Journal of Computer Science

Original Research Paper

WNBLI - A Multifarious Liability Examination for Wireless Sensor Networks

[1]Tamilarasi B and [2]UmaRani R

[1]Research Scholar (RM12CS44), SCSVMV University, Kanchipuram, India
[2]Associate Professor of Computer Science, Sri Saradha College for Women, Salem -16, India

Article history
Received: 07-01-2018
Revised: 05-02-2018
Accepted: 07-05-2018

Corresponding Author:
Tamilarasi B
Research Scholar
(RM12CS44), SCSVMV
University, Kanchipuram, India
Email: arasi21373@gmail.com

Abstract: Wireless communications has completely revolutionized the living environment, just as personal computers forever altered the way of working in the 1980's and the Internet dramatically changed the obtainment and accessing way of information in the 1990's. Wireless sensor networks are evolving to a new level. The Failure liability is a very prominent problem in Wireless Sensor Network (WSN) research. Failure liable nodes behave in different form and this faulty behavior is referred to as multifarious faults or heterogeneous faults in WSN. This paper presents a Multifarious Failure liability Investigation protocol called Wireless Nodal Behavior and Liability Investigation (WNBLI) protocol for wireless sensor networks. The proposed WNBLI protocol consists of three vital phases, such as Association phase, Apprehension phase and Assortment phase. The performance of the proposed Wireless Nodal Behavior and Liability Investigation protocol is evaluated using network simulator.

Keywords: Wireless Sensor Networks, Wireless Node Behavior, Multifarious Liability, Probabilistic Neural Network, Liability Examination

Introduction

Wireless Sensor Networks (WSNs) play a very important part in implementing a technology because they require low deployment funds. The deployment of WSNs requires less effort and it is versatile in nature. The power consumption by the nodes in the WSNs is a major problem. This power constraint issue must be eradicated to deploy fully converged WSNs. The power consumption is attributed to two main factors namely transmission flow and data packet generation. The consumption of power can be minimized by taking control of the transmission and optimizing the data packet size. If the transmission rate is increased, it will have a direct proportion to the bit error rate suppressing factors. On the other hand if more power is utilized, then it will drain the power resources. If the route path involved in the transmission uses less power and can achieve the required rate of transmission, then the utilization of high power rate is unnecessary. The size of the data packet in the transmission is directly related to the power allotment from the power energy source. Thus packet size plays a vital role in the data transmission optimization (Kurt *et al*., 2017). The nodes in the WSNs consume energy depending upon their utilization. This utilization factor is dependent upon the load of data packets that is transmitted and the node position. If the node is located in the best route path, the transmission factors will be high whereas less interested path require less or sometimes null energy. The power factor of the nodes depends upon the area of deployment. If the nodes have to collect data all day long for everyday, say years, then the power backup for the nodes must be taken into account. If the power supply is negligible, then the nodes operating in such areas will suffer a power consumption defect. The node efficiency is dependent on the power consumption which is directly proportional to the load of transmission. If the nodes are located in areas inaccessible to humans, then it will cause a seizure of node operations, thus making it vulnerable to various layers of attacks (Jan *et al*., 2017).

Wireless Sensor Networks (WSNs) have attracted lot of attention due to their pervasive nature and their wide deployment in Internet of Things, Cyber Physical Systems and other emerging areas. The limited energy associated with WSNs is a major bottleneck of WSN technologies (Shaikh and Zeadally, 2016). Recent breakthroughs in wireless technologies have greatly

spurred the emergence of Industrial Wireless Sensor Networks (IWSNs). To facilitate the adaptation of IWSNs to industrial applications, concerns about networks' full coverage and connectivity must be addressed to fulfill reliability and real-time requirements. Although connected target coverage algorithms in general sensor networks have been extensively studied, little attention has been paid to reveal both the applicability and limitations of different coverage strategies from an industrial viewpoint (Han *et al*., 2017). The growth in technology paved way for the advancement in Multimedia Wireless Sensor Networks (MWSNs). The main problem in deployment of MWSNs is the security and trust establishment between necessary nodes. The privacy factors also play a very crucial role in the MWSNs. The user authentication and authorization part, that is employed in the MWSNs are much different from the conventional sensor network topology. The limitation factors in MWSNs range from high power consumption to optimal route path selection failure due to the inability to separate the failure node points from the efficient nodes while determining the best route path (Usman *et al*., 2016). There are a number of security issues in the WSNs and the one that overwhelmingly studied and attracted the researcher in the recent years is the key management issues of network. Though researchers have proposed various key management schemes to enhance the secure information transfer, the existing works lack in the tradeoff between its efficiency and the cost in terms of the factors like structure formulation, energy utilization and key storage overheads (Swaminathan and Vivekanandan, 2017; Rault *et al*., 2014; Li *et al*., 2013). Wireless sensor networks can be deployed in any attended or unattended environments like environmental monitoring, agriculture, military, health care etc., where the sensor nodes forward the sensing data to the gateway node. As the sensor node has very limited battery power and cannot be recharged after deployment, it is very important to design a secure, effective and light weight user authentication and key agreement protocol for accessing the sensed data through the gateway node over insecure networks. Most recently, Turkanovic *et al*. proposed a light weight user authentication and key agreement protocol for accessing the services of the *WSNs* environment and claimed that the same protocol is efficient in terms of security and complexities than related existing protocols (Gutierrez *et al*., 2014; Nikolidakis *et al*., 2013; Li and Lin, 2015; Amin and Biswas, 2016).

Related Work

Research works pertaining to the concept of improvising the performance and efficiency factors of the wireless sensor networks have been published. In all the works, the study was focused on two main aspects which is network topology and best route path discovery. All the works have their own delicate point of supremacy which excels them for future study.

Amin and Biswas (2016) designed a novel architecture for the Wireless Sensor Network environment and basing upon which a scheme has been proposed and presented for user authentication and key agreement scheme. The security validation carried out in the proposed work follows a session management and key authorization methodology.

Zhang *et al*. (2016) proposed an asymptotically optimal algorithm based on the dual decomposition method and a sub optimal algorithm. This sub optimal algorithm was done with lower complexity. The working nature of the algorithm differs in the operational key management status which ensures maximum performance.

Baroutis and Younis (2017) presented a technique for preserving privacy with location attribute, which is the in-situ base-station. The technique injects deceptive transmissions aiming to even the traffic abundance across the network. The highlight of the proposed work is the trade-off between location privacy and network's performance.

Prabhu *et al*. (2016) proposed a detailed study of various distributed clustering approaches in the wireless sensor networks. This detailed research study is made on optimized cluster initialization based on jumping ant approach in order to avoid random cluster development. The operational status of the cluster heads is given more importance implicating its performance.

Gu *et al*. (2016) research work aims to address the gaping issues on the sink mobility problem occuring in the mobile networks. The way of depiction of the work is similar to the bisection of wireless sensor networks. The schemes proposed in this research work are divided into four categories: Uncontrollable Mobility (UMM), Path-Restricted Mobility (PRM), Location-Restricted Mobility (LRM) and Unrestricted Mobility (URM).

Weber (2017) proposed a multi-round Aloha-based protocol which illustrates the network topology as the sending and receiving node ends. The receiving nodes collect data from the transmitting nodes. The protocol takes the working part of the nodes performance into account and generates metrics which acts as a base for the node performance evaluation. The concentration of the nodes existence and the performance attribute collectively form the metric analyzing factor.

Prathima *et al*. (2016) proposed a Secure Data Aggregation for Multiple Queries (SDAMQ) in Wireless Sensor Networks. In this proposed work many links of nodes are formed into a aggregate queries from the depression node and are authenticated, then listed to the sensor nodes. The sensor nodes respond by collecting these data and making them into a single packet, thereby reducing the transmit metric factor. The proposed model consumes less energy.

Njoya *et al*. (2017) proposed and implemented an algorithm which is performance efficient and scalable. The proposed algorithm creates virtual sensors which reduces the need of physical sensors. This deployment of virtual sensors creates an array of functional nodes which carries the data from the source to the destination in a controlled way depicting a secure transmission process. The performance of the virtual sensors is similar to the physical nodes employed and the working part of the data transmission unit of the virtual nodes is scalable in both operations and fail safe nature.

Understanding the Structure of the Model

The following structure describes the Inferences, Network representation and Flow of Communication.

Inferences

The following inferences must be taken into account:

- The wireless sensor nodes are static and analogous in nature
- Each nodes in the WSN has an Unique Identification (Node-UID)
- There are supervisory nodes which are fail safe and are more efficient than other nodes
- The deployment of the supervisory nodes is controlled in the initial state but after the flow of communication happens, the node positions and their contributions are dependent upon their working functionality
- There are virtual links created to maintain loop free communication

Network Representation

The wireless sensor network can be represented as a graph $GR(nn,bin)$ where nn is the number of nodes and bin is the bay area in the network. The $node_a$, $node_b \in GR(nn)$ can communicate, if there is a virual link existing $bin_{a,b} \in GR(bin)$ between the two nodes. In the proposed methodology, NUM numbers of nodes are distributed in a region of REG_1 x REG_2 which is greater than the transmission ($Tran_{dis}$) range of the wireless sensor nodes. The virtual link $bin_{a,b}$ between the two nodes $node_a$ and $node_b$ exist, if the Euclidean distance between them is less than their transmission ($Tran_{dis}$) range of the wireless sens Li, S., Da Xu, L. and Wang, X. (2013). Compressed sensing signal and data acquisition in wireless sensor networks and internet of things. *IEEE Transactions on Industrial Informatics*, *9*(4), 2177-2186. Let (j_a,k_a,l_a) and $(j_b.k_b.l_b)$ are the Cartesian co ordiantes of the sensor nodes $node_a$ and $node_b$ respectively. Therefore, the Euclidean Distance ($EU\text{-}DIS(.)$) between the two wireless sensor nodes are calculated by:

$$EU - DIS\left(node_{a,}node_b\right) = \sqrt{\left(\left(j_b - j_a\right)^2 + \left(k_b - k_a\right)^2 + \left(l_b - l_a\right)^2\right)} \quad (1)$$

From the above equation, virtual link $bin_{a,b}$ exists between the two nodes $node_a$ and $node_b$, if $EU\text{-}DIS(node_{a,}\ node_b) \leq Tran_{dis}$, where $Tran_{dis}$ is the transmission range of the static and analogous nodes. Again all the virtual links of this network graph $GR(nn,bin)$ are undirected and it is represented as the following equations:

$$bin(a,b) \in GR(bin) \leftrightarrow bin(b,a) \in GR(bin) \quad (2)$$

$$bin(a,b) \in GR(bin) \leftrightarrow EU - DIS\left(node_{a,}node_b\right) \leq Tran_{dis} \quad (3)$$

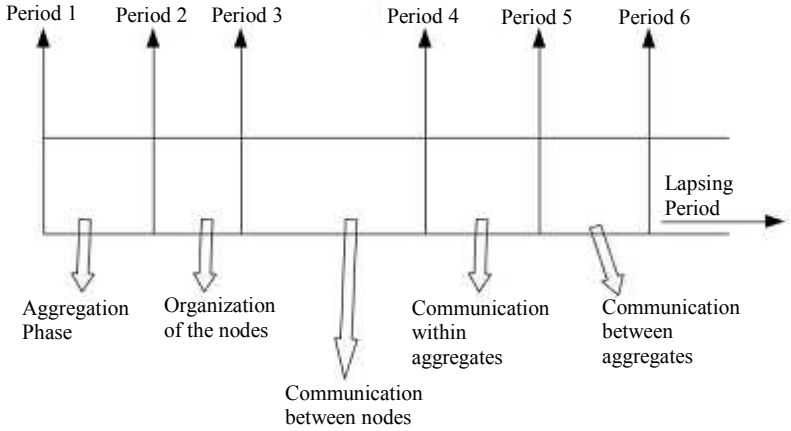

Fig. 1: Time Period depicting the entire flow of communication

Flow of Communication

The entire flow of communication between the nodes is clearly depicted in the following Fig. 1. Initially the nodes are aggregated to form a cluster. This cluster then organizes itself in a certain period of time. After the cluster organization, the nodes begin to communicate between them. This nodal communication phase is prominent and the communication flow of the nodes are detailed and initiated to the aggregate communication. The flow of information happens between aggregates and also within aggregates to provide enough information on the information flow.

Proposed Work

The proposed Wireless Nodal Behavior and Liability Investigation (WNBLI) protocol for wireless sensor networks is accomplished mainly in three major phases, such as: (i) Association phase, (ii) Apprehension phase and (iii) Assortment phase.

Association Phase

In this phase the nodes involved in the transmission process is analyzed deeply and is given a node identity which is unique in nature. The nodes are grouped together into a group which contains all the active nodes in them like $NN_g = \{node_1, node_2, node_3,..node_n\}$. There are numerous clusters that are formed as group of association head nodes where the number of nodes is always more than the association head nodes count like $NH_g = \{nh_1,nh_2,nh_3,....nh_m\}$. The nodes which falls under the transmission range of the Base supervisory node is taken into account and the actual count of the nodes participating in the transmission is noted down. The Base transmission node will always communicate with the supervisory node and the number of nodes involved with the supervisory nodes is taken in as an unit and this unit is unique to that region of transmission. The active nodes are associated with a supervisory node and this count is given by $NH_g = \{nh_1\}$ where NH_g is the count of association head nodes which contains the first set of associated nodes nh_1.This process of Association will take place till all the active nodes are assigned to a supervisory node. The node distribution is depicted clearly in the Fig. 2 given below. The working nature of the nodes with the supervisory nodes is directly proportional to the transmission range of the nodes with the supervisory node and also the range up to which the supervisory node can communicate with the base transmission node.

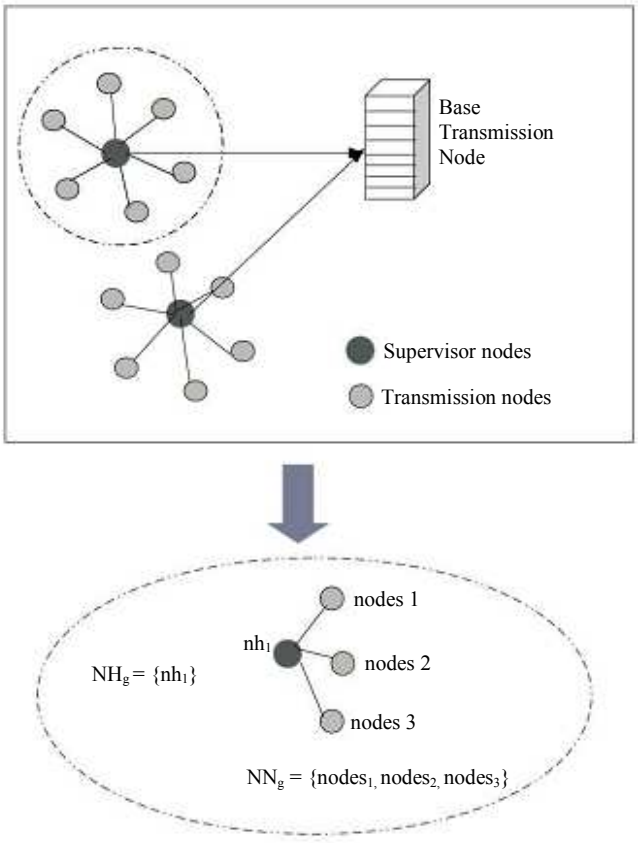

Fig. 2: Topology depicting Supervisor and Transmission nodes

a) Algorithm: Association Algorithm

(1) Initialized group of nodes: NN_g = {$node_1$, $node_2$, $node_3$,….$node_n$}
(2) Initialized group of association head nodes: NH_g = {nh_1,nh_2,nh_3,….nh_m} where n > m
(3) Initialized $nh(node_u)$: Set of all nodes which falls within the transmission range of the node $node_u \in NN_g$
(4) Initialized $dis(u,v)$ which gives the distance between $node_u$ and node association head nh_v
(5) Assign the node $node_u \in Nn_g$ to corresponding node association head group $nh_v \in nh(node_u)$ based on $dis(u,v)$
(6) Perform the Association technique based on allotment made with association head node groups NH_g = {nh_1,nh_2,nh_3,….nh_m}
(7) while $NN_g \neq$ Null
(8) Choose a association head node group nh_v from association from NH_g
(9) Assign the nodes $node_u$ to nl_v such that $nh_v \in nh(node_u)$ and $node_u$ is closer to nh_v
(10) Delete $node_u$ from NN_g
(11) Update the association array NH_g so that the load is minimum to the nh_v which is the next association head node group element
(12) end while
(13) Stop

Apprehension Phase

After the Association phase, the nodes that are defective are identified and suitable measures are undertaken by reacting to the imprecise node activity. In the Apprehension phase, the inaccurate nodes are identified. This phase is divided into two states depending upon the imprecise node characteristics. They are dense defective state and Elastic defective state.

a) Dense Defective State

In this state the permanently erroneous nodes are identified. In a node communication topology, there will be a base transmission node and this node communicates with all the supervisory nodes. The supervisor nodes present in the topology maintains a table called as Node Administration Table [NADT]. This NADT maintains all the information about the nodes connected to its domain like node state (active or passive), node activity (transmitting or dormant), node positions, etc.,. The status of the node ($node_i$) belonging to the group (NN_g) at the time (tp) is denoted as ($nod\text{-}stat^{tp}{}_i$). The possible values are 1,2 and 3 where 1 is node functioning normally, 2 is nodes that are likely to be unreliable and 3 is node that are completely erroneous.

Initially the $nod\text{-}stat^{tp}{}_i$ value is 1 as all the nodes belonging to that domain are considered to be functioning normally. The supervisor nodes will send Hello messages to all the participating nodes in its domain for each time period of tp. After receiving the Hello message, the participating nodes send Acknowledgment message (n-ACK) back to the supervisory nodes. The Acknowledgment messages (n-ACK) are received within a time period tp_{out}. If the n-ACK messages is *n* not received within the time period tp_{out}, then the node status value $nod\text{-}stat^{tp}{}_i$ of the NADT of the supervisory node is changed to 2 which means the node is likely to be unreliable and can become erroneous some time soon. The total time instance (tp) depends upon various aspects. The time out period tp_{out} is the total time taken for the complete exchange of a data packet and its corresponding acknowledgment packet. The time out period tp_{out} for node $node_i$ belonging to the group NN_g [$node_i \in NN_g$] is given by the formula:

$$tp_{out} = \left(2\ x\ tp_{propd}\right) + tp_{trand} + tp_{queued} + tp_{processd} \quad (4)$$

Where:
tp_{propd} = Data proliferation delay
tp_{trand} = Transmission delay-aware period
tp_{queued} = Queinng delay-aware period
$tp_{processd}$ = Processing delay period

b) Data Proliferation Delay

The data proliferation delay tp_{propd} for $node_i$ in the data transmission between $node_i$ and $node_k$ given by:

$$tp_{propd}(i) = \left(distance\left(node_i, node_k\right)\right) / m \quad (5)$$

where, m is the data proliferation speed in the given data transaction.

c) Transmission Delay-Aware Period

The transmission delay-aware $periodtp_{trand}$ for $node_i$ in the data transmission between $node_i$ and $node_k$ given by:

$$tp_{trand} = ln\left(pack_i\right) / cbw \quad (6)$$

where, $ln(pack_i)$ denotes the data packet length for $node_i$ and cbw denotes the channel bandwidth of the propagation that is happening.

d) Queinng Delay-Aware Period

The queinng delay-aware period is the time taken by the transmitted data packet to wait till it awaits its turn to get processed.

e) Processing Delay Period

The processing delay period is the time period taken for processing the data in the router.

At any given point of time, the value obtained for Queinng delay-aware period (tp_{queued}) and Processing delay period ($tp_{processd}$) is very negligible. So the values for tp_{queued} and $tp_{processd}$ are considered to be zero. So taking these considerations, Equation 4 can be reprocessed as:

$$\begin{aligned} tp_{out} &= \left(2\ x\ tp_{propd}\right) + tp_{trand} + tp_{queued} \\ &+ tp_{processd} = 2\ x\ tp_{(propd)i} + tp_{(trand)i} \\ &= 2\ x\left[\left(distance\left(node_i, node_k\right)\right)/\ m\right] + \left[ln\left(pack_i\right)/\ cbw\right] \end{aligned} \tag{7}$$

The channel access scheme of Time Division Multiple Access based Media Access Protocol is used where the slot duration is set as tp_{out}. The $nod\text{-}stat^{tp}{}_i$ value of node ($node_i$) for time instance tp = {1,2,3,4,5….m….tp} is computed and the status of the node ($node_i$) is given by:

$$nod - stat^{tp}{}_i\left(node_i\right) = \sum_{tp=1}^{tp} nod - stat^{tp}{}_i \tag{8}$$

Let the node $node_i$ is considered as unreliable node by the Node Administration table [NADT] which is obtained by checking the $nod\text{-}stat^{tp}{}_i$ value that is computed. If the value is 50% or more than 50% in failing to respond, then it is considered as completely erroneous node with tp denoting the total time slot is given by the equation:

$$nod - stat^{tp}{}_i\left(node_i\right) = \sum_{tp=1}^{tp} nod - stat^{tp}{}_i \geq [tp/2] \tag{9}$$

f) Elastic Defective State

In this state the elastic nodes, that is nodes with little faults are identified and the operational efficiency of the nodes are taken into account. Initially each sensor node $node_i$ belonging to the group $NN_g[node_i \in NN_g]$ sends its sensed value sen_j, j = {1,2,3….m} to its association head node $nh_k \in NH_g$ in its transmission range. The supervisor nodes detects the presence of erroneous nodes in its domain. A statistical based analysis method namely Analysis of Variance (ANOVA) method (O'brien, 1979; Cvijović *et al*., 2005; Zhang *et al*., 2012; Ozcelik and Erzurumlu, 2006; Azadeh *et al*., 2007) is used to analyze the actual sensor values and the erroneous sensor values and the unreliable nodes are picked out. The ANOVA method follows two types of hypothesis for testing namely Null Hypothesis (HY_0) and Alternative Hypothesis (HY_1). The Null Hypothesis (HY_0) states that there is no significant difference between the values given by the transmitting nodes. The Alternative Hypothesis (HY_1) states that there is at least one significant difference between the values given by the transmitting nodes. The ANOVA method steps are discussed as follows: Calculate the mean node values of each nodes $node_i \in NN_g$ where $node_i$ having sensor value {sen_1,sen_2,sen_3,…..s_m}. So the mean sensor value for the node $node_i$ is denoted as μ_i and is calculated in equation (10):

$$\mu_i = \frac{1}{m}\sum_{j=1}^{m} sen^j \tag{10}$$

Let the sensor nodes set NN_g = {$node_1$, $node_2$, $node_3$,…. $node_n$} have means μ_i ={μ_1, μ_2, μ_3,…..μ_i}, The overall mean μ for n number of nodes is calculated as Equation 11:

$$\mu = \frac{1}{n}\sum_{i=1}^{n^j} \mu i \tag{11}$$

Then calculate the sum squared difference between the nodes [ssq_{diff}]. The ssq_{diff} between n number of nodes having m number of sensor value $nod\text{-}stat^{tp}{}_i$ of the NADT of the supervisory node is changed s per node is given by the Equation 12:

$$sumsq_{diff} = \sum_{i=1}^{m^j} m_x\left(\mu_i - \mu\right)^2 \tag{12}$$

The degree of freedom (deg_{fr}) between n number of nodes is calculated in Equation 13:

$$deg_{fr} = n - 1 \tag{13}$$

Then the mean square value (msq_b) is calculated between the n number of nodes as depicted in Equation 14:

$$msq_b = sumsq_{diff}\ /\ deg_{fr} \tag{14}$$

Next the sum of squared difference (ssq_{diffr}) value within n number of sensor nodes having sensor value {sen_1,sen_2,sen_3,…..s_m} per node is calculated in Equation 15:

$$ssq_{diffr} = \sum_{i=1}^{n}\sum_{j=1}^{m} n\left(sen_j - \mu_i\right)^2 \tag{15}$$

The degree of freedom (deg_{fr}) within n number of nodes is calculated in Equation 16:

$$Degw_{fr} = n \times \left(m - 1\right) \tag{16}$$

Then the mean square value (msq_b) is calculated within the n number of nodes as depicted in Equation 17:

$$msqw_b = sumsqw_{diff}\ /\ degw_{fr} \tag{17}$$

Finally the Erroneous node ratio which is given as EN-ratio is calculated, which is defined by the ratio

between mean squared value between nodes (msq_b) and mean squared value within *nodes*($msqw_b$). These values are already computed in Equation 14 and 17:

$$EN - ratio = msq_b \,/\, msqw_b \tag{18}$$

From this, the error critical value (Er_{critic} = (deg_{fr}, $degw_{fr}$) is calculated where the significance level α varies from 5 to 90%. The Er_{critic} = (deg_{fr}, $degw_{fr}$) is compared with the EN-ratio value for finding the significant level α. If the EN-ratio > Er_{critic} = (deg_{fr}, $degw_{fr}$) condition is satisfied then, the Null Hypothesis (HY_0) is rejected and concluded that there is a significant difference between the sensor values of the sensor nodes. If the EN-ratio < Er_{critic} = (deg_{fr}, $degw_{fr}$) condition is satisfied then, the alternative Hypothesis (HY_1) is rejected and it is concluded that there is no significant difference between the sensor values of the sensor nodes.

g) Algorithm: Apprehension Algorithm

(1) Initialized the Node Administration table [NADT] for each node association lead node
(2) Initialized $nod\text{-}stat^{tp}{}_i$ value for each sensor node with time out period tp_{out} for node $node_i$ belonging to the group NN_g [$node_i \in NN_g$]
(3) For each time period moment = {1,2,3,….tp} do
(4) Supervisor nodes will send Hello messages to all the participating nodes in its domain for each time period of tp.
(5) After receiving the Hello message, the participating nodes send Acknowledgment message (n-ACK) back to the supervisory nodes.
(6) The $nod\text{-}stat^{tp}{}_i$ of the NADT of the supervisory node is changed according to the values (possible values are 1,2 and 3 where 1 is node functioning normally, 2 is nodes that are likely to be unreliable and 3 is node that are completely erroneous.)
(7) Valid $nod\text{-}stat^{tp}{}_i$ values are recorded
(8) end for
(9) Calculate value $nod\text{-}stat^{tp}{}_i$ ($node_i$) = $\Sigma^{tp}{}_{tp\,=\,1}$ $nod\text{-}stat^{tp}{}_i$
(10) if $nod\text{-}stat^{tp}{}_i$ ($node_i$) = $\Sigma^{tp}{}_{tp\,=\,1}$ $nod\text{-}stat^{tp}{}_i \geq$ [tp/2] then
(11) declare node as completely erroneous
(12) endif
(13) for i =1 to n do
(14) Calculate mean $\mu_i = \frac{1}{m}\sum_{j=1}^{m} sen^j$
(15) end for
(16) Calculate overall mean $\mu = \frac{1}{n}\sum_{i=1}^{n'} \mu i$
(17) Initialized the sum squared difference between the nodes [ssq_{diff}] which is taken as 0
(18) for i = 1 to n do
(19) Calculate the value for $sumsq_{diff} = \sum_{i=1}^{m'} m_x\,(\mu_i - \mu)^2$
(20) end for
(21) Calculate degree of freedom (deg_{fr}) between n number of nodes $deg_{fr} = n - 1$
(22) Calculate the mean square value (msq_b) which is $msq_b = sumsq_{diff}/deg_{fr}$
(23) for i = 1 to n do
(24) for j = 1 to m do
(25) Calculate $ssq_{diffr} = \sum_{i=1}^{n}\sum_{j=1}^{m} n\left(sen_j - \mu_i\right)^2$
(26) end for
(27) Calculate degree of freedom (deg_{fr}) which is $degw_{fr} = n \times (m\text{-}1)$
(28) Calculate mean square value (msq_b) which is $msqw_b = sumsqw_{diff}/degw_{fr}$
(29) Calculate the Erroneous node ratio which is EN-ratio = $msq_b/msqw_b$
(30) Calculate the Error critical value (Er_{critic} = (deg_{fr}, $degw_{fr}$)
(31) for α = 0.1:0.05:0.90 do
(32) $Er_{critic}(\alpha)$ = Error distribution (α, deg_{fr}, $degw_{fr}$)
(33) end for
(34) for α = 1 to r do
(35) if EN-ratio > Er_{critic} = (deg_{fr}, $degw_{fr}$)
(36) HY_1 ++;
(37) else
(38) HY_0 ++;
(39) endif
(40) end for
(41) if $HY_1 \leq [HY_0/2]$ then
(42) node n_i is declared as a erroneous node
(43) else
(44) node n_i is declared as a normal node
(45) endif
(46) Stop

Probabilistic Neural Network

The nodes that are identified as erroneous are classified in the assortment phase. Probabilistic Neural Network (PNN) (Wu *et al*., 2007; Mishra *et al*., 2008; Specht, 1988; Xiao *et al*., 2017; Silva *et al*., 2018) is used to assort the erroneous nodes. PNN is used for analyzing the probability of different erroneous situations for unreliable node assortment process. PNN is generally a classifier, which takes the input that is feed, analyze its pattern and then turn them into corresponding output for assortment. Compared to the operational efficiency of neural networks, the PNN mode of operation and performance is much higher in network training like back propagation technique. PNN lacks the capacity of including minimal problem which makes it a better player in assorting the nodes after the pattern is analyzed.

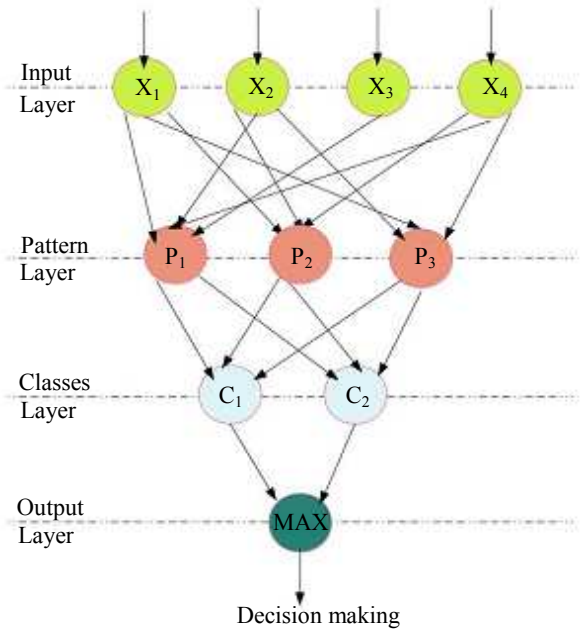

Fig. 3: Probabilistic Neural Network Architecture

The architecture of the Probabilistic Neural Network (PNN) has four layers namely (a) Input layer, (b) Pattern layer, (c) Classes layer and (d) Output layer which is depicted in Fig. 3. The PNN has its operations which link to Bayes-Parzen (Nandagopal *et al*., 2017; Fernandes *et al*., 2016; Syahputra *et al*., 2017) window probability density function (pdf) estimator. PNN uses the Parzen's method of calculating the density function for variables that are occurring spontaneously. According to the Bayes' theorem PNN is classified an unknown sample using the following:

$$CL_i.cost_i.p_{denfunci}(x) > CL_j.cost_j.p_{denfuncj}(x), j \neq i \quad (19)$$

where, CL is the class of the i^{th} population and $p_{denfunci}$ defines the density function of the sample x of the i^{th} population. Input layer contains the values that are got from the sensor nodes. These inputs are occurring spontaneously and they have to be assorted into any one type of population. The pattern layer will obtain the sensor values and bisect them into various classes which depends on the inaccuracy of the nodes. The classes layer is directly related to the pattern layer and the functionality of this layer depends upon the count of patterns of the pattern layer and make them to various classes with each classes having unique attribute of identification. This layer sums up all the patterns from each layer and fed that to the output layer. The output layer obtains all the patterns from the classes layer and this layer is responsible for the decision making process. Then the associated class label is determined by the following ways: (i) Sample values that are close to the unknown values should have a large contribution (increase the probability of classifying the unknown as that population) and (ii) Sample values that are present far away from the unknown should have a small contribution (decrease the probability of classifying the unknown as that population). So in every population the average value is calculated and then the unknown value is classified. The main advantage of employing PNN is that the unknown value measurements can be correlated and assorted using the probabilistic features which employs a polynomial mapping process.

Assortment Phase

After the Apprehension phase, each node association lead node performs the Probabilistic Neural Network computation. The collected values from the Apprehension phase is taken in to account and the erroneous nodes are identified and assorted. The Pattern layer collects the sensor values and these values can be normal functioning, likely to be unreliable or completely erroneous. This is assorted into three classes with each class representing one of the sensor values. The input layer collects the values and the pattern layer groups them based on the values obtained. Then each dendrites present in the patterns are watched and the average distance of the nodes are calculated. This distance subjected to Gaussian function gives the activation of the pattern layer dendrites. Each of the pattern layer values in the i^{th} class corresponds to a Gaussian function related to the i^{th} function. The activation of each class is taken by the dendrites associated with it. The classes layer adds all these values and finds the probability density function (pdf):

$$pdf(x) = \frac{1}{n}\frac{1}{(2\times 3.14)^{\frac{n}{2}}}\sum_{i=1}^{n}\exp-\frac{(x-xi)^2}{2\alpha^2} \quad (20)$$

where, x is the spontaneously occurring value and xi is the vector. N denotes the total number of assorting values and α is the balancing parameter. The main aspect of the assortment phase is that all the values must belong to only one class. All the three classes are computed for pdf value and are assigned a unique attribute namely gc_1, gc_2 and gc_3 respectively for class 1, class 2 and class 3. Therefore all the values will be evenly distributed among all the available classes which make spontaneous sample representation mandatory in PNN computation. All the three classes along with their sensor values are taken into account and the total value of computation is calculated as an standard average. The iteration aspect is absent in PNN so it can handle large memory loads and since computational weights are not mandatory, the execution process will be slow but the output will be unbiased.

a) Algorithm. Assortment Algorithm

(1) Initialized: Sample values and balancing parameter α for computation.
(2) Arrange the sample values into k sets ($k = 3$)
(3) Assign Class 1 (normal faults) ← $\{c_{11}, c_{12}, c_{13},\ldots.c_{1n}\}$
(4) Assign Class 2 (median faults) ← $\{c_{21}, c_{22}, c_{23},\ldots.c_{2n}\}$
(5) Assign Class 3 (Erroneous faults) ← $\{c_{31}, c_{32}, c_{33},\ldots.c_{3n}\}$
(6) Initialized: sample values for testing $\bar{x} \leftarrow \{x_1, x_2, x_3,\ldots.x_n\}$
(7) for each class $j = 1$ to k do
(8) for each sample value $i = 1$ to n do
(9) Compute Gaussian function value for GF_i
(10) end for
(11) end for
(12) for each class $j = 1$ to k do
(13) Calculate $pdf(x) = \frac{1}{n}\frac{1}{(2\times 3.14)^{\frac{n}{2}}}\sum_{i=1}^{n}\exp-\frac{(x-xi)^2}{2\alpha^2}$
(14) end for
(15) if $gc_1(x) > gc_2(x)$ && $gc_1(x) > gc_3(x)$ then
(16) sample values belong to class 1 error;
(17) else if $gc_2(x) > gc_1(x)$ && $gc_2(x) > gc_3(x)$ then
(18) sample values belong to class 2 error;
(19) else if $gc_3(x) > gc_1(x)$ && $gc_3(x) > gc_2(x)$ then
(20) sample values belong to class 3 error;
(21) else
(22) sample value belongs to spontaneously occurring error;
(23) endif
(24) Stop

Simulation Parameters

In this section, the performance of the proposed Wireless Nodal Behavior and Liability Investigation (WNBLI) protocol for wireless sensor networks is evaluated using Network simulator 2 (Issariyakul and Hossain, 2011). To get a clear working knowledge of the proposed protocol, it is simulated in the Network Simulator with the parameters suggested in the following Table 1. When a research work is conducted, it is always advisable to give a clear concise view of the operating scenario. It will help the experiment to take new forms when it comes to the evaluation part. The below listed parameters in Table 1 can be used to set up the operational environment which will express the proposed routing protocol better.

Initially 700 fault free sensor nodes are deployed in the network area. The heterogeneous faulty environment is created by adding different percentage of composite errors such as completely erroneous, soft unreliable nodes, intermittent errors and transient errors in the network. In this simulation, the soft unreliable nodes gives 81 to 100% erroneous results, the intermediate faulty node gives erroneous results for some random interval having 31 to 80% faulty values and the temporary error node gives erroneous results for a spike time interval having 5 to 30% faulty values. The composite error percentage are gradually increased from 5 to 40% in the network having 700 sensor nodes. The different faulty nodes percentage are shown in the Table 2.

Table 1: Simulation parameters

Parameters	Value
Number of sensor nodes	700
Deployment area	800×800 m^2
Resting time	1 sec
Simulation time	300 sec
Carrier sense range	200 m
Transmission range	90 m
Duration of control period	8 sec
Duration of control phase slot	0.3 ms
Duration of data slot	15 ms
Duration of RTS/CTS packet	0.12 ms
Duration of data packet	10.5 ms
Data packet size	32 bytes
Rx power	62.1 mW
Tx power	53.4 mW
Idle power	125 μW
Sleep power	20 μW
Channel rate	275 kbps
Source rate	8 pkt/s
Radio propagation model	Log-distance path loss
Path loss exponent (γ)	3-7
Shadowing deviation (σ_s)	7-4
Number of time tested	75

Table 2: Erroneous nodes in a network topology of 700 nodes

Error percentage	Number of malfunctioning nodes	Completely erroneous nodes	Soft unreliable nodes	Intermediate error nodes	Temporary error nodes
5	50	5	41	3	1
10	75	8	62	3	2
15	100	9	81	6	4
20	125	9	100	7	9
25	150	20	119	2	9
30	175	16	144	3	12
35	200	18	156	8	18
40	225	20	176	9	20

Results and Discussion

The proposed routing protocol is compared with the existing protocols in two ways. Firstly the characteristics of the proposed routing protocol are compared with the already existing protocols. Secondly the proposed protocol is compared with other protocols of same nature for its operational efficiency in the wireless sensor network environment along with their energy consumption nature. The operational efficiency of the proposed protocol is tested for its faulty node identification and self rectification. This self rectification is made possible by dissecting the working nature of the proposed algorithm into three distinct feature sets of dependent nature algorithms namely association, apprehension and assortment.

Comparison Based on the Routing Characteristics

Characteristics like their adjunct characteristics, routing path, route position and packet generation methodology is taken into account. The following table (Table 3) gives a clear comparison between various other protocols belonging to the same feature set as the proposed routing protocol.

Comparison Based on Fault Detection Accuracy and Energy Consumption

The proposed routing protocol is compared with two other algorithms namely Threshold sensitive Energy Efficient sensor Network (TEEN) protocol which is proposed by Manjeshwar and Agrawal (2001) and Base-station Controlled Dynamic Clustering Protocol (BCDCP) which is proposed by Muruganathan *et al*. (2005). The TEEN routing protocol is targeted towards reactive networks as the nodes operating with this routing protocol sense the network environment continuously. It can be applied to real time applications like intrusion detection and network explosion detection.

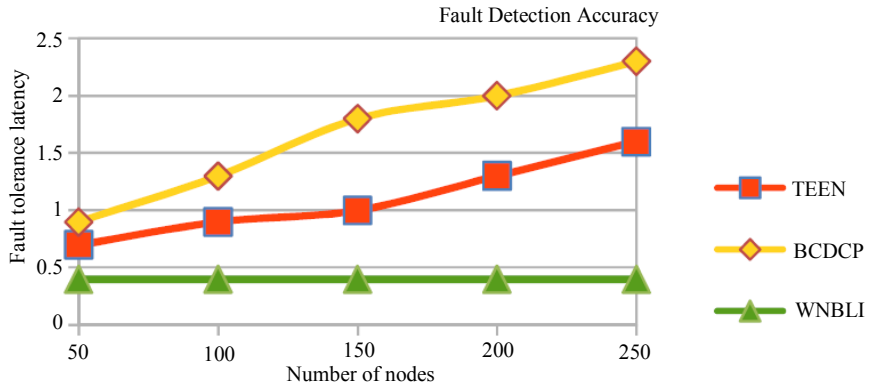

Fig. 4: Comparison of TEEN, BCDCP and WNBLI routing protocols for fault detection accuracy

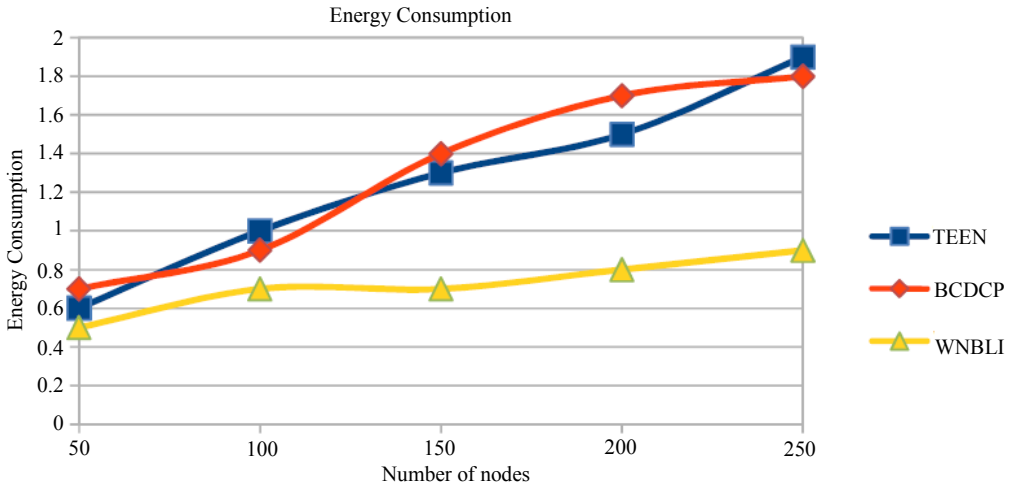

Fig. 5: Comparison of TEEN, BCDCP and WNBLI routing protocols for energy consumption

Table 3: Routing Evaluation of various Protocols

Protocols	Adjunct Characteristics	Route Path	Position	Packet
LEACH (Tyagi and Kumar, 2013)	Yes	Yes	No	No
EDAL (Yao *et al*., 2015)	Yes	Yes	Yes	Yes
ALBA-R (Petrioli *et al*., 2012)	Yes	Yes	Partial	Yes
R3E (Niu *et al*., 2014)	Yes	Yes	Yes	Yes
WNBLI	Yes	Yes	Yes	Yes

The BCDCP routing protocol distributes the energy abundance evenly among all sensor nodes to improve network lifetime and improve the average energy savings by the operating nodes. The proposed routing protocol is compared with TEEN and BCDCP routing protocols for Faulty node detection. The graph which compares the Fault detection accuracy among the routing protocols is depicted clearly in the Fig. 4.

The energy consumption criteria of the routing protocols are also taken into account for analysis. This energy consumption is compared with the number of faulty nodes existing in the wireless sensor networks. The graph which illustrates the energy comparison difference among the routing protocols is depicted in the Fig. 5.

Conclusion

In this research work, a static and analogous routing protocol with error detecting capability for wireless sensor network is proposed. The proposed Wireless Nodal Behavior and Liability Investigation (WNBLI) protocol is accomplished mainly in three major phases, such as Association phase, Apprehension phase and Assortment phase. A time division based multi channel media access control protocol is designed for data communication. The architecture of the proposed protocol is designed for a clustering method which is load balanced by nature. This working structure makes the proposed protocol consume less energy while operating in a network. The main thing to be noted in this research work is the point that the energy consumption factors play a very vital role in determining the outcome of a research. The node behavior is explained on the context of topology in which they belong so that the node operations can be analyzed for their utility. Once the node behavior is analyzed and an output is obtained, it is very easy to understand its operational functionality in the present topology. The proposed protocol is simulated in the network simulator and it is compared with other protocols operating at the same environment. The result shows that the proposed protocol is energy efficient and performance oriented for wireless sensor networks. By employing the proposed routing protocol, many prospects of the wireless sensor networks can be benefited.

Acknowledgment

The Corresponding author would like to thank the Research Department of SCSVMV University for the great support to conduct this research work under the guidance of Research Supervisor Dr. R. Uma Rani.

Authors Contributions

Tamilarasi B: Research plan, data collection, project plan, conducting the experiments, testing the results, manuscript written based on the work.

UmaRani R: Project plan, data analysis, testing the hypothesis, results of experiments and reviewing the manuscript written by the corresponding author.

Ethics

This research manuscript titled "WNBLI - A Multifarious Liability Examination for Wireless Sensor Networks" submitted to Journal of Science Publication is original and is not published in whole or in part elsewhere. There is no ethical issue involved in this research article.

References

Amin, R. and G.P. Biswas, 2016. A secure light weight scheme for user authentication and key agreement in multi-gateway based wireless sensor networks. Ad Hoc Netw., 36: 58-80. DOI: 10.1016/j.adhoc.2015.05.020

Azadeh, A., S.F. Ghaderi and S. Sohrabkhani, 2007. Forecasting electrical consumption by integration of neural network, time series and ANOVA. Applied Math. Comput., 186: 1753-1761. DOI: 10.1016/j.amc.2006.08.094

Baroutis, N. and M. Younis, 2017. Load-conscious maximization of base-station location privacy in wireless sensor networks. Comput. Netw., 124: 126-139. DOI: 10.1016/j.comnet.2017.06.021

Cvijović, Z., G. Radenković, V. Maksimović and B. Dimčić, 2005. Application of ANOVA method to precipitation behaviour studies. Mater. Sci. Eng. A, 397: 195-203. DOI: 10.1016/j.msea.2005.02.021

Fernandes, S.E.N., K.K.F. Setoue, H. Adeli and J.P. Papa, 2016. Fine-tuning enhanced probabilistic neural networks using metaheuristic-driven optimization. Bio-Inspired Comput. Applic. Image Process. DOI: 10.1016/B978-0-12-804536-7.00002-8

Gu, Y., F. Ren, Y. Ji and J. Li, 2016. The evolution of sink mobility management in wireless sensor networks: A survey. IEEE Commun. Surveys Tutor., 18: 507-524. DOI: 10.1109/COMST.2015.2388779

Gutierrez, J., J.F. Villa-Medina, A. Nieto-Garibay and M.A. Porta-Gandara, 2014. Automated irrigation system using a wireless sensor network and GPRS module. IEEE Trans. Instrument. Measurement, 63: 166-176. DOI: 10.1109/TIM.2013.2276487

Han, G., L. Liu, J. Jiang, L. Shu and G. Hancke, 2017. Analysis of energy-efficient connected target coverage algorithms for industrial wireless sensor networks. IEEE Trans. Industrial Informat., 13: 135-143. DOI: 10.1109/TII.2015.2513767

Issariyakul, T. and E. Hossain, 2011. Introduction to Network Simulator NS2. 2nd Edn., Springer Science and Business Media, ISBN-10: 1461414067, pp: 512.

Jan, M., P. Nanda, M. Usman and X. He, 2017. PAWN: A payload-based mutual authentication scheme for wireless sensor networks. Concurrency\Computation: Practice Experience.

Kurt, S., H.U. Yildiz, M. Yigit, B. Tavli and V.C. Gungor, 2017. Packet size optimization in wireless sensor networks for smart grid applications. IEEE Trans. Industrial Electron., 64: 2392-2401. DOI: 10.1109/TIE.2016.2619319

Li, M. and H.J. Lin, 2015. Design and implementation of smart home control systems based on wireless sensor networks and power line communications. IEEE Trans. Industrial Electron., 62: 4430-4442. DOI: 10.1109/TIE.2014.2379586

Li, S., L. Da Xu and X. Wang, 2013. Compressed sensing signal and data acquisition in wireless sensor networks and internet of things. IEEE Trans. Industrial Informat., 9: 2177-2186. DOI: 10.1109/TII.2012.2189222

Manjeshwar, A. and D.P. Agrawal, 2001. TEEN: A routing protocol for enhanced efficiency in wireless sensor networks. Proceedings of the 15th International Parallel and Distributed Processing Symposium, Apr. 23-27, IEEE Xplore Press, San Francisco, CA, USA, pp: 2009-2015. DOI: 10.1109/IPDPS.2001.925197

Mishra, S., C.N. Bhende and B.K. Panigrahi, 2008. Detection and classification of power quality disturbances using S-transform and probabilistic neural network. IEEE Trans. Power Delivery, 23: 280-287. DOI: 10.1109/TPWRD.2007.911125

Muruganathan, S.D., D.C. Ma, R.I. Bhasin and A.O. Fapojuwo, 2005. A centralized energy-efficient routing protocol for wireless sensor networks. IEEE Commun. Magazine, 43: S8-13. DOI: 10.1109/MCOM.2005.1404592

Nandagopal, M.G., E. Abraham and N. Selvaraju, 2017. Advanced neural network prediction and system identification of liquid-liquid flow patterns in circular microchannels with varying angle of confluence. Chem. Eng. J., 309: 850-865. DOI: 10.1016/j.cej.2016.10.106

Nikolidakis, S.A., D. Kandris, D.D. Vergados and C. Douligeris, 2013. Energy efficient routing in wireless sensor networks through balanced clustering. Algorithms, 6: 29-42. DOI:10.3390/a6010029

Niu, J., L. Cheng, Y. Gu, L. Shu and S.K. Das, 2014. R3E: Reliable reactive routing enhancement for wireless sensor networks. IEEE Trans. Industrial Informat., 10: 784-794. DOI: 10.1109/TII.2013.2261082

Njoya, A.N., C. Thron, J. Barry, W. Abdou and E. Tonye, *et al*., 2017. Efficient scalable sensor node placement algorithm for fixed target coverage applications of wireless sensor networks. IET Wireless Sensor Syst., DOI: 10.1049/iet-wss.2016.0076

O'brien, R.G., 1979. A general ANOVA method for robust tests of additive models for variances. J. Am. Stat. Assoc., 74: 877-880. DOI: 10.1080/01621459.1979.10481047

Ozcelik, B. and T. Erzurumlu, 2006. Comparison of the warpage optimization in the plastic injection molding using ANOVA, neural network model and genetic algorithm. J. Mater. Process. Technol., 171: 437-445. DOI: 10.1016/j.jmatprotec.2005.04.120

Petrioli, C., M. Nati, P. Casari, M. Zorzi and S. Basagni, 2014. ALBA-R: Load-balancing geographic routing around connectivity holes in wireless sensor networks. IEEE Trans. Parallel Distr. Syst., 25: 529-539. DOI: 10.1109/TPDS.2013.60

Prabhu, B., N. Balakumar and S. Sophia, 2016. Biologically inspired clustering mechanism in dense distributed wireless sensor networks. Int. J. Eng. Stud. Technical Approach.

Prathima, E.G., T.S. Prakash, K.R. Venugopal, S.S. Iyengar and L.M. Patnaik, 2016. SDAMQ: Secure data aggregation for multiple queries in wireless sensor networks. Proc. Comput. Sci., 89: 283-292. DOI: 10.1016/j.procs.2016.06.060

Rault, T., A. Bouabdallah and Y. Challal, 2014. Energy efficiency in wireless sensor networks: A top-down survey. Comput. Netw., 67: 104-122. DOI: 10.1016/j.comnet.2014.03.027

Shaikh, F.K. and S. Zeadally, 2016. Energy harvesting in wireless sensor networks: A comprehensive review. Renewable Sustainable Energy Rev., 55: 1041-1054. DOI: 10.1016/j.rser.2015.11.010

Silva, S., P. Costa, M. Gouvea, A. Lacerda and D. Leite, 2018. High impedance fault detection in power distribution systems using wavelet transform and evolving neural network. Electric Power Syst. Res., 154: 474-483. DOI: 10.1016/j.epsr.2017.08.039

Specht, D.F., 1988. Probabilistic neural networks for classification, mapping, or associative memory. Proceedings of the IEEE international conference neural networks, Jul. 24-27, IEEE Xplore Press, San Diego, CA, USA, pp: 525-532. DOI: 10.1109/ICNN.1988.23887

Swaminathan, A. and P. Vivekanandan, 2017. An Effective Lightweight Key Management (ELWKM) model for wireless sensor networks using distributed spanning tree structure. Asian J. Res. Soc. Sci. Humanities, 7: 749-770. DOI: 10.5958/2249-7315.2017.00126.5

Syahputra, M.F., I. Aulia and R.F. Rahmat, 2017. Hypertensive retinopathy identification from retinal fundus image using probabilistic neural network. Proceedings of the International Conference Advanced Informatics, Concepts, Theory and Applications, Aug. 16-18, IEEE Xplore Press, Denpasar, Indonesia, pp: 1-6. DOI: 10.1109/ICAICTA.2017.8090989

Tyagi, S. and N. Kumar, 2013. A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks. J. Netw. Comput. Applic., 36: 623-645. DOI: 10.1016/j.jnca.2012.12.001

Usman, M., M.A. Jan, X. He and P. Nanda, 2016. Data sharing in secure multimedia wireless sensor networks. Proceedings of the Trustcom/BigDataSE/ISPA, Aug. 23-26, IEEE Xplore Press, Tianjin, China, pp: 590-597. DOI: 10.1109/TrustCom.2016.0114

Weber, S., 2017. A slotted aloha message concentration protocol for wireless sensor networks. Proceedings of the IEEE Wireless Communications and Networking Conference, Mar. 19-22, IEEE Xplore Press, San Francisco, CA, USA, pp: 1-6. DOI: 10.1109/WCNC.2017.7925551

Wu, S.G., F.S. Bao, E.Y. Xu, Y.X. Wang and Y.F. Chang, 2007. A leaf recognition algorithm for plant classification using probabilistic neural network. Proceedings of the IEEE International Symposium Signal Processing and Information Technology, Dec. 15-18, IEEE Xplore Press, Giza, Egypt, pp: 11-16. DOI: 10.1109/ISSPIT.2007.4458016

Xiao, L., Q. Mao, P. Lan, X. Zang and Z. Liao, 2017. A Fault Diagnosis Method of Insulator String Based on Infrared Image Feature Extraction and Probabilistic Neural Network. Proceedings of the 10th International Conference Intelligent Computation Technology and Automation, Oct. 9-10, IEEE Xplore Press, Changsha, China, pp: 80-85. DOI: 10.1109/ICICTA.2017.25

Yao, Y., Q. Cao and A.V. Vasilakos, 2015. EDAL: An energy-efficient, delay-aware and lifetime-balancing data collection protocol for heterogeneous wireless sensor networks. IEEE/ACM Trans. Netw., 23: 810-823. DOI: 10.1109/TNET.2014.2306592

Zhang, H., H. Xing, J. Cheng, A. Nallanathan and V.C. Leung, 2016. Secure resource allocation for OFDMA two-way relay wireless sensor networks without and with cooperative jamming. IEEE Trans. Industrial Informat., 12: 1714-1725. DOI: 10.1109/TII.2015.2489610

Zhang, Z., M. Choi and G.E. Karniadakis, 2012. Error estimates for the ANOVA method with polynomial chaos interpolation: Tensor product functions. SIAM J. Scientific Comput., 34: A1165-A1186. DOI: 10.1137/100788859