AN EFFICIENT WAY TO DETECT FAULTY NODE IN WSN

Tanushri Chandraker[1], Santosh Kumar[2]
Computer Science & Engineering, NGF College of Engineering & Technology, Palwal, INDIA

Abstract: Detection of fault in wireless sensor network is tedious task, because of the large scale of sensor nodes in the network. Thus to identify the fault node, to collect information from individual nodes need to be process. Because of the large volume of the sensor nodes irrelevant data or inaccurate of data passed to the base center, median, mean mode approach are used .Neighbor coordination approach is used in this paper. This paper presents a novel approach for detecting sensors which produce faulty data in adistributed way as well as identifying the type of datafaults using trust concepts to gain a high degree ofconfidence. We validate our method with simulations results.

Keywords: Fault node detection, Accuracy detection, trust, Reputation

I. INTRODUCTION

The task of every sensor node is to observe the environment and send the current report to the center, which is called as Base Station. In a Wireless Sensor Node (WSN) each node monitor the environment and to report to the center. If any node sends irrelevant data to the center i.e. that has become faulty node. Based on the inaccurate data, the base station may take wrong decision; become the unproductive of the network.In the recent years, WSNs haveInfluenced our existence by providing various services like remote environmental monitoring, infrastructure management, target tracking, target localization, home and office security, public safety, event detection, event boundary identification, medicine, transportation and many more.[1,2] Detection of fault node is one of the big challenges, building up the dependable protocol. Faulty node is detected, remaining node can be insulated from the network, and computation can proceed [4].

TO collect the information from every sensor and notice the faulty node sensors need to be charged. Batteries are powerful recourse to charge the sensor. Batteries are very expensive for the center in every aspect, to get information from nodes, to identify the faulty node in centralized manner. In real time mode different application require fault detection with high throughput. So a localization and distributed generic algorithm are highly used in wireless sensor node.

Crash fault and soft fault are two main categories in sensor fault. A SN (sensor node) becomes unable to communicate with other nodes, called as crash fault, and that SN become inactive in the network. Where as in the network, a SN starts behaving randomly is called as soft fault. Therefore it becomes important task to identify the set of SNs in the network. There are many restrictions in SNs in terms of storage, proficiency, power efficient and communiqué capability due to their small range. However SNs have many reasons of fault i.e. mechanical or electrical problems like problems in their internal circuit, power supply degradation, even hostile tampering. If any SNs Installed incorrectly or harsh environment then SNs behaves operate autonomously. All are the reasons of arising fault frequently.

II. RELATED WORK

Methods of fault detection based on simplified Algorithms such that they could be implemented on a single sensor node are investigated in [3][1]. Both have approaches based on two relationships of correlation: the correlation of a node's measurement and its neighbor's measurements, and the correlation of a node's measurement and its own previous measurement. In [3], a naive Bayes algorithm employed which maintains counters of the number of times a particular pair occurs over the history of the sensor network. Several fault node detection and fault diagnosis techniques of distributed WSNs have been proposed in [5-10].

Krishnamachari et al. [5] have proposed a Bayesian fault identification approach to resolve the fault event disambiguation problem in WSNs. Koushanfar et al. [6] have presented a cross validation based approach for online detection of faulty SNs in WSNs. In this approach, statistical methods are used to detect the sensors which are having the highest probability of faults.

Ruiz et al. [7] have proposed an external manager based fault node identification approach for event driven based WSNs. Even though the external manager is capable of performing more complex tasks than the typical SNs, still there exist a problem of communication between the SNs and the external manager. However, there are some kinds of faults which require cooperative-diagnosis among a set of sensor nodes. A large portion of faults in WSNs are in this category. For example, Detection method proposed in [8] is to identify faulty sensor nodes in event detection application. The detection method is based on the assumption that sensor nodes in the same region should have similar sensed value unless a node is at the boundary of the event region. It takes measurements of all neighbors of a node and uses the results to compute the probability of the node being faulty.

Chen et al. [9] have presented a distributed fault node detection approach for WSNs. In this approach, local comparisons are done using a modified majority voting technique. In this approach, each SN compares its own sensed data with its neighboring node’s data and based on which, a decision has to be taken by taking all neighboring nodes in confidence. However, the approach becomes little complicated because the exchange of information between two neighboring nodes is done twice in order to reach a local decision of fault status which is based on a threshold value.

III. MODEL AND PROBLEM FORMULATION

The problem of producing reliable information can be diminished to a basic question: "how the sensor nodes trust each other?" Trust is the expectation of one entity about the actions of another [11].There is much confusion between trust and reputation. When entities face uncertainty, they tend to trust to entities that have high reputation. Reputation is not a physical quantity but it is a belief; it can only be used to statistically predict the future behavior of other nodes and cannot define deterministically the actual action performed by them.

Trust is a subjective expectation a node has about another node's future behavior. This can be obtained by taking the statistical expectation of the probability distribution representing the reputation between the two nodes. Note that, unlike reputation, the trust metric is simply a number.

We have used the RFSN method for calculating the reputation; more details on RFSN can be found in [11]. Fault node detection in WSNs has become prime area of study. Fault node detection approach can be categorized into two basic types: centralized approach and distributed approach. In this section, model for the distributed sensor network is implemented by using the open source software NS2 and the problem of fault node detection is investigated using distributed approach.

Distributed approach is based upon ultra-reliable SNs having high computation capacity and large storage. The SNs are connected with each other within a specified transmission range T. Here the connection of one node with other neighbor nodes is based upon disc that formed by transmission range T. All the nodes within the disc are connected with each other and share their sensed data among themselves. Each node compares its own sensed data with the data of its neighbors, which are inside a disk and hence identifies its own fault status within a WSN. Distributed WSNs consist of a large number of SNs which are presented using network simulator NS2.

We have implemented a distributed topology which consist of 50 nodes (N=50) and simulated in network simulator NS2 as shown in Fig. 1. Each SNs are treated as the vertices in a graph G = G (V,E) where V is the number of vertices (nodes) and E is the number of edges between two vertices in the graph G. Each node-n present inside a disc and its neighbor nodes are linked with each other. Each node communicates its sensed data with all other SN present inside the disc and thus identifies its fault status using data of all the neighboring SNs in WSN.

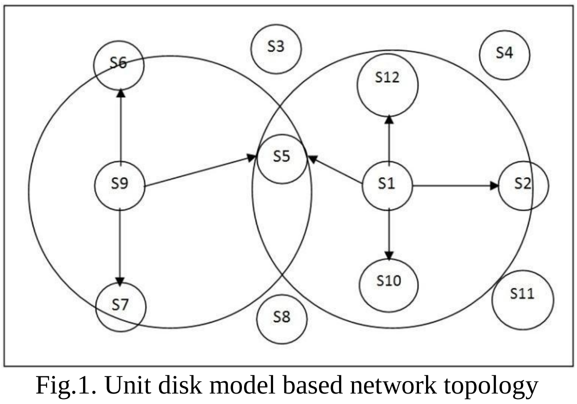


Fig.1. Unit disk model based network topology

A fault can be introduced into sensor at every point in the sensor network: from failures in the sensor itself, to software bugs and computational errors, to lossy communication [11]. Faults are application and sensor type dependent. There is no significant list of features to consider for fault modeling. The set of models to be used for fault detection contain models for both good and faulty data. Ni et also [11] categorize the existed faults as: outlier, stuck-at, spike, excessive noise, calibration fault, hardware fault, low battery, and clipping. Sharma et also [12] categorize faults differently as: NOISE, SHORT and CONSTANT.

The three types of faults that we intend to detect, as well as how they can be detected, are as follows.

SHORT: A sharp change in the measured value between two successive data points.

Detection method: we set an acceptable range for every hours of day and every season. If thereading value is outside this range and the difference between current and previous readingvalue is above a threshold, it is labeled as outlier.

NOISE: The variance of the sensor readingsincreases. Unlike SHORT faults that affect asingle sample at a time, NOISE faults affect anumber of successive samples.

Detection method: Compute the standarddeviation of sample readings within a window W. If it is above a certain threshold, the samples are affected by the NOISE fault.

CONSTANT: The sensor reports a constant value for a large number of successive samples. The reported constant value is either very high or very low compared to the "normal" sensor readings.

Detection method: if for several consecutive time intervals the read value by a sensor is invariant then we must check the read values by the sensor's neighbors. If most of the fault free sensors have variant values in these time intervals then this sensor has constant fault.

The statistical distributed fault node detection algorithm to detect the soft faulty nodes in WSN is described.The overall algorithm can be divided into three distinct major stages:-

1. Data Collection stage
2. Analysis stage
3. Decision stage

In various stages, every SN Ni does various tasks. In the initial stage, every SN Ni requests all their neighbors by transmitting their own sensed data X(i) at an instance of time. At the same instance of time, every node Ni collects sensed data X(j) from all their neighbor nodes Nj ϵ Neb(i) which are present in its transmission range T. Now in analysis phase,every SN Ni estimates its probable sensed value with the help of all of their neighbor's data.

In decision stage, after estimation of probable sensed data, every SN Ni equates its sensed data with the probable sensed data to evaluate its fault status. Based on this fault status, Ni will actively participate in the network operation. The perceptive descriptions of all events that occur during the three various stages of the suggested distributed fault detection algorithm are as follows:-

1) Data Collection Stage: In this stage, each SN Ni, transmits its own sensed data X(i) to all neighbor nodes within its transmission range T. Based on the received data, every SN Ni identifies its neighbor nodes Nj ϵ Neb(i). It is supposed that at the time of installation, every SN Ni deployed in network is good and free from faults. Hence initial fault

status of SN IFS(i)=0.

2) Analysis stage: Every SN Ni computes their estimated sensed data PX(i) from the data which is received from their neighbor nodes Nj which are present in its transmission range T. A Z-score test is a statistical approach, in which the dissemination of the test statistics under the null hypothesis can be estimated by a normal distribution. Since it is assumed that the data received by various SNs are independent of each other. Hence if the degree of the each node is high, the z-score test is preferable. The value of mean M(i) of the data received by neighbor node N(j) at SN Ni is calculated by using the formula.

$$M(i) = 1/(Degree\ (Ni))*\sum j=1* y(j) \quad (1)$$

Now, to evaluate the value of standard deviation S(i) of the data received by node N(i), the following formulas is used.

S(i)=rootof(∑ j=1*(square of (y(j)-μ)/ (Degree(Ni)-1)

Finally, by using z-score test and with the help of values of M(i) and S(i), we can calculate the value of PX(i) as follows.

$$SE(i)= root\ of(s/Degree(Ni)) \quad (3)$$

$$PX(i)= (y(i)-M/SE(i)) \quad (4)$$

Decision Stage: In this stage every SN Ni compares its own estimated data PX(i) with the actual sensed data y(i).

$$D(i) = |y(i)-PX(i)| \quad (5)$$

If difference between the sensed data X(i) and estimated data PX(i) is within the range of 2 to 3 then SN Ni declares itself as a fault free node otherwise it is decided as a faulty SN.

$$FFS(i)=\{0 \text{ if } 2 \le |D(i)| \le 3 \text{ (good)} \\ 1 \text{ if } |D(i)| < 2\ or\ |D(i)| > 3\ (faulty) \quad (6)$$

Since the algorithm is based on distributed WSN, so all the stages in the given algorithm A are to be executed by the every individual SN Ni of a WSN. The notations used for developing this algorithm are summarized in Table I.

Once again the accuracy detection (ad), fault rate alarm (fra) and Fault rate positive (frp) calculated for different probabilities of faulty nodes. But as the presence of outliers affect the actual mean (M) and standard deviation (S), estimated mean (M) and standard deviation (S) this leads to further investigation. We have analyzed the presence of faulty SN in WSNs by calculating the Z-score.

The neighbor coordination algorithms for detecting the faulty node in distributed WSNs can be described as follows:

A. Algorithm for detection of Faulty Node

Input: SN Position (xi, yi), sensed data X[i].

Output: detection accuracy (DA), fault alarm rate (FAR) and fault positive rate (FPR).

```
1. FOR each node i = 1 to n DO
2. Generate sensed data y[i]
3. Find all neighbors and keep them in set Nb[i]
4. Set IFS[i] = 0;
5. END FOR
6. Calculate F = n * pf
7. FOR j = 1 to f DO
8. random[j] = generate(n)
9. set UIFS[random[j]]=1
10. END FOR
11. Sum[i] = 0;
12. FOR j = 1 to |Nb[i]| and n[j] = Nb[i] DO
13. Sum[i]=Sum[i] + y[i];
14. END FOR
15. Mean[i] = Sum[i] / Degree(Ni);
16. csd[i] = 0;
17. FOR j = 1 to |Nb[i]| and n[j]=Nb[i] DO
18. csd[i] = csd[i] + (X[j] - Mean[i])2;
19. END FOR
20. sd[i] = sqrt (csd / (Degree(Ni) - 1));
21. SE[i] = sqrt (SD[i] / Degree(Ni));
22. PX[i] = y[i] – CRDN[i] / SE[i];
23. If (|y[i]-PX[i]| >= 2) and (|y[i]-PX[i]| <= 3) then
24. Sensor node Ni is detected as Fault-Free
25. FFS[i] = 0;
26. Else
27. Sensor node Ni is detected as Faulty
28. FFS[i] = 1;
29. FOR i = 1 to N DO
30. IF UIFS[i] == 1 && FFS[i] == 1 THEN
31. initializecount_ad = count_ad+1
32. END IF
33. IF UIFS[i] = = 0 && FFS[i] = = 1 THEN
34. initializecount_fra = count_fra+1
35. END IF
36. IF UIFS[i] = = 1 && FFS[i] = = 0 THEN
37. Initialize count_frp = count_frp+1
38. END IF
39. END FOR
40. Evaluate ad = count_ad / f
41. Evaluate fra = count_fra / n-f
42. Evaluate fpr = count_fpr / f
```

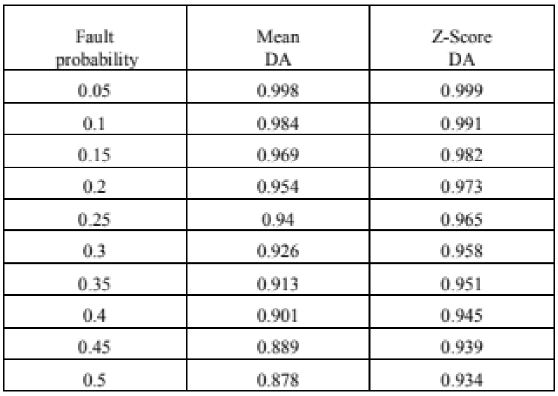
Here we have used neighbor coordination based distributed algorithm to evaluate the value of DA, FAR and FPR using mean. Then we have used neighbor coordination based distributed algorithm to evaluate the value of DA, FAR and FPR using z-score. The list of parameters which are used to develop the neighbor coordination based distributed algorithm for detection of faulty SN in WSNs in the given Table I.

Parameter	Description of Parameter
n	Number of sensor nodes
Ni Sensor node at location (xi, yi)	Sensor node at location (xi, yi)
T	Transmission range of sensor node Ni
y[i]	Sensed data of sensor node Ni

Nb[i]	Set of neighbor nodes of sensor node Ni
Degree[i] Degree of sensor node Ni	Degree of sensor node Ni
Sum[i]	Sum of sensed data of all neighbors nodes at sensor
node Ni	Mean[i] Estimated mean value of all neighbors nodes at sensor node Ni
csd[i]	Cumulative standard deviation of sensed data of all neighbors at sensor node Ni
sd[i]	Standard deviation of sensed data of all neighbor nodes at sensor node Ni
PX[i]	Z score test value of sensor node Ni
Pf	Probability of fault node
f	Total number of fault node
Generate(N)	Pseudo random number generation for fault node
random[i]	Array to store Pseudo random number for fault node i
IFS[i]	Initial fault status of the sensor node Ni
FFS[i]	Final fault status of the sensor node Ni
Count_ad	Counter for data accuracy(DA)
Count_fra	Counter for fault alarm rate(FAR)
Count_frp	Counter for fault positive rate (FPR)
da	Detection Accuracy
fra	Fault Alarm Rate
fpr	Fault Positive Rate

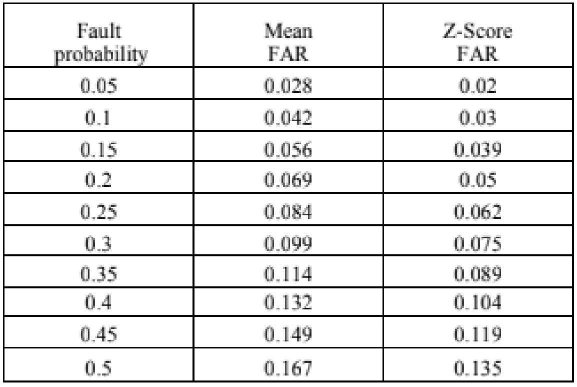
IV. EXPERIMENTAL RESULT

The performance measures of the proposed neighbor coordination based fault node detection algorithms are evaluated by using network simulation. We have build a distributed sensor network in network simulator tool (NS2) and network is simulated using the proposed algorithm described in the section II to evaluate the accuracy detection (ad), faultrate alarm (fra) and fault rate positive (frp) over different probability of fault nodes (pf). Where ad can be defined as the ratio of the number of faulty SNs detected to the total number of faulty SNs, fra can be defined as the ratio of the number of fault-free nodes detected as faulty to the total number of fault-free node and fpr can be defined as the ratio of the number of faulty SNs diagnosed as fault free to the total number of faulty SNs present in the network. Initially we have analyzed the faulty SNs in a WSN with algorithm-A to evaluate the ad, fra and fpr for 1024 nodes with the help of actual sensed data y(i) and standard deviation sd(i). Then we have used Z-score with algorithm-A to evaluate the ad, fra and frp with estimated PX(i) and standard deviation sd(i). Here the performance is improved as compared to the actual sensed data y(i) and standard deviation sd(i).



Fault probability	Mean DA	Z-Score DA
0.05	0.998	0.999
0.1	0.984	0.991
0.15	0.969	0.982
0.2	0.954	0.973
0.25	0.94	0.965
0.3	0.926	0.958
0.35	0.913	0.951
0.4	0.901	0.945
0.45	0.889	0.939
0.5	0.878	0.934

TABLE II.Accuracy Detection with Mean and Z-Value.



Fault probability	Mean FAR	Z-Score FAR
0.05	0.028	0.02
0.1	0.042	0.03
0.15	0.056	0.039
0.2	0.069	0.05
0.25	0.084	0.062
0.3	0.099	0.075
0.35	0.114	0.089
0.4	0.132	0.104
0.45	0.149	0.119
0.5	0.167	0.135

TABLE III. False rate alarm with Mean and Z- value

Table III shows the comparison of ad using Mean and ad using Z-Score, over various values of fault probability pf. With the increase in value of fault probability pf, the value of ad decreases. When the value of pf=0.05, 99.8% of the faulty nodes are accurately detected using Mean where as 99.9% of the faulty nodes are accurately detected using Z-score.

V. CONCLUSION

We have proposed a coordination based solution to detect the SNs in wireless Sensor Network. After simulation result shows that projected fault node detection algorithm is operative in terms of ad(Accuracy Detection), fra(False Rate Alarm), frp(False rate Positive). Planned algorithms are not using any complex operations and consume energy efficiently. Number of faulty nodes kept fixed during whole simulation process. So the work will be upgraded for dealing with variable number of faulty sensor nodes in a network. We will extend and modify the planned faulty node detection algorithm to tolerate transient faults in future work staff of CSE department for their technical support.

REFERENCES

[1] B. R. Badrinath, M. Srivastava, K. Mills, J. Scholtz,and K. Sollins.“Special issue on smart spaces and environments”. IEEE Personal Communications, Oct 2000.

[2] S. Lindsey, C. Raghavendra, and K.Sivalingam. “Data gathering in sensor networks using the energy delay metric”. In International Workshop on Parallel and Distributed Computing: Issues in

Wireless Networks and Mobile Computing, San Francisco, USA, Apr 2001.

[3] E. Elnahrawy and B. Nath,"Context-aware sensors", In Proceedings of the European Conference on Wireless Sensor Networks, pages 77-93, 2004.

[4] S. Chessa, P. Santi, "Comparison-Based System-Level Fault Diagnosis in Ad Hoc Networks", Proceedings of twentieth IEEE Symposium on Reliable Distributed Systems, 200 I, pp. 257-266.

[5] B. Krishnamachari, S. Iyengar, “Distributed Bayesian algorithms for fault tolerant event region detection in wireless sensor networks”, IEEE Transactions on Computers 53 (3) (2004).

[6] F. Koushanfar, M. Potkonjak, Sangiovanni-Vincentelli, “On-line fault detection of sensor measurements”, IEEE Sensors 2 (2003) 974–980.

[7] L.B. Ruiz, I.G. Siqueir a, L.B. Oliveira, H.C. W ong, J.M.S. Nogueira, A.A.F. Liureiro, “Fault management in event-driven wireless sensor networks”, in: MSWIM’04, 2004.

[8] Krishnamachari and Iyengar,"Distributed Bayesian algorithms for fault-tolerant event region detection in wireless sensor networks", IEEE Trans. Comput. Vol.53, no.3, pp. 241-250, March2004.

[9] J. Chen, S. Kher, A. Somani, “Distributed fault detection of wireless sensor networks”, in: Proceedings of 2006 Workshop DIWANS, pp. 65-72.

[10] M. Ding, D. Chen, K. Xing, X. Cheng, “Localized fault-tolerant event boundary detection in sensor networks”, IEEE Info com (2005) 902– 913.

[11] S. Ganeriwal, L. K. Balzano, and M .B. Srivastava,"Reputation-Based Framework for High Integrity Sensor Networks", ACM Transactions on Sensor Networks, Vol. V, No. N, March ,2007.

[12] Ni, K., et als., "Sensor Network Data Fault Types", ACM Transactions on Sensor Networks, in review.

[13] A. Sharma, L. Golubchik, R. Govindan, "On the Prevalence of Sensor Faults in Real-World Deployments", Proceedings of the forth Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Networks, pp. 213-222, 18-21 June 2007.