Ad Hoc Networks 25 (2015) 170–184

Contents lists available at ScienceDirect

Ad Hoc Networks

journal homepage: www.elsevier.com/locate/adhoc

Distributed self fault diagnosis algorithm for large scale wireless sensor networks using modified three sigma edit test

Meenakshi Panda *, P.M. Khilar

Department of CSE, National Institute of Technology Rourkela, 769008, India

ARTICLE INFO

Article history:
Received 6 May 2014
Received in revised form 6 October 2014
Accepted 8 October 2014
Available online 22 October 2014

Keywords:
Wireless sensor networks
Distributed self fault diagnosis algorithm
Soft fault
Hard fault
Three sigma edit test
Normal Gaussian distribution

ABSTRACT

Distributed self diagnosis is an important problem in wireless sensor networks (WSNs) where each sensor node needs to learn its own fault status. The classical methods for fault finding using mean, median, majority voting and hypothetical test based approaches are not suitable for large scale WSNs due to large deviation in inaccurate data transmission by different faulty sensor nodes. In this paper, a modified three sigma edit test based self fault diagnosis algorithm is proposed which diagnose both hard and soft faulty sensor nodes. The proposed distribute self fault diagnosis (DSFD) algorithm is simulated in NS3 and the performances are compared with the existing distributed fault detection algorithms. The simulation results show that the detection accuracy, false alarm rate and false positive rate performance of the DSFD algorithm is much better in adverse environment where the traditional methods fails to detect the fault. The other parameters such as detection latency, energy consumption and the network life time are also determined.

1. Introduction

Wireless sensor networks (WSNs) comprise of thousands of tiny sensor nodes having limited memory, less processing power, battery constraint is deployed in an environment such as underwater, underground for multimedia and terrestrial applications like navigation, environmental monitoring, agriculture, landslide monitoring and emergency rescue operations [1,43,33]. The sensor nodes are sensing and processing capability and interact with other sensor nodes using a wireless medium. The sensor nodes form an adhoc network without following any infrastructure. In fact, the communications between sensor nodes is a type of peer to peer communication.

Since the sensor nodes are deployed in hostile and human inaccessible environments, they are subjected to various kinds of faults. These faulty sensor nodes lead to produce erroneous results during their normal operation. In order to prevent the WSNs from the effect of faulty sensor node, a self fault diagnosis algorithm is proposed here to diagnose both hard [3,7,13] and soft faulty sensor nodes [6,17,42,2]. The fault free sensor nodes produce correct results during the network operation. The hard faulty sensor nodes do not respond, whereas the soft faulty sensor nodes respond with erroneous data skewed from the original data. The actual and erroneous data sensed by different sensor nodes is assumed to be a random variable which varies from one region to another region of the sensor network [18,34]. Distributed self fault diagnosis of WSNs has recently important due to their application in various sectors of society. The dependence on WSN encompasses to design and develop a robust fault diagnosis algorithm for WSNs, so that they can sustain for longer duration in the presence of faults.

The fault diagnosis techniques based on classical estimates like sample mean, variance, co-variance or

* Corresponding author. Tel.: +91 9439724318.
E-mail address: meenakshi.nitrkl@gmail.com (M. Panda).

http://dx.doi.org/10.1016/j.adhoc.2014.10.006
1570-8705/

correlations are adversely influenced by large deviation of data for a faulty sensor node [25,8,30]. In fact, these estimators are producing correct fault status when many sensor nodes are faulty within a particular region. Motivated by this, a modified three sigma edit test approach is adapted to diagnose the faulty sensor nodes present in WSNs. In the proposed approach, the performance of the diagnosis depends on neighboring node's data where each sensor node participates in the fault diagnosis process to identify itself as faulty (hard or soft) or fault free. The accuracy in finding the status of all the sensor nodes depends on the number of neighboring nodes. The proposed diagnosis algorithm performs better when more number of neighboring nodes are likely to be faulty.

It has been seen from the literature that the existing method leads to a large number of message exchanges over the network for data and status exchange which rapidly depletes the energy of the sensor nodes. It really puts a large overhead for the large scale WSNs. Due to poor performance and high energy overhead of the existing set of approaches, it is necessary to design and develop an efficient fault diagnosis algorithm for large scale WSNs.

This paper has following contributions: (i) Modified three sigma edit test based fault diagnosis algorithm is proposed to diagnose the faulty sensor nodes present in WSNs. (ii) The proposed method is compared with traditional comparison model and three sigma edit test. (iii) A distributed self-fault diagnosis (DSFD) algorithm where each sensor node diagnoses itself with high detection accuracy, low false alarm rate and false positive rate. (iv) Evaluation of the DSFD algorithm using standard simulator NS3 and compare the performance with the existing works in the literature given by Panda and Khilar [29], Chen et al. [6] and Jiang [17].

The remaining part of the paper is organized as follows. In Section 2, the related work which provides an exhaustive survey about the previous work is discussed. The network, fault and radio model are discussed in Section 3. The proposed DSFD algorithm is described in Section 4. The analysis of the algorithm along with modified three sigma edit test method is discussed in Section 5. The simulation results with discussions are given in Section 6. Finally, Section 7 concludes the paper.

2. Related works

The system and node level fault diagnosis has been traditionally known as a system level diagnosis after the work of the researchers Preparata et al. in [32]. Since then, all the system level diagnosis algorithms are feasible for multi-computer and multi-processor systems. A generalized theory for system level fault diagnosis is presented in [39]. The necessary and sufficient conditions are provided for any fault-pattern of any size to be uniquely diagnosable, under the symmetric, and asymmetric invalidation models with or without the intermittent faults. In [38], the authors have proposed distributed algorithms to diagnose faulty processors for regular interconnected structures in the presence of a large fault set. The algorithm is based on near neighbor communication and does not require broadcasting. The concept of multiple fault detection in multi-processors system is first introduced in [39,38]. These algorithms are not suitable for WSNs due to high energy overhead.

The fault diagnosis algorithms available in literature are classified into three types such as centralized, distributed diagnosis and self-diagnosis. In centralized approach [20,27,28], each sensor node sends the data to the central node which is usually sink node or base station. The central node identifies the fault status of each node by analyzing the received data and then disseminates the fault status to all nodes in the network.

In centralized diagnosis approach [20], an ultra-reliable central node with high computation capability and large storage is used. The other shortcomings of the centralized fault diagnosis algorithms are

1. For keeping status information of N sensor nodes deployed in the region of interest, the central node needs minimum $N(1 + \log_2 N + C)$ bits of memory. Where C bits are required for keeping the sensed data of the sensor node, one bit is required for keeping the fault status as a binary decision (faulty or good) and $\log_2 N$ bits are required for keeping the sensor node's identifier.
2. For transmitting the data to the central node, each sensor node needs multi hop communication as they are far away from the central node which depletes energy of the network quickly, especially the sensor nodes nearer to a central node.
3. The actual status of sensor node may change while central node such as sink node or base station acquire the status of the entire network in real time.
4. Here, all the sensor nodes (faulty or fault free) send their sensed data to the base station as they treat themselves as fault free before diagnosis. By doing this, the intermediate node depletes by transmitting faulty node's data unnecessary.
5. The detection latency is high as it consumes time to acquire data from all the sensor node using multi hop communication.
6. If the central node becomes faulty, it is difficult to find the status of all sensor nodes in the network.

Looking for the above disadvantages of the centralized approach, in literature the distributed fault diagnosis algorithms in WSNs [40,10,4] have been proposed where each sensor node participate in the diagnosis process but the final fault status is decided by the central node. Every sensor node acquires the data or output of giving task from the neighboring sensor node and find their probable fault status by adapting the neighbor coordination, comparison, or task based approaches as discussed in [7,41,35]. In these approaches, each sensor node acts as a tester as well as testing node. Each sensor node also tests its neighbors status and collects all test results known as a syndrome. Then, send the resultant syndrome to the central node to identify the list of faulty and fault free sensor nodes by adapting evolutionary algorithm or syndrome analysis approach as discussed in [26,9]. After identifying list of faulty and fault free sensor nodes, the central unit sends the status to all