

Hybrid Anomaly Detection using K-Means Clustering in Wireless Sensor Networks

Mohammad Wazid

Center for Security, Theory and Algorithmic Research,
International Institute of Information Technology, Hyderabad 500032, India
wazidkec2005@gmail.com, mohammad.wazid@research.iit.ac.in

Abstract. Security is the biggest concern in Wireless Sensor Networks (WSNs) especially for the ones which are deployed for military applications and monitoring. They are prone to various attacks which degrades the network performance very rapidly. Sometimes multiple attacks are launched in the network using hybrid anomaly. In this situation it is very difficult to find out which kind of anomaly is activated. In this paper, we have proposed a hybrid anomaly detection technique with the application of k-means clustering. The analysis of the network data set consists of traffic data and end to end delay data is performed. The data set is clustered using weka 3.6.10. After clustering, we get the threshold values of various network performance parameters (traffic and delay). These threshold values are used by the hybrid anomaly detection technique to detect the anomaly. During the experimentation, it has been observed that two types of anomalies are activated in the network causing misdirection and blackhole attacks.

Keywords: Hybrid Anomaly, Misdirection, Blackhole, K-Means Clustering, Hybrid Anomaly Detection Algorithm.

1 Introduction

A Wireless Sensor Network (WSN) is a collection of densely deployed sensor nodes, sensing physical phenomena such as temperature and pressure. The important applications of WSN include environmental monitoring, personal healthcare, enemy monitoring and so on. Sometimes the sensitive data is communicated through an insecure medium. WSN can be easily attacked by enemies who cause information loss, along with large energy expenditure. Therefore, securing the link is important in designing a WSN. Sometimes hybrid anomaly (multiple anomalies) is introduced in the network to degrade the network performance and to trouble the attack specific detection mechanism. For this work, we have considered two attacks (misdirection and blackhole) and the analysis of network performance is done, and a data set is created. The clustering of data set of existing anomalies is also done, and threshold values are computed. The proposed technique takes these threshold values as inputs. During the experimentation, it has been observed that two kinds of anomalies are activated in the network, one is causing misdirection attack and other is causing blackhole attack. The proposed technique can also be applied for other anomalies with some modifications in traffic and delay model.

The rest of the paper is organized as: Section 2 contains the literature survey. The problem definition is discussed in section 3 along with the related terminology. The related terminologies are defined in section 4. The methodology and the experiment design are explained in section 4 and 5 respectively. The work is concluded in section 6.

2 Literature Survey

In [1], authors have done a topological analysis of WSN in the presence of misdirection attack and an algorithm for the prediction of delay and throughput is proposed. In [2], an efficient technique that uses multiple deployed base stations to counter the impact of black holes on data transmission is proposed. In [3], a specification based intrusion detection system to detect blackhole attack in WSN is proposed. The proposed scheme tries to optimize the local information into global information, in order to compensate the communication pattern in network. In [4], the effect of blackhole attack on the performance of WSN is measured, and then a cluster based technique for the detection and prevention of blackhole attack is proposed. In [5], few key design principles related to the development of anomaly detection techniques in WSNs are discussed. The analysis and comparisons of the approaches that belong to a similar technique category are also represented.

In [6], a non parametric approach for traffic classification is proposed. The performance of classification can be improved by correlating the information during the classification. During the experimentation, it has been observed that the performance of traffic classification can be improved effectively by the proposed scheme (also under few training samples). A novel classification scheme of network traffic is presented in [7]. It can improve the performance of classification having few training samples, which is also proved by the help of experimentations. In [8], KNN-based anomaly detection (AD) scheme is proposed, in which hyper grid intuition based approach is applied. The computational complexity is reduced by redefining anomalies from hypersphere detection region to hypercube detection region. During the experimentations, it has been observed that the proposed method is effective and robust and can be applied without any human intervention in WSN applications. In [9], a hybrid detection framework that depends on data mining classification and clustering techniques is proposed. Random forests classification algorithm is used, in order to detect misuse by building intrusion patterns from training dataset. These patterns are then matched with network connections to detect network intrusions. K-means clustering algorithm is used to detect novel intrusions by clustering the network connections for anomaly detection. In [10], anomaly traffic detection system based on the entropy of network features and support vector machine (SVM) are compared. A hybrid technique, a combination of both entropy of network features and SVM is compared with the individual methods.

In [11], hybrid anomaly-based intrusion detection method is proposed that uses two methods. These methods are trained in supervised way. The authors have used following additional techniques to improve the performance of proposed approach: First, a feature selection technique using the entropy of features is used for extracting

optimized information from KDD data set and second, a novel method is proposed to combine the results of these two learning based methods. In [12], the attempt has been made to apply hybrid learning approach by combining k-Medoids based clustering technique followed by Naive Bayes classification technique. Because of the fact that k-Medoids clustering techniques represent the real world scenario of data distribution, the proposed enhanced approach will group the whole data into corresponding clusters more accurately than k-Means such that it results in a better classification. An experiment is carried out in order to evaluate performance, accuracy, detection rate and false positive rate of the classification scheme. In [13], a new framework based on a hybrid intrusion detection system for known and unknown attacks in an efficient way has been proposed. This frame work has the ability to detect intrusion in real time environment from the link layer. In [14], a hybrid IDS is proposed which uses the signature and anomaly information together. The proposed algorithm first explore those traffic features, which are changing during an intrusion activity and then based on a predefined threshold value the most prominent features related to attack are identified. These features are included in snort rule set to detect the anomalous traffic. This anomaly detection process is combined with existing signature of snort to produce the better detection.

The extended partitioning based k-means clustering technique is presented in [15]. It performs clustering when number of clusters and number of objects are increased. So it can be applied to dynamic database where the data changes frequently. This method can also be applied for clustering of large multidimensional dataset. In [16], intrusion detection system associated with high false alarm with moderate accuracy and detection rates is proposed. To overcome this problem the authors have proposed, a hybrid learning approach through the combination of k-means clustering and naive bayes classification. The proposed approach is used to cluster all data into the corresponding group before applying a classifier for classification purpose. In [17], the study of intrusion detection for wireless industrial sensor networks is done and then classification is performed. The selection of better methodologies against various intrusions is also done. In [18], a new hybrid intrusion detection system (HIDS) design principles and evaluation results are reported. This hybrid system has the advantages of low false-positive rate of signature-based intrusion detection system (IDS), and the ability of anomaly detection system (ADS) to detect novel unknown attacks.

Hybrid anomaly (presence of more than one attacker) can be easily introduced in the network, which can degrade the performance of WSN very rapidly. So the detection of hybrid anomalies has become important. During literature survey we have not found any clustering based solution for the detection of hybrid anomaly of WSN. If we apply, k-means clustering our task becomes easy because it can give us the threshold values of various network parameters, which can be further used by the hybrid anomaly detection technique. So this work is missing in previous work done by the various authors.

3 Problem Definition and Related Terminology

The wireless sensor network is prone to various attacks so the confidential information can be leaked or altered. When attack happens, in the network the performance

degrades i. e. increased end-to-end delay and decreased network throughput. So the information cannot reach to the destination within time. Sometimes multiple attacks are launched in the network to degrade the performance and to trouble the attack specific detection mechanism. In this situation, it becomes very crucial to find out which kind of anomaly is activated in the network. In this paper, we have proposed a hybrid anomaly detection technique using k-means clustering. The analysis of the network data set consists of traffic data and end to end delay data is performed. The data set is clustered using weka 3.6.10 and threshold values of parameters (end-to-end delay and traffic received) are computed. The computed threshold values are input for proposed hybrid anomaly detection technique.

The following are the related terms used in this paper:

3.1 Misdirection Attack

In misdirection attack the attacker routes the packet from its children to other distant nodes but not necessarily to its legitimate parent. This produces long delay in packet delivery and decreases the throughput of the network. The packets reach to the destination but from a different route which further produces long delay, thus decreasing the throughput of network.

3.2 Blackhole Attack

Blackhole attack occurs when an intruder captures and re-programs a set of nodes in the network to block the packets they receive, instead of forwarding them towards the base station. As a result any information that enters in the black hole region is captured. Black hole attacks are easy to constitute and they are capable of undermining network effectiveness by partitioning the network such that important event information do not reach the base stations. The network performance parameters i.e. throughput and end to end delay are affected in the presence of blackhole nodes.

3.3 Clustering

Clustering basically is the task in which the data points are divided into homogenous classes or clusters. Homogeneous means that they are similar. Items present within the same class are as much as similar. Thus this process can also be referred as Grouping.

K- Means Clustering

K-means was first used by MacQueen in 1967 is one of the simplest clustering method comes under unsupervised learning algorithms used to solve the well known clustering problem. It follows a simple and easy way to classify a given data set through a certain number of clusters (i.e. k clusters) fixed a priori.

K- Means clustering comes under the category of partitioning method in which a partition of a database D of n objects is done into a set of k clusters. Given a k , the

main task is of finding a partition of k clusters that optimizes the chosen partitioning criterion. That's why we have preferred k-means clustering for this work.

The input to this algorithm is k and task is to partition a set of n objects into k clusters so that the resulting intra cluster similarity is high but the inter cluster similarity is low. Cluster similarity is measured in regards to the mean value of the object in a cluster, which can be viewed as the cluster's centroid or center of gravity.

The algorithm has following steps:

- We place k points into the space represented by the objects that are being clustered. Initial group centroids are represented by these points.
- Assign each object to the group that has the closest centroid.
- After the assignment of all objects, recalculate the positions of the k centroids.
- Repeat Steps 2 and 3 until the centroids no longer move.

4 Methodology

A wireless sensor network, under normal flow and under attack is simulated. The results of the simulation are traffic data and end to end delay data, are clustered using weka. The k-means clustering technique is used, in computing the clusters of traffic and delay data, and threshold values for various performance parameters are calculated. Then these threshold values are used by hybrid anomaly detection technique, in order to find out the presented anomalies. Two type of analysis are done, one is using the traffic data, where we find the threshold values for blackhole nodes and second one is done using delay data, where we get the threshold values for misdirection nodes.

The two kinds of data traffic is used i.e. traffic sent and traffic received. In traffic sent analysis, we have detected Blackhole nodes. The traffic received analysis, helps to find the nodes working in collaboration with misdirection attackers, as they would have high value of traffic received under attack. The nodes with high delay are misdirection attacker nodes. To check which node is working in collaboration with which misdirection attacker node, we have to check the communication range of nodes. If a node is in the communication range, then it can work in collaboration with that misdirection attacker node, otherwise the collaboration is not possible.

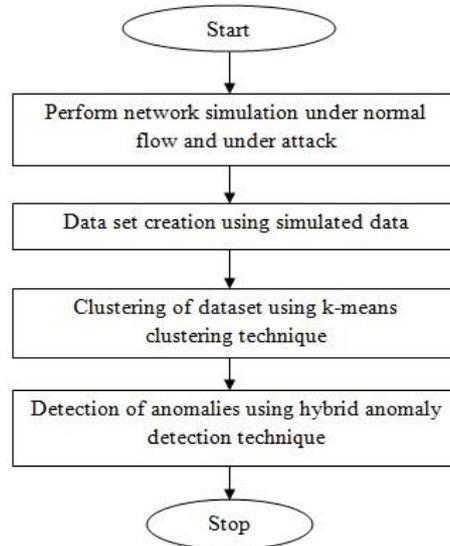


Fig. 1. Proposed Work

Figure 1 depicts the process of proposed work.

4.1 Mathematical Model

For the proposed hybrid anomaly detection scheme, we have developed following mathematical model.

Traffic Model

Traffic Sent

If a node is a blackhole node, then traffic sent under attack must be zero. Otherwise under normal flow it has some finite value:

$$\begin{aligned} Tr_{sent} &= 0, \text{ under blackhole attack} \\ Tr_{sent} &= \text{some finite value, under normal flow} \end{aligned}$$

Traffic Received

If a node is working correctly, then it receives traffic less than the calculated threshold value.

Suppose for a node the traffic received threshold value is x , then under normal flow it is equal to or less than this value.

$$\begin{aligned} Tr_{received} &\leq x, \text{ under normal flow} \\ Tr_{received} &> x, \text{ under misdirection attack, if a node is working in collaboration with} \\ &\quad \text{misdirection attacker node.} \end{aligned}$$

In this case traffic received for a node is the sum of its own traffic and traffic send by a misdirection attacker node. Suppose A is that node and B is misdirection attacker node, then traffic received will be:

$$\text{Tr}_{A \text{ received}} = \text{Tr}_{A \text{ received}} + \text{Tr}_{B \text{ received}} \text{ (Traffic received at node A under misdirection attack)}$$

Delay Model

In the presence of misdirection attack the delay is increased at some nodes.

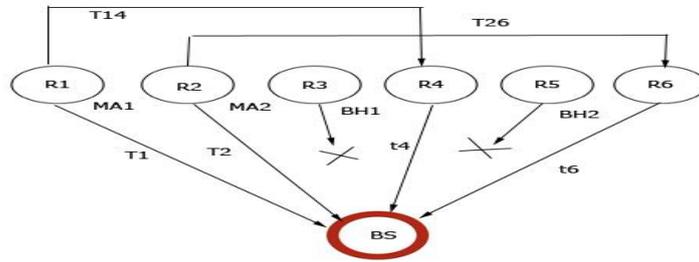


Fig. 2. Traffic flow under hybrid anomaly

Figure 2 shows traffic flow under hybrid anomaly. Suppose there are six nodes (R_1 , R_2 , R_3 , R_4 , R_5 and R_6) and a base station (BS) in the network, two of them are misdirection attacker nodes (MA_1 and MA_2) and two are blackhole attacker nodes (BH_1 and BH_2).

Under normal flow, let say T_1 and T_2 are delays for R_1 (MA_1) and R_2 (MA_2). If misdirection attack happens, then delay increases let say it becomes T_1' and T_2' for MA_1 and MA_2 respectively.

$$T_1' = T_{14} + t_4$$

Packets take T_{14} time from R_1 to R_4 and t_4 time from R_4 to base station (BS).

$$T_2' = T_{26} + t_6$$

Packets take T_{26} time from R_2 to R_6 and t_6 time from R_6 to base station (BS).

So under misdirection attack:

$$\begin{aligned} T_1' &> T_1 \\ T_2' &> T_2 \end{aligned}$$

4.2 Hybrid Anomaly Detection Algorithm

The proposed hybrid anomaly detection algorithm is given below:

```

//  $Tr_{sent}$  is traffic sent by a node
//  $btry\_rem\_node[i]$  is remaining battery backup of node
//  $T[i]$  is a time taken by a node in packet transmission to base
//  $T[i]_{thrsh}$  is the threshold value of time taken by a node in packet transmission
to the base also includes time window utilizes in case of network congestion
//  $Tr_{received}$  traffic received by a node
//  $X_j$  threshold value of traffic received by a node
//  $d_{ij}$  communication range distance between node i and node j
//  $2R$  limit of communication range of node i and node j
hybrid_anomaly_detection_algorithm ()
{
    At each intermediate node
    for i:=1 to n
    blackhole_node_detection ()
    {
        At Node i
        if ( $Tr_{sent} := 0$ ) then
            battery_remaining ( ) /* Used to check remaining battery backup at a
            node*/
            {
                if ( $btry\_rem\_node[i] := 0$ ) then
                    Node [i] failure
                otherwise
                    Node [i] is a blackhole attacker node
            }
        otherwise
            No intruder
    }
    misdirection_node_detection ()
    {
        At Node i
        if ( $T[i] > T[i]_{thrsh}$ ) then
            Node [i] is a misdirection attacker node
            comm_range ( ) /* Used to check the communication range of a
            node*/
            {
                if ( $d_{ij} < 2R$ ) then
                    if ( $Tr_{received} > X_j$ ) then
                        Node[ j] is working in collaboration with node [i]
                    otherwise
                        continue for other neighbor nodes
                otherwise
                    Node[j] and Node[i] are not in communication range
            }
        otherwise
            No intruder
    }
}

```

```
}  
}
```

5 Experiment Design and Results

5.1 Network Simulation Design

A wireless sensor network using opnet modeler is simulated and a dataset is created. Opnet tool has facility to export the data into some formats i.e. MS Excel. The data sheet of MS Excel is converted into .CSV format which is supported in weka. Then clustering is done using weka and threshold values of various parameters are computed.

The simulation scenario consists of 18 sensor nodes.



Fig. 3. Network Scenario under normal flow



Fig. 4. Network Scenario under attack

In figure 3, we have used 18 sensor nodes and build a scenario without any attacker showing a normal flow of traffic. In figure 4, we have used 18 sensor nodes and build a scenario with different attackers. R1, R2, R3 and R5 are attacker nodes. The various design parameters are listed in Table 1.

Experiment Design Parameters

Table 1. Common Parameters used in experimentation

Parameter		Value
Network simulation Tool		Opnet
Area		500x500 met
Network Size	Normal Flow	18 Sensor Nodes
		06 Routers with normal flow
		01 Coordinator
	Attacking Scenario	12 Sensor Nodes
		02 Routers with normal flow
		04 Router (attacker)
		01 Coordinator
	Topologies	
Simulation Time		60 Minutes
Packet Size (bits)		Constant (1024)
Clustering Tool		Weka 3.6.10
Clustering Technique		Simple K-Means
Clustering Data Set		Traffic sent
		Traffic received
		End to end delay
Maximum instances		600

5.2 Results

The clustering is done using Weka and the following results are obtained:

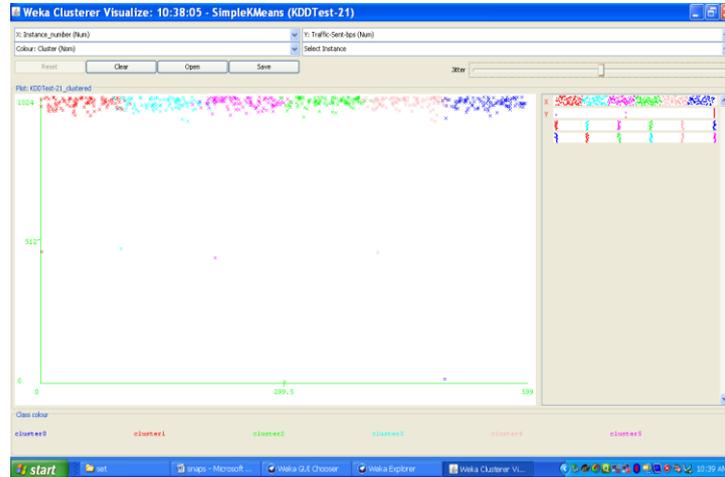


Fig. 5. Computed clusters of traffic sent under normal flow

Six clusters are obtained as shown in figure 4. All clusters are at the same level.

Table 5. Traffic sent under normal flow

Attribute	Cluster#					
	0	1	2	3	4	5
Traffic Sent (bps)	1018.31	1018.31	1018.31	1018.31	1018.31	1018.31
Class	R1	R2	R3	R4	R5	R6

Table 5 is drawn on the basis of clustering results, we can see that the value of traffic sent is same for all nodes (R1, R2, - - - R6). So the threshold value of traffic sent is 1018.31 bps for all nodes.

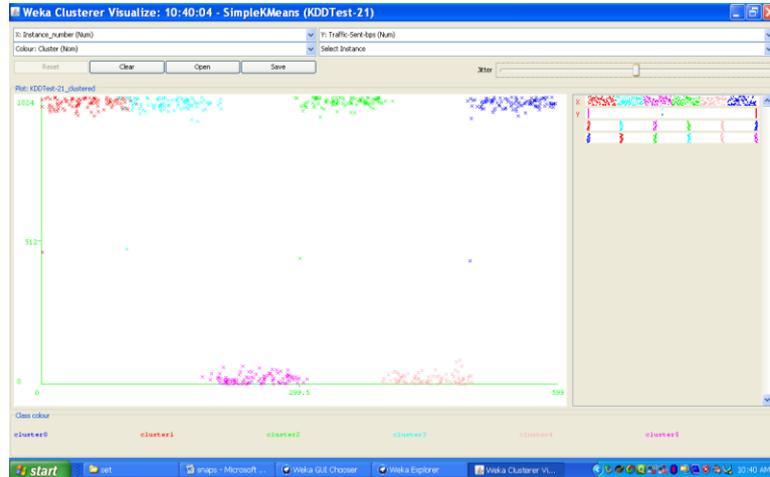


Fig. 6. Computed clusters of traffic sent under attack

Six clusters are obtained as shown in figure 6. Four clusters are at the same level and two clusters are at same level having almost zero value of traffic sent.

Table 3. Traffic sent under attack

Attribute	Cluster#					
	0	1	2	3	4	5
Traffic Sent (bps)	1018.31	1018.31	0.00	1018.31	0.00	1018.31
Class	R1	R2	R3	R4	R5	R6

Table 3 is drawn on the basis of clustering results, we can see that the value of traffic sent is 1018.31bps for R1, R2, R4 and R6 and 0.00 bps for R3, R5. But the threshold value of traffic sent for R3 and R5 is 1018.31 bps, so R3 and R5 nodes are detected as blackhole nodes.

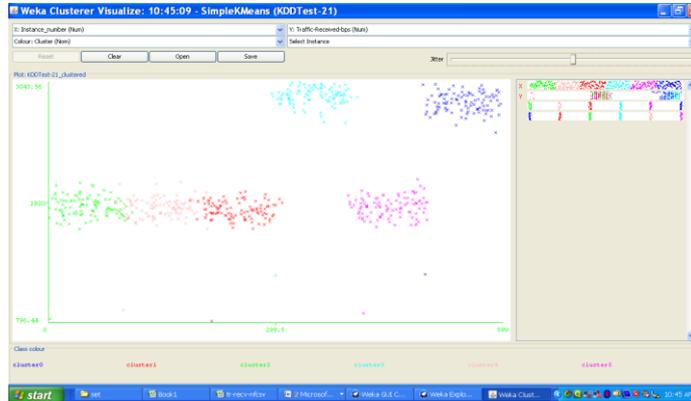


Fig. 7. Computed clusters for traffic received under normal flow

Six different clusters are formed as shown in figure 7.

Table 4. Traffic received under normal flow

Attribute	Cluster#					
	0	1	2	3	4	5
Traffic received (bps)	1868.80	1872.21	1839.22	2896.78	1867.66	2849.56
Class	R1	R2	R3	R4	R5	R6

Table 4 is drawn on the basis of clustering results again six different clusters are formed. The computed threshold values of traffic received for R1, R2, R3, R4, R5 and R6 are 1868.80, 1872.21, 1839.22, 2896.78, 1867.66 and 2849.56 bps respectively.

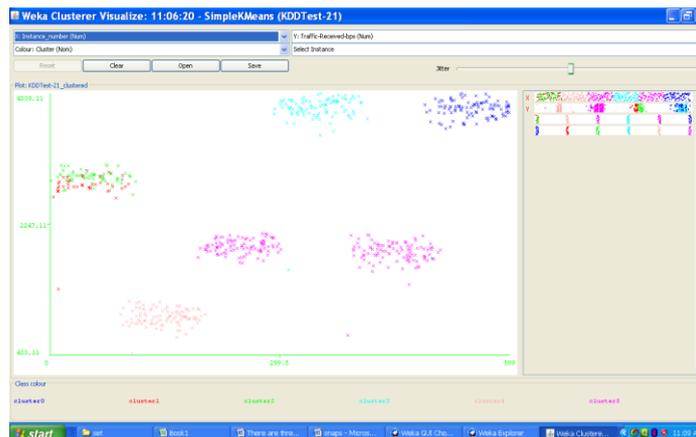


Fig. 8. Computed clusters for traffic received under attack

Six different clusters are formed as shown in figure 8.

Table 5. Traffic received under attack

Attribute	Cluster#					
	0	1	2	3	4	5
Traffic received (bps)	1868.80	1872.21	1839.22	3848.25	1867.66	3822.93
Class	R1	R2	R3	R4	R5	R6

Table 5 is drawn on the basis of clustering results again six different clusters are formed, two of them are having higher value of traffic received. The traffic received at R4 and R6 are 3848.25 and 3822.93 bps respectively which are very high as compared to the normal threshold values i.e. 2896.78 and 2849.56 bps respectively. Thus R4 and R6 are detected as the nodes working in collaboration with misdirection attacker nodes.

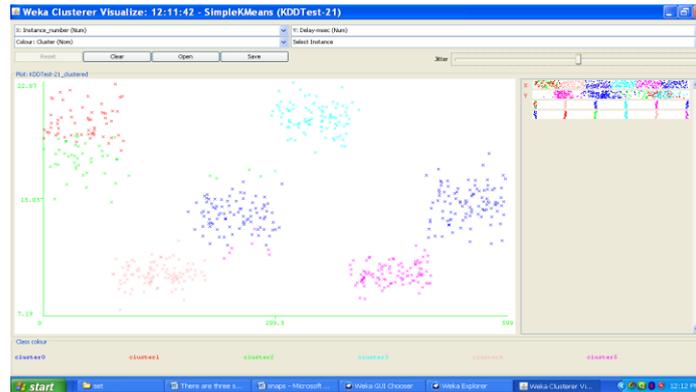


Fig. 9. Computed clusters for end-to-end delay under normal flow

Six different clusters are formed as shown in figure 9.

Table 6. End to end delay under normal flow

Attribute	Cluster#					
	0	1	2	3	4	5
End to end delay (ms)	19.09	24.56	13.92	20.37	10.01	14.73
Class	R1	R2	R3	R4	R5	R6

Table 6 is drawn on the basis of clustering results; the six different clusters are formed. The computed threshold values of delay for R1, R2, R3, R4, R5 and R6 are 19.09, 24.56, 13.92, 20.37, 10.01 and 14.73 ms respectively.

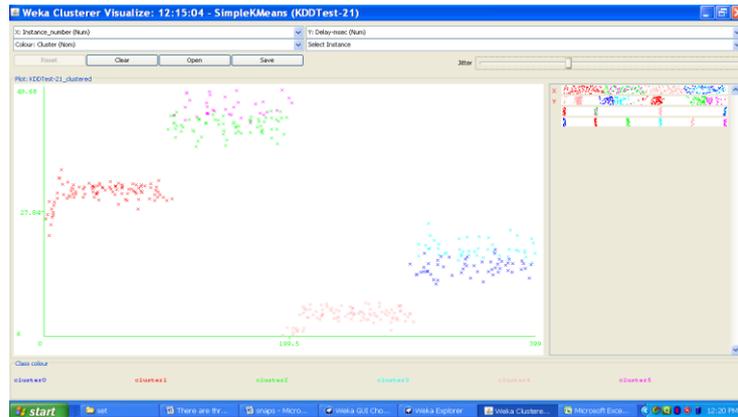


Fig. 10. Computed clusters for end to end delay under attack

Six different clusters are formed as shown in figure 10.

Table 7. End to end delay under attack

Attribute	Cluster#					
	0	1	2	3	4	5
End to end delay (ms)	31.01	29.59	13.92	20.37	10.01	14.73
Class	R1	R2	R3	R4	R5	R6

Table 7 is drawn on the basis of clustering results; six different clusters are formed. The normal threshold values of end to end delay for R1 and R2 are 19.09 ms and 24.56 ms respectively. These values are very less as compared to 31.01 and 29.59 ms obtained in this case. Thus R1 and R2 are detected as misdirection attacker nodes. Since R4 is nearer to R1 and R6 is nearer to R2, so R4 is the misdirection collaborating node of R1 (MA1) and R6 is the misdirection collaborating node of R2 (MA2)

5.3 Key Findings

During the experimentation following observations are made:

- R1 and R2 are detected as misdirection attacker nodes and nodes R4 and R6 are working in collaboration with these nodes. R1 misdirects traffic to R4 and R2 to R6 (Refer Table 5 and 7).

- Nodes R3 and R5 are blackhole attacker nodes (Refer Table 3).
- So two types of anomalies (hybrid anomaly) first misdirection attacker nodes and second blackhole attacker nodes are detected successfully.

6 Conclusion

The proposed technique is capable to detect hybrid anomaly exists in the wireless sensor network. The described method is capable of finding the blackhole nodes and misdirection nodes just by the analysis of two types of network performance parameters i.e. traffic data and end to end delay data. Blackhole nodes are the ones who don't forward the traffic and absorb all the packets reaching them. Thus their detection is done by comparing the traffic sent values which are generally zero for them. Misdirection nodes misdirect the traffic thus increasing the delay values sometime making it infinite also, these nodes are detected using end to end delay parameter values. During the experimentation, it has been observed that, the introduced hybrid anomaly is detected successfully by the proposed scheme.

This work can be extended further by adding more anomalies like sink hole, gray hole and other attacks in the network.

References

1. Roshan Singh Sachan, Mohammad Wazid, et.al, "Misdirection Attack in WSN: Topological Analysis and an Algorithm for Delay and Throughput Prediction", 7th International Conference on Intelligent Systems and Control (ISCO'13), 2013.
2. Satyajayant Misra, Kabi Bhattacharai, Guoliang Xue, "BAMBi: Blackhole Attacks Mitigation with Multiple Base Stations in Wireless Sensor Networks", IEEE International Conference on Communications (ICC), 2011.
3. Mukesh Tiwari, Karm Veer Arya, Rahul Choudhari, Kumar Sidharth Choudhary, "Designing Intrusion Detection to Detect Black hole and Selective Forwarding Attack in WSN based on local Information", 4th IEEE International Conference on Computer Sciences and Convergence Information Technology (ICCIT), 2009.
4. Mohammad Wazid, Roshan Sachan, et.al, "Detection and Prevention Mechanism for Blackhole Attack in Wireless Sensor Network", IEEE International Conference on Communication and Signal Processing (ICCSP), 2013.
5. Miao Xie, Song Han, Biming Tian, Sazia Parvin, "Anomaly Detection in Wireless Sensor Networks: A Survey", Elsevier Journal of Network and Computer Applications, vol. 34, Issue 4, July 2011, pp. 1302-1325.
6. Jun Zhang, Yang Xiang, Yu Wang, Wanlei Zhou, Yong Xiang, Yong Guan, "Network Traffic Classification Using Correlation Information", IEEE Transactions on Parallel and Distributed Systems, vol. 24, no. 1, pp. 104-117, 2013.
7. Jun Zhang, Chao Chen, Yang Xiang, Wanlei Zhou, and Yong Xiang, "Internet Traffic Classification by Aggregating Correlated Naive Bayes Predictions", IEEE Transactions on Information Forensics and Security, vol. 8, no. 1, pp. 5-15, 2013.
8. M. Xie, J. Hu, S. Han, and H.H. Chen, "Scalable Hyper grid k-NN-Based Online Anomaly Detection in Wireless Sensor Networks", IEEE Transactions on Parallel and Distributed Systems, vol. 24, no. 8, August 2013, pp. 1661-1670.
9. Reda M. Elbasiony, Elsayed A. Sallam, Tarek E. Eltobely, Mahmoud M. Fahmy, "A Hybrid Network Intrusion Detection Framework Based on Random Forests and Weighted k-Means", Elsevier Journal of Ain Shams Engineering, 2013.

10. Basant Agarwal, Namita Mittal, "Hybrid Approach for Detection of Anomaly Network Traffic using Data Mining Techniques", Elsevier Journal of Procedia Technology, vol. 6, 2012, pp. 996-1003.
11. K Qazanfari, M S Mirpouryan, H. Gharaee, "Novel Hybrid Anomaly Based Intrusion Detection Method", 6th IEEE International Symposium on Telecommunications (IST), 2012.
12. Roshan Chitrakar, Chuanhe Huang, "Anomaly Based Intrusion Detection Using Hybrid Learning Approach of Combining k-Medoids Clustering and Naïve Bayes Classification", 8th IEEE International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM), 2012.
13. A S Aneetha , T S Indhu, S Bose, "Hybrid Network Intrusion Detection System using Expert Rule Based Approach", 2nd ACM International Conference on Computational Science, Engineering and Information Technology (CCSEIT), 2012.
14. K V Arya, Hemant Kumar, "A Clustering Based Algorithm for Network Intrusion Detection", 5th ACM International Conference on Security of Information and Networks, 2012.
15. Sanjay Chakraborty, N K Nagwani, Analysis and Study of Incremental K-Means Clustering Algorithm", High Performance Architecture and Grid Computing, Communications in Computer and Information Science, vol. 169, 2011, pp. 338-341.
16. Z Muda, W Yassin, M N Sulaiman, N I Udzir, "Intrusion Detection Based on k-Means Clustering and Naive Bayes Classification", 7th IEEE International Conference on Information Technology in Asia, 2011.
17. Sooyeon Shin , Taekyoung Kwon ,Gil-Yong Jo , Youngman Park, "An Experimental Study of Hierarchical Intrusion Detection for Wireless Industrial Sensor Networks", IEEE Transactions on Industrial Informatics, vol. 6 , Issue 4, Nov. 2010.
18. Kai Hwang, Min Cai, Ying Chen, Min Qin, "Hybrid Intrusion Detection with Weighted Signature Generation over Anomalous Internet Episodes", IEEE Transactions on Dependable and Secure Computing, vol. 4, Issue 1, 2007.