DISTRIBUTED FAULT-TOLERANT EVENT DETECTION FOR NON-SYMMETRIC ERRORS IN WIRELESS SENSOR NETWORKS

Nandita Das1, B. Victoria Jancee2 and S. Radha3

1PG Student, Department of ECE, SSN College of Engineering, Chennai, India
2Associate Professor, Department of ECE, St. Joseph‟s College of Engineering, Chennai, India
3 Professor and Head, Department of ECE, SSN College of Engineering, Chennai, India

Abstract

Wireless sensor network (WSN) are powered by batteries to perform various sensing tasks in a given environment. The measurements made by the sensors are sometimes unreliable and erroneous due to noise in the sensor or hardware failure. For a large scale WSN to be economically feasible, it is important to ensure that the faulty node does not affect the overall behaviour of the system. In this paper a binary fault-tolerant event detection technique has been proposed for the non-symmetric errors and its performance has been analysed. Theoretical analysis and simulation show that almost 97 percent of faults can be corrected even when 10 percent sensor nodes are faulty.

KEYWORDS

Non-Symmetric Errors, Fault Tolerance, Event Detection

1. INTRODUCTION

WSN consists of network of autonomous sensors which are powered by batteries to perform various sensing tasks in a given environment. These networks are used in various applications like detection, estimation, monitoring, tracking etc [14],[19],[20].

Lot of effort has been made to develop the hardware and software architectures of sensor devices as per the requirements of the wireless sensing applications. The various challenges and design issues of WSN has been addressed in a number of works [14],[17],[18],[20]. In this work the problem of event detection is addressed. Event detection in the inaccessible environment is one important application. The measurements made by the sensor are sometimes unreliable and erroneous due to noise in the sensor or hardware failure. It is therefore mandatory to employ a fault-tolerance mechanism which can help avoid and correct the failure of any node and also, which does not affect the overall performance and behavior of the system.

DOI :10.5121/ijscai.2014.3203

2. EVENT REGION DETECTION MODEL

Sensor measurements in the operational regions are always spatially correlated while the sensor faults are likely to be stochastically uncorrelated. Having these two mail assumption we put forward an algorithm for event detection in a fault tolerant manner. To tackle faults in WSN, the system should follow two main steps. The first step is event detection. It is to detect that a specific functionality is faulty, and to predict that it will continue to function properly in the near future. Fault recovery is the second step to enable the system to recover from faults.

Normally, an event, if it happens, should be detected as "event" by sensors at the location. The faulty behaviour we consider occurs when the detection decision is converted to "no-event" due to sensor fault or vice versa.

The first step in event region detection is for the nodes to determine which sensor reading is interesting. Here interesting reading means the readings of interest. By using a threshold the node can determine whether their reading corresponds to an event. The threshold can be specified with a query or otherwise made available to the nodes during deployment.

A more challenging task is to disambiguate events from faults in the sensor readings since an unusually high reading could probably correspond to both. It is assumed that sensor faults are uncorrelated while the event measurements are correlated.

One of the key challenges in detecting event in a WSN is how to detect it accurately transmitting minimum information providing sufficient details about the event. For this reason a fault-tolerant event detection scheme has to be implemented. A possible solution can be given by providing high degree of redundancy to compensate for the faulty nodes. However, the cost sensitivity and energy limitations of sensor networks make such an approach undesirable. So a better and efficient approach is adopted by collaboration between neighbouring nodes. This increases the reliability of detection decisions. Here fault-tolerant event detection is addressed in context of distributed binary detection for non-symmetric errors.

3. FAULT RECOGNITION

Standard Wireless sensor deployment experiences show that the data collected is shown to be imprecise due to internal or external influences. So an early recognition of faults is necessary for the effective operation of the network as a whole. In an environment where the event readings are typically spread out geographically over multiple contiguous sensors, faults can be disambiguated from the events by examining the correlation in the readings of nearby sensors.

The real situation at the sensor node to be modelled by a binary variable is given as T_i. This variable T_i=0 if the ground truth is that the node is in normal region and T_i=1 if it is in an event region. The real output of the sensor is mapped into a binary variable S_i. This variable S_i=0 if the sensor measurement indicates normal value and S_i=1 if it measures an unusual value. Thus there can be four possible scenarios which are shown in the table below.

Table 1. Possible scenarios of sensors

S_i	T_i	**Scenario**
0	0	Sensor correctly reports a normal reading
0	1	Sensor faultily reports an unusual/event reading
1	0	Sensor faultily reports a normal reading
1	1	Sensor correctly reports an unusual/event reading

By implementing the fault recognition algorithm an estimate *Ri* can be determined of the true readings *Ti* after obtaining information about the sensor readings of the neighbouring sensors.

4. Decision Schemes For Fault Recognition

There are various decision schemes for fault recognition. Here three schemes are examined [1]. They are mentioned below.
1. Randomised decision scheme
2. Threshold decision scheme
3. Optimal threshold decision scheme

The detailed descriptions of these schemes are explained in the following sections.
Here the sensor fault probability *p* is assumed to be uncorrelated and non-symmetric.

$$P(Si=0|Ti=1)\neq P(Si=1|Ti=0) \qquad (1)$$

If $P(Si=0|Ti=1)=p1$ and $P(Si=1|Ti=0)=p2$ then p is such that $p=(p1+p2)$. Thus the probability that there is no fault in the sensor is $(1-p)$.

The binary model is obtained by applying threshold on the real-valued readings of the sensor. If *mn* is the mean of normal reading and *mf* is the mean of event reading, then a reasonable threshold for distinguishing between the two probabilities can be given as

$$\Theta= 0.5(mn+mf) \qquad (2)$$

The errors due to sensor faults and environmental fluctuations are modelled as Gaussian distribution with mean 0 and a standard deviation σ. The fault probability can thus be given as

$$p=Q((mf-mn)/2\sigma) \qquad (3)$$

Q function decreases monotonically, hence it can be said that the fault probability is low when mean normal and the event readings are not sufficiently distinguishable or when the standard deviation σ of the sensor measurement error is high. The assumption that the sensor failures are uncorrelated is a standard and reasonable assumption because these failures are primarily due to imperfections in manufacturing and not a function of nodes spatial deployment.

4.1. Randomised decision scheme

Let each node *i* have *N* neighbours and evidence $E_i(a,k)$ is that *k* of the neighbouring sensors report the same binary readings „*a'* as node *i*, while *N*-*k* of them report the reading '-*a*", then

$$P(R_i=a|E_i(a,k))= k/N \qquad (4)$$

The task of each sensor is to determine a value for R_i given information about its own sensor reading S_i and the evidence $E_i(a,k)$ reading the readings of the neighbour.

Assuming when, the error is symmetric. P_{aak} shows the statistics with which the sensor node makes the decisions about whether or not to disregard its own sensor reading S_i in face of the evidence from its neighbour.

$$P_{aak} = P(R_i=a|S_i=b,E_i(a,k)) \qquad (5)$$

The above expression can also be written as

$$P(R_i = a \mid S_i = b, E_i(a,k)) = \frac{P(R_i = a, S_i = b \mid E_i(a,k))}{P(S_i = b \mid E_i(a,k))} \qquad (6)$$

Thus the above expression can be simplified as follows

$$P_{aak} = \frac{(1-p)k_i}{(1-p)k_i + p(N_i - k_i)} \qquad (7)$$

Each node could incorporate randomization and announce if its sensor readings is correct with probability P_{aak}.

Then a random number *u* is generated such that u belongs to (0,1) . If $u<Paak$, then *Ri* is set to *Si* else *Ri is set to -Si*.

4.2. Threshold decision scheme

This scheme makes use of a threshold value θ which ranges from 0 to 1 i.e. 0<θ<1. If *Paak* > θ , then *Ri* is set to *a* and the sensor believes that the sensor reading is correct. If the metric is less than threshold then the node decides that the sensor reading is faulty and sets *Ri* to -*a*.

4.3. Optimal threshold decision scheme

The optimal decision threshold scheme also uses a threshold value. Here picking the threshold value θ is equivalent to picking an integer *kmin* such that the node decodes to a value *Ri=Si=a* if and only if at least kmin of its *N* neighbours report the same sensor measurements *a*.

5. ANALYSIS OF FAULT RECOGNITION ALGORITHM

The analysis of each of the Fault-Recognition scheme is given as follows.

5.1 Analysis of the Randomized Decision Scheme

Here an assumption is made that, if node *i* is in the event region, then all its neighbours are also in event region. And, if *i* is not in an event region, neither are any of its neighbours in event region. This assumption is valid everywhere except at nodes which lie on the boundary of an event region.

When the error is symmetric, *gk* is the probability that exactly *k* of node *i's N* neighbours are not faulty g_k is given as

$$g_k = \binom{N}{k}(1-p)^{k} p^{(N-k)} \quad (8)$$

Here *j1* and *j2* is integer. With the binary values possible for the three variables corresponding to the ground truth T_i, the sensor measurement S_i, and the decoded message R_i, there are eight possible combinations. The conditional probabilities corresponding to these combinations are useful metrics in analysing the performance of these fault recognition algorithms.
α gives the average number of errors after decoding.

$$\alpha = \left(1 - \sum_{k=0}^{N} P_{aak}(g_k - g_{N-k})\right) n \quad (9)$$

The reduction in average number of errors is thus given as *β* gives the average number of sensor faults corrected by the Bayesian fault recognition algorithm.

$$\beta = \left(1 - \sum_{k=0}^{N} P_{aak} g_{N-k}\right) np \quad (10)$$

A related metric is γ, which gives the average number of faults uncorrected.

$$\gamma = \left(\sum_{k=0}^{N} P_{aak} g_{N-k} \right) np \tag{11}$$

The Bayesian fault recognition algorithm has one drawback, though it can help us correct sensor faults, it may introduce new errors if the evidence from the neighbouring sensors is faulty. This effect can be captured by the metric δ, the average number of new errors introduced by the algorithm.

$$\delta = \left(1 - \sum_{k=0}^{N} P_{aak} g_k \right) (1 - p)n \tag{12}$$

5.2 Analysis of the Optimal Threshold Decision Scheme

In the optimal threshold decision scheme with a threshold value θ is equivalent to picking an integer k_{min} such that node i decodes to a value $R_i = S_i = a$ if and only if at least k_{min} of its N neighbours report the same sensor measurement a.

The optimal threshold value which minimizes α , the average number of errors after decoding , is $\theta = (1-p)$. this threshold value corresponds to $k_{min} = 0.5N$.

The performance metrics for the Optimal threshold decision scheme are listed below.
α : the average number of errors after decoding.

$$\alpha = \left(1 - \sum_{k=0}^{N} (g_k - g_{N-k}) \right) np \tag{13}$$

β : the average number of sensor faults corrected by the Bayesian fault recognition algorithm

$$\beta = \left(1 - \sum_{k=0}^{N} g_{N-k} \right) np \tag{14}$$

γ : gives the average number of faults uncorrected.

$$\gamma = \left(\sum_{k=0}^{N} g_{N-k} \right) np \tag{15}$$

δ : the average number of new errors introduced by the algorithm.

$$\delta = \left(1 - \sum_{k=0}^{N} g_k\right)\left(1 - p\right)^{n} \qquad (16)$$

5.2 Analysis of the Optimal Threshold Decision Scheme with Non-Symmetric Error

When the error is non- symmetric, g_k is the probability that exactly k of node *i's* N neighbours are not faulty g_k is given as

$$g_k = \left(\frac{N!}{k!*e!*t!}\right)\left(1 - p_1 - p_2\right)^{k} p_1^{\,e} p_2^{\,t} \qquad (17)$$

6. SIMULATIONS AND RESULTS

Some experiments to analyse the performance of the fault recognition algorithms. The simulation results of the randomized decision scheme and the optimal threshold decision scheme are analysed in detail.

The scenario consists of n=1024 nodes placed in a 32X32 square grid of unit area. The communication radius is set to $1/(\sqrt{n}-1)$ so that each node can communicate with its immediate neighbour in each cardinal direction. All sensors are binary: they report a "0" to indicate no event and a "1" to indicate there is an event. Thus each sensor has an independent probability of reporting a "0" or "1" or vice versa.

In figure 1, it is seen that, for $p<0.1$ (10 percent of the nodes being faulty on average), over 75 percentage of the fault can be corrected. However, this algorithm has a setback that, though it can correct sensor faults, it may introduce new errors if the evidence from the neighbour is faulty. Hence the number of new errors introduced δ is seen to increase steadily with the fault rate and starts to affect the overall reduction in error significantly after about p=0.1. From the figure 2, we can conclude that when the number of neighbourhood size is increased, then, for $p<0.1$, more than 75 percentage of fault can be corrected and the number of errors introduced can be reduced relatively.

6.1. Randomised Decision Scheme for Symmetric Error

The performance metrics for the randomized decision scheme with two different neighbourhood sizes and symmetric error has been analysed with respect to the sensor fault probability.The figures below shows the various metrics with respect to the sensor fault probability.

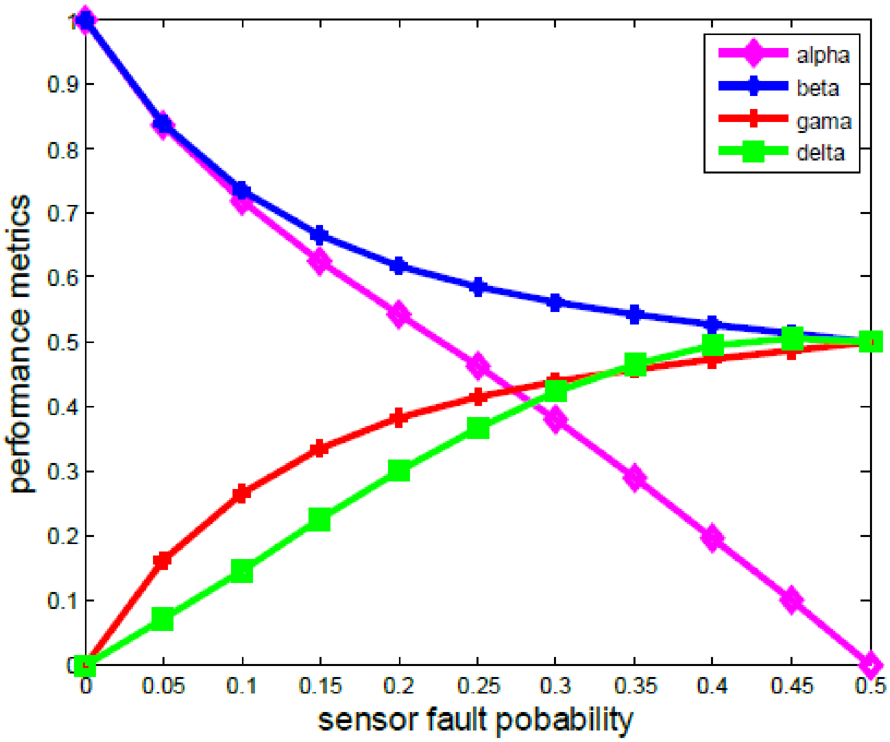

Figure 1. Metrics of randomized decision scheme for symmetric error (N=4)

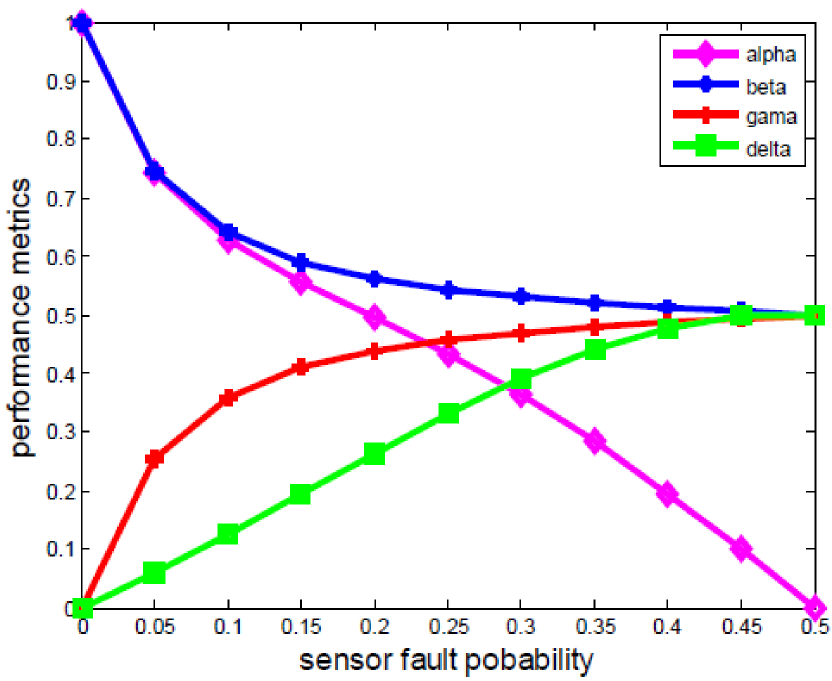

Figure 2. Metrics of randomized decision scheme for symmetric error (N=8)

6.2 Optimal threshold decision scheme for symmetric error

In the optimal threshold decision scheme with a threshold value θ is equivalent to picking an integer k_{min} such that node i decodes to a value $R_i=S_i=a$ if and only if at least k_{min} of its N neighbours report the same sensor measurement a.

The optimal threshold value which minimizes α, the average number of errors after decoding, is θ =(1-p). This threshold value corresponds to k_{min} = 0.5N. The optimal threshold value which minimizes α, the average number of errors after decoding, is θ=(1-p). This threshold value corresponds to kmin = 0.5N.

Here, if $k>k_{min}$ of its neighbour and also read the same value, the node decides on $R_i=a$, thus the P_{aak} term from (9) – (12) can be replaced by a step function.

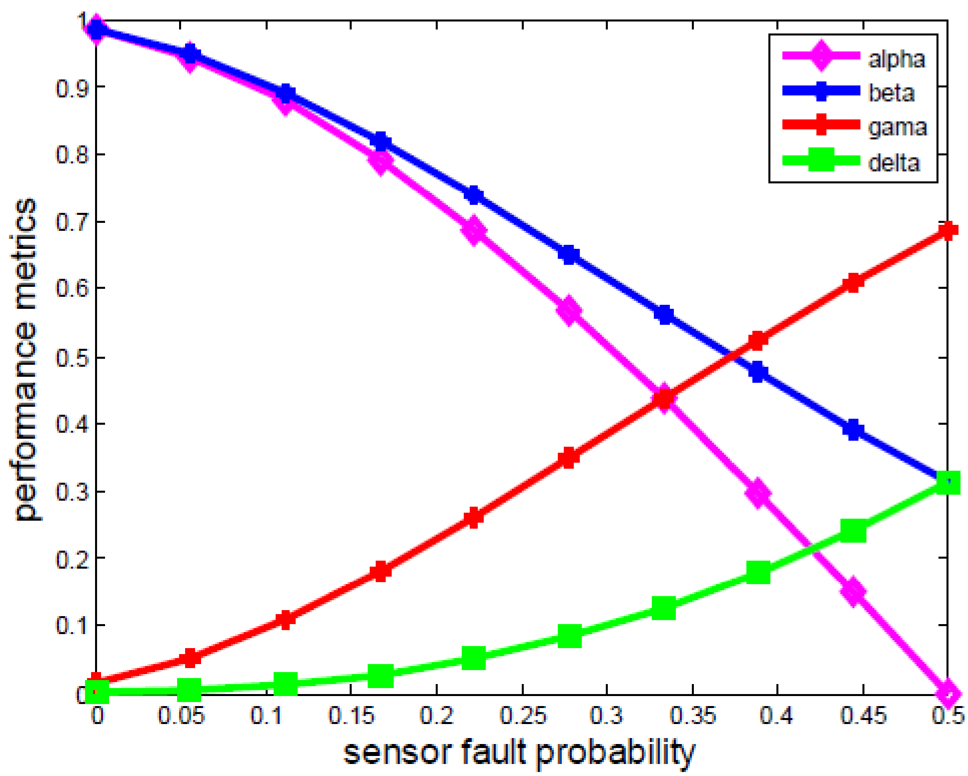

Figure 3. Metric of optimal threshold decision scheme for symmetric error (N=4)

The most significant way in which the simulations differ from theoretical analysis is that, the theoretical analysis ignores edge and boundary effects. At the edge of the deployed network, the number of neighbours per node is less than that in the interior and, also, the nodes at the edge of an event region are more likely to show erroneous reading if their neighbours provide wrong information. Such boundary nodes are likely to be sites of new error introduced by the fault recognition algorithm.

From the figure 3, it can be inferred that the number of new errors introduced in the optimal threshold decision algorithm is less than that of the randomized decision scheme. Thus it can be concluded that the best policy for each node is to accept its own sensor reading if and only if at least half of its neighbours have the same reading. This eases out the sensors work as it can help the sensor perform the optimal decision even without having to calculate the sensor error probability.

6.3. Optimal threshold decision scheme for non-symmetric errors

It is seen that the optimal threshold decision scheme is better than the randomized decision scheme**.** So implementing this scheme for the binary event detection with non-symmetric errors shows notable changes.

In the figure 4, it is seen that, for $p<0.1$ (10 percent of the nodes being faulty on average), more than 97 percentage of the fault can be corrected. However, this algorithm has a setback that, though it can correct sensor faults, it may introduce new errors if the evidence from the neighbour is faulty. Hence the number of new errors introduced δ is seen to increase steadily with the fault rate and starts to affect the overall reduction in error significantly after about $p=0.1$.

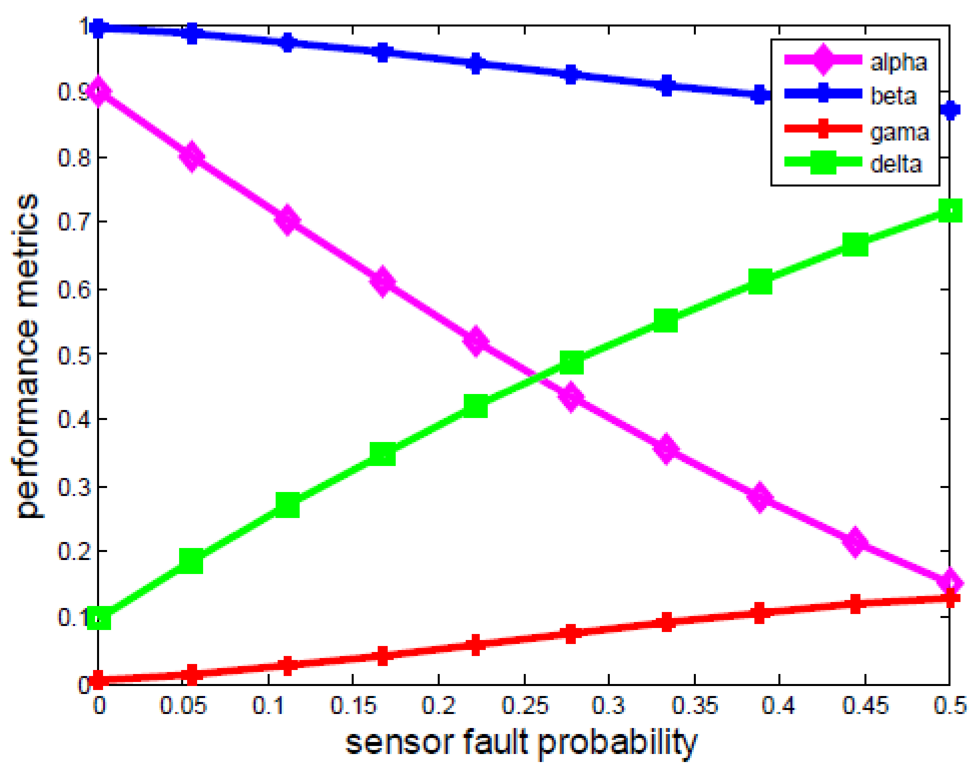

Figure 4. Metrics of optimal threshold decision scheme for non-symmetric error (N=4)

7. CONCLUSIONS

The randomised decision scheme and the optimal threshold decision scheme are thoroughly analysed. And the analysis showed that the optimal threshold decision scheme has a better performance in terms of minimising the error, and also, it introduced very lesser number of new errors (due to faulty evidence from the neighbouring sensors). It has also been observed that, in the optimal threshold decision scheme, the probability of detection error reduces if the number of neighbourhood nodes taken under consideration for the decision making process is increased. In the event detection technique for the non-symmetric error, more than 97 percent of faults can be corrected even when 10 percent sensor nodes are faulty**.** The limitation here is that the number of new error introduced is seen to increase steadily. In some practical applications like that of the real life environment monitoring, the chances of errors to be symmetric is very rare

because errors are always random in nature. In such scenarios, the technique of the fault-tolerant event detection for the non-symmetric errors can be applied.

ACKNOWLEDGEMENTS

The author would like to thank the co-authors for their constant support and innovative ideas and also their willingness to share the knowledge.

REFERENCES

[1] B. Krishnamachari and S. Iyengar, "Distributed Bayesian Algorithms for Fault-Tolerant Event Region Detection in Wireless Sensor Networks," IEEE Trans. Computers, vol. 53, no. 3, Mar. 2004.
[2] X. Luo, M. Dong, and Y. Huang, "On Distributed Fault-Tolerant Detection in Wireless Sensor Networks," IEEE Trans. Computers,vol. 55, no. 1, pp. 57-80, Jan. 2006.
[3] M. Li, Y. Liu, and L. Chen, "Non-Threshold Based Event Detection for 3D Environment Monitoring in Sensor Networks," Proc. 27th Int'l Conf. Distributed Computing Systems (ICDCS), 2007.
[4] EI. Moutapha Ould-Ahmed-Vall, B.H.Ferri, and G.F.Riley,"Distributed Fault-Tolerance for Event Detection Using Hetrogenous Wireless Sensor Networks", IEEE Trans on mobile computing, vol.11, no.12, Dec 2012.
[5] F. Kounshanfar, M.Potkonjak and A. Sanjiovanni-Vincentelli, "Fault Tolerance Techniques for Wireless Ad Hoc Sensor Networks", Proc.IEEE Sensors,2002.
[6] J. Chamberland and V.V.Veeravalli,"Distributed Detection in Sensor Networks", IEEE Trans. Signal Processing, vol.51, no.2, Feb.2003.
[7] L. Prasad, S.S.Iyengar, R.L.Rao and R.L.Kashyap,"Fault Tolerant Sensor Integration using multiresolution decomposition", Physical Rev.E,vol.49,no.4, Apr 1994.
[8] R.Vishwanathan and P.K Varshney,"Distributed Detection with Multiple Sensors:Part-I Fundamentals", Proc.IEEE, vol.85, no.1,1997.
[9] H.Li, P.Li, Z.Xie and M.Wu,"Event Region Fault-Tolerant Detection Based on Distributed Weight for Wireless Sensor Networks", Proc.IEEE,2008.
[10] R.Niu and P.K.Varshney," Distributed Detection and fusion in a large wireless sensor network of random size", EURASIP Journal on Wireless Communication and Networking 2005:4,462-472.
[11] M.P. Michaelides, C. Laoudias and C.G. Panayiotou,"Fault tolerant detection and tracking of multiple sources in WSNs using binary data", Joint 48th IEEE Conference on Decision and Control and 28th Chinese control conference, Dec 2009.

[12] A.A. Taleb, J. Matew, D.K. Pradhan and T. Kocak, " A Novel Fault Diagnosis Technique in Wireless Sensor Networks", International Journal on Advances in Networks and Services, vol. 2 no 4, year 2009.

[13] Q. Zhang, P.K. Varshney, and R.D Wesel,"Optimal Bi-Level Quantization of i.i.d. Sensor Observations for Binary Hyperthesis Testing." IEEE Trans. Information Theory, vol. 48, no. 7, pp. 2105-2111, 2002.

[14] G.J Pottie and W.J. Kaiser, "Wireless Integrated Network Sensors," Comm. ACM, vol. 43, no. 5, pp. 551-558, May 2000.

[15] J. Byers and G. Nasser, "Utility-Based Decision-Making in Wireless Sensor Networks" Proc. First ACM Int'l Symp. Mobile Ad Hoc Networking and Computing, pp.143-144, 2000.

[16] Q. Zhang, P.K. Varshney, and R.D Wesel,"Optimal Bi-Level Quantization of i.i.d. Sensor Observations for Binary Hyperthesis Testing." IEEE Trans. Information Theory, vol. 48, no. 7, pp. 2105-2111, 2002.

[17] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the World with Wireless Sensor Networks," Proc. Int'l Conf. Acoustics, Speech and Signal Processing (ICASSP 2001), May 2001.

[18] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Next Century Challenges: Scalable Coordination in Sensor Networks," Proc. ACM/IEEE Int'l Conf. Mobile Computing and Networks (MobiCom‟99), Aug. 1999.

[19] A. Cerpa et al., "Habitat Monitoring: Application Driver for Wireless Communication Technology," Proc. 2001 ACM SIGOMM Workshop Data Comm. in Latin America and the Caribbean, Apr. 2001.

[20] D. Estrin et al., "Embedded, Everywhere: A Research Agenda for Networked Systems of Embedded Computers," Nat'l Research Council Report, 2001