http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 10 (1): 80-88, 2011
ISSN 1812-5638

A Cluster-based Real-time Fault Diagnosis Aggregation Algorithm for Wireless Sensor Networks

Wei Wang, Bingwen Wang, Zhuo Liu and Lejiang Guo
Department of Control Science and Engineering,
Huazhong University of Science and Technology, Wuhan Hubei, 430074, China

Abstract: In Wireless Sensor Networks (WSNs), it is of great importance for fault diagnosis in the WSNs to ensure the accuracy of the information collected from the network, for the low cost and the deployment of a large number of sensor nodes in uncontrolled or harsh or hostile environments. In this study, we have presented a protocol, named CRFDA (Cluster-Based Real-time Fault Diagnosis aggregation algorithm for WSNs), which is based on the comparison approach aiming at achieving a correct and complete diagnosis for hierarchical WSNs. CRFDA is based on clustering in the network which can be carried out simultaneously in all cluster. Besides, it is a real-time diagnosis method and can be spontaneously implemented in each round. It is known from CRFDA that both hard and soft faults can be diagnosed correctly in the finite time. Algorithm correctness, communication complexity and time complexity proofs are presented in this study. Compared with other existing protocols, CRFDA performs better. We have also developed a simulator by NS-2 (using LEACH) to analyze the correctness of our algorithm. The simulation results show that our algorithm is a correct and efficient fault diagnosis algorithm in WSNs.

Key words: Wireless sensor networks, data aggregation, clustering topology, fault diagnosis, real-time diagnosis

INTRODUCTION

Wireless Sensor Networks (WSNs) are self-organizing, rapidly deployable, dynamic variability and require no fixed infrastructure. The significant advantage of WSNs is that it is the low-cost, low-power, multifunctional miniature wireless sensor nodes which consist of sensing, data processing and communication components (Akyildiz *et al*., 2002). Those nodes can easily be deployed into a expected remote of dangerous area by planes or other means and to perform automatically various monitoring tasks.

Due to the low cost and the deployment of a large number of sensor nodes in uncontrolled or even harsh or hostile environments, it is not uncommon for the sensor nodes to become faulty and unreliable. The networks must exclude the faulty sensors to ensure the network quality of service. Sensor nodes are powered by batteries, which are considered as limited resources (Paradis and Han, 2007). It is very expensive for the base station to collect information from every sensor and identify faulty sensors in a centralized manner. Different applications may require the fault detection to be conducted in a real-time mode with low latency or high throughput. Therefore, a Real-time fault diagnosis algorithm for each node is highly preferred in wireless sensor networks.

Research in wireless sensor networks, clustering routing is a key research direction, there are already many classical protocols such as LEACH (Heinzelman *et al*., 2000), HEED (Younis and Fahmy, 2004), LEACH-C (Heinzelman *et al*., 2002) etc. In these routing protocols, the network is usually divided into clusters, each cluster consists of a cluster head and some cluster members, of which, cluster head is responsible to communicate with the base station. Since WSNs face a big challenge is the problem of limited energy resources, we use data aggregation (Intanagonwiwat *et al*., 2003; Xue *et al*., 2005; Samuel *et al*., 2002; Yao and Gehkre, 2002) to solve these problems that improve the efficiency of information gathering and avoid communication bandwidth and energy consumption in data transmission. Data aggregation is mainly in accordance with certain rules to handle data from multiple nodes to form a better result to meet customer needs, it helps to reduce the amount of data transmitted between sensor nodes and the BS. Since, data aggregation is used within the network processing, to improve the accuracy of information

Corresponding Author: Wei Wang, Department of Control Science and Engineering,
Huazhong University of Science and Technology, Wuhan Hubei, 430074, China

obtained and avoid misleading results, it is very necessary to diagnosis the sensor nodes of the network.

Our major contribution of this study is the development of a localized cluster-based real-time fault diagnosis aggregation algorithm (CRFDA) for WSNs.

LEACH (Heinzelman *et al*., 2000) is one of the earliest proposed clustering routing protocol in WSNs,it is also a classical algorithm, It utilizes randomized rotation of cluster heads to evenly distribute the energy load among all of the sensor nodes in the networks, The operation of LEACH is divided into rounds. Each round begins with a set-up phase when the clusters are organized, followed by a steady-state phase when data are transferred from the nodes to the cluster head and on to the BS. The duration of the steady phase is longer than the duration of the set-up phase in order to minimize the overhead. HEED protocol (Younis and Fahmy, 2004), the cluster head node election took into account the residual energy and node distribution around the neighbor nodes. LEACH-C (Heinzelman *et al*., 2002) protocol with centralized control, the choice of cluster heads by the base station. TEEN (Manjeshwar and Agrawal, 2001) used in data transmission phase set of soft and hard two thresholds to control the number of times to send data. PEGASIS (Lindsey and Raghavendra, 2002) proposed a near optimal chain-based protocol that is an improvement over LEACH.

Chessa and Santi (2002) developed a crash fault detection mechanism which provided bounds on the maximum number of faults which could be tolerated and the minimum number of bits to be exchanged. Chessa and Santi (2001) presented a comparison based diagnostic model based on the one-to-many communication paradigm. Chen *et al*. (2006) proposed a distributed fault detection algorithm to locate the faulty sensors in the WSNs. It calculates the measurement difference between neighbor sensors at different time to find if the current measurement of a sensor is different from its previous measurement. Ruiz *et al*. (2004), a failure detection scheme called MANNA by using management architecture for WSNs, The scheme created a manager, which has the global vision of the network, to perform complex tasks such as retrieving the node state and detect node failure. Kuo-Feng *et al*. (2006) proposed a mechanism which can both tolerate and locate data inconsistency failures in sensor networks. Luo *et al*. (2006) proposed a fault-tolerant detection scheme that explicitly introduces the sensor fault probability into the optimal event detection process, in this scheme, the neighborhood size of fault correction is chosen based on the given detection error bound such that better balance between detection accuracy and energy usage is obtained.

SYSTEM MODEL

The network model: We consider a wireless sensor network consisting of a set of sensors and a Base Station (BS) that is randomly dispersed in the target area. The BS and locations of sensors are fixed. The BS knows the locations of all sensors a prior, which can be obtained by triangulation method or by using GPS-equipped sensors. A sensor can transmit data to any other sensor and can communicate directly with the BS. The sensors periodically monitor their vicinity and generate monitoring data. The Cluster-Head (CH) are gathered from sensors which can be in the same the cluster at each round, then the CH performs data aggregation to reduce the number of messages, finally, the CH sends to the BS for further processing.

The fault model: In the network, at time t, the topology described as an undirected graph G(t) = (V,L(t)), where, V is the set of sensor nodes, L(t) is the set of logical links at time t, Given any $u,v \in V$, edge $(u, v) \in L(t)$ if and only if v is in the transmitting range of u at time t. Each node is either faulty or fault-free and faults are either hard or soft, When a node is hard-faulted, it is unable to communicate with other nodes. Soft faults are subtle, since a soft-faulted node continues to operate and to communicate with the other nodes, although with altered specifications (Chessa and Santi, 2002). In this study we utilize the invalidation rule of the gMM model (Sengupta and Dahbura, 1992; Maeng and Malek, 1981), which is used to diagnose the state of the sensor nodes, the gMM model is summarized as follow Table 1.

In this study, we use the classic clustering algorithm LEACH (Heinzelman *et al*., 2000) to discuss in this study, its Clustering network structure diagram is shown as Fig. 1.

Leach can satisfy as following:

- **A1:** Each mobile has a unique identifier; cluster-heads know the total number of sensor nodes and their identities
- **A2:** Exist a MAC protocol to solve contentions over logical links and collision problems

Table 1: The invalidation rule of the gMM model

Comparison outcome of and w generated by u	w	v	u
0	Fault-free	Fault-free	Fault-free
1	Fault-free	Faulty	Fault-free
1	Faulty	Fault-free	Fault-free
1	Faulty	Faulty	Fault-free
x	Any	Any	Faulty

x means an unreliable outcome 0 or 1

- **A3:** All the nodes in the transmitting range of the same cluster-head belong to the same cluster and they are 1-hop neighbors of the cluster-head, That is, the communication protocol provides a 1-hop reliable broadcast primitive (Pagani and Rossi, 1997), called 1_rb(·). one cluster-head is not the neighbor of another one

There properties ensure that fault-free mobiles will be able to correctly identify the sender of a message, whereas, A3 guarantees that any message broadcasted from a fault-free node is correctly received on all other fault-free nodes in its neighborhood in bounded time.

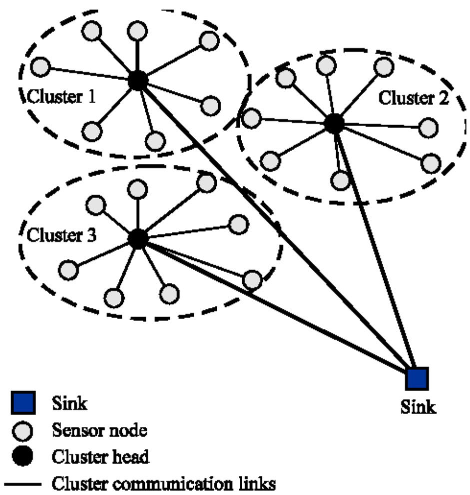

Fig. 1: The classic hierarchical ad-hoc network

The diagnostic model: Chessa and Santi (2002) presented the comparison-base diagnostic model in ad-hoc networks the model uses the comparison approach as a basis in order to help mobiles to self-diagnose each other. It exploits the shared nature of the communication in ad-hoc networks by allowing each fault-free mobile u to test its neighbors by sending them a test request. Based on the responses provided by u's neighbors, nodes can be diagnosed their state using the comparison-based fault model as shown in Table 1.

In this study, we assume that the network topology does not change during the self-diagnosis session, this means that if CH node u sends its test request at time t and given the timeout time called T_{out}, then $N_{CH}(u,t) = N_{CH}(u,t') = N_{CH}(u)$ for any $t<t'\leq T_{out}$, However, we should note that this assumption does not mean that the networks are static, rather than its topology does not change during diagnosis. In fact clustering routing protocol will have been two phases: set-up phase and steady-state phase we can also subdivide into three phase (Fig. 2) cluster head selection phase, cluster set-up phase, data transmission phase. And through the cluster set-up phase, the network topology remains stable in the data transmission phase; they will not change, namely, the nodes cannot migrate out of their neighbors' transmitting ranges. The diagnosis rules to which the sensor nodes should conform during a self-diagnosis session and called fixed topology comparison protocol, can be illustrated as follows:

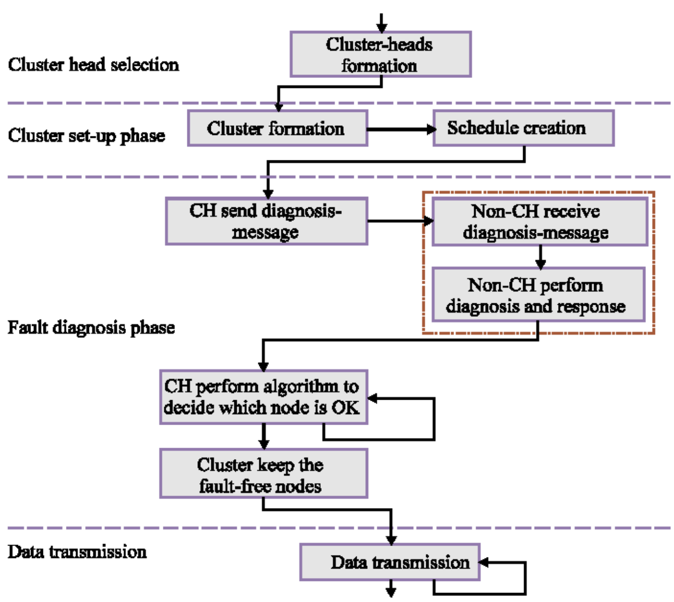

Fig. 2: The basic process of the CRFDA algorithm

(D1)Test request message generation: In the inception of the fault detection phase, at time t, CH node u generates a message of type Testing Msg. The test message is composed of a sequence number i and a test task T_i. at the same time, node u also perform the T_i by itself and generates the expected result R^u_i, then it sends the Testing Msg m = (TestingMsg,u,i,T_i) to $N_{CH}(u)$ and initiates a timer set to T_{out} and (u,i) is the header of the TestingMsg.CH node u expects to receive all its cluster nodes' responses within this bound. The T_{out} must be chosen such that all of u's fault-free cluster nodes are guaranteed to respond to the test request within the most that time hard faults are unable to respond within T_{out}.

(D2) Test request message reception: Given $v \in N_{CH}(u)$, i.e., v is a cluster node. Upon receiving the test request message m = (TestingMsg,u,i,T_i), v proceeds as follows. Firstly, v executes the test task T_i and generates the result R^v_i. Then, it sends a response message called RespondedMsg to CH node. That is it transmits m' = (ResponsedMsg,u,i,R^v_i) at time t', with $t<t'<t+T_{out}$ and (u,i) is the header of the Responsed Msg.

(D3) Response message reception: Unlike Chessa and Santi (2002) response message are received only by CH node, upon the reception of m' = (TestingMsg,u,i,T_i), CH node proceeds as follows: it compares R^u_i and R^v_i, If $R^u_i = R^v_i$, then the v is considered fault-free; otherwise, it is considered faulty. Fault-free and faulty nodes are stored in the FF(u) and F(u).

(D4) Diagnose result message reception: After CH node diagnoses all the cluster nodes, It generates a message of type Outcome Msg which is composed of (u,i) and F(u), m" = (OutcomeMsg,u,i,F(u)) then CH node distributes to the cluster, Non-CH nodes receive the Outcome Msg, if anyone own in F(u), then it automatically turns off all modules.

(D5) Timeout message reception:

- At time $t+T_{out}$, the CH node u receives the message of type TimeoutMsg from the timer; it diagnoses its cluster nodes that didn't respond within the bounded time as the faulty nodes. At this time, CH node u can know about the status of all its cluster nodes, that is, u includes two sets: FF_u denoted the set that contains all of u's fault-free nodes and F_u denoted the set that contains u's faulty nodes
- Chessa and Santi (2002), they didn't give specific methods for T_{out}, in the diagnosis of actual network, T_{out} is also difficult to accurately obtain, The value of T_{out} may also be dynamic for the following reasons: We assume the existence of faulty nodes in the network and the number of faulty nodes changes as the network's lifetime increases, this assumption exists, the number of nodes which need to diagnose in the network is decreasing. That is, T_{out} need to be adjusted dynamically. Therefore, T_{out}'s settings affect data transmission delay and energy consumption of the network. In clustering routing protocols, the value of T_{out} can been easily calculated according to the number of cluster. We let $T_{1\text{-diagnose}}$ denote the time of diagnosing one node, if there are *l* nodes in the cluster, then, $T_{out} = \alpha \times (l-1) \times T_{1\text{-diagnose}}$, α is the correction factor which can be made based on experience, under normal circumstances, should be more than one. Since, T_{out} can be accurately calculated, it is possibility to reduce the delay of the diagnosis

ALGORITHM DESCRIPTION

Here, we will Introduce a cluster-based real-time fault diagnosis aggregation algorithm for WSNs, called CRFDA, It references (Chessa and Santi, 2002), called CSFD algorithm, In the CSFD, each node eventually diagnoses all its neighbors as a testing node. Simply speaking, if there are n nodes in a transmitting range, any one of them is diagnosed repetitively n-1 times (by its n-1 neighbors) and diagnosis messages are transmitted by means of flooding, which bring about not only much bandwidth overhead for the wireless communication medium, but also much computation overhead for the energy of the nodes. CRFDA applies the clustering topology, independent operations in each cluster, to solve the above problems. the basic process of the algorithm is shown as Fig. 2.The whole process is divided into four phases: cluster head selection, cluster set-up phase, fault diagnosis phase and data transmission, which will be implemented in every round of the lifetime in the network.

As the clustering protocol algorithm, it does not exist gateway node between the two cluster, that node is part of a cluster, thus, it is not exist the problem that need diagnosis message to disseminate throughout the whole networks. In Fig. 2, cluster head selection, cluster set-up phase and data transmission are the same of LEACH, Fault Diagnosis Aggregation algorithm will action after Cluster Set-up Phase, i.e., start the Fault Diagnosis Phase. Firstly, CH node sends the test request message, while Non-CH nodes receive the test request message, they execute the algorithm to response the result of diagnosis, then, they send the response result messages to the CH node, CH node executes the CRFDA to determine whether the response nodes are normal after it received the response messages. If there are faulty nodes (where the fault is software failure, the nodes which are hardware failure can not be able to communicate and they will be considered dead), CH node distributes the information table which consists of faulty nodes. Then, Non-CH node receive the information table, if some one is diagnosed as a faulty one, then the faulty node automatically close all the modules of itself to stop working and enter the dead status, this ensures all nodes within the cluster are fault-free.

In CRFDA, cluster head node to bear the important task of networks diagnosis, it is important to ensure that the cluster head node as fault-free before diagnosing the cluster. We assume that the base station understand the nature of all nodes in the networks. And all of the nodes are homogeneous in this study. We can combine base station and all cluster-heads into one super cluster, the base station is the cluster-head. According to previously described methods for diagnosis of cluster head, If cluster head node CH_k was diagnosed as faulty node, the networks will implement the cluster-head election algorithm again, actually only need to re-select a cluster head in the cluster containing the CH_k.

Diagnosis algorithm and network routing protocols are closely related. Change the network topology will affect greatly the operating efficiency of diagnosis algorithm. In clustering protocol, the network topology is relatively static between set-up phase and steady-state phase. This provides the best opportunity to diagnose network. In LEACH, the number of cluster head node in each round is not fixed and the number of nodes in the cluster is also not fixed. CRFDA just can adapt to these changes, thus, clustering protocols can satisfy this requirement well.

Algorithm pseudo code is shown below:

```
// Data structures for sensor node u:
// N_CH(u) : set of cluster nodes at the time of CH
// V : set of all sensor nodes in the network
// F(u) : set of sensor nodes diagnosed as faulty,initialized to ∅
// T_1-diagnose : the elapsed time to execute diagnosis for a node
// l : the total number of N_CH(u)
// FF(u) : set of sensor nodes diagnosed as fault-free,initialized to ∅
// TestGenerated : boolean variable initialized to FALSE.
//                 It is TRUE if unit u generated the test request,
//                 FALSE otherwise;

Procedure Test_Initial(){
1: i = Generate a sequence number;
2: T_i = Generate a test task;
3: R_i^u = Compute the expected result of test task T_i;
4: 1_rb(m_u = (TestMsg,u,i, T_i));
5: T_out = α * (l − 1) * T_1−diagnose;
6: SetTimer(T_out);    // the T_out is for hard faults detection,
7 : TestGenerated = TRUE;
}

Procedure CH_Diagnose(){
  do{
     Receive(msg);
     if(CheckHead(msg,v) == FALSE) {
        F(u) = F(u) ∪ {v};
     } else{
        case msg.Type of {
           ResponseMsg: // msg = (ResponseMsg,u, j, R_j^v)
                // v has been diagnosed,go to next disgnose
                if(v ∈ (F(u) ∪ FF(u))) {
                   continue;
                }  else {
                   R_j^v = get the result for T_j;
                   if(R_j^v == R_j^u) {
                      FF(u) = FF(u) ∪ {v};
                   } else {
                      F(u) = F(u) ∪ {v};
                   }
                }
           Timeout:
                   F(u) = F(u) ∪ (N_CH(u) - FF(u));
        }
     }
  } while(F(u) ∪ FF(u)  ≠  N_CH(u))
  1_rb(m" = (OutcomeMsg,u,i,F(u)));
}

Procedure Response_Generate(){
     Receive(msg);
     if(CheckHead(msg,u) == FALSE) {
        exit;  // do nothing
     } else {
        case msg.Type of {
           TestingMsg: // msg = (TestingMsg,u, j, T_j)
                if(v ∈ N_CH(u)) {
                   R_j^v = generate the result for T_j;
                   1_rb (m = (ResponseMsg,u, j, R_j^v));
                }
           OutcomeMsg: // msg = (OutcomeMsg,u, j,F(u))
                if(itself IN F(u)) {
                   destory(msg);  // close all modules
                }
        }
     }
}
```

ALGORITHM EVALUATION

Here, firstly, we will prove that CRFDA protocol can carry out correct and complete self-diagnosis for hierarchical wireless sensor networks. And then, we will analyze CRFDA's communication complexity and time complexity.

Based on the discussion of the section system model, we can state the following theorem:

Theorem 1: In any round of the hierarchical wireless sensor networks, the whole network topology is fixed, assume that the CH node u generate a test request at time t. Then, at the end of time $t+T_{out}$, CH node u has correctly diagnosed the state of the cluster nodes in $N_{CH}(u)$. The life time of clustering network is calculated by round, so the diagnosis algorithm in the entire life of the network can correctly diagnose and it is real-time dynamic.

Algorithm correctness: The correctness proof is based on the following properties:

- **Correct clustered local diagnose:** The CH node can correctly diagnoses the state of its cluster node. This is a consequence of Theorem 1
- **Correct dissemination:** All nodes can be correctly diagnosed their state in the network in finite time

Lemma 1. (Correct dissemination): Let, G = (V,L) be the connected graph representing the system at the time of diagnosis, let $C = \{C_1, C_2, C_3, ..., C_k\}$ (k denote the number of clusters) denote the set of all the clusters. Each cluster(C_i) can work as an independent unit, every C_i can work in parallel and with no interference, therefore, the network diagnostics is working in parallel, they can diagnose the entire network in the finite time.

Proof: In clustering route protocol, we know each cluster is only dealing with matters within the cluster, that is, they work independently. After set-up phase, the network topology formed and relatively unchanged. Each cluster entered the fault diagnosis phase. Therefore, the diagnosis algorithm can be activated respectively by cluster head in each cluster and then they diagnose each cluster nodes. We let k denote the number of clusters; in fact, k-cluster can be as a cluster, because they simultaneously implement the diagnosis task. If the largest cluster complete the diagnosis task at time $T_{largest}$, we can always assert that the entire network was also completed the diagnosis task at time $T_{largest}$.

The following theorem states the correctness of our CRFDA which is based on the Lemma 1 and Theorem 1:

Theorem 2 (CRFDA correctness): Let, G = (V,L) be the connected graph representing the system at the time of diagnosis, CRFDA is correct and complete i.e. CRFDA can correctly and completely diagnose the state of all nodes in the network in finite time.

Communication complexity of algorithm

Theorem 3(Communication complexity): Let, G = (V,L) be the connected graph representing the system at the time of diagnosis, the communication complexity of algorithm is $O(d_{max})$.

Proof: In any cluster, given a fault-free CH node u, it can be easily seen that u generates one test request message TestingMsg. In turn, the test request message TestingMsg generates at most $|N_{CH}(u)| \leq d_{max}$ test responses, every fault-free cluster node generates one test response message ResponsedMsg. finally, u generates one diagnose result message OutcomeMsg. Let, k denote the number of clusters, n denote the total number of the networks, there are n/k nodes per cluster, every cluster simultaneously diagnoses its cluster. Thus, the communication complexity of the largest cluster (which contains the maximum number of nodes) is maximum, which is the communication complexity of the networks.

Time complexity of algorithm: Let, G = (V,L) be the connected graph representing the system at the time of diagnosis and let Δ_G denote, the diameter of G. Let, T_{gen} be an upper bound to the elapsed time between the reception of the first diagnostic message and the generation of the test request message.

Theorem 4 (Time complexity): The time complexity of the diagnosis protocol is $O(\Delta_G \cdot T_{gen} + T_{out})$

Proof: A fault-free node generates its test request message at most in time T_{gen} since the first diagnostic message is received. It means that at most in time $\Delta_G \cdot T_{gen}$ all the fault-free nodes generate their test request messages. After the test request message is generated, any CH node diagnoses its cluster nodes at most in time T_{out}.

PERFORMANCE EVALUATION

In this study, we used the NS-2 simulator for simulations of CRFDA. We did the project in 2010 between May and June and completed this project under the guidance of Professor Wang in control theory laboratory.

We used LEACH as the clustering route protocol, simulation environment is the same as LEACH, 100 sensor nodes are randomly distributed over the region of 100×100 m as shown in Fig. 3. The base station is located far away from the region, at point (50,175). In the simulation experiment, we consider node may be faulty at

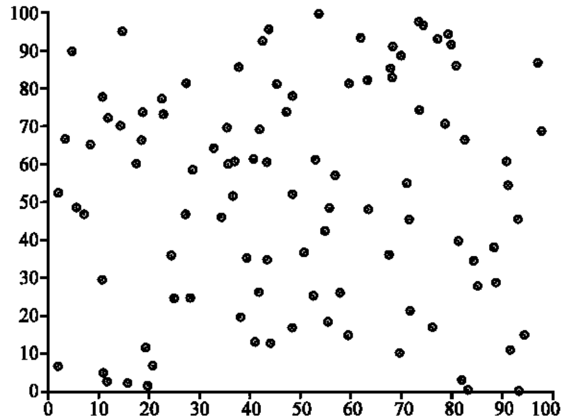

Fig. 3: 100 nodes random sensor network

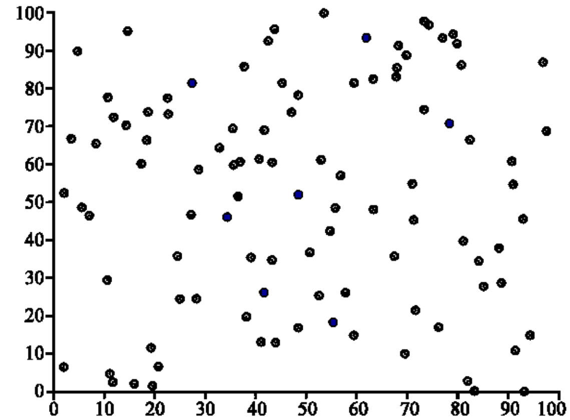

Fig. 4: Blue points denote the faulty nodes, they are eliminated from the network

any one time, which sends the data should be regarded as invalid data. Therefore, for convenience sake, we assume the following:

- At the initial phase of the network, all nodes can work correctly to ensure that when they start working status of all cluster head nodes are fault-free, this assumption is feasible
- If a node is detected as faulty node, it becomes dead node and it will be excluded from the network
- In the diagnosing process, the energy consumption for receiving and sending data is not as focused, so we ignore that

In the diagnosing process, we use the uniform rules to generate the test task T_j, it also ignores the energy consumed by the implementation of T_j.

At the same time, we assume that the network randomly generates two faulty nodes when it is running in the 100th, 200th, 300th, 400th round. Then we simulate the algorithm and mainly analyze the number of surviving nodes in the network, we hope to see that the number of nodes in the network decreases correspondingly two when the system is running in the previous descriptive rounds.

In Fig. 5, the red curve indicates the simulation result when there is no faulty node in the network and the green curve indicates another simulation result when the network exist faulty nodes and implemented the CRFDA algorithm. The figure can be clearly seen that the number of nodes is reduced 2 when the system is running in the 100th, 200th, 300th, 400th round. And in Fig. 4, the solid blue nodes denote the faulty ones which are diagnosed in the network. We can see CRFDA algorithm can correctly diagnose and eliminate the faulty nodes and let the system continue normal operation.

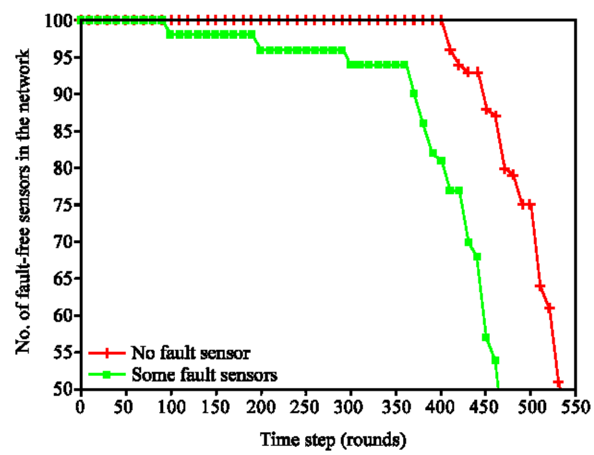

Fig. 5: CRFDA's result of diagnosis,there are two nodes be diagnosed to be faulty when the system is running, respectively, in the 100th, 200th, 300th, 400th round

It can also be seen from the Fig. 5, the number of fault-free nodes in the system will reduce when faulty nodes arise in the network. But the network lifetime will be reduced with the reduction in the number of nodes, after 350 rounds, the large number of deaths due to node energy consumption, this behavior is correct, which is determined by the LEACH. Therefore, It can be seen that the CRFDA algorithm is feasible in clustering route protocols.

In Fig. 6, these curves represent the amount of data will be sent by the faulty nodes after these nodes went wrong, for example, No. 24 node will send more than 13,000 bytes of data when the system run at 575th round. If these data are received by the cluster head, they will affect the cluster-head for data fusion operation and

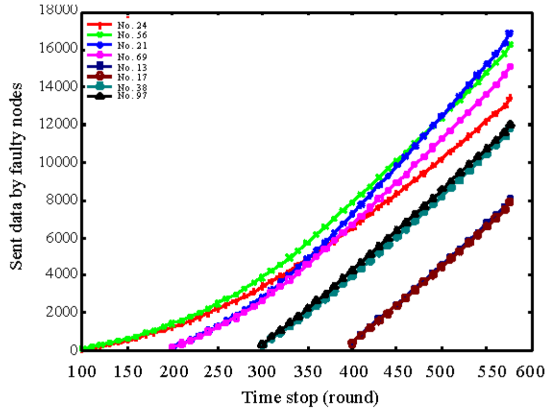

Fig. 6: These curves represent separately the additional amount of data, which are sent by these faulty nodes

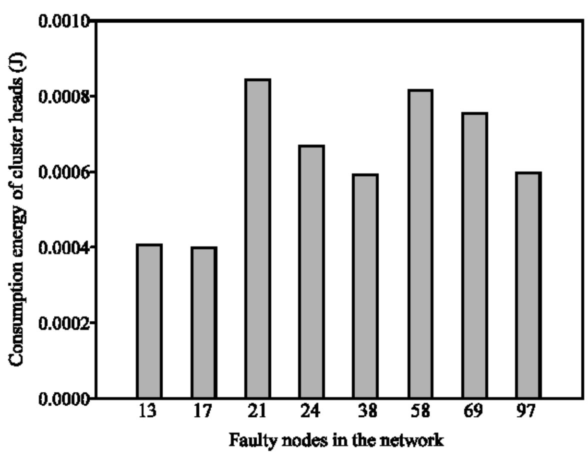

Fig. 7: The histogram represent that each faulty node has consumpted the total energy of the cluster nodes

lead to bias results of data aggregation, if the biased data are sent to the base station, which will seriously affect the application system for monitoring and analysis.

As can be seen from Fig. 7, when some faulty nodes come into being, if they do not be diagnosed and not be excluded from the system, they will consume additional energy of cluster heads, for example, No. 17 node will consume more than 0.0008 joule of energy of the cluster heads. Then, it will shorten the lifetime of the entire network.

Compared with the CSFD, first of all, In CRFDA algorithm, each cluster independently implement diagnosis, but for the whole network, they are parallel diagnosing, In this way, it greatly reduces the time spent diagnosing. Second, diagnosis task is activated by the cluster head which bears the responsibility of cluster diagnosis. Therefore, the diagnosis range significantly reduced without the diagnosis from node to node, The number of nodes used to send test messages greatly reduced, Obviously we can see, it greatly saves the energy of the network. Finally, the diagnosis work carried out by the cluster initiative, do not need to spread test message throughout the network and this improved network utilization and reduced the latency.

CONCLUSIONS

In this study, we have presented a Cluster-Based Real-time Fault Diagnosis aggregation algorithm for wireless sensor networks, called CRFDA that is based on the comparison approach in order to achieve a correct and complete diagnosis for hierarchical wireless sensor networks. Our algorithm is based on the network for clustering, it has the advantage of network diagnosis can be carried out simultaneously in all cluster. Compared with CSFD, it not only greatly simplified the complexity of the algorithm, reduced the detection time and improved the diagnostic efficiency ,but also significantly saved the energy consumption of network and prorogated the lifetime of the network. However, the most important feature of CRFDA is that it is a real-time diagnosis method and it can be spontaneously implemented in each round in the network, In this way, it effectively improved the accuracy of data fusion, ensured the authenticity and efficiency of the monitoring data from the network and provided the most valuable field data for the application system. By NS-2 for simulation analysis, our CRFDA algorithm is correct.

In future work, we will further study the real-time dynamic fault diagnosis, consider the feedback from the sensor node hardware, on this basis, we will conduct a new analysis to ensure effective investment by reducing the diagnosis error rate and improving the utilization of network resources.

ACKNOWLEDGMENTS

This study is supported by two grants from the National Natural Science Foundation of China (No. 60773190, No. 60802002)

REFERENCES

Akyildiz, I.F., W. Su, Y. Sankarasubramaniam and E. Cayirci, 2002. Wireless sensor networks: A survey. Comput. Networks, 38: 393-422.

Chen, J., S. Kher and A. Somani, 2006. Distributed fault detection of wireless sensor networks. Proceedings of the 2006 Workshop on Dependability Issues in Wireless Ad Hoc Networks and Sensor Networks, Sept. 26, Los Angeles, CA, USA., pp: 65-72.

Chessa, S. and P. Santi, 2001. Comparison-based system-level fault diagnosis in ad hoc networks. Proceedings of the 20th IEEE Symposium on Reliable Distributed Systems, Oct. 28-31, New Orleans, Louisiana, pp: 257-266.

Chessa, S. and P. Santi, 2002. Crash faults identification in wireless sensor networks. Comput. Commun., 25: 1273-1282.

Heinzelman, W.R., A. Chandrakasan and H. Balakrishnan, 2000. Energy-efficient communication protocol for wireless micro sensor networks. Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, (AHICSS'00), USA., pp: 3005-3014.

Heinzelman, W.B., A.P. Chandrakasan and H. Balakrishnan, 2002. An application-specific protocol architecture for wireless microsensor networks. IEEE Trans. Wireless Commun., 1: 660-670.

Intanagonwiwat, C., R. Govindan and D. Estrin, J. Heidemann and F. Silva, 2003. Directed diffusion for wireless sensor networking. IEEE/ACM Trans. Network, 11: 2-16.

Kuo-Feng, S., C.H. Chou, H.C. Jiau and W.T. Hu, 2006. Detection and diagnosis of data inconsistency failures in wireless sensor networks. Comput. Networks, 50: 1247-1260.

Lindsey, S. and C.S. Raghavendra, 2002. PEGASIS: Power-efficient gathering in sensor information systems. Proceedings of the IEEE Aerospace Conference, March 9-16, Los Angeles, CA, USA., pp: 3-1125-3-1130.

Luo, X., M. Dong and Y. Huang, 2006. On distributed fault-tolerant detection in wireless sensor networks. IEEE Trans. Comput., 55: 58-70.

Maeng, J. and M. Malek, 1981. A comparison connection assignment for self-diagnosis of multiprocessor systems. Proceedings of the 11th International Symposium Fault-Tolerant Computing, (ISFTC'81), USA., pp: 173-175.

Manjeshwar, A. and D.P. Agrawal, 2001. TEEN: A routing protocol for enhanced efficiency in wireless sensor networks. Proceedings of the 15th International Parallel and Distributed Processing Symposium, April 23-27, San Francisco, California, USA., pp: 30189-30189.

Pagani, E. and G.P. Rossi, 1997. Reliable broadcast in mobile multihop packet networks. Proceedings of the 3rdACM/IEEE International Conference on Mobile Computing and Networking, (ICMCN'97), New York, USA., pp: 34-42.

Paradis, L. and Q. Han, 2007. A survey of fault management in wireless sensor networks. J. Network Syst. Manage., 15: 171-190.

Ruiz, L.B., I.G. Siqueira, L. Oliveira, H.C. Wong, J.M. Nogueira and A.A.F. Loureiro, 2004. Fault management in event-driven wireless sensor networks. Proceedings of the 7th ACM international Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems, Oct. 04-06, Venice, Italy, pp: 149-156.

Samuel, M., J.F. Michael, M.H. Joseph and H. Wei, 2002. Tag: A tiny aggregation service for ad-hoc sensor networks. Proceedings of the 5th Symposium on Operating Systems Design and Implementation, Dec. 9-11, Boston, Massachusetts, pp: 131-146.

Sengupta, A. and A.T. Dahbura, 1992. On self-diagnosable multiprocessor systems: Diagnosis by the comparison approach. IEEE Trans. Comput., 41: 1386-1396.

Xue, Y., Y. Cui and K. Nahrstedt, 2005. Maximizing lifetime for data aggregation in wireless sensor networks. Mobile Networks Appli., 10: 853-864.

Yao, Y. and J. Gehrke, 2002. The cougar approach to in-network query processing in sensor networks. SIGMOD, 31: 9-18.

Younis, O. and S. Fahmy, 2004. HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Trans. Mobile Comput., 3: 366-379.