**Indonesian Journal of Electrical Engineering and Computer Science**
Vol. 21, No. 1, January 2021, pp. 516~523
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v21.i1. pp516-523 

# Data mining approach to analyzing intrusion detection of wireless sensor network

**Md Alauddin Rezvi, Sidratul Moontaha, Khadija Akter Trisha, Shamse Tasnim Cynthia, Shamim Ripon**
Department of Computer Science and Engineering, East West University, Dhaka, Bangladesh

**Article Info**

***Article history:***

Received May 5, 2020
Revised Jul 4, 2020
Accepted Jul 18, 2020

***Keywords:***

Class imbalance
DoS attack
Intrusion detection system (IDS)
Wireless sensor network (WSN)

**ABSTRACT**

Wireless sensor network (WSN) is a collection of wireless sensor nodes which are distributed in nature and a base station where the dispersed nodes are used to monitor and the physical conditions of the environment is recorded and then these data are organized into the base. Its application has been reached out from critical military application such as battlefield surveillance to traffic, health, industrial areas, intruder detection, security and surveillance. Due to various features in WSN it is very prone to various types external attacks. Preventing such attacks, intrusion detection system (IDS) is very important so that attacker cannot steal or manipulate data. Data mining is a technique that can help to discover patterns in large dataset. This paper proposed a data mining technique for different types of classification algorithms to detect denial of service (DoS) attacks which is of four types. They are Grayhole, Blackhole, Flooding and TDMA. A number of data mining techniques, such as KNN, Naïve Bayes, Logistic Regression, support vector machine (SVM) and ANN algorithms are applied on the dataset and analyze their performance in detecting the attacks. The analysis reveals the applicability of these algorithms for detecting and predicting such attacks and can be recommended for network specialist and analysts.



***Corresponding Author:***

Shamim Ripon
Department of Computer Science and Engineering
East West University
A/2 Jahurul Islam City, Dhaka, Bangladesh
Email: dshr@ewubd.edu

## 1. INTRODUCTION

Nowadays wireless sensors network (WSN) [1] is a standout amongst the most rising and quickly developing fields in the world [2-3]. At first, it was designed to accelerate and facilitate military operations but later its application has been extended to health, traffic, industrial areas and threat detection [4]. WSN consists of specialized transducers with a communications infrastructure which uses radio to observe and record physical or environmental conditions [5]. It is built of local scattered sensors with limited sensing capability and transmission rate which is deployed in monitoring region through wireless communication [6]. The sensors are responsible for exchanging the environment information to construct a model of the monitored region [7]. To transfer data over the network, each sensor node consumes some energy. So, the lifetime of the network depends on how much energy spent on each transmission. In order to extend the life time of WSN, routing protocols are designed for reducing energy consumption of sensors. Some of the well-known routing protocols are LEACH, PEGASIS, TEEN, APTEEN and HEED [8].

The dataset we use in this paper was built using LEACH protocol [9]. LEACH is a protocol which is clustering, adaptive, and self-organizing [10]. LEACH protocol is aimed to establish clusters of nodes for distributing the energy in all of the network nodes [11]. There is a Cluster Head (CH) node in each protocol which is responsible for gathering data from its cluster member nodes and forward them to the Base Station (BS) or Sink. The location of BS is far away from the sensor nodes. The LEACH protocol aims for the reduction of the consumption of energy that is necessary to maintain clusters to prolong the lifetime of a WSN [8]. The limitation of computation power, memory and battery lifetime of sensor nodes also increases the insecurity of WSN. The aim of most of the attacks in WSN is either limiting or eliminating network ability for performing its expected work procedures [12]. One of those attacks is DoS attack [13]. This attack is performed by hardware failure, bug, resource exhaustion, malicious broadcasting of high energy signals and minimize the performance of the network.

Many pieces of research [14] show that many existing prevention mechanisms are not sufficient to secure the data packets of the network and maintain the service of WSN [15]. They also cannot prevent the network from all the attacks experienced by them. For this reason, the detection based technique combined with a prevention based technique will be more efficient [16]. So, intrusion detection system is vital for monitoring suspicious activity in network traffic and issues alert if such activity is detected. Through intrusion detection system we can detect an attack early and save the network from various malicious attacks. The author in [13] demonstrates a process for detecting four types of DoS attacks in WSN. They have applied random forest classifier to detect attacks on a dataset and have achieved the best performance for Blackhole, Flooding, Grayhole, Scheduling (TDMA) attacks and Normal behavior. M. Ahsan Latif and M. Adnan in [17] have developed three different ANN-based model so that the behavior of the network traffic can be discovered. They have proposed a model that enjoying an intelligent agent for monitoring the traffic pattern at the level of the base station.

The dataset used in this paper contains four types of DoS [18] attack: Grayhole, Blackhole, Flooding and TDMA [19-20]. By applying suitable data mining techniques, it is possible to actively detect intrusion in network because it helps to discover patterns in large datasets [21]. Then different algorithms help to classify and predict intrusion. By using data mining, we have discovered patterns of attacks and also can predict whether it is an attack or not when a new request arrives. After applying various mining algorithms, it is possible to analyze the performance of intrusion detection as well as compare various techniques with each other to identify the best algorithm for intrusion detection.

In [22] the author has presented some DoS attack and have suggested some solution for detecting and solving certain attack. They have categorized some DoS attacks according to protocol stack layers and have described sinkhole, Hello flooding, wormhole, selective forwarding attack for network layer and flooding attack for the transport layer. Luigi Coppolino et al. [21] has proposed a hybrid intrusion detection system for WSN that uses for misuse and anomaly-based detection techniques. They have used decision tree as classification algorithm for the detection process.

The rest of the paper is organized as follows. Section 2 outlines the proposed method applied in this. Section 3 describes an overview of the dataset that is used in the experiment. Section 4 provides implemented part of our work. Section 5 presents results and analysis. Section 6 demonstrates how class imbalance problem can be handled for imbalance dataset. Finally, Section 7 draws a conclusion of the paper to summarize the work and to outline our future plan.

## 2. PROPOSED METHOD

After collecting the dataset, we performed preprocessing for the convenience of our experiment. At first, the attack names are converted into numeric values. The values 0, 1, 2, 3, 4 are used for normal(non-attack), Blackhole attack, Grayhole attack, Flooding and TDMA attack respectively. A binary classification has also been performed as attacks and non-attack (normal) type and assigns 1 and 0 respectively.

After processing the dataset, it has been divided into 60% and 40% as training and testing data applying 10-fold cross-validation. Then we have chosen some classification algorithms which are mostly used in various literature and apply them on the dataset to detect intrusion both for type of each attack and for binary classification. In this paper, we have applied KNN, Naïve Bayes, Support Vector Machine (SVM), and Logistic regression algorithm to see which algorithm can detect intrusion more accurately. ANN has also been applied to present a comparative analysis with the data source in [4]. Various standard metrics are used to measure the obtained result. The experimental result analysis can reveal important information regarding the types of algorithms and their suitability for detecting various types of attacks. The proposed framework is illustrated in Figure 1.

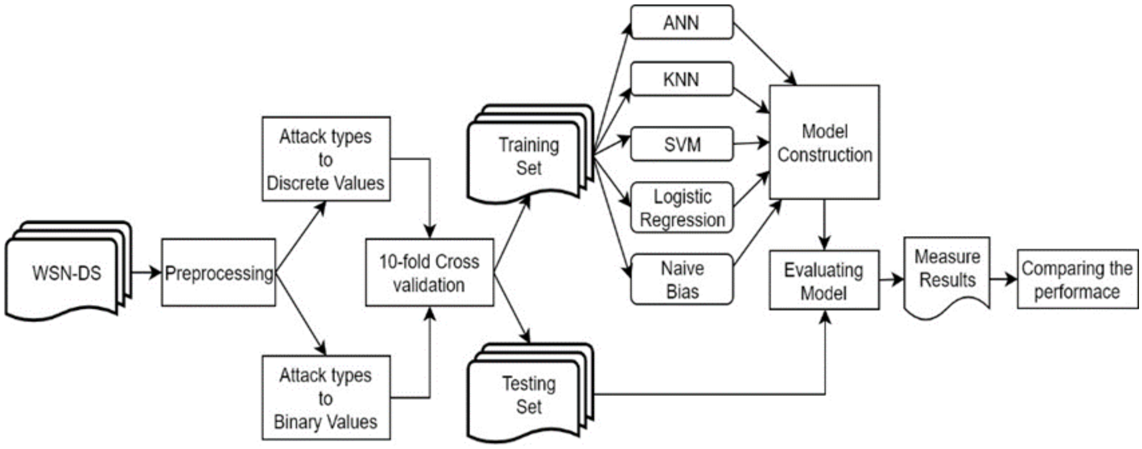


Figure 1. Proposed method

## 3. DATASET OVERVIEW

The dataset that has been collected from the work in [4]. It is called WSN-DS [9]. To develop a specialized dataset for the WSN intrusion detection system the author has used LEACH protocol. This IDS dataset is consistent with the characteristics of WSN and is amenable for detection and classification of four types of DoS attacks. The dataset contains 374661instances and 23 attributes. Among these 23 attributes RSSI, Max distance to CH, Average distance to CH, Current energy are not used for detecting DoS attacks. Table 1 shows the attributes and their types.

Table 1. Data type of dataset attribute

| Attribute name | Data type |
|---|---|
| Id, time, Is_CH, who_CH, ADV_S, ADV_R, JOIN_S, JOIN_R, SCH_S, SCH_R, Rank, DATA_S, DATA_R, Data_Sent_To_BS, Send_code | Integer |
| Dist_To_CH, Dist_CH_To_BS, Consumed energy | Float |
| Attack type | Polynomial |

The attacks that this dataset contains are, Grayhole, Blackhole, Flooding and TDMA. The distribution of these attacks and the non-attacks (normal) are shown in Figure 2. We have compared each attribute of the dataset with attack attribute to identify which attribute is responsible for detecting which type of attack. The result is illustrated in Table 2.

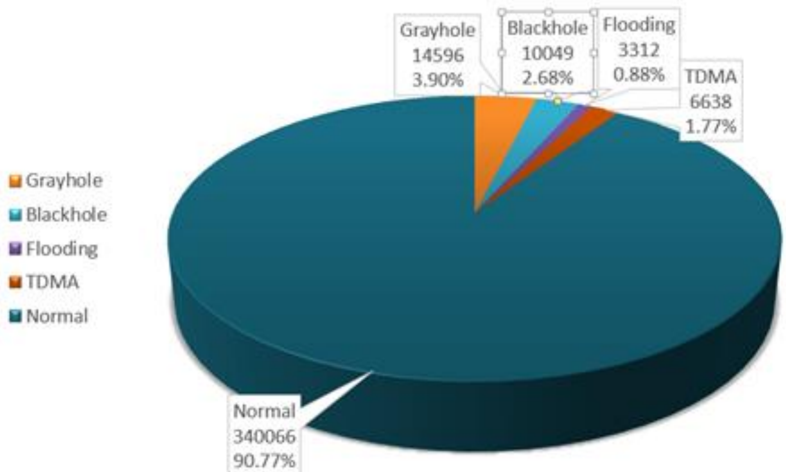


Figure 2. Attack statistics in dataset

Table 2. Summary of attributes of various attacks

| Attack name | Attribute name |
|---|---|
| Grayhole | Time, Is_CH, who_CH, adv_r, join_r, Data_R data_sent_to_bs |
| Blackhole | Time, Is_CH, who_CH, adv_r, join_r, Data_R, data_sent_to_bs |
| Flooding | Is_CH, who_CH, consumed_energy, adv_s, adv_r, data_sent_to_bs, Dist_CH_to_BS |
| TDMA | Time, Is_CH, join_r, sch_s, data_sent_to_bs |

## 4. IMPLEMENTATION

The machine that has been used for this work is a Core i3 CPU with speed 2 GHz, and Jupyter Notebook is used as a working environment. To predict each type of attack, the string data of attack type attributes are converted into numeric value as mentioned in the proposed method and apply various algorithms on the dataset. In this paper, first, we have trained the model then fit the model by a fit() function to predict accuracy, and derive confusion matrix from testing data. Table 3 represents some of the parameters that have been used in the mining algorithms.

*K-nearest neighbor:* A supervised machine learning algorithm which is mostly used for classification. KNN classifies new cases based on similarity measures. KNN is effective for large training data.

*Naïve Bayes:* A supervised classification algorithm that is particularly used for large data. A Naive Bayes classifier is a machine learning model and it is probabilistic.

*Support vector machine:* SVM is a supervised learning model which is used for classification and regression analysis.

*Logistic regression:* It is a classification algorithm used to predict probability of binary value based on one or more independent variables.

Table 3. Parameter used in the algorithms

| Algorithm | Parameter and Description | Value |
|---|---|---|
| KNN | n-neighbors - Number of neighbors to use by default | 3 |
| Naïve Bayes | Alpha - It is additive smoothing parameter (0 for no smoothing) | 1.5 |
| SVM | C – It means penalty parameter C of the error term | 1.0 |
| Logistic Regression | C – Inverse of regularization strength and must be a positive float | 1.0 |

### 4.1. Applying machine learning algorithms

Four widely used algorithms are applied in the experiment. The classification performances are measure by applying widely used metrics, such as, Accuracy, Precision, Recall, F1Score, and Error (1)-(4).

$$Accuracy = \frac{tp+tn}{tp+tn+fp+fn} \quad (1)$$

$$Precision = \frac{tp}{tp+fp} \quad (2)$$

$$Recall = \frac{tp}{tp+fn} \quad (3)$$

$$F1\;Score = 2 * \frac{precision*recall}{precision+recall} \quad (4)$$

KNN makes clusters to classify each and every attacking and non-attacking classes. In the testing part of the dataset, there are 4077 Blackhole attacks among which KNN can detect 3794 attacks, from 5938 Grayhole attacks it can detect 5169 attacks, 1006 Flooding attack can be detected from 1310, and from 2651 TDMA attack 2143 is predicted. Among the 135889 normal data, the detected number is 135323.

Naïve Bayes can detect only 1399 Blackhole attacks from 4077, from 5938 Grayhole attacks it can detect 2760 attacks, 1075 Flooding attacks are detected from 1310, for TDMA, 2148 attacks are predicted from 2651 and for normal data, 108690 are detected from 135889.

## 5. RESULT AND ANALYSIS

This section describes the application of various machine learning techniques on the training and testing set of data. Learning algorithms are applied in two phases. The first phase is a multi-class classification phase, where all the algorithms are applied to classify all the attacks in the dataset. In the second phase, binary classification is performed where all the attacks are considered one class (attack class) and non-attacks are the other class (normal).

Support vector machine (SVM) can detect only 879 attacks from 4077 Blackhole attack SVM, it cannot predict any Grayhole and Flooding attack, only 411 TDMA attacks are predicted from 2651 and normal data is detected 135883 from 135889. Logistic Regression predicts 4055 attacks from 4077 Blackhole attack, from 5938 Grayhole attacks it can detect 930 attacks, 905 Flooding attacks can be detected from 1310, 2194 TDMA attacks are predicted from 2651 and normal data is detected 135435 from 135889. Table 4 illustrates the comparison of all the predicted attacks by each of the algorithms. It can be found that no single technique is a winner here. However, LR performs better in most of the cases.

Table 4. Comparison of various attack prediction

| Attack Type | Total | Prediction KNN | NB | SVM | LR |
|---|---|---|---|---|---|
| Blackhole | 4077 | 3794 (93.06%) | 1399 (34.31%) | 879 (21.56%) | 4055 (99.46%) |
| Grayhole | 5938 | 5169 (87.05%) | 2760 (46.48%) | 0 (0.0%) | 930 (15.66%) |
| Flooding | 1310 | 1006 (76.84%) | 1075 (82.06%) | 0 (0.0%) | 905 (69.08%) |
| TDMA | 2651 | 2143 (80.84%) | 2148 (81.03%) | 411 (15.50%) | 2194 (82.76%) |
| Normal | 135889 | 135323 (99.58%) | 108690 (79.98%) | 135883 (100%) | 135435 (99.67%) |

## 5.1. Artificial neural network (ANN)

ANN is applied with 2 hidden layers and calculates the precision and accuracy to compare the results with the previously published result [9] using 10-fold cross-validation. The comparative analysis of the obtained results is illustrated in Table 5. Our experiment produces almost the same result as in the original work. For ANN with 2 hidden layers we get 98.56% accuracy for detecting attack and in previous work, the author of [6] found 98.53% accuracy for detecting attack.

Table 5. Comparison of results between our work and previous work [9]

| | TPR | | FPR | | FNR | | TNR | | Precision | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Present Work | Ref. | Present Work | Ref. | Present Work | Ref | Present Work | Ref. | Present Work | Ref. |
| Normal | 99.79% | 99.80% | 1.85% | 2.00% | 0.22% | 0.20% | 98.15% | 98.00% | 99.80% | 99.80% |
| Flooding | 98.97% | 98.50% | 0.09% | 0.10% | 1.03% | 1.50% | 99.91% | 99.90% | 90.30% | 90% |
| TDMA | 91.99% | 91.50% | 0.01% | 0.00% | 8.01% | 8.50% | 99.99% | 100.00% | 99.10% | 99.20% |
| Grayhole | 85.19% | 86.70% | 0.64% | 0.70% | 14.81% | 13.30% | 99.36% | 99.30% | 84.40% | 83.20% |
| Blackhole | 80.29% | 77.80% | 0.57% | 0.50% | 19.71% | 22.20% | 99.43% | 99.50% | 79.60% | 81% |

## 5.2. Binary classification

In binary classification, all the attack classes are considered as Attack and rest are considered Normal. Table 6 illustrates the number of predictions by each of the applied algorithms. It can be observed that Naïve Bayes shows the best performance among all the applied algorithms. Comparison of classification accuracies are illustrated in Figure 3.

Table 6. Comparison of number of attack prediction

| Attack Type | Total | Predicted KNN | NB | SVM | LR |
|---|---|---|---|---|---|
| Normal | 135889 | 135232 | 116268 | 135889 | 134907 |
| Attack | 13976 | 12657 | 13819 | 103 | 10008 |

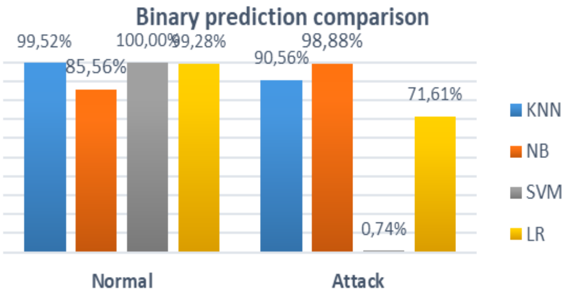


Figure 3. Binary attack prediction comparison

### 5.3. Performance for splitting data

Apart from analyzing the performances of various learning techniques, we have also examined whether the dataset size has any impact on the classification performance of the algorithms. We have experimented accuracy by splitting data sets into the sizes of 100K, 200K, 300K+ instances and applied all four algorithms. Figure 4 shows the comparison of accuracies for splitting data. It is observed that only the accuracy of KNN is increasing with the increase in data size. However, for SMV the accuracy is decreased for 200k data size. The other two algorithms do not have any considerable effect over the data seize.


Figure 4. Accuracy comparison for splitting data

### 5.4. Comparative analysis

After applying all the mentioned algorithms for each attack type the overall performance is evaluated by calculating accuracy, error, precision, recall and F1-Score for the dataset and illustrated in Table 7. From table, we can see that the accuracy of KNN is 98.4%, Naïve Bayes is 86.88%, Logistic Regression is 96.72%, SVM is 91.22% and ANN has 98.56%. In the experiment, ANN and KNN show better performance than the others. Figure 5 represents the comparison of the accuracy of applied algorithms.

Apart from calculating classification accuracies, Precision and Recall are also calculated. For this experiment, Recall is crucial as it is required to know how good an algorithm can predict each type of attack among the total number of attacks in the dataset. In the experiment, the performance of SVM is very poor while Naïve Bayes performs very well.

Table 7. Comparative analysis of all algorithms

| Algorithm | Accuracy (%) | Error (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|---|
| KNN | 98.4 | 1.6 | 95.69 | 91.50 | 93.55 |
| Naïve Bayes | 86.88 | 13.12 | 41.39 | 98.87 | 58.35 |
| Logistic Regression | 96.72 | 3.28 | 90.99 | 71.81 | 80.28 |
| SVM | 91.22 | 8.78 | 95.95 | 5.78 | 7.16 |
| ANN | 98.56 | 0.27 | 90.66 | 91.24 | 90.954 |

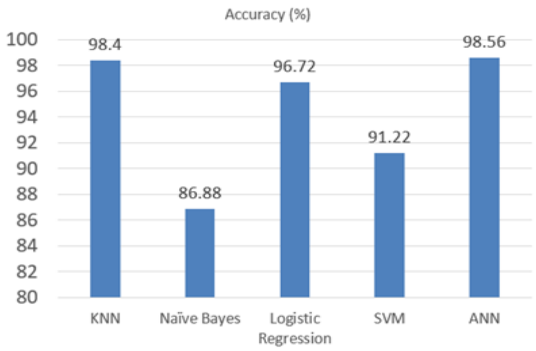


Figure 5. Comparison of accuracy of different algorithm

## 6. MANAGING CLASS IMBALANCE PROBLEM

Class imbalance problem [23] in a dataset refers to imbalance distribution of classes in the dataset. The WSN-DS dataset that has been used in the experiment suffers from such problem. It can be observed from Figure 2. that the total number of attacks in the dataset is far less than that of non-attacks. Having imbalanced distribution of classes in a dataset, the classification accuracy of an algorithm might be very high while the classifier may completely fail to predict the minority class. In the present work, intrusion detection is the main focus, however, in the dataset, it is the minority class. Hence, it is an essential task to handle the class-imbalance problem to properly predict the attacks. Among various available techniques, SMOTE (Synthetic Minority Over-sampling Technique) [24] is a method that has been applied very widely. We experimented to observe the effect of SMOTE [25] while classifying the minority classes (all types of attacks in this work).

Table 8. illustrates the experimental result. For the experimental purpose, only Naïve Bayes and Logistic Regression are applied here. It can be observed that in most of the cases, after applying SMOTE, the prediction rate of attacks is increased (except for Grayhole). Such experiment reveals that even if the dataset is imbalanced, the minority classes can now be classified properly after applying SMOTE. As in earlier experiments, Logistic Regression also performs better in this experiment.

Table 8. Effect of SMOTE on classifying attack types

| Attack Type | Total | Predictions | | | |
|---|---|---|---|---|---|
| | | Naïve Bayes | | Logistic Regression | |
| | | Before SMOTE | After SMOTE | Before SMOTE | After SMOTE |
| Normal | 135889 | 118167 (86.96%) | 113460 (83.49%) | 134958 (99.31%) | 113389 (83.44%) |
| Flooding | 1310 | 1143 (87.25%) | 1250 (95.42%) | 1053 (80.38%) | 1274 (97.25%) |
| TDMA | 2651 | 958 (36.14%) | 1610 (60.73%) | 1930 (72.80%) | 2064 (77.86%) |
| Grayhole | 5938 | 3082 (51.90%) | 2736 (46.08%) | 3003 (50.57%) | 3206 (53.99%) |
| Blackhole | 4077 | 3808 (93.40%) | 3940 (96.64%) | 2816 (69.07%) | 3826 (93.84%) |

## 7. CONCLUSION

Various data mining techniques for detecting DoS attacks from WSN-DS dataset have been presented in this paper. After using KNN, Logistic regression, SVM, Naïve Bayes and ANN we have found out that ANN (98.56%) and KNN (98.4%) performed better for intrusion detection in our dataset. A comparison has been made with the existing work from where the dataset has been collected. Our experiment shows similar result to that work and it confirms that our experiment has been conducted using an appropriate and standard setup. As the dataset is obtained by using LEACH protocol, selected algorithm is proper for this type of dataset. However, these data mining techniques can also be applied to the dataset obtained by other routing protocols. We can use KNN and ANN algorithm in IDS of a WSN which will help to detect attack in network and give alarm to start the prevention mechanism. This can be used in real time application. As the dataset is imbalance, we also show how to manage the imbalance the dataset and perform classifications. The experimental result shows that after applying SMOTE the algorithms can predict attacks than that without balancing the dataset.

In future we would like to apply other algorithms to find out which will have better performance than KNN and ANN algorithm. We will use different hybrid algorithms and also will improve the performance of Naïve Bayes, SVM and Logistic Regression. It is also beneficial to develop various association rules among the attributes that can support users to manage network requests to predict various types of attacks.

## ACKNOWLEDGEMENTS

We are grateful to Dr. Iman Almomani, (Associate Professor, Prince Sultan University, Women Campus, Riyadh, KSA) for proving us the dataset that has been used in our experiment

## REFERENCES

[1] I. F. Akyildiz and M. C. Vuran, "Wireless sensor networks," *John Wiley & Sons*, vol. 4, 2010.
[2] A. H. Farooqi and F. A. Khan, “Intrusion Detection Systems for Wireless Sensor Networks: A Survey,” *In International Conference on Future Generation Communication and Networking, Springer*, Berlin, Heidelberg, pp. 234–241, 2009.
[3] Ghosal, A., & Halder, S., "Intrusion detection in wireless sensor networks: Issues, challenges and approaches," *In Wireless networks and security*. Springer, pp. 329-367. 2013.

[4] A. H. Farooqi and F. A. Khan, “A survey of Intrusion Detection Systems for Wireless Sensor Networks,” *Int. J. Ad Hoc Ubiquitous Comput*., vol. 9, no. 2, pp. 69-83, 2012.

[5] Ardiansyah, A. Y., & Sarno, R., "Performance analysis of wireless sensor network with load balancing for data transmission using xbee zb module," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS),* vol. 18, no. 1, pp. 88-100, 2019.

[6] Q. Liao and H. Zhu, “An Energy Balanced Clustering Algorithm Based on LEACH Protocol,” *Appl. Mech. Mater*., vol. 341–342, pp. 1138–1143, Jul. 2013.

[7] S. Umrao, A. Kumar, and P. Umrao, “Security attacks and their countermeasures along with node replication attack for time synchronization in wireless sensor network,” in *International Conference on Advanced Nanomaterials & Emerging Engineering Technologies*, pp. 576–581, 2013.

[8] Jalil Jabari Lotf, M. Hosseinzadeh, R. M. Alguliev, “Hierarchical routing in wireless sensor networks: a survey,” in *2010 2nd International Conference on Computer Engineering and Technology*, pp. V3-650-V3-654, 2010.

[9] Almomani, I., Al-Kasasbeh, B., & Al-Akhras, M., "WSN-DS: A dataset for intrusion detection systems in wireless sensor networks," *Journal of Sensors*, 2016..

[10] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-efficient communication protocol for wireless microsensor networks,” in *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences,* vol. 2, p. 10, 2000.

[11] A. A. Hussien, S. W. Al-Shammari, M. J. Marie., "Performance evaluation of wireless sensor networks using LEACH protocol," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, no. 1, pp. 395-402, 2020.

[12] C. Ioannou, V. Vassiliou, and C. Sergiou, “An Intrusion Detection System for Wireless Sensor Networks,” in *2017 24th International Conference on Telecommunications (ICT)*, pp. 1–5, 2017.

[13] T.-T.-H. Le, T. Park, D. Cho, H. Kim, “An Effective Classification for DoS Attacks in Wireless Sensor Networks,” in *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 689–692, 2018.

[14] Zhang, R., & Xiao, X., "Intrusion Detection in Wireless Sensor Networks with an Improved NSA Based on Space Division," *Journal of Sensors*, 2019.

[15] H. Suhaimi, S. I. Suliman, I. Musirin, A. F. Harun, and R. Mohamad, "Network intrusion detection system by using genetic algorithm," *Indones. J. Electr. Eng. Comput. Sci. (IJEECS)*, vol. 16, no. 3, pp. 1593–1599, 2019.

[16] F. Josephin and N. Ramaraj, “Prevention and Detection of Selective Forwarding Attack in Wireless Sensor Networks,” *Int. J. Sci. Res. Innov.*, vol. 1, pp. 12–17, 2016.

[17] M. A. Latif, M. Adnan, "ANN-Based Data Mining for Better Resource Management in the Next Generation Wireless Networks," in *2016 International Conference on Frontiers of Information Technology*, pp. 35–39, 2016.

[18] A. P. Abidoye and E. O. Ochola, "Denial of Service Attacks in Wireless Sensor Networks with Proposed Countermeasures," *in Information Technology - New Generations*, pp. 185–191, 2018.

[19] H. Kalkha, H. Satori, and K. Satori, "Preventing Black Hole Attack in Wireless Sensor Network Using HMM," in *Procedia Computer Science*, vol. 148, pp. 552–561, 2019.

[20] A. Ahmed, M. I. Channa, and U. A. Khan, "Performance Evaluation of Wireless Sensor Network in Presence of Grayhole Attack," *Quest Res. J.*, vol. 13, no. 1, pp. 1–8, 2014.

[21] L. Coppolino, S. DAntonio, A. Garofalo, and L. Romano, "Applying Data Mining Techniques to Intrusion Detection in Wireless Sensor Networks," in *2013 Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, pp. 247–254, 2013.

[22] Ž. Gavrić and D. Simić, "Overview of DOS attacks on wireless sensor networks and experimental results for simulation of interference attacks," *Ingeniería e Investigación*, vol. 38, no. 1, pp. 130–138, 2018.

[23] N. Japkowicz and S. Stephen, "The class imbalance problem: A systematic study," *Intell. Data Anal.*, vol. 6, no. 5, pp. 429–449, Nov. 2002.

[24] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.

[25] N. Santoso, W. Wibowo, H. Hikmawati “Integration of synthetic minority oversampling technique for imbalanced class," *Indonesian Journal of Electrical Engineering and Computer Scienc (IJEECS)*, vol. 13, no. 1, pp. 102-108, 2019.