Knowledgeable Handling of Impreciseness in Feature Subset Selection using Intuitionistic Fuzzy Mutual Information of Intrusion Detection System

P.Sudha, R.Gunavathi

Abstract: *One of the most promising areas of domain in research field is security because of its exponential usage in everyday commercial activities. Due to prevalence diffusion of network connectivity, there is a high demand for protection against cyber-attack which necessitates the importance of intrusion detection system as a significant tool for network security. There are many intrusion detection models available to classify the network traffic s either normal or attack type. Because of huge volume of network traffic data, these classifier techniques fail to attain high detection rate with less false alarms. To overcome the above problem, this paper introduces the potential feature subset selection model using Intuitionistic Fuzzy Mutual Information (IFMI). This model efficiently selects the optimal set of attributes without loss of information even in presence of impreciseness among attributes. This is achieved by representing each attribute in the dataset in terms of degree of membership, non-membership and hesitation. To validate the performance of the IFMI its reduced feature subset is used for classification using random forest classifier. After analyzing the feature subset, the simulation results proved that the proposed model has improved the performance of classifier for predicting the network intrusion attempts. It also helps the classification model to achieve high classification rate and reduced false alarm rate in an optimized way.*

Keywords: feature subset selection, Intrusion detection, impreciseness, intuitionistic fuzzy mutual information, random forest classifier.

I. INTRODUCTION

With the advancement in network and its related applications, serious security issues have arisen [1]. Cyber security plays an important role in protecting computers, data and networks from intruders [2].

Revised Manuscript Received on October 30, 2019.
* Correspondence Author

Mrs. P.Sudha*, Department of Computer Science, Sree Saraswathi Thyagaraja College, Pollachi - 642 107, Tamil Nadu, India. Email: sudha.sabariananth@gmail.com

Dr. R.Gunavathi,, Head, Master of Computer Science and Applications, Sree Saraswathi Thyagaraja College, Pollachi - 642 107, Tamil Nadu, India. Email: hodmca@stc.ac.in

It comprised of antivirus software, firewall and Intrusion Detection System (IDS). Intrusion detection involves in discovering unauthorized traffic, logins, destructions of data and abnormal behavior. Still the existing IDs suffer from inability to prevent attacks by themselves due to frequent occurrence of false alarms. Thus, data mining approaches are used to overcome these complexities which help in understanding the pattern of huge volume of network data [3]. Hence, Data processing plays an important role in intrusion detection. Feature subset selection which is also known as variable or attributes subset selection is a data processing model in machine learning and pattern recognition system [4]. By applying feature subset selection, it can influence accuracy and generalization abilities of classifiers and promotes the learning models with reduced data dimensionality while handling high dimensional network data processing.

In contrast to feature extraction which generates a new set of features from original data features, feature selection involves in selecting the best and most relevant subset of features from the available original data features. It is generally divided into three different cadres namely filter method, wrapper and embedded method. Filter approach chose most relevant and useful features from the original feature set which doesn't depend on model type. But wrapper method validates the selected feature subset using learning algorithms. Embedded approach integrates both filter and wrapper method [5]. The reputation of feature selection is to reduce the problem size and ensuing search space for learning algorithms. There are many models and approaches have bee tired out for feature selection this paper proposed an enhanced genetic mutual information-based feature subset selection whose ultimate aim is to reduce the redundancy and increase the relevance of intrusion detection model.

II. RELATED WORKS

This section discusses about some of the existing works in feature selection in intrusion detection using various data mining approaches. Amiri et al [6] introduced an efficient feature selection model using mutual information technique, they compared the performance of the mutual information with linear correlation model and the result proved the accuracy of classification in different types of network attacks.

Retrieval Number: L31161081219/2019©BEIESP
DOI: 10.35940/ijitee.L3116.1081219
Journal Website: www.ijitee.org

Senthilnayaki et al [7] developed an IDS approach which uses the gain ratio of two random variables and they are validated using support vector machine. This kind of feature selection approach in mainly used for DoS attack classification based on their class labels. Farrahi and Ahmadzadeh[8] proposed k-means clustering model with multiple classifiers to determine its accuracy based on oneR, Naïve bayes and support vector machine. This intrusion detection model classifies whether the traffic is normal or attacking type. The Dos attacks rate of classification is higher while comparing probe, R2L and U2R attacks.

Saxena and Richariya [9] designed a gain ratiomodel which uses particle swarm optimization-based support vector machine to perform feature selection process. But computation time of SWM with PSO is not examined as it is an important factor while performing optimized feature selection method. Sumaiyaand Aswani Kumar [10] in their work used chi-square feature selection and support vector machine as multi classifier. In this approach performance of support vector machine is fine tuned using radial basis kernel function. Particle swarm optimization is used to optimized the kernel parameter using samples variance belong to both similar and different classes. This model decreases the training and testing time in intrusion detection system.

Saraet al [11] developed an IDS based feature selection model which integrates both filter and wrapper method. This work used linear correlation coefficient for feature grouping and cuttlefish algorithm correspondingly. Decision tree is used as the classifier in the proposed method. There are no proper proof to handle the possibility of indeterminacy, vagueness in overcoming redundant information produced by the features related to the class label and discovering more relevant features which can contribute the process of higher information about the class label. Two overcome these two issues this paper uses theory of intuitionistic fuzzy sets along with mutual information system to boost its ability to handle the uncertainty while acquiring information from multiple features in intrusion detection system.

III. MUTUAL INFORMATION

Mutual information is a kind of quantity measure which defines the amount of information shared among each variable, in this feature selection problem it is used to measure the relevancy among feature x and the class label cl [12]. The amount of information shared among two individual variables is termed as mutual information. It is characterized as shown in the equation (1).

$$MI(x,y) = \sum_{x \in X} \sum_{y \in Y} P(x,y) log \frac{P(x,y)}{P(x).P(y)} \qquad (1)$$

where p(x,y) is joint probability function of both X and Y , their marginal densities are p(x) and p(y) respectively. To discover how similar joint distribution of p(x,y) is to products of factored marginal distribution is known as mutual information.

In this feature selection in IDS dataset, the goal is to maximize the mutual information among selected subset of features Fs and the target class variable y.

$$\bar{s} = max_S \; MI(F_S, y) \qquad (2)$$

where $d = |s|$ is the number of variables that has to be selected and it is termed as joint mutual information

This journal uses double-blind review process, which means that both the reviewer (s) and author (s) identities concealed from the reviewers, and vice versa, throughout the review process. All submitted manuscripts are reviewed by three reviewer one from India and rest two from overseas. There should be proper comments of the reviewers for the purpose of acceptance/ rejection. There should be minimum 01 to 02 week time window for it.

IV. PROPOSED METHODLOGY

Proposed Intuitionistic Fuzzy Mutual Information Feature Subset Selection in Intrusion Detection System

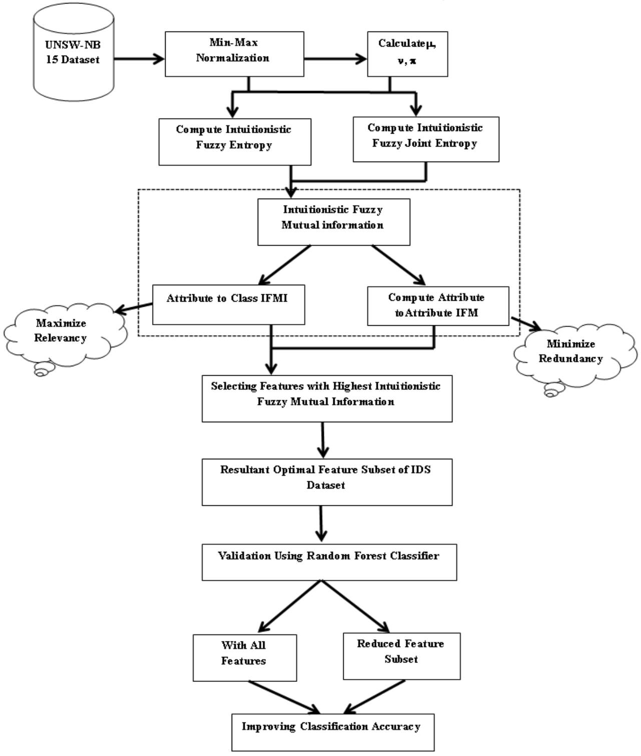
UNSW-NB 15 Dataset
Min-Max Normalization
Calculateμ, ν, π
Compute Intuitionistic Fuzzy Entropy
Compute Intuitionistic Fuzzy Joint Entropy
Intuitionistic Fuzzy Mutual information
Attribute to Class IFMI
Compute Attribute to Attribute IFM
Maximize Relevancy
Minimize Redundancy
Selecting Features with Highest Intuitionistic Fuzzy Mutual Information
Resultant Optimal Feature Subset of IDS Dataset
Validation Using Random Forest Classifier
With All Features
Reduced Feature Subset
Improving Classification Accuracy

Fig. 1.Proposed model feature subset selection

As shown in the Figure 1, in this proposed model feature subset selection is done using Intuitionistic fuzzy mutual information which handles impreciseness in selecting the optimal feature which contributes highest mutual information. In this work initially attribute-class mutual information is computed and selects the attributes that provides highest mutual information. In general, mutual information process selected attributes are removed from the original attribute set and added to the selected attribute subset. Next, the non-selected attributes attribute-class mutual information is calculated and attribute-attribute mutual information is determined among each of the selected attributes. From the obtained values, this algorithm selects a feature that has the highest attribute-class mutual information and minimum attribute-attribute mutual information.

In Intuitionistic fuzzy [15, 16] mutual information for intrusion detection dataset IDDS comprised of n attributes or features idf1, idf2, idf3,…, idfn which is signified by IDF, the delinquent is to choose a subset IDF' of m attributes (where m $\leq$ n). For a given class label CLi so that it achieves three major objective such as

i) $idfi' \in IDF'$ are more relevant for CLi

ii) IDF' is optimal

iii) Accuracy of classification is higher while using IDF' $idfi' \in IDF'$ than other subset of features

This proposed work used Intuitionistic fuzzy theory and mutual information to select features using best first approach. Intuitionistic Fuzzy Mutual information is defined as

$$IFMI(X,Y) = IFH(X) + IFH(Y) - IFH(X,Y) \qquad (3)$$

where X,Y are two intuitionistic fuzzy variables, IFH(X), IFH(Y) are Intuitionistic fuzzy entropy values for the values X and Y correspondingly whereas IFH (X, Y) is Intuitionistic fuzzy joint entropy for X and Y. To choose a feature it calculates two values they are attribute-class intuitionistic fuzzy mutual information and attribute-attribute intuitionistic fuzzy mutual information. Attribute-Class intuitionistic fuzzy information is applied to identify the correlation of an attribute with respect to its class label and choose an attribute which has highest correlation value. At the same time, attribute-attribute intuitionistic fuzzy mutual information is calculated to discover how similar those two attributes are. While finding high mutual information it reduces impreciseness among attributes importance and if there is zero mutual information then those two variables are considered as independent. Thus, in this work we select attributes in attribute-attribute intuitionistic fuzzy mutual information whose values are very low which means that redundancy among attributes can be rejected.

Let us assume that V is random variable with n number of elements such as {v1, v2 …, vn} and A and B are two intuitionistic fuzzy sets defined on V.

The Intuitionistic fuzzy membership value of k^th feature for i^th class represented as μ i,k, its non-membership value and hesitation degree is ϑ i,k, and π i,k, respectively. The equation for computing these three values are defined in the equation (4).

$$\mu_{i,k} = \left(\frac{\left\| \bar{v}_i - v \right\|}{d + \in} \sigma \right)^{-2/q-1} \qquad (4)$$

where q is the intuitionistic fuzzy coefficient, and $\in$ is a small value which is used to avoid singularity its value is greater than zero. σ is the standard deviation which is involved in computation of distance.

$\bar{v}_i$ represents mean of data instances that belong to class variable i and d signifies the radius of data which is calculated as d=max $\left(\left\| \bar{v}_i - v \right\| \sigma \right)$. By obtaining the membership value, non-membership $\vartheta_{i,k}$ and hesitation values are derived as shown in the equation 5 and 6 respectively.

$$\vartheta_{i,k} = \frac{1 - \mu_{i,k}}{1 + \tau \mu_{i,k}} \qquad (5)$$

$$\pi_{i,k} = 1 - \mu_{i,k} - \vartheta_{i,k} \qquad (6)$$

where $\tau > 0$ is a constant value.To determine the optimal feature subset intuitionistic fuzzy entropy is defined in the following equations 7 a) and b)

$$IFH(A) = -\frac{1}{n} \sum_{x \in X} [\mu A(x) \log \mu A(x) + \vartheta A(x) \log \vartheta A(x) - (1 - \pi A(x)) \log(1 - \pi A(x)) - \pi A(x)] \qquad 7(a)$$

$$IFH(B) = -\frac{1}{n} \sum_{x \in X} [\mu B(x) \log \mu B(x) + \vartheta B(x) \log \vartheta B(x) - (1 - \pi B(x)) \log(1 - \pi B(x)) - \pi B(x)] \qquad 7(b)$$

$$IFH(A \cup B) = \frac{1}{n} \sum_{x \in X} [(\mu A(x) v \mu B(x)] + [\vartheta A(x) . \vartheta B(x)]$$

$$\log[\vartheta A(x) v \vartheta B(x)] - [1 - \pi A(x) v \pi B(x) \log[1 - \pi A(x) v \pi B(x)] - \pi A(x) v \pi B(x)] \qquad (8)$$

Algorithm for Intuitionistic Fuzzy Mutual Information based Optimal Feature Subset Selection in Intrusion Detection System

Input: *Intrusion Detection Dataset IDDS with number of attributes m; IDF, set of attributes {idf1, idf2, idf3,…idfn}*

Output: *IDF', Potential feature subset*

Procedure Steps:

```
for t = 1 to n do
    calculate IFMI(idft, CL)
end
Choose the attribute idfi with maximum IFMI (idfi, CL)
IDF' = IDF' $\cup$ {idfi}
IDF = IDF - {idfi}
cnt = 1;
while cntm $\leq$ do
for each attribute idfs $\in$ IDF do
IFAAMI = 0;
for each attribute idfs $\in$ IDF'do
    IFAAMI = IFAAMI + calculate_ IFAAMI(idfi,idfj)
end
    Avg_IFAAMI = avg(IFAAMI)
IFACMI = Calculate_IFACMI(idfj,CL)
end
choose next feature idfj that has minimum IFAAMI but maximum IFACMI
IDF' = IDF' $\cup$ {idfj}
IDF = IDF- {idfj}
i=j
cnt = cnt + 1
end
Return potential feature set IDF'
```

V. RESULT AND DISCUSSION

The accuracy of the proposed model IFMIS-FS is evaluated on UNSW-NB15 dataset its detailed description is given in the following subsection. This proposed model is implemented using python with tensorflow.Tests were conducted on a personal computer with 2.53 GHZ CORETM i5 CPU and 4GB of memory under windows 10.

In course of experimentation dataset training and testing is done on the same ratio. The evaluation metric used for validating the performance of three different algorithms are done using accuracy, precision, recall, time taken and memory capacity.

Tests were conducted on a personal computer with 2.53 GHZ CORETM i5 CPU and 4GB of memory under windows7.

- True Positive (TP) : It represents the actual attack packets as correctly classified as attacks
- True Negative (TN): This represents actual normal packets correctly classified as normal
- False Negative (FN): Here the actual attack packets are classified as normal packets
- False Positive (FP) : This value denotes incorrect prediction where the normal packets are classified as attack

Accuracy refers to percentage of instances correctly classified. It is an important metric which is used to measure the accuracy of classification model.

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN)$$

Precision is the measure of positive samples correctly predicted by the classification model.

$$\text{Precision} = (TP)/(TP+TN)$$

Recall: It is a measure of true positive rate which is defined as the ratio of positive instances correctly classified as positive

$$\text{Recall}= (TP)/(TP+FN)$$

A. DATASET DESCRIPTION

The proposed Intuitionistic fuzzy mutual information feature subset selection uses UNSW-NB15 dataset published in 2015 [13, 14]. It was created by IXIA PerfectStorm tool in the Cyber Range Lab of the Australian Centre for Cyber Security (ACCS). This dataset includes nine dissimilar modern attack types with 49 features including class labels comprising of 2,540,044 records.

Complete list of attributes used in UNSW-NB15 is shown in the Table-I.

Table-I: List of attributes in UNSW-NB15 dataset

Feature with description	*Feature with description*
Srcip (Source IP address)	Stcpb -Source TCP base sequence number
Sport -Source port number	Dtcpb- Destination TCP base sequence number
Dstip - Destination IP address	Smeansz - Mean of the packet size transmitted by the src
Dsport - Destination port number	Dmeansz - Mean of the packet size transmitted by the dst
Proto - Transaction protocol	trans_depth - Represents the pipelined depth into the connection of http request/response transaction
State - Indicates to the state	res_bdy_len - Actual uncompressed content size of the data transferred from the server's http service.
Dur - Record total duration	Sjit - Source jitter (mSec)
Sbytes - Source to destination transaction bytes	Djit - Destination jitter (mSec)
Dbytes- Destination to source transaction bytes	Stime - record start time
Sttl-Source to destination time to live value	Ltime - record last time
Dttl-Destination to source time to live value	Sintpkt - Source interpacket arrival time (mSec)
Sloss-Source packets retransmitted or dropped	Dintpkt - Destination interpacket arrival time (mSec)
Dloss-Destination pkts retransmitted or dropped	Tcprtt - TCP connection setup round-trip time, the sum of 'synack' and 'ackdat'.
Service - http, ftp, smtp, ssh, dns, ftp-data ,irc	Synack - TCP connection setup time, the time between the SYN and the SYN_ACK packets.
Sload - Source bits per second	Ackdat - TCP connection setup time, the time between the SYN_ACK and the ACK packets.
Dload - Destination bits per second	is_sm_ips_ports - If source (1) and destination (3)IP addresses equal and port numbers (2)(4) equal then, this variable takes value 1 else 0
Spkts-Source to destination packet count	ct_state_ttl - No. for each state (6) according to specific range of values for source/destination time to live (10) (11).
Dpkts - Destination to source packet count	ct_flw_http_mthd - No. of flows that has methods such as Get and Post in http service.
Swin - Source TCP window advertisement	is_ftp_login - If the ftp session is accessed by user and password then 1 else 0.
Dwin - Destination TCP window advertisement	ct_ftp_cmd - No of flows that has a command in ftp session.
ct_srv_src - No. of connections that contain the same service (14) and source address (1) in 100 connections according to the last time (26).	ct_dst_sport_ltm - No of connections of the same destination address (3) and the source port (2) in 100 connections according to the last time (26).
ct_srv_dst - No. of connections that contain the same service (14) and destination address (3) in 100 connections according to the last time (26).	ct_dst_src_ltm - No of connections of the same source (1) and the destination (3) address in in 100 connections according to the last time (26).
ct_dst_ltm - No. of connections of the same destination address (3) in 100 connections according to the last time (26).	attack_cat- The name of each attack category. In this data set , nine categories e.g. Fuzzers, Analysis, Backdoors, DoS Exploits, Generic, Reconnaissance, Shellcode and Worms
ct_src_ ltm - No. of connections of the same source address (1) in 100 connections according to the last time (26).	Label - 0 for normal and 1 for attack records
ct_src_dport_ltm - No of connections of the same source address (1) and the destination port (4) in 100 connections according to the last time (26).	

Table-II: Feature subset Generated by four different Methods

Feature Subset Selection Methods	**No. of. Attributes Selected**	**Attribute Description**
Principal Component Analysis	12	dttl, dpkts, slosssinpkt, tcprtt, ct_dst_ltmmis_sm_ips_ports,ct_ftp_cmd,dbytes, djit, ct_src_dport_ltm,dloss, spkts
Greedy Stepwise Search	9	Spkts, sttl,dload, sinpkt, swin,Tcprtt,trans_depth,ct_dst_sport_ltm, ct_flw_http_mthd
Mutual Information	8	Spkts, sttl,dload, sinpkt, swin,Tcprtt, trans_depth, ct_dst_sport_ltm, ct_flw_http_mthd,dloss

Intuitionistic Fuzzy Mutual Information	7	'dttl','ct_dst_ltm', 'ct_src_dport_ltm', 'ct_dst_sport_ltm','ct_dst_src_ltm','ct_src_ltm', 'ct_srv_dst'

The following figure shows that the number of features selected for the methods PCA, GSS, MI, and IFMI (Intuitionistic Fuzzy Mutual Information).

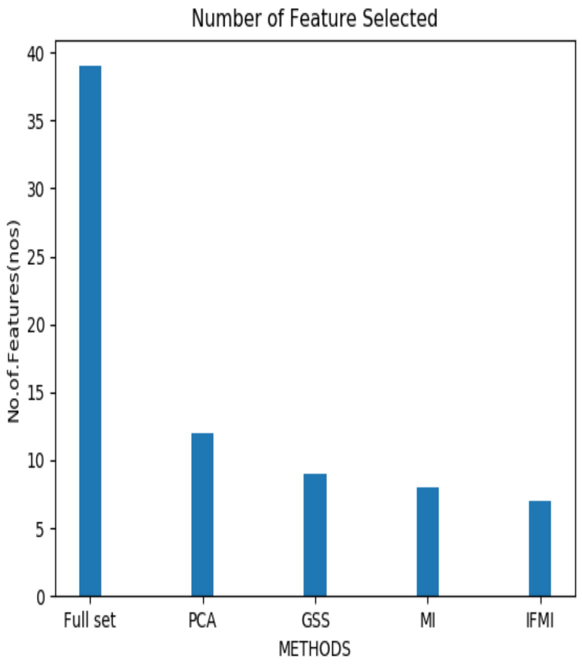

Fig.2.Subset selection for different model

Table II and Figure 2 describes about feature subset selection done by four different models namely principal component analysis, greedy stepwise search, mutual information and proposed intuitionistic fuzzy mutual information. Among 49 attributes of **UNSW-NB15 dataset, P**rincipal Component Analysis produces 12 attributes as feature subset, Greedy Stepwise Search produces 9 attributes as feature subset, mutual information-based feature subset selection produces 8 attributes as feature subset and finally Iintuitionistic Fuzzy subset produces 7 attributes as feature subsets. Based on their logics each of the models produces different number of feature sets as output which is highly related to their analysis technique.

Table- III: Performance comparison based using Random Forest Classifier

Feature Subset Selection Methods	**Accuracy**	**Precision**	**Recall**
With Full Attributes	89.07	85.03	87.98
Principal Component Analysis (PCA)	80.9	76.8	79.5
Greedy Stepwise Search (GSS)	86.5	82.3	83.42
Mutual Information (MI)	90.54	87.8	89.52
Intuitionistic Fuzzy Mutual Information (IFMI)	97.35	94.46	95.9

The table III shows the comparison of four different feature subset selection models-based classification of random forest algorithm. The metrics used to validate the feature subset generated by each of the four models are accuracy, precision and recall.

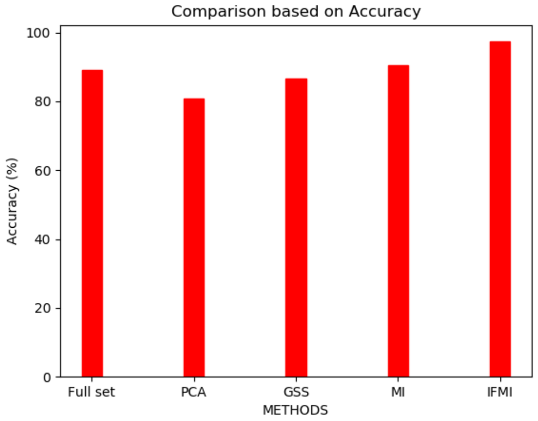

Fig.3. Accuracy Comparison

From the figure 3 shows the accuracy comparison of random forest classifier while using four different models. The proposed IFMI based feature subset selection produces highest accuracy while comparing the other models because it has the ability to represent each attribute mutual information in terms of membership, non-membership and hesitation degree. The impreciseness in selection of potential attributes are well handled using the proposed model. While Mutual information fails to handle the impreciseness when the given input dataset is vague and its random nature of feature selection approach.

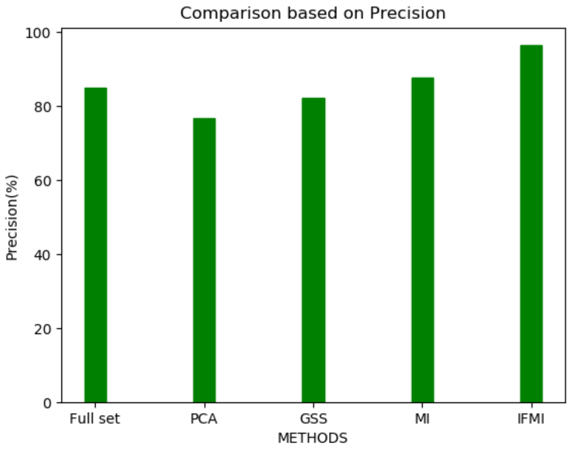

Fig.4. Precision for various methods

From the figure 4, It is proved that the highest true positive instances by finding precision are determined by random forest classifier to determine the intrusion detection on UNSW-NB15 dataset. While using the feature subset of IFMI as input compared to the whole feature subset, Mutual Information, PCA and greedy stepwise search. This is because the intuitionistic fuzzy reduces the redundancy of attributes and choose the attributes with highest merit score. Compared the whole feature set, reduced feature subset of IFMI achieves higher precision value in intrusion detection using random forest classifier.

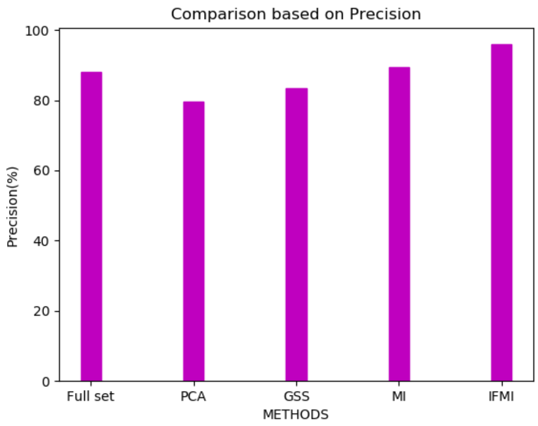

Fig.5. Performance of the proposed IFMI

The figure 5 proved that the performance of the proposed IFMI based feature subset selection achieves highest recall value compared to other models PCA, GSS and MI. The objective of intuitionistic fuzzy is to handle the impreciseness in **UNSW-NB15 dataset** classification of normal and attack instances, during feature subset selection which focuses on hesitation degree of each attribute as an important factor during the process of attribute selection done in both attribute to class mutual information and attribute to attribute mutual information computation.

VI. CONCLUSION

The main objective of this paper is to handle the impreciseness in selection of feature subsets which results in high classification accuracy of the intrusion detection process. The proposed model introduces intuitionistic fuzzy mutual information-based feature subset selection. In this each of the attributes are represented in the form of three degrees namely membership, non-membership and hesitation degree.The UNSW-NB15 dataset is used for classifying the instances as normal or attack type. Based on the score of intuitionistic fuzzy mutual information obtained by each attributes the highly scored attributes are considered as more independent variables and they produce more information about the class label. The performance of the four different models of feature subset selection is done by random forest classifier. The results proved that the feature subsets generated by proposed intuitionistic fuzzy mutual information produces more accuracy and less false alarms compared to mutual information, principal component analysis, greedy stepwise search and whole feature subset.

REFERENCES

1. Inayat Z, Gani A, Anuar N.B, Khan M.K, Anwar S, Intrusion response systems: Foundations, design, and challenges. Journal of Network Computing Appl. 2016, 62, 53–74.
2. Singh, R.; Kumar, H.; Singla, R.K.; Ketti, R.R. Internet attacks and intrusion detection system: A review of the literature. Online Inf. Rev. 2017, 41, 171–184
3. Wang, Z. Deep Learning-Based Intrusion Detection with Adversaries. IEEE Access 2018, 6, 38367–38384.
4. Karim, I.; Vien, Q.T.; Le, T.A.; Mapp, G. A comparative experimental design and performance analysis of Snort-based intrusion detection system in practical computer networks. MDPI Computers. 2017, 6, 6
5. Peng, H.C.; Long, F.H.; Ding, C. Feature selection for high-dimensional data: A fast correlation-based filter solution. In Proceedings of the 20th International Conference on Machine Learning, Washington, DC, USA,21–24 August 2003
6. Amiri, Fatemeh, Mahdi, Mohammad, Yousefi, Rezaei, 2011. Mutual information based feature selection for intrusion detection systems, J. Network Comput. Appl. 34, 1184–1199.
7. Senthilnayaki, Balakrishnan, Venkatalakshmi, K., Kannan, A., Intrusion detection system using feature selection and classification technique. Int. J. Comput. Sci. Appl. 3 (4), 145–151, 2014.
8. Farrahi, Vahid S., Ahmadzadeh, Marzieh, 2015. KCMC: a hybrid learning approach for network intrusion detection using K-means clustering and multiple classifiers. Int. J. Comput. Appl. 124 (9), pp. 18–23. Published by Foundation of Computer Science (FCS),NY, USA.
9. Saxena, Harshit, Richariya, Vineet, 2014. Intrusion detection in KDD99 dataset using SVM-PSO and feature reduction with information gain. Int. J. Comput. Appl. 98 (6), 25–29.
10. SumaiyaThaseen Ikram, Aswani Kumar Cherukuri, Intrusion detection model using fusion of chi-square feature selection and multi class SVM Journal of King Saud University – Computer and Information Sciences (2017) 29, 462–472
11. Sara Mohammadi, Hamid Mirvaziri , Mostafa GhazizadehAhsaeea , Hadis Karimipour, Cyber intrusion detection by combined feature selection algorithm, Journal of Information Security and Applications 44 (2019) 80–88.
12. H. Peng, H. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy." Pattern Analysis and Machine Intelligence,IEEE Transactions on 27.8, 2005, 1226-1238.
13. Moustafa, Nour, and Jill Slay. "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)."Military Communications and Information Systems Conference (MilCIS), 2015. IEEE, 2015.
14. https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/
15. Krassimir T. Atanassov, Fuzzy Sets and Systems, North-Holland, Volume 20 (1986), pages 87-96, ISSN 0165-0114
16. Krassimir T. Atanassov, Series "Studies in Fuzziness and Soft Computing", Volume 35, Springer Physica-Verlag, 1999, ISBN 3-7908-1228-5

AUTHORS PROFILE

Mrs.P.Sudha has completed her M.Phil. in Computer Science and pursuing Ph.D in Bharathiar University. Her research area is Data Mining and Big Data Analytics. She has 13 years teaching experience. She is currently working as Assistant Professor in Computer Science, Sree Saraswathi Thyagaraja College, Pollachi. She has 10 years of research experience. She has published around 20 research articles in the refereed International Journals with high impact factor and also presented many research papers in the National and International level Conference.

Dr.R.Gunavathi has completed her Ph.D. in Computer Science in Mother Teresa Women's University, Kodaikanel, and her research is on "Efficient Cluster head selection algorithms to improve the Quality of service in Mobile Ad hoc networks". She has 20 years of teaching experience and currently working as Associate Professor and Head, Department of MCA at Sree Sarawthi Thyagaraja College, Pollachi. She has 15 years of research experience. Her current research interest in Mobile Ad hoc Networks, Vehicular Ad hoc Networks and Big Data Analytics. She has published around 30 research articles in the refereed International Journals with good impact factor and also presented 25 research papers in the National and International level Conferences. She has organized many National level seminars, workshops and conferences.