Comput Syst Sci & Eng (2020) 5: 335–345

International Journal of
Computer Systems
Science & Engineering

A Frame Work for Categorise the Innumerable Vulnerable Nodes in Mobile Adhoc Network

Dr. Gundala Swathi*

Vit University, Vellore, 632014, India

Researches in wireless mobile ad hoc networks have an inherent challenge of vulnerable diagnosis due to the diverse behaviour pattern of the vulnerable nodes causing heterogeneous vtype1, vtype2, vtupe3 and vtype4 faults. This paper proposes a protocol for the diagnosis of vulnerability nodes with three-phases of clustering, vulnerable detection and vulnerable fault classification in wireless networks. This protocol employs the technique of probabilistic neural network for classification of vulnerable nodes and detects vulnerable nodes through timeout mechanism and vtype3, vtype4, vtype1, vtype2 nodes through the method of analysis variance. Network simulator NS-2.3.35 is employed for performance evaluation of the protocol.

Keywords: heterogeneous faults, clustering, probabilistic neural network, time out mechanism

1. INTRODUCTION

There is a dire need for designing a robust wireless mobile ad hoc network for gathering, processing and monitoring data from diverse sources in the general context of prevalent node failure and the consequent vulnerable information due to natural phenomenon such as lightning, earthquakes or sudden environmental changes, or other reasons such as vibrations, issues of coverage area, battery depletion and mote malfunctioning. Consequent these Errors in mobile nodes, erroneous results could be projected, drastically affecting the tasks of mobile network computations causing network failure. Hence, detecting the Vulnerable nodes, separating the vulnerability and vulnerable free nodes with proper vulnerable node classification become a priority.

Network faults are classified as hard and soft on the basis of the behaviour of faulty nodes [3,4] and as permanent, transient and intermittent due to persistence [5]. The nodes affected with hard faults fail to communicate with the other nodes in their network range and are also known as permanent hard faulty nodes [6]. The nodes with soft permanent faults give modified results in every individual time duration. The intermittent and transient faults are known as temporary faults and are very difficult to identify, as their effects are dynamic on the network. Generally external noise or errors cause these transient faults. These transient and intermittent faulty nodes demonstrate altered behaviour for some spike time or small time durations and normal behaviour in the other durations.

The researchers in general focus exclusively on isolated errors in the network, while the current context requires heterogeneous and comprehensive handling of Errors in networks. This motivates the proposal of a vulnerable diagnosis protocol capable of simultaneously handling heterogeneous errors. The following are the principal contributions of the paper:

a) Proposal of a heterogeneous wireless network protocol to diagnose heterogeneous faults

b) Employment of a multi-channel MAC protocol for inter-node communication within the network to reduce instances of end-to-end delay and packet-drop.

c) Use of load-balanced method of clustering to minimize consumption of energy.

d) Implementation of a status register mechanism based on timeout to detect vtyp1 and vtype2 network nodes.

*gundalaswathi@vit.ac.in

e) Employing a statistical method of analysis variance for detection of vtype3, and vtype4 nodes within the network.

f) Use of fault classification methods to identify categories of faults on the basis of probabilistic neural network (PNN) and the forwarded feedback.

g) Utilizing the network simulator, NS 2.35 for the performance evaluation of the proposed protocol for fault diagnosis.

The following is the organization of the paper: Section 1 gives the general introduction and motivation along with tracing the other contributions. Section 2 briefly discusses the related works, while Section 3 illustrates the proposed protocol and its system model. The proposed protocol and methodology for fault diagnosis are detailed in Section 4. Section 5 deals with the evaluation through simulations and the results of the testbed experiment pertaining to the considered protocol. Section 6 considers conclusion along with future research perspectives.

2. RELATED WORK

Several researches in fault diagnosis protocols in wireless networks have been briefly considered here.

A protocol of distributed fault diagnosis has been advocated by Panda and Khilar [9] deploying the testing method of using Neyman Pearson for the fault status prediction of the mobile nodes. This method effectively detects random and the byzantine network faults such as struck at zero, one and non-zero. They further proposed a protocol of self fault diagnosis for huge networks through the method of three stigma edit test, which diagnoses networks hard and soft permanent faulty nodes. Panigrahi et al. [10] have put forward a distributed algorithm for fault tolerant estimation in wireless networks deploying the method of diffusion adaptation for the permanent faulty nodes in the network both hard and soft. Several research proposals on fault diagnosis protocols depend on the degree of density of the neighbouring nodes within the range of communication. The protocols perform well in denser topology of higher degree and decreases in sparse density topologies. Mahapatro and Khilar [11] put forth an online distributed algorithm based on clusters for fault diagnosis of soft and hard permanent faulty nodes in networks. A similar distributed fault detection algorithm for wireless networks is proposed by Sahoo and Khilar [12] through comparison of timeout mechanism and the neighbour nodes for identification of intermittent and permanent faulty network nodes. An approach based on majority neighbours voting has been propounded by Chen et al. for distributed localized fault detection for WSN to detect soft permanent faulty nodes present in the network. These algorithms are further extended by Xianghua Xu et al. [14], who proposed a similar distributed approach based on local comparisons for the diagnosis of localized fault detection for the intermittent and soft permanent network faulty nodes. The proposed fault detection algorithm of Lee et al. [15] uses the method of comparisons of neighbour mobile nodes for the identifying the soft permanent faults as well as the mechanism of time redundancy, so that transient faults in the network can be tolerated. Researchers have proposed several fault detection protocols based on comparisons, majority voting and coordination of neighbour. Protocols based on comparisons require threshold settings, the functioning of which may be unstable in the dynamic environment. Faulty nodes and the neighbours are compared in the approaches based on majority voting or neighbour coordination resulting in the decreased performance of the network protocol when fault percentage increases.Based on fuzzy inference system (FIS) of Takagi–Sugeno–Kang (TSK), Khan et al. have put forth a fault detection model for wireless networks [16], where each mobile node is trained to comprehend the measurements of the neighbouring mobile nodes. This model detects the soft permanent type of faulty nodes and endures the transient faulty nodes existing in the network. The approach of back propagation neural network has been preferred by Mourad and Nayak [17] in their proposed system, where the faults are detected at the system level itself for the wireless and wired interconnected networks. Simple and generalized comparison methods are employed in this model for the detection of hard and soft permanent faults in the network. The proposed fault detection scheme of Zhang Ji et al. [18] for mobile networks, uses the evidence theory along with radial basis function neural network (RBFNN) for fusion of information in order to identify the soft permanent faults. Another approach for fault diagnosis and isolation has been proposed by Jabbari et al. [19] for transportation system, where two distinct algorithms of neural network have been used for residual generation and verification. The neural network architecture of generalized regression and the probabilistic neural network (PNN) are applied for residual generation and residual verifications respectively. Different categories of soft and hard faults present in the mobile network are classified using the probabilistic features and residual evaluation. Azzam et al. [20] have proposed yet another dynamic model to diagnose the soft permanent faults in WSNs, where the altered recurrent neural network (RNN) is applied to model the WSN into two phases of learning and production to identify the faulty network nodes. Zhu et al. [21] propounded an approach based on neural network and a principal component analysis (PCA) for the mobile network systems to detect soft permanent faults. Here, the PCA model adopts the neural network in conjunction with a credit assigned cerebellar model articulation controller (CA-CMAC) to diagnose the multi-fault mobile system. Faulty mobiles are identified through the method of the squared prediction error (SPE). Researchers have proposed several fault detection protocols based on neural networks and soft computing. The constraint in these protocols is the increase in computational cost corresponding to the increase in the quantity of nodes with probability of fault. Kamal et al. [22] put forth a fault detection model based on sequence for use in WSNs to identify failures of node or link and reboot of node. Fletcher checksum algorithm along with server-side storage-intensive computation are employed in this method for exploitation of intra-network packet tagging and network failure detection. Nitesh and Jana [23] have put forth a distributed algorithm for fault identification for twin-tiered WSNs based on clusters, which identifies the transient and permanent faulty relay nodes on the basis of neighbouring table status. Swain et al. [24] have advanced an efficient distributed graph theoretic protocol for simultaneous detection of faults and cuts present in the wireless networks. Though these protocols employ diverse

approaches for the diagnosis of failure, they fail to consider the network faults due to persistence. Swain and Khilar [25,26] employ neural networks in their proposed composite protocol for fault diagnosis in WSNs, which is capable of detection, diagnosis and isolation of transient, intermittent, soft permanent and hard permanent faults present in the network. This composite protocol for detection of faults could categorise diverse faults, but needs a skilled neural network in addition to the order of magnitude being slow. In the context of the failure of the previous existing protocols for wireless networks to simultaneously handle diverse faults, the current state-of-art initiates research in the direction of simultaneous detection of heterogeneous faults in the network.

3. SYSTEM MODEL

The system model of the proposed methodology is illustrated hereunder with the assumptions, the network, communication and vulnerable models:

3.1 Assumptions

The following are some of the assumptions in the study:

(i) The nature of mobile nodes is homogeneous and static.

(ii) There is a unique identification for every mobile node.

(iii) The coordinator nodes are empowered with higher power of energy, transmission Range, computation, storage than the other network nodes.

(iv) There are no errors in coordinator nodes.

(v) The network topology is aware of the deployment of the coordinator nodes initially.

(vii) The MAC layer protocol undertakes the virtual links, which are largely fault-free.

3.2 Network Model

A graphic depiction of Wireless network as $G(v, e)$ can be made with $G(v)$ and $G(e)$ denoting the sets of nodes and edges in the network respectively. If a virtual link exists between the nodes $K_i, K_j \in G(v)$ and $V_{i,j} \in G(e)$, they can inter-communicate. The proposed approach uniformly distributes N number of mobile nodes in an N1 $\times$ N2 region, exceeding their transmission range R. If the two nodes K_i *and* K_j have the Euclidean distance either less than or equal to the range of transmission R, then the virtual link $V_{i,j}$ exists between them. Let (x_i, y_i, z_i) and (x_j, y_j, z_j) be the Cartesian coordinates pertaining respectively to the mobile nodes, K_i *and* K_j Then, the computation of the Euclidean distance $E()$ between the two mobile nodes can be depicted as Eq. (1):

$$E\left(K_i, K_j\right) = \sqrt{\left(x_j - x_i\right)^2 + \left(y_j - y_i\right)^2 + \left(z_j - z_i\right)^2} \quad (1)$$

The equation above illustrates that the link l_{ij} exists between, K_i *and* K_j if $E(K_i, K_j) \leq t$, with t representing the range of transmission of the homogeneous mobile nodes. Since all the virtual links in this network graph $G(v, e)$ are by nature undirected, the representation could be as.

$$l_{ij} \in G(e) \Leftrightarrow l_{ji} \in G(e) \quad (2)$$

$$l_{ij} \in G(e) \Leftrightarrow E\left(K_i, K_j\right) \leq K_i \quad (3)$$

is known as the neighbouring node of a node $K_i \in N$ by being within its transmission range. Then, virtual links between them are established for inter-communication.

3.3 Communication Model

This section illustrates the communication model depicting the complete communication time First, the cluster set up, formation and head selection are done in a stipulated period of time, initiating the next phase of intra-node communication. In this phase, there is an inter-node communication between all the mobile nodes $k_i \in N$ and the neighbour nodes $k_j \in N$ located within the range of transmission t. The set of neighbour nodes of a mobile node $k_i \in N$ is defined in Eq. (4) as:

$$N_e\left(k_i\right) = \left\{K_j \in G(v), e_{ij} \in G(e) \mid d_{ij} \leq t\right\} \quad (4)$$

Intra-cluster communication occurs in the next phase between the cluster head or coordinator node, as referred to in this paper, and the mobile nodes or cluster members, within the range. The coordinator node has a transmission power of T_p, where $T_p > T_c$, T_c being the cluster member's transmission power. Inter-cluster communication occurs in the last phase between the base station and cluster head through multi-hop technique.

All the above intra-node, intra-cluster and inter-cluster communications make use of multi-channel time division multiple access (TDMA) as MAC protocol to share the data and to reduce the scope for packet collision and delay in end-to-end transmission [27–30]. There are two phases in the MAC protocol based on the proposed schedule of TDMA, namely the control phase and the data phase. The control phase takes care of the reservation regarding the schedule of inter-communication of data between the source and the destination. The data phase handles data communication along with the schedule of control phase. The control phase is sub divided into several minute slots $\{s_1, s_2, s_3 \ldots, s_n\}$, where the slots' size depend on the size of RTS/CTS packet. Similarly the data phase is subdivided into a few slots $\{d_1, d_2, d_3 \ldots, d_n\}$ where the slots' size depend on the size of the data packet. The size of the RTS/CTS packet is comparatively very small to the size of the data packet and the exchange of RTS/CTS packets occurs in the control phase. The RTS packet holds the id of the receiver node and the selected data's slot number pertaining to data phase. The node with data forwards the RTS packet to its respective destination, while the destination node also transmits a CTS packet containing the source id and the slot number of the selected data. In the control phase all the nodes remain active, initially. Then, they select a slot of any channel of the control phase at random to set a timer. On zeroing of the set timer, the node transmits the RTS packet. The receiver after receiving the packet, responds with a CTS within the period of timeout. All the other nodes are well aware of the data schedule on the successful handshake of

the packets of RTS/CTS. When the receiver gets an RTS packet and is aware that the chosen slot has already been reserved for another node, it transmits a negative CTS, to indicate the source to find an alternate slot in the data phase. The control phase is so scheduled for data communication in the unique allotted slot of data phase as to prevent any possible collision, which may occur if any two nodes select the same slot at the stage of control phase itself for advertisement of their data. In such instances, the nodes attempt to choose the next available free slot for their data advertisement. If the node is aware of the exchange transmission of RTS/CTS exchange in the control phase itself, it waits till the comprehensive handshake is finished. After the successful accomplishment of the exchange of RTS/CTS packets, both the source and destination nodes in the data phase are awake in their allotted slot for communication of data. The source and destination nodes in their reserved slot in data phase interchange the packets of data and acknowledgement between them. The pair of nodes are awake in their respective reserved data slot, while all the other nodes are in sleep mode in order to minimize the waste of energy and idle listening and enhance the performance. The proposed protocol comprehensively enhances its ratio of packet delivery through this scheme of medium access control.

Vulnerable model

The proposed protocol for fault diagnosis deals with four categories of faults namely, (i) vtype1 (ii) vtype2, (iii) vtype3 and (iv) vtype4. Regarding vtype1, the r node $k_i \in N, i = \{1, 2, 3, \ldots, N\}$, the time instance denoted as $\tau = \{1, 2, 3, \ldots t\}$, then the node k_i for all time instances $\forall\tau\{k_i(\tau)\}$ fails to respond to its neighbours. Regarding vtype2 the node $k_i \in N$ with mobile value s, τ for each time instance $\tau = \{1, 2, 3, \ldots, t\}$, so the value of the mobile node $k_i(m_\tau)$ for all time instances is not equal to the actual mobile value α_τ. The vtype2 depicted in Eq. (5) as:

$$\forall(\tau)k_i\,(m_\tau) \neq \forall(\tau)\,\{k_i\,(\alpha_\tau)\} \tag{5}$$

where $k_i(\alpha_\tau)$ the actual or ambient value of node $k_i \in N$ and $\tau = 1, 2, 3, \ldots, t$ is the time instance. Regarding vtype3, for the time instance $\tau 1 = 1, 2, 3, \ldots, t1$, the mobile node value $k_i(m_{\tau 1})$ is not equal to the actual mobile value $m_{\tau 1}$ and for the time instance $\tau 2 = 1, 2, 3, \ldots, t2$, the mobile node value $k_i(m_{\tau 2})$ is equal to the actual mobile value. $m_{\tau 2}$ Here $\{\tau 1, \tau 2\} \in \tau$ and $\tau = \{1, 2, 3, \ldots \tau\}$ is the total time instance of the mobile node $k_i \in N$. The vtype3 error is depicted in Eq. (6) as:

$$\forall(\tau)\{\exists_{\tau 1}\{k_i(m_{\tau 1})\} \neq \{k_i(\alpha_{\tau 1})\}\}\&\{\exists_{\tau 2}\{k_i(m_{\tau 2})\} = \{k_i(\alpha_{\tau 2})\}\} \tag{6}$$

Regarding vtype4 fault, for a small time instance $\tau_s = \{1, 2, 3, \ldots, t_s\}$, the mobile node value $k_i(m_{\tau s})$ is not equal to the actual mobile node value $\alpha_{\tau s}$ and for the regular time instance $\tau_r = \{1, 2, 3, \ldots, t_r\}$, the mobile node value $k_i(S_{\tau r})$ is equal to the actual mobile node value $\alpha_{\tau s}$. Here $\tau_s << \tau_r$ and $\{\tau_s, \tau_r\} \in \tau$, where $\tau = \{1, 2, 3, \ldots, t\}$ is the total time instance of the mobile node $k_i \in N$. The temporary fault is depicted in Eq. (7) as:

$$\forall(\tau)\{\exists_{\tau s}\{k_i(m_{\tau s})\} \neq \{k_i(\alpha_{\tau s})\}\}\&\{\exists_{\tau r}\{k_i(m_{\tau r})\} = \{k_i(\alpha_{\tau r})\}\} \tag{7}$$

where the condition of $\tau_s <<, \tau_r$ is also satisfied. The proposed protocol makes use of the normal distribution method for the design of vulnerable node model and to compute the dependabilit of the mobile node $k_i \in N$ [31,32]. Let the mobile node be $k_i \in N$, with the mobile value $m_i(\tau)$ at τth time instance, where $i = \{1, 2, 3, \ldots, N\}$ and $\tau = \{1, 2, 3, \ldots, t\}$. Analysis of all the values, i.e. $\{m_i(\tau)\}t_\tau = 1$, is required to estimate the vulnerable behaviour of node k_i. In the τth time instance, the mobile value $m_i(\tau)$ can be identified either as valid or erroneous data. Normal distribution procedure, represented as N $(\alpha_i(\tau) \cdot \sigma_i(\tau))$ is followed by the mobile data value $m_i(\tau)$, where $\alpha_i(\tau)$ denotes the node's actual data at τth time instance, while the erroneous data variance existing at mobile node k_i is denoted as $\sigma_i(\tau)$. The following can be the depiction of the mobile node value:

$$m_i(\tau) = \alpha_i(\tau) + e_i(\tau) \tag{8}$$

Where $\alpha_i(\tau)$ is the actual or ambient mobile data measured by the mobile node $k_i \in N$ at time instance τ (this actual data depends on the requirement of the user such as humidity, temperature and pressure) and $e_i(\tau)$ is the erroneous value at τth time instance owing to any distortion or fault. The probability density function of the normal distribution for the mobile value is $m_i(\tau)$. The vulnerable rate corresponds to the variance of the mobile nodes. The variance remains the same for the nodes which are fault-free, while it is very high for vulnerable

4. PROPOSED METHODOLOGY

Three major phases of clustering, fault detection and vulnerable classification, accomplish the proposed methodology. The diverse stages of clustering technique are depicted in algorithm 1. Such is the process of clustering formation of the mobile nodes to their cluster head. All the information related to the respective cluster members of the region are stored by the cluster head or the coordinator node. Coordinator node's process of fault detection among the nodes in the network, which is detailed in Algorithm 2. A detailed explanation of proposed methods are presented in 4.1, 4.2 and 4.3 respectively.

4.1 Clustering Phase

At the time of deployment of all the homogeneous mobile nodes in the terrain area $A1 \times A2$, they are fault-free in nature. Every mobile node gathers all the possible information from the neighbouring nodes located in its radio range and duly records the same in the neighbouring table. The Coordinate nodes are the specific nodes tested to be fault-free with GSP enabled greater range of transmission and having larger computational power. They are uniformly deployed in the network along with the homogeneous mobile nodes. These coordinator nodes use multi-hop technique to maintain connectivity with the base station. As has already been detailed earlier, these coordinator nodes have higher transmission range T_C than the range of the other

Algorithm 1 Clustering Algorithm.

1. N: recognised of mobile nodes $= \{k_1, k_2, k_3 \ldots \ldots k_n\}$
2. recognised of of cluster heads $h_c = \{h_1, h_2, h_3 \ldots \ldots h_i\}$, where $n > i$;
3. Initialized $h_c(k_i)$: group of cluster heads, which are within the broadcast array of node $k_i \in N$
4. Prepared $d(i, j)$: distance between node k_i to cluster head h_c (which is calculated by Eq. (11));
5. Allocate the mobile node $k_i \in K$ to the consistent cluster head $h_c \in h_c(k_i)$ depends on the $d(i, j)$;
6. Execute a sorting technique using cluster heads depends on number of prearranged mobile nodes & $h_c = \{h_1, h_2, h_3 \ldots \ldots h_i\}$.
7. while $N \neq Null$ do
8. Select a cluster head from the sorting array h_c;
9. Allocate the mobile nodes to k_i to $h_j =$, such that $h_j \in h(k_i)$ & k_i is the nearest of h_j;
10. Delete k_i from N;
11. Inform the sorting array h_c, so that the negligible loaded cluster head h_j is the next element;
12. end while
13. Stop

mobile nodes $T_r(T_C > T_r)$. Then all the mobile nodes transmit their ids to the proximate in-range coordinator nodes, which are known as cluster heads, denoted as h_i. After the transmission of the id signal, the distance between the coordinator node and the respective mobile node is computed employing the model of Friis free space propagation [33]. The receiving power R is computed by Eq. (10) as :

$$R = t_p \times g_t \times g_r \times \frac{w_l^2}{(4 \times \pi \times d(..))^2} \tag{9}$$

where R denotes receiving power, t_p the transmission power, g_r the receiver's gain, g_t the transmitter's gain, w_l the signal wavelength, and d the distance between the two nodes respectively. The inter-node distance of k_i and k_j, i.e. $d(k_i, k_j)$ is computed through Eq. (11) as:

$$d\left(k_i, k_j\right) = \sqrt{t_p \times g_t \times g_r \times \frac{w_l^2}{(4 \times \pi)^2} \times \frac{1}{R}} \tag{10}$$

The following are the fundamental design principles of the clustering phase [34,35]. First, all the mobile nodes allot a cluster head on the basis of distance. Let us assume that the mobile node $k_i \in N$ has been allotted to the cluster head. A sorting algorithm is then implemented to sort out all the existing cluster heads on the basis of their respective assigned mobile nodes (or cluster members). If t $h_c = h_1, h_2, h_3, \ldots \ldots h_i$ is considered the cluster head sort, then, h_1 is assigned with the least number of cluster members according to sorting. Node $k_i \in N$ is then assigned to the most proximate h_1, in such a way that $h_i \in h(k_i)$, where $h(k_i)$ becomes the set of all the cluster heads located within the range of transmission of the node k_i. The cluster head is then rearranged after the assignment of the respective node to the considered cluster head h_1 in such a way that least number of cluster members are allotted to the next element. In the same way, the nodes are picked up one after another to be assigned to the nearest cluster head located within the range of transmission. All the nodes are processed in this manner. In each iteration of this procedure, the mobile node $k_i \in N$ is allotted to its respective cluster head, so that h_ih consisting the least number of mobile nodes has been already allotted for clustering. Thus, balancing of load between cluster heads is accomplished, reducing energy consumption considerably. The diverse stages of clustering technique are depicted in algorithm 1. Such is the process of clustering formation of the mobile nodes to their cluster head. All the information related to the respective cluster members of the region are stored by the cluster head or the coordinator node.

This algorithm requires O(1) constant time for the initialization steps of 1, 2, 3 and 4. Assignment of n number of mobile nodes to their respective cluster heads requires O(n) time in Step 5. The efficient heap sort technique is employed to sort the cluster heads, m in number in Step 6, taking a time of O(m log m). Iteration takes a time O(n) in Step 7, with Steps 8, 9 and 10 taking constant time in this loop. Sorting array is updated in Step 11 by adjusting min- heap tree consuming a time of O(log m). The time taken by the total loop is O(n log m). Hence, the total required time complexity for the algorithm is O(1)+ O(n)+ O(m log m+ O(n log m) = O(n log m), where $n > m$.

4.2 Fault Detection Phase

Faulty nodes in the network are detected in the phase of fault detection, which follows the clustering phase. It comprises two phases of detection of (i) phase1 and (ii) phase2

4.2 phase1 detection phase: The phase identifies the permanent vulnerability nodes existing in the network. A coordinator table containing the status register pertaining to its cluster members is maintained by each coordinator nose. For instance, t_s denotes the status register related to node $k_i \in N$ at time instance t. First, the value of all the status registers is taken as zero (0). Every coordinator transmits a hello message in its cluster for each duration of time t. Then, the cluster members on receiving such messages transmit back acknowledgment (Ack) to their related coordinator nodes located within the range of communication. The Ack message is received within a specified timeout period from the cluster, failing which that specific node is identified as likely faulty (LF) and its value with status register changes to one (1). The coordinator table updates this information. Then, the coordinator node transmits a hello message in the next time instance and awaits the Ack message. The total time instance $t = 1, 2, 3, \ldots, m, \ldots$ The timeout period (T_O) can be understood as the time taken for the transmission of a data packet and corresponding packet of Ack within the range of communication.

$$T_O = (2 \times T_d) + T_r + T_q + T_P \tag{11}$$

T_P Represents the delay in propagation or the consumed time for the transfer of packet from the source to destination. Eq. (14) defines the T_P for node k_i as:

$$T_P = \frac{\text{dist}(k_i, k_j)}{s}, \tag{12}$$

Where dist(k_i, k_j) represents the distance between node k_i and k_j and s denotes the packet's speed of propagation. T_r Signifies the delay in transmission or the router's packet transfer time in the channel. Eq. (15) defines the T_r for the node k_i as:

$$T_r = \frac{\text{len}(p_m)}{b}, \tag{13}$$

Where len(p_m) depicts the packet length for node k_i and b denotes the channel's bandwidth.

T_P Represents the delay in processing delay or the processing time of a packet in the router. The waiting time of the packet in the queue or queuing delay is depicted as t_q The values of the T_P and t_q are very marginal and hence, negligible and often considered zero. Presume that the coordinator considers the node k_i as (LF) after due verification of the status register. If the node fails to respond in 50% or more time instances to its environment, then it is categorized as per permanently faulty.

Phase2 *detection phase:* After the identification of the (vtyp4) vulnerable node(s), the proposed method in this phase attempts to the vtype3 nodes in the network. Every mobile node $K_i \in N$ transmits its value v_j where $j = \{1, 2, 3, \ldots, m\}$ to its coordinator node $h_k \in h_c$, located within its range of communication. The coordinator node identifies the faulty nodes present in its cluster region. The faulty nodes are identified using the Analysis of Variance method [36–38] based on statistics to analyse the actual and faulty mobile values. The method of ANOVA pursues two ways of hypothesis testing, namely, H_o = Null hypothesis and (ii) H_1 = Alternative hypothesis. As per the description of the null hypothesis (H_0) considerable variation between the mobile nodes and mobile values are absent, while the description of alternative hypothesis (H_1) establishes at least one considerable change between them.

The following is the discussion of steps of the ANOVA method: The mobile values mean of each individual mobile node $k_i \in N$, is computed, where k_i contains the mobile values $\{m_1, m_2, \ldots \ldots m_n\}$. Hence, μ_i denotes the mean of the mobile values and is computed in

Eq. (15) as:

$$\mu_i = \frac{1}{n}\sum_{j=1}^{n} m_j \tag{14}$$

Let the set of mobile nodes be $m = \{m_1, m_2 \ldots \ldots m_m\}$ with means $\mu_i = \{\mu_1, \mu_2, \mu_3, \ldots, \mu_n\}$. Then, in Eq. (16) the computation of the overall mean μ of the mobile nodes n in number is represented as:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} \mu_i \tag{15}$$

Compute the sum of squared differences among the nodes. Then, in Eq. (17) the sum of squared differences s_d among n number of nodes containing m number of mobile values for each node is computed as:

$$S_d = \sum_{i=1}^{n} m \times (\mu_i - \mu)^2 \tag{16}$$

The degrees of freedom d_b among the mobile nodes n in number is computed in Eq. (18) as:

$$d_b = n - 1. \tag{17}$$

Algorithm 2 Vulnerable node Detection Algorithm.

1. Initialized the coordinator table T_c for each coordinator node $h_j \in h_c$;
2. Initialized the status register $s_i = 0$ for each mobile node $k_i \in N$ at time instance t;
3. for each time instance $t = 1, 2, 3, \ldots, T$ do
4. Each coordinator node $h_j \in h_c$ broadcast hello messages in its cluster region;
5. After receiving hello message, each cluster member reply an *Ack* message in time period, (T_o) is defined in Eq. (16);
6. After timeout period (T_o) is expired, coordinator node h_j identifies the likely faulty (LF) node as k_i by changing the status register $S_t = 1$ at time t;
7. end for
8. Calculate the status of node $k_i \in N$, $S(k_i) = \sum_{t=1}^{T} S_i^t$
9. if $s(k_i) \geq \frac{T}{2}$ then
   Status of node k_i at time $(t + 1)$, changes to $S_i^{t+1} = 2$ and node k_i is declared as permanently hard faulty node in the network;
   end if
10. For n number of mobile nodes, each mobile node k_i sends m number of values, i.e. $\{m_1, m_2 \ldots \ldots m_m\}$ to its coordinator node in its cluster region;
11. for $i = 1$ to n do 14:
    Calculate mean $\mu_i = \frac{1}{n}\sum_{j=1}^{n} m_j$
    end for
12. Calculate overall mean in eq 20
13. Initialized sum squared differences between nodes $S_d = 0$
14. for $i = 1$ to n do
    a. $S_d = S_d + m \times (\mu_i - \mu)^2$
    end for
15. Calculate degrees of freedom between nodes in eq 22
16. Calculate mean squared value between nodes, in eq 26
17. Initialized sum squared differences within nodes, $S_d = 0$
18. for $i = 1$ to n do
    for $j = 1$ to m do
    $S_d = S_d + (S(i, j) - \mu_i)^2$
    end for
19. end for
20. Calculate degrees of freedom within nodes in eq 25
21. Calculate mean squared value within nodes, in eq 26
22. Calculate F-ratio, in eq 27
23. for $\alpha = 0.1$ 0.05 0.90 do
    $f_\sigma(\alpha) = f$ distribution (α, s_d, s_b)
24. for $\alpha = 1$ to r do
25. if $f_r > f_\sigma(\alpha)$ then
    $H_1 + +$;
    else
    $H_0 + +$;
    end if
    end for
26. if $H_1 > H_0$
    then Node k_i is declared as vtype3 node in the network;
    else
    Node K_i is declared as fault-free node in the network;
27. end if
28. STOP

Then, the mean squared value m_b among the mobile nodes n in number is computed in Eq. (19) as:

$$m_b = \frac{s_d}{d_b} \tag{18}$$

Next, the sum total of the squared differences s_d within the n number of mobile nodes containing $\{m_1, m_2 \ldots\ldots m_m\}$ mobile values for each node is computed in Eq. (20) as:

$$S_d = \sum_{i=1}^{n} \sum_{j=1}^{m} (S_j - \mu_i)^2 \tag{19}$$

In Eq. (21) the level of freedom d_b within n number of mobile nodes is computed as:

$$d_b = n \times (m - 1) \tag{20}$$

Then, Eq. (22) gives the computation of the mean squared value m_b within n number of mobile nodes as:

$$m_s = \frac{s_d}{d_b} \tag{21}$$

Finally, F-ratio is computed and defined as the ratio between the mean squared values among the nodes and within the nodes. Definition of the values of *ms b* and (*ms w*) have already been done in Eq. (23):

$$f_r = \frac{s_d}{d_b} \tag{22}$$

The coordinator node implements the ANOVA test within the region of its cluster after it receives the data from the mobile node using data sensed either by its own self or by other nodes. Then, comparisons are made at diverse prominent levels (0.05 to 0.90). If the comparisons are effected for r number of times, H_0 is satisfied for k number of times for all the normal nodes. H_0 is rejected for k number of times in case of vtype3 faulty nodes, H_0 is satisfied for less than $\frac{k}{2}$ number of times for both intermittent faulty nodes and in case of transient faulty nodes, H_0 is rejected greater than equal to $\frac{k}{2}$ number of times the hypotheses. As per the results, a node is declared faulty in the network, if the condition of H_0 is rejected for greater than equal to $\frac{k}{2}$ times for it. This is the coordinator node's process of fault detection among the nodes in the network, which is detailed in Algorithm 2.

Steps 1 and 2 in the above algorithm are the initialization stages taking O(1) time. The time instance in Step 3 is the actual duration of time with the loop taking O(t) time. Steps 4, 5 and 6 occur within the loop and take a time of O(t). Step 8 considers n mobile nodes to compute the value of status register for every instance of time and requires a time of O(nt) $\cong$ O(n) as t is constant time. Comparison of status register value for the considered n mobile nodes is done in Step 9 taking O(n) time, depending on which Step 10 needs constant time. Steps and 14 illustrate the calculation of the mean values of the considered n mobile nodes with each node containing m mobile values taking O(nm) time, where $n > m$. Step 16 shows the computation of overall mean, which takes O(n) time. Constant time of O(1) is taken in step 17. while steps 18 and 19 need O(n) time. Steps 21, 22, and Type equation here.23 need O(1) constant time. Steps 24, 25, and 26 take O(nm) time. Steps 29, 30, and 31 take constant time O(1). Due to the computation of F-critical value for the purpose of significance level of r times, steps 32 and 33 require O(r) time. The loop runs for r number of times in steps 35 to 41 for the comparison of F-ratio value with F-critical value requiring O(r) time. All these statements require O(1) time as r is a constant value. Steps 42 to 46 require O(n) time for the comparison of hypothesis pertaining to the n number of mobile nodes. Thus, the total required time complexity of the above algorithm is O(1)+ O(t)+ O(n) + O(nm)+ O(r) = O(nm), where $n > m$. The best case of the algorithm's time complexity O(n), can be accomplished, if the mobile value number m is less and constant.

5. PERFORMANCE EVALUATION

NS–2.35 [42] simulator is employed for performance evaluation of the proposed protocol in comparison with the other three existing protocols as advocated by Panda et al. [9], Chen et al. [13], and Azzam et al. [20] respectively. A squared area of 700×700 is considered for uniform deployment of all the mobile nodes. If the number of nodes deployed is N in the network to which the number of added nodes as faulty is P, then, the fault-free nodes are computed as N-P, through the application of the fault detection algorithm. T is simulation considers 500 mobile nodes deployed uniformly in a specific area. Cluster algorithm is used for cluster formation as illustrated in 4.1. Multi-channel mode as illustrated in Section 3.3 is used for inter-mobile node communication. Table 1. Reflects the parameters of default simulation. Initial deployment of mobile nodes is largely fault-free in nature subsequently adding faulty nodes gradually. A comparative analysis of the existing protocols and the proposed protocol is made.

Fig. 1. graphically depicts the accuracy of vulnerable detection of vtype4 nodes against fault probability, reflecting a decrease in detection accuracy, when fault probability is increased more percent in all the considered four protocols after obtaining the average results from 50 runs.

Fig. 2. graphically illustrates the accuracy of fault detection of vtype3 nodes against fault probability reflecting a decrease in detection accuracy, when fault probability is increased in all the considered four protocols. The output observations indicate a superior performance of the proposed protocol against the other three protocols with about more detection accuracy. In Panda et al. every individual node performed the altered three sigma test in conjunction with its neighbouring values regarding the detection of soft permanent faulty nodes considering each node k_i and m number of mobile values.

Fig. 3. graphically represents the accuracy of fault detection regarding heterogeneous vulnerable nodes against fault probability. An increase in the fault probability correspondingly decreases the accuracy in detection in all the four referred approaches. The proposed protocol establishes about increased accuracy when compared to the other three protocols in existence. The existing protocols are inadequate to accurately identify the percentage of vulnerable nodes. The performance of the proposed protocol proves superior to the existing approaches in its fault diagnosis, using ANOVA to investigate and control the level variation in the mobile values.

In Fig. 4 false alarm rate and fault probability graph are juxtaposed. Enhancing fault probability correspondingly increases false alarm rate. Panda et al., and Chen et al., put forward a neighbour-based approach, where fault-free nodes

Table 1 Simulation parameter.

Parameter	Value
Number of nodes	500
Deployment area	700 ×700 m^2
Carrier sense range	150 m
Transmission range	75 m
Duration of frame	236.4 ms
Duration of control period	5 ms
Duration of control phase slot	0.1 ms
Duration of data slot	12 ms
Duration of RTS/CTS packet	0.9 ms
Duration of data packet	8.5 ms
Data packet size	32 bytes
Rx power	59.1 mW
Tx power	52.2 mW
Channel rate	250 kbps
Source rate	5 pkt/s
Simulation time 200 s	200 s

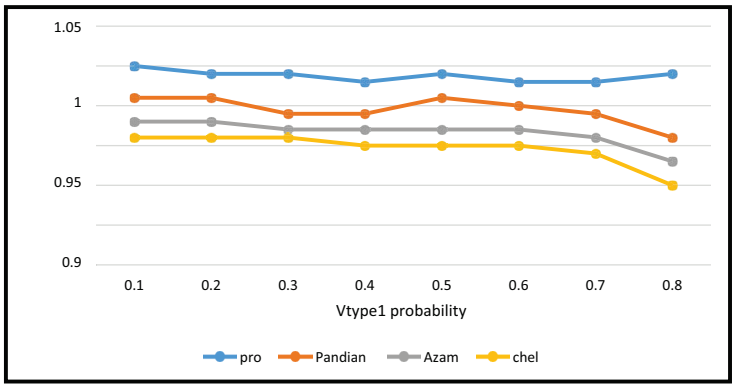

Figure 1 Vtype1 probability vs detection accuracy.

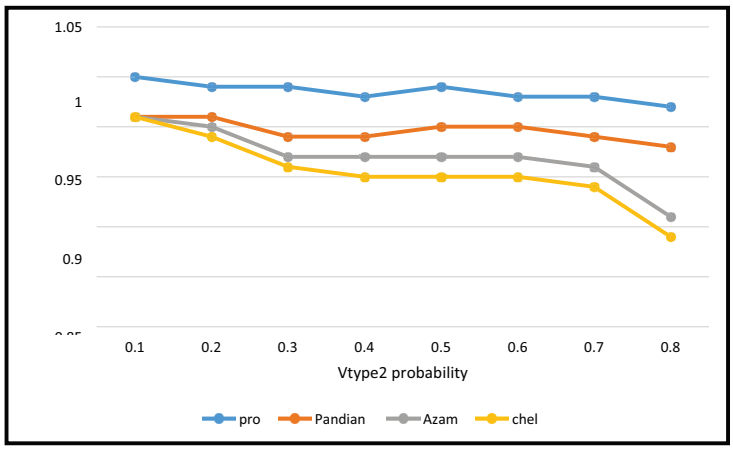

Figure 2 Vtyp2 probability vs detection accuracy.

located amidst more vulnerable neighbours are also identified as faulty. This rate of false alarm degraded the protocol.

Similarly, Fig. 5 graphically juxtaposes the rate of false positive and fault probability, where any increase in fault probability correspondingly increases the false positive rate regarding all the four protocols. As the proposed protocol can identify all categories of faulty nodes, it results in far less false positive rate in comparison to the other three existing approaches.

Fig 6. Graphically juxtaposes the energy consumption (EC) against the number of vulnerability nodes. The EC is computed on the basis of the total requirement of energy for the identification of faulty nodes existing in the network. Any increase in the number of faulty nodes correspondingly increases FDC and EC in the network pertaining to all the protocols.

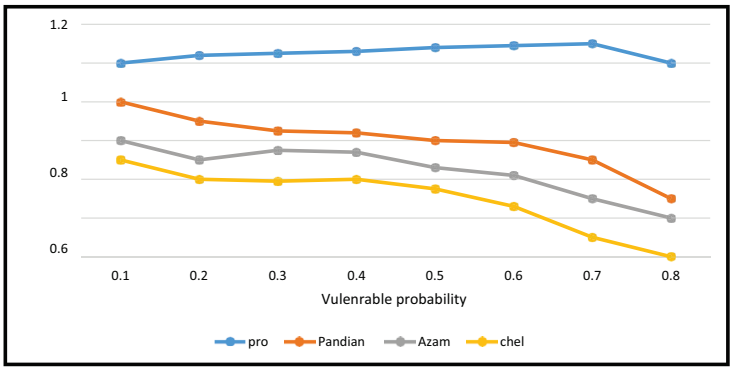

Figure 3 Detection accuracy vs vulnerable probability.

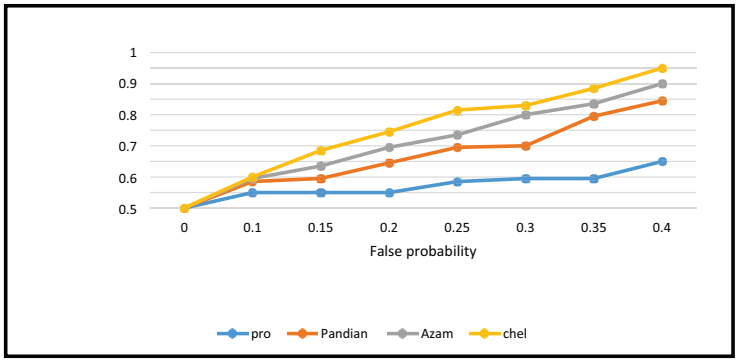

Figure 4 False alarm rate vs false probability.

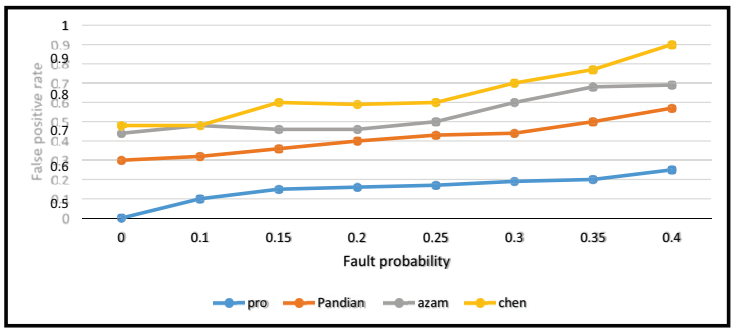

Figure 5 False positive rate vs false probability.

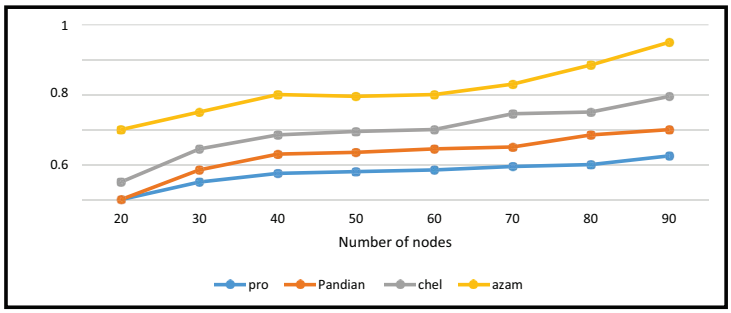

Figure 6 Energy conservation vs number of nodes.

6. CONCLUSIONS

This study proposes a protocol for heterogeneous fault diagnosis in adhoc network. The protocol functions in three phases of clustering, fault detection and fault classification. A multi-channel MAC protocol based on time division is developed for communication of data and the load balanced clustering method for reduced consumption of energy in the network. The network's vtype4 nodes are identified through timeout mechanism by checking the status register value for every time

instance. The test method of analysis of variance is used regarding the vtype3, vtype2, and vtype1 nodes for identification of significant levels of variation between the accurate mobile values and the values of the vulnerable nodes. The probabilistic neural network is used as model for the phase of fault classification to categorize the faulty nodes. Both simulation and the testbed environments are used for performance evaluation of the proposed protocol with probability variation from low to high. The output results of the evaluation establish the superiority of the proposed protocol over the existing ones as Panda et al. [9] , Chen et al. [13] , and Azzam et al. [20] as per the specific performance parameters of accuracy in fault detection and in false alarm rate, positive rate and classification rate. The methodology could further be used for diverse applications and mobile networks, such as body area mobile networks, vehicular adhoc networks, and under-water mobile networks etc.,.

REFERENCES

1. MouradElhadef AzzedineBoukerche HishamElkadiki, A distributed fault identification protocol for wireless and mobile ad hoc networks, 68(3) (2008) 323–55.
2. Michael Barborak, Anton Dahbura, Miroslaw Malek, The consensus problem in fault-tolerant computing, ACM Computing Surveys (CSUR), v. 25 n. 2, p. 172–20, June 1993.
3. S. Chessa, P. Santi, Comparison-based system-level fault diagnosis in ad hoc networks, in: Proceedings of the 20th IEEE Symposium on Reliable Distributed Systems (SRDS-2001), New Orleans, LA, USA, October 2001, pp. 257–266.
4. Arun Subbiah, Douglas M. Blough, Distributed Diagnosis in Dynamic Fault Environments, IEEE Transactions on Parallel and Distributed Systems, v.15 n.5, p.454–67, May 2004.
5. S. Misra, S. K. D, M. S. O, K. V., and U. G., *A low-overhead fault-tolerant routing algorithm for mobile ad hoc networks: A scheme and its* simulation analysis, Simulation Modelling Uractice and Theory, 2010, vol. 18, pp. 636–49.
6. M. Chouhan, M. N. Sahoo, U. M. Khilar **Fault Diagnosis in MANET** Communications in Computer and Information Science, 190 (2011), pp. 111–28.
7. S. Guo, Z. Zhong, T. He, Find: faulty node detection for wireless sensor net-works, in: Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems, ACM, 2009, pp. 253–266.
8. M. Panda, P. M. Khilar, Distributed byzantine fault detection technique in wire-less sensor networks based on hypothesis testing, Comput. Electr. Eng. 48 (2015) 270–285.
9. M. Panda, P.M. Khilar, Distributed self fault diagnosis algorithm for large scale wireless sensor networks using modified three sigma edit test, Ad Hoc Netw. 25 (2015) 170–184.
10. T. Panigrahi, M. Panda, G. Panda, Fault tolerant distributed estimation in wire-less sensor networks, J. Netw. Comput. Appl. 69 (2016) 27–39.
11. M. N. Sahoo, P. M. Khilar, Diagnosis of wireless sensor networks in presence of permanent and intermittent faults, Wirel. Pers. Commun. 78 (2) (2014) 1571–1591.
12. H. Guangjie, Huang Guangjie, G. Wei Guo Wei, S. Jian Su Jian, *A Novel Fault Diagnosis System for MANET Based on Hybrid GA-BP Algorithm*, 2008, pp, 1–4. M. R. Meybodi, H. Beigy **New Algorithms for Adaptation of Back propagation Algorithm Parameters** Iranian Journal of Science& Technology, Transaction B, 25 (B3) (2001), pp. 515–32.
13. H. Beigy and M. R. Meybodi, *Experimentation on Learning Automata Based Methods forAdaptation of BP Parameters*, Proceedings of Ninth Conference on Electrical Engineering, Power &Water Institute of Technology, 2001, pp. 4–1 to 4–9. 8–10.
14. W. Chen, J. Wan, R. Yu, Distributed fault diagnosis of wireless sensor networks, in: Proceedings of the 11th IEEE International Conference on Com-munication Technology, ICCT 2008, IEEE, 2008, pp. 148–151.
15. M.-H. Lee, Y.-H. Choi, Fault detection of wireless sensor networks, Comput. Commun. 31 (14) (2008) 3469–3475.
16. Azzam Moustapha and Rastko Selmic. 2007. Wireless sensor network modeling using modified recurrent neural networks: Application to fault detection. In Proceedings of IEEE ICNSC.
17. S. A. Khan, B. Daachi, K. Djouani, Application of fuzzy inference systems to detection of faults in wireless sensor networks, Neurocomputing 94 (2012) 111–120.
18. E. Mourad, A. Nayak, Comparison-based system-level fault diagnosis: a neu-ral network approach, IEEE Trans. Parallel Distrib. Syst. 23 (6) (2012) 1047–1059.
19. Z. Ji, W. Bing-shu, M. Yong-guang, Z. Rong-hua, D. Jian, Fault diagnosis of sensor network using information fusion defined on different reference sets, in: Proceedings of the International Conference on Radar, CIE'06, IEEE, 2006, pp. 1–5.
20. Jabbari R. jedermann, W. Lang Application of computational intelligence for sensor fault detection and isolation, WorldAcad Sci. Eng Technol 33 (2007) 262–70.
21. A. I. Moustapha, R. R. Selmic, Wireless sensor network modeling using modified recurrent neural networks: application to fault detection, IEEE Trans. Instrum. Meas. 57(5) (2008) 981–988.
22. D. Zhu, J. Bai, S. X. Yang, A multi-fault diagnosis method for sensor systems based on principle component analysis, Sensors 10 (1) (2009) 241–253.
23. A. R. M. Kamal, C. J. Bleakley, S. Dobson, Failure detection in wireless sensor networks: a sequence-based dynamic approach, ACM Trans. Sens. Netw. (TOSN) 10(2) (2014) 35.
24. K. Nitesh, P. K. Jana, Distributed fault detection and recovery algorithms in two-tier wireless sensor networks, Int. J. Commun. Netw. Distrib. Syst. 16(3) (2016) 281–296.
25. R. R. Swain, T. Dash, P. M. Khilar, An effective graph-theoretic approach towards simultaneous detection of fault (s) and cut (s) in wireless sensor networks, Int. J. Commun. Syst. 30 (13) (2017), doi:10.1002/dac.3273.
26. R. R. Swain, P. M. Khilar, Composite fault diagnosis in wireless sensor networks using neural networks, Wirel. Pers. Commun. 95(3) (2016) 2507–2548.
27. R. R. Swain, S. Mishra, T. K. Samal, M. R. Kabat, An energy efficient advertisement based multichannel distributed mac protocol for wireless sensor networks (ADV-MMAC), Wirel. Pers. Commun. 95(2) (2016) 655–682.
28. S. Ray, I. Demirkol, W. Heinzelman, Supporting bursty traffic in wireless sensor networks through a distributed advertisement-based TDMA protocol (ATMA), Ad Hoc Netw. 11(3) (2013) 959–974.
29. S. Mishra, R. R. Swain, T. K. Samal, M. R. Kabat, CS-ATMA: a hybrid single channel Mac layer protocol for wireless sensor networks, in: Computational Intelligence in Data Mining-Volume 3, Springer, 2015, pp. 271–279.
30. R. R. Swain, S. Mishra, T. K. Samal, M. R. Kabat, ADV-MMAC: an advertisement based multichannel mac protocol for wireless sensor networks, in: Proceedings of the ntemporary Computing and Informatics (IC3I), 2014, IEEE, 2014, pp. 347–352.
31. S. M. Kay, Fundamentals of Statistical Signal Processing, Prentice Hall PTR, 1993.T. Pham-Gia, T. Hung, The mean and median absolute deviations, Math. Com-put. Model. 34(7–8) (2001) 921–936.
32. H. T. Friis, A note on a simple transmission formula, Proc. IRE 34(5) (1946) 254–256.

33. P. Kuila, P. K. Jana, Energy e?cient load-balanced clustering algorithm for wire-less s networks, Proc. Technol. 6 (2012) 771–777.
34. H. Beigy and M. R. Meybodi, *Experimentation on Learning Automata Based Methods for Adaptation of BP Parameters*, Proceedings of Ninth Conference on Electrical Engineering, Power & Water Institute of Technology, 2001, pp. 4–1 to 4–9. 8–10.
35. S. Chessa and P. Santi, "Comparison-based system-level fault diagnosis in ad hoc networks", Proceedings of IEEE Symposium on Reliable Distributed Systems, 2001, pp. 252–66.
36. M. Elhadef, A. Boukerche and H. Elkadiki, "Performance Analysis of a Distributed Comparison-Based Self-Diagnosis Protocol for Wireless Ad-Hoc Networks" MSWiM'06, 2006.
37. K. Phanse and L. SaSilva, "Addressing the Requirements of QoS Management for Wireless Ad-Hoc Networks", Computer Communications, Volume 28, number 12, July 2003, pp. 1261–273.
38. N. Sridhar, "Decentralized Local Failure Detection in Dynamic Distributed Systems", Proceedings of the 25th IEEE Symposium on Reliable Distributed Systems, pp. 141–54, 2006.
39. Yalçin, S., Erdem, E. A mobile fault detection algorithm in heterogeneous wireless sensor networks: a bio-inspired approach. *Sâdhanâ* **45,** 4 (2020). https://doi.org/10.1007/s12046-019-1241-7.