Electronic Preprint for Journal of Information Processing Vol.20 No.1

Regular Paper

An Energy-Efficient Data Gathering Mechanism using Traveling Wave and Spatial Interpolation for Wireless Sensor Networks

Yoshiaki Taniguchi[1,2,a)] Akimitsu Kanzaki[2] Naoki Wakamiya[2]
Takahiro Hara[2]

Received: April 11, 2011, Accepted: September 12, 2011

Abstract: Wireless sensor network technologies have attracted a lot of attention in recent years. In this paper, we propose an energy-efficient data gathering mechanism using traveling wave and spatial interpolation for wireless sensor networks. In our proposed mechanism, sensor nodes schedule their message transmission timing in a fully-distributed manner such that they can gather sensor data over a whole wireless sensor network and transmit that data to a sink node while switching between a sleep state and an active state. In addition, each sensor node determines the redundancy of its sensor data according to received messages so that only necessary sensor data are gathered and transmitted to the sink node. Our proposed mechanism does not require additional control messages and enables both data traffic and control traffic to be drastically reduced. Through simulation experiments, we confirmed that with our proposed mechanism, the number of message transmissions can be reduced by up to 77% and the amount of transmitted data can be reduced by up to 13% compared to a conventional mechanism.

Keywords: data aggregation, data gathering, spatial interpolation, traveling wave, wireless sensor networks

1. Introduction

Wireless sensor networks (WSNs) have attracted a lot of attention in recent years [2]. In particular, monitoring applications such as surveillance or environmental monitoring are among the most promising applications of WSNs. Until now, there have been developed several WSN monitoring systems all over the world [3], [4], [5], [6], and the volume of information generated by WSNs is rapidly increasing. Thus, in view of the information explosion, WSN techniques should be developed [7]. Here, a WSN generally consists of a large number of battery-powered sensor nodes, each of which has a general purpose processor with limited computational capability, a small memory and a radio transceiver. Therefore, WSN control mechanisms to deal with the information explosion should be energy-efficient to prolong the lifetime of a WSN.

Many studies have been conducted on energy-efficient techniques for WSNs. In general, a sensor node has the capability of switching its modules on and off. For example, almost all commercial sensor nodes, such as MICAz [8], SunSPOT [9], can switch their mode between sleep mode and active mode. By entering sleep mode, energy consumption of a sensor node can be drastically reduced. Therefore, efficient control of switching between active and sleep mode (i.e., *sleep scheduling*) is necessary for energy-efficiency of WSNs [10].

To enable WSN sleep scheduling in a fully-distributed manner for periodical monitoring applications, we have proposed a traveling wave-based communication mechanism (WAVE) [11]. WAVE can organize a variety of periodic communications with sleep scheduling depending on dynamically changing application requirements. In the case of periodical data gathering from every sensor node to a sink node, WAVE message transmissions begin from the edge of the WSN and end at the sink node. When sensor nodes at the farthest hop distance from the sink node transmit messages, the sensor nodes that are closer to the sink node by one hop (i.e., the next-hop nodes) are scheduled to wake up to receive the messages. After a certain period, they also transmit messages under a timing such that their next hop sensor nodes awaken to receive the messages. At this same time, the farthest sensor nodes also receive the messages and go to sleep. As a consequence of such scheduling, a concentric circle-shaped message propagation (i.e., a *traveling wave*) from the edge of the WSN to the sink node is accomplished. In WAVE, each sensor node just periodically broadcasts a message, including its sensor data and a small amount of control information, in accordance with its own timer. The sensor node timer is adjusted appropriately according to the reception of messages from neighbor sensor nodes. WAVE does not require additional signaling. The energy-efficiency and adaptability to the dynamically changing environment of WAVE

1 Cybermedia Center, Osaka University, Toyonaka, Osaka 560–0043, Japan
2 Graduate School of Information Science and Technology, Osaka University, Suita, Osaka 565–0871, Japan
a) y-tanigu@cmc.osaka-u.ac.jp

This paper is an extended version of the paper which appeared in Proceedings of the 2011 International Conference on Communications, Computing and Control Applications (CCCA 2011) [1].

are shown through simulation evaluations [11], and its practicality is shown through implementation and experimental evaluations using commercial sensor nodes [12].

As for monitoring applications, *data aggregation* is also an important energy-efficient technique [13]. Data aggregation techniques reduce the amount of data to be transmitted by using the feature that sensor data generally have spatial correlations. Deleted sensor data by a data aggregation technique is estimated at the sink node using other sensor data gathered to the sink node. Since wireless communication generally consumes more energy compared with other operations such as sensing and processing, traffic reduction by data aggregation is effective in reducing energy consumption in WSNs. Until now, many data aggregation mechanisms have been proposed [13], [14], [15]; however, these proposals require control message exchanges in order to share information [14] or cannot guarantee that the differences between the results of approximation at the sink node and the actual sensor data value are smaller than a predetermined value [14], [15].

Unlike previous proposals, we have proposed an overhearing-based data aggregation mechanism using spatial interpolation (ODAS) [16]. In ODAS, each sensor node overhears messages and determines the redundancy of its sensor data by using sensor data overheard from neighbor nodes. A sensor node does not transmit its sensor data when it determines its sensor data to be redundant. Therefore, only the necessary sensor data are gathered and transmitted to the sink node. In addition, the maximum error between the estimated sensor data value at the sink node and the actual sensor data value is guaranteed. Since ODAS does not require control messages, both data traffic and control traffic can be drastically reduced. The effectiveness of ODAS is shown through simulation evaluations using some kinds of temperature distributions [16], [17]. However, ODAS mainly focuses on data aggregation and it requires some assumptions on the under layer, such as ideal time division multiple access (TDMA) scheduling, time synchronization among sensor nodes, and ideal routing between sensor nodes and the sink node.

In this paper, in order to further reduce energy consumption in monitoring applications, we propose a data gathering mechanism integrating our communication mechanism and data aggregation mechanism. In the proposed mechanism, we use WAVE as the under layer technique and ODAS as the higher layer technique. As in both these previous works, the mechanism proposed in the present work does not require signaling and control messages. The target application of our proposed mechanism is periodical monitoring from whole flat-surface areas where sensor nodes are sufficiently distributed in the monitoring area. Example applications are vineyard monitoring [18], habitat monitoring [19] and soil ecology monitoring [20]. Through simulation experiments, we evaluate our proposed mechanism against conventional mechanisms in terms of the number of non-redundant sensor data, consumed energy, and data gathering delay.

The remainder of this paper is organized as follows. In Section 2, we describe the assumptions upon which we rely in this paper. Next, we briefly describe the conventional mechanisms in Section 3. Then, in Section 4, we propose an energy-efficient data gathering mechanism with transmission reduction. After that, we evaluate our proposed mechanism in Section 5. Finally, in Section 6, we conclude this paper with an outlook for future research.

2. Assumptions and Definitions

In this section, we describe the assumptions upon which we rely in this paper. We assume a multi-hop WSN, comprised of a single sink node and a set of sensor nodes $\mathcal{N} = \{n_1, \cdots, n_N\}$ deployed in a flat region. All sensor nodes have the same circular radio transmission range (radius of R). Each sensor node knows its own location and the locations of its neighbor sensor nodes through the use of a global positioning system (GPS) or a localization mechanism [21]. Here, the neighbor sensor nodes of a sensor node are defined as the set of sensor nodes that exist within range R from the sensor node. The sink node knows the location of all sensor nodes.

The monitoring application requires the WSN to periodically provide information on the entire target region, for example, distribution of temperature. The cycle duration T and acceptable error range E are set in advance. At interval T (e.g., one hour), each sensor node senses some physical phenomena (e.g., temperature) and sends the sensor data to the sink node. The acceptable error range E (e.g., 0.1 degree) is the acceptable difference between the actual sensor data value and an estimated sensor data value for all sensor nodes. The estimated sensor data value is a sensor data value estimated at the sink node based on a data aggregation mechanism. Thus, the following condition should be satisfied for all sensor nodes:

$$|d_i - \hat{d}_i| \leq E \ \ (\forall n_i \in \mathcal{N}), \tag{1}$$

where d_i and $\hat{d}_i$ stand for the actual sensor data value and the estimated sensor data value of sensor node n_i, respectively.

3. Conventional Mechanisms

In this section, we briefly explain our previously proposed mechanisms.

3.1 WAVE [11]

3.1.1 Overview

We first introduce an overview of WAVE. As the medium access control (MAC) protocol, WAVE assumes a carrier sense multiple access (CSMA)-based protocol, which is widely implemented in commercial sensor units such as MICAz [8]. WAVE can organize two types of message propagation, *data diffusion* and *data gathering*, depending on dynamically changing application requirements. In WAVE, any sensor node can become a point, called a *core node*, from which messages are disseminated or to which messages are gathered. When there is no session, sensor nodes transmit messages at their own timing independently from the others. If there is a session, concentric circle-shaped message propagation, that is, a *traveling wave*, is generated in a self-organizing manner under which the direction of the traveling wave is from the core node to the edge of the WSN in data diffusion or the opposite in data gathering. **Figure 1** shows an overview of WAVE in the case of data gathering from all sensor nodes to the core node from the viewpoint of sensor nodes two hops away from the core node. To autonomously generate and

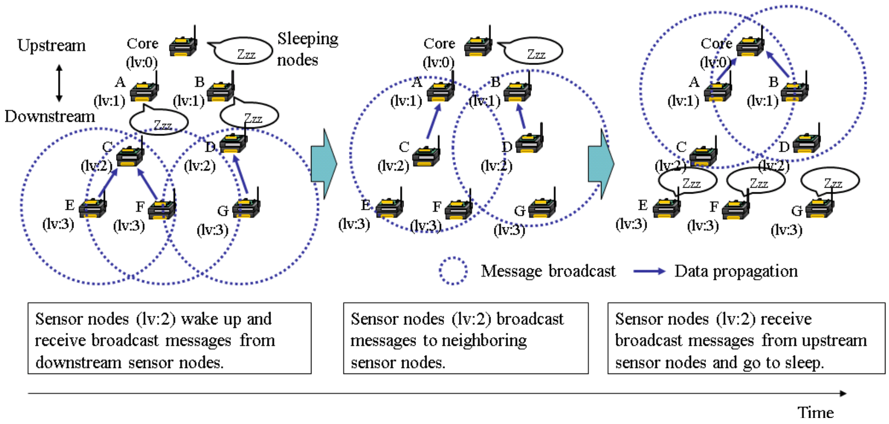


Fig. 1 Basic behavior of WAVE in the case of data gathering.

maintain concentric circle-shaped traveling waves, WAVE adopts a pulse-coupled oscillator model which explains a biological synchronization apparent in groups of flashing fireflies [22].

In the following, we explain the basic behavior of WAVE from the viewpoint of data gathering. In WAVE, each sensor node maintains a phased timer, a phase response curve (PRC) function, a level value, and some control parameters. The phase shifts toward T as time passes. When it reaches T, it expires and goes back to zero. The PRC function determines the amount of phase shift on receiving a message. By configuring the PRC function appropriately, the desired traveling wave appears regardless of the initial phase condition of sensor nodes [11]. The level value indicates the number of hops from a core node, and it is initialized when a new session starts.

A sensor node periodically broadcasts messages containing its control information and all sensor data in its local buffer whenever its phase expires. When a sensor node receives a message from a neighbor sensor node whose level value is smaller, it sets its level value as the received value plus one. Then, it is stimulated to generate and maintain a traveling wave. The stimulated sensor node shifts its phase based on the PRC function. To avoid being stimulated by delayed messages, a sensor node is not stimulated by messages from other sensor nodes with a smaller level value for a certain duration after it has already been stimulated. When a sensor node receives a message whose level value is larger, the sensor node deposits the received sensor data in its local buffer. Through such mutual interactions among neighbor sensor nodes, concentric circle-shaped traveling waves are autonomously generated. After traveling waves appear, a sensor node starts sleep scheduling by turning off its modules between successive receptions and transmissions.

3.1.2 Feature of WAVE

We should note here that each sensor node only broadcasts messages in accordance with its own timer. WAVE does not require additional signaling or control messages to generate and maintain traveling waves. However, WAVE does not consider transmission reduction of sensor data and all of the sensor data are gathered to the core node. If the application accepts a certain range of error among the sensor data value, the amount of traffic can be reduced. In addition, since WAVE is a broadcast-based communication mechanism, sensor data are forwarded to the core node in multi-paths, which generally consumes more energy in comparison with single-path forwarding. Thus, there is a room to reduce the energy consumption of WAVE.

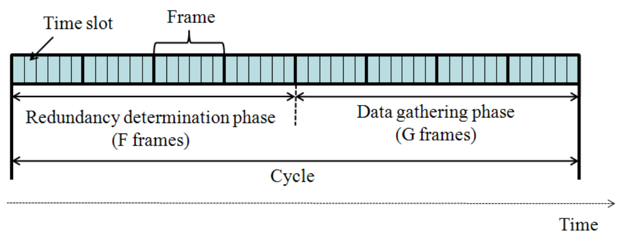


Fig. 2 Frame structure of ODAS in one cycle.

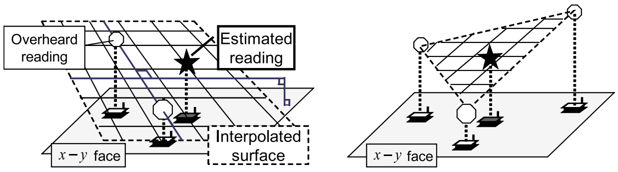
3.2 ODAS [16]

3.2.1 Assumptions

We next briefly introduce our data aggregation mechanism. In ODAS, TDMA is assumed as the MAC protocol. In TDMA, time is divided into time slots. Each sensor node is assigned a time slot and is allowed to transmit a message including a sensor data to a neighbor sensor node within its assigned time slot. Multiple time slots are grouped into a frame, with each time slot appearing within the period of the frame. In ODAS, it is assumed that an optimum time slot assignment is achieved in advance by considering radio interference relationships. It is also assumed that all sensor nodes know the time slots assigned to themselves and to their neighbor sensor nodes. Therefore, each sensor node can implement sleep scheduling by turning off its modules during unrelated time slots. In addition, the clock is assumed to be synchronized among all sensor nodes by applying some conventional time synchronization protocol [23].

3.2.2 Overview

In ODAS, each cycle consists of two phases, namely, the *redundancy determination phase* and the *data gathering phase*, as shown in **Fig. 2**. The redundancy determination phase consists of F frames. In the beginning frame, some sensor nodes first transmit messages without overhearing. Here, a message consists of the sensor node identifier and the sensor data of the node. These sensor nodes are selected to be evenly distributed within the target region. When a sensor node overhears a message from a neighbor sensor node, it determines the redundancy of its sensor data as described in Section 3.2.3. If a sensor node determines that its



(a) Using two sensor data. (b) Using three sensor data.

Fig. 3 Spatial interpolations in ODAS.

sensor data is not redundant, it transmits a message in a succeeding frame. Otherwise, the sensor node does not transmit a message in the cycle. Such message receptions and transmissions are repeated during the redundancy determination phase.

In the data gathering phase, such non-redundant sensor data are gathered to the sink node through a tree-shaped communication route that is constructed in advance. The data gathering phase consists of G frames. The number of frames G should be configured to the maximum hop distance between the sink node and the sensor nodes. During the data gathering phase, a sensor node that has non-redundant sensor data transfers the sensor data to its parent node. On the other hand, redundant sensor data are discarded at the sensor node. Such missing sensor data, however, can be restored at the sink node by satisfying Eq. (1) because the sink node can reenact the redundancy determination process performed by each sensor node.

3.2.3 Redundancy Determination Mechanism

To determine the redundancy of sensor data, ODAS uses three-dimensional spatial interpolation. Here, we assume an $x - y - z$ space in which the $x - y$ face corresponds to the target flat space for sensing; in other words, the x and y coordinates represent the sensor node's location, and the z coordinate corresponds to the sensor data value at the location. First, a sensor node calculates the estimated sensor data value by the following procedure:

- When the sensor node stores only one overheard sensor data, the sensor node only regards this sensor data value as its estimated sensor data value.
- When the sensor node stores two overheard sensor data, as shown in **Fig. 3**, the sensor node first derives a flat surface including the line containing the overheard sensor data value and its perpendicular, which is parallel to the $x-y$ face. Next, the sensor node chooses the value, whose x and y coordinates correspond to the ones of itself, from the derived surface as its estimated sensor data value.
- When the sensor node stores more than two overheard sensor data, as shown in Fig. 3 (b), the sensor node first chooses three nodes whose locations construct a triangle containing the location of the sensor node. If there are multiple candidates, the sensor node chooses ones in which the total distance between the three sensor nodes and itself is the smallest. On the other hand, if there is no set of sensor nodes that construct a triangle containing the sensor node, the sensor node chooses three sensor nodes in which the total distance between the three sensor nodes and itself is the smallest. Next, the sensor node derives a flat surface that contains the triangle constructed of the overheard sensor data value of the chosen three sensor nodes. Finally, it chooses a value whose x and y coordinates correspond to its own coordinates, from the derived surface as its estimated sensor data value.

After calculating the estimated sensor data value, each sensor node determines the redundancy of its sensor data value. In this procedure, sensor node n_i evaluates the following condition:

$$|d_i - \hat{d}_i| \leq E, \tag{2}$$

where d_i and $\hat{d}_i$ stand for the actual sensor data value and the estimated sensor data value of sensor node n_i, respectively. When this condition is false, sensor node n_i determines that its sensor data is not redundant. Otherwise, it determines that its sensor data is redundant.

3.2.4 Feature of ODAS

By the redundancy determination mechanism of sensor nodes, ODAS can drastically reduce the amount of message transmissions. However, ODAS makes some assumptions for the under layer as described in Section 3.2.1. Although ODAS itself does not need control messages, TDMA scheduling, time synchronization, and routing generally require overhead or control messages. To accomplish autonomous and energy-efficient scheduling, we apply WAVE for the under layer of ODAS in the following section.

4. An Energy-efficient Data Gathering Mechanism Using Traveling Wave and Spatial Interpolation

In this section, we propose a data gathering mechanism for WSNs which integrates WAVE and ODAS. The basic behavior of our proposed mechanism is based on WAVE by regarding the core node as the sink node.

4.1 Control Parameters

In our proposed mechanism, sensor node $n_i \in \mathcal{N}$ has a timer with phase $\phi_i \in [0, T]$ ($d\phi_i/dt = 1$). It maintains level value l_i, offset τ_i ($\tau_{\min} \leq \tau_i \leq \tau_{\max}$), sensing delay ϵ ($0 < \epsilon < \tau_{\min}$), PRC function $\Delta_i(\phi_i)$, parent node identifier n_i^p, a set of reference sensor data $\mathcal{D}_i^r$, and a set of transmission sensor data $\mathcal{D}_i^t$. The level value indicates the number of hops from the sink node. The level value of a sensor node is initialized to the maximum value (e.g., 255 when l consists of eight bits), and that of the sink node is set to zero. The offset defines the interval of message transmission between a sensor node of level $l - 1$ and that of level l. Offset τ_i is chosen randomly to avoid synchronous message transmission among sensor nodes of the same level. The maximum offset $\tau_{\max}$ is determined taking into account the density of sensor nodes over the whole WSN. Sensing delay ϵ is the maximum delay for reading sensor data value from its sensor devices. To autonomously generate concentric circle-shaped traveling waves for message propagation regardless of the initial phase condition, we use the following PRC function for all sensor nodes [11].

$$\Delta_i(\phi_i) = a \sin \frac{\pi}{\tau_i} \phi_i + b(\tau_i - \phi_i), \tag{3}$$

where a and b are parameters which determine the speed of convergence. The parent node identifier is used to reduce the amount of message transmissions due to the multi-path effect that arises as a consequence of broadcast-based communication. The set of

Table 1 Control information maintained at sensor node n_i.

Parameter	Content
ϕ_i	phase of timer ($\phi_i \in [0, T], d\phi_i/dt = 1$)
l_i	level value, i.e., number of hops from the sink node
τ_i	offset ($\tau_{\min} \leq \tau_i \leq \tau_{\max}$)
ϵ_i	sensing delay ($0 < \epsilon < \tau_{\min}$)
$\Delta_i(\phi_i)$	phase response curve (PRC), Eq. (3)
n_i^p	parent node identifier
$\mathcal{D}_i^r$	the set of reference sensor data (to use for redundancy determination)
$\mathcal{D}_i^t$	the set of transmission sensor data (to be gathered to the sink node)

Table 2 Information in a broadcast message from sensor node n_i.

Information	Content
l_i	level value
n_i^p	parent node identifier
n_i	sensor node identifier
$\mathcal{D}_i^t$	sensor data

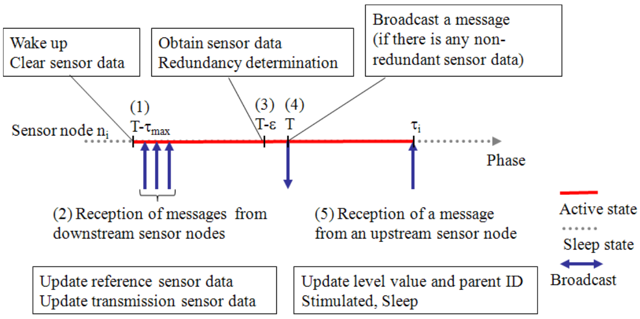


Fig. 4 Sensor node behavior in our proposed mechanism.

reference sensor data $\mathcal{D}_i^r$ is only used to determine the redundancy of sensor data. The set of transmission sensor data $\mathcal{D}_i^t$ is used both to determine the redundancy of sensor data and to transfer it to the upstream sensor nodes. These control parameters are summarized in **Table 1**.

4.2 Message Structure

In our proposed mechanism, sensor data are forwarded among neighboring sensor nodes by broadcast as WAVE. A broadcast message from sensor node n_i contains the set of transmission sensor data $\mathcal{D}_i^t$ and a small amount of control information: level value l_i, its identifier n_i, and parent node identifier n_i^p, as shown in **Table 2**. No control message is required in our proposed mechanism.

4.3 Behavior of Sensor Node

First, we should note that our proposed mechanism is an extension of WAVE. At the beginning, there is no concentric circle-shaped traveling waves in the WSN in our proposed mechanism since the phase of each sensor node is randomly initialized. After the elapse of certain cycles and mutual interactions among neighbor sensor nodes, the traveling waves are autonomously generated and a sensor node starts sleep scheduling as WAVE. The sensor node behavior in this stage is shown in **Fig. 4**. In the following, we describe the behavior of a sensor node after the traveling wave is generated and the sleep scheduling of sensor nodes is already started.

Sensor node n_i behaves in accordance with its phase ϕ_i and received information from neighbor sensor nodes as follows:

(1) When phase ϕ_i becomes $T - \tau_{\max}$, sensor node n_i wakes up. After waking up, sensor node n_i clears the set of reference sensor data $\mathcal{D}_i^r$ and the set of transmission sensor data $\mathcal{D}_i^t$. Then, it waits for message reception. Downstream sensor nodes are scheduled to broadcast messages when phase ϕ_i is between $T - \tau_{\max}$ and T from the viewpoint of sensor node n_i.

(2) When sensor node n_i receives a message from downstream sensor node n_j with $l_j = l_i + 1$ and $n_j^p \neq n_i$, sensor node n_i adds the received sensor data $\mathcal{D}_j^t$ to its set of reference sensor data $\mathcal{D}_i^r$. On the other hand, when sensor node n_i receives a message from downstream node n_j with $l_j = l_i + 1$ and $n_j^p = n_i$, sensor node n_i adds the received sensor data $\mathcal{D}_j^t$ to its set of transmission sensor data $\mathcal{D}_i^t$.

(3) When phase ϕ_i reaches $T - \epsilon$, sensor node n_i reads sensor data value d_i from its sensor device. Then, sensor node n_i determines the redundancy of its reading sensor data using the set of sensor data from downstream neighbor sensor nodes $\mathcal{D}_i^{det} = \{d_k \in \mathcal{D}_i^r \cup \mathcal{D}_i^t \mid l_k = l_i + 1\}$. The redundancy determination mechanism is the same as that in ODAS (see Section 3.2.3). If sensor node n_i determines that its sensor data is not redundant, it adds its sensor data to its set of transmission sensor data $\mathcal{D}_i^t$.

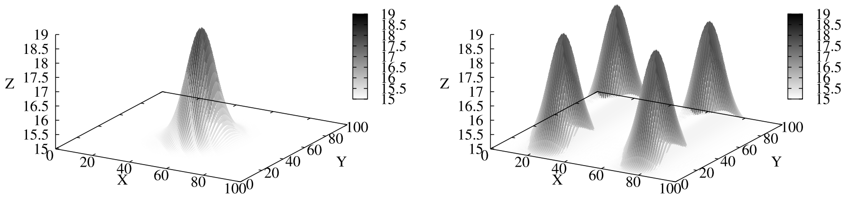
(4) When phase ϕ_i reaches T, sensor node n_i checks the set of transmission sensor data $\mathcal{D}_i^t$. If there is any sensor data in $\mathcal{D}_i^t$, sensor node n_i broadcasts a message, which is received by any awake sensor node in the range of radio communication. On the other hand, if there is no sensor data (i.e., $\mathcal{D}_i^t = \emptyset$), sensor node n_i does not broadcast a message. The phase ϕ_i, reaching T, goes back to zero.

(5) After the phase returns to zero, sensor node n_i stays awake and waits for message reception from an upstream sensor node. When sensor node n_i receives a message from sensor node n_j with $l_j < l_i$, it sets its level value $l_i = l_j + 1$ and parent identifier $n_i^p = n_j$. It then shifts its phase by an amount $\Delta_i(\phi_i)$ in Eq. (3) so as to maintain a traveling wave. After that, sensor node n_i goes to sleep.

In a steady-state situation, sensor node n_i receives the message from the upstream sensor node when its phase ϕ_i reaches τ_i. Therefore, sensor node n_i stays awake for the duration of $\tau_{\max} + \tau_i$ in one data gathering cycle T; in other words, the duty cycle becomes $(\tau_{\max} + \tau_i)/T \simeq 2\tau_{\max}/T$.

4.4 Discussion on the Redundancy Determination

In our proposed mechanism, the redundancy of each sensor datum is determined at the same time as data gathering, whereas ODAS assumes two phases (i.e., the redundancy determination phase and the data gathering phase). Since each sensor node is scheduled to broadcast a message before its upstream sensor nodes broadcast messages (see Fig. 4), the sensor data from upstream nodes cannot be used for redundancy determination. In addition, the traveling wave-based mechanism does not ensure the order of message transmission among sensor nodes of the same level value. Since the sink node should reenact the interpolation process performed by each sensor node to restore any missing sensor data, sensor data from same-hop nodes cannot also be



(a) Single distribution. (b) Combined distribution.

Fig. 6 Two-dimensional normal distribution of temperature.

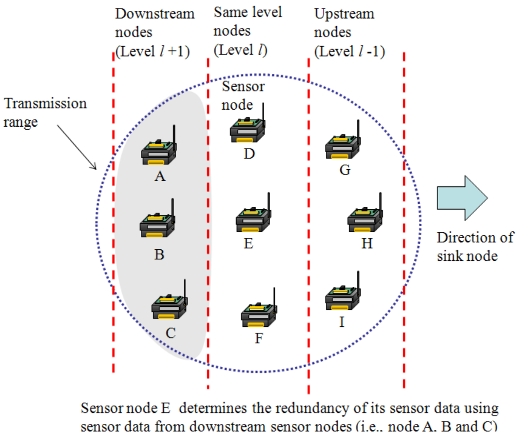


Fig. 5 Sensor data used in the redundancy determination process.

used for redundancy determination. For these reasons, in our proposed mechanism, each sensor node only uses sensor data from downstream nodes to determine the redundancy of its sensor data as shown in **Fig. 5**. Therefore, a sensor node cannot choose three sensor nodes whose locations construct a triangle containing the location of the sensor node as described in Section 3.2.3. As a result, the performance of redundancy determination differs between our proposed mechanism and the ODAS mechanism.

5. Simulation Experiments

In this section, we evaluate our proposed mechanism through simulation experiments.

5.1 Simulation Settings

In the simulation experiments, we consider a WSN of $N = 200$ sensor nodes randomly distributed in a 100 m × 100 m region. A sink node is located at a corner of the region. The radio communication range R is fixed at 20 m. We use two kinds of temperature distribution in this paper. The first one is a single two-dimensional normal distribution as shown in **Fig. 6** (a), and the second one is a combination of four two-dimensional normal distributions as shown in Fig. 6 (b). Hereinafter, we refer to these two distributions as the single distribution and the combined distribution. The size of the sensor data is 64 B. The monitoring application requires that the temperature distribution over the whole WSN be gathered at an interval of $T = 60$ s with an acceptable error range E. **Table 3** shows the energy consumption model of a sensor node [8]. Since wireless communication generally consumes more energy compared with other operations, we only consider the energy consumption of wireless communication and ignore the energy consumed by other operations such as computation, sensing, etc. For the parameters of our proposed mechanism, we use $\tau_{\min} = 0.2$ s and $\tau_{\max} = 0.6$ s [*1]. In addition, we use $a = 0.01$ and $b = 0.5$ for PRC in Eq. (3).

Table 3 Energy consumption model.

Operation mode	Current [mA]
Transmit	22
Receive (overhear)	22
Listen	2
Sleep	0.15

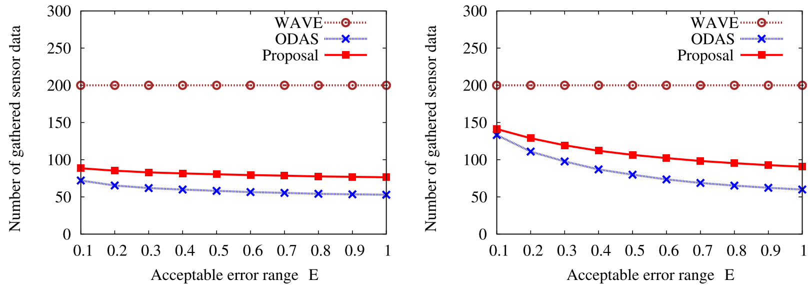
We conduct 100 simulations and evaluate the performance of our proposed mechanism in terms of *number of non-redundant sensor data*, *number of message transmissions*, *amount of transmitted data*, *consumed energy*, and *data gathering delay*. The number of non-redundant sensor data is the average number of sensor data which is determined to be non-redundant and is gathered at the sink node in a cycle. The number of message transmissions is the average number of message transmissions on the whole WSN in a cycle. The amount of transmitted data is the total amount of transmitted sensor data on the whole WSN in a cycle. The consumed energy is the total consumed energy on the whole WSN in a cycle. The data gathering delay is the duration between the time that a cycle begins and the time that the sink node finishes receiving sensor data from the WSN in the cycle.

To the best of our knowledge, there is no appropriate research work which considers sleep scheduling and data aggregation like our proposed mechanism. Therefore, we conduct simulation experiments for WAVE and ODAS in this paper for comparison purposes. In ODAS, we set the number of intermediate frames F at 2, and the number of data gathering frames G at the maximum hop number of the tree in the data gathering phase, since the performance of ODAS becomes the best in these settings [16]. The duration of a time slot is set as 20 ms. In addition, the optimal time slot assignment for all sensor nodes is attained in advance.

5.2 Evaluation of the Number of Non-redundant Sensor Data

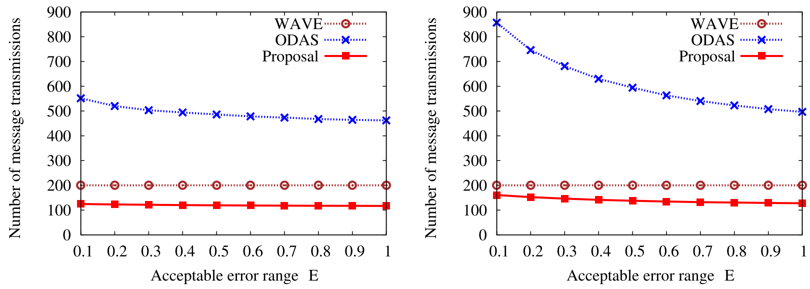
Figure 7 (a) shows the number of non-redundant sensor data against the acceptable error range E when the single distribution is used. Since WAVE does not determine the redundancy of sensor data, all sensor data are gathered to the sink node regardless of the acceptable error range. On the other hand, both our proposed mechanism and ODAS can drastically reduce the number of non-

*1 In our simulation settings, the maximum number of neighbor sensor nodes is around 40. When we assume 20 ms is needed to send one message between sensor nodes as ODAS assumes, at least 0.8 s is needed for waking up to receive messages from all neighbor sensor nodes. In our mechanism, the wake-up duration of a sensor node is between $\tau_{\max} + \tau_{\min}$ and $2\tau_{\max}$. In the case of above settings of offset, the wake-up duration satisfies 0.8 s.



(a) Single distribution. (b) Combined distribution.

Fig. 7 Number of non-redundant sensor data.



(a) Single distribution. (b) Combined distribution.

Fig. 8 Number of message transmissions.

redundant sensor data in comparison with WAVE because of their transmission reduction mechanism. For example, the number of non-redundant sensor data with our proposed mechanism is up to 61% lower than that with WAVE.

In both ODAS and our proposed mechanism, the number of non-redundant sensor data decreases when the acceptable error range increases, as shown in Fig. 7 (a). This is because a larger number of sensor data satisfy the condition Eq. (2) when acceptable error range E becomes larger.

A comparison between our proposed mechanism and ODAS reveals that the number of non-redundant sensor data is between 22% and 44% larger with our proposed mechanism than with ODAS. This is because of the geographical difference of sensor data for determining redundancy, as described in Section 4.4. A sensor node in our proposed mechanism can use only downstream sensor nodes to determine the redundancy of sensor data whereas a sensor node in ODAS can use three sensor nodes whose locations construct a triangle containing the location of the sensor node. This situation in our proposed mechanism makes satisfaction of the condition Eq. (2) more difficult in the simulation experiments.

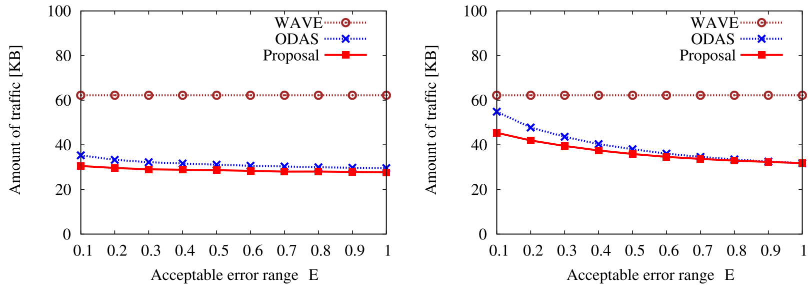
Figure 7 (b) shows the number of non-redundant sensor data when the combined distribution is used. The relationships among three mechanisms in the combined distribution have the same tendency of the single distribution. In ODAS and our proposed mechanism, since the combined distribution has higher spatial frequency than the single distribution, a larger number of sensor data are determined as non-redundant sensor data in the combined distribution as shown in Fig. 7. Since WAVE does not determine the redundancy of sensor data, the results in both the single distribution and the combined distribution are the same.

5.3 Evaluation of the Number of Message Transmissions and the Amount of Transmitted Data

Figure 8 (a) shows the number of message transmissions against the acceptable error range when the single distribution is used. In Fig. 8 (a), ODAS needs a much larger amount of message transmissions compared to WAVE and our proposed mechanism. From another viewpoint, the number of message transmissions in our proposed mechanism is up to 77% less than in ODAS. This is because in both WAVE and our proposed mechanism, sensor nodes broadcast a message only once in a cycle. On the other hand, in ODAS, some sensor nodes transmit messages in both the redundancy determination phase and the data gathering phase. In addition, ODAS assumes that each message contains a single sensor datum whereas a message in our proposed mechanism consists of more than one sensor datum. Therefore the number of message transmissions in ODAS becomes higher.

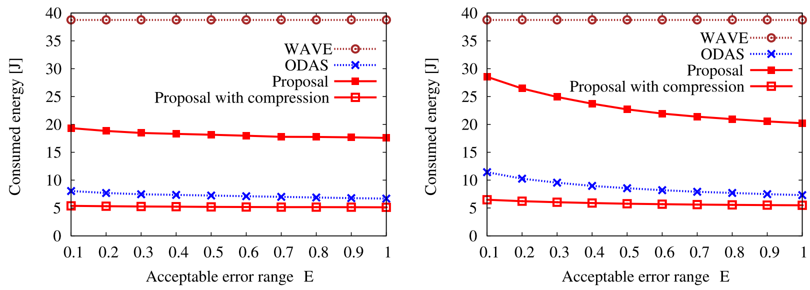
A comparison between our proposed mechanism and WAVE reveals that the number of message transmissions in our proposed data gathering mechanism to be up to 41% less. This is because a sensor node with no transmission sensor data stops broadcasting messages. When we compare the number of non-redundant sensor data as shown in Fig. 7 (a) and the number of message transmissions in our proposed mechanism as shown in Fig. 8 (a), we see that the number of message transmissions is higher. This shows that some sensor nodes forward sensor data from downstream sensor nodes to upstream sensor nodes although their sensor data is determined to be redundant.

Figure 8 (b) shows the number of message transmissions when the combined distribution is used. Since the number of non-redundant sensor data in the combined distribution is larger than that in the single distribution as described in Section 5.2, the number of message transmissions in the combined distribution becomes higher than that in the single distribution as shown in



(a) Single distribution. (b) Combined distribution.

Fig. 9 Amount of transmitted data.



(a) Single distribution. (b) Combined distribution.

Fig. 10 Consumed energy.

Fig. 8.

Figure 9 shows the amount of transmitted data against the acceptable error range. Among three mechanisms, our proposed mechanism achieves the smallest in both the single distribution and the combined distribution. For example, the amount of transmitted data in our proposed mechanism using the single distribution is up to 13% less than that in ODAS. Although the number of non-redundant sensor data in our proposed mechanism is higher than that in ODAS as described in Section 5.2, ODAS requires some sensor nodes to transmit messages in both the the redundancy determination phase and the data gathering phase as described in Section 5.3. As a result, the amount of transmitted data in our proposed mechanism becomes smaller.

5.4 Evaluation of Consumed Energy

Figure 10 (a) shows the consumed energy against the acceptable error range when the single distribution is used. As shown in Fig. 10 (a), the consumed energy in our proposed mechanism is up to 54% less than that in WAVE. This result is clear since our proposed mechanism is based on WAVE while integrating measures to reduce transmissions. Therefore, we can say that our proposed mechanism is more energy-efficient than WAVE.

When we compare the consumed energy in our proposed mechanism and that in ODAS, the consumed energy in ODAS is lower. Although the amount of transmitted data in our proposed mechanism is the smallest among the three mechanisms as described in Section 5.3, the number of sensor nodes that receive a message in our proposed mechanism is larger than that in ODAS. Since ODAS assumes ideal TDMA scheduling on the under layer, only sensor nodes related to message transmission/reception are awake. On the other hand, in our proposed mechanism, almost all the neighbor sensor nodes are awake when a sender node transmits a message. Therefore, those neighbor sensor nodes receive a

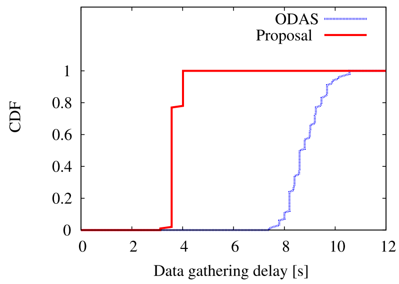


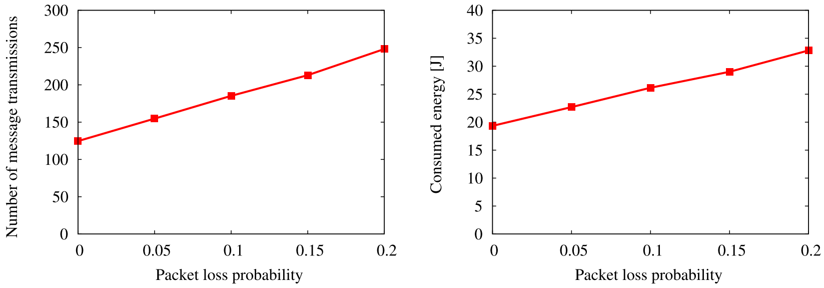
Fig. 11 CDF of data gathering delay.

message from the sender node even if the message is unnecessary.

However, when we consider an ODAS-related overhead such as TDMA scheduling, time synchronization, and routing as described in Section 3.2, we expect that the actual consumed energy would be much higher in a practicable ODAS situation. Furthermore, when we consider applying some data aggregation/compression mechanisms to our proposed mechanism at the timing of message transmission, the amount of traffic and consumed energy become much smaller. Here, ODAS can not apply such mechanisms directly since ODAS assumes that a message with single sensor datum is transmitted in a time slot. Fig. 10 also shows the consumed energy when our proposed mechanism applies a data compression mechanism which compresses multiple sensor data to the constant size of data. In this situation, the consumed energy in our proposed mechanism with compression is up to 33% less than that in ODAS. That is, by applying some data aggregation/compression mechanisms, our proposed mechanism would be more energy-efficient than ODAS.

5.5 Evaluation of the Data Gathering Delay

Figure 11 shows the cumulative distribution function (CDF) of data gathering delay for each mechanism. Since the distribution of data gathering delay with WAVE and that with our proposed



(a) Number of message transmission. (b) Consumed energy.

Fig. 12 Evaluation against packet loss probability.

mechanism are the same, we only show the distribution with our proposed mechanism. As shown in Fig. 11, the distribution of data gathering delay of our proposed mechanism is always lower than that of ODAS. Since ODAS consists of two phases (i.e., the redundancy determination phase and the data gathering phase), its data gathering delay is larger than that of our proposed mechanism. As a result, the data gathering delay with our proposed mechanism is 58% lower than with ODAS on average.

As shown in Fig. 11, the distribution of the data gathering delay of our proposed mechanism is a stepwise pattern since the data gathering delay of our proposed mechanism is determined by the maximum number of hops. On the other hand, the data gathering delay of ODAS is determined by the maximum number of hops and the results of time slot assignment.

5.6 Evaluation Against Packet Loss

Finally, to evaluate the fundamental performance of our proposed mechanism against packet loss probability, we conduct the following simulation evaluations.

In the simulation, we assume that a receiver node fails to receive a message from a sender node with certain probability p ($0 \leq p < 1$). When a sensor node fails to receive a message from a downstream node, the downstream node retransmits the same message. The retransmitted message is received at all awake neighbor sensor nodes. We assume that such retransmission is achieved by some under-layer mechanisms such as ACK or NACK mechanisms. We ignore the overhead related to achieve retransmission, since it is smaller than the overhead to retransmit a message. As temperature distribution, the single distribution is used. For the acceptance error range, we use $E = 0.1$.

Figures 12 (a) and (b) show the number of message transmissions and the consumed energy against packet loss probability p, respectively. When the packet loss probability is high, the number of retransmissions increases and the number of transmissions also increases as shown in Fig. 12 (a). At the same time, the consumed energy becomes higher due to the increase of retransmissions as shown in Fig. 12 (b). For example, when the packet loss probability is 0.1, the number of message transmissions increases by 50% and the consumed energy increases by 35%.

In this paper, to show the fundamental performance of our proposed mechanism, we do not conduct an in-depth evaluation considering radio characteristics such as instability and unreliability. These evaluations and extensions of our proposal, as well as experimental evaluations through implementation, will be the focus of future work.

6. Conclusions and Future Work

In this paper, we proposed an energy-efficient data gathering mechanism with data transmission reduction for WSNs. To drastically reduce energy consumption, we integrated our data gathering mechanism and our data aggregation mechanism. Through simulation experiments, we confirmed that with our proposed mechanism, the number of message transmissions can be reduced by up to 77% and the amount of transmitted data can be reduced by up to 13% compared to ODAS.

However, in this paper, we evaluated our proposed mechanism in an ideal environment. For example, in actual environments, radio transmission is unstable and unreliable. In addition, there are certain errors which may affect our proposed mechanism, such as sensing error, localization error, etc. Therefore, for future research, an in-depth evaluation and an extension of our proposed mechanism within dynamic and unstable environments are needed. We are also planning the implementation and practical experiments of our proposed mechanism within an actual environment such as an X-Sensor testbed [7].

Acknowledgments This work was partly supported by the Grant-in-Aid for Young Scientists (B) 23700079 of Japan Society for the Promotion of Science and the Grand-in-Aid for Scientific Research on Priority Areas 18049050 of the Ministry of Education, Culture, Sports, Science and Technology in Japan. The authors would like to thank anonymous reviewers for their helpful comments.

Reference

[1] Taniguchi, Y., Kanzaki, A., Wakamiya, N. and Hara, T.: Autonomous data gathering mechanism with transmission reduction for wireless sensor networks, *Proc. CCCA 2011* (2011).
[2] Akyildiz, I.F., Su, W., Sankarasubramaniam, Y. and Cayirci, E.: A survey on sensor networks, *IEEE Communication Magazine*, Vol.40, pp.102–114 (2002).
[3] Mainwaring, A., Culler, D., Polastre, J., Szewczyk, R. and Anderson, J.: Wireless sensor networks for habitat monitoring, *Proc. WSNA 2002*, pp.88–97 (2002).
[4] Werner-Allen, G., Johnson, J., Ruiz, M., Lees, J. and Welsh, M.: Monitoring volcanic eruptions with a wireless sensor network, *Proc. EWSN 2005*, pp.108–120 (2005).
[5] Stoianov, I., Nachman, L. and Madden, S.: PIPENET: A wireless sensor network for pipeline monitoring, *Proc. IPSN 2007*, pp.264–273 (2007).
[6] Ceriotti, M., Mottola, L., Picco, G.P., Murphy, A.L., Guna, S., Corra, M., Pozzi, M., Zonta, D. and Zanon, P.: Monitoring heritage building with wireless sensor networks: The torre aquila depelopment, *Proc. IPSN 2009*, pp.277–288 (2009).
[7] Kanzaki, A., Wakamiya, N. and Hara, T.: Energy-efficient task

scheduling and data aggregation techniques in wireless sensor networks for information explosion era, *Wireless Sensor Network Technologies for the Information Explosion Era*, Springer, pp.45–73 (2011).

[8] MEMSIC Inc.: MICAz, available from ⟨http://www.memsic.com/⟩.

[9] Oracle Corporation: SunSPOT, available from ⟨http://www.sunspotworld.com/⟩.

[10] Wang, L. and Xiao, Y.: A survey of energy-efficient scheduling mechanisms in sensor networks, *Mobile Networks and Applications*, Vol.11, No.5, pp.723–740 (2006).

[11] Taniguchi, Y., Wakamiya, N. and Murata, M.: A traveling wave based communication mechanism for wireless sensor networks, *Journal of Networks*, Vol.2, No.5, pp.24–32 (2007).

[12] Taniguchi, Y., Wakamiya, N. and Murata, M.: A communication mechanism using traveling wave phenomena for wireless sensor networks, *Proc. t2pWSN 2007*, pp.1–6 (2007).

[13] Rajagopalan, R. and Varshney, P.K.: Data-aggregation techniques in sensor networks: A survey, *IEEE Communications Surveys and Tutorials*, Vol.8, pp.48–63 (2006).

[14] Guestrin, C., Bodi, P., Thibau, R., Paski, M. and Madden, S.: Distributed regression: An efficient framework for modeling sensor network data, *Proc. IPSN 2004*, pp.1–10 (2004).

[15] Banerjee, T., Chowdhury, K. and Agrawal, D.P.: Distributed data aggregation in sensor networks by regression based compression, *Proc. MASS 2005*, pp.290–293 (2005).

[16] Iima, Y., Kanzaki, A., Hara, T. and Nishio, S.: Overhearing-based data transmission reduction for periodical data gathering in wireless sensor networks, *Proc. DMIEW 2009*, pp.1048–1053 (2009).

[17] Iima, Y., Kanzaki, A., Hara, T. and Nishio, S.: An evaluation of overhearing-based data transmission reduction in wireless sensor networks, *Proc. SeNTIE 2009*, pp.519–524 (2009).

[18] Brooke, T. and Burrell, J.: From ethnography to design in a vineyard, *Proc. DUX 2003* (2003).

[19] Mainwaring, A., Polastre, J., Szewczyk, R., Culler, D. and Anderson, J.: Wireless sensor networks for habitat monitoring, *Proc. WSNA 2002*, pp.88–97 (2002).

[20] Musăloiu-E., R., Tezis, A., Szlavecz, K., Szalay, A., Cogan, J. and Gray, J.: Life under your feet: A wireless soil ecology sensor network, *Proc. EmNets 2006* (2006).

[21] Wang, J., Ghosh, R.K. and Das, S.K.: A survey on sensor localization, *Journal of Control Theory and Applications*, Vol.8, No.1, pp.2–11 (2010).

[22] Goel, P. and Ermentrout, B.: Synchrony, stability, and firing patterns in pulse-coupled oscillators, *Physica D*, Vol.163, No.3–4, pp.191–216 (2002).

[23] Sundararaman, B., Buy, U. and Kshemkalyani, A.D.: Clock synchronization for wireless sensor networks: A survey, *Ad Hoc Networks*, Vol.3, No.3, pp.281–323 (2005).

Yoshiaki Taniguchi received his B.E., M.E., and Ph.D. degrees from Osaka University, Japan, in 2002, 2004, and 2008, respectively. He is currently an Assistant Professor at the Cybermedia Center of Osaka University. His research interests include wireless sensor networks, wireless mesh networks and object tracking systems. Dr. Taniguchi is a member of IEEE, IEICE, and IPSJ.

Akimitsu Kanzaki received his B.E., M.E., and Ph.D. degrees from Osaka University, Japan, in 2002, 2004, and 2007, respectively. He is currently an Assistant Professor at the Department of Multimedia Engineering of Osaka University. His research interests include wireless networks and communication protocols. Dr. Kanzaki is a member of IEEE, IEICE, IPSJ, and DBSJ.

Naoki Wakamiya received his M.E. and Ph.D. degrees from Osaka University, Japan, in 1994 and 1996, respectively. He is currently a Professor at the Department of Bioinformatic Engineering of Osaka University. His research interests include bio-inspired information communication technologies and self-organization of overlay networks, sensor networks, and mobile ad-hoc networks. He received the 2nd IEEE ComSoc Asia-Pacific Young Researcher Award in 2005. Dr. Wakamiya is a senior member of IEICE and a member of ACM, IEEE, and IPSJ.

Takahiro Hara received his B.E., M.E., and D.E. degrees from Osaka University, Japan, in 1995, 1997, and 2000, respectively. He is currently an Associate Professor at the Department of Multimedia Engineering of Osaka University. His research interests include distributed database systems in advanced computer networks, such as high-speed networks and mobile computing environments. Dr. Hara is a member of ACM, IEEE, IEICE, IPSJ, and DBSJ.