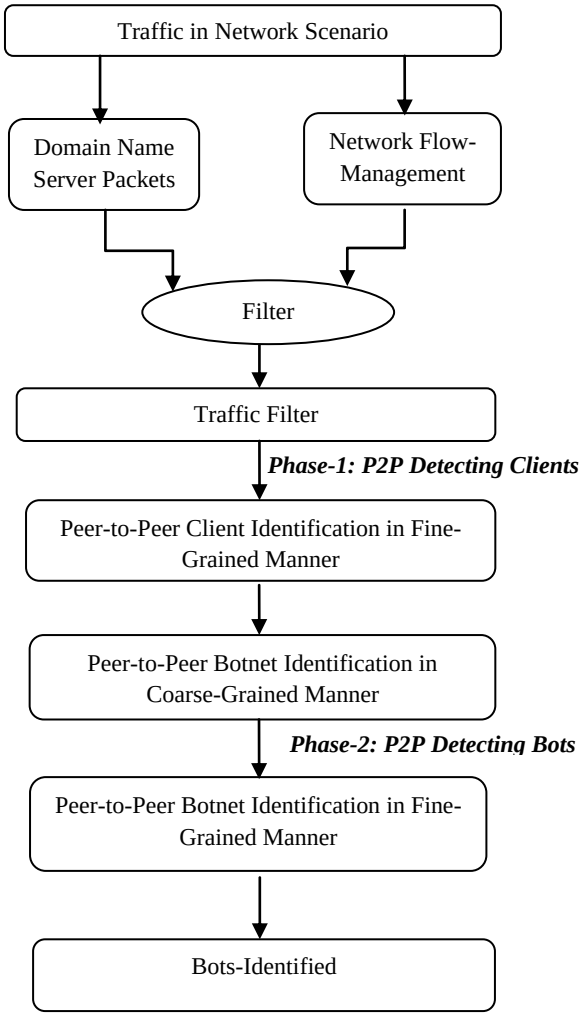

Fig.1. System Design Perspective

1.1. System Contributions

The contributions of the proposed approach are described as below:

(I) A replacement flow-clustering-based analysis approach to spot hosts that have interaction in P2P communications.

(II) Associate economical algorithmic rule for P2P traffic identification, wherever we have a tendency to build applied math fingerprints to profile numerous P2P applications and estimate their active time.

(III) A P2P botnet finding methodology which will effectively detect concealed P2P bots although the P2P botnet traffic is overlapped with traffic generated by legitimate P2P applications (for instance, Skype) running on an equivalent compromised machine.

(IV) A ascendable style supported associate economical detection algorithmic rule and parallelized computation.

(V) A example system and in depth analysis supported real-world network traffic, that has incontestable high detection accuracy (that is, a detection rate of 100% and zero.2% false positive rate) and nice quantifiability (that is, process eighty million flows in zero.8 hour) of our style.

2. RELATED STUDY

In the year of 2004, the authors "T. Karagiannis, A. Broido, M. Faloutsos, and K. Claffy" proposed a paper titled "Transport layer identification of P2P traffic" in that they described such as: Since the rise of shared (P2P) organizing in the late '90s, P2P applications have duplicated, developed and set up themselves as the main `growth application' of Internet movement workload. As opposed to original P2P systems which utilized all around characterized port numbers, current P2P applications can mask their reality using self-assertive ports. Therefore, solid appraisals of P2P activity require examination of bundle payload, a methodological landmine from lawful, protection, specialized, calculated, and financial points of view. For sure, access to client payload is frequently rendered unthinkable by one of these components, repressing reliable estimation of P2P activity development and elements. In this paper, we build up an orderly procedure to recognize P2P streams at the vehicle layer, i.e., in light of association examples of P2P systems, and without depending on bundle payload [5], [7].

The proposed approach is the main strategy for describing P2P activity utilizing just learning of system elements as opposed to any client payload. To assess the proposed approach the author build up a payload strategy for P2P movement distinguishing proof, by figuring out and breaking down the nine most well known P2P conventions, and show its viability with the revelation of P2P conventions . The outcomes show that P2P activity keeps on developing unabatedly, in spite of reports in the well known media.

In the year of 2008, the authors "T. Holz, M. Steiner, F. Dahl, E. Biersack, and F. Freiling" proposed a paper titled "Measurements and mitigation of peer-to-peer-based botnets: A case study on storm worm", in that they described such as: Botnets, i.e., systems of traded off machines under a typical control framework, are usually controlled by an assailant with the assistance of a focal server: all bargained machines interface with the focal server and sit tight for charges. In any case, the main botnets that utilization shared (P2P) systems for remote control of the bargained machines showed up in the wild as of late. In this paper, we acquaint a procedure with break down and alleviate P2P botnets. For a situation think about, we look at in detail the Storm Worm botnet, the most broad P2P botnet right now engendering in nature [11].

We could penetrate and dissect inside and out the botnet, which enables us to assess the aggregate number of traded off machines. Moreover, we exhibit two diverse approaches to disturb the correspondence channel amongst controller and traded off machines with a specific end goal to alleviate the botnet and assess the viability of these components.

In the year of 2007, the authors "G. Bartlett, J. Heidemann, C. Papadopoulos, and J. Pepin" proposed a paper titled "Estimating P2P traffic volume at USC", in that they described such as: With the ascent of distributed (P2P) record sharing appli-cations there has been an expanding enthusiasm for under-standing the fame and utilization of P2P. In this investigation, we take a gander at P2P use on the University of Southern California's grounds organize all through a 14-hour time span. We measure the volume of movement from P2P action and additionally the quantity of grounds IPs associated with P2P at USC. Since port-coordinating methods regularly come up short for P2P applications, we evaluate activity in light of both port-based and association design based techniques [12].

The system shouldn't approach bundle information thus these measures give just limits on P2P activity. Moreover, while distinguish P2P sharing, system cannot remark the sorts of information being shared (either music or information, confined or uninhibitedly accessible). It is found that 3– 13% of dynamic IPs on grounds take an interest in P2P, and that this activity represents 21– 33% of the bytes exchanged to and from our grounds.

3. RESULT AND ANALYSIS

The following table illustrates the performance of various implemented applications over P2P network.

TABLE-I. Analysis of Various P2P Network Schemas

Schemes	Duration	Iteration	Accuracy
Bittorent	14hrs	55	96.85%
Limewire	18hrs	58	96.99%
Emule	24hrs	65	96.97%
Skype	12hr	72	96.93%
Ares	15hrs	29	96.99%
Scalable Botnet Finding	14.5hrs	28	97.89%

The figure 2 illustrates the strategies of authentication levels of administrator over Peer-to-Peer network scenario.

Fig.2. Authentication Scenario of Administrator over P2P Network.

The figure 3 illustrates the Coarse Grained Peer-to-Peer Botnet Detection scenario and the figure 4 clearly explains the graphical view of Domain Analysis with the strategy for verifying the identified Botnets over the home page of the administrator.

The figure 5 shows the clear output nature of this system, that is the removal of Botnets from the network scenario.

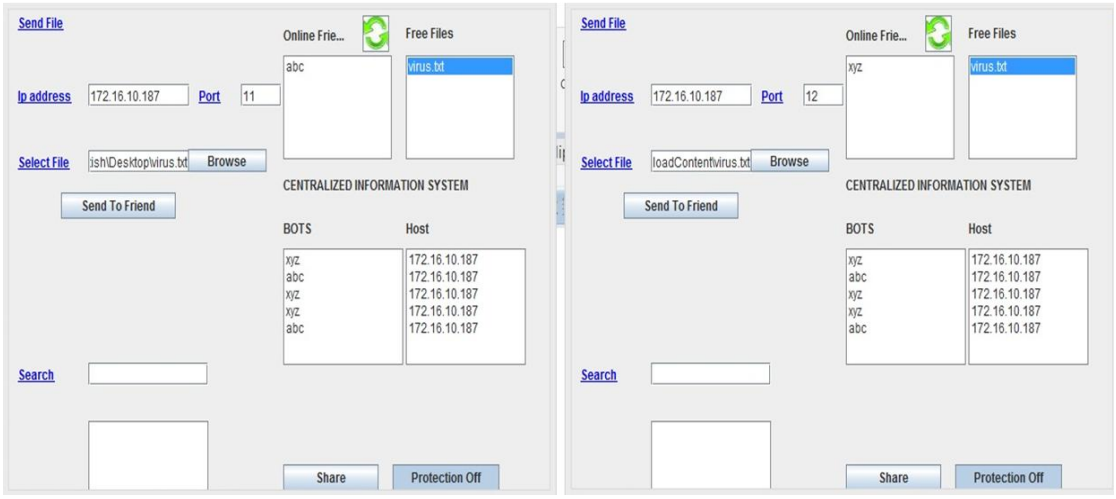

Fig.3.Coarse Grained Peer-to-Peer Botnet Detection scenario

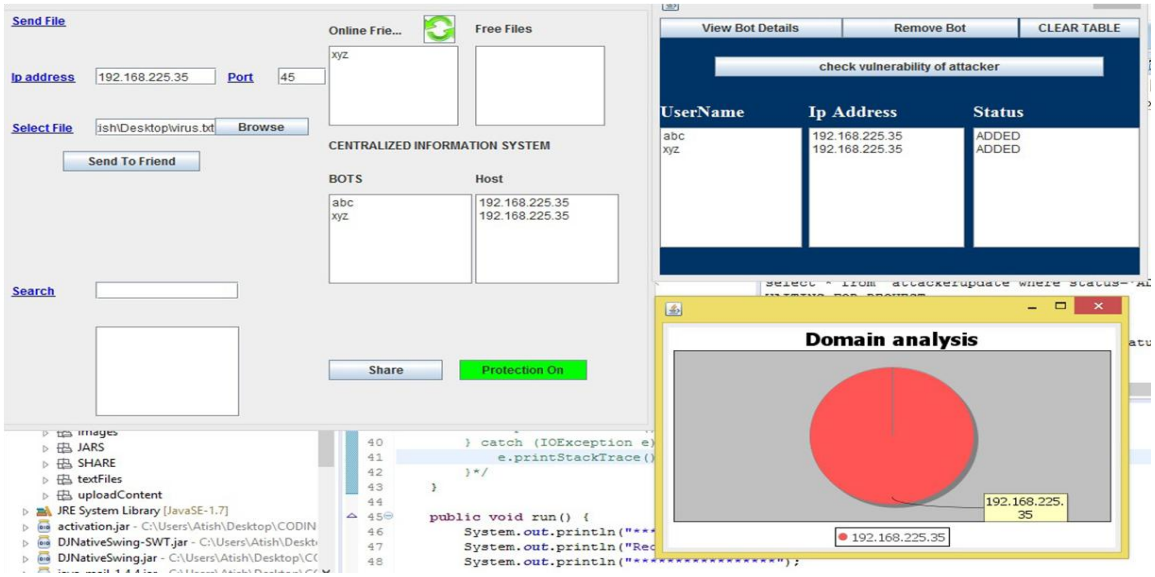

Fig.4. Domain Analysis with the strategy for verifying the identified Botnets over the home page of the administrator

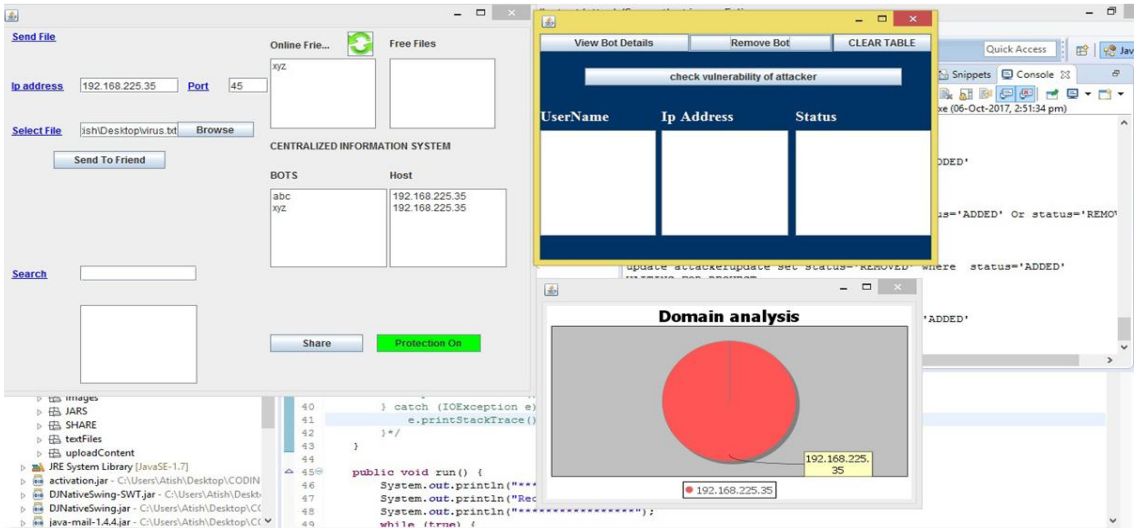

Fig.5. Removal of Botnets from the Network Scenario

4. CONCLUSION

The iterative botnet identification framework is proposed, that can recognize Peer-to-Peer botnets efficiently, and whose noxious approaches may not be discernible. To achieve this undertaking, a measurable fingerprints of the Peer-to-Peer interchanges are determined to first recognize Peer-to-Peer network scenarios as well as next-step of precedence is recognizing those activities, that are a piece of honest to goodness Peer-to-Peer systems (for instance, file sharing systems) and Peer-to-Peer bots. This approach also recognizes the execution bottleneck of our framework and enhance its adaptability. The assessment comes about exhibited that the proposed framework fulfills high precision on distinguishing effective Peer-to-Peer bots and extraordinary adaptability. Also in proposed approach there is scope for duration, iterations and the accuracy of the system.

5. REFERENCES

[1] J. Zhang, R. Perdisci, W. Lee, U. Sarfraz, and X. Luo, "Detecting stealthy P2P botnets using statistical traffic fingerprints," in Proc. IEEE/IFIP 41st Int. Conf. DSN, Jun. 2011, pp. 121–132.

[2] S. Saad, I. Traore, A. Ghorbani, B. Sayed, D. Zhao, W. Lu, et al., "Detecting P2P botnets through network behavior analysis and machine learning," in Proc. 9th Annu. Int. Conf. PST, Jul. 2011, pp. 174–180.

[3] D. Liu, Y. Li, Y. Hu, and Z. Liang, "A P2P-botnet detection model and algorithms based on network streams analysis," in Proc. IEEE FITME, Oct. 2010, pp. 55–58.

[4] W. Liao and C. Chang, "Peer to peer botnet detection using data mining scheme," in Proc. IEEE Int. Conf. ITA, Aug. 2010, pp. 1–4.

[5] T. Karagiannis, K. Papagiannaki, and M. Faloutsos, "BLINC: Multilevel traffic classification in the dark," in Proc. ACM SIGCOMM, 2005, pp. 229–240.

[6] S. Sen, O. Spatscheck, and D. Wang, "Accurate, scalable in-network identification of P2P traffic using application signatures," in Proc. 13th ACM Int. Conf. WWW, 2004, pp. 512–521.

[7] T. Karagiannis, A. Broido, M. Faloutsos, and K. Claffy, "Transport layer identification of P2P traffic," in Proc. 4th ACM SIGCOMM Conf. IMC, 2004, pp. 121–134.

[8] A. W. Moore and D. Zuev, "Internet traffic classification using Bayesian analysis techniques," in Proc. ACM SIGMETRICS, 2005, pp. 50–60.

[9] M. P. Collins and M. K. Reiter, "Finding peer-to-peer file sharing using coarse network behaviors," in Proc. 11th ESORICS, 2006, pp. 1–17.

[10] D. Stutzbach and R. Rejaie, "Understanding churn in peer-topeer networks," in Proc. 6th ACM SIGCOMM Conf. IMC, 2006, pp. 189–202.

[11] T. Holz, M. Steiner, F. Dahl, E. Biersack, and F. Freiling, "Measurements and mitigation of peer-to-peer-based botnets: A case study on storm worm," in Proc. USENIX LEET, 2008, pp. 1–9.

[12] G. Bartlett, J. Heidemann, C. Papadopoulos, and J. Pepin, "Estimating P2P traffic volume at USC," USC/Information Sciences Institute, Los Angeles, CA, USA, Tech. Rep. ISI-TR-2007-645, 2007.

[13] T. Zhang, R. Ramakrishnan, and M. Livny, "BIRCH: An efficient data clustering method for very large databases," in Proc. ACM SIGMOD, 1996, pp. 103–114.

[14] M. Halkidi, Y. Batistakis, and M. Vazirgiannis, "On clustering validation techniques," J. Intell. Inf. Syst., vol. 17, nos. 2–3, pp. 107–145, 2001.

[15] (2011). Argus: Auditing Network Activity [Online]. Available: http://www.qosient.com/argus/

[16] Z.Li, A.Goyal, Y.Chen, and A.Kuzmanovic, "Measurement and diagnosis of address misconfigured P2P traffic," in Proc. IEEE INFOCOM, Mar. 2010, pp. 1–9.

[17] (2011).Autoit Script [Online]. Available: http://www.autoitscript. com/autoit3/index.shtml

[18] (2011).Zeus Gets More Sophisticated Using P2P Techniques[Online].Available: http://www.abuse.ch/?p=3499

[19] A. Binzenhofer, D. Staehle, and R. Henjes, "On the stability of chordbased P2P systems," in Proc. IEEE Global Telecommun. Conf., vol. 2. Nov./Dec. 2005, pp. 884–888.

[20] S. Rhea, D. Geels, T. Roscoe, and J. Kubiatowicz, "Handling churn in a DHT," in Proc. Annu. Conf. USENIX Annu. Tech. Conf., 2004, pp. 127–140.