

RMER: Reliable and Energy-Efficient Data Collection for Large-Scale Wireless Sensor Networks

著者	DONG Mianxiong, OTA Kaoru, LIU Anfeng
journal or publication title	IEEE Internet of Things Journal
volume	3
number	4
page range	511-519
year	2016-01-12
URL	http://hdl.handle.net/10258/00009441

doi: info:doi/10.1109/JIOT.2016.2517405

RMER: Reliable and Energy-efficient Data Collection for Large-scale Wireless Sensor Networks

Mianxiong Dong, *Member, IEEE*, Kaoru Ota, *Member, IEEE*, and Anfeng Liu

Abstract— We propose a novel event data collection approach named RMER (Reliability and Multi-path Encounter Routing) for meeting reliability and energy efficiency requirements. The contributions of the RMER approach are the following: (a) Fewer monitor nodes are selected in hotspot areas that are close to the Sink, and more monitor nodes are selected in non-hotspot areas, which can lead to increased network lifetime and event detection reliability. (b) The RMER approach sends data to the Sink by converging multi-path routes of event monitoring nodes into a one-path route to aggregate data. Thus, energy consumption can be greatly reduced, thereby enabling further increased network lifetime. Both theoretical and experimental simulation results show that RMER applied to event detection outperforms other solutions. Our results clearly indicate that RMER increases energy efficiency by 51% and network lifetime by 23% over other solutions while guaranteeing event detection reliability.

Index Terms— wireless sensor networks, event monitor, network lifetime, reliability, multi-path encounter route.

I. INTRODUCTION

WIRELESS sensor networks (WSNs) are the most important component of the Internet of Things (IoT) [1][2][3][4]. However, designing energy-efficient and reliable sensor communications protocols for IoT remains a challenging issue [30][31]. WSNs have a wide range of applications, including in large-scale networks, wherein up to tens of thousands of network nodes can be deployed [1][2]. Sensor nodes sensing specific phenomena or previously defined measured values should send data in event-driven networks. Because of such networks being inherently resource constrained, for event data collection, energy efficiency and reporting reliability are crucial design parameters because replacing or recharging the batteries of sensor nodes is extremely difficult [1][4][5][6][7].

Reference [8] shows that a Sink will maintain high event detection reliability with increasing numbers of active reporting nodes (ARNs). However, large numbers of ARNs consume substantial amounts of energy, which decreases network lifetime. To the best of our knowledge, there is no solution that

provides not only high event detection reliability but also long network lifetime. Therefore, a novel event data collection approach named Reliability and Multi-path Encounter Routing (RMER) is proposed to meet long-term event detection reliability as well as high lifetime requirements. The main innovations of this work are as follows:

- 1) The RMER approach provides high event detection reliability. Because sensor nodes consume large amounts of energy in hotspots (referring to the area within a 1-hop range of the Sink) and less energy in non-hotspot areas, the RMER approach selects fewer representative nodes in hotspots and more representative nodes in non-hotspot areas which have sufficient energy. Consequently, the RMER approach can increase network event detection reliability under the premise of guaranteeing network lifetime.
- 2) The RMER approach provides a high network lifetime. In the RMER approach, to further improve the network lifetime, routes of many representative nodes that monitor an event are converged together outside non-hotspot regions and are sent to the Sink after aggregating data. In this case, we can greatly decrease energy consumption in hotspot regions and thus effectively increase network lifetime. Combining with the first innovation point above, the RMER approach can improve network lifetime and event detection reliability simultaneously.
- 3) Comparing with the latest study results, the RMER approach can increase the lifetime and event detection reliability by 10% to 20% simultaneously, and the energy utilization rate of the RMER approach is greater 95%, which was difficult to achieve in past studies.

The remainder of this paper is organized as follows: In Section II, related works are reviewed. The system model is described in Section III. In Section IV, a novel RMER approach is presented. Performance analyses of RMER are provided in Section V. Section VI includes experimental results and comparisons. We conclude in Section VII.

II. RELATED WORK

Because sensor nodes often run on batteries that are generally difficult to recharge once deployed, energy efficiency is a critical issue for prolonging network lifetimes [9]. The key to conserving energy is to place the nodes into sleep mode whenever possible because the energy consumption in such a

Manuscript received 15 April 2015; revised ****, 2015

Mianxiong Dong and Kaoru Ota are with Department of Information and Electronic Engineering, Muroran Institute of Technology, Japan. E-mail: {mx.dong, ota}@csse.muroran-it.ac.jp.

Anfeng Liu is with School of Information Science and Engineering, Central South University, Changsha, 410083 China. E-mail: afengliu@mail.csu.edu.cn

mode is significantly reduced compared to active mode [4][6][8]. Based on this concept, many MAC protocols suitable for sensor networks have been presented by researchers from different aspects. These protocols can be approximately divided into two categories [10]: random competition [11][12] and centralized dispatching [13][14][15]. S-MAC [13] is the most widely applied protocol based on competition mode. Through a consultative consistency of sleep schedule mechanism, this protocol forms virtual clusters and reduces energy consumption by controlling nodes into sleep mode to the greatest extent. T-MAC [14] has been proposed based on S-MAC. This protocol dynamically adjusts the active time based on network traffic flow, sends information using a paroxysmal method and reduces idle interception time, thereby further decreasing energy consumption.

In WSNs, nodes cooperate with each other and monitor specific physical phenomenon together. For most applications, to maintain a certain redundancy and reliability, a large number of sensor nodes are deployed in monitoring areas. When an event occurs, multiple adjacent active nodes can monitor the event simultaneously. Therefore, these adjacent nodes exhibit a spatial correlation for event detection. Researchers have conducted extensive studies on the spatial correlation of sensor nodes [15][16][17].

Once an event occurs, the nodes that monitor this event send data simultaneously and compete with each other for physical channels, which is called Spatially Correlated Contention [18]. However, in many applications, it is not necessary for all the monitor nodes to send data; only some of these nodes need to send data to the head node. Researchers have proposed many types of algorithms [8][15][18][19] to determine which nodes should be selected to send information. Such research attempts to reduce network communication to relieve or even eliminate channel collision between monitor nodes to decrease energy consumption and prolong network lifetime toward ensuring service quality. The CC-MAC protocol [15][19] divides a network into several sub-networks, and only one monitor node is allowed to send data, thereby reducing or eliminate channel collision.

Bouabdallah and others [8] note that the key performance metrics in WSNs are the network lifetime and the required reliability of the reported event. Therefore, a reliability-driven node selection scheme, named the SC-MAC protocol, that exploits the spatial correlation of the MAC layer to achieve further energy conservation as well as to meet the QoS requirements of event detection is proposed [8]. Their main objective is to provide a theoretical framework to calculate the average number of reports needed to be received at the Sink node from the N selected reporting nodes to achieve the required information reliability.

The goal of the above studies is determining how to reduce the energy consumption of event detection, and the main aspects for reducing energy consumption are the MAC layer, node selection (i.e., selected number of ARNs), and number of event monitor data packets to be sent. In addition to reducing energy consumption, the latest studies have been extended to determining how to guarantee the QoS of event detection,

mainly in terms of reliability of event detection. On the one hand, WSNs suffer from a lack of energy in hotspot areas. On the other hand, up to 90% energy is unused in non-hotspot areas. Thus, the network lifetime and reliability of event detection can be further improved by selecting more ARNs in non-hotspot areas and fewer ARNs in hotspot areas, which can lead to improved reliability of event detection and lifetime simultaneously.

Because many nodes monitor the same event, there exists a correlation between those data. Thus, aggregating data can reduce the amount of data sending, data collision and data transmission, which can greatly increase network lifetime. Therefore, this topic is the focus of substantial research.

According to the data aggregation method and goal, in [20], the correlated data gathering scheme can be roughly classified into three categories: routing driven, coding driven, and fusion driven. The routing-driven scheme was not originally designed for correlated data aggregation; however, correlated data can be aggregated when routed to one node and thus decreases the amount of transmitted data. This type of scheme [21][22][23] does not explicitly consider data fusion as an additional requirement when designing routes. The goal of such schemes is to minimize the total amount of transmitted data sent to the Sink. Data aggregation is only processed when routes encounter one another. Most routing-driven schemes adopt a full aggregation model, namely, all child nodal data are aggregated into one data packet at the parent node [24]. Directed Diffusion [21], LEACH [22], and PEGASIS [23] are representative examples of this type of algorithm.

Coding-driven schemes decrease the amount of transmitted data via network coding compression; such schemes are less concerned about the route as they are on designing efficient codes to minimize data amounts. This type of scheme [25][26] focuses on decreasing the amount of transmitted data via coding compression under the premise that only partial aggregation can be achieved. Fusion-driven schemes mainly focus on correlated data aggregation, which attempts to arrange highly correlated data meeting and obtains good data aggregation performance. Such schemes are less concerned with compression as they are with routing algorithms [20].

III. THE SYSTEM MODEL AND PROBLEM STATEMENT

A. The System Model

1) Network model

We consider a WSN consisting of sensor nodes that are uniformly and randomly scattered in a circular network, whose radius is R , with node density ρ , and nodes do not move after being deployed. The perception radius of a node is r_s , and the transmission range is r_c [8][27].

2) Event model

In this work, we consider event-detection-driven wireless sensor applications. In other words, communications are triggered by the occurrence of a pre-specified type of event. Once an event occurs, m reporting nodes must be selected from source nodes to continue generating n reports until the required event detection reliability Υ is achieved. Once the sink node

receives n reports, it instructs the sensor nodes to halt event reporting.

As in reference [8], we denote by f the network reporting frequency. The network reporting frequency is defined as the number of packets generated per unit of time by the network to report an event. Hence, given m ARNs, the reporting frequency of each sensor node must be set equal to $f_s = f/m$ to obtain the predefined network reporting frequency. The m reporting nodes continue generating reports at a rate f_s until the required event detection reliability Υ is achieved.

B. Data aggregation model

For data aggregation, we adopt the lossless step-by-step multi-hop aggregation model introduced in [20]. In such an aggregation model, the aggregation of κ multiple inputs with node s_i is performed sequentially; that is, incoming data are aggregated with existing data in order of arrival. \wp_i denotes the non-aggregated data (or origin data) packet of node s_i ; $\varphi(s_i, s_j)$ denotes the intermediate aggregation result of node s_i and node s_j , or simply φ_i to denote the current intermediate aggregation result of node s_i ; and ϕ_i denotes the final aggregation result of node s_i of all incoming nodal data and its own data.

When node s_i receives data ϕ_j from node j , node s_i aggregates ϕ_j with its own data (which may be the origin data \wp_i or the intermediate data φ_i). If the current data packet of node s_i is φ_i and if data from s_j is $\phi_j = \wp_j$, namely, the data to be aggregated are both origin data, then the aggregation formula is the following:

$$\varphi(s_i, s_j) = \max(\wp_i, \wp_j) + (1-c) \min(\wp_i, \wp_j) \quad (1)$$

In Eq. (1), C is the correlation coefficient [2]. If any data to be aggregated is not source data when being aggregated, the aggregation formula is the following:

$$\varphi(\varphi_i, \phi_j) = \max(\varphi_i, \phi_j) + \zeta(1-c) \min(\varphi_i, \phi_j) \quad (2)$$

In Eq. (2), ζ is called the forgetting factor and is a decimal in the range (0, 1) [20], and φ_i and ϕ_j respectively refer to the intermediate aggregation result and final result of child nodes, where there is at least one non-origin data packet in φ_i and ϕ_j .

C. Energy consumption model and relevant definition

Nodes adopt periodical sleep/active work modes and low duty cycles when no events occur [8]. When an event occurs, some representative nodes should be changed from the periodical sleep/active work mode to the continuous active work mode to monitor the event. Hence, the energy consumption of nodes can be described as three situations: (1) energy consumption of nodes in sleep mode, where ϖ_s denotes this energy consumption rate; (2) energy consumption of nodes when they are in the active work mode, where nodes in this status are called ARNs and ϖ_a denotes this energy consumption rate; and (3) energy consumption when nodes send and receive data.

The energy consumption of sending and receiving data is

TABLE I
NETWORK PARAMETER

Parameter	Value
Threshold distance (d_0) (m)	87
Sensing range r_s (m)	15
E_{elec} (nJ/bit)	50
e_{fs} (pJ/bit/m ²)	10
e_{amp} (pJ/bit/m ⁴)	0.0013
Initial energy (J)	0.5

calculated using a typical energy consumption model [9][27]. The energy consumption for sending data is given by Eq. (3), and the energy consumption for receiving data is given by Eq. (4).

$$\begin{cases} E_{member} = lE_{elec} + l\varepsilon_{fs}d^2 & \text{if } d < d_0 \\ E_{member} = lE_{elec} + l\varepsilon_{amp}d^4 & \text{if } d > d_0 \end{cases} \quad (3)$$

$$E_R(l) = lE_{elec} \quad (4)$$

where E_{elec} denotes the energy consumption of the transmitting circuit. If the transmission distance is less than a threshold distance d_0 , the loss of power amplifier follows a free space propagation model. If the transmission distance is equal to or greater than the threshold distance d_0 , the loss of power amplifier follows a multipath decline model. ε_{fs} and ε_{amp} denote the requisite power amplifier energy of these two models, and l denotes the number of bits. In this paper, the above sets of parameters are from references [9][27]; please see Table 1.

D. Problem statement

Research problems in this paper can be summarized into the two following aspects of optimization:

1) The maximization of event detection reliability

The event detection reliability Υ , when selecting m ARNs and a total number of reports n in the event area of an event source S , is the size of information distortion $D(m, n)$ that the Sink can experience when it rebuilds the event source. Recall that in [8], the expression for the distortion $D(m, n)$ was as follows:

$$\begin{aligned} D(m, n) = & \sigma_s^2 - \frac{\sigma_s^4}{n(\sigma_s^2 + \sigma_m^2)} \left(2 \frac{n}{m} \sum_{i=1}^m \rho(s, i) - 1 \right) \\ & + \frac{\sigma_s^6}{n(\sigma_s^2 + \sigma_m^2)^2} \frac{n-1}{m^2} \sum_{i=1}^m \sum_{j=1}^m \rho(i, j) \end{aligned} \quad (5)$$

where σ_s^2 and σ_m^2 are the variances of the event information S_i and the observation noise v_i of each sensor node n_i ($i = 1, \dots, m$), respectively; $\rho(s, i)$ denotes the correlation coefficient between the event source located at coordinate S and the sensor node n_i ($i = 1, \dots, m$); and $\rho(i, j)$ denotes the correlation coefficient between nodes n_i and n_j ($i, j = 1, \dots, m$) [8].

Obviously, with the decreased information distortion $D(m, n)$, the event detection reliability Υ can be improved. The first goal of this paper is to maximize event detection reliability, namely,

$$\max\{\Upsilon\} = \min\{D(m, n)\} \quad (6)$$

2) The maximization of network lifetime

The lifetime of the network can be defined as the time elapsed until the first sensor node in the network depletes its energy because once a sensor node dies, the sensing capability of the network begins to degrade, therein ceasing the network's function [9]. Hence, the definition of network lifetime in this paper is consistent with references [9][27] and is defined as the time elapsed until the first sensor node in the network depletes its energy. We denote E_i as the energy consumption of node i . The formula for maximizing network lifetime can be expressed as follows:

$$\max(T) = \min_{0 < i \leq n} \max(E_i) \quad (7)$$

Ultimately, we should also guarantee that the event detection reliability Υ is greater than an appointed threshold Φ , namely, $\Upsilon \geq \Phi$. In summary, the optimization goals in this paper are as follows:

$$\begin{cases} \max(T) = \min_{0 < i \leq n} \max(E_i), \\ \max\{\Upsilon\} = \min\{D(m, n)\}, \quad s.t. \Upsilon \geq \Phi \end{cases} \quad (8)$$

IV. RMER APPROACH DESIGN

A. RMER approach design

The main structure of the RMER approach is shown in Fig. 1. The main stages of the RMER approach are as follows:

1) Selection of ARNs.

After an event occurs, select m representative nodes to form m clusters in the event area of event source S . The event data in the clusters are represented by a representative node; then, the representative node switches from sleep/active work mode to the continuous active work mode to monitor the event (called ARNs). Meanwhile, other nodes in the cluster switch to sleep status to conserve energy. As shown in Fig. 1, a_1 , a_2 , a_3 , and a_4 are selected as ARNs. The approach to selecting ARNs is similar to the SC-MAC approach proposed in reference [8]; however, the difference from the former study is the number of ARNs selected by the RMER approach dynamically changes between different network areas. The principle behind ARN selection is as follows: the farther the m ARNs are from the Sink, the larger the number m will become, which can improve event detection reliability. In addition, there is more residual energy in the far Sink area, which can power more ARNs.

2) Stage of centrifugal routing.

All ARN routes deviate from the event source S during the initial routing to reduce collisions in the event area and the event detection delay. After leaving the event area (the radius of the event area is the same as the communication radius r_c), the

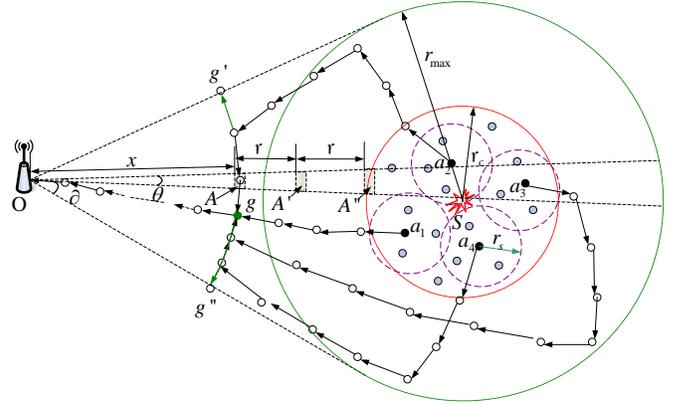


Fig. 1 Illustrate of the RMER approach

ARNs initiate "same-hop routing", or "Horizontal routing". "Horizontal routing" is a routing scheme wherein one node selects its next-hop node as having the same hop counts as itself (from the Sink) in its left-hand direction. In this routing method, each routing node has the same hop count from the Sink, hence its name "same-hop routing". The maximum distance in "same-hop routing" is limited as r_{max} (see Fig. 1).

Algorithm 1. The pseudo-code of the RMER approach

INPUT: event occurs at distance from sink of l m
OUTPUT: event data sent to Sink

- 1: compute the number of active reporting nodes m ;
- 2: select m active reporting nodes as in reference [6];
- 3: **For** each active reporting node a **Do** // centrifugal routing
- 4: node v is the neighbor of a with the farthest distance from event source S ;
- 5: node a sends its packet to node v ;
- 6: h is a random hop from $\{0, r_{max}/r\}$, $h_0=0$;
- 7: **While** $h_0 < h$ **Do**
- 8: v_i is the leftmost (rightmost) node farthest from node v and at a distance from the Sink of the same number of hops as node v ;
- 9: node v sends packet to node v_i ;
- 10: $h_0 = h_0 + 1$; let $v = v_i$;
- 11: **End while**
- 12: **End for**
- 13: **If** node v is the nearest node from Sink **then**
- //If the ARN nearest Sink
- 14: sends its packet to Sink using shortest route; //form main route
- 15: node g is random node selected from main route and node g is not in hotspots;
- 16: node g routes h_x^y hops to the left;
- // h_x^y is from Eq. (9), forms same-hop route path
- 17: node g routes h_x^y hops to the right;
- 18: **Else** //form main route
- 19: forwards its packet to Sink using shortest route before meeting the same-hop route path of node g ;
- 20: forwards its packet along the same hop route path to node g ;
- 21: **End if** // aggregate to node g
- 22: **When** all packets meet in node g **Do**
- 23: aggregate all packets;
- 24: send packet to sink;
- 25: **End when** // data fusion at node g

3) Stage of Multi-path encounter route

The first node that initiates routing to the Sink is the ARN that is closest to the Sink, namely, a_1 in Fig. 1. As in Ref. [8], each node is able to determine its own position, its distance to the Sink, and the node nearest to the Sink. Node a_1 adopts the shortest routing scheme to route to the Sink and forms the first route (named the main route). In the RMER approach, the most important difference from the former routing scheme is aggregating routes together before being sent to the Sink to aggregate relevant data, which can reduce energy consumption in hotspots and improve network lifetime. To aggregate the routes that monitor events, the following method is used in the RMER approach: denote an encounter node g (because of the multiple routes encountered at node g) according to the degree of the remaining energy of nodes in the routing path Oa_1 that forms the main route in the non-hotspot area. The encounter node g performs same-hop routing to the left (and right) based on its distance to the Sink and forms routing paths such as gg' and gg'' shown in Fig. 1. This makes it such that ARNs can encounter same-hop routings of gg' and gg'' with Shortest-path Routing [28] on their way to the Sink. When the routes of other ARNs encounter the same-hop routing, they aggregate the routes along the established same-hop routing to the main route of the encounter node g . This ensures a single route on the encounter node g of the main route and enables data fusion while reducing the number of data nodes to be routed. Clearly, if the distance from the Sink to the event source S is y and that to the encounter node is x , to aggregate all routing information of the same event, the encounter node g requires the following number of hops in same-hop routing:

$$h_x^y = \left\lceil \frac{x \times r_{\max}}{y \times r} \right\rceil \quad (9)$$

4) Stage of stable routing.

Every active reporting node with frequency f_s sends n event data packets to the Sink along the established routing path. Data packets are aggregated in the encounter node g of the main route and are finally sent to Sink.

The pseudo-code of the RMER approach is listed in **Algorithm 1**.

B. Determination of number of ARNs

1) Analysis of Network Energy Consumption Situation

In addition to aggregating relevant data of events, the other important difference of the RMER approach from previous schemes is based on residual energy, therein selecting more ARNs in areas where there is more residual energy to improve event detection reliability. Therefore, this section analyzes the selection of the number of ARNs in different network areas. We first analyze the residual energy situation. After an event occurs, if using the scheme adopted in reference [8], the number of selected ARNs is m , and they directly send data to the Sink. Then, the energy consumption situation under this scheme is given by theorem 1:

Theorem 1: Let the network radius be R , the random probability of event occurrence in the network be λ , and the total number of reports of each event be n . Every data packet is

sent to the Sink by adopting the multi-route scheme of shortest routing, and the node emission radius is r . For the node whose distance from the Sink is l , $l = hr + x$, and the number of data packets that this node will undertake is given as follows:

$$P_l = \left((z+1) + \frac{z(1+z)r}{2l} \right) n\lambda \quad |z = \text{Fix}((R-l)/r) \quad (10)$$

Proof: We have proved the following in reference [27]: if the random probability of event occurrence in a network is λ and if every event generates one data packet, where the distance from one node to the Sink is l , with $l = hr + x$, then the number of data packets that this node will undertake is

$$\left((z+1) + \frac{z(1+z)r}{2l} \right) \lambda \quad |z = \text{Fix}((R-l)/r) \quad (11)$$

Now, every event needs to send n data packets; therefore, the number of data packets that this node will undertake increases n times. Then, Eq. (10) is obtained. ■

Theorem 2: Consider a network radius of R , where an event occurs in the network randomly, the probability of which is λ . The total number of reports of each event is n , and the data of ARNs are aggregated and sent to the Sink when they are b hops from the event source. The node emission radius is r , and the distance from one node to the Sink is l , where $l = hr + x$. The number of data packets that this node will undertake is as follows:

$$P_l^a = \left(b+1 + \frac{(1+b)br}{2l} \right) n\lambda + \left(z-b+1 + \frac{(b+z)(z-b+1)r}{2l} \right) \times \lambda \{ (2-c) + (n-2)\zeta(1-c) \} \quad (12)$$

Proof: We select a small sector A with width $dx \rightarrow 0$ and angle range $\theta \rightarrow 0$ in the area that contains node i (as shown in Fig. 1). Then, the area of A is $\theta l dx$, and sector A receives the data of a node $l + kr$ $|k \in \{1..z\}$ away from itself, namely, data of A' , A'' ... in Fig. 1. In RMER, data of node i are divided into two parts. One part is the data for which the node is within b hops from node i , and these data are not aggregated. The other part is the data for which the node is not within b hops from node i , and these data are aggregated. The calculation of the first part of the data is given below. The number of nodes in area A is $\rho \theta l dx$, and the number of sending data packets is $n \rho \theta l dx \lambda$. The area of A' is $\theta(l+r)dx$, the number of nodes is $\rho \theta(l+r)dx$, the number of event occurrences is $\rho \theta(l+r)dx \lambda$, and the number of sending data packets is $n \rho \theta(l+r)dx \lambda$. By analogy, the number of sending data packets of area A is as follows:

$$P_l^1 = n \{ \rho \theta l dx \lambda + \rho \theta(l+r) dx \lambda + \dots + \rho \theta(l+br) dx \lambda \} \quad (13)$$

If the distance from area A^x whose data is undertaken by area A is more than b hops, then area A receives data from fusion; hence, its length is less than that of the original data. Because event data are aggregated in encounter node g , the aggregation data are calculated by the aggregation data scheme.

After data aggregation, the length of the first two arriving data packets is as follows:

$$\varphi(g) = \ell(p) + (1-c)\ell(p) = \ell(p)(2-c) \quad (14)$$

As for the subsequent $n-2$ data packets, after data aggregation, the length of the total data packets in the encounter node g is as follows:

$$\begin{aligned} \varphi(g) &= \ell(p)(2-c) + (n-2)\zeta(1-c)\ell(p) \\ &= \ell(p)\{(2-c) + (n-2)\zeta(1-c)\} \end{aligned} \quad (15)$$

We can calculate the number of the second part of the data packets that node area A receives as

$$P_i^2 = \left\{ \begin{array}{l} \rho\theta(l+(b+1)r)d_x\lambda + \\ \rho\theta(l+(b+2)r)d_x\lambda + \dots + \rho\theta(l+ZR)d_x\lambda \end{array} \right\} \quad (16)$$

$$\{(2-c) + (n-2)\zeta(1-c)\}$$

Area A transmits two parts of the above data, and area A includes $\rho\theta l d_x$ nodes; thus, the number of data packets that every node sends is obtained as Eq. (10). ■

2) Determination of number of ARNs

According to the event model in Section III, the network reporting frequency f is assured. If the total number of reports n of event collection is predetermined, then the time for one event detection is $\Gamma = n/f$. ARNs are always in the active status, therein consuming more energy compared to nodes in sleep status. Therefore, increasing numbers of ARNs results in increased energy consumption. Every time an ARN is increased in one event detection, additional energy is consumed as follows:

$$e_\Delta = \Gamma(\varpi_a - \varpi_s) = n(\varpi_a - \varpi_s)/f \quad (17)$$

Theorem 2 indicates the amount of data that nodes need to undertake for the same number of ARNs. We know that the closer to the Sink, the more data the node undertakes. In addition, the largest data amount for a node to undertake in a network is P_{\min} , in which min denotes the closest node to the Sink. Considering that the energy consumption of sending and receiving a data packet is e_u , the residual energy of the node at a distance l from the Sink is

$$E_i^{\text{left}} = (P_{\min} - P_l)e_u \quad (18)$$

From the above analysis, for a node whose distance from the Sink is l , the number of selected ARNs can be obtained from Theorem 3.

Theorem 3: Let the selected number of ARNs in hotspots be m_0 , and total number of reports is n . When event detection has occurred a distance l from the Sink, the number of ARNs selected is as follows:

$$m_l = m_0 + \frac{f(P_{\min} - P_l)e_u}{n(\varpi_a - \varpi_s)} \quad (19)$$

Proof: According to Eq. 16, the residual energy of one node a distance l from the Sink is E_i^{left} , and Eq. (17) shows that the increased energy consumption is e_Δ when increasing the number of ARNs by one. Hence, the increasing number of ARNs a distance l from the Sink is

$$m_\Delta = E_i^{\text{left}} / e_\Delta = \frac{(P_{\min} - P_l)e_u}{n(\varpi_a - \varpi_s)/f} = \frac{f(P_{\min} - P_l)e_u}{n(\varpi_a - \varpi_s)} \quad (20)$$

Therefore, the number of ARNs at a distance l from the Sink is

$$m_l = m_0 + m_\Delta = m_0 + \frac{f(P_{\min} - P_l)e_u}{n(\varpi_a - \varpi_s)} \quad (21)$$

V. PERFORMANCE ANALYSIS

Because the RMER approach uses a different number of ARNs in different network areas, event detection distortion is different in different areas of the network. The following theorem concerns event detection distortion in different network areas using the RMER approach.

Theorem 4: In the RMER approach, if the total number of reports is n and the number of ARNs in hotspots is m_0 , then the event detection distortion a distance from the Sink of l is

$$\begin{aligned} D(l, n) &= \sigma_s^2 - \frac{\sigma_s^4}{n(\sigma_s^2 + \sigma_m^2)} \left(2 \frac{n}{m_l} \sum_{i=1}^m \rho(s, i) - 1 \right) \\ &+ \frac{\sigma_s^6}{n(\sigma_s^2 + \sigma_m^2)^2} \frac{n-1}{m_l^2} \sum_{i=1}^m \sum_{j=1}^m \rho(i, j) \end{aligned} \quad (22)$$

Proof: According to Eq. 5, if the total number of reports is n and if the number of ARNs is m , then the event detection distortion is

$$\begin{aligned} D(m, n) &= \sigma_s^2 - \frac{\sigma_s^4}{n(\sigma_s^2 + \sigma_m^2)} \left(2 \frac{n}{m} \sum_{i=1}^m \rho(s, i) - 1 \right) \\ &+ \frac{\sigma_s^6}{n(\sigma_s^2 + \sigma_m^2)^2} \frac{n-1}{m^2} \sum_{i=1}^m \sum_{j=1}^m \rho(i, j) \end{aligned} \quad (23)$$

According to Eq. (17) in Theorem 3, the number of ARNs is m_l when the distance from the Sink is l . Then, Eq. (19) can be obtained with m_l substituted in the above equation. ■

Theorem 5: In the RMER approach, if the total number of reports is n and the number of ARNs in hotspots is m_0 , then the weighted event detection distortion for the entire network is

$$D_w(m_0, n) = \int_0^{2\pi} \int_0^R D(l, n) l d_\theta d_l \quad (24)$$

Proof: We can obtain a tiny random area whose distance from the Sink is l (such as area A in Fig. 1), where the area of A is $d_\theta l d_x$. According to Theorem 4, the event detection

distortion of this area is $D(l, n)$, and the weighted event detection distortion for the entire network is

$$D_w(m_0, n) = \int_0^{2\pi} \int_0^R D(l, n) l d\theta dl \quad (25)$$

VI. EXPERIMENTAL RESULT

OMNET++ [29] is used for experimental verification. The network scene is set as having a Network radius $R=500$ m, and the number of nodes is 1000. Using the energy consumption parameters of Table 1, the event is randomly generated in the network, and the length of each packet is 100 bits.

A. Event detection distortion

Fig. 2 shows, when the number of selected ARNs in hotspots is m_0 , a comparison of network average event detection distortion between the RMER approach and the SC approach for different m_0 . From the experimental results of Fig. 2, when m_0 remains constant, the average event detection distortion of the RMER approach is smaller than that of the SC approach because the RMER approach can select more ARNs in non-hotspot areas. As the experimental results show, the RMER approach can decrease the distortion degree by 16%~20%. Fig. 3 shows a comparison of the distortion degree between the RMER and SC approaches for $m_0=40$ and when the total number of reports n varies. Similarly, RMER can select more ARNs in far-Sink areas, thus generating less average event detection distortion than can the SC approach.

Fig. 4 shows, when no influence on the network lifetime is allowed, the number of ARNs that can be selected using the RMER approach for different distances from the Sink when the emission radius is 40, 60, and 80 m. We can see that, far from the Sink, more ARNs can be selected by the event, which can decrease the event detection distortion. Fig. 5 shows a comparison of event detection distortion between different detection approaches in different network areas. The experimental results show that, because RMER can make full use of surplus energy in non-hotspot areas and can select more ARNs, the farther from the Sink, the lower the event detection distortion. However, the SC approach selects the same number of ARNs over the entire network; thus, the event detection distortion remains invariant over the entire network.

Fig. 6 shows, for $n=10$, $m_0=10$ and different values of r , the comparison results of the average event detection distortion between the RMER approach and the SC approach. The results demonstrate that the RMER approach can decrease distortion degree by approximately 15%.

The following can be concluded from the experiments: (1) The experimental results are consistent with the theoretical analysis, indicating that our analysis is correct. (2) The experimental results also confirm that our RMER approach

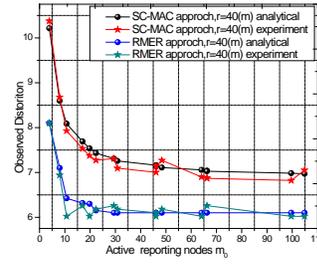


Fig. 2. Distortion under different m_0

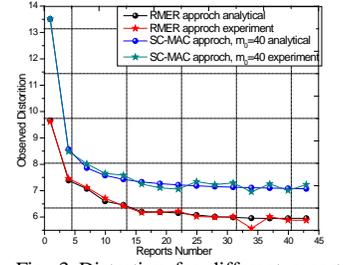


Fig. 3 Distortion for different report numbers

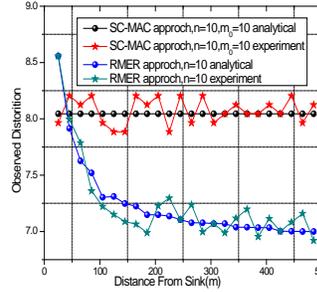


Fig. 4 RMER Active node number for different distances

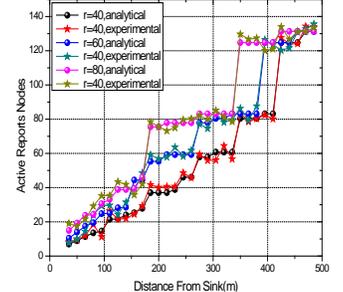


Fig. 5 Distortion for different distances

better guarantees reliability.

B. Energy and lifetime

Fig. 7 and Fig. 8 present 3D graphs of the energy consumption of the SC approach and the RMER approach, respectively. The graphs show the following: in the SC approach, the energy consumption in the near-Sink area is higher than that in other areas, which causes an energy hole and severely reduces network lifetime. In contrast, in the RMER approach, selecting more ARNs in non-hotspot areas balances energy consumption over the entire network because it can not only make full use of residual energy but also reduce network event detection distortion.

Fig. 9 shows a comparison of network maximal energy consumption for different numbers of reports n and for the same network average event detection distortion. Fig. 9 shows that energy consumption in the SC approach is 1.37~1.609 times higher than that of the RMER approach. Because the maximal energy consumption is linearly and inversely proportional to the network lifetime, the RMER approach can obtain a higher network lifetime. Fig. 10 shows the total network energy consumption for the different approaches. Because the RMER approach can make full use of energy in non-hotspot areas, even though the maximal energy consumption is lower than in the former approach, the total energy consumption is higher than that in the former approach, which demonstrates that the approach in this paper obtains a higher energy utilization ratio.

VII. CONCLUSION

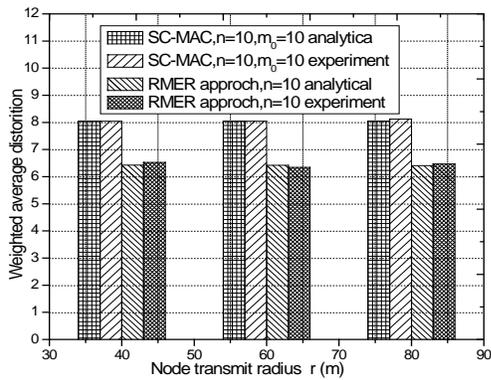


Fig. 6 Illustrate of the RMER approach

In this paper, we propose a novel routing scheme named RMER (Reliability and Multi-path Encounter Route) for meeting reliability, maximum lifetime as well as energy-efficiency requirements. Comparing with previous studies on the SC-MAC approach, the RMER approach ingeniously selects more ARNs far from the Sink, routes data packets generated by one event to a main route, and aggregates data before sending the data to the Sink. This approach can not only ensure event detection reliability but also reduce both redundant data transmission and the energy consumption of nodes. Thus, we can guarantee maximal network lifetime and network reliability. We theoretically and experimentally analyze and prove the RMER schemes. Comparing with previous studies, we provide substantial improvements in event detection distortion and network lifetime.

ACKNOWLEDGEMENT

This work is supported by JSPS KAKENHI Grant Number 26730056, 15K15976 and the A3 Foresight Program, the National Natural Science Foundation of China (61379110). Anfeng Liu is the corresponding author.

REFERENCES

- [1] A. Liu, D. Zhang, P. Zhang, G. Cui, and Z. 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, 2014. doi: 10.1007/s12083-012-0130-1.
- [2] S. He, J. Chen, F. Jiang, D. K. Y. Yau, G. Xing, and Y. Sun, "Energy provisioning in wireless rechargeable sensor networks", *IEEE Trans. on Mobile Comput.*, vol. 12, no. 10, pp. 1931-1942, 2013. doi: 10.1109/TMC.2012.161.
- [3] H. Ning, H. Liu, L. T. Yang, "Aggregated-proof Based Hierarchical Authentication Scheme for the Internet of things", *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 3, pp. 657-667, 2015. doi: 10.1109/TPDS.2014.2311791.
- [4] J. Wang, Y. Zhang, J. Wang, Y. Ma, and M. Chen, "PWDGR: pair-wise directional geographical routing based on Wireless Sensor Network", *IEEE Internet of Things Journal*, vol. 2, no. 1, pp. 14-22, 2015. doi: 10.1109/JIOT.2014.2367116.
- [5] D. Ebrahimi and C. Assi, "Compressive data gathering using random projection for energy efficient wireless sensor networks", *Ad Hoc Netw.*, vol. 16, pp. 105-119, 2014. doi: 10.1016/j.adhoc.2013.12.004.
- [6] L. Anchora, A. Capone, V. Mighali, L. Patrono, and F. Simone, "A novel MAC scheduler to minimize the energy consumption in a Wireless Sensor Network", *Ad Hoc Netw.*, vol. 16, pp. 88-104, 2014. doi: 10.1016/j.adhoc.2013.12.002.

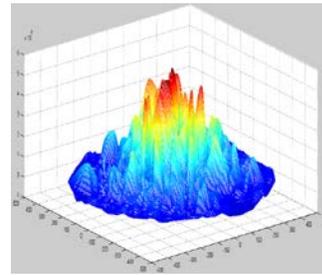


Fig. 7. Energy consumption of SC-MAC

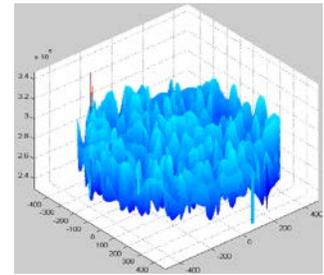


Fig. 8. Energy consumption of RMER

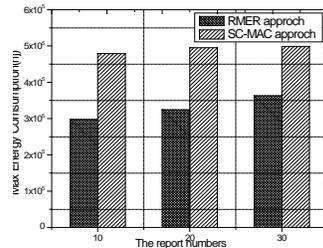


Fig. 9. Max Energy consumption for different n

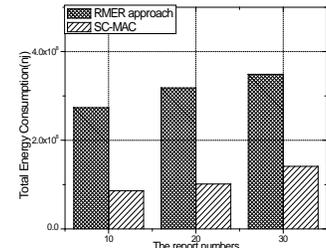


Fig. 10. Total Energy consumption for different n

- [7] R. Zhang, O. Berder, J. Gorce, O. Sentieys, "Energy-delay tradeoff in wireless multihop networks with unreliable links", *Ad Hoc Netw.*, vol. 10, no. 7, pp. 1306-1321, 2012. doi: 10.1016/j.adhoc.2012.03.012.
- [8] F. Bouabdallah, N. Bouabdallah, and R. Boutaba, "Efficient reporting node selection-based MAC protocol for wireless sensor networks", *Wireless Netw.*, vol. 19, no. 3, pp. 373-391, 2013. doi: 10.1007/s11276-012-0473-9.
- [9] A. Liu, Z. Liu, M. Nurudeen, X. Jin, and Z. Chen, "An elaborate chronological and spatial analysis of energy hole for wireless sensor networks", *Computer Standards & Interfaces*, vol. 35, no. 1, pp. 132-149, 2013. doi: 10.1016/j.csi.2012.06.005.
- [10] K. Langendoen and G. Halkes *Embedded Systems Handbook*. Boca Raton, FL, USA: CRC Press, 2005.
- [11] V. Rajendran, K. Obraczka, and J. J. Garcia-Luna-Aceves, "Energy-efficient, collision-free medium access control for wireless sensor networks", *Wireless Netw.*, vol. 12, no. 1, pp. 63-78, 2006. doi: 10.1007/s11276-006-6151-z.
- [12] K. Arisha, M. Youssef, and M. Younis, "Energy-Aware TDMA-based MAC for Sensor Networks," *Comp. Netw.*, vol. 43, no. 5, pp. 539-694, 2003.
- [13] W. Ye, J. Heidemann, and D. Estrin, "Medium access control with coordinated adaptive sleeping for wireless sensor networks", *IEEE/ACM Trans. Networking*, vol. 12, no. 3, pp. 493-506, 2004. doi: 10.1109/TNET.2004.828953.
- [14] T. V. Dam and K. Langendoen, "An adaptive energy-efficient MAC protocol for wireless sensor networks", in *Proceedings of the of the ACM Conference on Embedded Networked Sensor Systems (Sensys'03)*, 2003, pp. 171-180.
- [15] M. C. Vuran and I. F. Akyildiz, "Spatial correlation-based collaborative medium access control in wireless sensor networks", *IEEE/ACM Trans. Networking*, vol. 14, no. 2, pp. 316-329, 2006. doi: 10.1109/TNET.2006.872544.
- [16] M. C. Vuran, Ö. B. Akan, and I. F. Akyildiz, "Spatio-temporal correlation: theory and applications for wireless sensor networks," *Comp. Netw.*, vol. 45, no. 3, pp. 245-259, 2004.
- [17] M. Gastpar and M. Vetterli, "Power, spatio-temporal bandwidth, and distortion in large sensor networks", *IEEE J. Select. Areas Commun.*, vol. 23, no. 4, pp. 745-754, 2005. doi: 10.1109/JSAC.2005.843542.
- [18] K. Jamieson, H. Balakrishnan, and T. Y. C. Sift, *A MAC Protocol for Event-Driven Wireless Sensor Networks*, *Wireless Sensor Networks*, Springer Berlin: Heidelberg, 2006, pp. 260-275.
- [19] Q. Zhao and L. Tong, "Quality-of-service specific information retrieval for densely deployed sensor networks", in *Military Communications Conference, MILCOM*, 2003, vol. 1, 2003, pp. 591-596.
- [20] H. Luo, Y. Liu, and S. Das, "Routing correlated data in wireless sensor networks: A survey", *IEEE Network*, vol. 21, no. 6, pp. 40-47, 2007. doi: 10.1109/MNET.2007.4395109.

- [21] C. Intanagonwiwat, R. Govindan, D. Estrin, and J. Heidemann, F. Silva, "Directed diffusion for wireless sensor networking", *IEEE/ACM Trans. Networking*, vol. 11, no. 1, pp. 2-16, 2003. doi: 10.1109/TNET.2002.808417.
- [22] A. Scaglione and S. Servetto, "On the interdependence of routing and data compression in multi-hop sensor networks", *Wireless Netw.*, vol. 11, nos. 1-2, 149-160. 2005. doi: 10.1007/s11276-004-4752-y.
- [23] S. Patten, B. Krishnamachari, and R. Govindan, "The impact of spatial correlation on routing with compression in wireless sensor networks", *ACM Trans. Sen. Netw.*, vol. 4, no. 4, pp. 24-32, 2008. doi: 10.1145/1387663.1387670.
- [24] E. Zeydan, D. Kivanc, C. Comaniciu, U. Tureli, "Energy-efficient routing for correlated data in wireless sensor networks", *Ad Hoc Netw.*, vol. 10, no. 6, pp. 962-975, 2012. doi: 10.1016/j.adhoc.2011.12.009.
- [25] T. Wang, A. Vosoughi, W. Heinzelman, A. Seyedi, "Maximizing gathered samples in wireless sensor networks with slepian-wolf coding", *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 751-761, 2012. doi: 10.1109/TWC.2011.122211.110417.
- [26] R. Cristescu, B. Beferull-Lozano, M. Vetterli, R. Wattenhofer, "Network correlated data gathering with explicit communication: NP-completeness and algorithms", *IEEE/ACM Trans. Networking*, vol. 14, no. 1, pp. 41-54, 2006. doi: 10.1109/TNET.2005.863711.
- [27] A. Liu, Z. Zheng, C. Zhang, Z. Chen, and X. Shen, "Secure and energy-efficient disjoint multipath routing for WSNs", *IEEE Trans. Veh. Technol.*, vol. 61, no. 7, pp. 3255-3265, 2012. doi: 10.1109/TVT.2012.2205284.
- [28] L. A. Villas, A. Boukerche, H. S. Ramos, H. A. B. F. de Oliveira, R. B. de Araujo, and A. A. F. Loureiro, "Drina: A lightweight and reliable routing approach for in-network aggregation in wireless sensor networks", *IEEE Trans. Comput.*, vol. 62, no. 4, pp. 676-689, 2013. doi: 10.1109/TC.2012.31.
- [29] OMNet++ Network Simulation Framework. Available: <http://www.omnetpp.org/>.
- [30] X. Tian, Yi. Zhu, K. Chi, J. Liu, and D. Zhang, "Reliable and Energy-efficient Data Forwarding in Industrial Wireless Sensor Networks," *IEEE Systems Journal*, 2015, doi: 10.1109/JSYST.2015.2466696.
- [31] J. Liu, X. Jiang, H. Nishiyama and N. Kato, "Reliability Assessment for Wireless Mesh Networks Under Probabilistic Region Failure Model," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 5, pp. 2253-2264, 2011.

Department of Information and Electronic Engineering, Muroran Institute of Technology, Japan. From March 2010 to March 2011, she was a visiting scholar at University of Waterloo, Canada. Also she was a Japan Society of the Promotion of Science (JSPS) research fellow with Kato-Nishiyama Lab at Graduate School of Information Sciences at Tohoku University, Japan from April 2012 to April 2013. Her research interests include Wireless Sensor Networks, Vehicular Ad Hoc Networks, and Ubiquitous Computing.



Anfeng Liu is a Professor of School of Information Science and Engineering of Central South University, China. He is also a Member (E200012141 M) of China Computer Federation (CCF). He received his M.Sc. and Ph.D. degrees from Central South University, China, in 2002 and 2005, both in computer science. His major research interest is wireless sensor

networks.



Mianxiong Dong received B.S., M.S. and Ph.D. in Computer Science and Engineering from The University of Aizu, Japan. He is currently an Assistant Professor in the Department of Information and Electronic Engineering at the Muroran Institute of Technology, Japan. Prior to joining Muroran-IT, he was a Researcher at the National Institute of

Information and Communications Technology (NICT), Japan. He was a JSPS Research Fellow with School of Computer Science and Engineering, The University of Aizu, Japan and was a visiting scholar with BCCR group at University of Waterloo, Canada supported by JSPS Excellent Young Researcher Overseas Visit Program from April 2010 to August 2011. His research interests include Wireless Networks, Cloud Computing, and Cyber-physical Systems.



Kaoru Ota received M.S. degree in Computer Science from Oklahoma State University, USA in 2008 and B.S., Ph.D. degrees in Computer Science and Engineering from The University of Aizu, Japan in 2006, 2012, respectively. She is currently an Assistant Professor with