Rule Based Classification Approach towards Detection of Network Intrusions

G.V.Nadiammai[1], M.Hemalatha[2]

Abstract

Network security has become an important issue due to the evolution of internet. It brings people not only together but also provides huge potential threats. Intrusion detection technique is considered as the immense method to deploy networks security behind firewalls. An intrusion is defined as a violation of security policy of the system. Intrusion detection systems are developed to detect those violations. Due to the effective data analysis method, data mining is introduced into IDS. This paper brings an idea of applying data mining algorithms to intrusion detection database. Performance of various rule based classifiers like Part, Ridor, NNge, DTNB, JRip, Conjunctive Rule, One R, Zero R and Decision Table are compared and result shows that classification algorithm performs well in terms of accuracy, specificity and sensitivity. The performance of the model is measured using 10- fold cross validation.

Keywords: — Data Mining, Intrusion Detection, Machine Learning, Rule based Classifier and Function Based Classifier.

1. Introduction

Due to the rapid development of the internet, security has become a major problem. Information is processed all over the world via internet. So the secured network must be free from threat and violation. An attacker would compromise the confidentiality, integrity and availability of network resources. By compromising those resources, they may cause some sorts of violation. One possible protection over vulnerability is utilizing IDS. Intrusion detection system monitors the network packet to find out whether violation has been taken place. If not, it ignores otherwise it and raises an alarm that makes the administrator to handle the situation and also some possible prevention measures. Firewall fails to detect the network traffic that has been done by the specific port or from authorized user port. So IDS serves as a gateway for providing secure network. Tomake IDS effective and to develop the accuracy of intrusion detection, data mining concept was introduced. Data mining has been popularly recognized as an important way to analyze useful information from large volumesof data that are noisy, fuzzy & random [1].IDS includes two types of detection approaches like a) **Misuse/Signature detection-**It accurately identifies the known attacks with less false alarm rate but fails to detect the unknown attacks. Detection efficiency of this method is quite high. b) **Anomaly Detection-** Efficiently detects the new sorts of attacks but with high false alarm rate. Two different types of IDS are Host based and Network based IDS. IDS with data mining were implemented in automated model mining with regard of audit data for intrusion detection. Here ASCII values are changed into connection level information including attributed like duration, protocol, dst_types, and so on.

1. Research Scholar, Karpagam University,Coimbatore. gvnadisri@gmail.com,
2. Head, Dept. of Software Systems & Research, Karpagam University, Coimbatore. hema.bioinf@gmail.com

The rest of this paper is organized as follows, related work is presented in section 2 followed by a description of rule and function based classifier in section 3, dataset description in section 4, result & discussion in section 5and conclusion &future scope in section 6.

II. Related Work

In [2] author compared both ANN and SVM to perform the feature selection in IDS. Performance of SVM with all 41 features, some selected features are analyzed and the result shows that the SVM oriented IDS performs well than ANN based IDS. Decision tree based SVM solves the detection problem over IDS. Here the author [3] proposed an algorithm to formalize the multi class problem in intrusion detection system. Sandhya Peddabachigari et al., [4] present two hybrid approaches for modeling IDS. This model provides high accuracy and minimized computational complexity. Performance of three classifiers & ensemble approach has been considered. Naïve bayes and decision tree algorithms slightly differ in detection rate. In case of normal, dos and u2r attacks the decision tree generates high detection rate and naïve bayes performs well in r2l and probe attacks. B. Abdullah et al., [5] evaluated genetic algorithm in IDS to perform high detection rate and low false alarm rate. J48, Bayes net, and SMO have been compared with the result of genetic algorithm. Namita Shrivastava et al., [6] implemented Naïve bayes, SVM and NB-ACO to analyze the performance criteria. Efficiency of Naïve bayes is less than SVM. But NB-ACO performs well in both DR and FAR. In [7] author proposed the hybrid j48 and random forest combined k means algorithm and observed the recall and precision values in which the hybrid j48 and hybrid random forest performs well with the number of clusters 4,5& 6. In [8] random selection method is used to reduce the size of the data set. Result of SVM, SVM with Rocchio bundling and SVM with DGSOT are compared to identity the performance over IDS. In that, the proposed SVM with DGSOT is seems to be better than other two algorithms. Improved FP growth [9] algorithm has developed for associative of FCM network intrusion detection system with statistical binning. Detection rate of improved FP growth algorithm is good than existing. In [10] Rule based classifiers are compared with proposed ensemble approach of ADABOOST algorithm to build an effective network intrusion detection model. Results of ensemble approach with NNge and decision table shows better results in terms of DR, RR, FNR and F-value. Dewan Md.Farid [11] applied a new algorithm based on boosting & naïve bayes classifier to perform an ensemble approach towards intrusion detection which improves the accuracy percentage with the algorithms like KNN, C4.5,SVM,NN & GA.

III. Model Classification

Different rule based Classifiers are used in this work to evaluate the effectiveness of those classifiers in a classification problem. The Classifiers applied are:

A. DTNB

DTNB classifiers build and utilize decision table and naïve bayes classifier [12]. Evaluation takes place by dividing two disjoint subsets naïve bayes and decision table. A forward selection is applied for modeling both the decision table and naïve bayes algorithms. At every step, algorithm drops an attribute from the entire model.

B. JRIP

It implements [13] a rule learner and incremental pruning to produce an error reduction (RIPPER) proposed by William W. JRip to possess an efficient propositional rule to classify instances.

C. NNge

Brent in 1995 has introduced an non-nested generalized exemplars NNge [14] to perform the generalization on the basis of merging exemplars and forms hyper rectangle to determine conjunctive rules with internal disjoints. Algorithm specifies a generalization, when a new element is added it joins to its nearest neighbor of the same class.

D. One R

One R algorithm develops an individual rule for each attribute e[15]. The most recurrent class of an attribute is needed to generate the rule for the specific attribute. An attribute value may bound to the most recurrent class is said to be rule. It works as follows,

Pseudo-code for One R algorithm is:

For each attribute *A*,

For each value *VA* in the attribute include a rule as follows:

Add up how often each class appears Locate the most frequent class *Cf*

Generate a rule when $A=VA$; class attribute value $=Cf$

End For-Each

Compute the error rate of all rules End for-Each

Select the rule with the smallest error rate

If the number of instances in training have not compromise with their current attribute in the rule that would result an error rate. If they have same error rate then the particular rule is randomly selected.

A. PART

It is a class [16] that performs decision list of a PART. PART uses separate and conquer approach to build a C4.5 decision tree partially for every iteration that specifies the "BEST" among the rule.

B. RIDOR

RIDOR [17] is a type of classifier used to perform default rule. RIDOR stands for **Ripple Down Rule** Learners. An exception has been generated for each error rate with lower or best value depending upon the rule. It possess tree like structure with corresponding exception and the default rule without exception.

Class indulged five different types of inner classes.

The Ridor node class has a default class exception rule which implements at least one node in the Ridor tree.

With the help of REP, the Ridor rule class implements single exception rule. Remaining three classes involves Ridor Rule namely Antd, Numeric Antd and Nominal Antd. Numeric and Nominal Antds implements the corresponding abstract functions. The two sub class Numeric and Nominal Antds has functions in accordance with their antecedents of Numeric and Nominal Attributes respectively.

C. Zero R

It tests [18] the results of other rules. It chooses the common category and compares the result of other learners in order to represent their usefulness in the dominating category.

D. Conjunctive Rule

It is a decision making rule[19] in which the values are assigned for any number of factors and it can be replaced if it does not meet the minimum values of all factors. Conjunctive rules uses AND logical operation to correlate the attributes.

This type of learner chooses the antecedents by calculating the information gain of those antecedent and prunes the produced rule using reduced error pruning method depend on the number of antecedents.

E. Decision table

Decision table [20] majority and decision table local are the two types of classifiers involved in the decision table. Decision table implements majority of the training set. Decision table local returns the decision table entry with many attributes if the particular cell is empty.

IV. Dataset Description

KDD CUP 99 Dataset [7] is developed based on DARPA 98 dataset in a MIT Lincoln Laboratory. Protocol such as TCP, UDP and ICMP has been used in this dataset to evaluate the anomaly detection methods.The dataset contains 24 different training attacks and 14 types in the test data.Here 7500 records are selected for the study out of 3, 11,029 Corrected dataset. Attacks such as Probe, U2R, and R2L are found to be less.80% of data belongs to DoS attack respectively.

The attacks fall in one of the following four categories:

- **DoS Attack**

It is kind of attack in which the attacker makes traffic busy over the network in order to restrict the authorized users access to the systems.

- **User to Root Attack**

Here, the attacker access to a normal user account on the system by compromising it through sniffing password and gain access to the remote system.

- **Remote to Local Attack**

In R2L, an attacker sends packets to an appropriate machine over a network since who does not have any rights to access a system and make some violations.

- **Probe Attack**

This kind of attack is carried out to get the systems and network information to make some attack in future.

V. Performance Evaluation

Performance of rule and function based classifiers are evaluated using KDD Cup 99 dataset based upon following criteria,

5.1 10Fold - Cross Validation

Cross-validation is also known as rotation estimation. It is a model to access the results of a statisticalanalysis to generalizean independent data set. It is mainly used in goal setting based on prediction. 10-Fold cross-validation is normally used in K-fold crossvalidationthrough which the folds are selected to estimate the meanresponse value is roughly equal in all the folds.

5.2 Comparison Criteria

The performance of classifiers involves Accuracy, Sensitivity, Specificity, MAE and RMSE. The accuracy, sensitivity and specificity were estimated by True Positivemeasure, False Positive measure, False Negative measure and True Negativemeasure [1].

Accuracy is the possibility that the algorithms can correctly predict positive and negative instances.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Sensitivity is the possibility that the algorithms can correctly predict positive instances.

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

Specificity is the possibility that the algorithms can correctly predict negative instances.

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

Mean absolute error, is the average of the difference between predicted and actual value in all test cases; it is the average prediction error. Whereoutput is a, possible output is c.

$$MAE = (|a1 - c1| + |a2 - c2| + \ldots + |an - cn|) / n \quad (4)$$

Root Mean Square Error is frequently used values to predict a model or estimation technique through which the values are observed from the being modeled or estimated. Square root of MAE are considered as RMSE,

$$RMSE = \sqrt{\frac{(a_1 - c_1)^2 + (a_2 - c_2)^2 + \ldots + (a_n - c_n)^2}{n}} \quad (5)$$

Mean squared error is used for numeric prediction. This value is computed by analyzing the average of the squared differences between computed value and its corresponding correct value.

The accuracy of mean absolute and root mean squared error has been calculated for each machine learning algorithm.

VI. Result and Discussion

This work is performed using Machine learning tool to predict the effectiveness of all rule based and some function based classifiers. The performance of the various algorithms measured in classification accuracy, Sensitivity, Specificity, RMSE and MAE values are measured.Table1 Comparison among Rule based and Function based classifiers in terms of MAE, RMSE and Time complexity. Figure1 specifies the corresponding chart for the result obtained in table 1.

TABLE 1: COMPARISON BASED ON MAE, RMSE AND TIME

Classifiers	MAE	RMSE	Time taken (in Sec.)
DTNB	0.0201	0.0984	72.08
One R	0.0527	0.2295	0.11
JRIP	0.0154	0.0901	7.73
Part	0.0143	0.0901	0.78
Ridor	0.0096	0.0981	24.05
Zero R	0.1638	0.2861	0.02
Conjunctive Rule	0.1207	0.2458	0.16
Decision Table	0.0428	0.1313	4.42
NNge	0.1535	0.0945	1.17

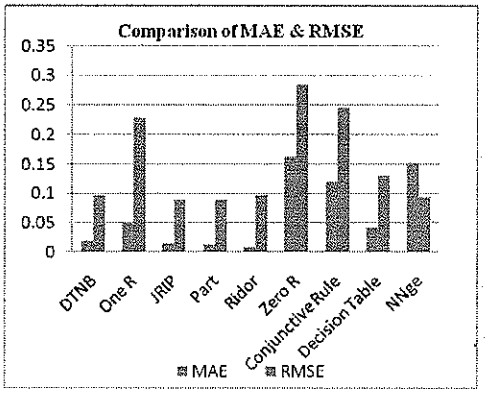

FIGURE 1: GRAPHICAL REPRESENTATION OF MAE & RMSE VALUES

Figure 2, illustrates the build time of all rule based classifiers and some function based classifiers. Zero R, the probabilistic classifier tends to learn more rapidly for the given dataset. But NNge takes less time and also provides better accuracy percentage in terms of all rule based classifiers.

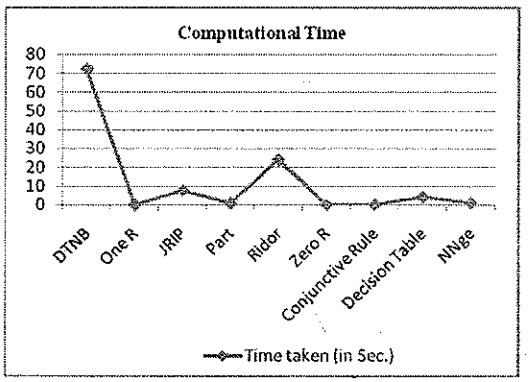

FIGURE 2: COMPUTATIONAL TIME TAKEN FOR RULE AND FUNCTION

Table 2 represents the accuracy, sensitivity and specificity values for all 12 classifiers. Based on values evaluated, accuracy of DTNB is 95.21%, the accuracy of One R is 76.29%, the accuracy of JRIP is 94.36%, the accuracy of Part is 95.67%, the accuracy of Ridor is 94.92%, the accuracy of Zero R is 42.14%, the accuracy of Decision Table is 91.14% and the accuracy of NNge is 96.83%. Finally, NNge Classifier took highest accuracy percentage compared to 9 rule based classifiers.

TABLE 2: COMPARISON BASED ON ACCURACY, SENSITIVITY & SPECIFICITY

Classifiers	Accuracy	Sensitivity	Specificity
DTNB	95.21	91.52	95.08
One R	76.29	73.02	85.21
JRIP	94.36	92.45	95.56
Part	95.67	94.03	95.55
Ridor	94.92	90.18	93.75
Zero R	42.14	60.15	78.59
Decision Table	91.4	80.34	91.15
Conjunctive Rule	60.35	68.45	75.25
NNge	**96.83**	**94.62**	**96.73**

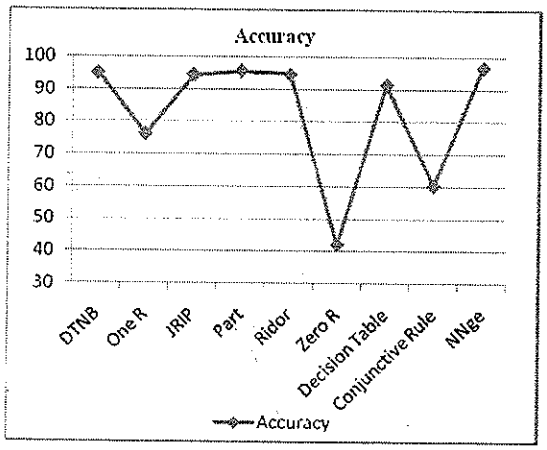

FIGURE 3: COMPARISON BASED ON CORRECTLY CLASSIFIED INSTANCES

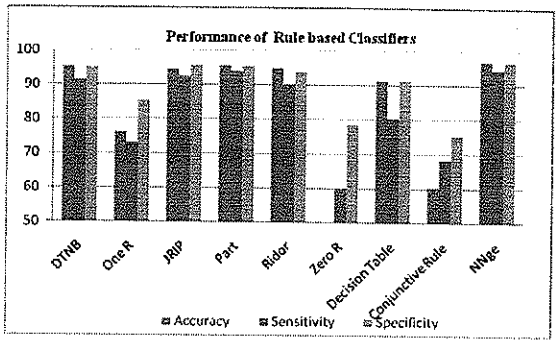

FIGURE 4: GRAPHICAL REPRESENTATION OF ACCURACY, SPECIFICITY & SENSITIVITY

The performance of learning method is highly dependent on the nature of the training data. The result indicates that, NNge classifier has got the first position in ranking followed by DTNB, One R, JRIP, Part and other classifiers as shown in figure 3 &4.

VII. CONCLUSION & FUTURE WORK

Due to the rapid growth of the network, new attack tends to happen. Intelligent IDS not only detect new kind of attacks but also has a low impact ratio. Data mining has been popularly recognized as an important means to mine useful information from large volumes of data which is noisy, fuzzy, and random. Machine learning algorithms can improve the efficiency of IDS. In this paper, 10-fold cross validation model has been used to evaluate the performance of rule and function based classifier over intrusion detection system. Among these classifiers, a NNge (Non-Nested Generalized Exemplars) outperforms from all the other algorithms used. Next comes and so on.

In future, a hybrid intrusion detection system can be developed based on data mining algorithms combine with optimization techniques which would be fast and robust in identifying the huge variety of new and unusual attacks.

VII. REFERENCES

[1]. Jiawei Han and Kamber," Data Mining: Concepts and Techniques, 2nd Edition", Morgan Kaufman Publishers, Elsevier Inc, 2006.

[2]. Srinivas Mukkamala1 & Andrew H. Sung, "Feature Selection for Intrusion Detection using Neural Networks and Support Vector Machines", 2002.

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

[4]. SandhyaPeddabachigari, AjithAbraham,"Intrusion Detection Systems Using Decision Trees and Support Vector Machines", Journal of Network & Application, 2005.

[5]. Abdullah B., Abd-alghafar I., Gouda I. Salama, "Performance Evaluation of a Genetic Algorithm Based Approach to Network Intrusion Detection System", Aerospace Sciences & Aviation Technology, 2009.

[6]. Namitashrivastava,VineetRichariya, "Ant Colony Optimization with Classification Algorithms used for Intrusion Detection, International Journal of Computational Engineering & Management, Vol.15, No.1, Jan 2012.

[7]. KusamKumariBharti, SanyamShukla, Sweta Jain, "Intrusion Detection using Clustering", International Journal of Computer & Communication Technology, pp 158 to 165, Vol.1, No.2, 2010.

[8]. Latifur Khan, MamounAwad, BhavaniThuraisingham.,"A new intrusion detection system using support vector machines and hierarchical clustering", The VLDB Journal, DOI 10.1007/s00778-006-0002, 2007.

[9]. Desheng Fu, Shu Zhou, Ping Guo, "The Design and Implementation of a Distributed Network Intrusion Detection System Based on Data Mining", IEEE, pp. 446 – 450, DOI: 10. 1109/ WCSE.2009.225, 2009.

[10]. Mrutyunjaya Panda,ManasRanjanPatra, "Ensembling Rule Based Classifiers For Detecting Network Intrusions", International Conference on Advances in Recent Technologies in Communication and Computing,IEEE,pp.19-22, DOI:10.1109/ARTCom.2009.121 ,2009.

[11]. Dewan Md. Farid, Mohammad ZahidurRahman , ChowdhuryMofizurRahman," Adaptive Intrusion Detection based on Boosting and Naïve Bayesian Classifier", International Journal of Computer Applications, pp.12-19, Vol. 24 No.3, June 2011.

[12]. Shengli Sheng, Charles X. Ling. (2005). "Hybrid Cost- sensitive Decision Tree, Knowledge Discovery in Databases", PKDD 2005, Proceedings of 9th European Conference on Principles and Practice of Knowledge Discovery in Databases. Lecture Notes in Computer Science 3721 Springer 2005, ISBN 3-540-29244-6.

[13]. http://weka.sourceforge.net/doc/weka/classifiers/rules/JRip.html

[14]. http://weka.sourceforge.net/doc/weka/classifiers/rules/NNge.html

[15]. Gaya Buddhinath& Damien Derry. "A Simple Enhancement to One Rule Classification".

[16]. C. Lakshmi Devasena, T. Sumathi, V.V. Gomathi and M. Hemalatha, "Effectiveness Evaluation of Rule Based Classifiers for the Classification of Iris Data Set", Bonfring International Journal of Man Machine Interface, Vol. 1, Special Issue, December 2011.

[17]. http://weka.sourceforge.net/doc/weka/classifiers/rules/Ridor.html

[18]. http://chem-eng.utoronto.ca/datamining/dmc/zeror.htm

[19]. http://weka.sourceforge.net/doc/weka/classifiers/rules/ConjunctiveRule.html

[20]. Ron Kohavi," The power of decision Tables", 8th European Conference on Machine learning, pp.174-189, 1995.

[21]. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

[22] M. Revathi, T Ramesh,"Network Intrusion Detection System Using Reduced Dimensionality", Indian Journal of Computer Science and Engineering, pp.61-67, Vol. 2 No. 1, 2011.

Authors Profile

Dr. M. Hemalatha completed M.Sc., M.C.A., M. Phil., Ph.D (Ph.D, Mother Terasa women's University, Kodaikanal). She is Professor & Head and guiding Ph.D Scholars in Department of Computer Science at Karpagam University, Coimbatore. Twelve years of experience in teaching and published more than hundred papers in International Journals and also presented more than eighty papers in various national and international conferences. She received best researcher award in the year 2011 from Karpagam University. Her research areas include Data Mining, Image Processing, Computer Networks, Cloud Computing, Software Engineering, Bioinformatics and Neural Network. She is a reviewer in several National and International Journals.

G.V.Nadiammai completed M.C.A., in the year 2006 from Bharathiar University, B.C.A in the year 2003 from Bharathiar University and currently pursuing Ph.D in computer science at Karpagam University under the guidance of Dr.M.Hemalatha, Professor and Head, Dept. of Software System, Karpagam University, Coimbatore. Published four papers in International Journals and presented three papers in international conference. Area of research is Data Mining, Network Security and Knowledge Discovery.