Ensemble Based Optimal Feature Selection Algorithm for Efficient Intrusion Detection in Wireless Sensor Network

Shyam Sundar S[1*], R.S. Bhuvaneswaran[1], and SaiRamesh L[2]
[1] College of Engineering Guindy, Anna University, Chennai-600025, India
[2] St. Joseph's Institute of Technology, Chennai, India
[e-mail: shyam24phd@gmail.com, bhuvan@annauniv.edu, sairamesh.ist@gmail.com]
[*]Corresponding author: Shyam Sundar S

Received August 15, 2022; revised December 21, 2023; revised April 9, 2024; revised June 10, 2024; accepted July 29, 2024; published August 31, 2024

Abstract

Wireless sensor network (WSN) consists of large number of sensor nodes that are deployed in geographical locations to collect sensed information, process data and communicate it to the control station for further processing. Due the unfriendly environment where the sensors are deployed, there exist many possibilities of malicious nodes which performs malicious activities in the network. Therefore, the security threats affect performance and life time of sensor networks, whereas various security aspects are there to address security issues in WSN namely Cryptography, Trust Management, Intrusion Detection System (IDS) and Intrusion Prevention Systems (IPS). However, IDS detect the malicious activities and produce an alarm. These malicious activities exploit vulnerabilities in the network layer and affect all layers in the network. Existing feature selection methods such as filter-based methods are not considering the redundancy of the selected features and wrapper method has high risk of overfitting the classification of intrusion. Due to overfitting, the classification algorithm fails to detect the intrusion in better manner. The main objective of this paper is to provide the efficient feature selection algorithm which was suitable for any type classification algorithm to detect the intrusion in an effective manner. This paper, the security of the network is addressed by proposing Feature Selection Algorithm using Chi Squared with Ensemble Method (FSChE). The proposed scheme employs the combination of decision tree along with the random forest classification algorithm to form ensemble classifier. The experimental results justify the feasibility of the proposed scheme in terms of attack detection, packet delivery ratio and time analysis by employing NSL KDD cup data Set. The obtained results shows that the proposed ensemble method increases the overall performance by 10% to 25% with respect to mentioned parameters.

Keywords: IDS, Ensemble method, feature selection, sensor network, filter method, decision tree.

1. Introduction

WSN is an infrastructure less network where the large number of sensor nodes are connected in ad-hoc manner which are used to monitor the objects, environmental conditions and track the vehicles etc. WSN is applicable to military applications [1], hence the secure data transmission is more important to WSN. Therefore, various authors [2] [3] [4] have proposed on techniques secured data transmission from the source nodes to destination the destination nodes. In WSN, a cyber-attack is such type of attack which is intended to perform the malicious attacks on the sensed data gathered by sensor nodes. The malicious nodes in the network launches the various types of attacks which compromise integrity of the network during data transmission. By doing so, the malicious nodes can cause various attacks namely Denial of Service attack and Probing attack, User to Root and Remote to Local attacks which can disrupt the process of data communication in the network. Intrusion detection system (IDS) plays a major role to secure a wireless sensor network from intruder. The term Intrusion is considered to compromise the assets on the computer network such as confidentiality, integrity, and availability.

The primary functions of IDS in WSN are to calculate energy consumption of each sensor by discovering malicious activities of the sensor networks and then produce alarm. The first IDS was proposed by [5] since then IDS matured a lot. Many researchers developed IDS with the aim of proposing the intrusion detection with high accuracy and better false positive rate. IDS are broadly classified into two major approaches namely Intrusion detection approach and protection system approach. The intrusion detection approach is further classified into Signature based intrusion detection, Anomaly based intrusion detection and protocol-based intrusion detection. The Protected system approach is further classified into Host based IDS, Network based IDS, and Hybrid IDS, 3. behavior after an attack: Active IDS and Passive IDS. Signature based IDS detect security threats based on the characteristics of known attacks but it fails to detect zero-day attacks. Anomaly based IDS detects intruders in specific traffic flow with unknown attacks but false positive rate is high between expected and anomaly behavior. Protocol based detection method compares predefined rules with the specific data flow of the protocol on the network. Host based IDS monitor system activities like files modifications, memory usage and it is responsible for checking the files and analyzing logs [6] [7]. Network based IDS monitor network activities and its communications and also audits packet information to protect sensor network from potential threats.

Hybrid IDS [8] combines two or more IDS for better detection accuracy. IDS can be classified based on the behavior of an attack such as active and passive IDS. In active IDS predetermined actions and automatically blocks the attacks before the security analysts are involved whereas passive IDS only monitor and analyses the network traffic activities and produces an alarm on potential attacks [9]. Various authors proposed various feature selection based IDS [10] [11]. However, the existing systems fail to detect and classify the malicious activities more accurately. Therefore, it can be achieved with the help of an efficient feature selection algorithm. Feature Selection is necessary in IDS to select relevant input features from total original input data to improve the classification accuracy and reduces the false positive rates of attacks.

1.1 Motivation

Most of the existing systems employs only two feature selection algorithms namely filter approach and wrapper approach for efficient selection of most relevant features [12] [13]. The

Filter methods select the features based on training data to select the optimal features from the given data set. However, this method suffers from computational cost when it is compared with wrapper methods. On the other hand, wrapper method involves in optimizing a learning classifier but it is more complex and time consuming. Sometimes it gives better results than filter method. Information Gain, Chi-squared, Odds Ratio and Correlation Coefficient are known metrics.

1.2 Contribution

In this work, the security issues of the network are addressed by proposing Feature Selection Algorithm using Chi Squared with Ensemble Method (FSChE). The proposed scheme employs the combination of decision tree along with the random forest classification algorithm to form ensemble classifier. The proposed ensemble based optimal feature selection algorithm improves the performance of IDS in WSN by efficient selection of optimal feature set and there by improves the classification process by reducing the overfitting problem. **Table 1** shows the abbreviated term commonly used in this article.

1.3 Organization

The remaining article was organized as follows: Section 2 describes the existing research work related with intrusion detection using machine learning and other technique with their limitations. Section 3 gives the detailed view of proposed ensemble model for feature selection and followed by the experimental results to show that the performance of proposed model in improving the classification accuracy of IDS. Section 5 conclude the article with summary and future direction of the proposed system.

Table 1. List of abbreviations

IDS	Intrusion detection System
WSN	Wireless Sensor Network
FSChE	Feature Selection Algorithm using Chi Squared with Ensemble Method
NSL-KDD	Neural Simulation Language Knowledge Discovery in Databases
DT	Decision Tree

2. Literature Review

Secure communication in vulnerable environment is always a challenging issue. Many researchers contributed to build secure environment for safe communication. Intrusion Detection System is one of the methods for secure communication in WSN. Many researchers have worked on IDS. Among them, [14] proposed an Anomaly based IDS for detecting anomalies in wireless sensor network using neuro fuzzy method. It is a lightweight protocol

which does not consumes much overhead and uses two-step detection method. One is to create trust value using fuzzy logic to categorize trust or untrusted nodes and another one is to separate the malicious node from genuine node using neural network. Therefore, this method minimizes the percentage of false positive rate. The limitation of this paper is that it doesn't consider the node mobility and density. A lightweight energy consumption-based IDS system for WSN was proposed [15]. This system uses mobile agents to collect energy readings in order to detect malicious nodes in the network based on sensor nodes energy. The energy consumption is predicted using linear regression model which provides better results by detecting malicious nodes with high accuracy. IDS using dynamic feature selection method has been proposed [16] called dynamic recursive feature selection algorithm where dynamically selects the optimal features to detect intruders by using their proposed scheme.

Network intrusion detection system employs two-stage deep learning model to protect computer networks from various threats and malicious attacks [17]. This model works in two phases. The initial phase deals on classifying the traffic to detect the network traffic is normal or traffic by employing the probability score value. The second phase is to detect normal state of the traffic and different types of the attacks. The main advantage of this system is to prevent a reconnaissance attack before it infiltrates the internal network. But this NIDS is not able to discover network threats against the host.

Survey on various IDS methodologies, types, and technologies were carried out [18]. Moreover, several machine learning techniques are proposed in this paper to detect zero-day attacks. Comparison study was carried out for the existing feature selection techniques such Information Gain, Correlation, Relief and Symmetrical Uncertainty with C4.5 decision technique [19]. Among this C4.5 with Information Gain based feature selection technique produced highest accuracy rate of 99.68% with selected 17 features to develop IDS. An IDS using Naive Bayes classifier was proposed for NSL-KDD dataset with various pre-handling techniques [20]. Optimal feature selection algorithm and enhanced classification algorithm was proposed for multi class KDD cup intrusion detection dataset using support vector machine (SVM) to reduce irrelevant available features in the data set [21].

Data pre-processing techniques are influencing to increase attack detection accuracy by using different classifying methods namely decision tree, naïve bayes and rule based classifier with NSL-KDD cup intrusion detection dataset. The author [22] analyzed data Pre-processing technique such as filter methods and wrapper methods to detect attacks accurately. Deep learning based IDS [23] to detect the intrusion in the network. This method provides high classification accuracy than traditional classification algorithms. IDS using Chi-Square optimal feature selection method and multi-class SVM proposed in [24]. This proposed model provides high detection rate and low false alarm rates than other traditional approaches.

Survey carried out on Anomaly-Based IDS by Machine Learning [25] models which improve detection accuracy by reducing false detection rates. The proposed feature selection method works with four different estimators. The first estimator is selected if correlated value 0.3 is greater than the target variable. The other variable estimators namely Chi-Square, ANOVA F-value, and extra-trees predictive methods are used to select 25 best features from the dataset. They developed two machine learning model such as an ensemble learning model and a convolutional neural network model for protecting networks from probing attacks. The performance of IDS depends on suitable feature selection methods and machine learning algorithm. Irrelevant features can decrease the detection accuracy of IDS. In [26], Information Gain, Gain Ration, Chi-squared, and Relief selection methods were utilized to examine J48, Random Forest, Naïve Bayes, and KNN algorithm to show the performance of IDS.

2.1 Research Gap

From the above discussion, it is noted that many existing works attained various feature selection based IDS. However, these works provides less attack classification accuracy for developing IDS to WSN environment. Therefore, the proposed work identifies the need for optimal feature selection algorithm for efficient intrusion detection in wireless sensor network.

3. Proposed Work

3.1 Feature selection

Feature selection is a process in machine learning algorithm to find best features to build accurate model. It is broadly classified as supervised feature selection: used to select the relevant features based on labelled data and unsupervised feature selection: used to select the relevant features based on labelled data. Supervised feature selection method is further classified into filter method, wrapper method and hybrid method. Features are selected based on the statistical measures called filter method. This method consists of various techniques such as information gain, chi-square χ^2 method, Fisher's Score, missing value ratio. In wrapper method, features are selected based on the search problems. The following techniques are considered as wrapper methods such as forward selection method, backward elimination method, exhaustive feature selection, and recursive feature elimination method. Combination of both filter and wrapper method is called hybrid method. Chi-Square test is considered in this work to find the relationship between the predicted and actual variables. Chi-Square χ^2 method is a filter based method to reduce the irrelevant and redundancy of the feature sets to select the relevant features and this method is applied to test the independence between the features in order to reduce the set of features.

$$\chi^2 = \sum_{i=1}^{n} \frac{(P_i - A_i)^2}{A_i} \qquad (1)$$

Where, P is a Predicted variable and A is an Actual variable.

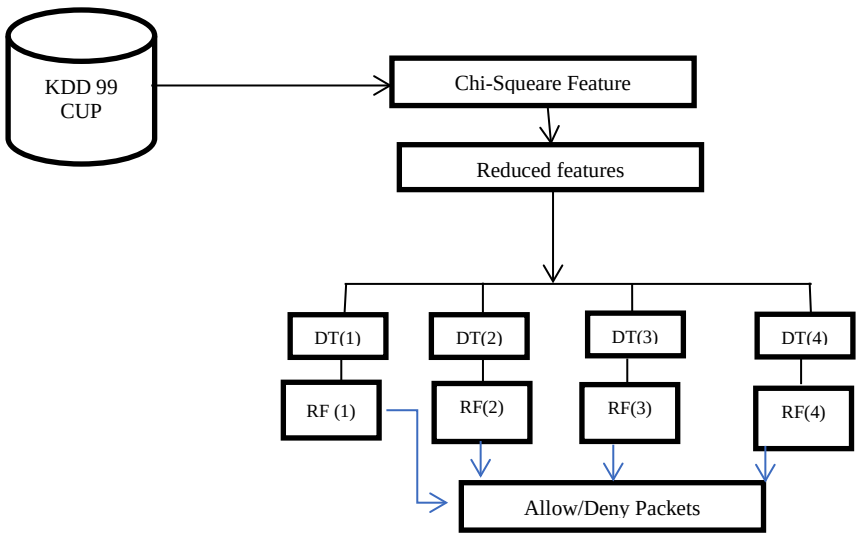

Fig. 1. Proposed Ensemble Classification with Chi-Square feature Selection

3.2 Ensemble method

An ensemble method is a ML technique which combines two or more models of ML algorithms to produce one optimal predictive model. In this paper, we considered Decision Tree, and Random Forest [9] and compared its accuracy to select the optimal model for predicting the Intrusions accurately.

Fig. 1 shows the flow the of the proposed ensemble classification approach for intrusion detection with Chi-square feature selection. Chi-square reduce the features from the initial feature of KDD cup data set. Then it will send to DT and the output of DT sent to Rf and finally the output mentioned as allow or deny the packet. Decision Tree (DT) is a tree structure based supervised algorithm to solve both classification and regression problems, but it is widely preferred for solving the classification problems. DT consists of set of tree structured based decision test with divide and conquer method. DT is constructed with (i) internal node which denotes attribute test value, (ii) branch that denotes the outcome of test and (iii) leaf node is the terminal node which holds class label. DT algorithms such as ID3, C4.5, and CART were intentionally developed for classification. The steps to be followed for creating Decision Tree (DT) for the Given Dataset are shown in **Table 2**.

Table 2. Algorithm for Creating Decision Tree for the Given Dataset

Algorithm 1 Creating Decision Tree for the Given Dataset

```
1: Create Decision Tree (DT)
2: Load R
3: Call attribute selection measure [21] to find best attributes
4: Divide the data set R into subsets S
        S ⊆ R
5: Generate DT nodes with best attributes
6: Make a new decision using S
7: Continue the process until leaf node n=0.
        for(i=n; i> 0; i--)
        {
            n = n - 1;
        repeat n!=0;
            n=0;
        }
      end for
8: Stop
```

Random Forest is a collection of DT and it is also one the ensemble method which is built using bagging with random attribute selection.

3.3 Feature Selection Algorithm with Ensemble method

In this paper, Feature Selection Algorithm using Chi Squared with Ensemble Method is proposed for finding the intruder efficiently for IDS in WSN shown in **Table 3**. The system gets original features from dataset and calculates the degree of freedom based on the contingency table with respect to the fixed threshold value and performed Chi-Square for each feature and rank the features which achieves the threshold value. Select the features from the ranked features instead of processing all features will reduce the energy conservation of the sensor node and increase the lifetime.

Table 3. Algorithm for proposed feature selection algorithm using chi squared with ensemble method.

Algorithm 2 Proposed Feature Selection Algorithm using Chi Squared with Ensemble Method

Input: Set of Original Feature, S={F1, F2, …, Fn}
Output: Set of Selected features, SF
Method:
1: Load set of features S={F1, F2, …, Fn}
2: Set Threshold Value Th = 0.95
3: Calculate Degree of Freedom based on Contingency table
4: For i=1 to n do
Begin
5: Perform Chi-Square test χ^2 using Equation (1) for each features 'F' in the training dataset

$$\chi^2 = \sum_{i=1}^{n} \frac{(P_i - A_i)^2}{A_i} \qquad (1)$$

6: If ($\chi^2 < Th$)
Select the features and rank the important features based on rules
SF=S-F;
Else
Continue
End For
7: Split the selected feature dataset SF into training and testing dataset.
8: Build a classifier with SF={SF1, SF2, …, SFn}
9: Call Ensemble method (Algorithm 3) and perform classification to detect attacks
10: Evaluate the performance metric

Table 4 shows the proposed ensemble algorithm for IDS using DT and Random Forest.

Table 4. Algorithm for Proposed Ensemble Method for IDS

Algorithm 3 Proposed Ensemble Method for IDS

Input: Set of factors {sf1…sf2…sfn} from Chi square feature.
Output: Intrusion detection accuracy for all selected classifier
Method:

```
Begin
1: Load the original data set.
        DSF={dsf1,dsf2……dsfN} //where DSF → data set features
2: For all DSF
     Do
        Split DSF and perform the ordering of the futures based on chi square
        Future selection method.
3:      For all ordering features in DSF
         Do
          Train the DT Classifier
4:          For all the ordered features from the Chi Square test
             Do
5:              Compute the ranking of each feature based on order by using
                  wrapper approach. Such that
                RF={RF1,RF2,….RFN}
6:              Compute the futures with the lowest ranking square.
                  Set the root node of the decision tree which has highest
                  ranking
7:              Construct the decision tree based on ranking of the features
                  from the given data set such as
                    RDS= {Rdf1,,Rdf2…..RdfN}
8:           End Do
9:         End For
10:      End Do
11:     End For
12:    End Do
13:  End For

// Random forest generating

14: For RDF {Rdf1..Rdf2…Rdfn}
      Do
15:     Read the first ranking feature and create the root node for the Random
          Forest.
16:    Load the selected features computed by the chi square feature selection
         algorithm.
```

```
Initialize two classes d1 and d2 with empty set
    Read the next ranking feature and find out the weightage of it's attributes based on ranking for all the features
Do
    If weightage of the feature is less than the root node append the feature in d1
    Else appended the feature in d2
    Proceed to create the nodes sub for root node d1 and d2.
Repeat the steps 1-7 until when features [Φ]
Return decision tree and random forest.
```

4. Results and Discussion

4.1 Performance Metrics

The performance evaluation of the Feature Selection Algorithm using Chi Squared with Ensemble Method is tested on NSL KDD dataset. This dataset contains five different main attribute classes namely Normal, DoS, Probe, U2R and R2L.

The performance evaluation of the proposed techniques is measured based on the performance metrics namely True_Positive Score (TPS) where the Normal nodes are classified correctly), True_Negative Score (TNS) where the Abnormal nodes are classified correctly), False_Positive Score (FPS) where the abnormal nodes are misclassified as normal nodes) and False_Negative Score (FNS) where the Normal nodes are misclassified as abnormal nodes.

The accuracy of the classifier is computed using Equation (2). The Precision(P) (exactly measures the positive instances of class label) shown in Equation 3 and Recall(R) (completeness of positive instances of class label) shown in Equation 4and F-measure (F) is shown in Equation 5. These four metrices mainly used to predict the classification accuracy of the proposed machine learning algorithms. These metrics are also frequently used for classification in any machine learning approaches. In addition to this, packet delivery analysis and delay time was also used as performance parameter which was common for network related research works.

$$Accuracy\ (A) = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (2)$$

$$P = \frac{T_P}{T_P + F_P} \quad (3)$$

$$R = \frac{T_P}{T_P + F_N} \quad (4)$$

$$F = \frac{2 * P * R}{P + R} \quad (5)$$

4.2 Confusion matrix

Confusion matrix is used to describe the performance of the classifier. It can be applied for both binary and multiclass classification algorithm is shown in **Table 5** and **Table 6** respectively and the confusion matrix represents the counts from predicted class and actual class.

Table 5. Confusion matrix for binary classification

	Anomaly	Normal
Anomaly	9426	3296
Normal	286	9531

Table 6. Confusion matrix for multi class experiments

Predicted class / **Actual class**	**Normal**	**DoS**	**Probe**	**U2R**	**R2L**
Normal	9467	89	246	3	7
DoS	1023	6239	97	1	127
Probe	242	170	2937	0	7
U2R	156	0	19	24	14
R2L	2978	0	13	8	720

4.3 Attack detection metrics

Table 7 shows the attack detection metrics of intrusion accuracy, precision, recall and F-measure. **Fig. 2** shows the DoS attack detection accuracy of the proposed scheme when it is compared with Filter and Wrapper based feature selection methods.

Table 7. Attack detection metrics

Attack	**Accuracy**	**Precision**	**Recall**	**F measure**
DoS	0.99638	0.99592	0.99605	0.99598
Probe	0.99185	0.98747	0.98567	0.98666
U2R	0.99752	0.87638	0.89640	0.87631
R2L	0.97559	0.96789	0.96186	0.96579

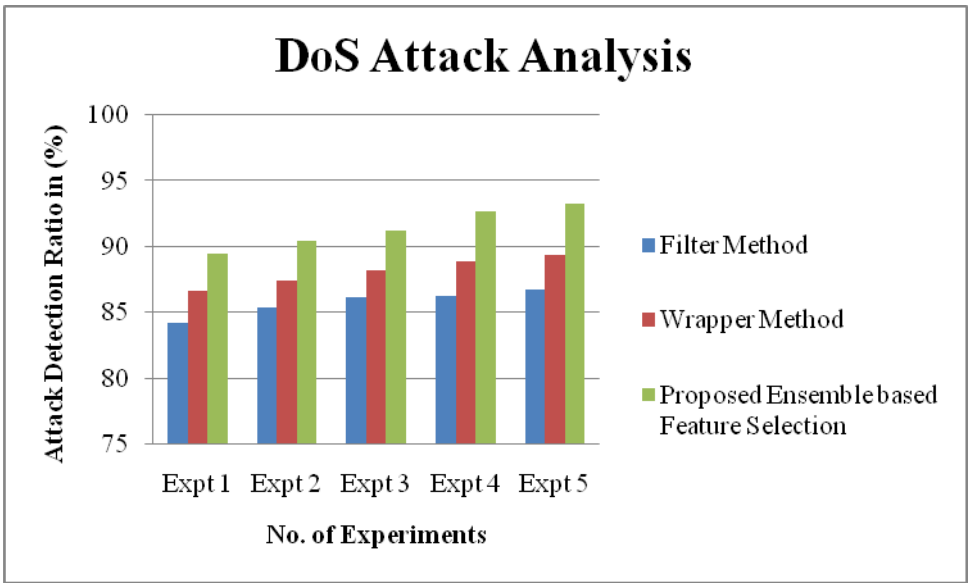

Fig. 2. DoS Attack detection accuracy using proposed ensemble based feature selection method.

From the **Fig. 2** it is observed that the DoS detection accuracy of the attacks are more in the proposed work when compared to the existing methods.

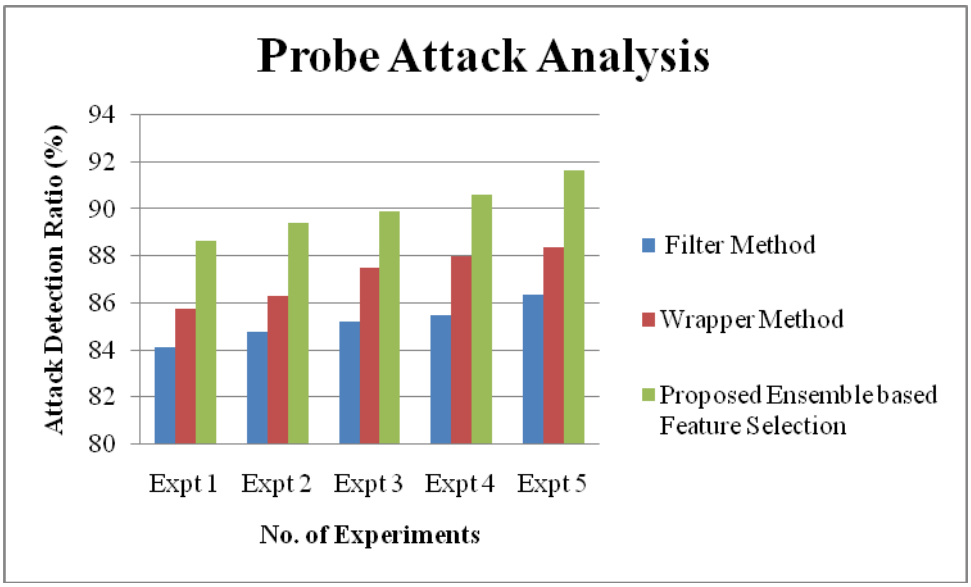

Fig. 3. Probe Attack detection accuracy using proposed ensemble based feature selection method

Fig. 3 shows the Probe attack detection accuracy of the proposed scheme when it is compared with Filter and Wrapper based feature selection methods. From the **Fig. 3** it is observed that the Probe detection accuracy of the attacks are more in the proposed work when compared to the existing methods.

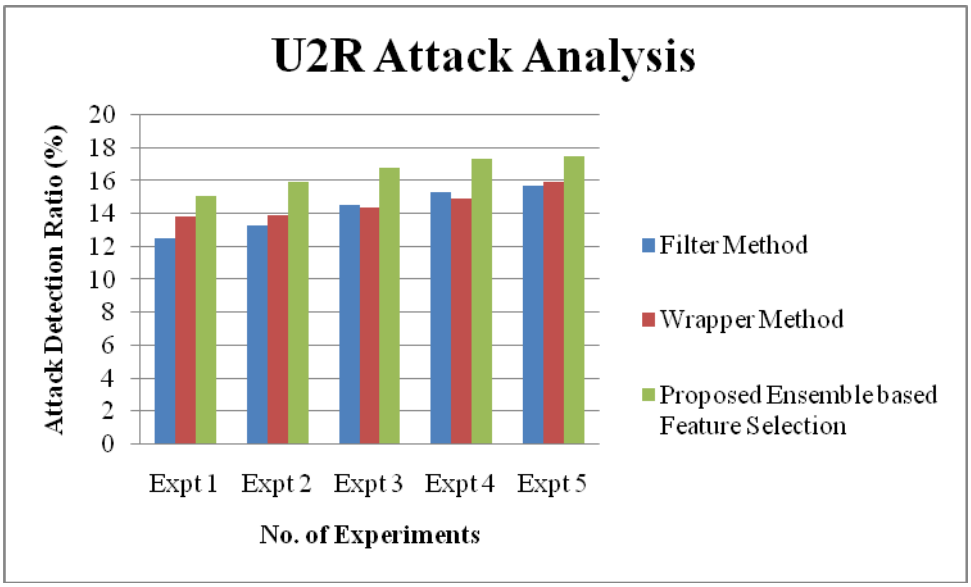

Fig. 4. U2R attack detection accuracy using proposed ensemble based feature selection method

Fig. 4 shows the U2R attack detection accuracy using proposed ensemble based feature selection method and it is compared with Filter and Wrapper based feature selection methods. From the **Fig. 4** it is observed that the U2R detection accuracy of the attacks are more in the proposed work when compared to the existing methods.

Fig. 5 shows the R2L attack detection accuracy using proposed ensemble based feature selection method and it is compared with Filter and Wrapper based feature selection methods. From the **Fig. 5** it is observed that the R2L detection accuracy of the attacks are more in the proposed work when compared to the existing methods.

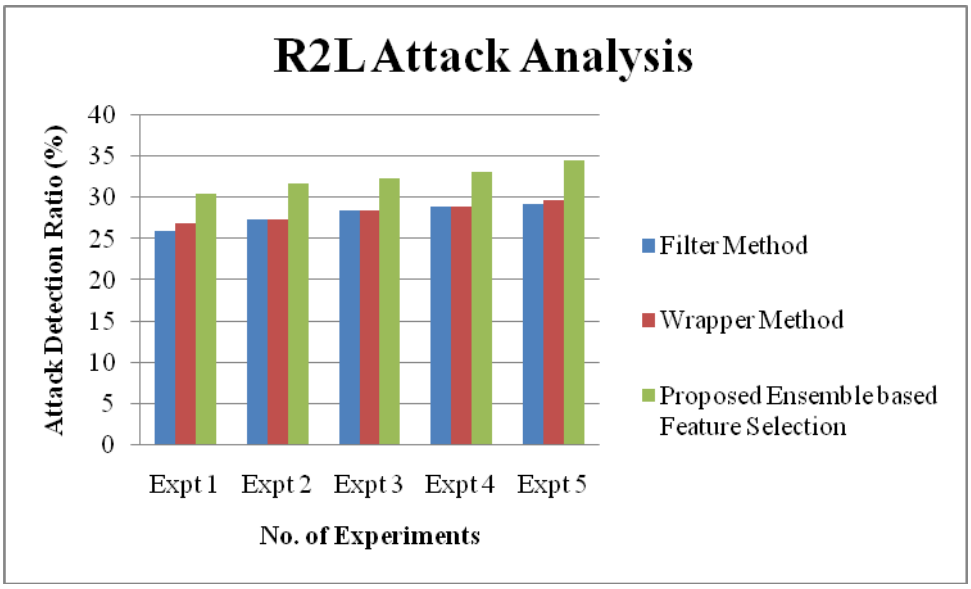

Fig. 5. R2L attack detection accuracy using proposed ensemble based feature selection method

Packet Delivery Ratio is another important factor to check the number packets are transmitted by the source and numbers of packets are received by the destination. **Fig. 6** shows the packet delivery ratio analysis for the proposed FSChE and it is compared with RNN-IDS [23] and DRFSA [16].

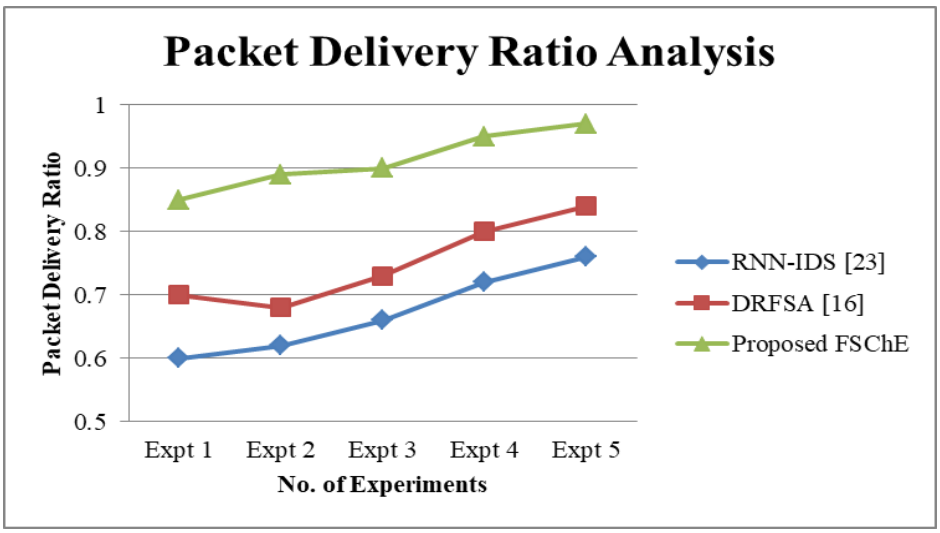

Fig. 6. Packet Delivery Ratio Analysis for ensemble based feature selection method

From **Fig. 6**, it is observed that the packet delivery ratio is more in proposed FSChE when compared to RNN-IDS [23] and DRFSA [16] because of combining Chi Squared feature selection with ensemble method.

Receiving packets on the time is a challenging issue in vulnerable WSN. This time delay analysis help us to find whether the packets are received with negligible delay or more delay.

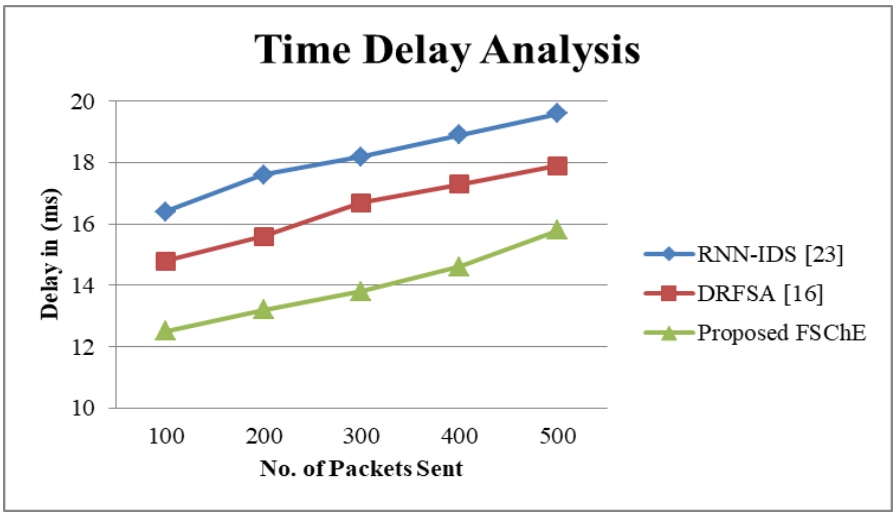

Fig. 7. Time Delay Analysis for ensemble based feature selection method

Fig. 7 shows the time delay analysis for the proposed FSChE and it is compared with RNN-IDS [23] and DRFSA [16]. From this analysis we observed that the time delay is reduced in the proposed FSChE when compared with other works such as RNN-IDS [23] and DRFSA [16].

5. Conclusion and Future Work

In this paper, an ensemble based optimal feature selection algorithm has been proposed for efficient intrusion detection in wireless sensor network. The proposed algorithm maintains the classification accuracy of the decision tree and random forest classifier and it uses a reduced set of features using chi-square algorithm from training data. The performance of the Feature Selection Algorithm using Chi Squared with Ensemble Method is tested on NSL KDD dataset. This dataset contains five different main classes: Normal, DoS, Probe, U2R and R2L. Attacks were detected with high accuracy when compared to the existing method. The delay time can need to be reduced further than the ensemble techniques proposed in this article. Future works in this direction is to ensemble other classifier algorithms with Random Forest and DT to achieve good classification accuracy and reduced delay time.

References

[1] S. Rajasoundaran, S. V. N. Santhosh Kumar, M. Selvi, S. Ganapathy, R. Rakesh, A. Kannan, "Machine learning based volatile block chain construction for secure routing in decentralized military sensor networks," *Wireless Networks*, vol.27, no.7, pp.4513-4534, 2021. Article(CrossRefLink)

[2] S. V. N. Santhosh Kumar, Y. Palanichamy, M. Selvi, S. Ganapathy, A. Kannan, S. P. Perumal, "Energy efficient secured K means based unequal fuzzy clustering algorithm for efficient reprogramming in wireless sensor networks," *Wireless Networks*, vol.27, pp.3873-3894, 2021. Article(CrossRefLink)

[3] M. Selvi, K. Thangaramya, S. Ganapathy, K. Kulothungan, H. K. Nehemiah, A. Kannan, "An Energy Aware Trust Based Secure Routing Algorithm for Effective Communication in Wireless Sensor Networks," *Wireless Personal Communications*, vol.105, pp.1475-1490, 2019. Article(CrossRefLink)

[4] K. Thangaramya, K. Kulothungan, S. I. Gandhi, M. Selvi, S. V. N. Santhosh Kumar, K. Arputharaj, "Intelligent fuzzy rule-based approach with outlier detection for secured routing in WSN," *Soft Computing*, vol.24, pp.16483-16497, 2020. Article(CrossRefLink)

[5] J. P. Anderson, "Computer Security threat monitoring and surveillance," Technical Report, James P. Anderson company, 1980. Article(CrossRefLink)

[6] I. Martins, J. S. Resende, P. R. Sousa, S. Silva, L. Antunes, and J. Gama, "Host-based IDS: A review and open issues of an anomaly detection system in IoT," *Future Generation Computer Systems*, vol.133, pp.95-113, 2022. Article(CrossRefLink)

[7] C.-M. Ou, "Host-based Intrusion Detection Systems inspired by Machine Learning of Agent-based Artificial Immune Systems," in *Proc. of IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA)*, pp.1-5, 2019. Article(CrossRefLink)

[8] S. Raj, K. N. Singh, N. K. Gupta, R. Nigam, B. Verma, S. Karsoliya, "High Accuracy of Hybrid IDS System using Evidence Theory and SVM ML Technique," in *Proc. of 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, pp.1261-1264, 2021. Article(CrossRefLink)

[9] W. T. Yue, M. Çakanyıldırım, "A cost-based analysis of intrusion detection system configuration under active or passive response," *Decision Support Systems*, vol.50, no.1, pp.21-31, 2010. Article(CrossRefLink)

[10] I. Dutt, S. Borah, and I. K. Maitra, "Immune System Based Intrusion Detection System (IS-IDS): A Proposed Model," *IEEE Access*, vol.8, pp.34929-34941, 2020. Article(CrossRefLink)

[11] S. Ganapathy, K. Kulothungan, S. Muthurajkumar, M. Vijayalakshmi, P. Yogesh, A. Kannan, "Intelligent feature selection and classification techniques for intrusion detection in networks: a survey," *EURASIP Journal on Wireless Communications and Networking*, vol.2013, no.1, 2013.

Article(CrossRefLink)

[12] S. Murugesan, R. S. Bhuvaneswaran, H. Khanna Nehemiah, S. Keerthana Sankari, Y. Nancy Jane, "Feature Selection and Classification of Clinical Datasets Using Bioinspired Algorithms and Super Learner," *Computational and Mathematical Methods in Medicine*, vol.2021, pp.1-18, 2021. Article(CrossRefLink)

[13] V. R. Elgin Christo, H. Khanna Nehemiah, B. Minu, and A. Kannan, "Correlation-Based Ensemble Feature Selection Using Bioinspired Algorithms and Classification Using Backpropagation Neural Network," *Computational and Mathematical Methods in Medicine*, vol. 2019, 2019. Article(CrossRefLink)

[14] S. Sinha, A. Paul, "Neuro-Fuzzy Based Intrusion Detection System for Wireless Sensor Network," *Wireless Personal Communications*, vol.114, pp.835-851, 2020. Article(CrossRefLink)

[15] M. Riecker, S. Biedermann, M. Hollick, "Lightweight energy consumption based intrusion detection system for wireless sensor networks," in *Proc. of SAC '13: Proceedings of the 28th Annual ACM Symposium on Applied Computing*, pp.1784-1791, 2013. Article(CrossRefLink)

[16] P. Nancy, S. Muthurajkumar, S. Ganapathy, S.V.N. Santhosh Kumar, M. Selvi, K. Arputharaj, "Intrusion detection using dynamic feature selection and fuzzy temporal decision tree classification for wireless sensor networks," *IET Communications*, vol.14, no.5, pp.888-895, 2020. Article(CrossRefLink)

[17] F. A. Khan, A. Gumaei, A. Derhab, and A. Hussain, "A Novel Two-Stage Deep Learning Model for Efficient Network Intrusion Detection," *IEEE Access*, vol.7, pp.30373-30385, 2019. Article(CrossRefLink)

[18] A. Khraisat, I. Gondal, P. Vamplew and J. Kamruzzaman, "Survey of intrusion detection systems: techniques, datasets and challenges," *Cybersecurity*, vol.2, 2019. Article(CrossRefLink)

[19] H. S. Hota and A. K. Shrivas, "Decision Tree Techniques Applied on NSL-KDD Data and Its Comparison with Various Feature Selection Techniques," in *Proc. of Advanced Computing, Networking and Informatics, Smart Innovation, Systems and Technologies*, vol.27, pp.205-211, 2014. Article(CrossRefLink)

[20] D. H. Deshmukh, T. Ghorpade, P. Padiya, "Intrusion detection system by improved preprocessing methods and Naïve Bayes classifier using NSL-KDD 99 Dataset," in *Proc. of 2014 International Conference on Electronics and Communication Systems (ICECS)*, pp.1-7, 2014. Article(CrossRefLink)

[21] M. S. Pervez, and D. Md. Farid, "Feature selection and intrusion classification in NSL-KDD cup 99 dataset employing SVMs," in *Proc. of the 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2014)*, pp.1-6, 2014. Article(CrossRefLink)

[22] N. Paulauskas, J. Auskalnis, "Analysis of data pre-processing influence on intrusion detection using NSL-KDD dataset," in *Proc. of 2017 Open Conference of Electrical, Electronic and Information Sciences (eStream)*, pp.1-5, 2017. Article(CrossRefLink)

[23] C. Yin, Y. Zhu, J. Fei, and X. He, "A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks," *IEEE Access*, vol.5, pp.21954-21961, 2017. Article(CrossRefLink)

[24] S. T. Ikram, A. K. Cherukuri, "Intrusion detection model using fusion of chi-square feature selection and multi class SVM," *Journal of King Saud University - Computer and Information Sciences*, vol.29, no.4, pp.462-472, 2017. Article(CrossRefLink)

[25] E. Tufan, C. Tezcan, and C. Acartürk, "Anomaly-Based Intrusion Detection by Machine Learning: A Case Study on Probing Attacks to an Institutional Network," *IEEE Access*, vol.9, pp.50078-50092, 2021. Article(CrossRefLink)

[26] L. Hakim, R. Fatma, Novriandi, "Influence Analysis of Feature Selection to Network Intrusion Detection System Performance Using NSL-KDD Dataset," in *Proc. of 2019 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)*, pp.217-220, 2019. Article(CrossRefLink)

Mr. S. Shyam Sundar currently pursuing his Ph.D from Anna University. He completed his M.Sc in Computer Science from Madurai Kamaraj University in 2002. He received fellowship under Visvesvaraya PhD Scheme for Electronics and IT for carry out his research work. His area of research is Wireless Networks and Machine Learning.

Dr. R. S. Bhuvaneswaran was a Professor in Ramanujan Computing Centre, College of Engineering Guindy, Anna University. He awarded with Ph.D. in wireless sensor network from Faculty of Information and Communication Engineering, Anna University in 2003. He carried out his Postdoctoral Research in Grid computing from faculty of information and communication engineering, Nagoya institute of technology, Japan from 2004 – 2006. He guided 15 research scholars and 5 currently under his supervisorship. He published more than 100 research articles in reputed International and National journals. His area of research was Mobile Computing, Wireless Sensor Networks, Image Security and Cryptography.

Dr. L. SaiRamesh is currently working as an Associate Professor in Department of Computer Science and Engineering, St. Joseph's Institute of Technology, Chennai, India. He completed his M.E. in CSE in 2007 and Ph.D. in 2015 from Anna University, Chennai, India. He has seventeen years of teaching and research experience. He published more than 90 research articles in reputed journal and international conferences. His area of research is Security in WSN, Cloud computing and Machine Learning.