

Maximum Data Generation Rate Routing Protocol Based on Data Flow Controlling Technology for Rechargeable Wireless Sensor Networks

Demin Gao^{1,2,*}, Shuo Zhang¹, Fuquan Zhang¹, Xijian Fan¹ and Jinchi Zhang^{1,*}

Abstract: For rechargeable wireless sensor networks, limited energy storage capacity, dynamic energy supply, low and dynamic duty cycles cause that it is unpractical to maintain a fixed routing path for packets delivery permanently from a source to destination in a distributed scenario. Therefore, before data delivery, a sensor has to update its waking schedule continuously and share them to its neighbors, which lead to high energy expenditure for reestablishing path links frequently and low efficiency of energy utilization for collecting packets. In this work, we propose the maximum data generation rate routing protocol based on data flow controlling technology. For a sensor, it does not share its waking schedule to its neighbors and cache any waking schedules of other sensors. Hence, the energy consumption for time synchronization, location information and waking schedule shared will be reduced significantly. The saving energy can be used for improving data collection rate. Simulation shows our scheme is efficient to improve packets generation rate in rechargeable wireless sensor networks.

Keywords: Wireless sensor networks, maximum data generation rate, rechargeable-WSNs.

1 Introduction

As a significant portion of the smart pervasive Internet of Things (IOT), Wireless Sensor Networks (WSNs) have witnessed a significant promising technology, in which a larger number of nodes equipped with limited energy supply device send their own sensed information to a sink or data-processing center for collecting the data of a wide range area. Recently, WSNs have been used widely in a wide range of fields, e.g., military field [Amdouni, Adjih and Plesse (2015)], environmental field [Gao, Yin and Liu (2015)], health field [Hackmann, Guo, Yan et al. (2010)] and smart home field [Wang, Lin, Siahaan et al. (2014)]. However, researchers have not gained great breakthrough progress in battery tech-

¹ College of Information Science and Technology, Nanjing Forestry University, Nanjing, 210037, China.

² Department of Computer Science and Engineering, University of Minnesota, Minneapolis MN, 55455, USA.

* Corresponding Authors: Jinchi Zhang. Email: jinchi@163.com;
Demin Gao. Email: gdmnj@163.com.

nology, a fundamental problem is the limited lifetime of sensors owing to limited available energy [Wan, Yahya, Taib et al. (2014); Xu, Qian, Gu et al. (2011)]. When the battery of a sensor runs out, the sensor will be usefulness and discarded because that it is unrealistic generally to replace the batter. Thus, the problem of battery have restricted significantly the the wide applications of WSNs.

At present, for extending the lifetime of nodes, they are equipped with rechargeable technologies [He, Chen, Jiang et al. (2011)], such as, rechargeable batteries or super-capacitor(in the order of a million recharge cycles [Sudevalayam and Kulkarni (2011)]), which converts sources (e.g., body heat [Siddique, Wang, Madeo et al. (2014)], foot strike [Ko and Yegin (2013)], finger strokes [Kymissis, Kendall, Paradiso et al. (2002)] and solar [Amruta and Satish (2013)]) into electricity. Assuming energy neutral operation [Sudevalayam and Kulkarni (2011)], a sensor node¹ can operate perpetually when the energy expenditure rate is lower than the harvested energy rate. In such Rechargeable WSNs (R-WSNs) or Energy Harvesting WSNs (EH-WSNs), although their lifetime is less of an issue, the amount of energy harvested by a sensor is limited due to the size of generation elements and limited battery capacity.

In R-WSNs, for one thing, a node usually operates in a low duty cycle due to limited energy supplement and energy storage capacity [Gu and He (2010); Gao, Wu, Liu et al. (2014)], for another, the available energy tend to dramatically over time owing to complicated and volatile environment conditions [Gu, Zhu and He (2009); Khan, Qureshi and Iqbal (2015)]. Given these characteristics, nodes must regulate their activities and adapt their duty cycle to available energy for regulating energy consumption [Karthi, Rao and Pillai (2016)]. Since the duty cycles of nodes will be adjusted dynamically, it is unpractical to sustain a routing path for a long time. Therefore, new routing paths should be reestablished for each data transmission process. Extra energy will be consumed for new working schedules shared and routing paths reestablished, which lead to low efficiency of energy utilization. Hence, these unique characteristics of R-WSNs pose a high challenge for maximizing data collection rate.

Henceforth, in this paper, in order to address the problem of maximum data generation routing protocol, we introduce an algorithm of data flow controlling technology for maximizing data generation rate in R-WSNs. Specifically, since a routing path can not be sustained for a long time, we don't seek to establish an optimal routing path for data transmission based on working schedules shared. We let the data traffics from a node to flow to its neighbors freely for saving energy of rebuilding routing paths. Therefore, the saving energy can be used for strengthening data generation rate and improve the performance of R-WSNs.

Our contributions are summarized as follows:

- We propose a routing protocol of maximum data generation rate base on data flow controlling technology. To the best of our knowledge, there is the first in literature to study the problem based on data flow controlling for R-WSNs.

¹ In this paper, we will use “node” and “sensor” and “sensor node” interchangeable if no confusion.

- We propose a simple yet effective scheme, where we never seek to establish an optimal routing path for data delivery, but we let the data traffic from a sensor flows to its neighbors freely to save the energy of reestablishing routing paths.
- Our algorithm is suitable for a scenario where it is difficult to achieve time synchronization, location information and waking schedule shared. Especially, our algorithm provides better performance in mobile application of R-WSNs, where routing paths will be adjusted dynamically.

The rest of this paper is organized as follows. A number of existing maximum data generation solutions is presented in Section II. In Section III specify the system model, Section IV we present our method and design. In section V, we analyze the theoretical performance and provide experimental results. Conclusions are presented in Section VI.

2 Related work

Numerous strategies have been introduced to maximize data generation rate in WSNs, e.g., energy usage efficiently [Wang, Ju and Yu (2018); Shrivastava and Pokle (2014)], transmission power controlling scheme [Incel and Krishnamachari (2008)], data aggregation technology [Hakoura and Rabbat (2012)], and tree structure [Shi, Huang, Ren et al. (2013)], etc. In Hung et al. [Hung, Bensaou, Zhu et al. (2006a)], the authors introduce an energy-aware fair routing protocol with maximum data collection in WSNs, which is formulated as a concave utility maximization and is solved distributively by a sub-gradient algorithm [Hung, Bensaou, Zhu et al. (2006b)]. In Zhang et al. [Zhang, Jue, Sanglu et al. (2008)], the authors study and analysis Bellman-Ford routing algorithm for wireless sensor networks. In Padmanabh et al. [Padmanabh and Roy (2006)], multicommodity flow algorithm based on golden ratio is utilized for maximizing network lifetime, which optimizes the flow through each node. In these works, to achieve high data collection rate and obtain more packets from monitoring field, energy saving and energy usage efficiently are prior considered due to limited energy supplement. Therefore, it is inevitable to encounter critical tradeoffs between data flow and network lifetime.

Other closely related works have been done focusing on improving data generation rate in connection with the characteristics of R-WSNs, e.g., [Zeng, Zhang and Dong (2014); Roseveare and Natarajan (2014)]. In Liu et al. [Liu, Fan, Zheng et al. (2011)], a centralized algorithm and two distributed algorithms are provided for computing the maximum data collection rate. In Gao et al. [Gao, Lin, Liu et al. (2016)], a maximum data generation rate routing protocol in R-WSNs with multiple sinks has been provided to improve data collection rate by relieving the pressure of data collection with one sink in monitoring field. Ren et al. [Ren, Liang and Xu (2013)] formulate the data collection maximization with multi-rate transmission mechanism and transmission time slot scheduling among the sensors. Renner et al. [Renner, Unterschitz, Turau et al. (2014)] present a lightweight algorithm for online load adaptation of energy-harvesting sensor nodes. He et al. [He, Chin and Soh (2018)] formulate a Mixed Integer Linear Program (MILP) to determine the subset

of nodes that if upgraded will maximize the minimum source rate. Mehrabi et al. [Mehrabi and Kim (2016)] provide an optimization model for maximizing data collection throughput using a mobile sink. In these works, the optimal routing paths have been established to improve data generation rate by improving energy usage efficiency, or the optimal data generation rates have been calculated theoretically by establishing optimization model. The maximum data generation rate is obtained based on data routing path established before data transmission. Considering the topology changes dynamically, extra energy will be consumed for routing path rebuilt.

These earlier works for maximizing network generation rate can be divided into two parts: 1) focus on the static battery-powered network in traditional WSNs; 2) provide the details about the algorithm to establish optimal link paths for data transmission in R-WSNs. Compared to these earlier works, the main difference of our work is that we seek to maximize the data generation rate without establishing any routing paths before data delivery for saving energy, and improve data collection rate using asynchronous transmission model without clock synchronization. In our work, we introduce a maximum data generation routing protocol based on data flow controlling technology for R-WSNs, where data traffic flows to neighbors freely with fewer packets from previous sensor node for channel detection.

3 System model

3.1 Network model

For an undirected graph $G = (V, A)$ of rechargeable wireless sensor network, where V denotes n sensors and k sinks, $V = n \cup k$. A is the set of links, $A = \{A | (i, j) \in A, i, j \in V\}$. Each sensor $i \in S$ is powered by a rechargeable battery and its energy is harvested from its surrounding environment (e.g., solar power). Each sensor i senses its vicinity with sampling data generation rate G_i . We assume that the nodes have sufficient buffer space. The set of nodes are then connected to node i by links denoted as S_i . We assume that the network graph is connected, i.e., there always exists a path between any pair of nodes i and j in V . The current remaining energy of node i is E_i .

3.2 Energy consumption model

For a sensor, the power consumption generally contains five parts: sensing and generating data, idling state, listening channel, receiving packets, and transmitting packets. For all nodes, we assume that e_g denotes the power expenditure for generating one bit of data. The idle and channel listening power consumed per unit time, are assumed to be the equivalent for all nodes and independent of traffic, and are denoted by e_i and e_l , respectively. The first-order radio model for power consumption in receiving and transmitting is adopted in Suzuki [Suzuki (2012)]. Specifically, a node needs $\varepsilon_{elec} = 50nJ$ for running the circuitry and $\varepsilon_{amp} = 100pJ/bit/m^2$ for the transmitting amplifier. The energy expenditure for receiving one bit of packet is given by $e_r = \varepsilon_{elec}$. The power consumption for transmitting one bit of data to a neighbor node j is given by $e_t(i, j) = \varepsilon_{elec} + \varepsilon_{amp} * d_{i,j}^n$, where n is the path loss exponent, which typically ranges between 2 and 4 for free-space and short-

to-medium-range radio communication. We (x_i, y_i) and (x_j, y_j) denote the coordinate of node i and j , respectively. Hence, $d_{i,j}$ represents the Euclidean distance of node i and j and can be formulated as:

$$d_{i,j} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (1)$$

Hypothesis, $f_{i,j}$ denotes the data traffic from node i to node j in per unit time. The energy consumption in receiving and transmitting $f_{i,j}$ for node i are $e_{tr}(i, j)$ and $e_{re}(i, j)$, respectively. We have:

$$e_{tr}(i, j) = e_t(i, j) * \sum_{i,j \in V, j \in S_i} f_{i,j} \quad (2)$$

$$e_{re}(i, j) = \varepsilon_{elec} * \sum_{i,j \in V, j \in S_i} f_{i,j} \quad (3)$$

The value of $e_{tr}(i, j)$ is determined by the distance and the amount of data traffics of two nodes, while $e_{re}(i, j)$ is irrelevant to distance between nodes and is only affected by data traffic rate. Let w_i denotes the power consumption of node i in per unit time, which can be formulated as:

$$w_i = e_g * g_i + e_i + e_l + e_{tr}(i, j) + e_{re}(i, j) \quad (4)$$

Where, the values of e_i and e_l are usually invariant constants and are only affected by the physical properties heavily. Therefore, these factors can be regarded as a coefficient for energy consumption model.

4 Method and design

4.1 Description of the data flow in a distributed network

In the work, considering the pervasive nature and wide deployment in Internet of Things of R-WSNs, a routing protocol is proposed for focusing on a distributed network with multiple sources and multiple sinks deployed. Without loss of generality, the protocol design can be also used in different scenarios, e.g., a scenario with single source or single sink deployed, or mobile networks. In a distributed network with multiple sinks, these sinks form a tree structure firstly, and a sink is selected as root (called root sink) generally and other sinks (called leaves sinks) are as leaves of the root. All packets from leaves sinks will be transmitted to root sink through pseudo links. Considering a sink equipped with high ability of data processing and energy storage capacity, the overhead for data transmission between sinks can be neglected generally. Fig. 1 shows a sample of data transmission with multiple sinks deployed.

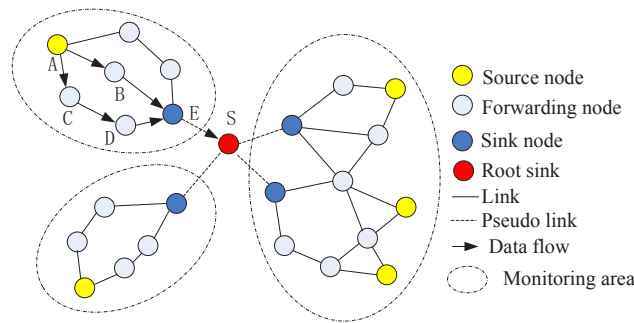


Figure 1: The routing paths were established with multiple sources and multiple sinks in distributed networks

4.2 Description of the flow controlling algorithm

We assume that the data generation rate is G_i for sensor i , which indicates that the sensor i will generate data traffic G_i per unit time in the network, where the unit time perhaps is one minute or ten minutes or one hour and so on. its value can be adjusted dynamically depending on application requirement. The duration of one unit time, it is enough for a sensor to forward packets at least one times. For each unit time, the packets sent to its neighbors are beyond the value $\varepsilon G_i/k$, where k denotes the number of paths established for data transmission from sensor i to its k neighbors, $\varepsilon > 0, \varepsilon \approx 0$. For illustrating the flow controlling algorithm, we assume the data traffic of sensor i and j are Q_i and Q_j , respectively, where, $Q_i > Q_j$. For sensor i , it will send $(Q_i - Q_j)/2$ data to sensor j and both of them have same amount of packets after data transmission. While, if $Q_j > Q_i$, sensor j will send $(Q_j - Q_i)/2$ data to neighbor i , as is shown in Fig. 2.

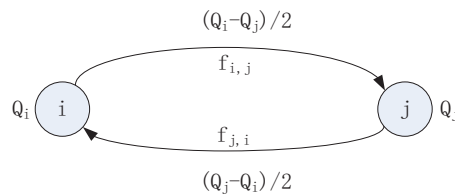


Figure 2: The routing path established based on data flow controlling technology

Theorem 1: There must be a time t , for a sensor i , the total of data traffic $\sum_{j=1}^k f_{i,j}$ from it to its neighbors meets the inequality:

$$\sum_{j=1}^k f_{i,j} \geq (1 + \varepsilon)G_i \tag{5}$$

Proof: We assume Eq. (5) is not proved, which indicates that the $\sum_{j=1}^k f_{i,j}$ is always

Algorithm 1: Flow controlling algorithm**Input:**Sensor i adds G_i traffic, $Q_i = Q_i + G_i$;Sensor j traffic Q_j ;**Procedure:**1: $\Delta Q_{i,j} = Q_i - Q_j$ 2: **if** $\Delta Q_{i,j} > 0$ *and* $|\Delta Q_{i,j}| > 2\varepsilon G_i/2$ **then**3: Sensor i sends $(Q_i - Q_j)/2$ data traffic to sensor j $Q_i = Q_i - (Q_i - Q_j)/2, Q_j = Q_j + (Q_i - Q_j)/2$ 4: **else if** $\Delta Q_{i,j} < 0$ *and* $|\Delta Q_{i,j}| > 2\varepsilon G_i/2$ **then**5: Sensor j sends $(Q_j - Q_i)/2$ data traffic to sensor i $Q_j = Q_j - (Q_j - Q_i)/2, Q_i = Q_i + (Q_j - Q_i)/2$ 6: **else**

7: the data transmission not happened

8: **End if**9: **End if****Output:****Return**{ Q_i, Q_j }

lower than $(1 + \varepsilon)G_i$ at anytime and we assume the current maximum $\sum_{j=1}^k f_{i,j}$ is Q_i^{max} , where k means that the data flow will be forwarded to next hops through k paths. For the next unit of time, sensor i will generate G_i data traffic and send total packets $f_{i,j}$ to its neighbors. So the current total packets of sensor i is $Q_i^{max} + G_i - \sum_{j=1}^k f_{i,j}$. Since the $\sum_{j=1}^k f_{i,j} < (1 + \varepsilon)G_i, \varepsilon \approx 0$, we assume sensor i sends the total data traffic G_i to its neighbor j , and sensor j obtains the same amount of data traffic with sensor i after data transmission from sensor i to sensor j .

For next unit time, sensor i also generates G_i data, and sends $G_i/2$ to its neighbor. Now, the total data traffic for sensor i is $Q_i^{max} + G_i/2$, obviously, the Q_i^{max} is not the maximum data traffic for sensor i , which is conflicted with our previous assumption. Since the data generation rate of sensor i is G_i , for ensuring the data flow decreases, the data transmission rate from sensor i to its neighbors should keep in the rate G_i all the time, which is unpractical obviously. Therefore, Eq. (5) is proved. The process of flow controlling algorithm is shown in Algorithm 1.

4.3 The flow controlling algorithm based on potential function

The packets generated by a source sensor will be forwarded to a sink through one or multiple routing paths with one-hop or multiple-hops data transmission. In the process of data traffic flowing to its destination, depending on the flow controlling algorithm proposed in the previous section, the amount of data traffic decreases due to some packets trapped in the forwarding nodes, which is shown in Fig. 2. For analyzing that the network throughput is affected by the data flow controlling algorithm, a potential function for sensors is introduced in the section.

Definition 1: The data flow potential function is set to $\phi(Q)$, which is differentiable convex function, where, Q denotes the amount of data traffics.

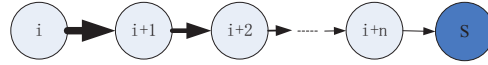


Figure 3: The data flow from a source to a sink with multiple-hops

When $Q = 0$, it means there is no packets in sensors. Generally, a sink always collects packets rather than sends these packets to other sensors. Hence, the amount of data for sinks can be set to 0, whose potential function is $\phi(0)$. The data flow potential function $\phi(Q)$ reflects the distribution of network data traffics. We assume $\phi(x) = e^{\alpha x}$, $\alpha > 0$. A source sensor generates packets and its potential function value increases, while a sink collects packets and its potential function value decreases.

To achieve maximum data generation rate, we calculate the maximum network throughput firstly and analyze the maximum sum of potential function values for all sensors in the networks. To illustrate the idea of flow controlling algorithm based on potential function, we analyze the total of data flow potential function belonging to a routing path from a source to a sink. For a source sensor i , its current data traffic is Q_i , its data generation rate is G_i . Depending on the theorem 1, for sensor i , when data generation rate is lower than data traffic sent to neighbors, its data traffic reaches a maximum. In other words, when the output rate begins to exceed the input rate, the data traffic reaches its maximum. We assume maximum data traffic of sensor i is $Q_i + G_i$, the potential function increase which can be given by:

$$\phi(Q_i + G_i) - \phi(Q_i) \quad (6)$$

Therefore, from Ineqt-1 of APPENDIX A, increase in potential function value can be given by:

$$\begin{aligned} & \phi(Q_i + G_i) - \phi(Q_i) \\ & \leq \phi(Q_i) + G_i \phi'(Q_i + G_i) - \phi(Q_i) \\ & = G_i \phi'(Q_i + G_i) \end{aligned} \quad (7)$$

Therefore, according to Eq. (7), when a sensor's data generation rate is set to be G_i , its maximum potential function increase reaching $G_i \phi'(Q_i + G_i)$. The packets generated by a source sensor reaches destination finally by one-hop or multiple-hops data transmission. In the process of forwarding data, potential function value decreases gradually from a source to a sink. We assume the data flow is $f_{i,j}$ from node i to j . For two ends of an edge (i, j) , as shown in Fig. 2, from Ineqt-2 of APPENDIX A, whose potential function increase which can be given by:

$$\begin{aligned} & \phi(Q_i) - \phi(Q_i - f_{i,j}) + \phi(Q_j) - \phi(Q_j + f_{i,j}) \\ & \geq f_{i,j} \phi'(Q_i) - 2f_{i,j}^2 \phi''(Q_i + f_{i,j}) - f_{i,j} \phi'(Q_j) \end{aligned} \quad (8)$$

Fig. 3 shows the data traffic flowing to sink from a source sensor. We assume there are $L - 1$ sensors and L edges in the routing path from source node to sink. In the process of data transmission from source to destination, some potential function values are canceled out. We assume the maximum amount of data flow in source sensor i is Q_i^{max} and its maximum potential function value is $\phi(Q_i^{max})$. After these data reaching a sink through k paths, $k \geq 1$, the potential function value in sink is 0 and its potential function value is $\phi(0)$. Therefore, in the process of the maximum amount of data flow generated by source sensor forwarded to a sink, the maximum potential function increases for all edges lying in the routing paths can be given as:

$$\sum_{j=1}^k f_{i,j} \phi'(Q_i^{max}) - f_{i,j} \phi'(0) - 2L f_{i,j}^2 \phi''(Q_i^{max} + f_{i,j}) \quad (9)$$

Hypothesis, $\phi(x) = e^{\alpha x}$, $\alpha = \frac{\varepsilon}{LG}$, $f_{i,j} = \frac{\varepsilon G_i}{k}$, where, $k \geq 1$, $L \geq 1$. Hence, Eq. (9) can be calculated as:

$$\sum_{j=1}^k f_{i,j} (\phi'(Q_i^{max}) - \alpha - 2 \times \alpha \times L \times f_{i,j} \phi'(Q_i^{max} + f_{i,j})) \quad (10)$$

where,

$$\begin{aligned} & 2 \times \alpha \times L \times f_{i,j} \phi'(Q_i^{max} + f_{i,j}) \\ &= 2 \times \alpha \times L \times f_{i,j} e^{(Q_i^{max} + f_{i,j})} \\ &\leq 2 \times L \times \frac{\varepsilon G_i}{k} \times \frac{\varepsilon}{LG_i} \times e^{\frac{\varepsilon G_i}{k}} \times \alpha \times e^{Q_i^{max}} \\ &\leq 2 \times \varepsilon^2 \times \phi'(Q_i^{max}) \end{aligned} \quad (11)$$

From Eq. (10) and Eq. (11), the maximum potential function increases of data flow for all edges lying in the routing paths from source to destination can be expressed as:

$$\sum_{j=1}^k f_{i,j} (\phi'(Q_i^{max})(1 - 2\varepsilon^2) - \alpha) \quad (12)$$

4.4 The maximum data generation rate for source

From the Eq. (6), we know $Q_i^{max} = Q_i + G_i$. If we consider that there are m source sensors in networks for data collection, the total potential function increases of the networks for all packets transmitted from all source to sinks which can be given as:

$$\sum_{i=1}^{i=m} \sum_{j=1}^k f_{i,j} (\phi'(Q_i + G_i)(1 - 2\varepsilon^2) - \alpha) \quad (13)$$

From Eq. (5), we know the maximum data flow between sensor i and j is $f_{i,j} = (1 + \varepsilon)G_i$. Hence, Eq. (13) can be rewritten as:

$$\begin{aligned}
 & \sum_{i=1}^{i=m} (1 + \varepsilon)G_i(1 - 2\varepsilon^2)\phi'(Q_i^{max}) - \alpha \sum_{i=1}^{i=m} (1 + \varepsilon)G_i \\
 &= (1 + \varepsilon - 2\varepsilon^2 - 2\varepsilon^3) \sum_{i=1}^{i=m} G_i\phi'(Q_i^{max}) - m\varepsilon L(1 + \varepsilon) \\
 &= \sum_{i=1}^{i=m} G_i\phi'(Q_i^{max}) + \sum_{i=1}^{i=m} (\varepsilon - 2\varepsilon^2 - 2\varepsilon^3)G_i\phi'(Q_i^{max}) \\
 & \quad - m\varepsilon L(1 + \varepsilon)
 \end{aligned} \tag{14}$$

From Eq. (14) and Eq. (7), $\sum_{i=1}^{i=m} G_i\phi'(Q_i^{max})$ indicates the sum of potential function increases for all source nodes. Therefore, $\sum_{i=1}^{i=m} (\varepsilon - 2\varepsilon^2 - 2\varepsilon^3)G_i\phi'(Q_i^{max}) - m\varepsilon L(1 + \varepsilon) \leq 0$, which ensures that the total of potential function of all sensors in network declines, hence:

$$\sum_{i=1}^{i=m} (\varepsilon - 2\varepsilon^2 - 2\varepsilon^3)G_i\phi'(Q_i^{max}) \leq m\varepsilon L(1 + \varepsilon) \tag{15}$$

Eq. (15) can be rewritten as:

$$\sum_{i=1}^{i=m} G_i\phi'(Q_i^{max}) \leq \frac{m(1 + \varepsilon)}{L(1 - 2\varepsilon - 2\varepsilon^2)} \tag{16}$$

From Eq. (7) and Eq. (16), we can observe that the maximum potential function increases is $G_i\phi'(Q_i^{max})$ for sensor i , whose maximum value is $\frac{(1+\varepsilon)}{L(1-2\varepsilon-2\varepsilon^2)}$. Therefore, in the process of data transmission from source to sink, if the data transmission rate meets $f_{i,j} = \frac{\varepsilon G_i}{k}$, based on Eq. (12), for ensuring the total potential function of network decreases, the inequality can be expressed as:

$$\varepsilon G_i(\phi'(Q_i^{max})(1 - 2\varepsilon^2) - \alpha) \geq \frac{(1 + \varepsilon)}{L(1 - 2\varepsilon - 2\varepsilon^2)} \tag{17}$$

where, $\alpha = \frac{\varepsilon}{L \cdot G} \approx 0, \varepsilon \approx 0$. Hence, Eq. (17) can be rewritten as $e^{\alpha Q_i^{max}} \geq 1/\varepsilon^2$. The maximum data traffic for sensor i can be given as:

$$Q_i^{max} \geq \frac{1}{\alpha} \ln \frac{1}{\varepsilon^2} = \frac{L \cdot G_i}{\varepsilon} \ln \frac{1}{\varepsilon^2} \tag{18}$$

In the work, if the data generation rate for a source sensor i is G_i and data transmission rate meets $f_{i,j} = \frac{\varepsilon G_i}{k}$, when the amount of data traffic for sensor i beyond that of Eq. (18), the total of potential function lying in the routing path decreases. Therefore, the maximum data traffic for sensor i is $\frac{L \cdot G_i}{\varepsilon} \ln \frac{1}{\varepsilon^2}$ and whose potential function decreases, which indicates that if data traffic of sensor i is lower than this maximum value, the data flow of whole sensors in network will keep in stable state. Therefore, we should ensure the maximum amount data traffic of each sensor will below the maximum value for all sensors in the network.

For $Q_i^{max} = \frac{L \cdot G_i}{\varepsilon} \ln \frac{1}{\varepsilon^2}$, what we need to declare is that the value is only the maximum data traffic to sensor i for total potential function lying in routing path decreasing, rather than the maximum amount data traffic to sensor i can obtain and is trapped in the sensor. In other words, a sensor maybe have more than $\frac{L \cdot G_i}{\varepsilon} \ln \frac{1}{\varepsilon^2}$ packets, which means a large number packets are trapped in the sensor. Especially, when a source sensor can not find its neighbors and no adjacent nodes receive its packets from sender, more and more packets will gathered in the sensor until its data buff is full. Obviously, more packets accumulated in sensors will cause high data transmission delay. Therefore, it is the purpose for an optimal routing protocol proposed to transmit data traffic to sink as soon as possible.

4.5 Maximum data generation rate for source sensor

According to Eq. (18), since the maximum data traffic for a source is $\frac{L \cdot G_i}{\varepsilon} \ln \frac{1}{\varepsilon^2}$, when its adjacent sensors' data is 0, the maximum data transmission between two sensors is $f_{i,j} = Q_i^{max} / 2$. According to Eq. (18), we have:

$$e_i + e_l + e_g G_i + (e_t(i, j) + \rho) \frac{L \cdot G_i}{2\varepsilon} \ln \frac{1}{\varepsilon^2} \leq E_i \tag{19}$$

From Eq. (19), for the data generation rate of sensor i , we have:

$$G_i \leq \frac{E_i - e_i - e_l}{e_g + (e_t(i, j) + \rho) \frac{L}{2\varepsilon} \ln \frac{1}{\varepsilon^2}} \tag{20}$$

where, we assume $\varepsilon = \Theta(y)$, there is a certain constant numbers $c1, c2, G_0$, for all sensors $G \leq G_0, 0 \leq c1 \cdot y \leq \varepsilon \leq c2 \cdot y$, when Eq. (20) is proved, the ε can be given as:

$$\varepsilon = \Theta\left(\frac{E_i - e_i - e_l}{e_g + (e_t(i, j) + \rho) \frac{L}{G_i} \ln \frac{1}{G_i}}\right)^{\frac{1}{2}} \tag{21}$$

According to Eq. (20), we can observe the maximum data generation rate for a source sensor is $\frac{E_i - e_i - e_l}{e_g + (e_t(i, j) + \rho) \frac{L}{2\varepsilon} \ln \frac{1}{\varepsilon^2}}$. Therefore, the maximum generation rate is affected by energy replenishment, energy expenditure and routing path length significantly. Energy replenishment and energy expenditure is influence of sensor distributed and hardware communication device significantly. If we reduce the routing path length L , the data generation rate will be improved, which means packets will be forwarded to sink with less hops. Now, a

sample is provided for illustrating the maximum data generation rate based on the data flow controlling technology, as is shown in Fig. 4.

After packets generated by a source sensor, these packets are always transmitted to destination with the shortest routing path based on data flow controlling technology. If there are fewer sensors lying in the routing path and packets will be forwarded to destination quickly generally. At the same time, for a sensor near a sink, its potential function value declines quickly and vice versa. If a routing path with more sensors and long distance from source to sink, it indicates that packets will be transmitted to sink with more time and more packets will be trapped in the path. The data flow controlling scheme is similar to water flowing downwards, which is the core idea behind our algorithm.

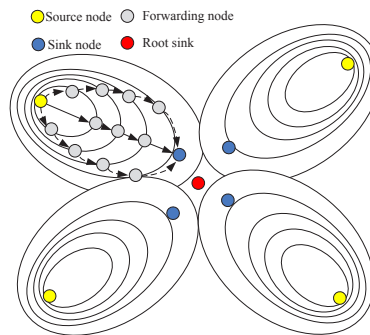


Figure 4: The gradient data flow from four sources to five sinks, which is similar to water flows downwards

5 Simulations

5.1 Experiment setup

In this section, we evaluate the performance in data generation rate of the proposed algorithm. Each sensor has a $20\text{ cm}^2 * 20\text{ cm}^2$ solar photovoltaic panel, and a rechargeable battery with $100J$ capacity. Matlab software is used for simulating our algorithm for R-WSNs, where 200 – 400 sensors and 1 – 10 sinks are randomly deployed in a $1000\text{ m} * 1000\text{ m}$ square area. Each node can communicate with adjacent nodes within 100 m .

5.2 System implement

When multiple sinks were deployed in the monitoring field, the performance for all algorithms is evaluated. Due to limited available energy, it is common to deploy multiple sinks for collecting information in realistic applications, as is shown in Fig. 5, which brings at least two advantages comparing to one sink. Firstly, in multiply sinks environment, since packets generated by sensors are only needed to be forwarded to anyone sinks, which is closest to it generally, the degree of routing path from source node to sink will be shorted due to higher sinks density. Secondly, considering sensors are distributed in scenario randomly, scarce of sensors are within the transmission range of sinks, which indicates that

these sensors have to undertake the tasks of data collection and data forwarding. It often represents a bottleneck when sensors around sinks run out their energy. Therefore, multiple sinks deployed are adopted widely in WSNs or R-WSNs for improving the performance in data collection and saving energy.

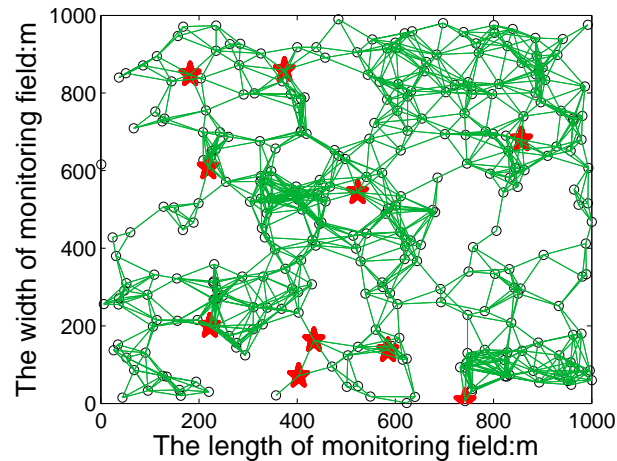


Figure 5: Connectivity graph with 300 nodes and 10 sinks

For further to express the communication procedure of our strategy in the large scale networks, we execute our algorithm with 300 sensors and 10 sinks randomly deployed in a $1000\text{ m} \times 1000\text{ m}$ square area. In the experiment scenario, sources (e.g., common sensors) sense ambient and generate information firstly. These sources will coach data or send these data directly to nearby sink with one or multiple-hops forwarding until these data reaching the destination. Fig.5 shows a connection graph with 300 sensor and 10 sinks deployed in monitoring field. In our algorithm, packets will be forwarded to a closer sink through all potential candidate intermediate nodes with high energy and low accumulation packets. Therefore, the energy expenditure for data forwarding will balance to all sensors deployed.

5.3 System performance comparison

For providing a insight of the performance of our algorithm (Maximizing Data Generation Rate, simplified as MDGR) under network settings, in this section, we provide an algorithm for performance comparison, which is provided in Mehrabi et al. [Mehrabi and Kim (2016)] and is one of the recognized widely forwarding protocols known to the community in low-duty-cycle WSNs. In Mehrabi et al. [Mehrabi and Kim (2016)], the authors propose an optimization model for Maximizing Data Collection Throughput (simplified as MDC-T), where the data collection problem is formulated as an optimization model. For a fair comparison, we introduce an improved version of the algorithm, where node's duty-cycle is determined periodically and locally. At the same time, multiple sinks are deployed for balancing the energy expenditure.

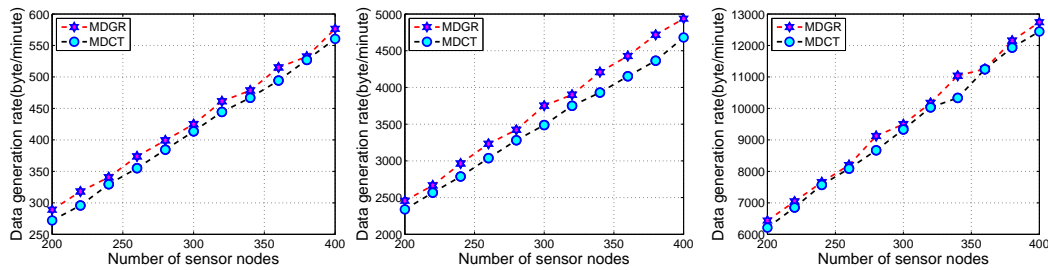


Figure 6: The data generation rate when 2 sinks are deployed and the duty cycle are set to 1%, 10%, 30%, respectively

We compare the data generation rate between our algorithm and MDCT scheme firstly under different number of sensors, where the average node's duty-cycles and the number of sinks are set to 1%, 10%, 30% and 2, 4, 6, 8, respectively, as shown in Fig.6 - Fig.9. From these Figures, we can observe that the data collection rate increases for both algorithms and our algorithm presents better performance comparing to that of MDCT scheme with the sensor density improved for all different node's duty-cycles and sinks. In our algorithm, a sensor adopts dynamic duty cycle mechanism. At the same time, it does not share its waking schedule to its neighbors and cache any waking schedules of other sensors. Therefore, scarce energy will be consumed for time synchronization and saving energy can be utilized for strengthening data collection. From Fig. 6, Fig. 7, Fig. 8, and Fig. 9. We can know that the data collection rates for our scheme are about 5%, 4%, 4% and 3% higher than that of the MDCT algorithm.

Theoretically, when more sensors are appended to experiment scenario, more packets will be generated and collected by sinks, which indicates that data generation rate is improved and energy expenditure will be balanced to all deployed sensors. it is a compulsive approach to improve data collection rate for a high sensor density network with low duty-cycles, where data transmission collision is scarce. Nevertheless, in real practice, when a large number of sensors are deployed in a smaller field relatively, data collision will be critical and retransmission is unavoidable, which will cause serious energy waste and high data transmission latency. In our work, an ideal condition with no collision and reasonable number sensor is considered for analyzing the data generation rate. In fact, the real data collection rate will be lower than that of our achieved.

6 Conclusion

In this work, for achieving maximum data generation rate in R-WSNs, we propose a data flow controlling technology to maximize the data collection rate. We do not seek to establish an optimal routing path for data delivery, but we let the data traffic from a sensor flows to its neighbors freely to save the energy utilized for routing path built. At the same time, a sensor does not share its waking schedule to its neighbors and cache any waking

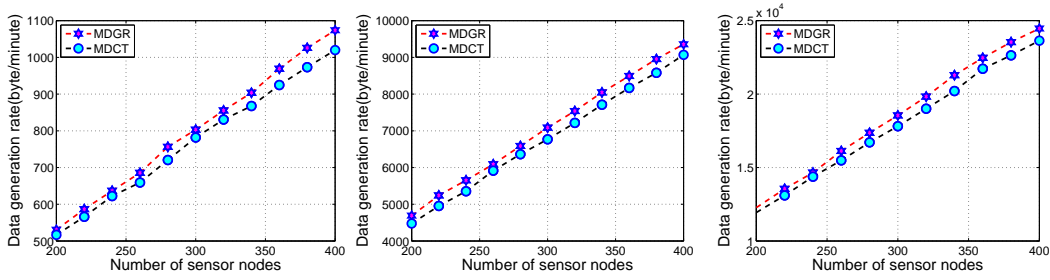


Figure 7: The data generation rate when 4 sinks are deployed and the duty cycle are set to 1%, 10%, 30%, respectively

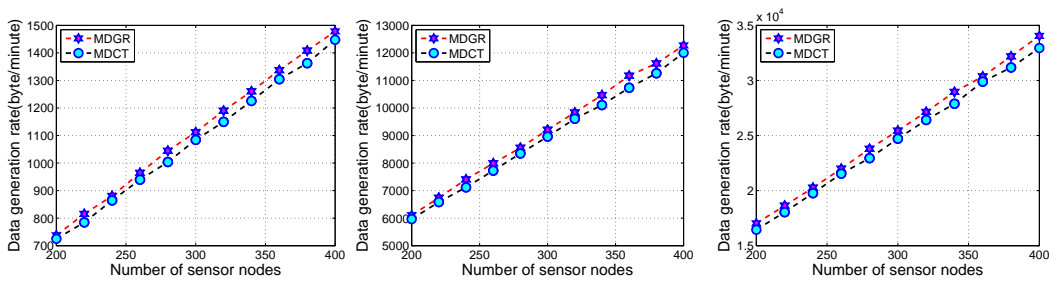


Figure 8: The data generation rate when 6 sinks are deployed and the duty cycle are set to 1%, 10%, 30%, respectively

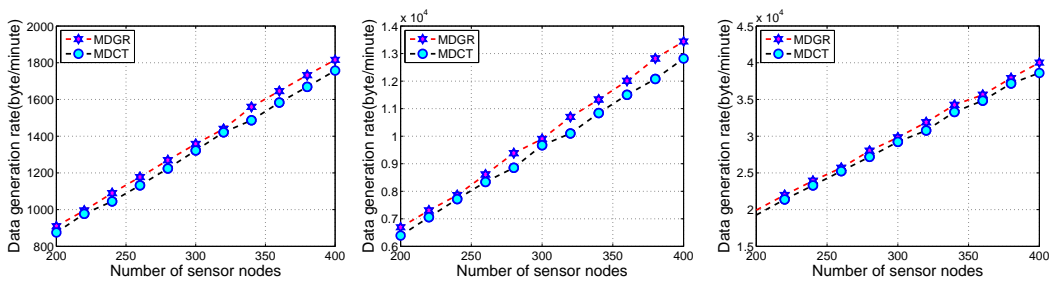


Figure 9: The data generation rate when 8 sinks are deployed and the duty cycle are set to 1%, 10%, 30%, respectively

schedules of other sensors. Therefore, the saving energy can be used for strengthening data generation rate and improve the performance of R-WSNs. Our algorithm is suited for a scenario that it is difficult to achieve time synchronization, location information and waking schedule shared, dynamic routing path with mobile application.

We define the network system and energy expenditure model firstly. Hereafter, the energy replenished and routing schemes are analyzed. Finally, the data flow controlling technology for maximum data generation rate is illustrated by an example in which a potential function is introduced for computing the maximum potential function and maximum amount of data flow in source sensor. Simulation and experiments show our algorithm provide a better performance for maximizing data collection rate in R-WSNs. There is one point needs attention that the E2E delay perhaps is serious in our scheme because of data recirculation.

Acknowledgement: This work was supported by The National Natural Science Fund of China (Grant No. 31670554) and The Natural Science Foundation of Jiangsu Province of China (Grant No. BK20161527). We also received three Projects Funded by The Project funded by China Postdoctoral Science Foundation (Grant Nos. 2018T110505, 2017M611828) and The Priority Academic Program Development(PAPD) of Jiangsu Higher Education Institutions. The authors wish to express their appreciation to the reviewers for their helpful suggestions which greatly improved the presentation of this paper.

References

- Amdouni, I.; Adjih, C.; Plesse, T.** (2015): Network coding in military wireless ad hoc and sensor networks: experimentation with gardinet. *International Conference on Military Communications and Information Systems*, pp. 1-9.
- Amruta, M. K.; Satish, M. T.** (2013): Solar powered water quality monitoring system using wireless sensor network. *International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing*, pp. 281-285.
- Gao, D.; Lin, H.; Liu, Y.; Guoxin, W. U.** (2016): Maximum data collection rate in rechargeable wireless sensor networks with multiple sinks. *China Communications*, vol. 13, no. 2, pp. 95-108.
- Gao, D.; Wu, G.; Liu, Y.; Zhang, F.** (2014): Bounded end-to-end delay with transmission power control techniques for rechargeable wireless sensor networks. *AEUE-International Journal of Electronics and Communications*, vol. 68, no. 5, pp. 395-405.
- Gao, D. M.; Yin, X. F.; Liu, Y. F.** (2015): Prediction of forest fire using wireless sensor network. *Journal of Tropical Forest Science*, vol. 27, no. 3, pp. 342-350.
- Gu, Y.; He, T.** (2010): Bounding communication delay in energy harvesting sensor networks. *IEEE International Conference on Distributed Computing Systems*, pp. 837-847.
- Gu, Y.; Zhu, T.; He, T.** (2009): Esc: energy synchronized communication in sustainable sensor networks. *IEEE International Conference on Network Protocols*, pp. 52-62.

- Hackmann, G.; Guo, W.; Yan, G.; Lu, C.; Dyke, S.** (2010): Cyber-physical codesign of distributed structural health monitoring with wireless sensor networks. *ICCPs'10*, pp. 119-128.
- Hakoura, B.; Rabbat, M. G.** (2012): Data aggregation in wireless sensor networks: a comparison of collection tree protocols and gossip algorithms. *25th IEEE Canadian Conference on Electrical & Computer Engineering*, pp. 1-4.
- He, S.; Chen, J.; Jiang, F.; Yau, D. K. Y.; Xing, G.** (2011): Energy provisioning in wireless rechargeable sensor networks. *IEEE INFOCOM*, pp. 2006-2014.
- He, T.; Chin, K. W.; Soh, S.** (2018): On maximizing min flow rates in rechargeable wireless sensor networks. *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 2962-2972.
- Hung, K. L.; Bensaou, B.; Zhu, J.; Nait-Abdesselam, F.** (2006): Energy-aware fair routing in wireless sensor networks with maximum data collection. *IEEE International Conference on Communications*, vol. 8, pp. 3438-3443.
- Hung, K. L.; Bensaou, B.; Zhu, J.; Nait-Abdesselam, F.** (2006): Maximum data collection least-cost routing in energy constrained wireless sensor networks. *IEEE Conference on Local Computer Networks*, pp. 177-184.
- Incel, O. Z. D.; Krishnamachari, B.** (2008): Enhancing the data collection rate of tree-based aggregation in wireless sensor networks. *IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, pp. 569-577.
- Karthi, J. S.; Rao, S. V.; Pillai, S. S.** (2016): Duty cycle adapted mac for wireless sensor networks with energy harvesting. *International Conference on Control Communication & Computing India*, pp. 685-690.
- Khan, J. A.; Qureshi, H. K.; Iqbal, A.** (2015): Energy management in wireless sensor networks: A survey. *Computers & Electrical Engineering*, vol. 41, no. C, pp. 159-176.
- Ko, G.; Yegin, K.** (2013): Footstep and vehicle detection using seismic sensors in wireless sensor network: Field tests. *International Journal of Distributed Sensor Networks*, vol. 9, no. 10, pp. 1380-1383.
- Kymissis, J.; Kendall, C.; Paradiso, J.; Gershenfeld, N.** (2002): Parasitic power harvesting in shoes. *Papers Second International Symposium on Wearable Computers*, pp. 132-139.
- Liu, R. S.; Fan, K. W.; Zheng, Z.; Sinha, P.** (2011): Perpetual and fair data collection for environmental energy harvesting sensor networks. *IEEE/ACM Transactions on Networking*, vol. 19, no. 4, pp. 947-960.
- Mehrabi, A.; Kim, K.** (2016): Maximizing data collection throughput on a path in energy harvesting sensor networks using a mobile sink. *IEEE Transactions on Mobile Computing*, vol. 15, no. 3, pp. 690-704.
- Padmanabh, K.; Roy, R.** (2006): Multicommodity flow based maximum lifetime routing in wireless sensor network. *International Conference on Parallel and Distributed Systems*, pp. 187-194.

Ren, X.; Liang, W.; Xu, W. (2013): Use of a mobile sink for maximizing data collection in energy harvesting sensor networks. *International Conference on Parallel Processing*, pp. 439-448.

Renner, C.; Unterschitz, S.; Turau, V.; Rmer, K. (2014): Perpetual data collection with energy-harvesting sensor networks. *ACM Transactions on Sensor Networks*, vol. 11, no. 1, pp. 1-45.

Roseveare, N.; Natarajan, B. (2014): An alternative perspective on utility maximization in energy-harvesting wireless sensor networks. *IEEE Transactions on Vehicular Technology*, vol. 63, no. 1, pp. 344-356.

Shi, F.; Huang, Y.; Ren, Z.; Wu, C. (2013): Design of adaptive tree-mesh hybrid wireless sensor networks for greenhouses. *Transactions of the Chinese Society of Agricultural Engineering*, vol. 29, no. 5, pp. 102-108.

Shrivastava, P.; Pokle, S. B. (2014): Energy efficient scheduling strategy for data collection in wireless sensor networks. *International Conference on Electronic Systems, Signal Processing and Computing Technologies*, pp. 170-173.

Siddique, M. R.; Wang, W.; Madeo, F.; Hayes, M.; O'Flynn, B. (2014): Body heat thermoelectric energy harvesting for self-powered wearable electronics. *International Conference on Body Area Networks*, pp. 236-239.

Sudevalayam, S.; Kulkarni, P. (2011): Energy harvesting sensor nodes: Survey and implications. *IEEE Communications Surveys & Tutorials*, vol. 13, no. 3, pp. 443-461.

Suzuki, M. (2012): Wireless power transmitting apparatus, wireless power receiving apparatus, and wireless power feeding system. *Journal of Computer Science & Technology*, vol. 5, no. 8, pp. 1-10.

Wan, I. S. W. D.; Yahya, S.; Taib, M. N.; Yassin, A. I. M.; Razali, R. (2014): Map: The new clustering algorithm based on multitier network topology to prolong the lifetime of wireless sensor network. *IEEE International Colloquium on Signal Processing & ITS Applications*, pp. 173-177.

Wang, J.; Ju, C.; Yu, G. (2018): A pso based energy efficient coverage control algorithm for wireless sensor networks. *Computers, Materials & Continua*, vol. 56, no. 3, pp. 433-446.

Wang, J. C.; Lin, C. H.; Siahaan, E.; Chen, B. W.; Chuang, H. L. (2014): Mixed sound event verification on wireless sensor network for home automation. *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 803-812.

Xu, X.; Qian, H. Y.; Gu, Y. L.; Du, J. (2011): A k-anycast routing protocol based on broadcast scheme for wireless sensor network. *Sensor Letters*, vol. 9, no. 4, pp. 1566-1570.

Zeng, Y.; Zhang, X. D.; Dong, Y. H. (2014): Effects of energy harvesting rate on lifetime and throughput capacity in wireless sensor networks. *Advanced Materials Research*, vol. 981, pp. 482-485.

Zhang, X.; Jue, H.; Sanglu, L.; Li, X.; Jiannong, C. (2008): Scoped bellman-ford geographic routing for large dynamic wireless sensor networks. *Journal of Computer Science & Technology*, vol. 23, no. 6, pp. 944-956.

Appendix: The inequalities based on Taylor's theorem

Ineqt-1: $\phi(x + \varepsilon) \leq \phi(x) + \varepsilon\phi'(x + \varepsilon)$

Ineqt-2: $\phi(x + \varepsilon) \geq \phi(x) + \varepsilon\phi'(x + \varepsilon) - \varepsilon^2\phi''(x + \varepsilon)$

Ineqt-3: $\phi(x + \varepsilon) \leq \phi(x) + \varepsilon\phi'(x) + \varepsilon^2\phi''(x + \varepsilon)$

