Content-based Wake-up Control for Wireless Sensor Networks Exploiting Wake-up Receivers

Junya Shiraishi[†], Hiroyuki Yomo[†], Kaibin Huang[‡], Čedomir Stefanović[*], and Petar Popovski[*]
[†]Graduate School of Science and Engineering, Kansai University, Japan
[‡]Dept. of EEE, The University of Hong Kong, Hong Kong
[*]Aalborg University, Denmark

***Abstract*—This paper proposes content–based control of wake–up receivers for data collection in wireless sensor networks. The wake–up procedure is designed with a goal of waking up only the subset of the sensor nodes which have the relevant data observations. This prevents the sensors with less relevant data from waking up and wasting energy, which is inevitable when employing conventional ID–based wake–up control. We apply the proposed content–based wake–up scheme to top–k query, where the sink attempts to collect information on the set of nodes that own top–k observations from the sensing field. Assuming medium access based on p–persistent CSMA, we design a content–based wake–up control scheme suited for the data collection of top–k query. We analyze the scheme theoretically in terms of data collection delay and energy–efficiency and compare it to the ID–based wake–up. The numerical results confirm the effectiveness of the proposed content–based wake–up control, especially when the number of sensor nodes is large.**

I. INTRODUCTION

Wireless sensor networks (WSNs) play a key role in supporting diverse Internet of Things (IoT) applications [1]. One of the key requirements on WSNs is *energy efficiency* that directly affects the lifetime of networks consisting of nodes operating on batteries.

Among the many solutions that have been proposed for energy efficient WSNs, this paper focuses on the concept of wake–up radio [2]. In this concept, each sensor node is equipped with two types of radio interfaces (I/Fs): one is a primary, main radio I/F used for data transmissions, and the other is a secondary radio, called wake–up receiver and dedicated to wake–up control. When there is no need to communicate, each node switches off its main radio I/F, and only the wake–up receiver is kept active. The wake–up receiver is designed so that its power consumption is much lower than that of the main radio I/F. When a node needs to communicate with the other node(s), it first transmits a wake–up signal to activate the main radio I/F(s) of target node(s). After successful wake–up signaling, data are transmitted through the main radio I/Fs. This realizes *on–demand* operations, where most energy is consumed by the main radio I/F only when needed, which improves the overall energy efficiency.

There have been mainly two types of wake–up signaling considered in the literature: *range–based* and *identity–based* wake–up [2]. With range–based wake–up, each wake–up receiver observes the signal level over a channel, and it activates the main radio I/F once it detects the level exceeding than a threshold. On the other hand, with identity–based wake–up, a unique ID is embedded into the wake–up signal, and if a wake–up receiver detects its own ID, its main radio I/F is activated. While the hardware configuration of range–based wake–up can be largely simplified, it is vulnerable to false wake–up, which can be avoided with identity–based wake–up at the cost of hardware complexity. The effectiveness of WSNs employing identity–based wake–up has been confirmed previously, e.g., in [3] and [4].

In this paper, we introduce a new type of wake–up signaling for realizing energy efficient data collection in WSNs: *content–based* wake–up. In many practical applications of WSNs, data collection from all sensor nodes may not be necessary. For instance, a user may only be interested in the set of sensor nodes that store anomalous data, e.g., temperature larger than a certain threshold. There have been also many studies on top–k query, in which users are interested in top–k observations of sensing data in a sensing field [5][6]. In these applications, only some of the sensor nodes store data in which users have interest. However, with the ID–based wake–up, the sink needs to wake up all sensor nodes and collect their data, since the sensing data of each sensor node is unknown for the sink when it conducts wake–up signaling. This makes sensor nodes storing undesired data to falsely wake up and waste energy. Therefore, in this paper, we propose a mechanism in which only sensor nodes owning the desired contents (data) wake–up and transmit their data. Specifically, the contributions of this paper are threefold:

- We propose a data collection method with content–based wake–up (CoWu), in which only sensor nodes storing sensing data larger than a specified threshold are woken up through wake–up signaling. Taking radio–on–demand sensor and actuator networks (ROD–SAN) [4] as an example of WSN employing wake–up receiver, we design the wake–up signal and wake–up mechanism, which realize the content–based wake–up.
- We apply CoWu to a scenario of top–k query, and propose an enhanced wake–up/data collection scheme called countdown–CoWu (CD–CoWu). The CD–CoWu attempts to collect sensing data of top–k observations by adapting the CoWu in a step–by–step manner.
- Assuming a p–persistent CSMA as a medium access control (MAC) protocol to resolve contention after the wake–up process, we theoretically analyze data collection delay and energy consumption of the proposed CD–CoWu and ID–based wake–up. As the exact calculations of the derived equations are highly complex, we also resort to an approximation based on the Markov Chain Monte Carlo (MCMC) method [7]. Our theoretical and simulation results confirm the validity of the derived equations as well as the effectiveness of the proposed CD–CoWu, especially when the number of sensor nodes is large.

Finally, we note that in the related literature, a wake–up process triggering nodes according to their sensed data for clustering sensors with similar reading is proposed in [8]. However, to the best of our knowledge, this is the first paper applying such a data (content)–based wake–up to data collections in WSNs with comparison to ID–based wake–up.

II. SYSTEM MODEL AND ID–BASED WAKE-UP SIGNALING

A. System Model

We consider a scenario where a number of sensor nodes are deployed over a sensing field, and a sink attempts to collect information observed by them. Each sensor node performs periodical measurements, storing the last observation for the potential reporting to the sink. In the rest of the text and without loss of generality, we assume that the sensors are measuring temperature. Each sensor node is equipped with two radio I/Fs: a main radio I/F used to transmit the observation to the sink and a wake–up receiver employed for wake–up signaling. While there are no communication requests from the sink, the sensor nodes switch their main radio I/Fs off and keep only their wake–up receivers active. This reduces energy consumed in stand–by state since the wake–up receiver consumes much less power than the main radio I/F. When the sink attempts to collect information from the sensing field, it sends wake–up request through dedicated wake–up signaling. When a sensor node detects a wake–up request at its wake–up receiver, it activates its main radio I/F. Then, the observation is transmitted by the main radio I/F with a single packet. We assume that the main radio I/F employs a p–persistent CSMA protocol for transmitting each packet, which has been shown to well–approximate the operation and performance of practical IEEE 802.15.4 MAC [9]. With our model of p–persistent CSMA, a channel is slotted and each node with a packet to transmit conducts carrier sensing at the beginning of a slot; hereafter, the length of time slot is denoted as δ $[s]$. If the channel is sensed to be free, the node transmits the packet with probability p. Otherwise, if the channel is detected to be busy, the node refrains from transmitting the packet and attempts to transmit it in the next slot. Once a collision happens, each node detects it through the absence of ACK and attempts to retransmit it with the above–mentioned operations. When detecting a successful transmission of packet by receiving an ACK from the sink node, each sensor node returns to a sleep state, i.e., it switches its main radio I/F off. For simplicity, we assume that packets can be lost only due to collisions, and ACK can be ideally transmitted from the sink to each sensor node. Furthermore, all nodes including sink are assumed to be located within communication/wake–up/carrier–sensing range of each other, i.e., we consider a single–hop network without the hidden terminal problem.

We assume that the sink is using top–k query [5][6] for collecting observations from the sensing field. In general, there can be two types of top–k query: one collecting information on k nodes with the highest readings, and the other collecting information on k highest values observed in the sensing field. In this work, we focus on the former type, where the sink attempts to collect information on k nodes with the highest readings. This can be useful, e.g., when we have limited resource to take some actions against sensing targets [5]. For simplicity, we assume in the paper that the temperature observed by each sensor node follows uniform distribution between some minimum and maximum values V_{min} and V_{max}, respectively, and that the probability for different nodes to observe exactly same value is negligibly small. In this case, data reported by top–k nodes correspond to top–k values observed in the sensing field.

B. ID–based Wake–up Signaling

We consider a wake–up control scheme that exploits the length of frame (i.e., the length of the energy burst) transmitted by the main radio I/F at a sink node. In this scheme, the wake–up receiver, which operates over the same channel as the main radio I/F, is designed in such a way that it only has capability to detect the length of frame observed over the channel. The detection of frame length can be realized with simple non–coherent envelope detection and on–off–keying (OOK) demodulation, which requires low power consumption [10]. This enables realization of wake–up signaling without adding an extra hardware to transmit the wake–up signal, as one can reuse the main radio I/F at the sink node as a transmitter of wake–up signal.

The conventional wake–up signaling is the ID–based wake–up (IDWu). For instance, a mapping can be made between different wake–up IDs and frame lengths, which is shared by all nodes. Then, the sink can transmit a frame with its length corresponding to the wake–up ID of the target wake–up node. For the detail of the wake–up signaling exploiting frame length detection, readers are referred to [4].

For IDWu described above, we can consider two types of IDs: Broadcast wake–up ID (BCWuID) and Unicast wake–up ID (UCWuID). BCWuID is a common ID to all sensor nodes, which triggers all nodes to wake up. On the other hand, UCWuID is unique to each node, and the sink can specify a single target node to wake up with an individual ID. We call the former type of wake–up as broadcast wake–up (BCWu), and the latter as unicast wake–up (UCWu).

C. Problem Definition

When we apply ID–based wake–up to the considered scenario of top–k query, the wasteful wake–up of nodes owning data out of top–k observations is inevitable. This is because the sink can only specify its target node to wake up with ID while it does not have any information on sensing data observed by each node. Therefore, the sink first needs to wake up all nodes by employing either BCWu or UCWu, and run data collection process through their main radio I/Fs. In this case, the nodes storing data that are out of top–k range also need to wake up at least once, which results in wasteful power consumption. In order to solve this problem, we propose a content–based wake–up, which enables the sink to only wake up nodes who have the desired data.

III. PROPOSED CONTENT–BASED WAKE–UP CONTROL

In this section, in order to solve the problem of ID–based wake–up described in Sec. II-C, we propose a wake–up control which activates nodes according to a condition on sensing data specified by a data collection node and actual sensed data at each node. Below, we first describe the basic operation of the proposed content–based wake–up (CoWu), followed by its extension to the scenario of top–k data collection.

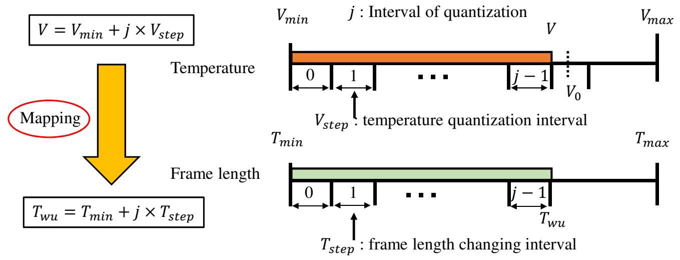

Fig. 1. Mechanism to decide wake–up frame length in CoWu.

A. *Content–based Wake–up*

CoWu is a wake–up control that embeds information on the desired sensing data into the wake–up signal transmitted by a data collection node, which enables only nodes storing the corresponding sensing data to wake up, thereby reducing wasteful wake–up of each node. In CoWu, each sensor node decides its wake–up frame length based on its own sensing data. As shown in Fig. 1, we prepare a mapping between the range of observable temperature [V_{min},V_{max}] and that of frame length employed for wake–up control [T_{min},T_{max}]. Note that T_{min} is selected so that it is longer than the frame length commonly employed for data transmissions. Each wake–up receiver regards only frames larger than T_{min} as wake–up frames in order to avoid false wake–up caused by background data traffic. We also define temperature quantization interval as V_{step}, and changing interval of frame length as T_{step}. Then, when the observed temperature V_o belongs to j-th interval, i.e., $V_{min} + j \times V_{step} \leq V_o < V_{min} + (j+1) \times V_{step}$, its wake–up frame length T_{wu} is set to $T_{wu} = T_{min} + j \times T_{step}$.

Fig. 2 shows an example of data collection with CoWu where the sink wants to collect data from nodes storing the observed data larger than V_{th}. As explained in Fig. 1, each node sets its wake–up frame length T_{wu} according to the observed temperature. Furthermore, the wake–up receiver of CoWu is designed so that only nodes satisfying $T_{wu} \geq T_{rx}$ wake up, where T_{rx} is the frame length detected at the wake–up receiver. The comparison between the detected and reference length of frames can be realized with the same hardware configuration as the wake–up receiver presented in [11], which requires low power consumption. The sink transmits a wake–up signal with its frame length of T_{th}, which corresponds to the temperature threshold of V_{th}. With these operations, in the example of Fig. 2, only sensor node 1, whose T_{wu} is longer than T_{rx}, wakes up, while sensor node 2, whose T_{wu} is shorter than T_{rx}, stays in a sleep mode. Thus, CoWu achieves activations of nodes with data satisfying the condition specified by the sink[1].

B. *Countdown Content–based Wake–up for top–k Query*

In this section, we explain our proposed scheme called countdown CoWu (CD–CoWu), which is designed so that CoWu can be applied to a scenario of top–k data collection. The basic idea of CD–CoWu is to gradually decrease the threshold of CoWu described in Sec. III-A until the sink manages to collect sensing data from top–k nodes.

Fig. 3 shows the operations of the proposed CD–CoWu. Our proposed CD–CoWu is detailed in Algorithm 1 and 2, which

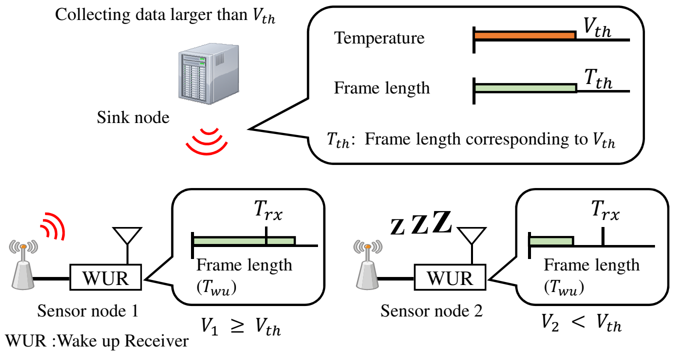

Fig. 2. Operations of the proposed CoWu.

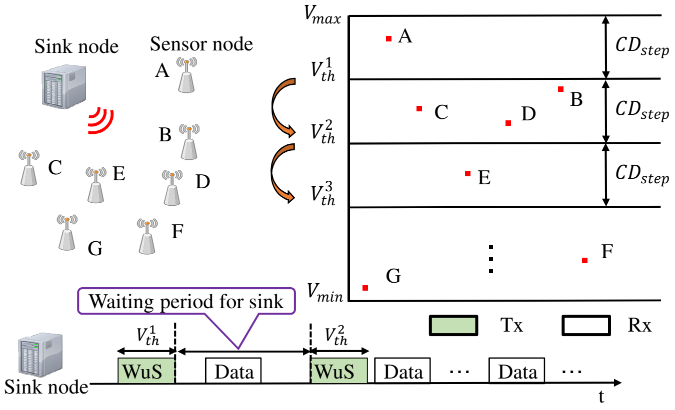

Fig. 3. Operations of the proposed CD–CoWu.

are the operation of the sink node and that of each sensor node, respectively. In CD–CoWu, the sink reduces the threshold of CoWu from an initial value by a parameter called countdown step (CD_{step}), by which it enlarges the range of temperature for data collections step-by-step. Here, the initial threshold of V_{th}^1 is set to $V_{max} - CD_{step}$. As shown in Fig. 3, the sink first generates a wake–up signal (WuS) whose length corresponds to V_{th}^1 and sends it to sensor nodes. The nodes which have T_{wu} larger than the length corresponding to V_{th}^1 (i.e., nodes storing temperature higher than V_{th}^1) wake up after successfully detecting the wake–up signal, and attempt to send a data packet to the sink. Note that, as mentioned in Sec. III-A, the length of the data frame is set to be shorter than T_{min}, therefore, it does not trigger wake–up of any other node. Here, the sink does not have any idea on how many nodes wake up with each wake–up trial, therefore, it needs to wait until it can make sure that there is no replying node. This can be achieved, e.g., by setting a time–out in a practical environment. The sensor nodes that have succeeded in data transmissions transit to a sleep state and are configured not to wake up for a certain period of time in order to prevent these nodes from being woken up again by the following wake–up signals transmitted by the sink. After the waiting period, the sink counts the number of packets it received, and sets it to n_k. Then, it checks whether the number satisfies $n_k \geq k$ or not. If not, the sink lowers the threshold to V_{th}^2 ($V_{th}^2 = V_{th}^1 - CD_{step}$), generates a new wake–up signal whose length corresponds to V_{th}^2, and sends it to sensor nodes. This operation will last until the last wake–up trial where the number of nodes with their data collected reach k, i.e., $n_k = k$. Through the above operations, the sink collects

---

[1] Note that, by employing different mappings between the observed data and frame length, we can flexibly change the condition for each node to wake up, e.g., nodes with data lower than a threshold or within some interval.

Algorithm 1 Operation of the sink in CD–CoWu

```
$V_{th} \leftarrow V_{max} - CD_{step}$
$n_k \leftarrow 0$
repeat
    Transmit a wake–up signal with the frame length $T_{th}$ that corresponds to $V_{th}$
    $T \leftarrow 0$
    Set a waiting period for sink to $T_{wait}$
    while $T < T_{wait}$ do
        if Data successfully detected then
            $n_k \leftarrow n_k + 1$
            Transmit an ACK frame to the corresponding node
        end if
        $T \leftarrow T + \delta$
    end while
    $V_{th} \leftarrow V_{th} - CD_{step}$
until $n_k \geq k$
```

Algorithm 2 Operation of each sensor node in CD–CoWu

```
loop
    In a sleep state
    Perform mapping between observed temperatures $V_o$ and $T_{wu}$
    Wait for detecting a wake–up frame $T_{rx}$
    if $T_{min} \leq T_{rx}$ and $T_{rx} \leq T_{wu}$ then
        Transit to the active state
        $Flag_{finish} \leftarrow 0$
        repeat
            if The results of CCA is free then
                Try to send a packet with probability $p$
                if Ack received within a ACK waitng period then
                    Transit to a sleep state
                    Suspend wake–up for a certain periods of time
                    $Flag_{finish} \leftarrow 1$
                end if
            end if
        until $Flag_{finish} == 1$
    end if
end loop
```

the information on top–k nodes with their temperatures and ID numbers without waking up nodes that do not belong to top–k range of temperatures specified by the sink. Note that the set of collected data with the above operations includes the set of top–k values that can be obtained by collecting information from all nodes and then selecting the top–k values since we assume in this work that all nodes observe different values as mentioned in Sec. II.

The parameter of CD_{step} affects the energy efficiency and data collection delay achieved by CD–CoWu. With larger CD_{step}, more nodes are simultaneously woken up with a single transmission of wake–up signal, which increases the congestion level and the number of wastefully activated nodes, thereby increasing the total energy consumption. This problem can be avoided by employing smaller CD_{step}, however, it increases data collection delay because we have more wake–up trials with no replying node. Therefore, CD_{step} should be optimized based on the target delay or energy efficiency, which will be conducted in Sec. V-B.

IV. THEORETICAL ANALYSIS OF IDWU AND CD–COWU

In this section, we analyze the data collection delay and energy efficiency achieved by CD–CoWu and IDWu, assuming that each sensor node operates with p–persistent CSMA MAC protocol to transmit data after the wake–up process. In [9], it is confirmed that p–persistent CSMA approximates well the characteristics of IEEE 802.15.4 MAC, which is one of the most common MAC protocols in practical environments of WSNs.

A. One-Shot Data Collection based on p–persistent CSMA

In BCWu and our proposed CD–CoWu, it can happen that multiple nodes attempt to transmit their individual packets to reply to the wake–up request from the sink. This type of traffic model is called one-shot data (OSD) model in [9], where each sensor node holds only a single packet to transmit. Under OSD model, each node that completes its transmission to the sink transits to a sleep state, and does not contend for the channel. In [9], data collection delay and energy consumption of nodes operating with p–persistent CSMA is theoretically analyzed when OSD traffic model is employed. According to [9], data collection delay $T_d(n_o)$ [s], which is defined as duration for n_o nodes ($n_o \geq 1$) with OSD model to complete their transmission, is

$$T_d(n_o) = \sum_{n=1}^{n_o} \frac{L - (L-1)(1-p)^n}{np(1-p)^{n-1}} \delta, \tag{1}$$

where L [slots] is the packet length in terms of slots. As a special case, we assume $T_d(0) = 0$. Further, the total energy consumption of the sensor nodes, $E_{total}(n_o)$ [J], can be computed as

$$E_{total}(n_o) = \sum_{n=1}^{n_o} \xi_R \delta \frac{L - (L-1)(1-p)^{n-1}}{p(1-p)^{n-2}} + \xi_T \delta \frac{L}{(1-p)^{n-1}}, \tag{2}$$

where ξ_R [W] and ξ_T [W] are the power consumptions of a node in the receive state and in the transmit state, respectively. We also assume $E_{total}(0) = 0$. By using the above results, we derive equations expressing data collection delay and energy consumption achieved by different wake–up control when they are applied to top–k data collection.

B. Analysis of conventional and proposed wake–up control

We first analyze delay and total energy consumption of conventional BCWu and UCWu, followed by the analysis of the proposed CD–CoWu. As already noted, the conventional BCWu and UCWu are required to activate all sensor nodes to collect top–k data since they are not able to conduct wake–up signaling according to requested/sensed data. Here, delay is defined as time required to collect top–k data from sensor nodes. As for energy consumption, considering that top–k query is repeated periodically by the sink, we focus on the total energy spent by all sensor nodes during a single cycle of top–k query. Considering that the wake–up receiver is always active during a cycle[2], the energy consumed by the wake–up receiver should be assumed to be same for all wake–up schemes. Therefore,

---

[2] We can also consider applying energy management, e.g., adaptive switch on/off of the wake–up receiver. However, the energy consumption of the wake–up receiver is so small that the impact of such an energy management on total energy consumption should be negligible.

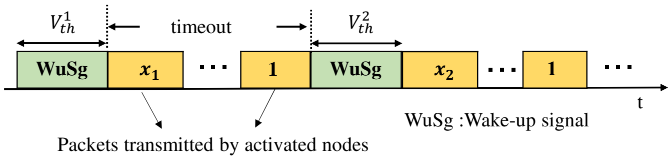

Fig. 4. Transmission model of CD-CoWu for theoretical analysis.

for simplicity, we neglect the energy consumed by the wake–up receiver, and only calculate energy consumed by the main radio. Furthermore, in the following analysis, all wake–up signals are assumed to be transmitted by the sink with $p = 1$ in the operation of p–persistent CSMA since no node is supposed to contend with the sink.

1) Analysis of BCWu scheme: With BCWu, the sink first sends a wake–up signal whose length corresponds to BCWuID in order to wake up all sensor nodes. Assuming that the total number of sensor nodes is N and all nodes detect the wake–up signal correctly, N nodes wake up and attempt to transmit their packets with p–persistent CSMA and OSD model. Therefore, delay of BCWu, $T_d^{BCWu}(N)$, is expressed as

$$T_d^{BCWu}(N) = T_d(N) + T_{BCWu}, \tag{3}$$

where T_{BCWu} [s] is the frame length corresponding to BCWuID. On the other hand, energy consumed by the main radios of sensor nodes is $E_{total}(N)$ [J] according to eq. (2). Therefore, energy consumption of BCWu is calculated as

$$E_{total}^{BCWu}(N) = E_{total}(N). \tag{4}$$

2) Analysis of UCWu scheme: In the case of top–k query employing UCWu, the sink individually activates each node by sending the wake–up signal whose length corresponds to UCWuID. In this way, the sink collects packets from sensor nodes one-by-one. For each wake–up signal sent by the sink, the number of wake–up nodes is one, therefore, data collection delay of UCWu is given by

$$\begin{aligned} T_d^{UCWu}(N) &= N{\cdot}T_d(1) + \sum_{i=0}^{N-1} T_{UCWu}(i) \\ &= N{\cdot}T_d(1) + T_{min}{\cdot}N + \frac{T_{step}}{2}N(N-1), \end{aligned} \tag{5}$$

where $T_{UCWu}(i)$ [s] is the frame length corresponding to the UCWuID assigned to sensor node i. Here, we design frame length of UCWu as $T_{UCWu}(i) = T_{min} + T_{step} \times i$ $(i = 0, 1, 2, \ldots)$. As for total energy consumption of UCWu, considering that N nodes separately transmit their data in reply to each wake–up request, it can be expressed as

$$E_{total}^{UCWu}(N) = N{\cdot}E_{total}(1). \tag{6}$$

3) Analysis of CD–CoWu scheme: With the proposed CD–CoWu applied to top–k query, as mentioned in Sec. III-B, the sink needs to transmit wake–up signal several times until it completes top–k data collections. For each wake–up signal, the number of activated and replying sensor nodes randomly varies depending on the distribution of sensed data and employed countdown step (CD_{step}). Thus, as shown in Fig. 4, different number of nodes activated by each wake–up signal try to send their packets to the sink with p–persistent CSMA. Here, for simplicity of analysis, we assume that timeout for each wake–up, during which the sink should wait for replies from the activated nodes, can be ideally set to a value that is required for all activated nodes to complete their transmissions. The sink continues wake–up requests until it collects data from more than k nodes, where the maximum number of possible wake–up trials W_{max} can be expressed as

$$W_{max} = \left\lceil \frac{V_{max} - V_{min}}{CD_{step}} \right\rceil. \tag{7}$$

Each node senses temperature which follows uniform distribution, therefore, the probability that a node is woken up at the n-th transmission of wake–up signal $P_{wake(n)}$ is given by

$$P_{wake(n)} = \begin{cases} \frac{CD_{step}}{V_{max}-V_{min}} & (n = 1, 2, \ldots, W_{max} - 1) \\ \frac{(V_{max}-V_{min})-(W_{max}-1)CD_{step}}{V_{max}-V_{min}} & (n = W_{max}). \end{cases} \tag{8}$$

Let random variables X_n (n = 1,...,W_{max}) denote the number of nodes to be activated in the n-th wake–up interval. Then, the variables X_n follow multinomial distribution, whose probability mass function is expressed as

$$\begin{aligned} &\mathbb{P}(X_1 = x_1, \ldots, X_{W_{max}} = x_{w_{max}}) \\ &= \begin{cases} N! \prod_{n=1}^{W_{max}} \frac{P_{wake(n)}^{x_n}}{x_n!} & (\sum_{n=1}^{W_{max}} x_n = N) \\ 0 & (\text{otherwise}), \end{cases} \end{aligned} \tag{9}$$

where N is the total number of sensor nodes. In CD–CoWu, the sink transmits wake–up signals until the number of nodes with their data collected exceeds k. Therefore, given the i-th realization (sample) of X_n as $s_i = \{x_1^i, x_2^i, \ldots, x_{W_{max}}^i\}$, the number of required transmissions of wake–up signal for s_i, defined as n_w^i, is minimum j $(1 \leq j \leq W_{max})$ satisfying

$$\sum_{n=1}^{j} x_n^i \geq k. \tag{10}$$

Then, by using eq. (1), mean total duration required for data transmissions with p–persistent CSMA (excluding duration for transmitting wake–up signal) can be calculated as

$$T_{CD-CoWu}^{Data}(N) = \sum_{i=1}^{M} \mathbb{P}(s_i){\cdot}\sum_{n=1}^{n_w^i} T_d(x_n^i), \tag{11}$$

where M is the total number of realizations of multinomial distribution given in eq. (9). On the other hand, as mentioned in Sec. III-B, the sink reduces the threshold of CoWu by CD_{step} for each wake–up trial. Then, the frame length of the n-th wake–up signal is set as

$$T_{CD-CoWu}^{WuS}(n) = T_{min} + T_{step}(W_{max} - n). \tag{12}$$

Here, for simplicity, we assume $CD_{step} = V_{step}$. Therefore, mean total duration required for transmitting wake–up signals is expressed as

$$T_{CD-CoWu}^{\Sigma WuS}(N) = \sum_{i=1}^{M} \mathbb{P}(s_i) \sum_{n=1}^{n_w^i} T_{CD-CoWu}^{WuS}(n). \tag{13}$$

Thus, total delay to complete top–k data collection with CD–CoWu is the sum of eqs. (11) and (13), and expressed as

$$T_d^{CD-CoWu}(N) = T_{CDCoWu}^{Data}(N) + T_{CD-CoWu}^{\Sigma WuS}(N). \tag{14}$$

Following the same derivation as above, total energy consumption of CD-CoWu, $E_{total}^{CD-CoWu}(N)$ can be calculated as follows

$$E_{total}^{CD-CoWu}(N) = \sum_{i=1}^{M} \mathbb{P}(s_i) \cdot \sum_{n=1}^{n_w^i} E_{total}(x_n^i). \tag{15}$$

V. NUMERICAL RESULTS AND DISCUSSIONS

A. Approximate analysis with MCMC

The exact calculation of delay and energy consumption of CD–CoWu, expressed by eqs. (14) and (15), requires consideration of all realizations of multinomial distribution given in eq. (9), whose number is $\binom{W_{max}+N-1}{N}$. This calculation becomes intractable as N becomes large, and/or W_{max} becomes large, which is the case when the value of CD_{step} is small. For this reason, in this paper, we resort to Markov Chain Monte Carlo (MCMC) method [7] to obtain approximate results. Specifically, we employ Metropolis algorithm [7] that generates sequence of samples (realizations) based on some probabilistic rules, as described below:

- STEP 0: [Set initial state] Set an initial state $X^{(0)} = \bigcup_{n=1}^{W_{max}} x_n^{(0)}$, where $x_n^{(0)}$ is the number of nodes belonging to the n-th wake–up interval under the initial state. We set the initial state so that the number of nodes in each interval is distributed as fairly as possible, with a constraint that the total number of nodes is N, i.e., $\sum_{n=1}^{W_{max}} x_n^{(0)} = N$. This constraint is the same in any arbitrary state (i.e., state $X^{(t)}$ is set s.t. $\sum_{n=1}^{W_{max}} x_n^{(t)} = N$).
- STEP 1: [Generate new sample] Create a new sample $X^{'}$ from the current state $X^{(t)}$. In each state, there are W_{max} intervals, from which we choose one interval i with its number of nodes n_i satisfying $n_i \geq 1$ and decrease n_i by 1. Then, we select another interval j with its number of nodes n_j satisfying $n_j < N$ and increase n_j by 1. Through these operations, new sample $X^{'}$ is generated.
- STEP 2: [Calculate transition cost] Calculate transition cost as $\frac{\mathbb{P}(X^{'})}{\mathbb{P}(X^{(t)})}$, by using probability mass function given by eq. (9)[3].
- STEP 3: [Update state] Generate a random number $r \in [0, 1]$, following uniform distributions, and decide the next state as follows

$$X^{(t+1)} = \begin{cases} X^{'} & (if\ r \leq \frac{\mathbb{P}(X^{'})}{\mathbb{P}(X^{(t)})}) \\ X^{(t)} & \text{(otherwise)}. \end{cases} \tag{16}$$

- STEP 4: [Go back to STEP 1].

Here, we define one round of STEP 1 to STEP 4, as Monte Carlo (MC) step, which is repeated Z times. For each MC step, by using the distribution of nodes represented by the state of $X^{(t)}$, delay and energy consumption are calculated through eqs. (10) – (15) and stored as $t^{(z)}$ and $e^{(z)}$, respectively. Then, after repeating MC step Z times, the approximate values of delay and energy consumption are respectively calculated as

$$\widehat{T}_{delay}^{CD-CoWu} = \frac{1}{Z}\sum_{z=1}^{Z} t^{(z)} \tag{17}$$

$$\widehat{E}_{total}^{CD-CoWu} = \frac{1}{Z}\sum_{z=1}^{Z} e^{(z)} \tag{18}$$

TABLE I
PARAMETERS EMPLOYED FOR NUMERICAL EVALUATIONS

Parameters	Values
Data transmission rate	100 kbps
Length of packet in time slots L	10
Time slot length δ	320 μsec
Power consumption in Transmit state ξ_T	55 mW
Power consumption in Receive state ξ_R	50 mW
Distribution of Temperatures $[V_{min}, V_{max}]$	$[0, 50]$
Interval of frame length T_{step}	0.16 msec [11]
Minimum frame length T_{min}	10.8 msec
k	5,10

A unique feature of Metropolis method is that each sample to calculate the expectations of a function as in eqs. (11), (13), and (15) is selected from the region where its probability is relatively high, while searching the other regions with probability r. After repeating a sufficient number of MC steps, we can expect to obtain approximate results that are close to exact solutions.

B. Performance evaluation of different wake–up schemes

In this section, we show numerical results of delay and energy consumption achieved by different wake–up schemes, which are obtained by the derived equations and computer simulations. The parameters employed for numerical evaluations are shown in Table I. The achievable performance of each wake–up scheme depends on different parameters such as p and CD_{step}. For instance, CD–CoWu can achieve high energy efficiency at the cost of delay by setting CD_{step} to small value with large p, which was confirmed with our preliminary study. Because of this trade-off relation, it is difficult to compare different wake–up schemes unless either target delay or energy is fixed. Therefore, in our numerical analysis, we investigate the effectiveness of CD–CoWu by comparing its total energy consumption with IDWu on condition that delay of CD–CoWu does not exceed that of IDWu. This upper bound on delay is obtained with the optimized transmission probability p of each IDWu (BCWu, UCWu), i.e., minimum data collection delay for different number of nodes within sensing field. We chose the optimal parameters as follows:

- UCWu: In UCWu, only a single node wakes up for each wake–up trial conducted by the sink, which means that there are no collisions. Therefore, we set the transmission probability of UCWu as $p = 1$. Here, we define the optimized data collection delay of UCWu for different number of sensor nodes N as $\tau_{UCWu}^{opt}(N)$.
- BCWu: In BCWu, the optimal transmission probability depends on the number of sensor nodes. In our preliminary study, we obtained the optimal probability p_{opt} for different number of sensor nodes N, which is employed in the following. Note that values of p_{opt} are obtained out of range from 0.01 to 0.25 with the step of 0.0001. Here, we define the optimized data collection delay of BCWu for different number of sensor nodes as $\tau_{BCWu}^{opt}(N)$.
- CD–CoWu: In CD–CoWu, the achievable performance depends not only on the transmission probability p but also on CD_{step}. With our preliminary study, we obtained the set of p and CD_{step} values, which minimizes the total energy consumption for different number of sensor nodes on condition that delay does not exceed $\tau_{UCWu}^{opt}(N)$

---

[3] More details on transition cost can be found in [7].

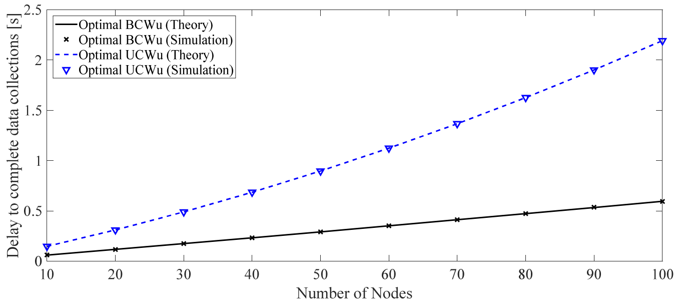

Fig. 5. Delay against the number of sensor nodes for BCWu and UCWu.

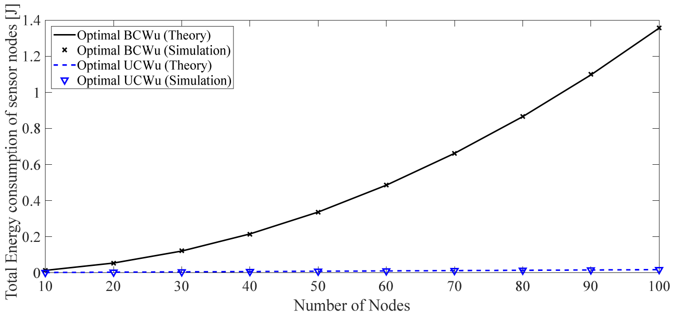

Fig. 6. Total energy consumption against the number of sensor nodes for BCWu and UCWu.

or $\tau_{BCWu}^{opt}(N)$. Note that the considered range of transmission probability p is the same as that of BCWu, and values of CD_{step} are varied from 0.5 to 5, with the step of 0.5, and 5 to 50 with the step of 5. Here, we define two sets of parameters for CD–CoWu $\{CD_{step}, p\}$: $\mathbb{U}(N)$, the optimized set for comparison with UCWu, and $\mathbb{B}(N)$, for comparison with BCWu, which are required to satisfy

$$\mathbb{U}(N) = \min_{CD_{step},p} \{E_{total}^{CD-CoWu}(N)\} \ s.t. \ \tau \ \le \ \tau_{UCWu}^{opt}(N) \ , \quad (19)$$

$$\mathbb{B}(N) = \min_{CD_{step},p} \{E_{total}^{CD-CoWu}(N)\} \ s.t. \ \tau \ \le \ \tau_{BCWu}^{opt}(N) \ , \quad (20)$$

where τ is data collection delay of CD–CoWu under the optimized set of parameters. These sets are employed in the following evaluations[4].

1) Comparison between BCWu and UCWu: Before investigating the efficiency of CD–CoWu, we evaluate and compare delay and total energy consumption achieved by BCWu and UCWu. Note that, with BCWu and UCWu, all nodes are activated, which means that its performance does not depend on k of top–k data collection. Figs. 5 and 6 respectively show delay and total energy consumption of BCWu and UCWu. First, from these figures, we can clearly see the agreement between theoretical results obtained with eqs. (3) – (6) and simulation results. Thus, we can confirm the validity of the derived equations. From Fig. 5, under the condition of optimal p, it can be seen that UCWu has worse performance than BCWu in terms of data collection delay. This is due to large amount of wake–up overhead, i.e., the transmission of wake–up signal is required before every data transmission in the case of UCWu. On the other hand, BCWu realizes better performance than UCWu because all sensor nodes can be woken up with only a single transmission of wake–up signal from the sink, and thanks to the optimization of p, nodes hardly experience collisions (note that collision is the major cause to deteriorate data collection delay with BCWu). Next, looking at Fig. 6, we can see that the total energy consumption of BCWu becomes significantly larger than that of UCWu as the number of nodes increases. This is because many nodes need to stay awake during contention process of p–persistent CSMA with BCWu while, with UCWu, only a single node wakes up and can transmit data and quickly go back to sleep without spending wasteful energy for contention process. From these results, we can confirm that BCWu is superior to UCWu in terms of delay while UCWu outperforms BCWu in terms of total energy consumption.

2) Efficiency of CD–CoWu: Figs. 7 and 8 show the comparison of total energy consumption of CD–CoWu with BCWu and with UCWu, respectively, when they are applied for top–5 and top–10 query. Note that, as described in Sec. V-B, delay of CD–CoWu is upper-bounded by that of BCWu and UCWu in those figures. From these results, we can first see that results of CD–CoWu obtained with our approximate analysis using the derived equations coincide with simulation results very well, which validates our approach. Here, we set the number of MC steps, Z, to 100,000. Next, from Fig. 7, it can be seen that total energy consumption of BCWu is much larger than that of CD–CoWu since, with BCWu, the sink needs to aggregate data from all sensor nodes with relatively small transmission probability (e.g., when the number of nodes $N = 100$, the transmission probability p needs to be set to 0.0111). On the other hand, CD–CoWu can achieve small energy consumption, especially when the number of nodes is large. This is because, in CD–CoWu, the sink can collect top–k data just by sending wake–up signals several times without waking up all sensor nodes. Under the constraint of data collection delay, total energy consumption of CD–CoWu is minimized by employing relatively small CD_{step} with high transmission probability p, which makes small number of activated nodes succeed in data transmissions quickly. Fig. 7 also shows that total energy consumption of CD–CoWu becomes larger as k increases due to the increase of the required number of top–k data. The value of energy consumption of CD–CoWu becomes larger with smaller number of nodes, due to the employed constraint of exceeding the data collection delay of the optimized BCWu. With smaller number of nodes, CD–CoWu needs to employ large CD_{step} in order to quickly collect top–k data from nodes sparsely spread within the temperature range. For instance, with top–10 query, the employed parameters of CD–CoWu with $N = 10$ are $CD_{step} = 50$ and transmission probability p is the same as that of BCWu. Therefore, CD–CoWu with $N = 10$ operates similarly to the optimized BCWu. Thus, we can interpret BCWu as the special case of CD–CoWu.

Finally, from Fig. 8, one can see that CD–CoWu achieves smaller energy consumption than UCWu when the number of sensor nodes is equal to or more than 20. As shown in Figs. 5 and 6, UCWu realizes higher energy efficiency compared with BCWu, however, it requires larger delay to complete data collection. Therefore, we can apply the small values of CD_{step} to CD–CoWu under the constraint of delay expressed

---

[4] We confirmed that top–k collections were completed against all sets of number of nodes while satisfying constraints of data collection delay of IDWu, which are given in eqs. (19) and (20).

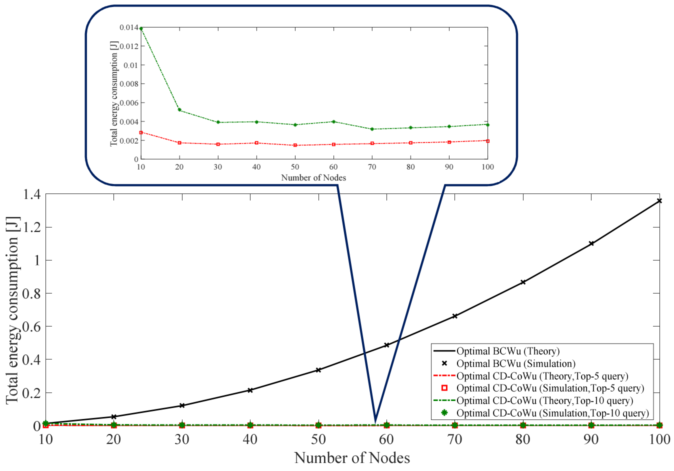

Fig. 7. Total energy consumption against the number of sensor nodes for BCWu and CD–CoWu.

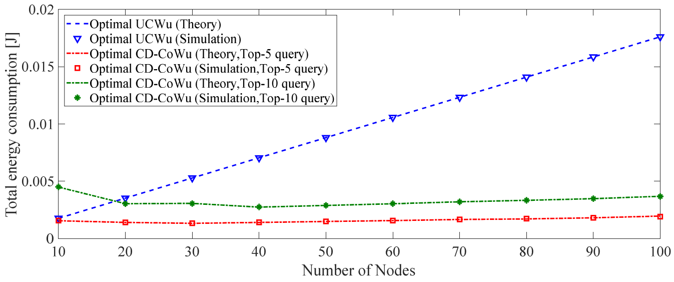

Fig. 8. Total energy consumption against the number of sensor nodes for UCWu and CD–CoWu.

in eq. (19). This can make the negative impact of congestion after the wake–up less significant. Furthermore, in CD–CoWu, unnecessary wake–up of nodes is suppressed while, in UCWu, all sensor nodes need to be woken up at least once. This makes it possible for CD–CoWu to achieve much smaller energy consumption than UCWu while delay constraint is satisfied.

Through the above discussions, we can confirm the superiority of CD–CoWu to IDWu, especially when the number of sensor nodes is large.

VI. CONCLUSIONS

In this paper, focusing on WSNs exploiting wake–up receivers, we have investigated how we can reduce wasteful wake–up when top–k data collections are considered. In ID–based wake–up, the sink needs to wake up all sensor nodes, which can result in large delay and energy consumption. In order to solve this problem, we have proposed CoWu and its enhanced scheme called CD–CoWu to be applied to top–k query. We have derived theoretical equations expressing data collection delay and total energy consumption of different wake–up schemes in order to analyze the effectiveness of our proposed scheme, assuming the employment of p–persistent CSMA as MAC protocol. In order to obtain numerical results on the proposed CD–CoWu, we have counted on an MCMC method. Through comparison between theoretical and simulation results, we have confirmed the validity of the derived equations and that of analysis based on the MCMC method. Our numerical results have shown that the proposed CD–CoWu achieves high energy efficiency, especially when the number of sensor nodes is large and top data to be collected is small.

In our future work, we will focus on the temperature distribution other than uniform distribution, e.g., exponential distribution, where observed data among sensors can be highly correlated, and consider a case where data collected from top–k nodes are different from top–k values. We will also investigate the effectiveness to dynamically change CD_{step} in accordance with distribution of observed data of sensor nodes. In addition, we will consider employing practical timeout mechanism for a sink in CD–CoWu.

ACKNOWLEDGEMENT

We would like to thank Dr. Seiji Miyoshi, Professor of Kansai University, for advice on approximate analysis using MCMC. This work was supported by JSPS KAKENHI Grant Number 19K04401. This work has been in part supported by the European Research Council (ERC) under the European Union Horizon 2020 research and innovation program (ERC Consolidator Grant Nr. 648382 WILLOW) and Danish Council for Independent Research (Grant Nr. 8022-00284B SEMIOTIC).

REFERENCES

[1] J. A. Stankovic, “Research Directions for the Internet of Things,” *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 3–9, March 2014.

[2] I. Demirkol, C. Ersoy, and E. Onur, “Wake-Up Receivers for Wireless Sensor Networks: Benefits and Challenges,” *IEEE Wireless Communications*, vol. 16, no. 4, pp. 88–96, August 2009.

[3] J. Oller et al., “Has Time Come to Switch From Duty-Cycled MAC Protocols to Wake-up Radio for Wireless Sensor Networks?” *IEEE Transactions on Networking*, vol. 24, no. 2, pp. 674–687, April 2016.

[4] H. Yomo, K. Abe, Y. Ezure, T. Ito, A. Hasegawa, and T. Ikenaga, “Radio-On-Demand Sensor and Actuator Networks (ROD-SAN): System Design and Field Trial,” *Proc. of Globecom 2015*, Dec. 2015.

[5] B. Malhotra, M. A. Nascimento, and I. Nikolaidis, “Exact Top-K Queries in Wireless Sensor Networks,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 10, pp. 1513–1525, Oct. 2011.

[6] C. Zhu, L. T. Yang, L. Shu, V. C. M. Leung, T. Hara, and S. Nishio, “Insights of Top-k Query in Duty-Cycled Wireless Sensor Networks,” *IEEE Transactions on Industrial Electronics*, vol. 62, no. 2, pp. 1317–1328, Feb. 2015.

[7] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[8] J.Blanckenstein, J.Klaue, and H.Karl, “Energy efficient clustering using a wake-up receiver,” *in Proc. of 18th Eur. Wireless Conf.*, pp. 1–8, April 2012.

[9] K. Yedavalli and B. Krishnamachari, “Enhancement of the 802.15.4 MAC Protocol for Scalable Data Collection in Dense Sensor Networks,” *Proc. 6th Int. Symp. Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT ’ 08)*, vol. 4, pp. 152–161, April 2008.

[10] H. Yomo, Y. Kondo, N. Miyamoto, S. Tang, M. Iwai, and T. Ito, “Receiver Design for Realizing On-Demand WiFi Wake-up using WLAN signals,” *in Proc. of IEEE GLOBECOM 2012*, Dec. 2012.

[11] H. Yomo, T. Kawamoto, K. Abe, Y. Ezure, T. Ito, A. Hasegawa, and T. Ikenaga, “ROD-SAN: Energy-efficient and High-response Wireless Sensor and Actuator Networks employing Wake-up Receiver,” *IEICE Transactions on Communications*, vol. E99-B, no. 9, pp. 1998–2008, Sept. 2016.