Journal of
Electrical & Electronic Systems

Research Article Open Access

Denial of Service (DoS) Attacks using PART Rule and Decision Table Rule

Aladesote O Isaiah*, Johnson OV and Ganiyu Mutiu

Department of Computer Science, Federal Polytechnic, Ile – Oluji, Ondo State, Nigeria

Abstract

Network Security has become a major and critical issue as a result of the vast growth in the field of Information Technology. This paper adopted the result of an existing extraction or attributes selection of KDD ’99 dataset. The dataset was run on data de-duplicated software developed using C# Programming Language and final mining analysis was carried out on Waikato Environment for Knowledge Analysis (WEKA) with the adoption of PART and Decision Table algorithms. The performance evaluation was carried out with some related existing works based on certain intrusion detection metrics. The Classification Rate of Decision Tree Rule, Part Rule and JRIP Rule are 98.14%, 99.4% and 99.1%, respectively. The False Alarm Rate of Decision Tree Rule, Part Rule and JRIP Rule are 0.86%, 0.43% and 0.55% respectively. The Sensitivity of Decision Tree Rule, Part Rule and JRIP Rule is 92.6%, 98.3% and 97.2% respectively while the Specificity of Decision Tree Rule, Part Rule and JRIP Rule is 99.1%, 99.6% and 99.4% respectively.

Keywords: Waikato Environment for Knowledge Analysis (WEKA); PART rule; Detection metrics; Decision table rule; Data deduplication

Introduction

Intrusion detection is an efficient method of dealing with network security related problems [1]. Network Security has become a serious concern due to the development and expansion in the field of Information Technology [2]. This appreciable improvement in network technologies has showed a way for invaders or hackers to devise an unauthorised means into a network system. Therefore, an effective and timely Intrusion Detection System, which helps to enhance the security of a network, is needed when attack(s) is/are noticed [3]. Intrusion detection is a security approach used to protect computer networks from unauthorised access [1].

An intrusion can be defined as any attempt that violates the basic elements of information security: confidentiality, integrity and availability [4]. There is necessity to apply data mining in Intrusion Detection System owing to the huge amount of existing intrusion dataset and also recently emerging network dataset [5]. There is need for effective and efficient intrusion system as conservative intrusion detection approach can no longer match the newly emerging dataset.

Coupled with enormous data available today with lots of record duplications, which to use for optimal data analysis becomes challenging. Data deduplication thus, helps to remove such bottlenecks, thereby leaving a copy of each record in a set of data; this leads to the reduction in the amount of data to be moved into the network [6].

Research Motivation

In the work of ref. [4], Hypothesis Testing was applied on KDD dataset. The significant attributes or features of the dataset were extracted; the records of the thirteen significant attributes were used in the research. The training set was run on an existing Decision Tree algorithm which resulted in some rules. The mean of each rule was determined and later used to form hypothesis. The accuracy of the system was tested using some detection metrics. Meanwhile there is the need to valid the accuracy of the existing result by applying data deduplication with other mining algorithm on the intrusion dataset to help offer more accurate classification.

Research Objective

The objectives of the research work are to develop deduplicated program, classify intrusion dataset using PART and Decision table Rules and also to carry out performance evaluation on the KDD dataset.

Methodology

Review of few existing works was carried out. The NSL-KDD dataset which is an improvement upon KDD ‘99 data was used. The records of Denial of Service (DoS) attacks and normal traffic based on the thirteen significant attributes were extracted, this contains Eighteen thousand, One hundred and Thirteen (18113) records. The dataset was run on data deduplication program developed using C#.

Decision table and PART Rules were used to classify the Denial of Service (DOS) attacks and normal traffic from WEKA data mining implementation. The performance of the system would be tested on the test data using classification rate, detection rate and false alarm rate, after which the comparative analysis would be carried out against the work of Oladunjoye [7].

Result and Discussion

Data deduplication

Table 1 shows the result obtained when the dataset was run on Data deduplicated program. 9711 records of Normal traffic were reduced to 7761, which amount to 20.1% reduction. 737 records of Apache2 were reduced to 440, which is 40.3% reduction. 359 records of Back were reduced to 65, which is 82% reduction. 7 records of Land were reduced to 3, resulting in 57.1%. 293 records of Mail bomb were reduced to 4, which amount to 98.6% reduction. 4557 records of Neptune were reduced to 295, which is 93.5% reduction. 41 records of Ping of Death (PoD) were reduced to 14, which equate to 65.8% reduction. 685 records of Processtable were reduced to 367, which is 46.4% reduction. 665 records of smurf were reduced to 10, which is equivalent to 98.5%

***Corresponding author:** Aladesote O Isaiah, Department of Computer Science, Federal Polytechnic, Ile – Oluji, Ondo State, Nigeria, Tel: +234-07039090114; E-mail: isaaladesote@fedpolel.edu.ng, lomaladesote@yahoo.com

Received March 22, 2016; **Accepted** April 24, 2017; **Published** April 29, 2017

Citation: Isaiah AO, Johnson OV, Mutiu G (2017) Denial of Service (DoS) Attacks using PART Rule and Decision Table Rule. J Electr Electron Syst 6: 220. doi:10.4172/2332-0796.1000220

Attacks/Normal Traffic	Before Deduplication	After Deduplication
Normal	9711	7761
Apache2	737	440
Back	359	65
Land	7	3
Mailbomb	293	4
Neptune	4557	295
PoD	41	14
Processtable	685	367
Smurf	665	10
Teardrop	12	2
Udpstorm	2	1
Warezmaster	944	180
Total	18113	9142

Table 1: Result obtained when the dataset was run on Data Deduplicated Program.

Attacks/Normal Traffic	TCC	TWC	TUC	TOTAL
Apache2	117	23	0	140
Back	23	0	0	23
Land	0	2	0	2
Mailbomb	0	1	0	1
Neptune	93	0	0	93
Normal	2303	20	0	2323
Ping of Death (PoD)	3	0	0	3
Processtable	107	0	0	107
Smurf	1	1	0	2
Teardrop	0	0	0	0
Udpstorm	0	0	0	0
Warezmaster	45	4	0	49

TCC: Test Correctly Classified; TWC: Test Incorrectly Classified; TUC: Test Unclassified

Table 2: Performance of Rules Generated on Test Data.

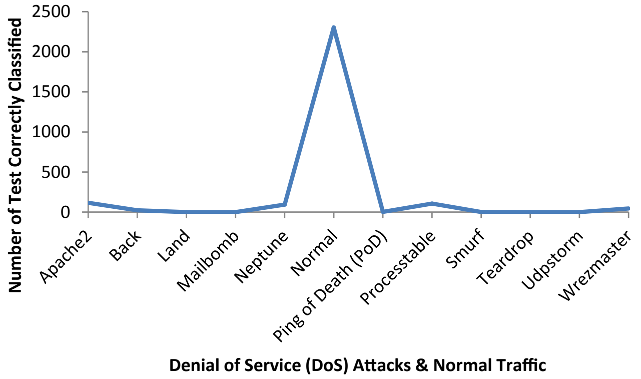

Figure 1: Graphical Representation of Decision Table rules on Test Data that are correctly classified.

reduction.12 records of teardrop were reduced to 2, which corresponds to 83.3% reduction. 2 records of teardrop were reduced to 1, which is 50% reduction while 994 records of warezmaster were reduced to 180, which is 80.9% reduction.

Performance of rules generated using decision table rules

The performance of rules generated on test data using Decision Table Rules from Table 2, Figures 1 and 2 show that out of 2303 records of Normal traffic, 2303 were correctly classified while 20 were wrongly classified. Out of 140 records of Apache2, 117 were correctly classified while 23 were wrongly classified. All records of Back, Neptune, PoD and processtable were correctly classified. A record of mail bomb was wrongly classified. Out of 2 records of Smurf, 1 was correctly classified while the remaining 1 was wrongly classified. Out of 49 records of warezmaster, 45 were correctly classified while 4 were wrongly classified. Teardrop and udpstorm have no record in the test data.

Performance of rules generated using part rules

The performance of rules generated on test data using PART Rules from Table 3, Figures 3 and 4 show that all records of Apache2, Back, Mail bomb, PoD processtable and Smurf were correctly classified. The 2 records of Land were wrongly classified. Out of 92 records of Neptune, 92 were correctly classified and 1 was wrongly classified. Out of 2323 records of Normal traffic, 2313 were correctly classified while 10 were wrongly classified. Out of 49 records of warezmaster, 45 were correctly classified while 4 were wrongly classified. Teardrop and udpstorm have no record in the test data.

Confusion matrix obtained from denial of service (dos) and normal traffic using decision table rules

Table 4 shows the confusion matrix obtained from the Decision Table Rules Classification when DOS attacks and Normal Traffic test data were used. Out of 140 records of Apache2, 117 were correctly classified, while 21 and 2 were incorrectly classified as Neptune and Normal respectively. All records of Back, Neptune, Ping of Death (POD) and processtable were correctly classified. The 2 records of Land were incorrectly classified as Neptune. A record of Mail bomb was incorrectly classified as Normal. Out of 2323 records of Normal

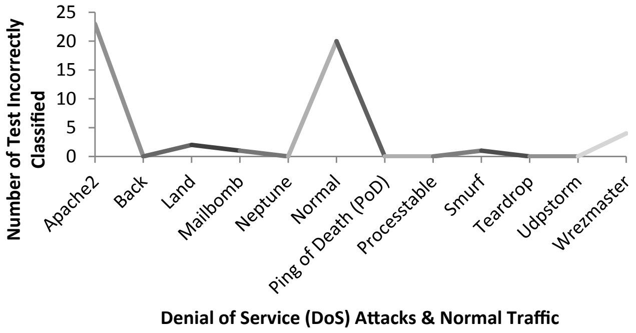

Figure 2: Graphical Representation of DecisionTable rules on Test Data that are incorrectly classified.

Attacks/Normal Traffic	TCC	TWC	TUC	TOTAL
Apache2	140	0	0	140
Back	23	0	0	23
Land	0	2	0	2
Mailbomb	1	0	0	1
Neptune	92	1	0	93
Normal	2313	10	0	2323
Ping of Death (PoD)	3	0	0	3
Processtable	107	0	0	107
Smurf	2	0	0	2
Teardrop	0	0	0	0
Udpstorm	0	0	0	0
Warezmaster	45	4	0	49

TCC: Test Correctly Classified; TWC: Test Incorrectly Classified; TUC: Test Unclassified.

Table 3: Performance of Rules Generated on Test Data.

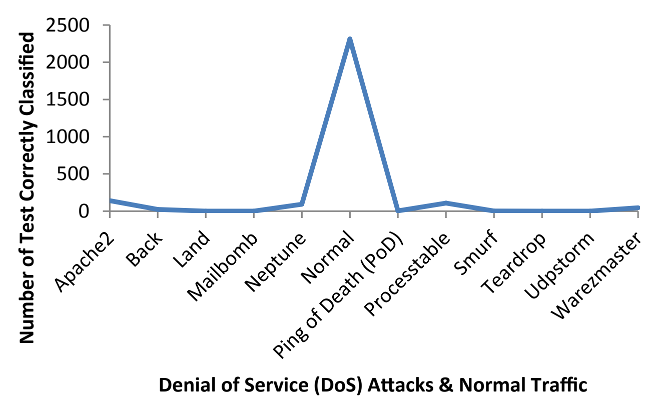

Figure 3: Graphical Representation of PART rules on Test Data that are Correctly Classified.

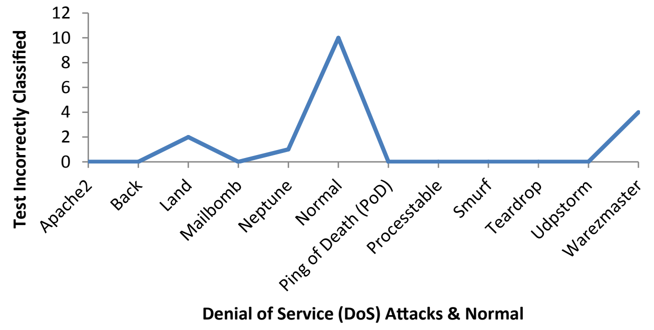

Figure 4: Graphical Representation of PART Rules on Test Data that are Incorrectly Classified.

	Ap	Ba	La	Ma	Nep	Nor	Pod	Pro	Smu	Tea	Udp	Wam
Ap	117	0	0	0	21	2	0	0	0	0	0	0
Ba	0	23	0	0	0	0	0	0	0	0	0	0
La	0	0	0	0	2	0	0	0	0	0	0	0
Ma	0	0	0	0	0	1	0	0	0	0	0	0
Ne	0	0	0	0	93	0	0	0	0	0	0	0
Nor	11	1	0	0	7	2303	1	0	0	0	0	0
Pod	0	0	0	0	0	0	3	0	0	0	0	0
Pro	0	0	0	0	0	0	0	107	0	0	0	0
Sm	0	0	0	0	0	0	1	0	1	0	0	0
Te	0	0	0	0	0	0	0	0	0	0	0	0
Ud	0	0	0	0	0	0	0	0	0	0	0	0
Wa	0	0	0	0	0	4	0	0	0	0	0	45

Table 4: Confusion Matrix obtained from decision table rules system on test data.

Traffic, 2303 were correctly classified while 11, 1, 7 and 1 were incorrectly classified as Apache2, Back, Neptune and Ping of Death (POD) respectively. 1 of the 2 records of Smurf was correctly classified while the other was incorrectly classified as POD. Out of 49 records of warezmaster, 45 were correctly classified while 4 were incorrectly classified as Normal.

TN=2303; FP=20; FN=21; TP=389

$$\text{Classification Rate }(\text{CR})=\frac{TP+TN}{TP+TN+FP+FN}=98.14\%$$

$$\text{False Alarm Rate }(\text{FAR})=\frac{FP}{TN+FP}=0.86\%$$

Sensitivity=(100 × TP/TP+FN)

=92.6%

Specificity=(100 × TN/TN+FP)=99.1%

Confusion matrix obtained from denial of service (dos) and normal traffic using part rules

Table 5 shows the confusion matrix obtained from the PART Rules Classification when DOS attacks and Normal Traffic test data were used. All records of Apache2, Back, Mail bomb, Ping of Death (POD), Processtable and Smurf were correctly classified. The 2 records of Land were incorrectly classified as Neptune. Out of 93 records of Neptune, 92 were correctly classified while 1 was incorrectly classified as Apache2. 2313 records of Normal were correctly classified out of 2323 while 1, 1, 1, 5 and 2 were incorrectly classified as Apache2, Back, Mail bomb, Neptune and Warezmaster respectively. Out of 49 records of Warezmaster, 45 were correctly classified while 4 were incorrectly classified as Normal.

NO*=Normal, WM*=Warezmaster, US*=Udpstorm, TD*=Teardrop, SM*=Smurf, PR*=Processtable, PD*=Pod, NE*=Neptune, MB*=Mailbomb, LA*=Land, BA*=Back, AP*=Apache2.

TN = 2313; FP = 10; FN = 7; TP = 413

$$\text{Classification Rate }(\text{CR})=\frac{TP+TN}{TP+TN+FP+FN}=99.4\%$$

$$\text{False Alarm Rate }(\text{FAR})=\frac{FP}{TN+FP}=0.43\%$$

Sensitivity=(100 × TP / TP + FN)=98.3%

Specificity=(100 × TN / TN + FP)=99.6%

Performance evaluation with existing system

Table 6 shows the number of records that are correctly classified incorrectly classified and not classified for each denial of services attacks and normal traffic.

Figure 5 reveals that the % of the record correctly classified using decision tree rules 98.14%, 99.43% when PART rules methods are

	Ap	Ba	La	Ma	Ne	No	Po	Pr	Sm	Te	Ud	Wa
AP*	140	0	0	0	0	0	0	0	0	0	0	0
BA*	0	23	0	0	0	0	0	0	0	0	0	0
LA*	0	0	0	0	2	0	0	0	0	0	0	0
MA*	0	0	0	1	0	0	0	0	0	0	0	0
NE*	1	0	0	0	92	0	0	0	0	0	0	0
NO*	1	1	0	1	5	2313	0	0	0	0	0	2
PD*	0	0	0	0	0	0	3	0	0	0	0	0
PR*	0	0	0	0	0	0	0	107	0	0	0	0
SM*	0	0	0	0	0	0	0	0	2	0	0	0
TD*	0	0	0	0	0	0	0	0	0	0	0	0
US*	0	0	0	0	0	0	0	0	0	0	0	0
WM*	0	0	0	0	0	4	0	0	0	0	0	45

Table 5: Confusion Matrix obtained from PART rules system on Test Data.

	Classification Rate (%)	False Alarm Rate (%)	Sensitivity (%)	Specificity (%)
DecisionTree Rules	98.14	0.86	92.6	99.1
PART Rules	99.4	0.43	98.3	99.6
JRIP Rules	99.1	0.55	97.2	99.4

Table 6: Performance evaluation with an existing work.

used and 99.1% for JRIP rules. It can be deduced that PART rules is competitively better with this type of classification than the other two methods.

Figure 6 shows the % of the normal connections that are not correctly classified in the training and testing sets. The result show that FAR is 0.86 when decision tree rules is applied, 0.43 when PART rules is used and 0.55 JRIP rules is used. This indication that the percentage of records that are misclassified is minimal when rules in PART used. Therefore, PART rules are preferably better in term of false Alarm rate for this type of classification.

Figure 7 show the % of the number of attacks connection that is correctly classified. The result indicates that the number of attacks that are correctly classified when decision tree Rules in used is 92.6%, 98.3% when PART rules in used whiles 97.2% when JRIP rules in used. PART rules perform better than the two other methods in term of sensitivity.

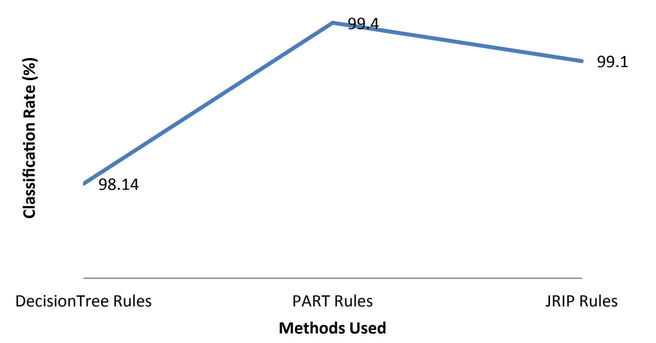

Figure 5: Graphical Representation of Classification Rate of different methods.

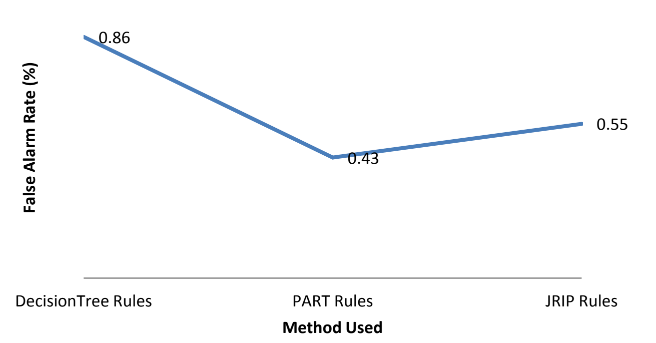

Figure 6: Graphical Representation of False Alarm Rate (FAR) of different methods.

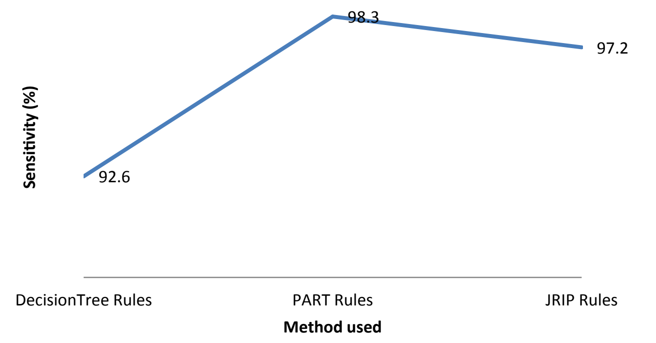

Figure 7: Graphical Representation of Sensitivity of different methods.

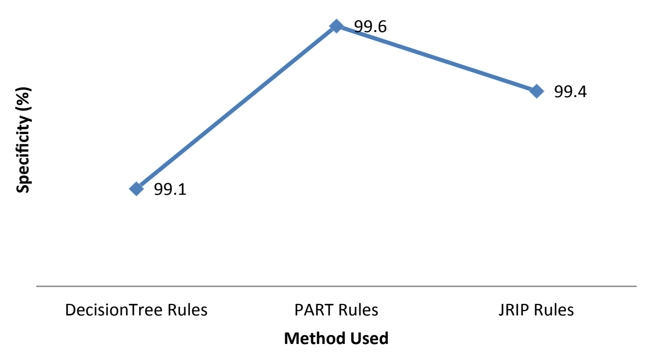

Figure 8: Graphical Representation of Specificity of different methods.

Figure 8 shows the specificity is 99.1% when decision tree is used, 99.6% when PART rules is used and 99.4% when a JRIP rule is used.

Conclusion

The system shows that PART Rules performed better than other methods in terms of Classification Rate, False Alarm Rate, Sensitivity and Specificity.

References

1. Manandhar P (2014) A Practical Approach to Anomaly-based Intrusion Detection System by Outlier Mining in Network Traffic. Masdar Institute of Science and Technology.
2. Amudha P, Karthik S, Sivakumari S (2015) A Hybrid Swarm Intelligence Algorithm for Intrusion Detection Using Significant Features. The Scientific World Journal.
3. Jaiganesh V, Sumathi DP, Mangayarkarasi S (2013) An Analysis of Intrusion Detection System using Back Propagation Neural Network. IEEE Computer Society Publication.
4. Aladesote OI, Boniface KA, Dahunsi F (2014) Intrusion Detection Technique using Hypothesis Testing. Proceedings of the World Congress on Engineering and Computer Science.
5. Shona D, Senthilkumar M (2016) International Journal of Applied Engineering Research 11: 4161-4166.
6. Meister D (2013) Advanced Data Deduplication Techniques and their Application. Johannes Gutenberg University, Mainz.
7. Oladunjoye F (2015) Intrusion Detection System using JRIP. Unpublished manuscript, Rufus Giwa Polytechnic, Owo, Nigeria.