Heliyon 10 (2024) e25998

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Multi-factor optimized mobile sink data collection framework for hybrid WSN-LTE assisted IoT network

Saranga Mohan *, Sunita Panda

Department of Electrical, Electronics, and Communication Engineering, GITAM School of Technology, GITAM (Deemed to Be University), Bengaluru, India

ARTICLE INFO

Keywords:
Clustering optimization
Graph theory
Internet of things
Multi-criteria quality of service
Mobile sink
Wireless sensor networks

ABSTRACT

The convergence of wireless sensor network-assisted Internet of Things has diverse applications. In most applications, the sensors are battery-powered, and it is necessary to use the energy judiciously to extend their functional duration effectively. Mobile sinks-based data collection is used to extend the lifespan of these networks. But providing a scalable and effective solution with consideration for multi-criteria factors of quality of service and lifetime maximization is still a challenge. This work addresses this problem with a hybrid wireless sensor network-Long term evolution assisted architecture. The problem of maximizing lifetime and providing multi-factor quality of service is solved as a two-stage optimization problem involving clustering and data collection path scheduling. Hybrid meta-heuristics is used to solve the clustering optimization problem. Minimal Steiner tree-based graph theory is applied to schedule the data collection path for sinks. Unlike existing works, the lifetime maximization without QoS degradation is addressed by hybridizing multiple approaches of multi-criteria optimal clustering, optimal path scheduling, and network adaptive traffic class-based data scheduling. This hybridization helps to extend the lifetime and enhance the QoS regarding packet delivery within the proposed solution. Through simulation analysis, the introduced approach yields a noteworthy increase of at least 6% and reduces packet delivery delay by 26% compared to existing methodologies.

1. Introduction

Wireless sensor network-assisted Internet of Things (WSN- IoT) is used in many applications like smart parking, health care monitoring, security surveillance, wildlife monitoring, etc. [1]. In these applications, the complexity of providing an Internet of Things (IoT) interface at each node is vast. Thus, the IoT interface is provided at only designated nodes (sinks), and the rest of the nodes communicate to these designated nodes using multi-hop forwarding. In these networks, sensor nodes are powered by batteries whose energy reduces due to transmission and reception. If the routing protocol uses the nodes' energy fairly, the node's energy can drain faster, reducing the network's lifetime. In unattended deployments, replacing the battery is complex, and the network becomes unreliable with holes in the network due to the energy loss of nodes [2].

Utilizing mobile sink-based data collection is the optimal method for mitigating energy consumption and enhancing the operational lifespan within WSN-IoT networks [3]. The data collection using a mobile sink operates in rendezvous [4] and cluster-based [5] modes. In rendezvous mode, data collection from the nearby nodes is taken up by the designated nodes situated along the trajectory path of the

* Corresponding author.
E-mail address: msaranga@gitam.edu (S. Mohan).

https://doi.org/10.1016/j.heliyon.2024.e25998
Received 8 May 2023; Received in revised form 30 December 2023; Accepted 6 February 2024
Available online 15 February 2024

mobile sink. This strategy yields energy savings on a macroscopic scale by diminishing the requisite number of data hops between nodes and the sink. However, this method exhibits a trade-off in data transmission delay. On the other hand, the clustering-based approach calls for choosing cluster leaders inside the network. The cluster members forward their data to the respective cluster head, which then orchestrates the routing of this data toward the mobile sink. This approach concedes energy efficiency in favor of reducing data transmission delays. The energy spent for the packet transmission from the cluster head to the sink can be even more reduced by the optimal placement of cluster heads, effective trajectory path for the sink, and integration of Long term evolution (LTE) mode in cluster heads. Integration of LTE provides a backhaul for data transfer to sink in case of congestion in the WSN interface for delay-sensitive applications.

No such hybridization involves multi-level grouping of nodes based on multifactor optimization, path planning across these groups to collect data using a mobile sink, and adaptive data forwarding decisions based on packet deadline and network dynamics. Such hybridization can improve the lifespan and the network QoS for packet delivery. Realizing this gap in the existing works, this work proposes a multi-criteria Quality of service (QoS) optimized mobile sink data collection framework for a hybrid WSN-LTE-assisted IoT network.

The proposed multi-criteria factor optimized mobile sink data collection framework solves the problem of optimizing multi-criteria factors as a two-stage optimization problem. To identify the ideal cluster heads, hybrid metaheuristics are used in the initial stage. The optimal path for data collection is planned in the second stage, applying the concepts of minimal Steiner points along with Particle swarm optimization (PSO). At each cluster head, the decision to use the WSN interface/LTE backhaul interface is made based on the priority of the packets and cost minimization of the LTE interface. The contributions of this paperwork are listed below.

(i) An innovative hybrid meta-heuristic approach for multi-criteria QoS optimized cluster head selection in a hybrid WSN-LTE supported IoT network. Bat algorithms with exploration capability and PSO with exploitation capability are integrated to strengthen the challenge of identifying optimal cluster heads. In this way, local minimal problems in using a single optimization algorithm and resulting sub-optimal cluster head selection are solved in the proposed solution. Among multiple factors for optimization, utmost emphasis is placed on the critical goal of energy consumption minimization. Thereby with the optimal cluster heads, a noticeable decrease in energy use and a corresponding notable extension of the network's operational lifespan is achieved.

(ii) A novel Graph theory integrated PSO algorithm for optimal trajectory path planning of mobile sink. Cluster head visits to collect data are solved as a Steiner minimal tree problem using a greedy heuristics algorithm. The data collection path with minimal delay is found, and the sink is scheduled on this path.

(iii) A multi-criteria LTE backhaul selection algorithm for data forwarding which balances between LTE and WSN interface usage based on delay deadline. Two different traffic classes, ordinary and emergency, are assigned to packet flows. Based on the traffic class and the projected delay in the path, a decision is made between saving the packets in the cluster head till the mobile sink collects the data or forward the packet on LTE backhaul. Delay projection is based on the probability distribution of delay over the past time interval. In this way, the delay for emergency packets is minimized. LTE integration with controlled use of it balancing between the data cost and the delivery deadline has not been considered in any of the earlier works.

The subsequent sections of this paper follow a structured sequence. Section II delivers an extensive overview of prior research efforts to improve QoS and lifetime in WSN-assisted IoT networks. In section III, the proposed multi-criteria factor optimized mobile sink data collection framework is presented. The empirical results of the suggested solution and a comparative analysis against established methodologies are conducted in section IV. Finally, Section V concludes the work by summarising the findings and identifying prospective future research directions.

2. Related work

The literature survey is presented as two categories: i) clustering-based solutions ii) sink optimization-based solutions, both of which are explored to assess their effectiveness in optimizing energy usage.

A. Clustering-based solutions

For WSN-assisted IoT networks, Shukla et al. [6] suggested a multi-tier clustering approach in which the network is partioned into subregions. As network size increases, the number of subregions is also increased to avoid long distance communication. Two types of aggregation nodes, Cluster heads, and Relay nodes are selected in the network. Data from nodes are sent via multi-hop forwarding to sink via CH. The network's lifespan decreases as the network grows because more hops must be made to reach the sink. Using a single mobile sink, Wei et al. [7] optimized clustering-based data collecting in WSN networks. Clustering is done using a modified LEACH clustering. Each node is scored based on the integrated energy density and motion performance function, which helps to select the cluster head. Toor et al. [8] presented a cluster-based routing system for WSN energy optimization. The clusters are divided as sectors, and each cluster's highest residual energy node is chosen as the cluster head in which a mobile sink collects data for each sector. The delay is higher due to the store and forward mode adopted for packet transmission. A spanning tree approach is used by Wang et al. [9], which is constructed using the sensor network with sensor nodes as vertices and the distance between nodes as weights. Data collection points along the spanning tree are selected to optimize the data collection time. The approach does not handle mobility. The data collection point selection is based on the sensor's sensing rate, which needs to be fixed and cannot be varied. Data collection using a mobile sink for sensor nodes was presented by Wang et al. [10], where the network is divided into sectors, and a cluster head is selected for each sector. The mobile sink moves at a constant angular velocity to collect data and forwards the data to the cluster head of its corresponding sector. The delay is more significant in this network, and the method cannot be scaled to massive networks. Fu et al. [11] presented a mobile sink-based data collection scheme for clustered network topology in which density-based clustering is used to cluster the network. A cluster head is selected which has the highest residual energy node within each cluster. PSO is utilised to plan the best route between cluster heads for data gathering. Cluster rebalancing is done using a splitting/merging strategy. The approach has a higher delay for data collection and creates buffer overflow at the cluster head. Yarinezhad et al. [12] suggested a mobile sink data-gathering technique for two-level clustering topology based sensor networks. Data is forwarded from nodes to their cluster head to a relay node and from there to the mobile sink. The delay is higher due to the two-level hierarchy and store/forwarding at the relay node. Zhang et al. [13] clustered the network and created a spanning tree connecting the cluster heads. Path maintenance effort is high over a long time due to re-clustering for energy balancing. A similar approach was presented by Wang et al. [14], where the network is clustered, and a data collection path connecting clustering heads is created using improved ant colony optimization approach. Path design does not consider the energy rebalancing of a cluster. Thomas et al. [15] considered the sensor network as a node and created a path connecting nodes by solving it as a traveling salesman problem. Though energy is optimized, the delay is high in this solution. Bencan et al. [16] clustered the network based on energy optimization. The path connecting cluster heads is optimized for delay using game theory. All the packets at the cluster head are forwarded directly to the sink, which increases the energy consumption. Kumar et al. [17] clustered the network using Fuzzy C means and selected cluster heads using an advanced LEACH algorithm. The approach was designed to maximize cluster head energy without considering QoS factors like delay.

B. Sink optimization-based solutions.

A mobile sink data-gathering strategy for wireless networks was put forth by Wang et al. [18]. Sink stations are data collection points and send data collection intention to neighbors. Nodes cache the data at themselves and forward the cached data upon receiving data collection intention from the sink. Gharaei et al. [19] deployed a data collection mechanism using two mobile sinks which travel in a spiral path in the opposite direction and advertise their data collection intention. Node forwards data to sink in a multi-hop manner on receiving data collection intention. Though the delay is reduced, it is not scalable, and the cost of an additional mobile sink is high. Najjar et al. [20] used a mobile sink to increase energy efficiency in data collection. The network is clustered using ant colony optimization and agglomerative clustering. The data collection path connecting the clustering head is created using a genetic algorithm. The delay in the solution is higher as clustering is considered only energy optimization. A comparable approach with a data-gathering path that is energy-efficient was put forth by Wen et al. [21]. Konstantopoulos et al. [22] addressed the delay problem

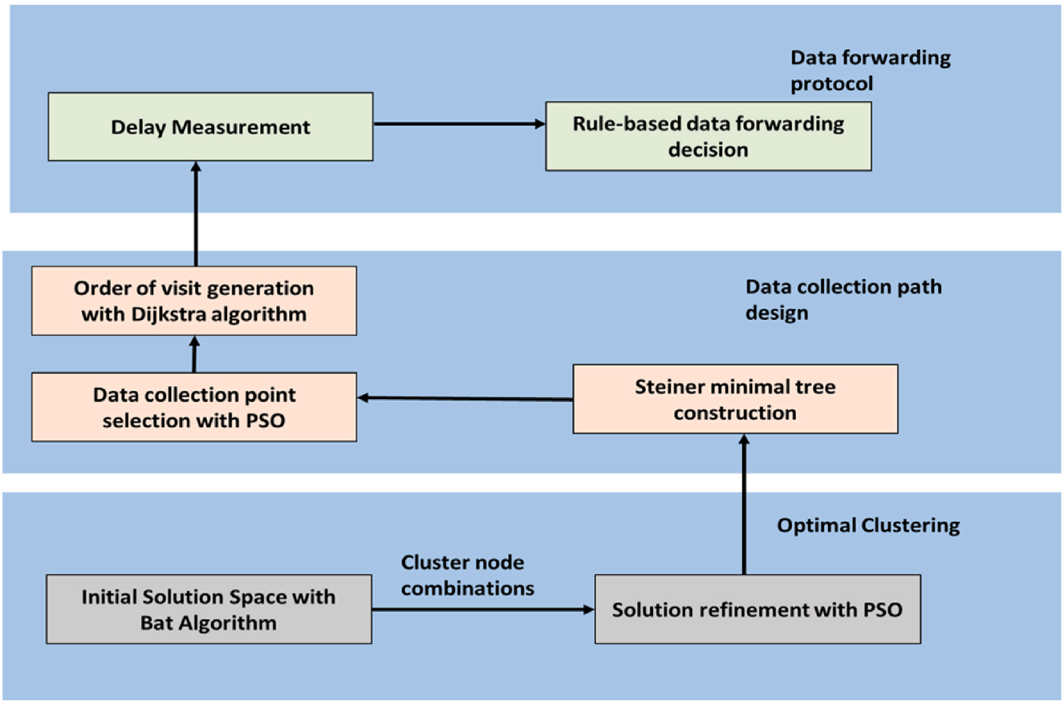

Fig. 1. Multi-factor optimized mobile sink data collection framework.

in mobile sink-based data gathering employing multiple mobile sinks. But the approach could be more scalable for large networks. Raj et al. [23] used game theory and improved ant colony optimization to choose the network's data gathering points. Once the data collection points are found, the path is planned using ant colony optimization. Since the data collection node is fixed, this network finds energy balancing impossible. Chowdary et al. [24] selected cluster heads in the network based on multiple criteria. A mobile sink path connecting the CH is formed without consideration for the optimal order of visits. Kumar et al. [25] selected the optimal data collection points in the network using an ant colony optimization algorithm. The path is chosen frequently to rebalance the energy. The approach considered only energy optimization and did not consider other factors like delay, network overhead, etc. Krishnan et al. [26] developed a mobile sink-based data-gathering path for dynamic clustering structures in sensor networks. The network is dynamically clustered based on energy optimization. A path connecting the cluster head is found using ant colony optimization, which has a higher maintenance overhead. Bouarourou et al. [27] combined clustering and multi-sink data collection to maximize the lifetime of the sensor network. The network is clustered using an improvised ant colony algorithm. Based on the cluster selected, multiple sinks are placed in the network. The path from nodes to sink is found with the objective of minimization of energy consumption by heuristics selection among numerous paths. The cost for data collection using multiple sinks is higher than a single sink, and the approach has a higher sink deployment cost for large networks.

From the survey, most sink-based data collection methods considered only energy optimization in the clustering step and delay in the path planning step. Multiple QoS criteria, like overhead, etc., were not considered. Differential handling of packets was not considered in the data forwarding decisions. In most solutions, data collection paths are not adaptive to network dynamics.

3. Methodology

The overall flow of the proposed solution is given in Fig. 1. The network is clustered, and the higher level of cluster heads, called Steiner point, is found by applying the Steiner minimal tree heuristics algorithm on a graph connecting cluster heads. The second level of grouping of Steiner points is then done to identify the data collection points using PSO. The data collection path among these points is found using the Dijkstra algorithm. Sink travels along these points to collect the data. For emergency messages, the data is passed without storing at clustering by using the LTE interface.

The network model for which the proposed MFO-MS is designed consists of nodes distributed across the network. Some percentage of total nodes have a dual interface of WSN and LTE. Using an LTE interface to send packets to the sink incurs costs and must be minimized as much as possible. A single mobile sink moves around to collect data. Mobile sink travels around the network and stations itself at specific designated locations. The sink advertises a hello broadcast at these locations and collects node data. These designated locations were sink stations to collect data, and the order of visits to these locations must be selected based on the optimization of multi-criteria factors. This work solves the problem of selecting designated locations and order of visits of locations as a two-stage optimization problem.

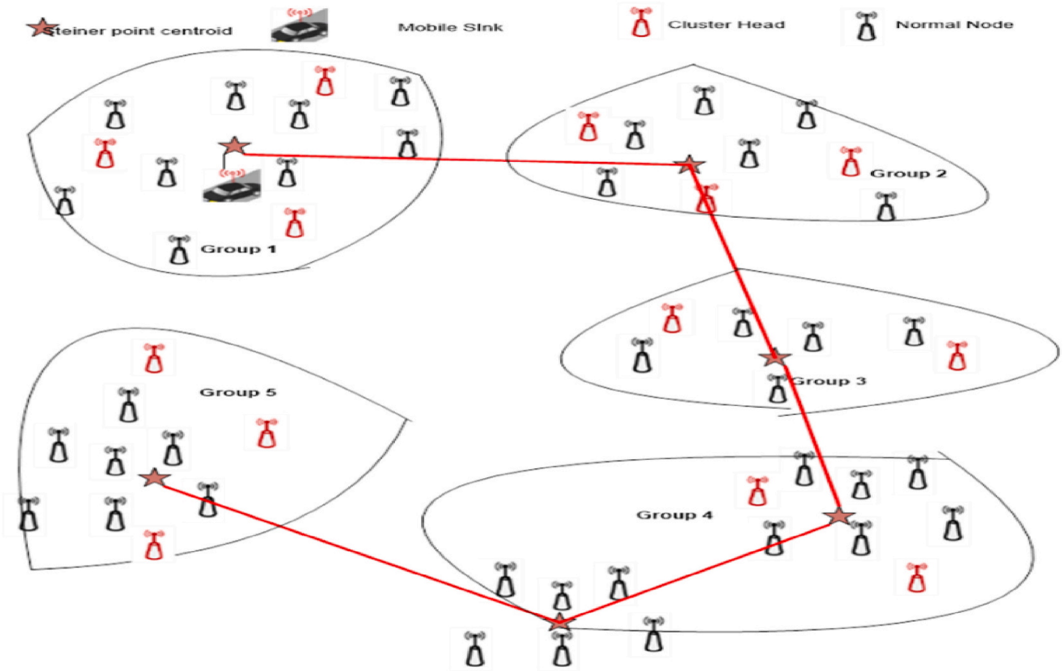

Fig. 2. Network topology of the proposed solution.

The network topology of the proposed solution is shown in Fig. 2. The network is clustered, and the Steiner point is found among the cluster heads. The Steiner points are grouped, and the centroid of it is chosen as data collection points. Sinks travel in the shortest path visiting the data collection points. The data is first forwarded from the node to the cluster head and then to the sink in three modes depending on packet priority and deadline criteria. In the first stage, the network is clustered based on optimizing multi-criteria QoS factors. Hybrid meta-heuristics are applied to optimize the multi-criteria QoS factors. In the second stage, designated locations are found using the Steiner minimal tree and particle swarm optimization concept from the cluster information. From designated locations, the order of visit locations is found. Over the multi-criteria QoS optimized mobile sink data collection framework, a data forwarding technique is proposed using either a WSN or LTE interface for routing a packet to sink based on multiple constraints of delay, reliability, and LTE interface cost. Each of the processing stages of the proposed multi-criteria QoS-optimized mobile sink data collection framework is explained in the subsections below.

3.1. Clustering

The clustering of the network is based on multifactor optimization using a hybrid meta-heuristics algorithm combining particle swarm optimization (PSO) with bat algorithm. The initial solution is found using the bat algorithm, which is further refined using PSO to get the final solution. The final solution is the cluster heads in the network. The use of hybrid meta-heuristics avoids the problem of getting into local minima and getting a near-optimal solution.

The "bat algorithm" is a search optimization technique, based on sonar characteristics of bats. In search of prey, they fly from position x_i with a random velocity v_i, frequency f, and loudness A_0. A bat's location, velocity, and pulse frequency are updated over successive iterations (t) as in eq. 1,eq. (2), and eq. (3).

$$f_i = f_{min} + (f_{max} - f_{min})\beta \tag{1}$$

$$v_i^t = v_i^{t-1} + \left(x_i^{t-1} - x_*\right) f_i \tag{2}$$

$$x_i^t = x_i^{t-1} + v_i^t \tag{3}$$

β is the random variable. Each bat is given a random wavelength between $[f_{min}, f_{max}]$ at the start. The frequency has been fine-tuned to enable diversification and a better study of the prey in the search space. PSO is a soft computing algorithm based on swarm intelligence. Due to multiple reasons like flexibility, simplicity to implement, and versatility, it has become the most popular algorithm for solving optimization problems. Each organism updates their position by moving randomly with different velocities. A particle represents the candidate solution. Every particle modifies itself (changes its velocity) based on its local best (p_{best}) value and global (g_{best}) particle. The particle's next position is updated based on the calculated velocity and distance to the current and global position. Velocity controls the motion of the particle in search space. The current position of the particle (X_i) and the current velocity (V_i) are updated over every time (t) as given in eq. (4) and eq. (5).

$$X_i(t+1) = X_i(t) + V_i(t+1) \tag{4}$$

$$V_i(t+1) = wV_i(t) + c_1 r_1 (p_{best}(t) - X_i(t)) +$$

$$c_2 r_2 (g_{best}(t) - X_i(t)) \tag{5}$$

c_1 and c_2 control the rate of acceleration towards local and global solutions. r_1 and r_2 are random numbers, w is the inertia weight.

The position is updated regularly until the maximum number of iterations has been reached or there hasn't been any movement in the position. The notations used in equations described further in this work are detailed in Table 1. Thus, this work uses a hybridized PSO with a Bat algorithm. The hybridization works as follows – initial solution to the problem is found using the Bat algorithm. This is taken as input to PSO where exploration is done to find the optimal solution.

Table 1
Variables used.

Variable	Detail	Equations used
f_{min}	Minimum frequency of bat	1
f_{max}	Maximum frequency of bat	1
β	Random variable	1
v_i^t	Velocity of i^t bat at time t	2
x_i^t	Position of ith bat at time t	2
x_*	Position of the best bat	2
$X_i(t)$	Position of the i particle at time t	4
$V_i(t)$	Velocity of the i particle at time t	5
$p_{best}(t)$	Local best solution at time t	5
$g_{best}(t)$	Global best solution at time t	5

The clustering is done based on the following factors.

1. Maximize the density (N_d).
2. Minimize the energy consumption (E).
3. Minimal distance to a relay (d_R).
4. Minimize cluster maintenance overhead (O).

The node density (N_d) is calculated as in eq. 6

$$N_d = \frac{1}{N_{num} + 1} \tag{6}$$

Where N_{num} is the number of one-hop neighbors of the node. The energy consumption (E) is calculated as transmission energy necessary for transmitting one packet from neighbor nodes to nodes considered for cluster head. It is calculated as in eq. (7).

$$E = NF * P_T * \varnothing_T \tag{7}$$

Where P_T is the transmitting power, $\varnothing_T$ is the transmission duration, and NF is the average number of frames to be transmitted for delivery of one packet of information. It is calculated as in eq. (8) [28].

$$NF = N_T * \frac{1 - \left(1 - (1-\rho)^2\right)^{R+1}}{1 - \left(1 - (1-\rho)^2\right)} \tag{8}$$

Where N_T is the total frames needed to transmit one packet, ρ is the error probability and R is the number of retransmissions. The number of retransmissions is estimated based on the distance between the node to cluster head node (d) and the density of the cluster head node (N_d) as in eq. (9) [28]

$$R = \frac{1}{\left(\frac{1}{d}\right) + N_d} \tag{9}$$

The distance to relay (d_R) is the distance to the nearest neighbor on a path to sink. It is calculated as minimal of distance of nearest cluster heads to sink node. The cluster maintenance overhead (O) is calculated for the stability of nodes in term of speed and distance change with respect to its neighbor and is given in eq. (10). It is calculated as number of neighbors moving with same speed (N_s) and direction as that of cluster head node out of total neighbor nodes (N_{num})

$$O = \frac{N_s}{N_{num} + 1} \tag{10}$$

The fitness function (F) for optimal clustering is given as in eq (11).

$$F = N_d + \frac{1}{E} + \frac{1}{d_R} + \frac{1}{O} \tag{11}$$

The hybrid meta-heuristics algorithm for clustering starts by finding the initial cluster heads using Bat algorithm. Each bat in bat algroithm or particle in PSO is modelled as bit array of size M (where M is the total number of nodes). In this bit array N (number of clusters) positions are 1 and the rest are 0. Intital value are filled randomly in the array. The bat algorithm creates N random combination of nodes as clusters and finds the best N combinations. Starting with the N best combinations, the PSO algorithm starts the iteration and tries to find the best of N combinations with higher fitness values calculated using equation (11). The sample pseudo-code of the hybrid-meta heuristics algorithm is given below.

Algorithm: Hybrid meta-heuristics clustering

Initialize each particle X_P with N solutions found by Bat algorithm.
While true
For i=1: n particle
Calculate fitness value using Eq.9 for each node in the particle and sum it to get total fitness.
End
Find the local best and global best solution for each particle
update the number of positions in local solution to global solution based on velocity.
update velocity as

$$X_i(t+1) = X_i(t) + V_i(t+1)$$

$$V_i(t+1) = wV_i(t) + c_1 r_1 + c_2 r_2 \left(g_{best}(t) - X_i(t)\right)$$

If no movement in global solution or max iteration reached
break;
End
Return X_P

The clustering process described in this subsection is summarized as a flowchart in Fig. 3.

3.2. Data collection path planning

A graph G is created with the cluster head nodes and one hop neighbors of cluster heads nodes with LTE interface as vertices and the distance between them as weights. Taking the graph as input, Graph iterated 1 Steiner algorithm (GI1S) [17] is applied to get the Steiner tree. The GI1S algorithm produces K Steiner points in the graph [29]. Once the Steiner points are found, they must be grouped based on multi QoS criteria and centroid of the grouped Steiner points are selected as designated locations for sink visit.

In order to create a Steiner minimal tree, S points called Steiner points must be connected in such a way that it incurs the least amount of cost and the Steiner points are close to vertices in its neighborhood. Steiner minimal tree constructed for a graph is given in Fig. 5. The Steiner points are grouped in the second level hierarcy to K groups. The grouping of Steiner points (SP) is done using PSO. Each particle is a binary string of size K (with each position belonging to a Steiner point). M binary strings, each corresponding to a particle are created. When the Steiner point is part of the group, it is encoded as 1, and it is encoded as 0 for the point not part of group. The fitness function (F) is designed with two conditions. (i) Maximizing the coverage i.e. number of SP in group (ii) minimization of delay traversing the SP in group. It is given as eq. 12

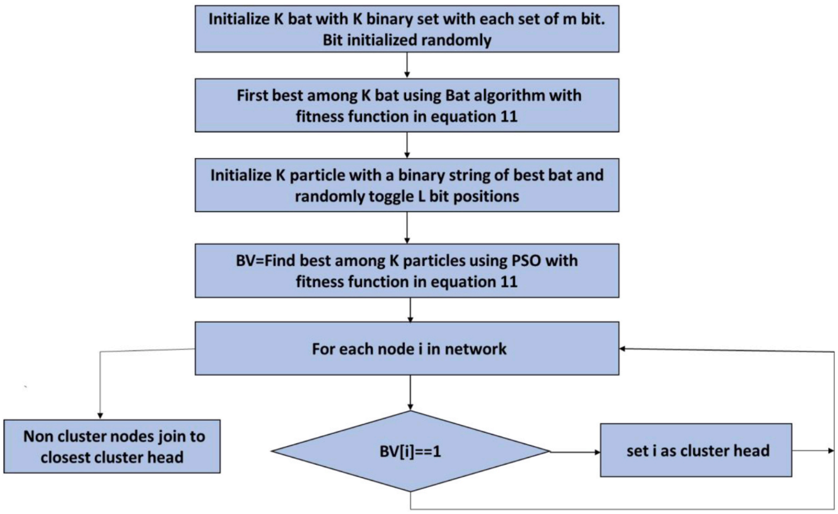

Fig. 3. Clustering flow.

$$F = \frac{\sum_{i=1}^{K} |G_i| + 1/SP(G_i)}{K} \tag{12}$$

In above equation, the number of SP marked 1 in binary string is denoted as $|G_i|$ and the shortest path length connecting the SPs in group is denoted as $SP(G_i)$.

The output of PSO algorithm is M groups of Steiner points. In each group of points, the centroid location is selected as the designated location for sink visit. From the M designated locations, the order of visit of the locations is found by solving the problem as Dijkstra's traveling salesman problem. The flow of data collection path computation process discussed in this section is summarized in Fig. 4.

From the M designated locations, the order of visits of the locations is found by solving the problem as Dijkstra's traveling salesman problem. The flow of the data collection path computation process discussed in this section is summarized in Fig. 5.

3.3. Data forwarding

The designated location's information is broadcast once, and nodes near designated locations store this information. When they move to a new location, they broadcast this information to one-hop neighbor nodes. A cluster head possesses the designated location information and broadcasts a JOIN message to its neighboring nodes. For any node receiving the JOIN message, check the nearest cluster head and join those clusters. Cluster maintenance effort is very minimal in the proposed solution. The node near to designated location information becomes the cluster head. This simple mechanism can effectively handle node failures or a small degree of node movement.

The Routing of packets in the proposed framework is handled using a deadline-guided data forwarding protocol. Each node with packets to send adds deadline time (d_t) and forwards the packets to its cluster head node. Based on the packet deadline, the cluster head makes the forwarding decision. The cluster head makes one of three decisions (i) store and forward, (ii) forward on WSN interface (iii) forward on LTE interface based on the packet deadline. Each cluster head calculates the packet delivery latency to sink (D) as a probability density function of past observations of round-trip time (R) as in eq. (13).

$$D = R/2 \tag{13}$$

$$R = \begin{cases} \sum_{i=0}^{\infty} f_i(a).f_i(b), x = 0 \\ \sum_{i=0}^{\infty} f_i(a).f_{2x+i}(b) + \sum_{i=0}^{\infty} f_i(b).f_{2x+i}(a), x > 0 \end{cases} \tag{14}$$

Where $f_i(a)$ the probability is mass function of delay in the direction of B to A. The rules for decision making at cluster head are given in Table 2.

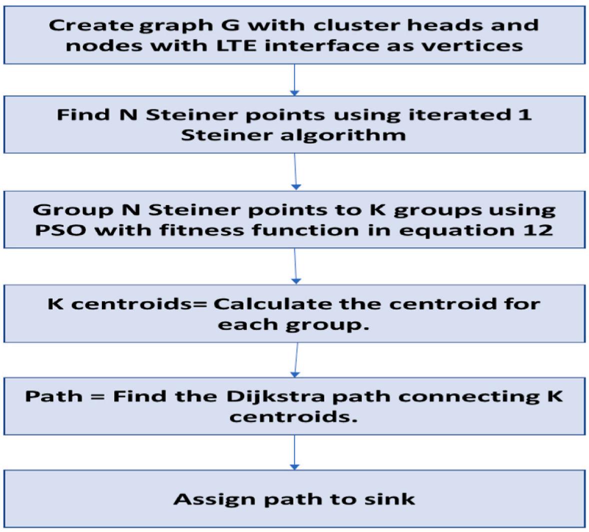

Fig. 4. Data collection path computation process.

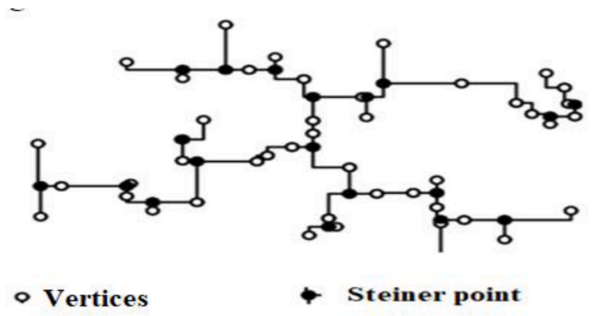

Fig. 5. Steiner path.

Table 2
Data forwarding decision rules.

Criteria	Decision
$d_t > D$ and packet priority is normal	Store and forward
$d_t > D$ and packet priority is not normal	Forward on WSN interface
$d_t < D$	Forward on LTE interface

Only packets whose deadline are very less and cannot be forwarded via WSN interface within deadline are sent on LTE interface. The rest of packets are sent on WSN interface. When the packet priority is normal and there is enough time, the packet is cached in the cluster head node, till the sink arrives at nearby designated location for data collection.

3.4. Network maintenance

The data collection relay node near the data collection path can exhaust the energy over the period and need to be replaced [30]. This process of replacement is initiated by the relay node. Once the relay node's energy falls below a threshold, it broadcasts a message for the next relay selection. Nodes near the data collection path reply with their current residual energy. The current relay node selects the node with highest residual energy and broadcasts the next relay information and relinquish its role as relay. This process reduces the risk to network lifetime and avoids holes in the network.

4. Results

Network simulator (NS2) was used to simulate the proposed solution.The simulation is conducted by varying the number of nodes for different configuration parameters. The trace file results of NS2 are analyzed using AWK scripts. These scripts calculate the various metrics of packet delay, lifetime, network overhead and energy consumption.

The performance is compared in terms of packet delay, energy consumption, life time, overhead. The average time required for a packet to deliver from its source node to the sink node is called as Packet delay. Energy consumption is measured by calculating the sum of energy decrement at each node from start to end of simulation. Lifetime is measured as the time taken by the first node to die. Network overhead is calculated as the number of packets exchanged in network after leaving the data packets. The simulation was run with the following parameters, shown in Table 3.

The performance of the proposed solution is compared to a cluster based mobile sink data collection scheme proposed by Wei et al. [6], energy balanced data collected with constrained path for mobile sink by Fu et al. [10] and multi sink clustering approach proposed

Table 3
Simulation configuration.

Parameters	Values
Number of Nodes	100 to 500
Communication range	100 m
Area of simulation	1000 m*1000 m
Spread of LTE interface nodes	20%
Placement of nodes	Random
Time for simulation	30 min
Length set for interface queue	50
MAC protocol	802.11
Normal to deadline packet ratio	0.1
Default node energy	100 J
LTE interface cost	5 to 1

by Bouarourou et al. [27]. Use of LTE interface as backhaul for emergency packet delivery is a salient feature of the proposed solution. But the compared works don't use LTE interface. For a fair comparison, LTE interface was used only minimal by keeping the normal to dead line packet ratio to 0.1 and keeping the LTE interface cost to maximum value 5. This limits the LTE inteface use in the proposed solution.

The average delay for packet transition from node to sink is measured by varying the number of nodes in the network and the results are given in Table 4.

The average packet delay in the proposed solution is 26% lower compared to Wei et al., 73% lower compared to Fu et al. and 80% lower compared to Bouarourou et al. The delay has been reduced in the proposed solution due to three different ways of data forwarding. Packets with tight deadlines are forwarded via LTE interface. Packets with tolerable delay are routed via multi hop to sink. Only packets of low priority without deadline constraints are considered for store and forward.

The multi hop forwarding is made effective in proposed solution with use of Steiner minimal path for sink traversal. This, along with reduced congestion due to store and forward has reduced the overall delay in proposed solution. Both Wei et al. and Fu et al. clustered the nodes based on energy optimization and did not consider optimal data collection. The delay was higher in Bouarourou et al. due to higher number of hops in routing path compared to other works. The average energy consumed in each node is calculated by changing the number of nodes and the result is given in Table 5.

The average energy consumption of nodes in the proposed solution has reduced by 29.6% compared to Wei et al., 38% compared to Fu et al., 22% compared to Bouarourou et al. The energy consumption has reduced in the proposed solution due to optimal cluster selection and reduction of hops for data transmission. Mobile sink trajectory is optimized on Steiner minimal path. Both Wei et al. and Fu et al. did not consider mobile sink trajectory optimization for energy consumption which increased the hop count for transmission and resulted in energy consumption at nodes. The higher hop count for transmission in Bouarourou et al. increased the energy consumption of all nodes in the path.

Table 8 shows the result of the lifetime measured in terms of first node death by changing the number of nodes.

The lifetime in proposed solution has increased by 13% compared to Wei et al., 11% compared to Fu et al. and 6% compared to Bouarourou et al. The lifetime has increased in proposed solution due to effective energy saving achieved at each node for packet transmissions. Table 6 shows the result of the overhead in clustering maintenance measured by varying the number of nodes.

The cluster maintenance overhead in proposed solution is on average 42% lower compared to Wei et al., 47% lower compared to Fu et al. and 57% lower compared to Bouarourou et al. The cluster maintenance overhead has been reduced in the proposed solution due to less messages needed for selection of new cluster heads compared to existing works. Once designated locations are established and communicated, nodes near the designated locations become cluster heads and messages are needed only for communication of this information to neighboring nodes. But in the existing works, messages are needed for re-electing new cluster heads and communication of new cluster heads. The overhead was higher in Bouarourou et al. due to need for more messages to re-establish the two-level hierarchy in clustering. The nodes were introduced to a small degree of movement (oscillation around it position) and the packet delay is measured fixing the number of nodes as 300. The results are given in Table 7.

The packet delay in proposed solution is 11.3% lower compared to Wei et al., 24% lower compared to Fu et al. and 4% lower compared to Bouarourou et al. Node movement especially cluster head affected the data collection cycle in existing works. But in the proposed solution it was managed effectively due to a shift in cluster head management role to node close to designated locations. The drift between cluster head and mobile sink distance became higher in existing works but it was managed effectively by fixing the designated locations in proposed solution. Due to this packet delay was comparatively lower in the proposed solution.

Lifetime was measured varying the speed for different number of nodes and results are given in Table 9. As in Fig. 6, the lifetime reduces with increase in speed as more messages are needed for cluster maintenance. But to comparatively lower cluster maintenance messages, the energy consumption in the node is lower and this has increased the lifetime in the proposed solution.

The energy histogram is the percentage distribution of nodes in various energy levels at the end of simulation. It is an indicator of how long the network can last. The energy histogram is plotted for various speeds and result is given in Fig. 7a to f. As the speed increases, the percentage of nodes in higher energy levels reduces and percentage of nodes in lower energy levels increases. But the proposed solution has more number of nodes in higher energy levels compared to existing works. This is due effective clustering and reduction of hops for routing in proposed solution.

The average current residual energy of the nodes for 500 nodes over the simulation time is measured and the result is given in Fig. 8.

The average residual energy in the proposed solution is alteast 11% higher at the end of 10 min, 41% higher at the end of 20 min and 54% higher at the end of 30 min compared to existing works. The average residual energy has reduced as the routing hop for the

Table 4
Comparison of packet delay.

Packet delay (msec)				
No of nodes	Proposed	Wei et al.	Fu et al.	Bouarourou et al.
100	18	22	30	31
200	17	21	29	30
300	16	20	27	29
400	15	19	27	28
500	14	19	26	26
Average	**16**	**20.2**	**27.8**	**28.8**

Table 5
Comparison of average energy consumption.

Average energy consumed (J)				
No of nodes	Proposed	Wei et al.	Fu et al.	Bouarourou et al.
100	27	36	37	33
200	26	34	36	32
300	26	33	35	31
400	25	32	35	31
500	24	31	34	30
Average	**25.6**	**33.2**	**35.4**	**31.4**

Table 6
Comparison of overhead.

Overhead (Kbps)				
No of nodes	Proposed	Wei et al.	Fu et al.	Bouarourou et al.
100	19	23	24	26
200	17	22	23	24
300	16	21	22	24
400	13	21	22	23
500	11	21	21	23
Average	**15.2**	**21.6**	**22.4**	**24**

Table 7
Packet delay vs speed.

Packet delay (milli sec) vs speed (m/sec)				
Speed	Proposed	Wei et al.	Fu et al.	Bouarourou et al.
4	20	23	25	22
8	22	25	29	23
12	25	27	31	26
16	27	30	33	28
20	29	32	35	30
Average	**24.6**	**27.2**	**30.6**	**25.8**

Table 8
Comparison of lifetime.

Lifetime (minutes)				
No of nodes	Proposed	Wei et al.	Fu et al.	Bouarourou et al.
100	18	16	17	18
200	19	17	18	19
300	22	19	19	20
400	23	20	20	21
500	25	21	21	22
Average	**21.4**	**18.6**	**19**	**20**

nodes has reduced in the proposed solution using mobile sink based data collection and LTE backhaul based forwarding.

The average number of hops used for routing the packets to sink is measured at end of the simulation time and the results are given in Fig. 9.

The average hop count in proposed solution is alteast 53% lower in the propsoed solution compared to existing works. The use of steiner point along with PSO based grouping has found optimal data collections and placed it uniformly across the network. Mobile sinks collects data from these points. This has reduced the hop count for trasmission.

The LTE interface cost is measured in the proposed solution for varied percentage of deadline constrained packets and the result is given in Fig. 10. The LTE interface cost increases linearly with increase in ratio of deadline constrained packets.

The delay for packet delivery is measured for varied percentage of deadline constrained packets in the proposed solution and the result is given in Fig. 11.

As the deadline ratio constrained ratio is increased, the packet delay increases. But the delay increament is low in case of lower LTE interface cost, as LTE backhaul is used for emergency packet delivery.

The closest architecture to the proposed solution is hybrid WSN with LTE based bachauling using MIMO proposed by Baniata et al.

Table 9
Lifetime vs speed.

Nodes	Proposed	Wei et al.	Fu et al.	Bouarourou et al.
Speed = 4 m/s				
100	17	15	16	17
200	18	16	17	18
300	20	18	18	19
400	21	19	19	20
500	23	20	20	21
Average	**19.8**	**17.6**	**18**	**19**
Speed = 8 m/s				
100	15	13	14	14
200	16	14	15	15
300	17	14.7	15.5	16
400	17.5	15	16	16,2
500	18	16	16.3	16.5
Average	**16.7**	**14.54**	**15.36**	**15.375**
Speed = 12 m/s				
100	14	13	12	13.2
200	15.3	13.4	12.4	13.9
300	16.1	13.9	13.1	14.1
400	16.8	14.1	13.4	14.6
500	17.2	14.3	13.6	15.3
Average	**15.88**	**13.74**	**12.9**	**14.22**
Speed = 16 m/s				
100	12	10	11	11
200	13.1	11.1	11.5	11.4
300	13.6	11.4	12.1	11.6
400	14	11.7	12.3	12.3
500	14.4	12.1	12.6	12.7
Average	**13.42**	**11.26**	**11.9**	**11.8**
Speed = 20 m/s				
100	11.2	10	10	11
200	11.6	10.3	10.4	11.2
300	12.1	10.7	10.9	11.7
400	12.4	11.1	11.2	12.1
500	12.6	11.4	11.6	12.3
Average	**11.98**	**10.7**	**10.82**	**11.66**

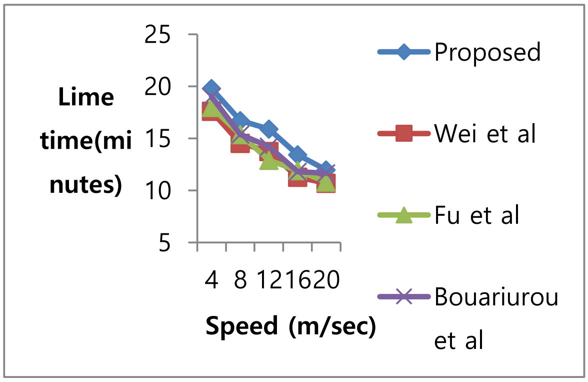

Fig. 6. Lifetime vs speed.

[31]. In this scheme, network is clustered based on residual energy. A probabilistic routing protocol selecting between multi hop routing and LTE interface based forwarding is proposed. The selection between network interface mode is made based on data transfer rate of the node. The performance is compared between the proposed solution for LTE interface cost of 1 and deadline constrained packet ratio of 0.5 and the solution proposed by Baniata et al. [31]. The performance is compared in terms of delay and energy consumption and LTE usage cost.

The delay has reduced in proposed solution on an average of 7% due to selection of effective cluster heads based on multiple criteria and use of minimal hop for routing the packet and the result is given in Fig. 12. The clustering is done based on residual energy without considering hop count and thus delay was higher in Baniata et al. [31].

The energy consumption is measured and the result is given in Fig. 13.

The energy consumption in proposed solution is atleast 16% lower comapred to Baniata et al. The energy consumption has reduced

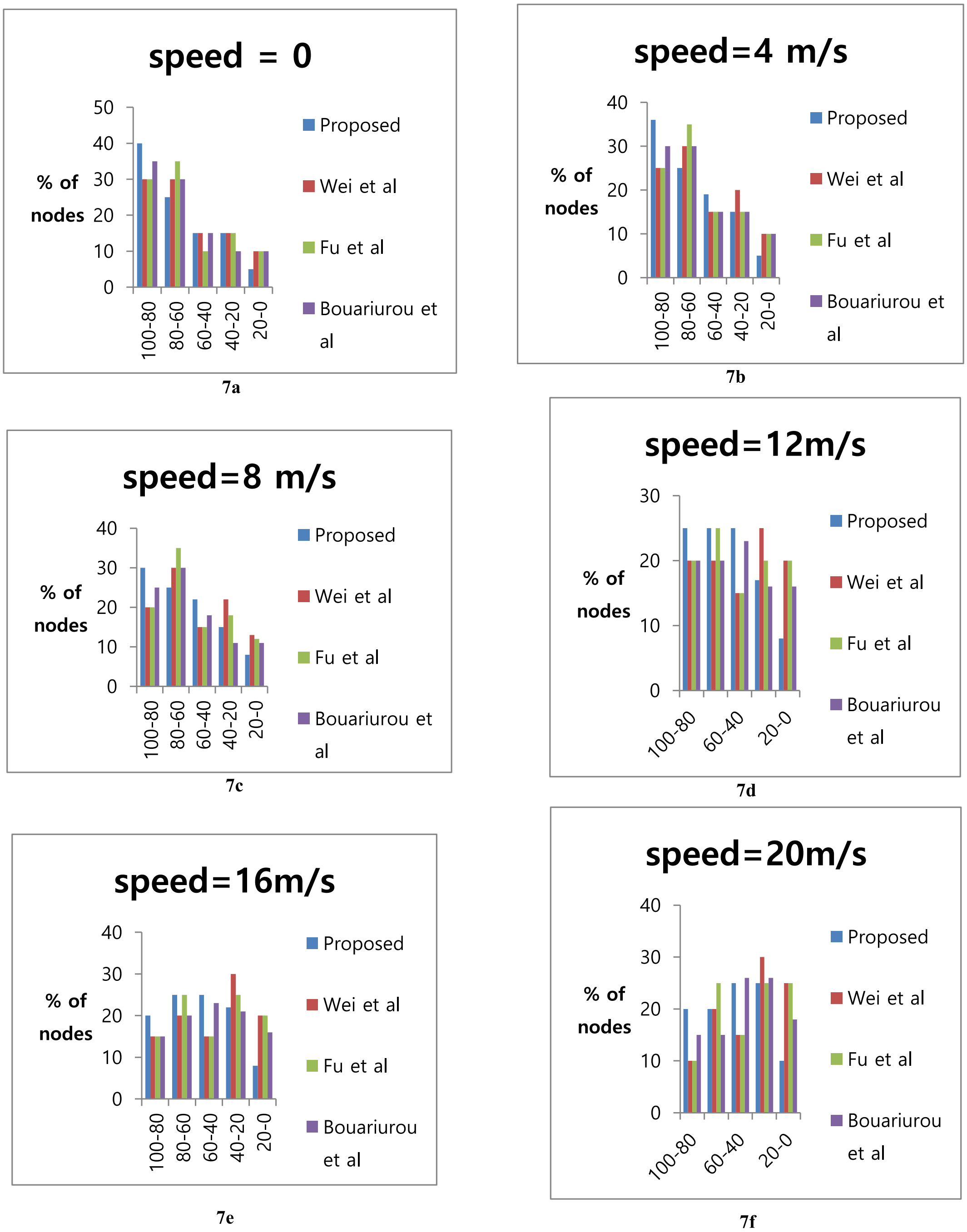

Fig. 7. Energy histogram vs Speed.

in the proposed solution due to minimal number of hop for data transfer and control on LTE interface. There were no such control on LTE interface in Baniata et al. and every node attempted to use LTE probabilistically. This increased the energy consumption in Baniata et al.

The LTE interface usage cost is measured and the result is given in Fig. 14.

The LTE interface cost in proposed solution is atleast 7.3% lower compared to Baniata et al. The LTE interface cost in proposed solution is lower due to effective control on it based on delay projection. But Baniata et al. used the LTE interface based on the current rate without considering the delay/congestion feedback from the network.

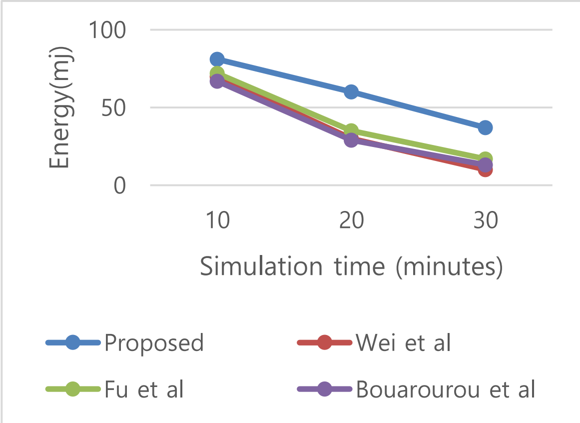

Fig. 8. Average residual energy.

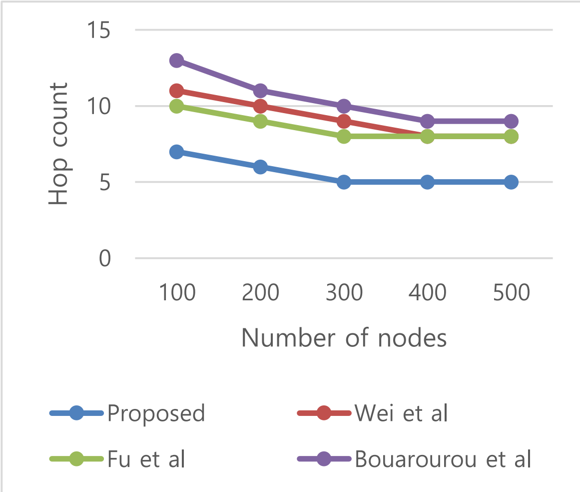

Fig. 9. Average Hop count.

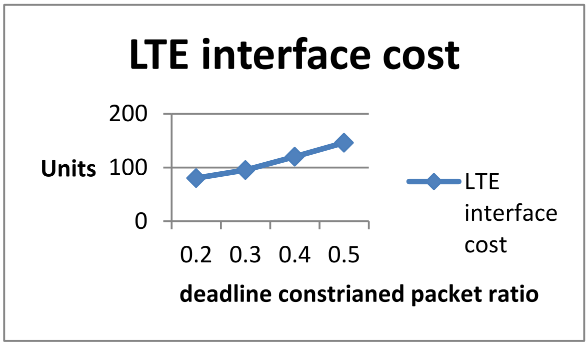

Fig. 10. LTE interface cost.

5. Discussion

One of the important advantage of the proposed solution is its hybridrization of multiple method to maximize life time and improve the QoS. Its hybridized multi level grouping of nodes based on multifactor optimization, path planning across these groups to collect data using mobile sink,adaptive data forwardsing decision based on packet deadline and network dynamics with LTE backhaul. Both

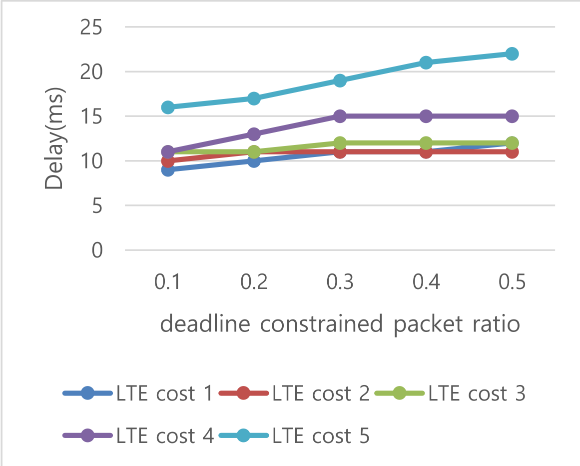

Fig. 11. Delay vs deadline constrained ratio for various LTE interface cost.

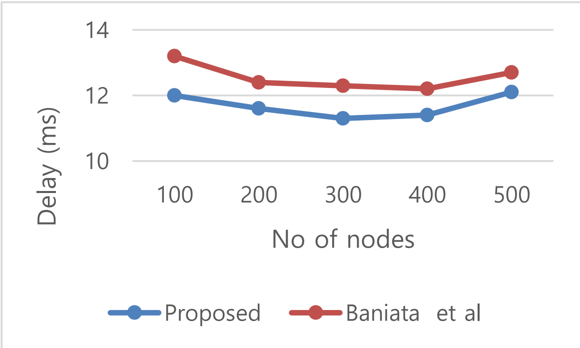

Fig. 12. Comparison of delay.

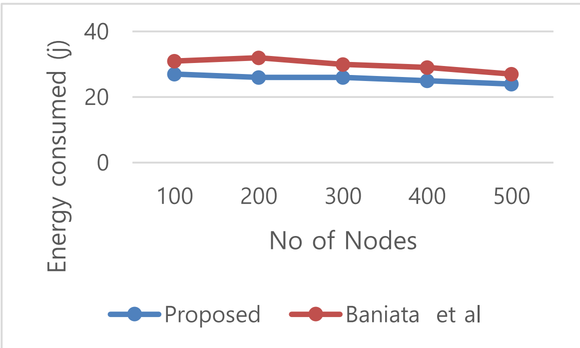

Fig. 13. Comparison of energy consumption.

Wei et al. and Fu et al. clustered the nodes based on energy optimization and did not consider optimal data collection and QoS of packet delivery. Bouarourou et al. though optimzied the energy consumption using multiple sinks, did not consider optimal routing to those sink in presence of congestion. But the proposed solution modelled the congestion using round trip delay and adapted the packet delivery based on congestion. This has improved the QoS in the proposed solution.

6. Conclusion

This work proposed a multifactor optimized mobile sink data collection framework for hybrid WSN-LTE assisted IoT network. The

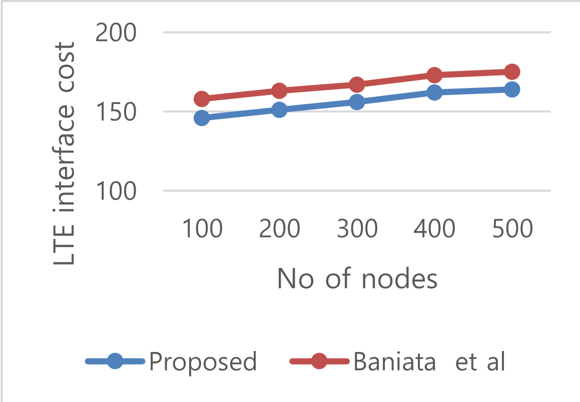

Fig. 14. Comparison of LTE interface cost.

framework solved the problem of optimal placement of data collection points and order of visit of the data collection points as a two-stage optimization problem. The first stage of optimization is solved using hybrid meta heuristics combining PSO with bat algorithm. The second stage is solved combining graph theory-based Steiner minimal tree with PSO. The proposed solution was able to reduce the average delay by 26%, reduce average energy consumption at nodes by 22%, reduce cluster maintenance overhead by 42% and increase lifetime by 6%. The network maintenance phase in the proposed solution is not optimal and improving it is in the scope of the future work.

Funding

This research received no external funding.

Data availability statement

Data sharing is not applicable to this article as no datasets were used during the study.

Additional information

No additional information is available for this paper.

CRediT authorship contribution statement

Saranga Mohan: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Sunita Panda:** Conceptualization, Supervision, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] C.W. Tsai, C.F. Lai, A.V. Vasilakos, Future Internet of Things: open issues and challenges, Wireless Network 20 (8) (2014) 2201–2217.
[2] K. Guleria, A.K. Verma, Comprehensive review for energy efficient hierarchical routing protocols on wireless sensor networks, Wireless Network 25 (3) (2019) 1159–1183.
[3] T.-S. Chen, W.-Q. Du, J.-J. Chen, Energy-efficient data collection by mobile sink in wireless sensor networks, in: 2019 IEEE Wireless Communications and Networking Conference, WCNC), 2019, pp. 1–6.
[4] L. Bagheri, M. Dehghan Takht Fooladi, A rendezvous-based data collection algorithm with mobile sink in wireless sensor networks, in: 2014 4th International Conference on Computer and Knowledge Engineering, ICCKE), 2014, pp. 758–762.
[5] G. Xie, F. Pan, Cluster-based routing for the mobile sink in wireless sensor networks with obstacles, IEEE Access 4 (2016) 2019–2028.
[6] A. Shukla, S. Tripathi, A multi-tier-based clustering framework for scalable and energy-efficient WSN-assisted IoT network, Wireless Network 26 (5) (2020) 3471–3493.
[7] Q. Wei, K. Bai, L. Zhou, Z. Hu, Y. Jin, J. Li, A cluster-based energy optimization algorithm in wireless sensor networks with mobile sink, Sensors 21 (7) (2021) 2523.
[8] A.S. Toor, A.K. Jain, Energy-aware cluster based multi-hop energy-efficient routing protocol using multiple mobile nodes (MEACBM) in wireless sensor networks, AEU-International Journal of Electronics and Communications 102 (2019) 41–53.
[9] Y.C. Wang, K.C. Chen, Efficient path planning for a mobile sink to reliably gather data from sensors with diverse sensing rates and limited buffers, IEEE Trans. Mobile Comput. 18 (7) (2018) 41–53.

[10] J. Wang, Y. Gao, W. Liu, A.K. Sangaiah, H.J. Kim, Energy efficient routing algorithm with mobile sink support for wireless sensor networks, Sensors 19 (7) (2019) 1494.
[11] X. Fu, X. He, Energy-balanced data collection with path-constrained mobile sink in wireless sensor networks, AEU-International Journal of Electronics and Communications 127 (2020) 153504.
[12] R. Yarinezhad, A. Sarabi, Reducing delay and energy consumption in wireless sensor networks by making virtual grid infrastructure and using mobile sink, AEU-International Journal of Electronics and Communications 84 (2018) 144–152.
[13] H. Zhang, Z. Li, Energy-Aware data gathering mechanism for mobile sink in wireless sensor networks using particle swarm optimization, IEEE Access 8 (2020) 177219–177227.
[14] J. Wang, J. Cao, R.S. Sherratt, J.H. Park, An improved ant colony optimization-based approach with mobile sink for wireless sensor networks, J. Supercomput. 74 (12) (2018) 6633–6645.
[15] M. Krishnan, S. Yun, Y.M. Jung, Dynamic clustering approach with ACO-based mobile sink for data collection in WSNs, Wireless Network 25 (8) (2019) 4859–4871.
[16] Bencan, Panpan Gong, Peng Dong, Chen, Dong, Ren," Evolutionary game–based trajectory design algorithm for mobile sink in wireless sensor networks, Int. J. Distributed Sens. Netw. 16 (3) (2020) 1–10.
[17] S. Thomas, T. Mathew, Intelligent path discovery for a mobile sink in wireless sensor network, Procedia Comput. Sci. 143 (2018) 749–756.
[18] J. Wang, Y. Gao, X. Yin, F. Li, H.J. Kim, An enhanced PEGASIS algorithm with mobile sink support for wireless sensor networks, Wireless Commun. Mobile Comput. 2018 (2018) 1–9.
[19] N. Gharaei, K. Abu Bakar, S.Z.M. Hashim, A.H. Pourasl, Inter- and intra-cluster movement of mobile sink algorithms for cluster-based networks to enhance the network lifetime, Ad Hoc Netw. 85 (2019) 60–70.
[20] S. Najjar-Ghabel, L. Farzinvash, S.N. Razavi, Mobile sink-based data gathering in wireless sensor networks with obstacles using artificial intelligence algorithms, Ad Hoc Netw. 106 (2020) 102243.
[21] W. Wen, S. Zhao, C. Shang, C.Y. Chang, EAPC: energy-aware path construction for data collection using mobile sink in wireless sensor networks, IEEE Sensor. J. 18 (2) (2018) 890–901.
[22] C. Konstantopoulos, N. Vathis, G. Pantziou, D. Gavalas, Employing mobile elements for delay-constrained data gathering in WSNs, Comput. Network. 135 (2018) 108–131.
[23] P.V.P. Raj, A.M. Khedr, Z.A. Aghbari, Data gathering via mobile sink in WSNs using game theory and enhanced ant colony optimization, Wireless Network 26 (4) (2020) 2983–2998.
[24] K.M. Chowdary, V. Kuppili, Enhanced clustering and intelligent mobile sink path construction for an efficient data gathering in wireless sensor networks, Arabian J. Sci. Eng. 46 (9) (2021) 8329–8344.
[25] P. Kumar, T. Amgoth, C.S. Annavarapu, ACO-based mobile sink path determination for wireless sensor networks under non-uniform data constraints, Appl. Soft Comput. 69 (2018) 528–540.
[26] M. Krishnan, S. Yun, Y.M. Jung, Enhanced clustering and ACO-based multiple mobile sinks for efficiency improvement of wireless sensor networks, Comput. Network. 160 (2019) 33–40.
[27] S. Bouarourou, A. Zannou, E.H. Nfaoui, A. Boulaalam, An efficient model-based clustering via joint multiple sink placement for WSNs, Future Internet 15 (2) (2023) 75.
[28] Y.M. Raghavendra, U.B. Mahadevaswamy, Energy efficient routing in wireless sensor network based on composite Fuzzy methods, Wireless Pers. Commun. 114 (2020) 2569–2590.
[29] L. Kou, G. Markowsky, L. Berman, A fast algorithm for steiner trees, Acta Inf. 15 (2) (1981) 141–145.
[30] Nitin Kumar, Preeti Rani, Vinod Kumar, Pawan Kumar Verma, Deepika Koundal, TEEECH: three-tier extended energy efficient clustering hierarchy protocol for heterogeneous wireless sensor network, Expert Syst. Appl. 216 (2023).
[31] M. Baniata, H.T. Reda, N. Chilamkurti, A. Abuadbba, Energy-efficient hybrid routing protocol for IoT communication systems in 5G and beyond, Sensors 21 (2) (2021 Jan 13) 537.