



ANALYSIS OF NON-BINARY FAULT TOLERANT EVENT DETECTION IN WIRELESS SENSOR NETWORKS

B Victoria Jancee¹, S Radha² and Nandita Das³

¹Department of Electronics and Communication Engineering,
St. Joseph's College of Engineering, India.

^{2,3}Department of Electronics and Communication Engineering,
Sri Sivasubramaniya Nadar College of Engineering, India.

Emails: bjjancee@yahoo.co.in, radhas@ssn.edu.in, and nanditadas101@gmail.com

Submitted: May 10, 2014

Accepted: July 26, 2014

Published: Sep. 1, 2014

Abstract- A distributed non-binary fault tolerant event detection technique is proposed for a wireless sensor network (WSN) consisting of a large number of sensors. The sensor nodes may be faulty due to harsh environment and manufacturing reasons. In the existing works on event detection, the detection of event is decided by only one threshold level. The objective of this paper is to extend the fault recognition and correction algorithm for non-binary event detection. The analysis presented here takes into account both the symmetric and non-symmetric error in a straightforward manner. In addition, simulation is done for symmetric error and 75 percentage of the errors can be corrected. The theoretical analysis shows that more than 95 percentage of symmetric errors can be corrected and almost 92 percentage of non-symmetric errors can be corrected (for $k=2$, i.e. half of the neighbors give correct decision), even when as many as 10 percentage of the sensor nodes are faulty.

Index terms: Distributed algorithm, fault-tolerance, non-binary event detection, non-symmetric error, symmetric error, wireless sensor networks.

I. INTRODUCTION

Wireless sensor network consists of network of autonomous sensors, which are powered by batteries to perform various sensing, data processing and communicating tasks in a given environment. Environment can be physical world, biological system, or information technological world. Because of the typical functions of the wireless sensor networks, these are potential for use in a wide variety of applications such as detection, estimation, monitoring, tracking etc [1], [2], [3]. One particular application that has received a growing amount of attention in the recent years is event detection.

In the detection application, the nodes are required to determine the occurrence of any event of interest. Event detection is commonly performed using a large number of unreliable low-cost sensor nodes. These nodes can have a very high probability of errors due to harsh environment and manufacturing reasons. It is, therefore, important to develop fault-tolerant mechanisms that can detect faulty sensor nodes and take appropriate actions. The sensors are generally deployed in high density manner and in large quantities. The measurements made by sensors are sometimes unreliable and erroneous. Some sensor node may fail or may be blocked due to lack of power, have physical damage or environmental interference. The failure of sensor nodes should not affect the overall task of the sensor network. This is the reliability or fault tolerance issue. Fault tolerance is the ability to sustain the sensor network functionality without any interruption due to sensor node failure [4].

There are wide varieties of applications in WSN. They include military applications, environmental applications, remote monitoring of physiological data, tracking and monitoring doctors and patients inside a hospital, drug administration, elderly assistance and home applications, environmental control in industrial and office buildings, inventory control, Vehicle tracking and detection. The works which are given below target these applications.

N.Samanta et. al. (2014) in [5] proposed a low power and minimally invasive multi sensor network which consists of current sensor node, motion sensor node and wearable heart rate sensor node. This system has been developed for health monitoring of elderly people. By using this proposed system, an average of 68% power saving has been achieved.

N.K.Suryadevara et.al. (2012) developed a mechanism to monitor the wellness of elderly people living alone by monitoring their daily activities everyday. The mechanism is based on the usage of household appliances connected through a range of sensing units. In order to determine the position of the elderly on performing essential daily deeds two wellness functions have been defined [6].

The authors Sachin Bhardwaj et.al (2008) focused on detection of arrhythmia disease, norm calculation, orientation calculation and fall detection to monitor an elderly person at home [7]. The analysis of ECG signals and fall detection of elderly patient could provide informative details to the doctors using PC/PDA.

Cheng Chunling et.al. (2013) proposed an algorithm for outlier detection in [8]. Outliers are the data sets which are inconsistent with other data observations. These data may convey some important information. A cluster algorithm based on flocking model has been developed to detect outliers. There by the accuracy and reliability of data collected by sensors could be enhanced for decision making.

There are a number of related works where fault tolerant event detection techniques have been proposed. Dyi-Rong Duh, Ssu-Pei Li, and Victor W Cheng (2013) [9] proposed a distributed and fault-tolerant algorithm to extend the uses of fault-event disambiguation. This algorithm can identify not only faulty and fault-free sensors but also the region where the event occurs. Simulation results demonstrated that it has high detection accuracy and low false alarm in any shape of event regions.

Xuefeng Liu et.al (2013) proposed a fault-tolerant and a distributed event detection scheme in structural health monitoring (SHM) to detect faulty sensor node [10]. In addition, a method called iterative faulty node detection is proposed to detect nodes with faulty readings by taking vector data as input.

E.Ould-Ahmed-Vall, B.H.Ferri and G.F.Riley (2012) [11], presented a general fault tolerant event detection scheme that allows nodes to detect erroneous local decisions by leveraging the local decisions reported by their neighbors. This approach considered the case where nodes can have different failure probability values. It also handled various types of failures such as noise related failures, biased measurement, drift over time, stuck-at failures, calibration-related failures and environmental related failures.

Sung-Jib Yim and Yoon-Hwa Choi (2010) in [12] presented a distributed adaptive fault-tolerant event detection scheme for wireless sensor networks. This scheme employed a filter for tolerating transient faults and achieved high performance for a wide range of sensor fault probabilities. Depending on the fault status of sensor nodes, the threshold for event detection is adjusted dynamically. The status of sensor nodes is maintained using confidence levels. Sensor nodes behaving incorrectly for an extended period of time are isolated from the network and replaced later if some necessary conditions on confidence levels are set up. Both high event detection accuracy and low false alarm rate can be maintained even with increasing number of faults.

In [13], Myeong-Hyeon Lee, Yoon-Hwa Choi (2008) exploited spatial and time information simultaneously to detect the sensor fault. In this, the status of each sensor node can be identified based on comparison of local neighboring sensors data with some thresholds. Using this scheme, faulty sensor nodes

are detected with high detection accuracy for a wide range of sensor fault probabilities, whereas the false alarm rate is maintained low. A fault-tolerant distributed decision fusion in the presence of sensor faults is addressed in [14] by Tsang-Yi Wang et.al (2008). In addition, they proposed collaborative sensor fault detection (CSFD) scheme for removing unreliable local decisions when performing distributed decision fusion.

Hong Li et.al (2008) in [15] proposed a distributed weighted fault-tolerant algorithm (DWFA) for nodes in regular deployment and an optimal fault-tolerant mechanism based on weighted distance (WDOFM) for nodes in irregular deployment. Both the algorithms exploited exchange of data between the sensor nodes and nearby nodes so that fault detection can be conducted on central nodes.

Tsang-Yi Wang and Qi Cheng (2008) explored the problem of event-region detection and the related problem of boundary-region detection in wireless sensor networks (WSNs) in [16]. Sensor nodes regarded here could observe heterogeneous regions, and the underlying event at sensor nodes can be correlated. Each node makes its decisions regarding which region it locates and/or if it is a boundary sensor. A fusion rule has been proposed considering communication constraints and possible sensor faults.

M.Li, Y.Liu, and L.Chen (2007) [17], proposed a non-threshold based approach for complex event detection in 3D environment monitoring applications. Event feature patterns are proposed to specify complex events and develop a pattern based event detection method on the obtained 3D gradient data map. Space OP model is proposed to describe the environment data distributions. Partial data maps are aggregated by merging OP regions with similar environmental data.

J.Chen, S.Kher and A.Somani (2007) [18], proposed and evaluated a localized fault detection algorithm to identify the faulty sensors. The implementation complexity of the algorithm is low and the probability of correct diagnosis is very high even in the existence of large fault sets.

X.Luo, M.Dong and Y.Huang (2006) [19], addressed two problems for distributed fault-tolerant detection in wireless sensor networks. Firstly, the noise related measurement error and sensor fault are addressed simultaneously in fault-tolerant detection. Secondly, an algorithm to choose a proper neighborhood size 'n' for fault correction of the sensor nodes is developed.

B. Krishnamachari and S. Iyengar (2004) [20], proposed a solution for binary detection of interesting environmental events and a solution to the fault-event disambiguation problem in sensor network. They exploited the notion that measurement errors due to faulty equipments are likely to be uncorrelated, while

environmental conditions are spatially correlated. They proposed a mechanism letting an individual sensor node communicate with its neighbors and using their binary decision to correct its own decision. A majority voting scheme has been shown to be the optimal decision scheme for fault tolerance in their work.

In binary fault disambiguation problem, when a node tries to decide whether a specific event is present or absent, a binary variable is used to code the decision. A value of 1 is shown when an event occurs and a value of 0 is shown to depict the absence of event [20]. The node uses the sensed data obtained by its local sensors as well as the decisions at its neighboring nodes.

In environment like that of Mushroom Cultivation, various physical factors have to be monitored for proper growth and yield of it. During the spawning and growth phase of the Mushroom Cultivation, the temperature of the cultivation room should be maintained between 75 degree Fahrenheit and 80 degree Fahrenheit. The cultivation of mushroom is very vulnerable to temperature fluctuations. In case when majority of the sensor nodes show even a slight temperature deflection from the required range, it can be understood that there may be change in the temperature provided in the cultivation area, so necessary measures can be taken. Not just temperature, other physical factors like relative humidity level, substrate moist level etc should also be monitored. Here the Wireless Sensor Networks may be used for monitoring these factors. In such scenarios, the binary fault disambiguation problem and the binary level fault tolerance cannot be used. Thus, there comes a necessity to use two threshold levels, i.e., a lower threshold level and an upper threshold level.

In this paper, a solution has been proposed for the non-binary event detection technique. The analysis has been done for two cases: symmetric error and non-symmetric error. Furthermore, simulation is done for symmetric error. In the symmetric error, the sensor fault probability is assumed to be uncorrelated and symmetric. i.e., Probability of sensor measurements indicating normal value given ground truth event value is equal to the probability of sensor measurements indicating event value given ground truth normal value. Moreover, in the non-symmetric error, the sensor fault probability is assumed uncorrelated and asymmetric. i.e., Probability of sensor measurements indicating normal value given ground truth event value is not equal to the probability of sensor measurements indicating event value given ground truth normal value. A majority-voting rule (k-out-of-n rule) is used to find out the nodes to agree on a decision before accepting the decision.

This research paper has been organized in the following manner: Section 2 defines the problem statement of the event detection technique. Section 3 discusses the analysis of the fault recognition algorithms for non-binary event detection in detail. The performance evaluation and results obtained are discussed in section 4. Finally, the concluding remarks and suggestions for further works are presented in Section 5.

II. PROBLEM STATEMENT

Consider a network of sensor nodes placed in an operational environment. The network is supposed to monitor the temperature of that particular environment within a range θ_1 to θ_2 , where θ_1 is the lower limit of the normal value and θ_2 is the upper limit of the normal value. The real output of the sensor is x . The non-binary model can result by placing two thresholds on the real valued readings of the sensor. The threshold may be specified with a query or made available to the nodes during deployment. The real situation at the sensor node to be modeled by a variable which is given as R_i . This variable $R_i=0$ if the ground truth is that the node is in normal region and $R_i=1$ or $R_i=2$ if it is in an event region. The real output of the sensor is mapped into variable M_i . This variable $M_i=0$ if the sensor measurement indicates normal value and $M_i=1$ if it measures a value which is lesser than the lower limit of the normal value and $M_i=2$ if it measures a value which is larger than the upper limit of the normal value. The mapping of real output of sensor x into the variable M_i is shown in the table below.

Table 1: Mapping of real output x of sensors to M_i

Output of the Sensor 'x'	M_i
$x < \theta_1$	1
$\theta_1 < x < \theta_2$	0
$x > \theta_2$	2

In this case, an event is said to have occurred when the sensor gives output as 1, or 2, instead of 0, because 0 represents the reading of the range of interest. Thus, there can be nine possible scenarios, which are shown in the table below.

Table 2: Possible scenarios of sensors

M_i	R_i	Scenario
0	0	Sensor correctly reports a normal reading
0	1	Sensor faultily reports a normal reading
0	2	Sensor faultily reports a normal reading
1	0	Sensor faultily reports an unusual/event reading
1	1	Sensor correctly reports a unusual/event reading
1	2	Sensor faultily reports an unusual/event reading
2	0	Sensor faultily reports an unusual/event reading
2	1	Sensor faultily reports an unusual/event reading
2	2	Sensor correctly reports a unusual/event reading

E_i is the decoded value which is an estimate of R_i . Therefore, by implementing the fault recognition algorithm, E_i can be determined from the true readings R_i after obtaining information about the sensor readings of the neighboring sensors. The discussion consists of two cases, first is that the sensor fault probability is uncorrelated and symmetric, and second is that the sensor fault probability is uncorrelated and non-symmetric. The table 3 summarizes the various notations, which are used in the discussion and analysis.

Table 3: Summary of notations used for analysis and discussion of fault recognition

Symbol	Definitions
N	Total number of deployed nodes
n_o	Number of nodes in the normal region
n_f	Number of nodes in the event region
N	The number of neighbors of each node
M_i	The variable indicating the ground truth at node i
R_i	The variable with decoded value
P	The sensor fault probability
c_k	The probability that k out of N neighbors of node i are not faulty
$E_{reduced}$	The average number of errors reduced after decoding
$E_{corrected}$	The average number of errors corrected
$E_{uncorrected}$	The average number of errors uncorrected
$E_{newerrors}$	The average number of new errors introduced

2.1 DECISION SCHEME FOR FAULT RECOGNITION

The average number of errors after decoding is minimized by accepting a node's own sensor reading if and only if at least half of its neighbors have the same reading. There are three decision schemes available for fault recognition. They are randomized decision scheme, threshold decision scheme and optimal-threshold decision scheme. The effect of uncorrelated sensor faults is minimized by using optimal-threshold decision scheme, which seems to be an exceedingly practicable mechanism [20]. So, here, only this scheme is considered for the analysis of the non-binary event detection algorithm.

The optimal threshold decision scheme consists of three basic steps. In the first step, the sensor readings M_j of all N_i neighbors of node i is obtained. Then k_i is determined, which is the number of node i 's neighbors j with $M_j=M_i$. Finally it is checked if $k_{min} \geq 0.5N$. $E_i=M_i=a$ if and only if at least k_{min} of its N neighbors report the same sensor measurements a .

III. ANALYSIS OF THE FAULT-RECOGNITION ALGORITHM

To begin with the analysis of the Fault-Recognition mechanisms, certain assumptions are made. They are, if the node i is in event region, then all its neighbors are also in event region and, if node i is not in event region, neither of its neighbors are in the event region. This assumption is valid only inside the region, i.e. it does not apply to the nodes, which lie in the boundary of the event region.

3.1 ANALYSIS OF FAULT-RECOGNITION ALGORITHM WITH SYMMETRIC ERROR

Suppose there is totally 'n' number of sensor nodes deployed. For a sensor node i , there are N numbers of sensors in its neighboring region. These neighboring sensors may show one among the three decision i.e. 1 or 2 or 0. If the number 0 depicts the proper functioning of sensor node, then, the sensor is said to be faulty if decision 1 or 2 appears. Let the sensor fault probability be p . The sensor fault probability is actually the sum of probability that decision 1 appears instead of 2, and the probability that decision 2 appears instead of 1. In other words, the sensor fault probability can be expressed as,

$$p = p_e + p_t \quad (1)$$

Here, p_e is the probability that decision 1 appears and p_t is the probability that decision 2 appears. Let e is the number of nodes which show decision 1 and t is the number of nodes which show decision 2. Thus it can be also said that,

$$m = N - k = e + t \quad (2)$$

where m is the number of nodes, whose decision is different from the reference node. Also,

$k = r + s$, where r is the number of nodes which has sensor fault probability of p_e and s is the number of nodes which has sensor fault probability of p_t .

Let c_k be the probability that exactly k out of N nodes of node i 's neighbors are not faulty and it can be given as

$$c_k = \binom{N}{k} P(M_i = 0 | R_i = 0)^k P(M_i = 1 | R_i = 0)^{(N-k)} \quad (3)$$

Since the error is symmetric, p_e is assumed to be equal to p_t . Thus, equation (1) can be rewritten as,

$$\begin{aligned} p &= 2p_t & \because p_e &= p_t \\ p_t &= 0.5p \end{aligned} \quad (4)$$

To simplify the analysis, an assumption is made that the number of nodes showing decision 1 is same as the number of nodes showing decision 2. Thus, equation (2) can be rewritten as,

$$N - k = 2t \quad \because e = t \quad (5)$$

Implementing the relation (4) and (5) in (3), the relation (3) can be expressed as,

$$\begin{aligned} c_k &= \binom{N}{k} (1-p)^k \left(\frac{p}{2}\right)^{\frac{N-k}{2}} \left(\frac{p}{2}\right)^{\frac{N-k}{2}} \\ &= \binom{N}{k} (1-p)^k \left(\frac{p}{2}\right)^{N-k} \end{aligned} \quad (6)$$

The R_i , E_i , and M_i can be combined in sixteen possible combinations. Some metrics may be obtained to analyze the fault recognition algorithm by using the conditional probabilities of the sixteen possible combinations. The algorithm estimates if there is any event present or not. The probability that there is no event reading when the sensor is not faulty and absence of such event is indicated and is given as

$$P(E_i = a | M_i = a, R_i = a) = \sum_{k=0}^N c_k \quad (7)$$

The probability that the sensor faultily reports a correct reading is given by,

$$P(E_i = -a | M_i = -a, R_i = a) = \sum_{k=0}^N c_{N-k} \quad (8)$$

Four performance metrics can be obtained from the relation (7) and (8). The first metric is the average number of errors after decoding denoted as $E_{reduced}$. The reduction in average number of errors is obtained by normalizing the average number of error after decoding by the total number of nodes and the sensor fault probability.

$$E_{reduced} = P(E_i = 0 | R_i = 0)pn_o + P(E_i = 1 | R_i = 0)pn_{f_1} + P(E_i = 2 | R_i = 0)pn_{f_2}$$

Here n_{f_1} and n_{f_2} are the number of nodes showing faulty decision in the region. It is also written as

$$n_f = n_{f_1} + n_{f_2} . \text{ Thus, } E_{reduced} \text{ can be rewritten as,}$$

**B Victoria Jancee, S Radha and Nandita Das, ANALYSIS OF NON-BINARY FAULT TOLERANT
EVENT DETECTION IN WIRELESS SENSOR NETWORKS**

$$\begin{aligned}
 E_{reduced} &= \left(\sum_{k=0}^N c_{N-k} + 1 - \sum_{k=0}^N c_k + \sum_{k=0}^N c_{N-k} \right) pn_o + \left(\sum_{k=0}^N c_{N-k} + 1 - \sum_{k=0}^N c_k + \sum_{k=0}^N c_{N-k} \right) pn_f \\
 &= \left[1 - \sum_{k=0}^N c_k + \left(2 \sum_{k=0}^N c_{N-k} \right) \right] np
 \end{aligned} \tag{9}$$

Though the algorithm aims at doing fault tolerant event detection, there is some possibility for new errors getting introduced into the system because of the real nature of the sensor nodes lying at the boundary of the region of interest which is not known. It is not clear if they display the correct reading or event reading.

Thus another metric, which gives an estimate of the number of new errors introduced in to the system, is given as,

$$\begin{aligned}
 E_{newerrors} &= P(E_i = 1 | M_i = 0, R_i = 0) pn_o + P(E_i = 0 | M_i = 1, R_i = 1) pn_{f_1} \\
 &+ P(E_i = 0 | M_i = 2, R_i = 2) pn_{f_2} \\
 E_{newerrors} &= \left[1 - \sum_{k=0}^N c_k \right] pn_o + \left[1 - \sum_{k=0}^N c_k \right] pn_f \\
 &= \left[1 - \sum_{k=0}^N c_k \right] (n_o + n_f) p
 \end{aligned} \tag{10}$$

The average number of errors corrected by the algorithm can be obtained from the relation (8) and (10), which is given as

$$E_{corrected} = \left[1 - \sum_{k=0}^N 2c_{N-k} \right] np \tag{11}$$

When $E_{uncorrected}$ represents the average number of errors uncorrected and $E_{newerrors}$ represents the number of new errors introduced, then $E_{reduced}$ can be the sum of the two. In other words,

$$E_{reduced} = E_{uncorrected} + E_{newerrors}$$

Therefore, the average number of errors uncorrected can be given as,

$$E_{uncorrected} = E_{reduced} - E_{newerrors}$$

$$= \left[2 \sum_{k=0}^N C_{N-k} \right] np \quad (12)$$

The diagnosis of the sensor nodes is said to be correct if the good sensors are not diagnosed as faulty and the faulty sensors are not diagnosed as good. The diagnosis is said to be successful if all the sensors are identified either as good or faulty in a stipulated time. The diagnosis of the sensors are said to be incorrect either when a good sensor is tagged as faulty or a faulty sensor is tagged as good. The diagnosis is said to be unsuccessful when any node's status remains unknown. Therefore, a complete and proper diagnosis is required [21]. The probability that a good sensor is working properly and giving correct result is given as

$$P_{g1g} = (1 - p_e - p_t) \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{(N-i)!e!t!} \right) p_e^e p_t^t (1 - p_e - p_t)^{(N-i)} \quad (13)$$

Implementing relation (4) and (5) in (13), the above relation can be written as,

$$P_{g1g} = (1 - p) \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{(N-i)!e!t!} \right) \left(\frac{p}{2} \right)^e \left(\frac{p}{2} \right)^t (1 - p)^{(N-i)}$$

$$= (1 - p) \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{(N-i)!(e+t)!} \right) \left(\frac{p}{2} \right)^{(e+t)} (1 - p)^{(N-i)} \quad (14)$$

The probability that a good sensor is not working properly and giving incorrect result is given as

$$P_{g1f} = (1 - p_e - p_t) \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{i!r!s!} \right) p_e^r p_t^s (1 - p_e - p_t)^i \quad (15)$$

Implementing relation (4) and (5) in (15), the above relation can be written as

$$P_{g1f} = (1 - p) \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{i!r!s!} \right) \left(\frac{p}{2} \right)^r \left(\frac{p}{2} \right)^s (1 - p)^i$$

$$= (1 - p) \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{i!(r+s)!} \right) \left(\frac{p}{2} \right)^{r+s} (1 - p)^i \quad (16)$$

The probability that a faulty sensor has a tendency to show correct result is given as,

$$P_{f1g} = p \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{i!r!s!} \right) p_e^r p_t^s (1 - p_e - p_t)^i \quad (17)$$

Implementing relation (4) and (5) in (17), the above relation can be written as,

$$\begin{aligned} P_{f1g} &= p \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{i!r!s!} \right) \left(\frac{p}{2} \right)^r \left(\frac{p}{2} \right)^s (1 - p)^i \\ &= p \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{i!r!s!} \right) \left(\frac{p}{2} \right)^{r+s} (1 - p)^i \end{aligned} \quad (18)$$

The probability that a faulty sensor has a tendency to show incorrect result is given as,

$$P_{ff} = p \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{(N-i)!e!t!} \right) p_e^e p_t^t (1 - p_e - p_t)^{(N-i)} \quad (19)$$

Implementing relation (4) and (5) in (19), the above relation can be written as,

$$\begin{aligned} P_{ff} &= p \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{(N-i)!e!t!} \right) \left(\frac{p}{2} \right)^e \left(\frac{p}{2} \right)^t (1 - p)^{(N-i)} \\ &= p \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{(N-i)!(e+t)!} \right) \left(\frac{p}{2} \right)^{e+t} (1 - p)^{(N-i)} \end{aligned} \quad (20)$$

The probability of a faulty sensor being diagnosed as good is thus given as,

$$P_{FG} = p \sum_{a=\frac{N}{2}}^N \left\{ \binom{N}{a} p_{f1g}^a \sum_{b=0}^{\frac{a-N}{2}} \left[P_{g1g}^b \sum_{c=0}^{N-a-b} (P_{ff}^{N-a-b-c} P_{gff}^c) \right] \right\} \quad (21)$$

Similarly, the probability of a good sensor being diagnosed as faulty is given as,

$$P_{GG} = (1 - p) \sum_{a=0}^{\frac{N}{2}-1} \left\{ \binom{N}{a} p_{f1g}^a \sum_{b=0}^a \left[P_{g1g}^b \sum_{c=0}^{N-a-b} (P_{ff}^{N-a-b-c} P_{gff}^c) \right] \right\} \quad (22)$$

The probability that the faulty sensors are successfully diagnosed as faulty in the whole network is given as,

$$P_{NFG} = (1 - p * P_{FG})^n \quad (23)$$

The probability that the good sensors are diagnosed as faulty in the entire network is given as,

$$P_{NGG} = (p - (1-p) * P_{GG})^n \quad (24)$$

The relation (23) and (24) give the detection accuracy and the false alarm rate of the network respectively. Since the WSN is densely deployed, the detection accuracy is high and the false alarm rate is low.

3.2 ANALYSIS OF FAULT RECOGNITION ALGORITHM FOR NON-SYMMETRIC ERROR

In the non-symmetric error, as mentioned before, the p_e is not equal to p_t . Let N represent the number of cardinal neighbors of node i . Each neighboring sensor measurement can result in any of three possible outcomes. From Table 1, the three possible outcomes are normal value, intermediate value and high value. These possible outcomes are represented as E_1 , E_2 and E_3 . Suppose, further, that each possible outcome can occur with probabilities $(1-p)$, p_e and p_t . Then, the probability that E_1 occurs k times, E_2 occurs e times and E_3 occurs t times is

$$c_k = \left(\frac{N!}{k!e!t!} \right) (1-p)^k p_e^e p_t^t$$

$$c_k = C(1-p)^k p_e^e p_t^t \quad (25)$$

Here C is known as the multinomial co-efficient. The relation (9)-(12) can be used to obtain the first four performance metrics of the fault recognition for non-symmetric error. The diagnosis is said to be successful if all the sensors are identified either as good or faulty in a stipulated time, and the faulty sensors should not rise a false alarm in the network. The probability that a good sensor is working properly and giving correct result is given as,

$$P_{g1g} = (1 - p_e - p_t) \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{(N-i)!e!t!} \right) p_e^e p_t^t (1 - p_e - p_t)^{(N-i)} \quad (26)$$

The probability that a good sensor is not working properly and giving incorrect result is given as,

$$P_{g1f} = (1 - p_e - p_t) \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{i!r!s!} \right) p_e^r p_t^s (1 - p_e - p_t)^i \quad (27)$$

The probability that a faulty sensor has a tendency to show correct result is given as,

$$P_{f1g} = p \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{i!r!s!} \right) p_e^r p_t^s (1 - p_e - p_t)^i \quad (28)$$

The probability that a faulty sensor has a tendency to show incorrect result is given as,

$$P_{ff} = p \sum_{i=0}^{\frac{N}{2}-1} \left(\frac{N!}{(N-i)!e!t!} \right) p_e^e p_t^t (1 - p_e - p_t)^{(N-i)} \quad (29)$$

The probability of a faulty sensor being diagnosed as good is thus given as,

$$P_{FG} = p \sum_{a=\frac{N}{2}}^N \left\{ \binom{N}{a} p_{f1g}^a \sum_{b=0}^{\frac{a-N}{2}} \left[P_{g1g}^b \sum_{c=0}^{N-a-b} (P_{ff}^{N-a-b-c} P_{g1f}^c) \right] \right\} \quad (30)$$

Similarly, the probability of a good sensor being diagnosed as faulty is given as,

$$P_{GG} = (1 - p) \sum_{a=0}^{\frac{N}{2}-1} \left\{ \binom{N}{a} p_{f1g}^a \sum_{b=0}^a \left[P_{g1g}^b \sum_{c=0}^{N-a-b} (P_{ff}^{N-a-b-c} P_{g1f}^c) \right] \right\} \quad (31)$$

In this case, also, the relations (23) and (24) give the detection accuracy and the false alarm rate of the network respectively. Since the WSN is densely deployed, the detection accuracy is high and the false alarm rate is low.

IV. PERFORMANCE ANALYSIS AND RESULT

In order to analyze the performance of the fault recognition algorithm, consider a 32 X 32 region with 1024 sensor nodes deployed in a square grid of unit area as shown in the Figure 1. Each blue color dot denotes one sensor node.

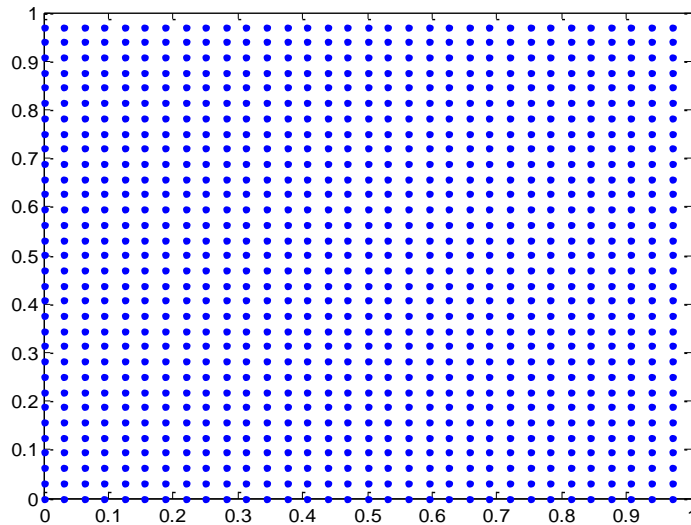


Figure 1. A 32X32 region with 1024 sensor nodes deployed in a square grid of unit area

The measurement parameter of the network is assumed to be x , which is the temperature to be monitored in the range of 75-80 degree Fahrenheit. All the sensors report '0' to indicate the normal range i.e. when x lies between 75-80 degree Fahrenheit, '1' to indicate the temperature below 75 degree Fahrenheit and '2' to indicate the temperature above 80 degree Fahrenheit. Sensor is said to be good if it displays the output '0', otherwise it is said to be faulty. Each node has an independent probability of error. In the algorithm two threshold values θ_1 and θ_2 are needed to map the sensors real output. Here θ_1 and θ_2 are set to 70 degree Fahrenheit and 85 degree Fahrenheit respectively. The average number of errors after decoding, the average number of sensor faults corrected, the average number of sensor faults uncorrected, the number of new errors introduced, detection accuracy and the false alarm rate are the six performance metrics used to evaluate the performance of the algorithm in test. In the theoretical analysis, the sensors are randomly chosen to be faulty with probabilities ranging from 0 to 0.50 with an interval of 0.05. The number of neighboring sensor chosen for the analysis is 4 i.e. only the neighbors, which are in direct connection with the reference sensors, are considered.

Case-I: Performance metrics for the fault recognition algorithm with symmetric error

The Figure 2 shows the first four parameters, i.e. the average number of errors after decoding, the average number of sensor faults corrected, the average number of sensor faults uncorrected, the number of new errors introduced, for symmetric errors.

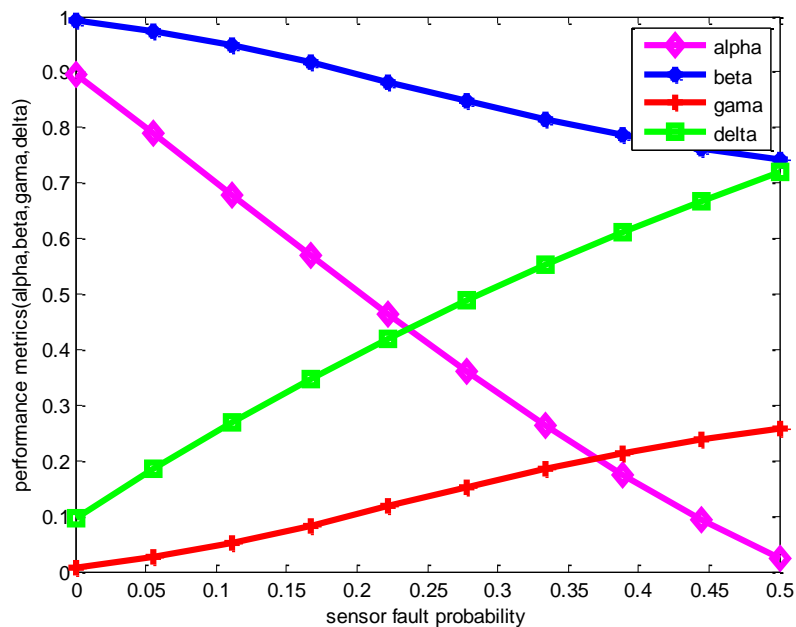


Figure 2. Performance metrics versus sensor fault probability for symmetric errors

Here, $\alpha = E_{reduced}$, $\beta = E_{corrected}$, $\gamma = E_{uncorrected}$ and $\delta = E_{newerrors}$

It is seen that for sensor fault probability as high as 10 percent, more than 95 percent of the errors are corrected. Since most of the errors are corrected, the number of uncorrected errors is very low, which is less than 5 percent for sensor fault probability less than 10 percent. The theoretical analysis ignores the edge and boundary effects. This affects the performance because at the edge of the deployed network the number of neighbors per node is lesser than that of interior and also the nodes at the edge are more likely to show erroneous decision if their neighbors provide conflicting information. Such nodes are the most vulnerable sites for new error introduced by the fault tolerance algorithm. Because of this the the new errors introduced is seen to increase steadily. The Figure 3 shows the detection accuracy of the faulty sensors against the sensor fault probability. The detection accuracy for four neighbors decreases when the fault probability increases. .

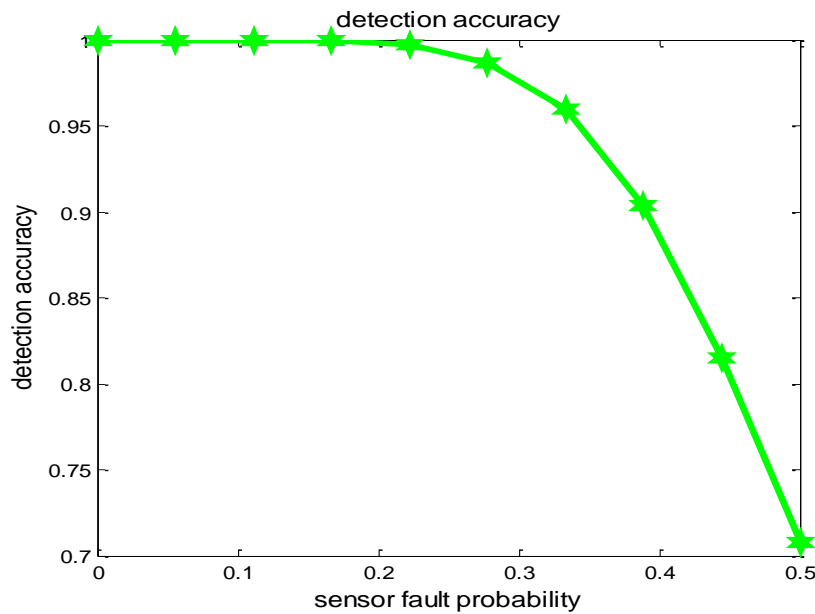


Figure 3. Faulty sensor detection accuracy versus sensor fault probability for four neighbors

Figure 4 shows the false alarm rate of the sensors versus the sensor fault probability. Higher the sensor fault probability, the higher will be the false alarm rate. This is because the faulty sensors test the good sensors and report them to be faulty, thus these good sensors are finally reported as faulty sensors. For 4 neighboring sensors and for sensor fault probability below 0.25 , the false alarm rate is as low as 0.

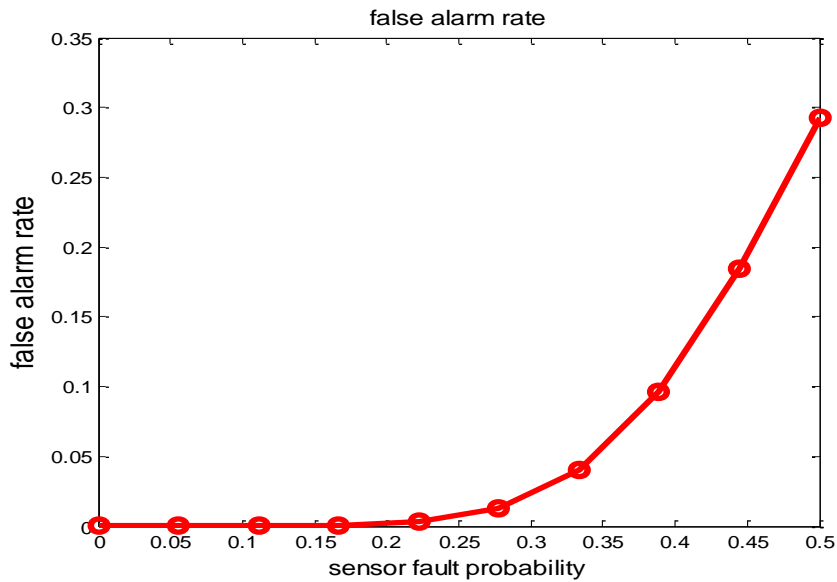


Figure 4. Fault alarm rate versus sensor fault probability for four neighbors

The overall performance of this algorithm for the symmetric errors is good. It has good detection accuracy and a low false alarm rate.

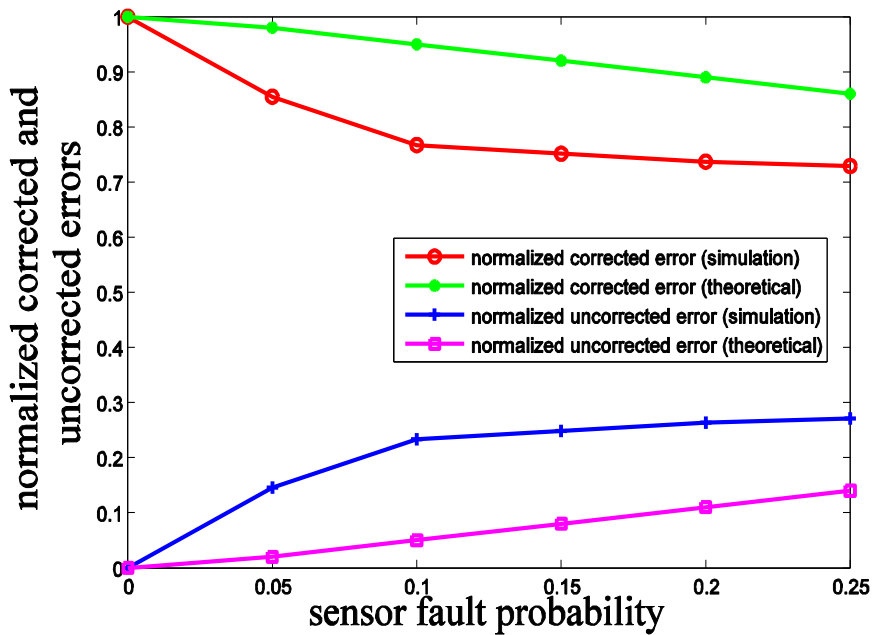


Figure 5. Normalized corrected and uncorrected errors

Figure 5 shows the graph for the fault recognition algorithm with the non – binary symmetric error from both the simulation as well as the theoretical equations. The main conclusion from these plots is that the simulation results are to some extent match with the theoretical predictions. The simulation results indicate that 75 percentage of the symmetric errors can be corrected even when the sensor fault rate is as high as 10 percentage. In addition, the theoretical results show that more than 95 percentage of the non - symmetric errors can be corrected for sensor fault probability more than 10 percent.

Case-III: Performance metrics for the fault recognition algorithm with non-symmetric error

To analyze the fault recognition algorithm with non-symmetric errors, the relation (9)-(12), (24) and (25) are used. Seeing the relations, it can be observed that, to obtain c_k the value of k should be varied from half of N to N . Thus for each k there are few combinations of e and t , in a way that,

$$N = k + e + t \tag{32}$$

Similarly, to obtain g_{N-k} the value of k is again varied from half of N to N , which in turn gives certain combinations of r and s such that,

$$k = r + s \tag{33}$$

The experiment is conducted for different combinations of r , s , e , and t for each k , then the combination, which yields the best result, is considered and the performance metrics has been obtained for it. The various values of r , s , e , and t are given in the Table 4.

Table 4: Combinations of e and t for various k

Number of neighboring nodes whose reading is same as the reference node i, k	Number of nodes which show decision 1, e	Number of nodes which show decision 2, t
2	1	1
	2	0
	0	2
3	0	1
	1	0
4	0	0

Table 5: Combinations of r and s for various k

Number of neighboring nodes whose reading is same as the reference node i, k	Number of nodes which have sensor fault probability of $p_{e, r}$	Number of nodes which have sensor fault probability of $p_{t, s}$
2	1	1
	2	0
	0	2
3	2	1
	1	2
	3	0
	0	3
4	0	4
	4	0
	3	1
	1	3
	2	2

The Figure 6 shows the performance metrics plotted against the sensor fault probability with $k = 4$. It is seen that almost all the errors are corrected when the sensor fault probability is as high as 50 percent. It can thus be said that the algorithm can correct maximum number of errors if more than half of the neighboring sensor nodes report the same reading as that of the reference sensor node. Since almost all the errors are corrected, the number of uncorrected errors is very negligible. Since there are three probable outputs for sensor, the probability for the boundary nodes to show conflicting results is higher. Hence, the number of new errors introduced increases with the sensor fault probability. The reduction in average number of errors is seen to decrease with increasing sensor fault probability.

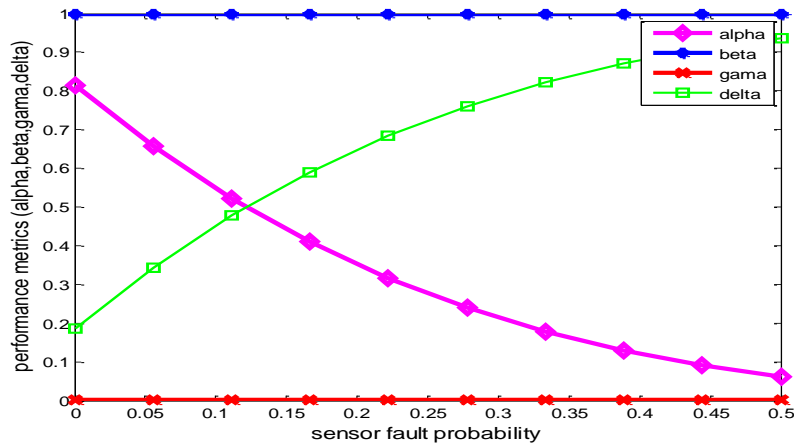


Figure 6. Performance metrics versus sensor fault probability with $k = 4$

Here, $\alpha = E_{reduced}$, $\beta = E_{corrected}$, $\gamma = E_{uncorrected}$ and $\delta = E_{newerrors}$

The Figure 7 and Figure 8 show the detection accuracy and false alarm rate respectively. It is seen that the detection accuracy is very high, i.e. 1 when 20 percent of the nodes are faulty. Beyond this, the detection accuracy decreases with the increase in the sensor faults. Similarly, the false alarm rate is as low as 0 for sensor fault probability as high as 20 percent, and it increases when most of the sensors become faulty.

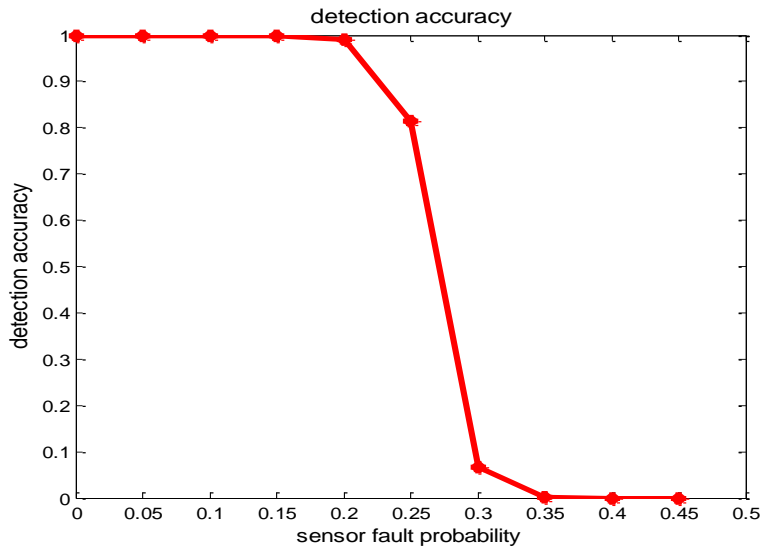


Figure 7. Faulty sensor detection accuracy versus sensor fault probability

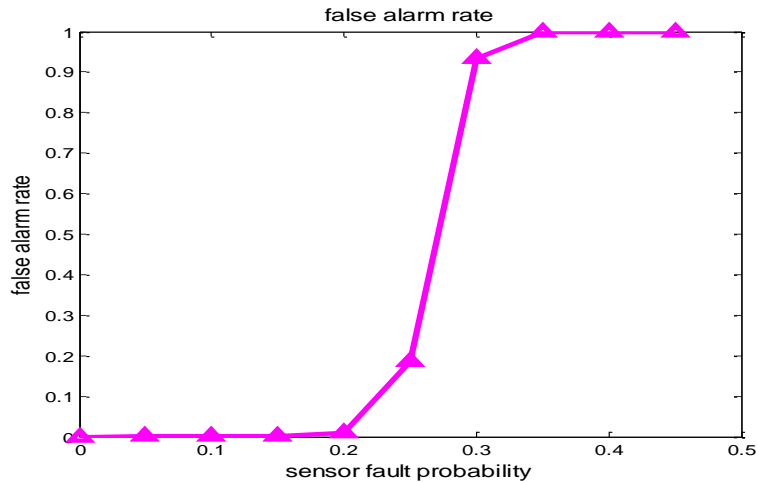


Figure 8. False alarm rate versus sensor fault probability

V. CONCLUSION AND FUTURE WORK

The algorithm discussed here is distributed and localized. Each sensor identifies its real output to be acceptable or not, by obtaining the information from the neighboring sensor and then making a comparison. The fault recognition algorithm is analyzed with the non – binary symmetric error from both the simulation as well as the theoretical equations. The simulation results match with the theoretical predictions to some extent. The simulation results indicate that 75 percentage of the symmetric errors can be corrected even when the sensor fault rate is as high as 10 percentage. Whereas, the theoretical results show that more than 95 percentage of the non - symmetric errors can be corrected for sensor fault probability more than 10 percent. Further, the algorithm is analyzed by using a probabilistic approach for both the symmetric errors and the non-symmetric errors. The analytical result shows that as much as 95 percent errors are corrected for fault rate as high as 10 percent with symmetric errors and for the same sensor fault probability, with non-symmetric errors almost 92 percentage errors are corrected. The results also show that with symmetric error, almost 98 percentage of detection accuracy can be obtained even when 25 percent of the nodes are faulty and false alarm rate is very low too. At the same time, for the non-symmetric error, 100 percent detection accuracy and zero false alarm rate is seen with up to 20 percent sensor fault probability. The analysis thus proves that the algorithm is quiet efficient. In addition here the performance of the algorithm is seen to be affected by the introduction of new errors due to conflicting decisions from the neighboring boundary nodes. Thus, this work may be extended towards finding a solution to reduce the number of new errors introduced due to the boundary nodes. Much of the work presented here can be implemented on event region detection in 3D environment monitoring using real sensor network hardware.

REFERENCES

- [1] J. Warrior, “Smart Sensor Networks of the future,” Sensors Magazine, Mar 2007.
- [2] G.J Pottie and W.J Kaiser, “Wireless Integrated Networks Sensors,” Communications of the ACM, Vol.43, No.5, pp.51-58, May 2000.
- [3] A.Cerpa et al., “Habitat Monitoring: Application Driver for Wireless Communications Technology,” Proc.2001, ACM SIGCOMM Workshop Data Communication in Latin America and the Caribbean, April 2001.
- [4] F. Kounshanfar, M.Potkonjak and A. Sanjiovanni-Vincentelli, “Fault Tolerance Techniques for Wireless Ad Hoc Sensor Networks”, Proc. IEEE Sensors, pp. 1491-1496, 2002.

**B Victoria Jancee, S Radha and Nandita Das, ANALYSIS OF NON-BINARY FAULT TOLERANT
EVENT DETECTION IN WIRELESS SENSOR NETWORKS**

- [5] N.Samanta, A.K.Chanda, C.RoyChaudhuri, “An Energy Efficient, Minimally Intrusive Multi-Sensor Intelligent System For Health Monitoring Of Elderly People”, *International Journal On Smart Sensing And Intelligent Systems* Vol. 7, No. 2, June 2014, pp. 762 -780.
- [6] N.K.Suryadevara, A. Gaddam, R.K.Rayudu and S.C. Mukhopadhyay, “Wireless Sensors Network based safe Home to care Elderly People: Behaviour Detection”, *Sens. Actuators A: Phys.* (2012), doi:10.1016/j.sna.2012.03.020, Volume 186, 2012, pp. 277 – 283.
- [7] Sachin Bhardwaj, Dae-Seok Lee, S. C. Mukhopadhyay and Wan-Young Chung, “Ubiquitous Healthcare Data Analysis and Monitoring Using Multiple Wireless Sensors for Elderly Person”, *Sensors & Transducers Journal*, Vol. 90, Special Issue, April 2008, pp. 87-99.
- [8] Cheng Chunling, Wu Hao, Yu Zhihu, Zhang Dengyin, Xu Xiaolong, “Outlier Detection Based On Similar Flocking Model In Wireless Sensor Networks”, *International Journal On Smart Sensing And Intelligent Systems*, Vol. 6, No. 1, February 2013 pp. 18-37.
- [9] Dyi-Rong Duh, Ssu-Pei Li, and VictorW. Cheng, “Distributed Fault-Tolerant Event Region Detection of Wireless Sensor Networks”, *International Journal of Distributed Sensor Networks*, Volume 2013, Article ID 160523, 8 pages.
- [10] Xuefeng Liu, Jiannong Cao and Shaojie Tang, “Fault Tolerant Complex Event Detection in WSNs: A Case Study in Structural Health Monitoring”, *INFOCOM, 2013 Proceedings IEEE* , 14-19 April 2013, pp. 1384-1392.
- [11] E.I Moutapha Ould-Ahmed-Vall, B.H.Ferri, and G.F.Riley, “Distributed Fault-Tolerance for Event Detection Using Heterogeneous Wireless Sensor Networks”, *IEEE Transaction on Mobile Computing*, Vol.11, No.12, December 2012, pp. 1994-2007.
- [12] Sung-Jib Yim and Yoon-Hwa Choi, “An Adaptive Fault-Tolerant Event Detection Scheme for Wireless Sensor Networks”, *Sensors* 2010, 10, pp. 2332-2347.
- [13] Myeong.-Hyeon Lee and Yoon Hwa Choi, “Fault detection of wireless sensor networks,” *Computer Communications*, vol. 31, no. 14, 2008, pp. 3469–3475.
- [14] Tsang-Yi Wang, Li-Yuan Chang, Dyi-Rong Duh, and Jeng-Yang Wu, “Fault-Tolerant Decision Fusion via Collaborative Sensor Fault Detection in Wireless Sensor Networks”, *IEEE Transactions on Wireless Communications*, Vol. 7, no.2, pp. 756-768, February 2008.

- [15] Hong Li, Zheng Xie, Min Wu, Ping Li, "Event Region Fault-Tolerant Detection Based on Distributed Weight for Wireless Sensor Networks", 4th International Conference on Wireless Communications, Networking and Mobile computing, 12-14 October 2008, pp. 1-4.
- [16] Tsang-Yi Wang, and Qi Cheng, "Collaborative Event-Region and Boundary-Region Detections in Wireless Sensor Networks," IEEE Transactions On Signal Processing, Vol. 56, No. 6, pp. 2547-2567, June 2008.
- [17] M. Li, Y. Liu, and L. Chen, "Non-Threshold Based Event Detection for 3D Environment Monitoring in Sensor Networks," Proc. 27th International Conference on Distributed Computing Systems (ICDCS), 2007.
- [18] J.Chen, S.Kher and A.Somani, "Distributed Fault Detection of Wireless Sensor Networks" DIWANS'06 Proceedings of the 2006 workshop on Dependability issues in wireless ad hoc networks and sensor networks, Pages 65-72, ACM New York, NY, USA, 2006.
- [19] X. Luo, M. Dong, and Y. Huang, "On Distributed Fault-Tolerant Detection in Wireless Sensor Networks," IEEE Transactions on Computers, vol. 55, no. 1, pp. 57-80, Jan. 2006.
- [20] B. Krishnamachari and S.S Iyengar, "Distributed Bayesian Algorithms for Fault-Tolerant Event Region Detection in Wireless Sensor Networks," IEEE Transactions on Computers, vol. 53, no. 3, pp. 241-250, March 2004.
- [21] Arun K., Somani, and V.K.Agarwal, "Distributed Diagnosis Algorithms for Regular Interconnected Structures", IEEE Transactions on Computers, vol. 41, no. 7, pp. 899-906, July 1992.