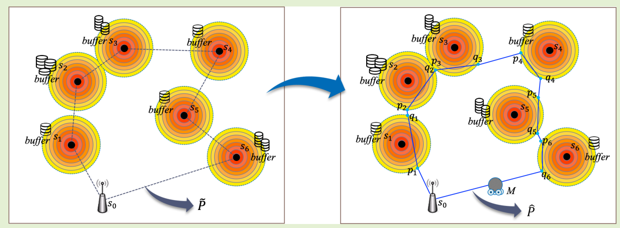
Multirate Data Collection Using Mobile Sink in Wireless Sensor Networks

Chih-Yung Chang, *Member, IEEE*, Shi-Yong Chen, I-Hsiung Chang, Gwo-Jong Yu, *Member, IEEE*, and Diptendu Sinha Roy

***Abstract*—Data collection is one of the most fundamental and important operations in applications of wireless sensor networks (WSNs). Recent research has shown that using mobile sink to collect data brings many advantages to WSNs, such as improving energy efficiency of sensor nodes and prolonging network lifetime. However, the limited speed of mobile sink leads to larger latency, which cause that the data is not fresh or the buffer overflow problem of static sensor nodes. Based on the adaptive rate control, this paper proposes a data collection algorithm which aims to reduce the length of path along which mobile sink collects data from static sensor nodes. The proposed algorithm not only guarantees the completeness of data collection from each static node but also reduces path length by adopting appreciate transmission rate. Through extensive simulations, experimental results show that our approach can notably reduce the path length of mobile sink, while the data of every sensor can be collected completely and efficiently.**



***Index Terms*—Mobile sink, data collection, wireless sensor network, machine type communication, transmission rate.**

I. Introduction

WIRELESS sensor networks (WSNs) have been applied to many applications, such as healthcare, agriculture, environmental monitoring, military and social networks. In a wireless sensor network, nodes are classified into two types, namely sensor node and sink node. A sensor node is used to monitor the environment condition, forward the collected data and relay data to the sink node, while the sink node aims to collect the data from all sensors, plays the role of interface between WSNs and the Internet.

Data collection problem is one of the most important issue in WSNs and has been widely discussed in the past few years. In literature, many studies [1]–[3] have developed data collection approaches, aiming to collect data from static sensors to the sink node such that the lifetime of a given WSN can be prolonged. These studies were mainly classified into two types: static sink and mobile sink.

In the static sink category, a number of studies [4]–[6] developed topologies which determine the route for each sensor to relay its data to the sink node. The most popular topology is the data collection tree rooted by the sink node. Each sensor in tree receives data from its children nodes and relays the data to its parent. However, the nodes closer to the sink have higher forwarding traffics and hence will deplete their energy much faster than others, causing the sensing hole or network disconnection problems.

Some other studies [7], [8] adopted mobile sink instead of static sink. They assume that the sensors are stationary in the network. Mobile sink visits each sensor along the constructed path and each sensor directly transmits data to mobile sink at fixed time. As a result, no sensor needs to relay data and hence the energies of sensors are balanced. These algorithms employ one or several mobile sinks to visit each sensor and collect data from the visited sensor. The method can balance the energy consumption among the static sensors and alleviate the sensing hole problem. However, how to construct the shortest path, while improving the efficiency of data collection still requires further improvement.

When visiting each sensor, the distance between the mobile sink and the visited sensor varies with time. The mobile sink and sensor can adopt different transmission rates to estimate the time required for the data collection and hence apply the multi-rate for reducing the length of data collection path. Table I gives the referenced transmission rates under differ-

Manuscript received March 2, 2020; revised March 14, 2020; accepted March 15, 2020. Date of publication March 18, 2020; date of current version June 18, 2020. This work was supported by the Ministry of Science and Technology, Taiwan, under Grant MOST 106-2221-E-032-015-MY3 and Grant MOST 108-2622-E-032-002-CC3. The associate editor coordinating the review of this article and approving it for publication was Prof. Reza Malekian. *(Corresponding author: Chih-Yung Chang.)*

Chih-Yung Chang, Shi-Yong Chen, and I-Hsiung Chang are with the Department of Computer Science and Information Engineering, Tamkang University, New Taipei City 25137, Taiwan (e-mail: cychang@mail.tku.edu.tw; dora@gms.tku.edu.tw; elite5931.tw@gmail.com).

Gwo-Jong Yu is with the Department of Computer Science and Information Engineering, Aletheia University, New Taipei City 25103, Taiwan (e-mail: yugj@mail.au.edu.tw).

Diptendu Sinha Roy is with the Department of Computer Science and Engineering, National Institute of Technology Meghalaya, Shillong 793003, India (e-mail: diptendu.sr@nitm.ac.in).

Digital Object Identifier 10.1109/JSEN.2020.2981692

TABLE I
MACHINE TYPE COMMUNICATION (MTC)

Rates(kbps)	Range(m)	Modulation
250	10	ASK
100	30	O-QPSK
20	75	BPSK

ent modulations, which refers to the spec. of 802.15.4 [9]. This paper proposes a novel path construction algorithm by considering the machine type communication (MTC).

The MTC shown in TABLE I indicates that the mobile sink adopts the lower level of modulation can have a larger distance to the sensor node but its data receiving rate will be reduced. Consider the case that one sensor has few data to be transmitted to the mobile sink. The mobile sink can move along the path which might be far away the visited sensor. This can also guarantee that the mobile sink can receive all data from that sensor since there are few data buffered in that sensor. On the contrary, consider that there is a large amount of data in one sensor. To guarantee that the mobile sink can receive all data from that sensor, the mobile sink should move along the path which is closed to the sensor location. Since the distance between mobile sink and sensor is short, based on the MTC table, the data receiving rate of mobile sink can be increased, which guarantees that mobile sink can receive all data from that sensor. Based on the MTC table, the mobile sink can calculate the lowest modulation that guarantees to receive all data from each sensor. This can get benefit that mobile sink can receive the data from sensor along a better path which allows the mobile sink to have a larger distance from the sensor. The rate adaption applied in the proposed path construction algorithm can further reduce the time period for maintaining the communication link and hence can further reduce the length of segment that passes through the communication range of the visited sensor.

This paper presents an algorithm to construct a path which passes through each sensor for mobile sink to collect data from each static sensor with a guarantee that all data of each sensor can be fully collected. To reduce the path length, the proposed algorithm then utilizes the triangle theorem, aiming at minimizing the data collection path length. Compared with the existing works [14]–[17], the contributions of this paper are itemized as follows:

A. Collecting Data While Moving

In [10] and [11], mobile sink stops at the sensor location when collecting data from that sensor. This paper implements the concept of "data collecting while moving" which can efficiently move along the path.

B. Without Passing Through the Center of Sensor

The mobile sink considers data volume and communication rate to construct the shortest segment within the communication range of the sensor, which reduces the total path length, as compared with the existing studies [12], [13].

C. Taking Into Account the Transmission Rate

The adoption of multi-rate can accurately construct the shortest chord which guarantees to receive all data from the visited sensor while reducing the path length.

D. The Guarantee of Complete Data Collection

A collection segment (the chord) is analyzed and constructed such that the mobile sink moving along the chord of each sensor can guarantee that all data of the visited sensor is completely collected.

E. Saving Energy and Reducing the Fresh Cycle of Data Collection

The path constructed by connecting the chords can reduce the path length for data collection and thus reduce the length of fresh cycle of data collection.

The rest of this paper is organized as follow: Section II presents the previous work related to data collection in WSNs. Section III presents the network environment and problem formulation. Section IV proposes the algorithms while Section V depicts the simulation results. Finally, the conclusion is drawn in Section VI.

II. RELATED WORK

In literature, a large number of approaches that adopted mobile sink to collect data have been proposed in the past few years. These studies can be mainly classified into two categories, depending on whether or not the mobile sink passes through the centers of visited sensor nodes. In the category that the mobile sink passes through the centers of visited sensor nodes, those approaches can be further divided into partial-data-forwarding [1]–[3] and no-data-forwarding [6]–[8]. The following reviews these studies and compares them with the proposed algorithm.

A. Stop at the Center of the Sensor Node

The approaches fall in this category can be further divided into two sets, including the no-data-forwarding set and partial-data-forwarding set. In the studies of no-data-forwarding set, the mobile sink visits and collects data from each sensor. On the other hand, in the studies of partial-data-forwarding set, mobile sink selects some sensors, called data collecting point, or CP, to be visited. All the other sensors that are not visited by mobile sink should forward their data to the closest CP. The following reviews these studies.

1) Partial-Data-Forwarding: In [1]–[3], sensor nodes were partitioned into a number of groups and sensors in each group were organized as a tree topology rooted by a collection point. In each tree, all static sensors send their data along the tree topology and the collection point stores the data until the mobile sink visits them. The mobile sink moves along the established path which passes through all the collection points aiming to collect data in a way that the network lifetime of the WSN can be prolonged as much as possible. However, the length of constructed path can be further reduced.

In [4], a weighted rendezvous planning (WRP) is proposed. Each sensor is assigned with a weight and the sensor nodes with the highest weight are considered as CPs. Then a traveling tour that passes all CPs is constructed for the mobile sink. In path construction, the WRP treats the sink as first CP. Then the weight of each sensor node is calculated by multiplying the number of packets that it forwards by its hop distance to the closest CP on the tour. Treating the highest weighted sensor node as CP, the mobile sink builds a path that passes through all CPs. Since the tour length is limited, the visited path length has to equal to or less than the maximum tour length. However, the path length from the current location to the next CP is not taken into consideration. This might raise the problems of fewer collected data or shorter network lifetime.

In [5], a novel dynamic clustering based mobile-to-cluster (M2C) scheme was proposed to optimize the service process for both sensor nodes and the mobile sink. By estimating the energy consumption of being a CP, the sensor nodes with residual energy close to the estimation are actively elected as CPs. However, the amount of data stored in each sensor is not considered. Furthermore, the length of path constructed for mobile sink can be further reduced.

2) No-Data-Forwarding: To reduce the energy consumption of the mobile sink, study [6], [7] proposed travel planning algorithm for data collection. The algorithm aims to shorten the path but can still visit all sensor nodes and balance energy consumptions of sensors. However, with an increasing number of sensor nodes, it will still become impractical due to the high computational complexity.

Study [8] presented two schemes for mobile sink to collect data, called Min Mobile Growth (MMG) and Mobile Node Movement Scheduling algorithm (MNMS). The centralized algorithm MMG reduces the number of required mobile nodes by global t-scan and optimizes their travel paths. The MNMS algorithm generates the paths, carefully schedules the movement of each mobile sink and makes sure that all data collected by mobile nodes can be transferred to the base station. Global t-scan is achieved by utilizing Min Mobile Growth and MNMS. However, the efficiency of the constructed path still can be improved.

Study [10] presented a variable length genetic algorithm to optimize data collection path of mobile sink. The mobile sink has to go through the whole network and collect data from each sensor by stopping at some predefined locations from where sensors are in its vicinity. However, the path construction is time consuming since the generic algorithm is applied. Furthermore, the length of path constructed for mobile sink can be further reduced.

B. Visiting Sensor Node With a Constant Speed

The basic concept of this category is that the mobile sink constructs the data collection path which passes the communication range of each sensor node. Compared with those studies fall in the first category, studies fall in the second category have the major advantage that the length of the constructed path can be reduced since the path needs not to pass the center of each sensor. Study [11] shown the mobility-assisted data collection problem with a progressive optimization approach, which considered the data-rate constraints of the communication between the mobile sink and the sensor nodes. It determined the proper locations around the collection site to conduct the data collection.

Study [12] presented a shorten Traveling Salesman Problem(TSP) route for mobile sink to collect data. Along the resultant route, energy efficiency was achieved for the mobile sink to collect data in a round. It proposed an approach which utilized a mobile sink to collect data while moving along a designated TSP route in wireless sensor environments. However, it did not consider the amount of data buffered in each sensor. Therefore, it did not provide a mechanism to guarantee that all data can be received from each sensor to the mobile sink. Furthermore, the length of the path can be further reduced.

TABLE II
SIMULATION SETTING

Parameter	Value
Simulator	MATLAB
Monitor region	1000*1000 unit^2
The number of sensor nodes	50-200
Mobile sink speed	0.5m/s
Transmission range of each CP	10m-75m
Consumed energy in transmitter circuit	0.18J/packet
Consumed energy at the receiver circuit	0.1J/packet
Consumed energy of sink movement	1J/unit

Study [13] proposed two versatile problems, the *k-traveling salesperson problem with neighborhood* (k-TSPN) and the *k-rooted path cover problem with neighborhood* (k-PCPN). The common goal of both problems is to find the trajectories of the k mobile sinks to collect data from n wireless sensor nodes on 2-D Euclidean space such that the data collection latency (the length of the longest trajectory among k trajectories) is minimized. However, the efficiency of mobile sink can be further improved.

The MSFLA [18] used a single mobile sink to collect data from all cluster headers. The whole region was partitioned into several sub-regions. Then the MSFLA selected one CP for each sub-region and allocated a mobile sink to collect data from the selected CPs. Similar to MSFLA, the existing CDCA [19] partitioned the sensor network into several areas and selected one CP from each area. Then the CDCA constructed a Hamiltonian route passing through the selected CPs.

Although the studies in the second category constructed shorter path, none of them considered the data amount stored in the sensor node. Hence, they did not guarantee that all data can be collected by the mobile sink. Moreover, to our knowledge, it is the first work which adopts different communication rates based on the distance between the mobile sink and the visited sensor. The adoption of multi-rate can accurately construct the shortest chord which guarantees to receive all data from the visited sensor and reduces the length of data collection path in advance. The following table compares the proposed algorithm and the related studies. As shown in Table II, the proposed algorithm has good features in terms of data collecting while moving, without stopping center of sensor and multiple transmission rates.

III. NETWORK ENVIRONMENT AND PROBLEM FORMULATION

This section initially introduces the network environment and assumptions of this work. Then, the problem formulation is proposed.

A. Network Environment

Assume that the monitoring region is O. Consider a base station s_1 and a set of n-1 static sensors s_i, $2 \leq i \leq n$. Let $S = \{s_1, s_2, \ldots, s_n\}$ denote the set of base station and n-1 sensors, which are randomly deployed in region O. Each sensor has limited buffer capacity and it stores new sensing data in the buffer with a constant rate. A mobile sink, denoted by M, moves with a constant velocity for periodically visiting each static sensor, aiming at collecting data from each sensor during a predefined period T.

Energy consumption is the paramount factor for determining the lifetime of a WSN. The lifetime of a WSN is measured by the time period starting from the time point when the network begins operation to the time point when the first sensor node runs out of energy, because a dead sensor is expected to be unable to relay or send data to the sink node. This paper presents a method to plan a tour for collecting data from each sensor. This helps each sensor minimize energy expenditure and hence maximize the network lifetime. The proposed strategy plans the path for data collection to minimize the power consumption of the static sensor nodes.

The following presents the energy consumption model of this paper. The sensor nodes mainly consume energy by operations of data sending and receiving. Assume that the size of each packet is κ bits. Let sender s_i and receiver s_j be a communication pair and consider that s_i sends κ bits to s_j. Let d_{ij} be the distance between s_i and s_j. The energy consumption of each receiver node for receiving one packet (κ bits) can be measured by (1):

$$E_R = \kappa\beta \tag{1}$$

where β is the energy consumption per bit incurred by the receiving circuit. The energy consumption for sender s_i to transmit one packet (κ bits) to its parent s_j is expressed as:

$$E_T = \kappa\varepsilon_1 + \kappa d_{i,j}^{\gamma}\varepsilon_2 \tag{2}$$

where ε_1 is a factor that represents the energy consumption per bit of the transmitting circuit, ε_2 is the energy consumption per bit of the amplifier circuit, and γ is the path-loss exponent (which typically ranges between 2 and 4, depending on the environment).

B. Problem Formulation

This paper aims to construct the shortest data collecting path for mobile sink M such that the buffer data of each static sensor can be fully collected. Let $\hat{P} = (\hat{s}_1, \ldots \hat{s}_i, \ldots, \hat{s}_n, \hat{s}_1)$ denote the data collection path which passes through the communication ranges of all sensors in set S, where s_0 denote the base station. Let location of $\hat{s}_i$ be presented by $(\hat{x}_i, \hat{y}_i)$. Let $d(\hat{s}_i, \hat{s}_{i+1})$ denote the distance of two adjacent sensors $\hat{s}_i$ and $\hat{s}_{i+1}$. Let $L_{\hat{P}}$ denote the length of path $\hat{P}$. The goal of this paper is to construct the path with minimal length for mobile sink M to collect data from each sensor s$_i$ in S. Equ. (3) reflects the goal of this paper:

$$min(L_{\hat{P}} = \sum_{i=1}^{i=n} d(\hat{s}_i, \hat{s}_{i+1})) \tag{3}$$

Some constraints which should be satisfied are given below when developing an algorithm for achieving the goal depicted in Equ. (3). The first one is the communication constraint which restricts mobile sink M to pass through the communication range r_c of each sensor in S.

1) Communication Constraint:

$$d(M, s_i) \le r_c \text{ for any } s_i \epsilon S \tag{4}$$

In fact, the path of M needs to intersect with the communication range of each sensor. This can guarantee that M can receive data from each sensor. Herein, we emphasize that M is not necessary to pass the location of each sensor. This might help reduce the length of the path $\hat{P}$.

The second constraint is the ***Buffer Constraint***, which restricts mobile sink to guarantee that it can receive all data stored in the buffer of each sensor. To achieve this, the mobile sink M should fall in the communication range of each sensor for a long enough time period such that it can receive all data stored in buffer of that sensor.

Let t_i^{start} and t_i^{end} denote the times that M enters into and exits the communication range of the sensor s_i, respectively. Let r_t denote the receiving rate of data sent by sensor s_i at time t, where $t_i^{start} \le t \le t_i^{end}$. Let D^{buffer} denote the maximal capacity of each sensor's buffer. The following *Complete Data Collection Constraint* guarantees that M can collect all data of each sensor s_i:

2) Complete Data Collection Constraint:

$$\int_{t_i{}^{start}}^{t_i{}^{end}} r_t dt \ge D^{buffer} \tag{5}$$

In addition to the *Complete Data Collection Constraint*, another constraint is the limited path length. The total time for mobile sink M visiting all sensors along the constructed path should be constrained. There are two main reasons to constrain the total visiting time. The first one is that the energy of the mobile sink M is limited. The second reason is that the time period between two visits to the same sensor should be constrained. A long period not to visit sensor s_i might lead to the situation that the buffer of sensor s_i is overflowed. Let T denote the total time required for travelling along path $\hat{P}$. Let T_{move} denote the time period that M moving without receiving data. Let $T_{collection}$ denote the time period that M is moving and collecting data. It is obvious that

$$T = T_{move} + T_{collection} \tag{6}$$

Assuming that the data generation rate of each sensor s_i is $\rho_i bit/s$. The following ***Path-Length Constraint*** prevents buffer of each sensor s_i from overflow.

3) Path-Length Constraint:

$$T^* \rho_i \le D^{buffer} \tag{7}$$

IV. Algorithms

The section presents the proposed data collection algorithm, which aims to reduce the length of data collection path while satisfying the demand for visiting all sensors. The algorithm mainly consists of two phases. The first phase aims to construct an initial path that passes thought all sensors. Based on the constructed path, all sensors have been ordered in a certain visited sequence. The second phase aims to further reduce the path length. The following presents the proposed algorithm.

A. Path Construction Phase

Let $S = \{s_1, s_2, \ldots, s_n\}$ denote the set of n sensors. Let (x_i, y_i) denote the location of sensor s_i. Let $S_{southeast}$ be the southeast point in S. That is:

$$S_{southeast} = \arg\min_{s_i \in S} (y_i) \tag{8}$$

Let $\hat{s}_1 = S_{southeast}$ will be the first visited point in the path. Let π be a horizontal line passing through $S_{southeast}$ and π has an infinite length. Turn π in a counterclockwise direction until it touches any point, say $\hat{s}_2$. Then sensor $\hat{s}_2$ will play the role of $S_{southeast}$ and repeatedly executes the operations described above to find the next point $\hat{s}_3$. Let $\hat{s}_k$ be the last point which find $\hat{s}_1$ as its next point. Then we have constructed a path $P^{init} = (\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_k)$ which forms convex polygon Q.

In case of $k<n$, it implies that there are n-k remaining points that are inside Q but are not included in P^{init}. Let S^{rem} denote the set of remaining n-k points. The point $s_{closest}$ in S^{rem} that satisfies the following condition represents that it joins the path P^{init} will increase the least length. Hence the selected $s_{closest}$ will be chosen as the first point to join path P^{init}.

$$s_{closest} = arg \min_{s_j \in S^{rem}, \hat{s}_i \in Q} \left[dist\left(s_j, \hat{s}_i\right) + dist\left(s_j, \hat{s}_{i+1}\right) - dist\left(\hat{s}_i, \hat{s}_{i+1}\right)\right] \tag{9}$$

Let the constructed polygon be $Q = (\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_i, s_{closest}, \hat{s}_{i+1}, \ldots, \hat{s}_k)$. Let sink node in Q. Repeated executing the abovementioned operations until all n-k remaining points have been joined the polygon. Reorder the sensors in P^{init}. We have

$$\tilde{P} = \{\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_n\} \tag{10}$$

After finishing the construction of initial data collection path that passes through each sensor, the path length might too long. The next phase aims to reduce the length of path $\tilde{P}$.

B. Path Reduction Phase

This phase aims to reduce the length of path $\tilde{P}$. To determine the data collection segment for each sensor and construct a path which connecting the segments of all sensors, two tasks should be performed. The first one is to analyze the length of the collection segment. The second one further connects segments of all sensors to form a data collection path.

Definition: ***Data Collection Segment*** g_i of sensor s_i

A data collection segment g_i of sensor s_i is a chord of the circle of communication range of sensor s_i. The length of this segment should long enough such that mobile sink moves along it can completely receive all data buffered in sensor s_i.

Fig. 1 depict an example of data collection segment g_i of sensor s_i. In this phase, we aim to construct the shortest data collection segment for each sensor such that the connection of these segments will be the shortest path. For example, as shown in Fig. 1, the three sensors s_{i-1}, s_i and s_{i+1} have their own data collection segments. The dash line connecting the three segments can form a path along which the mobile sink can completely collect all data buffered in the three sensors. As shown in Fig. 1, the segment g_{i+1} is closer to sensor s_{i+1} than that of sensor s_i. This occurs because that

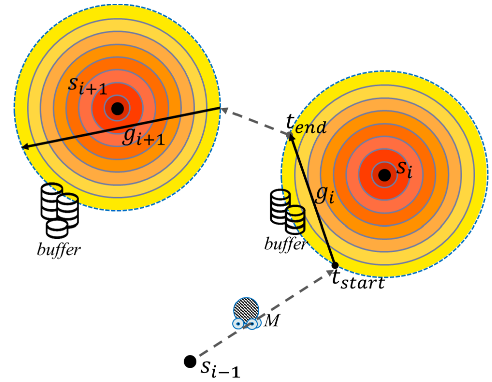


Fig. 1. The data collection segment g_i of sensor s_i.

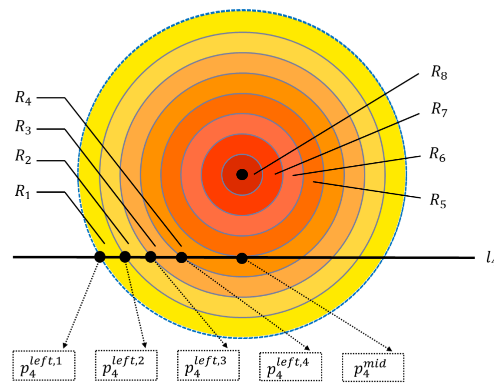


Fig. 2. Ring areas with different radiuses corresponding to different MTC types.

sensor s_{i+1} has more buffered data than sensor s_i. It is obvious that the length of segment g_{i+1} is larger than g_i. As a result, when the mobile sink moves along g_{i+1}, it can receive more data from sensor s_{i+1} which applies higher transmission rate to transmit its data.

One important demand when discussing the shortest segment length of g_i is that the mobile sink should receive all data stored in s_i when it moves along g_i. However, the data receiving rate highly depends on the distance of mobile sink M and s_i. Moving on the g_i, mobile sink M will change its modulation type to guarantee the performance of data receiving and hence its data receiving rate is changed accordingly. Assume that there are u types of MTC. The applied type of MTC is determined by s_i which selects one type based on the signal strength received by mobile sink. Since the distance between s_i and mobile sensor M changes with time, the type of MTC also changes accordingly. Assume that the mobile sink moves with constant speed. The following presents the design of the first task which aims to analyze the length of g_i.

C. Analyzes the Length of Data Collection Segment

The single strength received by mobile sink M highly depends on the distance of s_i and M. Since there are u types of MTC, we may partition the communication range of each sensor into u ring areas with the common center point (the location of s_i), each ring area with a radius corresponding to the distance of each MTC type. Fig. 2 gives an example of $u = 8$.

Let R_j, $1 \le j \le u$, denote the j-th ring area in each sensor, counting from the outermost ring. The communication range R of sensor s_i can be expressed by

$$R = \bigcup_{j \ge 1}^{u} R_j \tag{11}$$

Each ring area R_j consists of two boundary circles. Let C_j^{out} and C_j^{in} denote the outside and inside boundary circles of R_i, respectively. It is obviously:

$$C_{j-1}^{out} = C_j^{in}, \ \ for\ 2 \le j \le u \tag{12}$$

Consider that the mobile sink moves along g_i. The mobile sensor changes MTC type and applies different transmission rate since it moves crossing different areas. The following presents how to construct a data collection segment g_i for each sensor s_i. Let notation l_w denote the horizontal line that is tangent to C_w^{in}. Let Ω_w denote the set of ring areas that are passed by tangent l_w. We have

$$\Omega_w = \{R_1, \ldots, R_w\} \tag{13}$$

As shown in Fig. 2, the set of regions that are passed through by tangent l_4 is:

$$\Omega_4 = \{R_1, R_2, R_3, R_4\} \tag{14}$$

Let P_w be the ordered list of intersection points of l_w and Ω_w. Let p_w^{mid} denote the tangent point of l_w and C_w^{in}. Let P_w^{left} denote the set of intersection points that are located on the left side of the tangent point. Let $p_w^{left,i}$ denote the i-th intersection points on the left side of the tangent point. Consequently, we have

$$P_w^{left} = (p_w^{left,1}, \ldots, p_w^{left,w}) \tag{15}$$

Assume that the mobile sink moves along l_w with a constant speed. Let t_i denote the time point that the mobile sensor arrives intersection point $p_w^{left,i}$. Let

$$T_u = (t_1, t_2, \ldots, t_w) \tag{16}$$

denote the list of time points that M passes through the set of points P_u. The mobile sink M moving along l_w and collecting data from each sensor s_i should guarantee that the data stored in sensor s_i can be completely collected. Let $d(M, s_i)$ denote the distance between mobile sink and sensor s_i. Let P_i^{tx} and P_i^{rx} denote the transmitting and receiving power of sensor s_i, respectively. According to Friis's free space equation, the strength of power obtained by mobile sink M.

$$P_M^{rx} = \frac{G^{rx} G^{tx} \eta}{L^p} \left(\frac{\lambda}{4\pi(d(M,\ s_i) + \beta)}\right)^2 P_{s_i}^{tx} \tag{17}$$

where notations G^{rx} and G^{tx} denote the antenna gains of M and s_i, respectively and λ denotes by the amplitude, L^p denotes the polarity loss.

According to the MTC types, the mobile sink can adaptively select the most appreciate modulation level to have the best communication performance. Let $f(j).rate$ denote the receiving rate and strength of the j-th level modulation. That is, the data receiving rate of mobile sink located in ring area R_j is $f(j).rate$. Let t_{start} and t_{end} denote the time of mobile sensor arriving and leaving the points $p_i^{left,j}$ and $p_i^{left,j+1}$, respectively. When the mobile sensor get into R_j, the receiving rate of mobile sensor is $f(j).rate$. Since the path l_w passes through different regions R_j, the data receiving rate of mobile sink M changes accordingly. Equ. (18) evaluates the amount of data, denoted by δ_j^{max}, received by M.

$$\delta_j^{max} = 2 \sum_{R_j \in \Omega_w} \left(\int_{t_{start}}^{t_{end}} f(j).rate\ dt \right) \tag{18}$$

Because that mobile sink moves along path l_w and the l_w is tangent to R_j, the size of data received by the mobile sink totally depends on the R_j. That is, the size of data received by M is increased if the visited region is closer to sensor s_i. This occurs because that the mobile sink has higher transmission rate and longer time for receiving data. Let ϑ_i denote the size of data stored in s_i. The R_j that is tangent by l_w should satisfy Equ. (19).

$$R^{best} = \underset{R_j, 1 \le j \le u}{arg} \left(\delta_j^{max} \le \vartheta_i \le \delta_{j+1}^{max}\right) \tag{19}$$

Let R^{best} can be achieved when line l_w tangent to C_{best}^{in}. According to Equ. (19), the line l_w should be moved closer to s_i until it tangent to the C_{best}^{in} of region R^{best}.

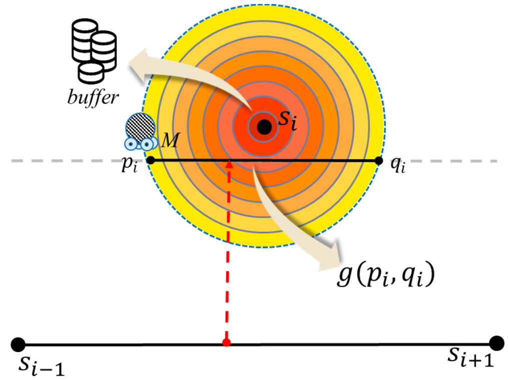


Fig. 3. The construction of segment g_i for sensor s_i.

The following presents ***Segment Construction*(*SC*) *Procedure*** for constructing data collection segment g_i of s_i. Let $\Delta s_{i-1} s_i s_{i+1}$ denote the triangle connecting sensors s_{i-1}, s_i and s_{i+1} as shown in Fig. 3. Let $l(s_{i-1}, s_{i+1})$ denote the line connecting s_{i-1} and s_{i+1}. Let line l_i denote the line parallel to the line $l(s_{i-1}, s_{i+1})$. Move line l_i closing to the location of sensor s_i until l_i tangent to the C_{best}^{in} of region R^{best} which can be obtained by applying Equ. (19). Then line l_i will satisfy the property of g_i. Let points p_i and q_i represent the left and right intersecting points of sensor s_i and the segment l_i, respectively. Let segment $g(p_i, q_i)$ denote the line connecting points p_i and q_i. That is, $g(p_i, q_i)$ has the important property that the mobile sink M moving along the segment $g(p_i, q_i)$ with a constant velocity υ can collect all data of sensor s_i.

Let O_i denote the i-th group, which is organized by $s_{(i-1)}$, s_i and $s_{(i+1)modn}$. Applying the abovementioned **SC Procedure**, the data collection segment $g_i = \bar{p_i q_i}$ of sensor s_i can be constructed.

D. Connects Segments of All Sensors to Form a Data Collection Path

This task aims to connect the data collection segment of each sensor to form a data collection path. These three

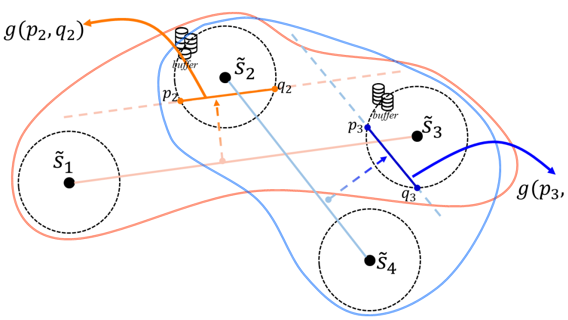


Fig. 4. Partitioning sensors into groups and constructing data collection segment for each sensor.

sensors in according to the abovementioned operations, each sensor s_i has constructed the collection segment g_i and has two intersection points with the communication circle of corresponding points p_i and q_i. Then the data collection path can be further constructed by connecting points q_{i-1} and p_i and connecting points q_i and $p_{(i+1)\text{mod}n}$, for $1 \leq i \leq n$. Fig. 3 gives the constructed path, called P^{best}. Recall that Phase I will construct a data collection path $\tilde{P} = \{\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_n\}$. This path will be the input of phase II. The phase II aims to reduce the path length by constructing a segment for each sensor and connecting all segments to form a shorter path.

Based on the given path $\tilde{P} = \{\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_n\}$, notation O_i denote the i-th group, which is organized by three consecutive sensors $\tilde{s}_{(i-1)}$, $\tilde{s}_i$ and $\tilde{s}_{(i+1)\text{mod}n}$. Since the task one of this phase will construct a data collection segment $g_i = \bar{p_i q_i}$ for each sensor $\tilde{s}_i$, the reduced circular path, denoted by $\hat{\boldsymbol{P}}$, can be further constructed by connecting each point n_i to point $m_{(i+1)modn}$. As a result, a new path

$$\hat{\boldsymbol{P}} = (p_1, q_1, p_2, q_2, \ldots, p_n, q_n, p_1) \tag{20}$$

can be obtained.

The following gives an example to illustrate the operations of path construction. As shown in Fig. 4, there are four sensors: $\tilde{s}_1, \tilde{s}_2, \tilde{s}_3$ and $\tilde{s}_4$. Initially, the sensors are partitioned into several groups. Each group consists of three sensors. In this example, sensors $\tilde{s}_1, \tilde{s}_2$ and $\tilde{s}_3$ belong to one group while sensors $\tilde{s}_2, \tilde{s}_3$ and $\tilde{s}_4$ belong to another group. Then the mobile sink constructs the data collection segment for each sensor. For instance, the mobile sink constructs a data collection segment $g_2 = g(p_2, q_2)$ which connects points p_2 and q_2. The constructed $g(p_2, q_2)$ should satisfy Equ. (18) to guarantee that mobile sink moves along the constructed segment $g(p_2, q_2)$ can completely receive all data from sensor $\tilde{s}_2$. Since there are many segments with different angles for sensor $\tilde{s}_2$, the next step aims to fix the angle such that the final segment can be connected to the two segments of sensors in the same group. To achieve this, a line $l(\tilde{s}_1, \tilde{s}_3)$ connecting locations of sensors $\tilde{s}_1$ and $\tilde{s}_3$ will be constructed. Let $g(\tilde{s}_1, \tilde{s}_3)$ denote the intersected segment of line $l(\tilde{s}_1, \tilde{s}_3)$ and the communication circle of $\tilde{s}_2$. The line $l(\tilde{s}_1, \tilde{s}_3)$ will be moved closer to the location of sensor $\tilde{s}_2$ until the length of $g(\tilde{s}_1, \tilde{s}_3)$ is equal to the length of g_2. Then the $g(\tilde{s}_1, \tilde{s}_3)$ will be the final data collection segment $g(p_2, q_2)$ of sensor $\tilde{s}_2$. Similarly, all data collection segments $g_i = g(p_i, q_i)$ of sensors $\tilde{s}_i$ can be constructed according to the abovementioned three steps. As shown in Fig. 5, all segments will be connected

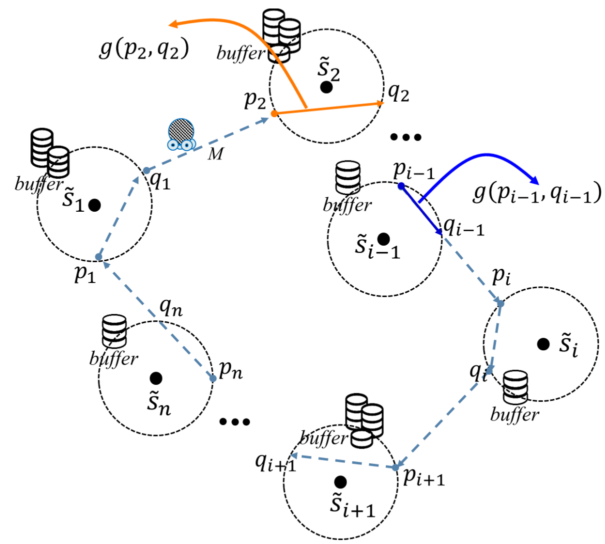


Fig. 5. The connection of all segments of sensors.

to form the data collection path $\hat{\boldsymbol{P}}$ by connecting points $p_1, q_1, p_2, q_2, \ldots, p_n, q_n, p_1$.

E. The Proposed MRDC Algorithm

This section presents the proposed MRDC algorithm. The proposed algorithm mainly consists of two phases: the path construction phase and the path reduction phase. In the path construction phase, an initial path which passes through each sensor will be constructed. As a result, each sensor has an visited order in the path. In the second phase, the legnth of initial path will be further reduced. In this phase, the 'data collection segment' of each sensor is firstly defined. Then the minimal length of data collection segment of each sensor will be derived to guarantee that mobile sink can completely receive data from the visited sensor. Finally, the data collection segments of all sensors will be connected according to its order to form a reduced path.

There are three inputs, including $S = \{s_1, s_2, \ldots, s_n\}$, the location of each sensor, and the size of data amount ϑ_i of sensor s_i, respectively. The output of MRDC is the established data collection path $\hat{P}$.

Steps 1 to 3 select the southeast point in S as the first visited point in the initial path. Steps 4 to 8 aim to construct a path with polygon shape. Steps 9 to 12 add the remaining sensors inside the polygon to the constructed path. Steps 13 to 15 intend to group every three consecutive sensors according to the ordered sequence. Steps 16 to 20 further construct the data collection segment for each sensor. Finally, step 21 interconnects the data collection segments of all sensors.

Herein, it is worth to notice that When the communication ranges of multiple sensors are overlapped, it is difficult to apply the proposed chord segment because the overlapping cases are very complicate. In this situation, the proposed MRDC will simply create a subpath which passes through the point fallen in the overlapped region. Since the amount of collected data along the constructed subpath can be calculated, the mobile sink will stop for a while when collecting data from the overlapped sensors if the data cannot be completed collected.

Algorithm 1 *Procedure:Multi-Rate Data Collection (MRDC) Algorithm*

Inputs:

1. A set of n static sensors $S = \{s_1, s_2, \ldots s_i, \ldots s_n\}$.
2. The location of each sensor is (x_i, y_i) .
3. The size of data stored in s_i is ϑ_i.

Outputs:

1. The collection path: $\hat{P}$.

/* **Initial Data Collecting Path Construction (IDCPC) Phase** */

1. $S_{southeast} = \arg\min_{s_i \in S} (y_i)$;
2. $\hat{s}_1 = S_{southeast}$;
3. $P^{init} = (\hat{s}_1)$;
4. Let π be a horizontal line passing through $\hat{s}_1$;
5. **Repeat**{
   i++;
6. Turn π in a counterclockwise direction until it touches any point, say $\hat{s}_i$;
   Add $\hat{s}_i$ to path P^{init};
   Move π to pass through $\hat{s}_i$;
7. **Until** $(\hat{s}_i == \hat{s}_1)$ }
   $k = i$; /* $P^{init} = (\hat{s}_1, \ldots, \hat{s}_k)$*/
8. **If** $(k < n)${
9. Let S^{rem} denote the set of remaining n-k points;
10. **For** ($j = 1$; $j \leq n$-k; j++){
11. Compute the point $s_{closest}$ in S^{rem} according to Equ.(9);} $\tilde{P}=(\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_i, s_{closest}, \hat{s}_{i+1}, \ldots, \hat{s}_k)$}
12. Return $\tilde{P}$;

/* **Path Reduction Phase***/

13. Let $\tilde{P} = \{\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_n\}$ denote P^{init};
14. **For** $(i = 2;\ i < n;\ i++)$
15. Group $\tilde{s}_{(i-1)}, \tilde{s}_i, \tilde{s}_{(i+1)mod n}$ as a group;
16. **For** $(i = 1$; i<n; i++){
17. Calculate δ_j^{max} for sensor $\tilde{s}_i$ according to Equ.(18);
18. Let $l(s_{i-1}, s_{i+1})$ denote the line connecting s_{i-1} and s_{i+1};
19. Move segment $l(s_{i-1}, s_{i+1})$ toward s_i until Equ. (18) satisfied;
20. Let p_i and q_i be the left and right intersection points of $l(s_{i-1}, s_{i+1})$ and R^{best}, respectively;
21. $\hat{P} = \hat{P} + (p_i, q_i)$; }
22. **Return** $\hat{P}$;

V. Performance Evaluation

This section compares the proposed algorithm, called **MRDC**, and the existing MSFLA [18] and CDCA [19] algorithms in terms of the path length, improvement rate as well as effective index. The CDCA algorithm [19] is selected to compare with the proposed MRDC algorithm because that it is the newest path construction algorithm which can flexibly allow that mobile sink needs not to pass through the center of each CP. This property allows the CDCA can be fairly compared with the proposed MRDC since the path constructed by the MRDC does not pass through the center of each sensor. The CDCA can be degenerated from multiple mobile sinks into single mobile sink. To fairly compare the existing CDCA and the proposed MRDC, only single mobile sink is considered in the experiments.

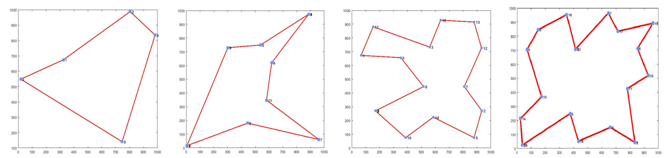


Fig. 6. The random deployed scenario.

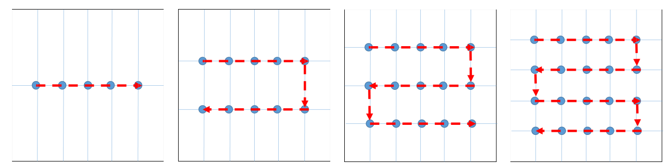


Fig. 7. The regular deployed scenario.

The existing MSFLA and CDCA algorithms selected CPs from sensors and all sensors transmit their data to the selected CPs. Then the existing MSFLA and CDCA algorithms constructed paths passing through the selected CPs for mobile sink to visit and collect data from CPs. Since the proposed MRDC constructs a path passing through all sensors, to fairly compare the proposed MRDC with the existing CDCA and MSFLA in the experiments, the proposed MRDC applies the CP selection mechanisms proposed in CDCA and MSFLA. Then the MRDC constructs the path which only passes through the selected CPs. In this way, the major difference between MRDC and the existing MSFLA and CDCA algorithms is the path construction technique. As presented in the last section, the proposed MRDC constructs the data collection segment for each CP by considering the data amount buffered in that CP and dynamically adopting the MTC, aiming to reduce the path length but guarantee the complete data collection from each CP to the mobile sink.

Two scenarios are considered in the experiments. The first one is a randomly deployed WSN while the second one is regularly deployed WSN. The communication ranges of all CPs are not intersected.

The following illustrates the parameters considered in the simulation environment. The considered monitor region is sized 1000*1000 unit^2. The number of sensors deployed in the monitor region is varied including 50, 100, 150 and 200. The initial energy of each sensor is 100J. The transmission range of each CP is set ranging from 10m to 75m. There are four levels of data volume offered to each sensor: 250 (KB) to 500 (KB), 500 (KB) to 1000 (KB), 1000 (KB) to 1500 (KB) and 1500 (KB) to 2000 (KB). Table II summarizes the parameters.

Figs. 8 and 9 depict the path lengths of the compared MSFLA, CDCA and MRDC algorithms in the random scenario. The common trend of the compared MSFLA, CDCA and MRDC algorithms is that the path lengths are increased with the number of CPs. In comparison, the proposed MRDC

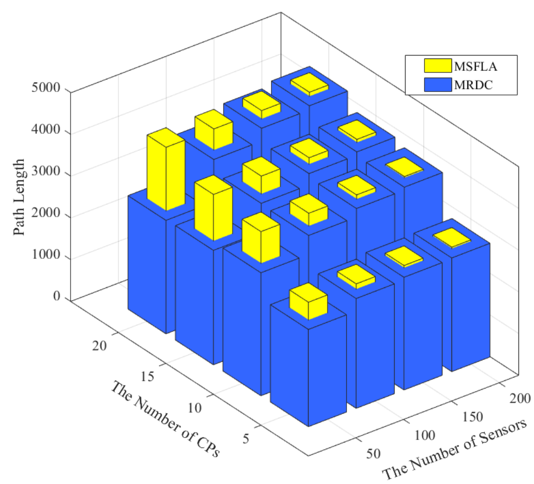


Fig. 8. The comparison of MRDC and MSFLA in terms of path length in the random scenario.

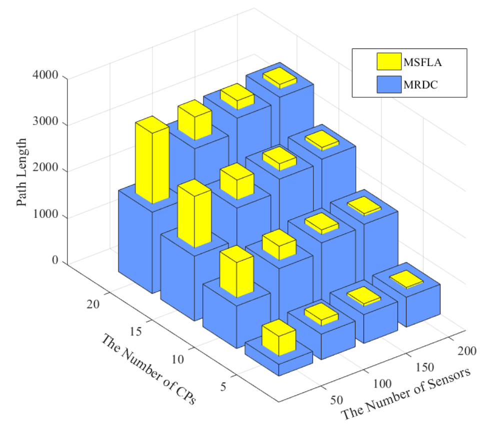


Fig. 10. The comparison of MRDC and MSFLA in terms of path length in the regular scenario.

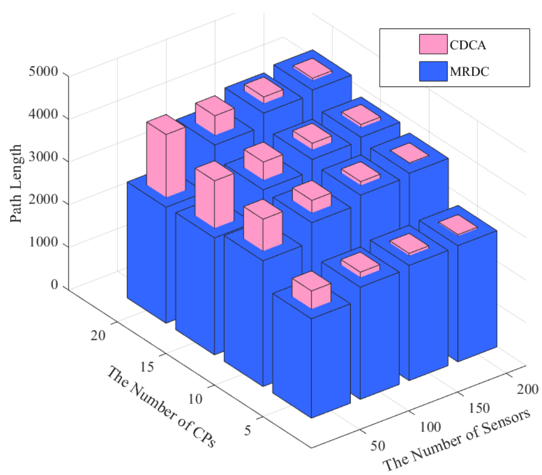


Fig. 9. The comparison of MRDC and CDCA in terms of path length in the random scenario.

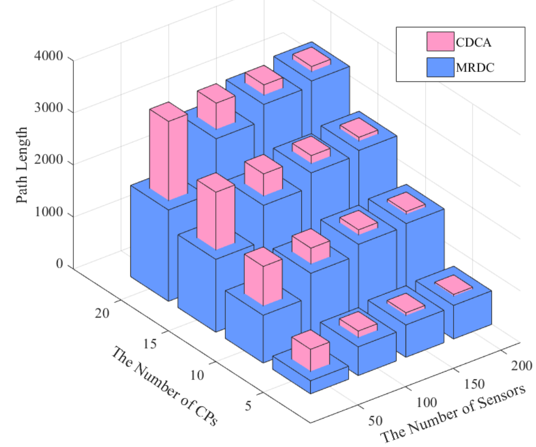


Fig. 11. The comparison of MRDC and CDCA in terms of path length in the regular scenario.

outperforms the other two algorithms in terms of path length. This occurs because that the proposed MRDC calculates the transmission rates, adopts multi-rate transmission, constructs a shorter chord and hence establishes a path for mobile sink. As a result, the path of MRDC needs not to pass through the center of each CP and is shorter than those of MSFLA and CDCA. As shown in Fig. 8, the height of yellow bar represents the path length reduced by the proposed MRDC, as compared with MSFLA. It is interesting that the height of yellow bar decreased with the number of sensor nodes. This occurs because that the data volume of each CP is increased with the number of sensor node, leading to the chord of each CP closer to its center when applying the proposed MRDC. As a result, the path lengths of MRDC and MSFLA are approached when the number of sensors is increased. This result is significant especially when the number of sensors is 200. The pink bars in Fig. 9 also depict similar trend when the number of sensors are changed.

Fig. 10 depicts the path lengths of the mobile sink by applying the compared MSFLA and MRDC algorithms in the regular scenario. The common trend of MRDC and MSFLA is that the path length is increased with the number of sensors. The proposed MRDC outperforms the MSFLA algorithm because that the path of MRDC does not pass through the center of each CP. However, the path lengths of MRDC and MSFLA are approached when the number of sensors is increased. This occurs because that the paths constructed in the regular scenario are almost straight lines, which makes no significant difference between MRDC and MSFLA. This result also can be observed in Fig. 11.

A good data collection algorithm will keep a higher effective index $\zeta_{effective}$, which is measured by the total data volume to the whole path length, as depicted in Equ. (21).

$$\zeta_{effective} = \frac{\sum_{i=1}^{i=n} \vartheta_i}{\hat{P}} \tag{21}$$

A large value of $\zeta_{effective}$ indicates that the data collection algorithm creates a short path but the mobile sink can collect massive data along this path. Fig. 12 compares the three algorithms in terms of effective index from the random scenario. The offered data for each CP is ranging from 250 ∼2000 KB. In general, the efficiency indices of CDCA, MSFLA and MRDC increase with the data volume. This occurs because the volume of data received by mobile sink is increased.

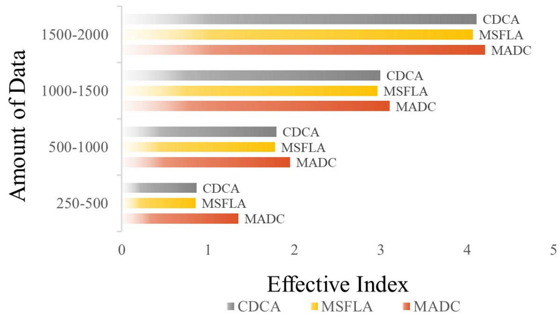


Fig. 12. The comparison of the effective index of three algorithms in case of 20 CPs.

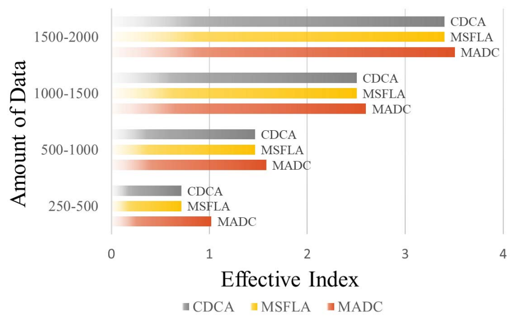


Fig. 13. The comparison of the effective index of three algorithms in case of 15 CPs.

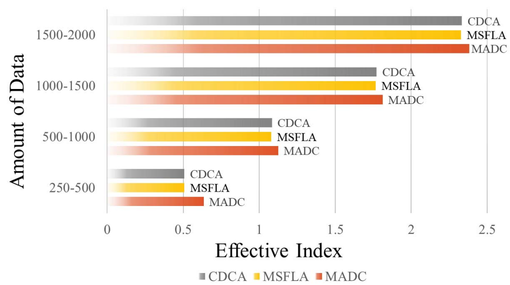


Fig. 14. The comparison of the effective index of three algorithms in case of 10 CPs.

The proposed MRDC outperforms the other two algorithms in all cases. The proposed MRDC calculates the data volume of each CP and constructs a chord for each CP. Since the chord length is shorter than the diameter of each CP, the proposed MRDC construct a shorter path and hence achieves higher effective index, as compared with the other two algorithms. Figs. 13, 14, 15 have similar results where the performance of the proposed MRDC outperforms the other two algorithms in terms of effective index.

Fig. 16 further compares the effective index of the three compared algorithms by varying the buffer size of each CP and the data volume in the random scenario. The number of deployed CPs is 20. The data volume is varied from 250 to 2000 KB while the buffer size of each sensor is varied from 500 to 2000 KB. One common trend of the three compared algorithms is that their effective indices are increased with data volume and buffer size. This occurs because that the larger buffer size reduces the probability of data overflow in each CP.

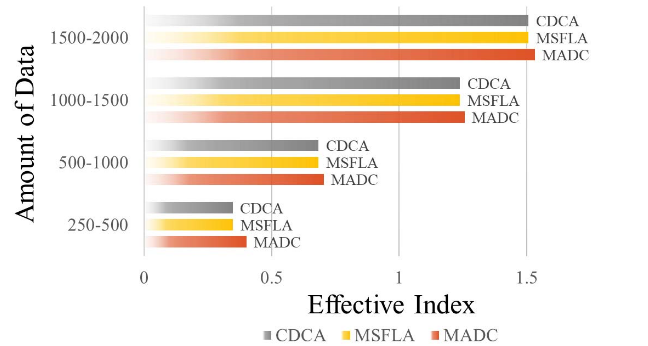


Fig. 15. The comparison of the effective index of three algorithms in case of 5 CPs.

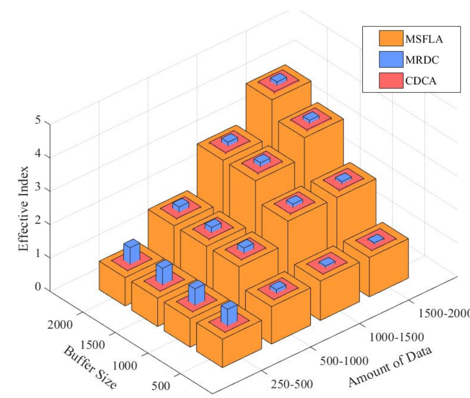


Fig. 16. The comparison of the effective index of three algorithms in case of 20 CPs. The buffer size and the amount of data are varied.

In comparison, the proposed MRDC outperforms MSFLA and CDCA. When the data volume grows, the effective indices of the three compared algorithms are approached. This occurs because that the chord constructed by the proposed MRDC will be closer to the diameter when the buffer size grows.

Fig. 17 depicts similar experiments as Fig. 16. In general, the two Figures have similar trend that the effective index is increased with amount of data and buffer size. The major difference between Figs. 16 and 17 is that the number of CPs has been reduced. That is, the numbers of CPs in Figs. 16 and 17 are 20 and 10, respectively. Compared with Fig. 16, the effective indices of Fig. 17 are reduced. This occurs because that the total data volume is reduced when the number of CPs is reduced.

Let notation $\tau_{improvement}(A)$ denote the improvement index of the proposed MRDC in terms of path length, as compared with existing algorithm A. Equ. (22) calculates the value of improvement index, where P^{RADC} and P^{A} denote the path lengths constructed by applying the proposed *MRDC* and existing algorithm A, respectively.

$$\tau_{improvement}(A) = 1 - \frac{P^{RADC}}{P^{A}} \qquad (22)$$

A large value of $\tau_{improvement}(A)$ indicates that the *MRDC* has a big improvement against the existing algorithm A. For examples, $\tau_{improvement}(MSFLA)$ and $\tau_{improvement}(CDCA)$ denote the improvements of the proposed MRDC algorithm against the existing algorithms

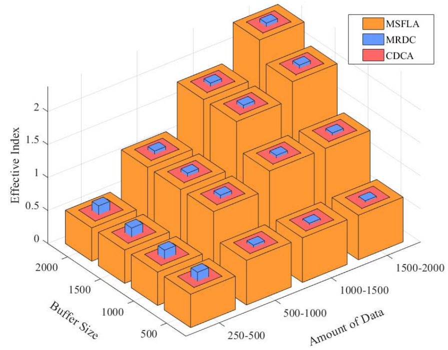


Fig. 17. The comparison of the effective index of three algorithms in case of 10 CPs. The buffer size and the amount of data are varied.

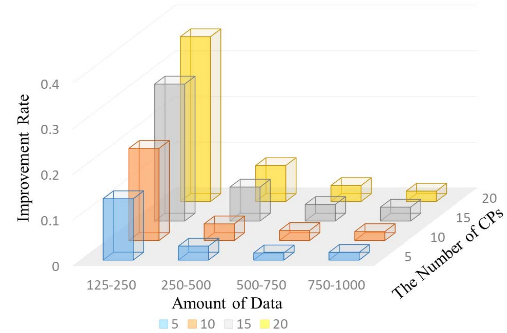


Fig. 19. The comparison of improvement ratio of MRDC to CDCA.

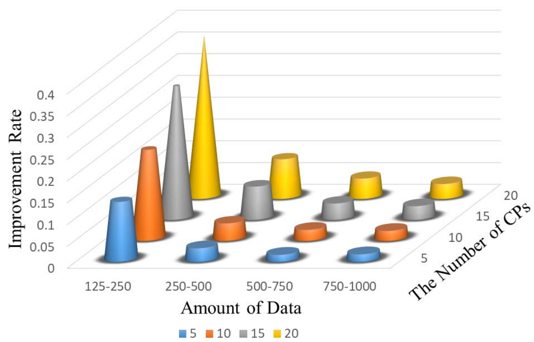


Fig. 18. The comparison of improvement ratio of MRDC to MSFLA.

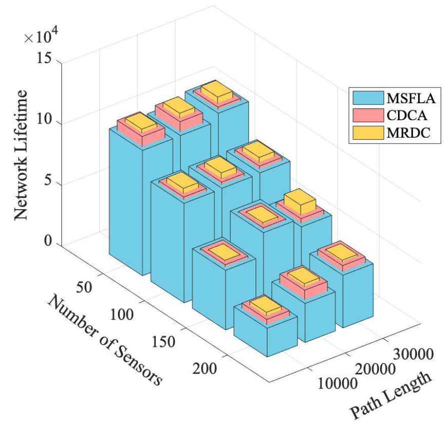


Fig. 20. The comparison of the three algorithms in terms of network lifetime.

MSFLA and CDCA, respectively. Figs. 18 and 19 measure the improvement indices $\tau_{improvement}(MSFLA)$ and $\tau_{improvement}(CDCA)$, respectively, in the random scenario. Two parameters, including the number of CPs and the amount of data are varied.

As shown in Fig.18, the improvement index is decreased with the amount of data but increased with the number of CPs. This occurs because that the existing MSFLA approach needs to pass through the center of each CP. As a result, the length of path constructed by applying MSFLA has significant incensement. Since the proposed MRDC does not need to pass through the center of each CP, it reduces larger length when the number of CPs grows, as compared with the existing MSFLA. On the other hand, when the amount of data grow, the chord constructed by applying the proposed MRDC is closer to the center of CPs, resulting in a small improvement, as compared with MSFLA.

Fig. 19 has similar result when the number of CPs and the data volume of each CP are changed.

Fig. 20 compares the performances of MRDC, CDCA and MSFLA in terms of network lifetime. Herein, we notice that the proposed MRDC constructs a path passing through all sensors. Different from the proposed MRDC, the existing CDCA and MSFLA select CPs and the path of mobile sink only passes through CPs. To fairly compare the three algorithms in terms of network lifetime of the sensor networks, a large number of sensors are deployed in the region and the mechanisms for selecting CPs in CDCA and MSFLA are applied in the proposed MRDC. Then the three algorithms are applied to construct different paths passing through the selected CPs. The number of sensors varies ranging from 50 to 200 while the path length varies from 10000 to 30000 units. In general, the network lifetime decreases with the number of sensors and increases with the path length. This occurs because that a large number of sensors can generate larger amount of sensing data, which increases the forwarding load of the CPs and hence reduces the network lifetime. However, a long path can increase the number of visited CPs and hence reduce the forwarding load, increasing the network lifetime. In comparison, the proposed MRDC achieves better performance than CDCA and MSFLA in all cases, as shown in Fig. 20. This occurs because that the proposed MRDC constructs a shorter path for passing through the same set of CPs. The remaining path of MRDC can further visit more CPs, reduce the forwarding loads of the CPs and hence improve the lifetime of the sensor networks.

Fig. 21 compares the performances of MRDC, CDCA and MSFLA in terms of energy consumption of the mobile sink. The number of sensors varies ranging from 50 to 200 while the number of CPs varies ranging from 5 to 20. A general trend of the three algorithms is that the energy consumption of the mobile sink increases with both the numbers of sensors and CPs. This occurs because that the WSN containing more sensors can generate a larger amount of data which consume more energy of the mobile sink for receiving the data. Similarly, when the number of CPs increases, the mobile sink increases

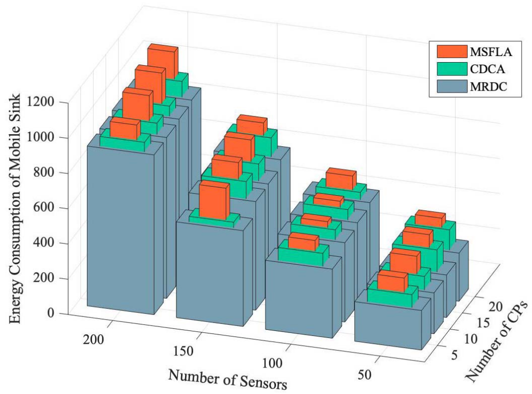


Fig. 21. The comparison of three algorithms in terms of energy consumption of mobile sink.

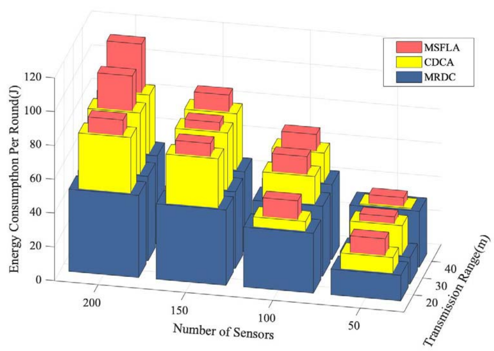


Fig. 22. The comparison of energy consumptions of three algorithms.

its energy consumption for movement. In comparison, the proposed MRDC outperforms the existing CDCA and MSFLA. This occurs because that the proposed MRDC constructs the data collection segment for each sensor by considering the amount of data and the MTC (machine type communication). As a result, a shorter path can be constructed, which reduces the energy consumption of mobile sink for moving along the path.

Fig. 22 further compares the total energy consumption of sensors in each round by applying the compared three algorithms. The number of sensor nodes varies ranging from 50 to 200, whereas the transmission range varies ranging from 20 m to 40 m. As shown in Fig. 22, the three algorithms exhibited a similar trend that the total energy consumption increases with the number of sensor nodes and the communication range. This occurred because more sensors can lead to a larger amount of data, increasing the forwarding loads of CPs. Similarly, when the transmission range is enlarged, each sensor increases its energy consumption. In comparison, the proposed MRDC yields the best performance. The main reason is that MRDC reduces the path length by considering chord as the sub-path within the communication range of each sensor, as compared with the existing CDCA and MSFLA.

VI. Conclusions

Data collection is one of the most important issues in wireless sensor networks. Mobile sink can help reduce and balance the energy consumption of sensor nodes since it can visit sensor nodes and collect data from them along a specific path. This paper aims to construct the shortest and energy efficient data collection path but guarantee that the data of each sensor can be totally received by the mobile sink. The main idea of this paper is to dynamically adopt the appreciate transmission rate according to the data amount of each sensor. This helps reduce the path length of the mobile sink and hence can reduce the fresh cycle of data collection. The proposed MRDC consists of two phases: Initialization Phase and Path Reduction Phase. The *Initialization Phase* aims to construct an initial collecting path that passes thought all sensors. Based on the constructed path, all sensors have been ordered in a certain visited sequence. Then the *Path Reduction Phase* further partitions the sensors into many groups, according to the ordered sequence, the MRDC further reduces the path length of each group by applying 'multi-rate transmission' and 'data collecting while moving' schemes. To guarantee that mobile sink can receive all data from each sensor, a chord, called data collection segment, is constructed for each sensor. Finally, the paths of all chords are interconnected to form the data collecting path. Performance evaluation shows that the proposed MRDC outperforms existing schemes in terms of the network lifetime, energy consumption of each sensor, path length, effective index as well as improvement index.

Future work will consider the support of multiple mobile sinks. The key issue will be the cooperation between mobile sinks, including the task partition, load balance and rendezvous points arrangement for data exchange between mobile sinks.

References

[1] G. Li *et al.*, "Energy efficient data collection in large-scale Internet of Things via computation offloading," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4176–4187, Jun. 2019.

[2] C. Zhang and S. Fei, "A data collection algorithm with mobile elements in wireless sensor networks," in *Proc. 33rd Youth Academic Annu. Conf. Chin. Assoc. Autom. (YAC)*, Nanjing, China, May 2018, pp. 1040–1044.

[3] J. Chen, N. Wang, D. Chen, and J. Wan, "An adaptive dynamic topology WSN data collection method based on compressive sensing," in *Proc. 2nd IEEE Adv. Inf. Manage., Commun., Electron. Autom. Control Conf. (IMCEC)*, Xi'an, China, May 2018, pp. 1376–1379.

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

[5] K. Liu *et al.*, "An active mobile charging and data collection scheme for clustered sensor networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 5100–5113, May 2019.

[6] N. Annuar, N. Bergmann, R. Jurdak, and B. Kusy, "Poster abstract: Energy efficient mobile data collection from sensor networks with range-dependent data rates," in *Proc. 17th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Porto, Portugal, Apr. 2018, pp. 122–123.

[7] M. Ma and Y. Yang, "Data gathering in wireless sensor networks with mobile collectors," in *Proc. IEEE Int. Symp. Parallel Distrib. Process.*, Apr. 2008, pp. 1–9.

[8] C. Wang and H. Ma, "Data collection with multiple controlled mobile nodes in wireless sensor networks," in *Proc. IEEE 17th Int. Conf. Parallel Distrib. Syst.*, Dec. 2011, pp. 489–496.

[9] *Wireless Medium Access Control (MAC) and Physical Layer (PRY) Specifications for Low Rate Wireless Personal Area Networks (LR-WPANs)*, IEEE Standard 802.15.4, May 2003.

[10] R. Anwit, "A variable length genetic algorithm approach to optimize data collection using mobile sink in wireless sensor networks," in *Proc. 5th Int. Conf. Signal Process. Integr. Netw. (SPIN)*, Feb. 2018, pp. 73–77.

[11] J. Tao, L. He, Y. Zhuang, J. Pan, and M. Ahmadi, "Sweeping and active skipping in wireless sensor networks with mobile elements," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Anaheim, CA, USA, Dec. 2012, pp. 106–111.

[12] T.-S. Chen, W.-Q. Du, and J.-J. Chen, "Energy-efficient data collection by mobile sink in wireless sensor networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Marrakesh, Morocco, Apr. 2019, pp. 1–6.

[13] D. G. Macharet, J. W. G. Monteiro, G. R. Mateus, and M. F. M. Campos, "Bi-objective data gathering path planning for vehicles with bounded curvature," *Comput. Oper. Res.*, vol. 84, pp. 195–204, Aug. 2017.

[14] J. Li, M. Zhao, and Y. Yang, "OWER-MDG: A novel energy replenishment and data gathering mechanism in wireless rechargeable sensor networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Anaheim, CA, USA, Dec. 2012, pp. 5350–5355.

[15] W. Wen, S. Zhao, C. Shang, and C.-Y. Chang, "EAPC: Energy-aware path construction for data collection using mobile sink in wireless sensor networks," *IEEE Sensors J.*, vol. 18, no. 2, pp. 890–901, Jan. 2018.

[16] D. Kim, B. H. Abay, R. N. Uma, W. Wu, W. Wang, and A. O. Tokuta, "Minimizing data collection latency in wireless sensor network with multiple mobile elements," in *Proc. IEEE INFOCOM*, Orlando, FL, USA, Mar. 2012, pp. 504–512.

[17] C. Liu, H. Du, and Q. Ye, "Utilizing communication range to shorten the route of sweep coverage," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017, pp. 1–6.

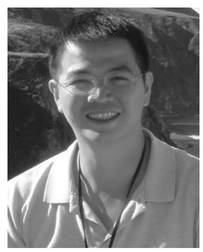
[18] X. Gu, X. Zhou, B. Yuan, and Y. Sun, "A Bayesian compressive data gathering scheme in wireless sensor networks with one mobile sink," *IEEE Access*, vol. 6, pp. 47897–47910, Aug. 2018.

[19] W. Wen, C.-Y. Chang, S. Zhao, and C. Shang, "Cooperative data collection mechanism using multiple mobile sinks in wireless sensor networks," *Sensors*, vol. 18, no. 8, pp. 2627–2646, Aug. 2018.

Chih-Yung Chang (Member, IEEE) received the Ph.D. degree in computer science and information engineering from National Central University, Taiwan, in 1995. He is currently a Full Professor with the Department of Computer Science and Information Engineering, Tamkang University, New Taipei, Taiwan. His current research interests include the Internet of Things, wireless sensor networks, artificial intelligence, and deep learning. He has served as an Associate Guest Editor for several SCI-indexed journals, including the *International Journal of Ad Hoc and Ubiquitous Computing* (IJAHUC) from 2011 to 2020, the *International Journal of Distributed Sensor Networks* (IJDSN) from 2012 to 2014, *IET Communications* in 2011, *Telecommunication Systems* (TS) in 2010, the *Journal of Information Science and Engineering* (JISE) in 2008, and the *Journal of Internet Technology* (JIT) from 2004 to 2008.

Shi-Yong Chen received the B.S. degree in multimedia and game science from the Lunghwa University of Science and Technology, Taiwan, in 2015, and the M.S. degree in information engineering from Tamkang University, Taiwan, in 2018, where she is currently pursuing the Ph.D. degree in computer science and information engineering. Her current research interests include wireless sensor networks and the Internet of Things.

I-Hsiung Chang received the Ph.D. degree in education from National Chengchi University, Taiwan, in 2012. He is currently pursuing the Ph.D. degree with the Department of Computer Science and Information Engineering, Tamkang University, Taiwan. His current research interests cover the joint of education and information technologies in AI-IoT, robot, and big data.



Gwo-Jong Yu (Member, IEEE) received the Ph.D. degree in computer science and information engineering from National Central University, Taiwan, in 2001. He is currently a Full Professor with the Department of Computer Science and Information Engineering, Aletheia University, New Taipei, Taiwan. His research interests include the Internet of things, 5G, deep learning, artificial Intelligence, and computer vision.

Diptendu Sinha Roy received the Ph.D. degree in engineering from the Birla Institute of Technology, Mesra, India, in 2010. He is currently an Associate Professor with the Department of Computer Science and Engineering, National Institute of Technology Meghalaya, India. His current research interests include the Internet of Things, big data, machine learning, and soft and evolutionary computing.