ISSN No. 0976-5697

Volume 2, No. 2, Mar-Apr 2011

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

# Automated Signature Generation for Internet attacks using Hybrid Intrusion Detection System

Prof. S.S.Manivannan
School of Information Technology & Engineering
VIT University, Vellore -632014.
Tamil Nadu, India.
manivannan.ss@vit.ac.in

*Abstract:* As the usage of internet systems get expanded, the probability of undetected intrusions from internet are increasing in day by day. This paper proposes a hybrid approach for detecting the internet intrusions that not only maximizes the detection rate of intrusions but also reducing the false alarm rate. Whenever the connection episode from internet exceeds certain standard boundaries, rules are automatically generated using SNORT and stored in the rules database. The generated signature is mapped with the rules database and the corresponding intrusion (s) is detected if the match exists. The overall false alarm rate occurring in detecting the intrusion (s) is reduced by fixing the medium level of threshold for boundary conditions. The entire system is implemented in real time using winPcap, javaPcap and SNORT tools**.**

***Keywords:*** intrusion detection; signature generation; false alarm rate; hybrid system; boundaries; SNORT rules.

## INTRODUCTION

Today internet plays vital role in education, research, entertainment and sports. Internet can be used for both good and bad purposes. The type of the attacking mechanism to attack the systems using internet is increasing in various ways. An important security product that has emerged is Intrusion Detection Systems. The IDS protects an organization from malicious attacks from the Internet if someone tries to break in through the firewall. An intrusion detection system monitors network traffic and monitors for suspicious activity and alerts the system or network administrator. In some cases the IDS may also respond to anomalous or malicious traffic by taking action such as blocking the user or source IP address from accessing the network. IDS come in a variety of flavors and approach the goal of detecting suspicious traffic in different ways. There are network based (NIDS) and host based (HIDS) intrusion detection systems. Hybrid IDS that detect based on looking for specific signatures of known threats and anomalous behavior of unknown threats.

The rest of the paper is organized as follows: In section 2, the related works of the proposed system is given. In section 3, architecture of the proposed system is carved. In section 4, the packet tracing is discussed. In section 5, SNORT rules are discussed. Next, the Pattern Matching algorithm is proposed in section 6. In section 7, Anomaly Detection is discussed. In section 8, the working principle of Hybrid IDS engine is presented. Section 9 the results of detected intrusions are described. Section 10 provides the conclusions and the future works.

## RELATED WORKS

A variety of techniques have been proposed for detecting the internet attacks. Krishnamoorthi, N.V. S.Reddy , Acharya [2], in their work proposed an hybrid approach for modeling intrusion detection system (IDS) i.e,the rule based classifier and simple K-means clustering are combined as a hybrid intelligent system. The initial prototype developed by the rule base classifier improves the performance of K-means clustering. Kemmerer, Vigna proposed the Hi-DRA, a network surveillance, analysis, and response system [3] for high-speed WANs that provides a framework for the modular development of intrusion detection sensors in heterogeneous, high-speed environments. In addition, the system provides an infrastructure that supports the dynamic configuration of the sensors and the collection and interpretation of their results. The system, as a whole, is able to provide fine-grained monitoring across WANs and, at the same time, is able to correlate the results of the analysis of the different sensors into a high-level expressive description of security violations.

Yasinsac.A, Childs proposed a novel tool for analyzing classical cryptographic protocols [4] can be used to model and analyze the more complex Internet security protocol families and modifications made in the tool to illuminate the flaws in the Transport Layer Security (TLS) protocol.

Intrusion Detection Systems (IDS) have improved steadily in the efficiency and effectiveness with which they detect intrusive activity. This is particularly true with signature-based IDS due to progress with intrusion analysis and intrusion signature specification. At the same time system complexity, overall numbers of bugs and security vulnerabilities have been on the increase. This has led to the recognition that in order to operate over the entire attack space, multiple heterogeneous IDS must be used, which need to interoperate with one another, and possibly also with other components of system security. This paper describes our research into developing algorithms for attack signature matching for detecting multi-stage attacks manifested by alerts from heterogeneous IDS. It describes also the testing and preliminary results of that research, and the administrator interface used to analyze the alerts produced by the tests and the results of signature matching.

Yong Tang, Shigang Chen proposed two algorithms based on Expectation-Maximization (EM) and Gibbs Sampling for efficient computation of PADS from polymorphic worm samples and double-honeypot system [5], which is able to automatically detect new worms and isolate the attack traffic and introduced position-aware distribution signature (PADS), which fits in the gap between the traditional signatures and the anomaly-based systems. The new signature is capable of handling certain polymorphic worms.

Hybrid Intrusion Detection systems, internet Frequent Episode Rules (FER) from internet connections that are generated .The known FER anomaly connection episode is converted to signature and stored in signature database. The signature matching engine detects intrusion if it finds a match between incoming signatures with stored signature. Anomalies are clustered based on the categories of attack such as DoS, U2R, R2L, probe [1].

## SYSTEM ARCHITECTURE

The overall architecture of Hybrid Intrusion Detection system is shown in the Fig.1. The GUI is designed to interface with the hybrid IDS engine. The Hybrid intrusion detection system combines the advantages both misuse detection system and anomaly detection system. In the misuse detection system the data packets coming from the internet connection n are traced using winPcap and SNORT tools. The SNORT rules are downloaded from the website and the SNORT rules are stored in the database. The various field values of traced packet are extracted and matched with the already stored SNORT rules field values to check for a match. If the match exists then the corresponding intrusions are detected.

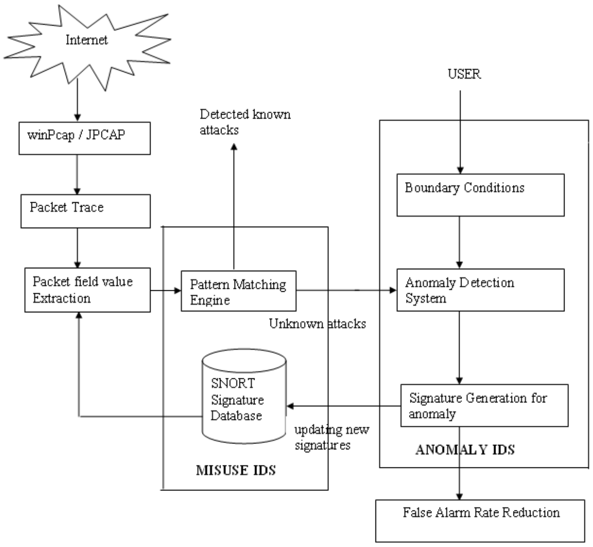


Figure.1 Architecture of Hybrid Intrusion Detection System

In anomaly detection system whenever the internet connection exceeds certain standard boundary conditions, the rules are automatically generated using SNORT and these rules are updated in the SNORT rules database.

## PACKET TRACING

Normally packets from Internet will reach Network Interface Card ( NIC ) and reach to operating system and finally the application service used by the user [6]. By means of placing winPcap ( windows Packet capture ) & JPCAP tools between operating system and NIC the packets from Internet are captured. Fig.2 shows the packet capture mechanism from internet connection with the use of winPcap and JPCAP with the use of SNORT tools.

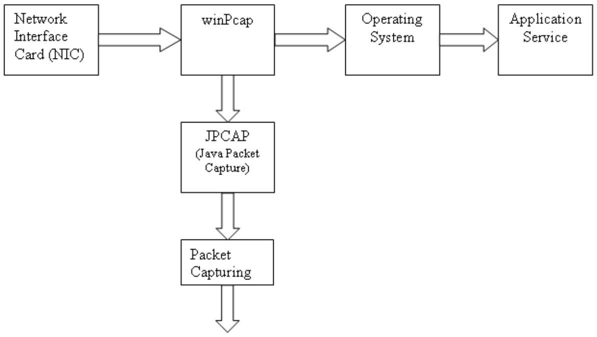


Figure. 2 Packet Tracing

The packets coming from the internet are traced with the fields such as source IP address, source port number, destination IP address , destination port number and data content etc…

## SNORT RULES

SNORT is the open source industrial standard intrusion detection and prevention systems. It performs real time traffic analysis, content searching or matching on various protocols. Variety of attacks like buffer overflow, DoS, U2R, R2L, probe can be detected. These rules can be downloaded from the website named www.snort.org. Some of the rules used in our proposed system are described below.

Rule format

alert ip any any -> any any (msg: “alert message";)

SNORT Rules

In http related attacks GET can be given as hexa decimal nos like 47, 45, 54

alert tcp 192.168.1.0/24 any -> ![192.168.1.0/24] any \ (content: "|47 45 54|"; msg: "GET matched";)

- To find large size packet

alert ip any any -> 192.168.1.0/24 any (dsize: > 6000; \ msg: "Large size IP packet found";)

- To find Source & Destination has same IP address

alert ip any any -> 192.168.1.0/24 any (msg: "Same IP address"; \ sameip;)

- To find the specific keyword matching

alert ip any any -> 192.168.1.0/24 any (content-list: \"porn"; msg: "Porn word matched";)

- To find the routing attacks

alert ip any any -> any any (ipopts: lsrr; \msg: "Loose source routing attempt";)

- To find the secret information

alert tcp 192.168.2.0/24 23 <> any any \(content: "confidential"; msg: "Detected confidential";)

## PACKET FIELDS EXTRACTION

From the traced entire packet content the various packet fields like source IP address, destination IP address, source port number, destination port number and data content are extracted separately. As well as from the updated SNORT rules source IP address, destination IP address, source port number, destination port number and data content are extracted and stored in the array.

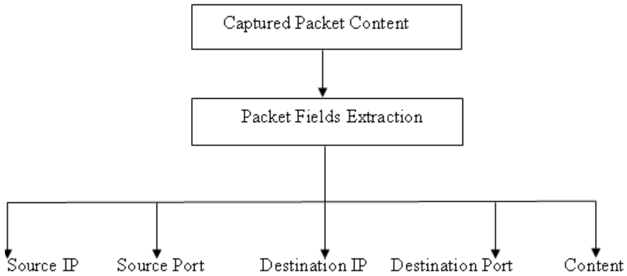


Figure. 3 Packet Fields Extraction

The extracted packet fields from the traced packet content and the SNORT rules are stored in different array.

## VII. PATTERN MATCHING ALGORITHM

The extracted packet field values from the entire packet content are matched with already stored SNORT rules and checks for a match using pattern matching algorithm [7]. If the match is found then the corresponding intrusions (s) is displayed in the daily created log files. The following algorithm matches the fields of captured packet content and SNORT rules.

```
Pattern Matching Algorithm
Input     : captured packet content.
Output    : corresponding intrusions of each captured packet.
Step 1    : Get captured packet content ( )
Step 2    : If (captured packet content = found)
      2.1 : Write the contents into a file
Step 3    : If (contents are writing in a file)
      3.1 : Extract the source IP, source port
      3.2 : Extract the destination IP, destination port, content
Step 4    :Get the fields from SNORT rules database
Step 5    : If (SNORT rules fields= fields of captured packets)
      5.1 : Display the alert message about Intrusion
Step 6    : End.
```

## VIII. ANOMALY DETECTION BY BOUNDARY

The behavior of the internet is continuously monitored in the graph created by the user interface. For every and every internet connection the minimum and maximum number of packets captured for a specified milli seconds (say 20 ms) is calculated. From the minimum and maximum number of packets captured, the average number of packets can be calculated and the line is drawn in the graph. Network behavior of an end host can be abstracted as a series of connections to/from that host.

### *A.* Feature Extraction

The following features are extracted from the internet connection.TCP connection; each UDP packet is a connection. Each connection can be represented by a vector of one-dimension variables: X=(X1, X2,… Xn)

Features are as follows

- Duration of the service
- Transport protocol type
- Type of Service
- Outgoing or Incoming packet
- Data size
- Time since last connection, if the remote host is visited before, etc
- Aggregated features of connections
- No of connections per minute
- Model of network behavior

### *B.* Boundary Conditions for Anomaly

For each and every internet connection certain standard boundary conditions are exists. If the connections exceed the boundary conditions then the corresponding anomalies are detected. Some of the boundary conditions used in the hybrid intrusion detection system are listed below.

- Max no of connections to the same source.
- Max no of bytes to be transferred (1500 bytes) from internet at a time
- Abnormal port references apart from standard port references.

e.g..21-FTP, 80 - HTTP, 25–SMTP, 23 –TELNET

- Packet payload size
- Packet window size
- Packet information content
- If EXPN (expand) is used in SMTP then there is a chance of buffer overflow.
- Port numbers greater than 255 is abnormal port references for TCP

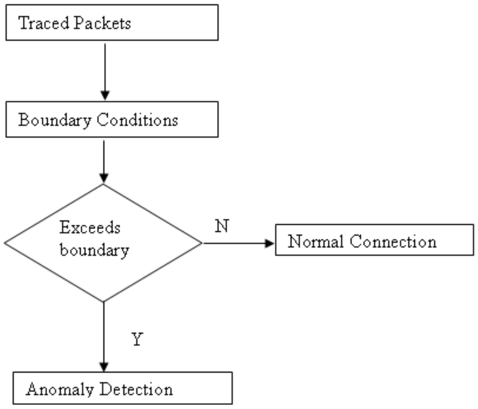


Figure.4. Anomaly Detection by Boundary Conditions

## IX. HYBRID IDS ENGINE

Hybrid IDS engine is detecting the intrusions by matching the generated signature with the SNORT rules for the detected anomaly.

### A. *Signature Generation*

Whenever the anomalies are detected by anomaly detection engine, the rules are automatically generated using SNORT and these rules are updated in the SNORT rules database. Some of the signatures generated when it is detected by the hybrid IDS engine are listed below.

- If the incoming IP packet exceeds 1500 bytes/ sec
  "large size IP packet"
- If the port number greater than 255 is used
  "Port no above 255 is used"

### B. *Hybrid IDS Engine*

Hybrid Intrusion detection engine combines the advantages of both statistical based anomaly detection and signature based misuse detection [8] and [9]. The following Fig.5 shows the working of the hybrid IDS engine. The attacks are clustered [10] as U2R, R2L,DoS and probe attacks based on their attack strategy.

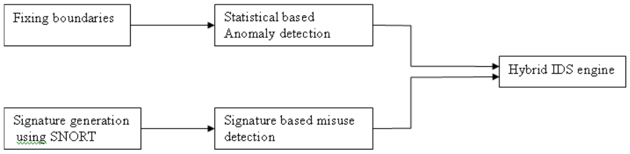

Fixing boundaries
Statistical based Anomaly detection
Hybrid IDS engine
Signature generation using SNORT
Signature based misuse detection

Figure.5 Hybrid IDS Engine.

### C. *Reducing the False Alarm Rate*

False alarm rate is the ratio of no of false attacks detected by the IDS system to the total no of attacks entered into the system. The overall false alarm rate is reduced by fixing the medium level of threshold value and ignoring the unwanted packets that are coming from the internet. The following Fig.6 shows the method of reducing the false alarm rate.

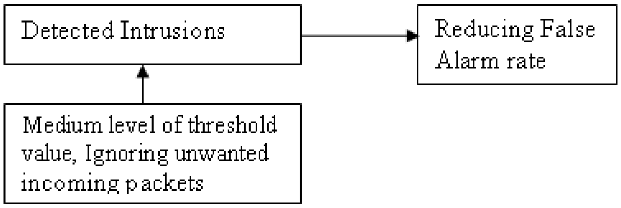

Detected Intrusions
Reducing False Alarm rate
Medium level of threshold value, Ignoring unwanted incoming packets

Figure.6 Reducing the False Alarm Rate

Some of the techniques used for reducing the false alarm rate are described below.

- KNN classifier is used to classify the false alarms from the right alarms by calculating support and confidence values and Fixing medium level of threshold value for boundaries.
- If threshold is made too low then it causes high no of false positive.

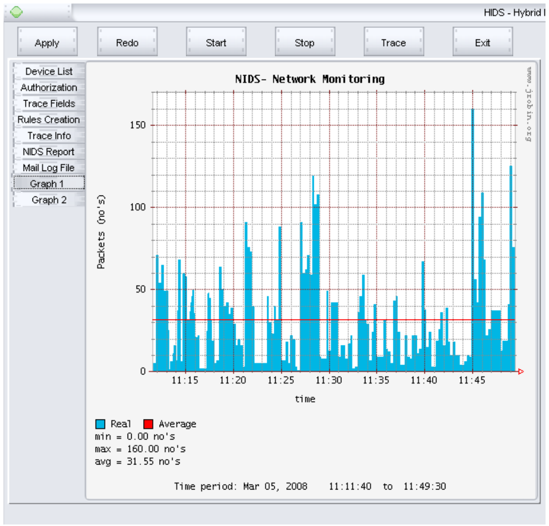
Ex. Max no of bytes can be transferred from internet is 1500 bytes / sec.

- If we made the threshold value is 1000 or 1200 bytes / sec then for each and every time it generates an alarm when transfer rate exceeds 1000 or 1200 bytes / sec
- Most of the scan traffics are due to connections to ports on which no service is running.
- These connections will not contain more than 3 packets & does not cause any attack (only disturbance). So it can be ignored.
- In contrast, majority of legitimate users tends to last longer than 3 packets
- Approximately 60% of all connections involved more than 3 packets.

## X. PROJECTION OF INTRUSION DETECTION RESULTS

### A. *Monitoring Internet Behavior*

The behavior of the internet is continuously monitored in the daily, hourly, minute basis through the graph. From the graph the abnormal behavior of the system is determined.



min = 0.00 no's
max = 160.00 no's
avg = 31.55 no's

Time period: Mar 05, 2008 11:11:40 to 11:49:30

Figure.7 Monitoring the Internet Behavior in GUI

From the above graph the internet behavior is monitored for per day , per hour and per  minute and the graph is shown in the Fig. 8

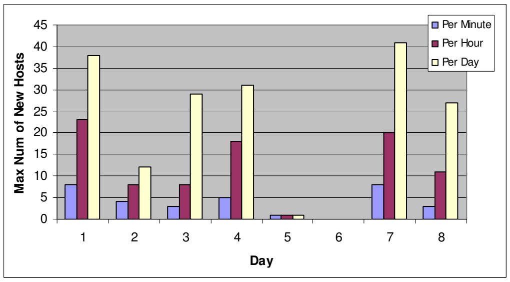


Fig.8  Internet Behavior (Day / Hour / Minute)

The TCP internet connection for week wise is analyzed using the graph.

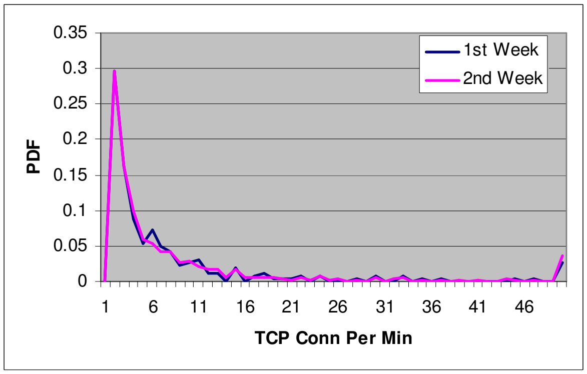


Figure.9  TCP Connection Behavior ( Week wise)

## B. *Updating SNORT rules*

The Snort rules are updated in a period basis. The Snort rules are updated once in every 2 months or 3 months. These rules can be obtained in the Snort website. These rules are copied and pasted into the importing the SNORT rules page. The various fields of the snort rules like source IP address, source port number, Destination IP address, port number and content are extracted and those fields are updated in the SNORT rules database.

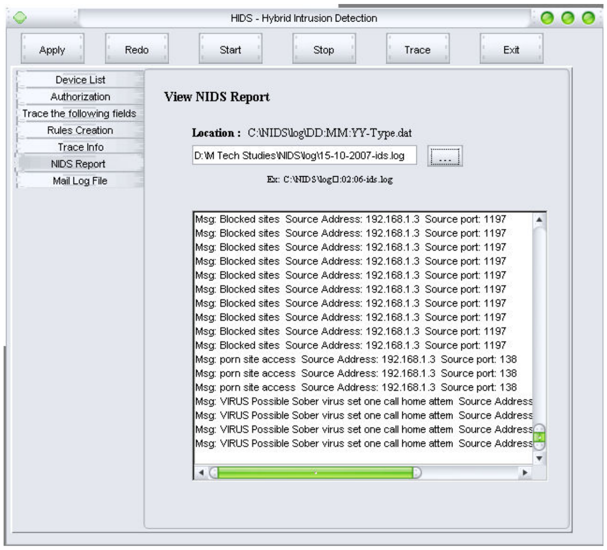
Table.1  Updating  SNORT rules

| S.No | Source IP | S.Port | Dest.IP | Dest Port | Content |
|---|---|---|---|---|---|
| 1 | 69.61.39.142 | Any | 64.111.206.194 | Any | Illegal site |
| 2 | 217.160.58.5 | Any | $HOME_NET | Any | Illegal site |
| 3 | 85.12.16.105 | Any | $HOME_NET | Any | Buffer Overflow |
| 4 | 64.111.206.194 | Any | $HOME_NET | Any | Sober Virus |
| 5 | $EXTERNAL_N | 20432 | $HOME_NET | Any | Sober Virus |

## C. *Intrusion Detection Results*

The entire system is implemented in real time using winPcap, JPCAP and Java. The incoming captured packet fields are matched with SNORT rules database field and checks for a match, if a match is found it will detect the corresponding attack(s). The attacks are stored in the daily created log files.

These attacks shows about the type of the attack , attackers source IP address, source port number. These attacks are categorized as Denial of Service ( DoS)  attack, Probe  attacks, User to Remote ( U2R) attacks, Remote to Local (R2L) attacks. From the detected attacks the detection ratio of DoS, U2R, R2L and Probe attacks are determined. The detection ratio graph is drawn for these attacks.



Location : C:\NIDS\log\DD:MM:YY-Type.dat

D:\M Tech Studies\NIDS\log\15-10-2007-ids.log

Ex: C:\NIDS\log□:02:06-ids.log

```
Msg: Blocked sites  Source Address: 192.168.1.3  Source port: 1197
Msg: Blocked sites  Source Address: 192.168.1.3  Source port: 1197
Msg: Blocked sites  Source Address: 192.168.1.3  Source port: 1197
Msg: Blocked sites  Source Address: 192.168.1.3  Source port: 1197
Msg: Blocked sites  Source Address: 192.168.1.3  Source port: 1197
Msg: Blocked sites  Source Address: 192.168.1.3  Source port: 1197
Msg: Blocked sites  Source Address: 192.168.1.3  Source port: 1197
Msg: Blocked sites  Source Address: 192.168.1.3  Source port: 1197
Msg: Blocked sites  Source Address: 192.168.1.3  Source port: 1197
Msg: Blocked sites  Source Address: 192.168.1.3  Source port: 1197
Msg: porn site access  Source Address: 192.168.1.3  Source port: 138
Msg: porn site access  Source Address: 192.168.1.3  Source port: 138
Msg: porn site access  Source Address: 192.168.1.3  Source port: 138
Msg: VIRUS Possible Sober virus set one call home attem  Source Address
Msg: VIRUS Possible Sober virus set one call home attem  Source Address
Msg: VIRUS Possible Sober virus set one call home attem  Source Address
Msg: VIRUS Possible Sober virus set one call home attem  Source Address
```

Figure.10   Detected Intrusions from Internet.

The following graph shows the detection ratio of various attack types.

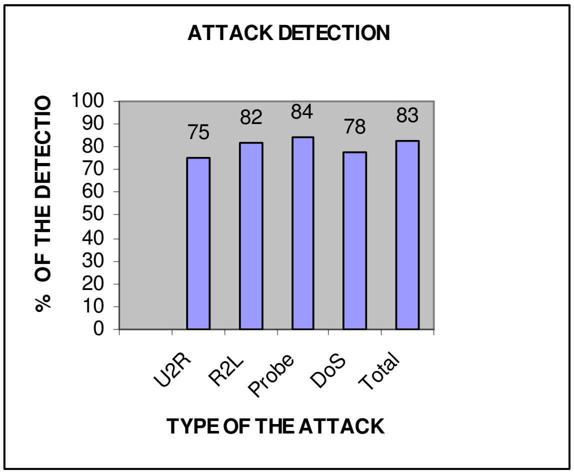


Figure.11  Detection Rate of attack types

From the above graph the detection rate of various attack types are determined. The total detection rate of the system is 83 % which is very high when compared to existing systems.

## XI. CONCLUSION AND FUTURE WORK

In order to implement the Hybrid Intrusion Detection Systems on Internet a Graphical user Interface is created. For the hybrid intrusion detection, whenever the packets are captured from internet exceeds certain standard conditions the rules and signatures are automatically generated using SNORT and are updated in the SNORT rules database. These generated signatures are matched with the SNORT rules database and the corresponding intrusion(s) are detected if the match exists. The overall false alarm rate occurring in the hybrid intrusion detection is reduced by fixing the medium level of threshold value and ignoring the unwanted packets.

In future work, the Intrusion Detection system which is used to implement in a distributed system environment with efficient usage of memory space and minimum false alarm rate of intrusion.

## XII . REFERENCES

[1] Kai Hwang, Min Cai, Ying Chen And Min Qin, "Hybrid Intrusion Detection With Weighted Signature Generation Over Anomalous Internet Episodes", IEEE Transactions On Dependable And Secure Computing , Vol. 4, No.1, Page No 1-15, January -March 2007.

[2] Krishnamoorthi,N.V..S.Reddy ,Acharya, ,"A Two-stage Hybrid Model for Intrusion Detection", International Conference on Advanced Computing and Communications, Page Nos : 163-165, Dec 2006.

[3] Kemmerer, R.A, Vigna, G, "Hi-DRA: intrusion detection for Internet security", Proceedings of the IEEE [4] Yasinsac.A, Childs, "Analyzing Internet security protocols", Sixth IEEE International Symposium on High Assurance Systems Engineering, Page(s): 149-159, March 2001.

[5] Yong Tang, Shigang Chen, "Defending Against Internet Worms: A Signature- Based Approach", Proceedings of the IEEE Conference Page(s): 1-11, October 2005.

[6] Khushboo Shah, Edmond Jonckheere, Stephan Bohacek, "Dynamic Modeling of Internet Traffic for Intrusion Detection", Proceedings of the IEEE Conference Page(s):1-19, April 2005.

[7] Nathan Carey, George Mohay and Andrew Clark, "Attack Signature Matching and Discovery in Systems Employing Heterogeneous IDS**",** Page(s): 1-15, March 2005.

[8] Amitava Biswas, Purnendu Sinha, "On improving the performance of Network Intrusion Detection Systems by efficient packet capturing", Network Operations and management Symposium, Page No 1- 4, Nov 2006.

[9] Iosif-Viorel Onut and Ali A. Ghorbani, "A Feature Classification Scheme for Network Intrusion Detection", International Journal of Network Security, Vol.5, No.1, Page No.1–15, July 2007.

[10] Shi zhong, taghi khoshgoftaar, Nameem seliya, "Clustering based Network Intrusion Detection", International Journal of Reliability, Quality & Security Engg ,Vol-14, No.2, Page Nos 169-187, March 2007.