# Soft set based Approach for Intrusion Detection using Data Mining Techniques

**Lokendra Singh Parihar**[1] ,**Akhilesh Tiwari**[2]
Department of CSE & IT
[1,2] Madhav Institute of Technology and Science Gwalior, MP

***Abstract-*** *The usage of internet is growing day-by-day, which has lead to the increase in concerns about the security of the internet. The Internet atmosphere has become untested and much complex. Systems of Enterprise networked are unavoidably exposed to the growing threats posed through hackers as well as malicious consumers internal to a network. Technology of IDS is one of the most significant tools used present days, to counter such threats. Detection of Intrusion has become a network administration serious component because of the huge various attacks determinedly threaten our computers. Classical method of intrusion detection are limited and do not give a full problem solution. They search for possible malicious activities on the traffics of network; they sometimes succeed to discover attacks of true security and anomalies. However, in the numerous cases, they fail to identify behaviors of malicious (false negative) or they are fire alarms when nothing wrong in network (false positive). This paper show a hybrid data mining method encompassing feature selection, classification, filtering. A method for classifying the attribute type of the attacks applying soft set method and then applying classification to classify the attacks based on the selected attributes. The IDS is introduced for the efficient attacks identification to attain high detection and also accuracy rate as well as low false alarm rate.*

**Keywords**- Intrusion Detection System, Feature Selection, Classification, KDDCup'99, Detection Rate, False Alarm Rate

## I. INTRODUCTION

### A. Intrusion Detection System

Intrusion detection is the analyzing and monitoring procedure the data and events happening in the network and/or system of computer in order to the attacks detect, vulnerabilities and also other different security issue [1]. IDS can be categorized data sources according into: detection of network-based and also detection of host-based. Detection of host-based, OS processes and data files of the host are directly monitored to define precisely which host resources are the specific attack targets. In contrast, detection systems of network-based monitor network traffic data applying sensors set attached to the network to capture any malicious activities. Networks security issue can vary extensively and can affect various security need including availability, authentication, integrity, and authorization. Intruders can cause various kinds of attacks for example DoS, scan, compromises, and viruses and worms[2,3].

**Techniques for Intrusion Detection:**

Each malicious activity or attack has a specific pattern. The patterns of only some of the attacks are known whereas the other attacks only show few deviation from normal patterns. Therefore, the techniques used for detecting intrusions are based on whether the attacks patterns are known or unknown. The two main techniques used are:

**A. Anomaly Detection:** It is based on the assumption that intrusions always reflect some deviations from normal patterns. The network normal state, breakdown, protocol, traffic load and packet size are defined by the system administrator in advance. Thus, anomaly detector compares the current state of the network to the normal behavior and behavior of malicious looks. It can detect both unknown and known attacks.

**B. Misuse Detection**: It is based on the knowledge of known patterns of previous attacks and system vulnerabilities. Misuse detection, always associates present activity to identified patterns of intrusion to confirm that any attacker is not attempting to exploit known vulnerabilities. To accomplish this task, it is need to define all intrusion pattern in the detail. It cannot detect unknown attacks[4].

**Table 1: Techniques of Intrusion Detection**

| | **Misuse Detection** | **Anomaly Detection** |
|---|---|---|
| **Characteristics** | Use patterns of well-known attacks (signatures) to identify intrusions. Any match with signatures is reported as a possible attack | Use deviation from normal usage patterns to identify intrusions. Any significant deviations from the expected behavior are reported as possible are reported as possible attacks |

| | | |
|---|---|---|
| **Drawbacks** | - False negatives<br>- Unable to detect new attacks<br>- Need signatures update<br>- Known attacks has to be hand-coded<br>- Overwhelming security analysts | - False positives.<br>- Selecting the right set of system features to be measured is ad hoc and based on experience<br>- Has to study sequential interrelation between transactions<br>- Overwhelming security analysts |

Between these two approaches, [5,6,7] only anomaly detection has the ability to detect unknown attacks, since misuse detection can only detect intrusions which contain known patterns of attack.

**B. Feature Selection**

It is based on the soft set method. Soft Sets represent a powerful tool for decision making about information systems, data mining and drawing conclusions from data, especially in those cases where some uncertainty exists in the data. Its efficiency in allocating with uncertainty issue is as a result of its parameterized concept. Recently, various researches had been done various works in theory and in practices. We recall some basic soft set theory notion introduced by Molodtsov (1999) and some useful definition from Maji et al., (2002; 2003). Here, □ to be an initial universal set and □ to be a set of parameters and □, □⊂□ [Pal &Mondal, 2011].

Definition 2.1 (Soft Set) A pair $(F, E)$ is called a soft set (over $U$) if and only if $F$ is a mapping of $E$ into the set of all subsets of the set $U$. In other different words, the soft set is a parameterized subsets family of the set $U$. Every set $(e)$, $e \in E$, from this family may considered as the e -approximate elements set of the soft set. Let us consider the following example.

. Example 2.1.1: A soft set $(F, E)$ describes the attractiveness of the bikes which Mr. X is going to buy [Pal &Mondal, 2011].

$U$ is the set of bikes under consideration. $E$ is the set of parameters. Each parameter is a word or a sentence.

E = (e1= stylish; e2 =heavy duty; e3= light; e4 =steel body; e5 = cheap; e6 = good mileage; e7 =easily Started; e8 = long driven; e9=costly; e10 = fibre body)

In this case, to define a soft set means to point out stylish bikes, heavy duty bikes, and so on.

Example 2.1.2 : Let $U = \{u1, u2, u3, u4, u5\}$ be a universal set and $E = \{x1, x2, x3, x4\}$ be a set of parameters. If $A = \{x2, x3, x4\}$ and then the soft set $FA$ is written by $F_A$={(x 2,{u 2,u 4}),(x 4,U )} .

Definition 2.2 (Operation with Soft Sets)
Suppose a binary operation denoted by *, is defined for all subsets of the set $U$. Let $(F, A)$ and $(G, B)$ be two soft sets over $U$. Then the operation * for the soft sets is defined in the following way: $(F, A) * (G,) = (H, A \times B)$ Where $( ,\beta) = F(\alpha) * G(\beta)$, $\alpha \in A$, $\beta \in B$ and $A \times B$ is the Cartesian product of the sets $A$ and $B$.

Definition 2.3 (NOT Set of a Set of Parameters) Let E= {e1,e2,e3,…en} be a set of parameters. The NOT set of E denoted by ⌝$E$ and is defined by ⌝$E$ = {⌝1 ,⌝$e$2 , ⌝$e$3 ,... ⌝$en$ } where ⌝$ei$= not $ei$ for all $i$. It may be noted that ⌝ and ⌝ are two different operations.

Definition 2.4 (Complement of a Soft Set) The complement of a soft set (F, A) is denoted by (F, A) c and is defined by (F, A) $^c$ = ($F^c$ , ⌝A) where $F^c$ : ⌝A → P(U) is a mapping which is defined by $F^c$ (α) =U - F(⌝α), for all α∈⌝A.

Definition 2.5 (Relative Complement of a Soft Set) The relative complement of a soft set (F, A) is denoted by (F, A)r and is defined by (F, A)r = (Fr , A) where F r :A→P(U) is a mapping given by F r (α)=U-F(α), for all α∈A.

Definition 2.6 (NULL Soft Set) A soft set (F, A) over U is said to be a NULL soft set denoted by Φ, if for all ε ∈ A, F (ε) = ϕ (null-set).

Definition 2.7 (Relative NULL Soft Set) A soft set $(F, A)$ over $U$ is said to be relative NULL soft set with respect to parameter set A denoted by Φ$A$ if ε ∈ A, F(ε) =ϕ (null set).

Definition 2.8 (Relative Whole Soft Set) A soft set $(F, A)$ over $U$ is said to be relative whole soft set (with respect to parameter set $A$) denoted by $UA$, if for all ε ∈ A, F (ε )=U.

Definition 2.9 (Absolute Soft Set) The relative whole soft set $(E)$ with respect to the universe set of parameters $E$ is called the absolute soft set over $U$.

Definition 2.10 (AND Operation on Two Soft Sets) If $(F,A)$ and $(G, B)$ be two soft sets then $(F, A)$ AND $(G, B)$ denoted by $(F,A) \wedge (G,B)$ and is defined by $(F, A)(G, B) = (H,A \times B)$ where $H(\alpha, \beta) = F(\alpha) \cap G(\beta)$ for all $(\alpha, \beta) \in A \times B$.

Definition 2.11 (OR Operation on Two Soft Sets) If $(F,A)$ and $(G, B)$ be two soft sets then $(F, A)$ OR $(G,B)$ denoted by $(F,A)$

$\vee$ $(G,B)$ is defined by $(F, A) \vee (G,B) = (O, A \times B)$ where $O(\alpha, \beta) = F(\alpha) \cup G(\beta)$ for all $(\alpha, \beta) \in A \times B$.

## II. RELATED WORK

Soft set theory was first introduced by Molodtsov in [8] as a new paradigm for mining uncertain data. Soft sets overcome the inadequacy of other techniques such as interval mathematics, fuzzy set theoriesand probability. Pei and Miao [9] explored information systems and soft sets in terms of relationship between them. The results of their experiments reveal that information systems and partition-type soft sets share a common formal structure. For instance fuzzy information systems and fuzzy soft sets are equal. Razak and Mohamad [8] proposed a group decision creating technique with criteria based on soft set based data mining method. The weight of each criterion is computed using a method known as AHP. The problem of group decision is solved using soft max – min decision making method.

Chetia and Das [10] extended Biswas's method for evaluation of answer scripts of students. They assumed five satisfaction levels in order to evaluate the performance of students. They include unsatisfactory, satisfactory, good, very good and excellent. They have developed an algorithm that takes student's statistics as input and build a soft set matrix before evaluating the performance of students.

Herawan et al. [11] presented an approach to reduce dimensionality of soft set. The existing solutions on soft set are Boolean – based. However, they may also have non-Boolean values. In case of multi-valued information systems, they presented an alternative approach for reducing attributes. They introduced the ideal of multi soft sets that are constructed from multi-valued information systems. Then they also used OR and AND operators on soft sets. They came to know from the experiments that the set of attributes (reduct) required in soft set theory are also same as that of rough set theory. The reduct approach was first introduced by Maji et al. [9]. They used it for decision making in soft set mining applications. Parameterization reduction is also possible in soft sets and related applications as presented by Chen et al. [12]. They further said that the approach followed by Maji was incorrect and also claimed that the reduct is not same for theory of soft set and also theory of rough set. Their idea for reduction of attributes in soft sets was based on the optimal choice concept that addresses the problems of sub-optimal solutions. This problem was also analyzed by Kong et al. [13] and defined actualparameter reduction that can overcome the problems of sub-optimality.

For decision making applying soft set where there is data deficiency Zou [13] proposed a novel technique. This technique is based on computation of weighted average as per the distribution objects.

XunGe and Songlin Yang [14] investigated operations on soft set. They explored the operations defined in the prior works. The results of their work help others to choose right operators and operational rules while working with soft sets. Rose et al. [16] Proposed two techniques to compare incomplete datasets. The techniques are based on aggregate and calculated support values and parity bits of supported set. When a dataset is downloaded or taken from a source, it might be an incomplete dataset due to VIRUS attacks or any software or hardware problems. As the processing of incomplete datasets will yield inconsistent results, it is essential to know whether the data sets are complete or incomplete prior to the applying them in data mining algorithms. The results of comparison help in finding missing attributes and take necessary steps to rectify the problems before actually processing the data.

Rajpoot et al. [15] proposed an association rule mining based on soft set approach using constraints with respect to initial support. The constraint is meant for filtering rarely occurred items and false frequent items. As the pruning reduces search space, the dataset is improved and it consumes fewer resources to mine association rules. Afterwards, the dataset is converted to Boolean – valued information system. The resultant dataset is known as soft set.

An approachof hybrid learning [17] through applying a combination of classification of naive bayes and K-means, cluster each data into the corresponding collection before using a classifier for purpose of classification. A system of hybrid anomaly detection [14] was proposed which combine two different classifiers and k-means: naive bayes and k-nearest neighbor. Firstly, it achieves the feature selection procedure from the of intrusion detection applying an entropy based feature selection algorithm which selects the significant attributes and eliminates the redundant attributes. The another level is cluster formation applying k-Means and then it additional classifies them through applying a hybrid classifier.

## III. METHODOLOGY

This part contains two main data mining method discussants: soft set method and J48Graft for classification approaches. The soft set approach helps in the feature selection process and reduce the attributes. Soft set in the literature has been extensively used algorithm because of its effectiveness measure. While the J48Graft algorithm in a classification not

only gives efficient classification results, but also contain the tree pruning with fast decision learning capability.

**A. Soft set Approach:** Soft set is a parameterized common mathematical tool which deals with an approximate descriptions set of the objects. All approximate description has two different parts, a predicate and an approximate value set. In the mathematics, a mathematical model of an object is created and describe the exact solution notion of this model. Commonly mathematical model is too complex and the exact solution is not simply obtained. So, the approximate solution notion is presented and the solution is calculated. In the theory of soft set, we have the opposite method to this problem. The first object description has an approximate nature, and we do not necessity to the exact solution notion present. The any restrictions absence on the approximate description of the soft set theory creates this theory most convenient and simply applicable in the practice. Any parameterization we prefer can be used with the words and sentences help, mappings, real numbers, functions and so on.

Soft sets could be regarded as neighborhood systems, and they are a particular context-dependent fuzzy sets case. In the theory of soft set the setting problem the membership function, among other different related problems, easily does not arise. This produces the theory most convenient and easy to using in practice.

**B. J48-Graft Algorithm:** J48-graft algorithm generates a grafted decision tree from a J48 tree algorithm. The grafting technique is an inductive procedure that enhances nodes to inferred decision trees. The grafting technique is an inductive procedure that adds nodes to the inferred decision trees with the purpose of reducing prediction errors. The J48-graft algorithm classify region of the multidimensional space of attributes not occupied through training examples [9]. This procedure is established to frequently progress predictive correctness. Special analysis might propose decision tree grafting is direct pruning reverse. To the contrary, it is argued that the two different procedures are complementary. This is because, for example standard tree rising methods, pruning uses only local knowledge, whereas grafting utilize non-local knowledge. The both pruning use and so grafting in the conjunction is demonstrated to gives the best common predictive accuracy over a representative learning tasks selection [18].

## IV. PROPOSED WORK

This section describes the system architecture for IDS based on the hybrid data mining methods.

A filter method is proposed for feature selection technique to reduce the noise and isolated points on the data set. It calculates the reducts based on the pairs that are being find out. Thus, helps in removing the redundant data.
We can easily divide our work in two phases:

1. Phase 1: Soft set approach for feature selection and removing the outliers.
2. Phase 2: Classification of the reducts that have been chosen from the dataset. After applying the classification, intrusions are classified and tested on various factors.

Fig. 1 depicts the system architecture for intrusion detection. It consists of selection of feature, filtering, classification, divide and merge, classification ensemble and normal and intrusion detection.

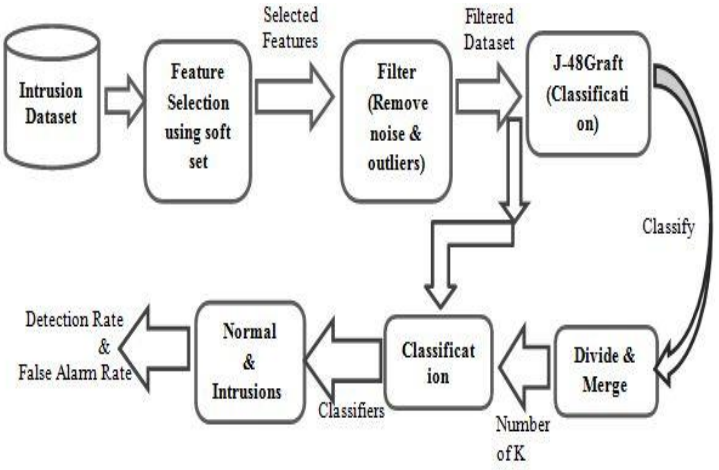


**Fig. 1: The System Architecture For Intrusion Detection**

**Description of the Proposed Model**

**KDD CUP 99 Data Set:** Since 1999, KDD'99 has been most widely used data set for determine of the detection of anomaly. This data set is organized by stolfo et al and is built based on the data captured in DARPA'98 IDS evaluation program. DARPA'98 is about 4 gigabytes of compressed draw (binary) tcpdump data of 7 weeks of network traffic, which can be processed into about 5 million connection records, each with about 100 bytes. The two weeks of test data have around 2 million connection records.

**The 10% of KDDCup'99 Data set**

Since the KDDCup'99 Data set consists of large amount of data records for the training and testing of IDS, trouble is faced for the analyzing dataset as a whole. Henceforth the

dataset utilized for the analysis of the proposed models is only 10% of the original dataset.

Kddcup,99 consist of two variations of training dataset; one is a full training set having five million connections and the another is 10% of this training set having 4942021 connections. The testing dataset is also having tow variations one is testing set with 311029 connections and the other is corrected test set which also have 311029 with class label.

This 10% KDDCup,99 dataset similar to the original KDDCup'99 consists of all 41 attributes along with the additional class label as the 42nd attribute, and 22 types of attacks. The applicability of the data usage depends upon the user who performs analysis of IDS that whether the 10% of the dataset is quite enough for the analysis or not.

The only difference between the whole KDD, 10% KDD and Corrected KDD datasets is of the total number of instances present for IDS. The vrience between the training and the corrected set of KDD is that the training set just comprises of the 22 types of attacks while the corrected set comprises of 39 types of attacks including 22 types of exiting attacks and 17 types of new attacks.

Pre-processing Phase: In the pre-processing phase, 10% of KDD dataset has been improved in order to create the classification simpler. The dataset here is preprocessed through classifying the prediction class into 4 various kinds attacks i.e. dos, probe, r2l and u2r. This categorization hence would help in creating the evaluation easier.

Dataset Splitter: In the dataset splitting phase, the KDDCup'99 dataset splits into two different parts: one is the set of training and the other is the set of test. The train set and test set is splited in the 3:2 ratio (i.e. the train set is the 66% and test set is the 34% of KDDCup'99 dataset) through the splitter. Splitter opt instances randomly for the model training from original dataset while the rest is presented for trained model testing. Therefore KDDCup'99 dataset utilize in the model contain 494021 instances which get separated randomly into the 326054 for training instances while remaining 167967 for testing instances.

Feature selection is significant if the data set contains various attributes. It selecting features include applying an knowledge gain feature selection technique which selects the significant attributes from the data set. A filter technique is proposed to decrease the isolated and noise points on the data set. After using filtering, initially the classification is done using J48Graft decision tree. The classification of data is done which helps in classifying the data available. The classification creates a wider range of data for processing thus, helping in making results better.

**Performance Evaluator:** The Performance evaluator phase assesses the REP based IDS model performance through calculating the following parameters:

a) True Positive Rate (TPR):

$$TPR = \frac{TP}{TP + FN}$$

b) False Positive Rate (FPR):

$$FPR = \frac{FP}{TN + FP}$$

Where, TP (True Positive), FN (False Negative), FP (False Positive) and TN (True Negative) can be defined as follows [10]:

• True Negative (TN): The percentage of valid records that are correctly classified.
• True Positive (TP): The percentage of attack records that are correctly classified.
• False Positive (FP): The percentage of records that were incorrectly classified as attacks whereas in fact they are valid activities.
• False Negative (FN): The percentage of records that were incorrectly classified as valid activities whereas in fact they are attacks.
These parameters described above can also be illustrated through Table I.

**Table 1: Confusion Matrix of TN,TP, FN and FP**

| | **Correctly Classified** | **Incorrectly Classified** |
|---|---|---|
| **Valid Record** | True Negative(TN) | False Positive(FP) |
| **Attack Record** | True Positive(TP) | False Negative(FN) |

Confusion Matrix is one of the other different parameters in the literature to the analyze the model performance. A confusion matrix is a tabular visualization of the algorithm performance. The column in the matrix represents the prediction class instances while the row represents the actual class instances.

**Visualization:** In this phase the presentation REP based IDS model results can be visualized by several means for example graph, text etc. On the results obtained basis in this phase, the model efficiency can be examined.

## V. SIMULATION ENVIRONMENT

We have done experiments on base and proposed algorithms. KDD cup'99 data set is used as dataset for anomaly detection. The simulation environment used here is MATLAB(Matrix Laboratory) and WEKA (Waikato Environment of knowledge analysis) is a most popular suite of machine learning software tool in data mining research field which written by java, developed at the university Waikato, New Zealand. MATLAB and WEKA is open source software available under GNU General Public license. All experiments were performed well and fully on Dell workstation with 4 GB RAM and 32-bit operating system, running windows 7.

## VI. RESULTS ANALYSIS

This part shows the experimental results find from soft set-J48Graft based model of IDS along with its comparison with the K2 based IDS method. The two different algorithms when compared, it has been experiential that more suitable outcomes are obtained through using soft set-J48Graft. Since both algorithms consequently results in the data classification as normal or the attack type accordingly, the taking time for data estimating is nominal in the soft set-J48Graft case. Also, the accuracy obtained for soft set-J48Graft is raised appreciably than K2. Hence it would not be irrelevant to say that soft set-J48Graft has proved itself to be more better classification method than K2.

Table II shows the comparison of TPR & FPR between K2 and Softset-J48Graft

| Class | K2 | Softset-J48Graft | K2 | Softset-J48Graft |
|---|---|---|---|---|
| | TPR | TPR | FPR | FPR |
| DoS | 0.988 | 1.000 | 0.000 | 0.000 |
| Probe | 0.978 | 0.984 | 0.005 | 0.000 |
| R2L | 0.959 | 0.971 | 0.001 | 0.000 |
| U2R | 0.810 | 0.583 | 0.005 | 0.000 |
| Normal | 0.985 | 0.999 | 0.002 | 0.000 |

Table III contains the confusion matrix of the attacks classified by Softset-J48Graft :

| | Normal | U2R | DoS | Probe | R2L |
|---|---|---|---|---|---|
| Normal | 33086 | 0 | 7 | 14 | 12 |
| U2R | 3 | 7 | 1 | 0 | 1 |
| DoS | 7 | 0 | 133075 | 0 | 0 |
| Probe | 15 | 0 | 7 | 1353 | 0 |
| R2L | 10 | 0 | 0 | 1 | 368 |

Fig. 2 and Fig. 3 show the comparison number of correctly classified instances and incorrectly classified instances between K2 based IDS model and proposed Softset-J48Graft based IDS model respectively.

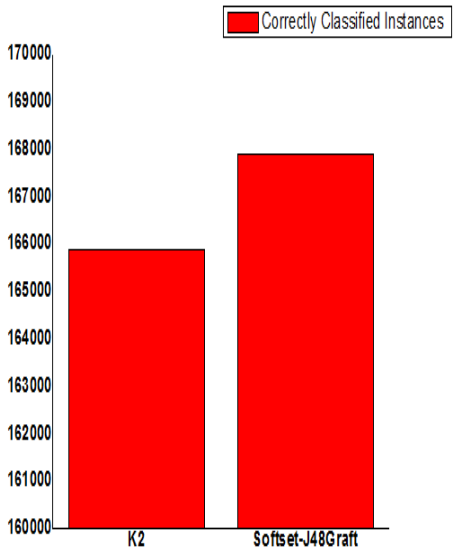


**Fig 2. Correctly Classified Instances**

The values obtained for correctly classified instances for J48Graft are 167889 while for K2 is 165873.
Also, for incorrectly classified instances, the values obtained for J48Graft is 78 while for K2 is 2094. These values have been shown with the help of graphs above.

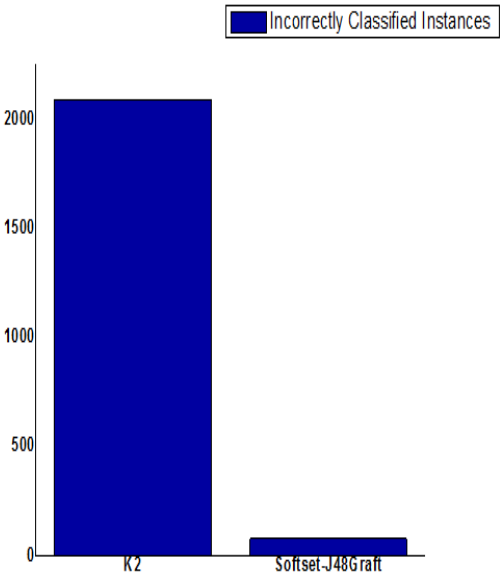


**Fig 3: Incorrectly Classified Instances**

The class wise comparison of accuracy between K2 based IDS model and proposed Softset-J48Graft based model are shown in Fig.4 to Fig.8. The values for accuracy for true positive rate are calculated in percentage. The proposed technique has better results.

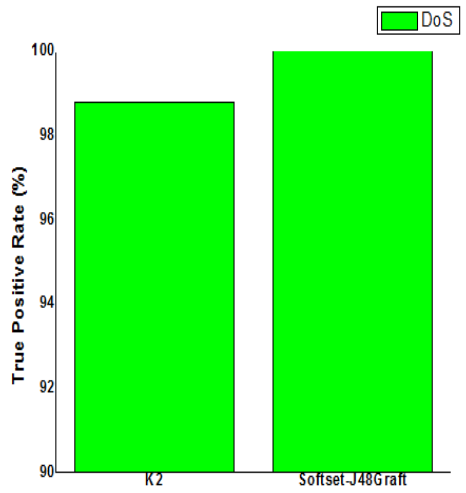


**Figure 4: DoS Attacks Detected by K2 and Softset-J48Graft**

It can be noticed than Softset-J48Graft has been extremely efficient in detecting each type of attacks than K2 and only lacks in effectively detecting U2r attacks. The TP rate obtained for J48Graft is 1.00 which is 100% while for K2 it is 98.8%.This has been shown in comparison graph.

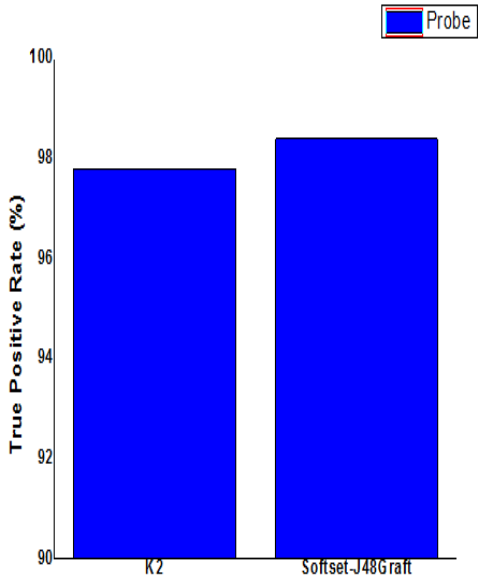


**Figure 5: Probe Attacks Detected by K2 and Softset-J48Graft**

The results for the various attacks are compared with the previously defined technique i.e. K2. The graphs below give a brief idea that how they are better in terms of TPR.

Fig 5 is the representation of probe attack comparison between K2 and J48Graft algorithm. The values of the probe attack for TP rate are compared in this graph. The values obtained are 98.4% for J48Graft while 97.8% for K2.

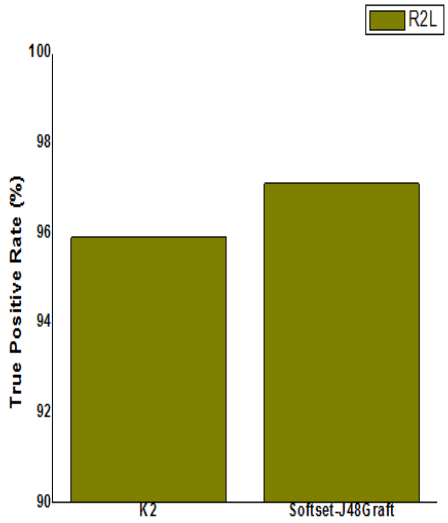


**Figure 6: R2L Attacks Detected by K2 and Softset-J48Graft**

Fig. 6 gives the brief comparison of the TP rate for R2L attack. The values are 97.1% and 95.9% for J48Graft and K2 respectively.

The attacks are firstly classified into groups and TPR are calculated.

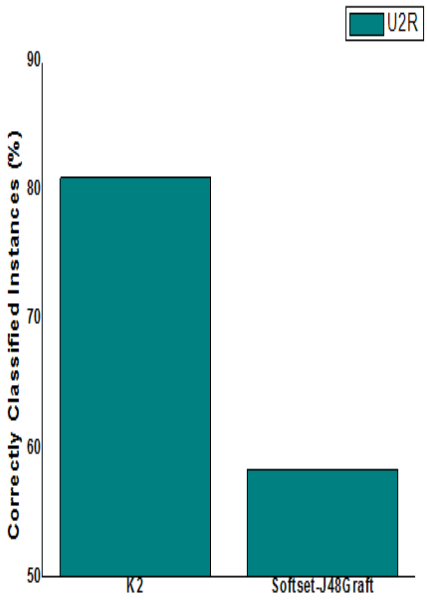


**Figure 7: U2R Attacks Detected by K2 and Softset-J48Graft**

Fig. 7 gives the brief comparison of the TP rate for U2R attack. The values are 58.3% and 81% for J48Graft and K2 respectively.

U2R attacks for the proposed algorithm are far less than the K2. The correctly classified instances for this attack should be less. Thus, it shows better results.

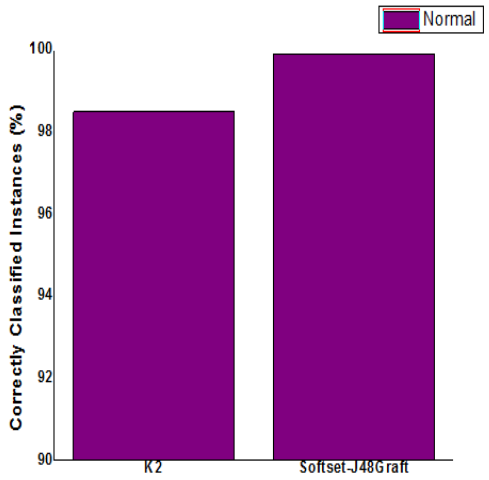


**Figure 8: Normal Attacks Detected by K2 and Softset-J48Graft**

Fig. 8 gives the brief comparison of the TP rate for Normal attack. The values are 99.9% and 98.5% for J48Graft and K2 respectively.

All the attacks have been compared above and have shown that the softest based J48Graft approach has proved to be a better approach in all aspects.

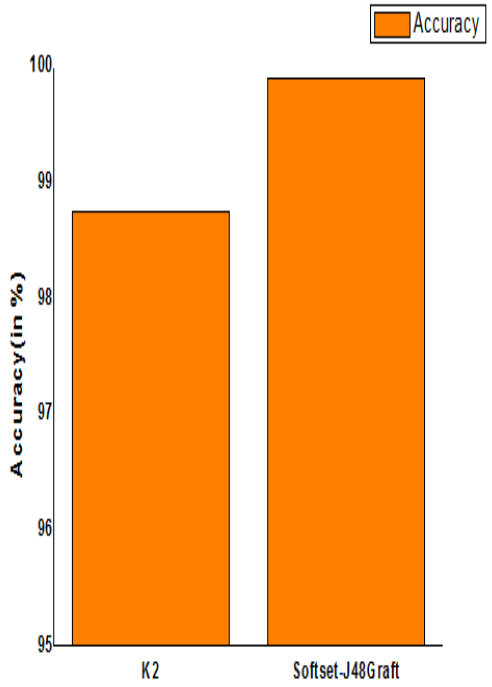


**Figure 9: Accuracy of K2 based IDS and Softset-J48Graft based IDS**

The comparison of overall accuracy between K2 based IDS and Softset-J48Graft based IDS models shows in Fig. 9. Again, Softset-J48Graft excels K2 with higher accuracy.

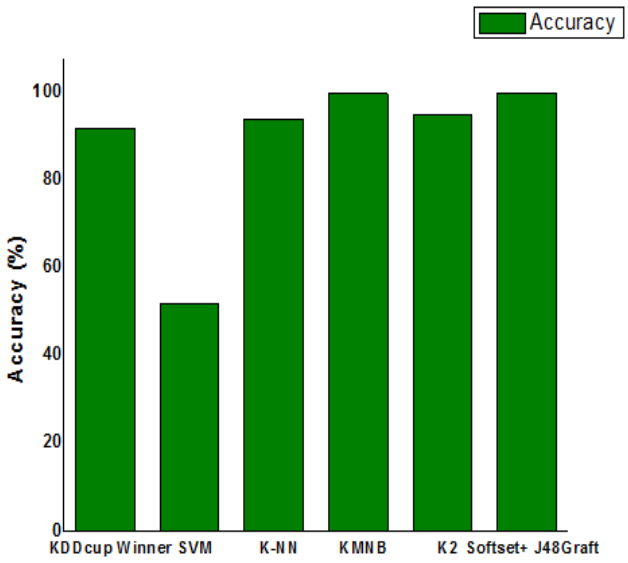


**Figure 10: Accuracy of various techniques based on IDS and Softset-J48Graft based IDS**

When the proposed Softset-J48Graft based IDS model is compared against the commonly preferred techniques of Data Mining which were used for intrusion detection, the Softset-J48Graft was found with the maximum accuracy result. The comparison can be visualized from Fig. 10:

## VII. CONCLUSION

The experimental analysis performed for intrusion detection has demonstrated the applicability of Softset-J48Graft based Model excelling the compared frequently used K2 algorithm. Also the comparative tables and graphs shows, Softset-J48Graft as the most efficient algorithm aiding in high accuracy and better detection in each type of attacks discussed in the paper. Also Softset-J48Graft along with greater classification performance leads to reduce the error rate by applying pruning on the dataset. Thus Softset-J48Graft is supposed one of the preferable algorithms of the future. The only shortcoming faced or a future scope can be expected in reducing the higher amount of data for performing pruning.

## REFERENCES

[1] Jiawei Han and. MichelineKamber, Data Mining: Concepts and Techniques, Morgan Kufmann, 2nd edition 2006, 3rd edition 2011.

[2] S.J. Stolfo, W. Lee. P. Chan, W. Fan and E. Eskin, "Data Mining – based Intrusion Detector: An overview of the Columbia IDS Project" ACM SIGMOD Records vol. 30, Issue 4, 2001.

[3] W. Lee and S.J. Stolfo, "Data Mining Approaches for Intrusion Detection" 7th USENIX Security Symposium, Texas, 1998.

[4] Conry-Murray, "Anomaly Detection On the Rise", June 2005, available on http://business.highbeam.com/787/article-1G1-132920452/anomaly-detection-rise-network-behavioranomaly-detection.

[5] Mazu Networks, "What You Can't See Can Hurt You: Ensuring Application Availability through Enterprise-Wide Visibility", November 2006. http://www.developertutorials.com/whitepapers/network-communications/

[6] Liebert, Chris, "Internal Threat Protection with Net-Based Detection, Prevention and Behavioral Systems", October 2006, http://www.mazunetworks.com/resources/analystreports/Internal_Threat_Protection_January_06.pdf.

[7] Enterprise Management Associates: Behavioral Analysis Enables a New Level of Network Security Awareness, technical White Paper, June 2004. http://security.ittoolbox.com/research/behavioralanalysis-enables-a-new-level-of-network-security-awareness-3755.

[8] D. Molodtsov (1999), "Soft Set Theory-First Results", Computers and Mathematics with Applications, Vol. 37, Pp. 19–31.

[9] P.K.Maji, A.R. Roy & R. Biswas (2002), "An Application of Soft Sets in a Decision Making Problem", Computers and Mathematics with Applications, Vol. 44, Pp. 1077–1083.

[10] P.K. Maji, R. Biswas& A.R. Roy (2003), "Soft Set Theory" Computer and Mathematics with Application, Vol. 45, Pp. 555–562.

[11] M. Pal & S. Mondal (2011), "Soft Matrices", African Journal of Mathematics and Computer Science Research, Vol. 4, No. 13, Pp. 379–388.

[12] D. Pei & D. Miao (2005), "From Soft Sets to Information Systems", Proceedings of the IEEE International Conference on Granular Computing, Vol. 2, Pp. 617–621.

[13] B. Chetia and P. K. Das, "Application of Vague Soft Sets in students' evaluation". Advances in Applied Science Research, 2011, 2 (6):418-423.

[14] TututHerawan,.Mustafa Mat Deris. (2009). A Direct Proof of Every Rough Set is a Soft Set. Third Asia International Conference on Modelling& Simulation.0 (0), p1-6.

[15] Zhi Xiao', Ling Chen', Bo Zhong', Shijie Ye.(2005). Recognition for Soft Information Based on the Theory of Soft Sets.IEEE. 0 (0), p1-3.

[16] Yan Zou, Yuke Chen. (0). Research on soft set theory and parameters reduction based on relational algebra.IEEE.0 (0), p1-5.

[17] XunGe and Songlin Yang, "Investigations on some operations of soft sets", World Academy of Science, Engineering and Technology 75 2011.

[18] VikramRajpoot, Prof. Shailendraku. Shrivastava, Prof. AbhishekMathur, "An Efficient Constraint Based Soft Set Approach for Association Rule Mining", International Journal of Engineering Research and Applications (IJERA).

[19] Ahmad NazariMohd. Rose, Mohd Isa Awang, Hasni Hassan , Mustafa Mat Deris, " Comparison of Techniques in Solving Incomplete Datasets in Softset", International Journal of Database Theory and Application Vol. 4, No. 3, September, 2011.

[20] Z. Muda, W. Yassin, M.N. Sulaiman, N.I. Udzir, "Intrusion Detection based on K-Means Clustering and Naïve Bayes Classification", In Proceedings of 7th International Conference on IT in Asia (CITA), IEEE, 2011. [11] Hari Om, AritraKundu, "A Hybrid System for Reducing the False Alarm Rate of Anomaly Intrusion Detection System", In Proceedings of 1st Int'l Conf. on Recent Advances in Information Technology (RAIT-2012),IEEE, 2012.

[21] D. Jachyra, K. Pancerz, and J. Gomula, "Classification of MMPI Profiles using Decision Trees" Concurrency, Specification and Programing, Poland, 2011, pp. 397.407.

[22] C.S. Trilok, J. Manoj, "WEKA Approach for Comparative Study of Classification Algorithm," International Journal of Advanced Research in Computer and Communication Engineering, 2013, pp.