CHAPTER 1

INTRODUCTION

Peer-to-peer network is a decentralized communication model in which the each party has the same capacities and either party can initiate a communication session. Unlike the client/server, client make a service request and the server fulfill. The peer-to-peer allow each node to service both as client and server.

P2P system can be used to provide anonymized routing of network traffic, massive parallel computing environments, distributed storage and other functions. Most P2P programs are focused on media sharing and P2P is therefore are often associated with software privacy and copyright violation.

Typically ,peer-to-peer applications allow users to control many parameters of operation:

- How many member connections to seek or allow at one time.
- Whose system to connect to or avoid.
- What service to offer and how many resources to devote to network.

In its simplest form, a peer-to-peer (P2P) network is created when two or more PCs are connected and share resources without going through a separate server computer. A P2P network can be an ad hoc connection—a couple of computers connected via a Universal Serial Bus to transfer files. A P2P network also can be a permanent infrastructure that links a half-dozen computers in a small office over copper wires. Or a P2P network can be a network on a much grander scale in which special protocols and applications set up direct relationships among users over the Internet.

Peer-to-Peer overlay networks are distributed systems consisting of interconnected nodes which self-organize into network topologies. They are built with specific purposes of sharing resources such as content, CPU cycles, storage and bandwidth, and have the ability to accommodate a transient population of nodes while maintaining acceptable connectivity and performance, without requiring the intermediation or support of a global centralized server or authority.

The construction of P2P networks is on the top of IP layer, typically with a decentralized protocol allowing ‘peers’ to share resources. In the earlier research it is noted to find only malicious packets coming to the network.But here we set a rule based approach to control only specified packets come to the network .By this we can set certain rules.Here we can even classify incoming packets. By classifying incoming packets we can check if the coming packets are malicious.A greedy based algorithm called FDPM is used for packet filtering. In this project , we develop a compromised router detection protocol that dynamically infers the precise number of congestive packet losses that will occur. Once the congestion ambiguity is removed, subsequent packet losses can be safely attributed to malicious actions. We believe our protocol is the first to automatically predict congestion in a systematic manner and that it is necessary for making any such network fault detection practical.

Anomaly score based botnet detection is proposed to identify the botnet activities by using the similarity measurement and the periodic characteristics of botnets. To improve the detection rate, the proposed system employs two-level correlation relating the set of hosts with same anomaly behaviors. The proposed method can differentiate the malicious network traffic generated by infected hosts (bots) from that by normal IRC clients, even in a network with only a very small number of bots. In the remainder of this project, we briefly survey the related background material, evaluate options for inferring congestion, and then present the assumptions, specification, and a formal description of a protocol that achieves these goals. We have evaluated our protocol in a small experimental network and demonstrate that it is capable of accurately resolving extremely small and fine-grained attacks. The experimental results show that, regardless the size of the botnet in a network, the proposed approach efficiently detects abnormal IRC traffic and identifies botnet activities.

CHAPTER 2

LITERATURE SURVEY

Literature review provides us with various techniques for detecting malicious packets coming to network.. Also, various classification techniques have been adopted by the authors.

L. Prabhu, K.Dhivya, V.Geetha propsed a paper titled" An efficient scalable system for peer-to-peer botnet detection" which analysis the behavioral characteristics of identifying P2P and it finds the difference between P2P botnet traffic and legal p2p traffic.

In this paper , proposed a novel scalable P2P botnet detection system that is able to identify stealthy P2P botnets. To perform this task statistical fingerprints of P2P communications have been derived to detect P2P clients and further distinguish between those that are part of legitimate P2P networks and P2P bots. The results shows that the proposed system accomplishes high accuracy on detecting stealthy P2P bots and great scalability.

L. Li, S. Mathur, and B. Coskun, proposed a paper titled" Gangs of the Internet: Towards Automatic Discovery of Peer-to-Peer Communities" which propose a method to discover P2P networks (both benign and malicious) from network flow records captured at the boundary of a tier-1 Internet backbone provider.

The basic idea is that flows belonging to P2P applications can be modeled as observations from a mixed membership statistical model, with P2P applications acting as latent variables. Hence the communication patterns of hosts (who-talks-to-whom), as measured at the edge of a large network, can be decomposed into constituent application-layer P2P communities without any human effort in selecting specific features. The proposed method can discover P2P communities independent of their specific protocol or topology, using only communication patterns, allowing:

(i) policy enforcement related to P2P traffic in private networks such as enterprises and schools,
(ii) discovery of new/evolving malicious P2P botnets, and
(iii) network service providers a means to monitor P2P traffic as separate communities at the application level.

H. Hang, X. Wei, M. Faloutsos, and T. Eliassi-Rad , propsed a paper “Entelecheia:
Detecting p2p botnets in their waiting stage,” which shows a graph-based approach called Entelecheia that is aimed at finding decentralized or peer-to-peer botnets, botnets that are in a non-active period known as the “Waiting” stage, and polymorphic bots that evade signature detection.

This propose a novel, simple, intuitive, yet effective approach for detecting P2P botnets during their Waiting stage by identifying their “social” behavior. Here operate under the following two requirements, assume no signatures or prior knowledge and assume no seed information through a blacklist of IPs. Key insight is to exploit the inherent behavior of botnets by examining long-lived and low intensity flows.

Tomas Isdal , Michael Piatek , Arvind Krishnamurthy ,Thomas Anderson proposed a paper titled” Privacy-preserving P2P data sharing with OneSwarm” which shows the reduce the performance cost of privacy and our measurements of the live system show that anonymized data transfers are performance competitive with unanonymized use.

OneSwarm’s novel lookup and transfer techniques yield more than an order of magnitude improvement in transfer speeds relative to Tor, another widely-used anonymization system. Here have built OneSwarm, a file sharing system designed to reduce the cost of privacy to the average user. It develop novel techniques for efficient, robust, and privacy-preserving lookup and data transfer. It provide users flexible control over their privacy by defining sharing permissions and trust at the granularity of individual data objects and peers. The OneSwarm client is publicly available for download on Linux, Mac OS X, and Windows, and it is in widespread use around the globe. Measurements with the live

OneSwarm deployment show that it delivers on its promise: privacy-preserving downloads on One Swarm are roughly as fast as a direct Internet transfer between the two nodes, and an order of magnitude faster than using Tor for the same operation.

T. T. Lu, H.Y. Liao, M .F. Chen proposed a paper" An Advanced Hybrid P2p Botnet 2.0" In this paper, we propose an advanced hybrid peer-to-peer (P2P) botnet 2.0 (AHP2P botnet 2.0) using web 2.0 technology to hide the instructions from botmaster into social sites, which are regarded as C&C servers.

This work presents an advanced hybrid peer-to-peer (P2P) botnet 2.0 mechanism using web 2.0 technology to instruct social sites. The approach is particularly suitable for hiding the encryption malware information.

Genevieve Bartlett, John Heidemann , Christos Papadopoulos, James Pepin proposed a paper" Estimating P2P Traffic Volume at USC" we estimate traffic based on both port-based and connection-pattern based techniques.

Here quantify the volume of traffic from P2P activity as well as the number of campus IPs involved in P2P at USC. Since port-matching techniques often fail for P2P applications, we estimate traffic based on both port-based and connection-pattern based techniques. In addition, while we identify P2P sharing, we cannot comment the types of data being shared (either music or data, restricted or freely available). Find that 3–13% of active IPs on campus participate in P2P, and that this traffic accounts for 2 1–33% of the bytes transferred to and from our campus.

Wernhuar Tarng, Cheng-Kang Chou and Kuo-Liang Ou proposed a paper" A p2p botnet virus detection system based on data-mining algorithms" Here a P2P botnet virus detection system based on data-mining algorithms is proposed.

In this study to detect the infected computers quickly using Bayes Classifier and Neural Network (NN) Classifier. The system can detect P2P botnet viruses in the early stage of infection and report to network managers to avoid further infection. The system adopts real-time flow identification techniques to detect traffic flows produced by P2P application programs and botnet viruses by comparing with the known

flow patterns in the database. In this study, a P2P botnet virus detection system is developed based on two data-mining algorithms, i.e., Bayes Classifier and NN Classifier. When the system is installed on the network, it can monitor the traffic flows to detect the infected computers and identify the P2P botnet viruses in real time. The results show that the system can identify known or unknown flows International Journal of Computer Science & Information Technology (IJCSIT) Vol 4, No 5, October 2012 65 produced by P2P botnet viruses correctly in a short time to achieve the goal of infection control. After trained by adjusting the system parameters using test samples, the accuracy of Bayes Classifier is 95.78% and that of NN Classifier is 98.71%. Both classifiers can identify known and unknown P2P botnet viruses within 5 minutes.

Ping Wang, Lei Wu, Ryan Cunningham, Cliff C. Zou proposed a paper" Honeypot Detection in Advanced Botnet Attacks" Attackers could detect honeypots in their botnets by checking whether compromised machines in a botnet can successfully send out unmodified malicious traffic. Based on this basic detection principle, we present honeypot detection techniques to be used in both centralized botnets and peer-to-peer structured botnets.

In this paper, introduced various means by which botmasters could detect honeypots in their constructed botnets based on this principle. Honeypot research and deployment still has significant value for the security community, but we hope this paper will remind honeypot researchers of the importance of studying ways to build covert honeypots, and the limitation in deploying honeypots in security defense.

CHAPTER 3

PROBLEM DEFINITION

From the study we can only find the malicious packet. Inspite of detecting the malicious packets, here in this project , we develop a compromised router detection protocol that dynamically infers the precise number of congestive packet losses that will occur. Once the congestion ambiguity is removed, subsequent packet losses can be safely attributed to malicious actions. We believe our protocol is the first to automatically predict congestion in a systematic manner and that it is necessary for making any such network fault detection practical. Anomaly score based botnet detection is proposed to identify the botnet activities by using the similarity measurement and the periodic characteristics of botnets.

To improve the detection rate, the proposed system employs two-level correlation relating the set of hosts with same anomaly behaviors.. In the remainder of this project, briefly survey the related background material, evaluate options for inferring congestion, and then present the assumptions, specification, and a formal description of a protocol that achieves these goals. Here evaluated protocol in a small experimental network and demonstrate that it is capable of accurately resolving extremely small and fine-grained attacks. The experimental results show that, regardless the size of the botnet in a network, the proposed approach efficiently detects abnormal IRC traffic and identifies botnet activities. In the proposed the system we use FDPM method for packet filtering.

CHAPTER 4

PACKET INTRUSION DETECTION

Flexible Deterministic Packet Marking (FDPM) which provides a defense system with the ability to find out the real sources of attacking packets that traverse through the network. While a number of other traceback schemes exist, FDPM provides innovative features to trace the source of IP packets and can obtain better tracing capability than others. In particular, FDPM adopts a flexible mark length strategy to make it compatible to different network environments; it also adaptively changes its marking rate according to the load of the participating router by a flexible flow-based marking scheme. Evaluations on both simulation and real system implementation demonstrate that FDPM requires a moderately small number of packets to complete the traceback process; add little additional load to routers and can trace a large number of sources in one traceback process with low false positive rates. The built-in overload prevention mechanism makes this system capable of achieving a satisfactory traceback result even when the router is heavily loaded. The motivation of this traceback system is from DDoS defense. It has been used to not only trace DDoS attacking packets but also enhance filtering attacking traffic. It has a wide array of applications for other security systems.

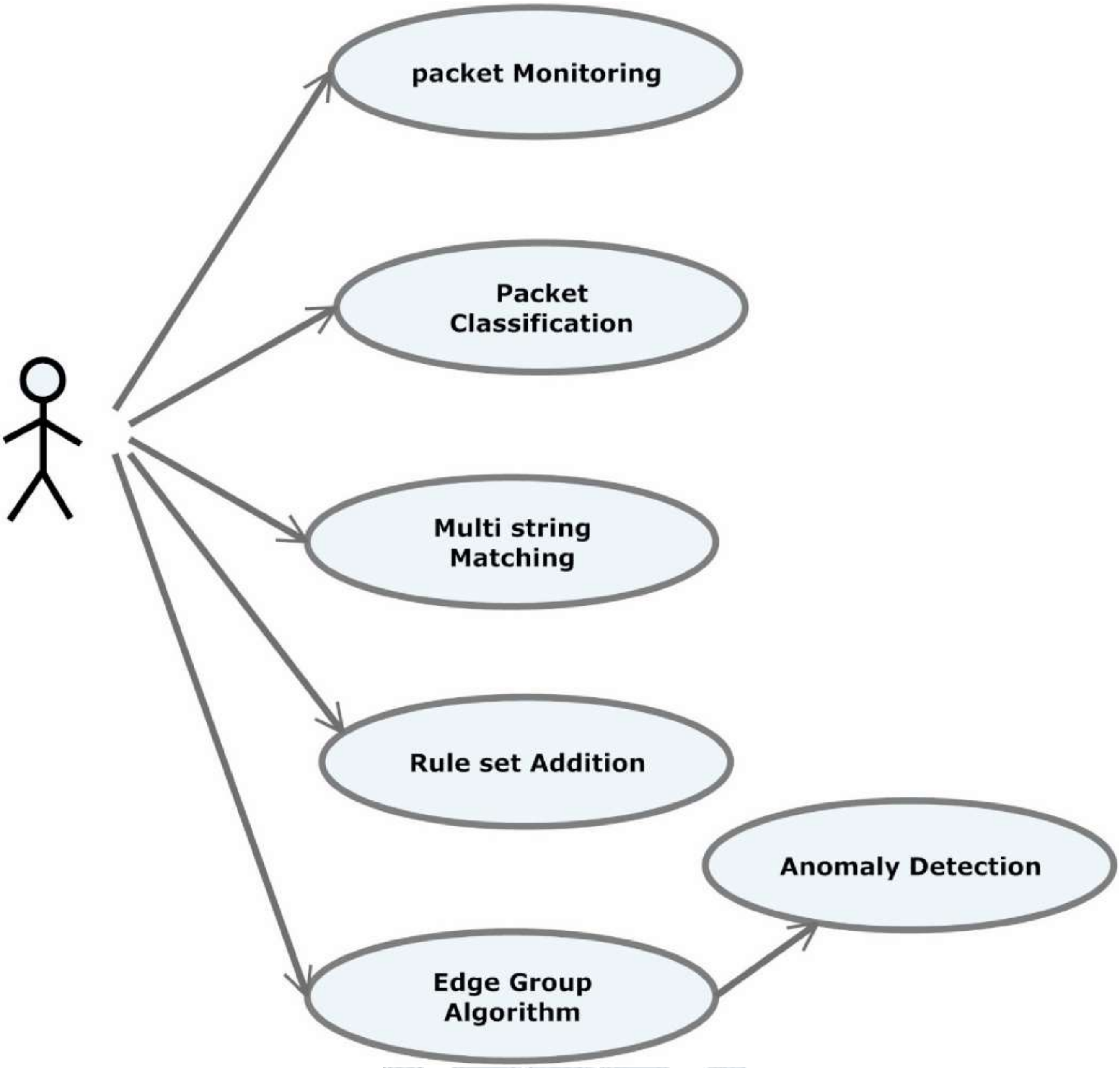


Fig 4.1 :use case diag of proposed system

4.1 DIFFERENT PHASES

In the previous system they only identify the known anomaly and also for a specified location. This is long term network statistical process.

To perform highly effective in providing a colorized visualization chart to network analysts in the presence of bursty network traffic. The implementation work can be divided into

1. Packet Monitoring
2. Port Scan Detection and Histogram Creation
3. Anomaly Detection
4. Rule Based detection

4.1.1 Packet Monitoring

Packet Analyzer

Description: It is used to analyze the incoming packets and to display the packet information. The display list contains the following:

Source IP	Identification
Destination IP	Flags
IP Version	Fragmentation Offset
Header Length	Time To Live (TTL)
Type of Service (TO S)	Protocol
Total Length of Header	Header Checksum

Table4. 1 contents in the incoming packets

We can identify TCP and UDP packets from the incoming IP packets by checking the protocol field of the IP header.

TCP Packets contains the following information which is separately displayed as a list:

Source Port	Flags
Destination Port	Window Size
Sequence Number	Checksum
Acknowledgment Number	Urgent Pointer
Header Length	

Table4.2 contents in incoming TCP packets

Source Port
Destination Port
Length
Checksum

Table 4.3 contents in UDP packets

- **Active Port Detection**

Description: Every system connected to a LAN will have some active ports (open ports). The system should have the following features:

- Get the IP address of the system to be scanned for active ports.
- Get the range of ports to be checked for activity.
- List the ports differentiating active and inactive ones.
- Display the service running on active ports.

4.2 Port Scan detection

Description: It is used to detect the port attacks from an intruder either from the network or from outside the network. If detected any port attack it will generate an alert. The system should display the following details related to the intrusion:

- IP Address of the intruder.
- IP Address of the system under attack.
- Time and Date of attack.
- The range of ports under attack.

Histogram Creation

Creation of histogram is based on joint values of multiple features like port number and range of IP addresses.

Important features are:

› SrcIP addr, ›
DstIP addr, › Src
port number,

› Dst port number

It is based on five minutes trace of data. Based on network scanning attacks that connected all host in the monitoring network .

4.3 Anomaly Detection

Here predicting the anomaly by checking the different port used by the system for getting connected. If different port will try to get connected in a instatenous time then this system will be considered as an anomaly. This section describes the results of the normal density approximation study and the prediction algorithm results. shows the period gram, correlation results, and histogram of a selected feature set; namely “Average port”, “High ports,” “Server factor,” and “Peered factor.” Similarity between the measured feature values and generated Gaussian data for all features are clearly visible except for the attribute “Peered factor.” The significant variation may be attributed to the highly stochastic nature of the traffic. In we show the auto- and cross-correlation plots of the selected features in order to determine the level of correlation for this measured feature set. Notice that, while and possess high correlation, the remaining pairs do not. This implies that features, and should be uncorrelated before they are used individually or collectively.

This will try to get connected to remote systems by using open port ,then it will scan for the port number if port number available then connection established and now he can communicate to the remote system. By selecting the list machine we will get the total number of ports used by the particular system for getting connected to the remote system. This will display all the details of the port, this is done for analysis any rapid change in the packet capturing in sudden.

The act of systematically scanning a computer's ports .Since a port is a place where information goes into and out of a computer, port scanning identifies open doors to a computer. Port scanning has legitimate uses in managing networks, but port scanning also can be malicious in nature if someone is looking for a weakened access point to break into your computer.

If attacker in the network then it will detect that file secretly by using the open port. This will a small program that will be run in the backend of the software. This will not literally not known to the attacker and he will detected very easily. This type program will be also act like attacker this will be a serious problem while this case of the project..

A backdoor is a mechanism surreptitiously introduced into a computer system to facilitate unauthorized access to the system. While backdoors can be installed for accessing a variety of services, of particular interest for network security are ones that provide interactive access. These are often installed by attackers who have compromised a system to ease their subsequent return to the system.

4.4 Rule Based detection and Analysis

In this module the possible conflicts on the rule segments are identified and corresponding action constraints are determined. This module has following sub modules:

a) Conflicting Segment identification & Correlation
b) Action Constraint Generation
c) Conflicting Rule Reordering

For conflict detection and resolution, conflicting segments are identified in the first step. Each conflicting segment associates with a policy conflict and a set of conflicting rules. Also, the correlation relationships among conflicting segments are identified and conflict correlation groups (CG) are derived. Policy conflicts belonging to different conflict correlation groups can be resolved separately;

thus, the searching space for resolving conflicts is reduced by the correlation process.The second step generates an action constraint for each conflicting segment by examining the characteristics of each conflicting segment. The third step utilizes a reordering algorithm, which is a combination of a permutation algorithm and agreed algorithm, to discover a near-optimal conflict resolution solution for policy conflicts. A rule in a firewall is redundant if and only if removing the rule does not change the function of the firewall, i.e., does not change the decision of the firewall for every packet. In this module, redundant rules that on deleting will not make any change on firewall policy is discovered and then eliminated.

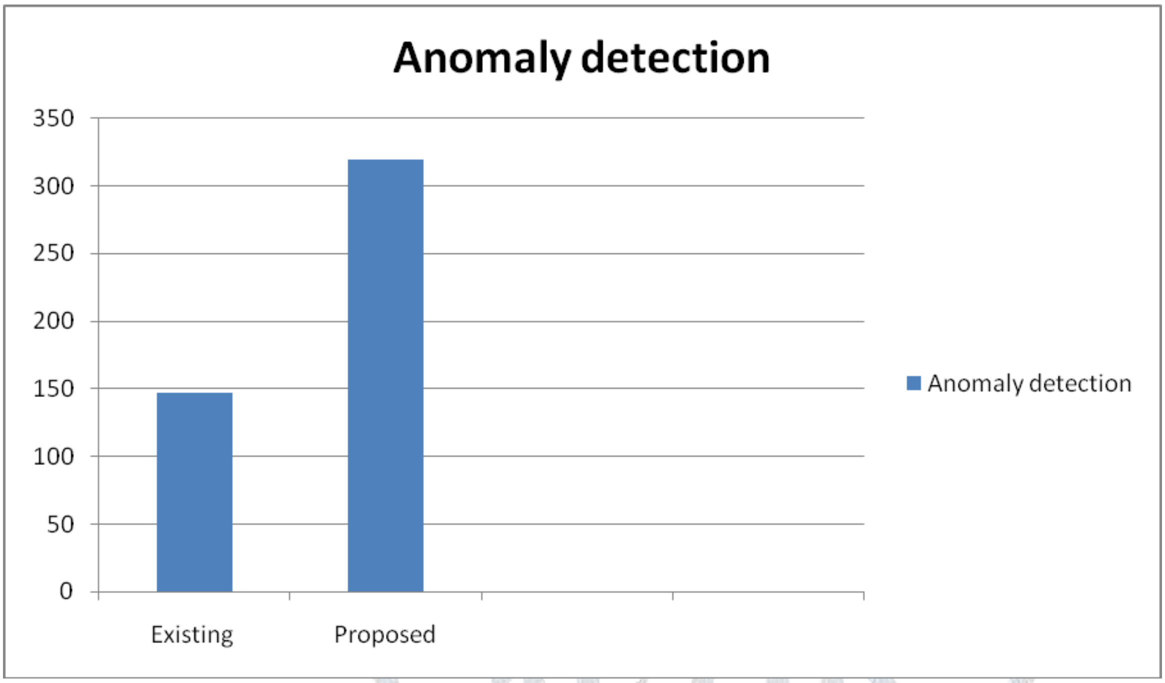
Policy Settings: Rules are the building blocks of Firewall and it works based on the rules or policies, similarly administrator can set certain rules in DFAME so that the incoming and outgoing packets can be allowed or blocked as per the administrator demand. Rule reordering can also be done here to avoid conflicts.

Packet Freezing: Another important feature of DFAME is freezing of packets. Here we can freeze those packets from specific IP address for some period of time. For example we can freeze packets from Mozilla Firefox until we unfreeze it.

CHAPTER 5
RESULT AND ANALYSIS

Anomalies are handled based on the category in which the anomaly belongs to. If the intrusion is from the network, then that should be prevented from entering in to the node and from forwarding to another node. We cannot delete that packet because the packet is created by some other node in the network .If any vampire is found inside the node that should be deleted immediately and should prevent from forwarding. For avoiding the entry of anomalies from the network to any packet, all the packets should satisfy no backtracking property. Port scan details can be availed by continuously monitoring all the open ports in the node.

Each individual computer runs on multiple ports. Port Scanning is the name for the technique used to identify open ports and services available on a network host. Hackers typically utilize port scanning because it is an easy way in which they can quickly discover services they can break into. In some cases, hackers can even open the ports themselves in order to access the targeted computer. At any time, there are open ports on one's personal computer, there is potential for the loss of data, the occurrence of a virus, and at times, even complete system compromise. It is essential for one to protect his or her virtual files, as new security risks concerning personal computers are discovered every Computer protection should be the number one priority for those who use personal computers.



Port scanning is considered a serious threat to one's PC, as it can occur without producing any outward signs to the owner that anything dangerous is taking place. If any foreign node is continuously trying to access any open port in a node then that should be suspected. Then packets from that particular IP address can be blocked .Entropy variation is an indication of the variation in packet arrival rate.

	Existing Method	Proposed method (Rule Based)
Trace Length (seconds)	3600	3600
Number of packets	874613	1074132
Avg packet rate(per second)	242.9	298.3
TCP Packets	303142	403433
UDP Packets	571471	670699
Anomaly Detection rate	147	320

Table 5.1 comparison of botnet detection in existing and proposed system.

When number of packet increase the rate of anomaly detection rate will increase .