

## Research Article

# Fault Detection Modelling and Analysis in a Wireless Sensor Network

Shuang Jia <sup>1</sup>, Lin Ma <sup>1</sup> and Danyang Qin <sup>2</sup>

<sup>1</sup>School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150080, China

<sup>2</sup>Department of Communication Engineering, Heilongjiang University, Harbin 150080, China

Correspondence should be addressed to Lin Ma; [malin@hit.edu.cn](mailto:malin@hit.edu.cn)

Received 26 May 2018; Accepted 6 September 2018; Published 17 October 2018

Guest Editor: Ioana Fagarasan

Copyright © 2018 Shuang Jia et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For the serious impacts of network failure caused by the unbalanced energy consumption of sensor nodes, hardware failure, and attacker intrusion on data transmission, a low-energy-consumption distributed fault detection mechanism in a wireless sensor network (LEFD) is proposed in this paper. The time correlation information of nodes is used to detect fault nodes in LEFD firstly, and then the spatial correlation information is adopted to detect the remaining fault nodes, so as to check the states of nodes comprehensively and improve the efficiency of data transmission. In addition, the nodes do not need to exchange information with their neighbor nodes in the detection process since LEFD uses the data sensed by the node itself to detect some types of faults, thus reducing the energy consumption of nodes effectively. Performance analysis and simulation results show that the proposed detection mechanism can improve the transmission performance and reduce the energy consumption of the network effectively.

## 1. Introduction

A wireless sensor network (WSN) consists of a large number of sensor nodes deployed in a specific area in a self-organized manner. There is no central control node in the network, and the end-to-end information transmission can be achieved by the intermediate nodes in a multihop forwarding way [1]. WSN with the flexible, distributed, and dynamic characteristics has a wide range of applications, such as battlefield, disaster relief, exploration, environmental threat detection, and other fields [2]. The sensor nodes, however, often suffer from various attacks and other external damage since they are usually deployed in severe environments. In addition, sensor nodes have the low manufacturing cost and limited resources and radio coverage. All the factors will cause failure to the nodes and thus will reduce the accuracy of monitoring data. Therefore, the network node fault detection is very important for ensuring the accuracy of monitoring results.

The fault detection algorithms of sensor nodes in WSN can be divided into centralized fault detection and distributed

fault detection according to different data processing methods [3]. The centralized fault detection algorithms usually require all the information being collected by a particular node and then determine the states of the other nodes. These algorithms can lead to many problems easily, such as single-node failure, information loss, and much energy consumption [4]. The distributed fault detection algorithms require each node to possess the ability to detect faults and adopt the data collected by itself or the surrounding nodes to determine their own faults [5]. At present, some problems existing in the fault detection algorithms are as follows [6]:

- (1) The network energy consumption is sacrificed for higher detection accuracy and lower false-positive ratio. In the existing detection algorithms, the sensor node needs to communicate with its neighbor nodes during fault detection, which will lead to higher energy consumption
- (2) The types of fault nodes are not fully considered. Therefore, the detection performance of these

algorithms will decline rapidly if the types of fault nodes increase

- (3) The ability of sensor nodes to collect data is not fully utilized, and only the spatial correlation of the sensor network is used to achieve fault detection so that the complexity of algorithms increases significantly

In order to solve the above problems, a low-energy-consumption distributed fault detection mechanism in a wireless sensor network is proposed in this paper. LEFD adopts the time correlation features of the data collected by sensor nodes to detect certain types of fault nodes and then removes them from the network. LEFD may reduce the time and energy consumption in the communication between neighbor nodes. Then, LEFD adopts the spatial correlation properties of WSN to detect the remaining fault nodes that are not detected during the initial detection phase. If the measured value of a node is the same or similar to that of the neighbor which is in the normal state, the node can be considered a normal node. Otherwise, the node is considered a faulty node. The algorithm also considers the transient faults in the sensor readings and corrects the fault data using the data collected in a short period of time when transient faults occur, which avoids mistaking the normal node as a faulty node.

## 2. Related Work

A distributed Bayesian algorithm for detecting and correcting node faults in WSN (BAFD) is proposed [7]. In BAFD, a sensor node exchanges information with its neighbor nodes to obtain the statistical probability of the event, and the failure ratio of the node is used to identify events and fault nodes. A fault detection scheme is proposed in [8], where each node detects any suspicious behavior using time correlation in its own reading, and the suspected node is required to communicate with the confident neighbor node to find the fault node. The algorithm has high detection accuracy and low communication overhead but does not take into account the impact of transient failure. The authors in [9] propose a distributed byzantine fault detection method based on hypothesis testing, the Neyman-Pearson test method is used to predict the fault states of each sensor node and adjacent sensor nodes, and then the final state of the node is determined by voting in this mechanism. A distributed fault detection method based on metric correlation is proposed in [10]. The algorithm detects the fault nodes through the internal metric correlation of sensor nodes. The computational complexity of the algorithm is low, but it does not consider the influence of transient faults. A distributed localized fault sensor detection algorithm for wireless sensor networks (DLFS) is proposed in [11], and the mutual test results between nodes and neighboring nodes are utilized to determine the states of nodes. DLFS has high detection accuracy and low computational complexity, but the algorithm requires at least two communications between adjacent nodes, thus resulting in much energy consumption. The authors in [12] propose a fault detection mechanism based on the hidden Markov

random field. The HMRF model is used to characterize the correlation between the measured value and the actual value of the sensor node, and then the parameters of the HMRF model are obtained through the variable error estimation method to determine the state of the node. The method has high detection accuracy and low false-positive ratio, but it also causes high computational complexity and much energy consumption. In [13], a fully distributed fault detection algorithm is proposed. In this algorithm, the nodes first collect the measurements of their neighborhoods and process them to determine whether they contain an exception value and broadcast the results. Then, nodes determine their own operational state autonomously. Therefore, the computational complexity of the algorithm is low, while the algorithm needs to communicate with its neighbor nodes several times, thus leading to much energy consumption. The authors in [14] propose a novel method for detecting a sensor that generates fault data in a distributed manner. The algorithm detects the fault nodes in the cluster locally through the cluster head and uses the trust concept to identify the type of data failure, which may reduce the influence of a fault node on sensor probability. However, this method leads to the uneven energy consumption of sensor nodes.

It can be seen that the existing detection algorithm does not fully consider the types of the faulty nodes, resulting in poor detection performance. In addition, the multiple communications between adjacent nodes generate much energy consumption.

Aimed at improving the detection accuracy and reducing the energy consumption of the existing detection algorithm, this paper presents a low-energy-consumption distributed fault detection mechanism in a wireless sensor network. The main contributions of this paper are summarized as follows:

- (1) A low-energy-consumption distributed fault detection mechanism is proposed, which uses the time correlation information of sensor nodes to detect the fault nodes in the initial detection stage. During the detection of the remaining nodes, the other nodes do not need to communicate with the detected fault nodes, thus reducing the communication traffic and network energy consumption
- (2) All kinds of fault nodes are fully considered to ensure the performance of the detection algorithm. In addition, the proposed algorithm also considers the nodes that may have transient faults in the sensor readings. For transient faults, the fault values will be promptly corrected by LEFD to avoid mistaking the normal node as a faulty node, thus improving the utilization of nodes and reducing false-positive ratio

## 3. System Model

*3.1. Network Model.* WSN is a special wireless communication system, which does not depend on any fixed communication facilities; it can be deployed in a complex environment for data communication rapidly, and the architecture of WSN is shown in Figure 1. Each node plays the role of a

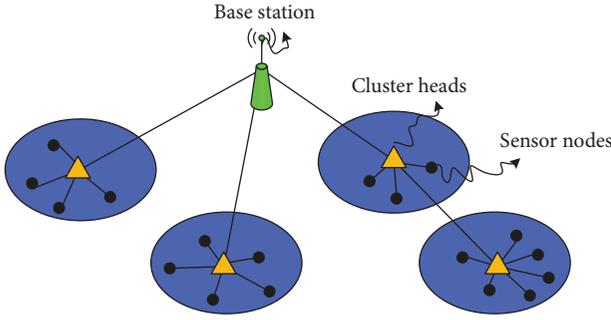


FIGURE 1: Network model.

router and endpoint and has access service and wireless backbone interface [15]. There is no absolute domination of the nodes in WSN, and each node is equal and independent. The data between nodes are transmitted to the destination node by intermediate nodes; that is, data transmission is carried out by multihop forwarding, which can guarantee the flexibility of network topology.

Assume that the number of nodes randomly deployed in a specific area is  $N$ . These sensor nodes have the same communication radius  $R_{\max}$ . A node stores at least  $k$  segments of data that has been collected before executing the fault detection algorithm.  $n_i$  represents the  $i$ -th node in WSN. The node in the node  $n_i$ 's communication radius is called the neighbor node of  $n_i$ .  $N(n_i)$  denotes all neighbor nodes of  $n_i$ , and  $\text{Num}(N(n_i))$  represents the number of neighbor nodes of  $n_i$ .  $d_i^t$  denotes the measurement data of the node  $n_i$  at  $t$ . It is assumed that the  $k$  segments of the data have been collected in the sensor and stored in the memory before  $t$ , that is,  $d_i^t, d_i^{t-1}, \dots, d_i^{t-k+1}$ . Node  $n_i$  and node  $N(n_i)$  are in the same or similar environment, which means that the neighbor nodes of  $n_i$  are also in the same event area if the node  $n_i$  is in the event area and the neighbor nodes of  $n_i$  are also in the same normal area if the node  $n_i$  is in the normal area. The remaining parameters are shown in Table 1.

**3.2. Fault Model.** Nodes could still receive, send, collect, and process data if the network is partially faulty, but the data collected by nodes is usually wrong. According to the abnormal data collected by nodes, the fault of sensor nodes can be divided into the following specific types [16]:

- (1) Fixed fault: a sensor with this fault collects data with the same reading, and the data is not affected by the environment
- (2) Random fault: node readings are random and uncertain
- (3) Offset fault: the node readings deviate from normal values, and the readings may change if the environment changes
- (4) Transient fault: transient fault may occur in a short time due to the hardware characteristics and the impact of the environment on the data collecting

TABLE 1: Notations and their definitions.

Notation	Definition
$n_i$	The $i$ -th node
$\text{Num}(N(n_i))$	Number of neighbors of $n_i$
$S_i$	Initial states of sensor nodes
$v_i^r$	Each row in matrix $\mathbf{Q}$
$E_c$	Total energy consumption
$N(n_i)$	All neighbor nodes of $n_i$
$d_i^t$	Measurement data of node $n_i$ at $t$
$FS_i$	Final states of sensor nodes
$k$	Number of data collected by node $n_i$
$R_{\max}$	Communication radius
DA	Detection accuracy
FPR	False-positive ratio

process, resulting in data anomalies occurring one or more times

In order to improve the utilization of the sensor nodes, this paper considers that the nodes with transient faults are normal nodes because the readings of these nodes are available at most of the time.

## 4. Proposed Fault Detection Model

**4.1. Detection Principle.** The data collected by sensor nodes in a short time is temporally relevant, which means that the collected data is the same or similar in a short time, and the change is not so great [17]. LEFD can detect some types of fault nodes based on this feature, such as random faults and transient faults. The value of the collected data in a short time is unstable when these faults occur. However, this paper will consider that the nodes with transient faults are normal nodes to improve the utilization of nodes, so only the collected data generated when the fault occurs is corrected; the normal node will not be mistaken for a fault node. The matrix  $\mathbf{Q}$  is established to determine whether there are transient faults or random faults based on the difference of the data collected by nodes. The faulty data will be replaced by the collected normal data at other times for transient faults; therefore, the false-positive ratio can be effectively reduced. However, it is not enough to use only the time correlation information. For example, the node's reading still satisfies the time correlation feature when fixed faults or offset faults occur, and this type of fault nodes cannot be detected just by using the time correlation feature, so the neighbor nodes are necessary. The node fails if the collected data of most neighbor nodes is not similar to the node's collected data; that is, the sensor nodes have spatial correlation property, which means that most sensor nodes have the same or similar readings in smaller areas.

The differences between LEFD and the existing algorithms can be summarized from the above analysis. Firstly, LEFD uses time correlation information to detect certain

types of fault nodes and corrects some values as needed, and then the spatial correlation property of nodes is adopted to detect the remaining fault nodes. However, the existing algorithms do not use time correlation information, or only the spatial correlation information is adopted to detect fault nodes, so there always are undetected fault nodes in the network. Secondly, the existing algorithms do not consider the transient faults of nodes so that the normal node is mistaken as a faulty node, thus reducing the utilization ratio of nodes.

**4.2. Detection Method.** The latest  $k$  segments of data can be obtained after the node  $n_i$  collects the data at time  $t$ . The matrix  $\mathbf{Q}$  is established according to

$$\mathbf{Q}_{m \times n} = \begin{cases} 0, & \text{if } (|d_i^m - d_i^n| \leq \beta_1), \\ 1, & \text{otherwise,} \\ m, n = \{0, 1, \dots, k-1\}. \end{cases} \quad (1)$$

For each row in matrix  $\mathbf{Q}$ ,  $v_i^r$  is calculated as

$$v_i^r = \begin{cases} 0, & \text{if } \left( \sum_{j=t-k+1}^t \mathbf{Q}_{ij} < \frac{k}{2} \right), \\ 1, & \text{otherwise,} \\ t-k+1 \leq r \leq t. \end{cases} \quad (2)$$

At time  $t$ , the value of  $v_i^r$  is corrected by  $v_i^t$ :

$$v_i^t = \begin{cases} 0, & \text{if } \left( \sum_{r=t-k+1}^t v_i^r < \frac{k}{2} \right), \\ 1, & \text{otherwise.} \end{cases} \quad (3)$$

Any measured value at other times when  $v_i^r = 0$  can be considered its value at time  $t$ .

Equation (4) is used to determine the initial states of sensor nodes:

$$S_i = \begin{cases} 0, & \text{if } \left( \sum_{r=t-k+1}^t v_i^r < \frac{k}{2} \right), \\ 1, & \text{otherwise.} \end{cases} \quad (4)$$

For the node  $n_i$  with state 0, the neighbor reading whose initial fault condition is 0 is obtained. Then, the final state of the node is determined according to (5) and (6):

$$v_{ij}^t = \begin{cases} 0, & \text{if } (|d_i^t - d_j^t| \leq \beta_2), \\ 1, & \text{otherwise.} \end{cases} \quad (5)$$

$$FS_i = \begin{cases} 0, & \text{if } \left( \sum_{n_j \in N(n_i), S_j=0} v_{ij}^t < \frac{\text{Num}(N(n_i), S=0)}{2} \right), \\ 1, & \text{otherwise,} \end{cases} \quad (6)$$

where  $\text{Num}(N(n_i), T=0)$  denotes the number of neighbor nodes of  $n_i$  and the number of nodes whose state may be normal.  $\beta$  represents the node failure thresholds.  $FS_i = 0$  represents that the node  $n_i$  is a normal node. Otherwise, the node  $n_i$  is a faulty node.

For example, assuming that  $k = 5$  and  $\text{Num}(N(n_i)) = 5$ , the  $k$  segments of data collected by the node  $n_i$  at time  $t$  and before  $t$  are  $\{d_i^t, d_i^{t-1}, d_i^{t-2}, d_i^{t-3}, d_i^{t-4}\}$ , and their corresponding values are  $\{60.12, 30.23, 31.54, 10.68, 30.87\}$ . The neighbor nodes' reading of node  $n_i$  is  $\{d_1^t, d_2^t, d_3^t, d_4^t, d_5^t\}$  at  $t$ , whose values are  $\{70.22, 31.35, 65.79, 30.84, 31.10\}$ , and  $\beta = 2$ . Then, the matrix  $\mathbf{Q}$  is established according to (1):

$$\mathbf{Q}_{5 \times 5} = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}. \quad (7)$$

There are  $v_i^r = \{0, 1, 0, 0, 1\}$  and  $t-4 \leq r \leq t (t \geq 4)$  according to (2). According to (3), the value of  $v_i^t$  is corrected to 0; then, there is  $v_i^r = \{0, 1, 0, 0, 0\}$ . The measured value at  $r$  is used to update  $d_i^t$  when there is  $v_i^r = 0$ , namely,  $d_i^t = 30.23$ . According to (4), the initial fault state of node  $n_i$  is considered  $S=0$ , which represents that the node  $n_i$  is a normal node. Then, the algorithm obtains the neighbor nodes' data of node  $n_i$  at time  $t$ . Equation (6) shows  $FS_i = 0$ , and the final state of  $n_i$  is normal.

As can be seen from this example, LEFD is a very effective detection method for transient faults and random faults. The detailed description of LEFD is shown in Algorithm 1.

The algorithm adopts the historical data sensed by nodes to determine the initial state of nodes. The node may be a normal node if the collected data is stable in a short time (almost no change). Otherwise, the node may be a faulty node. In other words, only the sensor node's own data can be used to identify some of the fault nodes. After determining the initial states of nodes, LEFD further determines that the initial states of their neighbor nodes are normal for the nodes with a normal initial state. A node will be determined as a normal node if its measured value is similar to that of most of its neighbor nodes. In the whole algorithm implementation process, the fault nodes that are identified in the initial detection process are no longer able to communicate with other normal nodes, and the algorithm adopts the data from nodes whose initial state is normal. This method not only consumes less energy but also reduces the error detection ratio. In addition, LEFD also considers the transient faults of nodes. The algorithm will correct the false readings when

```

(1) Begin
(2) for each node  $n_i$  in WSN ( $i = 1, 2, \dots, N$ )
/* The following method is adopted to establish  $Q^*$  */
(3)   for each  $k$  times before time  $t$  (including time  $t$ )
(4)     if  $|d_i^m - d_i^n| \leq \beta_1$ 
(5)        $Q_{mn} = 0$ 
(6)     else
(7)        $Q_{mn} = 1$ 
(8)     end if
(9)   end for
/* Generate test  $v_i^r$  */
(10)  if  $\sum_{j=t-k+1}^t Q_{ij} < k/2$ 
(11)     $v_i^r = 0$ 
(12)  else
(13)     $v_i^r = 1$ 
(14)  end if
/* Correct  $v_i^r$  */
(15)  if  $\sum_{r=t-k+1}^t v_i^r < k/2$  and  $v_i^t = 1$ 
(16)     $v_i^t = 0, d_i^t = d_i^{t-x} // ((t-x) \in [t-k+1, t])$  and  $v_i^{t-x} = 0$ 
(17)  else if  $\sum_{r=t-k+1}^t v_i^r \geq k/2$  and  $v_i^t = 0$ 
(18)     $v_i^t = 1, d_i^t = d_i^{t-x} // ((t-x) \in [t-k+1, t])$  and  $v_i^{t-x} = 1$ 
(19)  end if
/* Generate a state value  $S_i = 0$  based on the value of  $v_i^r$  */
(20)  if  $\sum_{r=t-k+1}^t v_i^r < k/2$ 
(21)     $S_i = 0$ 
(22)  else
(23)     $S_i = 1$ 
(24)  end if
/* For the nodes with status value  $S_i = 0$ , test each member of their neighbors to generate test  $v_{ij}^t \{0, 1\}$  by adopting the following way */
(25)  if  $|d_i^t - d_j^t| \leq \beta_2$ 
(26)     $v_{ij}^t = 0$ 
(27)  else
(28)     $v_{ij}^t = 1$ 
(29)  end if
/* Make the final decision of the nodes' state  $FS_i$  */
(30)  if  $\sum_{n_j \in N(n_i) \text{ and } S_j=0} v_{ij}^t < Num(N(n_i) \text{ and } S=0)/2$ 
(31)     $FS_i = 0$ 
(32)  else
(33)     $FS_i = 1$ 
(34)  end if
Process end

```

ALGORITHM 1: The pseudocode of low-energy-consumption distributed fault detection algorithm.

transient faults occur, which means that the algorithm adopts the reading at other times instead of the reading at this time to further improve the fault tolerance ability of sensor nodes to transient faults.

## 5. Simulation Experiment and Performance Analysis

**5.1. Performance Indicators.** The two indicators are usually adopted to evaluate the effect of fault node identification, namely, detection accuracy and false-positive ratio.

Detection accuracy (DA) refers to the ratio between the number of fault nodes that have been correctly identified and the total number of actual fault nodes:

$$DA = \frac{|F \cap A|}{|A|}, \quad (8)$$

where  $F$  represents the set of fault nodes detected by the algorithm and  $A$  represents the set of actual fault nodes.

The false-positive ratio (FPR) refers to the ratio between the number of normal nodes which are identified as fault nodes and the total number of normal nodes:

$$\text{FPR} = \frac{|F - A|}{N - |A|}. \quad (9)$$

Most of the energy consumption is caused by communication between nodes [18]. Thus, the total number of communications between nodes can be adopted to represent the total network. When the communication radius of node is  $R_{\max}$ , it is assumed that the average energy consumption when the node  $n_i$  communicates with its neighbor node once is  $\text{elec}_i$ :

$$E_c = \sum_{i=1}^N \text{elec}_i, \quad (10)$$

where  $E_c$  denotes the total energy consumption.

**5.2. Parameter Settings.** The performance of LEFD was analyzed using NS2 in this study [15]. In order to maintain the generality, it is assumed that the position of each node is known and all nodes have the same communication radius  $R_{\max}$ , and the reading of the nodes in the normal region is subject to the distribution of  $N(\mu, \sigma)$  ( $\mu = 35, \sigma = 1$ ). At least 5 segments of data ( $k = 5$ ) are stored in each sensor node. The value of  $k$  should not be chosen too high because the sensor nodes have limited storage capacity. The data may take up too much storage space if the value of  $k$  is too high. The node failure threshold  $\beta$  is 5, and the basic idea of the node failure threshold selection is to determine the node failure threshold according to the allowable deviation of the sensor node. The key step is designing an observer. The output of the observer and the output of the sensor node constitute a redundant signal, and then the two signals are compared to obtain the sensor residual sequence. The allowable error of the sensor node is selected as the node failure threshold [19]. Since the fixed fault is similar to the offset fault, the two types of faults are also regarded as offset faults. The results were obtained from the mean of 100 experiments. All of the simulation parameters are shown in Table 2.

**5.3. Experimental Results and Performance Analysis.** Figures 2(a) and 2(b) show the performance comparison results of different algorithms in terms of DA and FPR when only offset faults occur. It can be seen that the DA of DLFS is much higher when the sensor fault probability is less than 30% as shown in Figure 2(a), and the DA of LEFD proposed in this paper is similar to that of BAFD; the DA of DLFS is rapidly reduced compared with that of LEFD and BAFD when the sensor fault probability is higher than 30%. However, DLFS has low FPR, and the FPR of LEFD is between the FDR of DLFS and BAFD. Based on all the above factors, the performance of the LEFD algorithm is between the performances of DLFS and BAFD when only offset faults occur.

Figures 3(a) and 3(b) show the performance comparison results of different algorithms in terms of DA and FPR when

TABLE 2: Simulation parameters.

Parameter	Value
The type of nodes	Normal node, fault node
The number of nodes	1024
Network range	32m × 32m
The reading range of offset fault	61–70
The reading range of random fault	1–100
The reading range of transient fault	1–100
$R_{\max}$	2
$\text{elec}_i$	$10^{-5}$ J
$k$	5

only random faults occur. It can be seen that the sensor fault probability has a little effect on the DA, and the FPR increases with the increasing sensor fault probability as shown in Figure 3. However, LEFD also has good performance and always maintains high DA and low FPR even in the case of high sensor fault probability for random faults, since the LEFD algorithm first checks whether the nodes' reading is stable in a short time. The node may be faulty if its reading is unstable, because the data of the random fault sensor is random and unstable. Since the range of random faults is from 1 to 100, DLFS and BAFD are effective for this fault but are less efficient than LEFD. The DA of DLFS and BAFD are also more than 93%, and the DA may increase (such as that of BAFD) when the sensor fault probability increases. But the FPR of DLFS and BAFD also increase when the sensor fault probability increases. However, the FPR of LEFD is almost zero.

Figure 4 shows the relationship between the sensor fault probability and the FPR if only transient faults occur when  $k = 5$ . It can be seen that the FPR of all algorithms increase with the increasing sensor fault probability as shown in Figure 4. The FPR of LEFD is very low because it has a good ability to handle transient faults. For example, the FPR of LEFD is still less than 5% when the sensor fault probability is 50%, since LEFD determines whether the collected data is correct according to the  $k$  segments of data. LEFD will replace the current data with the data collected at another time to avoid the impact of transient faults if data is wrong. This is why LEFD has a good fault tolerance performance for transient faults. However, both DLFS and BAFD do not consider transient faults, so the FPR will continue to increase as the sensor fault probability increases.

Figures 5(a) and 5(b) show the relationship between the sensor fault probability and the DA/FPR when the offset faults, random faults, and transient faults occur randomly, respectively. The DA of DLFS and LEFD is almost the same as shown in Figure 5(a). The DA of DLFS and LEFD are higher than that of BAFD when the sensor fault probability is less than 35%. However, the DA of DLFS will decline rapidly when the sensor fault probability is greater than 35%. The FPR of LEFD is the lowest of the three algorithms. In short, the performance of the LEFD algorithm achieves our expectations in the event of mixed faults.

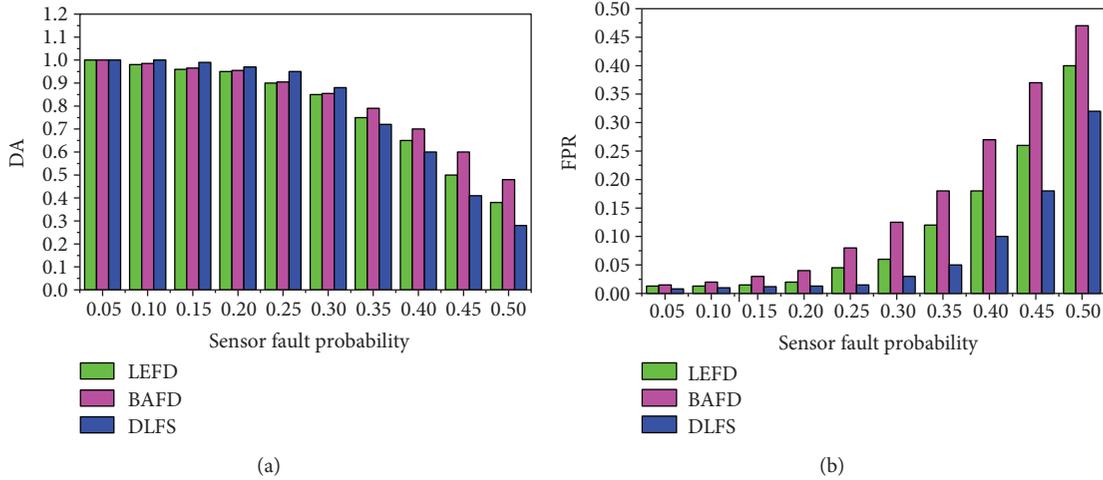


FIGURE 2: (a) Detection accuracy under offset faults. (b) False-positive ratio under offset faults.

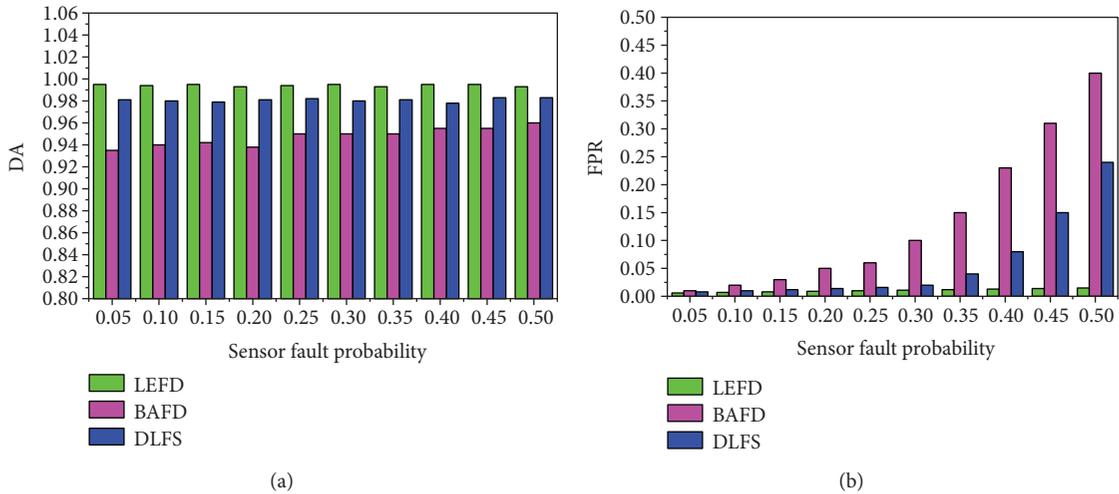


FIGURE 3: (a) Detection accuracy under random faults, (b) False positive ratio under random faults.

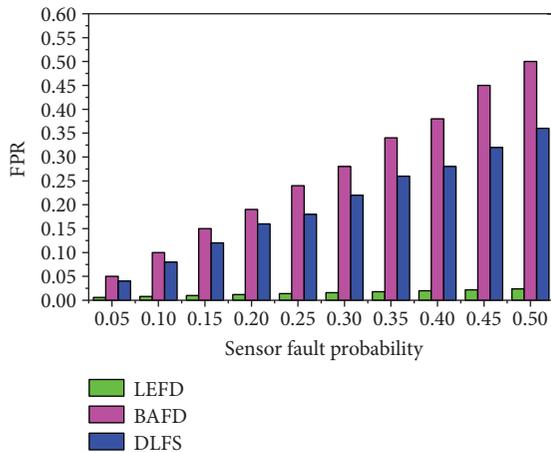


FIGURE 4: The relationship between the false-positive ratio and the sensor fault probability under transient faults.

Figure 6 shows the relationship between the energy consumption (EC) of DLFS, BAFD, and LEFD and the sensor fault probability when the transient fault ratio and the offset fault ratio are 1 : 1, the communication radius  $R_{max}$  is 2, and the nontransient faults occur. It can be seen that DLFS has much higher energy consumption as the sensor fault probability increases under the same conditions as shown in Figure 6, since each node needs to communicate with its neighbor nodes at least twice (the first communication is to exchange the initial data set, the second communication is to exchange the initial state of each node). However, nodes that have not yet determined the final state need to make the third communication. As a result, the energy consumption of DLFS is always relatively high. Each node only needs to communicate with its neighbor nodes once for BAFD, so its network energy consumption is moderate and does not change with the sensor fault probability. LEFD first adopts the time correlation information for initial fault detection, and each node does not need to communicate with its neighbor nodes in this process. Only nodes that have been detected

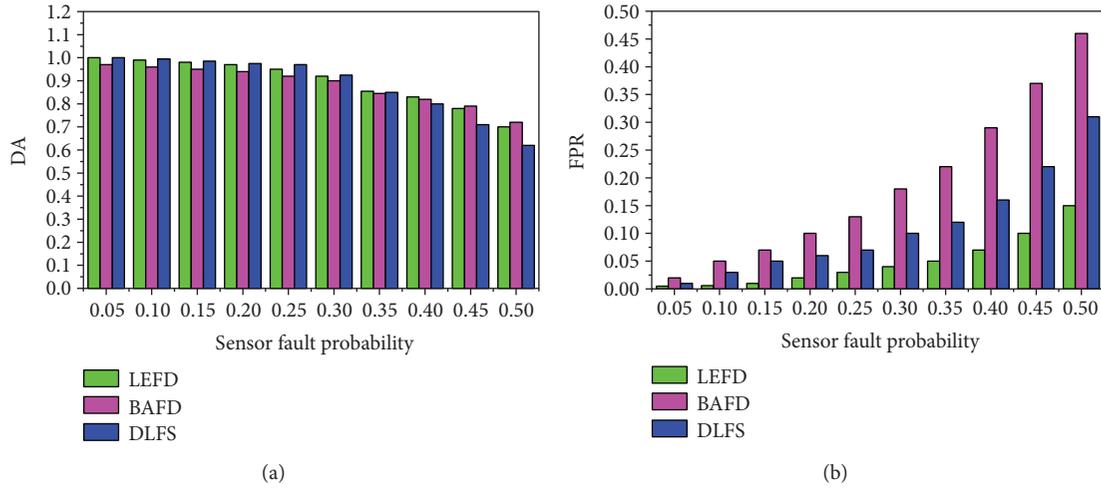


FIGURE 5: (a) The relationship between the detection accuracy and the sensor fault probability under mixed faults. (b) The relationship between the false-positive ratio and the sensor fault probability under mixed faults.

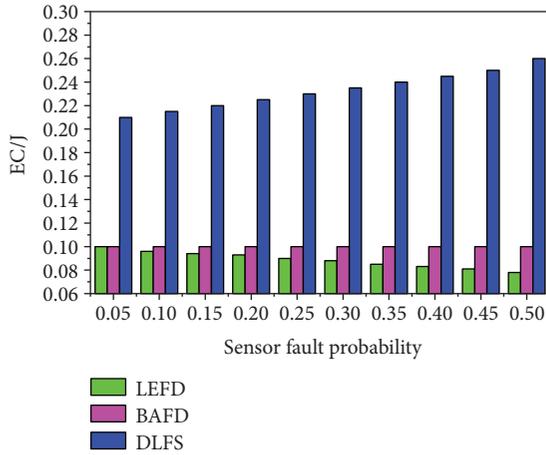


FIGURE 6: Energy consumption of different algorithms.

to have a normal state need to communicate with the neighbor nodes and consume additional energy. Therefore, most of the nodes will be detected as fault nodes by the LEFD algorithm in the case of high sensor fault probability, and the energy consumption of the network also decreases. In summary, the energy consumption of the LEFD algorithm is low.

## 6. Conclusions

WSN is an important component of modern mobile communication systems. However, network performance is seriously affected due to the breakage of data link and frequent changes of network topology. Therefore, a low-energy-consumption distributed fault detection mechanism in WSN is proposed in this paper. LEFD adopts the data sequence collected by the sensor node itself to detect a particular type of fault and then further uses the neighbor data to determine the states of nodes, thus reducing the communication traffic and network energy consumption. In addition, LEFD also considers the nodes that may have transient faults. For transient faults,

the fault values will be promptly corrected by LEFD to avoid mistaking the normal node as a faulty node, thus reducing false-positive ratio. The simulation results show that LEFD has high detection accuracy, low false-positive ratio, and less energy consumption for various faults. Future research will study a fault tolerance method for WSN, which may provide a new way for the effective transmission of data and ubiquitous routing.

## Data Availability

The data used to support the findings of this study are currently under embargo while the research findings are commercialized. Requests for data, 12 months after the publication of this article, will be considered by the corresponding author.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China (61571162, 61771186), Heilongjiang Province Natural Science Foundation (F2016019), Ministry of Education-China Mobile Research Foundation (MCM20170106), and University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province (UNPYSCT-2017125).

## References

- [1] Z. Chen, M. He, W. Liang, and K. Chen, "Trust-aware and low energy consumption security topology protocol of wireless sensor network," *Journal of Sensors*, vol. 2015, Article ID 716468, 10 pages, 2015.
- [2] M. P. Fanti, A. M. Mangini, and W. Ukovich, "Fault detection by labeled petri nets in centralized and distributed

- approaches,” *IEEE Transactions on Automation Science and Engineering*, vol. 10, no. 2, pp. 392–404, 2013.
- [3] M. Saihi, B. Boumedyen, A. Zouinkhi, and M. N. Abdulkrim, “A real time centralized fault detection application in wireless sensor networks,” in *International Conference on Automation, Control, Engineering and Computer Science (ACECS'14)*, pp. 95–101, Sousse, Tunisia, March 2014.
- [4] D. Kirby, N. Fischer, A. Kalra, and D. Haes, “Case study: centralized ground fault detection system for LADWP ungrounded distribution system,” in *The 41st Annual Western Protective Relay Conference*, pp. 1–10, Spokane, Washington, October 2014.
- [5] C. Lo, J. P. Lynch, and M. Liu, “Distributed reference-free fault detection method for autonomous wireless sensor networks,” *IEEE Sensors Journal*, vol. 13, no. 5, pp. 2009–2019, 2013.
- [6] R. Rafeh, “Proposing a distributed fault detection algorithm for wireless sensor networks,” *Soft Computing Journal*, vol. 2, no. 2, pp. 26–35, 2014.
- [7] B. Krishnamachari and 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, 2004.
- [8] K. P. Sharma and T. P. Sharma, “rDFD: reactive distributed fault detection in wireless sensor networks,” *Wireless Networks*, vol. 23, no. 4, pp. 1145–1160, 2017.
- [9] M. Panda and P. M. Khilar, “Distributed byzantine fault detection technique in wireless sensor networks based on hypothesis testing,” *Computers and Electrical Engineering*, vol. 48, pp. 270–285, 2015.
- [10] Q. Liu, Y. Yang, and X. Qiu, “A metric-correlation-based distributed fault detection approach in wireless sensor networks,” in *Proceedings of the 2015 IEEE Symposium on Computers and Communication (ISCC)*, pp. 526–531, Washington, DC, USA, July 2015.
- [11] J. Chen, S. Kher, and A. Somani, “Distributed fault detection of wireless sensor networks,” in *Proceedings of the Workshop on Dependability Issues in Wireless ad hoc Networks and Sensor Networks (DIWANS'06)*, pp. 65–71, Los Angeles, Fla, USA, September 2006.
- [12] J. Gao, J. Wang, and X. Zhang, “HMRF-based distributed fault detection for wireless sensor networks,” in *2012 IEEE Global Communications Conference (GLOBECOM)*, pp. 640–644, Anaheim, CA, USA, December 2012.
- [13] W. Li, F. Bassi, D. Dardari, and M. Kieffer, “Low-complexity distributed fault detection for wireless sensor networks,” in *IEEE International Conference on Communications*, pp. 6712–6718, Kuala Lumpur, Malaysia, May 2016.
- [14] Z. Taghikhaki and M. Sharifi, “A trust-based distributed data fault detection algorithm for wireless sensor networks,” in *2008 11th International Conference on Computer and Information Technology*, pp. 1–6, Khulna, Bangladesh, December 2008.
- [15] N. Marchang and R. Datta, “Light-weight trust-based routing protocol for mobile ad hoc networks,” *IET Information Security*, vol. 6, no. 2, pp. 77–83, 2012.
- [16] Q.-J. Wang and J.-L. Cheng, “Research on fault mode and diagnosis of methane sensor,” *Journal of China University of Mining and Technology*, vol. 18, no. 3, pp. 386–388, 2008.
- [17] P. Jiang, “A new method for node fault detection in wireless sensor networks,” *Sensors*, vol. 9, no. 2, pp. 1282–1294, 2009.
- [18] A. Cornejo, S. Vigar, and J. L. Welch, “Reliable neighbor discovery for mobile ad hoc networks,” *Ad Hoc Networks*, vol. 12, no. 6, pp. 259–277, 2014.
- [19] X. H. Wang, “Research on threshold selection principle based on sensor fault detection,” *Mechanical Engineering & Automation*, vol. 4, no. 2, pp. 105–106, 2009.



**Hindawi**

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

