Computers and Electrical Engineering 48 (2015) 270–285

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

Distributed Byzantine fault detection technique in wireless sensor networks based on hypothesis testing ☆

Meenakshi Panda *, P.M. Khilar

Department of Computer Science and Engineering, National Institute of Technology Rourkela, 769008, India

ARTICLE INFO

Article history:
Received 2 July 2013
Received in revised form 22 June 2015
Accepted 22 June 2015
Available online 18 July 2015

Keywords:
Wireless sensor networks
Distributed algorithm
Soft fault detection algorithm
Byzantine fault
Neyman–Pearson detector

ABSTRACT

Distributed fault detection in wireless sensor network is an important problem where every sensor node identifies its own fault status based on the information from its neighboring sensor nodes. This paper presents a novel distributed fault detection algorithm to detect the soft faulty sensor nodes in sparse wireless sensor networks. In the proposed scheme, every sensor node gathers the information only from their neighboring nodes in order to reduce the communication overhead. The Neyman–Pearson testing method is used to predict the fault status of each sensor node and the neighboring sensor nodes. A voting scheme is applied on the fault status information to obtain the final fault status of each sensor node. The generic parameters such as detection accuracy, false alarm rate, time complexity, message complexity, detection latency, network life time and energy consumption are considered to evaluate the performance of proposed scheme analytically as well as through simulation. The result shows that the proposed scheme significantly improves the performance over the existing algorithms.

1. Introduction

Wireless sensor networks (WSNs) comprise of a large number of small sensing and self-powered sensor nodes distributed in a geographical region. The sensor nodes gather information or detect special events and communicate in a wireless fashion [1]. Sensing, processing and communication are three key tasks whose combination in one tiny device gives rise to a vast number of remote sensing applications. Although WSNs provide endless opportunities, at the same time pose formidable challenges. Some of these challenges are low battery, less computational capabilities and inefficient use of communication resources. Among these impediments, the most difficult one is the mysterious data sent by an unknown faulty sensor node either to the fusion center (FC) such as base station (BS) or to the neighboring sensor node [2]. In WSNs, the accuracy of the observed data sent by a sensor node is important for the overall network's performance. Therefore, detection of faulty sensor nodes is an essential issue in WSNs [3].

A sensor node is said to be faulty if it is not functioning properly [4]. In the literature, the faults in WSNs are broadly classified into two types known as hard fault (permanent or static fault) and soft fault (or dynamic fault) [5,6]. The hard fault occurs if a sensor node fails to communicate with the rest of the sensor nodes in the network [7]. When the sensor node is able to communicate with the other sensor nodes, but transmits erroneous message, then such type of fault is known as soft

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Associate Editor Dr. Catalin Lacatus.

* Corresponding author.

http://dx.doi.org/10.1016/j.compeleceng.2015.06.024
0045-7906/

fault [8]. The soft faulty sensor nodes behave arbitrarily (i.e. Byzantine fault) and it is difficult to predict [2,9]. Therefore, the proposed work considers the Byzantine behavior of the sensor nodes such as stuck at zero, stuck at one, and random fault.

In the literature, both centralized and distributed soft fault detection methods are available [10–12]. The centralized algorithms required more communication, because all the sensor nodes send the information to the fusion center for fault detection. Where as in distributed case, each sensor node accumulates the data from the neighbors and detects its own fault status which reduces the communication. It is evident from the operation of WSNs that, the communication requires much higher energy than computation [12]. Therefore, distributed based approach is proposed here to detect the faulty sensor node in WSNs.

In the proposed work, to minimize the computational complexity and improve the accuracy, each sensor node initially tests the presence of faulty sensor nodes in its neighbor, if found, then predicts the probable fault status of them. For this, the Neyman–Pearson (NP) detection method is used. The sensor nodes shared the predicted probable fault status of the neighbors with them. Then, each sensor node uses a fusion scheme to take the final decision on its fault status.

The major contributions of this paper are (i) design and evaluation of an efficient distributed fault diagnosis algorithm for detecting soft faulty sensor nodes in large WSNs, (ii) the Neyman–Pearson (NP) detection method is used to detect the faulty sensor node (iii) the performance is compared with the existing distributed algorithms such as JSA [10] and Jiang [11], and (iv) the algorithms are implemented in NS3 [13].

The remaining part of the paper is organized as follows. In Section 2, the related work which provides an exhaustive view about the previous work is discussed. The network model used for the development of the distributed algorithm is provided in Section 3. The proposed distributed fault detection (DFD) algorithm is described in Section 4. The analytical model which proves the correctness of the algorithm is also given in this section. The simulation results are provided in Section 5. Finally Section 6, concludes the paper with discussion.

2. Related work

In this section, the related work on the soft fault detection in WSNs is briefly discussed. The basic emphasis is given to the distributed fault detection techniques based on comparison model.

Since long back many researchers have proposed different algorithms to find hard fault in processor [14,15]. However, these algorithms may not work to find the faults in WSNs, because the processor are not energy constrained whereas energy is the major constraint in WSNs. Lee and Choi [9] have introduced a fault detection algorithm in which each sensor node performs local comparison between own sensed data with its neighbor's data at a particular time instant and stores it locally in a table. This process is repeated for some constant c times (say) and at each time comparison results are stored in the table. Then, each sensor node identifies own fault status by analyzing the data stored in the table.

Similarly, in [10–12,16], the comparison based soft fault detection algorithms are discussed. In those approaches, each sensor node compares its own sensed data with the received data from the neighbors and shares the results with them unlike keeping in a local table [9]. Then, based on majority voting in the neighborhood, each sensor is tagged with a name likely fault free or likely faulty. Each likely fault free sensor node has identified as fault free sensor node by some rigid criteria. Finally, the remaining likely fault free or likely faulty sensor nodes are determined to be fault free or faulty with the help of the known fault free sensor nodes or its own tendency value respectively. The major disadvantage of these distributed approaches is high message complexity of cN, where N is the number of sensor nodes in WSNs and c is the number of times each sensor node shares message with the neighbors. Further, all the above comparison based fault detection algorithms depend on a threshold value. The appropriate choice of the threshold values is an important problem because, the accuracy of the algorithm depends on the threshold value.

In [17], Luo et al. presented the Bayesian framework for finding the faulty sensor nodes in WSNs. Here, each sensor node sends a query to the sink node to know how much noise is present in its own area. The noise calculation is based on the Bayesian and the Neyman Pearson performance criteria. After knowing own region's noise, each sensor node estimates the same for the neighbors. This scheme needs multi hop communication for estimating its noise which may change during the diagnosis period. Further, multihop communication adds high traffic in the network, which consumes more energy and make the algorithm energy inefficient.

Liang et al. in [18] introduced a new distributed fault detection scheme based on the comparison of its own data with the median of its neighbor's sensed data. Each sensor node arranges the neighbor's data in ascending order to calculate the median. If the comparison result between the median and own sensed data exceeds a threshold, then the sensor node is identified as faulty otherwise fault free. The algorithm needs more time to find the median and the accuracy of the algorithm depends on the threshold.

In [2], a Byzantine fault detection method has been proposed, where each sensor node sends a set of messages to a group of sensor nodes and also receives messages from the same group. If the number of messages sent is equal to the number of messages received, then the sensor node is identified as fault free otherwise faulty. This approach needs multi hop communication and requires coordination among the sensor nodes to identify the faulty one. Recently, Ssu et al. [19] presented a