Evaluation of Agent Based Host Intrusion Detection System (AHIDS) through Various Classification Techniques

Aumreesh Kumar Saxena[1], M. Arshad[2], Sitesh Sinha[3]
[1,2]Dept of CSE, SIRTS, Bhopal (M.P.) India.
[3]Dr. C.V. Raman Univcrsity, Vaishali (Bihar) India.

ABSTRACT
Intrusion-detection-system (IDS) is the necessary part of the system form security point of view. It comes in both hardware and software IDS. Primary works of IDS is the recognized and differentiate between usual and unusual things that are happing in the system and show unusual thing as intrusion. At present unusual activity in the system are growing every year so improvement in existing IDS is always required. This paper is presents two thing, one is the concept of new host based IDS which used agent based mechanism to find intrusion on a host so it become host based intrusion detection system (AHIDS) and second is the selection of good classification technique for IDS. Selection of good classification technique is necessary because it is provide more accurate and prediction analysis on large amount of record set and IDS uses such type of data set for finding intrusions. Proposed AHIDS using three agent like PE, RA and DB agent where PE is the packet capturing and extracting agent, RA is rule agent and DB is the database agent. All agents are work together that means PE agent pass data to DB agent and RA agent collect data from DB agent. For the selection of good classification technique, proposed AHIDS comparing five different classification techniques like Naïve-Bayes, K-nearest- neighbors, SVM, J48, and Random-Forest. NSL-KDD data set is used for results examination.

Keyword— Intrusion Detection System (IDS), Intrusion, Classification, Security, Agent.

I INTRODUCTION

The Internet is a global public and large scale open network. It has its own pros and cons. With expansion of the Internet and its capability, life is becoming easy for everyone [1]. It is helpful and beneficial for both business organization and individual user. In today's world, most of the work can easily done by Internet like e-business, e-ticket, e-shopping and many more. More and more public are associating with Internet to take benefit of these facilities. Now days, internetwork connectivity is becoming very important phase [2]. The widespread growth of Internet and increasing easily accessibility of tools for attacking computer network is one of the major concerns of network administrators and security providers for the security of computer systems and data in it.

Unusual events in system are called intrusion and they are tried to escape protection shield of the system. As we know that availability, integrity and confidentiality, are the basic security principal and it is suffer by the every attempt of intrusion [3].

Therefore, all unusual activity which is happening in the system is done by outside that means it is accessing system through network or insider that may by its authenticated person but both are trying to get control of the system so they can stop the security feature of the system or they can access the information of the system. [4-5].

Organizations of the paper are as follow: Introduction is in Section 1. Section 2 is present about intrusion detection system and data-mining classification technique. Literature surveys are discussing in Section 3 and section 4 is showing research finding. Section 5 gives proposed agent based host intrusion detection system. Section 6 discusses about the results and the section 7 is of conclusion.

II INTRUSION DETECTION SYSTEM AND DATA MINING CLASSIFICATION TECHNIQUE

IDS come in hardware or software form that examines a network or network node or computer for unfriendly movement [5]. Intrusion detection system works on the safety rules for identifying any abnormal activity. System administrator is the person who defines these policies according to the requirement of organization. Any action which breach these rules are consider as unusual activity. Every observed unusual activity is informed to the administrator through mails or any means. For efficient working of Intrusion detection system these security rules should be restructured regularly. Intrusion Detection System [5] is a system which examines a computer or group of computers on the network against any intrusion. Intrusion is any access or action in system which is not authorized. IDS are primarily used to identify any misuse of system or network. IDS identify the occurrence of an intrusion in the network and generate an alert [6]. Intrusion detection system performs three main roles: examining the systems, detection of any intrusion and raise an alert if any malicious activity is detected. Conceptualized IDS can be classified in two ways to defend the system from spiteful activities [7]. First move toward of build IDS for totally safe system through various cryptographic technique along with authorization techniques [8]. Figure 1 is showing the basic intrusion detection system mechanism. In this host are connected with various IDS and these IDS are handled by IDS admin if any intrusion are

identified by IDS then it reported to IDS admin for further action.

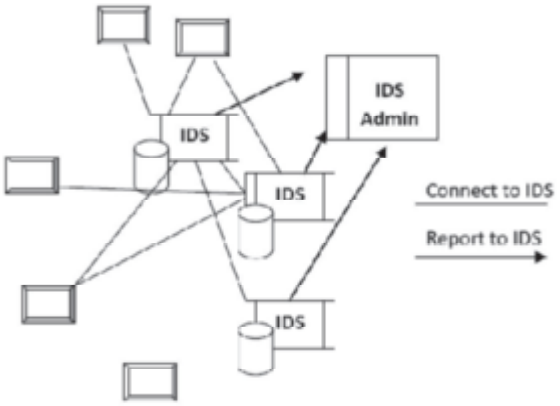

Fig. 1: Intrusion Detection System Mechanism

This approach is not providing fully safe system because each user has different type of vulnerability. So another way is required to protect system completely and IDS can be mechanisms which can be provide such type of environment where system can be fully safe and secure. There are four type of IDS available, named are information-based (knowledge-based), behavior-based, host-based, and net-based [11].

(a) Knowledge-Base IDS - Knowledge-Based or signature based IDS use predefine signature of the intrusion that is also known as patterns of the computer or network and its identified intrusion based on pre-knowledge. Its maintain record of previous signature of intrusion and vulnerabilities [12].

(b) Behavior Base IDS - Behavior based IDS work on the normal behavior of the System. It is continuously watch to the system and recorded normal activity of the system, on the basis of these normal activities, it is identify to the intrusion [13].

(c) Host Base IDS - The host-based-intrusion-detection system never monitors traffic of the network, moderately it checks out what is happening over the original targeted machines. This is done through monitoring the security event-logs or by monitoring the modification in the system, as an example modification to the complicated system-files or to systems-registry [14].

(d) Network Base IDS - Network-based IDS analyze network traffic to monitor entire computer networks. Controls the packets on network-wire and tried to introduce an intruder through matching attack-pattern for database of well known patterns of attack [14].

(e) Data-Mining for IDS: There are many data mining techniques for intrusion detection such as classification, clustering, frequent pattern mining and many more. Clustering is the technique of tagging data and give into cluster of comparable items exclusive of using recognized arrangement of data points. Members of similar group are like and instances of dissimilar groups are dissimilar from each other. Clustering technique can be categorized into groups of four: one is hierarchical algorithm second is partitioning algorithm, third is Grid based algorithm and fourth is density based algorithm [15]. Classification: Classification is the assignment of captivating each and every instances of record set in thought and transmission it to a exacting abnormal and normal class that means recognized arrangement is used for new instances. It can be efficient for together anomaly and misuse detection, but more regularly used for misuse detection [15]. Classification categorized the datasets into predetermined sets. Various classification method like Support vector machine, naive bayes classifier decision tree, K-nearest neighbors classifier, etc. are used in IDS [15].

III RELATED WORK

An intrusion detection system proposed by [16] which is the combination three classification techniques named are Naïve Bayes, decision trees and support vector machines (SVM). With the help of presented IDS they are trying to increase the accuracy of intrusion detection and trying to reduce false positive rate. Another IDS presented by [17] which is using is the combination of attribute selection, outlier detection, and enhanced multiclass SVM classification methods. They are trying to reduce its processing time and improving the intrusion detection accuracy. In [18] a mobile agent immune system (MAIS) intrusion detection system proposed. In this they have three agent named mobile and static agents with detector agents which is the major factor in MAIS-Intrusion detection system. Presented IDS of [18] has cloning, transmutation, immigration, association, as well as arbitrariness mechanism to improve intrusion detection accuracy. Another IDS presented by [19] which is using is the combination of Support Vector Machine (SVM) and Ant Colony Network (CAN).classification methods.

They are also trying to reduce its processing time and improving the intrusion detection accuracy. An network intrusion detection system proposed by [20] which is the combination three classification techniques named are support vector machine (SVM), K-Nearest-neighbors (KNN) and Decision Tree. In this they are trying to increase the accuracy of intrusion detection and trying to reduce false positive rate. In [21] a hybrid classifier model of IDS is proposed which is using genetic algorithm (GA) for feature selection to improving the accuracy of classification and support vector machine (SVM) which is a machine learning technique for classification of intrusions. Tree based data mining classification techniques such as Hoeffding tree, j48, Random Forest, Random Tree, Rep-Tree used in [22] on intrusion detection . They used KDD-99 data set with WEKA 3.9 tool to implement this model [22]. In [23] providing study and comparative analysis of various classification technique like Random-forest, Tree J.48, rules part, Neural network, Random-tree, logistic, SVM, Bayes-Net, decision table, Hoeffding tree, and Naïve-Bayes which is usable in IDS.

IV RESEARCH FINDING

Form the study of the various existing IDS [16-23] and the uses of classification technique for the evolution of IDS on parameters like false positive rate, precision, detection rate etc; it is very difficult to find the good classification technique to evaluate IDS because results are varying from IDS to IDS. As we know that classification techniques is the part of data mining and it can be apply on big data record set to assist more accurate and prediction analysis with this reason we focused on identification of good classification technique for IDS because IDS work on large data record set and accuracy and predication analysis is an important issue to evaluate any IDS.

V PROPOSED WORK

Proposed IDS "Agent Based Host Intrusion Detection System (AHIDS)" is an intrusion detection system (IDS), the proposed AHIDS is having three agents on host and the working of each agent is different from other. The purpose of proposed AHIDS structure is to boost, importunate disruption removal increasingly. The proposed AHIDS can be exploit to boost systematize way as it make menacing of communication as disruption to be predictable. Proposed AHIDS is the agent based HIDS which gratify position based efficacy. Figure-2 expresses the essential model of proposed AHIDS. Proposed AHIDS is a completely different from other HIDS or NIDS. It is software application which will install on particular hosts to be monitored. It will examine different types of modification over time on host which may signal safety problems. AHIDS examines the actions and activities of a host which is to be monitored and match up to with its abnormal behavior. By the proposed AHIDS it is examining and monitoring individual computer systems, software applications are installed on them. It only analyses the essential packets which is capturing by the PE agent. After monitoring packets by RA agent who are installed on workstations with classification technique, write down the data to database by DB agent for further action.

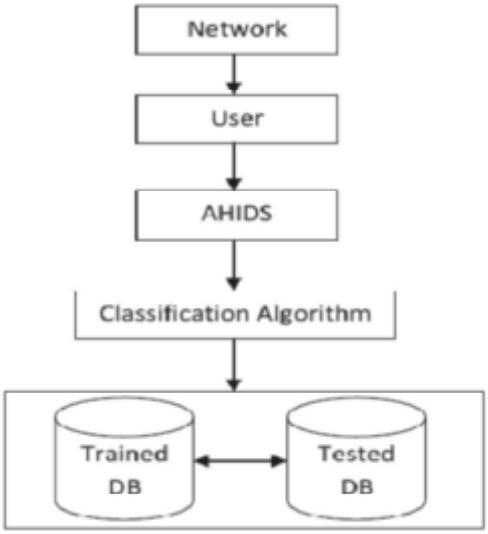

Fig. 2: Proposed Model of AHIDS

Proposed AHIDS can be execute real time, also the proposed AHIDS will grab real information at real time and it will move toward IDS. Figure-3 is showing the architecture of Proposed AHIDS which has three agents named is PE agent, RA Agent and DB agent. Workings of each agent are defined as follow:

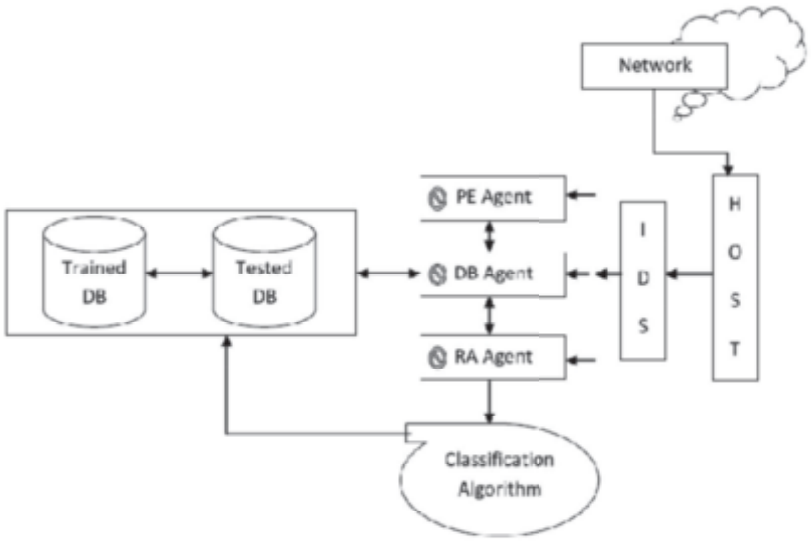

Fig. 3: Proposed Architecture of AHIDS

(a) **Agent-1: PE Agent -** PE Agent is the packet capture and extract agent. This agent captured packet form network and extract packet to fetch information from the packet which is passed to DB agent to store in tested data base for further uses. This is the first agent of the proposed IDS.

(b) **Agent-2: RA Agent -** RA Agent is the rule agent and it is very important agent because all rules which are design for intrusion detection are handled by this agent. Basically this agent is designing and implementing set of rules to find intrusion. Basically it used classification technique to identify intrusion. During intrusion detection this agent will retrieve data with the help DB agent from tested database and compare with each record of trained data base to find predefine pastern, if it is find then it marked that record as intrusion otherwise it is marked as normal.

(c) **Agent-3: DB Agent** - DB Agent is used to manage trained and tested data base. This agent will collect data from PE agent and store in the tested data base for further uses. DB agents provide data to RA agent on demand.

VI RESULTS

The proposed agent host based intrusion detection system has been implemented in Java along with classification techniques and tested on intel I5-2540M CPU with 2.60 Ghz and 4 GB ram on window 7 platform. The examination is arranged based on diverse constraint like True positive rate, false positive rate and precision and. True Positive (TP) that means rightly identifies normal which was predicted as normal [24]. False Positive (FP) that means predicated as intrusion but it is not an intrusion. Precision shows retrieved documents that are relevant to the query. There are various data record sets are available, but all have a few confines. The well-known data record set is KDD-99, but it has redundant information so it is not useful for the results evulsions [25]. In place of KDD-99 we used NSL-DD which is refine version of this data record set [25]. The NSL-KDD has two class of network packet one is normal and second is abnormal. In abnormal class various type of attack are shown like DOS, U2R, Probe and R2L [25]. Here we have used only 10% data record set of original NSL-KDD for results examination which is shown in table-1.

Table-2
Data Record Sets

Record Type	Instances
Normal/Intrusion	2000/3000

For the proposed IDS evaluation we have compared five classification techniques like Naïve-Bayes, K-nearest- neighbors, SVM, J48, and Random-Forest. Three parameters like True Positive Rate, Precision and False Positive Rate are participating in results by the 5 different classification techniques on predefine dataset which is shown by Table-1 and produced values are shown in Table-2.

Table- 2
Relative Examination of five classifications technique by Proposed IDS on True Positive Rate, Precision and False Positive Rate

S. No.	Algorithms	TPR	Precision	FPR
1	NaiveBayes	0.986	0.986	0.006
2	K nearest Neighbours	0.988	0.988	0.004
3	SVM	0.985	0.985	0.004
4	J48	0.942	0.942	0.069
5	Random Forest	0.976	0.977	0.048

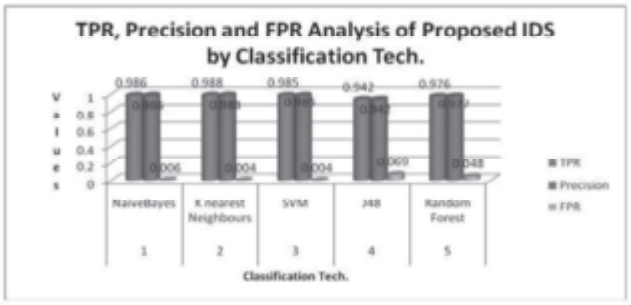

Fig. 4: Relative Examination of five classifications technique by Proposed IDS on True Positive Rate, Precision and False Positive Rate

Greater values of true positive rate and precision while lesser values of false positive rate are show the efficacy of the technique. From table 2 we get 0.988 true positive rates and precision for K-Nearest-neighbors while false positive rate of this is 0.004 similarly which is show that this classification technique is good of IDS. Similarly 0.986 true positive rate and precision for Naives Bayes while false positive rate of this is 0.006 which is the second good technique for IDS and furthermore 0.976 true positive rate and precision for Random Forest while false positive rate of this is 0.048 and 0.985 true positive rate and precision for Support Vector Machine SVM while false positive rate of this is 0.004 are also good for IDS but 0.942 true positive rate and precision for J48 while false positive rate of this is 0.069 are not producing effective results for IDS. Figure 4 is showing graphical analysis of presented results in table 3.

VII CONCULSION

Proposed Agent Based Host Intrusion Detection System (AHIDS) is an intrusion detection system (IDS) which is having three agents (PE, RA and DB) on host and the working of each agent is different from other. Basically proposed IDS are different from other HIDS because it's mechanism. The primary objective of this research is to find good classification technique for proposed IDS. To achieve this objective, we compared five classification techniques on three parameters like true positive rate, precision and false positive rate. Here results performance is completed on predefine NSL KDD data record set. As per results examination K-Nearest neighbors produced high true positive rate and precision where J48 classification technique produced low true positive rate and precision. False positive rate is good for K-Nearest-neighbors and support vector machine (SVM). It is found that K-Nearest-neighbors classification technique provides a high true positive rate and precision of above 98% with low false positive around 2%. So we can conclude K-Nearest-neighbors classification technique is appropriate and good for proposed agent based intrusion detection system when assessed with NSL-KDD dataset.

REFERENCES

[1] Firkhan Ali Bin Hamid Ali and Yee Yong Len "Development of Host Based Intrusion Detection System for Log Files" IEEE symposium on business, engineering and industrial application(ISBEIA) langkawi, malaysia Pp: 281-285, Dec. 2011

[2] Djemaa, B. ; Okba, K. "Intrusion detection system: Hybrid approach based mobile agent " IEEE International Conference on Education and e-Learning Innovations (ICEELI), Pp 1 – 6, 2012

[3] Ashutosh Gupta, Bhoopesh Singh Bhati, Vishal Jain, "Artificial Intrusion Detection Techniques: A Survey", IJCNIS, vol.6, no.9, pp.51-57, 2014. DOI: 10.5815/ijcnis.2014.09.07

[4] Jitendra S Rathore, Praneet Saurabh, Bhupendra Verma "AgentOuro A Novelty Based Intrusion Detection and Prevention System" Computational Intelligence and Communication Networks (CICN), Fourth International Conference, India, Pp: 695 – 699, Nov. 2012

[5] Audrey A. Gendreau; Michael Moorman " Survey of Intrusion Detection Systems towards an End to End Secure Internet of Things" 2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud), Austria, Pp 84 - 90, Sep-2016

[6] Safuan, H. ;; Cheah, Z.B. ; Lim, H.W. ; Chin, J.H. "Intrusion detection system based on mobile agent" Computers, Communications, & Signal Processing with Special Track on Biomedical Engineering. 1st International Conference on 14-16 Nov. Pp:266 – 270, 2005

[7] Bilal Maqbool Beigh,"A New Classification Scheme for Intrusion Detection Systems", IJCNIS, vol.6, no.8, pp.56-70, 2014. DOI: 10.5815/ijcnis.2014.08.08

[8] G Ramachandran and D Hart "A P2P Intrusion Detection System based on Mobile Agents" ACME '04, April 2-3, Huntsville, Alabame, USA Pp:186-190, 2004

[9] S Fenet, S Hassas "A distributed Intrusion Detection and Response System based on mobile autonomous agents using social insects communication paradigm" Electronic Notes in Theoretical Computer Science, Volume 63, Pp: 41-58. 2002

[10] N. B. Anuar, H. Sallehudin, A. Gani, O. Zakari, "Identifying false alarm for network intrusion detection system using hybrid data mining and decision tree", Malaysian journal of Computer Science, Vol. 21(2), 2008.

[11] C. Xiang, P.C. Yong, L.S. Meng, "Design of multiple-level hybrid classifier for intrusion detection system using bayesian clustering and decision trees", Pattern Recognition Letters 29, 2008.

[12] Janhavi Kaskar, Ruchit Bhatt, Rohit Shirsath "A System for Detection of Distributed Denial of Service (DDoS) Attacks using KDD Cup Data Set " (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3) , Pp: 3551-3555, 2014.

[13] Christopher Krügel Thomas Toth "Flexible, Mobile Agent Based Intrusion Detection for Dynamic Networks" Year: 2002 Pages:1-7

[14] Christopher Kr¨ ugel and Thomas Toth "Applying Mobile Agent Technology to Intrusion Detection" ICSE Workshop on Software Engineering and Mobility Year: 2001 Pages 1-5

[15] Rashmi Ravindra Chaudhari 1, Sonal Pramod Patil "Intrusion Detection System: Classification, Techniques And Datasets To Implement" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 04 Issue: 02 | Feb -2017

[16] S Khanum , M Usman and A Alwabel "Mobile Agent Based Hierarchical Intrusion Detection System in Wireless Sensor Networks" IJCSI International Journal of Computer Science Vol. 9, Issue 1, No 3, Pp 101-109, 2012

[17] S. Ganapathy,* P. Yogesh, and A. Kannan "Intelligent Agent-Based Intrusion Detection System Using Enhanced Multiclass SVM" Hindawi Publishing Corporation Computational Intelligence and Neuroscience Volume 12, Pp 1-10, 2012

[18] J Sen "A robust and fault-tolerant distributed intrusion detection system" 1st International Conference on Parallel Distributed and Grid Computing (PDGC), Pp 123-128, 2010

[19] Nilamadhab Mishra , Sarojananda Mishra "Support Vector Machine Used in Network Intrusion Detection" IOSR Journal of Engineering (IOSRJEN) ISSN (e): 2250-3021, ISSN (p): 2278-8719 PP 25-27 2018

[20] Nilesh B. Nanda and Dr. Ajay Parikh "Network Intrusion Detection System: Classification, Techniques and Datasets to Implement" International Journal on Future Revolution in Computer Science & Communication Engineering ISSN: 2454-4248 Volume: 4 Issue: 3 106 – 109