World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:7, No:12, 2013

A Proposed Optimized and Efficient Intrusion Detection System for Wireless Sensor Network

Abdulaziz Alsadhan, Naveed Khan

***Abstract*—**In recent years intrusions on computer network are the major security threat. Hence, it is important to impede such intrusions. The hindrance of such intrusions entirely relies on its detection, which is primary concern of any security tool like Intrusion detection system (IDS). Therefore, it is imperative to accurately detect network attack. Numerous intrusion detection techniques are available but the main issue is their performance. The performance of IDS can be improved by increasing the accurate detection rate and reducing false positive. The existing intrusion detection techniques have the limitation of usage of raw dataset for classification. The classifier may get jumble due to redundancy, which results incorrect classification. To minimize this problem, Principle component analysis (PCA), Linear Discriminant Analysis (LDA) and Local Binary Pattern (LBP) can be applied to transform raw features into principle features space and select the features based on their sensitivity. Eigen values can be used to determine the sensitivity. To further classify, the selected features greedy search, back elimination, and Particle Swarm Optimization (PSO) can be used to obtain a subset of features with optimal sensitivity and highest discriminatory power. This optimal feature subset is used to perform classification. For classification purpose, Support Vector Machine (SVM) and Multilayer Perceptron (MLP) are used due to its proven ability in classification. The Knowledge Discovery and Data mining (KDD'99) cup dataset was considered as a benchmark for evaluating security detection mechanisms. The proposed approach can provide an optimal intrusion detection mechanism that outperforms the existing approaches and has the capability to minimize the number of features and maximize the detection rates.

***Keywords*—**Particle Swarm Optimization (PSO), Principle component analysis (PCA),Linear Discriminant Analysis (LDA), Local Binary Pattern (LBP), Support Vector Machine (SVM), Multilayer Perceptron (MLP).

I. Introduction

THE importance of computer network is growing day by day due to its sharing nature; aim is to easily access the information in limited area. The idea was more précised by connecting these computer networks through internet to make it more comprehensive and form global village. The accessibility & flexibility of information sharing is made easy around the world, using internet but also raised some issues related to different types of attacks. The proper security mechanism might capable the system to prevent unauthorized access to the information or services. The security of the network is more important in order to ensure its accountability, confidentiality, availability, and integrity. It also imperative in other internal or external threats like DOS attacks, email based attacks, worms and Trojans etc.

To overcome such kind of attacks, different security techniques are available. The detection is mandatory for sensitive data in organization and governmental sectors. This illegal access may cause a severe damage to network, system users and also sensitive data can be lost. These systems are designed to implement and detect intrusions and allow the normal traffic packets to pass through the network known as intrusion detection system (IDS).

The performance of the IDS depends on the different parameters such as feature transformation algorithm like PCA, LDA, LBP etc.; optimal subset selection using stochastic techniques such as greedy search, back elimination, Particle Swarm Optimization (PSO), and optimal data sets can be further classified through classifier like SVM or MLP etc.

The performance of the IDS depends on detection rate and false alarm. A number of techniques have been introduced to increase the detection rate and minimize the false alarm. The key issue is to optimize the performance of the IDS by selecting a proper feature transformation algorithm, Apply a stochastic algorithm for optimal subset selection and finally classifier used to classify this dataset to get optimized results. Such issues have motivated the research for optimal feature selection, stochastic algorithm and classification mechanism presented in this paper.

The ideal intrusion detection system for classifying intrusive and normal data has always been an important subject for researchers. Different techniques have been implemented to get better and maximum performance.

There are some factors which can affect the performance of IDS like false positive and false negative and true positive and true negative. In detection, for feature transformation previously used technique like Principle Component Analysis (PCA) have some problem as highlighted in [1], [8] and [9] to select a large number of Principle Components (PCs) decrease training and testing efficiency, Loose sensitive features, and more complex architecture as PCs increased.

Furthermore to be more precise and get the more optimal subset heuristic search algorithm applied like Genetic Algorithm used. GA has limitation identified in [2] and [6] as computation overhead i.e. crossover and mutation, complex architecture, no quick convergence. The multilayer perceptron used to classify the optimal data sets. An issue related to this classifier is that it only solves problems that are linearly separable [3] and [7]. The other limitations were generalization ability, local minima and insufficient

Abdulaziz Alsadhan is with Department is with Software Engineering, College of Computer and Information Sciences, King Saud University, Saudi Arabia.

Naveed Khan is with Department is with Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia (e-mail: naveed@ksu.edu.sa).

knowledge of input samples. The KDD (Knowledge Discovery and data mining) data set were used mostly in experiments for classification of normal and intrusive packets. The deficiencies' found in KDD [4] data set were redundant records in the train set, duplicate records in the test data, selected records from each difficulty level group is inversely proportional to the percentage of records in the original KDD data set. These are the problems which can be solved by the proposed architecture.

The focus of this work based on the problems and issues with the previous approaches, described as below:

- How to improve current performance of Intrusion Detection System (IDS)?
- How to improve feature transformation process?
- How to enhance the feature subset selection?
- How to achieve the best possible classification accuracy?
- How to evaluate the performance of the system?

II. Literature Review

Different approaches were evaluated to achieve the required features of an intrusion detection system [11]. The general category of intrusion detection system is anomaly detection described in [12]. The mechanism used in this approach was to categorize activities, which differ from established patterns for users. It also comprised the conception of knowledge bases to make profiles for the scrutinized activities [12]. The benefit of this approach were unknown attacks can be detected through anomaly detectors, signature of novel intrusion and used detector to identify signature for misuse detectors. The drawback of this approach was high false positive rate and extensive data sets used to build profile of the system.

In [13] the authors explained the misuse detection technique for intrusion detection. In this technique to permeate a system, the evaluation of user's activities observed with the identified activities of attacker. The knowledge base of information can be exploits by misuse detection. The advantage of this approach were effective in detecting attacks without giving high false alarm rates, detect intrusions with known signatures and easily deployable in terms of state machine and signature analysis. The approach has the disadvantages of repository for limited signatures and maximum time consumption.

The approaches discuss in [12] and [13] thoroughly analyzed by researchers to evaluate the junction of anomaly detection approach and misuse detection approach. The combination of these approaches discussed in [14], which ratified a single intrusion detection system to monitor the internal and external attacks. The current intrusion detection system analyzed and provides protection to a system in the following scenario.

- The authorized and unauthorized activities by system users can be monitored.
- Identify those actions having indication for an attack on a system.
- The information attained from an intrusion detection system can be used to improve the overall security of the network.
- To enable system to make the analysis of an attack in real time.

III. Proposed System Architecture

Intrusion detection system (IDS) plays an important role in securing a network when deployed. It monitors and inspects malicious and suspicious activities of network traffic, and warns by initiating alarm to the security administrators. The inbound and outbound network traffic scrutinizes for potential security such as Trojans, virus, malware etc. while communicating on the internet [15]. The IDS architecture use sensors to identify malicious packets associating alert message and forward to server. A log file is created to analyze such packets and make it potential security threat. Generally two tools used by IDS network based and Host based. In network-based IDS numerous sensors used to deploy around the network, often beside the gateway router. A report sent to the central application to perceive anything that could hazardous or suspicious. The security of an individual system attained by deploying a host-based IDS.IDS performs a significant role in securing network because it actively guards the network 24 hours a day [15] as shown in Fig. 1.

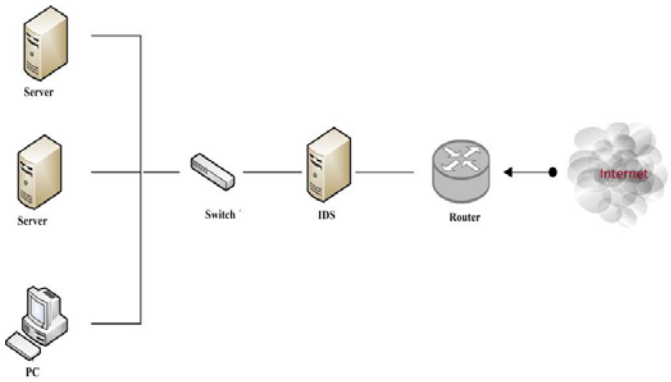

Fig. 1 IDS Architecture

The main goal of this research is to propose an optimized Intrusion detection mechanism using soft computing. The objective is to improve the performance and identify each activity in a robust way. The performance of IDS will be analyzed using following parameters:

- Feature transformation algorithm like PCA, LDA, and LBP etc.
- Optimal subset selection using stochastic techniques like greedy search, back elimination, Particle Swarm Optimization (PSO).
- Optimal data sets can be further classified through classifier like SVM or MLP etc.

In order to accomplish this goal, following objectives have been set to achieve.

- To design a suitable mechanism to enhance the performance by increasing detection rate and reducing false alarm.
- To implement and explore different transformation techniques like PCA, LDA and LBP.

- To implement & analyze different feature subset selection approaches like greedy search, back elimination, Particle Swarm Optimization (PSO)
- To implement SVM or MLP classifier to classify normal and intrusive packets.
- To train and test the system using standard testing parameters like ROC, Confusion Matrix etc.

IV. IDS Framework

The main detailed IDS evaluation framework is shown in Fig. 2.

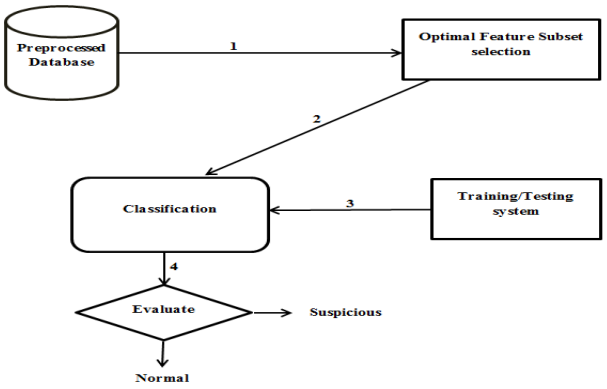

Fig. 2 Evaluation framework of IDS

A. Analyze the Knowledge Gap

The first step for proposing new research subject matter is the knowledge gap in existing approaches. The limitations in previous mechanism bring many new research issues and their explanation for contribution towards improved and optimal research solution. Intrusion detection systems and their performance mechanism has always been under discussion regarding their selection of datasets, their transformed and acceptable form for input, classifier architecture and its properties for training and testing the system in order to evaluate optimum results. For intrusion detection system (IDS), a new standard datasets KDD'99 cup has introduced by eliminating some deficiencies in old KDD cup. KDD'99 dataset [4] are available now with. The most important part of IDS is the optimal feature selection and classifier architecture. This research covers SVM and MLP classifier to be trained and tested with standard dataset in order to find optimal classifier in IDS with maximum accuracy and decreased false alarms.

B. Selection of Dataset for Experiments

The suitable selection of dataset is very important for evaluating the performance of the intrusion detection system. There are two standard datasets to be used in this research which are KDD'99 cup (includes four types of attacks namely DOS, Probe, R2L and U2R) and CAIDA (DDOS) datasets [5] and [10]. The performance of the Intrusion Detection System depends upon the accuracy of dataset provided in the training phase for learning and then the results of the dataset in testing phase to determine the classification ability of the proposed system [1].

C. Pre- Processing of the Selected Dataset

In this phase the selected dataset is pre-processed for our IDS mechanism before they are utilize in training phase. Selected datasets are transformed in feature space using PCA, LDA, or LBP and only required optimal feature set with high dimensional space are selected using PS), Greedy Search or Back Elimination. The purpose of preprocessing is that irrelevant data is discarded and thus reducing overhead and it can be directly input to our system [1].

D. Feature Transformation & Optimal Feature Subset Selection

In this phase two steps will be taken. In first step the feature transformation techniques will be used such as PCA, LDA or LBP. By using these techniques the raw feature will be transformed into principle features space and select the features based on their sensitivity. The new principle features will be more visible, organize arrange and sensitive that directly affect the performance of intrusion detection mechanism. Moreover these new principle features are further analyzed to get the optimal feature subset. To accomplish this feature subset selection approaches will be used like PSO, Greedy Search and Back elimination to get feature subsets randomly. These feature subsets are classified and get the optimal feature subset which can effectively improve the performance of the IDS system.

E. Classification Approach

The classification approaches will be used in order to find the normal and intrusive packets. Two types of approaches will be used for intrusion detection; multilayered perceptron (MLP) and Support Vector Machine (SVM). In the architecture of MLP, one hidden layer is equivalent to SVM. To classify network activity into normal and intrusive, the performance of both architectures can be compared in term of their discriminatory power and efficiency.

F. Training the System

In this phase the proposed system is trained with selected datasets for each classifier i.e. SVM and MLP. For every desired feature input there is relevant output pattern which is matched after the actual output is produced during the training phase. The purpose of training with benchmark datasets is to reduce difference between computed output and target output. After this training, the system performance will be tested for optimal classifier i.e. SVM or MLP.

G. Testing the System

Testing of the proposed classifier such as SVM and MLP is done by datasets selected for evaluating the performance of our proposed system, after the system is well trained with the selected inputs and desired outputs. The datasets are tested in two steps for the each of the classifier. The first step is verification step in which same training data is used to test the learning ability of the system in training phase. Then the next step is generalization step in which different testing dataset is provided to check the how well the system performs with its generalization capabilities.

H.Evaluating and Analyzing Results

In the methodology phase, the experimental results obtained for the selected datasets from the testing phase from each classifier. Each classifier is evaluated individually in order to find optimal classifier.

V. Conclusion

This research work is based on the performance of the of the intrusion detection system. The main contribution is to develop an intrusion detection system that performs excellent compared to previous approaches in term of high detection rate and minimize number of features.

The main impact and contribution of this research includes:

- Provide optimized performance in terms of reduce false alarms (false positive and false negative) and improve detection rate. Minimize training and computational overhead.
- Improved Architectural framework.
- Optimal intrusive analysis engine.
- Help in the intrusion detection mechanism for security implementers.

Moreover, in this work, an optimized intrusion detection system mechanism using soft computing techniques; PCA, LDA, LBP, PSO, Greedy Search, SVM and MLP will be proposed and implemented. The two standard datasets will be used in this research are Knowledge Discovery and data Mining (KDD'99) cup and the Cooperative Association for Internet Data Analysis (CAIDA), which are considered benchmarks for evaluating security detection mechanisms [1]. The proposed techniques like PCA, LDA, LBP, PSO, GreedySearch, Back Elimination, SVM and MLP will be analyzed and evaluated consequently, the result will show that the proposed approach will outperform than the existing approach having capability to reduce the number of features and increase the detection rates.

References

[1] Ahmad I, "Feature Subset Selection in Intrusion Detection Using Soft Computing Techniques, "Ph.D. dissertation", Dept., CIS, UTP., Tronoh, Perak, Malaysia, 2011.
[2] http://www.swarmintelligence.org/bibliography.php, [Accessed on 15th September, 2013].
[3] Leonardo N "Multilayer Perceptron Tutorial" School of Computing Staffordshire University Beaconside Staffordshire, 2005.
[4] http://iscx.ca/NSL-KDD/ [Accessed on 20th September, 2013].
[5] CAIDA: The Cooperative Association for Internet Data Analysis [online]. Available: http://www.caida.org [1st April, 2011].
[6] Ahmad I, Abdullah A, and Alghamdi AS. "Application of Artificial Neural Network in Detection of Probing Attacks" IEEE Symposium on Industrial Electronics and Applications (ISIEA). Kuala Lumpur, Malaysia. pp. 557 – 562, 2009.
[7] Ahmad I, Abdullah AB, and Alghamdi AS "Artificial Neural Network Approaches Intrusion Detection: A Review" Telecommunications and Informatics conference. Istanbul, Turkey, pp. 200-205, 2009.
[8] Bace R, and Mell P. "Intrusion Detection Systems. National Institute of Standards and Technology (NIST) Special Publication. pp. 1-51, 2001.
[9] Tang P, Jiang R and Zhao M "Feature Selection and Design of Intrusion Detection System Based on k-Means and Triangle Area Support Vector Machine". International Conference on Future Networks (ICFN), pp. 144-148, 2001.
[10] Zhang X "Intrusion Detection System Based on Feature Selection and Support Vector Machine". IEEE conference on Communications and Networking China, pp. 1-5, 2006.
[11] Fox KL, Henning RR, Reed JH, Simonian RP "Information Systems Security" International conference on Computer Security, pp. 124-134, 1990.
[12] Khan L, Awad M, and Thuraisingham B. "A New Intrusion Detection System Using Support Vector Machines and Hierarchical Clustering" International Journal on Very Large Data Bases 16(4):507–521,2010.
[13] Mukherjee B, Heberlein LT, and Levitt KN "Network Intrusion Detection", IEEE Network. pp. 26-41,1994.
[14] Pervez S, Ahmad I, Akram A, and Swati SU "A Comparative Analysis of Artificial Neural Network Technologies in Intrusion Detection Systems" WSEAS Transaction on Computers. pp. 175-180, 2007.
[15] Mike Meyers' "Managing and Troubleshooting Networks", Second Edition, 2009.