https://doi.org/10.15623/ijret.2018.0710003 Received: 24-07-2018, Accepted: 06-09-2018, Published: 27-09-2018

# PERFORMANCE EVALUATION OF ADVANCED INTRUSION DETECTION SYSTEM

**Kiran Bala[1], Narendra Kumar[2], Ashok Kumar Singh[3]**

[1,2]*Department of Computer Science, Magadh University, Bodh Gaya, India*
[3]*Department of Mathematics, Magadh University, Bodhgaya, India*

## Abstract

*An intrusion detection system (IDS) plays an important role in securing our computer networks. There are many techniques designed to help in detecting and preventing such attacks. Intrusion detection system (IDS) is used to observe unwanted action on network systems and individual computers. However due to the opposite effect of using IDS, using individual methods of IDS only misuse or anomaly attacks can be detected. This paper proposed a model that integrates both approaches signature and anomaly based on IDS to reduce obtained alerts and detects new attacks. False alert rate, area under ROC curve, accuracy, recall, F-score, and P-test are the parameters used to evaluate the effectiveness of hybrid IDS. This paper also considers KDD Cup 1999 information set keeping in mind the end goal to show the profits of the proposed intrusion detection system described in the KDD Cup record previously.*

***Keywords: -*** *Intrusion Detection Systems (IDS); Traffic generator; Performance evaluation; Snort; KDD.*

---------------------------------------------------------------------***---------------------------------------------------------------------

## 1. INTRODUCTION

Intrusion Detection Systems (IDS) has become one of the essential components for any organization's network. IDS is designed to detect any intrusion or hostile traffic in the network. Intrusion can be defined as an illegal mechanism that can be used to attack on a computer system resources or data. The attacker or designers are called intruders. Intruders can be classified to two primary categories Internal Intruders and External Intruders [1].

There are two main categories of IDS: signature based and anomaly based [2, 3] or integration between this two classes in purpose of developing their shared complementarities [4]. Signature based (Misuse detection) uses a database signature of known attacks to detect malicious activities, every single signature represents finger print of the attack; IDS raises the alarm when the traffic of the network matches one of the signatures, database system needs to be updated continuously in order to maintain the signature of new attacks. Whereas the second method is based on anomaly. This defines a model of normal behavior system, by creating a profile which describes normal usage as compared to the current behavior to detect deviations of raised alarms. Various techniques can be used to construct profile such as machine learning and data mining. One of the advantages of using anomaly based method is the ability to detect novel attacks. On the other hand, systems based on this method has a high rate of false alarm [5].

## 2. EVALUTION PARAMETERS FOR IDS

This section presents the parameters utilized prior for IDS execution assessment with their advantages and disadvantages for IDS assessment and examines them to perform proposed system performance evaluation.

### 2.1 Detection Rate (True Positive Rate) And False Alert Rate (False Negative Rate)

This is very commonly used performance parameter for IDS evaluation. Consider, TP be the count of attacks that are properly recognized and noticed by IDS, and FN be the count of attacks that are not noticed by IDS. So, we can consider as TN be the number of normal traffic packets that are correctly classified by IDS, and FP be the quantity of ordinary movement parcels that are effectively arranged by IDS, furthermore, for any IDS, it should satisfy security needs as well as the usability of IDS in the operating environment. Here, the security requirement can be analyzed with the help of TP rate and the convenience prerequisite is focused around the number of FPs as natural tradeoff between these two measurements.

The idea of discovering the ideal tradeoff of the measurements used to assess IDS is an example of the more general issue of multi criteria streamlining. To accomplish this, we need to either augment or minimize two amounts that are connected with one another by a tradeoff; we can straightforwardly pose as a viable rival the two or natural tradeoff between these two measurements. We can in this manner group the above characterized measurements into two conceivable approaches that will be investigated in whatever is left of this area: the tradeoff approach and the boost of a figure-of-merit esteem.

## 2.2 Area under ROC Curve (AUC)

AUC is the execution metric utilized for the ROC curve and is another good metric to measure the performance of IDS. A region of 0.5 is considered as great result while a perfect one will have a range of 1.ROC curves are utilized as a part of IDS to assess classifier execution over a scope of tradeoffs in the middle of TP $_{\text{rate}}$ and FP $_{\text{rate}}$. ROC bend is a diagram which has the x-axis as the false alert rate and y-axis as the location rate.

i.e.,

$$ROC = \{TP_{rate}\,; FP_{rate}\}$$

One of the profits of ROC chart is we can discover separate slips cost contemplations from the IDS execution. Also, it remains constant under varying class distributions in IDS solicitation. The demerit of this metric is, it causes drastic change with detection rate even with small variation in false alarm rate when the usual traffic flow inevitable compared to the violence stream of traffic in the system traffic flow.

## 2.3 Accuracy

The normally utilized IDS evaluation metric on a test data is the overall accuracy. It is normally utilized IDS assessment metric on test information is the general accuracy. It is calculated as Though, it is commonly used IDS evaluation parameter, is not a decent parameter for examination on account of system activity information since the genuine negatives flourish.

## 2.4 F-Score

F-Score is used to represent the balance in the middle of precision and recall. The F-score is utilized to quantify the exactness of a test. The F-score is harmonic mean estimation of review and accuracy, it is computed as: Precision, Recall and F-Score are the standard measures are derived based on probability theory and allows one to consider the natural varieties of execution estimation of any system. Out of all these parameters, F-score has the restriction for specifically applying tests of essentialness to IDS in request to focus the certainty level of the examination. The principle objective in IDS assessment is always the requirement for enhancement in together precision as well as recall values, so for this purpose the P-Test parameter can be used.

## 2.5 P-Test

It is another parameter which can be used to compare IDSs based on the tradeoff between precision and recall values with respect to attack detection of IDS as introduced by (Ciza Thomas (2009). Consider, IDS systems as A and B, then Consider, ($R_A$; $P_B$) and ($R_B$;$P_B$) be the estimations of review and accuracy regarding assault separately. Let A and B IDS distinguish $N_{pos}^{A}$ and $N_{pos}^{B}$ positives respectively and $N_{pos}$ be the aggregate number of positives in the gave test specimen. At that point the P-Test is connected as:

$$Z_R = \frac{R_A - R_B}{\sqrt{2R(1-R)\big/ N_{pos}}}$$

$$Z_P = \frac{P_A - P_B}{\sqrt{2P(1-P)\left(1\big/ N_{pos}^{A} + 1\big/ N_{pos}^{B}\right)}}$$

Where,

$$R = \frac{R_A + R_B}{2}\ ;\text{ and }\quad P = \frac{N_{pos}^{A}\,P_A + N_{pos}^{B}\,P_B}{N_{pos}^{A} + N_{pos}^{B}}$$

If $Z_R \geq 1.96$, then it is concluded that, $R_A$ value is significantly better than R$_B$ at 95 % Confidence Level.

If $Z_R \leq -1.96$, then it is concluded that, $R_B$ value is significantly better than R$_A$ at 95 % Confidence Level.

If Z$_R$ ≤1.96, then it is concluded that, $R_A$ value is being comparative to R$_B$. Similar test is available for evaluating P$_A$ and P$_B$.

Now, following criteria can be used to compare two IDSs *A* and *B*. If $IDS_A$ is better than$IDS_B$, then

$$R_A \geqslant R_B and P_A \sim P_B$$

$$R_A \geqslant R_B \quad and P_A \geqslant P_B$$

$$R_A \sim R_B and P_A \geqslant P_B$$

$$R_A \geqslant R_B and P_A \sim P_{B,}\ then\ A \sim B$$

## 3. SYSTEM MODEL

The experimental evaluation of proposed IDS was set with client server architecture. As stated earlier, it was possible to work with this system either in online as well as offline mode of operation. In online mode, five Pentium machines with windows operating system were used to capture the real-time network packet traffic with the help of Packet Sniffer tool installed on each of them. Even, network traffic data is also collected from existing network servers configured in the local network. The collected network traffic data then forwarded and collected at server side of the system for the further analysis. Similarly, in offline mode of operation, the KDD Cup 1999 data set is collected on server machine and performed further computation of the system to evaluate the IDS. The proposed system is integration of signature based IDS and anomaly based IDS, so along with proposed IDS, the following IDS are chosen to evaluate the IDS performance to examine the performance of proposed IDS with signature based as well as in anomaly based IDSs too.

https://doi.org/10.15623/ijret.2018.0710003 Received: 24-07-2018, Accepted: 06-09-2018, Published: 27-09-2018

The main data used for the evaluation is network packet attribute details with their values captured in online or collected in offline mode of operation. The Packet Sniffer tool used in online mode of operation to collect the evaluation data can extract the packet feature attribute values information to distinguish assaults from distinctive layers of system activity like bundle header peculiarity characteristics, parcel payload or content feature attributes and packet traffic feature attributes. So, it is quite flexible to get the required packet feature attribute values by running Packet Sniffer for IDS evaluation. Even, the selected KDD Cup 1999 data set is also collection of extracted information to detect attacks based on 41 packet feature attribute values. So this KDD Cup 1999 dataset is also suitable candidate to detect and to cover all kinds of attack classes on its evaluation. The details of usefulness ok KDD Cup 1999 dataset are explored in more depth in the next section. Existing IDSs also provides the evaluation based on such complementary collection of network traffic data. Take preference of such assessment, the accompanying two IDSs are chosen (Ciza Thomas (2009)):

1. PHAD is used to detect attacks by extracting the packet header information.
2. Snort collects data from both the header and the payload piece of each parcel on time based and on connection based way.

Packet Header Anomaly Detector (PHAD) that screens the 33 fields of the Ethernet, TCP, UDP and ICMP conventions is picked as one of the IDSs for the blend. Watching the header fields makes it effective to distinguish Probes what's more DoS attacks with low false alarm rate.

Snort is widely used an open source system interruption counteractive action and recognition framework focused around rule driven dialect, which gives the profits of mark, convention and peculiarity based location. As of now, Snort is the most generally sent interruption location and counteractive action innovation overall and has turned into the true standard for the security business. Grunt is proficient in catching the DoS assaults furthermore the U2R assaults with high location rate capabilities. So, it is suitable candidate to compare with the proposed IDS system. Since, the proposed system is constructed for intrusion detection as well as for intrusion prevention as like Snort.

### 3.1 Data Set

We provide simulations and experimental results on the real-world movement information and the KDD Cup 1999 information set which is originated from DARPA 1998 dataset in order to illustrate performance of the proposed system proposed in the parts three to five of this postulation. The fundamental purpose behind utilizing the KDD Cup 1999 information set is that we require significant information that can without much of a stretch and be imparted to different specialists, permitting them to assess, reuse. In general, intrusion recognition frameworks to claim great execution with real time movement make it hard to confirm and enhance past examination results, as the activity is never discharged and given to others of particular network of any organization because of privacy issues. We utilized both the KDD Cup 1999 data sets and the real world activity information collected from the local network. So, it is possible to compare and contrast the results against work based solely on the KDD Cup 1999 dataset with existing critics, and still allow us to compare the proposed work directly. As mentioned earlier, the KDD Cup 1999 dataset is proposed using DARPA 1998 dataset which is off-line evaluation data set generated by DARPA/Lincoln Laboratory and thus, comprehensive information set that can be utilized for IDS assessment. It is worth to dissect the bad marks furthermore qualities for such a basic assessment.

### 3.2 Usefulness of KDD Cup 1999 Data Set for IDS Evaluation

The KDD cup data set is originated based on the DARPA data set and mainly, it was used to construct a network intrusion detection as identification model which is able to cater between “bad” connection records, called attacks, and “good” called as normal records from the given input data. The section 6.5.1 gives the brief introduction about KDD Data Set includes its contents, number of test data records details provided, training data set details and the major attack class supported by this data set for known as well as for unknown attack detection. Further, the section 6.5.2 comments on the criticism against the KDD Cup 1999 Data Set due to its oldness and redundant data contents and section 6.5.3 describes the supporting facts for this data set to realize the usefulness of this data set still in today’s intrusion detection field, though it was developed and made publicly available since 1999.

### 3.3 Supporting Facts of the KDD IDS Evaluation Data Set

Another data set found and used by Ciza Thomas (2009) used for IDS assessment other than KDD Cup 1999 Data Set is the Defcon Catch the Flag (CTF) information set acquired from Defcon competition and tradition directed yearly. This information set has numerous properties that make it altogether different from this present reality system movement data expected by user. The differences lie with extremely high volume of assault movement, the foundation activity is occupied, also the not very many number of IP addresses are available in data set. So, other missing network traffic details are in this dataset give the opportunity and the initial decision to use DARPA data set as first choice for IDS performance evaluation.

Also, while handling the real data traffic, it is observed that there is always the lack of the data needed about the status of the activity. Additionally, it is not possible to get 100 % accurate results with intense analysis the predictions used in IDS performance evaluation because of changing, unpredictable behavior of attackers along with attack

https://doi.org/10.15623/ijret.2018.0710003 Received: 24-07-2018, Accepted: 06-09-2018, Published: 27-09-2018

sophistication methods used and involves high cost solution to deal with such real time network traffic data. Even, it is not possible to get proper detection rate or other evaluation metrics for comparisons in the real-time network traffic data.

The researchers Mahoney and Chan (1999) reported that if any progressive IDS proved unable perform and evaluate this data set accurately, then it could likewise not perform acceptably well on sensible information caught on the web. In this way, before discussing the demerits of KDD Cup 1999 Data Set for IDS evaluation, one can evaluate the IDS perform well with this data set and it covers all the attack spectrum and attack classes of this data set.

**Table 1:** List of Intrusions in *training_attack_types* File of KDD Cup 1999 Data Set

| sl no. | Attack Type | Attack Class | Sr. No. | Attack Type | Attack Class |
|---|---|---|---|---|---|
| **1** | back | DOS | **12** | buffer_overflow | U2R |
| **2** | ftp_write | R2L | **13** | guess_passwd | R2L |
| **3** | Imap | R2L | **14** | ipsweep | PROBE |
| **4** | land | DOS | **15** | loadmodule | U2R |
| **5** | multihop | R2L | **16** | Neptune | DOS |
| **6** | nmap | PROBE | **17** | Perl | U2R |
| **7** | phf | R2L | **18** | pod | DOS |
| **8** | portsweep | PROBE | **19** | rootkit | U2R |

## 4. NUMERICAL RESULT AND DISCUSSION

Fig. 1 gives the data preprocessing analysis of 10 % KDD Cup 1999 data set with respect to each attack class and their types mentioned in training file of intrusions or attacks as shown in Table 1. From Fig. 1(a), it is observed that after removing the duplicate instances by applying data preprocessing, in 10 % KDD Cup 1999 Data set, the number of records initially available were 4,93,967 comprising attack oriented and normal traffic records. After applying preprocessing, it is reduced for 1, 45,687 distinct records. Out of these, 1, 45,687 records, the attack prone records were 57,856 and normal traffic records were 87,831 as indicated in Weiss, Christian (2012) Fig.1(b). Fig. 1 (c), (d), (e) and (f) shows variation in number of records after data preprocessing of attack prone records which are further divided into specific attack class and their respective attack types namely, *DoS, Probe, U2R* and *R2L.*

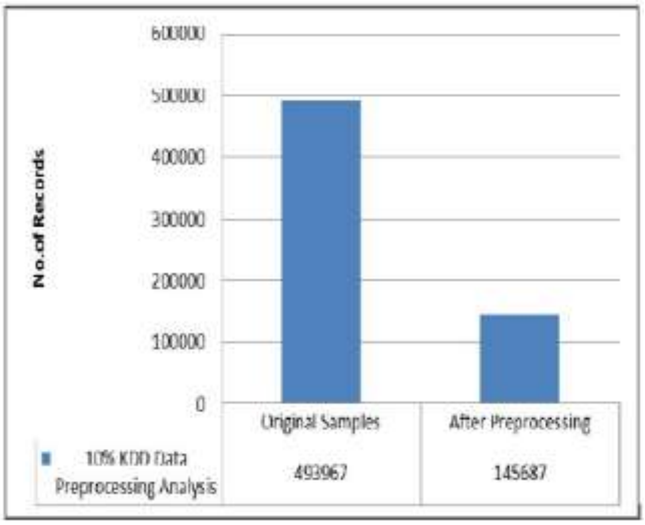

**Fig. 1(a):**10% KDD Cup 1999 Data Set Preprocessing for All Connection Records

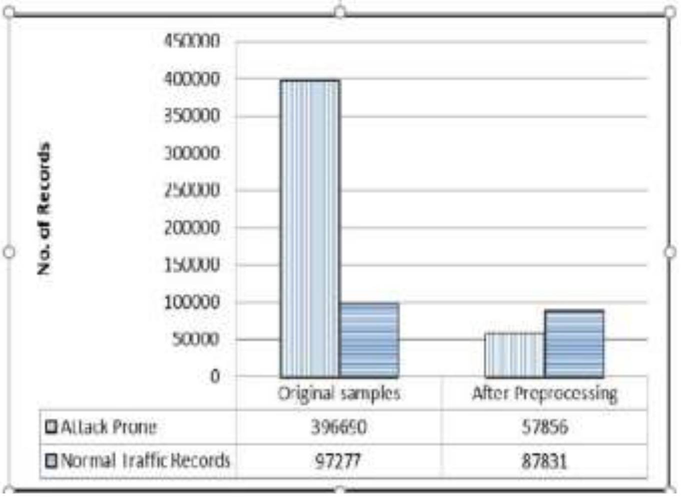

**Fig. 1(b):**10 % KDD Cup 1999 Data Set Preprocessing for Attack Prone and Normal Traffic Records

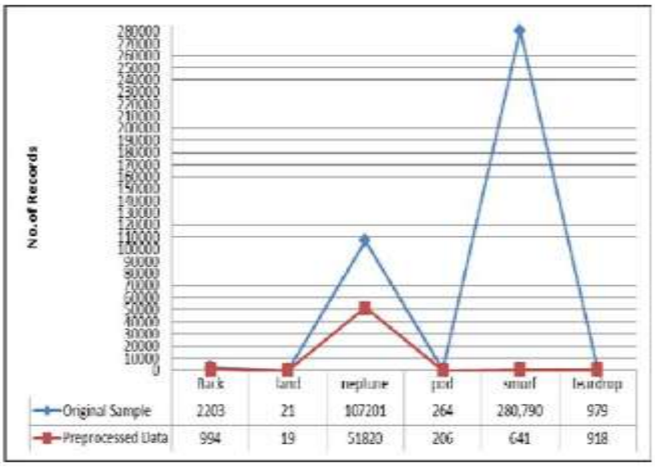

**Fig. 1(c):** Variation in *DoS* Attack Traffic Records after Preprocessing

https://doi.org/10.15623/ijret.2018.0710003 Received: 24-07-2018, Accepted: 06-09-2018, Published: 27-09-2018

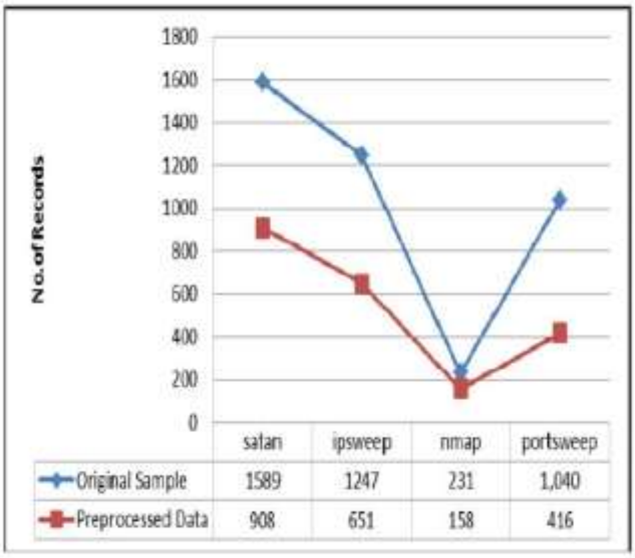

**Fig. 1(d):** Variation in *Probe* Attack Traffic Records after Preprocessing

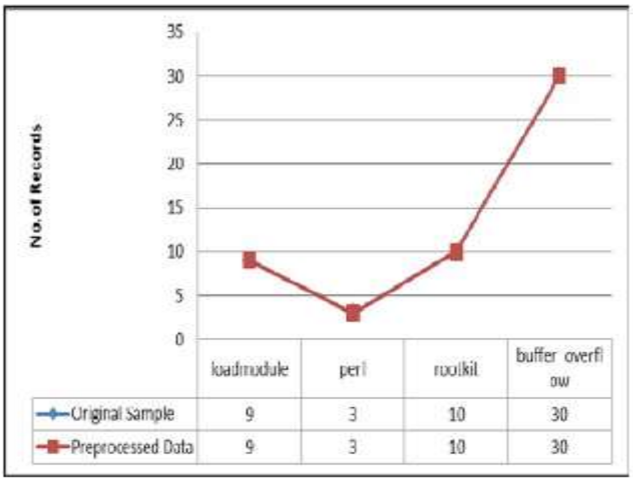

**Fig. 1(e):** Variation in *U2R* attack Traffic Records after Preprocessing

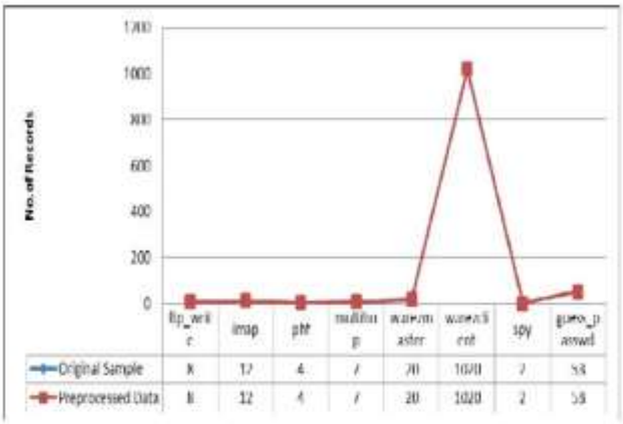

**Fig. 1(f):** Variation in *R2L* attack Traffic Records after Preprocessing

The proposed IDS has divided this 10 % KDD Cup 1999 dataset into the four different samples consisting of 1000,5000,10,000 and 1,00,000 records of constructed input file for the proposed IDS system evaluation from the original 10 % KDD Cup 1999 Data set. Based on the preprocessing module implemented as part of proposed advanced IDS system, it performed the preprocessing of these newly created data set files as per the following preprocessing time taken by each of these sample files and the percentage of data reduction is calculated based on the available connection records in each of the preprocessed file. The result is obtained on core i5 Intel series processor machine on windows operating system and it is clear from the obtained result, as the number of connection records increased, the time taken for preprocessing is increased as shown in Fig. 2 and the percentage of data reduction is also increased highly with number of records increased at large level as shown in Fig.3.

Before evaluating the performance of proposed advanced IDS, the selected other IDSs as per the test setup prepared for proposed system evaluation are examined separately on the KDD Cup 1999 test data set. Initially, the IDS Snort which performs the signature based intrusion detection was evaluated with this data set and obtained results are shown in Table 6.7. It is noticed that as shown in Table 6.7, few attacks for specific attack classes are detected whereas few are not detected by this selected IDS Snort, so they are appeared in both the classes of detection *i.e.* detected attacks and undetected attacks.

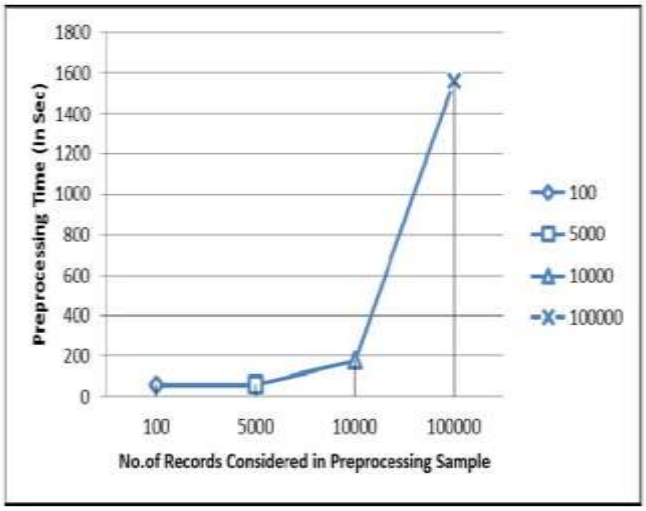

**Fig. 2:** Preprocessing Time of 1000,5000,10000 and 100000 Records of10 %KDD Data

https://doi.org/10.15623/ijret.2018.0710003 Received: 24-07-2018, Accepted: 06-09-2018, Published: 27-09-2018

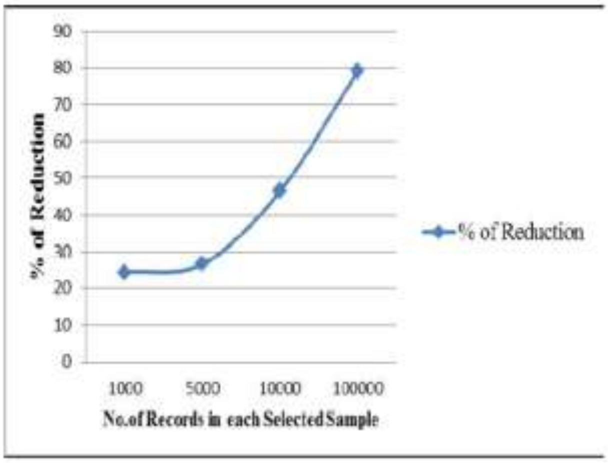

**Fig. 3:** Data Reduction in 1000,5000,10000 and 100000 Records of10 %KDD Data Set

## 5. CONCLUSION

During the study, it is found that, one parameter is more useful and gives the better results than other specified parameters, and then in such conflict cases, the non-probability based parameter like F-Score can be used rather than conducting evaluation based on such test parameters.

To break down and take care of the IDS assessment issues which are found in this proposal, assessment parameters, for example, detection rate, accuracy, review, and F-score have been represented in this area. Likewise, the P-test is introduced to compare two IDSs in innovative way comparing with other parameters as stated by Ciza Thomos (2009). Based on this study, measurements utilized for proposed IDS assessment like Precision, Recall, F-score what's more P-test are observed efficient to do evaluations of IDSs.

## REFERENCES

[1] M. A. Aydin, et al., "A hybrid intrusion detection system design for computer network security”, *Computers& Electrical Engineering*, vol. 35, pp. 517-526, 2009.

[2] Snort. The open Source network intrusion detection system [Online]. Available: http://www.snort.org

[3] S. Benferhat and K. Tabia, "Integrating Anomaly-Based Approach into Bayesian Network Classifiers," e-Business and Telecommunications, pp. 127-139, 2009.

[4] T .Elvis, et al., “A serial combination of anomaly and misuse IDSesapplied to http traffic”, *Proceedings of the 20th Annual Computer Security Applications conference*, 2004.

[5] N. Hubballi and V. Suryanarayan. “False alarm Minimization technique in signature-based intrusion Detection system”, *A Survey Computer Communication* Vol. 49, pp 1-17, 2014.

[6] X. Cui and F. Wang, "An Improved Method for K-Means Clustering," *in2015 International Conference on Computational Intelligence and Communication Networks (CICN*), pp. 756-759, 2015.

[7] F. H. M. Ali, Z. K. Safwan, Mawlood Hussein, "Evaluation Effectiveness of Hybrid IDS Using Snort with Naïve Bayes to Detect Attacks,*" Digital Information and Communication Technology and it's Applications (DICTAP)*, pp.256-260, 2012.

[8] M. L. Laboratory. DARPA dataset Available: http://www.ll.mit.edu/(2011).

[9] KDDCup99 Dataset, Available at http://kdd.ics.uci.edu/databases/kddcup99/.html. 1999

[10] A. M. Zainal, M.A.; Shamsuddin, S.M., "Feature Selection Using RoughSet in Intrusion Detection," presented at the 2006 IEEE Region 10Conference, 2006.