# AN EFFICIENT INTRUSION DETECTION SYSTEM USING COMPUTATIONAL INTELLIGENCE

J. Visumathi, Dr. K. L. Shunmuganathan

## ABSTRACT

Intrusion detection system is one of the widely used tools for defense in Computer Networks. In literature, plenty of research is published on Intrusion Detection Systems. In this paper we present a survey of Intrusion Detection Systems. We survey the existing types, techniques and approaches of Intrusion Detection Systems in the literature. Finally we propose a new architecture for Intrusion Detection System and outline the present research challenges and issues in Intrusion Detection System.

**Keywords:** Intrusion Detection, Neural Network, Fuzzy logic, Artificial Intelligence, Honeypot, Data mining.

## 1. INTRODUCTION

The purpose of Intrusion Detection Systems (IDS) is to detect and prevent electronic threat to computer systems. The extensive use of the computers and availability of the Internet increase the impact of problem in size. In today's world everyone is connected over networks, and many services are provided over the Internet. This global reach increases the risk of intrusion threats from unknown sources. According to the Computer Emergency Response Team CERT, 32,956 vulnerabilities were reported from many sources throughout 1995 until the first quarter of 2007[17]. Intruder can use these vulnerabilities to launch an attack against Computer Network or Servers. Two things are certain—intrusion detection is still a long way from being mature, and intrusion prevention technology is in its infancy.

Reasons for using Intrusion Detection System (IDS) are:

1. To protect network from attack and abuse.
2. To detect the violations in security and attacks on network.
3. To document the existing threat to an organization.
4. To get detail information about intrusions that occurred.

The rest of this paper is organized by its basic approaches for IDS which is described in Section 2. Section 3 of this paper describes the types of IDS and techniques are described in section 4. Proposed architecture and algorithm is covered in section 5 followed by challenges and issues in development of IDS in section 6. Last section covers the future works and conclusion.

## 2. BASIC APPROACHES FOR INTRUSION DETECTION SYSTEM

Approaches for Intrusion detection systems can be broadly classified as:

1. Signature based
2. Classification based
3. Anomaly based

### Signature Based (Misuse Detection) Approach [3] [19]

Most of the commercial IDSs are "misuse detection systems" which are designed to detect only known attacks. This approach uses a database of known attack signatures which is developed by experts and intrusion analyst. The traffic over the network or sequence of processes within the Computer is compared to the entries in this database. If there is a match with database entries, the IDS system generates an alert message. Even though such a system does not generate false positives alerts, these systems cannot identify new and novel attacks.

There are two advantages of Misuse Detection Approach:

1. It is very effective for detecting the attacks without generating an overwhelming number of false alarms
2. It can quickly and reliably diagnose the use of a specific attack tool

On the other hand, the disadvantages of Misuse Detection Approach are:

1. It can only detect those attacks that have been described in the database.
2. The database must be constantly updated with signatures of new attacks.

### Classification-based Intrusion Detection Approach

This approach uses normal and abnormal data sets of user behavior, and uses data mining techniques to train the IDS system. This creates more accurate classification models for IDS as compared to signature-based approaches and thus they are more powerful in detecting known attacks and their variants.
Disadvantage of Classification-based Intrusion Detection Approach: It is still not capable of detecting unknown attacks.

### Anomaly Intrusion Detection Approach

The basic assumption of Anomaly detection approach is that attacks are different from normal activity and thus they can be detected by IDS systems that identify these differences. Thus this approach begins with definition of desired form or behavior of the system and then distinguishes between that desired behavior and undesired or anomalous behavior. The main problem is, defining the boundary

between acceptable and anomalous behavior. So, the anomaly detector approach must be able to distinguish between the anomaly and normal.

There are two types of anomaly detectors:
*1. **Static anomaly detectors:*** It is based on the assumptions that there is a portion of the system being monitored that should remain constant.

***2. Dynamic anomaly detectors:*** To characterize normal and acceptable behavior a base profile is created by a dynamic anomaly intrusion system. Building the sufficiently accurate base profile is the main difficulty with the dynamic anomaly detection system.
The advantage of Anomaly Intrusion Detection approach is: It is possible to detect unknown attacks.

Disadvantages of Anomaly Intrusion Detection approach are: Produces a large number of false alarms due to the unpredictable behaviors of users and networks.
Therefore, large and accurate training data set is the major requirement of anomaly detection approaches to define the normal behavior patterns.

## 3. TYPES OF INTRUSION DETECTION SYSTEM

1. **Network-Based IDS:** Network-based IDS [3][4] monitors network traffic using techniques like packet sniffing to collect network traffic data and tries to detect malicious activity such as denial of service attacks; port scans or even attempts to crack into computers.
2. **Host-Based IDS:** Host-based IDS [17] monitors and analyzes system calls, application logs, file-system modifications and other host activities to identify the intrusion such as unauthorized remote login attempt, attempt to access unprivileged data. It normally works with Network-based IDS.
3. **Protocol-Based IDS:** Typically protocol-based IDS [16] are installed on a web server, and they are used for monitoring and analysis of the protocol in use of the computing system. If there is a deviation from intended behavior of protocol then it can be detected as intrusion.
4. **Graph-Based IDS:** Graph-based IDS [15] concerned with detecting intrusions that involve connections between many hosts or nodes. A graph consists of nodes representing the domains and edges representing the network traffic between them.

## 4. TECHNIQUES FOR INTRUSION DETECTION SYSTEM

1. **Neural networks (NNs)** [11] can be trained to recognize arbitrary

patterns in input data, and associate such patterns with an outcome, which can be a binary indication of whether an intrusion has occurred. Such models are only as accurate as the data used to train them.

2. **State transition tables** [2] [12] describe a sequence of actions an intruder does in the form of a state transition diagram. When the behavior of the system matches those states, an intrusion is detected.

3. **Hidden Markov Models (HMMs)** [12] are a stochastic version of the state transition techniques discussed above, where the states and transition probabilities are modeled as a Markov process with unknown parameters. A learning phase estimates these unknown parameters from the input data.
4. **Artificial Immune Systems** [14] are adaptive systems, inspired by theoretical immunology and observed immune functions, principles and models, which are applied to problem solving. The innate system of the human immune system can be compared with the misuse detection of the IDS; both uses pattern recognition respectively on memory cells and signatures database to detect intrusions. The adaptive system can be compared with the anomaly detection where both can detect yet unseen attacks and where their sensors have to go through a training phase.
5. **Genetic Algorithms (GAs)** – [1] Genetic algorithms mimic the natural reproduction system in nature where only the fittest individuals in a generation will be reproduced in subsequent generations, after undergoing recombination and random change.
6. **Decision Tree** [10] is a model of decisions and also can be used to show possible consequences for particular occurrences where there are conditional probabilities for each occurrence. Those occurrences of attacks form a tree-based structure that contains root node and a number of leaf nodes. Decision tree generally performs very efficiently even if dealing with a large amount of data.
7. **Bayesian Network** [13] Bayesian Network is a graphical representation of the joint probability distribution function over a set of variables. The network structure can be represented in Bayesian Network as a Directed Acyclic Graph where each node represent a random variable and each edge between nodes shows the relation between nodes (i.e. relationship between variables). Individual invents which occurs during attack are represented as nodes in the graph and relationship between those events are represented as edges of the graph and this graph is then used to detect the intrusion.

8. **Fuzzy logic** [4] is a set of concepts and approaches designed to handle vagueness and imprecision. A set of rules can be created to describe a relationship between the input variables and the output variables, which may indicate whether an intrusion has occurred
9. **Honeypot** [5] is an unreal network system designed to trap crackers and intruders. The honeypot is used as bait in the form of a vulnerable system to trap hackers and keep them away from accessing the critical information in the main system. In this technique alarming adversaries, initially detected by the IDS, will be rerouted to a honeypot network for a more close investigation. If as a result of this investigation, it is found that the alarm decision made by the IDS of the agent is wrong, the connection will be guided to the original destination in order to continue the previous interaction. This action is hidden to the user. Such a scheme significantly decreases the alarm rate and provides a higher performance of IDS.
10. **Data Mining** [3] [4] is an analytic process designed to explore data in search of consistent patterns and/or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new subsets of data.

## 5. PROPOSED ARCHITECTURE

Each type of IDS offers fundamentally different information-gathering, logging, detection, and prevention capabilities. Each technology type offers benefits over the others, such as detecting some events that the others cannot and detecting some events with significantly greater accuracy than the earlier technologies. In many environments, a robust IDS solution cannot be achieved without using multiple types of IDS technologies. For most environments, a combination of network-based and host-based IDS technologies is needed for an effective IDS solution. Thus in our architecture we combined Host-based and Network-based IDS. Network-based IDS is used to detect Dos, DDoS and Probing attacks and Host based IDS are used to detect R2L and U2R attacks.

Using IDS based on Data mining [3] [4] is an effective method. IDS based on Date Mining have a behavioral model through widely checking data. So it can accurately capture the actual invasion and normal behavior. This automated technique no longer needs manual analysis and manually coding the invasion mode and no longer needs to choose statistical methods by experience when build the normal behavior using model. The major advantage of the data mining technique is that, it can be applied to multiple data stream.

Many Researches have used fuzzy association rules effectively to design their NIDSs. Incremental Fuzzy-rule Mining can be very useful to meet the real-time

requirements of IDS because it can produce the new rules set while detection process is going on [4].

Data warehouse is the most suitable data store for storing the data records gathered online from network. This will increase the speed of incremental fuzzy-rule mining algorithm and is the most suitable data store to analyze multiple data streams [23].

Using the honeypot technique, the system is able to avoid many wrong decisions made by IDS. This will reduce the false alarm rate of the attack detection [5]. Figure 1 shows the block schematic of the proposed Network Intrusion Detection System.

**Feature Data Warehouse:** It is used to store packet information extracted by Feature Extractor, which is used to detect Intrusion.

**Known Attack Signature Database:** It is used to store Known Attack Signatures.

**Possible Attack Signature Database:** It is used to store possible attack signatures which are predicted by using Known Attack.

**Data Mining:** [2][21] [22] It uses Attack signature database and feature Data Warehouse along with Apriori algorithm to predict possible attack signatures using existing attack signatures.

**HIDPS Attack Signature Database:** Attack signatures for Host based IDS are centrally stored at machine running NIDS.

**Packet Sniffer:** It uses raw socket programming to fetch packets from network.

**Feature extractor:** It extracts information present within the packet such as, Source IP Address, Destination IP Address, values of flags present in Packet Header, etc... .

**Known Attack Detector:** Known Attack Detector module is used to detect network connections that correspond to attacks for which signatures are available.

**Possible Attack Detector:** It uses Possible Attack Signature database to detect whether traffic matches with possible attack signature generated by Data Mining unit. If there is a match it forward that connection to honeypot to detect whether there is an intrusion or not.

**Honeypot:** It is used to detect whether the connection is trying to do intrusion in the network or not.

Algorithms:

**Possible_Attack_Signature_Algorithm: Input:** Attack Signature Database (ASDb)

**Output:** Possible Attack Signature Database (PASDb)

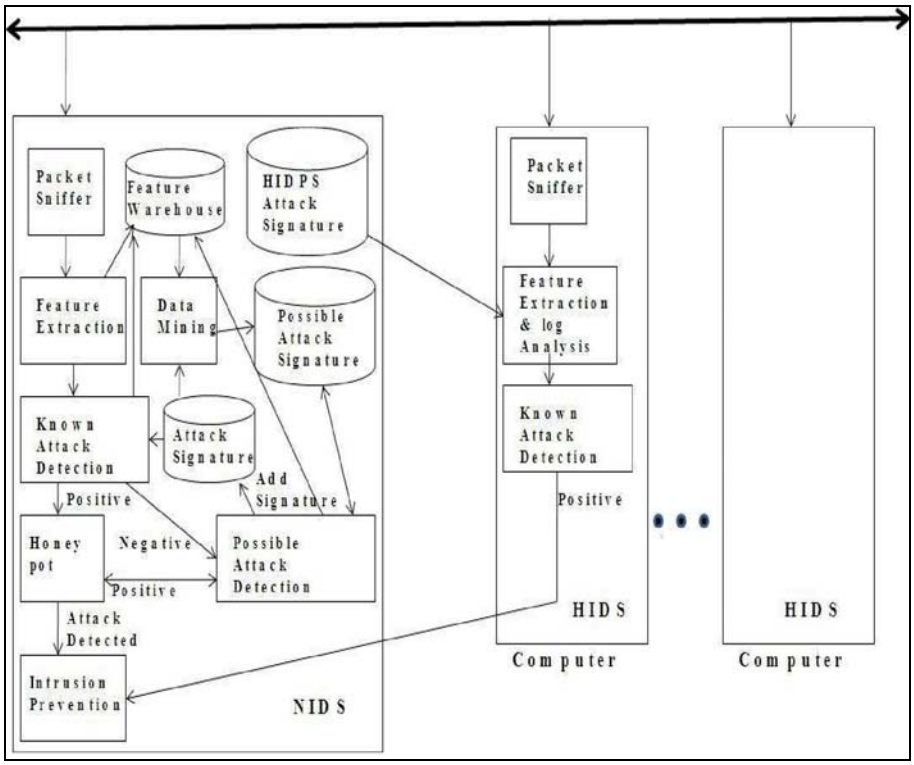


**Figure 1: System Diagram**

**Steps:**

1. [2][21][22]Apply Apriori algorithm on Feature Data Warehouse to generate Patterns Set

2. For each Pattern in Patterns set
    a. For each Signature in Known Attack Signature set
        i. Calculate Similarity between pattern and signature
        ii. If(Similarity > 0.9)
            1. Add pattern to Possible Attack Signature
3. Stop

**Known_Attack_Detection_Algorithm:**

**Input:** Network Traffic Feature, Attack SignatureDatabase

**Output:** Traffic Classification (Norma/Attack)

**Steps:**

1. For each Signature in Known Signature Set
    a. If(Traffic Feature matches with Signature)
        i. Forward corresponding Connection to Intrusion

Prevention module

ii. Mark corresponding entry in Feature Data Warehouse for attack

b. Else

i.Forward Network Traffic Feature to Possible Attack Signature detector

**Possible_Attack_Detection_Algorithm:**

**Input:** Network Traffic Feature, Possible Attack Signature Database

**Output:** Traffic Classification (Norma/Attack)

**Steps:**

1. For each Signature in Possible Signature Set
   a. If(Traffic Feature matches with Signature)
      i. Forward corresponding Connection to Honey pot module to detect Intrusion
2. If (Result from Honeypot is Positive)
   a. Remove Corresponding Signature entry from Possible Attack Signature Database
   b. Add removed Signature to Known Attack Signature Database

   Else
   c. Remove Corresponding Signature entry from Possible Attack Signature Database
3. Mark corresponding Network Traffic Feature entry in Feature Data Warehouse for attack

## 6. CHALLENGES AND ISSUES

1. With best of our knowledge many researchers have proposed new architecture for Intrusion Detection System but did not comment on how their architecture will accept in real time environment.
2. Further many of them did not marked that how much load their architecture will create on executing platform. (Future scope of our paper will include that part)

## 7. CONCLUSION AND FUTURE SCOPE

This paper reviews and tried to summarize different types, methods and approaches for Intrusion Detection System. Further this paper has proposed a new architecture for Intrusion Detection System which generates and test new signatures for Intrusion Detection without the interference of third party.

The proposed model is in its initial stage where an initial algorithm is proposed.

The future step for this proposal is under development where the real time analysis is going on.

## ACKNOWLEDGEMENT

We take immense pleasure in thanking our Chairman Dr. Jeppiaar M.A, B.L, Ph.D, the Directors of Jeppiaar Engineering College Mr. Marie Wilson, B.Tech, MBA.,(Ph.D) Mrs. Regeena Wilson, B.Tech, MBA., (Ph.D) and the Principal Dr. Sushil Lal Das M.Sc(Engg.), Ph.D for their continual support and guidance. We would like to extend our thanks to my guide, our friends and family members without whose inspiration and support our efforts would not have come to true. Above all, we would like to thank God for making all our efforts success.

## REFERENCES:

1. Suhail Owais, Václav Snášel, Pavel Krömer, Ajith Abraham "Survey: Using Genetic Algorithm Approach in Intrusion Detection Systems Techniques" CISIM 2008, IEEE, ISBN: 978-0-7695-3184-7
2. Rakesh Agrawal, Ramakrishnan S&ant "Fast Algorithms for Mining Association Rules", Proceddings of the 20th VLDB Conference Santiago, Chile, 1994
3. Hu Zhengbing1, Li Zhitang1,Wu Junqi, " A Novel Network Intrusion Detection System(NIDS) Based on Signatures Search of Data Mining" 2008 IEEE, Workshop on Knowledge Discovery and Data Mining
4. Ming-Yang Su, Kai-Chi Chang, Hua-Fu Wei, and Chun-Yuen Lin, "A Real-time Network Intrusion Detection System Based on Incremental Mining Approach", 2008 IEEE
5. Babak Khosravifar, Jamal Bentahar, "An Experience Improving Intrusion Detection Systems False Alarm Ratio by Using Honeypot", 2008 IEEE, 22nd International Conference on Advanced Information Networking and Applications
6. Marimuthu, Dr. A. Shanmugan, "Intelligent Progression for Anomoly Intrusion Detection", 2008 IEEE, ISBN: 978-1-4244-2106-0
7. ZHAN Jiuhua, "Intrusion Detection System Based on Data Mining", 2008 IEEE, Workshop on Knowledge Discovery and Data Mining
8. Youssif Al-Nashif, Aarthi Arun Kumar, Salim Hariri, Guangzhi Qu, Yi Luo, Ferenc Szidarovsky, "Multi-Level Intrusion Detection System", 2008 IEEE, International Conference on Automonic Computing
9. Bane Raman Raghunath, Shivsharan Nitin Mahadeo, "Network Intrusion Detection System", 2008 IEEE, First International Conference on Emerging Trends in Engineering and Technology
10. Joong-Hee Leet, Jong-Hyouk Leet, Seon-Gyoung Sohn, Jong-Ho Ryu, and Tai-Myoung Chungt, "Effective Value of Decision Tree with KDD 99 Intrusion Detection Datasets for Intrusion Detection System", 2008 IEEE, ISBN: 978-89-5519-136-3
11. Lgor Vinicius Mussoi de Lima, Joelson Alencar Degaspari, Jo˜ao Bosco Mangueira Sobral, "Intrusio n Detection Through Artificial Neural Networks", 2008 IEEE, ISBN: 978-1-4244-2066-7

12. Do-hyeon Lee, Doo-young Kim, Jae-il Jung, "Multi-Stage Intrusion Detection System Using Hidden Markov Model Algorithm", 2008 IEEE, International Conference on Information Science and Security .
13. Lu Huijuan, Chen Jianguo,d Wei Wei, "Two Stratum Bayesian Network Based Anomaly Detection Model for Intrusion Detection System", 2008 IEEE, Internation al Symposium on Electronic Commerce and Security
14. Divyata Dal, Siby Abraham, Ajith Abraham, Sugata Sanyal, Mukund Sanglikar, "Evolution Induced Secondary Immunity: An Artificial Immune System based Intrusion Detection System", 2008 IEEE, 7th Computer Information Systems and Industrial Management Applications
15. Amin Hassanzadeh, Babak Sadeghian, "Intrusion Detection with Data Correlation Relation Graph", 20 08 IEEE, The Third International Conference on Availability, Reliability and Security
16. S.Sangeetha, V. Vaidehi, N.Srinivasan, K.V. Rajkumar, S. Pradeep, N.Ragavan, C.Sri Sai Lokesh, I.Subadeepak, V.Prashanth, "Implementation Of Application Layer Intrusion Detection System Using Protocol Analysis", IEEE 2008, IEEE-International Conference on Signal processing, Communications and Networking
17. Mrs. P. Kola Sujatha Dr. A. Kannan S. Ragunath K. Sindhu Bargavi S. Githanjali,"A Behavior Based Approach to Host-Level Intrusion Detection using Self-organizing Maps", 2008 IEEE, First International Conference on Emerging Trends in Engineering and Technology
18. Robert, Richardson, "2007 Computer Crime and Security Survey", http://i.cmpnet.com/v2.gocsi.com/pdf/CSISurvey2007.p df
19. Ya-Li Ding, Lei Li, Hong-Qi Luo, "A novel signature searching for Intrusion Detection System using data mining", 2009 IEEE Eighth International Conference on Machine Learning and Cybernetics ISBN: 978-1-4244-3703-0
20. Juan Wang, Qiren Yang, Dasen Ren "An intrusion detection algorithm based on decision tree technology", 2009 IEEE Asia-Pacific Conference on Information Processing ISBN: 978-0-7695-3699-6
21. Rakesh Agrawal, Arun Swami, Tomasz Imielinski, "Mining Association Rules between Sets of Items in Large Databases" , Proceedings of the 1993 ACM SIGMOD Conference Washington DC, USA, May 1993
22. Heikki Manila, Hannu Toivonen, A. Inkeri Verkamo "Efficient Algorithm s for Discovering Association Rules", Knowledge Discovery in Databases (KDD'94) U.M. FAYYAD and R. Uthurusamy (EDS), AAAI Press, 1994, p. 181 – 192
23. "Creation and Deployment of Data Mining-Based Intrusion Detection Systemsin Oracle Database 10g " http://www.oracle.com/technology/products/bi/odm/pdf/odm_based_intrusion_detection_paper_1205. pdf