International Journal of Pure and Applied Mathematics
Volume 117 No. 20 2017, 513-535
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

Location and Distance Aware Node Failure Discovery Mechanism for Wireless Sensor Network

[1]K.B. Manikandan and [2]N. Sasirekha

[1] Rathinam College of Arts and Science,
Eachanari, Coimbatore.

[2]PG Department of Computer Applications,
Vidyasagar College of Arts and Science,
Udumalpet.

Abstract

Recent years have witnessed an increasing interest in wireless sensor networks (WSNs) for various applications such as environmental monitoring and military field surveillance. WSN have number of sensor nodes that communicate wirelessly and it deployed to gather data for various environments. The communication gets violated when there is any breakup in the network due to hardware or technical issues like node failure. Node failure of sensor nodes needs to be detected to gain communication link. In existing system, node failure detection and recovery mechanism based on clustering technique (NFDM-CT) is introduced to handle the node failure. However it has communication overhead issues which need to be resolved better. To overcome the abovementioned issues, in this research Location Tracking Algorithm-Hybrid Particle Swarm Optimization Fire Fly algorithm (LTA-HPSOFFA) is proposed. In this work, the clustering formation is performed by using Efficient K-Means Clustering (EKMC) which minimizes number of clusters. It is used to cluster the sensor nodes by clubbing the distant clusters together effectively. Then, LTA is proposed to exactly estimate the location of the sensor nodes by computing the minimum distance. To improve the energy consumption, the proposed HPSOFFA elects the best CH (CH) node with location information. Then the node failure detection is done by using probabilistic detection approach which reduces the number of node failures occurred in the given network. Data replication of nodes helps in node failure recovery process by preventing the data loss. Thus the LTA-HPSOFA model decreases the data loss, energy consumption and end-to-end delay significantly through the reduction of clusters and accurate location information. The experimental result proves that the proposed LTA-HPSOFFA is superior to existing algorithm in terms of throughput, network lifetime and lower energy consumption, end to end delay performance.

Index: Location tracking algorithm, hybrid particle swarm optimization firefly algorithm, CH selection, WSN.

1. Introduction

Wireless sensor networks (WSNs) have shown a lot of potential for large-scale data acquisition. However, it's only within the last decade that mobility has been used together with these WSN nodes. This new form of mobile wireless sensor networks (MWSNs) enabled applications hadn't been considered before and it is a key enabling technology in the future of ubiquitous computing [1].

The wide applicability of sensor networks in areas such as safety, research and military have gained more interesting the research community [2]. Sensor network's ability to sense phenomena without human presence, in potential harsh or hostile environments, make them an invaluable resource. Research topics such as MAC (medium access control) protocols, localization techniques, synchronization methods and routing protocols have all been studied in some detail within the scope of static sensor networks. However, the addition of mobility gives rise to new constraints and challenges, which calls for novel approaches to these problems [3].

MWSNs generally use the many-to-one communication style, in which data is gathered from the sensors and sent to the sink. The mobility of the network can cause frequent topology changes, which makes the routing of data difficult. Medium access is also a challenge since the number of nodes within transmission range will vary with time[4]. However, the total number of nodes in the network is usually fixed and less likely to suffer node failure.

Heterogeneous wireless sensor network consists of sensor nodes with different ability, such as different computing power and sensing range. Compared with homogeneous WSN, deployment and topology control are more complex in heterogeneous WSN. Role of heterogeneous nodes in the wireless sensor network decreases response time and improve battery life time. Heterogeneous node resources fall into three types as follows. Computational Heterogeneity has more complex processor and memory resulting in better performance of difficult tasks. Link Heterogeneity pose high bandwidth and long distant transceiver promising reliable transmission. Energy Heterogeneity node is line powered (its battery is replaceable). Out of the above the energy heterogeneity is the most important, since computation and link heterogeneity consumes more energy.

Computation and link heterogeneity results in reduced response time as effect of decreased waiting time. As a rule of thumb, if heterogeneity is properly used in a network, response time is tripled and the lifetime of the network can be increased by 5-fold [5].

Topology without a fixed infrastructure and topological structure allows mobile nodes to create a temporary communication network in the form clusters. Clustering is the division of the network into different virtual groups, based on rules in order to discriminate the nodes allocated to different sub-networks. The main goals of clustering are to attain communication scalability for a large number of nodes and high mobility, spatial reuse and coordination of resources and Virtual communication backbone. In Each cluster, a specific node is elected as a CH(CH) based on single metric or combination of metrics such as identity, degree, mobility, weight, density, etc. The possibility of clustering methods primarily determines the complicatedness of CH selection. The CH plays the role of coordinator within its substructure. Each CH acts as a temporary base station within its cluster and communicates with other CHs. A cluster is therefore composed of a CH, gateways and members node [6].

Node failure is experienced when an individual node fails to operate when they lose its contact with the cluster. Node failure occurs due to various reasons such as hardware failure or software crash, the loss of network connectivity or the failure of a state transfer. Node Failure discovery methods mostly rely on periodic transmission of node status data or inferring node status based on passive information collection. After the failure node discovery the failed node need to be reconstructed for ensured communication. There are two common ways to recovering a node from failure: (i) Recovery by node repositioning and (ii) the deployment of additional nodes to restore connectivity after failures have occurred [7].

In [8] focused on the recovery from an articulation node failure over WSNs. It used centralized solution to recover from the network partition and restore the lost connectivity. Moreover, the approach considers that the WSN uses a multi-channel communication to minimize the interference ratio. It uses a WSN reorganization technique to overcome the problem of connectivity loss. This method improves the WSN performance by using the node failure recovery. The sink is responsible for the WSN reorganization and then the channels re-allocation. The recovery information computed by the sink is communicated to the isolated WSN segments by the use of a rotation technique of some nodes. These nodes are chosen based on their locations, and they are rotated in a cascade.

The location tracking problem is addressed in a partially synchronized, heterogeneous WSN, comprised of sensor nodes (SN) that have very short transmission ranges and no sense of timing, and mutually synchronized absolute position routers (APR).

The hybrid location tracking scheme, which is based on the combination of time of arrival measurements between the target and the APRs, and RSS measurements between the target and the SNs. The advantage of the scheme exploits the both time of arrival and received signal strength. The time of arrival and received signal strength measurements are used for the observation of the movement of the target, without the traditional triangle methods [9].

2. Related Work

In [10] Wang et al (2006), studied a cost minimization for locating mobile users under delay constraints in mobile wireless networks. Specifically, a new location tracking algorithm is developed to determine the position of mobile terminals under delay constraints, while minimizing the average locating cost based on a unimodal property. Above model not only results in minimum locating cost, but also has a lower computational complexity compared to existing algorithms.

In [11] Ghumare et al (2015), formulated an energy saving algorithm which is Low Energy Adaptive Cluster Hierarchy algorithm. Various concern parameters considered are Packet Drop, Packet Delivery ratio and Throughput. Packet Drop during communication is sensitive area in wireless sensor network and elimination of packet drop is one of the important parameter considered here. Because of that packet delivery ratio is eventually high. Throughput is high during this operation of target tracking.

In [12] Latiff et al (2007), presented an energy-aware clustering for wireless sensor networks using Particle Swarm Optimization (PSO) algorithm which is implemented at the base station. A new cost function is defined with the objective of simultaneously minimizing the intra-cluster distance and optimizing the energy consumption of the network.PSO algorithm with cost function gives a higher network lifetime and delivers more data to the base station compared to LEACH and LEACH-C, additionally produces better clustering by evenly allocating the CHs throughout the sensor network area. PSO single hop routing among CH node requires more energy, so that multihop routing among the cluster nodes can be implemented to improve the energy efficiency.

In [13] Baskaranet al (2015), designed a novel firefly heuristic to avoid the local minimum problem. Firefly heuristic is based on the light intensity produced by fireflies. The intensity of light produced is mapped to the objective function and hence fireflies with low intensity are attracted towards fireflies with higher light intensity. Above hybrid firefly algorithm, synchronous firefly algorithm is based on (i) ranked sexual reproduction capability of select fireflies, (ii) the fireflies created by this method having the best genes from the ranked fireflies. The advantages of the synchronous firefly algorithm arefaster convergence and avoidance of multiple local optima.One of the three qualities of service parameters (packet loss rate, end to end delay, and remaining energy) can be considered to build a robust minimization problem.

In [14] Jinet al (2016), considers a probabilistic approach and formulates two node failure detection schemes that systematically combine localized monitoring, location estimation and node collaboration. The trade-offs of the binary and non-binary feedback schemes are demonstrated. Above method achieves high failure detection rates, low false positive rates, and low communication overhead with the scenarios of regular transmission range. Probabilistic approach does not work when location information is not available or there are communication blackouts.

In [15] Rehena et al (2014), presents low-cost agent-based fault detection (query based)approaches which work independently without creating any hindrance for actual data packet routing in WSNs. Nodes make decisions about the failure solely based on local view of the network. Although query path is initiated by the sink, dead information about the nodes is disseminated towards the neighboring nodes. The algorithms are also suitable for multiple-sink network where each manager or sink controls a sub-region of the network. Implementation complexities are involved for multi sink networks.

In [16] Essam et al (2015), presents a Recovery algorithm that forms a topology with Increased Robustness against recurrent failure (RIR). RIR tolerates the failure of multiple connectivity-critical nodes through repositioning of healthy nodes. The approach favors substituting a failed node with one with the highest residual energy in order to sustain the network connectivity for the longest time. RIR models the recovery as a Minimum Cost Flow problem to determine the best set of node relocations for repairing the network topology while minimizing the motion overhead of the recovery process.RIR for factor in the coverage loss caused by relocated nodes, the possibility of having rough terrain and the power consumption rate of candidate nodes are not considered.

In [17] Angadi et al (2016), designed a scheme for fault tolerance in wireless sensor networks by controlling the topology. The algorithm first detects the faulty node in the computed shortest path by considering the parameters mobility and buffer size. If the fault node is found, then the alternative shortest path excluding faulty node is identified for successful transfer of data. The sensor node fails due to the less buffer size and the high mobility of the sensor nodes. The fault sensor nodes can be identified and eliminated efficiently, and alternate path is computed. Above scheme works well for only single fault sensor nodes and it is not good for multiple fault sensor nodes.

In [18] Cao et al (2008), evaluated a routing optimization scheme based on graph theory and particle swarm optimization algorithm for multi-hop wireless sensor network. This scheme synthesized the intuitionist advantages of graph theory and optimal search capability of PSO.CHs election methods are based on maximum residual energy and in turns and by probabilities separately.The ways to reduce energy consumption and to optimize the network topology can be taken into account to for improving the network lifetime.

In [19] Pitchaimanickamet al (2014), formulated a hybrid approach involving Bacteria Foraging Algorithm (BFA) and Particle Swarm Optimization (PSO) applied to traditional clustering based protocol like LEACH-C that forms the k-optimal clusters by identifying the suitable CH. This algorithm searches the random direction in the tumble behavior of each bacteria for using the local best and global best position obtained by PSO. BFPSO LEACH-C maximizes the network life time by increasing the number of alive nodes for longer period of time and also reduced energy consumption for data transmission.

3. Proposed Methodology

In this section, Location Tracking Algorithm- Hybrid Particle Swarm optimization Fire Fly Algorithm (LTA-HPSOFFA) is proposed. This research contains modules such as clustering formation using Enhanced K-means clustering, Location Tracking Algorithm, CH using HPSOFFA and node failure detection. Data replication for node failure recovery. The overall proposed methodology is detailed in the below sections.

Cluster Formation using Enhanced K-Means Clustering (EKMC)

In this research, the Enhanced k-means clustering (EKMC) method is proposed for efficient cluster formation. Owing to easy implementation and fast convergence, k-means clustering is an applicable clustering method specifically in mobile wireless sensor networks. Its objective is to minimize the average squared Euclidean distance from their cluster centres. The first step of K-means is to select initial cluster centres K. The algorithm follows a simple way to sort out a specific data group through a distinct number of clusters (assume k clusters). The main idea is to determine centroids, each centroids belongs to one cluster [20]. K-means algorithm is executed for cluster formation with the target WSN. Assume that the WSN of *n* nodes is divided into *k* clusters. First, *k* out of *n* nodes are randomly selected as the CHs. Each of the remaining nodes decides its CH nearest to it according to the Euclidean distance. After each of the nodes in the network is assigned to one of *k* clusters, the centroid of each cluster is calculated.

$$d = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (1)$$

$$centroid(X, Y) = \frac{1}{s}\sum_{i=1}^{s} x_i, \frac{1}{s}\sum_{i=1}^{s} y_i \quad (2)$$

Algorithm 1

Input: k the number of initial clusters

k′ the number of resulting clusters

n set of mobile nodes

Output: set of k′ clusters

Stage 1: basic k means

1. random centroid selection

Centroids =select (n, k)

2. for each mobile node in n calculate the Euclidean distance using (1)

c=min (dis (n, centroids))

Assign (n, $cluster_i$) end for

3. for i=1 to k

Compute centroid using (2)

$centroids_i$ = mean ($cluster_i$)

end for

4. repeat 2, 3 until all centroids does not update

Stage: 2merging

5. (i, j) = min(dis($centroids_i$, $centroids_j$))
6. Merge($cluster_i$, $centroids_j$)
7. k=k-1
8. Repeat 5, 6, 7 until k =k′

To avoid the time complexity and energy consumption in this research two-stage enhanced k-means algorithm is formulated. The first stage is just a basic k-means algorithm, but as many as enough data points are selected to be the initial centroids. The second stage is a merge stage, i.e., merge the k intermediate clusters produced by stage one into the final k′ result clusters. The overall block diagram for the proposed system is illustrated in the Figure 1.

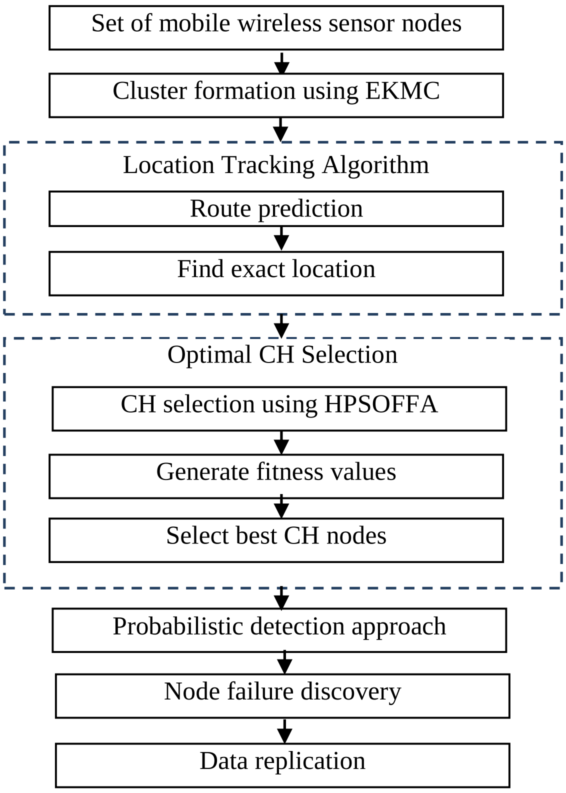

Figure 1: Overall block Diagram of the Proposed System

Location Tracking Algorithm (LTA)

In this section location tracking method is proposed for mobile nodes based on the route prediction and EKM clustering technique in the coverage of transmission range. As the location information is not available in the routing information additionally time of arrival and time difference of arrival techniques is used. The algorithm is developed to implement the concept of sensor node detection with the help of route prediction and clusters.

Route prediction technique predicts the route with the help of following properties.

Degree Centrality

No. of connections of node i with other nodes

$$k_i = C_d(i) = \sum_j^N x_{ij} \qquad (3)$$

Where N is number of nodes in network, C_d is degree of centrality and x_{ij} is distance between two source node o destination node

Closeness Centrality

Inverse sum of the shortest distance to all nodes from source

$$CC_i = \frac{1}{d_i} = \frac{N}{\sum_{j=1}^N d_{ij}} \qquad (4)$$

Where d_i is average distance from node i to all other nodes

Sensor node network checks for route prediction by the parameters degree centrality and closeness centrality. Above steps forms better clusters by clubbing the similar clusters and efficient CH selection minimizes the communication overhead by reducing packet transfer time delay to identify the location. Locations of source and destination nodes are recorded with time of arrival and time difference of arrival.

Let a destination node be denoted as A and sender node as B. The exact location of A is given as A(X, Y). Now node A can broadcast a notification of its presence to all other nodes in the communication range. Now node B estimates the position and reconfirms the presence of node A. The node B in proper functioning immediately updates it position B(X, Y) at the time of arrival of the broadcast notification of node A. The node B repeats this for N signaling periods. Finally, B applies the multilateral method to find the exact location of node A.

Algorithm 2

Let the current location of A be (X, Y)
Receive the first message from node A
Record the current location of node B as (x_0, y_0)
Record time of arrival as τ_0
While (B is sending the message)

```
{
        If (Time elapsed)
stop
else
   {
   Record the location of node A as (xi, yi)
   Record the time of arrival as τi
for (i=0; i<n; i++)
    {
     Calculate time difference of arrival τij
```

$$\tau_{ij} = \tau_i - (i-j) * \delta t$$

```
        }}}
Take four pairs of points of node A's instantaneous positions as:
(xi, yi) = [(x0, y0), (x1, y1), (x2,y2), (x3, y3)]
Ensure ((x0≠ x1≠x2≠x3) and (y0≠ y1≠y2≠y3);
```

After recording all these values, use the formula for getting the approximate value of the location of node A.

$$A_{i+1} = \frac{2}{c} * \left(\frac{x_{i+2}-x_i}{\tau_{i,i+2}} - \frac{x_{i+1}-x_i}{\tau_{i,i+1}}\right) \quad (5)$$

$$B_{i+1} = \frac{2}{c} * \left(\frac{y_{i+2}-y_i}{\tau_{i,i+2}} - \frac{y_{i+1}-y}{\tau_{i,i+1}}\right) \quad (6)$$

$$C_{i+1} = c * \left(\tau_{i,i+2}, \tau_{i,i+1}\right) - \frac{1}{c} * \left(\frac{x_{i+2}^2+y_{i+2}^2-x_i^2-y_i^2}{\tau_{i,i+2}} - \frac{x_{i+1}^2+y_{i+1}^2-x_i^2-y_i^2}{\tau_{i,i+1}}\right) \quad (7)$$

$$X_k = \frac{C_{i+1}*A_{i+2}-C_{i+2}*A_{i+1}}{B_{i+1}*A_{i+2}-B_{i+2}*A_{i+1}} \quad (8)$$

$$Y_k = \frac{C_{i+1}*B_{i+2}-C_{i+2}*B_{i+1}}{A_{i+1}*B_{i+2}-A_{i+2}*B_{i+1}} \quad (9)$$

Thus the location information of the source and destination nodes is estimated. In this LTA, the current locations are considered and messages are transmitted between source nodes to destination node.

The arrival time is recorded in a given time period and compute the location information using LTA. By recording all the transmission paths the distance of each and every node to destination is known. It helps to identify the nodes which are nearer to destination intentionally better parameter for CH selection. Thus the LTA is used to provide exact location for efficient communication over network.

CH Selection using HPSOFFA Algorithm

In this research, CH selection is performed by using HPSOFFA. It is focused to select the optimal CH for the resulted clusters. The parameters transmission range, energy and bandwidth availability which is computed in previous work when combined with additional location information leads to better selection of CH. While combining proposed HPSOFF algorithm with above metrics results in optimal CH selection.

1. Particle Swarm Optimization (PSO)

Inspired by the flocking and schooling patterns of birds and fish, Particle Swarm Optimization (PSO). Originally, these two started out developing computer software simulations of birds flocking around food sources then realized how well their algorithms worked on optimization problems.

The PSO is a computational approach that optimizes a problem in continuous, multidimensional search spaces. PSO starts with a swarm of random particles. Each particle is associated with a velocity. Particles' velocities are adjusted in order to the historical behavior of each particle and its neighbors during they fly through the search space. Thus, the particles have a tendency to move towards the better search space. The version of the utilized PSO algorithm is described mathematically by the following equations:

Each particle updates its own position and velocity according to formula (10) and (11) in every iteration.

$$v_{id}^{k+1} = \omega v_{id}^{k} + c_1\gamma_{1_1}\left(p_{id}^{k} - x_{id}^{k}\right) + c_2\gamma_{1_2}\left(p_{gd}^{k} - x_{id}^{k}\right) + \alpha(\mathrm{rand} - \tfrac{1}{2}) \qquad (10)$$

$$x_{id}^{k+1} = \begin{cases} 1 & s\left(v_{id}^{k+1}\right) > rand\ (0,1) \\ 0 & \text{else} \end{cases} \qquad (11)$$

where the $s\left(v_{id}^{k+1}\right)$is the sigmoid function $S(vid) = 1/(1 + \exp(-v_{id}))$,i = 1, 2, 3 ... m, m is the number of particles in the swarm, v_{id}^{k} and x_{id}^{k}stand for the velocity and position of the ith particle of the kth iteration, respectively. p_{id}^{k}denotes the previously best position of particle i, p_{gd}^{k} denotes the global best position of the swarm. ω is the inertia weight, c1 and c2 are acceleration constants (the general value of c1 and c2 are in the interval [0 2]), γ1 and γ2 are random numbers in the range [0 1].

Each feature subset can be considered as a point in feature space. The optimal point is the subset with least length and highest classification accuracy. The initial swarm is distributed randomly over the search space, each particle takes one position. The goal of particles is to fly to the best position. By passing the time, their position is changed by communicating with each other, and they search around the local best and global best position. Finally, they should converge on good, possibly optimal, positions since they have exploration ability that equip them to perform FS and discover optimal subsets.

The velocity of each particle is displayed as a positive integer; particle velocities are bounded to a maximum velocity V_{max}. It shows how many of features should be changed to be same as the global best point, in other words, the velocity of the particle moving toward the best position. The number of different features (bits) between two particles related to the difference between their positions.

After updating the velocity, a particle's position will be updated by the new velocity. Suppose that the new velocity is V. In this case, V bits of the particle are randomly changed, different from that of Pg. The particles then fly toward the global best while still exploring the search area, instead of simply being same as Pg. The V_{max} is used as a constraint to control the global exploration ability of particles. A larger V_{max} provides global exploration, while a smaller V_{max} increases local exploitation. When V_{max} is low, particles have difficulty getting out from locally optimal sections. If V_{max} is too high, swarm might fly past good solutions.

PSO which employed has lower convergence rate resulting in more number of iterations. So the searching efficiency over the entire network becomes the overhead by minimizing the optimal performance.

2. Fire Fly Algorithm (FFA)

FFA is one of the meta-heuristics swarm intelligence methods. Fireflies are the insects of Lampyridae family and they use bioluminescence to attract their mating partner and for getting prey.

The fireflies emit flashing lights which mainly act as a signaling system for attracting the other fireflies.

The light intensity of each firefly determine its brightness and hence its attractiveness. Attractiveness of the firefly is calculated using (12).

$$\beta\,(r) = \beta_0 e^{-\gamma r_{ij}} \qquad (12)$$

where$r_{ij} = d(x_i, x_j)$, a Euclidean distance between two data points iand j. In general, $\beta_0 \in [0,1]$, describes the fitness value o distance at $r = 0$, i.e., when two data points are found at the same point of search space S. The value of $\gamma \in [0,10]$ determines the variation of fitness value with increasing distance from communicated data points.

It is basically the light absorption coefficient and generally $\gamma \in [0,10]$.

The movement of the firefly i in the space which is attracted toward another firefly j is defined by using (14).

$$X_i = x_i + \beta_0 e^{-\gamma r_{ij}} + \alpha(\text{rand} - \tfrac{1}{2}) \qquad (13)$$

Where α is the randomization parameter in interval [0, 1] and rand is random number generator with numbers uniformly distributed in range [0, 1]. Parameter γ is controls the variation in attractiveness and define convergence.

The lower convergence rate of PSO is addressed using FFA in which the brightness and intensity behavior improves the PSO positions. Thus FFA increases the searching efficiency between source and destination nodes by reducing the number of iterations significantly.

3. Hybrid Particle Swarm Optimization Fire Fly Algorithm

(HPSOFFA)

To overcome the lower convergence rate of PSO the FF is hybrid and the HPSOFF is focused to improve the overall location tracking performance with better CH selection. FFA increases the convergence rate by concentrating on the brightness and intensity behavior and also used to increase the searching efficiency between source and destination nodes. Thus the number of iterations is reduced significantly. By using HPSOFFA, the optimal CH node is selected more effectively.

The pseudocode for combining swarm optimization with firefly algorithm is given below

Algorithm 3

Initializing a population with N individuals
Initialize the position and velocity of each particle (nodes) in the swarm
While Maximum Iterations is not reached do
Set algorithm factors:
higher packet delivery – H_{PDR}
lower end-to-end delivery – L_{EED}
lower energy consumption - L_{EC}
Objective function of f(x), from, where $x = (x_1, \ldots\ldots, x_d)^T$
Generate primary population of fireflies x_i $(i = 1, 2, \ldots, n)$
Describe light intensity of I_i at x_i via f (x_i)
While$(t < MaxGen)$ do
For $i = 1$ to n(all n fireflies (nodes));
For $j = 1$ to n(all n fireflies)
If $(I_j > I_i) f_i$ towards f_j
end if
Attractiveness vicissitudes with distance 'r' via Exp$[- r2]$
Estimation novel solutions and study light intensity;
End for j;
End for i;
Construct new CH

Evaluate the fitness of the new firefly solution which is directly proportional to its brightness
If the fitness value is better than its personal best (pBest)
Set current value as the new pBest
End
Choose the particle (node) with the best fitness value of all as gBest
For each particle (node)
Calculate particle velocity and update node position according (13) and (14)

Randomly selecting a $gbest$ for particle i from highest ranked solutions
Update the velocity and position of particle based on best firefly behavior (10) and (11)
Return most optimal CH
Update the pbest and gbest
Return the optimal CH
End

The above algorithm describes that the N number of nodes are taken for the given network. The objective function is considered as higher packet delivery ratio, lower energy consumption and lower end to end delay metrics. By using HPSOFF technique, the CH node is selected which has the capability of defined objective function. This hybrid algorithm generates better fitness function values and it selects the node as CH node which satisfies the threshold values. The PSO position is optimized by using fireflies' behavior and higher brightness as well as intensity values. Thus the HPSOFF provides optimal CH node to improve the packet transmission in the larger network.

Node Failure Detection using Probabilistic Detection Approach

In this research work, a probabilistic approach is adapted and proposes non binary feedback based node failure detection scheme that systematically combine localized monitoring, location estimation and node collaboration. The non-binary feedback scheme differs from the binary version in that A first gathers non-binary information from its neighbors and then calculates the conditional probability that B has failed using all the information jointly. Consider the case where none of A's neighbors has heard about B (otherwise, the case is trivial as we will describe soon). Specifically, suppose A receives responses from n – 1 neighbors about B. Without loss of generality, denote these n nodes (i.e., A and its n - 1 neighbors) as 1,….,n. For time t + 1, let $C_{i,j}$ denote the event that the i-th node does not hear the j-th heartbeat packet from B; let $P_{c,K}^{(i)}$ denote the probability that all the K heartbeat packets from B to node I are lost (K ≥ 1); let R_i denote the event that the i-th node is in the transmission range of B. Recall that D denotes the event that B fails at time t+1. Then A calculates the following probability:

$$P\left(D\middle|\bar{C}_{1,1},\ldots.\bar{C}_{1,K},\ldots.\bar{C}_{n,1},\ldots\bar{C}_{n,K}\right) = \overline{P\left(\bar{C}_{1,1},\ldots.\bar{C}_{1,K},\ldots.\bar{C}_{n,1},\ldots\bar{C}_{n,K}\right)} \quad (15)$$

Summarizing the above, in the non-binary scheme, A's neighbor, i, responds the following information to A. If it has heard from B at time t + 1, then it sends a single bit 0 to A (same as that in the binary feedback scheme). Otherwise, it sends $P_{c,K}^{(i)}$ and P (R_i) to A. If A receives a bit 0 from one of its neighbors, then it knows that B is alive. Otherwise, it obtains the probability that B has failed. If the probability is larger than thresholdθ, then A generates analarm that B has failed and sends it to the manager node.Algorithm 1summarizes the actions related to sending aquery message and the actions after hearing responses on the query. Algorithm 2 summarizes how a node responds toa query message. For the same reason as explained for the binary scheme, this scheme is insensitive to the choice of the detection threshold,θ and uses the same mechanism forforwarding the alarm to the manager node.

Algorithm 4

Non-binary feedback scheme (sending query)

Suppose A hears from B at t but not t+1

A calculates p, probability that B fails

If (p≥θ) then

A starts a timer with a random timeout value

If A has not heard a query about B when the timer times out then

A broadcast an inquiry about B

If A receives at least one response of 0 then

A does nothing (B is alive)

Else

A updates p based on feedbacks

If (p≥θ) then

A sends a failure alarm about B to the manager node

End if

End if

End if

End if

Algorithm 5

Non-binary feedback scheme (receiving query)

Suppose C receives a query message about B

If C has just heard from B then

C responds with 0

Else

C responds with the probability that all k messages from B to C are lost and the probability that C is in B's transmission range

End if

Data Replication for Node Failure Recovery

Node failure becomes a main concern when communication breakups in a group of nodes. The above methods detect the failure node in an efficient way but node failure recovery in a whole network still remains untouched. Node failure recovery is the process of restoring the data of a failure node to a fully functional state after one or more nodes in the system has experienced software or hardware related failure. It is concentrated for improving reliability, fault tolerance or its accessibility.

Recovering of failure node avoids the time delay when communication process takes place. The main purpose of recovery is to provide an immediate response to request that is to reduce time delay that involves route recovery to find the backup of the data incase in need of data recovery and to avoid the data loss which is generated at the time of failure. To recover a failure node memory replication scheme is introduced. Using the available location information of the destination nodes, replication of the data is made to node present in the route which has the needed resources or capacity.

To find the capacity or available resources of the nodes is found using the fuzzy rule [21].Fuzzy rule base is the collection of linguistic control rules where nodes make decisions on the basis of fuzzy control rules. The defuzzifier collects aggregated value and generates anon-fuzzy control which presents the types of nodes, such as appropriate node or inappropriate node. The input conditions are available memory and energy consumption. The available fuzzy conditions are

1. IF memory = high and energy = high THEN P_r(high)
2. IF memory = high and energy = low THEN P_r(medium)
3. IF memory = low and energy = high THEN P_r(medium)
4. IF memory = low and energy = low THEN P_r(low)

Where P_r is the probability of replication.

From the above rules it is interfered that high availability of memory and energy gives a higher probability of data replication to the node, combination of higher and lower energy and memory availability gives a medium probability of data replication and lower availability of energy and memory availability gives a lower probability of data replication.

After finding the eligible node to replicate then data replication is made on a compressed format. The data compression of the original node is performed for the lower resource consumption while storing the data. The main factors such as memory, energy are reduced much more when the data replication is done in a compressed manner. It also reduces the time and bandwidth which is required to transfer or replicate data from one node to another. The steps which are involved in replication of data is given below.

Algorithm 6

```
Node n1's original data x
n1 transfers x to n2
    At each ni
        If receive first replica of n1
        {
        Selects n2 based on fuzzy rule
        Transfers x to n2
        }
        If receive second replica of n1
        {
        do nothing
        }
```

The working of the given data replication pseudocode is, in a collection of sensor nodes, x is the original data of the node $n_1$initially node $n_1$have the compressed data of it. Then $n_1$transfers its data to n_2 to store it for future retrieval. Node $n_2$is selected based on the fuzzy rule which concentrates on the energy and memory availability. If node n_1 transfers the first replica of its own data then node $n_2$saves it. If it is second replica nothing is done further. This kind of data replication helps the sensor nodes in the network to recover the original data if any failures occur mainly by avoiding data loss which increases the reliability of the network of sensor nodes.

4. Experimental Result

In this section, a quantitative performance study is presented. In all the simulations, the nodes move in a 500m * 500m square area. The total number of nodes, N, is varied from 20 to 150. The initial locations of the nodes follow a 2D Poisson distribution. The transmission range of a node is circular with the radius, r, varied from 30m to 130m. The above combination of parameters lead to a wide range of neighborhood density for evaluating our approach The performance measures that are considered in this work evaluating its performance improvement over existing methodology, probabilistic NFDM-CTare listed as follows: detection rate, false positive rate, Communication overhead, energy consumption, end to end delay.

Detection Rate

Detection rate is defined as how well the proposed research scenario can find the number of nodes failed in the network environment for the varying number of nodes.

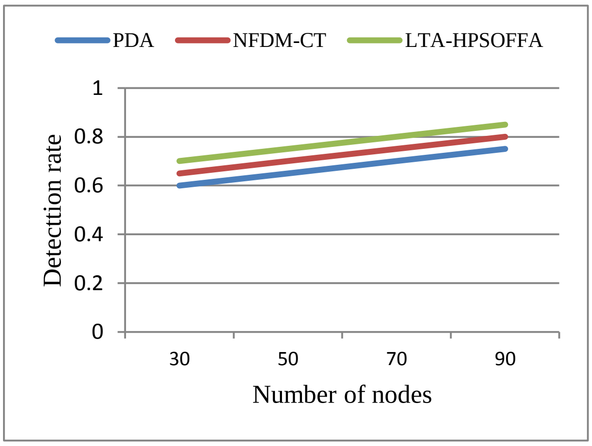

Figure 2: Detection Rate Comparison

Figure 2 shows the comparison of detection rate between the existing and proposed algorithm for node failure detection. The number of nodes is plotted in x-axis and detection rate is plotted in y-axis. Number of nodes varies from 30 to 90 for the existing and proposed approaches. The existing PDA and NFDM-CT methods provide lower detection rates whereas the proposed LTA-HPSOFFA method provides higher detection rates. LTA-HPSOFFA provides higher detection rate performance due to the coverage of sensor nodes over transmission range. Thus the result concludes that the proposed LTA-HPSOFFA has better location tracking performance rather than the existing algorithms.

False Positive Rate

False positive rate usually refers to the probability of falsely rejecting the null hypothesis for a particular test. The much lower false positive rate under our scheme is because of its ability to differentiate a node failure from the node moving out of the transmission range, while the existing scheme cannot differentiate these two cases.

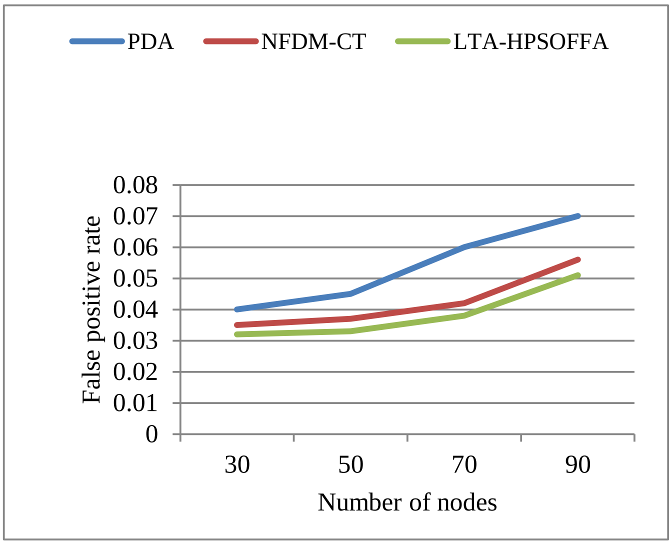

Figure 3: False Positive Rate

Figure 3 shows the comparison of false positive rate between the existing and proposed algorithm for node failure detection. The number of nodes is plotted in x-axis and false positive rate is plotted in y-axis. Number of nodes varies from 30 to 90 for the existing and proposed approaches. Proposed LTA-HPSOFFA provides lower false positive rate whereas PDA and NFDM-CT provides higher false positive rate. LTA-HPSOFFA provides lower false positive rate than the existing approaches due to the ability to differentiate a node failure from the node moving out of the transmission range. Better CH selection performance is achieved using the LTA-HPSOFA.

Energy Consumption

Energy consumption is the average energy required for sending, receiving or forward operations of a packet to a node in the network during the period of time.

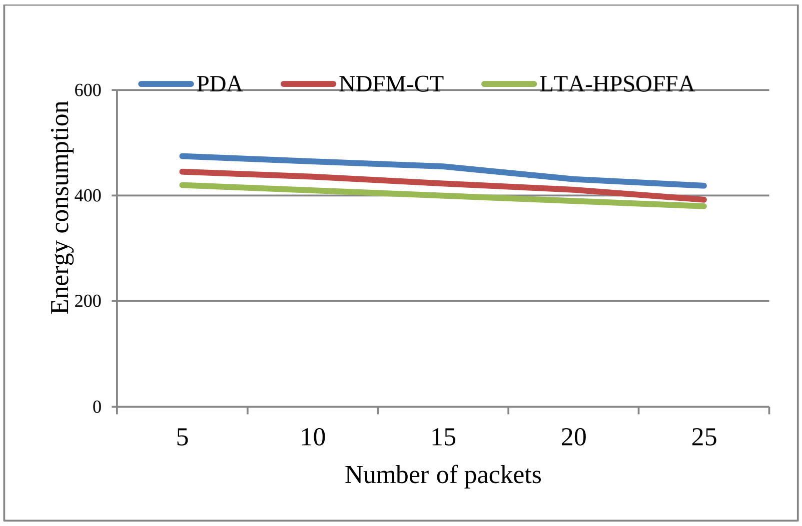

Figure 4: Energy Consumption

Figure 4 shows the comparison of energy consumption between the existing and proposed algorithm for node failure detection. The number of packet is plotted in x-axis and energy consumption is plotted in y-axis. PDA and NFDM-CT requires more energy compared to that of energy required in the proposed LTA-HPSOFFA. LTA-HPSOFFA provides reasonable lower energy consumption than the existing approaches because of the available location information with source nodes. LTA-HPSOFFA's Lower energy consumption increases the efficiency of the network.

End-to-End Delay

End-to-end delay: The average time which is incurred by a packet to be transmitted from source to destination through the network is known as the End to End delay.

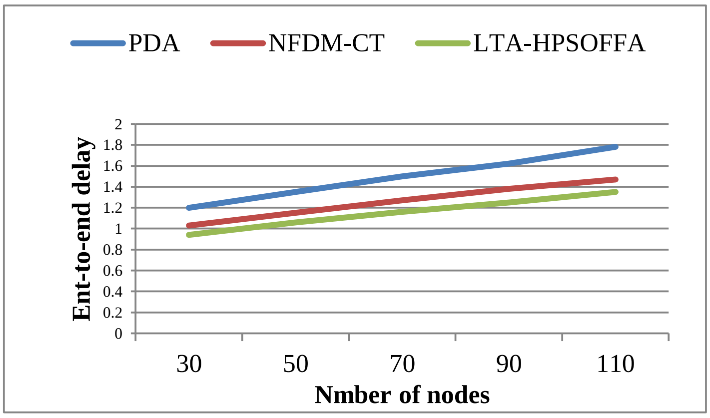

Figure 5: End-to-End Delay

Figure5 shows the comparison of end-to-end delay between the existing and proposed algorithm for node failure detection. The number of node is plotted in x-axis and end-to-end delay is plotted in y-axis. LTA-HPSOFFA reduces the End-to-end delay whereas PDA and NFDM-CT has higher end-to-end delay. LTA-HPSOFFA provides reasonable lower end-to-end delay than the existing approaches because the source node already knows the shortest path to forward the packet. Thus LTA-HPSOFFA increases the network usage in a better way.

Throughput

The rate with which the data packets get transmitted successfully over the network or communication links is defined as the throughput. It is measured in bits per second (bit/s or bps). It is also indicated by the units of information that are processed over a particular time slot.

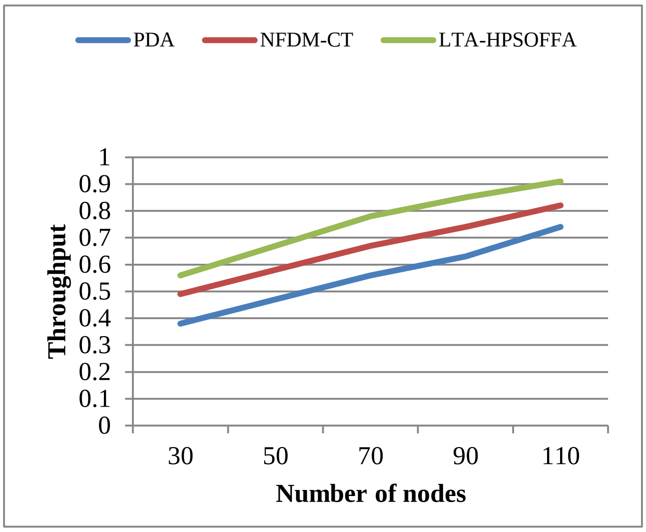

Figure 6: Throughput

Figure 6 shows the comparison of throughput between the existing and proposed algorithm. The number of node is plotted in x- axis and throughput is plotted in y- axis. LTA-HPSOFFA has higher throughput when compared with PDA and NFDM-CT method. The overall performance of the sensor network is increases gradually in proposed system than existing system since the location information and optimal CH node selection which is nearer to the destination.

5. Conclusion

In this research, location information and optimal CH election considered factors. In this network, node failure detection becomes the major task. Previous research handled node failure detection with node importance level clustering technique but the overall performance gets degraded due to large number of groups. To address the above issues enhanced clustering and location based CH selection is proposed. EKMC groups the similar clusters into single cluster thus minimizing the cluster count. Location estimation model records the location information of each node which helps the other nodes to forward packet with lower energy. Optimal CH selection is done with the location information and it improves the energy consumption, end to end delay and transmission range. Probalistic detection approach handles the node failure discovery which reduces the failure of nodes. Data replication minimizes the risks which are involved in the node failure recovery process by providing the replica of the data of the failed node. The overall performance of the sensor networks is improved in the terms of higher throughput, lower energy consumption, lower end-to-end delay, higher detection rate and lower false positive rate compared with existing methods. In this research, node failure detection is dealt in good way but the data which has lost due to node failure is not recovered. It may incur some security concerns to the networks, so secure data transfer with recovery mechanism can be further developed.

References

[1] Munir S.A., Ren B., Jiao W., Wang B., Xie D., Ma J., Mobile wireless sensor network: Architecture and enabling technologies for ubiquitous computing, In IEEE 21st International Conference on Advanced Information Networking and Applications Workshops (2007), 113-120.

[2] Aldeer M.M., A summary survey on recent applications of wireless sensor networks, IEEE Student Conference on Research and Development (SCOReD) (2013), 485-490.

[3] Mao G., Fidan B., Anderson B.D., Wireless sensor network localization techniques, Computer networks 51 (10) (2007), 2529-2553.

[4] Balobaid A., A survey and comparative study on different energy efficient MAC-protocols for Wireless Sensor Networks, IEEE International Conference on Internet of Things and Applications (IOTA) (2016), 321-326.

[5] Sasikumar P., Khara S., K-means clustering in wireless sensor networks, IEEE fourth international conference on Computational intelligence and communication networks (CICN) (2012), 140-144.

[6] Younis O., Krunz M., Ramasubramanian S., Node clustering in wireless sensor networks: Recent developments and deployment challenges, IEEE network 20 (3) (2006), 20-25.

[7] Younis M., Senturk I.F., Akkaya K., Lee S., Senel F., Topology management techniques for tolerating node failures in wireless sensor networks: A survey, Computer Networks 58 (2014), 254-283.

[8] Chouikhi S., El Korbi I., Ghamri-Doudane Y., Saidane L.A., Articulation node failure recovery for multi-channel wireless sensor networks, IEEE Global Communications Conference (GLOBECOM) (2015), 1-7.

[9] Song Y., Yu H., A new hybrid TOA/RSS location tracking algorithm for wireless sensor network, IEEE 9th International Conference on Signal Processing (2008), 2645-2648.

[10] Wang W., Xue G., A cost-minimization algorithm for fast location tracking in mobile wireless networks, Computer Networks 50 (15) (2006), 2713-2726.

[11] Ghumare S.S., Labade R.P., Gagare S.R., Importance of localization for tracking mobile target in wireless sensor network, IEEE International Conference on Information Processing (ICIP) (2015), 145-150.

[12] Latiff N.M.A., Charalampos C.T., Bayan S.S., Energy-aware clustering for wireless sensor networks using particle swarm optimization, IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications (2007).

[13] Baskaran M., Sadagopan C., Synchronous firefly algorithm for cluster head selection in WSN, The Journal of Scientific World (2015).

[14] Jin R., Wang B., Wei W., Zhang X., Chen X., Bar-Shalom Y., Willett P., Detecting Node Failures in Mobile Wireless Networks: A Probabilistic Approach, IEEE Transactions on Mobile Computing 15 (7) (2016), 1647-1660.

[15] Rehena Z., Mukherjee R., Roy S., Mukherjee N., Detection of node failure in wireless sensor networks, In Applications and Innovations in Mobile Computing (AIMoC) (2014), 133-138.

[16] Essam H., Younis M., Shaaban E., Minimum cost flow solution for tolerating multiple node failures in wireless sensor networks, IEEE International Conference on Communications (ICC) (2015), 6475-6480.

[17] Angadi B.M., Kakkasageri M.S., Kori, G.S., Topology control scheme for fault tolerance in wireless sensor network, IEEE International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES) (2016), 1355-1360.

[18] Cao X., Zhang H., Shi J., Cui G., Cluster heads election analysis for multi-hop wireless sensor networks based on weighted graph and particle swarm optimization, IEEE Fourth International Conference on Natural Computation (2008), 599-603.

[19] Pitchaimanickam B., Radhakrishnan S., A hybrid bacteria foraging using Particle Swarm Optimization algorithm for clustering in wireless sensor networks, IEEE International Conference on Science Engineering and Management Research (ICSEMR) (2014), 1-6.

[20] Meena Kowshalya A., Sukanya, A., Clustering algorithms for heterogeneous wireless sensor networks-A brief survey, International Journal of Ad Hoc, Sensor & Ubiquitous Computing 2(3) (2011).

[21] Ibrahim K., Weigle M.C., CASCADE: Cluster-based accurate syntactic compression of aggregated data in VANETs, IEEE GLOBECOM Workshops (2008), 1-10.