Fault-Tolerant Data Aggregation in Wireless Sensor Networks

K.Radhakrishnan
PG Student,
Applied Electronics,
Velammal Engineering College,
Chennai, India.

V .Latha
Professor,
Dept. of Electronics and communication,
Velammal Engineering College,
Chennai, India.

ABSTRACT

Accurate information is most important in Wireless Sensor Networks (WSNs). Incorrect (faulty data) information makes wrong decision; it decreases the reliability in communication. Fault tolerance is one of the issues in WSNs .Proposed fault -tolerant data aggregation scheme for identifying the faulty data sent by the sensors to the cluster head, where the aggregation is performed. Identified faulty data eliminated before aggregation. It improves the data aggregation accuracy. Consequently, the transmission overhead is reduced. The proposed scheme provides better performance in discovering faulty data and nodes. The incorrect data detected and eliminated in wireless sensor networks.

General terms

Wireless sensor networks, Rough set theory.

Keywords

Cluster Head (CH), Error detection, Sink, Fault-Tolerant, Sensor.

1. INTRODUCTION

Wireless Sensor Networks used in a variety of environments. Sensor networks are used to monitor the office buildings, wildlife tracking, traffic surveillance, geographical regions, health care, industrial plants and so on. Sensor network consist of a large number of the low-cost and low-powered sensor devices, which is called sensor nodes. Sensors are deployed inside the buildings. Aggregation is the process of gathering the information for the analysis. In wireless sensor networks aggregation is performed by the cluster head (CH). CH collects the data from various sensors and summarizing the data. Information collected by the sensor, continuously transmitted to the mobile or base station. Sink retrieves the information.

WSNs are employed in energy consuming operation and reduce the number of bits transmitted to each device [1]. The data aggregation scheme improves the bandwidth utilization and energy efficiency. Data aggregation protocol ESPDA prevents redundant data transmission [2]. In WSN domain, secure connections between the sensors are established by the key mechanism and SRDA scheme consumes less energy [3]. The operation of basic aggregators such as adding min, max, average and count used for the data processing operation. Latency is measured between the packet received at the sink and the data generated at the source [4] and [5]. In the harsh sensor environment, fault detection and correction are made through the embedded neural networks. Faulty data sent to the CH, CH send the result to base station. Fault-tolerant aggregation protocol reduces the wrong data transmission which increases the reliability of the network as in [6].

General purpose protocol like routing and clustering are employed in wireless sensor networks. Cluster-head co-ordinates transmission uses the time division multiple access (TDMA) technique [7]. Tree sampling and set sampling algorithms are sampling the aggregated data, hence it overcome the requirement of a key. Moreover, security level obviously increased as in [8]. Identify the faulty node in the sensor network and remove the faulty data before transmission. Hence, aggregation accuracy improved as in [9], [10] and [11]. The simulation shows that Correctness of gathering data [12] in a Trust based framework in wireless sensor networks. Digital signature algorithm investigates on a cryptographic protocol as in [14]. Our contribution in this paper is aggregation accuracy and transmission overhead. Data collection and aggregation is a wireless sensor network discussed in section 2. The idea of using the fault-tolerant aggregation protocol and Fault identification is discussed

2. NETWORK MODEL

Nodes can communicate by using the Radio Transceiver. Sensors are deployed in harsh environments. The possible way of Sensors being any one of the two states: faulty and fault-free. When the sensor is in unsafe and safe is called as the most faulty and faultless respectively. It is unnecessary to transmit the faulty data to the cluster head from the sensors. Hence, it consumes more energy. If the network fails to find the faulty sensor, then it’s very difficult to control the amount of faulty data transmission. Moreover, it increases the transmission overhead. In the second state, If the characteristics of the sensors are good (i.e. The sensor then safe) directly aggregates the collected information. Update the fault diagnosis protocol to improve the performance of accuracy. Once the information is gathered that to be on the table for the segregation.

The pre-processing stage involves rule generation and fault -diagnosis as shown in the Figure (1). However, fault-tolerant data aggregation protocol is used for detecting the faults and eliminating the faults. Henceforth, faulty data is discarded for the further proceedings.

Cluster head (CH) does the information aggregation. Aggregated data has been sent to the sink. Processing stage includes a Locality Sensitivity Hashing scheme which is the compact representation of data for encoding. Along with the sensor information locality sensitivity hashing codes sent to the aggregation unit. This compares the similarity between the LSH codes. So that improves the utilization of bandwidth and energy efficiency. Preliminaries in the aggregation network as in [8].

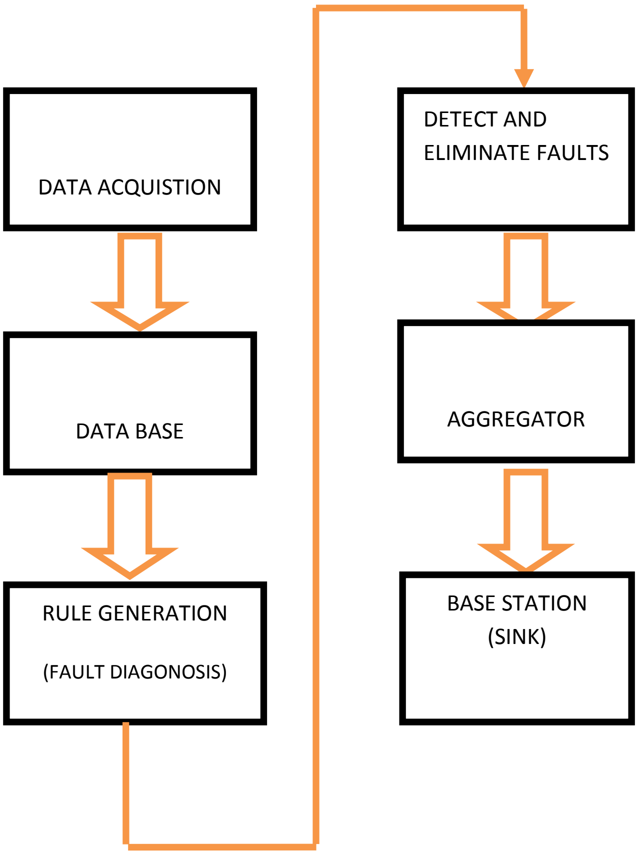

Figure 1: Flow Diagram for Error Detection in WSNs.

Similarity, distance metrics and local sensitivity hashing. Calculate the similarity between the data sets. Euclidean distance is the measure of distance between the pair of nodes. The cosine similarity of fault – tolerant aggregation is,

$$cos\big(\theta(Vi,Vj)\big) = \frac{Vi,Vj}{(||Vi||,||Vj||)} \qquad (1)$$

Hash function (Ꜧ) for any two point’s p, q ϵ Ꞧd as in [9]. It must satisfy the hashing function.

$$Pr[hr(u)] = hr(v) = 1 - \frac{\theta(u,v)}{\pi} \qquad (2)$$

The hamming distance calculated as,

$$b \times (1 - Pr) = Dn(LSHu, LSHv) \qquad (3)$$

The fault-tolerant scheme uses the LSH coding techniques and it increases the network lifetime of nthe network The hash function hour is defined as,

$$hr(u) = 1, if\ r \times u > 0 \ \text{And} hr(v) = 0, if\ r \times u < 0 \qquad (4)$$

Rough set concept has been explained in many applications of pattern recognition. An important aspect of the rough set concept is that decision making rules based on the measured sensor values. Rough set is a formal approximation of crisp set. Lower approximation of rough set classified as a member x.

$$\underline{B(x)} = \cup \{Ei\!: \in U^B\!: Ei \ \in Cx\ \} \qquad (5)$$

Upper approximation rough set classified as a member `x`

$$\overline{B(x)} = \cup \{E\!: \in U^B\!: Einx\ \{\}\} \qquad (6)$$

Fault class is defined as the difference between the upper approximation and lower approximation.

The basic aggregation scheme computes the variance and standard deviation. Every aggregator performs the summing operation sum(S),

$$S = \sum_{i=1}^{k} Xi \qquad (7)$$

Sum of the squares as in [1],

$$V = \sum_{i=1}^{k} Xi^2 \qquad (8)$$

3. FAULT-TOLERANT AGGREGATON SCHEME FOR ERROR CORRECTION

In data collection session, each time sensor senses the environment k times and store the values in terms of the temperature, pressure, voltage, current, power and so on. Let assume the sensed values are n bits and now sensor node data vector size (k*n) bits. Nodes transmit the (k*n) bit data to the aggregator. Hence the energy of battery level depleted. The sensor node generates sensitivity hashing code for each sensor data for their vectors to reduce the amount of bits transmitted.

The main advantage of the sensitivity hashing code is to represent the data in only less number of bits. Sensitivity hashing code applied to each sensor and obtain b bit sensitivity hashing code, but the value of b is less than (k*n). `B` bit code sends to the cluster- head for aggregation.

3.1 Fault Identification

In aggregation, the aggregator requests a sensor to transmit the sensitivity hashing code for processing. The sensor sends the sensitivity hashing code, but after this, aggregator does the comparison of pair of sensor nodes. There are two major strategies to follow before aggregation. In the first case, the similarity between the pair of nodes discovered from the sensitivity hashing code and similarity threshold. Sometimes the sensitivity hashing code will not match each other because of the event affected by the adversary and local outlier. Data aggregator shares the adversary list among the neighboring cluster to prevent the increment of count value of truthful data. Each cluster compares the hashing code with neighboring cluster and updates the count value. Moreover, Cluster exchange count list of local outlets.

In the second case, if the sensitivity hashing codes are perfectly matched during the comparison then increase the counter value by one. The cluster - head has the ability to find the sensor which has the same codes. After counting the faultless data the data aggregator performs data aggregation operation and aggregated data send to the base station. The important case of aggregation discards the redundant data transmission from the nodes. Using Rough set theory, when more than one sensor having the same code, the data aggregator takes any one data among the sensor nodes and sends the data. Furthermore it reduces the amount of transmission.

Algorithm for Faulty Data Elimination:

```
//Information gathering and sensitivity hashing code generation//
For all cluster-head session
    Gather data from set of sensor E.
    Generate hashing code for data set .end for
// Faulty-data Detection and Correction //
For all cluster-head session do
    Acknowledge the code from the sensor.
    Measure the similarities using codes.
    Find out the sensor data, the code which is perfectly
    Matching.
    Detect the faulty data (Redundant data).end for
// Data Aggregation //
For all cluster-head session do
    Remove the faulty data.
    Find the sensor node having the same value.
    Rough Set theory sends data from same value.
    Aggregate the received data.
    Send the aggregated data to the base station.
end for.
```

3.2 Redundant Data Elimination

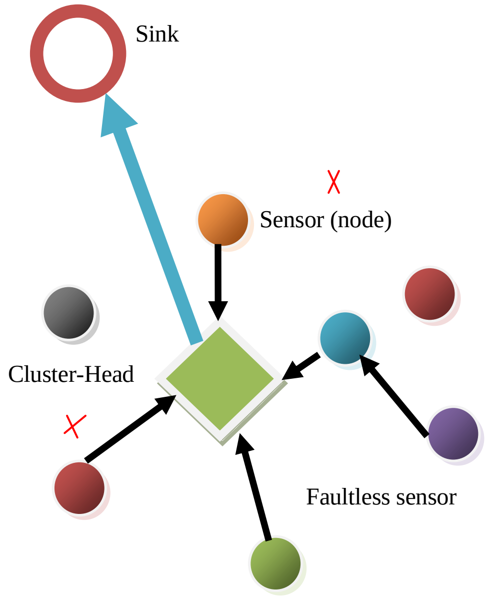

Figure 2(a): Normal case of Sensor Network.

Normal case of wireless sensor networks, probability of failure occurrence is more. When sensor nodes become incorrect it will not be identified, although incorrect data transmitted to the aggregator. Thereby increasing the transmission overhead and energy in the network as shown in Figure (2) a.

We have hashed code similarity and Rough set theory. Incorrect data consume a certain amount of energy, although further incorrect data will not send to sink or mobile station. Sensor having more than one similar data set, rough set theory will take only one set among the similar data set. Figure 2(b). It does not send incorrect data to sink. So that transmission overhead is reduced and longevity increased.

Data transmission paths demonstrated in the sensor networks. First sink sends a request to the cluster head in the corresponding region. Once the request is made depending upon the sensor location, cluster head delivers data to base station. If the sensor or node is not located in the region then the node is transferred to next cluster-head. If the nodes identified in the region of location it sends a data to the base sink. Then sink forms a selection path in order to send the data. In a cluster - head network, there are two acknowledgments one of the sensor node to the cluster-head and another from aggregator to the sink.

Simple aggregation algorithm tree construction of wireless sensor networks, as in [11]. The sensor nodes broadcast message from the leaf node to the parent node to all nodes receive the message. The network performs aggregation operation on parent node and its sub-node. Moreover, aggregation performed at various levels in the tree such as intermediate node, leaf node and leader node aggregation. Mobile station decrypts the aggregated.data and save it in the respective formats. After receiving the messages base station verify the authenticity of received data to ensure security

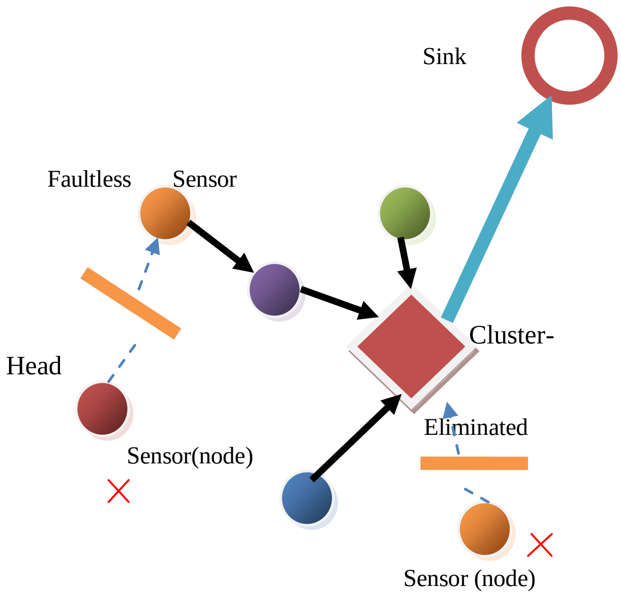

Figure 2 (b): Sensor with Error Correction network.

The cluster - head has the count of outliers and faultless data. Faultless data should send to cluster-head for aggregation. And the rest of outliers count eliminated. Finally the data aggregation process ensures that no redundant data transmission and no compromised node transmitted from the sensors. Sink receives the aggregated data without any fault. Fault- tolerant aggregation scheme provides better performance and has relieved from the intricacy.

4. SIMULATION RESULTS

Sensors are randomly distributed in terrain to measure the outer environment. The data aggregation scheme is simulated using the Network Simulator-2. Data set size of sensor is k=16 and data set of similarity code b=16 bits. The data size can be varied from different sizes 4, 8, 16, 24 and 32 bits. To carry out the operation 32 bit data transmitted.

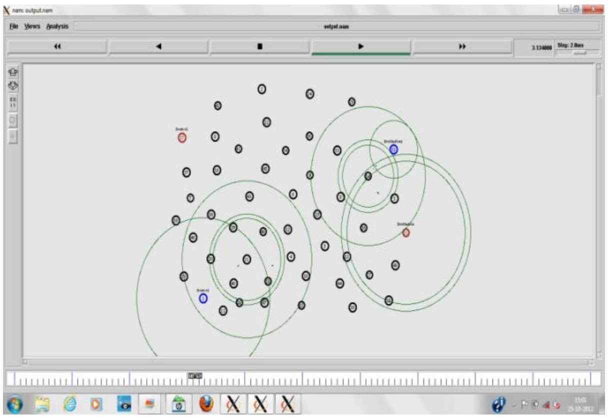

Figure 3: Total Data transmission aggregator to the sink

Aggregation accuracy is measured based on the outcome of the error in the cluster head. The error is completely eliminated in the cluster head one and only if the difference between the aggregated data computed by the CH and data sent by the sensor without fault. The result is presented in Figure (3) and compared with the neural network. As a consequence, BER (Bit Error Rate) is reduced with increasing faulty sensor. Fault tolerant aggregation scheme successfully detect adversary and outlier with high precision. Aggregator eliminates all outliers in the network. Data aggregator does not receive faulty data result in accurate aggregation. Assume that network has 49 sensors in the network which is randomly distributed in the environment. All the sensors try to send the data to the cluster-head, and use some amount of power and graph is plotted in Figure (4).

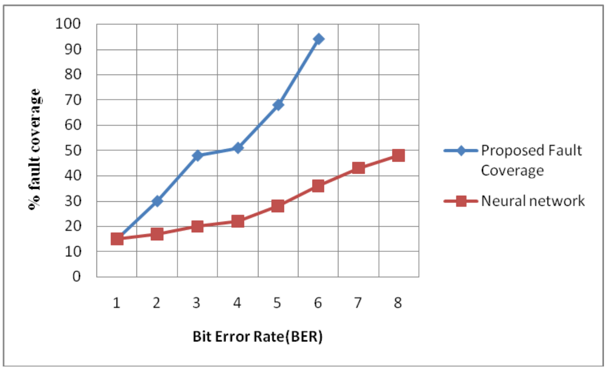

Figure 4: Fault Coverage

5. CONCLUSION

This paper presents the problem of Detecting and eliminating the faulty node in wireless sensor network with clustered architecture. Once the data are classified incorrect data can be removed and remaining data are to be aggregated in the CH and send it to base station. As a result redundant data are eliminated thereby increasing the data aggregation accuracy. Moreover, Transmission overhead is reduced. Another result is achieved by using this method is increase the longevity of wireless sensor networks.

6. REFERENCES

[1] C.Castelluccia, E.Mykletun and G.Tsudik Efficient aggregation of encrypted data in wireless sensor networks, Proc.Second Ann. Int`l Conf. Mobile and Ubiquitous systems, pp.109-107, July 2005.

[2] H.Cam, S.Ozdemir, P.Nair, D.Muthuavinashiappan and H.Ozgur sanli, Energy efficient secure pattern based data aggregation for wireless sensor networks, J.computer communication, vol.29.pp446-455, 2006.

[3] H.Salni, S.Ozdemir and H.Cam Secure reference based data aggregation protocol for wireless sensor networks, vol.7, and pp.4650-4654.sep.2004.

[4] J-Y.Chen, G.Pandurangan and D.Xu robust computation of aggregates in wireless sensor networks distributed randomized algorithm and analysis, IEEE Trans. Parallel distributed system vol.17,no.9,pp 987-1000, sep-2006.

[5] R.Rajagopalan and P.Varshney, Data aggregation techniques in wireless sensor networks, A survey, IEEE comm. survey tutorials, vol.8, no.4, pp.48-63,Oct-Nov.2006.

[6] Saeid Bahanfar, Helia Kousha, Ladan Darougaran, Neural networks for error correction and data aggregation in wireless sensor networks,vol.8,issue 5,no3, September 2011.

[7] L.J. Mpanza, T.Marwala, ant colony optimization of rough set for HV bushings fault detection.

[8] H.Yu secure and highly available aggregation queries in large scale sensor networks via set samping, proc.IEEE int`l conf. information processing in sensor networks.pp.1-12, 2009.

[9] Stefano chessa, paolo santi, Crash fault identification in wireless sensor networks.

[10] S.Basagni, M.Mastrogiovanni, A.panconesi, and C.Petrioli localized protocol for ad hoc clustering and backbone formation: a performance comparison, IEEE Trans, parallel distributed system, vol.17, no.4pp.292-306, apr 2006.

[11] Y.Yand, X.Wang, S.Zhu and G.Cao A survey hop by hop data aggregation protocol for sensor networks.ACM Trans. Information and security(TISSEC), vol,11,no.4,pp.1-43,2008.

[12] Yan sun, Hong Luo, Sajal K.Das, A true based frame work for fault tolerant data aggregation in multimedia wireless sensor networks, vol.9, no.6, 2012.

[13] D.Boneh, C.Gentry, B.Lynn and H.Shacham Aggregate and verifiably encrypted signatures from bilinear maps, proc.22nd int`l conf theory and application of cryptographic techniques, pp.416-432,2003.

[14] G.De Meulenaer, F.Gosset, F.X.Standaert and L.vandendorpe, on the energy cost of communication and cryptography in wireless sensor networks, proc.IEEE int conf. wireless and mobile computing, networking and comm, pp.580-58.