International Journal of Modern Engineering Research (IJMER)
www.ijmer.com Vol.2, Issue.1, pp-055-059 ISSN: 2249-6645

An Agent based Intrusion Detection System for Wireless Sensor Networks Using Multilevel Classification

K.Kulothungan[1], S.Ganapathy[2], P.Yogesh[3] and A.Kannan[4]
[1,2,3,4] Department of Information Science & Technology,
College of Engineering, Anna University, Chennai-25, Tamil nadu, India.

ABSTRACT
With the rapid growth of internet communication and availability of techniques to intrude the network, network security has become indispensable. In this paper, we propose a multilevel classification technique for intrusion detection that uses intelligent agents and a combination of decision tree classifier and Enhanced Multiclass Support Vector Machine algorithm for the implementation of an effective intrusion detection system in order to provide security to Wireless Sensor Networks. The main advantage of this approach is that the system can be trained with unlabeled data and is capable of detecting previously "unseen" attacks using agents. Verification tests have been carried out by using the KDD cup'99 data set. From the experiments conducted in this work, it has been observed that significant improvement has been achieved in intrusion detection rate and also in the reduction of false alarm rate.

Keywords*: Intrusion Detection, EMSVM, Multilevel Decision Tree, Intelligent Agents*

I. INTRODUCTION
Recently, security has become a vital concern in many application areas since computers have been networked together with a very large number of users and systems. The adoption of Wireless Sensor Networks (WSNs) has increased in recent years mainly due to their advantages in many applications. WSNs can be defined as a heterogeneous system that consists of nodes which are having tiny sensors and actuators. Sensors networks may consist of hundreds or thousands of low-power, low-cost nodes, fixed nodes deployed largely together to monitor and affect the environment. Intrusion detection and prevention techniques are necessary to provide security to WSNs because they are prone to various types of attacks from both insiders and outsiders. Denial of Service (DoS) attacks is an important attack that leads to more power consumption as well as the collapse of the entire network due to unnecessary flooding of packets.

An intrusion detection system can be used as a first line of defense in such a scenario in order to reduce possible intrusions and thereby reducing the risk of attacks. The current security mechanisms such as firewalls focus only external attacks. On the other hand, intrusion detection systems are capable of detecting both internal and external attacks. IDSs are classified, based on their functionality as misuse and anomaly intrusion detection system. A misuse intrusion detection system uses a set of well defined patterns for attack and they can be detected by matching these patterns against normal user behavior in order to detect intrusions effectively. Usually, misuse detection is harder than the anomaly intrusion detection since it is carried out by legitimate internal users who know the systems password and other credentials. In an anomaly intrusion detection system, the user behavior becomes different from the normal usage behavior. Therefore, it is necessary to provide an intelligent intrusion detection system which can find out both internal and external attacks.

However, the existing intrusion detection techniques, which are proposed by various researchers for misuse and anomaly detection [14,2], are generally not sufficient to provide the required security to WSNs because they have limited power and tiny structure. The attacks are carefully designed by the attackers and hence the application of the existing intrusion detection techniques causes a high false positive rate. Moreover, the existing Intrusion Detection techniques are capable of detecting only known intrusions since they classify instances by the rules they have acquired based on training from past data. However, it is necessary to build intelligent IDS with effective learning abilities in order to secure the network from both internal and external attacks.

In this paper, an intelligent agent based IDS that uses a multilevel classifier and also a decision maker agent for intelligently detecting the intruders in WSNs has been proposed and implemented so that it can provide effective security to WSNs. This intelligent system uses a combination of enhanced decision tree classifier and an Enhanced Multiclass SVM algorithm for binary classification of the past data as well as the current data.

We have enhanced the Support Vector Machines (SVM) for classification since SVM are the classifiers which are more effective in binary classification [19] [20]. In this work, we have combined SVMs with decision trees in order to design multiclass SVMs, which are capable of classifying the four types of attacks namely probing, DOS, U2R and R2L and normal data more accurately. The main focus of this paper is to provide a combined approach to detect the DDoS attacks which improves the training time, testing time and accuracy of IDS using this approach.

The reminder of this paper is organized as follows: Section 2 provides a survey of related works in the area of

misuse and anomaly detection, Decision trees and SVM. Section 3, depicts the architecture of the system proposed in this paper. Section 4 discusses about the proposed enhanced decision tree algorithm [12] and the enhanced multiclass SVM algorithm with agents. Section 5 shows the results obtained from this work and compares them with the existing works. Section 6 gives the conclusions on this work and suggests some possible future enhancements.

II. LITERATURE SURVEY

There are many works in the literature that deal with classification techniques [3] [16] [19]. For example, an algorithm called Tree structured Multiclass SVM has been proposed by Snehal A.Mulay et. al [15] for classifying data effectively. Their paper proposed decision tree based algorithms to construct multiclass IDS which are used to improve the training time, testing time and accuracy of IDS. However, the detection rate is not sufficient in the current internet scenario.

Multiple level tree classifiers were proposed by various researchers in the past [7] [8] [18] in order to design effective IDSs. In such systems, the data are split into normal DOS, PROBE and others (a new class label U2R and R2L). In the second level, the algorithm split the "others" into its corresponding U2R and R2L, while the third level classifies the attacks into its individual specific attacks. However, it is necessary to classify the DOS attacks with a special attention to improve the network performance.

Zeng and Wu et al [21] introduced a new anomaly detection approach based on multi-attribute decisional framework. The classification of data pattern is performed using K-nearest neighbour's method and SVM model. Experiments performed by them with KDD Cup 99 dataset demonstrate that their proposed method achieves good detection accuracy.

Kim and Reddy (2008) et al [22] introduced an anomaly IDS, which monitors packet headers of network traffic. It operates in postmortem but in real-time. The frequent attacks on network infrastructure, using various forms of DoS attacks, have led to an increased need for developing techniques for analyzing network traffic.

Cherkasova et al (2009) [23] proposed a novel framework that provides a powerful solution for automated anomaly detection and analysis of changes in application behavior. The online regression-based transaction model proposed in their work accurately detects a change in the Computational Power consumption pattern of the application and alarms about either observed performance anomaly or possible application change. One of the limitations of their work is that it cannot distinguish which of the transactions is responsible for a changed CPU consumption of the application. To complement the regression-based approach and to identify the transactions that cause the model change, they used the application performance signature that provides a compact model of runtime behavior of the application.

Techniques for the design and evaluation of Intrusion Detection models for Wireless Networks using a supervised classification algorithm and to evaluate the performance of the Multilayer Perceptron (MLP), and Support Vector Machine (SVM) has been provided by Aikaterini Mitrokotsa et. al [1]. The results provided by them point out that SVM exhibits high accuracy.

A novel architecture of Support Vector Machine classifiers utilizing binary decision tree (SVM -BDT) for solving the multiclass problems has been provided by Gjorgji Madzarov et. al [5]. This architecture provides techniques for achieving better classification accuracy.

In this paper, we propose an intelligent agent based multilevel classifier for IDS that uses a combination of decision tree classifier, enhanced C4.5 algorithm and intelligent agents for effective detection of intrusions in WSNs. This system applies the Enhanced Multiclass SVM algorithm for improving the training time, testing time and accuracy of IDS to reduce the false alarm rate. Comparing with existing works, the work proposed in this work different in many ways. First, this system uses intelligent agents for effective classification of DoS attacks. Second, this system uses a hybrid classification scheme for detecting intrusion. Finally, this system uses an enhanced C4.5 algorithm for effective classification.

III. SYSTEM ARCHITECTURE

The multilevel hybrid IDS architecture proposed in this paper is presented schematically in figure 1. This system consists of three modules where the tree classifier agent uses enhanced C4.5 algorithm with agent decision for constructing decision tree which is used to find misuse detection. The classification module uses Agent Multiclass SVM for unsupervised anomaly detection. Finally, for refined classification of anomaly detection, the agent based tree classifier has been used in this work.

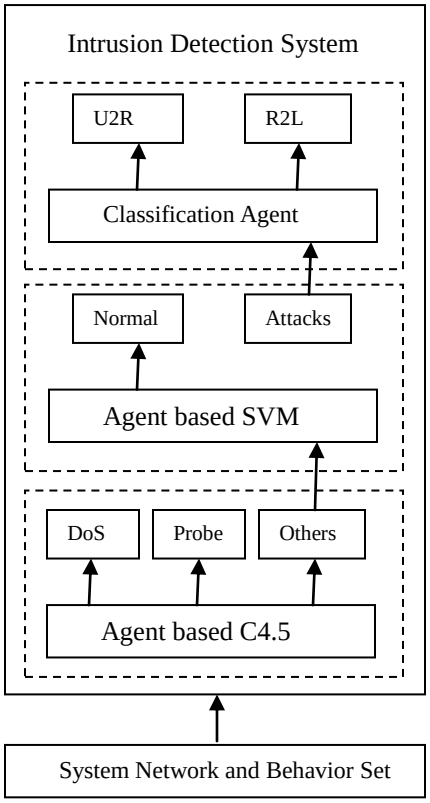

Figure 1 System Architecture

A. Classification Agent

This agent collects the KDD cup'99 data set and constructs a decision tree using the enhanced C4.5 algorithm where intelligent agents are used for making decision. The input data is mined and the specific attributes that have high information gain are only used in the construction of the decision tree where the KDD cup data is classified into DoS and the OTHERS categories. The DoS contains all specific Denial of service attacks like smurf, land, Neptune, back and teardrop. The others include the PROBE category as well. It contains the attacks including ipsweep, nmap, portsweep and Satan. All the other types of attacks and the normal connections are grouped into the OTHERS category.

The decision tree is first trained with the training data and decision tree is generated. The tree is pruned using agent to optimize the number of nodes in the decision tree. From the pruned decision tree, rules are formed. The rules are then applied to the test data and the input data is thus classified using the rules generated during the training phase.

- Stimulus/Response
  Stimulus: Collected training data
  Response: Decision tree with classified data

B. Behavior partition Module

In the Agent based Multiclass SVM algorithms, it is necessary to fix the number of classes are fixed in the beginning of classification. We use the agent based classification technology to determine the number of classes automatically. This Agent based Multiclass SVM algorithm is used to generate a classification whose outputs are normal and attack.

- Stimulus/Response

Stimulus: the part of the output from the decision tree classifier.
Response: two classes with are labeled normal and abnormal (attack).

IV. THE FRAMEWORK OF MULTILEVEL IDS

The main task of the Intrusion Detection System (IDS) is to discover the intrusions from the network packet data or system audit data. One of the major problems that the IDS might face is that the packet data or system audit data could be overwhelming. Moreover, some of the features of audit data may be redundant or contribute little to the detection process. Hence, agent based classification techniques has been used in this work to ease this task.
The network attacks fall into four main categories as discussed in [8].

• DoS (Denial of Service): Intrusions are designed to disrupt a host or network service, e.g. SYN flood;
• PROBE: Attacks include many programs which can automatically scan a network of computers to gather information or find known vulnerabilities as a possible precursor to more dangerous attacks.
• U2R (User to Root): Attacks correspond to a local user on a machine gaining privileges normally reserved for the UNIX root or super user.
• R2L (Remote to Local): Attacks correspond to an attacker who does not have an account on a victim machine, sends packets to that machine and gains local access, e.g. guessing password
In the next section, a brief introduction of the classification algorithms used in the hybrid IDS, i.e., the C4.5 algorithm for building decision trees and the Multiclass Support Vector Machine (SVM) are given.

Agent based Multiclass Support Vector Machine (EMSVM) Algorithm

In this section, we describe the intelligent Multiclass SVM algorithms, and illustrate how to apply this algorithm to generate anomaly type intrusion detection models. Figure 3 pictorially represents anomaly detection system discussed in this paper.

This agent based Multiclass Support Vector Machine (MSVM) algorithm is as follows: First, we first compute the distance between two classes of patterns and repeat it for each class of such patterns.

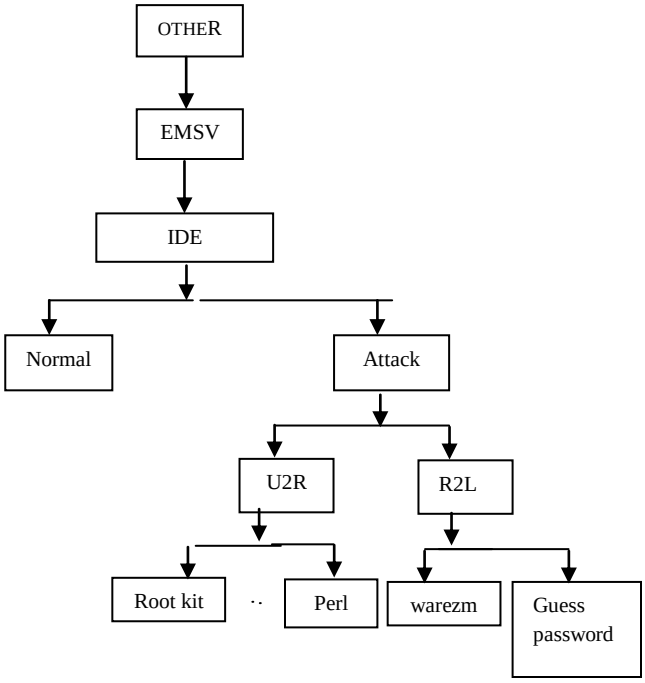

Figure 2. Behavior partition Module

where the distance between two classes is computed using the Minkowski Distance. According this method, the distance between two points

$$P = (x_1, x_2, \ldots, x_n) \text{ and } Q = (y_1, y_2, \ldots, y_n) \in R^n$$

is defined as

$$\left(\sum_{i=1}^{n} | x_i - y_i |^p \right)^{1/p}$$

where p is the order.
We find the center point of every class by using the formula

$$Ci = \sum^{nt}{}_{m=1} X^i{}_m \; / \; n_i$$

After this calculation, five classes obtained earlier are converted into two classes. For example let A, B, C, D and E be five classes. If the distance between any two classes are less than that of the other classes then that pair is replaced by 1(Normal). Otherwise, it is replaced by -1 (Attacker). So, at end of the repeated process, we have only 1's and -1's combinations. Since -1 classes are removed, the remaining classes are used to construct the tree.

The steps of the algorithms are as follows:
Algorithm: Search (E, n).

Input: Data set E, the number of sampling.

Output: Initial center (m1, m2)

[1] Sampling E, get S_1, S_2,…,S_n

[2] For i=1 to n do

Mi =Count_m(Si);

[3] For i=1 to n do

M= Count_m(mi);

[4] m1=m, m2=max(Sim(m,m_i));

[5] Check it with agent threshold to make final descision.

Enhanced Multiclass Support Vector Machine algorithm

[1] Confirm two initial cluster centers by algorithm search m.
[2] Import a new class C.
[3] Compute the Minkowski distance between two classes.
[4] if (dAB > dAC) then
B is assigned as Normal
Else C is assigned as Attacker.
[5] Find the min & max of the distance.
[6] If (dAB < threshold limit of the distance) then create a new cluster and this is the center of the new cluster.
Else
B is assigned as an Attacker.
[7] Repeat the operation until reduced the difference between the classes.
[8] Validate this difference using agent.

V. EXPERIMENTATION AND RESULTS

A. Training and Test Data

The dataset used in the experiment was taken from the Third International Knowledge Discovery and Data Mining Tools Competition (KDD Cup 99). Each connection record is described by 41 attributes. The list of attributes consists of both continuous-type and discrete type variables, with statistical distributions varying drastically from each other, which makes the intrusion detection a very challenging task.

B. Experimental Results

Table 1 shows the comparison between C 4.5 and agent based C 4.5 with respect to DoS, Probe and Other types of attacks. From this table, it can be observed that the intrusion detection rate is improved in agent based C 4.5 when it is compared with the existing C 4.5 algorithm. This result was obtained by carrying out the experiments 20 times and then by taking the average detection rate. The corresponding bar chart representation is shown in figure 3.

Table 1. Detection Rates (%) variation between C 4.5 and Agent based C4.5

Category	C 4.5	Agent based C 4.5
DoS	99.19	99.59
Probe	99.71	99.82
Others	66.67	68.06

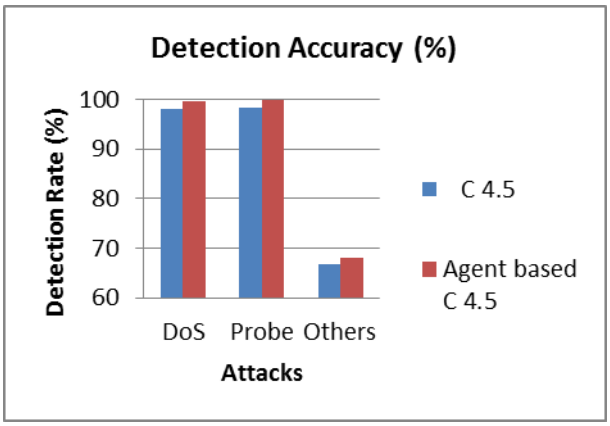

Figure 3 Comparison of Performance Analysis of C4.5 and Agent based C4.5

Table 2 shows the comparison of results obtained from Multiclass Support Vector machine(MSVM) and Agent based Multiclass Support Vector Machine algorithms. From this table, it is observed that the false positive rate has been reduced in the agent based MSVM when it is compared with MSVM. Moreover, the classification time is also reduced in the agent based MSVM due to the effective decisions made by the agent in classification.

Table 2. The Results Comparison between MSVM and AMSVM

Algorithm	TN	Accuracy (%)	R-error (%)	T-time (Sec)
MSVM	14756	83.5821	6.5147	846
Agent based MSVM	15332	92.4612	5.2136	223

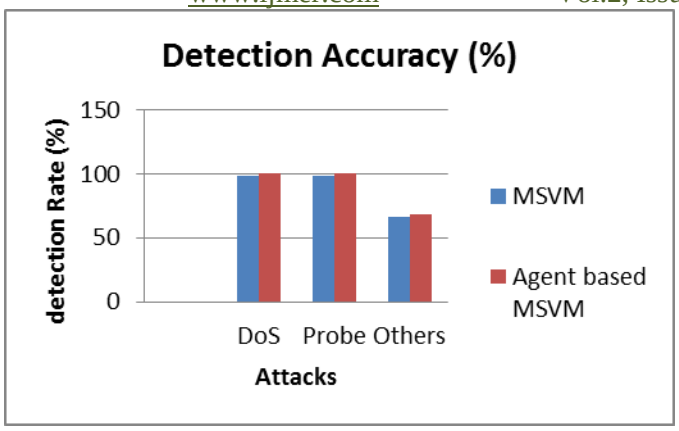

Figure 4. The Results Comparison between MSVM and Agent based MSVM

Figure 4 shows the comparison of intrusion detection rate between MSVM and Agent based MSVM algorithms. From this, it can be seen that the agent based MSVM performs well with respect to detection rate.

VI. CONCLUSION AND FUTURE WORKS

In this paper, an intelligent intrusion detection system using an agent based multi-level classification model combining Decision trees and an agent based Multiclass Support Vector Machines has been proposed. From the experiments conducted in this work, it can be concluded that the agent based intrusion detection system improves the detection accuracy by 7% and 5% when it is compared with C 4.5 and SVM algorithms for DoS attacks. Moreover, it reduces the false positive rate by 1% when it is compared with existing system. Further works in this direction could be the use of effective preprocessing techniques for attribute selection in the IDS to enhance the performance.

REFERENCES

[1] Aikaterini Mitrokotsa, Manolis Tsagkaris and Christos Douligeris, "Intrusion *Detection in Mobile Ad hoc Networks using Classification Algorithms*", 2009.

[2] Denning D E, "*An Intrusion Detection Model*",. *IEEE Transactions on Software Engineering, Vol. 51, no. 8, pp. 12-26, Aug. 2003.*

[3] Denning D E, "*An Intrusion-Detection Mode1*", *IEEE Transaction on Software Engineering*, Vol. 13, No. 3, pp. 222-232, April 2000.

[4] Denning.D.E., Neumann.P.G, "*Requirements and Model for IDES-A Real-Time Intrusion Detection System*", *Technical Report, Computer Science Laboratory, SRI International,Menlo Park, California, pp.58-63,1985.*

[5] Gjorgji Madzarov, Dejan Gjorgjevikj and Ivan Chorbev,"*A Multiclass SVM Classifier Utilizing Binary Decision Tree*", *Informatica33, 2009, pp. 233-241.*

[6] Jun GUO, Norikazu Takahashi, Wenxin Hu, "*An Efficient Algorithm for Multiclass Support Vector Machines*", *IEEE-2008.*

[7] *KDD Cup 1999 Data, Information and Computer Science, University of California, Irvine.* http://kdd.ics.uci.edu/databases/kddcup99/kddcup99. html

[8] Lippmann R.P., Fried D.J., Graf I., Haines J.W., Kendall K.R., McClung D., Weber D., Webster S.E., Wyschogrod D., Cunningham R.K., and Zissman M.A.,"*Evaluating Intrusion Detection Systems: The 1998 DARPA Off-Line Intrusion Detection Evaluation*", *in Proceedings of the 2000 DARPA Information Survivability Conference and Exposition (DISCEX), Vol. 2, IEEE Computer Society Press: Los Alamitos, CA, pp.12-26, 2000.*

[9] Lee.W and S. J. Stolfo, "*A Framework for Constructing Features and Models for Intrusion Detection Systems*," *ACM Transactions on Information and System Security*, Vol. 3, No. 4, pp. 227-261, Nov. 2000.

[10] Latifur Khan, Momoun Award, Bhavani Thuraisingam," *A new Intrusion Detection System using Support Vector Machines and hierarchical clustering*", The VLDB Journal DOI 10, 1007/s00778 – 006 – 0002, 2007.

[11] Mahbod Tavallare.Y, Ebraham Bagheri, Wei Lu and Ali A.Ghortban,"*A Detailed Analysis of KDD Cup'99 data set*", *Symposium on Computational Intelligence Security on Defence Applications (CISDA),IEEE-2009.*

[12] Quinlan.J.R, "*C4.5: Programs for Machine Learning*", Morgan Kauffman, 1993.

[13] Sandya Peddabachigari, Ajth Abraham, Crina Grosan, Johnson Thomas, "*Modeling Intrusion Detection Systems using Hybrid Intelligent Systems*", *Journal of Network and Computer Applications-2005.*

[14] Stefan Axelsson, "*Intrusion Detection Systems: A Survey and Taxonomy*", Technical Report No 9, Dept. of Computer Engineering, Chalmers, University of Technology, Sweden, pp. 9-15, 2000.

[15] Snehal A.Mulay, P.R. Devale, G.V. Garje, "*Intrusion Detection System using Support Vector Machine and Decision Tree*", International Journal of Computer Applications, Volume3-No.3, pp.0975-8887, June 2010.

[16] Teresa L, Ann T and Fred G, "*A Real-Time Intrusion Detection Expert System(IDES)*" Technical Report, Computer Science Laboratory, SRI International, Menlo Park, California, pp.158-163,1992.

[17] Xiaodan Wang, Zhaohui Shi, Chongming Wu and Wei Wang, "*An Improved Algorithm for Decision Tree based SVM*", Intelligent Control and Automation, 2006. (WCICA 2006) , pp. 4234 - 4238 , IEEE -2006.

[18] Xiang.C, M.Y.Chong and H.L.Zhu, "*Design of Multiple-Level Tree Classifier for Intrusion Detection System*", In *Proceedings of 2004 IEEE Conference on Cybernetics and Intelligent Systems*, Singapore, pp.872-877, Dec. 2004

[19] Zhi-song Pan, Songcan Chen, Gen-bao Hu, Dao-qiang Zhang, "*Hybrid Neural Network and C4.5 for Misuse Detection*", Proceedings of the Second IEEE International Conference on Machine Learning and Cybernetics, pp. 2-5 , November 2003.

[20] Zhi-xin Yu, Jing-Ran Chen and Tian-Qing Zhu, "A Novel Adaptive Intrusion Detection System Based on Data Mining", In proceedings of the fourth international Conference on Machine learning and Cybernetics, Guangzhou, pp. 2390-2395, August 2005.

[21] Zeng, Q. and Wu, S. "Anomaly detection based on multi-attribute decision", WRI Global Congress on Intelligent Systems, Vol. 2, pp. 394-398, 2009.

[22] Kim, S.S. and Reddy, A.L.N. "Statistical techniques for detecting traffic anomalies through packet header data", IEEE/ACM Transactions on Networking, Vol. 16, No. 3, pp. 562-575, 2008.

[23] Cherkasova, L., Ozonat, K., Symons, J. and Smirni, E. "Automated anomaly detection and performance modelling of enterprise applications", ACM Transactions on Computer Systems, Vol. 27, No. 3, pp. 1-32, 2009.