A Review on Hybrid Intrusion Detection System using Artificial Immune System Approaches

Pavitra Chauhan
Amity University
Sector-125, Noida
India

Nidhi Chandra
Amity University
Sector-125, Noida
India

ABSTRACT

With the growing advances in the technology the uses of computer systems and the internet is also growing at a rapid rate, and with the increase in their usage vulnerabilities and threats are also increasing tremendously. A large number of approaches have been proposed till now for improving the security of a host system and a network. One of the proposed approach is an Intrusion Detection System (IDS). An IDS works for a system is referred as Host IDS and the one that works for a network is referred as Network IDS. But their functionality is specific to particular host and a network, one does not work as an alternative to another. Thus, an IDS is needed that overcomes the drawbacks of both the systems and combines their advantages to form a Hybrid Intrusion Detection System. An Hybrid IDS captures both host and network data and thereby apply an analysis approach. In order to make these systems robust and effective biologically inspired Artificial Immune System (AIS) approaches can be used that makes the system flexible enough to work in every scenario. This paper provides a review of various IDS and application of various AIS approaches to them.

General Terms

Intrusion Detection System (IDS)

Keywords

Host Intrusion Detection System (HIDS), Network Intrusion Detection System (NIDS), Hybrid intrusion Detection System, Artificial Immune System.

1. INTRODUCTION

Intrusion Detection Systems (IDS) are one of the most efficient ways of identifying the intrusions in a computer system by monitoring the activities of the target system. IDS as defined in [1], are security tools that like other mechanisms such as antivirus software, firewalls, and access control mechanisms are intended to strengthen the security of information and communication systems [5]. These tools identify intrusions by differentiating between normal and abnormal activities of the system.

In general two types of IDS have been developed: Host IDS and Network IDS, that further uses any of the anomaly or misuse detection mechanisms. Anomaly based mechanism makes a profile of normal behavior of the system and based on this it identify possible intrusions. Misuse based mechanism uses rules which are developed manually by the administrator or are preconfigured and these rules are matched against signature of the intrusions.

Thus, in order to make these IDS effective a Hybrid Intrusion Detection System can be developed that uses the concepts of both network and host. An Hybrid IDS captures data from both host and network, which can be further analyzed for possible intrusions. Since these systems will work for both host and network they prove to be more efficient than traditional IDS.

Artificial Immune Systems (AIS) being one of the most effective information processing mechanisms which simulates human immune mechanism can be used for IDS to make them more robust and effective. Human Immune System (HIS) provides protection to human at various levels thereby protecting body from several known and unknown antigens. This characteristic feature of HIS can be used in intrusion detection system thereby protecting the system at various levels, thus providing security in every possible way.

The rest of the paper is divided into following sections. Section II deals into Intrusion Detection System, Section III gives a brief overview of Hybrid Intrusion Detection System, Artificial Immune System is discussed in Section IV, Section V discusses about related work.

2. INTRUSION DETECTION SYSTEM

The evolution of intrusion detection began long ago in 1970's. Initially, intrusion detection was performed by system administrators by simply sitting in front of the console and monitoring the user activities. This form of Intrusion Detection was ad-hoc and not scalable [2]. In the late 70's and early 80's, *audit logs* were prepared on fan-folded papers, but a huge amount of amount of data was collected at the end of the week which was difficult to analyze. These audit logs were then moved online but still it was very slow. In early 90's, real-time Intrusion Detection system was developed that used to analyze the audit logs as it was produced [2].

Intrusion detection System is an active field of research in last two decades, starting in 1980's with the publication of John Anderson's Computer Security Threat Monitoring and Surveillance, as one his pioneer work. Anderson identified various threats to the security of the computer system and classified the attackers into two main broad categories:

1. *External Penetrator:* External Penetrator is an illegitimate user of a computer system who gains access to the computer system without the consent of the owner. In addition to this, Anderson defined it as an a employee of an organization who has the physical access to the building having the computer system but is not authorized to access it.
2. *Internal Penetrator:* Internal Penetrators are more frequent than external penetrator in many installations. There are three classes of users that are identified:

i.) *Masquerader:* Masquerader is the user who has gained access to the computer system, can be a external penetrator or an

internal user who has full access to the computer system and wishes to exploit other legitimate user by impersonation. In this case, there is no direct way to differentiate between a legitimate user and a masquerader.

ii.) *Legitimate User:* Legitimate User can operate as a misfeasor, who has access rights to the computer system and misuses the information and data to exploit the security.

iii.) *Clandestine User:* Clandestine User has or gains administrative control of the computer system, and works below the normal auditing mechanisms. Thus, these are difficult to detect.

Intrusions are *incidents*, which are violations to computer security policies. Incidents have various causes such as, malware, attackers to gains unauthorized access to the computer system or the authorized users who misuses their privileges or tries to escalate their privileges for which they are not authorized. *Intrusion Detection* is the process of monitoring the host or the network for the possible intrusions. *Intrusion Detection System* is software that automates the process of intrusion detection. An IDS is used to examine the malicious activity by monitoring the resources of the target system and thereby analyzing them. The ability of such system is to make an analysis of an attack in real-time which in turn will allow it to limit the adverse effects which are penetrated on the system. These systems either shut down the system or generate alarm in case of any intrusions. IDS is often seem to work as House alarm system which it does like in the sense of providing an alert when a predefined event occurs. Although its functionality are quite different from house alarm system as in it provide when and where the type of event that occurred in network/server environment and the source of the intrusion. The following fig. shows the general workflow of an Intrusion Detection system:

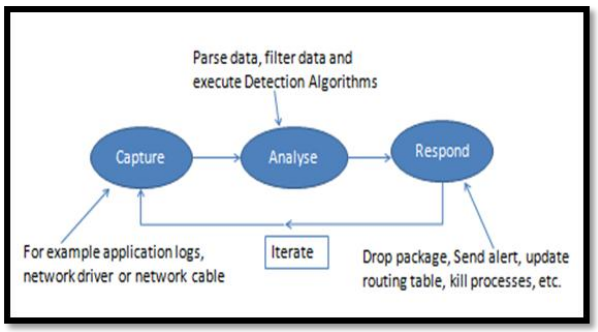

Fig 1. Basic Workflow of an Intrusion Detection System

The IDS can also be understood by studying its principle that gives the relationship between behaviors of the intruder and the authorized user. This relation is described with the help of a graphical representation, which is between probability density function across measurable behavior parameter [8].

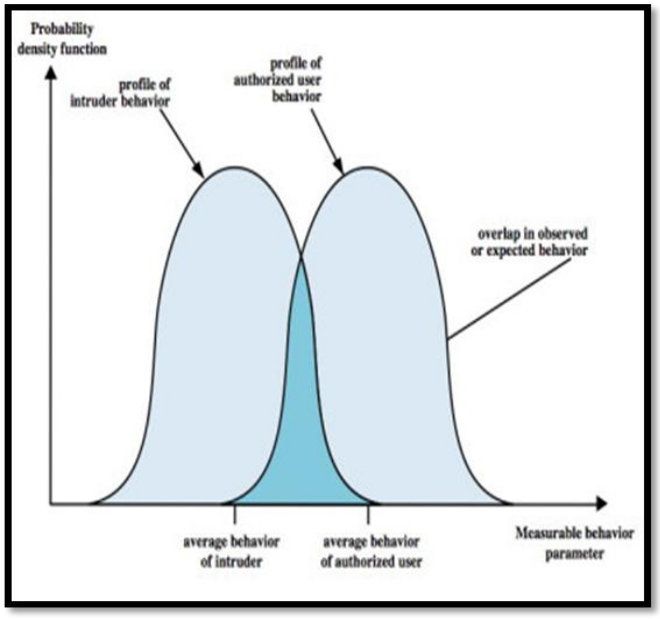

Fig 2: IDS Principle

These can be basically categorized based on analysis method or source of data used.

Based on analysis method, it can be:

i.) Misuse based IDS

ii.) Anomaly based IDS

Based on source of data used, it can be:

i.) Host Intrusion Detection system (HIDS)

ii.) Network Intrusion Detection System (NIDS)

The current state of IDS consist of two verities: Host Intrusion Detection System and Network Intrusion Detection System that uses any of the analysis methods for analyzing the data obtained through the various sources. The future state of IDS is an Hybrid Intrusion Detection System that combines the features of both HIDS and NIDS.

2.1 Host based Intrusion Detection System (HIDS)

The Host Intrusion Detection System abbreviated as HIDS deals with vulnerabilities at system or host end. This has been the first type of intrusion detection system which was designed for mainframe computer where outside interactions were infrequent.

Some common abilities of HIDS system include log analysis, event correlation, integrity checking, policy enforcement, rootkit detection and alerting. It has the capability of protecting the host systems from malicious or anomalous activity. HIDS act an agent that is installed on the Host machine which generates report on system configuration and application activity or as an agent that monitors whether anything or anyone, whether external or internal, has circumvented the system's security policy. These have the ability to baseline the system's behavior in order to identify any variation in the system configuration and application

activity. The Host based intrusion detection system requires small programs (or agents) to be installed on the individual system to generate report of individual systems revealing the systems malicious activity if found any.

To be effective in an environment with multiple host machines, HIDS is deployed in a manner that it is managed by central location. A policy is configured on the management system of the environment for the deployment to local host agents. This policy can be single for all the systems, but depending on the varying configurations of the systems it can be different for particular type of operating system, machine types, physical location and user types.

Once the policy is configured, it is then distributed and deployed to local host machines with agents installed. With this central configuration approach it has the ability to apply changes to various systems at once and thus creating a baseline for known system types. This central management architecture provides the advantage of central authentication, alerting and reporting.

The following diagram shows a small architecture which is defining how each host is secured by an agent based principle.

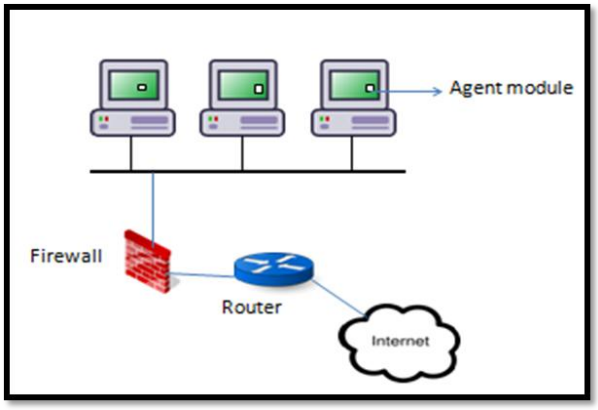

Fig 3: Generic Architecture of Host Intrusion Detection System

The various tools for HIDS includes Snare, eXpert BSM IDS, OSSEC, etc. The following table 1 [9] lists some of the advantages and disadvantages of HIDS:

Table 1: Advantages and Disadvantages of HIDS

ADVANTAGES	DISADVANTAGES
Ability to associate a user to an event.	Information provided by the HIDS becomes unreliable as soon as an attack on that host has been successful.
May detect attacks that are not detectable by NIDS.	When an OS is brought down by an attack, the HIDS goes down with the system.
Can analyze encrypted data that has been decrypted on the host.	In order to monitor several hosts, an HIDS would need to be placed on each host.
Ability to provide information about a host during an attack on that host.	HIDS are not able to detect network scans.
	HIDS may be ineffective during a DoS attack.
	HIDS require resources of the host in order to operate.

2.2 Network Intrusion Detection System (NIDS)

The network-based intrusion detection system is placed on the network for capturing the traffic, active services and servers for detecting, blocking and reporting of any unauthorized access of the resources. The network traffic is built at various layers of the network and they pass the data from one point to another. The system uses network packet sniffers (e.g., tcpdump) for capturing the network traffic. The captured data is further analyzed for any of the anomalous or misuse trends. NIDS basically consist of two components: appliance and software. A NIDS appliance is a piece of hardware that includes operating system (OS), network interface card (NIC) and software. The second component contains all the IDS software and at times the OS. The software-only NIDS are less expensive than appliance-based NIDS as the latter requires hardware configuration. It comprises of not only one device but a number of physically distributed components (e.g., sensors, database servers, console, etc.).It can detect several kinds of events from remediable that are less dangerous to the malicious.

A NIDS typically pick packets travelling in the network to which it is attached and compare it with database of so called signatures known to be malicious activity. Different types of signatures are used in NIDS and they are:
Header Signatures: masks the header information of the packet to identify suspicious or inappropriate information.
Port Signatures: masks the destination port number of the packets whether the destination port is pointing to legitimate server or not.
String Signatures: detect strings in the payload of the packet which known to be present in malicious code.
An NIDS is generally placed between the firewall and the system or the organisation structure to be protected. The following diagram shows an generic Netwrok Intrusion Detection System.

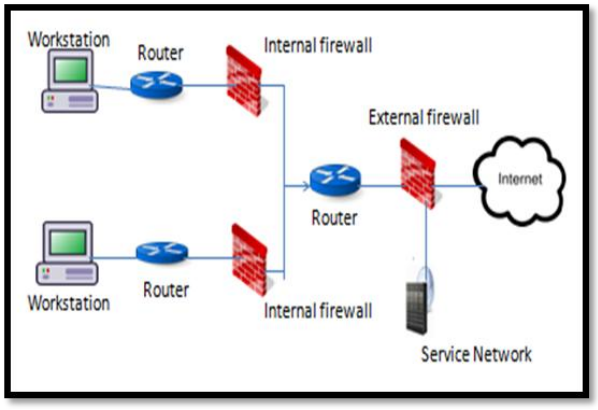

Fig 4: Generic Architecture of Network Intrusion Detection System

The various tools for NIDS includes SNORT, Sguil, etc. The following table 2 [9] lists some of the adavantages and disadvantages of NIDS:

Table 2: Adavantages and Disadvantages of NIDS

ADVANTAGES	DISADVANTAGES
Breadth of Coverage. An entire subnet may be covered by one NIDS.	False / Positive Alerts
Stealth	Cannot Analyze Encrypted Traffic
Minimal Install/Upgrade Impact to Network	NIDS only as strong as the latest signature update. New or variations in attack patterns will not register.
Avoid DoS that would otherwise affect a Host	Latency between time of attack and time of alert. By the time an alert is received the damage may have already occurred.
Ability to Identify Network Layer Errors	Difficulty in processing packets in a congested network.
Operating Environment Independent	Does not indicate whether the attack was successful.

3. HYBRID INTRUSION DETECTION SYSTEM

An Hybrid Intrusion Detection System combines the concepts of both Host Intrusion Detection System (HIDS) and Network Intrusion Detection System (NIDS) . It is developed by gathering data or information for analysis from both host and the network and then, an analysis methodology is applied. These systems will prove to be very efficient and effective in the future as it overcomes the drawbacks of both HIDS and NIDS.

An HIDS works efficiently for a host system by gathering the resource usage information of the system, this captured data is further analyzed based on behavior ie. Anomaly based detection which determines intrusions based on deviations from normal behavior using knowledge gathered over normal usage usage patterns of the resources of the system by the user, or based on Knowledge which uses may use database of signature of intrusions. An NIDS works effectively for any network by capturing and monitoring the activities of the network. It gathers various network parameters such as Source port number, Destination port number, Number of hosts connected, Hit count, Source IP address, Destination IP address,etc., and based on these header information obtained analysis mechanism is applied which can be Knowledge-based ie. Misuse detection which identify intrusions based on signatures, or Behavior-based ie. Anomaly detection which identify intrusions based on normal behavior pattern.

An Hybrid IDS takes advantage from both the approaches, and develops an IDS that works for both a host and a network. These systems are the future IDS. A lot of research is currently is carried out to overcome the drawbacks of HIDS and NIDS, and develop an Hybrid system that work for both host and network.

4. ARTIFICIAL IMMUNE SYSTEM

Artificial Immune System is highly inspired by our very own biological immune system, which defends the body from harmful infections. The human immune system provides protection at several layers, thereby protecting the body in every possible manner. The biological immune system consists of various terminologies that should be noticed like pathogen (a disease causing micro-organism), reservoir (place where pathogens is usually found), endemic (disease which is always present at the low level), epidemic (when the number of cases of a disease increases significantly) and vectors (organisms which carry pathogens between other organisms). The biological immune system is able to categorize all cells (or molecules) within the body as self-cells or non-self cells.

There are two major categories of immune system: innate immune system and adaptive immune system. Classification is given in the figure 2.6. Innate immune system is the one which detect and destroy the known pathogens in the body while adaptive immune system has special capability of destroying unknown pathogens in the body. Depending on the type of the pathogen, and the way it gets into the body, the immune system uses different response mechanisms (differential pathways) either to neutralize the pathogenic effect or to destroy the infected cells.

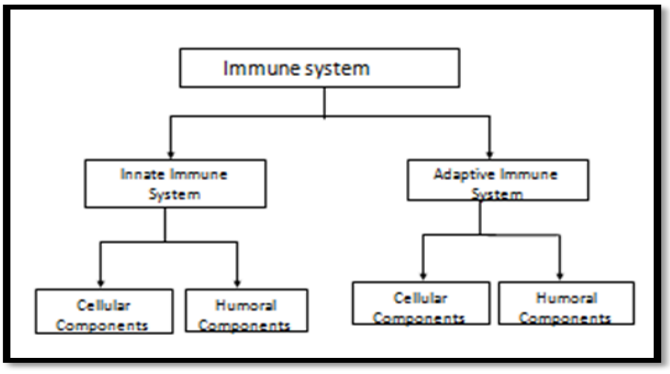

Fig 5: Classification of Immune System.

There are two most important-cells in this process which are white blood cells, called T-cells, and B-cells. They originate in the bone marrow, but T-cells pass on to the thymus to mature, before they circulate the body in the blood and lymphatic vessels. B-cells are the one responsible for the generation and excretion of antibodies. Each B-cell can produce exactly one kind of antibody.

This whole scenario has lead to development of Artificial Immune system. This method is based on Jerne's idiotypic network theory, which suggests that the immune system maintains a network of interconnected B-cells. There are basically four types of AIS techniques, which are listed as follows: Clonal Selection Algorithm, Negative Selection Algorithm, Immune Network Algorithm and Dendritic Cell Algorithm.

4.1 Negative Selection Algorithm

This algorithm differentiates between self cells and non-self cells that mean its focus is on the antibodies that generate negativity and instead of protecting the body it starts destroying self cells. If the body immune system discovers "self" cells then it would respond normally but if it comes across "non-self" cells it would result to abnormality. Negative selection algorithms are responsible to differentiate between such types of self and non-self cells. The following fig 6 shows the basic workflow of negative selection algorithm [7].

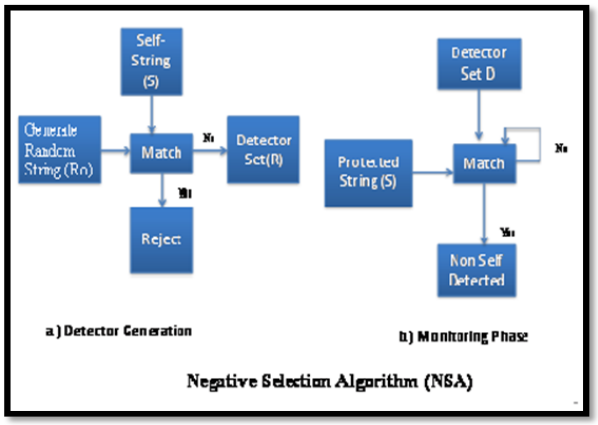

Fig 6: Basic workflow of Negative Selection Algorithm

4.2 Clonal Selection Algorithm

It shows how certain types of B and T lymphocytes are selected for the elimination of specific antigens invading the body. Such algorithms focus on the Darwinian attributes of the theory where selection is motivated by the closeness of antigen-antibody interactions, reproduction is inspired by cell division and variation is motivated by somatic hypermutation. The following fig 7 shows the basic workflow of clonal selection algorithm [7].

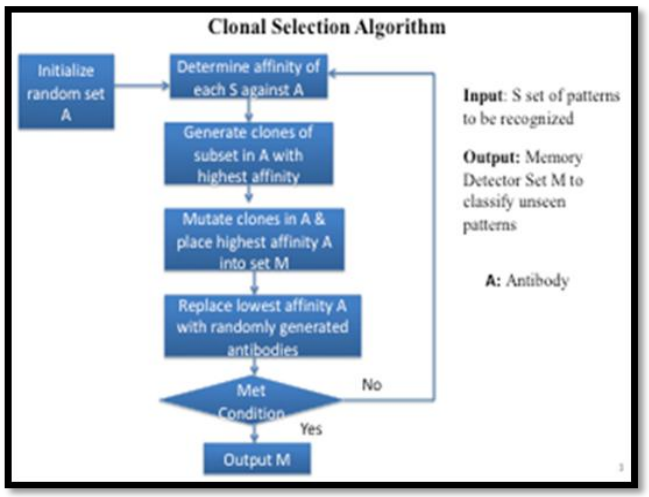

Fig 7: Basic workflow of Clonal Selection Algorithm

4.3 Immune Network Algorithm

The artificial immune network algorithm (aiNet) the input data are assumed to be unlabeled, thus resulting in a kind of competitive learning algorithm. Clustering of input data is performed with the help of Competitive learning so that MSE (mean square error) is iteratively reduced. The basic implementation issues that are essential for aiNet include: Data representation, network initialization, effect of stimulation on the development of the network, population control, termination condition along with mutation algorithm of the nodes. The following fig 8 shows the basic workflow of immune network algorithm [7].

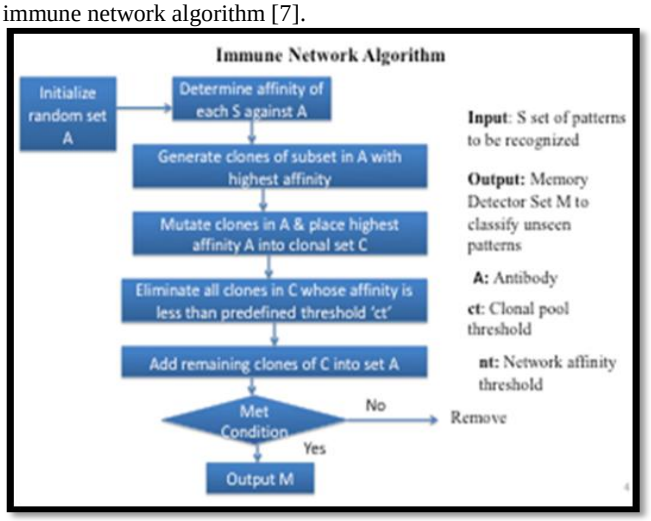

Fig 8: Basic workflow of Immune Network Algorithm

4.4 Dendritic Cell Algorithm

They are inspired by the dendritic cells of human immune system, which detects danger. In the same way dendritic cell algorithm responds to danger signals rather than. discriminating self from non-self. The main aim of the algorithm is to provide specific information by preparing dendritic cells how to classify normal and anomalous input patterns.

5. RELATED WORK

In [4] Li et. al uses the concept of the bionic theory of immune network and clonal selection which is inspired from natural immune system. The text clustering algorithms was

proposed based on the artificial immune system by Tang Na and Rao Vemuri, and it has better cluster quality than other algorithms through data compression, but it also has some disadvantages such as higher complexity time and the ability to find new clusters [4]. To overcome these disadvantages, text feature dimension reduction and clonal selection algorithm is used for the purpose of clustering. In this approach, first the calculation method for the affinity between antibody and antigens and the affinity of antibodies, the genetic operation factors were included; and lastly, clustering process and analysis were done based on the text set in a corpora.
The proposed concept can be used for clustering the network traffic into self and non-self categories based on the information of the various attributes of the network.

The artificial immune system (AIS) has been the active and vibrant field of research for a number of years, modelling natural immune system to tackle with wide variety of real world applications. AIS has various characteristics such as: distributed, robust, autonomous, flexible, uniqueness and many more. In [3] Zheng et. Al presents the survey of several current applications of AIS, and provides reflections on the prominent areas of the domain that can be used in future for dealing with real-world problems. It can use in different fields: machine learning, pattern recognition, computer virus detection, anomaly detection, optimization, robotics and etc [3]. In computer security, adaptability property of artificial immune system can be used to learn the intrusions and develop signatures of these intrusions. The most challenging issue with computer security to determine the difference between the normal and abnormal traffic. An efficient solution in this respect is in the form of biologically inspired computing, and particularly AIS.

The Dendritic Cell Algorithm (DCA) is the latest research of the danger theory has been applied into wide range of problems, particularly anomaly detection [5]. The analysis process of DCA is generally performed offline, when the entire information sequence has been generated. Thus, in order to provide real-time anomaly detection an algorithm is proposed by Rassam et. al. This algorithm, it is more easy to understand and easy to realise real-time anomaly detection. In the process of information generation whenever enough evidence of an antigen is obtained, the system will immediately analyze and output the detection results. The real-time analysis algorithm proposed in this paper, effectively improve the performance of the DCA in the field of anomaly detection.

Several Intrusion Detection System (IDS) have been developed till now. Signature based approach for anomaly detection requires the regular update with the latest signatures of unknown attacks and hence impractical. Anomaly based approach suffers from high false positive and low detection rates and requires labeled dataset for detection profile. In [6] Yuan et. al used a biologically inspired artificial immune network for clustering the DARPA KDD Cup 1999 dataset on which rough set method was applied to obtain most prominent features [6]. The detection rate is enhanced with this reduced set and artificial immune network proved to be robust in detecting novel attacks.

6. CONCLUSION

The various Intrusion Detection Systems developed till now works either for a host system or network and are capable of identifying only known intrusions. Thus, a Hybrid Intrusion Detection System needs to be developed that combines the concept of both Host based IDS and Network based IDS, and is capable of identifying known as well as unknown intrusions. An Hybrid IDS will captures both host and network data which will further analyzed for possible intrusions. Artificial Immune System (AIS) that is robust and flexible can be used for IDS, thereby making them capable enough to identify intrusions in every possible manner.

7. REFERENCES

[1] G. Teodoro, J. D. Verdejo, G. Macia-Fernandez, and E. Vazquez, 2009. Anomaly-based network intrusion detection.

[2] Richard A. Kemmerer, and G. Vigna, 2002. Intrusion Detection: A Brief History and Overview.

[3] J. Zheng, Y. Chen, and W. Zhang, 2010. A Survey of artificial immune applications.

[4] Ma Li, Yang Lin, Bai Lin, and W. Rongxi, 2012. Immune network based text clustering algorithm.

[5] M. A. Rassam, and Mohd. A. Maarof, 2012. Artificial immune network clustering approach for anomaly intrusion detection.

[6] Yuan, and Qi-juan Chen, 2012. A dendritic cell algorithm for real-time anomaly detection.

[7] Jon Timmis, 2007. Artificial immune systems-today and tomorrow.

[8] "Dawn of the new Security", Wordpress publication, 2009. Available: http://zulcap.wordpress.com/2009/10/27/lecture-9-intrusion-detection-system-ids/.

[9] Thomas Goeldenitz, 2002. IDS-Today and Tomorrow. SANS Institute InfoSec Reading room site.