

Fast and Simple On-Line Sensor Fault Detection Scheme for Wireless Sensor Networks

Jeng-Yang Wu¹, Dyi-Rong Duh^{1,*}, Tsang-Yi Wang², and Li-Yuan Chang³

¹ Department of Computer Science and Information Engineering
National Chi Nan University, Puli, Nantou Hsieh 54561, Taiwan
{s94321522, drduh}@ncnu.edu.tw

² Institute of Communications Engineering
National Sun Yat-Sen University, Kaohsiung 80424, Taiwan
tcwang@mail.nsysu.edu.tw

³ Department of Computer Science and Information Engineering
National Cheng Kung University, Tainan 70101, Taiwan
lance@csie.ncku.edu.tw

Abstract. Wireless sensor networks (WSN) are composed of a large number of sensor nodes and usually used to monitor a region of interest. The sensor nodes are very prone to damage due to low-cost design and random deployment. Additionally, faulty nodes may degrade the performance of the distributed hypothesis testing. This work addresses fault isolation in WSN where the fusion center attempts to identify faulty nodes through temporal sequences of received local decisions. Owing to the processor, memory, and power constraints in embedded systems, the employed method should be as simple as possible. Therefore, the primary goal of this investigation is to design a low-complexity sensor fault detection scheme, which can detect most sensor faults by using the majority voting technique. The simulation results show the proposed approach is effective in terms of identifying faulty members in a distributed sensor network.

Keywords: Wireless sensor networks, fusion technique, sensor fault detection, distributed detection; fault-tolerant detection; false testing.

1 Introduction

The problem of distributed decision fusion in wireless sensor networks (WSN) has received much attention because of many important applications [1, 4, 7, 8, 9, 10, 11]. Sensor nodes in WSN are deployed in the region of interest for collecting data. These sensor nodes, which consist of sensing, data processing, communicating, and power components, observe the phenomenon at each measured time step. After processing the observation, each node transmits individual local decision to a fusion center. The fusion center then makes a final decision based on these preliminary local decisions.

* Corresponding author.

WSN usually consist of a large number of sensor nodes, which are deployed in inaccessible and harsh environments. Furthermore, the sensor nodes are prone to damage as a result of low-cost design and random deployment. Additionally, placing sensor nodes in inaccessible areas makes them irreplaceable. Therefore, the design of distributed detection in WSN needs to be fault tolerant. The types of sensor faults in WSN may range from simple stuck-at faults to random sensor faults, which render prior failure probability models unsuitable for the design of distributed detection in WSN. For this reason, the primary goal of this study is to design an effective fault detection scheme, which can tolerate most sensor faults.

The fusion center may make a wrong decision when the combined effect of the number of faulty nodes and sensor fault types is high. To provide fault-tolerance capability in distributed detection, the detection system can remove the unreliable local decisions transmitted from faulty sensor nodes during the process of final decision making. This work considers the fault detection based on a collaboratively sequential detection scheme. The problem formulation in this study is the fusion center needs to identify faulty nodes at every time step. In the decision fusion process, the data sent from faulty nodes will be discarded for making more dependable final decisions. The considered scenario has many applications such as health monitoring and security surveillance. A deployed sensor network in each of these applications may have to report its decision at every measured time step; for this reason, an appropriate strategy can be immediately selected when an unexpected event occurs.

Some related investigations have addressed several variants of fault detection problems. Fault detection problems by central testing can be found in [2, 3]. The distributed fault detection problem for general nonlinear, non-Gaussian stochastic systems with multiple sensor nodes has been addressed [6]. The work in [9] applies the non-parametric statistics-based technique for identifying the faulty sensor nodes in a sensor network. For information assurance of the data fusion in WSN, a witness based approach has been demonstrated to verify the fusion data [8]. An improved witness-based approach using a direct voting mechanism has also been proposed to verify the fusion data without any cryptographic operation [10].

This work also considers the problem of sensor fault detection as follows. Assume that all sensors will have the same readings and make the same decision if they are fault-free and deployed in an area. The fusion center can identify a faulty node by judging whether its behavior is very different from the others since each node sends its local decision to the fusion center at every time step. Therefore, a sensor fault detection scheme with a record table, which records the history of all local decisions during the monitor process, is proposed. Because of the processor, memory, and power constraints in embedded systems, the employed approach should be as simple as possible. For this reason, the proposed scheme just applies the majority voting technique to differentiate between normal nodes and faulty nodes. Since the employed method is quite simple, applying the proposed sensor fault detection scheme in real applications is feasible.

The remainder of this investigation is organized as follows. Section 2 formally presents the system model and the problem formulation. The details of the proposed sensor fault detection scheme are described in Section 3. Section 4 shows the performance evaluation of the proposed approach by simulation. Finally, conclusions are drawn in Section 5.

2 System Model and Problem Formulation

This section first describes the system model and the problem formulation. In practice, the types of sensor faults in a sensor network are actually very diverse. Three considered sensor fault types in this study are presented finally.

2.1 Network Operation

A two-layer detection system is considered in this work, as illustrated in Fig. 1. The parallel fusion system, which consists of N identical sensor nodes and a fusion center, is used to determine whether an unknown binary hypothesis is H_0 or H_1 . The prior probabilities of H_0 and H_1 are assumed to be known. Each member of N sensor nodes is denoted by s_i , where $i = 1, \dots, N$. Let x_i^t denote the observation of the i th sensor node and u_i^t denote the binary decision of the i th node, where $i = 1, \dots, N$ and t represents the time index. The observations across sensor nodes are independent and identically distributed condition on phenomenon.

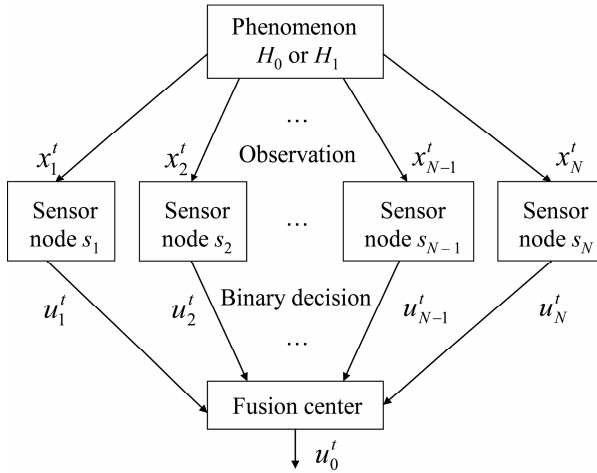


Fig. 1. System model of a parallel fusion network

Assume that an identical local decision rule is employed at each sensor node. Each node independently makes a binary decision based only on its observation. The local decision u_i^t of node s_i is obtained through the local decision rule γ as (1). Each sensor node reports its local decision to the fusion center at each time step. A decision ‘0’ is sent if the sensor node makes a decision in favor of H_0 ; otherwise, a decision ‘1’ is transmitted.

$$u_i^t = \gamma(x_i^t). \tag{1}$$

A parallel fusion network employing identical local decision rules at each node is asymptotically optimal based on error exponents when the number of nodes becomes very large. The error exponents of the identical local decision rules are equal to that of the non-identical local decision rules obtained by system-wide optimization [5]. For this reason, employing the identical local decision rule at each node is quite suitable for large-scale sensor networks.

Let's consider the fusion center is processing its information at time step t . All preliminary decisions up to time step t from all nodes are available at the fusion center. The fusion center begins to identify faulty members by utilizing the proposed sensor fault detection scheme. In the decision fusion process, the fusion center discards the data of faulty nodes for making a more believable final decision.

2.2 Sensor Fault Types

A sensor network is very likely to contain faulty nodes, because sensor nodes are usually low-cost and deployed randomly. Additionally, the sensor faults may include hardware or software damage resulting in all misbehavior; hence, the types of sensor faults are diverse.

Three types of sensor faults are considered in this work. In one fault type, a faulty sensor node is frozen to transmit a fixed local decision '0' to the fusion center regardless of the real observation. This type of sensor fault is named stuck-at-zero fault. Similarly, a fault type is called stuck-at-one fault when a faulty node always transmits a fixed decision '1'. The last sensor fault type is that a faulty sensor node reports decision '0' or '1' randomly regardless of the present hypothesis and called random fault. The fusion center does not know the sensor fault types in advance. The detection system identifies faulty nodes just according to the behavior of each node.

3 Sensor Fault Detection Scheme

The record table, which records the history of all local decisions during the monitor process, is introduced in this section. The record table will faithfully present the behavior of each node in a network. In fact, the record table plays a key role in the proposed fault detection scheme. The details of the proposed approach are described subsequently.

3.1 Record Table

This investigation considers that sensor nodes sequentially transmit local decisions to the fusion center. A sensor node can be reasonably assumed to be faulty when its behavior is very different from the majority of nodes. Recording the history of local decisions transmitted from all nodes is a method to represent the behavior of sensor nodes, since the fusion center receives each node's decision at every time step. Therefore, a record table is designed to record the behavior of each node at each time step. Let R_i^t denote the ratio of decisions '1' to all decisions which have been transmitted to the fusion center by the i th node at time step t as

$$R_i^t = \frac{1}{t} \sum_{j=1}^t u_i^j \tag{2}$$

As shown in (2), each R_i^t is independent of all other nodes and realistically shows the condition of local decisions transmitted by node s_i . For example, if a sensor node has ever sent three decisions ‘1’ to the fusion center at the first five time steps, its rate of decision ‘1’ is recorded as 3/5. For instance, a node has never transmitted decision ‘1’ to the fusion center at the first seven time steps and then its rate of decision ‘1’ is recorded as 0/7. The value of R_i^t is just between 0 and 1 obviously.

3.2 Proposed Scheme

The fusion center can identify faulty nodes through comparing each node’s behavior. The rates of decision ‘1’ of normal sensor nodes are similar since they have the same density function. A sensor node has the highest probability to be faulty when its rate is very different from the other nodes. For making a distinction between normal nodes and faulty nodes, the proposed scheme divides the entire rate value into q equal regions. Let p_i denote the range of each rate region as (3).

$$\text{range of } p_i = \left[\frac{i-1}{q}, \frac{i}{q} \right), \tag{3}$$

where $i = 1, \dots, q$. Exceptionally, the 1.0 rate value is included in the last rate region.

The R_i^t of each node can be corresponded to a rate region. At each time step, the number of nodes in each rate region is initialized to 0 first. After updating the record table, the R_i^t of each node is placed to the corresponding rate region. A rate region owning the maximum number of nodes can be easily found. The fusion center then identifies faulty nodes by using the majority voting technique. However, all normal nodes do not always exactly locate in the same rate region. Several normal nodes possibly locate in the neighbor regions. For this reason, the proposed scheme marks every three continuous rate regions to form a rate group except the first group and the last one. Let g_i denote the range of each rate group as (4).

$$\begin{aligned} g_1 &= p_1 + p_2, \\ g_i &= p_{i-1} + p_i + p_{i+1}, \\ g_q &= p_{q-1} + p_q, \end{aligned} \tag{4}$$

where $i = 2, \dots, q-1$. Rate groups are presently used to replace rate regions for lowering the probability of erroneous judgment. Similarly, the fusion center scans all rate groups for discovering a group which possesses the maximum count of nodes at each time step. A node is determined to be faulty by the fusion center if its rate does not locate in the group having the maximum number of sensor nodes.

The counts of nodes in different groups are sometimes equal. If two groups, g_a and g_b , have the same number of nodes, the fusion center will select g_a when rate region p_a has a larger count of nodes; if p_a and p_b also own the identical number of nodes, the fusion center will randomly select one rate group. Finally, the flow chart of the proposed sensor fault detection scheme at time step t is illustrated in Fig. 2.

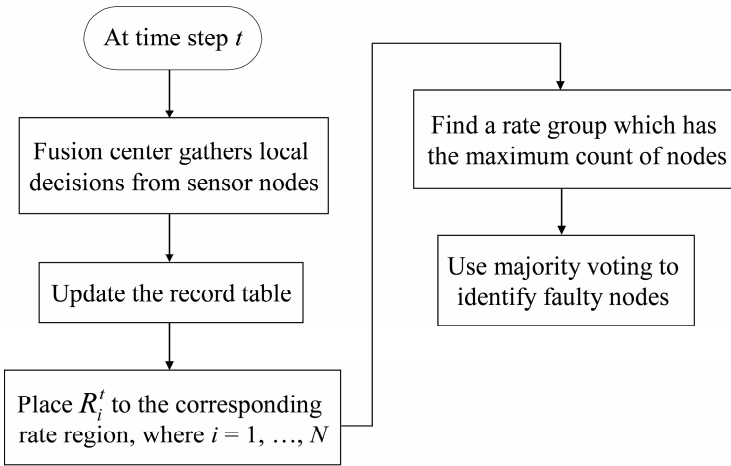


Fig. 2. Flow chart of the proposed scheme

4 Simulation Results

The error rate of fault detection in this study is described first in this section. The performance of the proposed sensor fault detection scheme is then evaluated. The sensor fault types and the actual number of faulty nodes are unknown in advance in these simulations.

4.1 Error Rate of Fault Detection

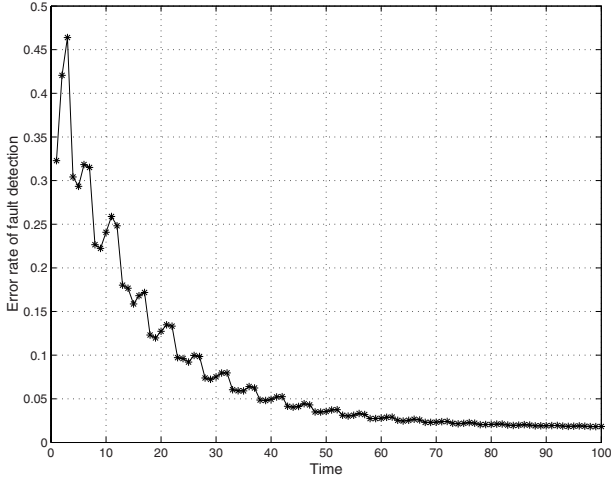
This investigation decides the error rate of fault detection through comparing the results detected by the proposed scheme with the real conditions. For instance, the proposed approach identifies two faulty sensor nodes at time step t , but all nodes are actually normal. The error rate at this time step is indicated as $2/N$. For example, there are three faulty nodes in fact, but the proposed scheme only detects two of them. The error rate in this condition is then indicated as $1/N$. Restated, the error rate of fault detection is the rate of difference between the detected result and the reality.

4.2 Simulation Setup

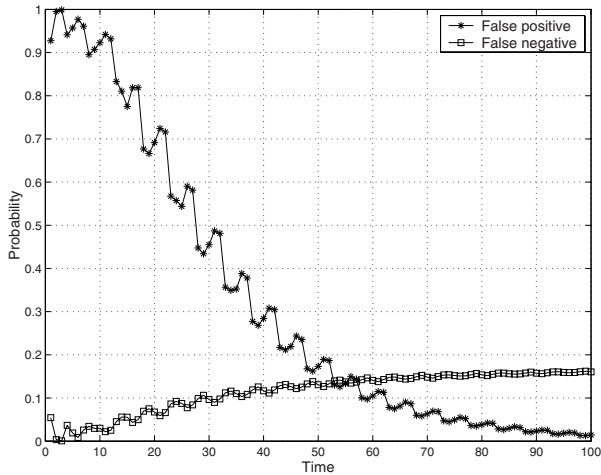
The detection of known signals in Gaussian noise is considered. This study lets the conditional densities at the sensor nodes be Gaussian with unit variance. Under H_0 and H_1 , the mean at all the sensor nodes is assumed to be zero and m respectively. Accordingly, the signal-to-noise ratio (SNR) can be defined as $20 \log_{10} m$. Additionally, each sensor node makes the local decision based on the bisection threshold $m/2$. The number of deployed sensor nodes N is set to 10 in all simulated conditions. In the following simulations, the number of rate partitions q is set to 10. Each simulated scenario is iterated 10,000,000 times to obtain the simulated performance. The true hypothesis and faulty sensor nodes are randomly chosen at the beginning of every iteration step.

4.3 Results and Analysis

In practice, a sensor network probably contains several types of faults at the same time and the fault types are unknown in advance. For convenience, a scenario, which simultaneously contains three kinds of fault sensor types including stuck-at-zero fault, stuck-at-one fault, and random fault, is investigated. The fault type of individual faulty node is randomly decided at the beginning of every iteration step.

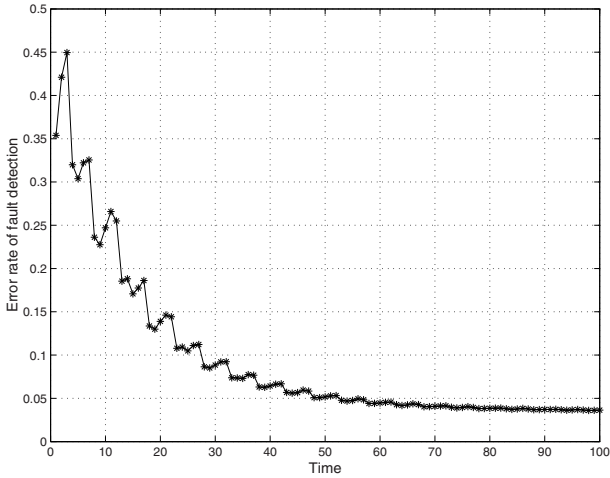


(a)

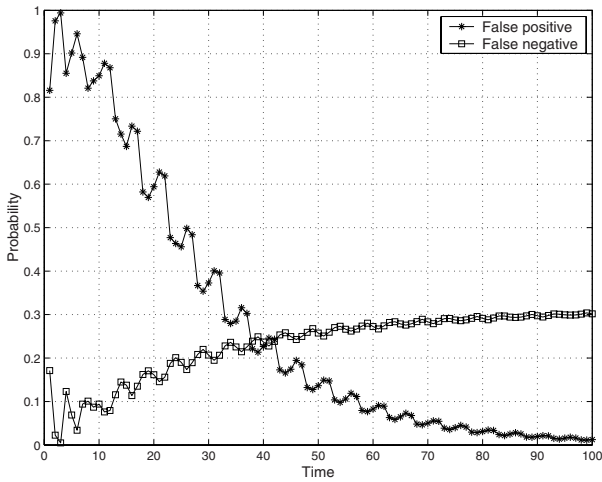


(b)

Fig. 3. (a) Error rate of fault detection and (b) probabilities of false positive and false negative in a network with one faulty node at 0 dB SNR



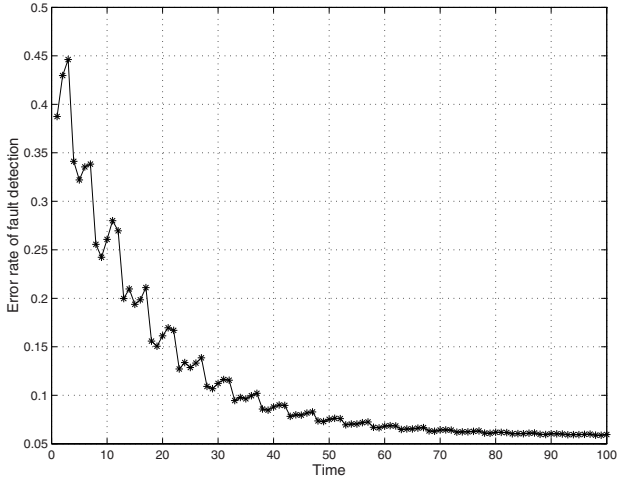
(a)



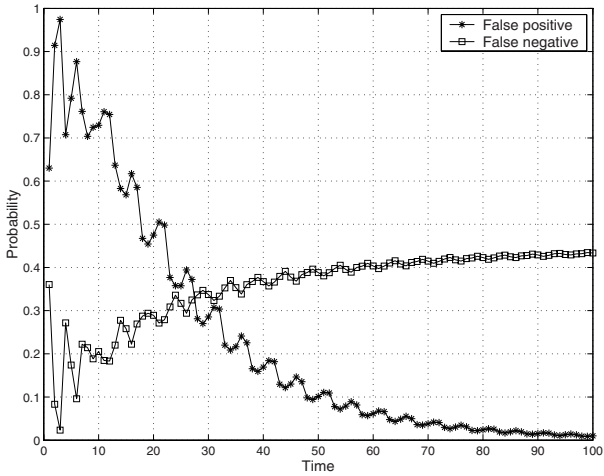
(b)

Fig. 4. (a) Error rate of fault detection and (b) probabilities of false positive and false negative in a network with two faulty nodes at 0 dB SNR

The performance of fault detection at 0 dB SNR in a situation with one faulty node is shown in Fig. 3(a). The error rate of fault detection rises in the first three time steps. The reason is no rate group can be formed in these three time steps because every possible value of rate is distinct from each other. Therefore, the fusion center identifies faulty nodes through comparing the number of nodes in each rate region. After time step three, several rate groups are formed and the probability of erroneous judgment is lowered.



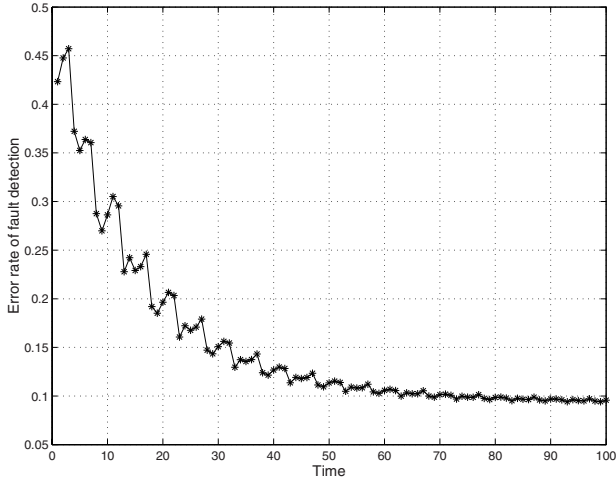
(a)



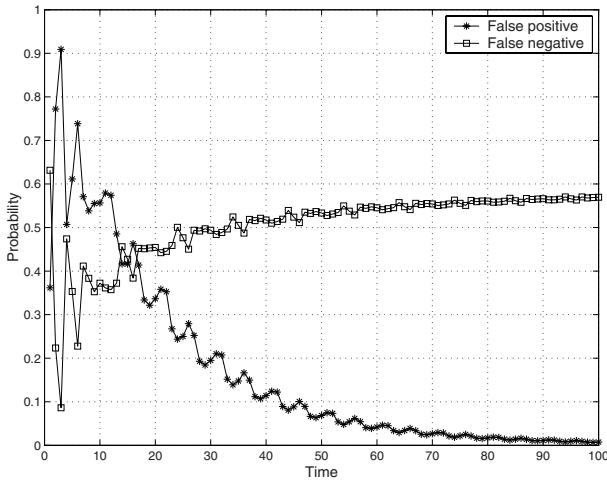
(b)

Fig. 5. (a) Error rate of fault detection and (b) probabilities of false positive and false negative in a network with three faulty nodes at 0 dB SNR

Fig. 3(b) shows the probabilities of reported false positive (fault-free node was indicated as faulty) and false negative (a node with fault was not detected) when one faulty node occur in a network. In the first several time steps, the rates of sensor nodes are unstable. Therefore, the corresponding rate regions of normal nodes are mutable. Several normal nodes are probably judged to be faulty at the moment. When time increases, the behavior of a normal node is gradually similar to other normal ones. Most normal sensor nodes will locate in a rate group which possesses the



(a)



(b)

Fig. 6. (a) Error rate of fault detection and (b) probabilities of false positive and false negative in a network with four faulty nodes at 0 dB SNR

maximum count of nodes. Therefore, the probability of false positive will gradually lower. On the other hand, the probability of false negative rises. The reason is that the behavior of random fault is not significantly different from that of a normal node in fact. When the rates of all nodes are more and more stable, random faults might be contained by the rate group having maximum number of nodes. Consequently, the probability of false negative rises owing to the erroneous judgment.

Figs. 4, 5 and 6 present the same evaluation information for the cases where two, three and four deployed sensor nodes were faulty, respectively. The descending error rate of fault detection shows the major part of faulty nodes can be identified by the fusion center apparently. For this reason, the proposed approach has the capability to assist the distributed detection system in making more dependable decisions by isolating most sensor faults. The reason for the trends of false positive and false negative in Fig. 6 is similar to the aforementioned statement. The rates of normal nodes become more and more stable and similar when time increases. The probability of false positive will gradually decline. Contrarily, the probability of failing in detecting random faults increases when the number of faulty nodes is large.

5 Conclusions

Faulty sensor nodes in WSN always report unreliable information. For this reason, the fusion center may make wrong decisions according to inaccurate local decisions. This study investigates the problem of detecting faulty sensor nodes in distributed sensor networks at every time step.

A sensor node whose behavior is very unusual may be faulty. In order to show the behavior of each node in a network, this work designs a record table for recording the history of all local decisions during the monitor process. Additionally, designing a simple fault detection scheme for WSN is necessary owing to the computing capability constraint. For this reason, this investigation applies the majority voting technique to identify faulty members. The simulation results show that the proposed scheme with a record table is effective in terms of fault detection. Most importantly, the proposed sensor fault detection scheme does not need complex operations. Therefore, the precious energy resource in WSN could be conserved. The number of faulty nodes in this investigation is fixed during the whole monitor process. This work will be continuously improved for dealing with the increasable number of faulty nodes in a network as the further work.

References

1. Aldosari, S.A., Moura, J.M.F.: Detection in decentralized Sensor Networks. In: Proc. of the 2004 IEEE Int'l. Conf. on Acoustics, Speech, and Signal Processing, vol. 2, pp. 277–280 (2004)
2. Basseville, M.: Detecting Changes in Signals and Systems-A Survey. *Automatica* 24(3), 309–326 (1988)
3. Basseville, M., Nikiforov, I.: *Detection of Abrupt Changes-Theory and Applications*. Prentice-Hall, Englewood Cliffs, NJ (1993)
4. Chamberland, J.F., Veeravalli, V.V.P.: Asymptotic Results for Decentralized Detection in Power Constrained Wireless Sensor Networks. *IEEE Journal of Selected Areas in Communications* 22(6), 1007–1015 (2004)
5. Chen, P.N., Papamarcou, A.: New Asymptotic Results in Parallel Distributed Detection. *IEEE Transactions on Information Theory* 39(6), 1847–1863 (1993)

6. Cheng, Q., Varshney, P.K., Michels, J., Belcastro, C.M.: Distributed Fault Detection via Particle Filtering and Decision Fusion. In: Proc. of the 8th Int'l. Conf. on Information Fusion, vol. 2, pp. 1239–1246 (2005)
7. D'Costa, A., Sayeed, A.M.: Data versus Decision Fusion for Distributed Classification in Sensor Networks. In: Proc. of the 2003 IEEE Int'l. Conf. on Acoustics, Speech, and Signal Processing, vol. 1, pp. 585–590 (2003)
8. Du, W., Deng, J., Han, Y.S., Varshney, P.K.: A Witness-Based Approach for Data Fusion Assurance in Wireless Sensor Networks. In: Proc. of GLOBECOM 2003, vol. 3, pp. 1435–1439 (2003)
9. Koushanfar, F., Potkonjak, M., Sangiovanni-Vincentelli, A.: On-Line Fault Detection of Sensor Measurement. In: Proc. of IEEE Sensors, vol. 2, pp. 974–979 (2003)
10. Pai, H.T., Han, Y.S.: Power-Efficient Direct-Voting Assurance for Data Fusion in Wireless Sensor Networks. IEEE Trans. on Computers (accepted)
11. Varshney, P.K.: Distributed Detection and Data Fusion. Springer, New York (1997)