Comparison of two Multi-Classification Approaches for Detecting Network Attacks.

A.Kumaravel[#1] , M.Niraisha[*2]
#1 Professor and dean, Department of Computer Science and Engineering, Bharath University
Selaiyur, Chennai-600073, India
drkumaravel@gmail.com

*2 PG student, Department of Computer Science and Engineering, Bharath University
Selaiyur, Chennai-600073, India
naga_sharaini@yahoo.co.in

***Abstract*-Extensive growth of the internet and increasingavailability of tools and tricks for intruding and attackingnetworks, have prompted intrusion detection to become acritical component of network administration. It is animportant attribute of defensive measure protecting computersystem and network traffic from abuses.In this paper we demonstrated thathigh attack detection accuracy can be achieved by usingclassificationtechniques and high performance is attained by the multi-classificationapproach. To test the results, we have used NSL-KDD datasets.We compared our proposed system with previous method which is lightly classified and tried to find which is more accurateand appropriate to detect intrusion.**

***Keywords*-Network Security, DataMining, Intrusion detection system,Feature Selection, Multi-Classification.**

1.INTRODUCTION

Intrusion detection process includes identifying a set of malicious actions that compromise theintegrity, confidentiality, and availability of information resources. Traditional methods for intrusion detection are based on extensive knowledge of signatures of known attack types. Monitored events are matched against the signatures to detect intrusions. Theseattacks are normally detected by tools known as intrusiondetection system[1].Detect intrusions by comparing the feature values to a set of attack signatures provided by human experts.

Data mining based intrusion detection techniques generally fall into one of two categories; anomaly detection and misuse detection. In misuse detection, each instance in a data set is labeled as 'normal' or 'intrusion' and a learning algorithm is trained over the labeled data.

In the proposed system, we have designed anomaly based intrusion detection usingmulti-classification. The input to theproposed system is KDDNSL dataset, which is divided into two subsets such as, training dataset andtesting dataset. The training dataset is classified into five subsets[2]so that, four types of attacks such as DoS(Denial of Service), R2L (Remote to Local), U2R (User to Root), Probe and normal data.

Classification is perhaps the most familiar and most popular data mining technique.Prediction can be thought of as classifying an attribute value into one of a set of possible classes.

The subject is introduced briefly as following,in section 2, the proposed method,insection 3, theexperimental results and analysis, in section 4, performance comparison with previouswork, We present the conclusion in section 5.

2.PROPOSED METHOD

The simulated attacks were classified, according to the goals of the attacker.Each attack type fallsinto one of the following four main categoriesDenial-of-Service (DoS) attacks have the goal oflimiting or denying services provided to the user,computer or network(e.g.teardrop).Probing or Surveillance attacks have the goal ofgaining knowledge of the existence orconfiguration of a computer system or network.Port Scans&sweeping of a given IP-address rangetypically fall in this category. (e.g. portsweep).User-to-Root (U2R) attacks have the goal ofgaining root or super-user access on a particularcomputer or system on which the attackerpreviously had user level access. These areattempts by a non-privileged user to gainadministrative privileges (e.g. Perl).Remote-to-Local(R2L) attack is an attack inwhich a user sends packets to a machine over theinternet, which the user does not have access to inorder to expose the machine vulnerabilities andexploit privilegeswhich a local user would haveon the computer (e.g. xclock).

2.1 Multi-Classification Intrusion Detection System

Our system is a modular network-based intrusion detection system that analyzes TCP dump data using data mining techniques to classify the network records to not only normal and attack but also identify attack type.The proposed system consists of two stages. First phase is for attack detection and the second phase is for attack classification. The data is input in the first phasewhich identifies if this record is a normal record or attack.

We train and test each layer to detect only a particular type ofattack. For example, first layer of our proposed model is trained todetect U2R[3] attacks only. When such a system is deployed online,other attacks such as can either be seen as normal or attack ,If R2Lattacks are detected as normal, then it must to be detected asattack at other layers[4] in the system. However if the R2L attacksare detected as U2R, it must be considered as an advantage sincethe attack is detected at an early stage. Hence, for four attackclasses, we have four independent multi-classes, which are trainedseparately with specific features to detect attacks belonging to thatparticular class. We represent the layered model in Figure 1

Our system has the capability of classifying network intruders into two stages. The first stage classifies the network records to either normal or attack. The second stage consists of four sequential Layers which can identify four categories/classes and their attack type. The data is input in the first stage which identifies if this record is a normal record or attack. If the record is identified as an attack then the module would raise a flag to the administrator that the coming record is an attack then the module inputs this record to the second stage which consists of foursequential Layers[5], one for each class type (R2L,U2R,Dos,Probe)[4]. Each Layer is responsible for identifying the attack type of coming record according to its class type.

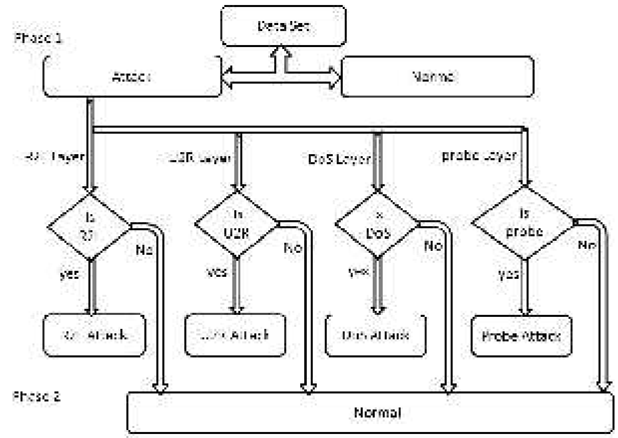


Fig.1 Proposed Layered-Model Approach System

Else the attack passes through the next layer.

Each layer act as a filters that classifies the attacks of each layer category which eliminate the need of further processing at subsequent layers but we took in consideration the propagation of errors as to simulate the real system and results be more accurate and real . We implement the Layered Approach to improve overall system performance as our layered intrusion detection model using JRipRule achieves high efficiency and improves the detection and classification with high rate of accuracy.

In previous work we consider the particular attack and normal data in that particular layer and avoided the rest of attacks. But in the proposed system we consider the rest of the attacks as normal which method is heavily classified. We compare the performance of our proposed approach with previous work in this field which is lightly classified shown in Fig 2.

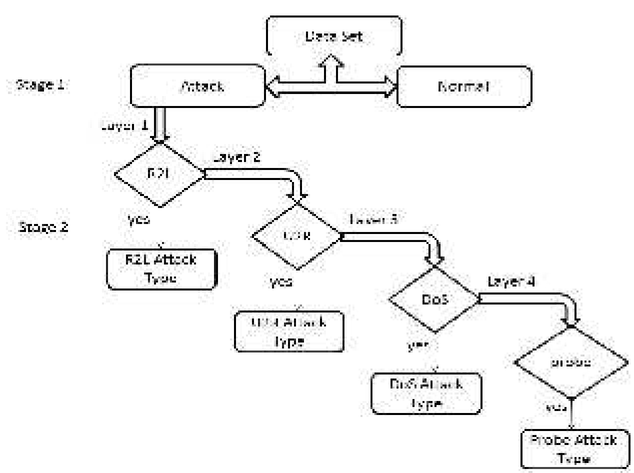


Fig.2 Previous Layered-ModelApproach System

3.EXPERIMENTAL ANALYSIS AND RESULTS

The data in the experiment is acquired from the NSL-KDD[6] dataset which consists of selected records of the complete KDD data set.Apply the dataset in wekatool[7] to findSelected Feature and Classification results.Experimental results havedemonstrated that our Multi-Classifier model ismuch more efficient in the detection of networkintrusions, compared to the other techniques

3.1 Dataset

3.1.1 DatasetDescription

Network based IDSs of nsl.cs.The two weeks of test data yielded around six thousand connection records. Each connection is labeled as either normal, or as an attack, with exactly one specific attack type,other attacks can be seen as normal. Actually 42 attributes are in dataset.

Attacks fall into four categories:

- DOS: denial-of-service, e.g. Neptune, back,ect.
- R2L: unauthorized access from a remote machine, e.g. warezclient,guessing password
- U2R: unauthorized access to local superuser (root) privileges, e.g., loadmodule, various ``buffer overflow" attacks;
- Probe: surveillance and other probing, e.g., IPsweep,nmap.

3.2Performance evaluation

During the analysis of intrusion detection we observe two mainchallenging issues in this system. First, the number of intrusionson the network is typically a very small fraction of the totaltraffic. Therefore the essential step is to reduce attributes of the various Layers. Second, the attacks are classified in their impact , it becomes necessary to treatthem differently.

To improve the minority attack detection rate, while maintaining areasonable overall detection rate. We proposed a layered model[8] withvarious classifiers(BayesNet, NavieBayes,DecisionStump[9] and rulesJrip) on values. In layered model we definefour layers that correspond to the four attack groups i.e. DoSlayerfor detecting DoS attacks, Probe layer for detecting Probe attacks,R2L layer for detecting R2L attacks and U2R layer for U2R attacks.

3.2.1 Feature Reduction

In this experiment, Weka tool is used for featurereduction. wekatool Evaluator: weka.attributeSelection.CfsSubsetEval with Best first approach is appliedon the training dataset to obtain the important features for theclassification process.Each subset is analysed using thecorrelation analysis for identifying the important features fora specific attack. This analysis result gives a set of particular featuresfor each subset which is sufficient to group the attack andnormal records. Thereduced features are considered as relevantfeatures for each attack in each layer.

Table 1 shows the weight calculation of the reduced attributes depends on its impact.

TABLE 1
SELECTED ATTRIBUTES

Layer. No	Layer	No. of attributes selected	Selected attributes
1	R2L Layer	9	1,5,10,11,22,27,31,33,36
2	U2R Layer	8	1,6,13,14,16,17,23,34
3	Dos Layer	7	5,6,10,19,31,37,41
4	Probe Layer	5	5,6,34,36,37

3.2.2 Classification with Phases

3.2.2.1 First Phase Results

Phase 1 duty is to classify whether coming record is normal or attack. It is observed that JRip has a significant detection rate for known and unknown attacks compared to BN,DS and NB.The results of Phase 1 are shown in table 2.

TABLE 2
FIRST PHASE CLASSIFICATION

Method	**Correctly classified**	**Incorrectly classified**
bayes.BayesNet	98.43%	1.57%
bayes.NaiveBayes	91.46%	8.54%
rules.Jrip	99.90%	0.09%
trees.DecisionStump	92.79%	7.21%

Table 2: First Phase classification

3.2.2.2 Second Phase Results

Records classified as attacks by the first Phase are introduced to second Phase which is responsible for classifying coming attack to one of the four classes (DOS, Probe, U2R and R2L) and identifying its attack type. Phase 2 consists of four sequential layers; a layer for each class which identify the class of each coming attack.

DoS Layer:

The results of Phase 2 DoS Layer are shown in table 3.

TABLE 3
CLASSIFICATION OF DOS LAYER

Method	Correctly classified	Incorrectly classified
bayes.BayesNet	99.68%	0.32%
bayes.NaiveBayes	93.66%	6.34%
rules.Jrip	99.98%	0.02%
trees.DecisionStump	93.47%	6.53%

Table 3: classification of DoS Layer

Probe Layer:

The results of Phase 2 Probe Layer are shown in table 4.

TABLE 4
CLASSIFICATION OF PROBE LAYER

Method	Correctly classified	Incorrectly classified
bayes.BayesNet	99.06%	0.93%
bayes.NaiveBayes	96.39%	3.61%
rules.Jrip	99.97%	0.03%
trees.DecisionStump	98.99%	1.01%

R2L Layer:

The results of Phase 2 R2L Layer are shown in table 5.

TABLE 5
CLASSIFICATION OF R2L LAYER

Method	Correctly classified	Incorrectly classified
bayes.BayesNet	98.44%	1.56%
bayes.NaiveBayes	97.08%	2.92%
rules.Jrip	99.90%	0.10%
trees.DecisionStump	99.69%	0.31%

U2R Layer:

The results of Phase 2 U2R Layer are shown in table 6.

TABLE 6
CLASSIFICATION OF U2R LAYER

Method	Correctly classified	Incorrectly classified
bayes.BayesNet	99.79%	0.21%
bayes.NaiveBayes	96.90%	3.10%
rules.Jrip	99.98%	0.02%
trees.DecisionStump	99.87%	0.13%

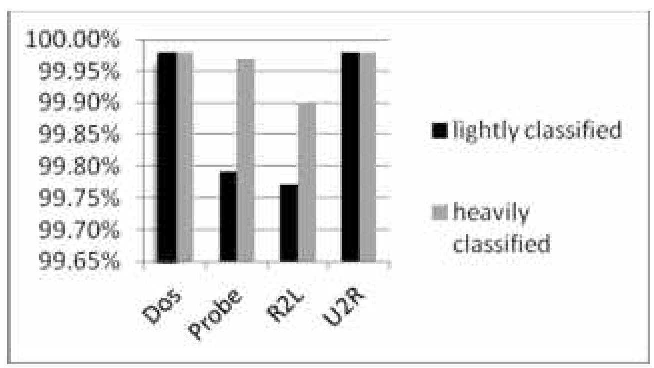
We compare this non-layered approach with the layered approach. We observe that the layered approach with featureselection is more efficient and more accurate in detecting attacks.

4. PERFORMANCE COMPARISONWITH PREVIOUS WORK

In this section, we compare the performance of our approachwithPrevious work[10] in this field which is lightly classified. This information is shown inTable 7.

TABLE 7
PERFORMANCE COMPARISON WITH LIGHTLY CLASSIFIED RESULT.

Layers	Lightly Classified	Heavily Classified
DoS	99.98%	98.2%
Probe	99.79%	99.97%
R2L	99.77%	99.90%
U2R	99.98%	99.98%



Graph.1: Performance Comparison With Lightly Classifiedresult.

According to the above table, proposed system has goodperformance that is competitive with previous work basedon classification rate which is shown in graph.1

5. CONCLUSIONS

A multi-classification intrusion detection system is developed to achieve high efficiency and improve detection and classification rate accuracy. The proposed system consists of two phases, first phase is defined between attacks and normalwherethe data is input in to the first phase which identifies if this record is a normal record or attack, the second phase is for attack classification, the identified attacks are layered. The advantage of the proposed multi-classification system is improve scalability as when new attacks of specific class are added, there is no need to train all the layers only the layer which is affected by the new attack.

Experimental results indicate that the proposed layered modelwithJRip classifier can result in better prediction ofProbe and R2L classes without hurting the prediction performance of theother classes.

ACKNOWLEDGEMENTS

The authors would like to thank the management of Bharath University for the support and encouragement for this research work.

REFERENCES

[1]Neelam Sharma, Saurabh Mukherjee “Layered Approach for Intrusion Detection Using NaïveBayes Classifier” ICACCI’12, August 3-5, 2012, Chennai, T Nadu, India.

[2] Ankita Gaur, VineetRichariya“A Layered Approach for Intrusion DetectionUsing Meta-modeling with ClassificationTechniques” International Journal of Computer Technology and Electronics Engineering (IJCTEE) Volume 1 , Issue 2

[3] P. GiftyJeya M. Ravichandran C. S. Ravichandran“Efficient Classifier for R2L and U2R Attacks” International Journal of Computer Applications (0975 – 8887)Volume 45– No.21, May 2012.

[4]Kapil Kumar Gupta, BaikunthNathandRamamohanarookotagiri, “A layered approach using conditional random fields for intrusion detection”, IEEETranc.on Dependence and secure computing, Vol.7, 2010

[5] OludeleAwodele, Sunday Idowu, OmotolaAnjorin, andVincent J. Joshua “A Multi-Layered Approach to the Designof Intelligent Intrusion Detection and Prevention System(IIDPS)” Issues in Informing Science and InformationTechnology Volume 6, 2009.

[6] “NSL-KDD dataset for network –based intrusion detectionsystems” available on http://iscx.info/NSL-KDD/

[7] Weka: http://www.cs.waikato.ac.nz/~ml/weka/

[8] SaharSelim, Mohamed Hashem and Taymoor M. Nazmy, "Hybrid Multi-level Intrusion Detection System , ” International Journal of Computer Science and Information Security (IJCSIS), pp. 23-29, Vol. 9, No. 5, May 2011

[9] HebaEzzat Ibrahim, Sherif M. Badr and Mohamed A. Shaheen," Phases vs. Levels using Decision Trees for Intrusion Detection Systems ," International Journal of Computer Science and Information Security, Vol. 10, No. 8, 2012

[10]A.Kumaravel, M.Niraisha, “Multi-Classification Approach for Detecting Network Attacks,”InternationalConferenceon Information and communication Technologies (ICT 2013),ICT545,4,2013.