Efficient Classifier for R2L and U2R Attacks

Efficient Classifier for R2L and U2R Attacks

P. Gifty Jeya
Department of Computer Science and Engineering
Sri Venkateswara College of Engineering
Chennai

M. Ravichandran
Department of Computer Science and Engineering
Sri Venkateswara College of Engineering
Chennai

C. S. Ravichandran
Department of Electrical and Electronics Engineering
Sri Ramakrishna Engineering College, Coimbatore

ABSTRACT

Intrusion Detection System (IDS) is an effective security tool that helps to prevent unauthorized access to network resources by analysing the network traffic and classifying the records as either normal or anomalous. In this paper, a new classification method using Fisher Linear Discriminant Analysis (FLDA) is proposed. The features of KDD Cup '99 attack dataset are reduced for each class of attacks using correlation based feature selection method. Then with the reduced feature set, discriminant analysis is done for the classification of records. Comparison with other approaches reveals that our approach achieves good classification rate for R2L (Remote-to-Local) and U2R (User-to-Root) attacks.

Keywords

Intrusion Detection System, R2L, U2R, Fisher Linear Discriminant Analysis, Feature reduction, SPSS, Weka, KDD Cup '99

1. INTRODUCTION

With the increased use of computers and ease of access to internet, the ways to attack and deceive a system has also increased. Though there are various ways to provide security such as cryptography, anti-virus, malwares, spywares, etc., it is not possible to provide complete secured systems. Therefore the need for Intrusion Detection System arose and has become the second line of defense[15]. To identify intruders, differentiating normal user behaviour and attack behaviour is essential. Efficient IDS can be developed by defining a proper rule set for classifying the network traffic log records into normal or attack pattern. Moreover, frequent abnormal traffic on backbone network requires more advanced technologies for monitoring and analysing the network traffic.

For traffic anomaly detection, there are two main approaches: signature-based approach and measurement-based approach. The signature-based approach applies previously established rules to the incoming traffic, while the measurement-based approach considers normal traffic characteristics such as traffic volume and the number of flows as well as link utilization, packet loss, distribution of IP addresses and port number for traffic anomaly detection [8].

The signature-based approach detects the known attacks when they occur. It uses predefined attack signatures and compares a current event against these signatures. Even though the approach shows high detection rate for known attacks, it is ineffective for novel attacks or slightly modified attacks whose signature is not available. To cope with novel attacks and unknown traffic pattern, it needs number of signatures and requires periodical updates with the latest rules. There are tools for this approach such as Snort [11] and Bro [10], and pattern matching is one of the well-known signature-based approaches. On the other hand, the measurement-based approach is designed to identify a source that exhibits deviating behaviour in a system. The construction of such system begins with developing a model for normal behaviour. A detection system can learn the normal behaviour by a training dataset collected over a certain time period with no intrusions. Since this is using traffic characteristics that can be observed through monitoring, it is more flexible and more sensitive than the signature-based approach, especially for detecting new anomalous traffic. However this approach needs to keep per-connection or per-flow state over a single link or node. Therefore, they require a lot of computing resources making their cost unaffordable for many Internet Service Providers. Several tools developed for this approach are ADAM, SPADE and NIDES [12].

Another crucial part of traffic anomaly detection is traffic monitoring. There are two main techniques for traffic monitoring: active monitoring and passive monitoring. Active monitoring monitors the network layer metrics such as delay, jitter, loss, bottleneck point and available bandwidth, by actively injecting probe packets into a network. Even though active monitoring may reduce system overhead by using small number of probe packets that have smaller sizes compared to real data packet, the performance measures may not be accurate for that reason. On the other hand, passive monitoring monitors up to application-layer that includes the user traffic condition such as the sizes of bandwidth, flow and packet, by analysing TCP packets. Since passive monitoring monitors a lot of data packets, it has system overhead problem.

2. RELATED WORK

There are many analytical researches on traffic anomaly detection. One of the analytical detection methods is applying the seasonal Auto-Regressive Integrated Moving Average Model (ARIMA). In [17] and [4], the authors tested the ARIMA model on a specific application such as HTTP of a network. Even though the ARIMA model is an effective time-series forecasting technique, it turns out that the anomalies cannot be well captured by this model [9].

Another approach to Network Intrusion Detection is investigated in [5] which is purely based on Self-Organizing Feature Maps (SOM). In this work, specific attention is given to the representation of connection sequence time and the hierarchical development of abstractions sufficient to permit direct labelling of SOM nodes with connection type. The overall classification rate of this approach is found to be 89%.

In [13], a Network Intrusion Detection system using fuzzy logic is experimented. This technique uses a set of fuzzy rules which are obtained from the definite rules using frequent items. The classification accuracy of this approach is above 90% for all types of attacks.

A study that analyses performance of some neural network when entire KDD dataset is used for training in order to classify and detect attacks is reported in [3]. The five types of neural networks that are studied: Multilayer Perception (MLP), Self-Organizing Feature Map (SOM), Jordan/Elman neural networks, Recurrent Neural Network and RBF neural network. Their results showed that Percent of Correct Classification (PCC) are 99.16%, 98.28%, 98.36%, 98.44% and 79.23% respectively.

The proposed system aims to increase the classification accuracy than the existing approaches explained above.

3. KDD CUP ’99 INTRUSION DETECTION DATASET

KDD Cup ’99 intrusion detection datasets [7] which are based on DARPA ’98 dataset provides labelled data for researcher working in the field of intrusion detection and is the only labelled dataset publicly available. The details of KDD dataset are given in the subsequent section. The KDD dataset is generated using a simulation of a military network consisting of three target machines running various operating systems and traffic. Finally, there is a sniffer that records all network traffic using the Tcpdump format. The total simulated period is seven weeks. Normal connections are created to profile that expected in a military network and attacks fall into one of the four categories:

- *Denial of Service (Dos):* Attacker tries to prevent legitimate users from using a service.
- *Remote to Local (r2l):* Attacker does not have an account on the victim machine, hence tries to gain access.
- *User to Root (u2r):* Attacker has local access to the victim machine and tries to gain super user privileges.
- *Probe:* Attacker tries to gain information about the target host.

There are 41 features for each connection, which are detailed in Table I. Specifically, “a connection is a sequence of TCP packets starting and ending at some well-defined times, between which data flows from a source IP address to a target IP address under some well-defined protocol”. Features are grouped into four categories:

- *Basic Features:* Basic features can be derived from packet headers without inspecting the payload.
- *Content Features:* Domain knowledge is used to access the payload of the original TCP packets. This includes features such as number of failed login attempts.
- *Time-based Traffic Features:* These features are designed to capture properties that mature over a 2 second temporal window. One example of such a feature would be the number of connections to the same host over the 2 second interval.
- *Host-based Traffic Features:* Utilize a historical window estimated over the number of connections instead of time. Host-based features are designed to access attacks, which span intervals longer than 2 seconds.

Table1: List of KDD Cup ’99 features with their descriptions

No	Feature	Description
1	duration	Duration of the connection
2	protocol type	Connection protocol (e.g. TCP, UDP, ICMP)
3	service	Destination service
4	flag	Status flag of the connection
5	source bytes	Bytes sent from source to destination
6	destination bytes	Bytes sent from destination to source
7	land	1 if connection is from/to the same host/port; 0 otherwise
8	wrong fragment	Number of wrong fragments
9	urgent	Number of urgent packets
10	hot	Number of “hot” indicators
11	failed logins	Number of failed logins
12	logged in	1 if successfully logged in; 0 otherwise
13	#compromised	Number of “compromised” conditions
14	root shell	1 if root shell is obtained; 0 otherwise
15	su attempted	1 if “su root” command attempted; 0 otherwise
16	#root	Number of “root” accesses
17	#file creations	Number of file creation operations
18	#shells	Number of shell prompts
19	#access files	Number of operations on access control files
20	#outbound cmds	Number of outbound commands in a ftp session
21	is hot login	1 if login belongs to the “hot” list; 0 otherwise
22	is guest login	1 if the login is the “guest” login; 0 otherwise
23	count	Number of connections to the same host as the current connection in the past 2 seconds
24	srv count	Number of connections to the same service as the current connection in the past two seconds
25	serror rate	% of connections that have “SYN” errors
26	srv serror rate	% of connections that have “SYN” errors
27	rerror rate	% of connections that have REJ errors
28	srv rerror rate	% of connections that have REJ errors
29	same srv rate	% of connections to the same service
30	diff srv rate	% of connections to different services
31	srv diff host rate	% of connections to different hosts
32	dst host count	Count of connections having the same destination host

33	dst host srv count	Count of connections having the same destination host and using the same service
34	dst host same srv rate	% of connections having the same destination host and using the same service
35	dst host diff srv rate	% of different services on the current host
36	dst host same src port rate	% of connections to the current host having the same src port
37	dst host srv diff host rate	% of connections to the same service coming from different hosts
38	dst host serror rate	% of connections to the current host that have an S0 error
39	dst host srv serror rate	% of connections to the current host and specified service that have an S0 error
40	dst host rerror rate	% of connections to the current host that have an RST error
41	dst host srv rerror rate	% of connections to the current host and specified service that have an RST error

The KDD '99 intrusion detection benchmark consists of three components, which are detailed in Table II. In the International Knowledge Discovery and Data Mining Tools Competition, only "10% KDD" dataset is employed for the purpose of training. This dataset consists of 22 attack types and is a more concise version of the "Whole KDD" dataset. It contains more examples of attacks than normal connections and the attack types are not represented equally. Because of their nature, DoS attacks account for the major of the dataset. On the other hand, the "Corrected KDD" dataset provides a dataset with different statistical distributions than either "10% KDD" or "Whole KDD" and contains 14 additional attacks.

Table2: Basic characteristics of KDD '99 Intrusion Detection Datasets in terms of number of samples

Dataset	DoS	Probe	U2R	R2L	Normal
10% KDD	391458	4107	52	1126	97277
Corrected KDD	229853	4166	70	16347	60593
Whole KDD	3883370	41102	52	1126	972780

In this paper, CorrectedKDD is used. There are 37 types of attacks in the dataset with varying percentage of different attacks. The various types of attacks in our experimental dataset which are classified into four categories are shown in Table 3.

Table 3: Attack types with their corresponding categories

Category	Attack types
Probe	ipsweep, mscan, nmap, portsweep, saint, satan
DoS	apache, back, land, mailbomb, neptune, pod, processtable, smurf, teardrop, udpstorm
U2R	buffer_overflow, loadmodule, perl, rootkit, ps, sqlattack, xterm
R2L	ftp_write, guess_password, imap, multihop, named, phf, sendmail, snmpgetattack, snmpguess, warezmaster, worm, xlock, xsnoop, httptunnel

4. PROPOSED WORK

Discriminant analysis is a statistical technique used to build predictive model of group membership based on observed characteristics of each case. The purpose of discriminant analysis is to classify objects (e.g. graduate, undergraduate, etc.,) based on attribute set which describe the objects (e.g. age, gpa, etc.,). The dependent variable (*y*) is the group and the independent variables (*x*) are the object features that describe the group.

Linear discriminant model can be used for groups that are linearly separable (i.e. the groups that can be separated by a linear combination of features that describe the objects). If there are only two features, the separators between object groups will become line. If the features are three, the separator is a plane and if the number of features is more than three, the separators become a hyper plane.

The proposed work comprises of 2 steps:

1.Feature reduction using correlation based analysis
2.Classification with reduced feature set

The architecture of the proposed system is shown in Figure 1. The Corrected KDD dataset is used. The preprocessing step involves the mapping of symbolic valued attributes into numeric valued attributes. Symbolic features like protocol type (3 different symbols), service (23 different symbols) and flag (7 different symbols) are mapped to integer values ranging from 0 to N-1 where N is the number of symbols. Then the KDD '99 dataset is fragmented into 4 subsets, each containing records of normal and a specific attack category.

.

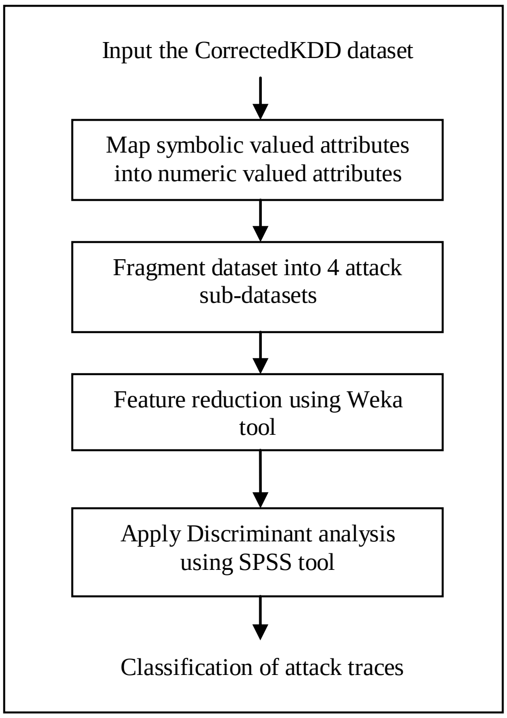
Input the CorrectedKDD dataset
Map symbolic valued attributes into numeric valued attributes
Fragment dataset into 4 attack sub-datasets
Feature reduction using Weka tool
Apply Discriminant analysis using SPSS tool
Classification of attack traces

Figure 1: Architecture of the proposed system

4.1. Feature reduction for KDD '99 dataset:

Feature reduction is a process of choosing a subset of original features so that the feature space is optimally reduced according to a certain evaluation criterion. In general, a feature is good if it is relevant to the class concept but not redundant to any of the other features. When the correlation between two variables is adopted as a goodness measure, as per the above definition the selected feature is considered good if it is highly correlated to the class.

The most well-known correlation measure is linear correlation coefficient. For a pair of variables *(x, y)*, the linear correlation coefficient *r(x, y)* is given by the formula (1):

$$r(x,y) = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left(n\sum x^2 - (\sum x)^2\right)\left(n\sum y^2 - (\sum y)^2\right)}} \dots\dots\dots\dots\dots (1)$$

In this experiment, Weka tool [16] is used for feature reduction. CfsSubsetEval with Best first approach is applied on the training dataset to obtain the important features for the classification process which is visualised as a two class categorization problem. Each subset is analysed using the correlation analysis for identifying the important features for a specific attack. This analysis result gives a set of features for each subset which is sufficient to group the attack and normal records. These features are considered as relevant features for each attack. The reduced features are tabulated in Table 4.

Table 4: List of features for which the class is selected most relevant

No	Attack Category	Reduced features
1	DoS	11, 12, 23, 29, 31, 37
2	Probe	12, 25, 27, 29, 33, 37, 38
3	U2R	9, 14, 18, 21, 33, 38
4	R2L	2, 3, 4, 5, 11, 12, 18, 19, 21, 22, 24, 27, 28, 31, 38, 40, 41

4.2 Classification with reduced feature set:

With the reduced features, discriminant analysis is done using SPSS tool [14] with Mahalanobis distance in Stepwise statistics. The output of the above step returns the discriminant value of the features along with classification summary. Classification and misclassification are based on actual and predicted group membership. Figure 2 shows the classification rate of all attack categories.

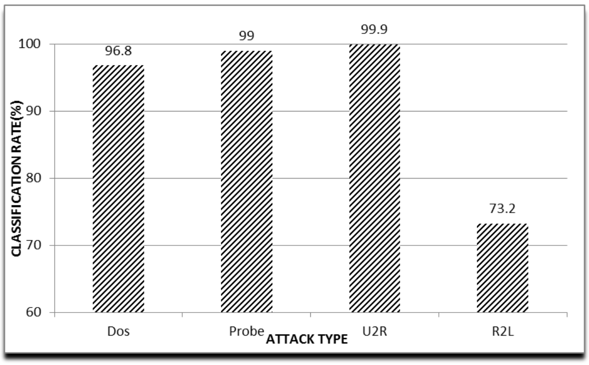

Figure 2: Classification rate of attack categories using FLDA

5. RESULTS DISCUSSION

Experimental results show that classification rate for all the four attack categories ie., DoS, Probe, U2R and R2L are better and even surpassing some of the other approaches. With respect to U2R and R2L category of attacks, our approach scored far better than the other approaches. It showed a classification rate of 99.9 % for U2R and 73.2 % for R2L.

However the R2L classification rate of 73.2% is low compared to the classification rate of other category of attacks. The reason behind this is explored in this section. It should be noted that most of the machine learning algorithms offered an acceptable level of classification rate for DoS and Probe attack categories as they exhibit multiple connections over a short period of time, while demonstrated poor performance for the R2L and U2R categories as these attacks are embedded in their data packets itself and do not form a sequential pattern unlike DoS and Probe attacks. This makes their detection by any classifier a difficult task. Inspite of this, our approach gained good classification rate for U2R category (99.9%).

6. 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 5.

Table 5: Comparison between some IDS techniques and our approach

Approach in [paper]	DoS	Probe	U2R	R2L
[2]	96.9%	73.2%	6.6%	10.7%
[1]	99.55%	100%	0%	99.981%
[6]	96.5%	72.8%	22.9%	11.3%
Our approach	96.8%	99%	99.9%	73.2%

According to the above table, proposed system has good performance that is competitive with other approaches based on classification rate.

7. CONCLUSION

In this paper, Fisher Linear Discriminant Analysis is carried out on KDD ’99 dataset. First, a feature relevance analysis is performed by applying correlation based feature selection. It analyses the involvement of each feature to classification and a subset of features are selected as relevant features. Then discriminant analysis is applied on reduced feature set and classification summary for each category is presented. As compared to the existing techniques, our proposed work fairly improves the classification accuracy for R2L and U2R attacks. Hence we can conclude that the Fisher Linear Discriminant Analysis proves to be an efficient classifier for R2L and U2R attacks.

REFERENCES

[1] Adel Jahanbani and Hossein Karimi, “A new approach for detecting intrusions based on the PCA neural networks”, Journal of Basic and Applied Scientific Reasearch, pp. 672-679, 2012.

[2] R. Agarwal and M. V. Joshi, “PNrule: A new framework for learning classifier models in data mining (A case study in Network Intrusion Detection), IBM research division technical report no. RC-21719, 2000.

[3] Beghdad. R, “Training all the KDD dataset to classify and detect attacks” in International Journal on Neural and Mass – Parallel computing and Information Systems, Vol. 17, March 2007.

[4] Y.W.Chen, "Traffic behaviour analysis and modelling of sub-networks", in International journal of network management, Vol. 12, pp. 323-330, September 2002.

[5] H. Gunes Kayacik, A. Nur Zincir Heywood and I. Heywood, "On the capability of an SOM based Intrusion Detection System", in Proceedings of the International conference on Neural Networks, Vol. 3, pp. 1808-1813.

[6] H. Gunes Kayacik, A. Nur Zincir Heywood and I. Heywood, "An hierarchical SOM-based intrusion detection" in Journal on Engineering Applications of Artificial Intelligence, Vol. 20, pp. 439-451, 2007.

[7] KDD Cup 1999 Intrusion Detection Dataset. [Online]. Available:http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

[8] S.S.Kim and A.L.N.Reddy, "Statistical techniques for detecting traffic anomalies through packet header data", IEEE/ACM Transaction on Networking, Vol. 16, no. 3, pp.562-575, January 2008.

[9] H. Moayedi and M.Masnadi-Shirazi, "Arima model for network traffic prediction and anomaly detection", in Proc. ITSim 2008, pp. 1-6, August 2008.

[10] V.Paxson, "Bro: A system for detecting network intruders in realtime", in Proc. USENIX Security Symposium, January 1998.

[11] M.Roesch, "Snort-lightweight intrusion detection for networks" in Proc. USENIX LISA 1999, pp.229-238, November 1999.

[12] R.Sekar, M.Bendre, D.Dhurjati and P.Bollineni, "A fast automation based method for detecting anomalous program behaviors", in Proc. IEEE Symposium on Security and Privacy, May 2001.

[13] R. Shanmugavadivu and Dr. N. Nagarajan, " Network intrusion detection system using fuzzy logic", in Indian Journal of Computer Science and Engineering, Vol. 2, pp. 101-111.

[14] SPSS Inc., "SPSS 13.0 Base User's Guide"

[15] Theuns Verwoerd, Ray Hunt, "Intrusion detection techniques and approaches", Computer Communications, Vol. 25, pp. 1356-1365, 2002.

[16] Weka tool. [Online]. Available: http://www.cs.waikato.ac.nz/ml/weka.

[17] Y. Zhang, Z.Ge, A.Greenberg and M.Roughan, "Network Anomography", in Proc. USENIX IMC 2005, p. 317-330, October 2005.