ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

Intrusion Detection Systems: A Survey and Analysis of Classification Techniques

V. Jaiganesh [1], S. Mangayarkarasi [2], Dr. P. Sumathi [3]

Assistant Professor, Department of Computer Science, Dr. N.G.P Arts and Science College, Coimbatore
Doctoral Research Scholar, Manonmaniam Sundaranar University, Tirunelveli, India[1]

M.Phil. Scholar, Department of Computer Science, Dr. N.G.P. Arts and Science College, Coimbatore, India[2]

Doctoral Research Supervisor, Assistant Professor, PG & Research Department of Computer Science, Government Arts College, Coimbatore, India[3]

Abstract: Today it is very important to provide a high level security to protect highly sensitive and private information. Intrusion Detection System is an essential technology in Network Security. Nowadays researchers have interested on intrusion detection system using Data mining techniques as an artful skill. IDS is a software or hardware device that deals with attacks by collecting information from a variety of system and network sources, then analyzing symptoms of security problems. This paper includes an overview of intrusion detection systems and introduces the reader to some fundamental concepts of IDS methodology. We also discuss the primary intrusion detection techniques. In this paper, we emphasizes data mining algorithms to implement IDS such as Support Vector Machine, Kernelized support vector machine, Extreme Learning Machine and Kernelized Extreme Learning Machine.

Keywords: SVM, KELM, Intrusion Detection System, Data Mining and IDS, ELM, Classification Techniques for IDS, KSVM

1. INTRODUCTION

Every computer is always at risk for unauthorized and intrusion, however, with sensitive and private information are at a higher risk. Intrusion Detection is a key technique in Information Security plays an important role detecting different types of attacks and secures the network system. Intrusion Detection is the process of observing and analysing the events arising in a computer or network system to identify all security problems. IDS provides three important security functions; *monitor*, *detect* and *respond* to unauthorized activities [1]. Intrusion Detection System monitors the operations of firewalls, routers, management servers and files critical to other security mechanisms. Intrusion Detection System can make the security management of system by non-expert staff possible by providing user friendly interface.

Intrusion detection systems usually provide the following services:

- Observing and analysing computer and/or network system activity
- Audit the system configurations and vulnerabilities
- Evaluating the integrity of critical system and data files
- Estimating abnormal activities

2. IDS TAXONOMY

IDSs are divided into two broad categories: host-based (HIDS) and network-based (NIDS) [2]. A host-based IDS requires small programs (or agents) to be installed on individual systems to be supervised. The agents monitor the operating system and write down data to log files and/or trigger alarms. A network-based Intrusion Detection System usually consists of a network application (or sensor) with a Network Interface Card (NIC) working in promiscuous mode and a separate management of interface. IDS is placed on a network segment or boundary and monitor all traffic on that segment. The current trend in intrusion detection is to combine both host based and network based information to develop hybrid systems that have more efficient.

2.1 Host Based Intrusion Detection (HIDS):

Host based intrusion detection (HIDS) refers to intrusion detection that takes place on a single host system. The data is collected from an individual host system. The HIDS agent monitors activities such as integrity of system, application action, file changes, host based network traffic, and system logs. By using common hashing tools, file timestamps, system logs, and monitors system calls and the local network interface gives the agent insight to the present state of the local host. If there is any unauthorized change or

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

activity is detected, it alerts the user by a pop-up, it alerts the central management server, blocks the activity, or a combination of the above three. The decision should be based on the policy that is installed on the local system. These host-based procedures are considered the ***passive*** component.

2.2 Network Based Intrusion Detection (NIDS):
A network-based intrusion detection system (NIDS) is used to monitor and analyse network traffic to protect a system from network-based threats where the data is traffic across the network. A NIDS tries to detect malicious activities such as denial-of-service (Dos) attacks, port scans and monitoring the network traffic attacks. NIDS includes a number of sensors to monitors packet traffic, one or more than servers for NIDS management functions, and one or more management relieves for the human interface. NIDS examines the traffic packet by packet in real time, or near to real time, for attempting to detect intrusion patterns. The analysis of traffic patterns to detect intrusions may be done at the sensors, at the management servers, or combination of the both. These network-based procedures are considered the ***active*** component.

2.3 Hybrid Intrusion Detection:
The current trend in intrusion detection is to combine both types host-based and network-based IDS to design hybrid systems. Hybrid intrusion detection system has flexibility and it increases the security level. It combines IDS sensor locations and reports attacks are aimed at particular segments or entire network.

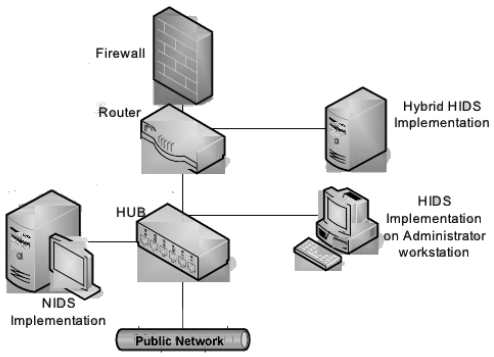

Figure 2.1: Types of Intrusion Detection System

3. INTRUSION DETECTION APPROACHES

There are currently a variety of approaches being utilized to accomplish the desirable elements of an intrusion detection system. There are two general approaches to intrusion detection:

- Anomaly detection
- Misuse detection

These approaches develop the core of several currently present intrusion detection techniques.

Anomaly Detection: Anomaly IDS trying to detect anomalies when any difference occur from the normal system. Anomaly detection is based on audit data gathered over a period of normal operation. Anomaly detection is an important tool for fraud detection, network based intrusion, and other unusual events that have great significance but they are hard to find. The importance of anomaly detection is due to the fact that anomalies in data translate to important actionable information in a huge variety of application domains. Anomaly detection is also sometimes referred to as *behaviour-based detection* because it associates with variations from user behaviour [3].

The advantage of anomaly detection approach is the ability to detect novel attacks or unknown attacks based on audit data.

The main drawback of the anomaly detection approach is that well-known attacks may not be detected.

Misuse Detection: Misuse IDS trying to detect abnormal behaviour by analysing the given traffic and go with several rules based on Analysis and comparison with the Rules the system can notice any attacks, such as matching signature pattern. Misuse detection is also sometimes referred to as *signature-based detection* because alarms are generated based on particular attack signatures. This kind of attack signatures encompass particular traffic or activity that is based on known intrusive activity.

The advantage of misuse detection is the ability to generate accurate result and having fewer false alarms.

The disadvantage of misuse detection approaches is that they will detect only the known attacks [4].

4. DATA MINING – CONCEPTS

Data Mining refers to a process of analysing data from different perspectives and summarizing it using useful information. It is the extraction of the hidden predictive information from large database.
Approaches to Data Mining Problems are:

- Sequential pattern discovery
- Discovery of patterns in time series
- Classification rules
- Neural Networks
- Genetic Algorithms

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

- Clustering and Segmentation

4.1 Goals of Data Mining:
Widely Speaking, the goals of Data Mining fall into the following groups: *prediction, identification, classification and optimization*

4.1.1 Prediction:
Prediction discovers the relationship between independent variables and relationship between dependent and independent variables. Data mining showing how particular attributes within the data will behave in future. In some application, business logic is used coupled with data mining.

4.1.2 Identification:
Data patterns are using to identify the existence of an item, an event, or an activity or some new patterns of customer behaviour. The area known as authentication is a layout of identification.

4.1.3 Classification:
Data Mining can separate the data so that different classes or categories can be recognized based on combination of parameters to find a clever say to display the data.

4.1.4 Optimization:
Data Mining can be optimized the use of limited resources such as time, space, money or materials and to maximize output variables under a given set of constraints.

4.2 Advantages of Data Mining:
Data mining applications are continuously developing in various industries to provide more hidden knowledge that enable to increase business efficiency and grow businesses. Data mining approaches plays an essential role in various domains. For the classification of security problems, a large amount of data has to be analysed containing historical data. It is difficult for human beings to find a pattern in such an enormous amount of data. Data mining, however, seems well-suited to overcome this problem and can therefore be used to discover those patterns.

5. DATA MINING AND IDS

Data mining techniques can be used to build up Intrusion Detection System in real time environment [5]. Data mining techniques can be dissimilated by their different model functions and representations, preference criterions, and algorithms.

There are some important things that contribute for an Intrusion detection implementation using data mining:

- Removing normal activity from alarm data for focusing real attacks
- Identifying false alarms and "awful" sensor signatures
- Finding abnormal activity that uncovers a real attack
- Identifying long and ongoing patterns

To achieve the above tasks, data miners using one or more of the following techniques:

- Data summarization: Summarizes evolutional data with statistics
- Visualization: Summarization of graphical data
- Clustering: collection of similar objects and that are dissimilar with objects in other clusters
- Association : Correlating the existence of a set of items with another range of values for another set of variables
- Classification: Predicting a hierarchy of class from an existing set of events or transactions
- Prediction: Showing how the certain attributes within the data will behave in the future
- Sequential Patterns: Sequence of actions or events is required

5.1 Data Mining assists for Intrusion Detection:
The goal of intrusion detection is to detect security violations in information systems. Intrusion detection is a passive approach to security as it monitors information systems and raises alarms when security violations are founded. Examples for security violations contain the abuse of privileges or the use of attacks to exploit software or protocol vulnerabilities.

Data mining helps to create successful applications in related domains, e.g., detecting fraud, fault/alarm management. Data mining having some maintain or update models on dynamic data.

6. SURVEY OF APPLIED CLASSIFICATION TECHNIQUES FOR INTRUSION DETECTION

Classification is classic data mining technique which is used to predict association for data instances. Classification techniques evaluate and classify the data into known classes. Each data sample is marked with a known class label. Also these techniques are used to learn a model using the training set data sample. This model is used to classify the data samples as anomalous behaviour data or the normal behaviour data [8].

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

The following section presents some of most popular classification algorithms for implementing Intrusion Detection System such as:

- Support Vector Machine (SVM)
- Kernelized Support Vector Machine (KSVM)
- Extreme Learning Machine (ELM)
- Kernelized Extreme Learning Machine (KELM)

6.1 Support Vector Machine (SVM):

Support vector machine proposed as an essential technique for intrusion detection system [3]. It is a machine learning algorithm which is used for both classification and regression. Some standard support vector machines (SVMs) which are powerful tools for data classification, classifies two-category points by assigning them to one of two disjoint half spaces in either the original input space of the problem for linear classifiers, or in a higher dimensional feature space for nonlinear classifiers [6].

Support vector machine (SVM) performs classification by constructing hyper planes in a multi-dimensional space that separates two classes. SVM tries to achieve maximum separation between the classes.

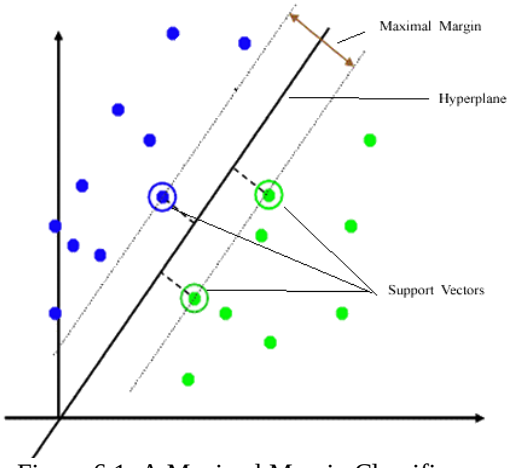

Figure 6.1: A Maximal Margin Classifier

Linear Support vector machine is composed by a set of given support vectors "z" and set of given weights "w".

The output with N support vectors z_1, z_2… Zn and weights w_1, w_2… w_N is given by:

$$F(x) = \sum_{i=1}^{N} w_i \langle z_i, x \rangle + b \qquad (1)$$

In the SVM study, a predictor variable is called an *attribute*, and a converted attribute that is used to describe the hyper plane is called a *feature*. The process of selecting the most suitable representation is known as *feature selection*.

A set of features that describes one case is called a *vector*. That is why the goal of SVM modelling is to find the optimal hyper plane that separates group of classes in a way that cases with one group of the target variable are on one side of the plane and cases with the other group are on the other size of the plane. The points falling in the bounding planes are called as *support vectors*. The figure below draws an overview of the SVM process.

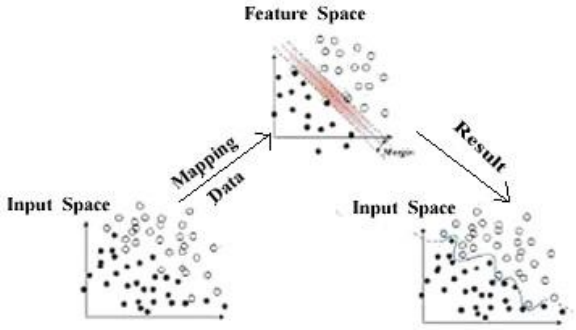

Figure 6.2: A SVM Algorithm

SVM is popular because it can be easy to use and this algorithm often has good generalization performance. And this same algorithm solves a variety of problems with little tuning.

6.2 Kernelized Support Vector Machine (KSVM):

Support Vector Machines (SVM's) can only perform as a linear classifiers and regressors. By using the *kernel trick*, SVM's are able to perform both non-linear classification and regression. Non-linear classifiers are created by applying the "kernel trick" to maximum-margin hyper planes. In the resulting algorithm, every "dot-product" positioned is replaced by a non-linear kernel function.

Kernel Trick:

The kernel trick depends on the "dot product" between two vectors. The kernel function represents an inner product in feature space and it is denoted by,

$$K(x, y) = \langle \varphi(x), \varphi(y) \rangle \qquad (2)$$

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

There are 4 standard kernels that are used to build a SVM classifier. They are,

1. Linear Kernel Function:

$$k(x, y) = x^T y + c$$

2. Polynomial Kernel Function:

$$k(x, y) = (\alpha x^T y + c)^d,$$

Where c and d are parameters defines the kernel's behaviour.

3. Gaussian Radial Basis Kernel (RBF):

$$k(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right),$$

Here **σ** is the width of the function.

4. Sigmoid Kernel (Hyperbolic tangent): $K(q, q') = tanh(aq.q' + b)$,

Where a and b are parameters defines the kernel's behaviour.

Even though the original input is not linearly separable in the input space, the data can be converted into linearly separable in feature space. Hence using kernel provides a way to obtaining nonlinear algorithms from algorithms in the past restricted to handling linear separable datasets [4].

Before fitting nonlinear curves to the data, SVM handles this process by using a *kernel function* to map the data into a different space where a hyper plane can be used for doing this separation.

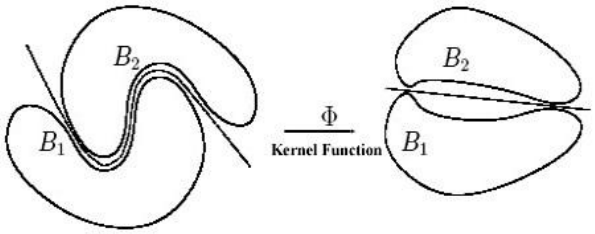

Figure 6.3: Separation of Non-linear region using SVM Kernel functions

The kernel function may convert the data into a higher dimensional space to make it possible to achieve the separation.

Using the kernel functions, the actual formulation for the SVM with support vectors z_1, z_2… Zn and weights w_1, w_2… w_N is now given by:

$$F(x) = \sum_{i=1}^{N} w_i\, k(z_i, x) + b \qquad (3)$$

Using kernel is the powerful methodology because it provides a bridge from linearity to non-linearity to any algorithm.

6.3 Extreme Learning Machine (ELM):

Extreme Learning Machine (ELM) is a new emergent technology which provides good generalization performance for both classification and regression problems at highly fast learning speed. Even though Support Vector Machine can produce better generalization performance, it has tow drawbacks as well: 1) the intensive computation involved in its training which is at least quadratic with respect to the number of training examples 2) For large complex applications, it generates large network size.

Both Neural Networks and Support Vector Machines face some challenging issues such as [11]:

- Intensive human appear
- Slow learning speed
- Poor scalability

Extreme **L**earning **M**achines (**ELM**) has worked for the ``generalized" single-hidden layer feed forward networks (SLFNs). The output function of the generalized SLFN is given by,

$$f(\boldsymbol{x}) = \sum_{i=1}^{L} \boldsymbol{\beta}_i\, \boldsymbol{h}_i(\boldsymbol{x}) \qquad (4)$$

Where $\boldsymbol{h}_i(\boldsymbol{x})$ is the output of the i^{th} hidden node.

6.3.1 ELM Algorithm:

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

Given a training set $\aleph = \{(X_k, t_k) | X_k \epsilon R^n, t_k \epsilon R^m, k = 1, \ldots, N\}$, an activation function g(x) and the number of hidden neurons $\tilde{N}$,

(i) Randomly assign input weights wi and biases bi according to some continuous probability density function.

(ii) Calculate hidden layer of output matrix H.

(iii) Calculate the output weights

Extreme Learning Machine (ELM) has several features:

- It is easy to use
- Faster learning speed
- High generalization performance
- Suitable for all non-linear activation functions
- Suitable for complex activation functions

6.4 Kernelized Extreme Learning Machine (KELM):
This section reviews kernel based ELM. If the hidden layer feature mapping h(x) is unknown to users, users can be described a kernel function for ELM [12].

ELM Kernel function is given by,

$$K_{ELM}(x_i, x_j) = 1/H\ f(x_i) \cdot f(x_j) \quad (5)$$

That is, the data has feed through the ELM hidden layer to obtain the feature space vectors, and their co-variance is then calculated and scaled by the number of hidden units. The main difference is that where ELM explicitly generates the feature space vectors, but in SVM or another kernel method only similarities between feature space vectors are used.
The entire above mentioned can be used to apply in regression, binary and multi-label classification applications directly. Kernel ELMs can be applied to complex space as well.

7. CONCLUSION

Internet and local networks have become everywhere. So organizations are increasingly employing various systems that monitor IT security breaches because intrusion events are growing day by day. This paper describes the different types of intrusion detection system and highlights techniques of intrusion detection. In this paper, we draw attention to data mining algorithms which is used to implement Intrusion Detection System (IDS) such as Support Vector Machine, Kernelized support vector machine, Extreme Learning Machine and Kernelized Extreme Learning Machine. It draws the conclusions on the basis of implementations accomplished using various data mining algorithms. Combining more than one data mining algorithms may be used to eliminate disadvantages of one another.

REFERENCES

[1] Defeng Wang, Yeung, D.S., and Tsang, E.C., "Weighted Mahalanobis Distance Kernels for Support Vector Machines", IEEE Transactions on Neural Networks, Vol. 18, No. 5, Pp. 1453-1462, 2007.
[2] Glenn M. Fung and O. L. Mangasarian, "Multicategory Proximal Support Vector Machine Classifiers", Springer Science and Business Media, Machine Learning, 59, 77–97, 2005.
[3] Guang-Bin Huang, Dian Hui Wang and Yuan Lan, "Extreme learning machines: a survey", Published: 25 May 2011_ Springer-Verlag, 2011.
[4] Hyeran Byun and Seong-Whan Lee, "Applications of Support Vector Machines for Pattern Recognition: A Survey", Springer-Verlag Berlin Heidelberg, 2002
[5] G. Jacob Victor, Dr. M Sreenivasa Rao and Dr. V. CH. Venkaiah, "Intrusion Detection Systems Analysis and Containment of False Positives Alerts", International Journal of Computer Applications *(0975 – 8887)*, Volume *5– No.8,* August *2010.*
[6] Joseph,J.F.C., Bu-Sung Lee, Das,A., and Boon-Chong Seet, "Cross-Layer Detection of Sinking Behavior in Wireless Ad Hoc Networks Using SVM and FDA", IEEE Transactions on Dependable and Secure Computing, Vol. 8, No. 2, Pp. 233-245, 2011.
[7] Kyaw Thet Khaing, "Enhanced Features Ranking and Selection using Recursive Feature Elimination (RFE) and k-Nearest Neighbor Algorithms in Support Vector Machine for Intrusion Detection System", International Journal of Network and Mobile Technologies, Vol. 1,No. 1, Pp. 8-14, 2010.
[8] Manish Joshi, "Classification, Clustering and Intrusion Detection System", International Journal of Engineering Research and Applications (IJERA), ISSN: 2248-9622, pp.961-964, Vol. 2, Issue 2, Mar-Apr 2012.
[9] Mohammad Sazzadul Hoque, Md. Abdul Mukit and Md. Abu Naser Bikas, "An Implementation Of Intrusion Detection System Using Genetic Algorithm", International Journal of Network Security & Its Applications (IJNSA), Vol.4, No.2, March 2012.
[10] R. Rajesh and J. Siva Prakash, "Extreme Learning Machines - A Review and State-of-the-art", International Journal of Wisdom Based Computing, Vol. 1(1), 2011.
[11] Reema Patel, Amit Thakkar and Amit Ganatra, "A Survey and Comparative Analysis of Data Mining Techniques for Network Intrusion Detection Systems", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-1, March 2012.
[12] Sandip Sonawane, Shailendra Pardeshi and Ganesh Prasad, "A survey on intrusion detection techniques", World Journal of Science and Technology, 2012.
[13] **Wenke Lee and Salvatore J. Stolfo, "**Data Mining Approaches for Intrusion Detection", Pp. 79–94 of the Proceedings, 7th USENIX Security Symposium, 1998.

BIOGRAPHY

V. JAIGANESH is working as an Assistant Professor in the Department of Computer Science, Dr. N.G.P. Arts and Science College, Coimbatore, Tamilnadu, India and Doing Ph.D., in Manonmaniam Sundaranar University, Tirunelveli. Tamilnadu, India. He

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

has done his M.Phil in the area of Data Mining in Periyar University. He has done his post graduate degrees MCA and MBA in Periyar University, Salem. He has presented and published a number of papers in reputed conferences and journals. He has about twelve years of teaching and research experience and his research interests include Data Mining and Networking.

S. MANGAYARKARASI received her MCA degree from Kalasalingam University, krishnankoil, Tamilnadu, India and pursuing M.Phil from Dr. N.G.P. Arts and Science College, Coimbatore, India. Her field of interest is Data Mining.

Dr. P. SUMATHI is working as an Assistant Professor, PG & Research Department of Computer Science, Government Arts College, Coimbatore, Tamilnadu, India. She received her Ph.D., in the area of Grid Computing in Bharathiar University. She has done her M.Phil in the area of Software Engineering in Mother Teresa Women's University and received MCA degree at Kongu Engineering College, Perundurai. She has published a number of papers in reputed journals and conferences. She has about Sixteen years of teaching and research experience. Her research interests include Data Mining, Grid Computing and Software Engineering.