GRDR: A Novel Data Gathering and Dissemination Scheme for WSNs

Jinhuan Zhang, Jun Long, Anfeng Liu, Guihu Zhao

School of Information Science and Engineering, Central South University, China
zhang_jh81@163.com, jlong@csu.edu.cn, afengliu@mail.csu.edu.cn, guihuzhao.csu@gmail.com

Abstract

Data gathering and dissemination is a key issue for WSNs. Data collection should take energy efficient, reliability as well as delay into consideration. To address these problems, a novel reproduction packets routing scheme named Green, Reliability and Delay based Route (GRDR) is proposed. The GRDR scheme integrates three core phases, namely, packets reproduction, direction dispersity and multi-route. The key idea of the GRDR is to generate different number of new packet copies after certain steps according to the distance to sink, thus obviating the energy efficiency and latency minimizing for data collection. We formulate the optimization problem as to maximize lifetime under end-to-end delay and given reliability constraints by controlling the system parameters. To demonstrate the effectiveness of the scheme, we conduct extensive theoretical analysis and simulations to evaluate the performance of GRDR. The analysis and simulations show that GRDR is more energy efficient and lower delay than the existing scheme.

Keywords: GRDR, Network lifetime, Reliability, Delay, Reproduction routing

1 Introduction

For most applications in wireless sensor networks (WSNs), it is essential to guarantee reliable communications [1-6]. Many applications require that each data packet is successfully delivered to sink with statistical probability $\delta < 1$, such as 60-95% [7-8], which is sufficient for these applications such as environment (temperature, humidity), agriculture (water tank, irrigation) [9]. In such cases, 100% reliable communications are costly and unnecessary [1]. But some of the reports reveal that wireless link packet error rate may be as high as 30% in real WSNs which are far from being satisfactory [10]. In addition, delay is also an important metric for data collection in sensor networks. It plays a vital role to the application to transport the detected information quickly to sink, in order to make a rapid response to the event. Therefore, the delay is generally defined as the time required for sensor nodes to send the sensed data to sink [11]. At last, due to wireless sensor networks nodes' limited battery capacity and normally it is not supplemented, to prolong the lifetime of the network as far as possible is an extremely important key issue for ensuring application.

There are two kinds of solutions to ensure reliability based on packet-loss recovery, which are Automatic Repeat-reQuest (ARQ) protocol and proliferation routing (PR) scheme [1]. In the ARQ, the data transmission reliability is guaranteed by retransmitting data. And in the PR approach, the source node has multiple copies of data packets and transmits them concurrently through multiple routes. The packet loss is recovered by the in-middle recovery or packets reproduction of PR. By comparison, the delay is generally larger because it resends data packet after perceiving the data loss for ARQ. While PR mechanism has lower delay because it employs multiple paths routing and its delay is closer to the ideal case where no data is lost. But, to guarantee the network reliability, the number of multiple routing paths is often calculated according to the worst case. This leads to more energy consumption and affects the network lifetime. These facts highlight the fact that energy efficient (Green), data Reliability, end-to-end Delay (GRD) in WSN are major concerns requiring further investigation. According to our current study, there are few researches on the comprehensive problem of energy efficiency, reliability and delay.

To address these issues, a novel routing scheme is proposed in this paper. It integrates the consideration of energy efficient (Green), data Reliability, end-to-end Delay and is called GRDR (Green, Reliability and Delay based Route, GRDR). Compared with the previous study, we make the following contributions:

A novel GRDR scheme is proposed to achieve longer network lifetime. As can be seen from the previous studies, the network lifetime depends entirely on the energy cost of nodes in hotspots [12-16]. Therefore, to reduce energy cost of nodes in hotspots can effectively improve network lifetime and to increase the energy cost of nodes in non-hotspots dose not necessarily affect the network lifetime. Accordingly,

*Corresponding Author: Jun Long; E-mail: jlong@csu.edu.cn
DOI: 10.3966/160792642018031902011

the data collection scheme named GRDR is proposed. It is similar with PR that in order to ensure data transmission success rate, packet needs to be reproduced to compensate for those lost ones when it successfully finishes a certain number of steps. The process is repeated until the data reach the destination. However, the core idea of the GRDR scheme is to make longer routes and more data copies for nodes in the area far from the sink, and shorter routes and the less number of data copies for nodes in the area close to the sink. In this way, the network delay and network lifetime could be optimized with the guarantee of network reliability.

It can effectively minimize end-to-end delay in wireless sensor networks. It is very challenging to reduce end-to-end delay with the guarantee of network reliability and network lifetime. The proposed GRDR scheme is similar with PR because PR has better advantage of delay than ARQ protocol. The network delay is reduced by reducing the number of renew times. Because packet reproduction needs more time than that required for packet transmission in the whole journey to sink. The reduction of network delay is realized by using longer routes and more data copies for nodes in the area far from the sink. In this way, the network delay and network lifetime could be optimized with the guarantee of network reliability.

Theoretical analysis on energy consumption is offered and presented. And the trade-off among parameters of GRDR is formulated as a multi-objective optimization problem under the reliability constraint. A corresponding trade-off method is also given.

Comprehensive simulation experiments are conducted to verify the effectiveness of the GRDR scheme. The results show that the proposed GRDR could obtain the goal of optimization of network lifetime and end-to-end delay. In the simulation, the maximum network lifetime could be improved by 20% and the end-to-end delay decreased by 15.62% simultaneously compared with PR, which presents the superiority of the strategy.

The rest of this paper is organized as follows: Section 2 reviews the related work. The system model is described in Section 3. The novel GRDR scheme is presented in Section 4 including parameters optimization. Performance evaluations through simulations are presented in Section 5. We conclude in Section 6.

2 Related Work

A great deal of research efforts have been devoted to this field to provide a reliable transmission service in WSNs because of the unreliable links. Liu et al. [1] point out that the existing works mainly fall into two categories: packet-loss avoidance and packet-loss recovery. Packet-loss avoidance (e.g. [9-10]) attempted to reduce the occurrence of packet loss and packet-loss recovery tried to recover the packet loss when it happened. Because packet-loss avoidance methods need to pay a high price and from the cost consideration the mechanism widely used in network is based on packet-loss recovery. While the reliable protocol based on packet-loss recovery can be dived into two categories. One way is the retransmission after packets loss scheme, whose representative protocol is Automatic Repeat-reQuest (ARQ). Another way is the packets reproduction strategy. In this mechanism, it can not guarantee the reliability for WSNs to exploit a single path routing. So the way of multiple path routing is often employed. Network delay has an important influence on WSNs applications. So there is a great deal of research in this field. Jechan Han et al conduct analysis on WSNs transport delay of NAK-based SR-ARQ [17]. There is also much research on SW E2E and SW H2H, e.g. Ref. [15, 17]. However, it is worth attention that most of the existing studies do research on simple linear network. And analysis on network delay of the flat networks which are widely applied in the real world is still relatively rare. In addition, network lifetime is hardly considered in most of the studies.

Moreover, recently, for flat network, Chilukuri Shanti and Sahoo [18] presented a new contention-free TDMA-based integrated MAC and routing protocol named DGRAM. Considering the unique phenomenon of data aggregation in WSNs, Huang et al. [19] proposed a centralized scheduling algorithm with the delay bound of $23R+\Delta+18$ time slots, where R is the network radius and Δ is the maximum node degree. Yu et al. [20] proposed a distributed scheduling method generating collision-free schedules with delay at most $24D+6\Delta+16$ time slots, where D is the network diameter. Xu et al. [21] theoretically prove that the delay of the aggregation schedule generated by their algorithm is at most $16R+\Delta-14$ time slots.

Network lifetime is an important performance metric and needs to be considered in all WSNs studies. There is also a great deal of research in the field, e.g. [13, 15-16]. However, as far as we know, most of the research only focuses on one performance metric of WSNs because of the network complexity and research difficulty, e.g. reliability [5, 8-9], network delay [17-21]. Some research integrates network delay and reliability, e.g. [1], or integrates network lifetime and delay, e.g. [22-24]. But there is few research taking network reliability, delay as well as lifetime into consideration, which is the main focus of the paper.

3 Problem Formulation

3.1 The System Model

(1) Consider a wireless sensor network consisting of sensor nodes that are deployed randomly (following uniform distribution) in a flat network with the radius

of R, the area of W and the node density of ρ. And nodes do not move after being deployed. On detecting an event, a sensor node will generate a message and forward to sink.

(2) The sensor nodes are assumed to know their relative locations and the sink location. We also assume that each sensor node has the knowledge of its adjacent neighboring nodes. The information about the relative location of the sensor domain may also be broadcasted through this network for routing information update.

3.2 Energy Consumption Model

Following the typical energy consumption model in [25], the expected energy cost for transmission is as

$$\begin{cases} E_t = lE_{elec} + l\varepsilon_{fs}d^2, & if \quad d < d_0 \\ E_t = lE_{elec} + l\varepsilon_{amp}d^4, & if \quad d \geq d_0 \end{cases} \quad \textbf{(1)}$$

And the energy consumption for receiving packet is computed by

$$E_r(l) = lE_{elec} \quad \textbf{(2)}$$

Where, E_{elec} denotes the transmitting circuit loss energy and d_0 denotes the threshold. The parameters ε_{fs} and ε_{amp} respectively represent the energy required by power amplifier. And l indicates the packet length (bits). Following parameters in [25], the network energy consumption parameters and the corresponding values are shown in Table 1.

Table 1. Parameters and corresponding values

Parameter	Value
threshold distance (d_0) (m)	87
sensing range r_s (m)	15
E_{elec} (nJ/bit)	50
ε_{fs} (pJ/bit/m^2)	10
ε_{amp} (pJ/bit/m^4)	0.0013
initial energy (J)	0.5

3.3 Problem Statement

Definition 1. End-to-endt Delay: it is defined as the time from a packet's first transmission until its successful arrival at the sink [11]. Let Γ_i denote the end-to-end delay from node v_i to sink.

Definition 2. Network Lifetime: it is defined as the first node death time and denoted by T.

Definition 3. Network Reliability: it represents the statistical probability of packet successfully forwarded to sink from one node in QoS level. Let δ_i denote the network reliability of packets transmitted from node v_i to sink. Each sensory data is received by the sink with a probability not less than δ.

Definition 4. Packet Lifetime: it is defined as the number of steps needs to traverse before reproduction and represents the route length. It is also called packet TTL.

Following the network model in [13], assume the transmission reliability of node v_i is ε_i and also there are h hops from the source node v_i to the sink. Let P_i denote the routing path from node v_i to the sink, where P_i ={ v_0^i , v_1^i , ..., v_h^i }, v_h^i denotes the node whose distance to the sink is h hops in the routing path of v_i. The reliability and the delay at each hop are denoted by ε_i^T = [ε_0 , ε_1 , ..., ε_h] and τ_i^T = [τ_0 , τ_1 , ..., τ_h], respectively. Therefore, the network reliability and end-to-end delay of packet transmitted to sink from node v_i are respectively $\delta_i = \left[1 - \prod_{k=0}^{h}(1-\varepsilon_k)\right]$ and $\Gamma_i = \sum_{k=0}^{h}\tau_k$. The focus of the paper is to meet the requirements that after passing through several relays, the packet should still satisfy targeted end-to-end reliability δ. That is $\delta_i \geq \delta$. And in the mean time, the network lifetime is maximized and end-to-end delay is minimized. To sum up the above, the network optimization goal can be expressed as the following, that is, as for any node v_i in the network it meets the following formula.

$$\begin{cases} \max(T) = \min \max_{0<i\leq n}(e_i), \min(\Gamma) = \min \max_{0<i\leq n}(\Gamma_i) \\ e_i = S_i^1 E_t + S_i^2 E_r, \Gamma_i = \sum_{k=0}^{h}\tau_k \\ s.t. \quad \delta_i = 1 - \prod_{k=0}^{h}(1-\varepsilon_k) \geq \delta \end{cases} \quad \textbf{(3)}$$

Where, S_i^1 and S_i^2 respectively represent the amount of sending and receiving data of node v_i. E_t and E_r denote the energy cost for transmitting and receiving one packet respectively. And e_i denotes the energy consumption of node v_i. Hence, the optimization goal is minimizing the largest node energy consumption and the maximum end-to-end delay under the guarantee of transport service quality, e.g. $\delta_i \geq \delta$.

4 GRDR Design

The proposed GRDR scheme consists of three phases: (1) packets reproduction, (2) direction dispersity and (3) multi-route. As an illustration of the methods, these phases will be presented in the following two sections. And then the trade-off among parameters is analyzed and discussed. At the end of the section, how to optimize parameters to achieve a trade-off between network delay and network lifetime is presented.

4.1 The Overall Approach

In this section, the overall approach is described. An example is illustrated in Figure 1. The figure depicts the reproduction procedure of a source node S with the distance of d to sink sending its initiated packets toward sink. During the process, the source node S initiates several data copies firstly. Then it disperses the packets over the network and the packets traverse the network by several steps. In the following phase, packets follow the shortest path routing to complete the rest journey. Because of the unreliable link or the bad link propagation, some packets are lost in the journey. So when a packet successfully finishes its lifetime it begins to reproduce. So the proposed scheme integrates three components: (1) packets reproduction; (2) direction dispersity; (3) multiple path routing. The whole transmission from the source node to the destination sink consists of many of the three periods.

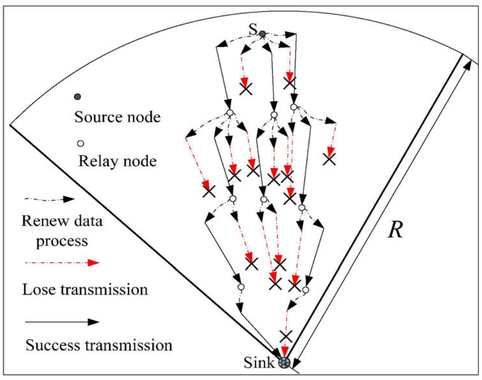


Figure 1. Illustration of the GRDR scheme

Packets reproduction. In order to ensure transmission reliability, the source node generally initiates M copies and sends them toward destination of sink. Some packets may fail during the transmission because of unreliable link. So, the λ copies of packets could be reduced to $\lambda^{'}$ copies after several steps and that is $\lambda^{'} < \lambda$. To compensate for the lost packets, it will reproduce several new copies after a certain number of steps, e.g. the number of new copies is N. Then $N + \lambda^{'}$ data copies go on forwarding to the destination.
Direction dispersity. The objective of dispersity is to disperse the reproduced copies of packets around the source node. The data copies are dispersed by direction dispersity and sent to the destination nodes, which are called Intermediate nodes I.

Multiple path routing. When data packets are dispersed to Intermediate nodes I, they will finish their rest journey by the traditional multi-path routing schemes. In this paper, the shortest path routing is adopted in the period.

4.2 GRDR Scheme

In the part, more details on the implementation are presented.
Each initiated packet carries the information (ID, TTL, DELAY), where ID describes the information associated with the packet. The packet carries the value of *TTL* to record the time to live. And DELAY denotes the transmission delay. In the transmission, the packet needs time to finish one hop journey or packet reproduction. The total time needed for the transmission and reproduction is recorded to make the final statistics. Each node knows the hop distance to sink and also knows those of its neighbors.

Assume that the number of renewed packets is λ, the routing length or the packet TTL is ϑ hops, the node reliability is ε and the end-to-end delay of each hop is τ. So after ϑ steps, the reliability of one packet is ε^{ϑ}. There are λ data copies or routing paths. So the total reliability δ is shown as the formula (4). According to the above relation among the network reliability δ, node reliability ε, data copies λ and route length ϑ, the number of data copies is calculated by formula (5). From the Figure 2, it can be seen that when the lifetime of packet generated by source or relay nodes increases the corresponding data copies should be increased to ensure the transmission reliability.

$$\delta = 1\text{-}(1\text{-}\varepsilon^{\vartheta})^{\lambda} \quad \textbf{(4)}$$

$$\lambda = \frac{\log(1-\delta)}{\log(1-\varepsilon^{\vartheta})} \quad \textbf{(5)}$$

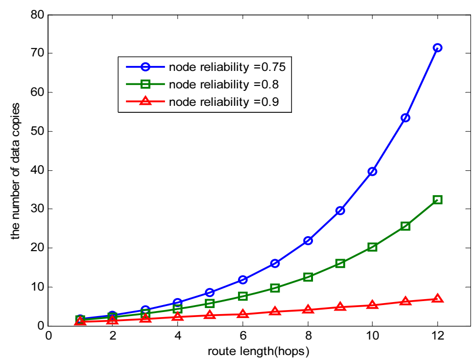


Figure 2. Relation between number of data copies and packet lifetime under the given node and network reliability

Direction dispersity. Let the distance from one source node s to sink is h hops. And the disperse range is up to X_h hops. The objective of direction dispersity is to transmit the renewed data copies to λ nodes of the kind of Intermediate nodes I and with the distance uniformly distributed in [h - X_h , $h+X_h$].

In the GRDR scheme and during the dispersity implementation process, X_h denotes the disperse distance to source node along the direction of X in the first stage. And Y_h represents the disperse distance along the direction of Y in the second stage. And in the first stage, X_h subtracts one after the packet is transmitted one hop along the direction of X. When X_h =0, the first dispersity stage is finished. In the second stage, dispersity is done along the direction of Y and it is the same as the first stage.

Here, more details about the dispersity are given. One value from [- X_h , X_h] is randomly selected for each data packet. If we assume the randomly generated number is Δ and $\Delta \in \{-X_h, X_h\}$, it represents that the packet will be dispersed to one destination node with distance of $h+\Delta$ to sink. In order to finish the first stage of dispersity, the safety is not high if the packet is only dispersed along the X direction. Therefore, it needs to be dispersed along the X direction and Y direction simultaneously as illustrated in Figure 3. In the Figure 3, there are four data packets at the end of dispersity dispersed to four relay nodes, which are Intermediate nodes I of S1, S2, S3 and S4. They are randomly distributed around the source node and with distance to sink within the range of [h - X_h , $h+X_h$].The following method is exploited to achieve this goal. Let the number of packet copies is λ . Then randomly select λ angles from 360 degree, which respectively represents the disperse direction. If we assume one selected angle is θ_1 , it means that the packet is transmitted Y_h = $\tan\theta_1$ hops along the Y direction whenever it is forwarded by one hop along X direction, which is illustrated by the red routing path in Figure 3.

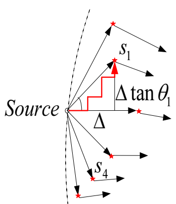


Figure 3. Illustration of packet dispersed to Intermediate nodes I

In the Figure 3, it shows the disperse process in the case of dispersing angle $0 \le \theta \le 90^0$. With regard to $90^0 \le \theta \le 180^0$, the disperse process is the packet forwarded by $\tan(\theta - 90^0)$ hops along Y direction whenever it is transmitted one hop along the opposite X direction to sink. For $180^0 \le \theta \le 270^0$ and $270^0 \le \theta \le 360^0$, they are the similar as $90^0 \le \theta \le 180^0$ except along the left or right X direction.

Shortest path routing. After packet dispersed to Intermediate nodes I, the typical shortest path routing is implemented to transmit the packet to the destination sink or the next reproducer.

Packet reproduction. In the transmission, the packet may be damaged or lost because of various reasons. So after certain steps, the packet needs to be reproduced to guarantee the reliability. Each packet carries the value of TTL and when a packet successfully finishes its lifetime, e.g. TTL =0, it arrives at a reproducer and begins to reproduce. The reproducer is called relay node. And the packet reproduction is named packet relay. Indeed, the reproduction is triggered by the packets not nodes. The packet lifetime is the metric of hops between two adjacent reproducers. After packet reproduction, the direction dispersity is triggered.

In the GRDR scheme, the number of generated copies is different according to the source or relay node's distance to sink. If the source or relay node has longer distance to sink, it would have more data copies and the packet carries a longer TTL, which means the packet has longer path route before reproduction. Otherwise, there are less data copies and shorter packet TTL. It needs to pay attention that if the relay or source node has one or less than one hop distance to sink, the packet does not need to be copied and transmitted to sink directly.

Each node knows the hop distance to sink and also knows those of its neighbors. Under the given network reliability δ , the node reliability ε and the network partition, the GRDR protocol is implemented as shown in Algorithm 1 in detail.

4.3 The Parameter Optimization of GRDR

The key issue in GRDR scheme is how to select optimal parameters to reduce network delay and prolong network lifetime under the guarantee of transmission reliability. In this part, the parameter optimization issue is discussed.

4.3.1 The Energy Consumption and Network Lifetime

In this part, the number of packets that each node needs to transmit in one round of data gathering is discussed. If the network segment width or packet lifetime ϑ is fixed, the number of packet generated by source or relay node is λ . According to the computation result of the packet number of each node needs to transmit in Ref. [17-19], the amount of data transmitted by the node with distance d to sink is approximately equal to the result in [17-19] multiplied by λ (Actually, it should be smaller than this value).

Algorithm 1: A green, reliability and delay based algorithm for WSNs

Algorithm1: A green, reliability and delay based algorithm for WSNs
Input: the network structure, the network reliability δ, the node reliability ε, and packet TTL ϑ
Output: transport routes from a source node to sink
Stage 1: Initialize and create packet
  (1) If the distance of source node to sink d <=1 hop
      | the number of data copies λ = 1 and is directly sent to sink;
    Else
      | The value of packet TTL is obtained according to d;
      | The number of data copies λ is computed following formula (5)
    End
Stage 2: Packets dispersity
  (2) For k =1: λ
      | One packet is generated;
      | Packet is dispersed to its intermediate nodes;
      | If the packet is not lost
      | | The value of packet TTL is reduced by 1;
      | End
    End
Stage 3: Shortest path routing
  (3) For each packet in intermediate nodes
      | If the distance to sink <=1 hop
      | | The packet is transported to sink directly;
      | Else
      | | It is further routed to new location following the shortest path routing;
      | | If the packet is not lost
      | | | The value of packet TTL is reduced by 1;
      | | End
      | End
    End
Stage 4: Packets reproduction
  (4) For each packet in the new location
      | If the packet current TTL is more than 0 hop
      | | Goto Stage 3: shortest path routing
      | Else
      | | If the distance to sink <=1 hop
      | | | The packet is transported to sink directly;
      | | Else
      | | | Packet's TTL τ is obtained according to the location;
      | | | The number of reproduced data copies λ is computed following formula (5);
      | | | Goto Stage 2: dispersity stage;
      | | End
      | End
    End
  (5) Output results;

So, the energy cost is different. There are more energy cost in hotspots and less energy consumption in non-hotspots. Therefore, the characteristic could be fully exploited to generate more data copies and make them have longer path journey or longer lifetime to reduce network delay without influence on network lifetime. In order to facilitate the analysis and computation, the flat network is partitioned into rings and the ring width is with proportional change. The proportional coefficient is denoted by q. In this case, we compute the amount of packets transmitted by one node with distance d to sink.

Theorem 1. Assume the transmission range of sensor is r and the network circular width or the packet lifetime ϑ has the proportional change with the proportional coefficient q. And each node finds its next forwarding node by the shortest path routing. Assume that the distance of one node to sink is d and $d = hr + x$. Then the amount of packet transmitted by the node is as the following.

$$\mu_d = \frac{\sum_{m=0}^{k}(d+m\cdot r)\cdot\lambda_m}{d} = \sum_{m=0}^{k}(1+\frac{mr}{d})\cdot\lambda_m \quad \textbf{(6)}$$

Where, λ_m denotes the number of data copies generated by source or relay nodes in different network circulars. And it is calculated by the following formula (7).

$$\lambda_m = \begin{cases} 1 & d+m\cdot r \le r \\ \log(1-\delta) \big/ \log(1-\varepsilon^{\vartheta_m}) & q^n \cdot r < d+m\cdot r \le q^{n+1}\cdot r \end{cases} \quad \textbf{(7)}$$

Proof. When the network partition width or the packet lifetime ϑ has the proportional change and based on the Theorem 1 in Ref. [17-19], the number of packets μ_d needs to be transmitted by node with the distance d to sink is as the following.

$$\mu_d = \frac{d\cdot\partial_x\cdot\rho\cdot\lambda + \sum_{m=1}^{k}(d+m\cdot r)\cdot\partial_x\cdot\rho\cdot\lambda_m}{d\cdot\partial_x\cdot\rho}$$
$$= \frac{\sum_{m=0}^{k}(d+m\cdot r)\cdot\partial_x\cdot\rho\cdot\lambda_m}{d\cdot\partial_x\cdot\rho} = \frac{\sum_{m=0}^{k}(d+m\cdot r)\cdot\lambda_m}{d} \quad \textbf{(8)}$$
$$= \sum_{m=0}^{k}(1+\frac{mr}{d})\cdot\lambda_m$$

Where, λ_m denotes the number of data copies generated by source or relay nodes in different network circulars. According to the proposed GRDR scheme, λ_m =1 when $d+m\cdot r \le r$. If $q^n\cdot r < d+m\cdot r \le q^{n+1}\cdot r$ and according to formula (5), the number of data copies initiated or reproduced by source nodes or relay nodes is calculated as the following.

$$\lambda_m = \log(1-\delta) \Big/ \log(1-\varepsilon^{\vartheta_m}) \quad \textbf{(9)}$$

Therefore, the final result is obtained:

$$\mu_d = \sum_{m=0}^{k}(1+\frac{mr}{d})\cdot\lambda_m,$$

$$where \quad \lambda_m = \begin{cases} 1 & d+m\cdot r \le r \\ \log(1-\delta)\Big/\log(1-\varepsilon^{\vartheta_m}) & q^n\cdot r < d+m\cdot r \le q^{n+1}\cdot r \end{cases} \quad \textbf{(10)}$$

From the above, it can be seen that the longer packet lifetime or routing journey in the region far from sink would cause more data copies for source or relay nodes in the area in order to ensure the transmission reliability, which leads to more energy consumption for network peripheral nodes. But, it does not necessarily lead to the increase of number of packets forwarded by nodes in hotspots.

4.3.2 The Relation between the Number of Renew Times and Delay

In this part, the relation between the number of relay times and network delay is discussed. In order to facilitate the analysis, the network is partitioned into circular rings and the circular width is with proportional change. Accordingly, the lifetime or the routing steps of the packet generated by source or relay node in different rings are with the corresponding proportional change. Let q represents the proportional coefficient. An example is illustrated in Figure 4. As can be seen from the Figure 4, the source or relay node in far region from sink has more data copies and longer packet lifetime than that in area close to sink.

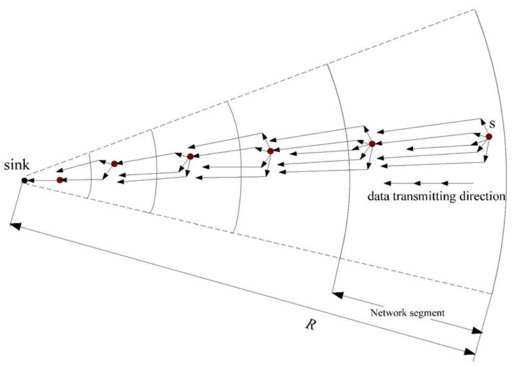


Figure 4. Network partition with proportional change

Theorem 2. Consider a network with the radius of R and the number of rings is k under the identical width partition. If the network circular width has the proportional change and the proportional coefficient is q, the number of relay or renew times in the whole transmission is calculated by

$$(k-x)\lg q = \lg(1+k(q-1))$$

$$|\text{where } x \ge 0 \text{ and is integer (the same below)} \quad \textbf{(11)}$$

Proof. If the network circular width is identical and the width is η hops, the number of rings is k and the total length of network radius is κ hops, that is $R=\kappa r$, we can get:

$$k\eta = \kappa \quad \textbf{(12)}$$

If the network circular width is with proportional change and the proportional coefficient is q (as illustrated in Figure 4), we have the following:

$$\eta + \eta q + \eta q^2 + ... + \eta q^{k-x} = k\eta = \kappa$$

$$\text{(where } x \ge 0 \text{ and it is an integer)} \quad \textbf{(13)}$$

Further simplification, $\eta\frac{q^{k-x}-1}{q-1} = k\eta = \kappa$ is derived.

That is $\frac{q^{k-x}-1}{q-1} = k$. So it is $q^{k-x} = 1+k(q-1)$.

Therefore, the following conclusion can be obtained.

$$(k-x)\lg q = \lg(1+k(q-1)) \quad \textbf{(14)}$$

It can be seen that when the network circular rings are with the proportional change, the number of relay times is reduced by one at least compared with the original case. So in the whole transmission from source node to sink, the number of packet reproduction times is reduced. And it means that the network delay is reduced because packet reproduction needs more time than that required for packet transmission in the whole journey to sink. We can reduce network delay by reducing the number of renew times. Therefore, in order to reduce network delay, we should reduce the number of packet reproduction times and prolong the packet routing path or lifetime. However, to ensure the transmission reliability we should increase the routing path number, which means increasing data copies. And it may lead to more energy cost.

4.3.3 An Approach for Parameter Optimization of GRDR

The increase of packet lifetime could lead to the change of number of relay times in the whole transmission to sink. And the change of number of relay times to sink may have obvious influence on network delay. At the same time, the increase of packet lifetime causes the increase of data copies, which leads to more energy consumption and may affect the network lifetime. So, the key issue is how to select the parameters of GRDR to realize a trade-off between network delay and lifetime under the guarantee of network reliability.

The network optimization goal in the paper is expressed as the following,

$$\begin{cases} \xi = \min(\alpha * \max_{0<i\le n}(e_i) + \beta * \max_{0<i\le n}(\Gamma_i)) \\ e_i = S_i^1 E_t + S_i^2 E_r, \Gamma_i = \sum_{k=0}^{h} \tau_k \\ s.t. \delta_i = 1 - \prod_{k=0}^{h} (1-\varepsilon_k) \ge \delta \end{cases} \quad \textbf{(15)}$$

Where, α and β are the proportional coefficient and the sum is equal to 1, e.g. α =0.6, β =0.4. In Formula (15), all other parameters are the same as in Formula (3).The network delay and energy cost is normalized to dimensionless metric. And then the weighting method is exploited to assign them the different weights.

In order to facilitate the analysis, the network is also partitioned into rings whose widths are with proportional change according to the distance to sink. Let q denote the proportional coefficient. In the case, the optimization is to select the coefficient q from sets of coefficient $\{q_s,...\}$ satisfying the constraint of formula (15). The enumeration method is applied to find the optimal q to ensure the optimization goal.

5 Simulation Results

In this section, the simulation results are presented to evaluate the GRDR scheme. The simulation scenario is described firstly, focusing more on the network model applied in the simulation experiments. Then the evaluation results are presented to reflect the impact of control parameters on the routing performance. Finally, the comparison with existing mechanism is demonstrated.

5.1 Evaluation on the Proposed Scheme

The performance evaluation mainly focuses on three metrics including the reliability, maximum energy cost and maximum end-to-end delay. For one round of data gathering, we assume the source data packet of each node is generated synchronously and each sensory data need to be forwarded to sink following the same procedure and routing paths. Therefore, we need only compare the proposed GRDR scheme with PR scheme in one round of data gathering, and the conclusion holds for the whole running time of the network.

Consider a wireless sensor network consisting of N =900 sensor nodes that are uniformly and randomly scattered in a flat circle with the radius of R =400. The sensing and transmitting range are all r =50 meters. The packet size is fixed 100 bits. And assume the time for one hop transmission is 1 ms and for packet reproduction is 3 ms. All the nodes except sink have the identical initial energy of 1J. The node and network reliability are the same value 0.85. Other parameters of the network are as shown in table 1. As illustrated in above sections, if the distance from source or relay nodes to sink is just one or less than one hop, the packets are transmitted directly to sink without packet reproduction. In order to facilitate the issue, the network is partitioned into circular rings. And the rings' width is with proportional change and assuming the coefficient value is an integer. That means packets generated by nodes in different segments have different lifetime. The data copies are calculated according to the packet TTL by formula (4). Table 2 and Table 3 respectively show the relation among the distance of source or relay nodes to sink, the packet lifetime and the number of data copies under the cases of network segment length with proportional changes of q = 2 and q = 4.

Table 2. Network segment length with proportional change of q = 2 for GRDR

distance to sink (m)	0-50	50-150	150-350	350-400
Packet TTL(hop)	1	2	4	5
number of copies	1	3	5	6

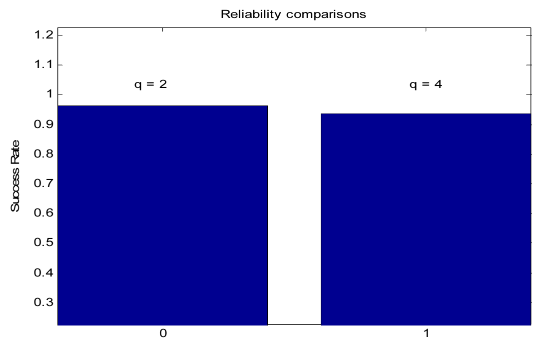
Table 3. Network segment length with proportional change of q = 4 for GRDR

distance to sink (m)	0-50	50-250	250-400
packet TTL(hop)	1	4	5
number of copies	1	4	6

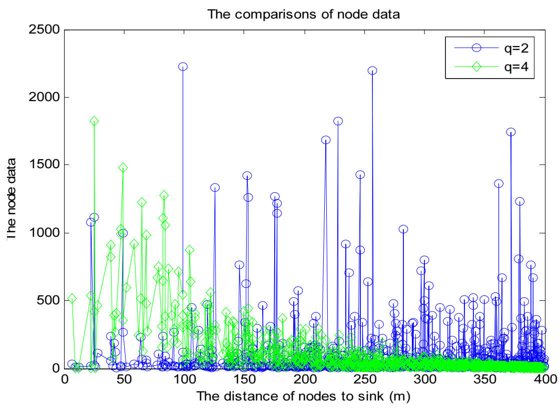
We do simulation on metrics of reliability, energy cost and delay under the two cases of network circular width with proportional changes of q = 2 and q = 4. The simulation results are shown in Figure 5.

Figure 5(a) presents the network transmission success rates under the corresponding proportional coefficient of q =2 and q =4 after one round of data gathering. They meet the network statistical reliability requirement. The number of packets transmitted by each node is shown in Figure 5(b). As can be seen from the figure, there are more data packets to be forwarded for nodes in the area close to sink as a whole. Figure 5(c) illustrates the energy cost of each node after one round of data gathering. In the simulation, it needs energy cost only in sending and receiving data packets. So it is with the similar variable trend as the forwarding amount of data packets. The end-to-end delay when one round of data gathering finishes is shown in Figure 5(d). Overall, the delay is longer when the distance to sink becomes much far. From the Figure 5(d), it can be seen that the delay is reduced when the network partition proportional coefficient is becoming larger. This is because that the number of relay times is reduced and packet reproduction takes more time than the transmission.

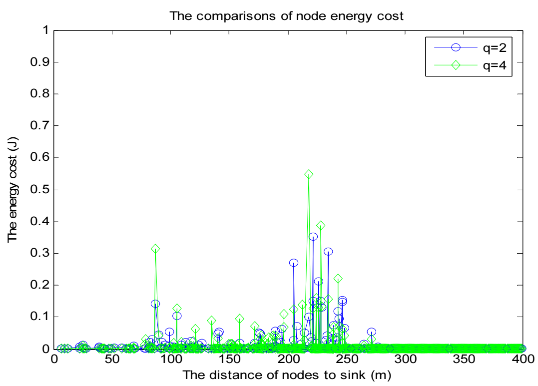
As can be seen from the Figure 5, the bigger proportional coefficient leads to smaller number of packet reproduction times in the journey of source node to the destination sink. And the end-to-end delay is reduced by reducing the number of relay times, which, however, may result in additional energy cost.



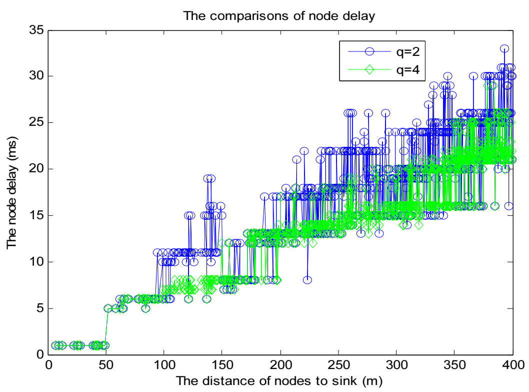
(a) reliability



(b) data load of each node



(c) energy cost

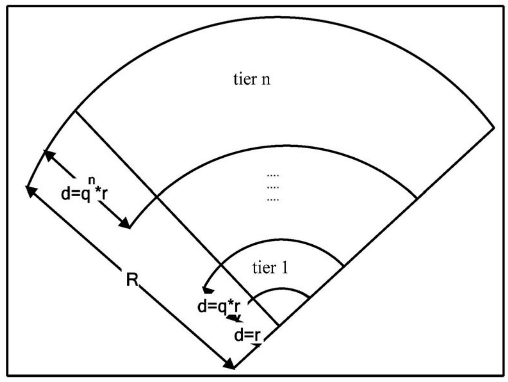


(d) end-to-end delay

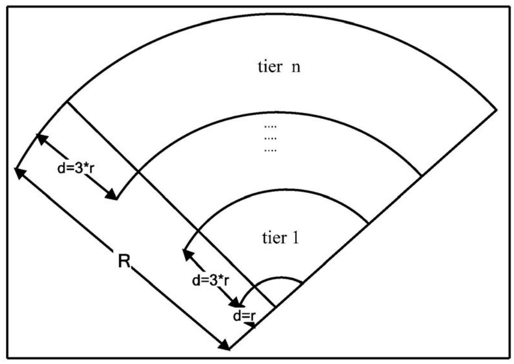
Figure 5. Simulation results for the algorithm GRDR

5.2 Comparison with PR

In this part, the proposed GRDR routing is compared with the existing PR scheme through the above mentioned flat circle network, focusing on energy cost performance. The parameters of linear and flat networks are the same as description in section 5.1. For GRDR strategy, the network circular width change is set with the proportional coefficient q =2, e.g. the circular width $d=\{r,qr,q^2r,...\}$. In order to facilitate the analysis, we assume the identical circular width of the partition in PR scheme is an integral multiply of r. Accordingly, in order to have the same number of circulars, the network circular width is fixed to $3*r$ in the PR routing. They are illustrated in Figure 6.



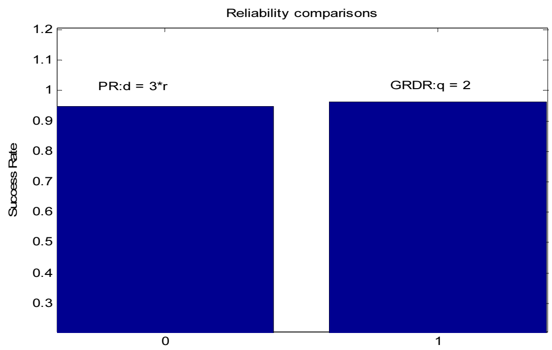
(a)



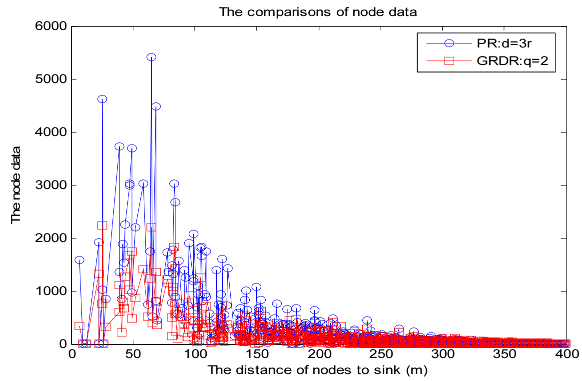
(b)

Figure 6. Illustration of the network partition (a) GRDR (b) PR

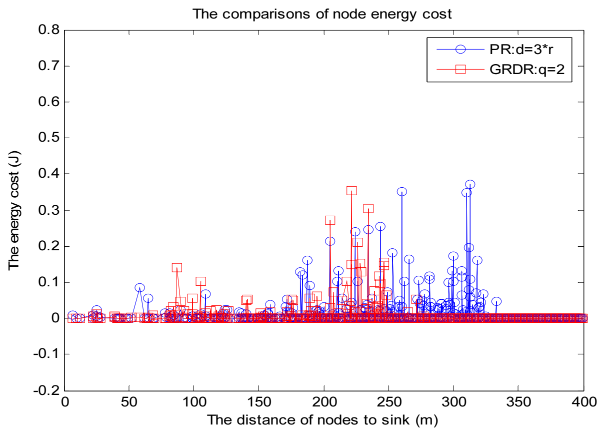
Figure 7(a) are the comparisons of reliability between GRDR and PR for flat circle network when one round of data gathering finishes. From the results, it can be seen that there are not obvious differences of reliability. The comparisons of the amount of transmitting data of each node are shown in Figure 7(b). On the whole, there are more data packets for nodes close to sink. Compared with PR routing, in the proposed GRDR scheme there are less amount of transmitting packets for nodes close to sink. This is attributed to the core idea of GRDR. It has more data copies for source or relay nodes far from sink to fully



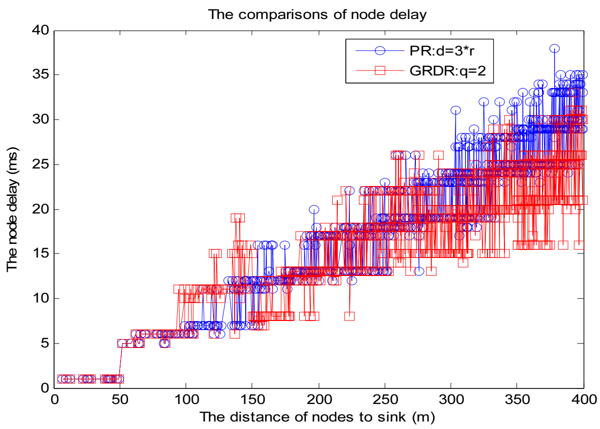
(a) reliability



(b) data load of each node



(c) energy cost



(d) end-to-end delay

Figure 7. Comparisons of performances between GRDR and PR

exploit the excessive energy and less data copies for nodes in the region close to sink to reduce the total amount of transmitting data packets. Accordingly, the packet lifetime is longer in the region far from sink than that in area close to sink.

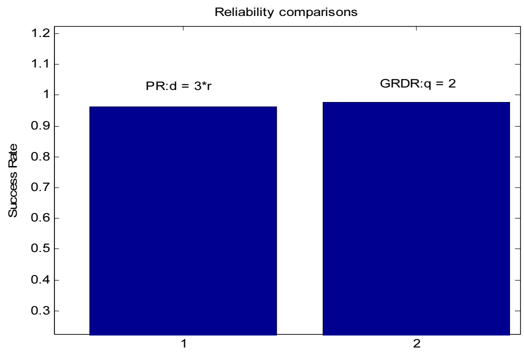
When one round of data gathering finishes, the energy cost for GRDR and PR in flat circle network are presented in Figure 7(c). From the figures, it can be seen that there is less energy cost for GRDR compared with PR. The maximum energy cost is reduced by 5.56% (see Figure 7(c)).

Figure 7(d) shows the end-to-end delay comparison between the proposed GRDR and the existing PR scheme after one round of data gathering of flat circle network. It is evident that the end-to-end delay changes alternately for nodes with different distances to sink. But for the maximum end-to-end delay, it is smaller in GRDR than that in PR. The maximum network delay could be reduced by 15.79%. This is because that the packet lifetime in GRDR is longer than that in PR and there is less number of relay times in the whole transmission to sink compared with PR.

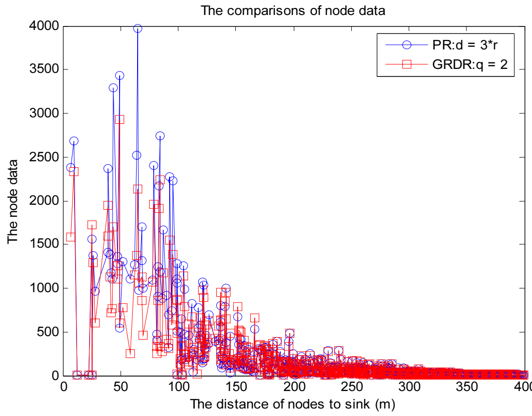
Figure 8, Figure 9 and Figure 10 respectively present the compared result between GRDR and PR when changing the communication range r, the number of nodes N and network area size R. Figure 8 demonstrates the comparisons of performances when r = 60, R = 400, N = 900. And Figure 9 demonstrates the comparison results when r = 60, R = 400, N = 800. When r = 50, R = 400 and N = 800, the comparison results are described in Figure 10. As shown in the figures, though the comparison of statistical reliability, data load and energy cost of each node, and end-to-end delay of each node are described, the performance metrics including reliability, maximum energy consumption and E2E delay of the network are obviously demonstrated. More importantly, as can be seen from the figures, it holds the same conclusion as that from Figure 7.

5.3 Evaluating the Optimization Performance

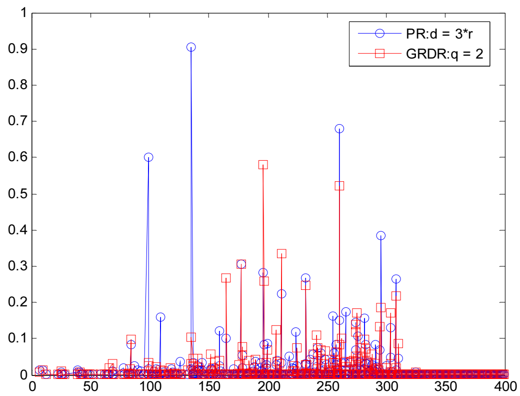
When the proportional coefficient q of network partition is 2, 3 and 4 respectively and if the parameters are set α =0.5, β =0.5 in the weighting method, the optimized q to satisfy the trade-off optimization constraint of formula (21) is discussed. After one round of data gathering, the statistical results of flat circle network when r = 50, R = 400, N = 900 are shown in table 4. It needs for dimensionless processing to eliminate the influence of index dimension because the energy cost and end-to-end delay have different dimensions, forms and effects on the total goal. Here, the proportion method is exploited to do index dimensionless processing.



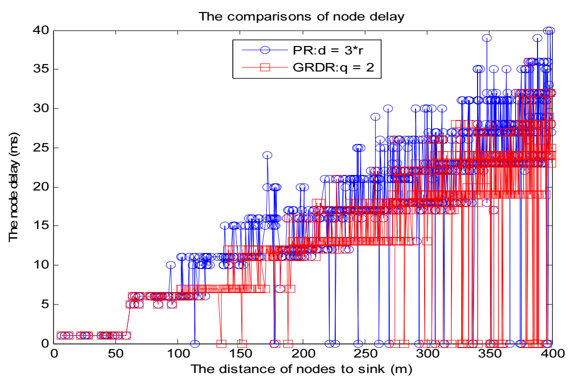
(a) reliability



(b) data load of each node

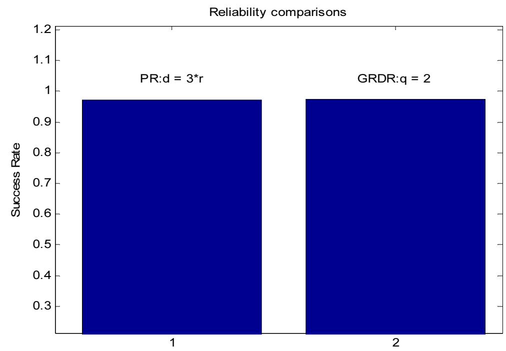


(c) energy cost

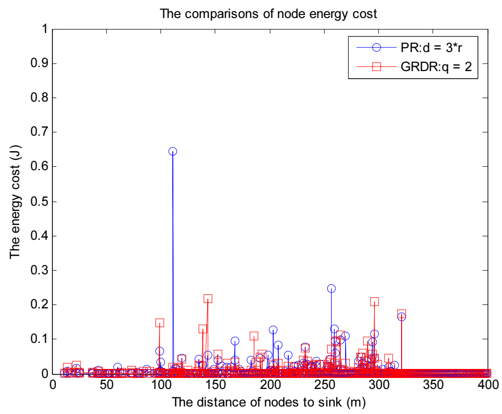


(d) end-to-end delay When $r = 60$, $R = 400$, $N = 900$(a)

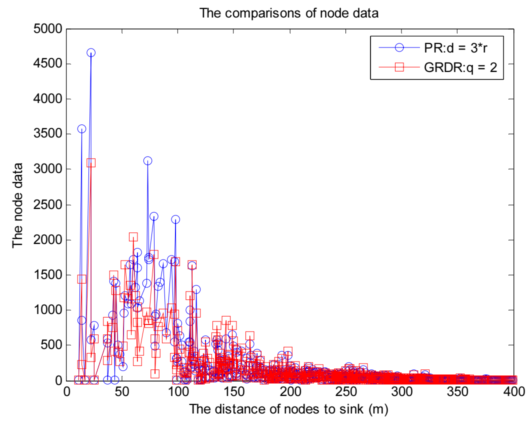
Figure 8. Comparisons of performances between GRDR and PR



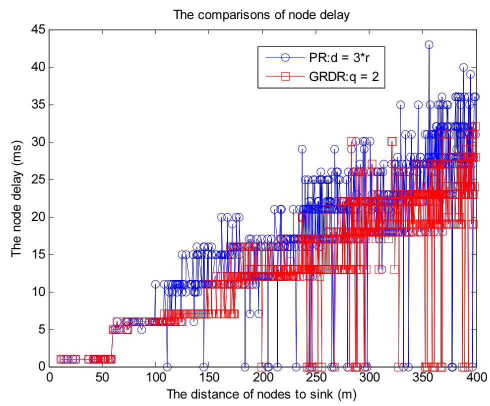
(a) reliability



(b) data load of each node

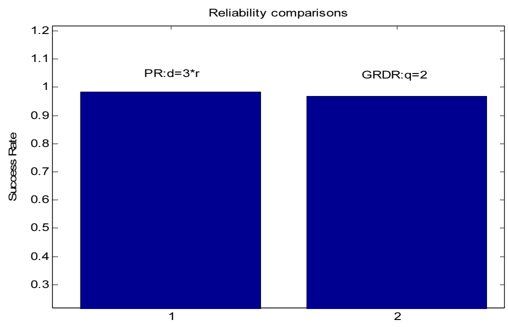


(c) energy cost

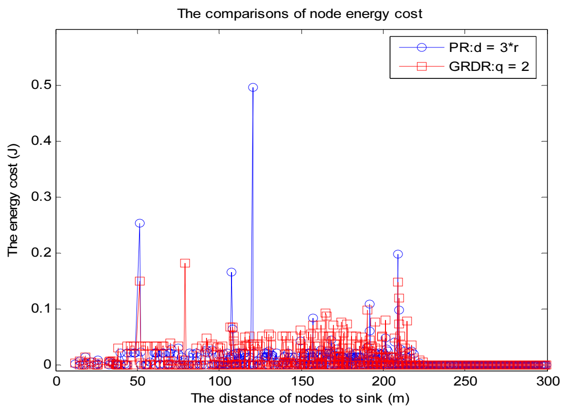


(d) end-to-end delay When $r = 60$, $R = 400$, $N = 800$

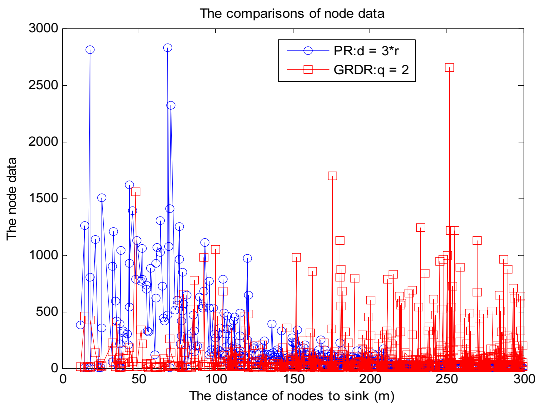
Figure 9. Comparisons of performances between GRDR and PR



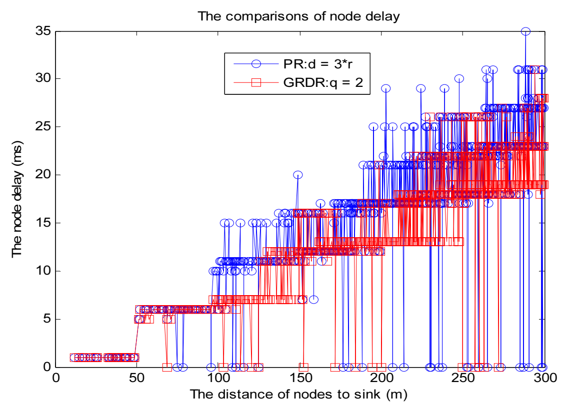
(a) reliability



(b) data load of each node



(c) energy cost



(d) end-to-end delay When $r = 50,\ R = 300,\ N = 900$

Figure 10. Comparisons of performances between GRDR and PR

From Table 4, when q =2 it obtains the optimal target value for flat circle network. The results further indicate that the end-to-end delay could be reduced by reducing the number of relay times, which does not necessarily lead to energy cost increase. Under the constraint of network reliability, we could select the optimized proportional coefficient q of network partition to obtain the optimization trade-off between network delay and energy consumption.

Table 4. Results for flat circle network

coefficient	The max energy cost (J)	The max end-to-end delay (ms)	The normalized target value
q =2	0.3537	33.0000	α *0.2381+ β *0.3474 =0.2928
q =3	0.5846	33.0000	α *0.3935+ β *0.3474 =0.3705
q =4	0.5473	29.0000	α *0.3684+ β *0.3053 =0.3368

6 Conclusion

In this paper, the problem of energy efficient (Green), data Reliability, end-to-end Delay (GRD) in WSNs is studied. Data collection should take energy efficient, reliability as well as delay into consideration. To address these problems, a novel reproduction packets routing scheme named GRDR is proposed. The GRDR consists of three phases. The theoretical analysis and simulation results show that the method has a better performance in terms of maximum energy consumption and E2E delay compared with PR. The optimized parameters can be obtained to ensure the reduction of end-to-end delay without reduction of the network lifetime. In the simulation, we can see that the network delay could be reduced by 15.79%. In many cases, the optimization of network lifetime and end-to-end delay could be obtained at the same time. In the simulation, the maximum network lifetime could be improved by 20% and the end-to-end delay be decreased by 15.62% simultaneously.

In addition, we formulate the optimization problem as to maximize lifetime under energy and given security constraints by controlling the system parameters. Carefully designed system parameters grant GRDR more energy efficient and lower end-to-end delay under given network reliability. Adaptive adjustment of these parameters is more preferable. All these issues are worth further research studies.

Acknowledgment

This work was supported in part by the Sub-Project of National Natural Science Foundation of China (U1201253, U08035003), National Natural Science

Foundation Emergency Management Project (M1450004), Ministry Education Foundation of China (MCM20121031), the National High Technology Research and Development Program of China (863 Program) (2012AA010105), National Science and Technology Major project (2012ZX09304-003-005) and the National Basic Research Program of China (973 Program) (2014CB046305).

References

[1] Y. Liu, Y. Zhu, L. M. Ni, G. Xue, A Reliability-Oriented Transmission Service in Wireless Sensor Networks, *IEEE Transactions on Parallel and Distributed Systems*, Vol. 22, No. 12, pp. 2100-2107, December, 2011.

[2] S. Boulfekhar, M. Benmohammed, A Novel Energy Efficient and Lifetime Maximization Routing Protocol in Wireless Sensor Networks, *Wireless Personal Communications*, Vol. 72, No. 2, pp. 1333-1349, September, 2013.

[3] M. R. Senouci, A. Mellouk, A. Aissani, Random Deployment of Wireless Sensor Networks: A Survey and Approach, *International Journal of Ad Hoc and Ubiquitous Computing*, Vol. 15, No. 1-3, pp. 133-146, August, 2014.

[4] L. Karim, N. Nasser, Reliable Location-aware Routing Protocol for Mobile Wireless Sensor Network, *IET Communications*, Vol. 6, No. 14, pp. 2149-2158, September, 2012.

[5] S. K. Bhoi, P. M. Khilar, Vehicular Communication: A Survey, *IET Networks*, Vol. 3, No. 3, pp. 204-217, September, 2014.

[6] N.-G. Kim, S.-J. Seok, A Sensor Link Control Protocol for Providing End-to-End Reliability in Sensor Networks, *Journal of Internet Technology*, Vol. 16 , No. 3, pp. 379-389, May, 2015.

[7] G. Han, J. Jiang, L. Shu, J. Niu, H.-C. Chao, Management and Applications of Trust in Wireless Sensor Networks: A Survey, *Journal of Computer and System Sciences*, Vol. 80, No. 3, pp. 602-617, May, 2014.

[8] W. R. Claycomb, D. Shin, A Novel Node Level Security Policy Framework for Wireless Sensor Networks, *Journal of Network and Computer Applications*, Vol. 34, No. 1, pp. 418-428, January, 2011.

[9] M. A. Mahmood, W. K. G. Seah, I . Welch, Reliability in Wireless Sensor Networks: A Survey and Challenges Ahead, *Computer Networks*, Vol. 79, pp. 166-187, March, 2015.

[10] T. Shu, M. Krunz, S. Liu, Secure Data Collection in Wireless Sensor Networks Using Randomized Dispersive Routes, *IEEE Transactions on Mobile Computing*, Vol. 9, No. 7, pp. 941-954, July, 2010.

[11] H. Yousefi, K. Mizanian, A. H. Jahangir, Modeling and Evaluating the Reliability of Cluster Based Wireless Sensor Networks, *Proceedings of the 24th IEEE International Conference on Advanced Information Networking and Applications (IEEE AINA'10)*, Perth, Australia, 2010, pp. 827-834.

[12] B. Rudra, O. P. Vyas, Exploration of Access Control Mechanisms for Service-oriented Network Architecture, *International Journal of Internet Protocol Technology*, Vol. 9, No. 1, pp. 1-11, January, 2015.

[13] L. S. Jiang, A. F. Liu, Y. L. Hu, Z. G. Chen, Lifetime Maximization through Dynamic Ring-Based Routing Scheme for Correlated Data Collecting in WSNs, *Computers & Electrical Engineering*, Vol. 41, No. 1, pp. 191-215, January, 2015.

[14] Y. Liu, A. Liu, Z. Chen, Analysis and Improvement of Send-and-Wait Automatic Repeat-reQuest protocols for Wireless Sensor Networks, *Wireless Personal Communications*, Vol. 81, No. 3, pp. 923-959, April, 2015.

[15] A. F. Liu, D. Y. Zhang, P. H. Zhang, G. H. Cui, Z. G. Chen, On Mitigating Hotspots to Maximize Network Lifetime in Multi-hop Wireless Sensor Network with Guaranteed Transport Delay and Reliability, *Peer-to-Peer Networking and Applications*, Vol. 7, No. 3, pp. 255-273, September, 2014.

[16] A. Liu, X. Jin, G. Cui, Z. Chen, Deployment Guidelines for Achieving Maximal Lifetime and Avoiding Energy Holes in Sensor Network, *Information Sciences*, Vol. 230, pp. 197-226, May, 2013.

[17] J. Han, J. Lee, Analysis Model for the Transport Delay of NAK-based SR-ARQ with a Finite Retransmission, *23rd International Technical Conference on Circuits/systems, Computers and Communications (ITC-CSCC2008)*, Shimonoseki, Japan, 2008, pp. 1709-1712.

[18] C. Shanti, A. Sahoo, DGRAM: A Delay Guaranteed Routing and MAC Protocol for Wireless Sensor Networks, *IEEE Transactions on Mobile Computing*, Vol. 9, No. 10, pp. 1407-1423, October, 2010.

[19] S. Huang, P. Wan, C. Vu, Y. Li, F. Yao, Nearly Constant Approximation for Data Aggregation Scheduling in Wireless Sensor Networks, *Proc. IEEE INFOCOM*, Barcelona, Spain, 2007, pp. 366-372.

[20] B. Yu, J. Li, Y. Li, Distributed Data Aggregation Scheduling in Wireless Sensor Networks, *Proc. of the 28th Conf. on Computer Communications (INFOCOM'09)*, Rio de Janeiro, Brazil, 2009, pp. 2159-2167.

[21] X. Xu, X.-Y. Li, X. Mao, S. Tang, S. Wang, A Delay-Efficient Algorithm for Data Aggregation in multihop Wireless Sensor Networks, *IEEE Transactions on Parallel and Distributed Systems*, Vol. 22, No. 1, pp. 163-175, January, 2011.

[22] A. Sinha, D. K. Lobiyal, A Multi-Leve l Strategy for Energy Efficient Data Aggregation in Wireless Sensor Networks, *Wireless Personal Communications*, Vol. 72, No. 2, pp. 1513-1531, September, 2013.

[23] P. Zhang, G. Xiao, H.-P. Tan, Clustering Algorithms for Maximizing the Lifetime of Wireless Sensor Networks with Energy-harvesting Sensors, *Computer Networks*, Vol. 57, No. 14, pp. 2689-2704, October, 2013.

[24] A. Abdulla, H. Nishiyama, J. Yang, N. Ansari, N. Kato, HYMN: A Novel Hybrid Multi-Hop Routing Algorithm to Improve the Longevity of WSNs, *IEEE Transactions on Wireless Communications*, Vol. 11, No. 7, pp. 2531-2541, July, 2012.

[25] P. Gupta, P. R. Kumar, The Capacity of Wireless Networks, *IEEE Trans. Information Theory*, Vol. 46, No. 2, pp. 388-404, March, 2000.

Biographies

Jinhuan Zhang received her Ph.D. degree in Communication and Information Systems from Wuhan University of Technology in 2009. Since July 2009, she has been an assistant professor at the School of information science and engineering, Central South University.

Jun Long is a professor in the School of information science and engineering, Central South University. He received the M.Sc. and Ph.D. degrees from Central South University, China. He is a visiting professor at the Columbia University in 2012-2013.

Anfeng Liu is a professor of School of Information Science and Engineering of Central South University, China. He is also a Member of China Computer Federation (CCF). He received the Ph.D. degrees from Central South University, China, 2005, in computer science. His major research interest is wireless sensor network.

Guihu Zhao received the B.Sc. and M.Sc. degrees in Computer Science from the Central South University, China, in 2007 and 2010 respectively. He is a Ph.D. Candidate at the Central South University. His major interests are multimedia applications, mobile networks, complexity theory, fractals and brain MRI image proccessing.