Time Efficient Data Collection with Mobile Sink and vMIMO Technique in Wireless Sensor Networks

Ye Miao[*†], Zhili Sun[†], Ning Wang[†], Yue Cao[†‡] and Haitham Cruickshank[†]

Abstract—Data collection is a fundamental yet challenging task of Wireless Sensor Networks (WSN) to support a variety of applications, due to the inherent distinguish characteristics for sensor networks, such as limited energy supply, self-organizing deployment and QoS requirements for different applications. Mobile sink and virtual MIMO (vMIMO) techniques can be jointly considered to achieve both time efficient and energy efficient for data collection. In this paper, we aim to minimize the overall data collection latency including both sink moving time and sensor data uploading time. We formulate the problem and propose a multihop weighted revenue (MWR) algorithm to approximate the optimal solution. To achieve the trade-off between full utilization of concurrent uploading of vMIMO and the shortest moving tour of mobile sink, the proposed algorithm combines the amount of concurrent uploaded data, the number of neighbours, and the moving tour length of sink in one metric for polling point selection. The simulation results show that the proposed MWR effectively reduces total data collection latency in different network scenarios with less overall network energy consumption.

Index Terms—multi-hop wireless sensor networks, data collection, vMIMO, mobile sink, delay minimization.

I. INTRODUCTION

WSNs have emerged as a promising technology with numerous and various military and civil applications, such as environment monitoring [1], disaster management [2], tactical surveillance [3], etc. Recently, mobile sinks are proposed and explored as an alternative solution to achieve load balancing and uniform distribution of energy consumption in sensor networks [4–7]. As the sink moves, hotspot nodes change and thus the high energy consumption zones around sinks shift. However, mobile sink introduces sink moving latency in data collection, which is large due to the wide-spread network topologies. On the contrary, data transmission time is largely reduced due to concurrent transmission of the same data from different independent data streams by applying virtual multiple-input multiple-output (vMIMO)[8, 9]. That is, several sensor nodes equipped with one or more antennas can emulate a multi-antenna node and achieve multiple independent transmit-receive paths[10].

Due to a large number of sensor nodes or large amount of sensing data in some specific data collection scenarios, the data uploading time of sensors could be large enough to be comparable with the sink moving time and become a dominant factor in the total data collection latency. The previous works have been mainly focused on the sink moving time only and the problem about total data collection latency including both the data uploading time of sensors and sink moving time hasn't been addressed properly. To investigate the joint design of mobile sink and concurrent uploading vMIMO technique,we aim to minimise the total data collection latency.

On one hand, the sink moving time depends on the trajectory design of mobile sink which is about the selection of polling point (PP) where the mobile sink stops to collect sensing data. The sink moving time can be reduced by decreasing the total number of polling points (PPs) and shorting the sink moving tour length. On the other hand, the overall data uploading time is decided by data uploading associations. The two sensors emulating a two-antenna node to perform concurrent data uploading are called a compatible sensor pair. With the selection of compatible sensors, the other sensors can transmit their sensing data to the compatible sensor pairs. One challenge in achieving full utilization of concurrent data uploading is to evenly distribute the associated data to the compatible sensors. Thus, the main issues this paper addresses are how to improve the utilization of vMIMO to shorten data uploading time, how to design the sink moving tour to shorten the sink time and how to jointly consider the two aspects to achieve the shortest total data collection latency.

The main contributions of this paper can be summarised as: (i) To investigate the joint design of mobile sink and concurrent uploading vMIMO technique. (ii) To formulate the delay minimization for multihop data collection (DM-MDC) problem. (iii) To propose a multihop weighted revenue (MWR) algorithm for the DMMDC problem. (iv) To carry out extensive simulations to evaluate the proposed MWR algorithm. The remainder of this paper is organized as follows. Section II provides the related works of data collection in WSNs applying vMIMO and mobile sink techniques. Section III presents system model and problem formulation. Section IV illustrates the proposed MWR algorithm. Simulation results and performance analysis are presented in Section V. Finally, Section VI concludes the paper.

II. RELATED WORKS

Jayaweera et al [11] firstly proposed Vertical Bell Laboratories Layered Time (VBLAST) based vMIMO communication architecture in WSNs. Taking into account both transmission and circuit power consumption, they analysed the network energy consumption and showed that the VBLAST based vMIMO architecture can offer significant energy savings over traditional Single-Input and Single-Output (SISO) communication based WSNs. Nguyen et al [12] compared energy consumption performance from different vMIMO techniques: VBLAST and distributed space-time coded (DSTBC) [13]. They illustrated that DSTBC is more energy efficient that VBLAST for long communication distances by utilizing higher diversity gain. These papers study the energy consumption by exploring the high transmit data rate and low transmission duration. That is, the circuit energy consumption is saved by the time saving of VBLAST. Yet, the data uploading time consumption itself has not been investigated. Besides, the analysis is applied only to single hop networks.

There are a few papers applying both mobile sink and vMIMO communication in WSNs. Zhao et al [14] firstly proposed a three-layer framework for mobile data gathering in concurrent data uploading based WSNs, including the sensor layer, cluster head layer, and mobile collector layer. The cluster heads were formed to emulate as multiple antenna sensor and upload sensing data to mobile sink concurrently. The objective was to achieve good scalability, long network lifetime and low data collection latency. Guo et al [15] proposed a data gathering cost minimization framework with concurrent data uploading, which is constrained by flow conservation, energy consumption, link capacity, compatibility among sensors and the bound on total sojourn time of the mobile collector at all anchor points. They evaluated the relations of energy consumption and sojourn time at anchor points, and also provided comparison of sojourn times under different data gathering cost.

Zhao [16] formalized the mobile data gathering problem with space-division multiple access (SDMA). Total data gathering time consists of sink moving time and data uploading time. Three algorithms were proposed to find a set of polling points to achieve the maximum number of compatible pairs among sensors while the moving tour containing all the selected polling points is the shortest. Besides, the problem of minimizing the maximum data gathering time among different regions was considered and a region-division and tour-planning algorithm was proposed for multiple sinks scenarios. However, the data gathering was only limited with single hop behaviour in the paper, which was inefficient for sink moving tour shortening. As mentioned, the sink moving time can be reduced by decreasing the total number of polling points, which can be effectively achieved by associating sensors via multihop communications.

III. SYSTEM MODEL AND PROBLEM FORMULATION

In some environment monitoring and military scenarios, the sensing data could be large enough so that the data uploading time may effectively affect the total data collection time [16].
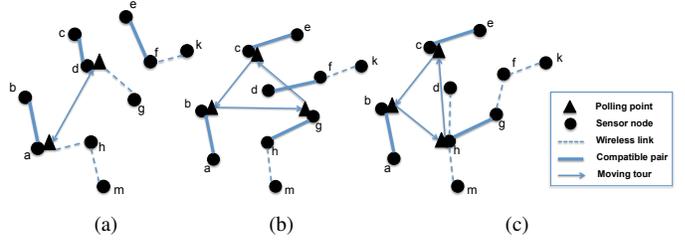

Figure 1. Three possible movement patterns for a mobile sink.

The most essential information comes from the time-sensitive data and the data collection delay could be of vital importance. Hence, the total data collection latency minimization problem needs to be addressed. To jointly consider concurrent data uploading based VBLAST and mobile sink, the potential trade-off between fully utilization of VMIMO and shortest moving tour is the key issue for delay minimization.

There are normally limited resources that can be utilized in such scenarios, and there is no chance for the independent set of locations or even a clear pro-knowledge about the area. Thus, there is a set of sensors randomly deployed in the filed and the polling points are selected from the sensors. The polling points can be either part of the compatible pairs or not. The sensors which are associated with the compatible sensors or non-compatible PPs are called association sensors. Once selected as a polling point, except for delivering its own sensing data, the sensor is also responsible for aggregating, buffering and transferring data from its associated sensors to the mobile sink. Therefore, there are different ways for sensing data to be collected by the sink:

(i). Compatible sensors upload their sensing data concurrently and directly to the mobile sink when it arrives.

(ii). Association sensors send their sensing data to the associated compatible sensors and polling points to buffer possibly via multiple hops. Upon the arrival of the mobile sink, the polling points and other compatible sensors upload their buffered data by vMIMO or SISO communications.

(iii). Non-association sensors upload their sensing data to mobile sink by one hop SISO when it arrives within the transmission range.

Figure 1 shows three possible association patterns with different compatible pairs and corresponding two moving tours of the sink. In Figure 1(a) two sensors (a and d) are selected as PPs and three compatible pairs are formed among the sensors (a, b); (c, d); (e, f). Sensor nodes h, g and k are associated with a, d and f respectively by one hop distance. Sensor node m is associated with a by two hops via h.

In Figure 1(b), there are also three compatible pairs formed during the sink moving tour, and three sensor nodes (b, c and g) are selected as PPs. In Figure 1(c), three compatible pairs are formed with three PPs being selected. Thus, for the three case, case (a) selects the minimum number of PPs and could get the shortest sink moving time. Case (b) forms the maximum number of compatible sensors with three PPs. This results in longer moving time, but it achieves more concurrent uploading benefit which leads to less data uploading time. Even though

TABLE I
FORMULATION NOTATIONS.

Indices:	
$S = \{S_i;\ (i = 1, ..., N)\}$	A set of sensors.
Constants:	
$R \geq 0$	The amount of sensing data for each sensor.
$D_{ij} \geq 0$	Distance between two sensors Si, Sj.
$H > 1$	Maximum hop boundary for multihop transmission.
$B > 0$	Buffer size of each sensor.
$V_m > 0$	Velocity of mobile sink.
$V_r > 0$	Effective data uploading rate.
Variables:	
$a_i = \{0, 1\}$	If sensor S_i is selected as a polling point, $a_i = 1$, otherwise, $a_i = 0$.
$k_i = \{0, 1\}$	If a polling point S_i is non-compatible, $k_i = 1$, otherwise (PP i is part of compatible pairs), $k_i = 0$.
$u_{ij} = \{0, 1\}$	If the sensors S_i and S_j are formed as a compatible pair, $u_{ij} = 1$, otherwise, $u_{ij} = 0$.
$e_{ij} = \{0, 1\}$	If the moving tour contains the segment between S_i and S_j, $e_{ij} = 1$, otherwise, $e_{ij} = 0$.
$x_{mih} = \{0, 1\}$	If the sensor S_m is associated with sensor S_i in h hop distance, $x_{mih} = 1$, otherwise, $x_{mih} = 0$.

case (c) forms three compatible pairs which is less than that in case (b), but it may achieve less overall data uploading time. This is attributed to that different amount of data for the two compatible sensors has to be uploaded in SISO way in case (b). While in case (c) sensor h and g buffer and upload the same amount of data from its associated sensors. All the sensing data in this case can be uploaded concurrently, hence, it achieves high utilization of vMIMO and small data total uploading latency.

In the phase of association design, with different locations of sensor nodes, the sensors could be responsible for uploading different amount of data to the mobile sink for other nodes due to the multihop communications. In this case, even though the sensors are formed as compatible nodes, the data uploading time may not be saved due to various amount of data from compatible sensors. Thus, it is challenging to evenly distribute the associated data to compatible sensors. Meanwhile the selection of compatible sensor pairs is closely related to the selection of PPs.

Therefore, to achieve the minimum total data collection delay does not necessarily mean to form the maximum number of compatible pairs or to establish the shortest moving tour, it should also consider the amount of concurrently uploaded data. Hence, how to jointly utilize vMIMO and organize the selection of polling points to achieve the minimum total data collection latency becomes our focus. We study the delay minimization for multihop data collection (DMMDC) problem. Our objective is to minimize the total data collection latency. Thus, the resulting optimal solution does not necessarily have the shortest moving tour length or the maximum number of compatible pairs, but instead, it is a trade-off between the moving tour and utilization of sensors compatibility.

Moreover, multihop transmission costs more energy for delivering sensing data to the sink. In order to limit the energy consumption for sensor nodes, the maximum number of hop distance is bounded. Due to the technique limitation of sensors, the buffer size is also bounded. The cost of sharing control information for vMIMO transmission in the

data gathering is ignored in this work. The reason is that the control packet is relatively short compared with the data packet and thus the energy consumption of additional data exchange will not greatly impact the energy consumption of vMIMO communication. The channel state information (CSI) is assumed to be perfectly known [17].

Given a set of sensors $S = 1, 2, ...N$ deployed over a sensing field, the DMMDC problem is to determine the selection of polling points, the compatible pairs, and the multihop associations between sensors to achieve the minimum data collection delay for sensors. Due to the limited resources of sensors, the buffer size of each sensor is bounded with B and the maximum distance for multihop transmission is bounded with H. The amount of sensing data for each sensor in one data collection cycle is R $(R < B)$. For a clear presentation, the notations used in the formulation are summarised in Table. I.

The total data collection latency minimization problem can be formulated as:

$$\min_{x_{mih}, u_{ij}, e_{ij}} \{N - \sum_{i=1}^{|S|} \sum_{j=1}^{|S|} [u_{ij}$$

$$+ \min(\sum_{m=1}^{|S|} \sum_{h=1}^{H} x_{mih}, \sum_{n=1}^{|S|} \sum_{h=1}^{H} x_{njh}) \cdot u_{ij}]/2\} \cdot R/V_r$$

$$+ \sum_{i=1}^{|S|} \sum_{j=1}^{|S|} D_{ij} \cdot e_{ij}/V_m \qquad (1)$$

subject to

$$\sum_{m=1}^{|S|} \sum_{h=1}^{H} x_{mih} \cdot R \leq B \qquad (i = 1, ..., |S|) \qquad (2)$$

$$\sum_{i=1}^{|S|} \sum_{h=1}^{H} x_{mih} + \sum_{j=1}^{|S|} u_{mj} + k_m = 1 \quad (i = 1, ..., |S|) \qquad (3)$$

$$x_{mih} \leq \sum_{j=1}^{|S|} u_{ij} + k_i \qquad (i = 1, ..., |S|, h = 1, ..., H)$$
$$\qquad (4)$$

$$\sum_{i=1}^{|S|} e_{ij} = a_i \qquad (i = 1, ..., |S|) \qquad (5)$$

$$\sum_{j=1}^{|S|} e_{ij} = a_j \qquad (i = 1, ..., |S|) \qquad (6)$$

Given the notation in Table. I, the DMMDC problem has been formulated as an integer linear program labelled from Eqn. (1) to Eqn. (6).

The objective function Eqn. (1) minimizes the total data collection latency which includes both data uploading time and sink moving time. The part of $\min(\sum_{m=1}^{|S|} \sum_{h=1}^{H} x_{mih}, \sum_{n=1}^{|S|} \sum_{h=1}^{H} x_{njh})$ specifies that the data collection time can only be saved by the concurrent uploading data from compatible sensor pairs.

Constraint (2) guarantees that the overall buffering data in any sensor does not exceed the sensor buffer limit. Constraint (3) - (4) guarantees that each sensor should be formed as part of the compatible pairs or be selected as non-compatible PP

or be associated with one of them, so that its sensing data can be collected during the moving tour. Constraints (5) - (6) guarantee that the mobile sink enters and departs each polling point only once.

The objective of the problem is to find a tour and the association relations between sensors, such that (i) all the sensors are either formed as a compatible sensor or associated with one, (ii) the total data collection time for each sensor is minimized, (iii) the total buffering data and the maximum hop distance are within sensors' constraints. DMMDC is NP-hard and it can be shown by a reduction from the well-known Travelling Salesman Problem (TSP). The total data collection time includes overall data uploading time which is affected by the total amount of concurrent-uploaded data and sink moving time which depends on the sink moving tour. In a special case where the network is super sparse. Assume that the sensing area is sufficiently large such that no two sensors are able to form compatible pairs and no sensor can be associated with other sensors. In this case, all the sensors have to be selected as polling points and the mobile sink visits all of them. Since the data uploading time for each sensor is the same with the amount sensing data, the overall data uploading time is proportional to the number of sensors. Thus, to achieve the minimum overall data collection time, the moving time for mobile sink should be minimized. Hence, the solution is to find the optimal shortest moving tour in order to visit all the sensors once, which forms a minimum distance TSP problem. Hence, DMMDC is NP-hard.

IV. MULTIHOP WEIGHTED REVENUE (MWR) ALGORITHM

In this section, we develop a heuristic multihop weighted revenue algorithm (MWR) to approximate the minimized data collection delay. The total data collection latency includes data uploading time and the moving time of the sink. Thus, the MWR is designed, on one hand, to minimize the moving tour delay of mobile sink, on the other hand, to maximize the amount of data that can be uploaded concurrently. The proposed algorithm runs at a central node and assists the underlying routing protocol.

There are different ways to maximize the utilization of vMIMO and maximize the amount of concurrent uploaded data. A direct way to increase the concurrent uploading data is to form as many compatible sensors as possible. However, this could lead to high number of polling points which may cause long sink moving delay. Another way is to increase the amount of data that the compatible sensors buffer. That is to say, to associate as many as sensors with the compatible pairs, and distribute the data as evenly as possible for two sensors in a compatible pair. In this case, the overall number of polling points can be limited. A good algorithm should jointly consider the two ways to form the association relations, so that the total delay can be minimized.

MWR selects PPs based on a combined metric of the number of compatible pairs, the number of uncovered neighbours and the sink moving time. Each sensor that is potentially selected as a PP is assigned a revenue metric. This *weighted revenue metric*, denoted by w, is formed of a weighted sum of normalized metrics. The weighted revenue metric $w(i)$ for a selected polling point i is designed in a way to capture the following three factors that affect a sensor's overall quality:

(i). The capability of sensor i to serve the compatibility among those non-associated sensors, denoted by $w_c(i)$. It is defined as the maximum number of possible uncovered compatible sensors divided by the number of uncovered sensors: $w_c(i) = \frac{N_c}{N_u}$.

(ii). The capability of sensor i to cover those non-associated neighbours, denoted by $w_n(i)$. It is defined as the ratio of the number of h-hop uncovered neighbours and the number of total uncovered sensors. An h-hop neighbour of a sensor is that the sensor can be reached by h-hop distance. The neighbour-covering revenue is $w_n(i) = \frac{\sum_{h=1}^{H} N_{nh}}{N_u}$, where H is the maximum hop boundary for multihop transmission.

(iii). The performance of sensor i to decrease the moving tour length, denoted by $w_d(i)$. It is defined as the minimum distance between the sensor i and the PPs divided by the maximum distance between any two sensors: $w_d(i) = 1 - \frac{d_{min}}{d_{max}}$.

Thus, the weighted metric is:

$$w(i) = w_c(i) \cdot \alpha + w_n(i) \cdot \beta + w_d(i) \cdot \gamma$$
$$= \frac{N_c}{N_u} \cdot \alpha + \frac{\sum_{h=1}^{H} N_{nh}}{N_u} \cdot \beta + (1 - \frac{d_{min}}{d_{max}}) \cdot \gamma \quad (7)$$

where $\alpha + \beta + \gamma = 1$, $0 < \alpha, \beta, \gamma < 1$ are the weighting factors.

The compatibility revenue $w_c(i)$ is the maximum possibility that the sensors in i's neighbour set can be formed into compatible pairs. This is a preferred metric to represent the utilization of vMIMO. The higher the vMIMO communication is utilized, the more concurrent data-uploading sensors are, which leads to more uploading time saving. The neighbour-covering revenue $w_n(i)$ is the capability of sensor i to cover the sensors in its h-hop neighbour set. This metric contributes to the objective in two ways: (1) For each sensor, higher number of h-hop neighbours potentially increase the possibility that the sensors are formed into high number of compatible pairs, hence decreasing total uploading time. (2) To select the sensor with large number of neighbours potentially leads to smaller number of polling points, hence decreasing the sink moving time. The distance-shoring revenue $w_d(i)$, which is to select the sensor that is close to the polling points, guarantees that the increase of the moving tour length is minimized and hence reduce the total length of the tour. The short moving tour leads to less sink moving time.

MWR chooses the sensor with the highest weighted metric value as the polling points, and associates the sensors evenly with the compatible sensors to maximize the amount of concurrent uploaded data. Algorithm 1 shows how the MWR algorithm works. It takes a set of sensors as input, and it outputs the selected polling points, the compatible sensor pairs and the association relations between sensors.

In lines 3 - 9, it checks each sensor for revenue values of the weighted metric (Eqn. (7)): maximum number of

---

Algorithm 1 Multihop based weighted revenue (MWR) algorithm.

Inputs:
    Set S contains N sensors;
Outputs:
    Set S', C, P are all subsets of S, contains the association sensors, the compatible sensors and the selected polling point sensors respectively;
    Association relations between all sensors;
RB Algorithms:
1: iter = 0;
2: **while** $(iter \leq N)$ **do**
3:    **for** each sensor $i \in (S \setminus C \setminus S' \setminus P)$ **do**
4:        Find the number of uncovered h hop neighbour sensors $N_{nh}(i)$;
5:        Find the minimum distance d_{min}, between sensor i and all sensors contained in P;
6:        Find the maximum number of compatible pairs $N_c(i)$ among uncovered sensors in $(S \setminus C \setminus S' \setminus P)$;
7:        Use $c(i)$ to record the sensors in those compatible pairs;
8:        Calculate the weighted revenue value $w(i)$ based on Eqn. (7);
9:    **end for**
10:    Find the sensor $i \in (S \setminus C \setminus S' \setminus P)$ that has maximum w(i);
11:    **if** $(S \setminus C \setminus S' \setminus P = \varnothing || N_c(i) = 0)$ **then**
12:        break;
13:    **else**
14:        Add sensor i into P, $P_c = i$; Add sensors in $c(i)$ compatible pairs into C; iter = iter + 1;
15:    **end if**
16:    Find the number of uncovered h hop neighbour sensors, denoted as $N_{nh}(j)$ $j \in (S \setminus C \setminus S' \setminus P)$;
17:    Find the sensor j with the maximum $\sum_{h=1}^{H} N_{nh}(j)$;
18:    **if** $(\sum_{h=1}^{H} N_{nh}(j) \neq 0)$ **then**
19:        Add the compatible pair j and P_c into C;
20:    **end if**
21:    **for** (each new-added pair of compatible sensors (i, j) in C) **do**
22:        **if** $(min(N_{n1}(i), N_{n1}(j)) \geq \frac{B}{R})$ **then**
23:            Associate $\frac{B}{R}$ 1-hop neighbours with each of i and j;
24:            Add the association sensors into S';
25:        **else**
26:            Associate number of $min(N_{n1}(i), N_{n1}(j))$ 1-hop neighbours with each of i and j;
27:            Add the association sensors into S';
28:            Check $\sum_{h=1}^{H} N_{nh}(i)$ and $\sum_{h=1}^{H} N_{nh}(j)$ with $h = 2, 3, ..., H$, associate at most overall $\frac{B}{R}$ neighbour sensors to each of i and j;
29:        **end if**
30:        Update the record of number of association sensors for each sensor i, denoted as $A(i)$;
31:    **end for**
32: **end while**
33: **while** $(S \setminus C \setminus S' \setminus P \neq \varnothing)$ **do**
34:    **if** (sensor m is in h-hop neighbour set of any PP $i \in C \cup P$ && $A(i) < \frac{B}{R}$) **then**
35:        Associate sensor m with i, update $A(i)$;
36:        Add sensor m into S';
37:    **else**
38:        Find the $i \in (S \setminus C \setminus S' \setminus P)$ that has maximum w(i);
39:        Add corresponding i into P;
40:        Associate all the sensors that in $(S \setminus C \setminus S' \setminus P)$ and in i's h-hop neighbour set with the i, update $A(i)$;
41:        Add the associated sensors into S';
42:    **end if**
43: **end while**

---

compatible sensor pairs, number of n-hop uncovered sensors, total number of uncovered sensors and the minimum and maximum distances between the sensor and all selected PPs.

In lines 16 - 20, all the uncovered sensors are check and the one with maximum number of h-hop neighbours is selected as the compatible sensor for the currently selected polling point. They are recorded as a compatible pair. If there is no compatible available for the selected PP, it is recorded as a non-compatible polling point.

In lines 21 - 31, for all the compatible pairs which are recently added to set C, the uncovered sensors are evenly associated with two sensors in each pair. In this stage, it is critical to keep the input amount of data at the same level for the two sensors in each compatible pair, so as to fully utilize vMIMO diversity gain and save uploading time. Moreover, the total association data to each compatible sensor can not exceed its buffer limit. In addition, to take into account the network energy consumption, the association sensors are chosen starting from 1-hop neighbours to H-hop neighbours.

Lines 33 - 43 guarantee that all the sensors are associated so that the sensing data can be collected by the mobile sink. The algorithm terminates when all the sensors are formed as compatible sensors, or selected as polling points, or associated with one of them.

Then the last step of MWR algorithm is to run an approximate algorithm for the TSP problem to find the shortest moving tour of the mobile sink to visit all the polling points. The time complexity of the proposed algorithm is dependent on how to find the maximum compatible pairs among the uncovered sensors. For a network with a total of N sensors, the worst case is when all the sensors are selected as polling points, which means the maximum information check will be $O(N^2)$ times. For each compatibility-check process, to find the approximate maximum compatible pairs among the uncovered sensors, the time complexity is $O(N^2)$. To find the uncovered h-hop neighbour set, the time complexity is $O(N)$. To find the distance between a sensor and the polling points, the time complexity if $O(N)$. To find the shortest tour on polling points, the time complexity is $O(N^2)$. Combing the information updating progress, the overall time complexity is $O(N^4 + N^3 + N^3 + N^2)$. Hence, the time complexity of the proposed MWR algorithm is $O(N^4)$.

V. PERFORMANCE EVALUATION

In this section, we evaluate the performance of the proposed MWR algorithm. MWR is firstly evaluated and compared with the optimal solution as formulated in Sec. III. Secondly, MWR is compared with two existing algorithms with simulations in different network scenarios. Thirdly, MWR is evaluated considering different settings of simulation parameters. Except for the optimal solutions, the simulations are performed using Matlab. Different kinds of network topologies have been evaluated with regard to the number of sensor nodes and the side length of the area.

A. Performance evaluation with optimal solution

To examine the performance of MWR, we first compare the results of the proposed MWR algorithm with the optimal solution results and a SISO data gathering scheme. For the SISO based algorithms, the overall data collection latency varies with the different MS moving times. Hence, the shortest moving tour (SMT) algorithm is chosen as the SISO benchmark. The optimal solution results (Optimal-MH)
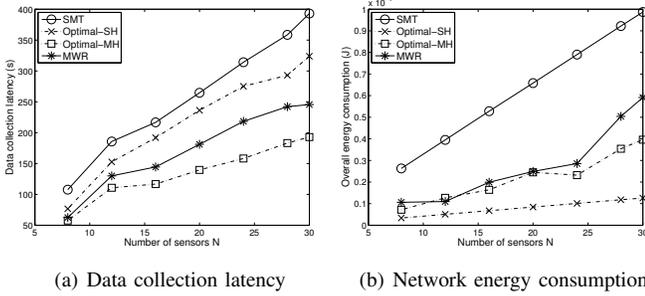
(a) Data collection latency (b) Network energy consumption

Figure 2. Performance comparison between optimum solutions and proposed MWR in small scale networks.

(a) Data collection latency (b) Network energy consumption

Figure 3. Performance evaluations with different number of sensors.

are obtained by solving the formulated problem in Sec. III by using the CPLEX [18]. However, due to the NP-hardness of the DMMDC problem, the results of optimal solution can only be achieved for a few small scale networks. As the third competitor, the single hop based data collection problem with vMIMO and mobile sink (Optimal-SH) is also formulated and solved with CPLEX. We consider a network with 8 to 30 sensors randomly deployed over an area of 100m × 100m. Any of the sensors can be selected and act as a PP with a limited buffer of $B = 5R$ [19], where R is the amount of sensing data for each sensor. The transmission range of sensors is set to be 30m. The weighting factors α, β and γ in MWR are set as 0.3, 0.3 and 0.4 respectively. The amount of sensing data for each sensor is $R = 1Mb$ and the effective data uploading rate is $V_r = 80Kbps$. The velocity of the mobile sink is $V_m = 0.8m/s$. In this set of simulations, the outputs are the minimum data collection latency and the overall network energy consumption. The results for performance evaluation are the average of 40 simulation experiments. Figure 2 demonstrates the comparison results: the data collection latency for different solutions follows the trend that Optimal-MH < MWR < Optimal-SH < SMT (Figure 4(a)), and the overall network energy consumption for different solutions follows another trend that Optimal-SH < Optimal-MH < MWR < SMT. Without achieving utilization and benefiting from vMIMO technique, there is neither time saving nor energy saving in SISO based algorithm SMT, it is reasonable to achieve the highest data collection time and network energy consumption for SMT. By allowing multihop transmission in the network, Optimal-MH enables more sensing data to upload to mobile sink via vMIMO transmission than Optimal-SH, thus, saving more data uploading time. Moreover, with more sensors being associated with the same PP, the number of PPs can decrease and this leads to less sink moving time. Summing up the two parts, the Optimal-MH scheme potentially saves more data collection time than that Optima-SH does, and achieves the lowest total data collection delay (Figure 4(a)). However, the multihop behaviour increases the energy consumptions for data transmission. Hence, the Optimal-SH achieves the best energy efficient (Figure 4(b)).

It is noticed that all the results are quite close when the number of sensor is less than 10, and the differences become larger as the increase of number of sensors in Figure 4(a). This is reasonable since when the number of sensors is small, the
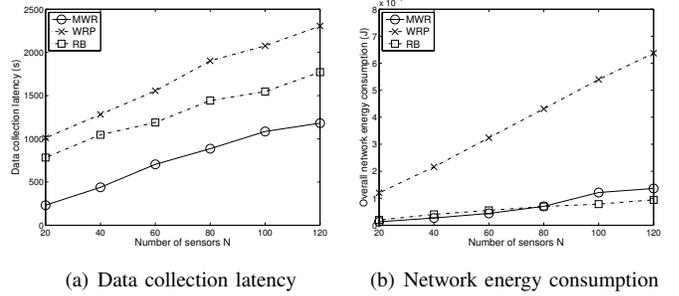
time saving from the utilization of vMIMO is limited, and the sink moving time dominants the overall data collection delay. The moving time is decided by the length of sink moving tour which can be similar for the four algorithms.

The proposed MWR scheme performs better than SMT and Optimal-SH and achieves very close performance to Optimal-MH with regard to the data collection delay. Besides, compared to SMT and Optimal-SH, MWR is overall more stable as the number of sensors increases. That is to say, the multihop behaviour helps to utilize the vMIMO especially when there is high number of sensors in the area (Figure 4(a)). Figure 4(b), on the other hand, shows that the multihop behaviour aggregates the energy consumption when comparing to Optimal-MH, MWR and Optimal-SH. With the increase of the number of sensors, the energy consumption increases intensively for MWR. Benefiting from the vMIMO, the energy consumptions for all Optimal-MH, MWR and Optimal-SH are much less than that of SMT (Figure 4(b)).

As mentioned, CSI is assumed to be known in the simulation. Since only a finite set of sensor locations the mobile sink can stop, it is feasible in practice to estimate the CST between each location and its neighbour sensors. This, however, introduces additional overhead of control information. The control message is neglected here because that a control message is relatively short compared with data packets. In practice, to obtain network topology information, a number of control messages will be introduced. The higher the signalling frequency is, the more control packets are. The transmission cost of such control messages should be taken into consideration when the number of control messages becomes significant.

B. Performance evaluation with other methods

In this part, we evaluate MWR by comparing its performance with other data collection algorithms. In order to show the benefits of both vMIMO and multihop behaviour, the algorithms for both vMIMO based single hop mobile data collection and SISO based multihop mobile data collection are chosen as the benchmarks:

- Revenue-Based (RB) algorithm [16], which is a vMIMO based single hop data collection algorithm, aims at minimum data gathering latency. By considering both compatible pairs and the MS moving tour in its weighted metric, RB utilizes vMIMO gains to an extend to save data uploading time.

- Weighted Rendezvous Planning (WRP) algorithm [20], which is a SISO based multihop data collection algorithm, aims to achieve the trade-off between data collection delay and energy consumption.

In this scenario, N sensors are randomly deployed over a $200m \times 200m$ area. Any sensor could be chosen as the polling point. The transmission range of sensors is set to be 30m. The weighting factors α, β and γ in MWR are set as 0.3, 0.3 and 0.4 respectively. We assume the amount of sensing data of each collection round is $R = 1Mb$, the effective data uploading rate is $V_r = 80Kbps$ and the data buffer size of each sensor is $B = 5R$. The velocity of the mobile sink is $V_m = 1m/s$. N varies from 20 to 120. To restrain the overall energy consumption, the maximum hop distance in multihop transmission scenarios is set as $H = 3$. In this set of simulations, the outputs are the data collection latency and the overall network energy consumption. The results for performance evaluation are the average of 40 simulation experiments.

Figure 3 shows the comparison results for the three algorithms. The results demonstrate stable performance trend of the data collection latency: MWR < RB < WRP (Figure 3(a)). With the increase of the number of sensors, the data collection delay increases stably for all three algorithms. It is clear that without any utilization of vMIMO, the WRP algorithm presents the highest data collection delay and highest network energy consumption. Benefiting from the multihop transmission behaviour, MWR achieves much lower data collection delay than that of RB. On average, compared to RB, MWR decreases data collection latency by 45 percentage.

In Figure 3(a), the delay tends to be stable with the increase of N for both MWR and RB. This is reasonable since when the network density reaches a certain level as the increase of N, the selected PPs are sufficiently enough to cover the increased sensors in the field and the increased sensing data can be more possibly uploaded concurrently. RB addresses the increased sensors by forming more compatible pairs for concurrent data-uploading. Except for forming more compatible pairs, MWR can also associate the increased sensors with the existing compatible sensors through multihop behaviour.

Both MWR and RB achieve dramatically lower energy consumption than that of WRP (Figure 3(b)). The energy consumptions for MWR and RB are quite close when the number of sensors is small, and MWR costs slightly higher energy than RB with the increase of N. With the increase of the number of sensors, MWR associates more sensors to perform multihop transmission to increase the amount of data that can be transmitted benefiting diversity gain. Hence, the network energy consumption of MWR becomes more aggressive with the higher number of sensors (Figure 3(b)).

Figure 4 shows simulation performance for the three algorithms with different side length of the area L considering the same number of sensors ($N = 60$) in the field. Thus, the sensor density is decreased with the increase of L. The data collection time is prolonged largely with the increase of L for all three algorithms. One important reason is that the MS moving length increases largely due to the longer distance

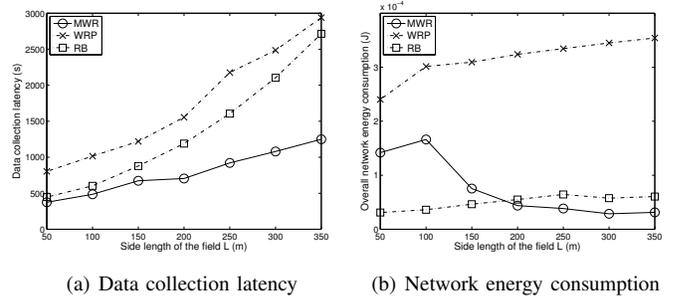(a) Data collection latency   (b) Network energy consumption

Figure 4. Performance evaluations with different side lengths of sensing area.

between sensors. MWR and WRP achieve the lowest and the largest data collection latency respectively (Figure 4(a)). MWR outperforms RB and this trend of superiority becomes even more remarkable as L increases. With the decrease of the density of network, to achieve more compatible pairs, RB has to deploy more PPs, which causes longer moving tour distance, hence longer moving time. Besides, for some far-away sensors, RB is more likely to select them as the non-compatible polling points. MWR, on the other hand, is able to associate those far-away sensors with the selected PPs or other compatible pairs via multihop behaviour.

MWR and RB lower the network energy consumption dramatically compared to WRP (Figure 4(b)). For MWR, the amount of network energy consumption rises when L is less than 100m and drops with the increase of L after that. This is attributed to that the number of multihop association sensors reduces with the network becoming sparser. In high density networks, MWR is more likely to associate the sensors and the associations become less and less with the decrease of network density. This is consist with the results in Figure 3(b) that the energy consumption increases as the network becomes dense (as the increase of N). The network energy consumption tends to be stable and slightly increase for RB. It is noticed that MWR consumes even less energy than RB does when L is larger than $200m$. The reason can be that with the wider network size, less sensors are able to form as the compatible pairs and benefit the energy efficiency from the concurrent data uploading for RB. Thus, less utilization of vMIMO leads to the increase of overall network energy consumption. To conclude from Figure 3 and Figure 4, the proposed MWR provides the best performance for time-sensitive applications with considerable energy consumption in all kinds of network densities. Algorithm RB works the best for energy-sensitive but relatively delay-tolerance applications when the network is sparse.

To evaluate how the MWR algorithm is affected by the application parameters, Figure 5 shows the performance of MWR with different parameter settings. MWR describes the results aforementioned in this section. With higher effective data uploading rate (MWR-1: $V_r = 160Kbps$), the total data collection delay largely decreases, and the decrease becomes more remarkable with the increase of N. In MWR-1, the data uploading time becomes sufficiently short and the sink moving time dominates the total data collection delay. Thus,
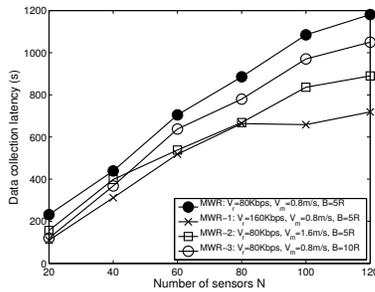
Figure 5. Performance comparison for MWR with different parameter settings.

the performance of MWR-1 is similar with the trend of moving tour length: The result tends to be stable with the increase of N. When the number of selected polling points reaches a certain level, most of the increased sensors in the field can be associated with the existing PPs, and thus the moving tour remains stable. The stable level is related to the network topology setting, such as the side length of the area L. As the sink moving velocity increases from $V_m = 0.8m/s$ (MWR) to $V_m = 1.6m/s$ (MWR-2), the total data collection delay generally decreases due to the reduction of sink moving time. The increase slope of the performance of MWR-2 is faster than that of MWR-1, due to the larger effect of the increase of N. Compared to MWR, MWR-3 rises the sensors' buffer size limit ($B = 10R$), so that more sensors are able to be associated with a same node which could lead to less number of polling points and more concurrent uploading of data. Thus, MWR-4 decreases the total data collection delay than MWR, while maintaining similar performance trend. Overall, all three parameters can be adjusted to meet different levels of network performance requirements. In sparse networks, the effective data uploading rate improves the network performance more critically than the other parameters. The sink moving velocity, on the contrary, shows major impact in dense networks.

VI. CONCLUSION

In this paper, we focus on minimizing the total data collection latency in multihop networks. We formulate the delay minimization problem for multihop data collection and propose a multihop weighted revenue (MWR) algorithm which jointly considers the amount of concurrent uploaded data and the moving tour length. It combines the number of compatible sensors, the number of h-hop neighbours, and the moving distance of sink in a weighted metric, to accurately account for these factors when ranking the available sensors for selection of polling points. Moreover, in order to achieve full utilization of concurrent uploading technique, MWR also emphasises on even associations of sensors to the compatible sensors.

Extensive simulation results show the effectiveness of the proposed algorithm. Compared to other algorithms, MWR effectively reduces the total data collection delay in different scenarios. Moreover, it requires less network energy consumption with least energy consumption in relatively sparse networks.

REFERENCES

[1] M. Li, Y. Liu, and L. Chen, "Nonthreshold-based event detection for 3D environment monitoring in sensor networks," *IEEE Trans. on Knowl. and Data Eng.*, vol. 20, no. 12, 2008.

[2] N. Aziz and K. Aziz, "Managing disaster with wireless sensor networks," in *Proc. 2011 13th International Conference on Advanced Communication Technology (ICACT'01)*, Feb 2011, pp. 202–207.

[3] E. Ben Hamida and G. Chelius, "Strategies for Data Dissemination to Mobile Sinks in Wireless Sensor Networks," *IEEE Wireless Communications*, vol. 15, no. 6, pp. 31–37, 2008.

[4] M. I. Khan, W. N. Gansterer, and G. Haring, "Static vs. mobile sink: The influence of basic parameters on energy efficiency in wireless sensor networks," *Comput. Commun.*, vol. 36, no. 9, pp. 965–978, May 2013.

[5] W. Liu, K. Lu, J. Wang, G. Xing, and L. Huang, "Performance analysis of wireless sensor networks with mobile sinks," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 6, pp. 2777–2788, July 2012.

[6] I. Krontiris, M. Langheinrich, and K. Shilton, "Trust and privacy in mobile experience sharing: future challenges and avenues for research," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 50–55, Aug 2014.

[7] M. Chen, S. Gonzalez, and V. Leung, "Applications and design issues for mobile agents in wireless sensor networks," *IEEE Wireless Communications*, vol. 14, no. 6, pp. 20–26, December 2007.

[8] S. Jayaweera, "V-BLAST-based virtual MIMO for distributed wireless sensor networks," *IEEE Transactions on Communications*, vol. 55, no. 10, pp. 1867–1872, Oct 2007.

[9] D. N. C. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1859–1874, 2004.

[10] M. Dohler, E. Lefranc, and H. Aghvami, "Virtual antenna arrays for future wireless mobile communication systems," *Publishing House of Electronics Industry*, vol. 2, pp. 501 – 505, 2002.

[11] S. Jayaweera, "An energy-efficient virtual MIMO architecture based on V-BLAST processing for distributed wireless sensor networks," in *Proc. 2004 First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON'04)*, Oct 2004, pp. 299–308.

[12] D. Nguyen and M. Krunz, "Cooperative MIMO in wireless networks: recent developments and challenges," *IEEE Network*, vol. 27, no. 4, pp. 48–54, July 2013.

[13] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theor.*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[14] M. Zhao and Y. Yang, "A framework for mobile data gathering with load balanced clustering and MIMO uploading," in *Proc. INFOCOM'11*, April 2011, pp. 2759–2767.

[15] S. Guo and Y. Yang, "A distributed optimal framework for mobile data gathering with concurrent data uploading in wireless sensor networks," in *Proceedings IEEE INFOCOM'12*, March 2012, pp. 1305–1313.

[16] M. Zhao, M. Ma, and Y. Yang, "Efficient data gathering with mobile collectors and space-division multiple access technique in wireless sensor networks," *IEEE Transactions on Computers*, vol. 60, no. 3, pp. 400–417, March 2011.

[17] M. E.-H. N. W. Lajos L. Hanzo, Osamah Alamri, *Near-Capacity Multi-Functional MIMO Systems: Sphere-Packing, Iterative Detection and Cooperation.* Wiley-IEEE Press, 2009.

[18] "Cplex package," http://www.ibm.com/software/cplex-optimizer, accessed: 08/2014.

[19] D. Malone, H. Qi, D. Botvich, and P. Patras, "802.11 Buffers: When Bigger Is Not Better?" in *Proceedings of International Workshop on Wireless Access Flexibility (WiFlex)*, 2013, pp. 37–48.

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

Ye Miao is a research engineer in State Key Laboratory of Space-Ground Integrated Information Technology, Beijing, China. She received her BSc in Electronic Information Science and Technology from China Agricultural University in 2010, her MSc in Mobile and Satellite Communication and PhD from University of Surrey, UK in 2011 and 2015, respectively. Her research interests are quality of service (QoS) and routing solutions in Mobile Ad-hoc networks (MANETs), integrated MANET-Satellite networks and Wireless Sensor Networks (WSNs).

Zhili Sun is a Professor with the 5G Innovation Centre, Institute of Communication Systems (ICS), University of Surrey. He has been with University of Surrey since 1993. He got his BSc degree in Mathematics from Nanjing University, China, in 1982, and PhD degree in Computer Science from Lancaster University, UK, in 1991. He worked as a postdoctoral research fellow with Queen Mary University of London from 1989 to 1993. He has been principle investigator and technical co-coordinator in many projects within the EU framework programs, ESA, EPSRC, and industries, and published over 200 papers in international journals, book chapters, and conferences. His research interests include wireless and sensor networks, satellite communications, mobile operating systems, traffic engineering, Internet protocols and architecture, QoS, multicast and security.

Ning Wang received his B.Eng. (honours) degree from Changchun University of Science and Technology, Changchun, China, in 1996, his M.Eng. degree from Nanyang University, Singapore, in 2000, and his Ph.D degree from University of Surrey, UK, in 2004, respectively. He is a Reader at the 5G Innovation Centre, Institute for Communication Systems (ICS), University of Surrey, UK. His research interests mainly include information centric networking (ICN), content and data caching management, network optimisation techniques and quality of service (QoS).

Yue Cao joined the Institute for Communication Systems (ICS) formerly known as Centre for Communication Systems Research (CCSR), at University of Surrey, Guildford, UK in 2009 and obtained his PhD degree in 2013. He is currently a Research Fellow at the ICS. His research interests focus on Delay/Disruption Tolerant Networks, Electric Vehicle (EV) communication, Information Centric Networking (ICN) and traffic offloading for cellular systems.

Haitham Cruickshank is a Senior Lecturer at the 5G Innovation Centre (5GIC), Institute for Communication Systems (ICS), formerly known as Center for Communication Systems Research (CCSR), at University of Surrey, Guildford, UK. He obtained his PhD degree from Cranfield University, Cranfield, UK, in 1989. His main research interests are network security, satellite network architectures, DTN security, and quality-of-service (QoS) provisioning.