# Mobility Assisted Adaptive Clustering Hierarchy for IoT Based Sensor Networks in 5G and Beyond

Biswanath Dey[1,*], Sivaji Bandyopadhyay[1], and Sukumar Nandi[2]

[1] National Institute of Technology Silchar, Silchar 788010, India
[2] Indian Institute of Technology Guwahati, North Guwahati 781039, India
*Correspondence: bdey33@yahoo.com (B.D.)

*Abstract*—**One of the massive machine type communication (mMTC) applications for monitoring and sensing in 5G cellular network is the Internet of Things (IoT) based wireless sensor network (WSN). Non uniform battery usage by the nodes in these networks often results in creating energy holes or voids in the network making the network disconnected. An effective solution is to deploy multiple mobile nodes throughout the network, however finding optimal path for these mobile nodes is reported to be an NP hard problem. This paper proposes MAACH (Mobility Assisted Adaptive Clustering Hierarchy), an efficient mobility assisted clustering and routing framework for IoT based sensor network in 5G and beyond. Also, an elaborate method to calculate the exact path optimally for multiple mobile nodes is presented to alleviate non uniform energy dissipation of the sensing nodes. Simulation results show that our algorithm effectively finds the optimal trajectory of multiple mobile nodes in a distribute manner and also improves network stability period by 60-70% and the network lifetime by 70-90% across multiple network deployments.**

*Keywords*—**IoT based sensor network, mobile nodes, clustering, routing, energy efficiency, network lifetime, 5G and beyond**

## I. Introduction

Technological innovation and socioeconomic change are transforming the 5G cellular network business, which is expected to carry information quickly and support many applications. One use case for the 5G network is the massive Internet of Things (IoT) [1]. Massive Machine Type Communication is included in this (mMTC). One of the mMTC applications is the Wireless Sensor Network (WSN) for monitoring and sensing [2, 3]. In WSN, energy efficiency becomes a major problem. The limited power of each sensor node limits the utilisation of WSN. The research community is becoming increasingly interested in the scalability and load balancing problems for Wireless Sensor Networks (WSN), which serve as the enabling perception layer for Internet of Things (IoT) networks, as a result of the IoT applications' unprecedented growth [4, 5].

As nodes in WSNs operate on finite battery power and have a finite transmission range, energy-efficient methods of relaying data from the network's sensing nodes to the base station is crucial [6]. Uneven battery usage of the sensing node, can result in the channel or a portion of network area being disconnected owing to energy constraints and prevent WSN from successfully transmitting end-to-end data, creating a network hotspot or energy hole within the target area of interest of the IoT deployment [7]. For multi-hop communication, eg., the nodes nearer to sink needs to forward data from the nodes far apart from the sink. As a result, they are the first nodes to deplete their battery power, whereas the other sensor nodes still have plenty of energy. In case of direct transmission or one hop communication from sensing nodes to the sink, nodes far apart from the sink depletes their energy more quickly than the other nodes due to considerable path loss of the transmission signal due to the long distance involved [8]. Such non uniform energy usage creates energy hole after few rounds of network operation that in turn causes network partitioning further restricting the coverage of the whole network for its entire operational lifetime [9].

In IoT based WSN, hierarchical communication paradigm emerged as a practical solution to deal with scalability, the node's limited energy and computational capabilities as hundreds of nodes are distributed over an area of interest to sense the environment and then report on it [10]. Such a hierarchical communication framework for 5G IoT based WSN is illustrated in Fig. 1.

In hierarchical communication framework viz. clustering, one node assumes the role of the coordinator or cluster head (CH) during each round [11-13]. The CH then gathers the data packets from each cluster member (CM) node, assembles them, and transmits them to either a distant base station (BS) or the next hop node on the path to the BS. The role of the CH is frequently shared across several cluster nodes in order to reduce this uneven energy consumption. Although such cluster-based schemes significantly minimize the overall energy usage by the individual nodes, overhead of cluster formation and optimal cluster head selection is a challenge in these

Manuscript received July 13, 2022; revised September 23, 2022; accepted November 23, 2022.

protocol [14]. Moreover, considerable computational requirement and energy wastage occurring due to the overhead involved in cluster formation and maintenance remains a major concern. Also, the problem of uniform energy usage by all nodes in the network still could not solved adequately by the clustering protocols because of the static nature of the cluster head nodes [15].

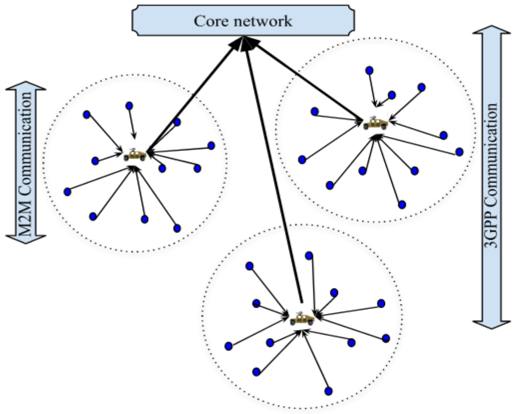


Figure. 1. Hierarchical network of IoT nodes in 5G.

One solution to mitigate the above problem involves each sensor node workout a suitable method to choose its next relay node towards the base station depending upon its remaining energy and distance from the best station or relay node [10]. Another approach is to using one or more mobile sink(s) or data collector(s) to collect sensed data from the sensing nodes and relay the collected data to the distance base station [16, 17]. While the former approach further burdens the limited battery powered energy constrained sensing nodes with huge computational and communication overhead, because of the many constraints must be considered, determining the optimal trajectory of the mobile data collector (s) and successively finding and appropriate routing strategy in itself post a difficult challenge [18-20]. Also, though such MDCs (Mobile Data Collectors) shares some load of the CHs and aids in conservation of the cluster head's energy reserve by relieving CHs from long distance packet transmission to the BS, synchronizing the communication between the CH and the MDCs are often become the bottle neck for such communication framework [21, 22].

With the development of M2M (Machine-to-Machine) communication and 3GPP communication, hierarchical communication framework in IoT based WSN where a set of static sensing nodes senses and a set of mobile nodes collects and sends the data to distance bases station becomes feasible and promising [23]. However, calculating the optimal path for the mobile nodes while ensuring uniform energy usage by all the static sensing nodes has been reported to be a NP-hard problem [24].

Several schemes and interesting works were reported in literature, that enables one or more MDCs to move around the network for data collection and reporting [15, 16, 22, 25-27], though no work shows exactly how the path for such mobile collectors can be calculated optimally in a distributed manner.

To the best of our knowledge, this is the first attempt to use mobile and static sensing nodes for IoT data collecting in 5G without clustering methods. Also, this is the first attempt to establish the exact ideal path of multiple mobile nodes in a distributed manner. Our approach assists efficient data gathering by the mobile nodes while maintaining uniform energy dissipation at the static sensing nodes. In our framework the mobile nodes self-propagate to maximize network coverage and minimize interference based on local information and network state.

Such framework can be effectively be utilized for Internet of Things (IoT)-based precision agriculture, where static sensing nodes could monitor soil parameters like humidity [28] and mobile sensing nodes could be deployed in vehicles like tractors and drones to collect data [29]. A mobile robot might collect data from stationary sensors in an Internet of Things (IoT)-based health monitoring system, such as one in a hospital, and deliver it to a base station in a doctor's office.

The paper is organized as follows. In section II, we discuss the related literature review. In section III we describe our proposed framework and methodology with network model, energy model and description of the optimal path calculation formulation and analysis. Section IV presents the simulations results. Finally in section V we conclude.

## II. Related Work

IoT based wireless sensor network has been the focus of research for recent years due to their common potential in numerous applications. Mobility in wireless environment is now considered with great depth as an advantage rather than being considered as a disturbance [12].

In [8] authors proposed Low energy adaptive clustering hierarchy (LEACH) which is the first distributed clustered-based communication protocol where CH selection is done out using a random probabilistic model and the entire network is divided into many non-overlapping clusters.

Later, a number of deterministic clustering and routing protocols that use different network characteristics for the topology setup have been presented [30]. Compared to the conventional LEACH protocol, the author of [4] significantly reduces network energy usage by combining the concepts of LEACH, mobile sinks, and rendezvous nodes. However, because it uses a single-hop data forwarding strategy, this type of methodology results in increased transmission energy.

In the author of [24], authors suggested a heuristic technique to calculate directions and distances. The proposed method drives the data collectors in an undesired direction—that is, toward the nodes that produce the most data packets.

In the authors of [31], authors proposed PEGASIS (Power Efficient Gathering in Sensor Information

Systems) which improves on the energy required to receive packets from multiple member nodes in the cluster. In PEGASIS forms nodes communicate to their nearest neighbor using greedy search forming a chain among the nodes and one of the nodes serves as the cluster head to send the packet finally to the base station. However, as the chain is formed selecting the neighbour nodes successively algorithm, it will generate long-chain causing considerable delay in forwarding packets for the nodes at the far end of the chain.

In the author of [32] Proposed a location and lifetime biased clustering algorithm for large scale sensor network. They select and rotate the cluster head position based on the distance of the node from the base station and the elapsed operational lifetime of the network with respect to the expected lifetime of the network.

When three-tier architecture was developed in [33], the idea of mobile collectors was first offered. The top tier is made up of WAN-connected devices, the middle tier uses data MULEs i.e. (Mobile Ubiquitous Local area network Extensions), or mobile transport agents, to establish connectivity in sparse WSNs, and the bottom tier is stationary sensor network nodes. MULEs used to move erratically, gather data from stationary sensor nodes, and then dump the information to the base station. Mobility of the collector is now introduced, which lowers traffic due to relaying of the traffic from sensor network, to solve the hotspot problem and balance the energy consumption among the sensor nodes. The key goal in this situation is to work with the original nodes in order to conserve energy and extend their lifespan without causing any network disconnections.

In the authors of [34], the base station is mobile and the cluster head is chosen from additional common nodes rather than being part of a separate collection of bodies. In that paper, the BS (Base station) roams the CHS (Cluster Head Station) in a random fashion, gathering data about the positions, and then delivers the data in a table. Now, the routing of the path is based on the priority mentioned in the table, and the BS channels the communication when it is in range of the CH.

In the authors of [35] proposed a distributed algorithm for mobile sink. Their algorithm is executed locally at each node. As it is a distributed algorithm the computational overhead is low, although their algorithm works for only for single sink and depends of computation of the sink trajectory by the resource constrained sensing nodes and substantial message exchange overhead between the sensing node and mobile sink.

A sink relocation strategy based on the Queen Honey Bee migration process is proposed in the authors of [36]. Their algorithm works for cluster-based network whereby energy consumption balancing is attempted through an active scanning phase during the cluster setup. However, their algorithm considers restriction of too many variables for example confidence factor etc. which are calculated in a random manner.

In the authors of [37] proposed a strategy for gathering data upon determining the energy expense for each sensor node, a genetic algorithm is used to select sink sojourn points. Though their algorithm is simple to implement, suitable for high density WSN, it entirely depends on base station to run GA algorithm for calculating next sink sojourn point and communicate the same to the mobile sinks in every round of network operation.

In the authors of [38], the authors used ant colony optimization algorithm (ACOA) for calculating the path for mobile sinks for collecting data from the cluster based WSN with satisfactory energy consumption. They proposed to use multiple mobile sink one in each cluster for the network which clustered using LEACH algorithm. Then using ACO they find trajectory for these mobile sinks that move to and fro with respect to some starting point to collect data from a designated set of cluster heads. While the clustering the network itself is randomized, using computationally intensive ACO to computer the trajectory for the MS once and for all do not take in to account the dynamic nature of the network.

In the authors of [39] proposed MEEC (Multiple data sink-based Energy Efficient Cluster-based routing protocol) where network is first clustered where the cluster head is selected depending on the node density, distance and remaining energy of the nodes. then multiple mobile sinks are deployed to share the burden of data forwarding by the relaying cluster head nodes.

In the author of [40], authors proposed dynamic relay-assisted clustering (DRAC) where sink deploys a mobile relay node that locates isolated nodes and chooses a new CH among them to preserve communication and coverage when a significant number of CHs in the system die or become unavailable. The system will reconfigure the network once it has attempted count times.

A fuzzy logic-based MDC data collection technique is also used in the authors of [41]. The network is divided into zones in this case, and each zone has an MDC for data collection. To determine the competition radius of an alternative CH, the use fuzzy logic approach with various inputs viz. the closest node, energy, and density. As a result, CHs that are nearer to the MDC trajectory have larger clusters.

However, most of the reported works resorted to restrictive path planning instead of finding a continuous optimal trajectory on the fly based on the dynamic state of the network. Thus, given a static deployment of the sensing network, calculating the optimal trajectory of these MNs that ensures the uniform energy usage across all nodes in the network over the entire operational network lifetime remains an open research problem.

## III. Proposed Methodology

In this section, we describe our proposed framework viz. Mobility Assisted Adaptive Clustering Hierarchy (MAACH) for IoT based Sensor Networks in 5G. Our framework consists of two types of nodes in the network, viz. a large number of static sensing nodes and a smaller number of mobile nodes that were distributed over the entire network to collect and process the data from the static sensing nodes and dispatch the data to base station via multi hop communication among the mobile nodes. The whole network is divided into multiple grids and initially the mobile nodes were transported to the approximate center of the grids before the network starts to operate. We also describe the detailed method for distributed calculation of exact optimal trajectories that is followed by the multiple mobile nodes distributed over the entire network in order to equalize the energy consumed by the static sensing nodes for communicating their data to the mobile nodes over the successive rounds of network operation.

## *A. Network Operational Model*

The network consists of $n$ number of static sensing nodes (SNs) distributed over a region of interest using uniform random distribution and $m$ number of mobile nodes (MNs) ($n > m$). We assume the SNs are limited battery powered and disposable. The MNs do not have any such power restrictions or MNs have higher battery reserve than the static SN. The network is divided into multiple squared grid cells and each grid is assumed to be serviced by one MN. The SNs have a finite fixed constant delectable energy $E^0$. Each SN will have a minimum fixed threshold energy $E^{th}$, below which it will be considered as a dead node. The SNs can tune their radio and invest only the amount of energy that is needed to reach the nearest MN, while transmitting. The base station (BS) is situated somewhere outside the network, and all sensed information by all the SNs are sent to the base station via the MNs in their respective network grid cells.

Before the start of network operation the MNs are dispatched and placed to the approximate center of each grid. The network operation is distributed in rounds, $r = \{1,2,3, \dots, L\}$, where L is the expected operational lifetime of the deployment. On reaching at each sojourn point the MNs broadcast a "ADV(Advertisement)" packet to all static sensing nodes of the grid. The ADV packet contains the ID of the MN, and its updated location of MN. In response to the ADV packet, the SNs reply with a "RESPONSE" packet send to the MN by single hop or multi-hop transmission depending upon its range and remaining energy level. The "RESPONSE" packet contains ID of the SN, its location and current remaining energy level. On receiving the "RESPONSE" packet from the SNs in the grid, the MN broadcast a "Schedule" packet to all SNs.

The SNs sends their data packets to respective MNs as per the "Schedule" received. The MNs after collect the data packets from all SNs in the network grid cell, aggregate the packet and sends to the base station (BS) or to the next hop MN on the path towards the BS.

The MNs normally move to the next sojourn point at every $k/L$ round [42], where $k$ is the number or SN in its grid cell being serviced by the respective MN, and L is the expected operational lifetime of the deployment. In special circumstances if a SNs remaining energy level reaches to a level such that it is unable to send its packet to the MN, then MN also moves to next sojourn point closer to that particular SN.

The path followed by the MNs are calculated and governed by a greedy cost pruning method with the objective to minimize the variance in the average squared distance between all SNs in a network grid cell or cluster from the respective MN over the entire operational lifetime of the network, as well as to maximize the distance of an MN from the other MNs in the network to avoid interference among the MNs.

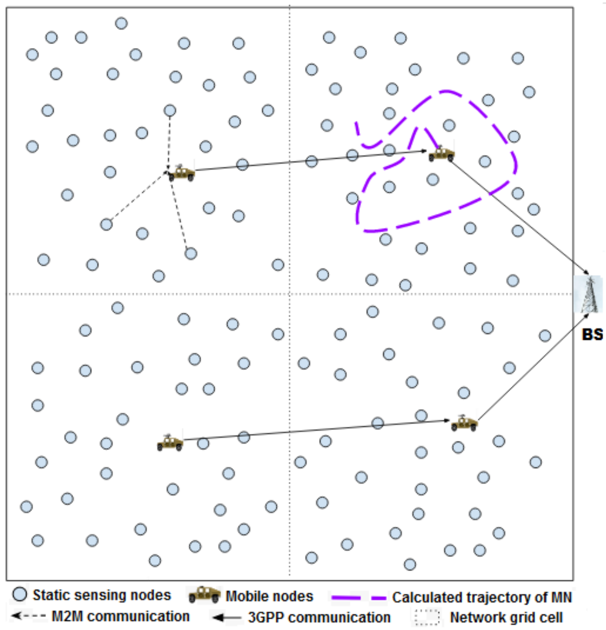


Figure. 2. Network data packet flow diagram

Fig. 2 depicts the node deployment, data flows in the network in our framework along with the mobility patterns of the mobile nodes.

Energy model

Considering first order radio model, the energy expended by a node to transmit a k bit packet over a distance d is given by Eq. (1) as below:

$$E_{tx}(k,d) = \begin{cases} k\left(E_{elec} + \varepsilon_{fs} d^2\right), d \le D_0 \\ k\left(E_{elec} + \varepsilon_{mp} d^4\right), d > D_0 \end{cases} \quad (1)$$

where $\varepsilon_{mp}$ and $\varepsilon_{fs}$ are constant coefficients for multi-path & free-space propagation respectively.

If the distance between the transmitter and receiver, is smaller than the threshold distance, $D_0$ free space propagation model is used, otherwise multi-path propagation model is opted.

The energy expended to collect a *k-bit* packet, is given by Eq. (2) as below:

$$E_{rx}(k) = k * E_{elec} \quad (2)$$

The energy expended to transmit a *k-bit* packet in a hop to hop multi path fashion can be calculated by Eq. (3) as shown belo::

$$E_{tx}(n, n+1) = \sum_{n=1}^{N} 2k * E_{elec} + k * \varepsilon_{mp} d^4 \quad (3)$$

where $n$ is the node that wants to send a k bit packet to the next node $n+1$.

Given d as the distance between two nodes, the optimal hop count for data transmission from SN to MN and MN to the BS can be calculated by Eq. (4) as in the author of [36]:

$$H^{opt} = \sqrt[4]{d \frac{3\varepsilon_{mp}}{2E_{elec}}} \quad (4)$$

*B. Calculation of Optimal Trajectories of the Mobile Nodes*

Trajectory finding involves the plotting of the route between two points [43]. Optimal path finding is not the same as simple path finding. Optimal path finding takes factors other than the shortest path into account [44]. Other factors include presence of obstructions, energy considerations, node fidelity, etc.

We divide the problem of finding the next optimal sojourn point for MN route at $r^{th}$ round of network operation into two sub problems. The first problem is, for any $r = \{1,2,3,\ldots,L\}$, where L is the expected operational lifetime of the deployment, we find the destination sojourn point of the MN after the $r^{th}$ round such that each MNs successively moves closer to the SNs who are at far apart in the previous rounds and move farther from the SNs that are closer to the MNs in the previous rounds, so that the squared distance between the MNs and SNs at each network grid cell can be equalized as far as possible over the entire operational network lifetime. The 2nd problem is to find the optimal path between two given points viz. current coordinate of the MNs and next destination sojourn points of the respective MNs, such that the MN should avoid coming in the wireless range of other MNs [45].

Our problem domain consists of a problem of finding the shortest path between two coordinates in a plane.

Let $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ be two points on the Cartesian plane, then the distance between P and Q is given by the following Eq. (5):

$$PQ = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (5)$$

Energy E is dissipated by SNs according to the following relations as in Eq. (6):

$$E(d[MN, SN]) = k_1 (d[MN, SN])^w + k_2 \quad (6)$$

where $d$ is the squared Euclidean distance between the MN and SN, $w$ is the path-loss exponent, and $k_1$ and $k_2$ are the parameters determined by the characteristic of the transceiver design and the channel [46].

$d$ can be defined as in Eq. (7):

$$d[\mathrm{MN, SN}] = [\,|\,x(t) - a_i\,|^2 + |\,y(t) - b_i\,|^2]^{1/2} \quad (7)$$

where $x(t)$, $y(t)$ are the two dimensional coordinates of the MNs at time instant t and $a_i$ and $b_i$ are the two dimensional coordinates of the static SNs.

Let, in any particular network grid cell, there are m static sensing nodes (SN) and the energy spent by these SNs for the $r^{th}$ round be $E_{1r}$, $E_{2r}$ ,….,$E_{mr}$. Suppose the coordinates of the SNs be $SN_1(x_1, y_1)$, $SN_2(x_2, y_2)$,…, $SN_m(x_m, y_m)$. We define the Center of Energy (CE) within a network grid cell or cluster as the Center of Gravity equivalent of the energy spent by an SN for communication /transmission with the MN at $r^{th}$ round as obtained from the transmission energy equation described in the energy model. Thus the coordinate for Center of Energy (CE) within the network cell at the $r^{th}$ round among all the SNs is represented by the following Eqs. (8)-(9):

$$CE_{rx} = \frac{E_{1(r-1)}x_1 + E_{2(r-1)}x_2 + \cdots + E_{m(r-1)}x_m}{E_{1(r-1)} + E_{2(r-1)} + \cdots + E_{m(r-1)}} \quad (8)$$

$$CE_{ry} = \frac{E_{1(r-1)}y_1 + E_{2(r-1)}y_2 + \cdots + E_{m(r-1)}y_m}{E_{1(r-1)} + E_{2(r-1)} + \cdots + E_{m(r-1)}} \quad (9)$$

Thus the optimal sojourn point coordinate for the mobile node in the grid cell at the $(r+1)^{th}$ round is given by Eq. (10) as:

$$(x_{optimal}, y_{optimal}) = (CE_{rx}, CE_{ry}) \quad (10)$$

That is the next sojourn point coordinate for the mobile node in the grid cell at the $(r)^{th}$ round biased towards the SNs which have spent maximum energy in communicating with the MN in the $(r\text{-}1)^{th}$ round of network operation. This ensures that the SNs which were far from the MN, and hence had to spend more energy while transmission will be closer to the MN in the next round and vice versa. Thus, the average energy spent by the SNs will be uniform, which solves our 1st problem.

The 2nd problem involves finding an optimal path between the source coordinate of the MN found at $(r\text{-}1)^{th}$ round and the destination coordinate of the MN found at $(r)^{th}$ round, such that the distance between the MNs of all the network grid cells is maximized to minimize the interference among the MNs [47]. For solving this, we use a cost based mathematical function and Dijkstra's path finding algorithm.

After we get the source & destination sojourn coordinate points of the MN at a particular round of network operation, we construct a rectangle such that the two points become the diagonal points. After constructing the rectangle matrix of dimension $n{\times}m$, we divide the matrix into many individual cells of dimension $k{\times}l$, such that $k << n$ and $l << m$. Now starting from the source sojourn point coordinate, we follow Dijkstra's algorithm to reach from the source sojourn point coordinate to the destination sojourn point coordinate.

Supposing that there are some m fixed points for the places where other MNs are present, centred at points $E_t = (a_t, b_t)$, $t$ = 1,..., m, on same xy-reference plane. Additionally, supposing that there is MN moving in this plane at a fixed velocity v, tracing a curve $z(t) = [x(t), y(t)]$, which is dependent on time t.

The benefit derived per unit incremental time at time t with respect to existing MN location i, by virtue of this path, is given by some function $f(D([E, z(t)]))$ of the distance $D(\cdot)$ between the current MN and the other MNs, where, as obtained from Eq. (2),

$$D\big(E, z(t)\big) = [(x(t) - a_i)^2 + (y(t) - b_i)^2]^{\frac{1}{2}} \quad (11)$$

Hence, the total cost derived over some time framework $\tau$ is given by

$$\int_{\tau} f(D\big(E,\ z(t)\big)dt \quad (12)$$

In our framework, for the nature of the problem situation and the constraints placed on the MN (that it needs to be as far as possible from the other MNs at all times of network operation), it is more appropriate to consider a normalized objective function, namely, the total cost per unit time as below,

$$\frac{\int_{\tau} f(D(E, z(t))dt}{\int_{\tau} dt} \quad (13)$$

Now, supposing that with respect to same frame of reference the path traced by the MNs can be represented as $Z(x) = [x,\ y(x)]$, and let Z denote the feasible set of functions $Z(\cdot)$ that satisfy any functional form restrictions and boundary conditions imposed on the path.

We assume that $y(x)$ is a twice continuously differentiable function of $x$. Now considering the MN moves with a uniform velocity $v$ we have, $v = ds/dt$, where ds is the incremental distance travelled in the incremental time dt. Thus, we obtain.

$$ds = \sqrt{(dx)^2 + (dy)^2} = \sqrt{1 + \big(y'(x)\big)^2}\ dx \quad (14)$$

where $y'(x) = dy\,/dx$.

Let $X$ denote the permissible variation in $x$, the total cost maximization problem based on (12) may be formulated as follows (where the constant scaling by v has been ignored).

$$Maximize_{z(x)\in z}$$

$$\int_X f(D\big(E,\ z(t)\big)\sqrt{1 + \big(y'(x)\big)^2}\ dx \quad (15)$$

In a similar manner, using (13), we can formulate the problem of maximizing the total cost per unit time (or distance) as follows:

$$Cost\ function\ F =$$

$$Maximize_{z(x)\in z}\ \frac{\int_X f(D\big(E, z(t)\big)\sqrt{1+\big(y'(x)\big)^2}\ dx}{\int_X \sqrt{1+\big(y'(x)\big)^2}\ dx} \quad (16)$$

Now, according to the problem in hand, we consider the cost problem using distance measures, when the path is restricted to begin and end at certain designated end-points. More specifically, we restrict the path function to lie in the following set, where $\theta > 0$, and $h$ are given constants. Thus,

$$Z = \{Z(X)\} = [x, y(x)]: 0 \le x \le \theta, \quad (17)$$

where $y(0) = 0$, $y(\theta) = h\theta$, and $y(\cdot)$ is twice continuously differentiable.

Note that the analysis of this problem includes the situation in which MNs needs to "optimally" traverse from one point location to another, in some designated sequence. We consider the cost problem.

As before, the existing coordinate of MNs are assumed to be located at the coordinates $(a_i,\ b_i)$, $i = 1, \dots, m$. The algorithm we proposed finds an optimal polynomial path for the MNs of degree n. Then we iteratively perturbs the path found via piecewise linear functions, using a progressively finer perturbation discretization. Each perturbed problem is equivalent to a shortest path problem, and is readily solved. From Eqs. (16)-(17), the cost function becomes

$$\frac{\int_0^{\theta} f(D(E, z(t))\sqrt{1+\big(y'(x)\big)^2}\ dx}{\int_0^{\theta} \sqrt{1+\big(y'(x)\big)^2}\ dx} \quad (19)$$

where, $y(0) = 0$, $y(\theta) = h\theta$. The above equation may be represented as

$$\frac{\int_0^{\theta} [(x-a)^2 + [(y(x)-b]^2 + K]\sqrt{1+\big(y'(x)\big)^2}\ dx}{\int_0^{\theta} \sqrt{1+\big(y'(x)\big)^2}\ dx} \quad (20)$$

Consider a discretization of x in the interval $[0,\ \theta]$ given by $x = k\Delta$ for k = 0,1,..., N, N + 1, where $(N + 1)\Delta = \theta$ for some integer $N \ge 1$. Next, select some odd integer $M \ge 3$, and for each k = 1,..., N, choose M function perturbation values $y_{kj}$, $j = 1, \dots, M$, about the value $y(x_k)$, where y(x), $0 \le x \le \theta$.

The solution obtained, $y_{kj} = y(x_k) + \delta[\frac{M+1}{2} - j]$ for $j = 1, .., M$, and where $\delta > 0$ is some perturbation parameter.

For example, with $M = 5$, we have $y_{k1} = y(x_k) + 2\delta$, $y_{k2} = y(x_k) + \delta$, $y_{k3} = y(x_k)$, $y_{k4} = y(x_k) - \delta$, and $y_{k5} = y(x_k) - 2\delta$, for all $k = 1, \dots N$.

The problem posed is to determine the best piece wise linear function $y_\delta(x)$ as a solution of the problem of cost, where $y_\delta(0) = 0, y_\delta(\theta) = h\theta$, and $y_\delta(x_k) = y_{kj}$, for some $j \in [1, \dots, M]$, for each $k = 1, \dots, N$, are the breakpoints of $y_\delta(x)$, and where the integral in equation (20) is evaluated separately for each linear segment.

The above problem is essentially the following shortest path problem.

Consider a graph $G_\delta$ in which the coordinate $(x_0, y_{01}) = (0,0)$, represents the starting sojourn point 0, the coordinates $(x_k,\ y_{kj})$ for $j = 1, \dots, m$ and $k = 1, \dots, n$ represent mn intermediate sojourn points, and the coordinate $(x_{n+1}, y_{(n+1)1}) = (\theta, h\theta)$ represents the terminal sojourn point t, $t \approx (mn+1)$. Then, construct a directed arc from point 0 to each of the intermediate points representing the coordinates $(x_1, y_{1j})$, $j = 1, \dots, m$, a directed arc from each points representing $(x_{k+1}, y_{(k+1)j})$, $j = 1, \dots, m$, for $k = 1, \dots, m-1$, and a directed arc from each sojourn point representing $(x_n, y_{nj})$, $j = 1, \dots m$ to terminal sojourn point $t$.

Because of the structure of the graph, Dijkstra's algorithm, or equivalently, a dynamic programming routine, can be used to solve this problem in polynomial time of complexity in $O(M^2N)$.

Hence, the principal task here is to compute the value of $C_k[y_{kp}, y_{(k+1)q}]$ via Eq. (20).

Let $s_{kpq} \approx [y_{(k+1)q} - y_{kp}]/\Delta$ denote the slope of the corresponding straight line segment of $y\_\delta(\cdot)$, we have from Eq. (20) that

$$C_k[y_{kp}, y_{(k+1)q}] = \int_0^{\Delta} \left[ (k\Delta + x - a)^2 + \left(y_{kp} + s_{kpq}x - b\right)^2 + K\right] \sqrt{1 + s_{kpq}^2}\, dx \tag{21}$$

Thus by integration we get:

$$C_k[y_{kp}, y_{(k+1)q}] = \left(\left\{\frac{[(k+1)\Delta - a]^3}{3} - \frac{[k\Delta - a]^3}{3} + K\Delta + \frac{\Delta}{3}\left[3\left(y_{kp} - b\right)^2 + s_{kpq}^2\Delta^2 + 3s_{kpq}\Delta\left(y_{kp} - b\right)\right]\right\}\sqrt{1 + s_{kpq}^2}\right) \tag{22}$$

We keep calculating the cost of each of the probable next sojourn point paths, choose the one with the maximum cost as with Eq. (22), and select it as the next point to go. We continue doing this until the destination sojourn point is reached.

## IV. Performance Evaluation

To evaluate performance of our algorithm, we simulated along different similar important protocols recently reported in literature. We have used Matlab Simulink [48] for simulation of our algorithm under different deployment scenario for various different network parameters. Table 1 below enlists the network parameters used.

We compared our framework viz. MAACH with various algorithms recently reported in literature having similar notion and network parameters setup, viz. Enhanced clustering and ACO-based multiple mobile sinks algorithm (ECACO) [38], Enhanced LEACH (EnLEACH) algorithm [30], Multiple data sink-based Energy Efficient Cluster-based routing protocol (MEEC) [39] and Dynamic Relay Assisted Clustering (DRAC) [40]. We tested our protocol along the various performance metrics like, Average squared distance of static sensing nodes from the mobile nodes, Average number of packets delivered or Throughput of the algorithms, Operational Network lifetime in terms of First Node Dies (FND), Half Node Dies (HND), and Last Node Dies (LND), and Network stability period. Figs. 3-10 reveals the performance comparison of our protocols with respect to other protocols having similar objectives, for different node distribution patterns and network deployments. While uniform normal distribution of nodes is applied mostly for various normal monitoring applications that requires continuous monitoring data, Gaussian distribution is more suitable for applications that requires special attention for some particular portion in the network for example to monitor a critical unit in a processing plant [19, 29].

Figs. 3-4 shows the Average squared distance of static sensing nodes from the MNs for different protocols for random and Gaussian distribution of sensing nodes respectively. As seen in the figure although during the start of network operation the average squared distance between the sensing nodes and the MN in our framework or CH in various protocols is governed by the network size and deployment patterns of the sensing nodes, as the network starts to operate the average squared distance in case of MAACH is lower than the other protocols in successive rounds of network operation. Also the average distance is more stable than the other protocols as in the case of MAACH, the respective MNs choose their successive sojourn points towards the center of energy of the network grids cluster of sensing nodes, whereas in other protocols data collectors actually tries only to minimize their distance with respect to respective CH and makes no attempt to minimize their average distance with respect to all sensor nodes in the cluster thereby aiding uniform energy dissipation among the static nodes.

TABLE I. SIMULATION PARAMETERS

| Network Parameter | Different Scenario / Values |
|---|---|
| Length x Width of network | 200×200 ($m^2$) |
| No of Sensing Nodes | 100 |
| Number of mobile nodes | 3-6 |
| Sensing range of each node | 20 m |
| Radio range of nodes | 50m |
| Initial energy of sensing nodes | 2 J |
| Node distribution strategy | Random, Gaussian |
| Packet size | 2000 bits |
| Energy overhead of amplifier ($E_{amp}$) | 100 pJ/bit/$m^2$ |
| Path loss exponent | 2-4 |

Figs. 5-6 shows Average number of packets delivered (Throughput) till different rounds of network operation by different algorithms for random and Gaussian distribution of sensing nodes respectively. As evident there is a substantial increase in the throughput in case of our framework protocol viz. MAACH, with respect to the existing protocols. The various other protocols use mostly TDMA schedule for collecting sensing information packets from the member nodes at the cluster head (CH) node which then aggregates and sends the packet to data collectors, while MAACH do not use any cluster head as such, instead the member sensing nodes in that zone sends their sensing information packet directly to their respective MN.

Fig. 7 shows the FND, HND, LND for different protocols for various deployment scenario for random distribution of sensing nodes. Fig. 8 shows the same for Gaussian distribution of sensing nodes. It can be seen that, MAACH shows better performance for each of the metrics viz. FND, HND and LND for both the deployment scenario viz. Random and Gaussian deployment of sensing nodes.

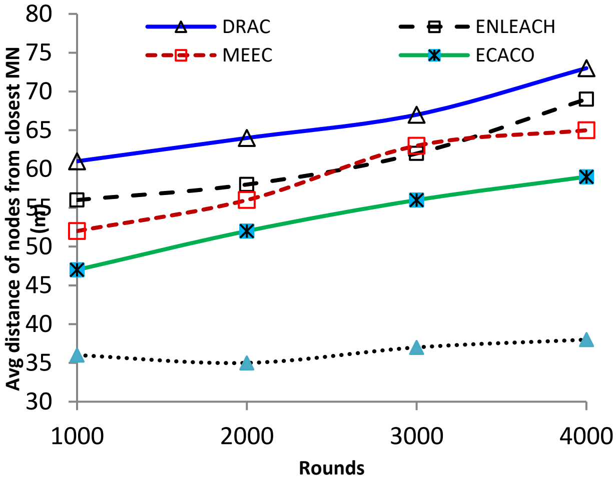


Figure 3. Average squared distance of sensing nodes from the MDCs for different protocols for Random distribution of sensing nodes.

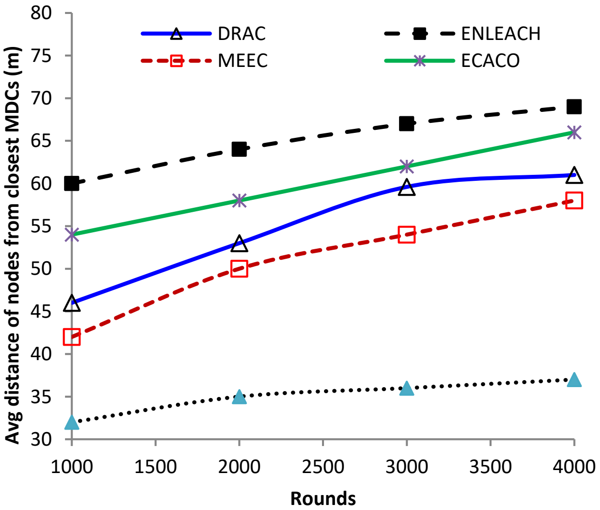


Figure 4. Average squared distance of sensing nodes from the MDCs for different protocols for Gaussian distribution of sensing nodes

For random deployment it is however observed that while MAACH outperforms more than 100% with respect to other protocols as far as HND is concerned, the same is not happened for LND metrics. It can be attributed to the uniform energy expense by the various sensing nodes that enables maximum number of nodes remain alive for higher number of rounds during HND period while most of the nodes dies almost at the same time span during LND period. The situation can be further observed in our next experiment on network stability period analysis.

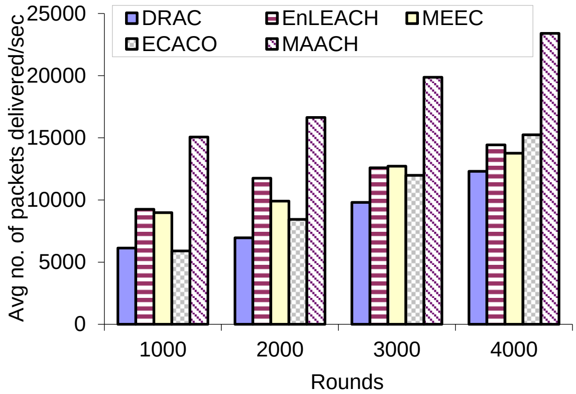


Figure 5. Average number of packets delivered (Throughput) for different protocols for Random distribution of sensing nodes

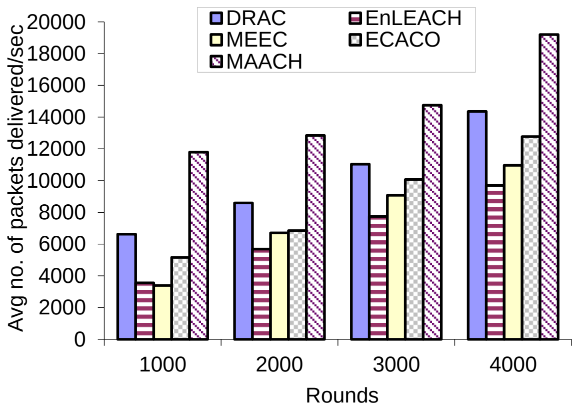


Figure 6. Average number of packets delivered (Throughput) for different protocols for Gaussian distribution of sensing nodes.

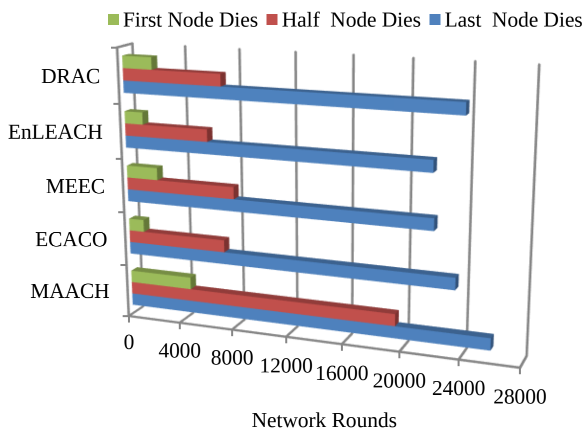


Figure 7. FND, HND, LND for different protocols for Random deployment.

Figs. 9-10 shows the operational network stability period where maximum number of nodes remained alive for minimum rounds of operational network lifetime for different algorithms for random and Gaussian distribution of sensing nodes respectively. It can be seen that, for both the deployment scenario viz. random and Gaussian deployment patterns around 80-90% of the nodes remained alive until the last node dies in case of

MAACH, while for the other protocols although they achieved nearly equivalent network rounds of operation, only around 20-30% of the network rounds have maximal number of active sensing nodes.

This is because in MAACH, the mobile nodes in every step judiciously propagated towards the sensing nodes with lower remaining energy reserve so that every node even with minimum remaining energy level still remains connected to at least one of the mobile nodes to forward its data successfully.

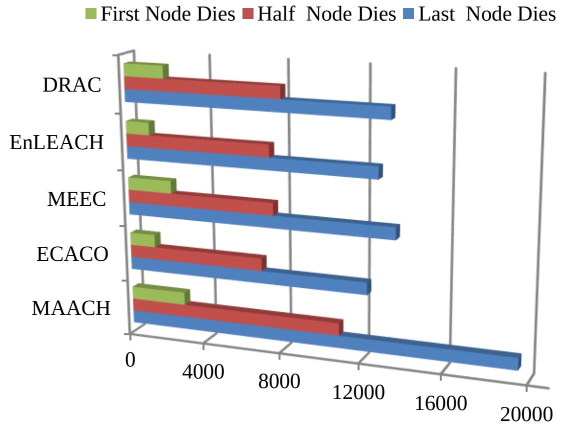


Figure 8. FND, HND, LND for different protocols for Gaussian deployment.

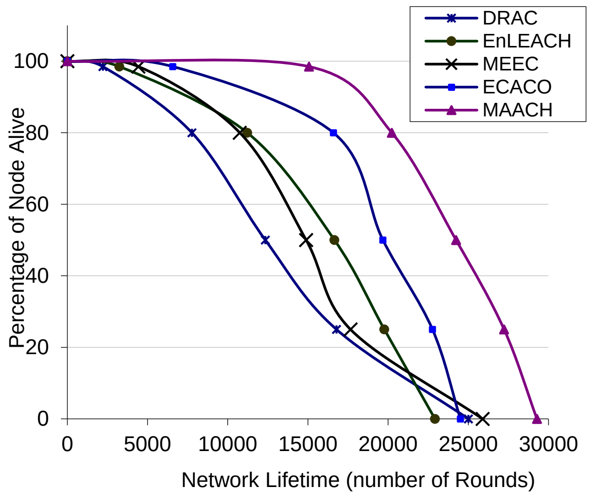


Figure 9. Network stability period for different protocols for Random deployment of sensing nodes

Such strategic movement of the mobile nodes delays the time period until a sensing node become unreachable from any of the mobile nodes and eventually declared dead, resulting in more number of nodes remain alive or connected to maximum portion of the operational network lifetime. Also it can be seen that MAACH increases the actual network lifetime with respect to other protocols. This is because in MAACH there is no clustering over head as such. As the mobile nodes successively drives towards calculated sojourn points such that distance of the sensing nodes from respective mobile nodes remains additively uniform throughout entire lifetime of the network and there is no cluster formation overhead as the sensing node sends their packets to their closest mobile nodes, so energy wastage of cluster head selection, members joining the cluster or cluster setup is significantly minimized which in turn also helps in achieving increase in the overall network lifetime of the deployment.

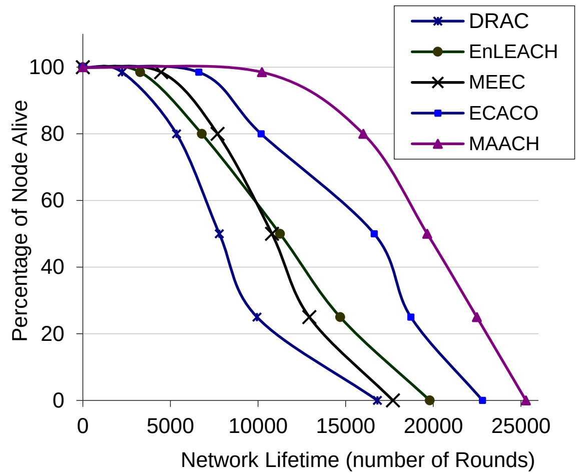


Figure 10. Network stability period for different protocols for Gaussian deployment.

## V. Conclusion and Future Work

In this work we have formulated MAACH, a mobility assisted adaptive clustering hierarchy for IoT based sensor network in 5G and next generation networks as well as a distributed computational method to find online the optimal trajectory of multiple mobile nodes aiding data collection from the static sensing nodes in an IoT based large scale sensor network. Analysis and experimental results show our approach effectively solves the non-uniform energy usage of the static sensor nodes mitigating energy hole and network partitioning problem. The MNs efficiently computes their optimal path through fine-tuned calculation of their next sojourn points in a completely distributed manner minimizing the variance of average squared distance between the static sensing nodes and the mobile nodes as well as maximizing the distance among the mobile nodes thus avoiding interference among the mobile nodes. In future works we plan to augment our algorithm to work on more restrictive predefined path-map to accommodate more real world, smart city applications.

## Conflict of Interest

The authors declare no conflict of interest.

## Author Contributions

Biswanath Dey and Sukumar Nandi conceptualize idea. Biswanath Dey and Sivaji Bandyopadhyay conducted the research. Biswanath Dey, Sukumar Nandi analyzed the data. Biswanath Dey & Sivaji Bandyopadhyay wrote the paper. All authors had approved the final version.

## REFERENCES


[1] L. Atzori, A. Iera, and G. Morabito, “The internet of things: A survey, computer networks,” *Computer Networks*, vol. 54, no. 15, pp. 2787-2805, 2010.

[2] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka *et al*., “Scenarios for 5G mobile and wireless communications: The vision of the METIS project,” *IEEE Communication Magazine*, vol. 52, 2014, pp. 26-35.

[3] M. Agiwal, A. Roy, and N. Saxena, “Next generation 5G wireless-networks: A comprehensive-survey,” *IEEE Communications surveys and Tutorials*, vol. 18, 2016, pp. 1617-1655

[4] S. Mottaghi and M. R. Zahabi, “Optimizing LEACH clustering algorithm with mobile sink and rendezvous nodes, AEU,” *International Journal of Electronics and Communications*, vol. 69, no. 2, 2015.

[5] L. Xu, R. Collier, and G. M. O‘Hare, “A survey of clustering techniques in WSNs and consideration of the challenges of applying such to 5G IoT scenarios,” *IEEE Internet Things J.*, vol. 4, no. 5, 2017, pp. 1229–1249.

[6] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: A survey, computer networks,” *Computer Networks*, vol. 38, no. 4, pp. 393-422, March 2002.

[7] B. Dey, S. Nandi, and G. Das, “Mobility assisted efficient coverage control in cluster-based sensor networks,” in *Proc. 2011 Second International Conference on Emerging Applications of Information Technology (EAIT)*, 2011, pp. 243–246.

[8] W. R. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, “An application-specific protocol architecture for wireless microsensor networks,” *IEEE Transactions on Wireless Communication*, vol. 1, no. 4, 2002, pp. 660-670.

[9] G. Anastasi, M. Conti, M. D. Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey,” *Ad Hoc Networks*, vol. 7, no. 3, May 2009.

[10] S. M. A. Oteafy and H. S. Hassanein, “Resource re-use in wireless sensor networks: Realizing a synergetic internet of things,” *Journal of Communications*, vol. 7, no. 7, pp. 484-493, 2012.

[11] R. M. B. Hani and A. A. Ijjeh, “A survey on LEACH-based energy aware protocols for wireless sensor networks,” *Journal of Communications*, vol. 8, no. 3, pp. 192-206, 2013.

[12] M. Ma, Y. Yang, and M. Zhao, “Tour planning for mobile data-gathering mechanisms in wireless sensor networks,” *IEEE Transactions on Vehicular Technology*, vol 62, no. 4, pp. 1472-1483, 2013.

[13] A. Slimani and M. Redjimi, “Weighted density center clustering protocol for wireless sensor networks,” *Informatica*, vol. 42, 2018, pp. 245–251

[14] P. Rawat and S. Chauhan, “A survey on clustering protocols in wireless sensor network: taxonomy, comparison, and future scope,” *Journal of Ambient Intelligence and Humanized Computing*, 2021.

[15] K. L. M. Ang, J. K. P. Seng, and A. M. Zungeru, “Optimizing energy consumption for big data collection in large-scale wireless sensor networks with mobile collectors,” I*EEE Syst. J*., vol. 12, no. 1, pp. 616–626, Mar.2018.

[16] Y. E. Khamlichi, Y. Mesmoudi, A. Tahiri, and A. Abtoy, “A recovery algorithm to detect and repair coverage holes in wireless sensor network system,” *Journal of Communications*, vol. 13, no. 2, pp. 67-74, 2018.

[17] Y. C. Wang, W. C. Peng, and Y. C. Tseng, “Energy-balanced dispatch of mobile sensors in a hybrid wireless sensor network,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 12, pp. 1836-1850, 2010.

[18] R. Zhang *et al*., “NDCMC: A hybrid data collection approach for large-Scale WSNs using mobile element and hierarchical clustering,” *IEEE Internet Things J*., vol. 3, no. 4, 2016, pp. 533-543.

[19] C. F. Cheng and C. F. Yu, “Data gathering in wireless sensor networks: a combine–TSP–reduce approach,” *IEEE Trans. Veh. Technol*., vol. 65, no. 4, pp. 2309–2324, Apr.2016.

[20] Y. Miao, Z. Sun, N. Wang, Y. Cao, and H. Cruickshank, “Time efficient data collection with mobile sink and vMIMO technique in wireless sensor networks,” *IEEE Syst. J.*, vol. 12, no. 1, pp. 639–647, Mar.2018.

[21] H. Salarian, K. W. Chin, and F. Naghdy, “An energy-efficient mobilesink path selection strategy for wireless sensor networks,” *IEEE Trans. Veh. Technol*., vol. 63, no. 5, pp. 2407–2419, Jun. 2014

[22] S. Yu, B. Zhang, C. Li, and H. T. Mouftah, “Routing protocols for wireless sensor networks with mobile sinks: A survey,” *IEEE Commun. Magazine*, vol. 16, no. 7, pp. 150–157, Jul.2014.

[23] M. R. Palattella, M. Dohler, A. Grieco, G. Rizzo, J. Torsner, T. Engel *et al*., “Internet-of-things in the 5G era: enablers, architecture, and business models,” *IEEE Journal on Selected-Areas in-Communications*, vol. 34, no. 3, 2016, pp. 510-527

[24] K. Akkaya, M. Younis, and M. Bangad, “Sink repositioning for enhanced performance in wireless sensor networks,” *Computer Network*, vol. 49, 2005.

[25] C. Tunca, S. Isik, M. Y. Donmez, and C. Ersoy, “Distributed mobile sink routing for wireless sensor networks: A survey,” *IEEE Commun. Surveys Tutl*., vol. 16, no. 2, pp. 877–897, 2014.

[26] M. D. Francesco, S. K. Das, and G. Anastasi, “Data collection in wireless sensor networks with mobile elements: A survey,” *ACM Trans. Sensor Netw*., vol. 8, no. 1, pp. 1–31, Jan.2011.

[27] M. Zhao, M. Ma, and Y. Yang, “Efficient data gathering with mobile collectors and space-division multiple access technique in wireless sensor networks,” *IEEE Trans. Comput*., vol. 60, no. 3, pp. 400–417, Mar.2011

[28] J. J. Estrada-López, A. A. C. Atoche, J. Vázquez-Castillo, and E. Sánchez-Sinencio, “Smart soil parameters estimation system using an autonomous wireless sensor network with dynamic power management strategy,” *IEEE Sensors J.*, vol. 18, no. 21, pp. 8913–8923, Nov. 2018.

[29] X. Yu *et al*., “A survey on wireless sensor network infrastructure for agriculture computer standards and interfaces,” *Computer Standards and Interfaces*, vol. 35, 2013, pp. 59-64

[30] A. O. A. Salem and N. Shudifat, “Enhanced LEACH protocol for increasing a lifetime of WSNs,” *Personal and Ubiquitous Computing*, vol. 23, no. 5, pp 901–907, 2019.

[31] S. Lindsey, C. Raghavendra, and K. M. Sivalingam, “Data gathering algorithms in sensor networks using energy metrics,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 13, 2002.

[32] B. Dey and S. Nandi, “Distributed location and lifetime biased clustering for large scale wireless sensor network,” in *Proc. Distributed Computing and Networking*, vol 4308, 2016.

[33] R. C. Shah, S. Roy, S. Jain, and W. Brunette, “Data MULEs: Modelling and analysis of a three-tier architecture for sparse sensor networks,” *Ad Hoc Networks*, vol. 1, pp. 215–233, Sep. 2003

[34] Y. H. Wang, C. H. Tsai, and H. J. Mao, “HMRP: Hierarchy-based multipath routing protocol for wireless sensor networks,” *Tamkang Journal of Science and Engineering*, vol. 9, no. 3, pp. 255-264, 2006

[35] Y. Yun, Y. Xia, B. Behdani, and J. C. Smith, “Distributed algorithm for lifetime maximization in a delay-tolerant wireless sensor network with a mobile sink,” in *Proc. IEEE Transactions on Mobile Computing*, vol. 12, pp. 1920–1930, Oct. 2013.

[36] G. J. Jong, A. Hendrick, and G. J. Horng, “A novel queen honey bee migration (QHBM) algorithm for sink repositioning in wireless sensor network,” *Wireless Personal Communications: An International Journal*, vol. 95, no. 3, August 2017, pp 3209–3232.

[37] N. Gharaei, K. A. Bakar, S. Z. M. Hashim, A. H. Pourasl, and S. A. Butt, “Collaborative mobile sink sojourn time optimization scheme for cluster-based wireless sensor networks,” *IEEE Sensors Journal*, vol. 18, no. 16, pp. 6669–6676, Aug. 2018

[38] M. Krishnan, S. Yun, and Y. M. Jung, “Enhanced clustering and ACO-based multiple mobile sinks for efficiency improvement of wireless sensor networks,” *Computer Networks*, vol. 160, no. 4, 2019.

[39] S. Verma, N. Sood, and A. K. Sharma, “A novelistic approach for energy efficient routing using single and multiple data sinks in heterogeneous wireless sensor network,” *Peer-to-Peer Networking and Applications*, vol 12, no. 5, 2019.

[40] N. Gupta, P. M. Pawar, and S. Jain, “Improve performance of wireless sensor network clustering using mobile relay,” *Wireless Personal Communications*, vol. 110, 2020.

[41] K. Koosheshi and S. Ebadi, “Optimization energy consumption with multiple mobile sinks using fuzzy logic in wireless sensor networks,” *Wireless Networks*, vol. 25, 2019.

[42] P. Neamatollahi, M. Naghibzadeh, S. Abrishami, and M.-H. Yaghmaee, “Distributed clustering-task scheduling for wireless

sensor networks using dynamic hyper round policy," *IEEE Trans. Mob. Comput*., vol. 17, no. 2, 2018.
[43] M. Berg, M. Kreveld, M. Overmars, and O. C Schwarzkopf, *Computational Geometry*, Berlin, Germany: Springer, 2000.
[44] S. Gao, H. Zhang, and S. K Das, "Efficient data collection in wireless sensor networks with path-constrained mobile sinks," *IEEE Trans. Mobile Comput*., vol. 10, no. 4, pp. 592–608, Apr.2011.
[45] F. D. Tolba and P. Lorenz, "Multi path routing based energy-efficient for extending mobile wireless sensor networks lifetime," *Journal of Communications*, vol. 14, no. 12, pp. 1224-1228, 2019.
[46] G. Han, X. Yang, L. Liu, M. Guizani, and W. Zhang, "A disaster management-oriented path planning for mobile anchor node-based localization in wireless sensor networks," *IEEE Trans. Emerging Topics Comput*., vol. 8, no. 1, pp. 115–125, Jan.–Mar.2020.
[47] I. Sohn, J. Lee, and S. H. Lee, "Low-energy adaptive clustering hierarchy using affinity propagation for wireless sensor networks," I*EEE Communications Letters*, vol. 20, pp. 558–561, 2016.
[48] MATLAB 2014a, The MathWorks, Natick, 2014