Multi-Stream Fused Model: A Novel Real-Time Botnet Detecting Model

Jae Moon Lee and Thien Nguyen Phu

Abstract--- *In the current computer era, spam, DDoS and phishing are familiar complications on the Internet. Once, attackers tended to make use of centralized high bandwidth associations to achieve their tasks. At present, even home users have high bandwidth internet connections, attackers have started infecting and using these home computers for their attacks. Attacking from distributed places, attackers are harder to catch or prevent and typically have more bandwidth to abuse. New schemes are required to sense the forming of these widespread networks of infected hosts, particularly now that it seems attackers have discovered the peer-to-peer (P2P) technology. They develop new features like P2P Command and Control (C&C), which makes conventional detection methods no longer efficient for indicating the existence of the bots. Here, a system is proposed that accurately and competently detects the existence of storm botnet. In this paper, based on a number of new P2P botnet characteristic properties, a novel real-time detecting model – MSFM (Multi-Stream Fused Model) is proposed. MSFM considers multiple categories of packets' unique characteristics and handle them with equivalent strategies. Experiment results demonstrate that this model can accurately detect botnet with comparatively low false-positive and false-negative rates.*

Keywords--- *Centralized Botnet, Discrete Kalman Filter, Multi-Chart CUSUM, P2P Command and Control.*

I. INTRODUCTION

A. An Overview

A BOTNET is a collection of compromised computers, each of which is known as a 'bot', connected to the Internet.

When a computer is compromised by an attacker, there is often code within the malware that commands it to become part of a botnet. The "botmaster" or "bot herder" controls these Compromised computers via standards-based network protocols such as IRC and http (4). Botnets have recently been identified as one of the most important threats to the security of the Internet. Traditionally, botnets organize themselves in a hierarchical manner with a central command and control location. This location can be statically defined in the bot, or it can be dynamically defined based on a directory server. Presently, the centralized characteristic of botnets is useful to security professionals because it offers a central point of failure for the botnet. In the near future, they believe attackers will move to more resilient architectures. In particular, one class of botnet structure that has entered initial stages of development is peer-to-peer based architectures.

- Peer-to-peer: A peer-to-peer network is a network in which any node in the network can act as both a client and a server.
- Bot: A bot is a program that performs user centric tasks automatically without any interaction from a user.
- Botnet: A botnet is a network of malicious bots that illegally control computing resources.

B. Objective of Study

The objective of the proposed system is to efficiently and accurately detect the presence of storm botnet. In this system we propose a novel real-time detecting model- MSFM (Multi Stream Fused Model). In this model, we incorporate different detecting methods based on unique characteristics of network flow packets. Firstly, MSFM lays emphasis on the UDP flow, which is most related with the botnet C&C (Command and Control). We use the Hurst Parameter to retrieve the whole condition, and detect the abnormality based on the UDP packets' self-similarity. Secondly, MSFM applies the discrete Kalman filter to find ICMP and SMTP packets' anomaly and Multi-chart CUSUM acts as the amplifier to make the abnormality clearer. Finally, we consider the impact on the botnet detection which web applications generate, especially the P2P applications, and use the properties of TCP flow to analyze that the abnormalities are caused by the botnets or the P2P applications. In addition, this paper uses the Kaufman algorithm to dynamically adjust the threshold to minimize the false positives and false negatives. After series of experiments, the results prove that the model can detect the Storm botnet in a relatively high precision with both low false-positive and false-negative rate.

II. BOTNET ATTACK CONCEPTS

A. Botnet

Malicious botnets are networks consisting of large numbers of bots. Bot is actually short for robot. Symantec defines a bot as: Bots are similar to worms and Trojans, but earn their unique name by performing a wide variety of automated tasks on behalf of their master.

Bots are also called "zombies" because a computer (infected with a bot) performs a task given by its master. A botnet herder (also called a botnet master) is a person or a group, who control the whole botnet: they can give instructions or upload data to the botnet. Botnets are used to

Jae Moon Lee, Gifu University, Japan.
Thien Nguyen Phu, Graduate School of Dongguk University, Korea.
DOI:10.9756/BIJDM.8331

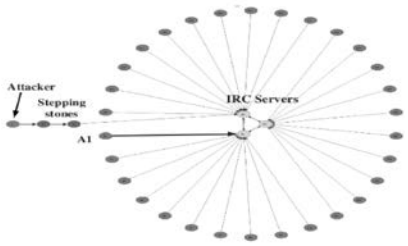
perform a wide variety of tasks but some of the most popular ones are sending spam or coordinating a DDoS attack. There are different ways to infect a computer with a bot. Often it searches the Internet to look for \unprotected" computers to infect by exploiting known software vulnerabilities but it can also being sent through email, or hidden in another program or installed by a malicious website, this software is also known as malware.

B. *Botnet Architectures*

There are a few topologies described in Botnet architecture. They differ from each other in how big they can get how easy they can be detected and disrupted and how they exchange messages and commands (command and control).

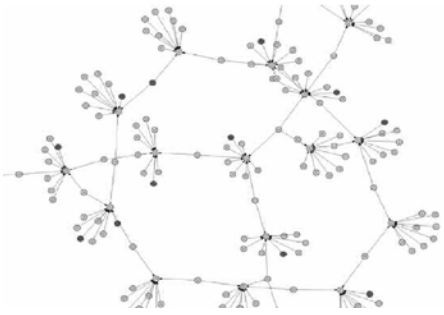
C. *Centralized*

The oldest type of topology is the centralized form, see figure 1 on the following page. There is one central point that forwards commands and data between the botnet herder and his botnet clients. The big advantage is that there is little latency. The biggest drawback is that they can be more easily detected then decentralized, since all the connections lead to a few nodes. Also when a IRC server gets disrupted or disconnected, an entire branch or perhaps the whole botnet could be taken offline because the botnet herder cannot pass any messages to the bots anymore. Passing commands and data in centralized systems go through a central point: in most cases they all connect to a central C&C node or a central IRC server where they will receive their commands.



Centralized Topology

D. *Decentralized*

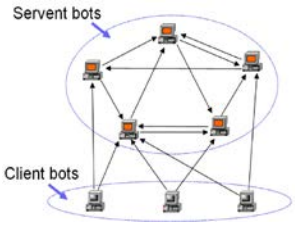


Decentralized Peer-To-Peer Network

A peer-to-peer (or P2P) computer network uses a partial or full mesh topology and the cumulative bandwidth of network participants. Peer-to-peer networks are typically used for connecting a very large number of nodes via ad hoc connections. These networks differ from the traditional client-server model because the peer nodes (both client and server) are all equal. The decentralized topologies do not have one central node which can just be taken down to disable the botnet. Instead, they are all connected to each other in a way. Some protocols use centralized servers for search operations or define \super peers" for maintaining an index or acting as a broker

E. *Hybrid*

There are a few distinct differences with the peer-to-peer and centralized topology. First, the bot communicates via a peer list contained in each bot. Each bot has it's own fixed and limited list, which it will not reveal to other bots. When a bot is compromised, only a limited number of nodes are exposed. This way, the botnet does not require bootstrapping, it can connect to the botnet directly. There is a disadvantage about this method: only bots with a static IP can be chosen to appear in the list because it's hard coded. Each servant bot listens on a self determined port for incoming connections and uses a self-generated symmetric encryption key for incoming traffic. This makes it very hard to detect because each bots traffic uses a different encryption key. The botnet master issues commands to his servant bots (listed in the peer list). This servant will forward the commands to the other nodes.



Hybrid Peer-To-Peer Botnet

III. PROPOSED SYSTEM WORK

Spam, DDoS and phishing are common problems on the Internet nowadays. In the past, attackers tended to use centralized high bandwidth connections to accomplish their tasks. Now that home users have high bandwidth internet connections, attackers have started infecting and using these home computers instead to for their attacks. Attacking from distributed locations, attackers are harder to catch or stop and often have more bandwidth to abuse. New methods are required to detect the forming of these widespread networks of infected hosts, especially now that it seems attackers have discovered the peer-to-peer (P2P) technology (6). They evolve new features such as P2P Command and Control(C&C), which makes traditional detection methods no longer effective for indicating the existence of the bots. In this paper, based on several of the new P2P botnet characteristic properties, we propose a novel real-time detecting model – MSFM (Multi-Stream Fused Model). MSFM considers multiple types of packets' unique characteristics and handle them with corresponding strategies. Extensive experiment results show that our model can accurately detect peer-to-peer botnet with relatively low false-positive and false-negative rates.

In the proposed system we propose a novel real-time detecting model- MSFM (Multi-Stream Fused Model). In this model, we incorporate different detecting methods based on unique characteristics of network flow packets. Firstly, MSFM

lays emphasis on the UDP flow, which is most related with the botnet C&C (Command and Control). We use the Hurst Parameter to retrieve the whole condition, and detect the abnormality based on the UDP packets' self-similarity. Secondly, MSFM applies the discrete Kalman filter to find ICMP and SMTP packets' anomaly and Multi-chart CUSUM acts as the amplifier to make the abnormality clearer. Finally, we consider the impact on the botnet detection which web applications generate, especially the P2P applications, and use the properties of TCP flow to analyze that the abnormalities are caused by the botnets or the P2P applications. In addition, this paper uses the Kaufman algorithm to dynamically adjust the threshold to minimize the false positives and false negatives. After series of experiments, the results prove that the model can detect the Storm botnet in a relatively high precision with both low false-positive and false-negative rate.

The major disadvantage of the existing systems is that the researches try to find the existence of bots keeping the data packet as a whole. Above all, researches in decentralized botnets detecting are still in a beginning period. It is difficult for most of the studies to identify that whether the net flow characteristics and the host behaviors are led by botnets or normal activities.

Proposed system deal with different types of packets in different methods: detecting increasing number of UDP, ICMP and SMTP packets by discrete Kalman filter; amplifying the results by multi-chart CUSUM; detecting UDP packets using self-similarity; and improving detection precision using Kaufman algorithm to dynamically adjust the threshold. This considered the web applications generated impacts on the botnet detection, especially the P2P applications, and differentiated normal flow with botnet communications**.**

IV. EXPERIMENT

Fighting botnets is often a matter of finding their weak spot: their central point of command, or command-and-control server. This is usually an IRC, Internet Relay Chat, network where all compromised computers connect to, but with the use of P2P technology, this central point of command is nowhere to find: the hosts connect to each other and the attacker only has to become one of the peers to broadcast his commands over the network. A new detection and fighting method is required to prevent or stop such hazardous networks.

A. Methods Are

- UDP Flow
- ICMP and SMTP Flow
- TCP Flow

B. UDP Flow

UDP flow is most related with the C&C (Command and Control) of botnet. In udp flow uses self-similarity to find botnet. The self-similarity refers to a scale invariance property, which intuitively means that plots of traffic intensity at different time-scales look very similar. Storm botnet causes abnormities in network flows: the number of UDP packets will increase because of communication between bots, which will weaken the self-similarity of UDP flow. Because the Hurst Parameter represents the degree of self-similarity, the change of Hurst Parameter will reveal the abnormities of UDP flow. Thus, we sampling the UDP flow, and calculate the Hurst Parameter of it.

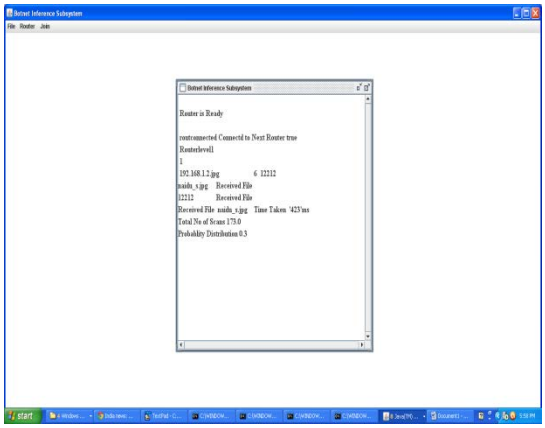
C. ICMP and SMTP Flow

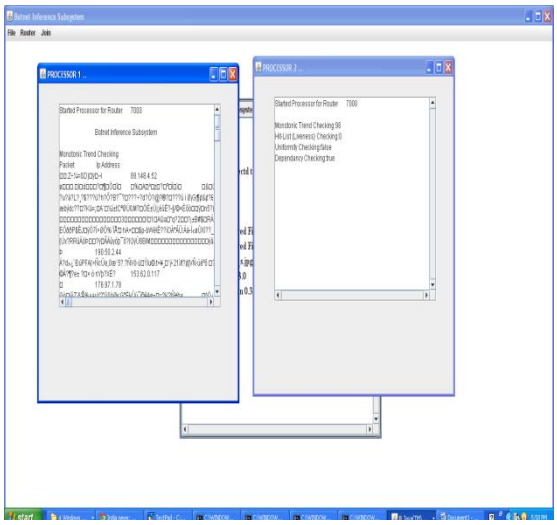
Discrete Kalman filter to find ICMP and SMTP flow anomaly, and Multi-chart CUSUM acts as the amplifier to make the abnormality clearer. The discrete Kalman filter estimates a process by using a form of feedback control: the filter estimates the process state at some time and then obtains feedback in the form of (noisy) measurements. So, the equations for the Kalman filter fall into two groups: time update equations and measurement update equations. The time update equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain a priori estimate for the next time step. The measurement update equations are responsible for the feedback for incorporating a new measurement into the priori estimate to obtain an improved posteriori estimate. Multi-chart CUSUM Internet traffic could be viewed as a complex random model: any abnormity in the traffic will bring obvious changes. However, since the observation series are blurred in the Internet security issues, it is hard to build up a specific model. For this reason, a nonparametric CUSUM (NP-CUSUM) that uses minimum a priori information is needed for abnormity detection. Moreover, more details in net flow changing are needed in the botnet detection, so multi-chart NP-CUSUM detection algorithm is used for multi-factor detection in the network.

D. TCP Flow

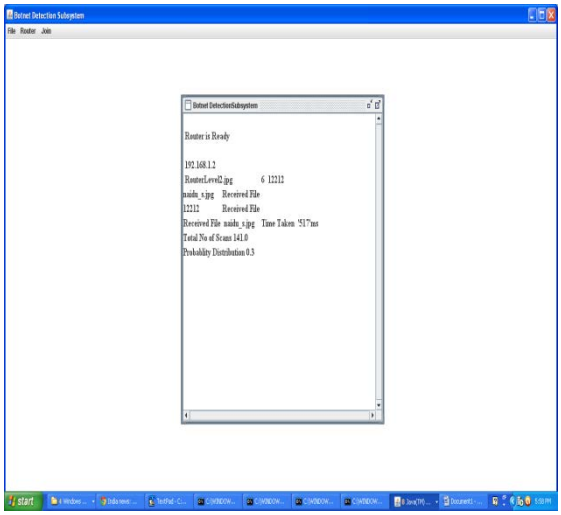
The new decentralized botnets are very similar to the traditional P2P applications, so the abnormalities mentioned above are also possibly caused by the traditional P2P applications in the network. Thus, some method is needed to analyze that the abnormalities are caused by the botnet or the P2P applications. The normal P2P applications always use the UDP packets to transfer the control information, and use the long TCP packets (the size is always bigger than 1300 bytes) to transfer the data. And the only data transmission of botnet is the "secondary injection" process, which is transferred by the HTTP protocol and the data volume is small.

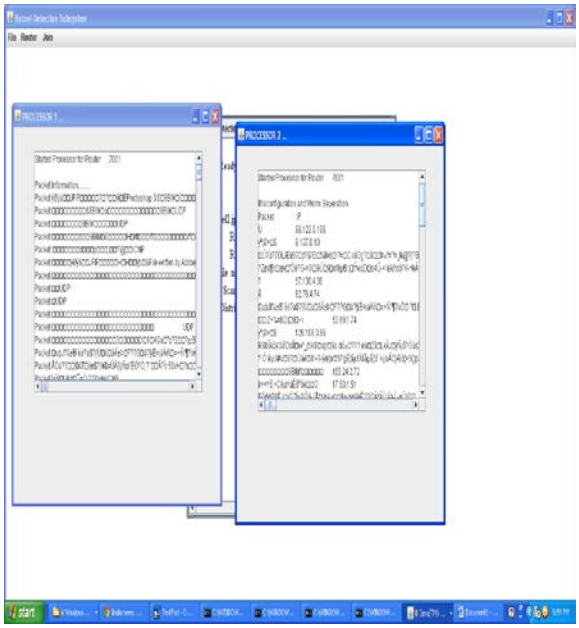
E. Inference System



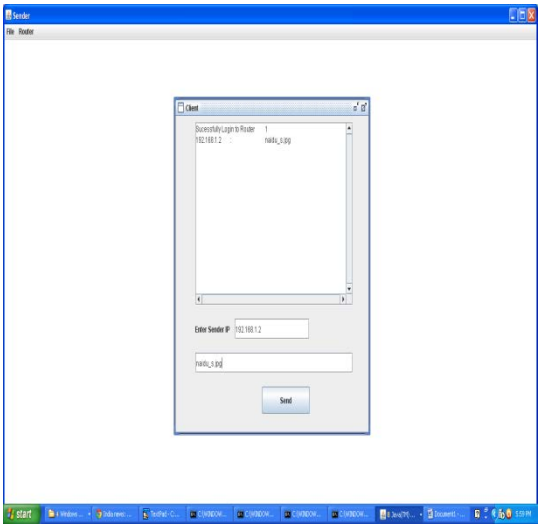


F. Detection System

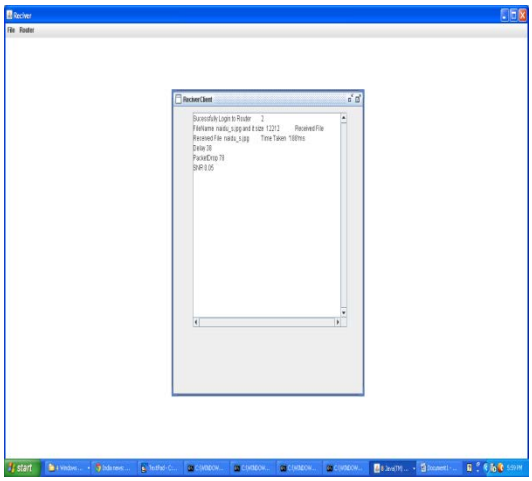




G. Sender



H. Receiver



V. CONCLUSION

Hence, in this paper, we have proposed and simulated a secure Mobile IP Communication Protocol to provide security in Mobile IP Registration as well as in Communication. Then we focused on simple implementation of proposed scheme for roaming devices in home networks. From the above simulation results, we have studied over the wireless characteristics of the moving mobile node from one residential network to another residential network. Using residential gateway as a tunnel end point is efficient and secure method. However, the residential gateways seem to be overloaded by both Mobile IP and IPSec tunnels for every communication of roaming devices. We have performed this simulation with less number of residential devices in residential network. In future, we would also like to study over the impact of more devices in this proposed method. For additional security overhead in residential gateway with more devices, we would like to focus on implementing some QoS mechanisms.

REFERENCES

[1] J. Stewart, "Storm Worm DDOS Attack", Secure Works, Inc., Atlanta GA, 2007.

[2] J. Grizzard, V. Sharma, C. Nunnery, B. Kang and D. Dagon, "Peer-to-Peer Botnets: Overview and Case Study", HotBots conference, 2007.

[3] S. Sarat and A. Terzis, "Measuring the Storm Worm Network" Technical Report 01-10-2007, HiNRG Johns Hopkins University, 2007.

[4] P. Wang, B. Aslam and C.C. Zou, "Peer-to-peer botnets: The next generation of botnet attacks", Electrical Engineering, 2010.

[5] A. Nummipuro, "Detecting P2P-Controlled Bots on the Host", Seminar on Network Security, Espoo, Helsinki, 2007.

[6] M. Steggink and I. Idziejczak, "Detection of peer-to- peer botnets", University of Amsterdam, Netherlands, 2007.

[7] P. Porras, H. Saidi and V. Yegneswaran, "A Multiperspective Analysis of the Storm (Peacomm) Worm", Technical report-Computer Science Laboratory, SRI International, 2007.

[8] C.R. Davis, J.M. Fernandez, S. Neville and J. McHugh, "Sybil attacks as a mitigation strategy against the storm botnet", Proc. 3rd Int. Conf. on Malicious and Unwanted Software, 2008.

[9] B. Kang, E. Chan-Tin, C. Lee, J. Tyra, H. Kang, C. Nunnery, Z. Wadler, G. Sinclair, N. Hopper, D. Dagon and Y. Kim, "Towards complete node enumeration in a peer-to-peer botnet", ACM Symposium on Information, Computer & Communication Security, 2009.

[10] Z. Li, B. Wang, D. Li, H. Chen, F. Liu and Z. Hu, "The Aggregation and Stability Analysis of Network Traffic for Structured-P2P-based Botnet Detection", Journal of Networks, Vol. 5, No. 5, Pp. 517- 526, 2010.

[11] H.R. Zeidanloo and A.B.A. Manaf, "Botnet Detection by Monitoring Similar Communication Patterns", International Journal of Computer Science and Information Security, Vol. 7, No. 3, Pp. 36-45, 2010.

[12] P. Maymounkov and D. Mazieres, "Kademlia: A peer-to-peer information system based on the XOR metric", 1st International Work shop on Peer-to-Peer Systems, 2002.