A Large-scale Multi-track Mobile Data Collection Mechanism for Wireless Sensor Networks

Guoqiang Zheng [1], Lei Fu [1], Jishun Li [2], Ming Li [1]
[1] School of Electronic Information Engineering, Henan University of Science and Technology
Luoyang, 471003, China
[e-mail: lyzgq@mail.haust.edu.cn, fuleifox@gmail.com, limin0919@163.com]
[2] Henan Key Laboratory for Machinery Design and Transmission System, Henan University of Science and Technology, Luoyang, 471003, China
[e-mail: xnzhzhangliwei@126.com]
*Corresponding author:Lei Fu

Received June 18, 2013; revised August 26, 2013; revised November 4, 2014; accepted March 12, 2014; published March 31, 2014

Abstract

Recent researches reveal that great benefit can be achieved for data gathering in wireless sensor networks（WSNs）by employing mobile data collectors. In order to balance the energy consumption at sensor nodes and prolong the network lifetime, a multi-track large-scale mobile data collection mechanism (MTDCM) is proposed in this paper. MTDCM is composed of two phases: the Energy-balance Phase and the Data Collection Phase. In this mechanism, the energy-balance trajectories, the sleep-wakeup strategy and the data collection algorithm are determined. Theoretical analysis and performance simulations indicate that MTDCM is an energy efficient mechanism. It has prominent features on balancing the energy consumption and prolonging the network lifetime.

Keywords: Wireless sensor networks, multi-track, energy-balance, mobile sink, data collection

This research was supported by the National Hi-Tech Research Development Program (863 Program) (no.2009AA044902), the Major national science and technology projects (2012ZX04004011), Henan Province basic and frontier technology research project (122300413209), the Education Department of Henan province science and technology research project (12B510010), the Henan University of Science and Technology Scientific Research Start-up Foundation for Dr.(NO.09001550), the Henan University of Science and Technology Innovation Fund Project for Postgraduates (CXJJ-YJS-Z011).

http://dx.doi.org/10.3837/tiis.2014.03.008

1. Introduction

WSNs have a wide range of applications [1][2], such as target tracking, environmental monitoring, industrial and agricultural managing. Because of its great application value, WSNs have aroused great concern of governments and academic institutions. Constituted by the sensor nodes, WSNs have gradually developed into an important application platform. Therefore，the data is the center of these applications. Considering the limited node energy, an energy efficient data collection mechanism is an urgent problem to solve.

In the traditional data collection mechanisms, the sensor nodes are randomly deployed in the monitored region. For the restricted communication radius, the generated packets are usually sent to the static sink node or static base station [3][4][5] in the one hop or multi-hop mode. And then it will cause the network energy-unbalance problem for the excessive data relay. The nodes near the base station/sink quickly run out energy. It causes energy holes and leads to the death of the network. Different from the static sink data collection mechanism, some mobile sink data collection mechanisms have been widely researched in recent years. The main idea of these mechanisms is using mobile data collection device or mobile sink(MS) to collect data. The optimal energy-saving way is letting MS traverse each node [6], but the relatively low speed of MS will result in long time delay in the data collection process. These solutions tend to be restricted in some delay sensitive applications.

So the hybrid mechanisms of mobility and data relay have frequently been proposed in recent years. Sheu et al. proposed an IDGP [7] (Infrastructure based Data Gathering Protocol) mechanism with the hierarchical multi-hop data relay ideas creatively solved excessive time delay problem. But it still accompanies with the network energy consumption unbalance problem. Naveen et al. proposed a TGF [8] (Tunable Locally-Optimal Geographical Forwarding) mechanism. TGF uses a sleep-wakeup mechanism to reduce the network energy consumption, but requires the nodes location information or all the nodes themselves having the locating devices. No doubt, it increases the utility cost and reduces the practicability.

Inspired by the IDGP annular MS moving path, for the view of network energy-balance performance and practicability, this paper proposes a multi-track large-scale wireless sensor networks mobile data collection mechanism. One MS is used in MTDCM and it is made to move along the defined energy-balance path to collect data. And the network works in a controlled sleep-wakeup model. Simulation results show that MTDCM has great advantages on energy-balance performance and practicability.

The MTDCM mechanism proposed in this paper is applicable to various applications, especially in environmental monitoring, weather monitoring and other periodic non real-time monitoring applications. The MS concerned in MTDCM needs sufficient energy and the geographic information of the motioned region. The contributions of this paper can be summarized as follows：

1. The designed energy-balance multi-track data collecting path is the core of this paper. The data relay consumption of all sensor nodes can be observably balanced.
2. To further reduce the network energy consumption and prolong the network lifetime, an efficient sleep-wakeup strategy is designed in this paper through the MS dada collection path.

The remainder of this paper is given as follows. Section 2 describes the related work. Section 3 shows the system model and Section 4 discusses the details of MTDCM. Section 5 describes the simulation analysis. Finally Section 6 gives a conclusion to this paper.

2. Related Work

In this section, some recent mobile data collection mechanisms are briefly reviewed. Based on the controllable performance, the mobile data collection mechanisms are divided into two categories. They are the controllable mechanisms and the uncontrollable mechanisms. And the controllable mechanisms can be subdivided into three subcategories. The following paper will make a brief description of the review.

The first category is the uncontrollable mechanisms [9][10][11] that the mobile data collection devices randomly run in or outside the monitored region. Shashidhar et al. proposed a mobile data collection mechanism [9] based on the static sink mechanism. It uses some randomly moving nodes for data collection. Ekici et al. [10] improved the algorithm in [9] by applying a data collection path to optimize the network connectivity and the network lifetime. In [11], the raccoon location tracking work also used the mobile devices. The common characteristics are the high stability and reliability. And the maintenance of the system is relatively simple. But their shortcomings are the lack of flexibility, and they can not adapt to the distributed networks and the dynamic environment.

The second category is the controllable mechanisms [12][13][14] [15][16][17]. This type of mechanisms use mobile data collection devices which can freely move to any positions and the movement path can also be set based on the specific target. This broad category can be further subdivided into three subcategories.

In the first subcategory, the mobile data collection device is controlled to traverse through each node or along a straight line in the region. All of the nodes upload information in a single-hop way. Kansal et al. [12] proposed a network application in which MS traverses along a straight line with a lack of flexibility. Ma et al. [13] studied the route of the mobile device to reduce data loss which is caused by the buffer overflow. The mobile data collection device moves along a defined path to visit each node. However, the length of the path will be increased in large-scale WSNs and resulted in great data delay latency.

The second subcategory has planed the data collection path of the mobile data collection devices and the nodes upload data in a multi-hop way. Richard et al. [14] improved the algorithm in [12] that the MS moving path is optimized from straight-line to curve. All nodes upload data according to the broadcasted information of MS.

The last subcategory considers the data transmission model and path planning model. Lin et al. [15] raised a cluster-based data collection mechanism to solve the buffer overflow problems in which MS collects data from the clusters. Zhu et al. [16] proposed a data collection mechanism based on clustered path planning. A minimum set is used to optimize the number of relay hops and the energy consumption of data collection. Zhao et al. [17] improved the algorithm in [13]. Some nodes are selected as the root nodes and the data collection trees are built based on the root nodes. Through the periodic accesses of the root nodes MS can execute the data gathering work in a multi-hop way. Compared with the mechanism in [13], the structure proposed in [17] greatly reduces the length of the MS data collection path. But the excessive data relay of the root nodes still results in an energy unbalance problem. Furthermore, the sleep-wakeup strategy hasn't been employed to reduce the energy consumption. For the practical operation of the entire network, MTDCM is proposed. The goal is to balance node energy consumption and prolong the network lifetime.

3. System Model and Problem Formulation

A brief description of the problems related to MTDCM is provided and the issues concerned in this paper are formulated in this section.

3.1 System Model

The applications of MTDCM are broad. The data collection path of MS can be defined by MTDCM in most terrains. Pursuing with the simple description and expandability, a regular hexagon two-dimensional monitored area is chosen as an example. In the entire WSNs, N nodes are randomly and uniformly deployed in the area with a practical density d. The distance from the center to the vertex is R. For saving the cost of actual applications, the communication radiuses of all sensor nodes and MS (or other Mobile data collection devices) are r.

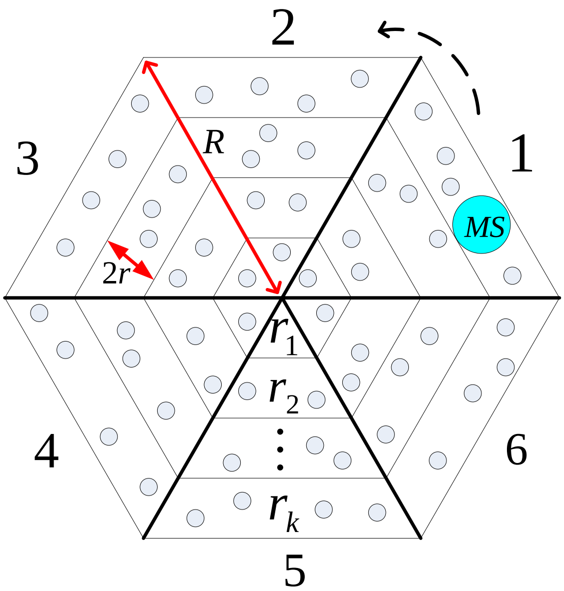


Fig. 1. MTDCM system model

As shown in **Fig. 1**, according to the node communication radius r, the application region is divided into k rings, the width of each ring is $2r$. S_i represents the nodes located in ring r_i and $|S_i|$ is the number of S_i. MS runs in the middle of each ring to collect data. Sensors in different rings transmit data to MS in a multi-hop way. According to a suitable angle $\theta = \pi / 3$, the monitored area is divided into $2\pi / \theta = 6$ sectors. So each sector has k layers. The angle θ can also be set as other values in different applications. The entire sensor nodes are deployed in different sectors and know their sector values. To focus the work on building the energy-balance path and the energy saving algorithm, the multi-hop routing algorithms [18][19] are used in the mechanism. Therefore, when MS completes the movement for one round in any ring, it can collect the information generated by all nodes.

3.2 Problem Formulation

The core of this mechanism is to build an energy-balance multi-track data collection path of MS. In addition, an effective sleep-wakeup strategy is designed to prolong the network lifetime. The data collection path is defined as $T_{MTDCM} = (x_k, x_{k-1}, \cdots x_1)$, x_i is the number of rounds that MS travels in ring r_i. So MS runs ring by ring with different rounds in the application area. According to this path, the energy consumption of all nodes can be balanced in the data collection period. To match most practical applications, MS moves inwardly from the boundary of the monitoring region.

Assuming that $T = (x_k, x_{k-1}, \cdots x_1)$ represent all the MS data collection paths, MS first runs in ring r_k for x_k rounds and then the inner ring r_{k-1} for x_{k-1} rounds. Finally, after x_1 rounds in ring r_1, MS completes the data collection path $T = (x_k, x_{k-1}, \cdots x_1)$.

The symbol e_i^{total} is the total energy consumption of any node s_i when MS completes the data collection path $T = (x_k, x_{k-1}, \cdots x_1)$. Assuming that S represents all the nodes in the monitoring region, $\forall s_a, s_b \in S$, if $e_a^{total} = e_b^{total}$, the path $T = (x_k, x_{k-1}, \cdots x_1)$ can be called an energy-balance path T_{EB}. In the regular hexagon two-dimensional monitored area, the length of the path in ring r_i is $l_i = 6(2ir - r)$. The total length of $T_{EB} = (x_k, x_{k-1}, \cdots x_1)$ is $L = \sum_{i=1}^{k} 6x_i\,(2ir - r)$. Then T_{MTDCM} is the shortest one of all the energy-balance paths. That is:

$$T_{MTDCM} = \min\left(\left|L_i\right|\right) = \min\left(\sum_{i=1}^{k} 6x_i\,(2ir - r)\right), \forall T_{MTDCM} \in T_{EB} \qquad (1)$$

In the application environment of MTDCM, all nodes work periodically. Assuming that the data collection period of the whole application region is t_0 in the actual applications, t is the actual time needed for MS in each ring. That is, MS runs with different speed in different rings. Each node performs the monitoring task in the time period t and produces a unit of packet with the size of q, so the total amount of data generated in each period is q.

The moving time for MS in the ring $r_{collect}=r_i$ is t. For the data uploading need of actual requirement, the time costs in $r_{collect}=r_i$ should be not more than t_0, so $t \le t_0$. $c_{i,ms}$ is the data transmission rate of sensor s_i in ring r_i. By Shannon theory: $c_{i,ms} = B \times \log_2(1 + SNR)$, where B is the channel bandwidth, SNR represents the signal-to-noise ratio in the Gaussian noise. Because each node produces a unit of packet with the size of q, then the total amount of data generated by all nodes is qN, where N is the total number of all sensor nodes. So the theoretical data transmission time is measured by the ratio of the total amount of messages and the total data transmission rate in ring r_i:

$$t_{\text{i,theo}} = \frac{N \times q}{\sum B \times \log_2(1 + SNR)}, \forall s_j \in S_i \qquad (2)$$

$t_{i,theo}$ is the theoretical data transmission time in ring r_i and there is k rings in the monitoring region. So at least k values are got and the maximum one is set as t_{theo}. In order to ensure that MS is able to receive the data generated by all nodes, t should be not less than t_{theo}. So the boundaries of the MS running time t are defined as:

$$\max\left(\frac{N\times q}{\sum B\times\log_2(1+SNR\ )}\right)\leq t\ \leq t_0,\forall s_j\in S_i,1\leq i\leq k \tag{3}$$

4. MTDCM Data Collection Mechanism

The details of the proposed MTDCM mechanism are described in this section. MTDCM contains two phases: the energy-balance phase and the data collection phase. In the energy-balance phase, MTDCM calculates the energy consumption of all nodes when MS completes a round in different rings. Then an energy-balance MS data collection path $T_{MTDCM}=(x_k,x_{k-1},\cdots x_1)$ is created by the results. In the data collection phase, a sleep-wakeup strategy is applied. And then MS will calculate and execute the data collection plans that include the speed of MS in each ring, the moving rounds in each ring, the starting point and the starting time. In the next paragraphs, the two phases will be discussed in details.

4.1 Energy-Balance Phase

The energy-balance problem can be defined as a linear programming problem. As shown before, e_i^{total} is the total energy consumption of s_i after MS completes the energy-balance data collection path T_{MTDCM}.

In order to meet the needs of the energy-balance performance and achieve the energy efficient purpose, this paper sets up an energy-balance function $f\left(e_N^{total},e_{N-1}^{total},\cdots,e_1^{total}\right)$ by the energy-balance theory [20], where N is the total number of the deployed sensors. The function can be normalized as follows:

$$f\left(e_N^{total},e_{N-1}^{total},\cdots,e_1^{total}\right)=\frac{\left(\sum_{i=1}^{N}e_i^{total}\right)^2}{N\times\sum_{i=1}^{N}\left(e_i^{total}\right)^2} \tag{4}$$

Assuming that $\forall s_i\in S$, the total energy consumption e_i^{total} have the same value. The energy-balance function will get the optimal value 1. The result shows that the greater the function value is, the better energy-balance performance will be. So the energy-balance goal of this paper is to maximize the value of the energy-balance function.

The energy-balance phase is aimed at balancing the energy consumption of all nodes. The nodes in $r_{collect}$ need to generate data and relay the data from the more remote rings. The energy consumption is larger than other nodes and results in an energy-unbalance problem. In order to balance the energy consumption, the following paragraphs will describe the formulation work and then build the energy-balance data collection path T_{MTDCM} of MS.

Assuming that S_m represents all the nodes in a random ring r_m, $|S_m|$ is the number of S_m. $\overline{S}_m$ shows the area of ring r_m. Each node generates q data in the time period t. Q_{mn} is the total amount of data transmitted by all nodes in ring r_m when MS has completed one round in ring r_n. So Q_{mn} can be calculated as follows:

$$|S_m| = \bar{S}_m \times d, \bar{S}_m \in \bar{S}, 1 \leq m \leq k \tag{5}$$

$$Q_{mn} = \sum |S_m| \times q \tag{6}$$

The symbol $\sum |S_m|$ in the formula above indicates the total number of the nodes in ring r_m and the nodes in the relayed rings. For example, assuming that the outmost ring is r_7, $r_{collect} = r_3$ and Q_{mn} is marked as Q_{53}. So the relayed rings ranges from r_6 to r_7, considering $r_m = r_5$ itself, $\sum |S_m|$ can be described as $\sum_{m=5}^{7} |S_m|$. Assuming that e_{mn} is the energy consumption of a random node s_m in ring r_m after MS finished the movement in ring r_n for one round, e_{unit} is the energy consumption of transmitting every unit of data. E_{mn} is the total energy consumption of all nodes in ring r_m. Then E_{mn} can be defined as follows:

$$E_{mn} = \sum |S_m| \times q \times e_{unit} \tag{7}$$

So the energy consumption e_{mn} can be further described as:

$$e_{mn} = \frac{E_{mn}}{|S_m|} = \frac{\sum |S_m| \times q \times e_{unit}}{|S_m|} \tag{8}$$

As shown before, e_m^{total} is the total energy consumption of s_m when MS completes the energy-balance data collection path T_{MTDCM} for one time. And because x_n is the moving rounds of MS in ring r_n, e_m^{total} can be calculated as follows:

$$e_m^{total} = \sum_{n=1}^{k} e_{mn} \times x_n \tag{9}$$

Assuming that t_{MTDCM} is the time consumption of MS when MS completes T_{MTDCM} for one time. While the time needed in each ring is always *t*, t_{MTDCM} can be described as follows:

$$t_{MTDCM} = x_k \times t + x_{k-1} \times t + \cdots + x_1 \times t = t \times \sum_{n=1}^{k} x_n \tag{10}$$

Because MS conducts the data collection work with the energy-balance path $T_{MTDCM} = (x_k, x_{k-1}, \cdots x_1)$, each node has a same energy consumption on the whole. To simplify the description, a random node s_o in the outmost ring r_k is selected as an example. So e_o^{total} is the total energy consumption of s_o after MS completes $T_{MTDCM} = (x_k, x_{k-1}, \cdots x_1)$ for one time and $e_{o.unit}^{MTDCM}$ is the energy consumption of s_o for per unit of time. From formulas (9) and (10), $e_{o.unit}^{MTDCM}$ can be calculated as follows:

$$e_{o.unit}^{MTDCM} = \frac{e_o^{total}}{t_{MTDCM}} = \frac{e_{kk} \times x_k + e_{k(k-1)} \times x_{k-1} + \cdots + e_{k1} \times x_1}{t \times \sum_{n=1}^{k} x_n} \tag{11}$$

And because the energy consumption of s_o can be obtained by dividing the total energy consumption of all nodes with the total number of the nodes, so $e_{o.unit}^{MTDCM}$ can be obtained as:

$$e_{o.unit}^{MTDCM} = \left(\frac{E_{kk} \times x_k + E_{k(k-1)} \times x_{k-1} + \cdots + E_{k1} \times x_1}{t \times \sum_{n=1}^{k} x_n} \right) \left(\frac{1}{|S_k|} \right) \tag{12}$$

In addition, the total energy consumption of the nodes in ring r_n can also be set by the total amount of data transmission and the energy consumption of per unit data transmission, from formula (6) and (12):

$$e_{o.unit}^{MTDCM} = \left(\frac{Q_{kk} \times x_k + Q_{k(k-1)} \times x_{k-1} + \cdots + Q_{k1} \times x_1}{t \times \sum_{n=1}^{k} x_n} \right) \left(\frac{e_{unit}}{|S_k|} \right) \tag{13}$$

And each node generates q data in each period, we can see:

$$e_{o.unit}^{MTDCM} = \left(\frac{x_n \sum_{n=1}^{k} |S_n| + x_{k-1} \sum_{n=k}^{k} |S_n| + \cdots + x_2 \sum_{n=k}^{k} |S_n| + x_1 \sum_{n=k}^{k} |S_n|}{t \times \sum_{n=1}^{k} x_n} \right) \left(\frac{q \times e_{unit}}{|S_k|} \right) \tag{14}$$

The formula (14) shows that the average energy consumption of $s_o \in S_k$ during per unit of time is affected by many factors. But the packet size, the energy consumption of transmitting every unit data and the deployment density are restricted by the hardware devices or the practical applications, so this paper chooses the MS moving rounds x_n to design the MS energy-balance path T_{MTDCM}.

As shown in formula (4), the energy-balance goal of this paper is to maximize the value of the energy-balance function $f\left(e_N^{total}, e_{N-1}^{total}, \cdots, e_1^{total}\right)$. In order to meet this target, MS need to move with the path T_{MTDCM}. And the energy consumption of each node must be identical when MS completes T_{MTDCM} for one time. Each ring a random sensor node is selected and the k nodes are marked from 1 to k. On account of the energy-balance data collection T_{MTDCM}, each one of the k nodes will has the same energy consumption after MS completes T_{MTDCM} for one time, that is $e_1^{total} = e_2^{total} \cdots = e_k^{total}$. From formula (8) and (14), the relationship can be derived as follows:

$$\begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1k} \\ e_{21} & e_{22} & \cdots & e_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ e_{k1} & e_{k2} & \cdots & e_{kk} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} e_1^{total} \\ e_2^{total} \\ \vdots \\ e_k^{total} \end{bmatrix} = \begin{bmatrix} e_1^{total} \\ e_1^{total} \\ \vdots \\ e_1^{total} \end{bmatrix} \tag{15}$$

Formula (16) describes the values of the variables x_i . In order to satisfy the formula (1) the optimal path is shown as:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1k} \\ e_{21} & e_{22} & \cdots & e_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ e_{k1} & e_{k2} & \cdots & e_{kk} \end{bmatrix}^{-1} \begin{bmatrix} e_1^{total} \\ e_1^{total} \\ \vdots \\ e_1^{total} \end{bmatrix} \tag{16}$$

In practical applications, e_1^{total} can be taken from the appropriate experimental values or unit 1. When the results are not integers, according to the formula (1), the nearest integer values are chosen as the moving rounds of MS. Finally the energy-balance data collection path $T_{MTDCM} = (x_k, x_{k-1}, \cdots x_1)$ of MS can be obtained that satisfies the practical applications.

4.2 Data Collection Phase

The implementation of MTDCM is discussed in this chapter. The previous sections focus on the energy-balance performance of the network. On the basis of the energy-balance data collection path $T_{MTDCM} = (x_k, x_{k-1}, \cdots x_1)$, this part adds a node sleep-wakeup mechanism which does not affect the real-time performance. The sleep-wakeup mechanism can deeply reduce the network energy consumption and prolong the network lifetime. The specific algorithm is shown in **Table 1.**

Table 1. Data collection algorithm

function_MS: MS sends the location information contains the sector values I_i *function_state*: contains the *active* state and the *sleeping* state *active* : sensors conduct the monitoring task, generate & transmit data *sleeping*:sensors can only receive the corresponding control message from the neighbor sensors or MS
1. for each sensor 2. initializes *function_state= sleeping*; 3. if the received control message I_i = the initialized sector value I_i' 4. sets *function_state= active*; 5. for *timer* $\le$ *t/6* 6. sends out the received control message for one time; 7. conducts the monitoring task, generates & transmits DATA ; 8. end for 9. sets *function_state= sleeping*; 10. end if 11. end for

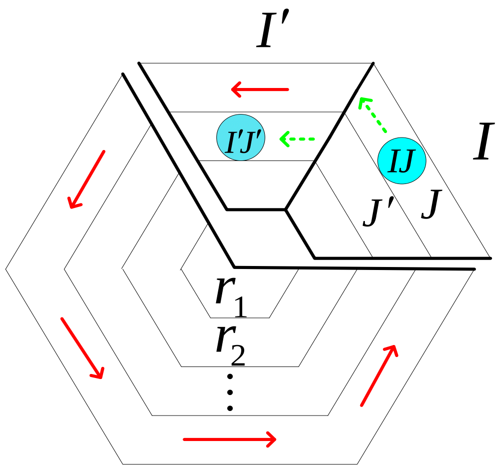


Fig. 3. Data collection schematic

As shown in **Fig. 3**, when MS has finished x_J rounds in ring *J*, it turns to the movement in the next ring J'. While T_{MTDCM} is completed for one time, MS then comes into the next cycling of T_{MTDCM} from the current position. At last it stops the data collection work when the death of the network arrives.

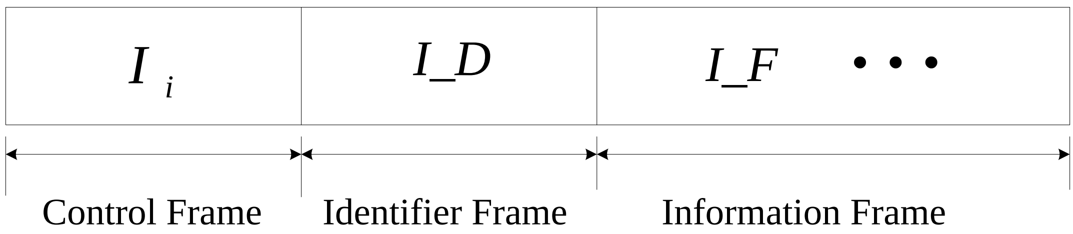


Fig. 4. Format of message in MTDCM

In the data collection phase, MS starts the data collection work from the boundary of the application area and sets a sector as the data collection unit. As shown in **Fig. 4**, the message format of MTDCM is composed of three parts. The control frame is the current location of MS, namely the current sector value I_i. The identifier frame is set to distinguish the control frame and the information frame. By formula (2), the time consumption of MS in each ring is much larger than that of the data transmission, so the activation time is relatively negligible. The time duration in each sector is here set as the data collection time.

As described before, all nodes are deployed with the knowledge of their sector value I_i', and I_i' is set as the control key. In the data collection process, MS broadcasts its current location in the movement. Namely, the wakeup signal is broadcasted: MS is now moving in sector I_i. Only the sensor in the communication radius of MS can receive the signal. As shown in **Table 1**, when the received location information I_i equals the memorized value I_i', the nodes turn into the *active* state. And at the same time they send out the received information for one time. If the other nodes find the received location information equals the memorized one, they repeat the work above. Because the running time of MS in each ring is *t* and MA is divided into 6 sectors, so the time duration of data collection in each sector is t/6. t/6 time later all nodes will come into the *sleeping* state and wait for the activation control message.

It is predictable that all nodes in sector I_i are activated in a short time. The nodes in other sectors are still staying in the *sleeping* state, for they haven't received the suitable control message. The *active* state is the normal working state and the nodes located in sector I_i upload the information to MS in a multi-hop manner. Moreover, MS can collect all the information after a round of movement in any ring and balance the energy consumption of nodes in this ring. Furthermore, when MS completes T_{MTDCM} for one time, the energy consumption of all nodes can be balanced.

5. Performance Evaluation

In this part, simulations are used to evaluate the performance of the proposed MTDCM mechanism. Inspired by the annular MS moving path of IDGP [7], we compare MTDCM with the existing static sink mechanisms classified as STATIC in [3][4][5] and the mobile sink mechanisms IDGP [7]. To further compare the energy saving performance of MTDCM mechanism, we add the MTDCM_NON mechanism which is the MTDCM mechanism without the sleep-wakeup strategy. Next part gives the simulation parameters.

Table 2. Simulation parameters

Network scale R	800m
The number of nodes	200-1200
Mobility model of MS	Proposed MTDCM Mechanism
Node initialization energy	1000J
Data upload period t	3-11h
Packet transmission energy consumption	7mJ/s
Packet reception energy consumption	3.5mJ/s
Sleeping state energy consumption	1mJ/s

Fig. 5 and **Fig. 6** show the comparison of the four mechanisms on the network lifetime performance. The network lifetime is defined by the time duration from the beginning of the network to the time the first monitoring hole appears. We here use the number of nodes and the data upload period as the variables.

As shown in **Fig. 5**, in the STATIC mechanism the network lifetime decreases dramatically with the increase of the number of nodes. Because the base station is located in the center position, the bottleneck effect will be exacerbated with the network scale and results in a significant lifetime decrease. The IDGP mechanism optimizes the data upload hops to reduce the data relay consumption. It relatively extends the lifetime of the network. Compared with MTDCM_NON, MTDCM hasn't exponentially increased the network life. The nodes in MTDCM_NON are designed to stop the data collection automatically when the buffer overflow comes. The data collection work is restarted when the buffer is wiped after uploading. The network lifetime of MTDCM and MTDCM_NON also slightly decreases with the increased relay work for the increase of the network scale. However, under the same conditions, the lifetime of MTDCM is 2.8 times longer than STATIC, 1.6 times longer than IDGP and 1.3 times longer than MTDCM_NON. MTDCM proves outstanding performance on the network lifetime.

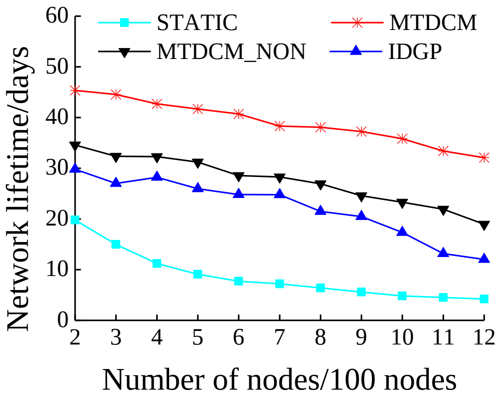


Fig. 5. Performance comparison on network lifetime with different number of nodes

As shown in **Fig. 6**, the network lifetime increases linearly with the increase of the data upload period. Because the data relay frequency is reduced by the increased data upload period. Thereby, the energy consumption of nodes is reduced and the network lifetime is prolonged. For the use of the energy-balance path, MTDCM and MTDCM_NON have a clear superiority on the network lifetime.

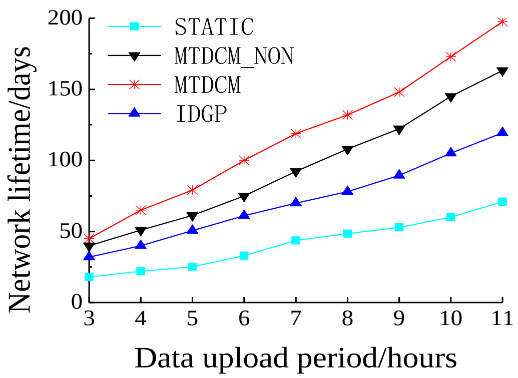


Fig. 6. Performance comparison on network lifetime with different data upload periods

Fig. 7 shows the stability of the network coverage ratio. The network coverage ratio can be got through the comparison between the current monitored area and the initial monitored area.

AS shown in **Fig. 7**, due to the relatively sufficient initial energy, the four data collection mechanisms show 100% coverage ratio in the first 25 days. 25 days later, the bottleneck effect appears and results in a deep decrease of the network coverage ratio. In the IDGP mechanism, MS conducts the data collection work on the planned path and all nodes upload the generated data in a multi-hop way. Because MS can communicate with more nodes than the STATIC mechanism, the coverage ratio decreases relatively slowly. Because of the energy-unbalance problem, some nodes in IDGP conduct little relay work, so the network indicates a low coverage ratio for some time. However, compared with MTDCM_NON which adopts the energy-balance path, it is obviously faster. And the performance of MTDCM is better than MTDCM_NON for adopting the sleep-wakeup strategy.

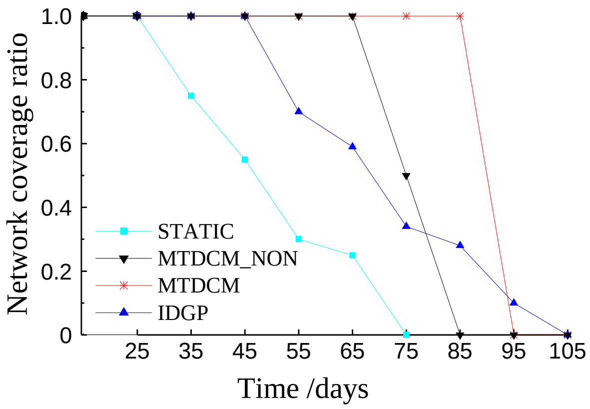
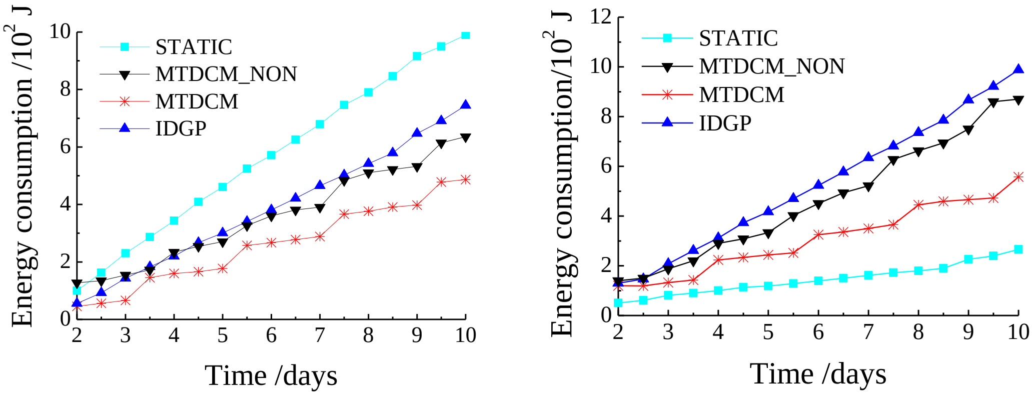


Fig. 7. Performance comparison on network coverage ratio.

Fig. 8 (a) and **(b)** indicate the total energy consumption of the nodes. We randomly select a node in the center ring and the other one in the boundary ring. The figures are drawn by the average energy consumption by 50 simulations.



(a) Center node (b) Boundary node

Fig. 8. Performance comparison on the total energy consumption of one node

As shown in **Fig. 8 (a)** and **(b)**, in the STATIC mechanism, the base station is often located in the center of the monitored area. The center nodes need to generate and relay a great deal of data, but the boundary ones only need to generate data. So the energy consumption of the center nodes is the highest of the four mechanisms. On the contrary, the boundary one is the lowest. In IDGP, the design of a reasonable data collection path and the limit of the dada transmission hops relatively reduce the energy consumption. However, compared with MTDCM_NON which adopts the energy-balance data collection path, there still exist a gap in the energy consumption of the center nodes and the boundary nodes. Furthermore, MTDCM which adopts the sleep-wakeup strategy greatly reduces the burden of the nodes and proves high superiority in reducing the total energy consumption.

As shown in **Fig. 9**, the energy-balance performance is compared. A center node and a boundary node are randomly selected. The difference of the average energy consumption is compared between the selected nodes .

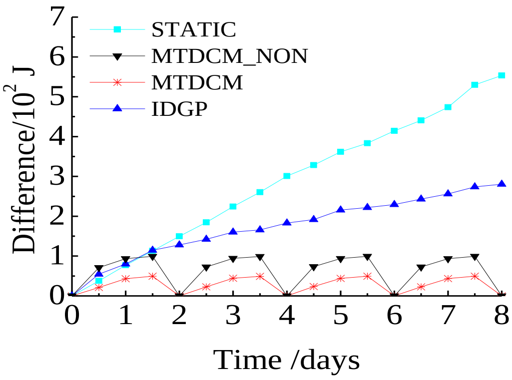


Fig. 9. Performance comparison on the difference of the average energy consumption

As shown in **Fig. 9**, because of the bottleneck effect, the difference between the center node and the boundary node increases with the operation of the network in the STATIC mechanism. In IDGP, although the path planning and the hops limiting algorithms are used to reduce the total energy consumption, the deference of the energy consumption between the center node and the boundary node still increases in a relatively lower speed. However, for the use of the energy-balance data collection path T_{MTDCM}, MTDCM_NON and MTDCM appear perfect on the energy-balance performance. In MTDCM, when MS finishes a round of movement in any ring, the energy consumption of the nodes in this ring is balanced. And when MS completes T_{MTDCM} for one time, the energy consumption of all nodes can be balanced. So the differentces of the energy consumption periodically comes to zero. Because the MTDCM_NON mechanism hasn't employed the sleep-wakeup strategy, the energy consumption is relatively high. Moreover, the proposed MTDCM mechanism obviously outperforms other data collection mechanisms on the comprehensive performance.

6 Conclusion

Based on the energy-balance and the practicability performance, a multi-track large-scale mobile data collection mechanism is proposed. MTDCM can determine the energy-balance data collection path of MS, the sleep-wakeup mechanism and the data collection strategy. The energy consumption of the entire network can be balanced and the network lifetime can be prolonged. MTDCM is an efficient and practical data collection mechanism and has outstanding features on the practicability and energy-balance performance.

References

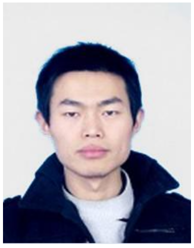
[1] H. Y. Jeong, B. H. Hong, "Monitoring system with wireless sensor network: a survey," *Applied Mechanics and Materials*, vol. 300, pp. 490-493, December, 2013. Article(CrossRef Link).

[2] H. C. Lee, Y. M. Fang and B. J. Lee, "The Tube: A Rapidly Deployable Wireless Sensor Platform for Supervising Pollution of Emergency Work," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 10, pp. 2776-2786, October, 2012. Article (CrossRef Link).

[3] A. Bereketli and O. B. Akan, "Event-to-sink directed clustering in wireless sensor networks," in *Proc. of IEEE Wireless Communications and Networking*, pp. 1-6, April 8-9, 2009. Article (CrossRef Link).

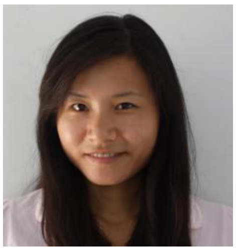
[4] J. Q. Guo, Z. Wei and S. L. Huang, "Data collection algorithm for clustered large-scale wireless sensor networks using virtual nodes," in *Proc. of IEEE International Instrumentation and Measurement Technology*, pp. 512-517, May 3-6, 2010. Article (CrossRef Link).
[5] C. T. Cheng, C. K. Tse and C. M. F. Lau, "A delay-aware data collection network structure for wireless sensor networks," *IEEE Sensors Journal*, vol.11, no.3, pp. 699-710, March, 2011. Article (CrossRef Link).
[6] S. E. Chu, D. W. Kang and J. N. Kim, "Routing and data dissemination over large-scale wireless sensor networks with multiple sources and mobile sinks," in *Proc. of 7th Int. Conf. on Advanced Language Processing and Web Information Technology*, pp.444-449, July 23-25, 2008. Article (CrossRef Link).
[7] J. P. Sheu, K. S. Prasan, C. H. Su and W. K. Hu, "Efficient path planning and data gathering protocols for the wireless sensor network," *Computer Communications*, vol.33, no.3, March, pp. 398–408, 2010. Article (CrossRef Link).
[8] K. P. Naveen and A. Kumar, "Tunable locally-optimal geographical forwarding in wireless sensor networks with sleep-wake cycling nodes," in *Proc. of IEEE INFOCOM 2010*, pp.1-9, March 14-19, 2010. Article (CrossRef Link).
[9] S. R. Gandham, M. Dawande, R. Prakash and S. Venkatesan, "Energy efficient schemes for wireless sensor networks with multiple mobile base stations," in *Proc. of IEEE Global Telecommunications*, pp. 377-381, December 1-5, 2003. Article (CrossRef Link).
[10] E. Ekici, Y. Y. Gu and D. Bozdag, "Mobility-based communication in wireless sensor networks," *IEEE Communications Magazine*, vol.44, no.7, pp.56-62, July, 2006. Article (CrossRef Link).
[11] T. Takeuchi, R. Matsuki and M. Nashimoto, "GPS cell phone tracking in the Greater Tokyo Area: A field test on raccoon dogs," *Urban Ecosystems*, vol.15, no.1, pp.181-193, January, 2012. Article (CrossRef Link).
[12] A. Kansal, A. A. Somasundara, D. D. Jea, M. B. Srivastava and D. Estrin, "Intelligent fluid infrastructure for embedded networks," in *Proc. of 2nd Int. Conf. on Mobile Systems, Applications and Services*, pp.111-124, June 6-9, 2004. Article (CrossRef Link).
[13] M. Ming and Y. Y. Yang, "SenCar: An energy-efficient data gathering mechanism for large-scale multihop sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol.18, no.10, October, 2007. Article (CrossRef Link).
[14] R. W. N. Pazzi and A. Boukerche, "Mobile data collector strategy for delay-sensitive applications over wireless sensor networks," *Computer Communications*, vol.31, no.5, pp.1028-1039, May, 2008. Article (CrossRef Link).
[15] P. L. Lin and R. S. Ko, "An efficient data-gathering scheme for heterogeneous sensor networks via mobile sinks," *International Journal of Distributed Sensor Networks*, vol.2012, 2012. Article (CrossRef Link).
[16] R. B. Zhu, Y. Y. Qin, J. Q. Wang, "Energy-aware distributed intelligent data gathering algorithm in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol.2011, 2011. Article (CrossRef Link).
[17] M. Zhao and Y. Y. Yang, "Bounded relay hop mobile data gathering in wireless sensor networks," *IEEE Transactions on Computers*, vol.61, no.2, pp.265-277, February, 2012. Article (CrossRef Link).
[18] A. Jarry, P. Leone, S. Nikoletseas and J. Rolim, "Optimal data gathering paths and energy-balance mechanisms in wireless networks," *Ad Hoc Networks*, vol.9, no.6, pp.1036-1048, June, 2011. Article (CrossRef Link).
[19] Y. Liu, X. Chen, C. Z. Long and N. R. Zhou, "Improvement of hierarchical routing algorithm for heterogeneous wireless sensor networks," *Journal of Computational Information Systems*, vol.8, no.10, pp.4143-4150, October, 2012. Article (CrossRef Link).
[20] R. Jain, D. Chiu and W. Hawe, *A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer Systems*, Eastern Research Laboratory, New York, 1984. Article (CrossRef Link).

Guoqiang Zheng received the Ph.D. degree in Communication and information systems professional from Xi'an Jiaotong University, China, 2008. He is a professor at Henan University of Science and Technology on College of Electronic Information Engineering, China. His research interests include wireless communication technology, network communication protocol and software radio theory.



Lei Fu received the B.S. degree in information engeering from the Henan University of Science and Technology, Luoyang, China in 2011. He is currently working towards M. S. degree on School of Electronic Information Engineering from Henan University of Science and Technology, Luoyang, China. His research interests include WSNs and wireless routing protocols.

Jishun Li received the Ph.D. degree in Mechanical manufacturing and automation from Shanghai Jiaotong University, China, 1996. He is a professor at Henan University of Science and Technology on school of Electromechanical Engineering, China. He is currently the chief of Henan Key Laboratory for Machinery Design and Transmission System, China. His research intrests include Mechanical and electrical integration, precision measurement theory, reverse engineering, signal acquisition and processing, numerical control technology and CAE.



Ming Li received the B.s. degree from Luoyang Institute of Science and Technology, Luoyang, China, in 2008. She is currently working toward the M.S degree with the Electronic Information Engineering, Henan University of Science and Technology, Luoyang, China. Her current research interests include synchronization techniques, cooperative communications for wireless sensor networks and wireless communication systems.