DEARER: A Distance-and-Energy-Aware Routing with Energy Reservation for Energy Harvesting Wireless Sensor Networks

Yunquan Dong†, Jian Wang†, Byonghyo Shim†, *Senior Member, IEEE*,
and Dong In Kim‡, *Senior Member, IEEE*
†Dept. of Electrical & Computer Engineering, Seoul National University, Korea
‡School of Information & Communication Engineering, Sungkyunkwan University, Korea
ydong@snu.ac.kr, jwang@islab.snu.ac.kr, bshim@snu.ac.kr, dikim@ieee.org

Abstract

In an energy harvesting wireless sensor network (EH-WSN), an energy-harvesting unit is used at each sensor node to collect energy from ambient environments. This ensures a perpetual energy supply to the network. However, due to limitation of energy intensity and randomness of energy arrivals, there is a mismatch between the real-time energy supply and the actual demand at each sensor node, resulting in irregular and sporadic energy shortage events. With an aim to combat the energy shortage issue in EH-WSNs, we develop a cluster-based routing protocol, referred to as the distance-and-energy-aware routing with energy reservation (DEARER). In DEARER, all nodes take turns to serve as cluster head (CH) nodes to forward packets, which are collected from non-CH nodes, to the sink. In particular, those nodes with large energy arrival rate or being closer to the sink are endowed with high chance to serve as the CH nodes. When a node works in the non-CH mode (i.e., without relaying), it reserves a portion of the harvested energy for future use in the CH mode. The proposed strategy cuts down energy-consuming transmissions from nodes to the sink while ensuring nodes more energy for the CH mode, thereby reducing energy shortage events at the CH nodes. By theoretical analysis and numerical simulations, we study the performance of DEARER in terms of the network energy efficiency, transmission ratio, and CH-outage probability. Our results demonstrate that the proposed DEARER protocol outperforms direct transmission and also approaches the genie-aided routing where CH nodes are selected based on real-time energy information of nodes.

Index Terms

Routing, energy harvesting, energy reservation, energy efficiency, outage probability.

I. INTRODUCTION

Wireless sensor network (WSN) has been widely deployed for various purposes, such as remote environment monitoring [1], health-care [2], and air quality monitoring [3]. By employing a large number of low-power sensors, a WSN can collect desired information in a distributed and self-organized manner. Such feature has made the WSN a key enabler to realize the Internet-of-Things (IoT) and green communication era [4], [5]. One well-known concern in WSNs is that the network lifetime is greatly restrained by the battery capacity of sensor nodes [1]–[3]. To address this issue, many efforts have been made in recent years, focusing on either exploiting sustainable energy to diversify the energy supply or greening network to reduce energy expenditure. Among these, energy harvesting has received considerable attention owing to its ability in extending the lifetime of WSNs [6]. By utilizing an energy harvesting unit and an energy buffer, each sensor node of a WSN can harvest energy (e.g., solar and wind power) from environments, thereby ensuring an unlimited energy supply to each node. This type of network is referred to as the energy harvesting WSN (EH-WSN) [6]. Sustainable energy supply techniques have also been used in hyper cellular networks [7]. The applications can be further extended with wireless energy transfer technologies [8].

As an alternative means to prolong the network lifetime of WSNs, greening network with energy efficient routing methods has been studied extensively [9]–[14]. The methods roughly fall into two categories: approaches relying on flat routing protocols [9], [10] and schemes based on hierarchical routing protocols [11], [12]. While each node plays the same role in the flat protocols, some nodes in the hierarchical protocols are used as cluster head (CH) nodes to forward packets of non-CH nodes to the sink. In general, approaches based on the hierarchical protocols outperform those using the flat protocols. Also, by incorporating more information of the network (e.g., geographic information and real-time energy status), further improvement of performance can be achieved [13], [14]. Such protocols, typically employing a central controller, are particularly preferable in the EH-WSN scenarios since the central controller can optimize routing paths by using the real-time energy status of nodes [15], [16]. In addition, protocols without a central controller yet achieving the centralized performance have also been proposed. For example, the DEHAR protocol identifies the most energy-efficient route with the aid of real-time node-to-node energy status exchanges [19]. However, due to extensive message exchanges, these protocols often suffer from heavy signaling overhead [19] and also are inefficient in coping with the randomness of energy arrivals in the EH-WSN, resulting in occasional energy shortage events. To be specific, since the available ambient energy is susceptible to environmental factors (e.g., light intensity for solar power and temperature for thermal power), the harvested energy at each node is irregular and random, and hence may not meet the demand. In fact, since CH nodes transmitting packets to the remote sink is more energy-consuming than non-CH nodes transmitting packets to adjacent CH nodes, energy shortage often occur to those CH nodes.

The primary purpose of this paper is to propose a new routing protocol, referred to as the *distance-and-energy-aware routing with energy reservation* (DEARER), to overcome the energy shortage problem and also improve the energy efficiency of EH-WSNs. Our approach is inspired by the observation that energy shortage often occurs at the CH nodes. In the DEARER paradigm, the network operates in rounds. In a round, each node works either in the non-CH mode or in the CH mode, where a node is said to be in the CH (non-CH) mode if it acts as a CH (non-CH) node. When a node is in the non-CH mode, a portion of the harvested energy is stored as a "strategic reserve" for future use. In this way, each node can reserve more energy for the upcoming CH mode, thereby reducing the probability of energy shortage. Another important feature of DEARER to reduce the energy shortage of CH nodes is to balance the energy consumption over the network. To be specific, in each round, "*enabler*" nodes, which are closer to the sink (i.e., lower energy consumption in transmission) or have higher energy arrival rate (i.e., more resistant to energy shortage), are endowed with more chance to serve as CH nodes. This essentially lowers the average energy consumption and provides CH nodes with more energy, thus reducing the probability of energy shortage. As a result, more packets can be delivered to the sink successfully, and hence the overall energy efficiency of the network is improved.

The main contributions of this paper are as follows:

- We propose a cluster-based routing protocol called DEARER. By employing an energy-and-distance-aware head selection algorithm and a proactive energy reservation policy, DEARER reduces the energy shortage events at CH nodes and improves the network energy efficiency.
- We present a model for evaluating the performance of routing algorithms in EH-WSNs. In the model, we use the *network energy efficiency* to characterize how many packets can be successfully delivered to the sink for given amount of energy. We also parameterize the transmission efficiency and reliability at each node in terms of the *transmission ratio* and

CH-outage probability. For these metrics, we provides explicit analytical expressions and efficient numerical evaluations.

- We compare the performance of the DEARER protocol with direct transmission (i.e., the worst strategy) and genie-aided routing (i.e., the best strategy). Our numerical and simulation results demonstrate that DEARER outperforms direct transmission and approaches genie-aided routing for most of realistic scenarios.

The rest part of the paper is organized as follows. In Section II, we present the network model, multiple access model, and energy harvesting model. We describe the DEARER routing protocol in Section III. The performance analysis of DEARER, including transmission ratio, CH-outage probability, and network energy efficiency, are provided in Section IV. Numerical and simulation results are presented in Section V. Finally, we conclude our work in Section VI.

II. System Model

A. Network Model

We consider an EH-WSN with N sensor nodes and one sink node. The sensor nodes collect information independently and send it to the sink according to a given routing protocol. Each sensor node is equipped with an energy harvesting unit so that it can also collect energy (e.g., solar energy and wind energy) from ambient environments. The harvested energy is assumed to be used only for the radio frequency (RF) unit. The sensor nodes are deployed uniformly in a square area of $2L \times 2L$ m^2 (see Figure 1). Let (x_n, y_n) and (x_s, y_s) be the position of an arbitrary node and the position of the sink, respectively, the distance between them is $d_n = \sqrt{(x_n - x_s)^2 + (y_n - y_s)^2}$. For two nodes with position (x_{n_1}, y_{n_1}) and (x_{n_2}, y_{n_2}), the piecewise distance is $d_{n_1 n_2} = \sqrt{(x_{n_1} - x_{n_2})^2 + (y_{n_1} - y_{n_2})^2}$.

In this paper, we consider the cluster-based routing protocol for an energy hungry EH-WSN, in which sensor nodes are classified into two categories: CH nodes and non-CH nodes. Each non-CH node sends information to the nearest CH node, and the CH node forwards the collected information to the sink. A CH node together with its non-CH members forms a *cluster*. In the cluster-based routing, the network works in *round*s and each round contains two phases: the *setup phase* and the *steady phase*. In the setup phase, the CH nodes are selected and the clusters are formed. In the steady phase, each non-CH node transmits the collected information to its CH node. Upon receiving all the packets from its non-CH members, a CH node aggregates

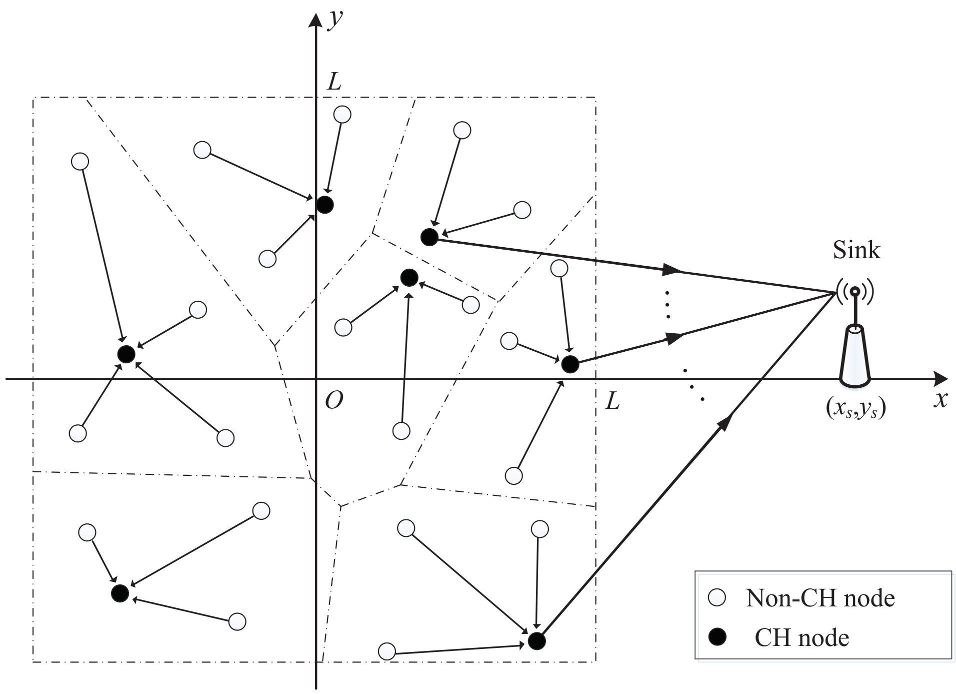


Figure 1. The DEARER based network model.

the packets into one and forwards it to the sink. This type of hierarchical routing is beneficial in various aspects: First, a large portion of direct transmissions to the sink, which are energy-consuming, are prevented, and hence the quick energy drainage of nodes can be avoided. Second, by aggregating information at CH nodes, the amount of information to be transmitted to the sink can be reduced significantly. For example, a CH node may compress the information collected from adjacent non-CH nodes with compressed sensing (CS) techniques [17], [18] and send only the compression to the sink. Third, periodic re-assignment of CH nodes can balance the energy consumption among nodes and thus improve the network lifetime. This is especially important for EH-WSNs since sensors can ensure time to accumulate the needed energy. Finally, the hierarchical routing strategy can be readily applied to large-scale networks.

B. Multiple Access Model

We assume that the time is slotted and each node communicates with its destination in a time division multiple access (TDMA) manner.[1] As shown in Figure 2, each frame consists of N slots

[1]Our results are readily extendable to other multiple access systems, such as the frequency division multiple access (FDMA) and the code division multiple access (CDMA) systems, by using sub-band or codeword (instead of time slot) as resource blocks.

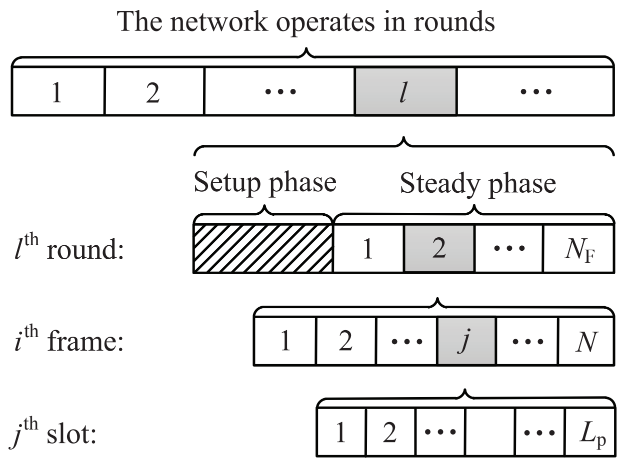


Figure 2. The frame structure under DEARER strategy.

and each of them is uniquely assigned to a node. In addition, the length of a slot is t_s during which a packet of L_p bits is transmitted. As mentioned, a round consists of a setup phase and a steady phase. Although the setup phase is an overhead to the network operation, the portion of this overhead is short compared to the steady phase that consists of N_F frames of information transmissions. After receiving all packets from its non-CH members, a CH node transmits the aggregated packet to the sink as long as it has enough energy.

C. Energy Model and Channel Model

In this subsection, we describe the energy model of how sensor nodes harvest, store, and use the energy. Before we proceed, we provide key assumptions used in our analysis.

i) The capacity of battery is assumed to be infinity. For example, the capacity of a small button battery is more than 200 milliampere hour (mAh), which is large enough for most of energy harvesting scenarios [21].

ii) Energy arrivals may occur at any time but the harvested energy can be used only from the next frame.

iii) Packets received at each CH node are aggregated into one.

iv) The energy arrival at each node follows a stationary and ergodic Poisson process with parameter λ_n [22]:

$$\Pr\left\{\frac{E_n^{\mathrm{h}}(i)}{e_0}=k\right\}=\frac{\lambda_n^k}{k!}e^{-\lambda_n},$$

where $E_n^{\mathrm{h}}(i)$ is the harvested energy by node n in the i-th frame and e_0 is the unit of the arriving energy.

v) The energy arrival rate λ_n of each node is different from each other. Let $F_\lambda(x)$ be the cumulative distribution function (CDF) of λ_n of the network. Then $\mu_\lambda = \mathbb{E}_\lambda[\lambda_n]$ is the average energy arrival rate over the network.[2]

We consider both free space (d^2-power loss) and multipath fading (d^4-power loss) channel models. Specifically, we use the free space (fs) model if the distance is shorter than d_0 and the multipath fading (mp) model otherwise. At the RF unit, the amount of energy to transmit one-bit information to the receiver is

$$e_{\mathrm{RF}}(d) = \begin{cases} \epsilon_{\mathrm{fs}} & d \leq 1, \\ \epsilon_{\mathrm{fs}} d^2 & 1 < d \leq d_0, \\ \epsilon_{\mathrm{mp}} d^4 & d_0 < d, \end{cases}$$

where ϵ_{fs} and ϵ_{mp} are positive constants and the unit of energy is Joule (J) [11]. Other than this, an energy e_{elec} for performing signal processing (e.g., the digital coding, modulation, filtering, and spreading) is needed. Thus, the total energy to transmit a packet from a non-CH node n to the corresponding CH node m is

$$E_n^{\overline{\mathrm{CH}}} = L_{\mathrm{p}}(e_{\mathrm{elec}} + e_{\mathrm{RF}}(d_{mn})). \tag{1}$$

Let $E_n^{\mathrm{h}}(i)$ and $E_n^{\mathrm{r}}(i)$ be the harvested energy and the remaining energy of node n at the i-th frame, respectively. A non-CH node n transmits a packet to its CH node in the assigned slot if

$$E_n^{\mathrm{r}}(i) \geq E_n^{\overline{\mathrm{CH}}}.$$

Otherwise, the node needs to wait and accumulate more energy. At the end of a frame, the remaining energy of node n is updated by

$$E_n^{\mathrm{r}}(i+1) = E_n^{\mathrm{r}}(i) + E_n^{\mathrm{h}}(i) - E_n^{\overline{\mathrm{CH}}} I_{E_n^{\mathrm{r}}(i) \geq E_n^{\overline{\mathrm{CH}}}}, \tag{2}$$

where I_A is the indicator function (1 if A is true and 0 otherwise).

[2]The scenario where nodes have the same energy arrival rate can be included by setting $F_\lambda(x) = \delta(0)$ as a unit impulse function.

For CH node m, the consumed energy $E_m^{\text{CH}}(i)$ in a frame consists of 1) the energy to receive signals from its cluster members, 2) the energy to aggregate the signals, and 3) the energy to transmit the aggregated packet to the sink. That is,

$$E_m^{\text{CH}}(i) = \left(L_{\text{p}} \sum_{k=1}^{\kappa_m - 1} I_{\text{Tx}_k(i)} (e_{\text{elec}} + e_{\text{da}}) + L_{\text{p}} e_{\text{da}} I_{\kappa_m > 1} + e_n \right) I_{\text{Tx}_m(i)}, \tag{3}$$

where κ_m is the number of nodes in the cluster associated with CH node m, e_{da} is the energy to aggregate each bit, $\text{Tx}_n(i)$ is the event that node n transmits a packet in frame i, and

$$e_n = L_{\text{p}}(e_{\text{elec}} + e_{\text{RF}}(d_n)) \tag{4}$$

is the required energy for node n to directly transmit a packet to the sink.

After receiving all the packets from its cluster members, the CH node performs some signal processing operations to aggregate these packets. If the CH node has enough energy, i.e., if $E_m^{\text{r}}(i) \geq E_m^{\text{CH}}(i)$, it will forward the aggregated packet to the sink. Otherwise, an outage occurs. At the end of a frame, the remaining energy of CH node m is updated by

$$E_m^{\text{r}}(i+1) = E_m^{\text{r}}(i) + E_m^{\text{h}}(i) - E_n^{\text{CH}}(i) I_{E_m^{\text{r}}(i) \geq E_n^{\text{CH}}(i)}.$$

We also need a measure characterizing the capability of the network in harvesting energy from ambient environments.

Definition 1: The *subsistence ratio* of an EH-WSN is the ratio between the sum of the harvested energy by the network in a frame and the total energy required for every node to directly transmit a packet to the sink:

$$\rho = \frac{\sum_{n=1}^{N} \lambda_n e_0}{\sum_{n=1}^{N} e_n}. \tag{5}$$

For most of realistic scenarios, ρ is chosen neither too small nor too large in consideration of operational practicability and cost savings.

III. The Distance-and-Energy-Aware Routing with Energy Reservation

A. Routing Design Considerations

For a cluster-based routing, most of nodes act as energy-saving non-CH nodes and transmit packets to their respective CH nodes. However, CH nodes remain energy-consuming since they need to transmit packets to the remote sink. Thus, it is highly likely that a CH node may suffer from energy shortage. Since the energy shortage event at a CH node blocks the transmission of the collected information from non-CH members, it degrades the network energy efficiency significantly. The reduction of energy shortage events at CH nodes is, therefore, crucial for improving energy efficiency of the network.

The DEARER protocol exploits two key features to address the energy shortage issue of CH nodes. First, CH nodes are re-assigned periodically according to a head selection rule where "enabler" nodes, which are in nature resistant to energy shortage, are encouraged to serve as CH nodes. Second, a portion of the harvested energy during the non-CH mode is reserved for future use in the CH mode. By doing so, the random energy arrivals are reshaped towards the desired energy profile that matches with the energy demand.

In the CH selection rule, both the energy harvesting capability and the energy-consuming level of each node are considered. Specifically, we characterize the relationship between the energy harvesting capability and the energy-consuming level using two parameters: 1) the average energy arrival rate λ_n and 2) the required energy e_n for a node to transmit a packet to the sink. It is clear that a node adjacent to the sink (i.e., with smaller e_n) is more energy efficient in forwarding packets to the sink. Also, a node with large energy arriving rate λ_n is preferred to serve as the energy-consuming CH node. In view of these points, we define the *priority* of each node serving as a CH node as

$$q_n = \frac{\frac{\lambda_n}{e_n}}{\sum_{k=1}^{N} \frac{\lambda_k}{e_k}}, \quad n = 1, 2, \cdots, N. \tag{6}$$

Moreover, to confront with the randomness of energy arrivals, each node saves a portion p_n^{sv} of the harvested energy in the non-CH mode, where the portion p_n^{sv} is referred to as *energy reservation ratio*. The saved energy $\widehat{E}_n^{\text{sv}}$ is stored into the energy buffer and will be used only when the node becomes a CH node so that each CH node is more resistant to energy shortage.

B. The Routing Protocol

The overall algorithm is depicted in Table 1. In the network initialization phase, each node is informed of its slot assignment and priority q_n through signaling exchange between nodes and the sink. In the setup phase of each round, CH nodes are selected and the corresponding clusters are formed (see Section III-C for details). Next, each CH node measures two types of energy in the buffer: 1) the remaining energy $E_n^{\text{r}}(N_{\text{F}})$ corresponding to the residual energy at

Algorithm 1 The DEARER protocol

Initialization:

```
Sink broadcasts the slots allocation, $M$, $\epsilon_{\text{fs}}$, $\epsilon_{\text{mp}}$, $e_{\text{elec}}$, $\epsilon_{\text{da}}$ at power $P_s$;
Each node decodes the slots allocation, $M$, $\epsilon_{\text{fs}}$, $\epsilon_{\text{mp}}$, $e_{\text{elec}}$, and $\epsilon_{\text{da}}$;
Each node measures the received power to determines $d_n$ and $e_n$;
Each node sends $\vartheta_n = \frac{\lambda_n}{e_n}$ to the sink;
The sink calculates and broadcasts $\sum_{n=1}^{N} \vartheta_n$;
Each node decodes $\sum_{n=1}^{N} \vartheta_n$ and calculates its own priority $q_n$ by (6);
```

Iteration:

```
for each round $l$ do
    /* Setup Phase: */
        Head selection: using Algorithm 2;
        Cluster formation: each node joins its nearest CH node;
        Energy update: each CH nodes updates remaining energy $E_n^{\text{r}}(0)$ by (7);
    /* Steady Phase: */
    for each frame $1 \le i \le N_{\text{F}}$ do
        Each non-CH node transmits a packet if $E_n^{\text{r}}(i) \ge E_n^{\overline{\text{CH}}}$;
        Each CH node receives packets from its non-CH members, aggregates them
        into one packet, and forwards it to the sink if $E_m^{\text{r}}(i) \ge E_m^{\text{CH}}$;
        Each node harvests energy and each non-CH nodes reserves energy;
        Each node updates remaining energy $E_n^{\text{r}}(i)$ by (2) or (5).
    end for
end for
```

the end of previous round and 2) the reserved energy $\widehat{E}_n^{\text{sv}}$ when the node acts as a non-CH node in previous rounds. Thus, the available energy at the beginning (of the first frame) of a round is

$$E_n^{\text{r}}(0) = E_n^{\text{r}}(N_{\text{F}}) + \widehat{E}_n^{\text{sv}}. \tag{7}$$

C. *Head Selection*

In DEARER, we use the *desired number* M of CH nodes as a key controlling parameter. To be specific, the average number of CH nodes in network can be adjusted by varying M.

Let C_n be a flag variable indicating whether node n has recently served as a CH node. Given C_n, we denote $p_n^{\text{cond}}(l)$ as the probability of node n being a CH node in the l-th round. At the beginning of the operation, C_n is initialized as zero. We set $C_n = 1$ once a node n is elected as a CH node and reset $C_n = 0$ after every $\max\{\lceil \frac{1}{Mq_n} - 0.5 \rceil, 1\}$ rounds. If $\frac{1}{Mq_n} \le 1.5$, it is clear that $\max\{\lceil \frac{1}{Mq_n} - 0.5 \rceil, 1\} = 1$, which means that node n elects itself as a CH node with

Algorithm 2 The head selection process

Initialization:

if $l \bmod \lceil \frac{1}{Mq_n} - 0.5 \rceil = 0$ **then**
    Set $C_n = 0$;
end if

Iteration:

for each node n **do**
    if $C_n = 0$ **then**
        generates a uniformly distributed number $x \sim \mathcal{U}(0, 1)$;
        if $x \le p_n^{\text{cond}}(l)$ **then**
            Elects itself as CH node;
            Set $C_n = 1$;
            Broadcasts its node ID n using a beacon message in its assigned slot;
        end if
    end if
end for

probability $p_n^{\text{cond}}(l) = 1$ in each round. Otherwise, node n serves as a CH node with probability

$$p_n^{\text{cond}}(l) = \begin{cases} \dfrac{1}{\lceil \frac{1}{Mq_n} - 0.5 \rceil - l \bmod \lceil \frac{1}{Mq_n} - 0.5 \rceil} & C_n = 0, \\ 0 & C_n = 1. \end{cases}$$

In particular, $p_n^{\text{cond}}(l)$ is the probability of node n being a CH node on condition that node n was not a CH node from round $\left\lfloor \frac{l}{\lceil \frac{1}{Mq_n} - 0.5 \rceil} \right\rfloor \left\lceil \frac{1}{Mq_n} - 0.5 \right\rceil + 1$ to round $l - 1$. Clearly $p_n^{\text{cond}}(l)$ varies from round to round. Moreover, we denote $\overline{p_n^{\text{cond}}}$ as the *unconditional probability* of node n being a CH node in an arbitrary round.

D. Energy Reservation Policy

As mentioned earlier, in order to prepare more energy for the energy-consuming CH mode, each node saves a portion of the harvested energy during the non-CH mode. Upon harvesting a unit of energy, node n immediately puts it into its energy buffer and updates the harvested energy $E_n^{\text{h}}(i)$ (i.e., $E_n^{\text{h}}(i) = E_n^{\text{h}}(i) + e_0$). Also, with probability p_n^{sv}, the reserved energy $E_n^{\text{sv}}(i)$ is updated (i.e., $E_n^{\text{sv}}(i) = E_n^{\text{sv}}(i) + e_0$). At the end of frame i, the harvested energy is reset as

$$E_n^{\text{h}}(i) = E_n^{\text{h}}(i) - E_n^{\text{sv}}(i).$$

Note that the saved energy $E_n^{\text{sv}}(i)$ will not be used until node n is selected as a CH node, at which time node n measures the total energy of the energy buffer and then updates the remaining energy $E_n^{\text{r}}(0)$ (see step 5 in Algorithm 1).

It is worth mentioning that the choice of p_n^{sv} affects the network throughput greatly. To be specific, if p_n^{sv} is very small, the node may transmit more packets in the non-CH mode, but this also results in less energy reservation (and thus higher possibility of energy shortage) in the CH mode. On the other hand, if p_n^{sv} is very large, the energy shortage events during the CH mode are reduced; however, at the same time the transmission of packets in the non-CH mode is also reduced, which degrades the network throughput. The details on how to properly determine the energy reservation ratio will be given in Section IV.

IV. Performance Analysis

In this section, we analyze the overall efficiency of the DEARER protocol in terms of the network energy efficiency. We also investigate the transmission efficiency and reliability of each sensor node in terms of the CH-outage probability and transmission ratio.

A. Definition of Metrics

Our main goal is to investigate how many packets can be successfully delivered to the sink for the given amount of energy in the network.

Definition 2: The *network energy efficiency* η is the normalized number of packets that a node can deliver successfully to the sink per frame per Joule of energy. That is,

$$\eta = \lim_{t\to\infty} \frac{1}{tN_{\text{F}} \sum_{k=1}^{N} \lambda_k e_0} \sum_{l=1}^{t} \sum_{n=1}^{N} \sum_{i=1}^{N_{\text{F}}} I_{\text{Tx}_n(i)} I_{\text{Tx}_m(i)}, \tag{8}$$

where node m is the CH node of the cluster to which node n belongs.

We are also interested in how many packets a node actually transmits.

Definition 3: The *transmission ratio* of a node is the proportion of frames in which the node has enough energy to transmit a packet:

$$r_n = \lim_{t\to\infty} \frac{1}{tN_{\text{F}}} \sum_{l=1}^{t} \sum_{i=1}^{N_{\text{F}}} I_{\text{Tx}_n(i)}. \tag{9}$$

Moreover, the *network transmission ratio* is the averaged transmission ratio over the network

$$r_{\text{net}} = \frac{1}{N} \sum_{n=1}^{N} r_n. \tag{9}$$

Note that the unit of r_n and r_{net} is packets per frame. In essence, the transmission ratio characterizes how many packets on average a node, either in the non-CH mode or the CH mode, can transmit in a frame, regardless of whether the packets are finally successfully delivered to the sink. However, if a CH node does not have enough energy to forward the packet to the sink, then an *outage* occurs. Thus we also need a metric to measure the outage event.

Definition 4: The *CH-outage probability* p_n^{out} of node n is the probability that the energy of a node in the CH mode is inadequate to deliver the aggregated packet to the sink. That is,

$$p_n^{\text{out}} = \lim_{t\to\infty} \frac{1}{N_{\text{F}} \sum_{l=1}^{t} I_{\text{CH}_n(l)}} \sum_{l=1}^{t} I_{\text{CH}_n(l)} \sum_{i=1}^{N_{\text{F}}} (1 - I_{\text{Tx}_n(i)}), \tag{10}$$

where $\text{CH}_n(l)$ is the event of node n being a CH node. Moreover, the *network outage probability* is the averaged CH-outage probability over all nodes in the network:

$$p_{\text{out}} = \frac{1}{N} \sum_{n=1}^{N} p_n^{\text{out}}. \tag{11}$$

In addition, to describe how often a node serves as a CH node, we use the following metric.

Definition 5: The *inter-CH time* T_n is the waiting time for node n to act again as a CH node.

The metric not only specifies the time for a node to accumulate energy for the upcoming CH mode, but also will play an important role in our analysis.

B. Inter-CH Time

In this subsection, we present the statistic characterization of inter-CH time. In the proposed head selection algorithm, a node satisfying $\frac{1}{Mq_n} \leq 1.5$ is chosen as a CH node in every round. In this case, the inter-CH time of the node is $T_n = 1$.

We now focus on the nodes satisfying $\frac{1}{Mq_n} > 1.5$. Let

$$\tau_k \in \left\{ k\left\lceil \frac{1}{Mq_n} - 0.5 \right\rceil + 1, \cdots, (k+1)\left\lceil \frac{1}{Mq_n} - 0.5 \right\rceil \right\}$$

be the round in which node n is a CH node where $k = \left\lfloor \frac{l}{\lceil \frac{1}{Mq_n} - 0.5 \rceil} \right\rfloor$. Then, according to Definition 5, $T_n = \tau_{k+1} - \tau_k$. In the following, we provide a useful lemma regarding τ_k and $p_n^{\overline{\text{cond}}}$ (see the definition of $p_n^{\overline{\text{cond}}}$ in Section III-C).

Lemma 4.1: Round τ_k is uniformly distributed in $\left[k\lceil \frac{1}{Mq_n} - 0.5 \rceil + 1, (k+1)\lceil \frac{1}{Mq_n} - 0.5 \rceil \right]$. For any round, the unconditional probability of node n being a CH node is

$$p_n^{\overline{\text{cond}}} = \frac{1}{\lceil \frac{1}{Mq_n} - 0.5 \rceil}. \tag{12}$$

Proof: See Appendix A. ∎

The lemma indicates that each node serves as a CH node once every $\lceil \frac{1}{Mq_n} - 0.5 \rceil$ rounds if $\frac{1}{Mq_n} \geq 1.5$. Using this lemma, we obtain the following proposition.

Proposition 4.2: Let μ_{T_n} and $\sigma^2_{T_n}$ be the mean and variance of T_n of node n, respectively. Then, if $\frac{1}{Mq_n} \leq 1.5$, node n works as a CH node in every round and $\mu_{T_n} = 1$. Otherwise,

$$\mu_{T_n} = \left\lceil \frac{1}{Mq_n} - 0.5 \right\rceil \text{ and } \sigma^2_{T_n} = \frac{\left\lceil \frac{1}{Mq_n} - 0.5 \right\rceil^2 - 1}{6}.$$

Proof: See Appendix B. ■

It is interesting to compare the proposed head selection method with the pure random head selection. For the pure random head selection method, each node elects itself as a CH node with a fixed probability $p_n^{\overline{\text{cond}}}$ (see (12)), regardless of whether the node has recently served as a CH node. It can be shown that

$$\Pr\{T_n^{\text{r}} = k\} = (1 - p_n^{\overline{\text{cond}}})^{k-1} p_n^{\overline{\text{cond}}},$$

where T_n^{r} is the inter-CH time for the pure random head selection method. Hence, the mean and variance of T_n^{r} are given by, respectively,

$$\mu_{T_n^{\text{r}}} = \left\lceil \frac{1}{Mq_n} - 0.5 \right\rceil \text{ and } \sigma^2_{T_n^{\text{r}}} = \left\lceil \frac{1}{Mq_n} - 0.5 \right\rceil \left(\left\lceil \frac{1}{Mq_n} - 0.5 \right\rceil - 1 \right).$$

It can be seen that $\mu_{T_n^{\text{r}}} = \mu_{T_n}$ and $\sigma^2_{T_n^{\text{r}}} \gg \sigma^2_{T_n}$, which means that the inter-CH time of the proposed head selection method suffers less fluctuation compared to that of the pure random head selection method. Noting that the inter-CH time is essentially the time to accumulate energy for the use in the next CH mode, the inter-CH time with smaller variance clearly ensures more stable energy supply to the CH mode, thereby reducing the probability of energy shortage.

C. Energy Consumption of Non-CH Nodes

Let $H = \sum_{n=1}^{N} I_{\text{CH}_n}$ be the total number of CH nodes in the network. If we assume that the average number of CH nodes $\mathbb{E}[H]$ is an integer, then we have

$$\mathbb{E}[H] = \mathbb{E}\left[\sum_{n=1}^{N} I_{\text{CH}_n}\right] = \sum_{n=1}^{N} p_n^{\overline{\text{cond}}} = \sum_{m=1}^{\mathbb{E}[H]} \left(\sum_{n=1}^{\kappa_m} p_n^{\overline{\text{cond}}} \right). \tag{13}$$

From (13), it is clear that for each cluster, $\sum_{n=1}^{\kappa_m} p_n^{\overline{\text{cond}}} = 1$ on average. To simplify the analysis, we assume that the nodes in the same cluster have a same energy arrival rate μ_λ. Then the priority of node n being a CH node can be expressed as $q'_n = \frac{\mu_\lambda}{\lambda_n} q_n$. Also, the probability of node n being a CH node is $p'_n = \min\{M q'_n, 1\}$. Thus, $1 = \sum_{n=1}^{\kappa_m} p'_n = \kappa_m p'_n$ (i.e., $\kappa_m = \frac{1}{p'_n}$).

Since each node is assumed to be uniformly distributed in the network, the area of cluster m consists of κ_m nodes is $\frac{4L^2\kappa_m}{N}$ on average. Under the assumption that the cluster has circular shape, the radius of cluster m is $R_m = \sqrt{\frac{4L^2\kappa_m}{\pi N}}$. Moreover, since all nodes are uniformly distributed, the density of a node over the cluster is $\rho(x,y) = \frac{1}{\pi R_m^2}$. Thus, the average squared distance $\overline{d}_{nm}^2$ between a non-CH node and CH node m is

$$\overline{d}_{nm}^2 = \iint \rho(x,y)\sqrt{x^2+y^2}\,dxdy = \int_0^{2\pi}\int_0^{R_m} \frac{r^2 r dr d\theta}{\pi R_m^2} = \frac{1}{2}R_m^2,$$

from which we obtain

$$\overline{d}_{nm} = \sqrt{\frac{2L^2\kappa_m}{\pi N}}.$$

By approximating the radius of the cluster to which node n belongs with the radius of the cluster where node n acts as the CH node, we have $\kappa_m = \kappa_n$. Hence, from (1) the average energy that a non-CH node consumes to transmit a packet to its CH node can be given by

$$\mu_{E_n^{\overline{\text{CH}}}} = L_\text{p}(e_\text{elec} + e_\text{RF}(\overline{d}_{nm})) = L_\text{p}\left(e_\text{elec} + e_\text{RF}\left(\sqrt{\frac{2L^2\kappa_n}{\pi N}}\right)\right).$$

In what follows, we use $\mu_{E_n^{\overline{\text{CH}}}}$ to approximate the energy for node n to transmit a packet to its CH node.

D. Transmission Ratio

Let $r_n^{\overline{\text{CH}}}$ and r_n^{CH} be the transmission ratio of node n in the non-CH mode and the CH mode, respectively. Then we have

$$r_n^{\overline{\text{CH}}} = \lim_{n\to\infty} \frac{\sum_{l=1}^{t}(1 - I_{\text{CH}_n})\sum_{i=1}^{N_\text{F}} I_{\text{Tx}_n(i)}}{\sum_{l=1}^{t}(1 - I_{\text{CH}_n})} \quad \text{and} \quad r_n^{\text{CH}} = \lim_{n\to\infty} \frac{\sum_{l=1}^{t} I_{\text{CH}_n}\sum_{i=1}^{N_\text{F}} I_{\text{Tx}_n(i)}}{\sum_{l=1}^{t} I_{\text{CH}_n}}.$$

Proposition 4.3: Let μ_{T_n} be the average inter-CH time of node n, κ_n be the number of nodes in the cluster associated with CH node n, and p_n^{out} be the CH-outage probability defined in (10). Then, the transmission ratio $r_n^{\overline{\text{CH}}}$ of node n in the non-CH mode and the transmission ratio r_n^{CH} in the CH mode satisfy

$$\begin{aligned} r_n^{\overline{\text{CH}}} &= \min\left\{\frac{\mu_{T_n}\lambda_n e_0 - (e_n + L_\text{p} e_\text{da} I_{\kappa_n>1})(1-p_n^{\text{out}})}{(\mu_{T_n}-1)\mu_{E_n^{\overline{\text{CH}}}} + L_\text{p}(\kappa_n - 1)(e_\text{elec}+e_\text{da})(1-p_n^{\text{out}})}, 1\right\}, \\ r_n^{\text{CH}} &= 1 - p_n^{\text{out}}. \end{aligned} \tag{14}$$

Proof: See Appendix C. ■

In essence, $r_n^{\overline{\text{CH}}}$ characterizes the best achievable transmission ratio of node n in the non-CH mode for a given CH-outage probability p_n^{out}. It offers guidelines for determining how much energy can be used for transmitting packets in the non-CH mode and how much energy should be saved for the upcoming CH mode (i.e., determining p_n^{sv}). To be specific, although $r_n^{\overline{\text{CH}}}$ and r_n^{CH} cannot be maximized simultaneously, the best trade-off can be achieved by choosing for each node a p_n^{sv} such that each CH node has exactly the required energy on average to forward all the information transmitted by its non-CH members. We can also gain useful insights by investigating the relationship between $r_n^{\overline{\text{CH}}}$, r_n^{CH}, and transmission ratio r_n. From Definition 3,

$$\begin{aligned} r_n &= \lim_{n\to\infty} \frac{1}{tN_{\text{F}}} \sum_{l=1}^{t} \sum_{i=1}^{N_{\text{F}}} I_{\text{Tx}_n(i)} = \lim_{n\to\infty} \frac{1}{tN_{\text{F}}} \left(\sum_{l=1}^{t} \left(I_{\overline{\text{CH}}_n(l)} + I_{\text{CH}_n(l)} \right) \sum_{i=1}^{N_{\text{F}}} I_{\text{Tx}_n(i)} \right) \\ &= \left(\frac{\mu_{T_n} - 1}{\mu_{T_n}} \right) r_n^{\overline{\text{CH}}} + \left(\frac{1}{\mu_{T_n}} \right) r_n^{\text{CH}}, \end{aligned} \tag{15}$$

where $\overline{\text{CH}}_n(l)$ is the event of node n being a non-CH node and the last equality follows from $p_n^{\overline{\text{cond}}} = \lim_{n\to\infty} \frac{1}{t} \sum_{l=1}^{t} I_{\text{CH}_n}$ and $\mu_{T_n} = \frac{1}{p_n^{\text{cond}}}$. Clearly if $\mu_{T_n} = 1$, node n acts as a CH node in every round. In this case, the transmission ratio becomes $r_n = r_n^{\text{CH}}$ and $r_n^{\overline{\text{CH}}}$ can be ignored.

E. Energy Reservation Ratio

For node n, the energy reservation ratio p_n^{sv} can be obtained by investigating its energy consumption in the non-CH mode.

Proposition 4.4: The energy reservation ratio of node n is

$$p_n^{\text{sv}} = 1 - \frac{\mu_{E_n^{\overline{\text{CH}}}} r_n^{\overline{\text{CH}}}}{\lambda_n e_0}. \tag{16}$$

Proof: Under our energy reservation policy, only a portion of the harvested energy is used for packet transmission in the non-CH mode while the rest is saved for the use in the next CH mode. Thus, the energy arrival process, which is a Poisson process with intensity λ_n, can be split into two independent subprocesses, more precisely, Poisson processes with intensities $(1-p_n^{\text{sv}})\lambda_n$ and $p_n^{\text{sv}}\lambda_n$. This allows us to formulate the energy consumption in the non-CH mode. Specifically, in the period of $T_n - 1$ rounds when node n works in the non-CH mode, the total consumed energy is the difference between the total harvested energy and the saved one, that is,

$$\sum_{l=1}^{T_n-1} \sum_{i=1}^{N_{\text{F}}} I_{\text{Tx}_n(i)} E_n^{\overline{\text{CH}}}(i) = \sum_{l=1}^{T_n-1} \sum_{i=1}^{N_{\text{F}}} E_n^{\text{h}}(i)(1 - I_{\text{sv}_n(i)}). \tag{17}$$

By taking expectation on both sides of (17) and eliminating common terms, we have

$$\mu_{E_n^{\overline{\mathrm{CH}}}} r_n^{\overline{\mathrm{CH}}} = (1 - p_n^{\mathrm{sv}})\lambda_n e_0, \tag{18}$$

from which we can directly obtain p_n^{sv}. ■

Since the consumed energy $\mu_{E_n^{\overline{\mathrm{CH}}}} r_n^{\overline{\mathrm{CH}}}$ in the non-CH mode is smaller than the totally harvested energy $\lambda_n e_0$, p_n^{sv} is always positive. Moreover, the larger λ_n is, the more energy a node can save, and the smaller the corresponding CH-outage will be. In fact, we can uniquely determine the CH-outage with $r_m^{\overline{\mathrm{CH}}}$ and p_n^{sv}.

F. The CH-Outage Probability

As mentioned, an outage event occurs at a CH node if it does not have enough energy to forward the aggregated packet to the sink. The distribution of the available energy at a CH node is difficult to analyze. Nevertheless, since the available energy at a CH node is the sum of Poisson random variables, which belong to the exponential family, the distribution can be well approximated by Gamma distribution, which facilitates proving the following proposition.

Proposition 4.5: The outage probability p_n^{out} of node n in the CH mode is

$$p_n^{\mathrm{out}} = \gamma\left(k_n, \frac{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}}{\theta_n}\right) - \frac{k_n\theta_n}{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}}\gamma\left(k_n+1, \frac{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}}{\theta_n}\right). \tag{19}$$

where $\gamma(k,x) = \frac{1}{\Gamma(k)}\int_0^x t^{k-1}e^{-t}\mathrm{d}t$ is the lower incomplete Gamma function and $\mu_{E_n^{\mathrm{CH}}}$ is the average energy that node n consumes per frame in the CH mode,

$$\mu_{E_n^{\mathrm{CH}}} = L_{\mathrm{p}}\left((\kappa_n - 1)(e_{\mathrm{elec}} + e_{\mathrm{da}})r_n^{\overline{\mathrm{CH}}} + e_{\mathrm{da}} I_{\kappa_n > 1}\right) + e_n, \tag{20}$$

and $k_n,\ \theta_n$ are given, respectively, by

$$k_n = \frac{N_{\mathrm{F}}\lambda_n(p_n^{\mathrm{sv}}(\mu_{T_n}-1)+1)^2}{N_{\mathrm{F}}\lambda_n(p_n^{\mathrm{sv}})^2\sigma_{T_n}^2 + p_n^{\mathrm{sv}}(\mu_{T_n}-1)+1} \quad \text{and} \quad \theta_n = e_0 + \frac{N_{\mathrm{F}}\lambda_n e_0 (p_n^{\mathrm{sv}})^2\sigma_{T_n}^2}{p_n^{\mathrm{sv}}(\mu_{T_n}-1)+1}.$$

Proof: See Appendix D. ■

For a node n acting as a CH node in a certain round, we can approximate the distribution of its total available energy E_n^{total} (see (D.24)) using Gamma distribution $\Gamma(k_n, \theta_n)$ (see (D.25)). By doing so, the average of E_n^{total} can be given as $\mu_{E_n^{\mathrm{total}}} = k_n\theta_n$. Suppose $Z \sim \Gamma(k_n, 1)$ is a Gamma-distributed random variable, we have $\mathbb{E}[Z] = k_n$ and

$$\gamma\left(k_n, \frac{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}}{\theta_n}\right) = 1 - \Pr\left\{Z > \frac{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}}{\theta_n}\right\} \overset{(a)}{\geq} 1 - \frac{\mathbb{E}[Z]}{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}/\theta_n} = 1 - \frac{\mu_{E_n^{\mathrm{total}}}}{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}},$$

where (a) uses Markov's inequality $\Pr\{Z > t\} \leq \mathbb{E}[Z]/t$ [23].

If the network is energy insufficient, i.e., if the subsistence ratios ρ is small, $\mu_{E_n^{\text{total}}}$ is much smaller than the average energy $N_{\text{F}}\,\mu_{E_n^{\text{CH}}}$ required for node n to transmit a packet in every frame of a round when it acts as a CH node, i.e., $\mu_{E_n^{\text{total}}} \ll N_{\text{F}}\,\mu_{E_n^{\text{CH}}}$. Thus, we have $\gamma(k_n, N_{\text{F}}\,\mu_{E_n^{\text{CH}}}/\theta_n) \approx 1$ and $\gamma(k_n+1, N_{\text{F}}\,\mu_{E_n^{\text{CH}}}/\theta_n) \approx 1$. In this case, the CH-outage probability of node n becomes

$$p_n^{\text{out}} = 1 - \frac{k_n\theta_n}{N_{\text{F}}\,\mu_{E_n^{\text{CH}}}}, \tag{21}$$

which is relatively simple to interpret.

From Proposition 4.3, 4.4 and 4.5, we can determine $r_n^{\overline{\text{CH}}}$, r_n^{CH}, p_n^{sv}, and p_n^{out}. Using these results, we can further compute the network energy efficiency η of the network, which incorporates both the transmission efficiency of non-CH nodes and the forwarding reliability of CH nodes.

G. Network Energy Efficiency

Recall that the network energy efficiency η is the ratio between the number of successfully delivered packets in a frame and the total energy harvested by the network in the frame (see Definition 2). By ergodicity of the energy-arrival process and the law of large numbers, we have

$$\begin{aligned}
\eta &= \lim_{t\to\infty} \frac{1}{tN_{\text{F}}\sum_{k=1}^{N}\lambda_k e_0} \sum_{l=1}^{t}\sum_{n=1}^{N}\sum_{i=1}^{N_{\text{F}}} I_{\text{Tx}_n(i)} I_{\text{Tx}_m(i)} \\
&= \frac{1}{N\mu_\lambda e_0}\sum_{n=1}^{N}\left(\frac{\mu_{T_n}-1}{\mu_{T_n}} r_n^{\overline{\text{CH}}}(1-p_m^{\text{out}}) + \frac{1}{\mu_{T_n}} r_n^{\text{CH}}\right) \\
&= \frac{1}{N\mu_\lambda e_0}\sum_{n=1}^{N}\frac{r_n^{\text{CH}}}{\mu_{T_n}}(1+(\mu_{T_n}-1)r_n^{\overline{\text{CH}}}),
\end{aligned}$$

where in the last equation we have used the approximation $p_m^{\text{out}} = 1 - r_n^{\text{CH}}$.

V. Numerical Results

A. Simulation Setup

We consider $N = 1,000$ nodes uniformly distributed over an area of 140×140 m^2, (i.e., $L = 70$). The sink locates at point $(0, 80)$. The slot length is $t_s = 0.0016$ s, in which a packet of $L_{\text{p}} = 400$ bits is transmitted. The steady phase of each round contains $N_F = 100$ frames and each frame consists of N slots. Following [11], parameters in the energy model are set to $e_{\text{elec}} = 50$ nJ/bit, $e_{\text{da}} = 5$ nJ/bit/signal, $\epsilon_{\text{fs}} = 10$ pJ/bit/m^2, $\epsilon_{\text{mp}} = 0.0013$ pJ/bit/m^4, and $d_0 = 87$

m. As mentioned, we assume that the energy arrival of each node follows the Poisson process. Without loss of generality, we assume that the energy arrival rate λ_n of each node is different from each other and follows Rayleigh distribution with parameter $\nu_\lambda = 1$.

B. Direct Transmission and Genie-aided Routing

In direct transmission, each node directly transmits a packet to the sink in the allocated slot if it has enough energy, in other words, the remaining energy $E_n^{\mathrm{r}}(i)$ is larger than $e_n = L_{\mathrm{p}}(e_{\mathrm{elec}} + e_{\mathrm{RF}}(d_n))$. At the end of each frame, node n updates the remaining energy by $E_n^{\mathrm{r}}(i+1) = E_n^{\mathrm{r}}(i) + E_n^{\mathrm{h}}(i) - e_n I_{E_n^{\mathrm{r}}(i) \geq e_n}$. The transmission ratio of node n is evaluated by $r_n^{\mathrm{dt}} = \lim_{t\to\infty} \frac{1}{tN_{\mathrm{F}}} \sum_{l=1}^{t} \sum_{i=1}^{N_{\mathrm{F}}} I_{\mathrm{Tx}_n(i)}$, and the network transmission ratio is $r_{\mathrm{net}}^{\mathrm{dt}} = \frac{1}{N} \sum_{n=1}^{N} r_n^{\mathrm{dt}}$. Note that there is no outage event in direct transmission since a node will not transmit any packet unless it has enough energy. However, for comparative purpose, we define $p_{\mathrm{out}}^{\mathrm{dt}} = \frac{1}{N} \sum_{n=1}^{N} (1 - r_n^{\mathrm{dt}})$ as the network outage probability of direct transmission.

Similar to DEARER, genie-aided routing is also a cluster-based routing protocol. In the genie-aided routing, instead of re-assigning CH nodes and establishing clusters once a round, CH selection and cluster formation is performed in each frame. With the aid of a genie that knows the real-time remaining energy of each node as well as the distance from each node to the sink, the CH nodes are selected among the most energy-saving and most energy-sufficient ones. More precisely, the genie updates the priority for each node being a CH node as

$$q_n^{\mathrm{ge}}(i) = \frac{\frac{E_n^{\mathrm{r}}(i)}{e_n}}{\sum_{k=1}^{N} \frac{E_k^{\mathrm{r}}(i)}{e_k}}.$$

Based on the computed priority, the CH nodes are chosen accordingly. Particularly, the number of CH nodes in the genie-aided routing is chosen to be the number that maximizes the network energy efficiency (i.e., the number that leads to the most packet delivery to the sink). After the clusters being formed, non-CH nodes start to transmit packets to their respective CH-node. At the same time, a CH node will transmit the aggregated packet to the sink if it has the enough energy to do so.

C. The Calculation of Metrics

It is worth mentioning that parameters p_n^{out}, $r_n^{\overline{\mathrm{CH}}}$, r_n^{CH}, and p_n^{sv} are related (with each other) by equations (14), (16), and (19), so that we cannot evaluate them individually. In Table 3, we

Algorithm 3 Calculation of p_n^{out}, $r_n^{\overline{\text{CH}}}$, r_n^{CH}, and p_n^{sv}

Initialization:

1: Set maximum iterations K, relative error ε;
2: Set $p_n^{\text{out}}[1] = 0^1$;

Iteration:

3: **repeat**
4: Solve $r_n^{\overline{\text{CH}}}[k]$ and $r_n^{\text{CH}}[k]$ from (14), using $p_n^{\text{out}}(k)$;
5: Solve $p_n^{\text{sv}}[k]$ from (16);
6: Update $\mu_{E_n^{\text{CH}}}[k]$ according to (20);
7: Update $p_n^{\text{out}}[k+1]$ according to (19);
8: **until** $k = K$ or $|p_n^{\text{out}}[k] - p_n^{\text{out}}[k+1]| \leq \varepsilon p_n^{\text{out}}[k+1]$
9: **Output:** $r_n^{\overline{\text{CH}}}[K]$, $r_n^{\text{CH}}[K]$, $p_n^{\text{sv}}(K)$, $p_n^{\text{out}}[K+1]$.

present an efficient algorithm to calculate these parameters. The algorithm outputs the solution in a few iterations. In particular, it terminates either when the maximum iterations K is reached or when the difference between the output p_n^{out} of two adjacent iterations falls below a pre-specified relative error ε. In our simulations, we set $K = 200$ and $\varepsilon = 0.005$.

D. Numerical and Simulation Results

In Figure 3(a), we plot network energy efficient η for three protocols under test as a function of the subsistence ratio ρ. Recall that the desired number M of CH nodes is a controlling parameter in DEARER. We consider $M = 10$, 100, and 500 in the simulation. However, since the direct transmission method is independent of M, its curves of network energy efficiency for different M coincide with each other. Likewise, since the number of CH nodes in the genie-aided routing is not controlled by M (in fact, it is chosen to maximize η), the curves of the genie-aided routing for different M also coincide with each other.

In Figure 3(a), we observe that for all protocols, the network is operated efficiently if $\rho \in (0.3, 0.4)$. In particular, DEARER largely outperforms direct transmission in this region. The gain of DEARER is mainly because 1) the cluster-based routing cuts down long-distance transmissions, and 2) the energy reservation policy reschedules the available energy to match with the energy profile on demand. We also observe that, when ρ is larger than 0.45, the network energy efficiency of all transmission schemes decreases monotonically with ρ. This is because when $\rho > 0.45$, the transmission ratio r_{net} already approaches one packet/node/frame (see Figure 3(c))

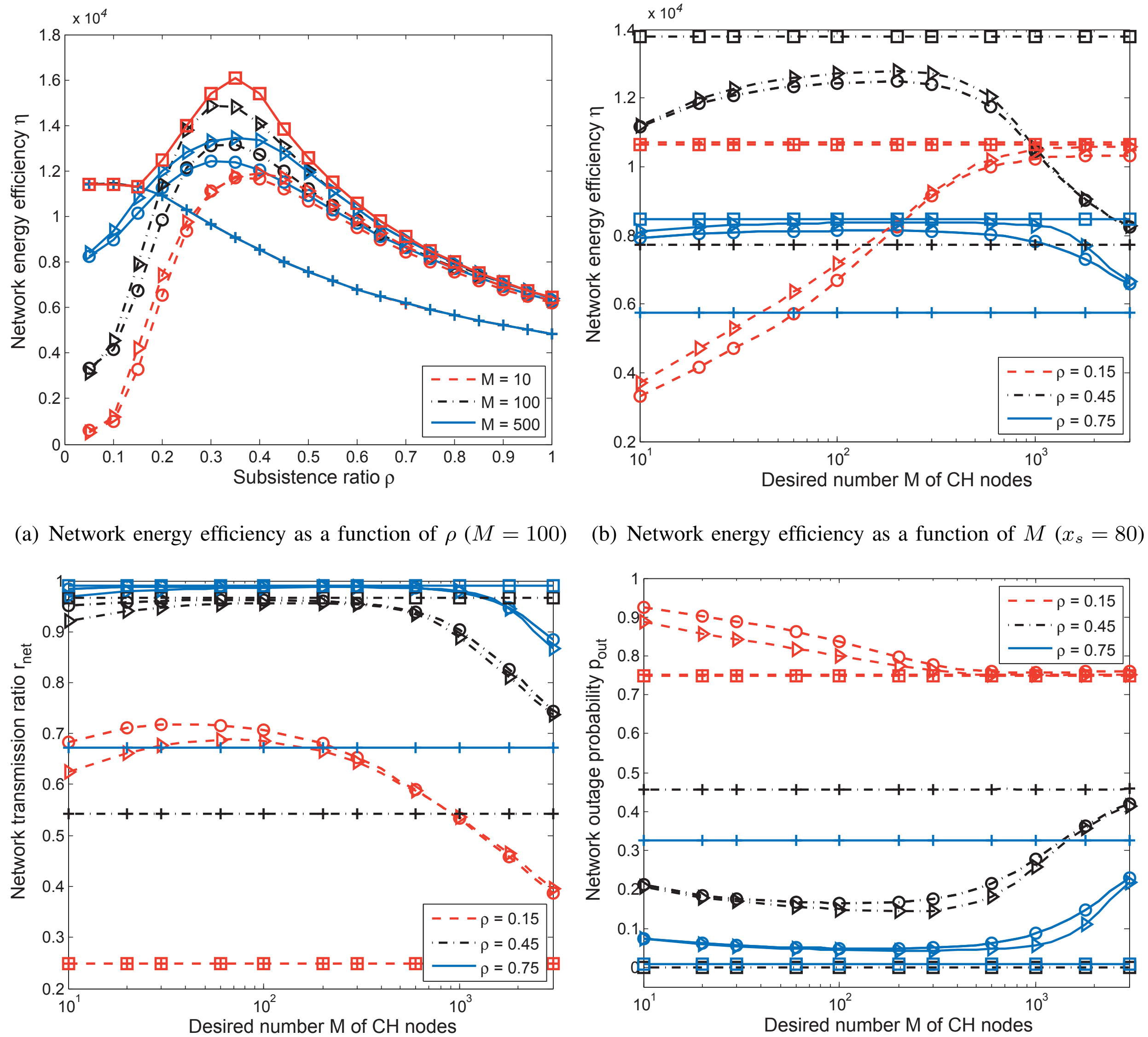
[1]In Algorithm 3, $x[k]$ is the value of x in the k-th iteration.

so that further gain of r_{net} is marginal irrespective of ρ. Since the network energy efficiency is normalized by the totally harvested energy, it decreases with ρ. In addition, we observe that direct transmission is better than DEARER and even performs close to the genie-aided routing when ρ is very small (i.e., when the network is seriously energy insufficient). This is mainly due to two reasons. First, if the available energy at each node very little, then for cluster-based routing protocols, the overhead of packet reception and aggregation is non-negligible. Second, the advantage of packet aggregation at CH nodes becomes limited because very few packets could actually be transmitted by non-CH nodes in this case.

Figure 3(b) depicts how the controlling parameter M affects the network energy efficiency. It can be seen that when the network is not very energy insufficient (e.g., $\rho = 0.45$ and $\rho = 0.75$), DEARER with properly chosen M achieves much higher energy efficiency than direct transmission, demonstrating advantages of the cluster-based routing and energy reservation policy. Note that the number of CH nodes in the network increases with M. If M is small, the number of clusters is small. As a result, the size of each cluster becomes large so that the traffic at each CH node will be heavy. On the other hand, if M is large, there will be a large number of CH nodes transmitting packets directly to the sink. In this case, DEARER essentially degrades to the direct transmission method. Thus, to achieve high network energy efficiency, a moderate M is preferred. However, when the network is seriously energy insufficient, (e.g., $\rho = 0.15$), DEARER has lower network energy efficiency when compared with the direct transmission method, as shown in Figure 3(a). Nevertheless, the improvement in energy efficiency can be achieved by choosing very large M. In doing so, the network energy efficiency of DEARER approaches that of direct transmission.

We present more details on the network transmission ratio ($r_{\text{net}} = \frac{1}{N}\sum_{n=1}^{N} r_n$) and network outage probability ($p_{\text{out}} = \frac{1}{N}\sum_{n=1}^{N} p_n^{\text{out}}$) in Figure 3(c) and 3(d), respectively. As is clear from the figure, DEARER has much higher transmission ratio and also much smaller CH-outage probability than direct transmission when $\rho = 0.45$ and $\rho = 0.75$. This means that under the DEARER protocol, the sensor nodes transmit more packets, and the packet forwarding at CH nodes is also more reliable. When $\rho = 0.15$, while the transmission ratio of DEARER is higher than that of direct transmission (see Figure 3(c)), its CH-outage probability is larger (see Figure 3(d)), which results in low network energy efficiency (see Figure 3(a)).

Figure 4 displays the energy reservation ratio p_n^{sv} of each node. The p_n^{sv} of nodes are sorted in



(a) Network energy efficiency as a function of ρ ($M = 100$) (b) Network energy efficiency as a function of M ($x_s = 80$)

(c) Network transmission ratio as a function of M ($x_s = 80$) (d) Network outage probability as a function of M ($x_s = 80$)

Figure 3. Performance comparison between direct transmission, genie-aided routing, and the proposed DEARER protocol, where the curves marked by o, $\triangleright$, $+$, and $\square$ correspond to DEARER in theory, DEARER by Monte Carlo, direct transmission, and genie-aided routing, respectively.

an ascending order. When the subsistence ratio ρ is very small, the available energy is limited for most of nodes so that there is little energy to save. Thus, the energy reservation ratio of most nodes is zero. It is seen that when the energy subsistence ratio is increased from $\rho = 0.15$ to $\rho = 0.45$ (or from $\rho = 0.45$ to $\rho = 0.75$), p_n^{sv} for most of nodes becomes larger. Moreover, if M is increased from $M = 10$ to $M = 100$, p_n^{sv} also becomes larger. This is due to the fact that when M increases, each node serves as a CH node more frequently and hence needs more

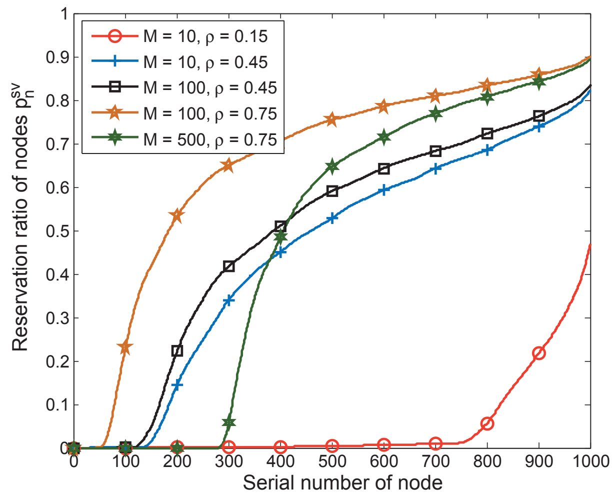


Figure 4. The energy reservation ratio of DEARER where $x_s = 80$ m.

energy for the CH mode. When $\rho = 0.75$ and we increase $M = 100$ to $M = 500$, however, the p_n^{sv} becomes smaller and even zero. This is because in these cases most of nodes act as CH nodes every round and hence do not save energy.

VI. Conclusion

In this paper, we have proposed a routing protocol called DEARER for EH-WSNs. By encouraging enabler nodes to serve more frequently as CH nodes and to save a portion of the harvested energy for the CH mode, DEARER outperforms direct transmission and approaches the genie-aided routing for most of the realistic scenarios. We have established a mathematically tractable model and presented a detailed analysis of DEARER by employing statistical approximations. These approximations are reasonable as the underlying quantities converge to their respective averages when the operation time goes to infinity.

In DEARER, the desired number M of CH nodes can be tuned to improve the network energy efficiency. From numerical and simulation results, we have shown that when the network is not very energy insufficient (i.e., the substance ratio ρ is not very small), with properly chosen M, DEARER largely outperform direct transmission. Whereas, when the network is seriously energy insufficient, the direct transmission method performs better. In fact, direct transmission performs as well as the genie-aided routing in this case.

We have also investigated how much energy a node should save in its non-CH mode in terms of the energy reservation ratio. To make full use of the harvested energy, we choose the energy reservation ratio such that in an average sense, each non-CH node reaches its best achievable transmission ratio while each CH node has the exact amount of energy to transmit the aggregated packet to the sink. We would like to point out that in the DEARER protocol, the energy reservation ratio is constant over time. In many application scenarios, however, real-time energy status is accessible (e.g., through message exchanges). By utilizing the real-time energy status, a dynamic energy saving (or an online energy scheduling) may lead to further improvement of the network energy efficiency. We would also like to point out that our model and analysis of the DEARER protocol are built upon approximations in the average sense and do not incorporate the game-playing among nodes. Note that selfish nodes may achieve their optimal individual transmission ratio if they evade serving as CH nodes so that the transmission scheme would degrade to the direct transmission. We believe that a game-theoretic formulation of the cluster-based routing may bring new insights into the design of routing algorithms in EH-WSNs. Our future research will be directed towards these avenues.

APPENDIX A
PROOF OF LEMMA 4.1

Proof: For simplicity, we consider the head selections from round 1 to round $\lceil \frac{1}{Mq_n} - 0.5 \rceil$. Let A_l be the event that node n is a CH node in round l and $\bar{A}_l$ be its complementary event. It is seen that $p_n^{\text{cond}}(l) = \frac{1}{\lceil \frac{1}{Mq_n} - 0.5 \rceil - l}$ is actually the probability of the event $\{A_l | \bar{A}_1, \bar{A}_2, \cdots, \bar{A}_{l-1}\}$. Therefore, the probability distribution of τ_1 can be given by

$$\begin{aligned}
\Pr\{\tau_1 = l\} &= \Pr\{A_l, \bar{A}_1, \cdots, \bar{A}_{l-1}\} = \Pr\{\bar{A}_1\} \Pr\{\bar{A}_2 | \bar{A}_1\} \cdots \Pr\{A_l | \bar{A}_1, \cdots, \bar{A}_{l-1}\} \\
&= \left(1 - \frac{1}{\lceil \frac{1}{Mq_n} - 0.5 \rceil - 1}\right)\left(1 - \frac{1}{\lceil \frac{1}{Mq_n} - 0.5 \rceil - 2}\right) \cdots \\
&\quad \left(1 - \frac{1}{\lceil \frac{1}{Mq_n} - 0.5 \rceil - (l-1)}\right) \frac{1}{\lceil \frac{1}{Mq_n} - 0.5 \rceil - l} \\
&= \frac{1}{\lceil \frac{1}{Mq_n} - 0.5 \rceil},
\end{aligned}$$

which implies that τ_1 follows the discrete uniform distribution. Likewise, one can also prove that τ_k is a discrete uniformly distributed random variable.

Noting that $\Pr\{A_l\} = \Pr\{A_l, \bar{A}_1, \cdots, \bar{A}_{l-1}\} + \Pr\{A_l, \overline{\bar{A}_1, \cdots, \bar{A}_{l-1}}\}$ during the period of round 1 to round $\lceil \frac{1}{Mq_n} - 0.5 \rceil$, we have

$$p_n^{\overline{\text{cond}}} = \Pr\{A_l\} = \Pr\{A_l, \bar{A}_1, \cdots, \bar{A}_{l-1}\} + \Pr\{A_l, \overline{\bar{A}_1, \cdots, \bar{A}_{l-1}}\} = \frac{1}{\lceil \frac{1}{Mq_n} - 0.5 \rceil}, \quad \text{(A.22)}$$

where $\Pr\{A_l, \overline{\bar{A}_1, \bar{A}_2, \cdots, \bar{A}_{l-1}}\}$ is zero because node n cannot serve as a CH node more than once in this period. In addition, the head selection for each node n during the period of round $k\lceil \frac{1}{Mq_n} - 0.5 \rceil + 1$ to round $(k+1)\lceil \frac{1}{Mq_n} - 0.5 \rceil$ is independent of the head selections before round $k\lceil \frac{1}{Mq_n} - 0.5 \rceil + 1$. Thus, we can conclude that (A.22) holds for any $k \geq 0$. ■

APPENDIX B
PROOF OF PROPOSITION 4.2

Proof: For simplicity, we consider the head selections during the period of round 1 to round $\lceil \frac{1}{Mq_n} - 0.5 \rceil$ and the inter-CH time $T_n = \tau_2 - \tau_1$. From Lemma 4.1, τ_1 and τ_2 are variables uniformly distributed in $\left[1, \lceil \frac{1}{Mq_n} - 0.5 \rceil\right]$ and $\left[\lceil \frac{1}{Mq_n} - 0.5 \rceil + 1, 2\lceil \frac{1}{Mq_n} - 0.5 \rceil\right]$, respectively. Denoting $\tau_1' = \lceil \frac{1}{Mq_n} - 0.5 \rceil - \tau_1$ and $\tau_2' = \tau_2 - \lceil \frac{1}{Mq_n} - 0.5 \rceil$, we have $T_n = \tau_1' + \tau_2'$, and

$$\begin{aligned}
\Pr\{\tau_1' \leq j\} &= \Pr\left\{\lceil \frac{1}{Mq_n} - 0.5 \rceil - \tau_1 \leq j\right\} = \Pr\left\{\tau_1 \geq \lceil \frac{1}{Mq_n} - 0.5 \rceil - j\right\} \\
&= 1 - \frac{\lceil \frac{1}{Mq_n} - 0.5 \rceil - j - 1}{\lceil \frac{1}{Mq_n} - 0.5 \rceil} \\
&= \frac{j+1}{\lceil \frac{1}{Mq_n} - 0.5 \rceil},
\end{aligned}$$

which means that τ_1' is discrete and uniformly distributed in $\left[0, \lceil \frac{1}{Mq_n} - 0.5 \rceil - 1\right]$. Since τ_1 is independent of τ_2, τ_1' is also independent of τ_2'. Note that the mean and variance of a discrete uniformly distributed random variable $X \sim \mathcal{U}(a, b)$ is $\mathbb{E}[X] = \frac{a+b}{2}$ and $\mathbb{D}[X] = \frac{(b-a+1)^2-1}{12}$. Thus,

$$\mathbb{E}[T_n] = \mathbb{E}[\tau_1'] + \mathbb{E}[\tau_2'] = \left\lceil \frac{1}{Mq_n} - 0.5 \right\rceil \quad \text{and} \quad \mathbb{D}[T_n] = \mathbb{D}[\tau_1'] + \mathbb{D}[\tau_2'] = \frac{\lceil \frac{1}{Mq_n} - 0.5 \rceil^2 - 1}{6},$$

which completes the proof. ■

APPENDIX C
PROOF OF PROPOSITION 4.3

Proof: For an arbitrarily chosen period of T_n rounds, node n serves as a non-CH node for $T_n - 1$ rounds and as a CH node for one round. Using (10) together with ergodicity of energy harvesting process and the law of large numbers, we have $r_n^{\text{CH}} = 1 - p_n^{\text{out}}$.

In an energy reservation policy that makes the best use of the harvested energy (without any waste or shortage of energy), the totally consumed energy of nodes in both the non-CH and CH modes equals to the totally harvested energy. That is,

$$\sum_{l=1}^{T_n-1}\sum_{i=1}^{N_{\mathrm{F}}} I_{\mathrm{Tx}_n(i)} E_n^{\overline{\mathrm{CH}}}(i) + \sum_{i=1}^{N_{\mathrm{F}}} I_{\mathrm{Tx}_n(i)} E_n^{\mathrm{CH}}(i) = \sum_{l=1}^{T_n}\sum_{i=1}^{N_{\mathrm{F}}} E_n^{\mathrm{h}}(i). \tag{C.23}$$

By taking expectation on both sides of (C.23), we have

$$(\mu_{T_n} - 1) r_n^{\overline{\mathrm{CH}}} E_n^{\overline{\mathrm{CH}}}(i) + E_n^{\mathrm{CH}}(i) r_n^{\mathrm{CH}} = \mu_{T_n} E_n^{\mathrm{h}}(i).$$

Since the transmission ratio of a node depends largely on its distance to the sink, the nodes within the same cluster have the same transmission ratio on average. Thus, using the approximation $E_n^{\overline{\mathrm{CH}}}(i) \approx \mu_{E_n^{\overline{\mathrm{CH}}}}$, we have

$$\Big(L_{\mathrm{p}}(\kappa_n - 1)(e_{\mathrm{elec}} + e_{\mathrm{da}}) r_n^{\overline{\mathrm{CH}}} + L_{\mathrm{p}} e_{\mathrm{da}} I_{\kappa_n > 1} + e_n\Big) r_n^{\mathrm{CH}} + (\mu_{T_n} - 1)\mu_{E_n^{\overline{\mathrm{CH}}}}) r_n^{\overline{\mathrm{CH}}} = \mu_{T_n} \lambda_n e_0,$$

where κ_n is the number of nodes in the cluster associated with the CH node n. By rearranging terms, we obtain the first term in the parenthesis of (14).

Since a node can transmit at most one packet in a frame under the DEARER protocol, the transmission ratio is upper bounded by one, irrespective of the energy arrival rate, which establishes the proposition. ■

APPENDIX D
PROOF OF PROPOSITION 4.5

Proof: By ergodicity of the energy arrival process, head selection process, and energy saving process, we have

$$p_n^{\mathrm{out}} = \sum_{j=0}^{N_{\mathrm{F}}-1} \left(\frac{N_{\mathrm{F}} - j}{N_{\mathrm{F}}} \Pr\left\{ j\,\mu_{E_n^{\mathrm{CH}}} < E_n^{\mathrm{r}}(0) + \sum_{i=1}^{N_{\mathrm{F}}} E_n^{\mathrm{h}}(i) < (j+1)\,\mu_{E_n^{\mathrm{CH}}} \right\} \right),$$

where $\mu_{E_n^{\mathrm{CH}}} = \mathbb{E}[E_n^{\mathrm{CH}}]$ is the average energy consumption of CH node n in a frame.

To prove Proposition 4.5, we first need to find the probability distribution of the energy available for a CH node. The Gamma distribution and the moment matching method are presented in the following.

1) Gamma distribution and moment matching: We denote the Gamma distributed random variable with shape $k > 0$ and $\theta > 0$ by $X \sim \Gamma(k, \theta)$. The probability density function (pdf) using the shape-scale parametrization is [20]

$$f_X(x) = x^{k-1}\frac{c^{-x/\theta}}{\theta^k \Gamma(k)},$$

where $\Gamma(k)$ is the gamma function evaluated at k. The first and second moments and variance are

$$\mathbb{E}[X] = k\theta, \quad \mathbb{E}[X^2] = k(k+1)\theta^2, \text{ and } \mathbb{D}[X] = k\theta^2.$$

It is well known that a distribution with $\mu = \mathbb{E}[Y]$, $\mu^{(2)} = \mathbb{E}[Y^2]$ and variance $\sigma^2 = \mu^{(2)} - \mu^2$ can be approximated with a Gamma distribution $\Gamma(k, \theta)$ by matching the first and second order moments where k and θ are given, respectively, by

$$k = \frac{\mu^2}{\sigma^2} \text{ and } \theta = \frac{\sigma^2}{\mu}.$$

This method is known as *moment matching* [20].

2) Probability distribution of the energy available at CH nodes: When node n is selected as a CH node, the total amount of available energy for the operation in the CH mode is given by

$$E_n^{\text{total}} = E_n^{\text{r}}(0) + E_n^{\text{h}}(1:N_{\text{F}}), \tag{D.24}$$

where $E_n^{\text{r}}(0) = \sum_{l=1}^{T_n-1}\sum_{i=1}^{N_{\text{F}}} E_n^{\text{sv}}(i)$ is the saved energy when node n is in the non-CH mode and $E_n^{\text{h}}(1:N_{\text{F}}) = \sum_{i=1}^{N_{\text{F}}} E_n^{\text{h}}(i)$ is the harvested energy in the CH mode. In particular, $\{E_n^{\text{sv}}(i), i = 1, 2, \cdots N_{\text{F}}\}$ for each round is a Poisson process with parameter $p_n^{\text{sv}}\lambda_n$.

Note that if X is a Poisson distributed random variable with parameter λ, then for any constant $c \neq 0$, the moment generating function (MGF) of cX is $G_{cX}(s) = \mathbb{E}\left[e^{scX}\right] = \exp\left(\lambda(e^{cs} - 1)\right)$. Thus, the MGF of $E_n^{\text{sv}}(i)$ is $G_{E_n^{\text{sv}}}(s) = \exp\left(p_n^{\text{sv}}\lambda(e^{e_0 s} - 1)\right)$ and the MGF of $E_n^{\text{r}}(0)$ is

$$\begin{aligned}
G_{E_n^{\text{r}}(0)}(s) &= \mathbb{E}\left[\exp\left(s\sum_{l=1}^{T_n-1}\sum_{i=1}^{N_{\text{F}}} E_n^{\text{sv}}(i)\right)\right] = \mathbb{E}_{T_n}\left[\prod_{l=1}^{T_n-1}\prod_{i=1}^{N_{\text{F}}}\mathbb{E}_{E_n^{\text{sv}}}\left[e^{sE_n^{\text{sv}}(i)}\right]\right] \\
&= \mathbb{E}_{T_n}\left[\exp\left((T_n-1)N_{\text{F}}p_n^{\text{sv}}\lambda_n(e^{e_0 s} - 1)\right)\right] \\
&= \exp(\varphi(s))G_{T_n}(\varphi(s)),
\end{aligned}$$

where $\varphi(s) = N_{\text{F}}p_n^{\text{sv}}\lambda_n(e^{e_0 s} - 1)$ and $G_{T_n}(s)$ is the MGF of inter-CH time T_n. Since T_n is the sum of two independent uniformly distributed random variables ($T_n = \tau_1' + \tau_2'$), we have

$$G_{T_n}^s(s) = e^t\left(\frac{e^{t(\lceil\frac{1}{Mq_n} - 0.5\rceil - 1)} - 1}{t(\lceil\frac{1}{Mq_n} - 0.5\rceil - 1)}\right)^2.$$

Given the MGF of a random variable X, the moments can be obtained as $\mathbb{E}[x^{(k)}] = G_X^{(k)}(0)(s)|_{s=0}$. Thus, the first and second moments of $E_n^{\mathrm{r}}(0)$ can be given by

$$\mathbb{E}[E_n^{\mathrm{r}}(0)] = c_n(\mu_{T_n} - 1) \quad \text{and} \quad \mathbb{E}[(E_n^{\mathrm{r}}(0))^2] = c_n^2(\mu_{T_n}^{(2)} - 2\mu_{T_n} + 1) + c_n e_0(\mu_{T_n} - 1),$$

respectively, where $c_n = N_{\mathrm{F}} p_n^{\mathrm{sv}} \lambda_n e_0$, $\mu_{T_n} = (G_{T_n}^r(s))'|_{s=0} = \mathbb{E}[T_n]$, and $\mu_{T_n}^{(2)} = (G_{T_n}^r(s))''|_{s=0} = \mathbb{E}[T_n^2]$. This implies that the variance of $E_n^{\mathrm{r}}(0)$ is

$$\mathbb{D}[E_n^{\mathrm{r}}(0)] = c_n^2 \mathbb{D}[T_n] + c_n e_0(\mu_{T_n} - 1).$$

Likewise, the MGF of $E_n^{\mathrm{h}}(1:N_{\mathrm{F}})$ is

$$G_{E_n^{\mathrm{h}}(1:N_{\mathrm{F}})}(s) = \mathbb{E}\left[e^{sE_n^{\mathrm{h}}(1:N_{\mathrm{F}})}\right] = \exp\left(N_{\mathrm{F}}\lambda_n(e^{e_0 s} - 1)\right),$$

and its mean and variance are

$$\mathbb{E}[E_n^{\mathrm{h}}(1:N_{\mathrm{F}})] = N_{\mathrm{F}}\lambda_n e_0 \quad \text{and} \quad \mathbb{D}[E_n^{\mathrm{h}}(1:N_{\mathrm{F}})] = N_{\mathrm{F}}\lambda_n e_0^2.$$

Since $E_n^{\mathrm{r}}(0)$ and $E_n^{\mathrm{h}}(1{:}N_{\mathrm{F}})$ are independent of each other, the mean and variance of E_n^{total} are

$$\mu_{E_n^{\mathrm{total}}} = \mathbb{E}[E_n^{\mathrm{r}}(0)] + \mathbb{E}[E_n^{\mathrm{h}}(1:N_{\mathrm{F}})] \quad \text{and} \quad \sigma^2_{E_n^{\mathrm{total}}}\mathbb{D}[E_n^{\mathrm{r}}(0)] + \mathbb{D}[E_n^{\mathrm{h}}(1:N_{\mathrm{F}})].$$

According to the moment matching method, the distribution of E_n^{total} can be approximated by a Gamma distribution $\Gamma(k_n, \theta_n)$ with the same moments where

$$k_n = \frac{\mu^2_{E_n^{\mathrm{total}}}}{\sigma^2_{E_n^{\mathrm{total}}}} \quad \text{and} \quad \theta_n = \frac{\sigma^2_{E_n^{\mathrm{total}}}}{\mu_{E_n^{\mathrm{total}}}}. \tag{D.25}$$

3) CH-outage probability: By ergodicity of the energy arrival process, the head selection process, and the energy saving process, one can rewritten the CH-outage probability in (10) as

$$\begin{aligned}
p_n^{\mathrm{out}} &= \lim_{t\to\infty} \frac{1}{N_{\mathrm{F}}\sum_{l=1}^t I_{\mathrm{CH}_n(l)}} \sum_{l=1}^t I_{\mathrm{CH}_n(l)} \sum_{i=1}^{N_{\mathrm{F}}} (1 - I_{\mathrm{Tx}_n(i)}) \\
&= \lim_{t\to\infty} \frac{t}{\sum_{l=1}^t I_{\mathrm{CH}_n(l)}} \lim_{t\to\infty} \frac{1}{tN_{\mathrm{F}}} \sum_{l=1}^t I_{\mathrm{CH}_n(l)} \sum_{i=1}^{N_{\mathrm{F}}} (1 - I_{\mathrm{Tx}_n(i)}) \\
&\overset{(a)}{=} \frac{\mu_{T_n}\overline{p_n^{\mathrm{cond}}}}{N_{\mathrm{F}}} \sum_{i=1}^{N_{\mathrm{F}}} \mathbb{E}_{T_n, E_n^{\mathrm{h}}(i), E_n^{\mathrm{sv}}(i)}\left[I_{\{E_n^{\mathrm{r}}(i) < E_n^{\mathrm{CH}}(i)\}}\right] \\
&\overset{(b)}{=} \sum_{j=0}^{N_{\mathrm{F}}-1} \frac{N_{\mathrm{F}} - j}{N_{\mathrm{F}}} \mathbb{E}_{E_n^{\mathrm{total}}}\left[\Pr\left\{j\mu_{E_n^{\mathrm{CH}}} < E_n^{\mathrm{total}} < (j+1)\mu_{E_n^{\mathrm{CH}}}\right\}\right] \\
&= \sum_{j=0}^{N_{\mathrm{F}}-1} \frac{N_{\mathrm{F}} - j}{N_{\mathrm{F}}\theta_n^{k_n}\Gamma(k_n)} \int_{j\,\mu_{E_n^{\mathrm{CH}}}}^{(j+1)\,\mu_{E_n^{\mathrm{CH}}}} x^{k_n - 1} e^{-\frac{x}{\theta_n}} \mathrm{d}x
\end{aligned}$$

where (a) follows the weak law of large numbers [23] and in (b), we have used $E_n^{\mathrm{CH}} = \mu_{E_n^{\mathrm{CH}}}$ to approximate the energy consumed by CH node n in a frame.

By rearranging the items, we have

$$\begin{aligned} p_n^{\mathrm{out}} &= \frac{1}{\theta_n^{k_n}\Gamma(k_n)} \int_0^{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}} x^{k_n-1} e^{-\frac{x}{\theta_n}} \mathrm{d}x - \sum_{j=0}^{N_{\mathrm{F}}-1} \frac{\int_{j\mu_{E_n^{\mathrm{CH}}}}^{(j+1)\mu_{E_n^{\mathrm{CH}}}} j\mu_{E_n^{\mathrm{CH}}} x^{k_n-1} e^{-\frac{x}{\theta_n}} \mathrm{d}x}{N_{\mathrm{F}}\theta_n^{k_n}\Gamma(k_n)\mu_{E_n^{\mathrm{CH}}}} \\ &\geq \gamma\left(k_n, \frac{N_{\mathrm{F}}\mu_{E_n^{\mathrm{CH}}}}{\theta_n}\right) - \frac{k_n\theta_n}{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}} \gamma\left(k_n+1, \frac{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}}{\theta_n}\right), \end{aligned} \tag{D.26}$$

where $\gamma(k,x) = \frac{1}{\Gamma(k)}\int_0^x t^{k-1}e^{-t}\mathrm{d}t$ is the lower incomplete Gamma function.

On the other hand,

$$\begin{aligned} p_n^{\mathrm{out}} &= \frac{N_{\mathrm{F}}+1}{N_{\mathrm{F}}\theta_n^{k_n}\Gamma(k_n)} \int_0^{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}} x^{k_n-1} e^{-\frac{x}{\theta_n}} \mathrm{d}x - \sum_{j=0}^{N_{\mathrm{F}}+1} \frac{\int_{j\,\mu_{E_n^{\mathrm{CH}}}}^{(j+1)\,\mu_{E_n^{\mathrm{CH}}}} (j+1)\mu_{E_n^{\mathrm{CH}}} x^{k_n-1} e^{-\frac{x}{\theta_n}} \mathrm{d}x}{N_{\mathrm{F}}\theta_n^{k_n}\Gamma(k_n)\,\mu_{E_n^{\mathrm{CH}}}} \\ &\leq \frac{N_{\mathrm{F}}+1}{N_{\mathrm{F}}} \gamma\left(k_n, \frac{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}}{\theta_n}\right) - \frac{k_n\theta_n}{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}} \gamma\left(k_n+1, \frac{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}}{\theta_n}\right). \end{aligned} \tag{D.27}$$

Using (D.26) and (D.27), and also noting that N_{F} is generally large, we have

$$p_n^{\mathrm{out}} = \gamma\left(k_n, \frac{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}}{\theta_n}\right) - \frac{k_n\theta_n}{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}} \gamma\left(k_n+1, \frac{N_{\mathrm{F}}\,\mu_{E_n^{\mathrm{CH}}}}{\theta_n}\right).$$

This completes the proof. ∎

REFERENCES

[1] J. K. Hart and K. Martinez, "Environmental Sensor Networks: A revolution in the earth system science?" *Earth-Science Reviews*, vol. 78, no. 3, pp. 177–191, Oct. 2006.

[2] P. E. Ross, "Managing care through the air," *IEEE Spectrum*, vol. 41, no. 12, pp. 26–31, Dec. 2004.

[3] K. K. Khedo, R. Perseedoss, and A. Mungur, "A wireless sensor network air pollution monitoring system," *Int. J. of Wireless and Mobile Netw.*, vol. 2, no. 2, pp. 31–45, May 2010.

[4] L. Mainetti, L. Patrono, and A. Vilei, "Evolution of wireless sensor networks towards the Internet of Things: A survey," in *Proc.* 19th *Int. Conf. on Software, Telecommun. and Comput. Netw.* (*SoftCOM'*11), Split, Croatia, Sep. 2011, pp. 1–6.

[5] A. Kouche, "Towards a wireless sensor network platform for the internet of things: Sprouts WSN platform," in *Proc. IEEE Int. Conf. on Commun.* (*ICC*12), Ottawa, ON, Canada, Jun. 2012, pp. 632–636.

[6] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: survey and implications," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 3, pp. 443–461, Mar. 2011.

[7] Z. Niu, X. Guo, S. Zhou, and P. R. Kumar, "Characterizing energy-delay tradeoff in hyper-cellular networks with base station sleeping control," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 4, pp. 641–649, Apr. 2015.

[8] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[9] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks," in *Proc. ACM Mobi-Com 2000*, Boston, MA, Aug. 2000, pp. 56–67.

[10] D. Braginsky and D. Estrin, "Rumor routing algorithm for sensor networks," in *Proc.* 1st *Wksp. Sensor Netw. and Apps.*, Atlanta, GA, Oct. 2002, pp. 22–31.

[11] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 660–670, Oct. 2002.

[12] A. Manjeshwar and D. P. Agarwal, "TEEN: a routing protocol for enhanced efficiency in wireless sensor networks," in *Proc.* 1*st* *Int. Wksp. on Paral. and Distrib. Comput. Issues in Wireless Netw. and Mobile Comput.*, San Francisco, CA, Apr. 2001.

[13] F. Kuhn, R. Wattenhofer, and A. Zollinger, "Worst-case optimal and average-case efficient geometric Ad hoc routing, in *Proc.* 4th *ACM Intl. Conf. Mobile Comput. and Netw.*, New York, NY, USA, Jun. 2003, pp. 267–78.

[14] J. Li and D. Liu, "DPSO-based clustering routing algorithm for energy harvesting wireless sensor networks," in *Proc. Int. Conf. on Wireless Commun.* & *Signal Process.* (*WCSP*'15), Nanjing, China, Oct. 2015, pp. 1–5.

[15] Z. A. Eu, H.-P. Tan and W. K. G. Seah, "Opportunistic routing in wireless sensor networks powered by ambient energy harvesting," *Comput. Netw.*, vol. 54, no. 17, pp. 2943–2966, Dec. 2010.

[16] Z. A. Eu and H.-P Tan, "Adaptive opportunistic routing protocol for energy harvesting wireless sensor networks" in *Proc. IEEE Int. Conf. on Commun.* (*ICC*12), Ottawa, ON, Canada, Jun. 2012, pp. 318–322.

[17] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[18] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 6–18, Jan. 2006.

[19] M. K. Jakobsen, J. Madsen, and M. R. Hansen. "DEHAR: a distributed energy harvesting aware routing algorithm for Ad-hoc multi-hop wireless sensor networks," in *Proc. IEEE Int. Symp. World of Wireless Mobile and Multimedia Netw.* (*WoWMoM*'10), Montreal, QC, Canada, Jun. 2010, pp. 1–9.

[20] R. W. Heath, M. Kountouris, and T. Bai, "Modeling heterogeneous network interference using poisson point processes," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4114–4125, Aug. 2013.

[21] Wikipedia contributors, "Button cell," *Wikipedia, The Free Encyclopedia*, https://en.wiki pedia.org/w/index.php?title=Button_cell&oldid=689713099.

[22] J. M. Jornet and I. F. Akyildiz, "Joint energy harvesting and communication analysis for perpetual wireless nanosensor networks in the terahertz band," *IEEE Trans. Nanotech.*, vol. 11, no. 3, pp. 570–580, May 2012.

[23] G. Grimmet and D. Stirzaker, *Probability and Random Processes*, Oxford University Press, 2001.