Hindawi
Wireless Communications and Mobile Computing
Volume 2019, Article ID 7367281, 11 pages
https://doi.org/10.1155/2019/7367281

Research Article

A Heterogeneous Energy Wireless Sensor Network Clustering Protocol

Mengjia Zeng[1,2] **Xu Huang**[1,3] **Bo Zheng**[1] **and Xiangxiang Fan**[1]

[1]School of Information Engineering, Huzhou University, Huzhou, Zhejiang Province 313000, China
[2]Qiuzhen School of Huzhou Teachers College, Huzhou, Zhejiang Province 313000, China
[3]College of Control Science and Engineering, Zhejiang University, Hangzhou, Zhejiang Province 310058, China

Correspondence should be addressed to Xu Huang; hx@zjhu.edu.cn

Received 10 October 2018; Revised 6 March 2019; Accepted 19 May 2019; Published 26 May 2019

Academic Editor: Gianluigi Ferrari

The Low-Efficiency Adaptive Clustering Hierarchical (LEACH) protocol, a hierarchical routing protocol, has the advantage of simple implementation and can effectively balance network loads. However, to date there has been a lack of consideration for its use in heterogeneous energy network environments. To solve this problem, the Energy-Coverage Ratio Clustering Protocol (E-CRCP) is proposed, which is based on reducing the energy consumption of the system and utilizing the regional coverage ratio. First, the energy model is designed. The optimal number of clusters is determined based on the principle of "minimum energy consumption", and the cluster head selection is based on the principle of "regional coverage maximization". In order to balance the network load as much as possible, in the next iteration of cluster head selection, the cluster head with the lowest residual energy and the highest energy consumption is replaced to prolong the network's life. Our simulated results demonstrate that the proposed method has some advantages in terms of longer network life, load balancing, and overall energy consumption in the environment of a heterogeneous energy wireless sensor network.

1. Introduction

Wireless sensor networks are complex and changeable working environments, and they require a large number of sensor nodes to complete measurement tasks cooperatively [1]. The reasonable placement of nodes and the optimization of parameters according to different environments help improve the overall efficiency and reduce wireless sensor network (WSN) costs. In recent years, researchers have tried to solve the energy consumption and life optimization problem of wireless sensor networks from different angles, and they have put forward many effective methods. Data routing is a problem that must be considered in wireless sensor networks, and one of the most important goals of data routing is energy saving [2]. Therefore, many researchers have tried to explore this issue from the aspects of low computational complexity, energy balance, and efficient routing. Reference [3] proposes a data decoding and fusion scheme for wireless sensor networks, which achieves data fusion in resource-constrained scenarios with low computational complexity. This method has explicit form of state estimation and residuals and is suitable for online computing. However, this method is aimed at the application of CEO (central evaluation officer) scenarios, which limits the migration of the algorithm. Document [4] studies the expansion state problem in data aggregation based on mobile agents, but the algorithm needs to calculate the dynamic migration path of mobile agents and deal with the fault and passive nodes, which increases the computational cost and hardware cost. Specific WSN software architecture design is important for maximizing network lifetime [5]. The token-based wireless sensor network cluster communication architecture in document [6] is to achieve energy-saving goals from this aspect, but the cost factor is introduced in the next hop node selection process, which increases the computing cost. Data volume in wireless sensor networks tends to grow continuously in both input and output [7]. Literature [8, 9] discussed the problem of reducing energy consumption from aspect of reducing the scale of data fusion, but literature [8] only considered the correlation of data and did not consider the correlation of adjacent sensor nodes.

Literature [9] used broadcasting mode to reduce the delay rate and prolong the life but did not discuss the flooding problem caused by broadcasting mechanism. The PED and PAD protocols proposed in literature [10] are event-driven and query-driven. The network load only becomes heavier when triggering data conditions, and most other time is lighter. Therefore, the energy consumption of network load is reduced to a certain extent. However, dynamic switching of network reporting schemes is needed in the implementation process, which increases the control overhead. Reference [11] proposes an energy-efficient cluster adaptive time division multiple access protocol EA-TDMA, which is a communication protocol between sensors in railway transportation system. This protocol improves energy efficiency by collecting information about future data packets rather than dispatching data packet exchanges in the competition stage [12], it is especially suitable for high-flow load characteristics of train operation [13], but its universality needs further verification.

The above-mentioned literature has actively explored and addressed the energy consumption of wireless sensor networks from different perspectives, while other researchers have studied this issue from the perspective of hierarchical mechanisms. The use of hierarchical mechanisms can optimize data delay to increase network scalability, reduce data redundancy and communication load, and optimize a network's lifespan [14]. Reference [15] studied the impact of uniform and nonuniform clustering on the performance of cluster sensor networks using numerical methods. It is concluded that uniform clustering has lower probability of decision-making errors than nonuniform clustering. Reference [16] implemented an efficient clustering protocol for wireless sensor networks from the perspective of fuzzy search to dynamically generate the optimal cluster number in each round using decentralization mechanism. However, the recalculation of the number of clusters per round increases the computational overhead. The EACA protocol proposed in [17] achieves better system lifetime prolongation effect, but it only considers the energy consumed by cluster head transmission and does not consider the condition that the base station is located at the far end of the network and the transmission energy consumption between cluster head and intracluster nodes. Document [18] introduces cluster sender for data transmission. Cluster heads are only responsible for allocating the time slots of TDMA within the cluster, which reduces the transmission burden of cluster heads and prolongs the network lifetime. But this method increases the election cost of cluster heads and cluster senders. From the above literatures, it can be seen that layering or clustering is an important technical means to reduce network energy consumption and prolong network life.

The typical representative of the hierarchical clustering routing protocol is the LEACH protocol [19] proposed by W. R. Heintzelman. The basic idea of this algorithm is that the cluster head nodes are randomly selected in a cyclic way, and the energy load of the whole network is equally distributed to each sensor node, which results in a reduction of the energy load and low network energy consumption that improves the overall network lifetime. The LEACH algorithm is divided into three parts: (1) cluster head election; (2) cluster members join clusters; and (3) cluster routing. At the time of election, each node generates a random number between 0 and 1. If the number is less than the threshold $T(n)$, the node will become the cluster head. The formula to calculate $T(n)$ [19] is

$$T(n)_{LEACH} = \begin{cases} \dfrac{p}{1 - p(\theta \bmod (1/p))} & \mathit{if}\ n \in G \\ 0 & \mathit{otherwise} \end{cases} \tag{1}$$

where p represents the cluster head ratio in the network; that is, the ratio of the number of cluster head nodes to the total number of nodes in the network. θ represents the current number of iterations. The set G indicates the set of nodes during the first $1/p$ iterations that are not cluster heads. From (1), we know that all nodes have the same probability of being selected as a cluster head. The energy consumption of all nodes in the system is balanced, which prolongs the life cycle of the system. Unfortunately, the LEACH algorithm has the following shortcomings: (1) the cluster head election is completely random, which may cause the cluster heads to be distributed unevenly in the monitored area. In turn, this will create an uneven global energy consumption distribution, especially for the node farthest away from the base station which may die early; (2) the scalability is poor, and communication between the cluster head and the base station is "single-hop", which is unsuitable for large-scale network applications; and (3) poor adaptability, due to diverse applications of wireless sensor networks, where the requirements of each may not be the same. As the LEACH algorithm adopts a unified, whole-network sampling and transmission period, it cannot be applied to heterogeneous networks [20–22].

In view of the shortcomings of the LEACH algorithm, considerable research has been done to address them. In [23], based on the LEACH protocol, energy-efficient and cooperative target tracking was regarded as a utility function of a cross-layer cluster optimization problem, which can obtain better simulation results, but does not involve any discussion of heterogeneous networks. A distributed algorithm proposed in [24] was used as an extension of the LEACH clustering algorithm. Although the lifetime of a network is longer than that of a LEACH network, the LEACH algorithm is only applicable to the problem of unit and nonunit circles, which has its limitations. Hence, the NEAP (Novel Energy Adaptive Protocol) energy-adaptive protocol was proposed [25]. Unlike the LEACH protocol, the threshold of NEAP is a function of the residual energy of the nodes, as shown in [25] (2):

$$T(n)_{NEAP} = \frac{p}{1 - p(\theta \bmod (1/p))} \times \left[\frac{E_{cur}}{E_{ini}} + \left(r_s \operatorname{div} \frac{1}{p}\right)\left(1 - \frac{E_{cur}}{E_{ini}}\right)\right] \tag{2}$$

where E_{cur} and E_{ini} represent the current and initial energy of the nodes, respectively, and r_s indicates the iterations where nodes have not been selected as cluster heads. Compared with the LEACH protocol, the NEAP protocol has better performance when selecting cluster heads, but to date there

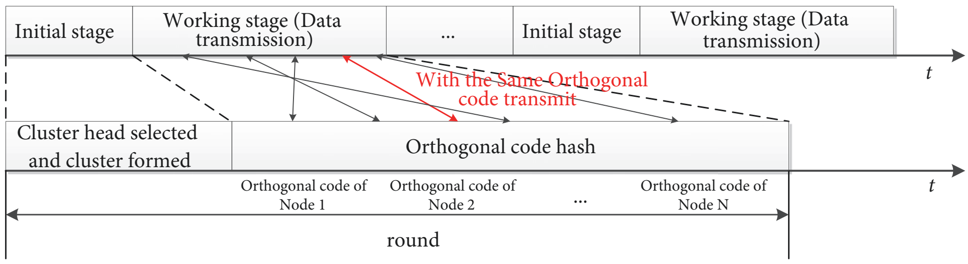

Figure 1: Graph of system timing.

has not been any discussion regarding energy optimization when selecting cluster heads. The authors of [26] proposed the Distributed Energy Efficient Cluster (DEEC) protocol, which adopts a heterogeneous, two-level energy structure network model, where each node chooses its cluster head based on its own residual energy. However, it does not consider the problem of system energy balance. The authors of [27] proposed an improved DEEC algorithm called the DDEEC (developed DEEC) protocol, which dynamically changes the standard for selecting cluster heads and then balances the energy consumed by the nodes. The authors of [28] also proposed an improved DEEC protocol called SEP (the Stable Election Protocol). The basic idea of SEP is based on the different initial energies of the nodes, which are divided into two categories: ordinary nodes and advanced nodes. Advanced nodes have a higher initial energy than ordinary nodes, and the probability of an advanced node being selected as a cluster head is higher. Reference [29] discussed the clustering protocol based on residual energy and distance information. Although [28, 29] carried out some useful explorations on the life cycle of wireless sensor networks with heterogeneous energy, the heterogeneous characteristics of node energy are not universal enough.

In view of the above problems, we propose a clustering protocol that can be applied to a heterogeneous energy wireless sensor network, the Energy-Coverage Ratio Clustering Protocol (E-CRCP). E-CRCP is an improvement of the LEACH protocol in terms of the selection of cluster heads, where we consider the lowest energy consumption in each communication iteration and the maximum cluster head coverage ratio. E-CRCP balances both the cluster head selection and the system energy load to extend the system's life cycle.

2. System Model

In order to simplify the problem, the following assumptions are made in our study: (1) the wireless sensor network is composed of a large number, N, of fixed sensor nodes; that is, once the sensor nodes are arranged in a monitoring area, the locations of the sensor nodes are no longer changed; (2) the nodes arranged in the monitoring area are subjected to a certain method to get their positional information (such as GPS); (3) all nodes are basically synchronized in second precision; (4) only one base station exists in the monitoring area, and its position is fixed in region A, which is at the center; (5) the N sensor nodes have heterogeneous energy levels; i.e., they have different initial energies; (6) the system routing model is based on a hierarchical routing protocol cluster that consists a cluster head (CH) node and several noncluster head (non-CH) nodes, which are called normal nodes. First, the normal nodes transmit their sensing data to their respective CHs, where each CH node is responsible for fusing the data from the normal nodes and forwarding it to the base station (BS).

2.1. System Timing. In this paper, we divide the system timing into several rounds, where a cycle is called an iteration. The initial and working stages are set for each iteration, as shown in Figure 1. At the initial stage, CHs are selected, and clusters are formed. Data are transmitted at the working stage. Data transmission from non-CH to CH nodes follows the principle of Code-Division Multiple Access (CDMA); that is to say, all nodes share spectrum resources in the form of orthogonal address codes; only nodes with the same orthogonal code can transmit information between sender and receiver. Nodes transmit data do not interfere with each other.

2.2. Energy Model. Each sensor node in the system needs to receive and send information in the process of data transmission. From the point of view of energy consumption, the sensor node is simplified to consist of only a receiver and an emitter, in which the emitter consists of an emitting component and a power amplifier. During data transmission, the sensor node will switch between emitting and receiving states, which means that the node is in an emitting or receiving state at any given moment. When a sensor node emits or receives data, it consumes energy. The emitter consumes energy when it runs the emitter components and power amplifiers. Assuming that the receiving and sending ends are placed at a distance d away from each other, if d is small, the free space transmission model is adopted. When d is large, the multipath fading channel model is adopted. Figure 2 shows the system radio energy consumption model.

Based on [30], we use the energy consumed by transmitting a *qbits* message between the transmitter and receiver of d is

$$E_{tr}(q,d) = \begin{cases} q \times E_{el} + q \times E_{frs} d^2 & if\ d < d_0 \\ q \times E_{el} + q \times E_{tworay} d^4 & if\ d \geq d_0 \end{cases} \quad (3)$$

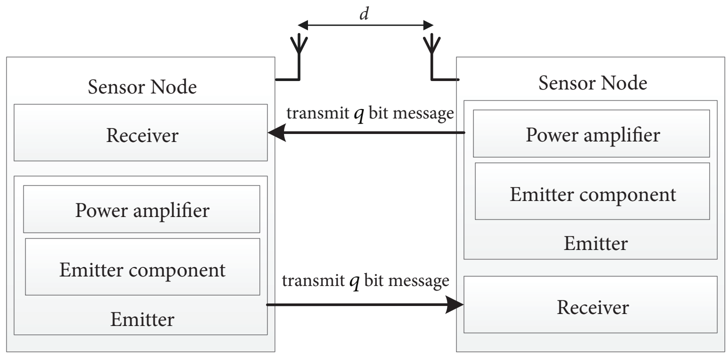

Figure 2: Radio energy consumption model.

where E_{el} represents the energy consumed per bit when the emitter components are running, E_{frs} and E_{tworay} represent the energy consumed by the unit power amplifier in the free space and the double path propagation model (two-ray ground model), respectively, and d_0 is as follows:

$$d_0 = \sqrt{\frac{(4\pi)^2 \times l \times h_t^2 \times h_r^2}{\lambda^2}} = \sqrt{\frac{E_{frs}}{E_{tworay}}} \quad (4)$$

In (4), h_t and h_r are the respective ground clearance of the sending and receiving ends and λ is the wavelength. Correspondingly, the energy consumed by receiving a *qbits* message is

$$E_{re}(q) = q \times E_{el} \quad (5)$$

2.3. Determination of the Optimal Number of Clusters. Cluster-based hierarchical routing protocols first divide nodes in the network into different clusters. How to assemble the cluster and select CH nodes is the problem that needs to be solved. The optimal probability p_{opt} of a node being a CH is an important embodiment of the clustering results. The authors of [31] proved that if the cluster number is nonoptimal, the energy consumption of the system will increase exponentially. Therefore, in this paper, we first calculate the optimal number of clusters from the perspective of the minimum energy consumption of the system. Routing protocols follow point 6 at the beginning of Section 2.

The number of nodes in the network and the initial energy of each node are different. We assume that, in iteration *Mth*, *C*-number of CHs are generated, and there are $N/C-1$ cluster member nodes in each cluster. Then, the member nodes send a *q bits* control message to the CH. The basic principle of a cluster is that the communication cost of all nodes in the cluster should be as low as possible. Generally, the CH is located at the center of a cluster, and the distance from other nodes to the CH is small. So, all member nodes that have a close distance to a CH are generally added to that cluster. It is assumed that the member nodes transmit the data to the CH based on the free space channel model. Therefore, we can get the energy consumption of each cluster member node that sends a *qbits* control message to the CHs, $E_{normal-CH}$, as

$$E_{normal-CH} = q \times E_{el} + q \times E_{frs} d_{CH}^2 \quad (6)$$

where d_{CH} is the average distance between the cluster members and the CH.

Next, let us consider a general situation where a CH is located far from the BS, and the message transmission model is a multipath fading channel model. The CH receives *qbits* control messages from the $N/C-1$ cluster member nodes and performs data fusion. Then, the total energy consumed by the transmission of the message to the base station is E_{CH-BS}:

$$E_{CH-BS} = q \times E_{el} \times \left(\frac{N}{C} - 1\right) + q \times E_{data} \times \left(\frac{N}{C} - 1\right) + q \times E_{tworay} d_{BS}^4 \quad (7)$$

The first term in (7) is the energy consumed by the *qbits* control message from the $N/C - 1$ cluster member nodes. The second term is the energy consumed when fusing the data, and the last term is the energy used to transmit the data to the base station. Among them, E_{data} indicates the energy consumed when fusing a one-bit message, and d_{BS} indicates the average distance from the CH to the base station. Therefore, the total energy of a cluster message in a communication iteration is $E_{cluster}$:

$$E_{cluster} = E_{CH-BS} + \left(\frac{N}{C} - 1\right) \times E_{normal-CH} \quad (8)$$

Then, in the *Mth* iteration, the total energy consumed by the network is E_{round}:

$$\begin{aligned} E_{round} &= C \times E_{cluster} \\ &= C \times E_{CH-BS} + (N - c) \times Enormal - C \\ &= q \times N \times E_{el} + q \times N \times E_{data} + q \times C \\ &\quad \times E_{tworay} \times d_{BS}^4 + (N - C) \times q \times E_{el} \\ &\quad + (N - C) \times q \times E_{frs} \times d_{CH}^2 \\ &= 2 \times q \times N \times E_{el} + q \times N \times E_{data} + q \times C \\ &\quad \times \left(E_{tworay} \times d_{BS}^4 - E_{el} - E_{frs} \times d_{CH}^2\right) \end{aligned} \quad (9)$$

It is hoped that the energy consumed by the network is the lowest in every iteration. Therefore, we can obtain the partial derivative C through (9) and set it to 0; this will give the optimal number of clusters C_{opt}. Accordingly, the optimal cluster head ratio p_{opt} is

$$p_{opt} = \frac{C_{opt}}{N} \tag{10}$$

2.4. Cluster Head Selection Based on the Maximum Coverage Ratio. Although implementation of the LEACH algorithm is simple, the random selection of CHs may result in a high density of them in one area, with other regions containing few or even no CHs. Hence, the CHs may be distributed unevenly throughout the system. In the process of CH selection, we should take account of the coverage ratio to prevent the uneven distribution of CHs. Coverage generally refers to the area coverage ratio [32]. Although all sensors may operate in a system, it is difficult to ensure that the coverage ratio of the target area is 100% [33]. In practical applications, small monitoring vulnerabilities have little impact on the system and are deemed acceptable. A coverage mechanism is used to ensure that nodes are kept active while meeting coverage expectations, and in this work we use the cluster coverage ratio. Based on previous work [34], the CH selection process is as follows.

2.4.1. Coverage Problem Description. Suppose the monitoring area is a rectangle with a length of h meters and a width of w meters, and the area is $h * w$ m^2. Taking h as the ordinate and w as the abscissa when establishing the two-dimensional coordinate system, we can get the coordinates of the N sensor nodes in the two-dimensional coordinate system. Let us further suppose that the sensing radius of each sensor is r, and the communication radius is R. In order to ensure network connectivity and wireless interference, $R = 2r$ [35]. Using $c_i = \{x_i, y_i, r\}$, it is shown that a circle with a radius of r is the center of the node coordinates $\{x_i, y_i\}$ $(i \in 1, \ldots, N)$. Assuming that the monitoring target coordinates are (x, y), the distance between the target and sensor nodes is $d(c_i) = \sqrt{(x_i - x)^2 + (y_i - y)^2}$. The event that the monitoring target is covered by the sensor node is e_i, and the probability $P\{e_i\}$ of the event is the probability that the target (x, y) is covered by the sensor node c_i. Next, we consider the monitoring environment and noise interference, where the probability distribution [36] of the sensor node measurement model in the actual application is given as

$$P_{cov}(x, y, c_i) = \begin{cases} 1, & if\ d \le r - r_e \\ e^{-\alpha_1 \varphi_1 \beta_1 / \varphi_2 \beta_2 + \alpha_2}, & if\ r - r_e < d < r + r_e, \\ 0, & otherwise \end{cases} \quad (i \in 1, \ldots, N) \tag{11}$$

where r_e $(0 < r_e < \mathrm{r})$ is the measurement reliability parameter of the sensor nodes and α_1, α_2, β_1, and β_2 are measured parameters related to the sensing node characteristics. φ_1 and φ_2 are input parameters:

$$\varphi_1 = r_e - r + d(c_i), \quad (i \in 1, \ldots, N) \tag{12}$$

$$\varphi_2 = r_e + r - d(c_i), \quad (i \in 1, \ldots, N) \tag{13}$$

To improve the probability that a target is measured, multiple sensor nodes are used to measure targets simultaneously. The combined measurement probability is as follows [29]:

$$P_{cov}(Cov) = 1 - \prod_{c_i \in Cov} (1 - P_{cov}(x, y, c_i)), \quad (i \in 1, \ldots, N) \tag{14}$$

The monitoring area is a rectangle of $h * w$ m^2 and is discretized into pixels. The pixel size is determined according to the actual application scenario. Whether each pixel is covered or not is measured by the joint measurement probability of node set $P_{cov}(Cov)$. In this paper, the area coverage $R_{area}(C)$ of node set C is defined as the ratio of the coverage area of node set C to the total monitoring area:

$$R_{area}(C) = \frac{\sum_{i=1}^{N} P_{cov}(Cov)}{h \times w} \tag{15}$$

Assuming that the monitoring area is a square of 20 m ∗ 20 m, it is divided into 100 pixels of equal size, and 20 sensor nodes are put into the area. A diagram of the monitoring area is shown in Figure 3, which shows the location of sensor nodes in the area. The coverage problem is described as follows: (1) use (11)-(13) to calculate the coverage of a sensor node to each pixel; (2) use (14) to calculate the joint coverage of the sensor nodes to each pixel; (3) repeat steps (1) to (2) to calculate the joint coverage rate from the sensor nodes to the pixel points; and (4) use (15) to calculate the area coverage and consider (15) as the optimization objective function of the coverage control algorithm.

2.4.2. Cluster Head Selection Algorithm. The N sensors in the monitoring area are numbered as 1-N, and we randomly select a node as the CH, assuming that the selected node is K. According to the optimal number of CHs calculated before, we need to select $C_{opt} - 1$ nodes from the remaining N-1 nodes as CHs. The selection principle is to compute the node coverage rate according to the steps mentioned above, followed by the maximum coverage rate, which is determined using (15). In the process of data communication, the energy consumption of all CHs is recorded as $E_{con_{i\vartheta}}$ $(i \in 1, \ldots, C_{opt}, \vartheta \in 1, \ldots, \theta)$, which is calculated according to (3)–(7). The cumulative energy consumption of the first θ iterations is recorded as E_{con_i} $(i \in 1, \ldots, C_{opt})$, and the residual energy of each CH is recorded as E_{rem_i} $(i \in 1, \ldots, C_{opt})$, where

$$E_{con_i} = \sum_{\vartheta=1}^{\theta} E_{con_{i\vartheta}}, \quad (i \in 1, \ldots, C_{opt}, \ \theta \in 1, \ldots, R_n), \tag{16}$$

$$E_{rem_i} = E_{0_i} - E_{con_i}, \quad (i \in 1, \ldots, C_{opt}), \tag{17}$$

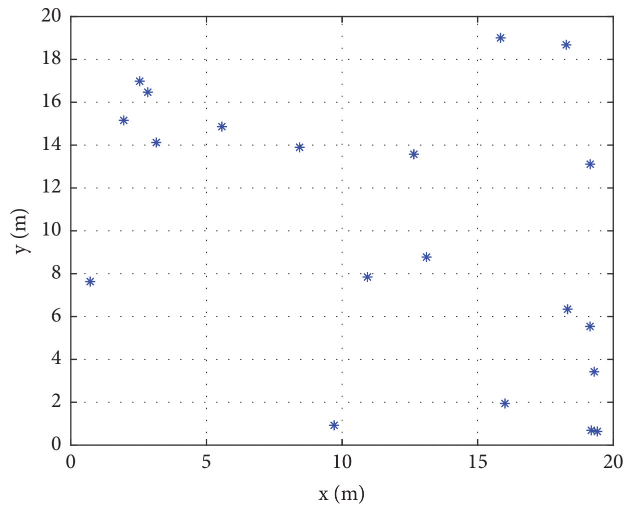

Figure 3: Random distribution of sensor nodes in 20 m ∗ 20 m monitoring area.

where E_{0_i} is the initial energy of the sensor node, ϑ is the current number of iterations, and R_n is the number of iterations at the point where the energy of the system is exhausted or the maximum number of cycles assumed by the algorithm has been reached.

Next, the CH vitality parameter η_i is introduced, which is defined as follows:

$$\eta_i = \frac{E_{rem_i}}{E_{con_i}}, \quad \left(i \in 1, \ldots, C_{opt}\right) \tag{18}$$

With the same energy consumption, the more the remaining energy is, the greater η_i is, and the higher the vitality of the CH is, the longer will be the life cycle. For the same residual energy, the more energy consumed by the current iteration is and the smaller the η_i is, the lower will be the vitality of the CH and the shorter will be the life cycle. After the end of a communication iteration, η_i of C_{opt} CHs are sorted from small to large. In order to extend the life cycle of the system as long as possible, we hope that the greater the vitality of CH nodes, the better.

The above algorithm is a one-iteration CH selection algorithm. After a data communication iteration is finished, the next iteration of CHs is selected. Because a CH needs to collect data from common nodes first and then fuse the data to send to the base station, the energy consumption of the CHs will be far greater than that of ordinary nodes. In order to balance the energy load of the whole system, we should try to let every node have the chance of becoming a CH. So, when the next iteration of CHs is selected, we need to replace the local CH and retain as many of the strongest CHs as possible. The replace proportion ρ ($\rho \in (0,1]$) is a pure decimal in the (0,1] interval, where $\rho = 1$ indicates that the C_{opt} CHs of the current iteration are all replaced, and all the next iteration CHs are selected from the noncluster head set G. A value of $\rho = 0$ indicates that all CHs remain unchanged in the next iteration, which is not exist in this model, so $\rho \neq 0$. The number of replaced CHs is $C_{rep} = C_{opt} \times \rho$, so, we replace the first C_{rep} CHs with η_i from small to large, selecting the C_{rep} CHs from the non-CH set G to maximize the region coverage in (15) and complete the next CH selection iteration. Because the data transmission of this model follows the CDMA, it can be seen from Section 2.1 that each communication iteration is divided into the initial stage of cluster head selection and clustering and the data transmission process in the working stage. The process of CH selection and clustering in the initial stage is actually a computing process. By selecting the control nodes with computing power to complete this process, the time required in the application process is fast and meets the needs. However, to ensure system efficiency, the time proportion of CH selection to the clustering process should not be too large in any communication iteration, or else system throughput will be affected. Therefore, in addition to the normal process of CH selection and clustering calculation, the system sets a time upper limit t_c for the initial stage, which is the average time for CH selection and clustering in the first θ communication iterations. If the initial stage is not completed when the t_c arrives, the system randomly generates the remaining CHs that should have been generated, but have not yet been generated. At this time, part of the algorithm degenerates to the LEACH algorithm. Figure 4 shows a flowchart of the CH selection algorithm.

In each CH selection iteration, the calculation process for each CH is mainly divided into area coverage calculation and low-energy CH replacement. In the process of calculating the regional coverage, the joint measurement probability $P_{cov}(Cov)$ is correlated with N^2 by (14), and the regional coverage $R_{area}(C)$ is also correlated with N^2 by (15), and therefore the computational time complexity of the regional coverage calculation is $O(N^2)$. In the process of CH selection and replacement, (18) shows that the calculation process is linearly related to N. Therefore, the time complexity of the whole algorithm is $O(N^2)$.

2.4.3. Clustering Process and Working Stage. After the C_{opt} CH nodes in one communication iteration have been elected, the CHs broadcast request messages to other normal nodes to join the cluster. After the non-CH nodes receive the message, they choose the nearest cluster to join until the cluster process terminates when the number of cluster nodes reaches N/C_{opt} or there are no remaining nodes. Because the value of N/C_{opt} may not be an integer, the number of nodes in the last cluster may be less than N/C_{opt}.

The working stage is also known as the data transmission stage. The CH broadcasts a CDMA data stream to notify its member nodes to start the data-acquisition process. The cluster member nodes send data to the CH according to the system timeing in Figure 1, where the CH collects the node data and then transfers them to the base station. After the data transmission is completed, the algorithm will enter the next iteration of CH selection and form a new cluster. In the working stage, data acquisition begins with the CH sending a CDMA broadcast to its member nodes. The cluster members send the collected data to their respective CHs during the CDMA process. After receiving all the data, the CHs integrate them to reduce the noise in the signal and then send the

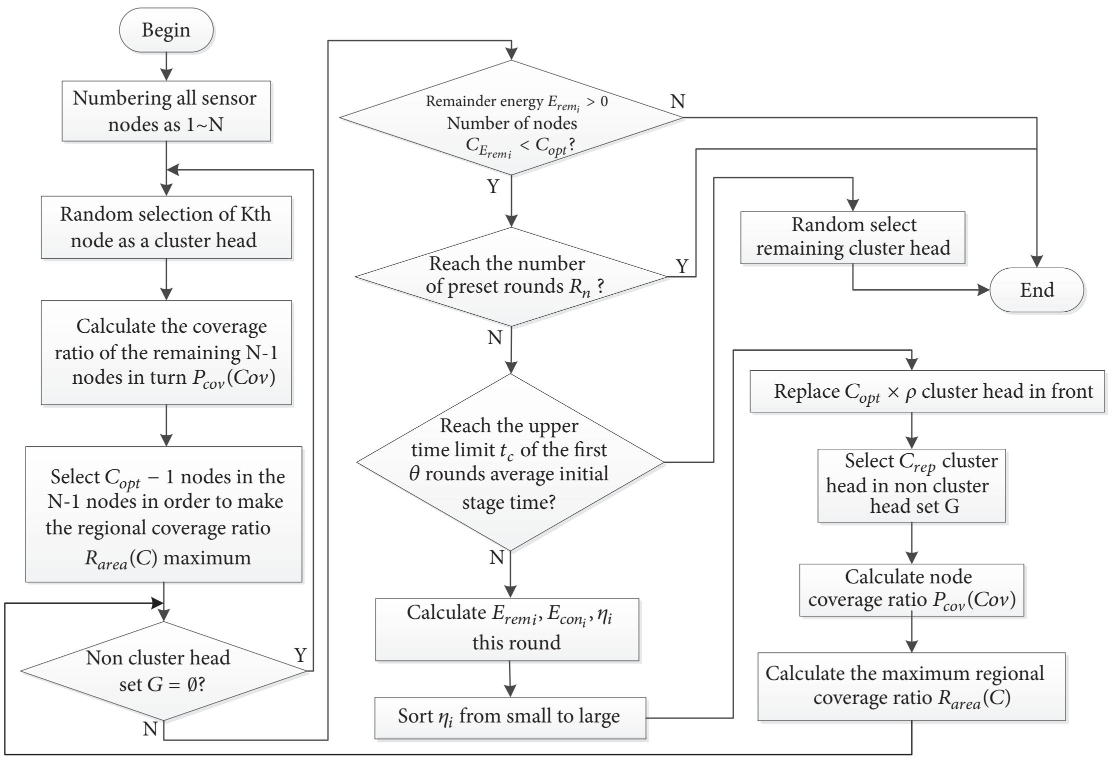

Figure 4: Cluster head selection algorithm flow chart.

Table 1: Simulation parameters.

Parameter	Value	Parameter	Value
Size of monitoring area $h \times w$	100 m×100m	Node energy coefficient α_i	Random number in (0,10), i ∈ [1, N]
Number of nodes N	100	E_{el}	50 nJ/bit
Node distribution	Random distribution	E_{frs}	10 nJ/bit/(m^2)
Sensing radius r	10 m	E_{tworay}	0.013 pJ/bit/(m^4)
Communication radius R	20 m	E_{data}	5 nJ/bit/signal
BS position	(50,50)	Size of data package	525 bytes
Node initial essential energy E_0	0.5 J	Size of CDMA package	25 bytes

integrated data to the base station in a single-hop or multihop manner. Then, the network begins to choose new CHs and form new clusters in the next iteration. When all nodes have become CHs, the next cycle will start.

3. Simulation Analysis

3.1. Simulation Parameter Hypothesis. In this section, we evaluate the performance of the proposed E-CRCP. The simulation was built using MATLAB R2016b and then compared with the LEACH, DDEEC, and SEP protocols. The specific simulation parameters are shown in Table 1.

In this experiment, the size of the monitoring area was fixed at 100m × 100m, the coordinate axis range was [(0,0), (100,100)], the number of sensor nodes N was 100, and the position coordinates were obtained by coordinate axis if they were randomly distributed in the monitoring area. The base station was located in the center of the monitoring area (50, 50). If the basic initial energy $E_0 = 0.5J$ and the energy coefficient α_i $(i \in [1, N])$ is a random number in (0, 10), then the initial energy of each node is

$$E_{0i} = E_0 \times \alpha_i \quad (19)$$

where E_{el}, E_{frs}, and E_{tworay} follow the parametric interpretation in Section 2.2 and E_{data} follows the parametric interpretation in Section 2.3.

3.2. Impact of the Cluster Head Replacement Ratio ρ on the Network Life Cycle. The purpose of the introduced CH replacement ratio ρ is to replace some low-vitality CHs in each cluster to prolong the network life cycle and balance the energy load of the whole system. The life cycle of the network, namely, the number of network lives, is expressed by the number of iterations. The value is equal to the iteration number when the last node in the network dies. The influence

Table 2: Comparison of several algorithms in terms of stability duration and network life cycle.

	Stability duration				Network life cycle			
	LEACH	DDEEC	SEP	E-CRCP	LEACH	DDEEC	SEP	E-CRCP
1	934	1255	1399	1689	5530	5860	8621	8650
2	955	1253	1396	1679	5532	5858	8625	8648
3	969	1258	1402	1682	5535	5861	8628	8651
4	936	1261	1392	1699	5540	5863	8629	8650
5	978	1251	1395	1692	5538	5864	8624	8649
6	933	1258	1403	1695	5534	5858	8626	8656
7	922	1247	1398	1698	5536	5869	8628	8655
8	955	1262	1400	1690	5538	5870	8629	8658
9	948	1263	1402	1688	5535	5865	8630	8652
10	972	1260	1398	1695	5539	5868	8625	8654

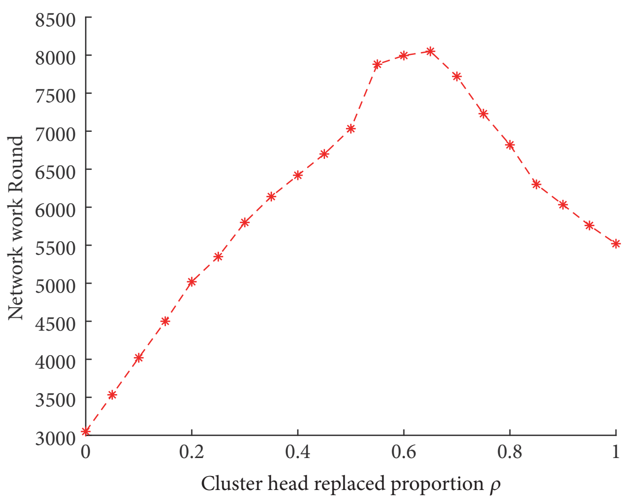

Figure 5: Impact of the cluster head replacement ratio ρ on the network life cycle.

of ρ on the life cycle of the system network is shown in Figure 5.

From Figure 5, we can see that ρ has a great influence on the life cycle of the network. A value of $\rho = 0$ indicates that all CHs in the next iteration are used in the local CH. In this case, the lifetime of the network is only maintained at about more than 3000 iterations, which is approximately half the highest point that is more than 8000 iterations. This shows that the life cycle of the CH nodes in the network determines the life cycle of the system. If the CHs are not replaced by other nodes, once all the CHs in the network are dead, the network will no longer work. It can be seen that the network without considering load energy consumption balance is not applicable. From the graph, the number of network iterations reaches the highest point near $\rho = 0.65$, which will prolong the life cycle of the network. A value of $\rho = 1$ indicates that all CHs in the current iteration are replaced, and the algorithm is identical to the LEACH algorithm. It is seen that the value of ρ first rises, reaches a maximum near $\rho = 0.65$, and then decreases, which shows that a too-large or too-small ρ has no positive influence on the life cycle of the network. Only the appropriate ρ value can prolong the network life cycle.

3.3. Comparison of Several Algorithms in Terms of Network Life Cycle. In this section, we analyze the performance of the proposed E-CRCP algorithm in two aspects: stability time and network lifetime. The stability time and network lifetime are represented by the number of iterations. The stability time is equal to the number of iterations from the initial time to the iteration when the first node dies. The network lifetime is equal to the number of iterations from the initial time to the iteration when the last node dies. First, we analyze the stability time and network lifetime of four kinds of protocols: LEACH, DDEEC, SEP, and E-CRCP. Table 2 lists the experimental data of the 10 tests. From Table 2, we can see that the stability duration of LEACH, DDEEC, SEP, and E-CRCP is 950.2, 1256.8, 1398.5, and 1690.7, respectively. The average life spans of the networks are 5535.7, 5863.6, 8626.5, and 8652.3, respectively. These indicate that the proposed E-CRCP protocol can effectively extend the stability period and the network lifetime.

From Table 2, we can see that the E-CRCP proposed in this paper can extend the stability time and network lifetime better than LEACH, DDEEC, and SEP. The reason is that the LEACH protocol does not consider the residual energy of the node and simply gives the same opportunity to each node, whereas DDEEC only considers the residual energy of the node, and SEP only considers the node energy level. However, these considerations are not conducive to the selection of a good CH. The proposed E-CRCP dynamically adjusts the replacement ratio of the CHs, which promotes the E-CRCP to select the best CH, reduces energy consumption, and thus prolongs the stability time and network life. A prolonged network lifetime means that more nodes can collect data, which helps the base station receive more data packets.

3.4. Relationship between the Number of Active Nodes and the Network Lifetime. An active node is a working node. Once the energy of the node is exhausted, the node will no longer work (i.e., it dies). As time goes on, there will be fewer and fewer active nodes in the system. This section compares the lifetime of different protocols in the network

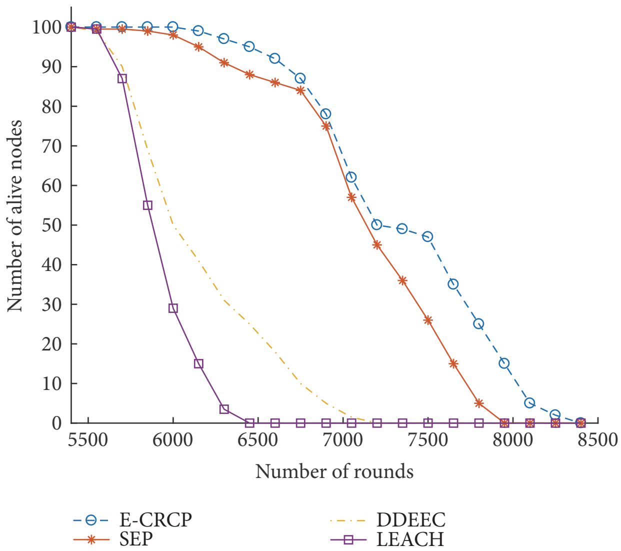

Figure 6: Relationship of the number of active nodes in the network and the network lifetime.

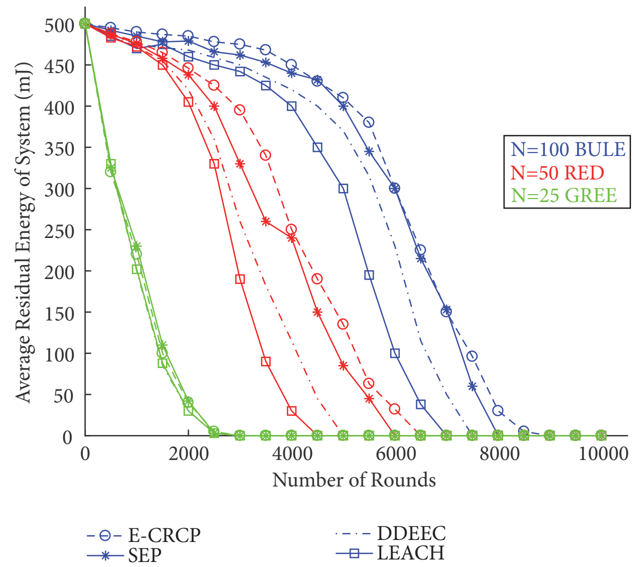

Figure 7: Influence of the number of sensor nodes on network lifetime and energy consumption.

when the number of active nodes is different. Figure 6 reflects the influence of the number of active nodes on the network lifetime.

Figure 6 illustrates the relationship between the number of active nodes and the network lifetime. As we can see from the graph, with decrease of the number of alive nodes in the network, LEACH's lifetime is about 6500 rounds, DDEEC's lifetime is about 7200 rounds, and SEP's lifetime is about 8000 rounds. That is to say, LEACH, DDEEC, and SEP to prolong the lifespan are failed. Conversely, the E-CRCP protocol can prolong the life span as the number of nodes decreases. This is because LEACH, DDEEC, and SEP do not take the cluster coverage mechanism into account. Each node needs to send all the collected environmental information to its CH nodes, including redundant information, thus increasing the energy consumption. The E-CRCP protocol not only considers the network energy load balance, but also takes the maximum coverage in the cluster into account, so that network life increases with the number of nodes decrease.

3.5. Influence of the Number of Sensor Nodes on Network Lifetime and Energy Consumption. To reveal the influence of the number of sensor nodes on network lifetime and energy consumption, the number of sensor nodes was chosen as $N = 100$, 50, and 25, and the system was given the same average initial energy (assumed to be 0.5 J) under all three models, with the other parameters set as in Table 1. Network life was still measured by the number of iterations of network operation. Figure 7 shows the relationship between network lifetime and average residual energy for several protocols.

As can be seen from Figure 7, when the number of nodes in the system is 100, 50, and 25, the average residual energy of each protocol decreases as the number of iterations increases. When $N = 100$ and 50, the change trends of several protocols are very similar. The curve of each protocol shows a convex trend, but the curve shows a concave trend when $N = 25$. When $N = 100$, the number of iterations to termination of the E-CRCP protocol was about 8500, whereas the number of iterations to termination when $N = 50$ was about 6500. In other words, the average residual energy of the system does not decrease synchronously as the number of nodes decreases exponentially but moves horizontally on the coordinate axis. This trend also exists for several other protocols. Moreover, from the graph, the performance of the proposed E-CRCP protocol is optimal when $N = 100$ or $N = 50$. This shows that there should be an appropriate range to determine the number of nodes in the monitoring area. The appropriate number of sensor nodes in the monitoring area is conducive to extending the life of the network system. However, when N drops to 25, the performance of several protocols is similar, and the average residual energy of the system shows a rapid downward trend. The reason for this is that when the number of nodes in the system is small, the communication distance between nodes is generally longer. The energy consumption transmission model of the nodes therefore changes from a free space transmission model to a multipath fading transmission model, which makes the data transmission need more energy and makes the system energy rapidly decay to 0. In addition, when N is small, the E-CRCP algorithm proposed in this paper does not have any advantages in the calculation of regional coverage and optimal cluster number, which makes its performance not necessarily superior to other protocols. Hence, when N is appropriately chosen, the performance of the E-CRCP algorithm proposed in this paper is better than that of other algorithms.

3.6. Impact of Coverage Area on Network Life. To simplify the model, it was assumed that the monitoring area was equal to the coverage area, which was a square area with equal length and width, and that the base station was located in the center

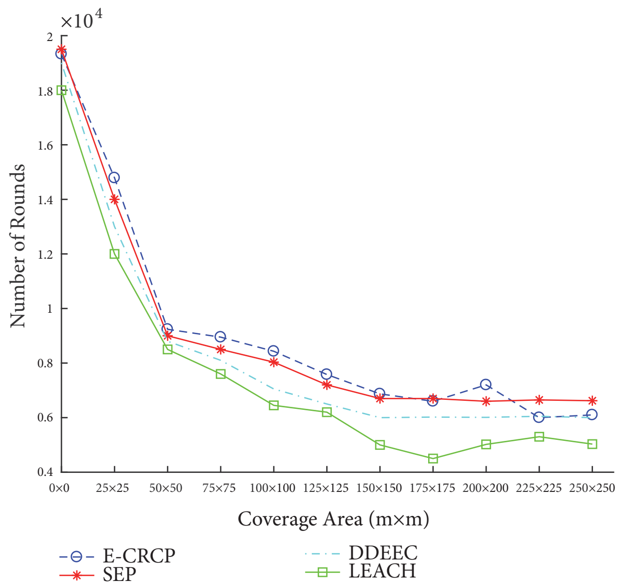

Figure 8: Impact of coverage area on network life.

of the area. The other parameters are listed in Table 1. Figure 8 shows the influence of changes in coverage area on network lifetime.

From Figure 8, it is clear that, in the range of [(0,0), (50,50)], changes in coverage have a strong influence on network lifetime, but in the range of [(50,50), (150,150)], this influence diminishes. In these two ranges, the influence of coverage on network lifetime shows a linear trend. As the coverage area continues to increase, there is no longer a stable linear trend between coverage area and network life. The reason for this is that the larger coverage area leads to a change in the data transmission model between nodes, making the state of the model unstable. It can be seen that the appropriate choice of coverage area has a positive impact on system stability.

4. Conclusions

In this paper, a CH selection protocol (E-CRCP), which is effectively applied to heterogeneous energy wireless sensor networks, is proposed as the solution to the CH selection problem in wireless sensor networks. First, a system-wide energy consumption model is established. The optimal number of system clusters is determined in the case of minimum energy consumption. Next, CH nodes are selected under the condition that the CH coverage is at a maximum, and CH nodes that consume a large amount of energy are replaced in the next communication iteration. The remaining members of each cluster join their nearest cluster and send their own data to the CH node. The CH node then sends the data to the base station after the data of each member node are fused, thus completing a single communication iteration. Our simulated results show that the algorithm proposed in this paper has obvious advantages over the LEACH, DDEEC, and SEP protocols in terms of the network lifetime of heterogeneous energy network applications. In the process of CH selection, E-CRCP reduces the overall energy consumption of the network, balances the network load, and prolongs the network life.

Data Availability

The data used to support the findings of this study are included within Table 2 of the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partly supported by Zhejiang Province Soft Science Foundation of China (2019C35006), major project of ZJFHSSC (2015Z004), major project of Center for Research in Regional Economic Open and Development of Zhejiang (16JDGH011), Zhejiang Province Natural Science Foundation of China (LY18G020008; LQ18F020002), postdoctoral fund in China (2017M621489), National Natural Science Foundation of China (61202290), Huzhou University's Scientific Research Foundation in 2018 (2018XJKJ63), and Huzhou City Zhejiang Province Key Industry Project of China (2018GG29).

References

[1] J. Yick, A. Bharathidasan, G. Pasternack, B. Mukherjee, and D. Ghosal, "Optimizing placement of beacons and data loggers in a sensor network - A case study," in *Proceedings of the 2004 IEEE Wireless Communications and Networking Conference, WCNC 2004*, pp. 2486–2491, USA, March 2004.

[2] M. H. Anisi, A. H. Abdullah, S. A. Razak, and M. A. Ngadi, "An overview of data routing approaches for wireless sensor networks," *Sensors*, vol. 12, no. 4, pp. 3964–3996, 2012.

[3] G. Ferrari, M. Martalò, and A. Abrardo, "Information fusion in wireless sensor networks with source correlation," *Information Fusion*, vol. 15, no. 1, pp. 80–89, 2014.

[4] G. P. Gupta, M. Misra, and K. Garg, "Towards scalable and load-balanced mobile agents-based data aggregation for wireless sensor networks," *Computers and Electrical Engineering*, vol. 64, no. 10, pp. 262–276, 2017.

[5] T. Cevik, "ECMTADR: energy conservative multitier architecture with data reduction for cluster-based wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2015, no. 1, pp. 1–11, 2015.

[6] T. Çevik and A. H. Zaim, "EETBR: energy efficient token-based routing for wireless sensor networks," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 21, no. 2, pp. 513–526, 2013.

[7] H. Teng, X. Liu, A. Liu et al., "Adaptive Transmission Power Control for Reliable Data Forwarding in Sensor Based Networks," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 2068375, 22 pages, 2018.

[8] H. Harb, A. Makhoul, D. Laiymani, and A. Jaber, "A distance-based data aggregation technique for periodic sensor networks," *ACM Transactions on Sensor Networks*, vol. 13, no. 4, pp. 1–40, 2017.

[9] S. Saginbekov and A. Jhumka, "Many-to-many data aggregation scheduling in wireless sensor networks with two sinks," *Computer Networks*, vol. 123, pp. 184–199, 2017.

[10] B.-D. Lee and K.-H. Lim, "An energy-efficient hybrid data-gathering protocol based on the dynamic switching of reporting schemes in wireless sensor networks," *IEEE Systems Journal*, vol. 6, no. 3, pp. 378–387, 2012.

[11] G. M. Shafiullah, S. A. Azad, and A. B. M. S. Ali, "Energy-efficient wireless mac protocols for railway monitoring applications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 649–659, 2013.

[12] M. Molodova, Z. Li, A. Nunez, and R. Dollevoet, "Automatic detection of squats in railway infrastructure," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 1980–1990, 2014.

[13] Y. Chu, S. Kosunalp, P. D. Mitchell, D. Grace, and T. Clarke, "Application of reinforcement learning to medium access control for wireless sensor networks," *Engineering Applications of Artificial Intelligence*, vol. 46, pp. 23–32, 2015.

[14] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Communications of the ACM*, vol. 43, no. 5, pp. 51–58, 2000.

[15] G. Ferrari, M. Martalò, and R. Pagliari, "Decentralized detection in clustered sensor networks," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 959–973, 2011.

[16] O. M. Alia, "A dynamic harmony search-based fuzzy clustering protocol for energy-efficient wireless sensor networks," *Annals of Telecommunications*, vol. 73, pp. 353–365, 2018.

[17] O. Zytoune, Y. Fakhri, and D. Aboutajdine, "A fairly balanced clustering algorithm for routing in wireless sensor networks," *Sensor Review*, vol. 30, no. 3, pp. 242–249, 2010.

[18] F. Bajaber and I. Awan, "An efficient cluster-based communication protocol for wireless sensor networks," *Telecommunication Systems*, vol. 55, no. 3, pp. 387–401, 2014.

[19] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd Annual Hawaii International Conference on System Siences (HICSS '00)*, vol. 2, pp. 55–60, IEEE, Washington, DC, USA, January 2000.

[20] Y. Li, N. Yu, W. Zhang et al., "Enhancing the performance of LEACH protocol in wireless sensor networks," in *Proceedings of the IEEE Conference on Computer Communications*, pp. 16–21, 2011.

[21] Y. Gui, "Research on HEDSA data aggregation of wireless sensor network," *Computer Engineering*, vol. 37, no. 7, pp. 160–164, 2011.

[22] K. S. Ahn, D. G. Kim, B. S. Sim et al., "Balanced chain-based routing protocol(BCBRP) for energy efficient wireless sensor networks," in *Proceedings of the 9th IEEE International Symposium on Parallel and Distributed Processing with Applications Workshops*, pp. 12–21, Republic of Korea, 2011.

[23] L. Arienzo, "An information-theoretic approach for energy-efficient collaborative tracking in wireless sensor networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, pp. 1–14, 2010.

[24] C.-F. Huang and Y.-C. Tseng, "The coverage problem in a wireless sensor network," in *Proceedings of the International Conference on ACM Wireless Sensor Networks and Applications*, pp. 115–121, ACM Press, New York, NY, USA, September 2003.

[25] S.-S. Wang and Z.-P. Chen, "LCM: A link-aware clustering mechanism for energy-efficient routing in wireless sensor networks," *IEEE Sensors Journal*, vol. 13, no. 2, pp. 728–736, 2013.

[26] L. Qing, Q. Zhu, and M. Wang, "Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks," *Computer Communications*, vol. 29, no. 12, pp. 2230–2237, 2006.

[27] S. H. Kang and T. Nguyen, "Distance based thresholds for cluster head selection in wireless sensor networks," *IEEE Communications Letters*, vol. 16, no. 9, pp. 1396–1399, 2012.

[28] G. Smaragdakis, M. Ibrahim, and A. Bestavros, "SEP: a stable election protocol for clustered heterogeneous wireless sensor networks," in *Proceedings of the International Workshop on SANPA*, pp. 13–21, 2004.

[29] X. Limou, L. Changyun, and M. Junfeng, "Residual energy and distance based clustering protocol for heterogeneous WSN," *Application Research of Computers*, vol. 33, pp. 2763–2766, 2016.

[30] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application on specific protocol architecture for wireless micro sensor networks," *IEEE Trans on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.

[31] C.-Y. Wan, A. T. Campbell, and L. Krishnamurthy, "PSFQ: A reliable transport protocol for wireless sensor networks," in *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, pp. 1–11, ACM Press, New York, NY, USA, 2002.

[32] L. Zhen, M. Yingchi, and C. Dong, "Effective data gathering scheme in heterogeneous energy WSN," *Application Research of Computers*, vol. 27, no. 6, pp. 2267–2270, 2010.

[33] Z. Lin, Y. Feng, and L. Yu, "Research on the strategy of wireless sensor networks coverage by the particle optimization evolutionary," *Chinese Journal of Sensors and Actuators*, vol. 22, no. 6, pp. 873–877, 2009.

[34] Y. Mao, Z. Wang, and Y. Liang, "Energy aware partial coverage protocol in wireless sensor networks," in *Proceedings of the International Conference on Wireless Communications, Networking and Mobile Computing IEEE*, pp. 2535–2538, China, September 2007.

[35] D. Tian and N. D. Georganas, "A coverage-preserving node scheduling scheme for large wireless sensor networks," in *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications (WSNA '02)*, pp. 32–41, September 2002.

[36] S. Li, C. Xu, W. Pan, and Y. Pan, "Sensor deployment optimization for detecting maneuvering targets," in *Proceedings of the 7th International Conference on Information Fusion*, pp. 1629–1635, IEEE Press, Piscataway, NJ, USA, 2005.