Classification of DoS Attacks in Wireless Sensor Network with Artificial Neural Network

Munawar Hussain, Jiadong Ren, and Awais Akram
(Corresponding author: Munawar Hussain)

Department of Information and Engineering, Yanshan University
438 Hebei Street West Section, Haigang, Qinhuangdao, Hebei, China
(Email: munawarjut@yahoo.com)
(Received June 12, 2019; Revised and Accepted Dec. 3, 2019; First Online Feb. 28, 2020)

Abstract

Due to deployment in sensitive military areas and other security applications, wireless networks are becoming a famous research spot in the field of computer science. To ratify the security and reliability in such kinds of application intrusion detection system can play an important role. There is a need of intrusion detection system, which has the capability to detect a large number of possible threats in wireless sensor networks. This article contains a customized dataset for wireless sensor network that can be categorized into four types of DoS attacks (Blackhole, Grayhole, Flooding & Scheduling attacks). For the experimental purpose, since it is highly used in WSN, Low Energy Aware Cluster Hierarchy (LEACH) protocol has been used in this research work. Using NS 2 network simulator to model a scheme which define to collect network traffic and create the dataset. Artificial Neural Network has been applied to train the dataset to classify it into different DoS attacks. Experimental work performed here gives high classification rate and accuracy for mentioning attacks with the help of proposed dataset. In future, suggested method can be useful for more attacks like Sybil/Wormhole presented in datalink layer as DoS attacks.

Keywords: Artificial Neural Network; DoS Attacks; LEACH Protocol; Machine Learning; Multilayer Perceptron (MLP); Wireless Sensor Network

1 Introduction

Wireless Sensor Network (WSN) is becoming more imperative research area due to its large employment in diverse kind of applications, for example health care, building monitoring, smart cities deployment and sensitive military areas *etc.* [14, 18]. The deployment of WSN nodes are usually done in wide areas to collect the data and to transmit the data in intelligent way to the base station [8]. The transmission of data in WSN is based on dedicated WSN protocols. Securing WSN is becoming main challenge because of limited resources of WSN like processing power, memory and battery backup [2]. Generally WSN nodes are deployed in an unprotected and hostile environment. This makes nature of WSN more vulnerable [13].

These limitations put constraint on existing security methods like cryptography, which are not always enough for this type of networks. WSN is always surrounded of attacks due to its deployment in open areas with limited resources. In WSN during frequent transmission of the packets, attackers can compromise with sensor node and may cause drop of the message, inject and alteration of the message, interrupt the integrity of the data as well as waste the energy resources. Denial of Services (DoS) attack is most common and dangerous attack for WSN. This attack may cause to interrupt and hang the WSN services [6].

Due to weak security prevention arrangements an Intrusion Detection System it is required to detect the unwanted attacks (known and unknown). IDS is complex for WSN due to its small size nodes and limited hardware. We have also found out that there is no specific dataset which have normal and abnormal traffic logs for these kind of experimental work [17].

In the light of above mentioned challenges IDS designing need two main characteristics. First IDS must be highly accurate for any attack detection secondly it must have low weight with minimum cost or overhead. In this paper we have built a specific dataset for WSN which have attack model which can be able to categorize DoS attack. LEACH protocol used to combat the WSNs? energy consumption challenges.

The remaining paper is distributed as mentioned. Section II included the idea of LEACH protocol and related work. Section III shows the dataset description and creation. Section IV mentioned experimental results and discussion. Last section summarized the conclusion and future work.

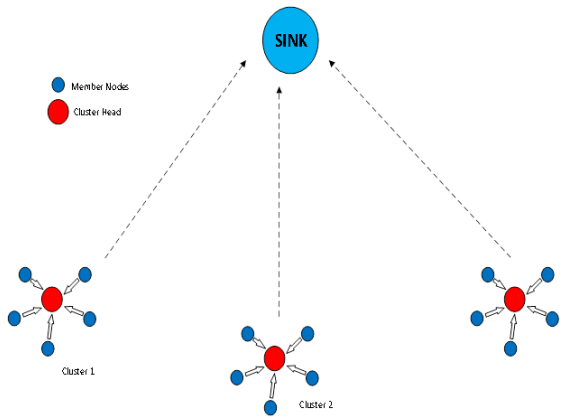


Figure 1: Nodes structure in LEACH routing protocol

2 Idea of LEACH Protocol & Related Work

The main Idea of the proposed LEACH Protocol was to reduce the energy consumption and to enhance the lifetime of sensor network, it?s a protocol based on clustering with self-organizing capability [9]. In LEACH usually BS (Base Station) is located far away from the nodes and nodes are similar to each other with limited space and power. Figure 1 shows the arrangement of nodes in LEACH which contain two steps on every round. Round one is step round where cluster are made and second round known as steady round where sensed data transferred to sink node. LEACH is still being mentioned and explored by many researchers. Authors in [4, 20] reviewed many clustering routing methods depending upon LEACH protocol for wireless sensor networks with detailed discussions and judgements. Another author in [12] presented improved versions of LEACH protocol. Authors have compared and discussed few features of LEACH protocol. In [3] authors have developed mechanisms to reduce the amount of Redundant data transmission in WSN to reduce the overall energy of the network. In [15] author evaluated and developed new clustering based principles for heterogeneous WSN based on LEACH protocol. Author [19] introduced LEACH Inner Clustering Election method depends upon LEACH. To improve clustering LEACH-ICE select new cluster head (CH) into the cluster which energy of current CH become lower as mentioned threshold. In the research [7] author planned an algorithm to select CH selection for wireless sensor network by managing distance among CH uniform distribution of CHs performed. In experimental area new method showed good performance for consumption of energy and life time of network.

3 Attack Model

The main aim of DoS is to create commotion in services by trying to limit the access to a service or machine instead of destabilizing the service itself. DoS attacks include

Black holes are places in a network when the network's incoming or outgoing traffic starts getting dropped or discarded silently or without getting the sources informed. The source are not even notified that their packets are not reaching the intended destinations. The only way to detect the Black holes in a network topology is by monitoring the lost traffic [16]. Algorithm 1 used for this type of attack. Flooding as its name states is made by intentionally sending a very high volume of traffic to a network node such that it can not allow or examine the permitted network traffic. For example in a TCP connection establishment during the three way handshaking an attacker can repeatedly, may be huge amount of times, initiate the handshake SYN and not respond afterwards. Since the listening server waits for the third handshake step that is SYN-ACK every time and dedicating small amount of resources to keep that session open. Eventually the server cannot handle any more resulting in no legitimate users to be able to log on [16]. Algorithm 3 used for this kind of attack.

Grayhole attacks are special case of DoS attacks and are kind of selective forwarding attacks. The nodes selectively discards or drops few, not all, packets which were supposed to be forwarded along the path [11]. Algorithm 2 used for this type of attack.

Wormhole attacks rre again a special type of DoS attack that misleads routing operations. The malicious nodes create tunnels through which the packets are replayed to malicious nodes hence corrupting the network routing procedures [10]. Algorithm 4 used for this kind of attack.

In recent days many researchers are trying to find the solutions for DoS attacks in WSNs but mainly they have found one or two or partial solutions with high energy consumption [21, 22]. So a suitable system should be developed to mention and classify different attitude of DoS of attacks.

IDS has become essential part of security mechanism while IDS application for WSN is still challengeable. There are two basic part of IDS feature extraction and modelling algorithm. Feature extraction id responsible to describe measured attributes which associates to IDS function. Modelling attributes responsible for accuracy and proficiency in IDS. Few important IDS are described as Monitoring, Analysis, Detection and three component described as action in IDS including Logging, Alarming and Prevention as shown in Figure 2.

4 Dataset Description & Creation

The main aim in this paper to create a dataset to classify the DoS attacks. Many works has been done to improve the IDS strategies. Important factor involved in this process is the data used for testing and training of the detection model. The better the data quality the better

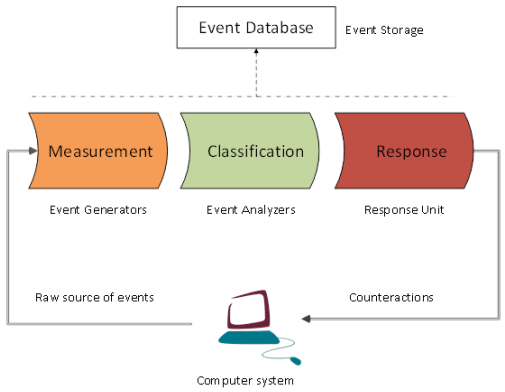


Figure 2: Nodes structure in LEACH routing protocol

the offline intrusion detection system can be analysed. In the research of IDS techniques the KDD data set (Knowledge Discovery and Data Mining) is a well-known benchmark [1]. Mostly KDD dataset has been constructed for LAN environments only, these dataset were not specific for wireless environment although many researchers used it for intrusion detection purposes. There is a need to develop a dataset for WSN to observe normal and abnormal behaviour. In order to create the dataset for WSN we needed the dataset related to wireless environment which can help in detection classification of DoS attacks. In this study distribution of the traffic among the nodes and every node is responsible in monitoring. Dataset attributes are described in Table 1.

5 Mathematical Model for LEACH

For WSN dataset construction a mathematical model developed to ensure the correctness of dataset, the term used in this model are following Table 2.

This model used to analysis in different phases including,

Advertisement phase: This step used to calculate the number of advertisements send by CHs and receive by CMs.

Cluster setup: This step responsible to calculated the join requests message send by nodes and received by related CHs.

Transmission of data: This step responsible to calculate the sensed data delivered to BS.

6 Experimental Results and Discussion

For collection of data NS2 simulation software used with following parameters settings as shown in Table 2. In this paper LEACH protocol is used to create the dataset,

Algorithm 1 Blackhole attack algorithm (Algorithm 3):

```
Begin
if rSNi ¡ TSNi then
   SNi=CH
else
   SNi=CM.
   if CHj, J = NC then
      X CMs join CHj
      CHj create TDNA schedule
   end if
   if CHj=MN  then
      Execute attack by dropping packets
   else
      Send data to BS

   end if
end if
End
```

Algorithm 2 Grayhole attack algorithm (Algorithm 3):

```
Begin
if rSNi ¡ TSNi then
   SNi = CH
else
   SNi = CM
   if CHj, J NC then
      X CMs join CHj
      CHj create TDNA schedule
   end if
   if CHj=MN  then
      CHj = Adv_CH (with high broadcasting)
   else
      CHj = Adv_CH (normal broadcasting)

   end if
end if
if xCMs join CHj  then
   CHj = Adv_CH (with high broadcasting)
else
   CHj create TDMA schedule
   xCMs send data to related CHj in associated
   TDMA time slice
end if
End
```

Dataset contain 374668 records which can be categorized into four kinds of DoS attack (Blackhole, Grayhole, flooding and Scheduling). Multilayer Perceptron (MLP) Artificial neural network classifier configuration used in this paper. MLP is highly famous ANN variation which enables to configure more than one layer ANN, and useful to build complex relation among input and output values.

Due to different performance confusion metrics we have used following performance matrix. These performance matrix are True Positive Rate (TPR), True Negative (TN), False Positive (FP), False Negative (FN) and Pre-

Table 1: Dataset attributes

Node ID	Matchless ID for differentiate the sensor node
Time	Simulation time
RSSI	Received signal strength indication
Space to CH	Distance between nodes
Max distance to CH	Maximum space between Cluster Head
Average distance to CH	Average of distance between cluster head
Current energy	Recent energy status of node
Energy consumption	Volume of energy to be used in last round
ADV CH send	Broadcast of messages by CH for sending
ADV CH receives	Received messages by CH
Join REQ send	Joining messages to CH by nodes
Join REQ receive	Received messages by CH for joining nodes
ADV SCH send	TDMA (Time Division Multiple Access) messages to nodes
ADV SCH receives	TDMA messages received from CH
Rank	Node order in TDMA
Data sent	Nodes data packages sent to CH
Send Code	Cluster code
Rank	Node order in TDMA
Data received	Data packages received from CH
Attack Type	Classification of attacks

Table 2: Description of mathematical model for LEACH

Term	Description
N	Number of nodes in network
Si	Node i
NC	CH numbers
CM	Cluster members
ADV CH Sent	Advertisement messages by CH
ADV-CH-RCVD	Advertisement received by nodes
Join Req Sent	Join requests by nodes
Join Req Rcvd	Join requests received by CH
TDMA Sent	TDMA routine send by CH
TDMA RCVD	TDMA routine received by nodes
NO PKT	Number of data packets received by CH

Algorithm 3 Scheduling attack algorithm (Algorithm 3):

```
Begin
if rSNi ¡ TSNi then
    SNi = CH
else
    SNi = CM
    if CHj, J NC then
        X CMs join CHj
        CHj create TDNA schedule
    end if
    if CHj process attack by structuring TDMA routine, give same time slice to all
    node for send the data then
        CHj = Adv_CH (with high broadcasting)
    else
        CHj structure normal TDMA routine

    end if
end if
if xCMs join CHj then
    CHj = Adv_CH (with high broadcasting)
else
    CHj xCMs send data to CHj in related TDMA time slice
    CHj send data to BS
end if
End
```

cision. Formulas 01 to 03 taken from [5]. Few MLP ANN structure developed in this paper with one, two and three hidden layers. Dataset was separated into training 60 % and testing set 40% as showing in Table 4.

We use two kind of cross validation method to train the dataset. Cross validation is used to evaluate the model. The problem with residual evaluations is that they do not give an indication of how well the learner will do when it is asked to make new predictions for data it has not already seen. One way to overcome this problem is to not use the entire data set when training a learner. Some of the data is removed before training begins. Then when training is done, the data that was removed can be used to test the performance of the learned model on "new" data. This is the basic idea for a whole class of model evaluation methods called cross validation. Dataset divided 70% training set and 30 testing sets for experiment purpose as mentioned below table.

Table 3: Dataset attributes

Parameters setting	Numbers values
Data packets size	500 bytes
Network length	100 x 100 meters
Packet headers size	25 bytes
Protocol	LEACH
Experimental time	1 hr
Cluster numbers	Five
Transmission limitation	200 meters

Holdout cross validation: This is simple type of cross validation by setting the data into train and test. By using approximator function only for train set. All data is randomly divided into same equal size data sets. The advantage of holdout method is mostly common to residual method and take short time for calculation.

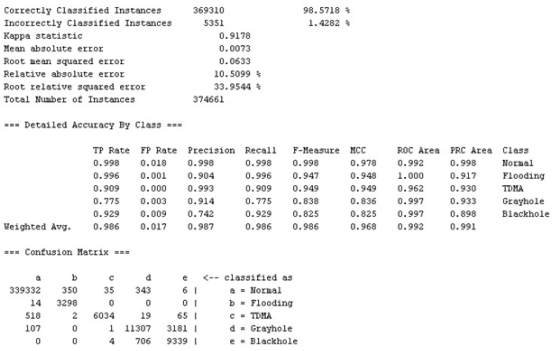
K-fold cross validation: This is an improved form for previous method where dataset divvied into given number for example 10 and holdout method recurrent k number of time. Advantage of this method is calculated variance is reduced as k number is increased.

Table 4: Dataset was separated into training 60 % and testing set 40%

Attack Types	Train Data	Test Data
Scheduling	3982	2656
Flooding	1988	1324
Grayhole	8758	5838
Blackhole	6029	4020
Normal	204039	136027

In this article dataset constructed using LEACH protocol in WSN. Dataset have 374668 records by representing four kinds of DoS attacks. Multilayer Artificial neural network machine learning (MLP) method is used to train dataset by having more than one hidden layer. Due to different performance confusion metrics we have used following performance matrix. These performance matrix are True Positive Rate (TPR), True Negative (TN), False Positive (FP), False Negative (FN) and Precision. Formulas 01 to 03 taken from [5]. Following Tables 4, 5, 6 and 7 describes the output summery of holdout cross validation method for TPR, FPR, FNR, TNR under hidden Layers I, II and III.

Classification is best for all type of attacks but not satisfactory for Grayhole and Scheduling attacks. Figure 3 for FPR having lower ratio and the good performance under hidden layer 1. Figures 4 and 5, shows good performance except for FNR. Figures 6 and 7 shows the simulation process by using single hidden layer. Figures 8, 9 and 10 shows the output performance with up to three



```
Correctly Classified Instances      369310               98.5718 %
Incorrectly Classified Instances      5351                1.4282 %
Kappa statistic                          0.9178
Mean absolute error                      0.0073
Root mean squared error                  0.0633
Relative absolute error                 10.5099 %
Root relative squared error             33.9544 %
Total Number of Instances           374661

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   MCC      ROC Area  PRC Area  Class
                 0.998     0.018     0.998     0.998     0.998     0.978     0.992     0.998     Normal
                 0.996     0.001     0.904     0.996     0.947     0.948     1.000     0.917     Flooding
                 0.909     0.000     0.993     0.909     0.949     0.949     0.962     0.930     TDMA
                 0.775     0.003     0.914     0.775     0.838     0.836     0.997     0.933     Grayhole
                 0.929     0.009     0.742     0.929     0.825     0.825     0.997     0.898     Blackhole
Weighted Avg.    0.986     0.017     0.987     0.986     0.986     0.968     0.992     0.991

=== Confusion Matrix ===

      a      b      c      d      e   <-- classified as
 339332    350     35    343      6 |      a = Normal
     14   3298      0      0      0 |      b = Flooding
    518      2   6034     19     65 |      c = TDMA
    107      0      1  11307   3181 |      d = Grayhole
      0      0      4    706   9339 |      e = Blackhole
```

Figure 3: Simulation output with single layer

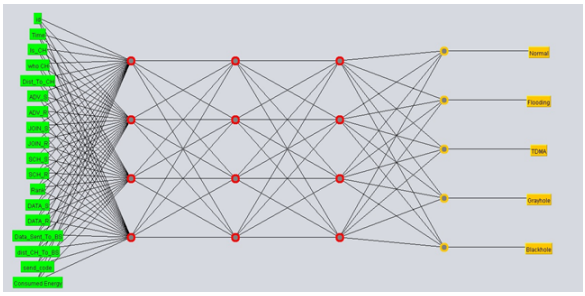


Figure 4: GUI simulation process with three hidden layers

hidden layer by using 10 cross validation method. Its observes that by using cross validation method CV method batter with single hidden layer for classification of normal and blackhole attacks.

So it is noted that cross validation architecture highly batter in the classification for denial of service attack in WSN. By obtaining the result with collected WSN dataset and applying MLP ANN is the possibility of highly accurate output for classification of denial of service attacks.

7 Conclusions

The idea of this paper to create the structure of Intrusion Detection System which have ability to work against DoS

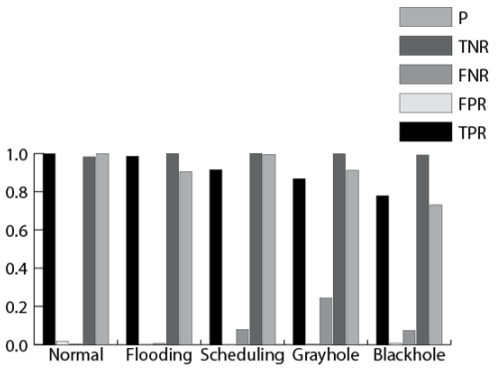


Figure 5: 10 cross validation method with one hidden layer

Table 5: True positive ratio summery with holdout cross validation

Hidden Layer	Normal	Flooding	Scheduling	Grayhole	Blackhole
I	0.99	1	0.98	0.92	0.34
II	0.99	1	0.96	0.71	0.81
III	0.98	0.95	0.97	0.57	0.98

Table 6: False positive ratio summery with holdout cross validation

Hidden Layer	Normal	Flooding	Scheduling	Grayhole	Blackhole
I	0.008	0.003	0	0.006	0.611
II	0.004	0.002	0	0.003	0.004
III	0.316	0.003	0.001	0.001	0.616

Table 7: False negative ratio summery with holdout cross validation

Hidden Layer	Normal	Flooding	Scheduling	Grayhole	Blackhole
I	0.004	0	0.014	0.679	0.657
II	0.004	0	0.16	0.286	0.182
III	0.005	0.011	0.027	0.424	0.011

Table 8: True negative ratio summery with holdout cross validation

Hidden Layer	Normal	Flooding	Scheduling	Grayhole	Blackhole
I	0.997	0.997	1	0.997	0.997
II	0.992	0.996	1	0.995	0.987
III	0.988	0.996	0.990	0.998	0.983

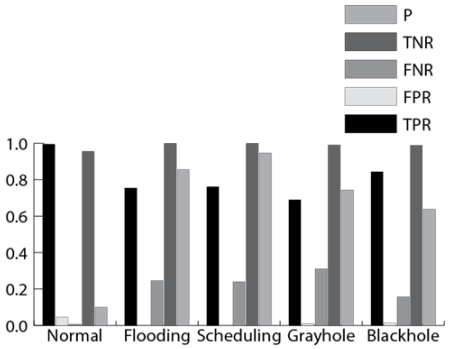


Figure 6: 10 cross validation method with three hidden layer

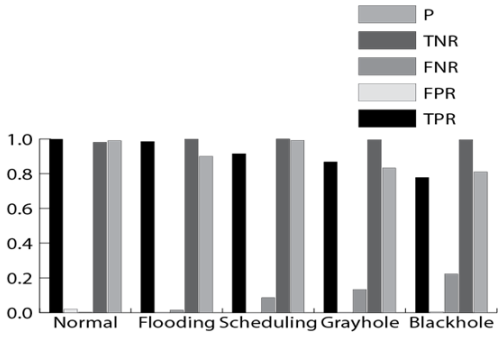


Figure 7: 10 cross validation method with two hidden layer

attacks with in affordable cost. The conclusion can be summarized with considering the results which are successfully classified as the DoS attacks with high detection rate. The future work will be extended with results having more hidden layers as well as able to classify more network attacks like wormhole attack or Sybil.

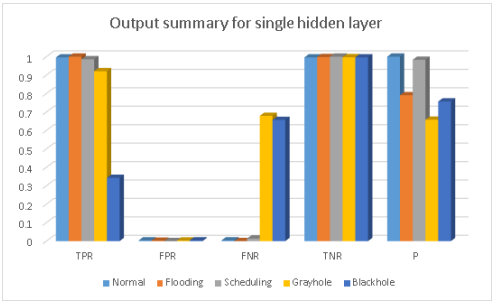


Figure 8: Holdout method Results with single hidden layer

References

[1] P. Aggarwal, S. K. Sharma, "Analysis of KDD dataset attributes - Class wise for intrusion detection," *Pro-*

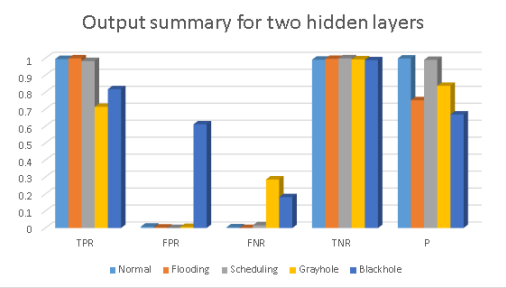


Figure 9: Holdout method Results with two hidden layer

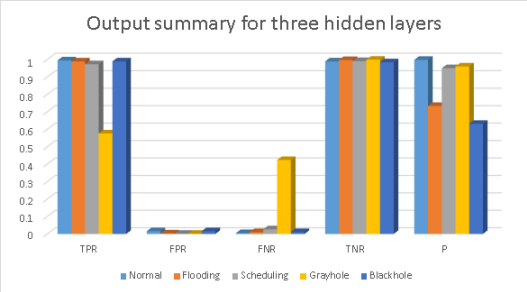


Figure 10: Holdout method Results with three hidden layer

cedia Computer Science, vol. 57, pp. 842-851, 2015.

[2] I. Butun, S. D. Morgera, and R. Sankar, "A survey of intrusion detection systems in wireless sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 266–282, 2014.

[3] R. N. Enam, N. Ismat, F. Farooq, "Connectivity and coverage based grid-cluster size calculation in wireless sensor networks," *Wireless Personal Communications*, vol. 95, pp. 429-443, July 2017.

[4] R. N. Enam, M. Tahir, R. Qureshi, "A survey of energy conservation mechanisms for dynamic cluster based wireless sensor networks," *Mehran University Research Journal of Engineering & Technology*, vol. 37, no. 2, pp. 279, 2018.

[5] M. Everingham, S. M. A. Eslami, L. V. Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge: A retrospective," *International Journal of Computer Vision 111*, no. 1, pp. 98-136, 2015.

[6] N. Farooq, I. Zahoor, S. Mandal, and T. Gulzar, "Systematic analysis of DoS attacks in wireless sensor networks with wormhole injection," *International Journal of Information and Computation Technology*, vol. 4, no. 2, pp. 173–182, 2014.

[7] A. Garofalo, C. D. Sarno, and V. Formicola, "Enhancing intrusion detection in wireless sensor networks through decision trees," in *Dependable Computing*, pp. 1–15, 2013.

[8] V. C. Gungor, B. Lu, and G. P. Hancke, "Opportunities and challenges of wireless sensor networks in smart grid," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 10, pp. 3557– 3564, 2010.

[9] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd IEEE Annual Hawaii International Conference on System Sciences*, pp. 1–10, 2000.

[10] S. K. Jangir, N. Hemrajani, "A comprehensive review on detection of wormhole attack in MANET," in *International Conference on ICT in Business Industry & Government (ICTBIG'16)*, pp. 1-8, 2016.

[11] R. Kaur, P. Singh, "Review of black hole and grey hole attack," *The International Journal of Multimedia & Its Applications*, vol. 6, pp. 35-45, 2014.

[12] D. Kumar, "Performance analysis of energy efficient clustering protocols for maximizing lifetime of wireless sensor networks," *IET Wireless Sensor Systems*, vol. 4, no. 1, pp. 9–16, 2014.

[13] M. Kumar, K. Dutta, I. Chopra, "Impact of wormhole attack on data aggregation in hieracrchical WSN," *International Journal of Electronics and Information Engineering*, vol. 1, no. 2, pp. 70–77, 2014.

[14] N. Marriwala and P. Rathee, "An approach to increase the wireless sensor network lifetime," in *Proceedings of the World Congress on Information and Communication Technologies (WICT'12)*, pp. 495–499, 2012.

[15] Y. M. Miao, "Cluster-head election algorithm for wireless sensor networks based on LEACH protocol," *Applied Mechanics and Materials*, vol. 738-739, pp. 19–22, 2015.

[16] S. Patil, S. Chaudhari, "DoS attack prevention technique in wireless sensor networks," *Procedia Computer Science*, vol. 79, pp. 715-721, 2016.

[17] M. A. Rassam, M. A. Maarof, and A. Zainal, "A survey of intrusion detection schemes in wireless sensor networks," *American Journal of Applied Sciences*, vol. 9, no. 10, pp. 1636–1652, 2012.

[18] R. Singh and M. S. Manu, "An energy efficient grid based static node deployment strategy for wireless sensor networks," *International Journal of Electronics and Information Engineering*, vol. 7, no. 1, pp. 32-40, 2017.

[19] S. Taneja, "An energy efficient approach using load distribution through LEACH-TLCH protocol," *Journal of Network Communications and Emerging Technologies (JNCET'15)*, vol. 5, no. 3, pp. 20–23, 2015.

[20] S. Tyagi and N. Kumar, "A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks," *Journal of Network and Computer Applications*, vol. 36, no. 2, pp. 623–645, 2013.

[21] S. S. Wang, K. Q. Yan, S. C. Wang, and C. W. Liu, "An integrated intrusion detection system for cluster-based wireless sensor networks," *Expert Systems with Applications*, vol. 38, no. 12, pp. 15234–15243, 2011

[22] J. Xu, J. Wang, S. Xie, W. Chen, and J. U. Kim, "Study on intrusion detection policy for wireless sensor networks," *International Journal of Security and Its Applications*, vol. 7, no. 1, pp. 1–6, 2013.

Biography

Munawar Hussain biography. Munawar Hussain is pursuing his PhD Degree from Yanshan university, China. He completed his master degree from University of Education Lahore, Pakistan. His interested area is complex network, Network security, Data mining and big data.

Jiadong Ren biography. Jiadong Ren is a professor at Yanshan University, Qinhuangdao, China. He is also a supervisor of doctoral and master students. His research area is data mining, network security and software security.

Awais Akram biography. Awais Akram currently pursuing his PhD Degree from Yanshan University China, He completed his Master degree from Islamia University Bahawalpur. His research interested area is Complex network security and data mining.