Multi-Classification Approach for Detecting Network Attacks.

A.Kumaravel , Professor and Dean, M.Niraisha,PG Student, Department of Computer Science and Engineering.
Bharath University, Selaiyur, Chennai-600073, India
drkumaravel@gmail.com , naga_sharaini@yahoo.co.in

Abstract

Intrusion Detection System (IDS) has increasingly become a crucial issue for computer and network systems. Intrusion poses a serious security risk in a network environment. The ever growing new intrusion types pose a serious problem for their detection. The acceptability and usability of Intrusion Detection Systems get seriously affected with the data in network traffic. A large number of false alarms mean a lot in terms of the acceptability of Intrusion Detection Systems[1].In this paper we consider the dataset with multi classes and propose the classification for each type of attacks in a separate layer. In this work, a multi-classification approach for detecting network attacks is designed and developed to achieve high efficiency and improve the detection and classification rate accuracy [6].

Keywords:**Network Security, DataMining, Intrusion detection system,Classification,Layered approach.**

1. INTRODUCTION

An Intrusion Detection System (IDS) provides an additional layer of security to network's perimeter defense, which is usually, implemented using a firewall. The goal of IDS is to collect information from a variety of systems and network sources, and then analyze the information for signs of intrusion and misuse. IDSs are implemented in hardware, software, or a combination of both.

On theother hand, our computers are under attacks and vulnerable to many threats. There is anincreasing availability of tools and tricks for attacking and intruding networks. An intrusion canbe defined as any set of actions that threaten the security requirements (e.g., integrity,confidentiality, availability) of computer/network resource (e.g., user accounts, file systems,and system kernels). Intruders have promoted themselves and invented innovative toolsthat support various types of network attacks. Hence, effective methods for intrusion detection (ID) have become an insisting need to protect our computers from intruders. In general, thereare two types of Intrusion Detection Systems (IDS); misuse detection systems and anomalydetection systems.

The problem of designing IDSs to work effectively and yield higher accuracies for minor attacks even in the mix of data has been receiving serious attention in recent times. The imbalance in data degrades the prediction accuracy. In most of the available literature this is overcome by resampling the training distribution. The resampling is done either by oversampling of the minority class or by under sampling of the majority class.

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, Insection 2, formulates the problem. Insection 3, theexperimental results and analysis. We present the conclusion in section 4.

2. PROBLEM STATEMENT

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 stage is for attack detection and the second stage is for attack classification in multi-layer [1]. The data is input in the first Stage.

2.1The Proposed Layered-Model Intrusion Detection System

The main characteristics of our system:

First, our system has the capability of classifying network intruders into two stages[2]. 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 four sequential Layers, one for each class type (R2L,U2R,Dos,Probe)[3]. Each Layer is responsible for identifying the attack type of coming record according to its class type.

978-1-4673-5758-6/13/$31.00 © 2013 IEEE

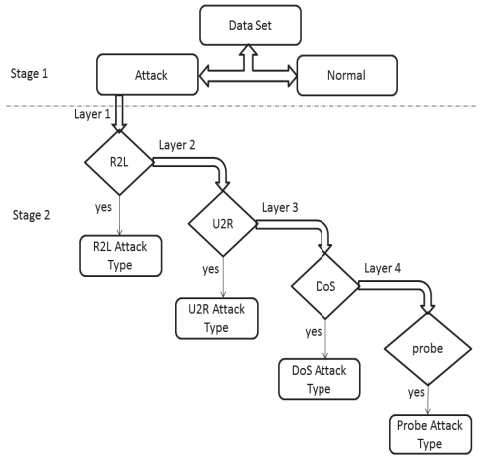


Fig.1 Layered-Model Approach System

Else the attack passes through the next layer.

Second, it takes less training time and even decrease in each layer where we use the whole dataset for training stage 1 then instage 2 we use only the attacks for training excluding the normal records. Then 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 .

In many situations, there is a trade-off between efficiency and accuracy of the system and there can be various avenues to improve system performance. 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.

3. EXPERIMENTAL ANALYSIS AND RESULTS

In this section, first we collect the dataset.Apply the dataset in Wekatool [5] to find Classification results.The datasets for these experiments are from NSL-KDD [6] network based IDSs. Although, the proposed data set still suffers from some of the problems and may not be a perfect representative of existing real networks, because of the lack of public data sets for network-based IDSs, but still it can be applied as an effective benchmark data set to help researchers compare different intrusion detection methods.

3.1 Dataset

3.1.1 DatasetDescription

Network based IDSs of nsl.cs.The raw data was about four gigabytes of compressed binary TCP dump data from seven weeks of network traffic. This was processed into about five million connection records. Similarly, the two weeks of test data yielded around two million connection records. A connection is a sequence of TCP packets starting and ending at some well-defined times, between which data flows to and from a source IP address to a target IP address under some well-defined protocol. Each connection is labeled as either normal, or as an attack, with exactly one specific attack type. Each connection record consists of about 100 bytes. Actually 42 attributes are in dataset.7000 instances that training dataset.

Attacks fall into four main categories:

- DOS: denial-of-service, e.g. syn flood;
- R2L: unauthorized access from a remote machine, e.g. guessing password;
- U2R: unauthorized access to local superuser (root) privileges, e.g., various ``buffer overflow'' attacks;
- Probing: surveillance and other probing, e.g., port scanning.

3.1.2 AttributeDescription

Duration	length (number of seconds) of the connection
protocol type	type of the protocol, e.g. tcp, udp, etc.
Service	network service on the destination, e.g., http, telnet,etc.
src_bytes	number of data bytes from source to destination
dst_bytes	number of data bytes from destination to source
Land	1 if connection is from/to the same host/port; 0 otherwise
wrong_fragment	number of ``wrong'' fragments
Urgent	number of urgent packets
srv_count	number of connections to the same service as the current connection in the past two seconds
srv_serror_rate	% of connections that have ``SYN'' errors
srv_rerror_rate	% of connections that have ``REJ'' errors
srv_diff_host_rate	% of connections to different hosts

3.2 Performance 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 select attributes of

the various Layers. Second, the attacks are classified in their impact and hence, it becomes necessary to treat them differently.

To improve the minority attack detection rate, while maintaining the overall detection rate. We proposed a layered modelwith various classifiers (BayesNet, NavieBayes,DecisionStump 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 for U2R attacks.

3.2.1 Attribute Selection

Attribute selection usingweka tool. Table 1 shows the weight calculation of the selected 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 Stages

3.2.2.1 First Stage Results

Stage 1 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 Stage 1 are shown in table 2.

TABLE 2:
First stage classification

Method	correctly classified	IN correctly classified
bayes.BayesNet	98.73%	1.26%
bayes.NaiveBayes	92.41%	7.58%
rules.Jrip	99.89%	0.10%
trees.DecisionStump	94.65%	5.34%

3.2.2.2 Second Stage Results

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

DoS Layer:

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

TABLE 3:
Classification of DoS Layer

Method	correctly classified	IN correctly classified
bayes.BayesNet	99.79%	0.21%
bayes.NaiveBayes	96.03%	3.97%
rules.Jrip	99.98%	0.02%
trees.DecisionStump	94.01%	5.99%

Probe Layer:

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

TABLE 4:
Classification of Probe Layer

Method	correctly classified	IN correctly classified
bayes.BayesNet	99.17%	0.83%
bayes.NaiveBayes	97.93%	2.07%
rules.Jrip	99.79%	0.21%
trees.DecisionStump	98.34%	1.66%

R2L Layer:

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

TABLE 5:
Classification of R2L Layer

Method	correctly classified	IN correctly classified
bayes.BayesNet	95.22%	4.78%
bayes.NaiveBayes	94.99%	5.01%
rules.Jrip	99.77%	0.23%
trees.DecisionStump	95.67%	4.33%

U2R Layer:

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

TABLE 6:
Classification of U2R Layer

Method	correctly classified	IN correctly classified
bayes.BayesNet	99.54%	0.46%
bayes.NaiveBayes	97.68%	2.32%
rules.Jrip	99.98%	0.02%
trees.DecisionStump	99.77%	0.23%

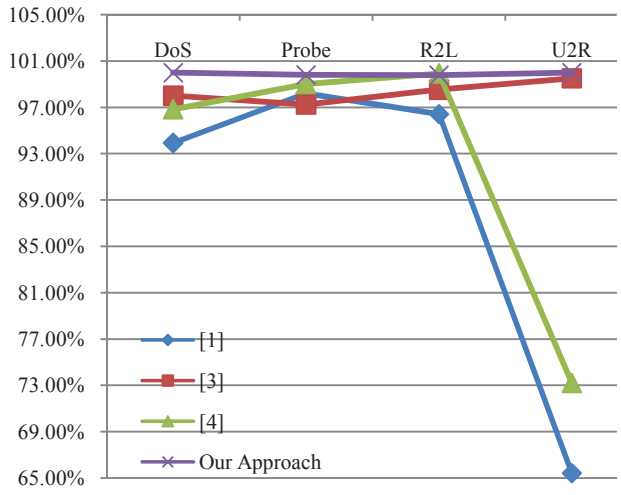
We compare this non-layered approach with the layered approach. We observe that the layered approach with featureselection is more efficient and more accurate in detecting attacks.

4. PERFORMANCE COMPARISON WITH EXISTING APPROACHES

In this section, we compare the performance of our approach with other works in this field. This information is shown in Table 7.

TABLE 7:
Performance comparison with existing system.

Existing Paper	DoS	Probe	R2L	U2R
[2]	93.9%	98.2%	96.4%	65.4%
[3]	96.8%	99.0%	99.9%	73.2%
[4]	97.99%	97.23%	98.52%	99.49%
Our Approach	99.98%	99.79%	99.77%	99.98%



GRAPH.1 Performance comparison with existing system.

According to the above table, proposed system has good performance that is competitive with other approaches based on classification accuracy which is shown in graph.1

5. CONCLUSIONS

A multi-Layer intrusion detection system has been developed to achieve high efficiency and improve detection and classification accuracy. The proposed system consists of two stages. First stage is for attack detection and the second stage is for attack classification. The data is input in the first Stage which identifies if this record is a normal record or attack. Experimental results indicate that the proposed layered model with JRip classifier can result in better prediction of minority classes without hurting the prediction performance of the majority class. It gives result for DoS Layer- 99.98%, Probe Layer- 99.79%, R2L- 99.77%, U2R- 99.98%.

ACKNOWLEDGEMENTS

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

REFERENCES

[1] *OludeleAwodele, Sunday Idowu, OmotolaAnjorin, andVincent J. Joshua, "A Multi-Layered Approach to the Design of Intelligent Intrusion Detection and Prevention System(IIDPS)," Issues in Informing Science and InformationTechnology Volume 6, 2009.*

[2] *Neelam Sharma, Saurabh Mukherjee, "Layered Approach for Intrusion Detection Using NaïveBayes Classifier," ICACCI'12, August 3-5, 2012, Chennai, T Nadu, India.*

[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] *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.*

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

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