Article

Maximum Data Collection Rate Routing Protocol Based on Topology Control for Rechargeable Wireless Sensor Networks

Haifeng Lin [1], Di Bai [2,*], Demin Gao [1,3] and Yunfei Liu [1]

[1]   College of Information Science and Technology, Nanjing Forestry University, Nanjing 210037, China;
      haifeng.lin@njfu.edu.cn (H.L.); gdmnj@163.com (D.G.); lyf@njfu.edu.cn (Y.L.)
[2]   College of Engineering, Nanjing Agricultural University, Nanjing 210009, China
[3]   School of Computer Science and Engineering, Southeast University, Nanjing 211189, China
*   Correspondence: baidi000@njau.edu.cn; Tel.: +86-25-8542-7464

Abstract: In Rechargeable Wireless Sensor Networks (R-WSNs), in order to achieve the maximum data collection rate it is critical that sensors operate in very low duty cycles because of the sporadic availability of energy. A sensor has to stay in a dormant state in most of the time in order to recharge the battery and use the energy prudently. In addition, a sensor cannot always conserve energy if a network is able to harvest excessive energy from the environment due to its limited storage capacity. Therefore, energy exploitation and energy saving have to be traded off depending on distinct application scenarios. Since higher data collection rate or maximum data collection rate is the ultimate objective for sensor deployment, surplus energy of a node can be utilized for strengthening packet delivery efficiency and improving the data generating rate in R-WSNs. In this work, we propose an algorithm based on data aggregation to compute an upper data generation rate by maximizing it as an optimization problem for a network, which is formulated as a linear programming problem. Subsequently, a dual problem by introducing Lagrange multipliers is constructed, and subgradient algorithms are used to solve it in a distributed manner. At the same time, a topology controlling scheme is adopted for improving the network's performance. Through extensive simulation and experiments, we demonstrate that our algorithm is efficient at maximizing the data collection rate in rechargeable wireless sensor networks.

Keywords: wireless sensor networks; maximum data collection Rate protocol; topology control; data aggregation; routing protocol; rechargeable-WSNs

1. Introduction

Recent years have witnessed the emergence of a significant promising technology, Wireless Sensor Networks (WSNs), containing numerous inexpensive sensor nodes randomly scattered over the area of interest to collect information on entities of interest. WSNs have been used for wide-ranging applications, such as in the military [1], environmental [2], health [3] and smart home fields [4]. In WSNs, sensor nodes are generally equipped with limited energy battery and limited power storage with low power, small size and low cost [5]. Sensor nodes are usually deployed in remote or dangerous areas. They can sense ambient conditions and generate data to reflect the real-world scenario [6]. Sensors can communicate with each other over a short distance and relay information on a wireless transmission device. Through multiple hop transmission, this information can be transmitted to a remote terminal and processed by other devices.

In WSNs, since researchers have not made breakthrough progress in battery technology, a fundamental problem is the limited lifetime of sensor nodes due to limited and fixed energy

supplement [7,8]. Note, it is generally accepted that the usefulness of a wireless sensor expires when its battery runs out. A significant amount of work has been carried out across the protocol stack to prolong the sensor's lifetime by saving energy [9,10], including duty-cycling strategies [11,12] and Medium Access Control (MAC) protocols [13,14]. However, it is inevitable that the energy will be exhausted and sensors have to be discarded because it is generally impractical to render servicing for sensors. Meanwhile, it is inevitable that the accumulation of discarded batteries will contaminate the environment.

Currently, Energy Harvesting or Rechargeable Wireless Sensor Networks (EH-WSNs or R-WSNs) have attracted more and more attention. They benefit from extending the lifetime of sensor nodes by equipping them with rechargeable technologies [15,16], which convert sources, such as body heat [17], foot strike [18], finger strokes [19], and solar [20] into electricity. A sensor can operate perpetually by using supercapacitors (with virtually unlimited recharge cycles) to store the harvested energy [21]. Note, a harvesting node is said to achieve energy-neutral operation if the energy used is always to a lesser degree than the energy harvested [22]. The energy rechargeable device and energy storage device as the energy supplement components of R-WSNs replace the common accumulator battery that is used in traditional WSNs (T-WSNs), as is shown in Figure 1, in T-WSNs, the the system architecture contains four components: battery, low-power sensor node, micro-controller and wireless transceiver device.
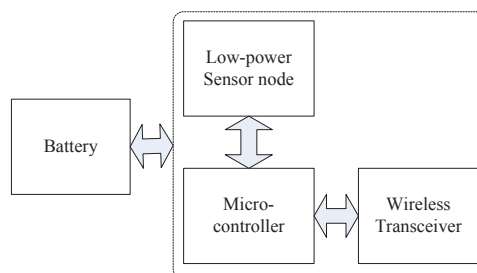

Figure 1. Components of a battery-powered sensor node.

In the R-WSNs, the system architecture contains five parts: rechargeable system, battery, low-power sensor node, Micro-controller and Wireless Transceiver device. The rechargeable system is utilized to extract the energy from ambient conditions, which is the key difference between the new WSNs and T-WSNs. The system architecture of a rechargeable sensor node is illustrated in Figure 2. Although their lifetime is less of an issue, a sensor node has to confront two issues. In R-WSNs, the energy available is usually very dynamic, which is influenced by the environment and can vary significantly over time. First and foremost, owing to the restriction of development of conversion technology, the sensor has to stay in a dormant state most of the time in order to recharge the battery and use the energy prudently. Moreover, due to the limited energy storage capacity, a node cannot always be beneficial at conserving energy when a network can harvest excessive energy from the environment [23,24]. Since more energy can be extracted from the ambient environment in R-WSNs, the harvested energy should be consumed as soon as possible [25,26]. Therefore, the surplus energy of a node can be utilized for strengthening packet delivery efficiency and improving network data collection rate.

Data aggregation can reduce data traffic significantly by removing data redundancy [27]. Since different sensor nodes partially monitor the same spatial region, sensing information is generally correlated [28]. To saving power, data should already be processed as it flows from the information source to the sink [29], which will decrease the accuracy of information around the monitoring area due to some relative packets being removed [30]. Therefore, we should let the sink collect more data only if its energy can support the action [31]. However, more packets indicate that more energy will be depleted for generating and forwarding packets. Hence, data aggregation technology can be utilized to deal with packets compression prudently [32]. More packets collected at the sink means more

accuracy of the information and more energy needed. However, data aggregation causes less packet transmission, energy depletion and accuracy. Energy expenditure and data accuracy will be traded off before utilizing data aggregation for multiple-hop transmission in WSNs and R-WSNs.
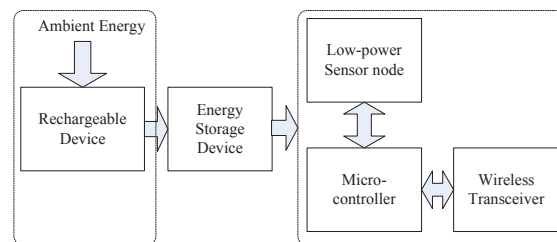

Figure 2. Components of a energy-powered sensor node.

Topology control is used for energy conservation of nodes or energy utilization reasonably, by shifting the roles of nodes from active to sleeping state and active state [33], which is significantly different to traditional WSNs. In T-WSNs, the topology control technique also is used for handling the limited energy of sensor [34], for example, if a sensor node has not been assigned new tasks, it will shut down the wireless transceiver device and stay in an idle state to save energy [35]. When a task coming, the sensor node wakes up itself and shifts into an active state for dealing with the issues [36]. However, it is inevitable to cause serious data transmission latency due to frequent shifts into the idle state in T-WSNs. In contrast, in R-WSNs, a sensor node does not shift into sleeping state when no task is assigned, but only if it has enough energy to support its operation. Therefore, in R-WSNs, topology control is utilized flexibly compared to the use of energy in T-WSNs.

In summary, we propose a maximum data collection rate routing protocol based on data aggregation for packets communications in R-WSNs. To the best of our knowledge, this is the first generic work that studies the use of distribute algorithm for maximizing data collection rate based on data aggregation technology in R-WSNs. In the work, we compute an upper data generation rate by maximizing it as an optimization problem and energy definition as constraint condition for a network, which is formulated as a linear programming problem. Subsequently, a dual problem by introducing Lagrange multipliers is constructed, and subgradient algorithms are used to solve it in a distributed manner. At the same time, a topology controlling scheme is adopted for improving network performance. The remainder of this paper is organized as follows. In Section 2 we present a number of existing routing protocol based on data aggregation, while in Section 3 specify the network model and assumption. In Section 4 we present our method and design. Section 5 contains experimental results. Conclusions are presented in Section 6.

2. Related Work

The data collection rate is affected significantly by the energy extracting rate from the environment. Therefore, a large number of researchers pay the utmost attention to the characteristics of R-WSNs [37,38], especially regarding the features of energy supplement. At present, solar is the main energy source considered firstly for providing clean power in R-WSNs and a solar panel is generally installed on the outside to recharge an energy storage device for a sensor node [39]. On the one hand, considering the cost factors and economic feasibility, the size of the solar panel always is small, which results in limited energy being gained [40]. On the another hand, considering the weather factor and the different angles of the solar panel, each sensor node will obtain distinct power. Thus, it is not realistic the achieve data transmission synchronously in R-WSNs. For maximizing data collection rate, we analyze the techniques provided in traditional WSNs.

Numerous works have been done focusing on maximizing data generation rate by saving energy in WSNs; e.g., [41–43]. A degree-constrained routing tree is introduced by minimizing hop counts with transmission power control to enhance the data collection rate, which can avoid the high-degree

bottleneck effectively [44]. In [45], the authors present the optimal convergecast scheduling algorithms by the maximum degree in the routing tree for which information is collected fast. Chen et al. [46] provide an iterative linear programming solution for finding the maxmin optimal rate assignment in a low-rate sensor network. Su et al. [47] formulate the optimal rate allocation with data aggregation as a network utility maximization problem and apply duality theory to decompose it into a rate control subproblem, other examples can be found in [48,49], and references therein. Despite the maximum data generation rate achieved in these algorithms, they encounter critical tradeoffs between data flow and lifetime due to energy constraint. In R-WSNs, the lifetime is less of an issue, which can be maximized by operating in an energy neutral mode for nodes.

Regarding the characteristics of R-WSNs, several protocols discuss several aspects of power management or MAC schemes to improve energy efficiency and maximum data collection rate [50,51]. A centralized algorithm with line programming is proposed to compute the lexicographically maximum data collection rate and routing paths for each node [52]. Sadlapur et al. [53] provide a distribute algorithm for jointly determining the routing structure and amount of flows on each link with flow adjustment to achieve an optimal data collection rate. Peng et al. [54] propose real time adaptive energy management policies based on observed information for optimal throughput. Prabhakar et al. [55] propose four throughput enhancement schemes from a simple naive scheme with low complexity to a probabilistic probing scheme incorporating advanced methods to appropriately use the harvesting energy. However, the ideal energy replenished precondition is used and data aggregation has not been considered in these protocols.

Data aggregation is adopted widely in WSNs for improving data collection rate by removing redundant information [56]. The main idea behind data aggregation or data fusion is that, rather than sending individual data items from sensors to sinks, multiple data items are aggregated as they are forwarded by the sensor network [57]. In [44], the authors introduce two strategies to maximize the data collection rate: an aggregation tree is contracted to minimize the number of time slots needed and combine the scheduling with transmission power control to reduce the effects of interference. Demin et al. [58] introduce an algorithm for maximum data collection rate in rechargeable wireless sensor networks with multiple sinks. In [59], the authors present a fuzzy-based data fusion approach for increasing the QoS whilst improving the data collection rate and reducing the energy consumption of the sensor network.

The topology control mechanism is utilized widely for reducing energy consumption to prolong the network lifetime by predicting the topology control messages in WSNs [60]. For coping with the data collection with delay-constraint, a load-aware power-increased topology control algorithm is presented to heuristically solve the problem, which reaches the O(1)-approximation ratio for the linear networks [61]. A new topology control mechanism, called hierarchical neighbor graphs, is proposed to build a structure for data collection depending on a single parameter, which requires only local knowledge at each node to be formed and repaired [62]. In [63], the authors propose approaches to maximize the topological network lifetime of the WSNs by considering a two-tiered Wireless Sensor Network deployed around strategic locations and base-stations. These techniques with topology controlling schemes are utilized for saving the energy and prolonging the network lifetime. Very few papers in the literature focus on the maximum data collection rate with data aggregation and topology control simultaneously in R-WSNs.

In contrast to earlier works, which either focus on static battery-powered network or maximum data generation rate without considering data fusion and topology control, in this work, we present a maximum data collection rate based on data aggregation and topology control for R-WSNs. In summary, on observing the lack of data aggregation and topology control techniques for improving data flow in existing routing protocols, we introduce the first generic routing protocol algorithm with the data aggregation and topology scheme in R-WSNs. To the best of our knowledge, there is no prior work studying the problem of maximum data generation rate routing protocol with data aggregation and topology control for R-WSNs.

3. System Model

3.1. Network Model

Consider a static rechargeable wireless sensor network modeled as an undirected graph $G = (V, A)$, where V is the set of n rechargeable sensor nodes and sink nodes within the wireless sensor network. A is the set of links, $A = \{A | (i, j) \in A, i, j \in V\}$. G consists of a finite nonempty vertex set V and edge set A of ordered pairs of distinct vertices of V. Two nodes i and j are connected by a link if they can transmit a packet to each other with a transmission power less than the maximum transmission power at each node, where a sensor node generally obtains a maximum transmission range depending on its maximum transmission power and neighbors only within the transmission range can receive the packets from the node. The maximum transmission power is affected by the hardware of wireless radio transmitting device. Considering the condition of energy supplement, at present, the maximum transmission range is set to be less than 200 m in the real scenario.

In the networks, all links are assumed to be bi-directional. This assumption is not necessary for the convergence of the distributed algorithms; however, it can make the presentation clearer. A graph is simple if it has no loops and no two of its links join the same pair of vertices. An acyclic graph is one that contains no cycles. The set of nodes are connected to node i by links denoted as S_i, which indicates that all sensor nodes in the set S_i can communicate with sensor i directly without forwarding other sensor nodes. 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. Isolated nodes are not considered in the work, even though they exist. The data generation rate and the remain energy of node i are expressed as g_i, e_i, respectively.

3.2. Energy Expenditure Model

The power consumption of a sensor node consists of five parts: sensing and generating data, idling, listening, receiving, and transmitting. The the power for generating one bit of data, from an idle state consumed by a node, is denoted by e_g and e_{idle}, respectively. They are assumed to be the same for all nodes and independent of traffic. e_l denotes the power expenditure for a sensor listening channel at each time point before sending a packet. For power consumption in receiving and transmitting, the first order radio model is adopted in [64]. Specifically, a node needs ε_{elec} = 50 nJ for running the circuitry and ε_{amp} = 100 pJ/bit/m^2 for the transmitting amplifier. Therefore, the power consumption for receiving one bit of data is given by $e_r = \varepsilon_{elec}$. The power consumption for transmitting one bit of data to a neighbor node j is given by $e_t = \varepsilon_{elec} + \varepsilon_{amp} \times d_{i,j}^n$, where n is the path loss exponent, which typically ranges between 2 and 4 for free-space and short-to-medium-range radio communication. $d_{i,j}$ denotes the Euclidean distance of the node i and j. If the coordinate of node i and j are (x_i, y_i) and (x_j, y_j), respectively, the $d_{i,j}$ can be formulated as:

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

We assume the the data traffic from node i to j per unit time is $f_{i,j}$. The per unit time can be set to one second or one minute or one times data transmission in the work, which is decided by the requirement of experiment. The different units selected will not affect the performance and evaluation of the system. The energy consumption for node i in receiving and transmitting data flow $f_{i,j}$ are $e_t(i, j)$ and $e_r(i, j)$, respectively, which can be expressed as:

$$e_t(i, j) = e_t \times \sum_{i,j \in V, j \in S_i} f_{i,j} \tag{2}$$

$$e_r(i, j) = e_r \times \sum_{i,j \in V, j \in S_i} f_{i,j} \tag{3}$$

Let w_i denote the total fraction of power consumption for node *i* per unit time. Considering the energy deplete for five parts, w_i can be formulated as:

$$w_i = e_g \times g_i + e_{idle} + e_l + e_t(i,j) + e_r(i,j) \tag{4}$$

3.3. Energy Replenish Model

In R-WSNs, the energy extracting from ambient conditions varies with environment and weather conditions. Even though an energy source cannot be controlled to yield energy at the desired times, its behavior can be modeled to predict the expected availability at a given time. For a sensor power relying on solar power, as is shown in Figure 3, its energy cannot be controlled. However, models for its dependence on diurnal and seasonal cycles are known and can be used to predict availability. For a region, its weather parameter is usually similar to the history data of the same period. Suppose the power output from the energy source is $P_i(t)$ at time *t*.

Figure 3. A sensor with solar panel is deployed in outside.

In many instances, the energy generation profile may be very different from the consumption profile. To help support this scenario, consider a device that has a mechanism to store any energy that is harvested. The stored energy may be used at any time later. The energy buffer is defined to be a device that can store limited energy, has an inefficient storage in charging, and does not leak any energy over time. In this case, the energy harvesting from ambient for all non negative values of *T* can be calculated as:

$$E_i = \int_0^T P_i(t)d(t) + B_i, \forall T \in [0, \infty) \tag{5}$$

where, B_i is the initial energy stored in the energy buffer and it always is below the storage capacity, $B_i \leq B$. Each sensor has limited energy buffer with equal storage capacity, which should be enough to provide at least one packet delivery or acceptance. The energy storage capacity is determined by physical material and storage technology, which is not the main point of our consideration in the work. Harvesting energy will be stored in an energy buffer at first, if a node is in the ideal state, until a packet arrives or if the buffer is full.

In the work, $P_i(t)$ indicates the ability of a harvesting energy device extracting energy from the environment. A high $P_i(t)$ value means more energy was extracted from the ambient conditions. For a sensor with solar panel as an energy supplement, the $P_i(t)$ is affected by three factors: lighting conditions, the size of panel, and the energy conversion efficiency. The energy conversion efficiency is influenced by energy conversion technology, for instance, at present, the solar-energy conversion efficiency is about 10%–30% [16]. We hypothesize that the power density (per square centimeter),

the size of panel (square centimeter), and energy conversion efficiency are λ_t, s_{panel}, γ, respectively. Therefore, the $P_i(t)$ can be concluded as:

$$p_i(t) = \lambda_t \times s_{panel} \times \gamma \tag{6}$$

For a solar panel as the energy source, the power density varies with weather conditions. Obviously, more power can be provided for a sensor on a sunny day and the power density can reach nearly 250,000 lux. Table 1 shows the power density in different weather conditions.

Table 1. The power density in different weather condition.

Weather Condition (at 12:00 noon)	The Power Density
cloudy day inside	4000–7000 lux
cloudy day outside	8000–12,000 lux
sunny day inside	30,000–50,000 lux
sunny day outside	>100,000 lux

In R-WSNs, a sensor is expected to achieve energy-neutral operation, its energy used is always less than the energy harvested. The following equation should be satisfied for energy consumption.

$$w_i \leq E_i \tag{7}$$

3.4. Data Aggregation Model

To incorporate data aggregation into the geometric routing model, we adopt the foreign-coding model [65] scheme. Specifically, we assume a node i is able to compress the data originating at its upstream neighbor j using its local data. The compression ratio depends on the data correlation between node i and j, which is denoted by the correlation coefficient $\rho(i,j) = 1 - H(i|j)/H(i)$, where $H(i)$ is the entropy coded data rate of the information at node i, and $H(i|j)$ is the conditional entropy coded data rate of the same information $H(i)$ at node i given the side information $H(j)$. Examples of correlation models include the Gaussian random field model [66] which assumes that the correlation coefficient $\rho(i,j)$ decreases exponentially with the distance between nodes, or $\rho(i,j) = \frac{1}{1+d(i,j)}$, and the inverse model [65] which assumes the data correlation is inversely proportional to the Euclidean distance between nodes, or $\rho(i,j) = exp(-\alpha \times d^2(i,j))$, where, α is data correlation parameters. Higher α means smaller data correlation, and vice versa.

3.5. Data Transmission Model

Let $S_i = \{j|d_{i,j} \leq R, j \in N\}$, where R is the radius of the node i' transmission range. According to the geometric routing, only those neighbors that are closer to the sink node s can serve as the downstream nodes. Let us denote this set of downstream neighbors as $\overrightarrow{S_i} = \{j|d_{j,s} < d_{i,s}, j \in S_i\}$. Similarly, the set of upstream neighbors is denoted as $\overleftarrow{S_i} = \{j|d_{j,s} \geq d_{i,s}, j \in S_i\}$. Note that in case a node has no neighbors that are closer to the sink node than itself, we encounter a problem known as "local maximum" where the node fails to find routing path to the sink node according to geometric routing. A few solutions have been proposed for this problem [67,68]. However, the consideration of these solutions is beyond the scope of this paper. In the following, we assume that the downstream neighbor set $\overrightarrow{S_i}$ non-empty for all $i \in N$.

To further illustrate the process of data aggregation, an example is provided, which is shown in Figure 4. The data aggregation and routing work as follows. A sensor node j is ready to send data to its downstream neighbor node s. The data from node j contains two parts, one parts is aggregation data or transit data passed from node i (which has been encoded by node i or its upstream nodes), no further encoding is performed, which is directly forwarded to the downstream neighbors. Another part is the

raw data of node *j*, it is encoded with the local information. All this transit traffic is forwarded to the downstream neighbors.
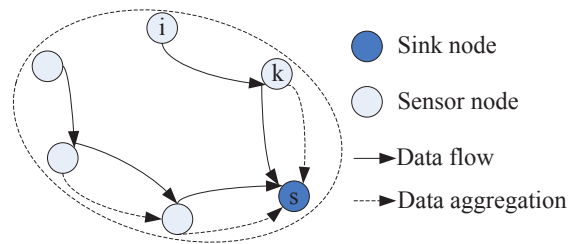

Figure 4. The process of data aggregation.

3.6. Topology Control Model

The topology control model is utilized for achieving two purposes in this work. First, the data transmission latency from source node to destination will be reduced or minimized by adjusting the transmission range of sensors flexibly. There is one more point, we should touch on: the energy utilization efficiency will be improved or maximized by neighbors' selection strategies, rather than by saving power of the sensor node. Considering the limited energy storage ability, a sensor cannot always be able to conserve energy. Therefore, since the energy can be replenished continually, a sensor will consume the energy as soon as possible only if its energy can support its operation.

To further illustrate the operation of the topology control model, an example is provided, which is shown in Figure 5. The sensor *i* senses the surroundings and generates packets, which will be forwarded to the destination sink *S* through multiple-hop data transmission. In traditional WSNs, by the Equation (2), we know that a sensor always select a closer neighbor as its next-hop node for data receiving in order to save energy. Therefore, data packets will be transmitted to sink *S* along the routing path $\{i, a, c, e, S\}$. While, in R-WSNs, considering the energy of sensor *i* can be replenished continually, it may select sensor *c* as its net-hop node only if it has enough energy to support this operation. The new routing path $\{i, c, S\}$ can be established for data transmission. A packet is forwarded to sink *S* through two hops in the new path, while it is four hops in the original path. Therefore, the End to End delay for data transmission is reduced dramatically depending on the topology control model, which is significantly different to traditional WSNs.
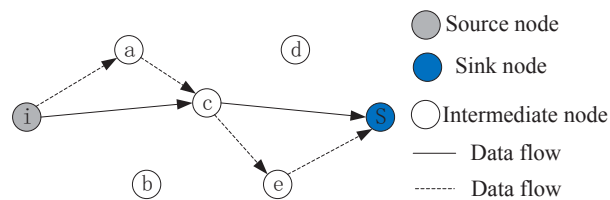

Figure 5. Topology control for data transmission.

4. Design and Method

4.1. Description of the Data Aggregation Problem

In R-WSNs, in the process of data transmission from packets generated by source nodes to these packets accepted by sink, packets will be encoded and aggregated in each sensor with local packets for reducing the data traffic. We let λ_i denote the data aggregation rate of node *i* and $\varphi(j, i)$ denotes the aggregated transit traffic rate at node *i* from its upstream node *j*. The aggregated transit traffic of node *i* is a superposition of two parts: the transit traffic passed from the upstream nodes, and the raw data

originated from the upstream nodes that is to be encoded using the local information, as shown in Figure 6. That is,

$$\lambda_i = \sum_{i,j \in V, j \in \vec{S}_i} (\lambda_j \varphi(j,i) + g_j(1 - \rho(j,i))) \tag{8}$$

where, g_j and $\rho(j,i)$ denote the data generation rate of node j and the data correlation between node i and j, respectively. For sensor j, the data traffic $\sum_{i,j \in V, j \in \overleftarrow{S}_i} \lambda_j \varphi(j,i) + g_j$ will be transmitted to its downstream sensor i, while these packets will be aggregated by sensor i firstly and send new data traffic $\sum_{i,j \in V, j \in \vec{S}_i} (\lambda_j \varphi(j,i) + g_j(1 - \rho(j,i)))$ to next-hop sensor along the path subsequently. The next-hop sensor that receives the packet will compress the packets by the Equation (8), so forth, until it reaches the destination.
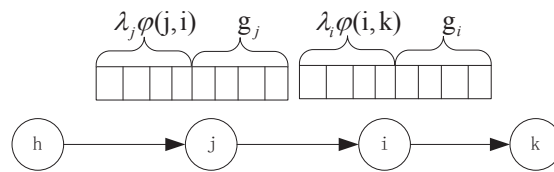
Figure 6. The aggregated transit traffic from a sensor to another sensor.

In this work, multiple sinks are deployed for balancing the power of the networks, which is recognized and used wildly in WSNs. Generally, these sinks obtain the same function and status in the work. The data from a source node will be transmitted to any sink finally by one or multiple-hops. For each node, the outflow equals or is less than the inflow and generation data from it due to some redundant information removed from data traffic. In the process of packets transmission, a data leak is not considered. Therefore, the data flow traffic for node i can be expressed as:

$$\sum_{i,j \in V, j \in \overleftarrow{S}_i} (\lambda_j \varphi(j,i) + g_j(1 - \rho(j,i))) = \sum_{i,j \in V, j \in \vec{S}_i} \lambda_i \varphi(i,j) \tag{9}$$

where, $\rho(j,i) \in [0,1]$, when a sensor only forward packets for its upstream neighbor to next-hop sensor without any packet generation, data fusion does not occur and $\rho(j,i) = 0$, while the two sensors have the exact same information and $\rho(j,i) = 1$, which indicates that the packets are duplicated from that of another sensor. Let w_i denote the fraction of power consumption for node i in each time unit. We have

$$w_i = e_g \times g_i + e_{idle} + e_l + e_r \sum_{j \in S_i} (\lambda_j \varphi(j,i) + g_j) + e_t \sum_{j \in S_i} (\lambda_i \varphi(i,j) + g_i) \tag{10}$$

4.2. Description of the mAximum Data Collection Rate

In R-WSNs, the data collection rate of a sensor is affected directly by the amount of remaining energy extracted from ambient conditions. More energy supplement means that more packets will be collected. According to the Equations (7) and (10), for a sensor i, the maximum data collection rate per unit of time is formulated as:

$$g(i) = \frac{E_i - (e_{idle} + e_r \sum_{j \in \overleftarrow{S}_i} (\lambda_j \varphi(j,i) + g_j))}{e_g} - \frac{e_l + e_t \sum_{j \in \vec{S}_i} (\lambda_i \varphi(i,j) + g_i)}{e_g} \tag{11}$$

where, e_l is used for a sensor to listen channel before data delivery with a fixed value for each time transmission. Hence, it is affected significantly by data flow, i.e., a sensor does not need to broadcast its

existence when the channel is idle if there are no packets to deliver. Now, we let $e_l = k \times \sum_{i,j \in V, j \in S_i} f_{i,j}$. The energy depletion for listening and transmission can be expressed as:

$$e_l + e_t(i,j) = (k + e_t) \times \sum_{i,j \in V, j \in S_i} f_{i,j} \tag{12}$$

Since k is fixed and e_t is variable, we can let e_{tr} represents the $k + e_t$. In fact, the energy utilization in idle state is a fixed value and small with less than 1 mW in per unit of time, which is only about 1/100 that of transmission and neglected in our work. Hence, the Equation (11) can be rewritten as:

$$g(i) = \frac{E_i - e_r \sum_{j \in \overleftarrow{S_i}} (\lambda_j \varphi(j,i) + g_j)}{e_g} - \frac{e_{tr} \sum_{j \in \overrightarrow{S_i}} (\lambda_i \varphi(i,j) + g_i)}{e_g} \tag{13}$$

From the Equation (13), we can observe the data collection rate for a sensor is affected significantly by the amount of its received data. The available energy is mainly consumed for at least two parts: receiving and transmitting. When more packets are accepted by a sensor, its data generation rate declines due to limited power supplement, which is an objective and obvious phenomenon in WSNs and R-WSNs. However, completely distinct strategies are adapted in them and a sample is provided for elaborating different routing schemes, as shown in Figure 7.
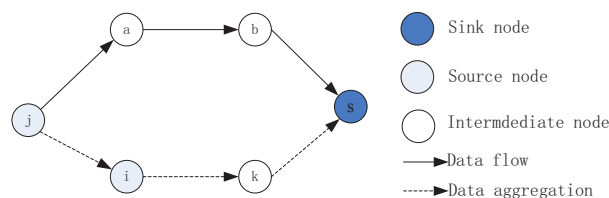

Figure 7. Different data transmission model.

We assume node j and node i are source nodes and generate packets by sensing the environment periodically. There are at least two transmission schemes for delivering packets to the sink s. One of them is that packets can be forwarded along the link path $\{j, i, k, s\}$, where sensor i aggregates the packets from an upstream neighbor j with its self generated data and forwards it to the next-hop adjacent sensor until reaching the sink s. This scheme is widely and recognized and used in traditional WSNs, which can reduce the amount of packets by removing redundant information and reducing energy consumption, effectively contributing to extended network lifetime.

Another transmission scheme is that data from sensor j and i are delivered along paths $\{j, a, b, s\}$ and $\{i, k, s\}$, respectively. Both raw data generated by node j and i will be transmitted to destination without considering data aggregation. The main differences to the two schemes can be summarized separately. For the first strategy, the advantage is energy consumption reduced with fewer packets by aggregating data. However, on the one hand, the data generation rate of node i declines. On the other hand, nodes a, b do not participate in the data transmission procedure, although their power is replenished repeatedly. For the second strategy, even though total energy consumption increases, the data collection rate of sensor i and sensors usage efficiency for node a, b are improved. Therefore, for improving data collection rate, the latter data delivery method is more suitable to R-WSNs with better performance compared to the first method.

The data collection rates vary over time as ambient condition is changed and are different to each other, and are significantly affected by energy harvesting rate and energy depletion rate. It is crucial to improve the data flow traffic for a sensor with minimum data collection rate. Therefore, we firstly find the minimum data collection rate of sensors. Secondly, we use the topology control and data fusion schemes. Lastly, we chose appropriate routing protocol to improve the data generation rate. Precisely, the objective is to find an optimal maximum minimum data collection rate, which will be defined in the following section.

4.3. Distribute Algorithm for Maximizing Data Collection Rate

The goal of the maximum data collection rate routing protocol is to deliver all the data packets generated by sensor nodes to base stations as soon as possible, subject to node harvesting energy, node capacity and link capacity constraints such that the data flow until the first (set of) sensor node with minimum data generation rate is maximized followed by data flow until the second (set of) sensor node with minimum value is maximized, and so on. Now, we formulate this problem as a line programming, which is given by:

$$Max \quad g_i$$
$$Subject \quad to:$$

$$\sum_{j \in \overleftarrow{S_i}} (\lambda_j \varphi(j,i) + g_j(1 - \rho(j,i))) = \sum_{j \in \vec{S_i}} \lambda_i \varphi(i,j) \tag{14}$$

$$e_r \sum_{j \in \overleftarrow{S_i}} (\lambda_j \varphi(j,i) + g_j) + e_{tr} \sum_{j \in \overrightarrow{S_i}} (\lambda_i \varphi(i,j) + g_i)$$

$$+ e_g g_i \leq E_i \tag{15}$$

$$\lambda_i = \sum_{i,j \in V, j \in \vec{S_i}} (\lambda_j \varphi(j,i) + g_j(1 - \rho(j,i))) \tag{16}$$

$$\sum_{j \in \overleftarrow{S_i}} \varphi(j,i), \sum_{j \in \overrightarrow{S_i}} \varphi(i,j) \leq R_i, \forall i,j \in V \tag{17}$$

$$\varphi(j,i), \varphi(i,j) \geq 0, \forall i,j \in V, j \in S_i \tag{18}$$

$$0 \leq g_i \leq R_i \tag{19}$$

$$E_i \leq B, E_1 = W_1 \tag{20}$$

$$\lambda_1 = g_1 \tag{21}$$

The first set of constraints in Equation (14) ensures that the inflow after aggregating with raw generation data equals the outflow. The second constraint in Equation (15) ensures that nodes do not consume more energy than they collect, which includes the energy consumption in sensing, packet transmissions, and packet receptions. The third constraint in Equation (16) formulates the λ_i of node *i*. Equations (17)–(18) state that the data flow and available energy do not go below zero and do not go above the link capacity or battery capacity. The set of constraints in Equation (19) ensures that the data collection rate of a node is positive and finite. Equation (20) states that the available energy in the system initialization is the initial battery level. The last Equation (21) states the data aggregation rate equals data collection rate for a leaf or edge node without upstream neighbors.

According to the Equation (12), the first set of constraints models the data flow conservation at each node. We change the variable to $q = 1/g$, which indicates the generation time for each unit of the packet. We define decay $(q(i))$ as the inverse of data collection rate $(q(i) = 1/g(i), g(i) > 0)$. Therefore, maximum data collection rate can be converted to the minimum time per unit of packet generation. We obtain an equivalent linear programming formulation.

Here again, the flow conservation and power conservation constraints should be satisfied first. There are additional constraints that enforce all q_i to be equal. Again we consider a quadratic objective function that is strictly convex in the q_i. Also, to ensure that the dual function is differentiable, we restrict the domain to $0 \leq q_i \leq q$, for some loose upper bound q. In addition, we use a simple approach similar to that used in [69]. We change the primal objective function to q_i^2, since minimizing q_i is the same as minimizing q_i^2. This is the optimization problem such that the maximum problem to maximize data collection rate is converted to the minimum problem to minimize data generation time for per unit of byte information, which will be solved in a distributed manner. We can interpret the above problem as minimizing the maximum ratio of time elapsed to collet a packet at a node. The linear

programming problem for Equation (22) is NP-hard and considerably difficult to solve directly. Hence, a dual model will be utilized for replacing the original problem.

$$
\begin{aligned}
&Min \quad q_i^2 \\
&Subject \quad to: \\
&\sum_{j\in S_i}(\lambda_i\varphi(i,j) - \lambda_j\varphi(j,i)) = \frac{1-\rho(j,i)}{q_j} \\
&h \le q_i(E_i - \sum_{j\in S_i}(e_r\lambda_j\varphi(j,i) + e_{tr}\lambda_i\varphi(i,j))) \\
&q_j(\lambda_i - \sum_{j\in \vec{S}_i}\lambda_j\varphi(j,i)) = 1 - \rho(j,i) \\
&h = e_r + e_{tr} + e_g \\
&q_i = q_j \\
&\sum_{j\in \overleftarrow{S}_i}\varphi(j,i), \sum_{j\in \overrightarrow{S}_i}\varphi(i,j) \le R_i, \forall i,j \in V \\
&\varphi(j,i), \varphi(i,j) \ge 0, \forall i,j \in V, j \in S_i \\
&0 \le q_i \\
&E_i \le B, E_1 = W_1 \\
&\lambda_1 = 1/q_1
\end{aligned}
\tag{22}
$$

4.4. Lagrangian Dual Problem

We construct the dual problem by introducing Lagrange multipliers v_i for the flow conservation constraint and μ_i for the data collection rate constraint at each node i. This results in the Lagrangian.

$$
\begin{aligned}
L(Q,f,\mu,v) &= q_i^2 \\
&+ \sum_{i\in V} v_i\{\sum_{j\in S_i}(\lambda_i\varphi(i,j) - \lambda_j\varphi(j,i)) - \frac{1-\rho(j,i)}{q_j}\} \\
&+ \sum_{i\in V} \mu_i\{q_i - q_j\} \\
&= q_i^2 + \sum_{i\in V}(q_i(\mu_i - \mu_j) - \sum_{j\in S_i} v_i\frac{1-\rho(j,i)}{q_i}) \\
&+ \sum_{i\in V}\sum_{j\in S_i} v_i(\lambda_i\varphi(i,j) - \lambda_j\varphi(j,i))
\end{aligned}
\tag{23}
$$

The dual function is given by

$$
\begin{aligned}
&g(\lambda,v) \\
&= \inf_{0\le q_i\le q}\{L(Q,f,\mu,v) \mid h \le q_i(E_i - \sum_{j\in S_i}(e_r\lambda_j\varphi(j,i) \\
&+ e_{tr}\lambda_i\varphi(i,j)))\} \\
&= \inf_{0\le q_i\le q,\lambda_i}\{q_i^2 + \sum_{i\in V}(q_i(\mu_i - \mu_j) - \sum_{j\in S_i} v_i\frac{1-\rho(j,i)}{q_i}) \\
&+ \sum_{i\in V}\sum_{j\in S_i} v_i(\lambda_i\varphi(i,j) - \lambda_j\varphi(j,i)) \mid h \le q_i(E_i \\
&- \sum_{j\in S_i}(e_r\lambda_j\varphi(j,i) + e_{tr}\lambda_i\varphi(i,j)))\}
\end{aligned}
\tag{24}
$$

4.5. Subgradient Algorithm

We will use the subgradient algorithm to solve the dual problem. During the k'th iteration of the subgradient algorithm, each node solves the following convex quadratic program with variables q_i, λ_i, for $i \in V, j \in S_i$.

$$
\begin{aligned}
Min \quad & q_i^2 + \sum_{i \in V} (q_i(\mu_i - \mu_j) - \sum_{j \in S_i} v_i \frac{1 - \rho(j, i)}{q_i}) \\
& + \sum_{i \in V} \sum_{j \in S_i} v_i(\lambda_i \varphi(i, j) - \lambda_j \varphi(j, i)) \\
Subject \quad to : \\
& h \le q_i(E_i - \sum_{j \in S_i} (e_r \lambda_j \varphi(j, i) + e_{tr} \lambda_i \varphi(i, j)))
\end{aligned}
$$

The subgradient of $-d$ at (μ^k, v^k) is given by

$$
\begin{aligned}
\gamma_{\mu_i}^{(k)} &= q_j^{(k)} - q_i^{(k)} \\
\delta_{v_i}^{(k)} &= 1 - \rho(j, i) - \sum_{j \in S_i} (\lambda_i^{(k)} \varphi(i, j) - \lambda_j^{(k)} \varphi(j, i))
\end{aligned}
\tag{25}
$$

Hypothesis: the $\mu^{(0)}, v^{(0)}$ are initial values of Lagrange multiplier, after k times of iteration.

$$
\begin{aligned}
\mu_i^{(k+1)} &= (\mu_i^{(k)} - \theta_k \gamma_{\mu_i}^{(k)})_+, \forall i = 1, 2, \ldots, m \\
v_i^{(k+1)} &= (v_i^{(k)} - \theta_k \delta_{v_i}^{(k)})_+, \forall i = 1, 2, \ldots, p
\end{aligned}
\tag{26}
$$

where, θ_k denotes the iteration step. For any $\theta > 0$, it exist a value r, when $\theta_k \in (0, r)$.

$$
\theta_k > 0, \theta_k \to 0, \sum_{k=1}^{\infty} \theta_k = \infty
\tag{27}
$$

The algorithm converges to the maximum value $d(\mu, v)$. If it is a strong duality, $d^{(k)} = d^*(\mu^{(k)}, v^{(k)})$ converges to an optimal solution for the original problem. The subgradient algorithms are used to solve it in a distributed manner. The algorithm is utilized for calculating the upper bound of data flow under ideal conditions rather than establishing a detailed routing path. In actual application, the network throughput will be lower than the result value of our method.

5. Results and Discussion

Simulation of our algorithm for R-WSNs was done by Matlab software, with up to 200–400 nodes and 3–8 sinks being randomly deployed in a 500 m × 500 m to 1000 m × 1000 m square field. The maximum communication range of each node is set to be 80 m. All nodes' energy devices are rechargeable with 20 cm^2 square size solar panel, and transmission powers are adjustable. The solar panel will be affected by cloudy or sunny environments, as well as the angle of sunlight, as solar radiation may change at any time, along with time or climate. Every two nodes can communicate with each other directly within their transmission range. Energy leakage and the case of signal loss of sensor are not considered in our work. $\theta_k \approx 1/3500$, $\rho(i, j) = exp(-\alpha \times d^2(i, j))$, $\alpha \in [0.001, 0.01]$, low value of α indicates high correlation, vice versa. Every data point in simulation figures is obtained by averaging 50 runs with different random seeds, node deployment and node working schedules, and some related simulation parameters are provided in Table 2. In the simulation, part of the information of the sensor stems from an actual environment, such as, the signal intensity, signal interference, etc.

Table 2. The simulation parameters.

Parameter	Value
area	500 m $\times$ 500 m to 1000 m $\times$1000 m
packet	512 byte
sinks number	3–8
sensors number	200–400
ε_{elec}	50 nJ
ε_{amp}	100 pJ/bit/m^2
the size of panel s_{panel}	20 cm
energy conversion efficiency γ	30%
transmission range R	80 m
Batteryrange($Bmin; Bmax$)	0, 10 MJ

What we use consists of a solar panel optimized for outdoor use, two eZ430-RF2500T target boards and one AAA battery pack, which is rechargeable and can be recharged repeatedly. The target board comprises the TIMSP430 microcontroller, CC2500 radio transceiver and an on-board antenna. The CC2500 radio transceiver operates in the 2.4 GHz band with data rate of 250 kbps and is designed for low power wireless applications. The harvested energy is stored in EnerChip, a thin-film rechargeable energy storage device with low self-discharge manufactured by Cymbet, as is shown in Figures 8 and 9.

Figure 8. The outdoor scenes for two sensors.

Figure 9. A sensor node is deployed in outside.

5.1. System Implement

Due to available energy constrained, it is a common practice to deploy multiple sinks for collecting packets from all sources in realistic application, which means there are at least two advantages compared with just one sink. Firstly, in the multiple sinks scenario, since packets generated by sensors are only needed to be forwarded to any one sink, the degree of routing path from source node to sink will be shorted significantly. Secondly, a limited number of sensors are distributed around the sink and it often represents a bottleneck for data transmission because all packets from sources have to be forwarded to these sensors before reaching a destination, while it will relieve this pressure for data forwarding with more than one sink. These two strong points will save energy and improve data generation rate.

Figures 10 and 11 show the routing paths established using different communication models when 20 sensors and 2 sinks deployed in experiment scenario with a 200 m × 240 m square area, which contains 6 source nodes, a sensing signal and other sensors undertaking forwarding tasks. Figure 10 presents the routing paths for all sources to reach at least one sink based on maximum data collection rate of our algorithm. By considering the limited energy supplement, a source node still chooses a closer base station as its destination; however, the packets from sources may travel through all intermediate nodes and the energy consumption will be balanced to all nodes in order to improve sensor usage efficiency.
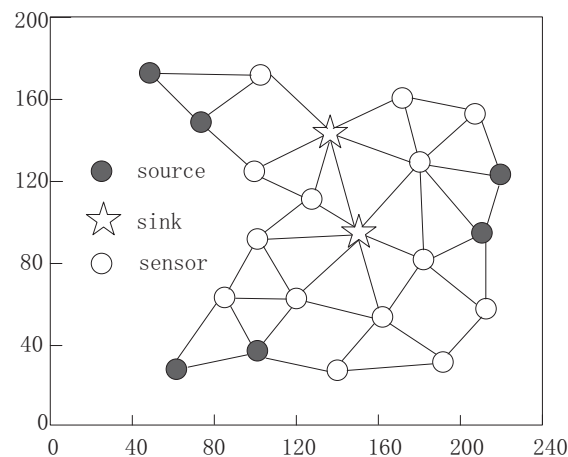
Figure 10. The routing process using maximum data collection rate scheme.
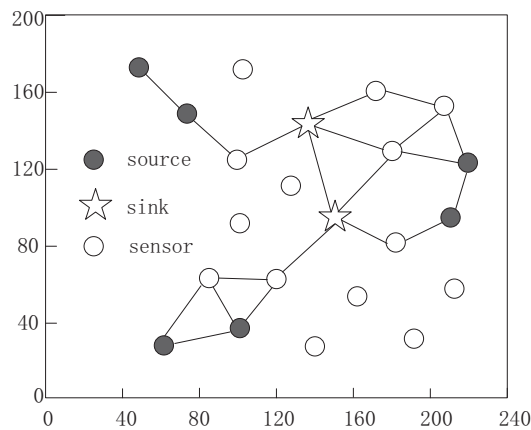
Figure 11. The routing process using minimum energy consumption scheme.

From the Figure 11, we can observe that only some of the nodes participate in the routing path formation. The Figure 11 shows that the routing process is based on the minimum energy consumption [70]. A source node will create a route path to the nearest base station with smallest link distance between each of the two adjacent nodes in order to reduce energy consumption, which is recognized and widely used in traditional WSNs. When a source sends its data to another source, data aggregation occurs and total packets are reduced, which is attributed to low power expenditure. There are at least four advantages of this scheme, such as, high data delivery speed, better system stability, low data transmission delay and low energy expenditure. However, it is designed for traditional WSNs based on fixed battery with constrained energy, which must be correct in R-WSNs.

5.2. System Performance Comparison

In order to further understand of the performance of our algorithm for Maximizing data collection rate (Mdcr) under network settings, in this section, we provide two schemes for performance comparison, including an optimal distributed lexicographic rate assignment (Dlex) designed for R-WSNs [53] and an Energy-efficient data aggregation tree (Edat) applied in WSNs. In [53], the authors propose distributed algorithms for determining the routing structure and amount of flows on each link without considering data fusion. It involves updated rate computation using optimal lexicographic rate assignment. In this method, nodes first compute their maximum rate using an initial rate procedure. Subsequently, they send a control packet containing the flow id and the maximum achievable rate to their next hop nodes.

In [71], Edat is constructed for energy efficiency, in which energy management is the main task and data aggregation is a secondary aim for energy efficiency in routing design. There is limited literature on maximum data generation rate routing based on data aggregation in R-WSNs. At the same time, the two methods were widely recognized and used for maximum data collection rate in R-WSNs and data fusion in WSNs, which can be attributed to the decision to select them for performance comparison in our work.
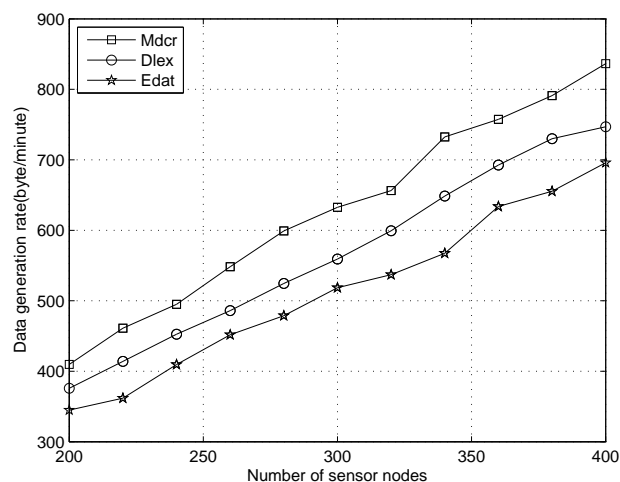

Figure 12. Data generation rate with number of nodes increasing when duty cycle is 1%.

We first compare the data generation rate between our algorithm and other two schemes under the distinct number of sensors and only one sink, where the average node's duty-cycles are 1%, 10%, 30%, respectively, as shown in Figures 12–14. From Figure 12, we can observe the data collection rates increase for all algorithms and that of our algorithm is slightly higher than that of other algorithms with the sensor density being improved for all different node's duty-cycles. Our algorithm is utilized for calculating the upper bound of data flow under ideal conditions rather than establishing a detailed

routing path. In actual application, the network throughput will be lower than the result value of our method. From Figure 12, we can see that the data generation rate of our algorithm is about 9% and 15% higher than that of the Dlex and Edat scheme, respectively.
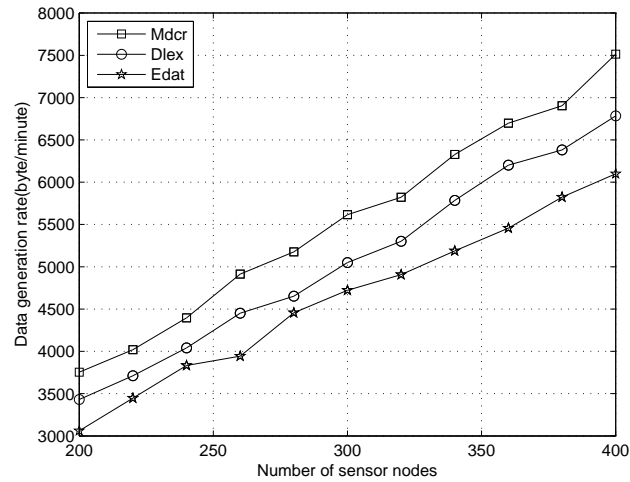

Figure 13. Data generation rate with number of nodes increasing when duty cycle is 10%.
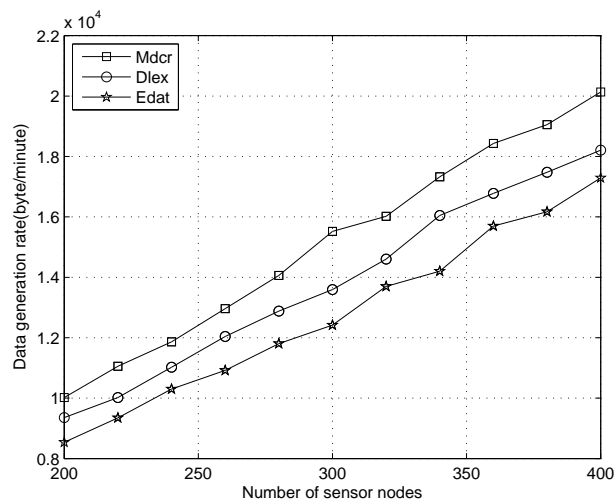

Figure 14. Data generation rate with number of nodes increasing when duty cycle is 30%.

Since high sensor density indicates that more packets will be sensed and forwarded to the destination, the data collection rate will be improved, and power expenditure will be balanced between all sensors if we append more sensor nodes to the monitoring field in a reasonable number of sensors. Despite this, sensors usually keep in a low duty cycle and collision is scarce: if too many sensors are deployed in a smaller area, the collision will be critical and retransmission is more frequent, which causes serious energy waste. In this work, we only consider an ideal condition that a reasonable number of sensors were deployed outdoors and collision was ignored.

Comparing Figure 12 with Figure 14, when the duty cycles exceed 30 times in Figure 14 compared with that in Figure 12, we can see that the data collection rate is improved significantly and it is about 25 times. Higher data collection rate is obtained with improved duty cycles, which indicates that a sensor will be in an active state more times and data transmission occurs more frequently. The deeper reason for higher duty cycle is that more energy is extracted from the environment. Therefore,

a sensor will generate more packets utilizing extra energy, from which we observe that more data will be collected on a sunny day than in cloudy weather.

We next analyze the data collection rate for all sensors under three data correlation settings ($\alpha = 0.001$, 0.005 and 0.01), as shown in Figures 15–17. From Figure 15, we can observe that the data flow rate in three algorithms increases with the number of nodes increasing in three different scenarios, with our algorithm performing a little better than the Dlex and Edat. As the number of nodes increases in the area, it affects the node density and delivery ratio. The overall raw data rate is proportional to the number of nodes in the network. More packets were generated in the network and the data aggregation rate improved greatly. At the same time, the nodes' density increased, which also drives the network topology from sparse to dense and the data correlation between neighboring nodes becomes higher, so more redundant information can be removed through data aggregation.
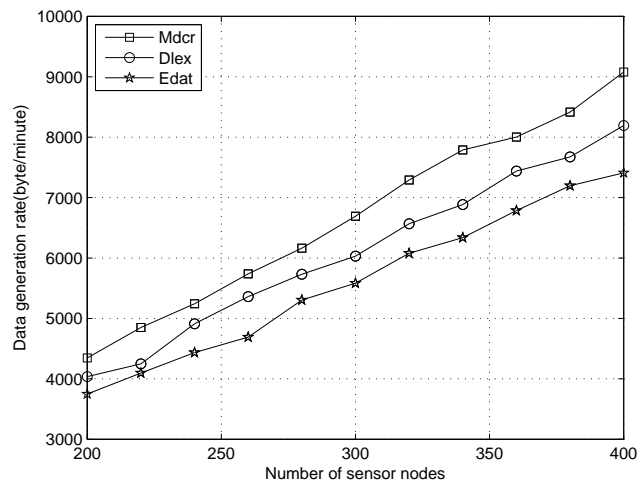
Figure 15. Data generation rate with number of nodes increasing when data correlation $\alpha = 0.001$.
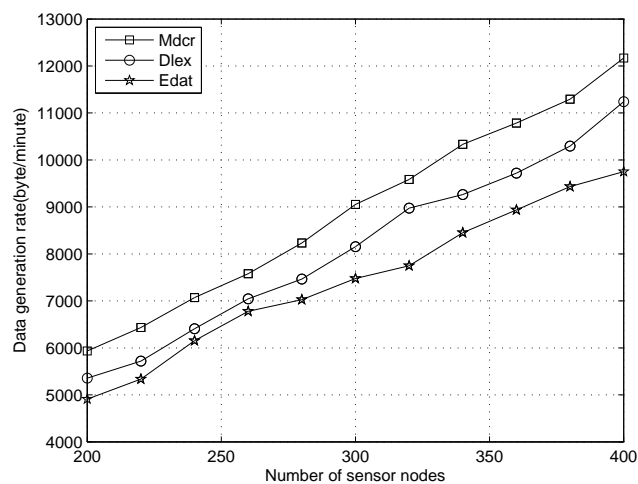
Figure 16. Data generation rate with number of nodes increasing when data correlation $\alpha = 0.005$.

With improving α, the data correlation and the data redundancy decrease, which indicates that the data variance increases, such as, packets come from two sensors with long distance or different scenario. Hence, more packets will be transmitted to the sink and data aggregation rate is also improved. When $\alpha = 0.001$, the data correlation coefficient and data redundancy are relatively high, packets can be

compressed with high ratio and data collection rate increases due to a decrease in the power being used for data forwarding. While $\alpha = 0.01$, the data correlation coefficient is low, which indicates that the packets are very different to each other and data is hard to compress. Therefore, original raw data will be forwarded to the sink and there will be more power expenditure for this procedure, which causes low data generation rate.
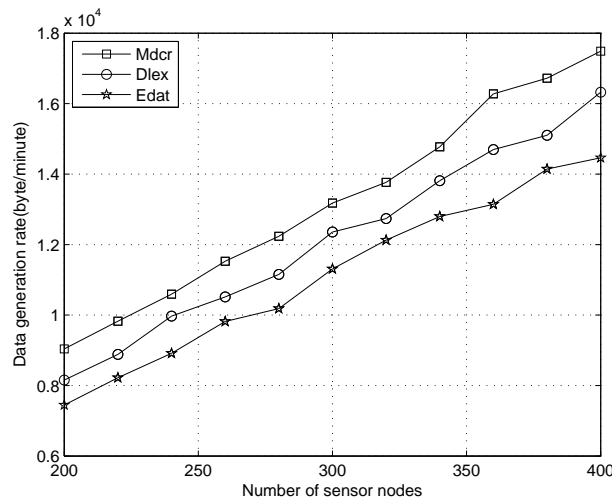

Figure 17. Data generation rate with number of nodes increasing when data correlation $\alpha = 0.01$.

To further illustrate the performance of data collection based on the topology controlling for our scheme we introduce a new algorithm that uses a cooperative repeater to improve data collection in low power generation for solar-powered wireless sensor networks [72]. In [72], the authors propose a repeater role to some nodes for improving the data collection performance, which only relays data from the other nodes without transmitting its own data. The reason for selecting the methods for performance comparison is that high data collection rate is the purpose and the scene is also subject to solar-powered WSNs. At the same time, the selected method is recognized and widely used for improving data collection rate in WSNs.
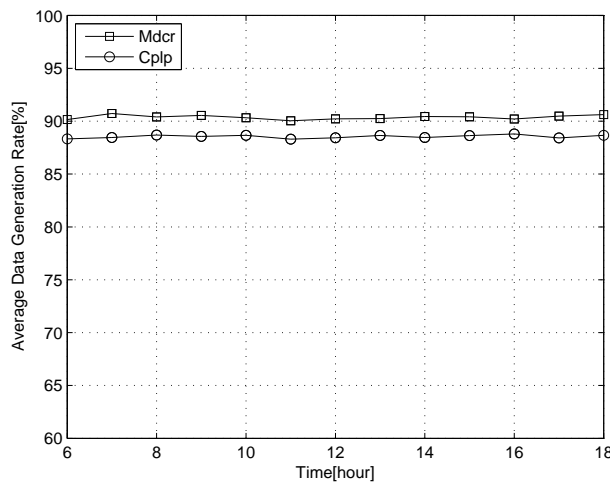

Figure 18. Comparison of the average data collection rate for a day when the probability of sunny day.

Figure 18 shows the comparison of the average data collection rate for a day when the probability of a sunny day is high in our scheme and in [72], which can be abbreviated as Cplp (Cooperative repeater for low power WSNs). We can see that the average data collection rate of the scheme sustains a high lever compared to Cply on a sunny day. This is because the topology controlling techniques are utilized for adjusting its communication paths continually, while this method has been considered in Cplp.

We analyze the total energy depletion for all sensors with the number of nodes increasing when the duty cycle is 1%, 10%, 30%, as shown in Figures 19–21, respectively. From Figure 19, we can observe that the energy depletion increases gradually when more sensors are added to the scenario, and our algorithm will consume more energy than other methods for any number of sensors. This does not mean that energy is wasted and there is a network lifetime reduction, which is the point of traditional WSNs. On the contrary, it demonstrates an outstanding performance for our algorithm because the power of each node can be replenished continually, which is significantly different to traditional WSNs, where saving energy is considered prior to routing design, e.g., the Edat scheme obtains better performance with respect to saving energy.
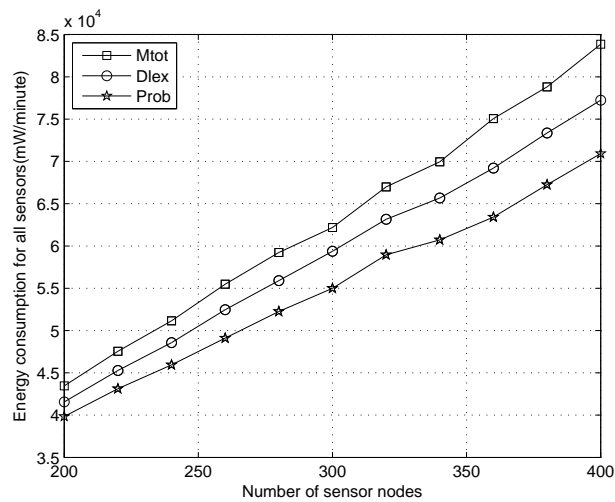

Figure 19. Energy consumption with the number of sensor nodes increasing when the duty cycle is 1%.
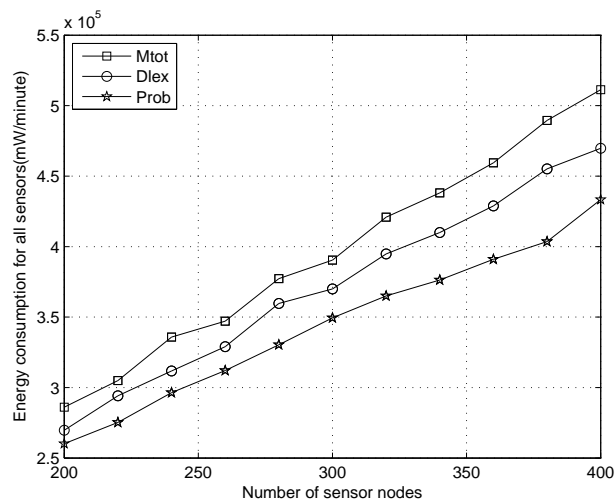

Figure 20. Energy consumption with the number of sensor nodes increasing when the duty cycle is 10%.

Comparing Figure 19 with Figure 21, when the duty cycles are raised 30 times in Figure 21 compared to Figure 19, we can see that the total energy expenditure increases significantly, i.e., about 20 times. When the duty cycles improved due to data collection rate increases, the total energy consumption will improved, which indicates that more energy is extracted from the environment. From another point of view, when more energy is supplied for a sensor, it will be more lively in the active state and undertake more tasks, such as, increased packet generation. Therefore, on a sunny day, we will improve the duty cycles of sensors with more energy and consume the energy as soon as possible.
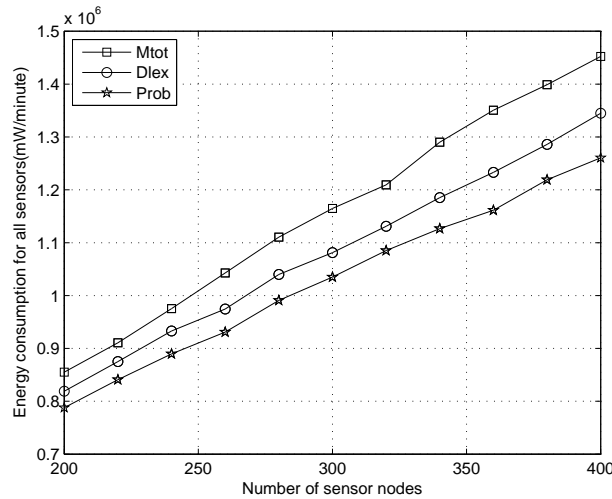

Figure 21. Energy consumption with the number of sensor nodes increasing when the duty cycle is 30%.
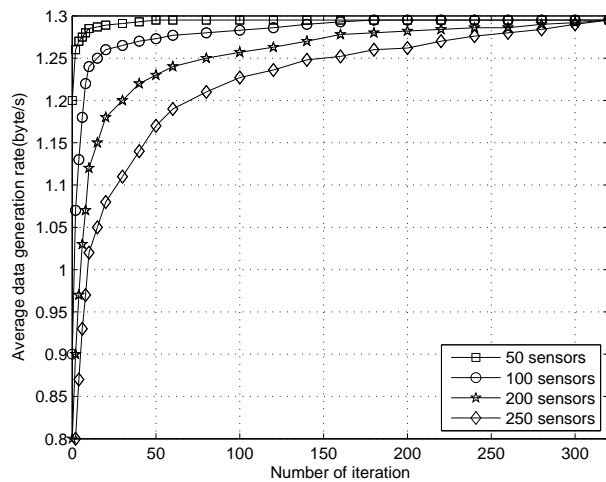

Figure 22. The iterations for data generation rate converges to maximum value.

Figure 22 shows that the data generation rate for each sensor converges to maximum value in our algorithm when the number of sensors are 50, 100, 200, respectively. For fewer nodes, the algorithm converges faster to the maximum generation rate. As the number of nodes increase, convergence of iterations needs to increase gradually. This is because, with the expansion of a network scale, the total sample space is growing, and algorithms often need to go through more iterations to achieve convergence. As it can be seen from Figure 22 , when the number of nodes is 50, 100, 200 and 250, the number of times to reach 95% of maximum generation rate value needs: 55, 120, 220, 300 iterations to

be performed, repectively. Overall, it will converge to maximum value with iterations no more than double the size scale of the networks in all scenarios, which is affordable in terms of the computing ability of the sensors.

The connectivity graph is provided in Matlab for expressing the routing paths for all sensors with connection to any one sink, where 100 sensors and 5 sinks are deployed in a 500 m $\times$ 500 m square field, which is shown in Figure 23. By considering the limited energy supplement, a source node still chooses a closer base station as its destination. However, the packets from all sources may travel through all intermediate nodes in the network area before they reach a sink, and the communication energy consumption will be balanced to all nodes in order to improve sensor utilization efficiency.
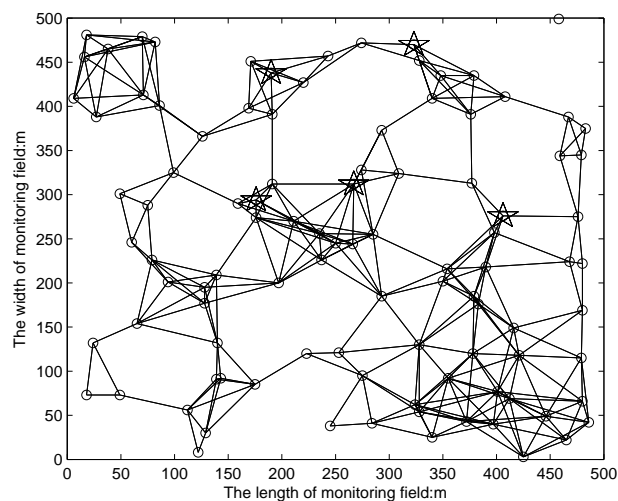

Figure 23. Connectivity graph with 100 nodes and 5 sinks deployed in a 500 m $\times$ 500 m.

6. Conclusions

In this work, we first define the network system, energy expenditure and replenished model, data aggregation scheme, and elaborate on the packet delivery process. The unique characteristics of R-WSNs pose a big challenge for packet's transmission. On the one hand, data collected by many sensors is based on common phenomena, and hence there is a high probability that this data has some redundancy. On the other hand, due to energy being replenished continually and limited energy storage capacity, it might not always be beneficial to conserve energy when a network can harvest excessive energy from the environment. Therefore, reasonable energy usage can improve data collection rate and sensor usage efficiency.

In order to improve data generation rate, we propose an algorithm to compute an upper data generation rate based on data fusion that maximizes it as an optimization problem for a network. First and foremost, we formulate it as a linear programming problem subject to the flow and energy conservation constraints. On top of that, a dual problem by introducing Lagrange multipliers is constructed. Last but not least, a subgradient algorithm is used to solve it in a distributed manner. Through extensive simulation and experiments, we demonstrate that our algorithm is efficient to maximize data collection rate in rechargeable wireless sensor networks.

References

1. Gupta, R.; Sultania, K.; Singh, P.; Gupta, A. Security for wireless sensor networks in military operations. In Proceedings of the 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, India, 4–6 July 2013.

2. Gao, D.; Liu, Y.; Zhang, F.; Song, J. Anycast Routing Protocol for Forest Monitoring in Rechargeable Wireless Sensor Networks. *Int. J. Distrib. Sens. Netw.* **2013**, *2013*, 1–14.

3. Hackmann, G.; Guo, W.; Yan, G.; Sun, Z.; Lu, C.; Dyke, S. Cyber-physical codesign of distributed structural health monitoring with wireless sensor networks. *IEEE Trans. Parallel Distrib. Syst.* **2014**, *25*, 63–72.

4. Wang, J.C.; Lin, C.H.; Siahaan, E.; Chen, B.-W.; Chuang, H.-L. Mixed sound event verification on wireless sensor network for home automation. *IEEE Trans. Ind. Inform.* **2014**, *10*, 803–812.

5. Akyildiz, I.F.; Su, W.; Sankarasubramaniam, Y.; Cayirci, E. Wireless sensor networks: A survey. *Appl. Mech. Mater.* **2013**, *300–301*, 490–493.

6. Hodge, V.J.; O'Keefe, S.; Weeks, M.; Moulds, A. Wireless Sensor Networks for Condition Monitoring in the Railway Industry: A Survey. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 1088–1106.

7. Din, W.I.S.W.; Yahya, S.; Taib, M.N.; Yassin, A.I.M.; Razali, R. MAP: The new clustering algorithm based on multitier network topology to prolong the lifetime of wireless sensor network. In Proceedings of the 2014 IEEE 10th International Colloquium on Signal Processing & its Applications (CSPA), Kuala Lumpur, Malaysia, 7–9 March 2014; pp. 173–177.

8. Xu, X.; Qian, H.Y.; Gu, Y.L.; Jie, D. A k-Anycast Routing Protocol Based on Broadcast Scheme for Wireless Sensor Network. *Sens. Lett.* **2011**, *9*, 1566–1570.

9. Saghar, K.; Kendall, D.; Bouridane, A. Application of formal modeling to detect black hole attacks in wireless sensor network routing protocols. In Proceedings of the 2014 11th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, 14–18 January 2014; pp. 191–194.

10. Shao, X.; Wang, C.; Xiang, H.; Huang, J. Network coding based Energy Efficient Multicast Routing for wireless sensor network. In Proceedings of the 2013 IEEE 4th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, 15–17 November 2013; pp. 293–296.

11. Guo, S.; He, L.; Gu, Y.; Jiang, B.; He, T. Opportunistic Flooding in Low-Duty-Cycle Wireless Sensor Networks with Unreliable Links. *IEEE Trans. Comput.* **2014**, *63*, 2787–2802.

12. Faheem, Y.; Boudjit, S. Duty-cycle SN-multi-point relay algorithm for mobile sink wireless sensor networks. *IET Wirel. Sens. Syst.* **2015**, *5*, 219–227.

13. Oller, J.; Demirkol, I.; Casademont, J.; Josep, P.; Gerd, U.G.; Leonhard, R. Wake-up radio as an energy-efficient alternative to conventional wireless sensor networks MAC protocols. In Proceedings of the ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, Barcelona, Spain, 3–8 November 2013; pp. 173–180.

14. Ahmed, M.M.; Bari, S.M.S.; Habaebi, M.H.; Khan, S.; Anwar, F. Traffic Aware Wireless Sensor Networks MAC Protocol for Smart Grid Applications Using Spiral Backoff Mechanism. In Proceedings of the IEEE International Conference on Computer and Communication Engineering, Kuala Lumpur, Malaysia, 23–25 September 2014; pp. 1177–1197.

15. Gao, D.; Wu, G.; Liu, Y.; Zhang, F. Bounded end-to-end delay with Transmission Power Control techniques for rechargeable wireless sensor networks. *AEU-Int. J. Electron. Commun.* **2014**, *68*, 395–405.

16. Sudevalayam, S.; Kulkarni, P. Energy harvesting sensor nodes: Survey and implications. *IEEE Commun. Surv. Tutor.* **2011**, *13*, 443–461.

17. Starner, T. Human-powered Wearable Computing. *IBM Syst. J.* **1996**, *35*, 618–629.

18. Xie, L.; Cai, M. An In-shoe Harvester with Motion Magnification for Scavenging Energy from Human Foot Strike. *IEEE/ASME Trans. Mechatron.* **2015**, *20*, 3264–3268 .

19. Gummeson, J.; Priyantha, B.; Liu, J. An energy harvesting wearable ring platform for gestureinput on surfaces. In Proceedings of the MobiSys'14, Bretton Woods, NH, USA, 16–19 June 2014.

20. González, A.; Aquino, R.; Mata, W.; Ochoa, A.; Saldaña, P.; Edwards, A. Open-wise: A solar powered wireless sensor network platform. *Sensors* **2012**, *12*, 8204–8217.

21. Eu, Z.A.; Tan, H.P.; Seah, W.K.G. Opportunistic routing in wireless sensor networks powered by ambient energy harvesting. *Comput. Netw.* **2010**, *54*, 2943–2966.

22. Kansal, A.; Hsu, J.; Zahedi, S.; Srivastava, M.B. Power Management in Energy Harvesting Sensor Networks. *ACM Trans. Embed. Comput. Syst.* **2007**, *6*, 32.

23. Wan, Z.G.; Tan, Y.K.; Yuen, C. Review on energy harvesting and energy management for sustainable wireless sensor networks. In Proceedings of the 2011 IEEE 13th International Conference on Communication Technology (ICCT), Ji'nan, China, 25–28 September 2011; pp. 362–367.

24. Seah, W.K.G.; Tan, Y.K.; Chan, A.T.S. *Research in eNergy Harvesting Wireless Sensor Networks and the Challenges Ahead*; Springer Berlin Heidelberg: Berlin, Germany, 2013.

25. Giuppi, F.; Niotaki, K.; Collado, A.; Georgiadis, A. Challenges in energy harvesting techniques for autonomous self-powered wireless sensors. In Proceedings of the IEEE 2013 European Microwave Conference (EuMC), Nuremberg, Germany, 6–10 October 2013; pp. 854–857.

26. Olds, J.P.; Seah, W.K.G. *Power Considerations for Very Low duty Cycle Wireless Sensor Networks Powered by Energy Harvesting*; Victoria University of Wellington: Wellington, New Zealand, 2011.

27. Koutsopoulos, I.; Halkidi, M. Measurement Aggregation and Routing Techniques for Energy-Efficient Estimation in Wireless Sensor Networks. In Proceedings of the 8th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt 2010), Avignon, France, 31 May–4 June 2010.

28. Guha, A.; Psota, E.T.; Perez, L.C. Non-binary joint network-channel decoding of correlated sensor data in wireless sensor networks. In Proceedings of the 2011 IEEE Wireless Communications and Networking Conference (WCNC), Detroit, MI, USA, 24–28 July 2011; pp. 950–955.

29. Pandey, V. A review on data aggregation techniques in wireless sensor network. *J. Electron. Electr. Eng.* **2010**, *1*, 1–8.

30. Jongsoo, J.; Jaeseok, K.; Woosuk, C.; Kim, H.; Kim, S.; Mah, P. A QoS-Aware Data Aggregation in Wireless Sensor Networks. In Proceedings of the 12th International Conference on Advanced Communication Technology: ICT for Green Growth and Sustainable Development (ICACT), Phoenix Park, UK, 7–10 Feburay 2010; pp. 156–161.

31. Maraiya, K.; Kant, K.; Gupta, N. Wireless sensor network: A review on data aggregation. *Int. J. Sci. Eng. Res.* **2011**, *2*, 1–6.

32. Ozdemir, S.; Cam, H. Integration of False Data Detection With Data Aggregation and Confidential Transmission in Wireless Sensor Networks. *IEEE/ACM Trans. Netw.* **2010**, *7*, 736–749.

33. Qureshi, H.K.; Iqbal, A.; Asif, W. Topology Control for Harvesting Enabled Wireless Sensor Networks: A Design Approach. *Wirel. Pers. Commun.* **2015**, *82*, 81–101.

34. Chu, X.; Sethu, H. Cooperative Topology Control with Adaptation for improved lifetime in wireless sensor networks. *Ad Hoc Netw.* **2015**, *30*, 99–114.

35. Torkestani, J.A. An Energy-Efficient Topology Control Mechanism for Wireless Sensor Networks Based on Transmit Power Adjustment. *Wirel. Pers. Commun.* **2015**, *82*, 2537–2556.

36. Fereydooni, M.; Sabaei, M.; Babazadeh Eslamlu, G. Energy Efficient Topology Control in Wireless Sensor Networks with Considering Interference and Traffic Load. *Adhoc Sens. Wirel. Netw.* **2015**, *25*, 289–308.

37. Guo, S.; Wang, C.; Yang, Y. Joint Mobile Data Gathering and Energy Provisioning in Wireless Rechargeable Sensor Networks. *IEEE Trans. Mob. Comput.* **2014**, *13*, 2836–2852.

38. Mao, S.; Man, H.C.; Wong, V.W.S. Joint Energy Allocation for Sensing and Transmission in Rechargeable Wireless Sensor Networks. *IEEE Trans. Veh. Technol.* **2014**, *63*, 2862–2875.

39. Zou, T.; Lin, S.; Feng, Q.; Chen, Y. Energy-Efficient Control with Harvesting Predictions for Solar-Powered Wireless Sensor Networks. *Sensors* **2016**, *16*, 53.

40. Li, X.; Chen, H.; Zhao, F.; Li, X. A range-based sleep scheduling algorithm for desired area coverage in solar-powered wireless sensor networks. In Proceedings of the 2014 Sixth International Conference on Wireless Communications and Signal Processing (WCSP), Hefei, China, 23–25 October 2014.

41. Patel, R.; Pariyani, S.; Ukani, V. Energy and throughput analysis of hierarchical routing protocol (LEACH) for wireless sensor network. *Int. J. Comput. Appl.* **2011**, *20*, doi:10.5120/2420-3247.

42. Walsh, M.; Hayes, M. A robust throughput rate control mechanism for an 802.15. 4 wireless sensor network-an anti-windup approach. In Proceedings of the IEEE American Control Conference (ACC'07), New York, NY, USA, 9–13 July 2007; pp. 3071–3076.

43. Boubiche, S.; Boubiche, D.E.; Bilami, A.; Toral-Cruz, H. An Outline of Data Aggregation Security in Heterogeneous Wireless Sensor Networks. *Sensors* **2016**, *16*, 525.

44. Incel, O.D.; Krishnamachari, B. Enhancing the data collection rate of tree-based aggregation in wireless sensor networks. In Proceedings of the 5th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON'08), San Francisco, CA, USA, 16–20 June 2008; pp. 569–577.

45. Incel, O.D.; Ghosh, A.; Krishnamachari, B.; Chintalapudi, K. Fast data collection in tree-based wireless sensor networks. *IEEE Trans. Mob. Comput.* **2012**, *11*, 86–99.

46. Chen, S.; Fang, Y.; Xia, Y. Lexicographic maxmin fairness for data collection in wireless sensor networks. *IEEE Trans. Mob. Comput.* **2007**, *6*, 762–776.

47. Su, L.; Gao, Y.; Yang, Y.; Cao, G. Towards optimal rate allocation for data aggregation in wireless sensor networks. In Proceedings of the Twelfth ACM International Symposium on Mobile Ad Hoc Networking and Computing, Paris, France, 16–20 May 2011.

48. Hakoura, B.; Rabbat, M.G. Data aggregation in wireless sensor networks: A comparison of collection tree protocols and gossip algorithms. In Proceedings of the 2012 25th IEEE Canadian Conference on Electrical & Computer Engineering (CCECE), Montreal, QC, Canada, 29 April–2 May 2012; pp. 1–4.

49. Li, X.Y.; Wang, Y.; Wang, Y. Complexity of data collection, aggregation, and selection for wireless sensor networks. *IEEE Trans. Comput.* **2011**, *60*, 386–399.

50. Zeng, Y.; Zhang, X.D.; Dong, Y.H. Effects of Energy Harvesting Rate on Lifetime and Throughput Capacity in Wireless Sensor Networks. *Adv. Mater. Res.* **2014**, *981*, 482–485.

51. Roseveare, N.; Natarajan, B. An Alternative Perspective on Utility Maximization in Energy Harvesting Wireless Sensor Networks. *IEEE Trans. Veh. Technol.* **2013**, *63*, 344–356.

52. Liu, R.S.; Fan, K.W.; Zheng, Z.; Sinha, P. Perpetual and fair data collection for environmental energy harvesting sensor networks. *IEEE/ACM Trans. Netw.* **2011**, *19*, 947–960.

53. Sadlapur, A.; Pushpa, P.V. Computing optimal data collection rate for energy harvesting sensor networks. In Proceedings of the 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Mysore, India, 22–25 August 2013; pp. 1468–1472.

54. Peng, S.; Low, C.P. Throughput optimal energy neutral management for energy harvesting wireless sensor networks. In Proceedings of the 2012 IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, China, 1–4 April 2012; pp. 2347–2351.

55. Prabhakar, T.V.; Iyer, M.; Prakruthi, K.; Prasad, R.V.; Niemegeers, I.; Sathyanarayanan, S.P. Throughput schemes for energy harvesting sensor networks. In Proceedings of the 2012 Fourth International Conference on Communication Systems and Networks (COMSNETS), Bangalore, India, 3–7 January 2012; pp. 1–7.

56. Kumar, T.P.; Reddy, K.N. A competent approach of data aggregation in tree based wireless sensor networks. *Int. J. Comput. Electron. Res.* **2013**, *2*, 80–82.

57. Solis, I.; Obraczka, K. In-network aggregation trade-offs for data collection in wireless sensor networks. *Int. J. Sens. Netw.* **2006**, *1*, 200–212.

58. Demin, G.; Lin, H.; Liu, Y.; Wu, G. Maximum data collection rate in rechargeable wireless sensor networks with multiple sinks. *China Commun.* **2016**, *13*, 95–108.

59. Izadi, D.; Abawajy, J.H.; Ghanavati, S.; Herawan, T. A data fusion method in wireless sensor networks. *Sensors* **2015**, *15*, 2964–2979.

60. Meseguer, R.; Molina, C.; Ochoa, S.F.; Santos, R. Energy-aware topology control strategy for human-centric wireless sensor networks. *Sensors* **2014**, *14*, 2619–2643.

61. Xu, H.; Huang, L.; Liu, W.; Wang, G.; Wang, Y. Topology control for delay-constraint data collection in wireless sensor networks. *Comput. Commun.* **2009**, *32*, 1820–1828.

62. Bagchi, A.; Madanm, A.; Premi, A.; Sankhla, S. Hierarchical neighbor graphs: A topology control mechanism for data collection in heterogeneous wireless sensor networks. *Ad Hoc Sens.Wirel. Netw.* **2015**, *26*, 171–191.

63. Pan, J.; Hou, Y.T.; Cai, L.; Shi, Y.; Shen, S.X. Topology control for wireless sensor networks. In Proceedings of the International Conference on Mobile Computing and Networking (MOBICOM 2003), San Diego, CA, USA, 14–19 September 2003; pp. 263–274.

64. Yin, Y.; Liu, F.; Zhou, X.; Li, Q. An Efficient Data Compression Model Based on Spatial Clustering and Principal Component Analysis in Wireless Sensor Networks. *Sensors* **2015**, *15*, 19443–19465.
65. Yin, J.L.; Yang, Y.; Lei, W. An Adaptive Data Gathering Scheme for Multi-Hop Wireless Sensor Networks Based on Compressed Sensing and Network Coding. *Sensors* **2016**, *16*, doi:10.3390/s16040462.
66. Cristescu, R.; Beferull-Lozano, B.; Vetterli, M. On network correlated data gathering. In Proceedings of the Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004), Hong Kong, China, 7–11 March 2004.
67. Kuhn, F.; Wattenhofer, R.; Zollinger, A. Asymptotically optimal geometric mobile ad-hoc routing. In Proceedings of the 6th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications (DIALM '02), Atlanta, GA, USA, 28 September 2002; pp. 24–33.
68. Javaid, N.; Ilyas, N.; Ahmad, A.; Alrajeh, N.; Qasim, U.; Khan, Z.A.; Liaqat, T.; Khan, M.I. An Efficient Data-Gathering Routing Protocol for Underwater Wireless Sensor Networks. *Sensors* **2015**, *15*, 29149–29181.
69. Xiao, L.; Johansson, M.; Boyd, S. Simultaneous routing and resource allocation via dual decomposition. *IEEE Trans. Commun.* **2004**, *52*, 1136–1144.
70. Wang, J.; Li, B.; Xia, F.; Kim, C.S.; Kim, J.U. An energy efficient distance-aware routing algorithm with multiple mobile sinks for wireless sensor networks. *Sensors* **2014**, *14*, 15163–15181.
71. Zhang, J.; Shen, X.; Tang, S.; Dai, G. Energy efficient joint data aggregation and link scheduling in solar sensor networks. *Comput. Commun.* **2011**, *34*, 2217–2226.
72. Ishitani, N.; Kobayashi, K.; Okada, H.; Katayama, M. Cooperative repeaters to improve data collection in low power generation for solar-powered wireless sensor networks. In Proceedings of the 12th Annual IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, USA, 9–12 January 2015; pp. 361–366.