A Survey Intrusion Detection with KDD99 Cup Dataset

[1]Shreya Dubey, [2]Jigyasu Dubey

Department of Information Technology
Rajiv Gandhi Proudyogiki Vishwavidyalaya Indore, India

Department of Information Technology
Shri Vaishnav Institute of Technology & Science
Indore, India

Abstract: In this paper we discuss about Intrusion Detection System with their basic terminology. It also focuses on services and characteristics of detection system. The main function of IDS is to distinguishing normal and abnormal pattern of input data. An IDS is a software or hardware device that deals with attacks by collecting information from a variety of system and network sources and then analyse security problems.

Keywords: detection rate, false alarm rate, intrusion detection, KDD Cup 99, KDD attributes, confusion matrics.

I. INTRODUCTION

Intrusion Detection System (IDS) can be defining as a system which detects intrusive activities. These intrusive activity compromises security principles like integrity, confidentiality and availability of resources. To control intrusive activity, detection systems are employed thus it named as Intrusion Detection System [1]. The main function of IDS is to distinguishing normal and abnormal pattern of input data. An IDS is a software or hardware device that deals with attacks by collecting information from a variety of system and network sources and then analyse security problems.

Every computer is always at risk for unauthorized and intrusion, however, with sensitive and private information are at a higher risk. Detecting an intrusion is a key technique in information security. It plays an important role in detecting different type of attacks and secures the system. Intrusion detection is the process of observing and analysing the events arising in a computer or network system to identify all security problems. IDS provide three important security functions: monitor, detect & response. IDS monitor the operation of firewalls, routers and other security mechanisms. [4]

Intrusion Detection system usually provides the following services:

- Observing and analyzing computer and/or network system activity.[4]
- Audit the system configurations and vulnerabilities.[4]
- Evaluating the integrity of critical system and data files.[4]
- Estimating abnormal activities.[4]

Characteristics of Intrusion Detection Systems:

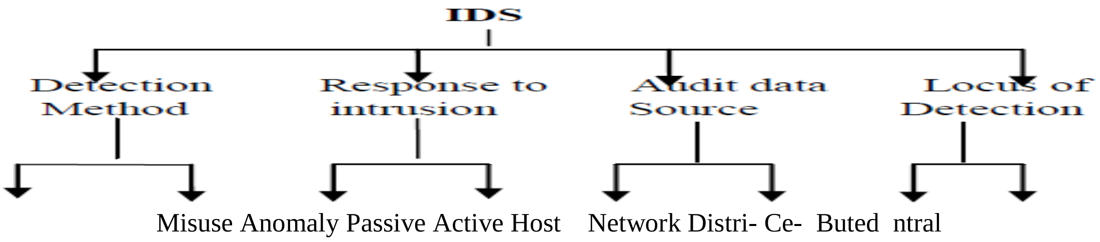


Figure 1.1: Classify IDS [6]

Basic Components of IDS:

Modern IDSs are extremely diverse in the techniques they employ to gather and analyze upon data. Basic Component of IDS comprises: a detection module which gathers data that may contain evidence of intrusions, an analysis engine which processes this data to identify intrusive activity, and a response component reports intrusions. To understand the complete detection system the basic flow of IDS is describe in figure 1.2. [5]

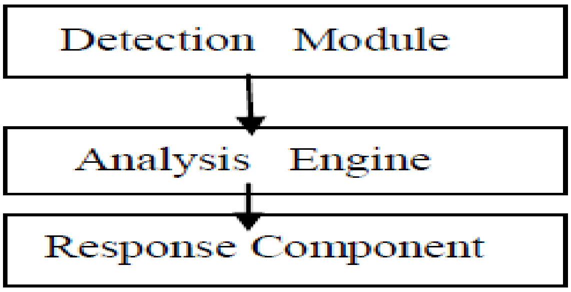


Figure 1.2: Basic Components of IDS [5]

II. IDS TAXONOMY

IDSs are divided into two broad categories; host-based (HIDS) and network-based (NIDS). A host-based IDS requires small programs (or agents) to be installed on individual systems to be supervised. The agents monitor the operating system and write down data to log files and/or usually consist of a network application (or sensor) with a Network Interface Card (NIC) working in promiscuous mode and a separate management of interface. IDS is placed on a network segment or boundary and monitor all traffic on that segment. The current technology in intrusion detection is to combine both host based and network based information to develop hybrid systems that have more efficient. [4]

Types of IDS generally based on their usage. We can install IDS on a network, on a host or combination of both. Figure 2.1 shows types of IDS. [4]

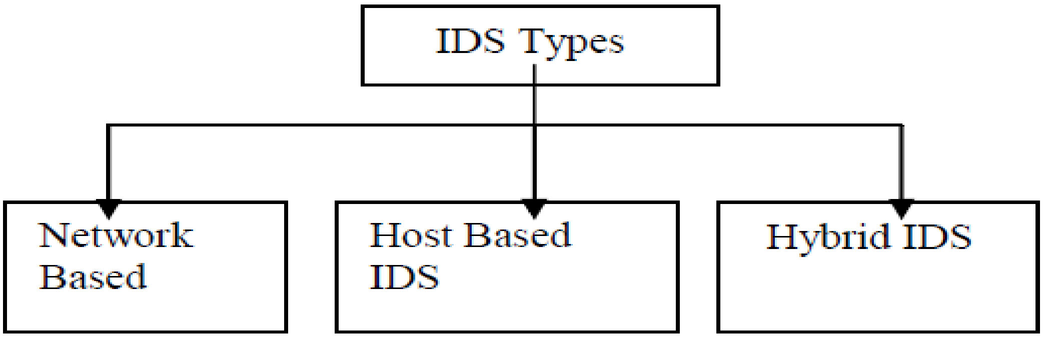


Figure 2.1: Types of IDS [2]

Network Based IDS

NIDS are IDS that operate as stand-alone devices on a network. It uses monitoring a port, when placed next to a networking device like hub, switch. NIDS works on the principle of signature matching i.e. comparing attack patterns to known signatures in their data base. Types of NIDS include Snort, Cisco NIDS and Netprowler. [2]

Host Based ID

SHIDS are IDS that operate on a single workstation. It monitors traffic on its host machine by utilizing the resources of its host to detect attacks. Types of HIDS include Tripwire, Cisco HIDS and Symantec ESM. It Work on the principle of configuration and change management. An alert is triggered when file attribute change, new file created or existing files deleted. [2]

Hybrid IDS

HIDS and NIDS can be combined to form a separate hybrid class of Network Node IDS (NNIDS). In NNIDS agents are deployed on every host within the network being protected. A NNIDS operates much like a hybrid per-host NIDS since a single agent processes the network traffic directed to the host it runs upon. [2]

IDS Detection Techniques

There are two general approaches to intrusion detection: misuse based detection and anomaly based detection which is described in figure 2.2. These approaches develop the core of several currently present intrusion detection techniques. [3]

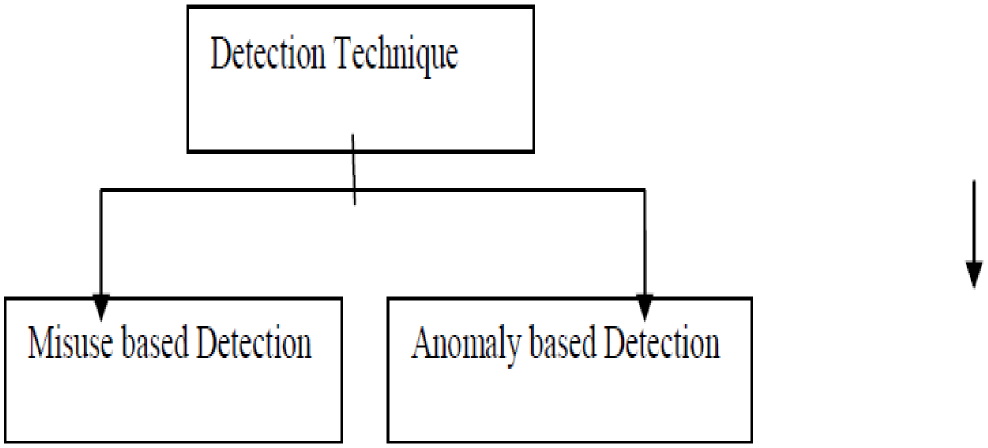


Figure 2.2: Types of Detection [3]

Misuse Based Detection

Misuse Detection centers on using an expert system to identify intrusions based on a predetermined knowledge base. These systems are capable of attaining high level of accuracy. It is also referred to as signature based detection because alarms are generated based on specific attack signatures.

The advantage of misuse detection is the ability to generate accurate result and having fewer false alarms. [4]

The disadvantage of misuse detection is that, it is incapable of detecting intrusions that are not represented in its knowledge base. They will detect only the known attacks. [4]

Anomaly Based Detection

Anomalies also known as outliers, exceptions or peculiarities are patterns in data that do not conform to a well defined notion of normal behavior of system. It can be either Static or Dynamic. In static, it is assumed that the portion of data or system behavior remains constant or static. It can be represented as a binary bit string such as files [3]. If this portion ever deviates from its original form, either an error has occurred or an intruder has altered the static portion of the system.

In Dynamic, the definition of behavior is included. System behavior is defined as a sequence of distinct events. Example, Audit Records produced by OS. Anomaly detection is an important tool for fraud detection, network based intrusion and other unusual events that have great significance but they are hard to find. Anomaly detection is also sometimes referred to as behavior-based detection because it associates with variations from user behavior [4].

The advantage of anomaly detection approach is the ability to detect novel attacks or unknown attacks based on audit data. [4]

The main drawback of the anomaly detection approach is that well-known attacks may not be detected. [4]

III. KDD DATA SET

KDD cup'99 data set is widely used and publically available data set for Network-based anomaly detection. This data set is prepared by Stolfo et al. and is built based on the data captured in DARPA'98 IDS evaluation program. DARPA'98 is about four gigabytes of compressed raw (binary) tcp dump data. KDD training dataset consists of approx 4,900,000 single connection vectors, each of which contains 41 features and is labeled as either normal or an attack. The datasets contain a total number of 24 training attack types, with an additional 14 types in the test data only. Attacks fall in four categories as – DoS, U2R, R2L and Probing. [1]

These attacks are as follows [1]:

- **Denial of Service (DoS)**

  The first category of attacks is DoS attacks. This type of attacks is that attackers attempt to disrupt a host or network resource in order to make legitimate users not be able to access to that computer service. The victim machines can be any network system such as web server, domain name system server, mail server, and so on. In the KDD99 data set, many common forms of DoS attacks are included. For example, over 70% attacks in this category are smurf attack [14].

  In short, the attacker makes some computing resources too busy or memory resources too full to handle authorized requests, or **DENIES** unauthorized user access to a machine. DoS attacks are classified based on services like apache2, land, mail, back etc [1].

- **Remote to Local (R2L)**

  The attacker who does not have an account on a remote machine send packets to that machine over a network and exploits some vulnerability to gain local access as a user of that machine which include send-mail & lock. [1]

  This type of attacks is that the unauthorized attackers through networks gain local access as a user of local machine and then exploit the machine's vulnerabilities. Totally 15 types of R2L attacks are included in the KDD99 data set. For example ftp_write attack is that the attackers create rhost file to make anonymous FTP directory writable and finally obtain local login to the system. The guess_passwd is that the attackers try to gain access to a user's account by repeatedly guessing the possible passwords. [14]

- **User to Root (U2R)**

  The attacker starts out with access as a normal user on the system and becomes a root User by exploiting vulnerabilities to gain root access to the system. [1]

  The attacker pretends as a legitimate user of a system without authorization and then exploits the system's vulnerabilities to get root access to that system. The KDD99 data set consists of eight different types of U2R attacks and the most common seen buffer_overflow attack is one of them. [14]

- **Probing**

  An attacker with a map of the machines and services that is available on the network can use this information to look for exploits [1].

  By using programs, to automatically scan a large amount of network IP addresses, the attacker can explore vulnerabilities of computers. Once any vulnerability is found, the attacker can thus gain the access to the system and start to gather information without authorization. The KDD99 data set collects six scanning attacks of this category. They are ipsweep, mscan, nmap, portsweep, saint, and Satan. [14]

IV. KDD ATTRIBUTES

41 attributes of KDD is divided into 3 groups; intrinsic attributes, content attribute and traffic attributes. The content of these attributes is described next: [8]

4.1 Intrinsic attributes: These attributes are extracted from the headers' area of the network packets as shown in table 4.1.

Table 4.1: Intrinsic Attributes [8]

No.	Name	Type	Description
1	duration	integer	duration of the connection
2	protocol_type	nominal	protocol type of the connection: TCP, UDP and ICMP
3	service	nominal	http, ftp, smtp, telnet... and other (if not much used service)
4	flag	nominal	Connection status. The possible status are this: SF, S0, S1, S2, S3, OTH, REJ, RSTO, RSTOS0, SH, RSTRH, SHR
5	src_bytes	integer	bytes sent in one connection
6	dst_bytes	integer	bytes received in one connection
7	land	binary	if source and destination IP addresses and port numbers are equal then, this variable takes value 1 else 0
8	wrong_fragment	integer	sum of bad checksum packets in a connection
9	urgent	integer	sum of urgent packets in a connections as activated UB

4.2 Content attributes: These attributes are extracted from the contents area of the network packets based on expert person knowledge. This is shown in table 4.2.

Table 4.2: Content Attributes [8]

NO.	**Name**	**Type**	**Description**
10	hot	integer	sum of hot actions in a connection such as: entering a system directory,creating programs and executing programs
11	num_failed_logins	integer	number of incorrect logins in a connection
12	logged_in	binary	if the login is correct then 1 else 0
13	num_compromised	integer	sum of times appearance “not found” error in a connection
14	root_shell	binary	if the root gets the shell then 1 else 0
15	su_attempted	binary	if the su command has been used then 1 else 0
16	num_root	integer	sum of operations performed as root in a connection
17	num_file_creations	integer	sum of file creations in a connection
18	num_shells	integer	number of logins of normal users
19	num_access_files	integer	sum of operations in control files in a connection
20	num_outbound_cmds	integer	sum of outbound commands in a ftp session
21	is_hot_login	binary	if the user is accessing as root or adm
22	is_guest_login	binary	if the user is accessing as guest, anonymous or visitor

4.3 Affic attributes: These attributes are calculated taking into account the previous connections. These are divided into two groups: (1) time traffic attributes (2) machine traffic attributes. The difference between one group and the other is the mode to select the previous connections. [8]

***4.3.1 Time traffic attributes*:** To calculate these attributes we considered the connections that occurred in the past 2 seconds.

Table 4.3.1: Time traffic attributes [8]

NO.	**Name**	**Type**	**Description**
23	count	integer	sum of connections to the same destination IP address
24	srv_count	integer	sum of connections to the same destination port number
25	serror_rate	real	the percentage of connections that have activated the flag (4) s0, s1, s2 or s3,among the connections aggregated in count (23)
26	srv_serror_rate	real	the percentage of connections that have activated the flag (4) s0, s1, s2 or s3,among the connections agg. in srv_count (24)
27	rerror_rate	real	the percentage of connections that have activated the flag (4) REJ, among the connections agg. in count (23)
28	srv_error_rate	real	the percentage of connections that have activated the flag (4) REJ, among the connections agg. in srv_count (24)
29	same_srv_rate	real	the percentage of connections that were to the same service, among the connections agg. in count (23)
30	diff_srv_rate	real	the percentage of connections that were to different services, among the connections agg. in count (23)
31	srv_diff_host_rate	real	the percentage of connections that were to different destination machines agg. in count (24)

4.3.2 Machine traffic attributes: To calculate these attributes we took into account the previous 100 connections.

Table 4.3.2: Machine traffic attributes [8]

NO.	**Name**	**Type**	**Description**
32	dst_host_count	integer	sum of connections to the same destination IP address

33	dst_host_srv_count	integer	sum of connections to the same destination port number
34	dst_host_same_srv_rate	real	the percentage of connections that were to the same service, among the connections aggregated in (32)
35	dst_host_diff_srv_rate	real	the percentage of connections that were to different services, among the connections agg. in (32)
36	dst_host_same_src_port_rate	real	the percentage of connections that were to the same source port, among the connections agg. in (33)
37	dst_host_srv_diff_host_rate	real	the percentage of connections that were to different destination machines, among the connections agg. in (33)
38	dst_host_serror_rate	real	the percentage of connections that have activated the flag (4) s0, s1, s2 or s3, among the connections agg. in (32)
39	dst_host_srv_serror_rate	real	the percent of connections that have activated the flag (4) s0, s1, s2 or s3, among the connections agg. in (33)
40	dst_host_rerror_rate	real	the percentage of connections that have activated the flag (4) REJ, among the connections agg. in (32)
41	dst_host_srv_error_rate	real	the percentage of connections that have activated the flag (4) REJ, among the connections agg. in (33)

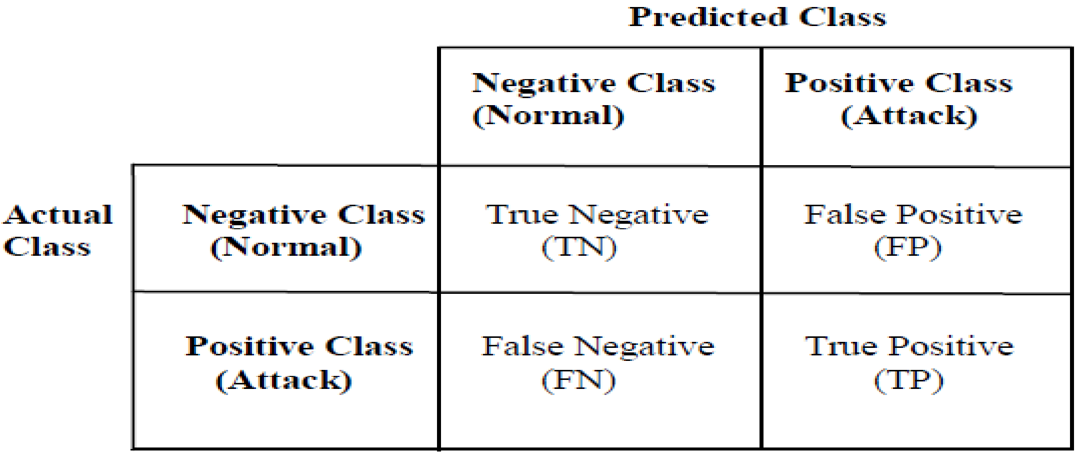
4.4 Class attribute:

The 42 attribute is the class attribute; it indicates which type of connections is each instance: normal or which attack. The values it can take are the following (view Table5): anomaly, dict, dict_simple, eject, eject-fail, ffb, ffb_clear, format, format_clear, format-fail, ftp-write, guest, imap, land, load_clear, loadmodule, multihop, perl_clear, perlmagic, phf, rootkit, spy, syslog, teardrop, warez, warezclient, warezmaster, pod, back, ipsweep, neptune, nmap, portsweep, satan, smurf and normal. [8]

V. EVALUATION APPROACH

Confusion Matrix

The effectiveness of IDS is evaluated by its ability to give a correct classification. According to the real nature of a given event and the prediction from IDS, four possible outcomes are shown in table below, which is known as the confusion matrix. True negatives and true positives correspond to a correct operation of the IDS; that is, events are successfully labeled as normal event or attacks, respectively; false positives taking as normal events being classified as attacks; false negatives are attack events incorrectly classified as normal events. [6]



		Predicted Class	
		Negative Class (Normal)	Positive Class (Attack)
Actual Class	Negative Class (Normal)	True Negative (TN)	False Positive (FP)
	Positive Class (Attack)	False Negative (FN)	True Positive (TP)

Figure 5.1: Confusion Matrix [6]

Based on the above confusion matrix, the evaluation mainly applies the following criteria to measure the performance of IDSs. [6]

1. True Negative Rate(TNR) Or Specificity: $\frac{\textbf{TN}}{\textbf{TN+ FP}}$

2. True Positive Rate(TPR) Or Detection Rate(DR) or Sensitivity: In information retrieval, this is called Recall

$$\frac{\textbf{TP}}{\textbf{TP + FN}}$$

3. False Positive Rate(FPR) Or False Alarm Rate(FAR) Or 1-specificity: $\frac{\textbf{FP}}{\textbf{TN+FP}}$

4. False Negative Rate(FNR) Or 1-sensitivity: $\frac{\textbf{FN}}{\textbf{TP+FN}}$

5. Accuracy: $\frac{\textbf{TN+TP}}{\textbf{TN+TP+FN+FP}}$

6. Precision, which is another information retrieval term and often is paired with Recall : $\frac{\textbf{TP}}{\textbf{TP+FP}}$

Where, TN = True Negative.
FP = False Positive.
TP = True Positive.
FN = False Negative.

The most popular performance metrics is detection rate (DR) together with false alarm rate (FAR). IDS should have a high DR and a low FAR. Other commonly used combinations include Precision and Recall, or Sensitivity and Specificity. [6]

VI. RELATED WORK

Author	**Publication & year**	**Classifier Name**	**Classifier Performance**
Dorothy E. Denning	IEEE transaction on Software Engineering, Vol. SE-13, No. 2, February 1987	Monitoring system's audit records, IDES model. [1]	IDES model provides a sound basis for developing powerful real-time intrusion detection. [1]
Z. Muda, W. Yassin, M.N. Sulaiman, N.I. Udzir	7th International Conference on IT in Asia (CITA), 2011	Intrusion Detection based on K-means Clustering and Naïve Bayes Classification [10]	Accuracy 99.65%,DR 99.8%,FA0.5% [10]
Te-Shun Chou, Tsung-Nan Chou	IEEE-Seventh Annual Communication Networks and Services Research Conference, 2009	Hybrid Classified Systems for Intrusion Detection [14]	DR 92.30%, FPR 3.13%, [14]
N.B. Amor, S. Benferhat, and Z. Elouedi	25-29 July, 2004 Budapest, Hungary	Qualitative Classification and Evaluation in Possibilistic Decision Trees. [15]	Classification using decision trees of objects characterized by uncertain attribute values and it is represented by qualitative possibilistic framework.
T.S. Chou, K.K. Yen, and J. Luo, Niki Pissinou and Kia Makki	IEEE 2007	Network Intrusion Detection Using Feature Selection of Soft Computing Paradigms. [16]	Their approach achieves the highest averaged accuracies with reducing the size of dataset. [16]
Mukkamala S., Janoski G., and Sung A.H.	IEEE 2002	Intrusion detection using neural networks and support vector machines. [17]	High accuracy when they compare their result with other neural-SVM based IDS systems. [17]

J. Zhang and M. Zulkernine	1st International Conference on Availability, Reliability and Security (ARES), IEEE 2006	A Hybrid Network Intrusion Detection Technique Using Random Forests. [18]	DR 94.7%, FPR 2%. [18]
Hao Wang, Yan Zhang, Danyun Li	7th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), IEEE 2010	Network Intrusion Detection based on Hybrid Fuzzy C-mean Clustering. [19]	Hybrid algorithm based on gradient descent of FCM, makes the algorithm a strong global searching capacity.

VII. CONCLUSION

In this paper we discuss about Intrusion detection system with their services and characteristics. We also discuss about types of IDS and their detection approaches. This survey paper includes KDD99 cup dataset. This data set is prepared by Stolfo and is built based on the data captured in DARPA'98 IDS evaluation program. KDD dataset have attributes which we discussed. For calculation, we have confusion matrices and various formulas. We also include those researchers who have done their work on IDS.

ACKNOWLEDGMENT

I especially thank to Mr. Jigyasu Dubey for stimulating and helpful discussion on IDS and helping me to clarify my ideas.

REFERENCES

[1]. D.E. Denning "An Intrusion Detection Model" IEEE Transaction on Software Engineering, VOL. SE-13, No. 2, February 1987.

[2]. Karthikeyan. K.R. and A. Indra "Intrusion Detection Tools and Techniques- A Survey," International Journal of Computer Theory and Engineering, Vol. 2, No. 6, December, 2010.

[3]. Douglas J. Brown, Bill Suckow and Tianqiu Wang, "A Survey of Intrusion Detection Systems".

[4]. V. Jaiganesh, S. Mangayarkarasi, Dr. P. Sumathi, "Intrusion Detection Systems: A Survey and Analysis of Classification Techniques", International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 4 april 2013.

[5]. Douglas J. Brown, Bill Suckow and Tianqiu Wang, "A Survey of Intrusion Detection Systems", Department of Computer Science, University of California, San Diego, USA.

[6]. Shelly Xiaonan Wu Wolfgang Banzhaf, "The use of Computational Intelligence in Intrusion Detection Systems: A Review", November 2008.

[7]. A.K. Jain, M.N. Murty and P.J. Flynn, "Data Clustering: A Review," ACM Computing Surveys, vol. No. 31, No. 3, Sept 1999.

[8]. Automatic Classification and Parallelism (ALDAPA) group (http://www.sc.ehu.es/acwaldap/) of Computer Architecture and Technology (KAT/ACT) department of University Of Basque Country (EHU/UPV).

[9]. Hari Om, Aritra Kundu "A Hybrid System for Reducing the False Alarm Rate of Anomaly Intrusion Detection System," 1st International Conference on Recent Advances in Information Technology RAIT-2012.

[10]. Z. Muda, W. Yassin, M.N. Sulaiman, N.I. Udzir, " Intrusion Detection Based on K-means Clustering and Naïve Bayes Classification", 7th International Conference on IT in Asia (CITA), 2011.

[11]. Sanjay Kumar Sharma, Pankaj Pandey, Susheel Tiwari and Mahendra Singh Sisodia,"An Improved Network Intrusion Detection Technique Based on K-means Clustering Via Naïve Bayes Classification," IEEE on advances in Engineering, Science and Management (ICAESM-2012) March 30,31, 2012.

[12]. V. Jaiganesh, S Mangayarkarasi, Dr. P. Sumathi, "Intrusion Detection Systems:A Survey and Analysis of Classification Techniques", International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 4th April 2013.

[13]. A.M. Chandrasekhar, K. Raghuveer, "Intrusion Detection Technique by using K-means, Fuzzy Neural Network and SVM Classifier", International Conference on Computer Communication and Informatics (ICCCI), Jan 04-06, 2013, Coimbatore, INDIA.

[14]. Te-shun Chou and Tsung-Nan Chou, "Hybrid Classified Systems for Intrusion Detection," 2009 IEEE Seventh Annual Communication Networks and Services Research Conference.

[15]. N.B. Amor, S. Benferhat, and Z. Elouedi, "Naïve Bayes vs. Decision trees in intrusion detection system," or "Qualitative Classification and Evaluation in Possibilistic Decision Trees," 25-29 July, 2004 Budapest, Hungary.

[16]. T.S. Chou, K.K. Yen, and J. Luo, "Network Intrusion Detection Design Using Feature Selection of Soft Computing Paradigms".

[17]. Mukkamala S., Janoski G., and Sung A.H., "Intrusion detection using neural networks and support vector machines," In IEEE International Joint Conferences on Neural Networks, 2002.

[18]. J. Zhang and M. Zulkernine "A Hybrid Network Intrusion Detection Technique Using Random Forests".

[19]. Hao Wang, Yan Zhang, Danyun Li "Network Intrusion Detection based on Hybrid Fuzzy C-mean Clustering," In seventh International Conference on Fuzzy Systems and Knowledge Discovery FSKD 2010.