Journal of Innovative Engineering and Research (JIER)
Vol.- 6, Issue -1, April 2023, pp. 26-30 (5 pages)

A Study on Fault Detection Algorithm in WSN

MANIDIPA ACHARJYA[1] AND DR. RAHUL MISHRA[2]

Department of Electronics & Communication Engineering, Dr. A. P. J. Abdul Kalam University, Indore (M.P.), 452010- India
Corresponding Author Email: manidipa.976@gmail.com

Abstract: A wireless sensor network (WSN) is a network of distributed functionally autonomous sensors that are used to monitor physical conditions of the environment or a geographical location, such as pressure, sound, temperature, etc. and to transfer the sensed data by multiple hops through the network to a central location or sink node. The networks used nowadays are bi-directional, which enables the controlling of sensor activities. The development of WSN was initially motivated due to applicability in military applications like battlefield surveillance but now-a-days such networks are used in various fields ranging from various industrial and consumer applications to industrial processes monitoring to health monitoring at hospitals and so on.

I. INTRODUCTION

The WSN is built of sensors or nodes which range from a few in number to several thousand, wherein each node is connected to one or sometimes several other sensors within their transmitting range. Each such WSN node has typically several parts: a radio transceiver with an antenna that is responsible for data transmission, a microcontroller used as an electronic circuit for sensor interfacing and an energy source, usually a battery since they are deployed in unreachable conditions or an embedded form of energy harvesting. A sensor node might vary in size from that of a box down to the size of a dust particle, although functioning "motes" of genuine microscopic dimensions are yet to be developed. The cost of nodes also similarly varies, ranging from a few pennies to hundreds and thousands of dollars, depending upon the complexity of the architecture of sensor nodes. Size and cost constraints on sensor nodes result in corresponding constraints on resources such as memory, computational speed, energy and communication bandwidth. The topology of the WSNs can vary according to functionality from a simple star network to an advanced multi-hop wireless mesh network. The propagation technique between hops of a network can be by routing or flooding.

Objective of the Study:

1. To describe and determine a fault detection algorithm
2. To evaluate the performance of proposed algorithm using the given parameters
   - Threshold values
   - Connectivity
   - Time redundancy
3. To compare the proposed algorithm with the existing algorithm

Structure of A WSN

The nodes in a WSN can be connected in a Mesh Topology, Peer to Peer, Tree or Star topology depending upon the requirements. The size of a node is as small as a coin. The wireless sensor nodes do not communicate with a central node directly but rather through their surrounding local nodes. There are many constraints associated in designing of a node. The nodes have an embedded processor which has to implement complex networking protocols with a memory of few kilobytes.

As the size of the device grows smaller and so does the power source. The nodes include both hardware and an operating system such as TinyOS (a widely used operating system for WSN). The most advanced hardware platform used is single chip CMOS device.

Clustering in WSN

In this type of topology the nodes form clusters with their neighbouring nodes and communicate with each other in order to support data aggregation through efficient network organization. Here, each cluster has a co-ordinator, reffered to as cluster head and a number of member nodes.

Characteristics:

The main characteristics of a WSN are:

•**Power consumption**: it is always a major constraint in the deployment of sensor nodes since they are deployed in inaccessible areas and need function accurately. So the only source of power to them is through battery which is obviously non-rechargeable.

•**Heterogeneity** : mostly the type of nodes deployed differ by the type of operation, and hence homogeneity or heterogeneity also depends upon the type of operations and area in which they are deployed

•**Scalability**: usually the performance of the network depends upon energy levels of the nodes, apart from that the network performance doesn't depend upon the size of the network

•**Nodal mobility**: if the nodes are mobile, then they need different algoruthms for functioning

•**Communication failures**: occour due to lack of battery power and if a node goes out of transmission range

•Ability to withstand harsh environmental conditions: should be able to operate in harsh conditions

Basic Architecture of WSN Nodes

RAM

The RAM in a sensor node is used to store the current readings sensed by the node. The data may be hampered in case there is power supply disruption.

ROM
The ROM is used to store the programs used in implementing the WSN, including data aggregation, fault detection and routing.
Transmitters
The transmitters are used in half-duplex mode for both receiving and sending operations. It has 4 states:
_ RECEIVE
_ TRANSMIT
_ IDLE
_ SLEEP
Power Supply
The basic objective here is to provide maximum power in minimum size. As they are deployed in human inaccessible conditions, batteries cannot be charged in their usual way, energy has to be obtained from other sources like photo voltaic cells, temperature gradient etc.

WSN Applications:
1: military applications
Military services now require a high quality effectively functioning WSN foe communication purposes, to identify the friendly forces set targets, survey the battle field, detect arms and ammunition and presence of weapons of all kinds like nuclear or biological.
2: Environmental Applications
It is mainly used in weather forecasting, especially used for agricultural purposes also to detect humidity, rain and temperature. It can also be deployed in forests to detect oforest fires or underneath rivers to detect floods.it can also be used to determine the contamination levels in air and water.
3: Health Applications
They are mainly used for surveillance systems in the hospitals to track down patients and doctors but also can be used in drug administration process.
4: Automotive Applications
Can be used to reduce wiring in automobiles and increase user-friendliness. It is also used in monitoring technical components and make devices automatic.

Sensor Node Components
- Sensing Unit
- Power Unit
- Transceiver Unit
- Processing Unit

Protocol Used
Zigbee: ZigBee is a type of IEEE 802.15 standard used for routing. Although being low- powered, ZigBee devices can transmit data over long distances by passing data through intermediate devices to reach more distant ones, creating a mesh type of network; i.e., a network with no centralized node to control others or high-power transeiver able to connect all of the networked devices directly. The decentralized nature of these wireless ad hoc networks make them suitable for applications where a central node can't be relied upon.
Zig Bee is used in wireless applications that require a low data rate, long battery life, and secure networking. Zig Bee has a defined rate of 250 Kbit/s, that is suits best for periodic or intermittent data or a single signal transmission from a sensor or input device.
Faults
Definition: Faults are a kind of a unsatisfactory feature or attribute that leads to errors in the system.
Types of Faults
Based on persistence, faults are categorized into the following types:
Soft faults: These faults occur very less frequently or rarely and gets removed without any external intervention. These are mostly caused due to noise.
Permanent Faults: The permanent fault is permanent for a node. The intermittently faulty node gradually becomes a permanent faulty node in due course of time. They are caused mainly due to power failure in the node.
Causes of Faults
Due to the fragile nature of sensor nodes and also because of the depletion of their limited power source, faults may occur. Due to harsh environments where nodes are being deployed, the nodes may receive and transmit incorrect sensor readings. In WSN, the links are also prone to faults. Also when nodes are embedded or mobile nodes sometimes go out of range of communication. Faults are also caused due to multi-hop communication as it takes several hops to deliver the data to sink. Failure of single intermediate node may lead to a total erroneous data being collected.
Congestion which occurs due to large number of nodes transmitting the same time which may also lead to packet loss.

Fault Event Disambiguation System
Here, the event region is presented to be a zone or area with a cluster having a different reding from the other nodes of WSN. They may report an unusual reading even though it may belong to an event region. Here, we present probabilistic decoding mechanisms with each cluster having a different area code that exploits the fact that sensor faults are likely to be stochastically uncorrelated, while event measurements are likely to be correlatedto their neighbours under same event zone. In analysing these schemes, it is shown that the impact of faults can be reduced greatly, even for reasonably high fault rates. We assume environments in which event readings are typically spread out geographically over a large area of contiguously deployed sensors. In such a scenario, we disambiguate faults from events by examining the correlation in the reading of nearby sensors of the same event zone.

II. LITERATURE REVIEW

Remote sensor systems are presently-a-days developing as stages for observing different natural conditions including remote geological districts, office structures, and mechanical plants . They are made out of an expansive number of little sensor hubs furnished with constrained processing and communication competencies. Since minimal effort-sensor hubs are frequently conveyed in an uncontrolled and harsh geographical domain, are inclined to have flaws. It is accordingly implicit to discover, find

the broken sensor hubs, and avoid them from the system throughout typical operation unless they could be utilized as communication sensors devoid of senor functionality.
The execution of the localized diagnosis, be as that may, is limited because of the random nature of degree/connectivity of the deployed sensors. Fault-tolerating event detection algorithm is also additionally been proposed in[1]. Luo et al have proposed a fault-tolerant energy-efficient event detection algorithm for WSN[3]. For a given detection error bound, minimum neighbours are selected to minimize the communication volume in the network[1].

ERROR DETECTION

An algorithm for detecting and logically isolating faulty sensor nodes in WSN has been proposed. The fault-free sensors employ local comparisons of sensed data between neighbours and dissemination of the test results to enhance the accuracy and correctness of diagnosis. Transient faults in sensor reading are tolerated by using time redundancy in the algorithm. Both the network connectivity and accuracy of diagnosis are taken into account to set the threshold values. Here, nodes with malfunctioning sensors are allowed to act as communication nodes for routing purposes, but they are logically isolated from the network as far as sensing or fault detection is concerned.

FAULT EVENT DIASMBIGUATION

The first step in this process is for the nodes to determine the zone and readings are of interest. In general, the readings are thought of as a real number. There is some prior work on systems that understand the normal conditions or readings over time so that they recognize unusual eventful readings. Instead, a reasonable assumption can be made that a threshold that enables them to determine whether their reading corresponds to an event has been specified with the query or otherwise made available to the nodes during deployment.[7]
A more challenging task is to disambiguate events from faults in the sensor measurements since an unusually high reading could potentially correspond to both a node being faulty or a node being in an event region. Conversely,it may also be concluded that a faulty node reports a low measurement value even though it is in an event region[8]. At this point, it is considered to be faulty. In this paper, the whole network is divided into small clusters of event zones. And nodes within that cluster are considered as neighbours for error-detection process. As e result,accuracy of the system is preserved as well as the faulty nodes are identified in an accurate way. It is believed that the event nodes are spatially correlated whereas the erroneous nodes are not so.

IMPLEMENTATION METHODOLOGY

In this section, we will go through the fault detection algorithm in detail. I studied the algorithm and picked out the weaknesses that could be modified. It is seen that the algorithm given in the paper, works out well if connectivity of the nodes is maintained constant. But in practical scenarios, with the passage of time connectivity changes constantly as the faulty nodes are removed from the network. The proposed algorithm calculates the connectivity of each node and calculates the number of faulty connected nodes as a percentage of total connectivity and hence prevents the detection accuracy being reduced with time. The nodes with permanent faults are removed from the network whereas the nodes with transient faults are kept if they report erroneous readings for a minimal time below a certain threshold. Else they are treated as intermittent faulty and removed from the network.
Now, we extent this work to an event disambiguation system, by dividing the network into clusters of zones which are expected to have some kind of reading by introducing another variable to store the event region of the node. In case the node is present in the event region, and then it is compared to its neighbouring nodes present in that event region and so on.

NETWORK MODEL

It is assumed that in the WSN, the sensor nodes are haphazardly deployed in the area under surveillance. The area is very dense and all the sensors are considered to have a common range of transmission. And the sensor nodes located in this range of transmission of a particular node is called its neighbours. There might be a fault occurring in any of these sensor nodes at a particular instant of time. Nodes with permanent faults like power failure or faulty sensors are to be identified.
Sensor nodes which generate incorrect sensing data or fail in communication intermittently are treated as usable nodes, and thus are diagnosed as fault-free unless they report erroneous readings very often. Sensor nodes with malfunctioning sensors could participate in the network operation since they are still capable of routing information. Only those sensor nodes with a permanent fault in communication (including lack of power) are to be removed from the network.

DATA MODEL

It is assumed that for each fault-free node, its neighbouring fault-free nodes should to have similar measurement values if they belong to the same zone or event, due to proximity. Let two nodes vi and vj be neighbours of each other and xi denote the measured value at node vi. Then the condition to be satisfied by vi and vj is that $|xi - xj| < \delta$, where δ is the threshold of accuracy in readings which may vary depending on the application field in which it is used. In the case of temperature, for example, a sensor node and its neighbours are expected to have almost same temperatures because of their geographic proximity. Hence δ is expected to be a very small number that differentiates the faulty nodes from fault-free nodes . If the reading is binary in the system then, the value of threshold, δ(limit of tolerance in a binary reading) must be set to 0.

COMMUNICATION MODEL

In the proposed system communication is between only

the neighbouring sensor nodes. A sensor node is considered as a neighbour of another sensor node if it lies within the transmission range of the first sensor. So, the set of nodes that a node n can communicate is given as:
E(s)={ni such that ni lies within the transmission range of n}
We have assumed a full duplex mode of communication, so the communication is bidirectional between the nodes. So, if there is a communication link between ni and nj there is also a link between nj and ni.

FAULT RECOGNITION

Let the real scenario at the sensor node be depicted by a binary variable Ti that stores the event code of that region.in case of a two region topology, the total area is divided into two regios, I,e., the event region and the normal region. The variable Ti = 0 if it is assumed that the node belongsto a normal region and Ti = 1 if it is assumed that the node belongs to an event region. We check the real output of the sensor into an abstract variable taken as Si. This variable Si = 0 if the sensor measurement generates a normal value and Si =1 if it reads an unusual value. In this simplest case, there are thus four possible outcomes: Si = 0; Ti = 0 that occurs when sensor belongs to a normal region and correctly reports a normal reading, Si =0; Ti = 1 when sensor faultily generates a normal reading, Si = 1; Ti = 1 when sensor correctly reports an event reading or is in an event region and is fault-free, and Si = 1; Ti = 0 occurs when sensor faultily reports an unusual reading i.e., it is in a normal region but gives an abnormal reading. While each node is aware of its value of Si, in the presence of a significant probability can occur when in a faulty reading, it can happens to be that Si ≠ Ti.

III, RESULTS AND ANALYSIS

The initial algorithm was implemented in a 100 X 100 unit network topology and the detection accuracy & False Alarm Rate were determined. The shortcomings of the algorithm were identified and overcome by using clustering in the network.

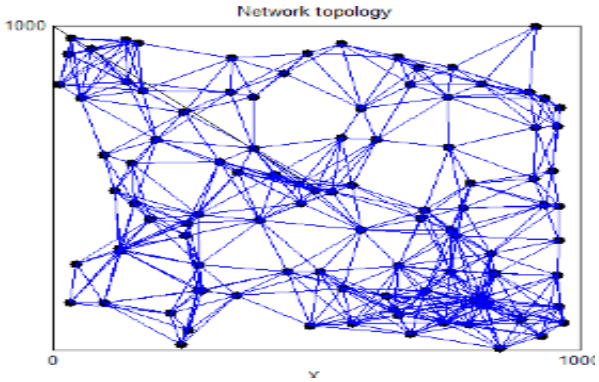


Fig 1: network topology for a network size 100

Fig 2: detection accuracy and false alarm rate for the initial algorithm

Detection accuracy for percentage of fault probability, where θ1=0.75, θ2=0.8,δ=1, no of reading taken=3, connectivity=10
This shows that there is a gradual decrease in the detection accuracy and increase in the false alarm rate(FAR) because of increase in arbitrary sensor readings that come under the threshold(mainly in case of permanent faulty nodes where reading is 0).

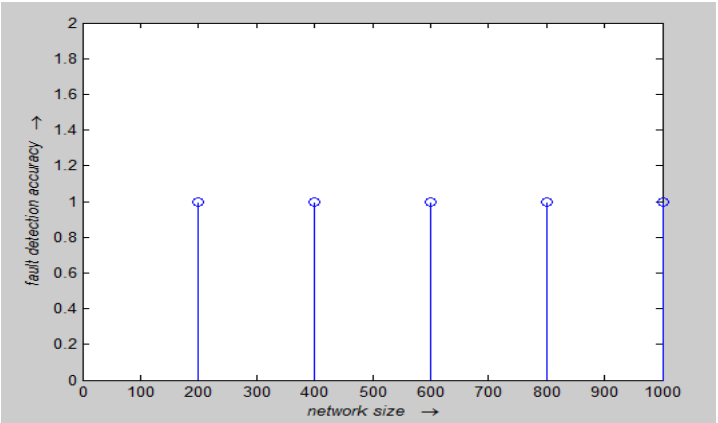


Fig 3: scalability of the algorithm

Detection accuracy for varying node size with %of faulty nodes=20%, θ1=0.75, θ2=0.8, δ=1, no of reading taken=3, connectivity=10
From the above diagram we conclude that algorithm is scalable i.e, it is applicable for networks of all sizes without the loss of functionality or generality.

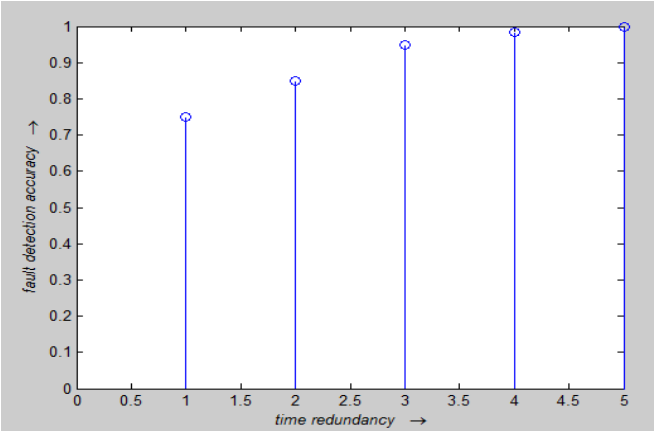


Fig 4: effect of time redundancy on the accuracy of the algorithm

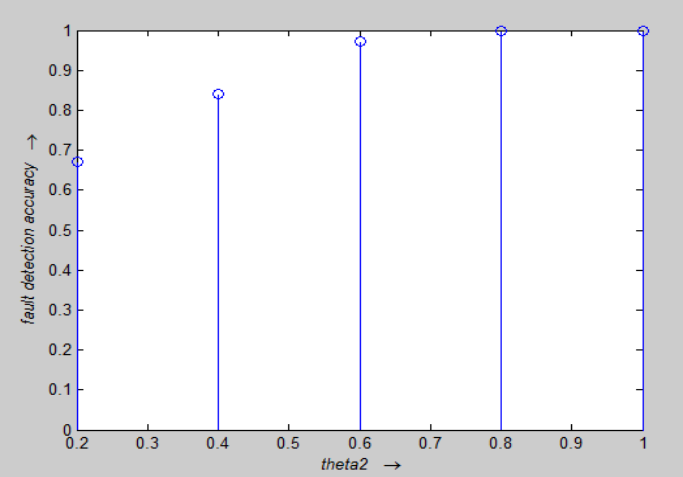


Fig 5: effect of parameter θ2 on detection accuracy

Detection accuracy for time redundancy with %of faulty nodes=20%, θ1=0.75, θ2=0.8 network size=100, connectivity=10. Fault detection accuracy increases with number of readings taken since transient faults are tolerated by increasing the time redundancy. But, increasing time redundancy beyond a certain limit makes no difference.

Detection accuracy for varying values of θ2with %of faulty nodes=20%, θ1=0.75, network size=100, readings taken=3connectivity=10. As θ2 increases the effect of transient faults on accuracy is neglected and hence the detection accuracy increases of the algorithm increases.

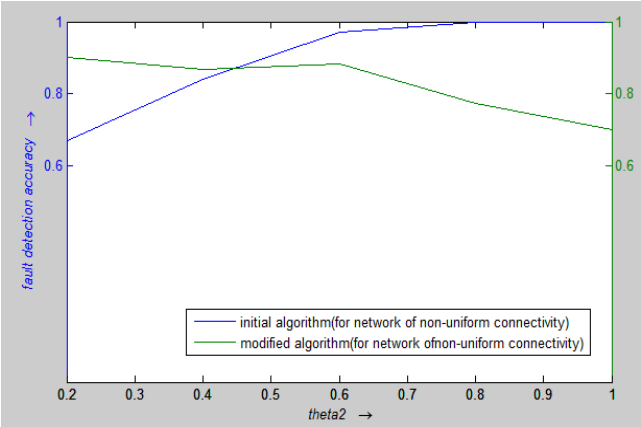


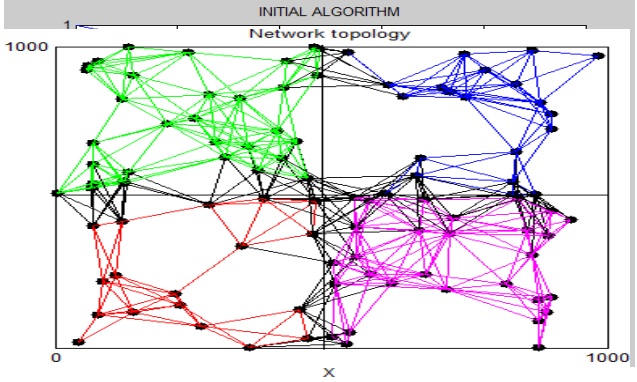
Fig 6: detection accuracies of the old and the modified algorithm prior to applying event disambiguation model

Detection accuracy for the two algorithms for non-uniform connectivity with %of faulty nodes=20%, θ1=0.75, network size=100

Here theta2 is the ratio of faulty nodes detected and total connectivity of a node. Initially the fault detection accuracy is low due to low value of theta2 which should increase with it. But, in the latter case, detection accuracy reduces due to un-uniform connectivity which is prevented using the modified algorithm.

Fig 7: network topology after clustering

Here, the WSN is shown as a cluster of 100 nodes which are distributed over an area of 1000X1000 units. Since the area is very large and the readings of nodes from in on extreme may vary from another extreme and may be



considered faulty, we divide the nodes into clusters. And the nodes in this cluster execute the algorithm independently. As a result, the threshold(δ) can be kept low as well as gradient of readings can be developed.

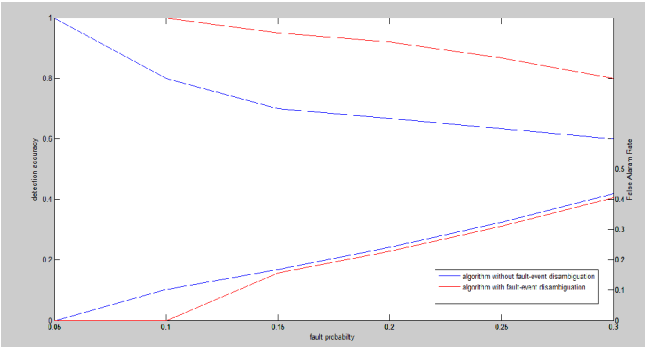


Fig 8: FAR and DA of initial and modified algorithms after applying clustering

In this figure, the blue line shows the fault detection accuracy and false alarm rate of the initial algorithm

So, we conclude that clustering improves the performance of the WSN by taking into consideration different values under consideration.

IV. CONCLUSION

Here, we have proposed an algorithm for fault detection and event disambiguation in WSN that takes into account the eventful areas and normal areas under two different categories. It accordingly forms two clusters and compares the readings of the nodes within their group. From the simulation, we have shown that initially, due to taking number of nodes as parameter, the detection accuracy of the network reduced with time because of un-uniform connectivity. In the proposed algorithm, we have overcome this problem by taking fault tolerance as a measure of percentage of total connectivity. We then have introduced another variable called fault disambiguation variable. This confirms that nodes of same zone form a cluster and consider readings of the neighbors hence formed. This algorithm is useful for fault detection when an event occurs because otherwise the data would be considered as faulty.

REFERENCES

1. S. Jia, "Fault Detection Modelling and Analysis in a Wireless Sensor Network," Journal of Sensors, Hindwai, Volume 2018 | Article ID 7935802.