IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received August 5, 2021, accepted August 25, 2021, date of publication August 27, 2021, date of current version September 3, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3108469

Enhancing Harvested Energy Utilization for Energy Harvesting Wireless Sensor Networks by an Improved Uneven Clustering Protocol

QIAN REN[1] **AND GUANGSHUN YAO**[1,2]

[1]School of Computer and Information Engineering, Chuzhou University, Chuzhou, Anhui 239000, China
[2]School of Computer Science and Engineering, Southeast University, Nanjing, Jiangsu 211189, China

Corresponding author: Guangshun Yao (yaogshun@chzu.edu.cn)

This work was supported in part by China Postdoctoral Science Foundation under Grant 2018M632209, in part by Anhui Provincial Natural Science Foundation under Grant 1808085MF177, in part by the Key Project of Anhui University Science Research under Grant KJ2019A0642, and in part by Anhui Provincial Excellent Talents Cultivation Funding Project under Grant gxgnfx2020103 and Grant gxyq2021216.

ABSTRACT The energy limitation in traditional wireless sensor networks is effectively ameliorated by equipping energy harvesting modules and rechargeable batteries on nodes in energy harvesting wireless sensor networks (EH-WSNs). However, enhancing the harvested energy utilization is still a challenge. In this paper, we proposed an improved uneven clustering protocol to enhance the harvested energy utilization of EH-WSNs. The protocol contains cluster establishment and data collection stages. To reduce the energy consumed for cluster head (CH) selection and reserve more energy for data transmission, a novel CH selection scheme is proposed to select nodes with better performance as CHs. To further enhance the harvested energy utilization, a dynamic transmission power adjustment scheme is designed for both CHs and cluster members in the data collection stage under the limited capacity of rechargeable batteries. A series of experiments are conducted to verify the effectiveness of the proposed algorithm, and the results demonstrate that the proposed algorithm utilizes the harvested energy more efficiently and performs better than the corresponding competitors.

INDEX TERMS Energy harvesting wireless sensor networks, cluster head, data transmission, energy utilization.

I. INTRODUCTION

Owing to the low cost and convenience, wireless sensor networks (WSNs) have received growing attention and have been applied in diverse areas [1], [2]. In traditional WSNs, sensor nodes are typically powered by batteries. Due to the finite stored energy in batteries and the inconvenience of battery replacement, energy efficiency has become a critical issue for WSNs [3]. Recently, many research efforts have been devoted to design energy-efficient algorithms to prolong the network lifetime [3]–[5], among which include cluster-based routings. Clustered WSNs are typically composed of a sink node and a certain number of clusters. Each cluster contains a cluster head (CH) and cluster member (CM) nodes. The CHs are responsible for receiving data from CMs, aggregating the received data, and then forwarding the data to the sink. Therefore, CHs consume more energy than CMs. To balance the energy consumption among the nodes, the CH is usually rotatory selected in a cluster. Many cluster-based routings have been designed to prolong the lifetime of WSNs [6]–[8]. However, as long as the battery capacity is finite, energy exhaustion of the batteries in sensors is inevitable.

Recently, energy harvesting (EH) technology has been integrated into WSNs to ameliorate energy limitations [9]. WSNs powered by energy harvesting devices are called energy-harvesting wireless sensor networks (EH-WSNs). The sensor nodes in EH-WSNs equipped with energy harvesting modules can extract energy from external sources, such as solar, thermal, vibration, and RF energy. In theory, they can work permanently until hardware failures occur if the consumed energy is less than the harvested energy [10]. However, EH-WSNs have unique features, that is, the uncontrollability and dynamics of available ambient energy, heterogeneous energy harvesting efficiencies among nodes, and limited

The associate editor coordinating the review of this manuscript and approving it for publication was Jie Tang.

capacity of rechargeable batteries. For example, sensor nodes equipped with solar panels harvest more energy on sunny days than on rainy or cloudy days. They generally receive more energy in sunny areas than in shady areas. Furthermore, the harvestable energy is uncontrollable and dynamic owing to dynamic external sources. The harvested energy is discarded if the rechargeable batteries are fully charged. Moreover, the nodes distributed in the monitored area have different metrics, such as node density, distance to the sink, and signal-to-noise ratio (SNR) of data transmission. Owing to the above features, cluster-based protocols for traditional WSNs cannot be adopted in EH-WSNs directly [4], [11]. Recently, many cluster-based protocols have been designed for EH-WSNs [12], [13]. However, the harvested energy utilization efficiency is ignored when designing data transmission schemes in these cluster-based protocols. Furthermore, the competition-based CH selection scheme adopted in these protocols consumes considerable energy and seriously degrades the performance of EH-WSNs.

In this paper, we consider solar as the ambient energy for sensor nodes and propose an improved cluster-based protocol called CPMHE to enhance the harvested energy utilization of EH-WSNs. It enhances the utilization by reducing the energy consumed for CH selection and dynamically adjusting the transmission power of both CHs and CMs for data transmission. The main contributions of this study are as follows.

(1) To reduce the energy consumed for CH selection and reserve more energy for data transmission, a novel CH selection scheme is proposed to select nodes with better performance as CHs. Many parameters that affect the performance of CHs have been considered.

(2) A dynamic transmission power adjustment scheme is proposed for both CMs and CHs to reasonably adjust their transmission power for data transmission under the limited capacity of rechargeable batteries.

(3) Extensive experiments were conducted, and the experimental results verified the effectiveness of CPMHE in using harvested energy.

The remainder of this paper is organized as follows. Related work is discussed in Section II. Section III describes the relative models, including the network model, energy consumption model, and harvest energy prediction model. Section IV presents the proposed CPMHE method. The evaluation of our algorithm and the analysis of the obtained results are included in Section V. Finally, we present the main conclusions and future work in Section VI.

II. RELATED WORK

Owing to the advantages in terms of both energy efficiency and scalability, cluster-based routings have received increasing attention and have been applied in increasing applications. The low-energy adaptive clustering hierarchy (LEACH) [14] is a well-known cluster-based data transmission approach. To prevent fast battery draining, every node in a cluster can get a chance to become the CH based on a predefined probability for the rotation of the head role among sensor nodes in LEACH. After a sensor node is selected as the CH, the other nodes are CMs and are connected to the CH with the minimal energy needed to reach the CH. Thereafter, the CH receives data from CMs in the same cluster. Based on LEACH, several algorithms have been proposed in recent decades [6], [8], [15]. However, all these algorithms are designed for traditional WSNs and cannot be directly adopted by EH-WSNs.

By equipping energy harvesting modules and rechargeable batteries on sensor nodes, nodes in EH-WSNs can harvest ambient energy and sustain a perpetual lifetime [16]. In the last few decades, many cluster-based protocols have been proposed for EH-WSNs [12], [13]. For CH selection, both Yujia *et al.* [17] and Sah and Amgoth [18] selected CHs based on the residual energy and predicted harvested energy of the sensor nodes. Different methods were adopted to predict the harvested energy (or energy harvest ratio). Yujia *et al.* [17] adopted a long short-term memory (LSTM) neural network to predict harvested energy, whereas Sah and Amgoth [18] used the exponential weighted moving average (EWMA). Sharma and Bhondekar [19] also considered traffic heterogeneity factors when selecting CHs in addition to the residual energy and energy harvesting rate of nodes in EH-WSNs. In contrast to the above algorithms that focus on the residual energy and energy harvesting rate of nodes for CH selection, Tang *et al.* [10] selected CHs based on the consumed energy and the estimated harvested energy of nodes. To address the mismatch between the energy harvesting process and the real energy demand, Dong *et al.* [20] proposed a cluster-based routing protocol called DEARER. The DEARER protocol selects nodes with high energy arrival rates or being close to the sink to serve as the CH nodes. Recently, Haq *et al.* [21] proposed a cluster-based algorithm called E^2-MACH for EH-WSNs. It selects CHs based on a weighted function defined by multiple attributes, such as link statistics, neighborhood density, current residual energy, and the rate of energy harvesting of nodes. However, the data transmission scheme was not introduced in [21]. For all the above cluster-based algorithms for EH-WSNs, similar to LEACH, the competition scheme was adopted to select the CHs. The candidate nodes need to transmit a large amount of status information for CH competition. As a result, a large amount of energy is consumed, and the performance of EH-WSNs is degraded. Furthermore, the limited capacity of the rechargeable battery is rarely considered when designing cluster-based routings in these algorithms. Therefore, the harvested energy is discarded after the rechargeable battery is fully charged, which decreases the harvested energy utilization efficiency.

In contrast to the above competition-based algorithms, some researchers have adopted other novel schemes for CH selection and proposed cluster-based routing protocols. Recently, Ren and Yao [22] proposed an energy-efficient CH selection scheme for EH-WSNs. The scheme divides all nodes in a cluster into three types: CH, CM, and the scheduling node (SN). The SN is used to monitor and store the real-time residual energy of all nodes, including the CMs and the CH in the same cluster. During the CH selection phase, the SN specifies a CM as a new CH according to the monitored results. The energy consumed by CH selection is reduced. However, the nodes in a cluster, including the CH and CMs, must transmit many messages to the SN for residual energy confirmation. This process also consumes a large amount of energy and degrades the performance of the EH-WSNs. In contrast to the scheme in [22], Zhang *et al.* [4] and Bozorgi *et al.* [23] adopted the waiting time for CH selection in EH-WSNs. When CH selection is triggered, each node gets a waiting time by considering some parameters and then starts a timer. Nodes with better parameters, such as more residual energy and more harvested energy, can get less waiting time. The node that reaches its waiting time first sends a CH win message and then changes its role as a CH. The other nodes that are still waiting and listening can receive the CH win message. Thereafter, they stop their timers and change their roles as CMs. The sensor nodes do not transmit any status information for CH selection. Therefore, the energy consumed for CH selection is reserved, and the performance of the EH-WSNs is improved. The parameters used to design the waiting time have a significant influence on the CH selection results. Zhang *et al.* [4] considered the ratio of the residual energy and energy gain when designing the waiting time. Bozorgi *et al.* [23] considered the energy level and amount of energy harvesting. However, some other vital metrics, such as the distance to the sink, node density, predicted harvest energy prediction, and SNR, are ignored in these algorithms. Furthermore, although the limited capacity of the rechargeable battery is considered in both [4] and [22], the limitations for adjusting the transmission power are too loose when designing cluster-based routings. As a result, some nodes adjust their transmission power irrationally and then drain their energy rapidly, which prevents the EH-WSNs system from working normally and then decreasing the harvested energy utilization efficiency. To the best of our knowledge, the problem considered in this study has not yet been studied, even though it occurs widely in practice.

III. NETWORK AND ENERGY CONSUMPTION MODEL FOR EH-WSNs

A. NETWORK MODEL FOR EH-WSNs

The EH-WSNs considered in this study contain N stationary sensor nodes and a resource-rich sink. Each sensor is powered by a capacity-limited rechargeable battery and a photovoltaic panel, which enables the nodes to harvest energy from the solar. The entire EH-WSNs are divided into multiple clusters, and each cluster contains a CH and several CMs. The main tasks performed by CMs are to sense and transmit the collected data to their CHs. Thereafter, the CHs aggregate the received data from the CMs or other CHs, and then transmit the aggregated data to the sink. We assume that the EH-WSNs system considered in this study has the following characteristics:

(1) It is a static network in which the nodes cannot move after deployment.

(2) Every node has a unique ID $i(1 \leq i \leq N)$ and knows its own position and that of the sink. The location can be obtained by GPS at deployment or localization protocols, which are beyond the scope of this paper.

(3) The transmission range of the sink can cover the entire deployment area, and the wireless transmission power of each node can be adjusted based on the distance between the receiver and itself. The maximum and minimum distances between the node and sink are d_{max} and d_{min} respectively.

(4) The limited rechargeable battery capacities of all nodes are identical and are denoted as E_{cap}. The residual energy of node i is expressed as E_i. The initial energies of all the nodes are identical and are denoted as E_{ini}.

B. ENERGY CONSUMPTION AND HARVESTING PREDICTION MODEL

We adopted the first-order radio model in this study. Let d_{ij} denote the distance between nodes i and j. The energy consumed for node i sending k-bit data to node j is

$$E_{tx}\left(k, d_{ij}\right) = k \times E_{ele} + \begin{cases} \varepsilon_{fs} \times d_{ij}^2 & d_{ij} < d_0 \\ \varepsilon_{mf} \times d_{ij}^4 & d_{ij} \geq d_0 \end{cases}, \tag{1}$$

and the energy consumed for node j receiving k bit data is

$$E_{rx}\left(k, d_{ij}\right) = k \times E_{ele}. \tag{2}$$

where ε_{fs} and ε_{mf} are the propagation loss coefficients. The value of n is determined by the transmission distance d_{ij} and a predefined threshold, $d_0 = \sqrt{\varepsilon_{fs}/\varepsilon_{mf}}.n = 2$ if $d_{ij} < d_0$ and $n = 4$ otherwise.

In EH-WSNs, harvest energy prediction is a significant issue because ambient resources are uncontrollable and change dynamically [18]. An accurate harvest energy prediction algorithm can effectively improve network performance [24]. Many solar energy prediction models have been proposed in recent decades, such as the LSTM neural network [16], EWMA [18], [25], Accurate Solar Energy Allocation (ASEA) [25], and Profile Energy Prediction (Pro-Energy) [26]. Because of the advantages of adaptability to dynamic change, low computational complexity, and fewer sample requirements, ASEA is adopted to predict the harvested energy in this study. The expected value of the harvested energy of node i in time slot $t+1$, i.e., $EH_i^{exp}(t+1)$, is calculated as follows in ASEA:

$$EH_i^{exp}(t+1) = \alpha \times EH_i^{exp}(t) + (1-\alpha)\, EH_i^{rel}(t), \tag{3}$$

where $EH_i^{exp}(t)$ and $EH_i^{rel}(t)$ denote the expected value of harvested energy and the real harvested energy of node i in time slot t respectively, α is the weight parameter between 0 and 1.

Thereafter, the predicted harvested energy of node i in time slot $t+1$, i.e., $EH_i^{pre}(t+1)$, is obtained by revising $EH_i^{exp}(t+1)$ as follows according to the influence of

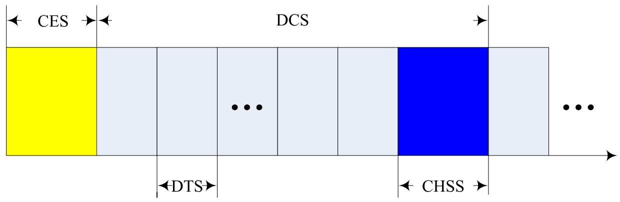

FIGURE 1. The block diagram of CPMHE.

weather factors on energy acquisition in the short term,

$$EH_i^{pre}(t+1) = \varphi_i(t) \times EH_i^{exp}(t+1), \tag{4}$$

where $\varphi_i(t)$ denotes the revision factor of node i in time slot t, and can be calculated using the following equation:

$$\varphi_i(t) = \frac{EH_i^{rel}(t)}{EH_i^{exp}(t)}. \tag{5}$$

IV. ALGORITHM IMPLEMENTATION

The working process of the CPMHE is divided into two stages: the cluster establishment stage (CES) and the data collection stage (DCS), as shown in Fig. 1. The main task of CES is to divide the entire monitored area into multiple uneven clusters and select an initial CH for each cluster. The DCS is divided into the data transmission stage (DTS) and CH selection stage (CHSS). The DTS is used to collect data from the CMs. Owing to the unstable and uneven harvested energy, CMs adopt different sampling rates for sustainable working in the CPMHE. The sampling rate of CMs is decided by some relative algorithms, such as the algorithm proposed in [27]. The CHs in this stage keep working in listening and receiving data. At the end of each data reception cycle, the CHs aggregate the received data and then forward the data to the sink. The CHSS is triggered to select a new CH for each cluster when the residual energy of the working CH is less than a predefined threshold. In the following sections, we elaborate on the details of these stages.

A. THE CLUSTER ESTABLISHMENT STAGE (CES)

This stage is designed for the initial clustering. Its main task is to divide the entire EH-WSNs into multiple uneven clusters and select an initial CH for each cluster. The CES procedure is formally described in Algorithm 1.

After all nodes are deployed, the sink sends a message *Partion_Cluster*(d_{max}, d_{min}) to all nodes for clustering. As the transmission range of the sink can cover all the deployed nodes, as mentioned in Section III.A, all nodes can receive *Partion_Cluster*(d_{max}, d_{min}). Thereafter, node i calculates its distance to the sink, i.e., d_{is}, according to the strength of the received signal $RSSI_i$. Then, node i sends a message *Cluster*(i, d_{is}, E_i, R_i^c) for clustering, where R_i^c is the competitive radius for node$i$ and can be calculated by the following equation:

$$R_i^c = (1 - \beta\frac{d_{max} - d_{is}}{d_{max} - d_{min}})R^C \tag{6}$$

Algorithm 1 CES()

The sink sends *Partion_Cluster*(d_{max}, d_{min});
for *(each node i)* ***do***
  if (node i has received *Partion_Cluster*(d_{max}, d_{min})) ***then***
    Calculate d_{is} according to $RSSI_i$;
    Get R_i^c by (6);
    Send a message *Cluster*(i, d_{is}, E_i, R_i^c) for clustering;
    Get its node density des_i;
    Calculate its signal-to-noise ratio (SNR_i) by (7);
  endif
end for
Adopt the method proposed in [28] for clustering;
All nodes work in listening state;
The sink sends *CH_Select()*;
for *(each node i)* ***do***
  if (node i has received *CH_Select()*) ***then***
    Calculate WTC_i^m by (8);
    Start a wait timer and work in listening state;
  endif
end for
for *(each node i)* ***do***
  if (receive *CH_Win(j,m)*) ***then***
    Stop its wait timer;
    Change its role as CM;
  else
    if (the timer reaches the wait time WTC_i^m) ***then***
      if (only node i has WTC_i^m) ***then***
        Change its role as CH;
        Broadcast *CH_Win(i,m)*;
      else
        Select node i' with the most residual energy as CH;
        Broadcast *CH_Win(i',m)*;
      endif
    endif
  endif
end for
for *(each node i in cluster m)* ***do***
    Decide the sampling rate by the method proposed in [27];
    Send *Sam_Rat*(i, j, m) to CH (node j);
end for
if (the CH (node j) in cluster m has received sampling rate messages from all CMs) ***then***
  Decide its server time slot for each CM;
  Broadcast the server time slots to all CMs;
endif

where β is a constant coefficient between 0 and 1, and R^C is the predefined maximum competition range of all the nodes.

Thereafter, the method proposed in [28] is adopted to divide the entire monitored area into multiple uneven clusters.

The clusters closer to the sink are smaller than those farther from the sink. As a result, the CHs closer to the sink consume less energy for intra-cluster data transmission and reserve some energy for inter-cluster data transmission. More details regarding uneven clustering are given in [28]. By receiving the clustering message from neighbors in the above clustering, node i is able to obtain the node density des_i in its deployment area. Node$i$ also obtains its signal-to-noise ratio (SNR_i) as follows:

$$SNR_i = 10\log_{10}(\frac{Power_i^{sig}}{Power_i^{noi}}), \tag{7}$$

where $Power_i^{sig}$ is the effective power of the signal and $Power_i^{noi}$ is the effective power of the noise.

After clustering, the sink sends a message *CH_Select()* to all nodes for starting the CH selection. After receiving *CH_Select()*, each node triggers a wait timer for CH selection and works in the listening state. For node i in cluster m, the wait time WTC_i^m is defined as the time interval between the time when node i receives *CH_Select()* and the time when node i sends the CH competition message. WTC_i^m is calculated by

$$WTC_i^m = \delta_1 \times \frac{E_i}{E_{cap}} + \delta_2 \times \frac{EH_i^{pre}(t+1)}{\max(EH_i^{pre})} + \delta_3 \times \frac{d_{is}}{d_{max}} + \delta_4 \times \frac{des_i}{\max(des)} + \delta_5 \times \frac{SNR_i}{\max(SNR)}, \tag{8}$$

where δ_1, δ_2, δ_3, δ_4, and δ_5 are constant coefficients between 0 and 1, and $\delta_1 + \delta_2 + \delta_3 + \delta_4 + \delta_5 = 1$. In (8), $\frac{E_i}{E_{cap}}$ leads to nodes that have more residual energy having more likely to be selected as CHs. $\frac{EH_i^{pre}(t+1)}{max(EH_i^{pre})}$ results in nodes with higher predicted harvested energy having more chances to be selected as CHs. Distance is an important parameter that affects energy consumption in wireless communication, as shown in (1). In (8), $\frac{d_{is}}{d_{max}}$ introduces the system to select nodes that are closer to the sink as CHs. $\frac{des_i}{max(des)}$ leads to the nodes at the center of their neighbors to be selected as CHs with more chances. Furthermore, the signal transmission quality has an important influence on the system performance in EH-WSNs. The last part in (8), i.e., $\frac{SNR_i}{max(SNR)}$, leads the system to select nodes with higher SNR as CHs.

After getting WTC_i^m and triggering the timer, node i works in the listening state and waits for the end of its wait time. During the waiting process, node i stops its wait timer and changes its role as a CM if it receives a CH win message *CH_Win*(j, m) from its neighbors in the same cluster. Otherwise, it announces itself as a CH at the end of its wait time. It also sends a CH win message *CH_Win*(i, m) to its neighbors in cluster m. If multiple nodes have equal wait times, node i' with the most residual energy is selected as the CH, because CH requires more energy for data transmission.

Owing to the unstable and uneven harvested energy, each CM in a cluster adopts a different sampling rate for sustainable working. After receiving the CH win message, each CM node i adopts the method proposed in [27] to determine its sampling rate, and then forwards the sampling rate message *Sam_Rat*(i, j, m) to the CH node j in the same cluster. Thereafter, the CH receives all sampling rate messages from the CMs in the same cluster and then decides its server time slot for each CM. Furthermore, each CH decides its data forwarding cycle according to its residual energy, storage capacity, and the sampling rate of CMs in the same cluster.

Algorithm 2 DCS()

Node i collects data periodically;
if ($E_i \geq \varphi \times E_{cap}$) ***then***
  Node i calculates $E_{tx}(l, d_{is})$ and $EH_i^{pre}(t+1)/T$;
  if ($EH_i^{pre}(t+1)/T \geq E_{tx}(l, d_{is})$) ***then***
    Node i adjusts its transmission power and sends data to the sink directly;
  else
    Node i sends data to the CH node j;
  endif
endif
CH j aggregates the received data periodically;
if ($E_j \geq \varphi \times E_{cap}$) ***then***
  CH j calculates $E_{tx}(l, d_{js})$ and $EH_j^{pre}(t+1)/T$;
  if ($EH_j^{pre}(t+1)/T \geq E_{tx}(l, d_{js})$) ***then***
    CH j adjusts its transmission power and sends the data to the sink directly;
  else
    CH j sends the aggregated data to the sink by multi-hop routing;
  endif
endif
if ($E_j < \gamma \times E_{ini}$) ***then***
  CH j broadcasts *CH_Adj*(j, m);
  Execute lines 14-43 in Algorithm 1 to select a new CH;
endif

B. THE DATA COLLECTION STAGE (DCS)

This stage is divided into data transmission stage (DTS) and CH selection stage (CHSS). In DTS, each CM wakes up in its working time slot and sends the collected data to the CH in the same cluster. If the residual energy of a CH in a cluster is less than a predefined threshold, the CHSS is triggered to select a new CH for this cluster. All clusters take the same process in this stage. Therefore, we only consider one cluster as an example to introduce the DCS process, which is formally described in Algorithm 2.

During the DTS stage, the CMs send the collected data to the CHs. However, they work at different sampling rates to maintain sustainable working, as mentioned in Section IV.A. Therefore, the CHs need to keep working in the listening state and receive data from CMs. The CHs also aggregate the received data and then forward the data to sink periodically. In cluster-based WSNs, multi-hop routing is usually adopted to save energy for CHs [3]–[5]. In contrast to the finite energy of nodes in traditional WSNs, the energy of nodes in EH-WSNs can be replenished by harvesting energy from external sources. However, the capacity of rechargeable batteries is limited. The harvested energy is discarded if the battery is fully charged. To enhance the harvested energy utilization efficiency, we propose a dynamic transmission power adjustment scheme for both CMs and CHs.

After each data transmission, the CM node i in cluster m checks its residual energy, E_i. If $E_i \geq \varphi \times E_{cap}$ (φ is a constant coefficient between 0 and 1), it calculates the energy consumption $E_{tx}(l, d_{is})$ for sending packets with l bit data to the sink directly with distance d_{is}. Thereafter, it compares $E_{tx}(l, d_{is})$ with $EH_i^{pre}(t+1)/T$, where T is the number of times the data are transmitted in the last energy harvesting cycle. Node i adjusts its transmission power and sends the data to the sink directly if $EH_i^{pre}(t+1)/T \geq E_{tx}(l, d_{is})$, and transmits data to the CH node j otherwise. The transmission power adjustment scheme for CHs is similar to that for CMs. However, the subscript i needs to be replaced by j if the above process is performed by the CH node j and multi-hop routing is used if $EH_j^{pre}(t+1)/T < E_{tx}(l, d_{js})$.

During the working process of the DTS stage, if the residual energy E_j of CH node j in cluster m is less than $\gamma \times E_{ini}$ (γ is a constant coefficient between 0 and 1), the CHSS stage is triggered and the CH broadcasts a request message for CH adjustment *CH_Adj(j, m)*. To reserve sufficient energy for node j to work continuously, we set $\gamma = 0.3$. Thereafter, the CH selection scheme (lines 14-43 in Algorithm 1) is adopted to select a new node as the CH of this cluster. Note that the message *CH_Select()* in Algorithm 1 needs to be replaced by *CH_Adj(j, m)* in this stage.

V. PERFORMANCE EVALUATION

A. SIMULATION SETUP

NS-3 was selected as the platform to evaluate the performance of the proposed CPMHE. For the simulation, we distributed 300 energy-harvesting sensor nodes in a two-dimensional area (500 m × 500 m). We also deployed the sink node at (250 m, 250 m). Each sensor node is equipped with a rechargeable battery with a maximal capacity of 100 J and a solar panel with dimensions of 10 mm × 10 mm. The updated National Solar Radiation Database statistical summaries [29], which holds solar and meteorological data for 1454 locations in the United States, is used as the solar power harvesting characteristic during the simulation. We also assume that 20% of nodes are randomly selected for deployment in shady areas. We set the energy-harvesting rate of nodes in shady areas to be 30% of the harvesting rate in sunny areas. All the parameters for the simulations are listed in Table 1.

TABLE 1. Simulation parameters setting.

Parameter	Value
Size of network	500 m × 500 m
Number of sensor nodes	300
Location of the sink	(250m, 250m)
E_{ini}	60J
E_{cap}	100J
E_{ele}	50nJ/bit
ε_{fs}	10pJ/bit/m^2
ε_{mf}	0.0013pJ/bit/m^4
Size of each packet	80 bytes
$\alpha, \beta, \gamma, \varphi$	0.5, 1/3, 0.3, 0.7
$\delta_1, \delta_2, \delta_3, \delta_4, \delta_5$	0.2, 0.2, 0.2, 0.2, 0.2
Size of solar panel	10mm × 10mm
The percent of nodes in the shaded area	20%

As the cluster-based routing protocols for traditional WSNs cannot be used in EH-WSNs directly, we only compare the proposed CPMHE with the cluster-based routing protocols designed for EH-WSNs. We compared CPMHE with the Uneven Clustering protocol for EH-WSNs (UCEH) proposed in [17], which is one of the state-of-the-art cluster-based routing protocols for EH-WSNs. At the beginning of each round, UCEH selects some nodes as tentative CHs by comparing the predefined threshold with a randomly generated numbers. Thereafter, UCEH adopts a competition-based scheme for CH selection from tentative CHs based on the residual energy and predicted harvested energy of these nodes. In the data transmission phase, UCEH always adopts multi-hop routing without considering the residual energy of the nodes and the capacity limitation of the rechargeable battery. We also compare CPMHE with CREW [4], as CREW adopts the similar waiting time scheme for CH selection in EH-WSNs. Furthermore, CREW also considers the limitations of the rechargeable battery when designing the routing protocol. The following metrics were used to evaluate the performance of all algorithms:

(1) The residual energy ratio after new CH selection (RERA) is defined as the ratio between the residual energy of a newly selected CH and the average residual energy of CMs in the same cluster with this CH, reflecting the energy quality of newly selected CHs, as the CHs play an important role in cluster-based routing, as mentioned in [6], [12], and [13].

(2) The residual energy ratio before new CH selection (RERB) is defined as the ratio between the residual energy of a CH and the average residual energy of CMs before new CH selection in the same cluster, reflecting the energy quality of the CH before selection [4]. In EH-WSNs, nodes can harvest energy from the environment, and the residual energy of the nodes is dynamically changed. After the CH completes its task as a CH role in a given time slot, it needs to keep working as a CM or a CH in the next time slot. Therefore, each CH must reserve sufficient energy for future working.

(3) The ratio of packet loss (RPL) is defined as the ratio between the number of packets sent by CMs and the number of packets received by the sink, reflecting the reliability of data transmission.

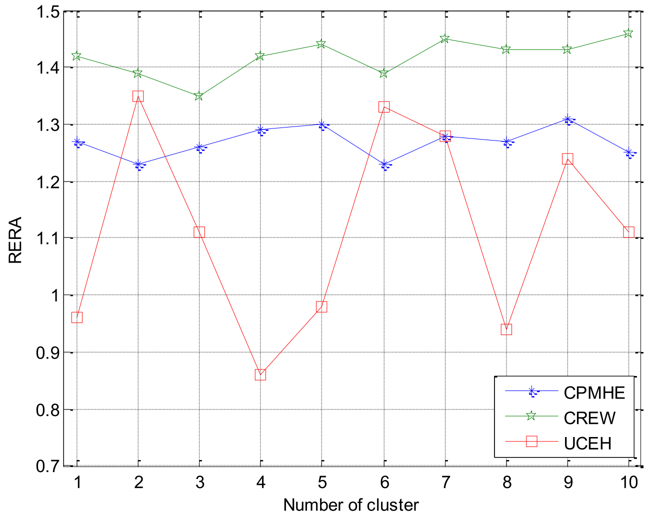

FIGURE 2. **Experimental results on RERA.**

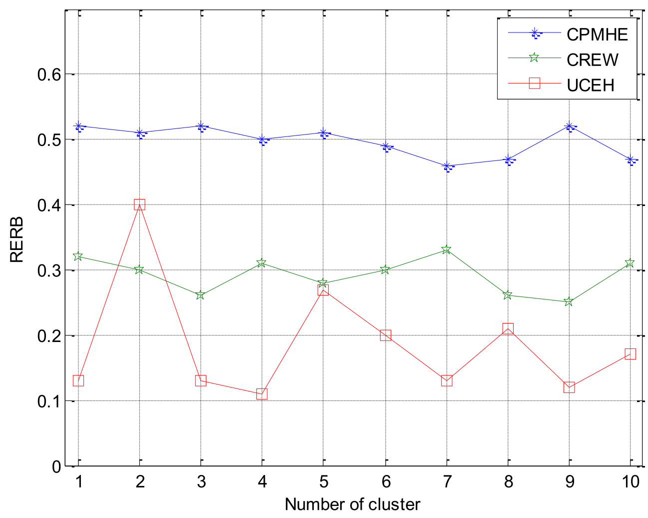

FIGURE 3. **Experimental results on RERB.**

(4) The average delay of packet delivery (ADPD) is defined as the average delay taken by all packets delivered from their initial sensor nodes to the sink, reflecting the tardiness of information transmission.

(5) The ratio of available energy utilization (RAEU) is defined as the ratio of the harvested energy by all nodes over the ambient harvestable energy to all nodes during each time quantum, reflecting the acquisition rate of the ambient harvestable energy.

(6) The harvested energy utilization efficiency (HEUE) is defined as the ratio of the number of packets received by the sink in one of the compared algorithms over the packets received by the sink in LEACH under the same energy consumed by the EH-WSNs system, reflecting the harvested energy utilization efficiency.

B. EVALUATION OF EXPERIMENTAL RESULTS

First, we compare the RERAs of all algorithms, and the results are shown in Fig. 2, which indicates that both CREW and CPMHE can provide stable RERAs while UCEH produces unstable results. Furthermore, CREW obtains the best results for RERA among all the compared algorithms, and the proposed CPMHE obtains moderate results. This phenomenon is due to the CH selection schemes adopted in these algorithms. Both CREW and CPMHE adopt the waiting time scheme for CH selection, whereas UCEH uses a competition-based scheme based on a randomly generated number. The different parameters considered for calculating the waiting time between the CREW and our proposed CPMHE leads to different RERAs. Only the residual energy and the energy gain of the nodes are considered in CREW for calculating the waiting time. In the proposed CPMHE, more factors, such as residual energy, predicted harvested energy, distance to the sink, node density, and signal transmission quality, are considered. As a result, the weight of the residual energy for CH selection in CREW is higher than that in CPMHE, and CREW can select nodes with more residual energy as new CHs. Although CREW obtains better RERAs than CPMHE, CPMHE provides better results on other compared metrics, as shown in the following sections. Furthermore, by adjusting the weights of δ_1, δ_2, δ_3, δ_4, and δ_5, the proposed CPMHE can select the nodes with more residual energy as CHS. For UCEH, some nodes are first selected as tentative CHs by comparing the predefined threshold with a randomly generated number. The selection process has a great deal of randomness, and some nodes with less residual energy are also selected as tentative CHs when they generate higher random numbers. If the nodes with less residual energy are selected as tentative CHs, the next competition-based scheme for CH selection can only be completed among these nodes and cannot select nodes with more residual energy as CHs. The randomness in selecting tentative CHs leads to unstable results of the RERA for UCEH.

Second, we focus on the compared results of RERBs of these algorithms, and the compared results are shown in Fig. 3, which shows that the RERBs of all algorithms are less than 1. This is because more tasks are completed by CHs. In cluster-based WSNs, the CHs need to receive data from CMs, aggregate the received data, and then forward the data. The CMs only sense the environmental data and transmit the data to the corresponding CH. More tasks consume more energy. Therefore, less energy was left in the CHs before the new CH selection. Fig. 3 also shows that CPMHE obtains the best results about RERBs among all the compared algorithms, while UCEH has unstable RERBs. This phenomenon is due to the CH selection scheme and the trigger scheme for new CH selection. First, the different CH selection schemes in these algorithms select nodes with different residual energies as CHs. As explained in Fig. 2, both CPMHE and CREW adopt waiting time schemes and select nodes with more residual energy as CHs. UCEH uses a competition scheme for CH selection based on a randomly generated number. As a result, the residual energy of the selected new CHs in the UCEH has a great deal of randomness, as shown in Fig. 2. Therefore,

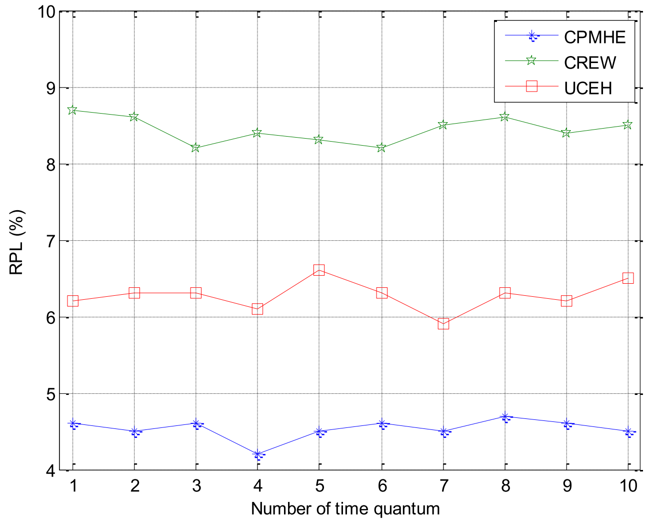

FIGURE 4. Experimental results on RPL.

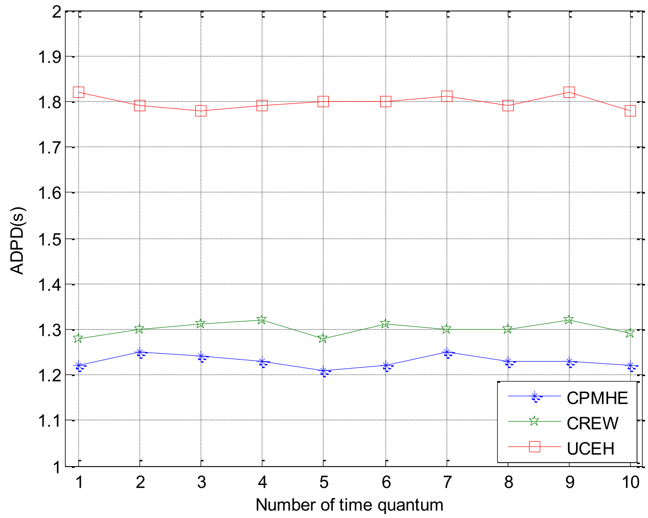

FIGURE 5. Experimental results on ADPD.

the energy of the newly selected CHs is heterogeneous in these algorithms, and the residual energy of these selected CHs is also heterogeneous after working as CHs for a while. Second, the trigger scheme for new CH selection also influences the results of RERBs. CPMHE triggers new CH selection when the residual energy of the CHs is less than $\gamma \times E_{ini}$. Therefore, CPMHE can reserve sufficient energy for CHs to work in the next time slot. Both CREW and UCEH trigger new CH selection after multiple rounds without considering the reserve energy for CHs working in the future. Therefore, CPMHE obtains the best RERBs, and UCEH obtains unstable results.

Third, we turn to the RPLs of all the algorithms, and the experimental results are shown in Fig. 4. It shows that CPMHE obtains the best results among all algorithms, and CREW obtains the worst RPLs. Generally, the RPL is proportional to two factors: the residual energy of the two communicators and the corresponding distance. Among all compared algorithms, only the proposed CPMHE considers these two factors to select CHs, as shown in (8), which contains five factors for CH selection. The third factor in (8), i.e., $\frac{d_{is}}{d_{max}}$, introduces the nodes near the sink to be selected as CHs. This reduces the distance from the CHs to the sink and then decreases the packet loss in the inter-cluster transmission. At the same time, the fourth factor, $\frac{des_i}{max(des)}$, leads to nodes at the center of their neighbors to be selected as CHs with more chances. This reduces the distance from the CMs to their CHs and decreases the packet loss in intra-cluster transmission. Furthermore, the SNR is considered in (8). Generally, nodes with better performance on the SNR transmit data more efficiently. Both CREW and UCEH only consider the residual energy for CH selection and ignore the distance between the sending and receiving nodes and the SNRs of the nodes. For data transmission, all of these algorithms adopt a multihop transmission scheme. Furthermore, both CPMHE and CREW can adjust the transmission power of the nodes. However, CPMHE adopts more stringent constraints than CREW. In CPMHE, the nodes adjust their transmission power and send data to the sink directly only when their residual energy is no less than $\varphi \times E_{cap}$ and $EH_i^{pre}(t+1)/T \geq E_{tx}(l, d_{is})$. In CREW, there is no limitation on the residual energy of the rechargeable battery, and the nodes transmit data to the sink directly if the energy consumed by sending data to the sink directly is less than the harvested energy. In other words, transmission over a long distance under insufficient energy is adopted in CREW. Therefore, long-distance transmission is used more frequently in CREW than in CPMHE. Therefore, CREW loses more packets than CPMHE does.

Next, we compare the ADPDs of these algorithms, and the results are shown in Fig. 5. This indicates that the proposed CPMHE has the best ADPDs among all algorithms, and both CPMHE and CREW obtain better results than UCEH. The ADPD is influenced by two factors: the ratio of packet loss and the number of hops used by the different algorithms to forward data. The first factor has been discussed in the previous paragraph. Fig. 4 shows that CPMHE obtains the best results on the RPL among all algorithms, and CREW obtains the worst RPLs. For the second factor, UCEH only adopts multi-hop routing, whereas both CPMHE and CREW can forward data to the sink directly in addition to multi-hop routing. Therefore, UCEH uses more hops for data transmission than CPMHE and CREW, which increases its ADPDs. For CPMHE and CREW, although CREW has more relaxed constraints on transmitting data to sink directly than CPMHE and has fewer hops for data transmission, more packet loss of CREW, as shown in Fig. 4, generates more re-transmission, which decreases its ADPDs.

We also evaluated the REAUs of these algorithms. The results are shown in Fig. 6, which shows that the proposed CPMHE obtains similar results about REAU with CREW, and both of them are better than UCEH. UCEH adopts only multi-hop routing for data transmission. When the batteries are fully charged, the harvestable energy is discarded if no changes are adopted for data transmission. Therefore, UCEH

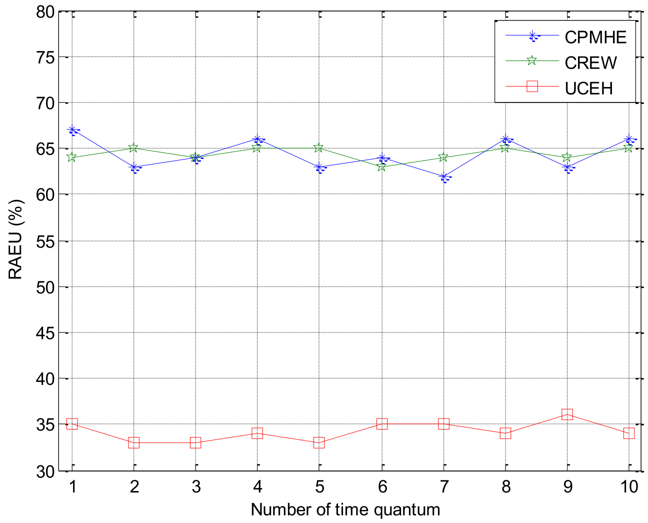

FIGURE 6. Experimental results on REAU.

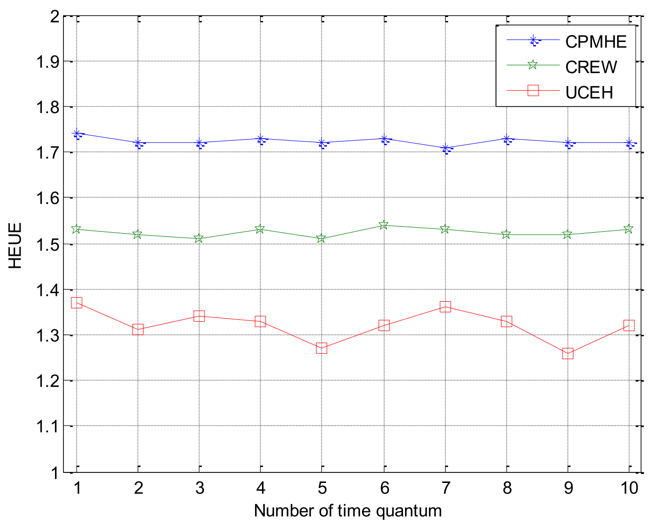

FIGURE 7. Experimental results on HEUE.

obtains lower values for REAU. For CPMHE and CREW, the sensor nodes, both CHs and CMs, transmit data to the sink directly as long as the corresponding constraints are satisfied, and thus they can harvest more energy from the environment.

Finally, we compare the HEUEs of all algorithms, and the results are shown in Fig. 7, which indicates that the proposed CPMHE obtains the best results among all the compared protocols. This phenomenon occurs because CPMHE harvests more ambient harvestable energy, as shown in Fig. 6, and transmits data with less RPL and less ADPD as shown in Fig. 4 and Fig. 5, respectively. Furthermore, the novel waiting time scheme for CH selection also decreases the energy consumption. As a result, more energy is reserved for data transmission, and the HEUE is improved.

The above simulations show that CPMHE obtains the best results on RERB, RPL, ADPD, and HUEU metrics. It also obtains moderate results on RERA and RAEU metrics. When considering all compared metrics simultaneously, it can be concluded that the proposed CPMHE has better performance than its competitors.

VI. CONCLUSION

In this paper, we propose an improved uneven clustering protocol called CPMHE to enhance the harvested energy utilization of EH-WSNs. It contains cluster establishment and data collection stages. CPMHE adopts a novel waiting time scheme for CH selection by integrating many factors. It decreases the energy consumed by CH selection and reserves more energy for data transmission. To further enhance the harvested energy utilization, CPMHE also adopts a dynamic transmission power adjustment scheme for data transmission in the data collection stage. The simulation results show that the proposed CPMHE can enhance the harvested energy utilization of EH-WSNs and have superior performance compared to its competitors.

In this paper, an improved uneven clustering protocol is proposed to enhance the harvested energy utilization of EH-WSNs. However, only one sink was considered. Recently, mobile devices, such as unmanned aerial vehicles, have received increasing attention. Therefore, we intend to integrate some mobile devices into EH-WSNs as mobile sinks and to design a corresponding routing protocol for EH-WSNs with multiple mobile sinks. Moreover, fault tolerance is not considered in this study, especially for the CHs. If some CHs fail owing to unpredictable reasons, the corresponding clusters cannot work normally. Therefore, we plan to design a fault-tolerant cluster-based routing for EH-WSNs based on this work in the future.

REFERENCES

[1] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Comput. Netw.*, vol. 52, no. 12, pp. 2292–2330, Aug. 2008.

[2] S. Li, Z. Li, J. Li, T. Fernando, H. H.-C. Iu, Q. Wang, and X. Liu, "Application of event-triggered cubature Kalman filter for remote nonlinear state estimation in wireless sensor network," *IEEE Trans. Ind. Electron.*, vol. 68, no. 6, pp. 5133–5145, Apr. 2021.

[3] P. Zhang, G. Xiao, and H.-P. Tan, "Clustering algorithms for maximizing the lifetime of wireless sensor networks with energy-harvesting sensors," *Comput. Netw.*, vol. 57, no. 14, pp. 2689–2704, Oct. 2013.

[4] X. Zhang, Y. Ding, G. Yao, and K. Hao, "An adaptive clustering routing algorithm for energy harvesting-wireless sensor networks," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Vancouver, BC, Canada, Jul. 2016, pp. 4699–4704.

[5] J. Singh, R. Kaur, and D. Singh, "A survey and taxonomy on energy management schemes in wireless sensor networks," *J. Syst. Archit.*, vol. 111, Dec. 2020, Art. no. 101782.

[6] N. Merabtine, D. Djenouri, and D.-E. Zegour, "Towards energy efficient clustering in wireless sensor networks: A comprehensive review," *IEEE Access*, vol. 9, pp. 92688–92705, 2021.

[7] D. W. Sambo, B. Yenke, A. Förster, and P. Dayang, "Optimized clustering algorithms for large wireless sensor networks: A review," *Sensors*, vol. 19, no. 2, p. 322, Jan. 2019.

[8] K. Guleria and A. K. Verma, "Comprehensive review for energy efficient hierarchical routing protocols on wireless sensor networks," *Wireless Netw.*, vol. 25, no. 3, pp. 1159–1183, Apr. 2019.

[9] T. Sanislav, G. D. Mois, S. Zeadally, and S. C. Folea, "Energy harvesting techniques for Internet of Things (IoT)," *IEEE Access*, vol. 9, pp. 39530–39549, 2021.

[10] C. Tang, Q. Tan, Y. Han, W. An, H. Li, and H. Tang, "An energy harvesting aware routing algorithm for hierarchical clustering wireless sensor networks," *KSII Trans. Internet Inf. Syst.*, vol. 10, no. 2, pp. 504–521, 2016.

[11] M. S. Bahbahani and E. A. Alsusa, "A cooperative clustering protocol with duty cycling for energy harvesting enabled wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 101–111, Jan. 2018.

[12] D. K. Sah and T. Amgoth, "Renewable energy harvesting schemes in wireless sensor networks: A survey," *Inf. Fusion*, vol. 63, pp. 223–247, Nov. 2020.

[13] D. Gurusamy and S. Abas, "Modified clustering algorithms for energy harvesting wireless sensor networks—A survey," in *Proc. 21st Int. Arab Conf. Inf. Technol. (ACIT)*, Giza, Egypt, Nov. 2020, pp. 1–11.

[14] 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.

[15] S. Tyagi and N. Kumar, "A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 36, no. 2, pp. 623–645, Mar. 2013.

[16] J. Zheng, Y. Cai, X. Shen, Z. Zheng, and W. Yang, "Green energy optimization in energy harvesting wireless sensor networks," *IEEE Commun. Mag.*, vol. 53, no. 11, pp. 150–157, Nov. 2015.

[17] G. Yujia, N. Yurong, and C. Yi, "Maximizing information transmission for energy harvesting sensor networks by an uneven clustering protocol and energy management," *KSII Trans. Internet Inf. Syst.*, vol. 14, no. 4, pp. 1419–1436, 2020.

[18] D. K. Sah and T. Amgoth, "A novel efficient clustering protocol for energy harvesting in wireless sensor networks," *Wireless Netw.*, vol. 26, no. 6, pp. 4723–4737, Aug. 2020.

[19] D. Sharma and A. Bhondekar, "An improved cluster head selection in routing for solar energy-harvesting multi-heterogeneous wireless sensor networks," *Wireless Pers. Commun.*, vol. 108, pp. 1–16, Oct. 2019.

[20] Y. Dong, J. Wang, B. Shim, and D. I. Kim, "DEARER: A distance-and-energy-aware routing with energy reservation for energy harvesting wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3798–3813, Dec. 2016.

[21] I. U. Haq, Q. Javaid, Z. Ullah, Z. Zaheer, M. Raza, M. Khalid, G. Ahmed, and S. Khan, "E2-MACH: Energy efficient multi-attribute based clustering scheme for energy harvesting wireless sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 16, no. 10, Oct. 2020, Art. no. 155014772096804.

[22] Q. Ren and G. Yao, "An energy-efficient cluster head selection scheme for energy-harvesting wireless sensor networks," *Sensors*, vol. 20, no. 1, p. 187, Dec. 2019.

[23] S. M. Bozorgi, A. S. Rostami, A. A. Hosseinabadi, and V. E. Balas, "A new clustering protocol for energy harvesting-wireless sensor networks," *Comput. Elect. Eng.*, vol. 64, pp. 233–247, Nov. 2017.

[24] D. K. Noh and K. Kang, "Balanced energy allocation scheme for a solar-powered sensor system and its effects on network-wide performance," *J. Comput. Syst. Sci.*, vol. 77, no. 5, pp. 917–932, Sep. 2011.

[25] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Trans. Embedded Comput. Syst.*, vol. 6, no. 4, p. 32, Sep. 2007.

[26] A. Cammarano, C. Petrioli, and D. Spenza, "Pro-energy: A novel energy prediction model for solar and wind energy-harvesting wireless sensor networks," in *Proc. IEEE 9th Int. Conf. Mobile Ad-Hoc Sensor Syst. (MASS)*, Las Vegas, NV, USA, Oct. 2012, pp. 75–83.

[27] S. Peng, T. Wang, and C. P. Low, "Energy neutral clustering for energy harvesting wireless sensors networks," *Ad Hoc Netw.*, vol. 28, pp. 1–16, May 2015.

[28] G. Chen, C. Li, M. Ye, and J. Wu, "An unequal cluster-based routing protocol in wireless sensor networks," *Wireless Netw.*, vol. 15, no. 2, pp. 193–207, 2009.

[29] (2011). *The National Solar Radiation Database*. [Online]. Available: https://rredc.nrel.gov/solar/old data/nsrdb/1991-2010/

QIAN REN received the M.S. degree from Hefei University of Technology, Hefei, China, in 2012. She is currently a Lecturer with the College of Computer and Information Engineering, Chuzhou University, China. Her research interests include wireless sensor networks and the Internet of Things.

GUANGSHUN YAO received the Ph.D. degree from Donghua University, China, in 2016. He is currently an Associate Professor with the College of Computer and Information Engineering, Chuzhou University, China. He is also working as a Postdoctoral Researcher with Southeast University, China. He is the author or coauthor of more than 20 academic articles, some of which have been published in international journals, such as IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Services Computing, *Information Sciences*, *Knowledge-Based Systems*, and *Soft Computing*. His research interests include cloud computing and the Internet of Things.