



Article

A Simultaneous Wireless Information and Power Transfer-Based Multi-Hop Uneven Clustering Routing Protocol for EH-Cognitive Radio Sensor Networks

Jihong Wang *, Zhuo Wang and Lidong Zhang

School of Electrical Engineering, Northeast Electric Power University, Jilin 132012, China; z1751852174@126.com (Z.W.); zhanglidong1010@126.com (L.Z.)

* Correspondence: wangjihong@neepu.edu.cn

Abstract: Clustering protocols and simultaneous wireless information and power transfer (SWIPT) technology can solve the issue of imbalanced energy consumption among nodes in energy harvesting-cognitive radio sensor networks (EH-CRSNs). However, dynamic energy changes caused by EH/SWIPT and dynamic spectrum availability prevent existing clustering routing protocols from fully leveraging the advantages of EH and SWIPT. Therefore, a multi-hop uneven clustering routing protocol is proposed for EH-CRSNs utilizing SWIPT technology in this paper. Specifically, an EH-based energy state function is proposed to accurately track the dynamic energy variations in nodes. Utilizing this function, dynamic spectrum availability, neighbor count, and other information are integrated to design the criteria for selecting high-quality cluster heads (CHs) and relays, thereby facilitating effective data transfer to the sink. Intra-cluster and inter-cluster SWIPT mechanisms are incorporated to allow for the immediate energy replenishment for CHs or relays with insufficient energy while transmitting data, thereby preventing data transmission failures due to energy depletion. An energy status control mechanism is introduced to avoid the energy waste caused by excessive activation of the SWIPT mechanism. Simulation results indicate that the proposed protocol markedly improves the balance of energy consumption among nodes and enhances network surveillance capabilities when compared to existing clustering routing protocols.

Keywords: cognitive radio sensor networks; RF energy harvesting; simultaneous wireless information and power transfer; uneven clustering routing protocol



Citation: Wang, J.; Wang, Z.; Zhang, L. A Simultaneous Wireless Information and Power Transfer-Based Multi-Hop Uneven Clustering Routing Protocol for EH-Cognitive Radio Sensor Networks. *Big Data Cogn. Comput.* **2024**, *8*, 15. <https://doi.org/10.3390/bdcc8020015>

Academic Editors: Han-Chieh Chao and Mohamed Elhoseny

Received: 21 November 2023

Revised: 19 January 2024

Accepted: 29 January 2024

Published: 31 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, the emergence of Internet of Things (IoT) technologies like smart homes and industrial automation has made daily life more convenient [1,2]. As a novel paradigm of IoT, cognitive radio sensor networks (CRSNs) enhance wireless sensor networks (WSNs) by integrating cognitive radio (CR) technology, effectively addressing the issue of spectrum scarcity in WSNs through the intelligent utilization of idle channels belonging to primary users (PUs) [3]. Nonetheless, the intrinsic energy limitations of sensor nodes are exacerbated with the introduction of CR functionalities. Once a node's energy is exhausted, it ceases to function, thereby posing a challenge to the network's efficient and enduring stability [4]. While clustering protocols alleviate this through cluster heads (CHs) that process and aggregate information from cluster members (CMs), thereby reducing data redundancy and communication distances, they do not fundamentally overcome the issue of limited node energy [5]. To guarantee sustainable and stable network operation even with significant energy demands, downlink radio frequency energy harvesting (RF EH) and simultaneous wireless information and power transfer (SWIPT) technologies have been adopted in CRSNs to supplement and equalize node energy. RF EH, distinct from traditional EH technologies that depend on fluctuating sources like solar or wind power, offers

a stable and controllable energy source without the need for complex energy collection devices such as solar cells or wind turbines [6]. SWIPT is an innovative solution for addressing the issue of energy holes in multi-hop networks, allowing for the simultaneous transmission of signals and energy [7], i.e., supplying power to wireless devices while interacting with them. Due to their advantages in energy and spectral efficiency, EH-CRSN nodes can be deployed on industrial equipment in place of traditional WSNs nodes to collect and upload various environmental information, thereby further advancing industrial automation and intelligence.

Conventional clustering protocols for CRSNs are typically divided into two types: uniform clustering protocols, where CHs closer to the sink deplete energy quickly, leading to potential energy holes or network fragmentation [8,9], and uneven clustering protocols that compute cluster radii under a predetermined uneven clustering coefficient using a linear expression of the CHs' Euclidean distance to the sink to further equilibrate network energy consumption [10]. They do not account for the variations in node energy due to SWIPT and EH when selecting CHs or relay nodes, failing to fully leverage the performance benefits of SWIPT and EH in CRSNs. Clustering protocols that include SWIPT or EH are mostly tailored for WSNs, leaving dynamic channel availability out of consideration. Therefore, there is a critical need for clustering protocols in EH-CRSNs that employ SWIPT technology to coordinate node operations and extend network lifespan.

This paper presents a CRSNs multi-hop uneven clustering routing protocol based on downlink RF EH, intra-cluster and inter-cluster SWIPT technologies. The protocol permits CRSN nodes to accumulate energy from the sink or PUs through RF EH and to replenish the energy of nodes in critical positions through intra-cluster and inter-cluster SWIPT mechanisms, thus equalizing energy expenditure within and between clusters and extending the number of rounds during which the network maintains effective monitoring capabilities. The innovations of this paper are summarized as follows:

1. Taking into account the dynamic changes in the residual energy of CRSN nodes caused by energy collection from PUs occupying channels or the sink, an EH-based energy state function is proposed, along with the selection criteria for CHs and relay nodes. Additionally, intra-cluster and inter-cluster SWIPT mechanisms are introduced to allow for the immediate energy replenishment for CHs or relays with insufficient energy while transmitting data, thereby preventing data transmission failures due to energy depletion.
2. To prevent the death of CRSN nodes due to insufficient energy, and to avoid the energy waste caused by excessive activation of the SWIPT mechanism, an energy status control mechanism is introduced. When the residual energy of nodes falls below the dormancy threshold, nodes only engage in EH and do not participate in data transmission or forwarding. Simulation results indicate that the proposed EH- and SWIPT-based multi-hop uneven clustering routing protocol (ES-MUCRP) markedly prolongs the network lifespan, improves the balance of energy consumption among nodes, and enhances the network monitoring capabilities when compared to existing clustering routing protocols for CRSNs.

2. Related Works

Current clustering routing protocols are broadly classified into non-EH clustering routing protocols and EH-based clustering routing protocols.

2.1. Non-EH Clustering Routing Protocols

Non-EH clustering routing protocols are further categorized by cluster radius into uniform and uneven clustering protocols. Uniform clustering protocols often see CHs near the sink tasked with more data transfers, leading to quicker energy depletion and potential energy holes or network fragmentation. To address this, uneven clustering protocols that adaptively adjust cluster radius sizes have been proposed. Such protocols divide the

network into clusters of varying sizes, with smaller-radius clusters located closer to the sink and larger-radius clusters situated farther away.

2.1.1. Uniform Clustering Routing Protocols

Uniform clustering routing protocols can be further categorized into centralized, distributed, and hybrid types based on their clustering approach.

1. Centralized protocols, such as Fuzzy C-means [11], ions motion optimization-based clustering routing protocol (IMOCR) [12], CogLEACH-C [13], and artificial bee colony clustering protocol [14], utilize CHs to reduce energy use and improve data efficiency while simplifying network administration, though they are limited by potential single point of failure and communication bottlenecks, which can hamper network scalability.
2. Distributed protocols like CogLEACH [15], distributed spectrum-aware clustering (DSAC) protocol [16], spectrum-aware cluster-based energy-efficient multimedia protocol [17], network stability-aware clustering (NSAC) protocol [18], and energy aware cluster based routing protocol [19] allow for more flexible CHs selection based on node energy levels, thus enhancing reliability and scalability, reducing bottlenecks and failures, and enabling more even energy consumption distribution across the network.
3. The spectrum-aware clustering algorithm based on weighted clustering metric (WCM) [20] and energy efficient spectrum aware clustering algorithm based on reinforcement learning (EESA-RLC) [21] are representative hybrid clustering routing protocols. Specifically, WCM achieves optimal clustering by solving optimization problems, selecting CHs and CMs based on temporal-spatial correlation, confidence levels, and residual energy, while minimizing energy consumption by limiting spectrum sensing to CHs. EESA-RLC protocol introduces a reinforcement learning-based clustering algorithm for spectrum sensing, establishing an energy consumption model that takes into account channel sensing, intra- and inter-cluster data transmission, and formulates the clustering process as a Markov decision problem to achieve optimal cluster configuration. This algorithm performs well in terms of energy efficiency, channel sensing accuracy, and computational complexity.

2.1.2. Uneven Clustering Routing Protocols

Uneven clustering routing protocols are divided into single-hop and multi-hop clustering routing protocols, depending on whether they can address inter-cluster routing issues.

1. In single-hop uneven clustering routing protocols, CHs far away from the sink will consume more energy to transmit their data, which will result in imbalanced energy consumption among nodes. In [22,23], the size of the cluster radius is determined by Equation (1). Candidate CHs are chosen based on the number of accessible channels, and those with substantial residual energy within the cluster radius are elected to become CHs [22]. In [23], an uneven clustering protocol is put forward based on particle swarm optimization, which seeks to minimize energy expenditure within clusters. This is achieved by implementing a multi-objective function encompassing residual energy of nodes, neighboring node count, and proximity to the sink for CHs selection within a defined circular coverage area. Additionally, it proposes two fitness functions to determine the optimal location of CHs within the circular area for comprehensive network coverage.

$$R_c = \left(1 - c \frac{d_{\max} - d(i, \text{sink})}{d_{\max} - d_{\min}} \right) R_c^0, \quad (1)$$

where R_c^0 signifies the maximum competitive radius, and c denotes the uneven clustering constant. d_{\max} and d_{\min} represent the maximum and minimum Euclidean distances

from all nodes to the sink, respectively, with $d(i, \text{sink})$ indicating the Euclidean distance from candidate CH i to the sink.

2. Multi-hop uneven clustering routing protocols can balance the energy consumption among CHs by controlling cluster radius, thereby extending the network lifespan. Reference [24] considers the nodes' residual energy when selecting candidate CHs and uses ant colony optimization to search for inter-cluster paths, which lightens the load on CHs, and the cluster radius is also determined by Equation (1). Reference [25] further factors in the residual energy of nodes, the number of neighbors, and the probability of idle channels when calculating the cluster radius. Protocols [24,25] offer an improvement in balancing energy consumption and prolonging network lifetime compared to uniform clustering protocols, but the precise value of c is not tailored to the specific network configurations. In [26], an algorithm is presented that leverages grid clustering for efficient multi-hop routing, with the explicit goal of minimizing energy consumption. The algorithm considers various parameters including network area, node location, and node energy, and introduces communication nodes to alleviate the burden of inter-cluster communications in multi-hop routing protocols. Our previous work [27] designs a multi-hop uneven clustering routing protocol for CRSNs based on intra-cluster SWIPT (S-MUCRP). It establishes criteria for selecting CHs and relay nodes based on an energy level function, assessing nodes' transmission capabilities by considering both energy availability and spectrum access, thereby facilitating the selection of qualified CHs and relay nodes for improved energy equilibrium and network connectivity. While uneven clustering protocols have successfully tackled issues related to energy holes and hot spots, they have not yet rectified the fundamental limitation of energy scarcity in sensor nodes.

2.2. EH-Based Clustering Routing Protocols

Reference [28] introduces an innovative EH approach that allows for simultaneous data and energy transmission over RF links—SWIPT, which aims to supplement and equalize the remaining energy among nodes through EH. Based on the target network, EH-based clustering routing protocols can be more specifically divided into clustering routing protocols for EH-WSNs and for EH-CRSNs.

2.2.1. Clustering Routing Protocols for EH-WSNs

Energy potential LEACH protocol [29] enhances the LEACH protocol by incorporating an energy potential function to assess nodes' EH capabilities. This enhancement has led to increased network lifespan and throughput compared to the original LEACH protocol. In the multi-hop energy-neutral clustering (MENC) routing protocol [30] for WSNs, the network is segmented into rings of equal size, and within each ring, CHs are selected based on nodes' residual energy to form clusters. MENC defines energy neutrality constraints from the energy expenditure of data transfers within and between clusters, enabling nodes to function in an energy-neutral mode. It maximizes the network information rate by optimizing the number of rings, clusters per ring, and the minimal data transmission interval, thus achieving network perpetuity. Adaptive energy harvesting aware clustering routing protocol [31] for EH-WSNs selects CHs based on the EH rate and the nodes' remaining energy, exhibiting improvements in node survival and network throughput over conventional WSNs clustering algorithms. Clustering protocols for EH-WSNs often choose CHs based on nodes' residual energy without accounting for dynamic spectrum accessibility. In contrast, EH-CRSNs clustering protocols require intra-cluster nodes to comply with a groupwise constraint, ensuring at least one common channel is available for use. Consequently, clustering routing protocols designed for EH-WSNs are not suitable for EH-CRSNs.

2.2.2. Clustering Routing Protocols for EH-CRSNs

Addressing the joint resource allocation challenge in EH-CRSNs, with constraints on EH efficiency and communication outage probabilities, [32] aims to maximize network throughput and introduces an allocation algorithm for sub-channels, power, and leasing time that increases throughput and network robustness in complex communication settings. Reference [33] proposes a clustering routing protocol for RF EH-CRSNs based on residual energy and channel quality. This protocol features a dual-level node classification algorithm: at the first level, nodes are categorized into either transmission nodes or EH nodes based on their remaining energy, and subsequently, the algorithm pairs transmission nodes with their ideal transmission channels. At the second level, transmission nodes that can deliver data packets on determined channels within the specified time perform data reporting, while others engage in EH. Studies related to EH-CRSNs clustering protocols tend to concentrate on the distribution of network resources and address only the single-hop intra-cluster communication challenge, with the complex multi-hop routing between clusters not yet being effectively addressed. The characteristic analysis and comparison of the aforementioned clustering routing protocols is shown in Table 1.

Table 1. Comparison of the characteristics of existing clustering routing protocols.

References	Target Network	Protocol Type	Considering Control Overhead	Applying SWIPT	Applying Downlink EH	Single-Hop/Multi-Hop
[11,13,14]	CRSNs	Uniform (Centralized)	×	×	×	Single-hop
[12]	CRSNs	Uniform (Centralized)	✓	×	×	Single-hop
[15]	CRSNs	Uniform (Distributed)	×	×	×	Single-hop
[16,17,19]	CRSNs	Uniform (Distributed)	×	×	×	Multi-hop
[18]	CRSNs	Uniform (Distributed)	×	×	×	Single-hop
[20]	CRSNs	Uniform (Hybrid)	×	×	×	Single-hop
[21]	CRSNs	Uniform (Hybrid)	×	×	×	Multi-hop
[22,23]	CRSNs	Uneven	×	×	×	Single-hop
[24,25]	CRSNs	Uneven	×	×	×	Multi-hop
[26]	WSNs	Uneven	×	×	×	Multi-hop
[27]	CRSNs	Uneven	✓	✓	✓	Multi-hop
[28]	—	—	—	✓	✓	—
[29]	WSNs	Uniform (Centralized)	×	×	✓	Single-hop
[30]	WSNs	Uneven	×	×	✓	Multi-hop
[31]	WSNs	Uniform (Distributed)	×	×	✓	Single-hop
[32]	CRSNs	—	×	×	✓	—
[33]	CRSNs	Uniform (Centralized)	×	×	✓	Single-hop
Ours	CRSNs	Uneven	✓	✓	✓	Multi-hop

3. SWIPT-Based Multi-Hop Uneven Clustering Routing Protocol Design for RF EH-CRSNs

3.1. Network Model

The sink is located at the center of a circular network with a radius of R , where K identical CRSN nodes and m PUs are evenly and randomly scattered. The network is segmented into z concentric rings, each with a width of R_t , and the rings are sequentially arranged from the innermost to the outermost as ring 1, ring 2, ..., ring z . A semi-Markov ON/OFF model is employed for mimicking the dynamic spectrum utilization patterns of PUs [34], while the energy consumption quantification for CRSN nodes is based on the established energy usage model in wireless communications [35].

It is presumed that CRSN nodes adhere to the following conditions:

1. Each CRSN node can perform linear EH at the beginning of each round from either the sink or the PUs engaged in communication. The sink consistently provides RF energy to CRSN nodes during a fixed period t_1 , while the PUs only supply energy to CRSN nodes when they are occupying the channels. The quantity of energy harvested by CRSN node j is indicated by Equation (2).

$$E_{EH}(j) = \begin{cases} P_{PU} \times \frac{G_T \times G_R \times \lambda^2}{16\pi^2 d_{toPU}(j)^2} \times t_1 \times \eta & \text{if } d_{to\ sink}(j) \geq \sqrt{\frac{5}{2}} \times d_{toPU}(j) \\ P_{sink} \times \frac{G_T \times G_R \times \lambda^2}{16\pi^2 d_{to\ sink}(j)^2} \times t_1 \times \eta & \text{otherwise} \end{cases} \quad (2)$$

where P_{PU} denotes the PU's transmit power, which is 40 W; P_{sink} indicates the transmit power of the sink, which is 100 W; G_T and G_R are the gains of the transmitting and receiving antennas, respectively; λ is the carrier wavelength, valued at 1/3 m; $d_{toPU}(j)$ denotes the Euclidean distance from CRSN node j to the PU; $d_{tosink}(j)$ represents the Euclidean distance from node j to the sink; and η signifies the efficiency of the linear EH process.

2. Once deployed, CRSN nodes do not change their locations.
3. CRSN nodes identify and exchange information such as their locations and remaining energy with adjacent nodes through common control channel (CCC).
4. CRSN nodes are capable of perfect spectrum sensing and opportunistic access to vacant channels for communication.
5. Relay nodes can transfer a portion of their remaining energy to the next hop while forwarding data.

3.2. Design Details of ES-MUCRP

ES-MUCRP protocol facilitates CRSNs to periodically collect and forward monitoring data to the sink. Specifically, ES-MUCRP protocol consists of four stages: spectrum sensing, linear EH, cluster formation and route establishment, and data transmission. The cluster radius in uneven clustering determines the range of control information exchange between nodes. In order to ensure the smooth execution of ES-MUCRP, we use theoretical derivation methods to determine the cluster radius in Section 3.3 and provide complexity analysis of ES-MUCRP in Section 3.4.

3.2.1. Spectrum Sensing Stage

CRSN node j monitors the channels used by PUs to identify which ones are available for use, and the count of accessible channels is denoted as $C(j)$.

3.2.2. Linear EH Stage

Each CRSN node j determines the RF energy source that can provide the maximum energy collection based on its location information, which could be either the sink or a PU currently occupying the channel, and then performs EH over the linear EH duration t_1 .

3.2.3. Cluster Formation and Route Establishment Stage

The cluster formation and route establishment stage of ES-MUCRP protocol can be subdivided into three sub-stages: CHs selection, cluster construction, and route selection, detailed as follows:

1. CHs selection sub-stage: The surviving CRSN nodes in the first ring become CHs directly, while surviving node j in other rings determines the number of neighboring nodes in the adjacent outer ring $Next(j)$ that share common available channels within the maximum communication range R_t , as well as the count of neighbors within the same ring and cluster radius $num(j)$ via control information exchanges. Subsequently, node j computes the overall energy $E_{intra}(j)$ consumed in processing

data from neighbors within the same ring and cluster radius, the energy $E_{forward}(j)$ used in forwarding data for outer rings, and the energy $E_{control}(j)$ expended in control information exchange, as shown in Equations (3)–(5). These calculations, alongside the node's remaining energy $E_{res}(j)$ and the energy harvested $E_{EH}(j)$, are used to derive the EH-based energy state function $EH_ESF(j)$, as indicated in Equation (6).

$$E_{control}(j) = 3L_1 \times (E_{elec} + E_{fs} \times R_{r(j)}^2) + 3L_1 \times E_{elec} \times (N_{r(j)} - 1), \quad (3)$$

$$E_{intra}(j) = (N_{r(j)} - 1) \times E_{elec} \times L_2 + N_{r(j)} \times E_{DA} \times L_2 + (E_{elec} + E_{fs} \times d_{CH(r(j)) \rightarrow CH(r(j)-1)}^2) \times L_2, \quad (4)$$

$$E_{forward}(j) = \frac{\sum_{k=r(j)+1}^z \frac{A_k}{S_k}}{\frac{A_{r(j)}}{S_{r(j)}}} \times [E_{elec} \times L_2 + (E_{elec} + E_{fs} \times d_{CH(r(j)) \rightarrow CH(r(j)-1)}^2) \times L_2], \quad (5)$$

$$EH_ESF(j) = \begin{cases} E_{res}(j) + E_{EH}(j) - E_{control}(j) - E_{forward}(j) & \text{if } r(j) = 1 \\ E_{res}(j) + E_{EH}(j) - E_{control}(j) - E_{intra}(j) - E_{forward}(j) & \text{otherwise} \end{cases}, \quad (6)$$

where L_1 and L_2 denote the sizes of the control and data packets, respectively; E_{elec} represents the energy consumed by electronic circuitry to send/receive a single bit of data; Under the assumption that the signal transmission adheres to the free-space path loss model, E_{fs} is the power amplifier's energy consumption per bit under this model; E_{DA} is the energy used to aggregate a bit of data; and $N_{r(j)}$ denotes the average node count in a single cluster of ring $r(j)$. $A_{r(j)}$ is the area of ring $r(j)$, with $A_{r(j)} = (2r(j) - 1)\pi R_t^2$. $S_{r(j)}$ indicates the average area of a single cluster in ring $r(j)$, with $S_i = \pi R_{r(i)}^2$. Here, $R_{r(j)}$ is the cluster radius of ring $r(j)$, and $A_{r(j)}/S_{r(j)}$ gives the total cluster count in ring $r(j)$, which corresponds to the total CHs count in that ring. Similarly, A_k/S_k signifies the total CHs count in ring k . $\sum_{k=r(j)+1}^z A_k/S_k$ represents the number of data packets from outer rings that need to be forwarded by ring $r(j)$. $d_{CH(r(j)) \rightarrow CH(r(j)-1)}$ measures the average distance from the CHs of ring $r(j)$ to the relay CHs located in ring $r(j) - 1$ or to the sink, and the specific calculation is shown in Equation (7).

$$d_{CH(r(j)) \rightarrow CH(r(j)-1)} = d_{CHtosink(r(j))} - d_{CHtosink(r(j)-1)}, \quad (7)$$

where $d_{CHtosink(r(j))}$ and $d_{CHtosink(r(j)-1)}$ represent the average distances from the CHs of ring $r(j)$ and ring $r(j) - 1$ to the sink, respectively, with $d_{CHtosink(r(j))}$ specifically illustrated in Equation (8).

$$d_{CHtosink(r(j))} = \frac{1}{\pi r^2(j) R_t^2} \int_{\theta=0}^{2\pi} \int_{r=0}^{r(j)R_t} r^2 dr d\theta = \frac{2}{3} r(j) R_t \quad (8)$$

Utilizing $EH_ESF(j)$, node j calculates the EH-based CHs selection weight $EH_W(j)$, as indicated in Equation (9):

$$EH_W(j) = \begin{cases} [\alpha \times EH_ESF(j)]^2 \times \sqrt[3]{C(j)} \times \sqrt{\frac{1}{d_{tosink(j)}}} \times \sqrt{\frac{1}{Next(j)}} \times \sqrt[3]{num(j)} & \text{if } r(j) \neq 1 \cap Next(j) \neq 0 \\ 0 & \text{if } Next(j) = 0 \\ \alpha \times EH_ESF(j)^2 \times \sqrt[3]{C(j)} \times \sqrt{\frac{1}{d_{tosink(j)}}} \times \sqrt{\frac{1}{Next(j)}} & \text{if } r(j) = 1 \cap Next(j) \neq 0 \end{cases} \quad (9)$$

where α is the weight factor that adjusts the influence of the energy state function. Once the CHs selection weight $EH_W(j)$ is determined, nodes with non-zero residual

energy outside the first ring disseminate their CHs weights within the cluster radius. Nodes receive these weights from neighbors and engage in a comparison of CHs weights. If a node's own weight is less than that of a neighboring node, it broadcasts a message to withdraw from the competition, and surrounding nodes receive this message; if the node has the highest weight among all its neighbors, it becomes a CH and broadcasts a CHs announcement message on CCC, leading to the withdrawal from the competition by the neighboring nodes. This process is repeated until all nodes have either become CHs or withdrawn from the competition.

2. Cluster construction sub-stage: Ordinary nodes that are not yet clustered look for the CH within their own transmission range that has common available channels and the highest CHs selection weight, to which they send a join request, thereby marking themselves as clustered. CHs acknowledge these join requests from the ordinary nodes and list them as their CMs. Ordinary nodes that fail to identify a CH become CHs by default. CHs that do not receive any join requests form clusters independently. Once all ordinary nodes have identified their CHs, cluster construction is completed, and the process moves to the route selection sub-stage.
3. Route selection sub-stage: Specifically, nodes in the first ring can reach the sink in a single hop, enabling them to send packets directly to the sink. Due to communication range limitations, all CHs beyond the first ring must select appropriate relay nodes to assist in forwarding data packets until they reach the sink. CH j , located in the third ring and beyond, selects two inner-ring CHs, a and b , that maximize the competition value $Compet(j)$ and record them; if no suitable nodes are discovered, CH j seeks assistance from its CM k to locate the subsequent hops and eventually determines two relay nodes, a and b , that maximize $Compet(j)$. CH j in the second ring selects the first-ring CH a that maximizes $Compet(j)$ as the next-hop relay; if none are found, then CH j uses its CM k to search for the next hop, and finally records two-hop relays, k and a , that maximize $Compet(j)$, completing the route selection. The expression for the competition value $Compet(j)$ is as shown in Equation (10).

$$Compet(j) = \begin{cases} EH_W(a) & \text{if } r(j) = 2, CM\ k \notin \mathbf{relay} \\ EH_W(a) \times EH_W(k) & \text{if } r(j) = 2, CM\ k \in \mathbf{relay} \\ EH_W(a) \times EH_W(b) & \text{if } r(j) \geq 3, CM\ k \notin \mathbf{relay} \\ EH_W(a) \times EH_W(b) \times EH_W(k) & \text{if } r(j) \geq 3, CM\ k \in \mathbf{relay} \end{cases} \quad (10)$$

where \mathbf{relay} denotes the relay set composed of CMs.

3.2.4. Data Transmission Stage

After the completion of cluster formation and route establishment, nodes enter the data transmission stage. However, frequent data transmission and relay will result in substantial energy depletion of CRSN nodes. To prevent CRSN nodes from dying prematurely due to insufficient energy, which would lead to data transmission failures, and to avoid the energy waste caused by excessive activations of the SWIPT mechanism, ES-MUCRP protocol has incorporated an energy status control mechanism that manages node states during data transfer. Specifically, this mechanism divides the state S_j of CRSN node j into three categories according to its residual energy $E_{res}(j)$: the dead state S_{death} , the sleep state S_{sleep} , and the active state S_{active} . When $E_{res}(j)$ falls below E_{death} , the node is energy-depleted and in the dead state S_{death} , unable to perform any functions and losing its environmental monitoring capabilities. When $E_{res}(j)$ is higher than or equal to E_{death} but below the dormancy threshold $E_{dormancy}(j)$, to prevent energy exhaustion, the node enters the sleep state S_{sleep} , only engaging in linear EH and not participating in data transmission, relay, or similar operations. When $E_{dormancy}(j) \leq E_{res}(j) \leq E_{max}$, the node has sufficient residual energy and is in the active state S_{active} , capable of conducting energy-intensive operations. The dormancy threshold $E_{dormancy}(j)$ for CH j consists of the energy consumed in the control

information exchange, reception, aggregation, and forwarding of intra-cluster data, as well as assistance in relaying data from outer layers per round, as shown in Equation (11).

$$E_{dormancy}(j) = 3L_1 \times (E_{elec} + E_{fs} \times R_{r(j)}^2) + 3L_1 \times E_{elec} \times (N_{r(j)} - 1) + (N_{r(j)} - 1) \times E_{elec} \times L_2 + N_{r(j)} \times E_{DA} \times L_2 + L_2 \times (E_{elec} + E_{fs} \times d_{CH(j) \rightarrow route(r(j)+1)}^2) + \frac{\sum_{k=r(j)+1}^z N_{CH}(k) \times (2E_{elec} + E_{fs} \times d_{CH(r(j)) \rightarrow CH(r(j)-1)}^2) \times L_2}{N_{CH}(r(j))} \quad (11)$$

where $d_{CH(j) \rightarrow route(r(j)+1)}$ signifies the Euclidean distance between CH j and its next-hop relay; $N_{CH}(r(j))$ quantifies the CHs in the same ring as CH j ; and $\sum_{k=r(j)+1}^z N_{CH}(k) / N_{CH}(r(j))$ is the quantity of data packets from outer rings that CH j aids in relaying.

The dormancy threshold $E_{dormancy}(k)$ for CM k comprises the energy required for control information exchange and for data delivery to the CH per round, as presented in Equation (12):

$$E_{dormancy}(k) = 3L_1 \times (E_{elec} + E_{fs} \times R_{r(k)}^2) + 2L_1 \times E_{elec} \times (N_{r(k)} - 1) + 2L_1 \times E_{elec} + L_2 \times (E_{elec} + E_{fs} \times d_{CM(k) \rightarrow CH}^2) \quad , \quad (12)$$

where $R_{r(k)}$ is the cluster radius size of the layer where CM k is located; and $d_{CM(k) \rightarrow CH}$ is the Euclidean distance from CM k to its CH.

Utilizing the aforementioned energy status control mechanism to identify the node's status, nodes in the active state carry out data transmission. Data transmission encompasses both intra-cluster and inter-cluster data transfers, detailed as follows:

1. The intra-cluster data transmission process is specifically depicted in Figure 1. It determines whether node j is a CH or a CM, and CHs that have CMs schedule time division multiple access (TDMA) time slots for their CMs, who receive the schedule information and decide whether to transfer energy to the CH through SWIPT while transmitting data based on their energy levels.

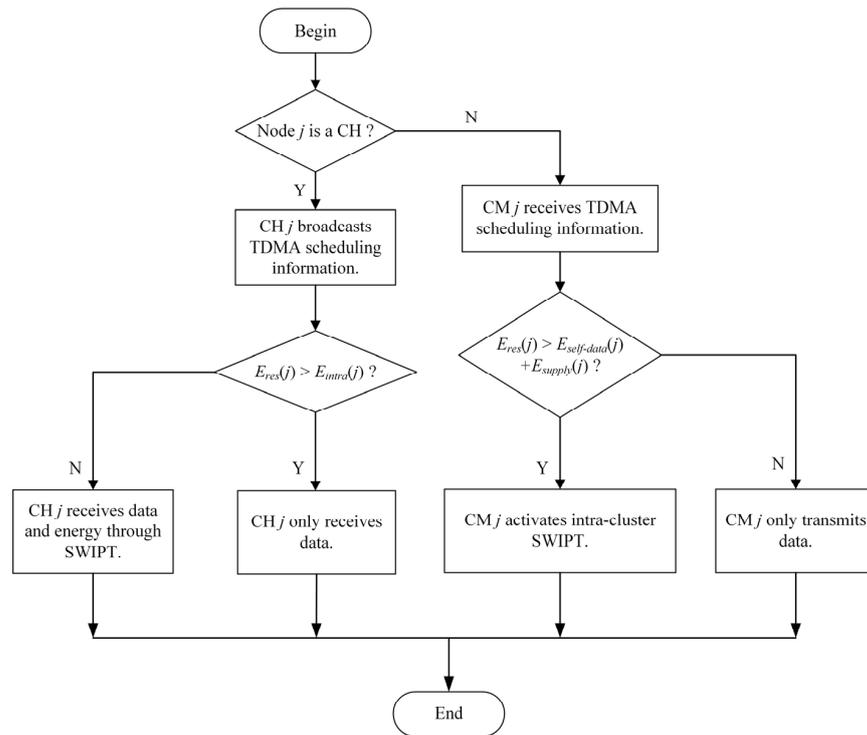


Figure 1. Flowchart of the intra-cluster data transmission process in ES-MUCRP.

Based on $E_{intra}(j)$, as shown in Equation (4), and the residual energy $E_{res}(j)$, it is determined whether to employ SWIPT technology. If $E_{res}(j)$ exceeds $E_{intra}(j)$, then all CMs are required only to transmit information; otherwise, each CM k must calculate the energy $E_{supply}(k)$ supplemented to its CH through SWIPT for intra-cluster data processing, as indicated in Equation (13).

$$E_{supply}(k) = \left(P_{thresh} + \frac{E_{intra}(j)}{n_{CM}(j)} \times \frac{1}{\eta} \right) \times d_{CM(k) \rightarrow CH}^2 \quad (13)$$

where P_{thresh} is the minimum threshold for SWIPT information transmission, set at 6.8 nW or -51.66 dBm; and $n_{CM}(j)$ is the number of CMs belonging to the cluster of CH j . CM k decides whether to perform intra-cluster SWIPT based on its remaining energy. If $E_{res}(k) - E_{self-data}(k)$ exceeds $E_{supply}(k)$, then CM k transfers energy $E_{supply}(k)$ to CH j while transmitting data using the SWIPT mechanism. In this case, the effective energy received by CH j is $E_{intra}(j)/n_{CM}(j)$. $E_{self-data}(k)$ is the energy consumption of CM k for transmitting its own data, as shown in Equation (14). If $E_{res}(k) - E_{self-data}(k)$ is less than or equal to $E_{supply}(k)$, then the CM only transmits data.

$$E_{self-data}(k) = \left(E_{elec} + E_{fs} \times d_{CM(k) \rightarrow CH}^2 \right) \times L_2 \quad (14)$$

2. Prior to inter-cluster data transmission, following the established routing path, node j assesses whether the residual energy $E_{res}(s)$ of the next hop s exceeds the energy required to forward its data $E_{relay}(s)$, as presented in Equation (15). If $E_{res}(s)$ is greater than $E_{relay}(s)$, node j transmits data directly to the next hop s , which relays the data packet. Otherwise, an assessment must be made according to Equation (16) to see if $E_{res}(j)$ surpasses the combined requirements of the intra-cluster data processing energy $E_{intra}(j)$, the inter-cluster supplemented energy $E_{S-supply}(j)$ for the next hop to relay data packets, and the dormancy threshold $E_{dormancy}(j)$. If it does, node j supplements the residual energy of the next hop while transmitting data using inter-cluster SWIPT.

$$E_{relay}(s) = \left(2E_{elec} + E_{fs} \times d_{CH(r(s)) \rightarrow CH(r(s)+1)}^2 \right) \times L_2, \quad (15)$$

where $d_{CH(r(s)) \rightarrow CH(r(s)+1)}$ represents the average relay distance from the CHs in the ring where relay s resides to their next hops.

$$E_{S-supply}(j) = \left(\frac{\left| E_{res}(s) - \left(2E_{elec} + E_{fs} \times d_{CH(r(s)) \rightarrow CH(r(s)+1)}^2 \right) \times L_2 \right|}{\eta} + P_{thresh} \right) \times d_{CH(j) \rightarrow s}^2 \quad (16)$$

3.3. Theoretical Derivation of Cluster Radius

This paper derives the relationship between the cluster radii of adjacent rings, aiming to equalize the net energy expenditure of CHs in adjacent rings. With the cluster radius of the outermost ring as the initial condition, the cluster radii of other rings are ascertained.

3.3.1. Derivation of the Relationship between Cluster Radii of Adjacent Rings

The average energy expenditure of a CH in ring i consists of three components: energy spent on control overhead, energy consumed for data processing within the cluster, and energy utilized for data relaying between clusters, as specifically shown in Equation (17).

$$E_{conCH(i)} = 3L_1 \times (E_{elec} + E_{fs} \times R_i^2) + 3(N_i - 1)E_{elec} \times L_1 + (N_i - 1)E_{elec} \times L_2 + N_i \times E_{DA} \times L_2 \\ + L_2 \times \left(E_{elec} + E_{fs} \times d_{CH(i) \rightarrow CH(i-1)}^2 \right) + \frac{\sum_{j=i+1}^z \frac{A_j}{S_j} \times (2E_{elec} + E_{fs} \times d_{CH(i) \rightarrow CH(i-1)}^2) \times L_2}{\frac{A_i}{S_i}} \quad (17)$$

The energy received by each CH in ring i through intra-cluster SWIPT is the sum of the energy transferred to it by its CMs based on their residual energy, as shown in Equation (18).

$$E_{harCH(i)} = (N_i - 1) \times E_{elec} \times L_2 + N_i \times E_{DA} \times L_2 + (E_{elec} + E_{fs} \times d_{CH(i) \rightarrow CH(i-1)}^2) \times L_2 \quad (18)$$

Similarly, the average energy consumption and the energy received through intra-cluster SWIPT for each CH in ring $i + 1$ per round can be deduced. Assuming uniform initial energy levels for all CRSN nodes, to ensure a balanced residual energy among CHs across different rings, it is necessary to calculate the energy variation for each CH per round, that is, the net energy consumption $E_{netCH(i)}$. From Equations (17) and (18), the $E_{netCH(i)}$ for an individual CH in ring i can be ascertained.

$$E_{netCH(i)} = 3(E_{elec} + E_{fs} \times R_i^2) \times L_1 + 3(N_i - 1) \times E_{elec} \times L_1 + \frac{\sum_{j=i+1}^z \frac{A_j}{S_j} L_2 (2E_{elec} + E_{fs} \times d_{CH(i) \rightarrow CH(i-1)}^2)}{\frac{A_i}{S_i}}. \quad (19)$$

Following a similar procedure as Equation (19), the net energy consumption of a single CH in ring $i + 1$ can be calculated. By equating it to Equation (19), the relationship between the cluster radii of ring i and its adjacent outer ring j ($j = i + 1$) can be determined.

When $j = z$, the relationship between the cluster radii of adjacent rings can be derived from equal net energy consumption of individual CHs in adjacent rings, as depicted in Equation (20):

$$R_{z-1} = \sqrt{(3L_1 m_1 R_z)^2 + 3L_1 m_1 m_2} / \left(3L_1 m_1 + \frac{m_2}{R_z^2}\right), \quad (20)$$

where

$$m_1 = E_{fs} + E_{elec}/50, \quad (21)$$

$$m_2 = \frac{(18E_{elec} \times L_2 + 4E_{fs} R_i^2)(2z - 1)}{9 \times (2z - 3)}. \quad (22)$$

When $j = z - 1$, the relationship between the cluster radii of adjacent rings is as indicated in Equation (23):

$$3L_1 m_1 (R_{z-2}^2 - R_j^2) + \left(\frac{2z-3}{R_j^2} + \frac{2z-1}{R_z^2}\right) m_3 \times L_2 \times \frac{R_{z-2}^2}{2i-1} - \left(\frac{2z-1}{R_z^2}\right) m_3 \times L_2 \times \frac{R_j^2}{2i+1} = 0, \quad (23)$$

where

$$m_3 = 2E_{elec} + 4E_{fs} R_i^2 / 9. \quad (24)$$

By inserting Equation (20) into Equation (23), the derived outcome is as stated in Equation (25).

$$R_{z-2} = \sqrt{\frac{\frac{3L_1 m_1 R_z^4 (3L_1 m_1 R_z^2 + m_2)}{2m_2^2 + 3L_1 m_1 m_2^2 + 3L_1 m_1 R_z^4}}{3L_1 m_1 + \left(\frac{(2i-1) \left(3L_1 m_1 + \frac{m_2}{R_z^2}\right)^2}{(3L_1 m_1 R_z)^2 + 3L_1 m_1 m_2} + \frac{2z-1}{R_z^2}\right) \times \frac{m_3 L_2}{2i-1} - \left(\frac{2z-1}{R_z^2}\right) \times \frac{m_3 L_2}{2i+1}}}} \quad (25)$$

Following the same logic, the relationship between the cluster radius of any given ring and that of the outermost ring can be established. Therefore, to determine the cluster radius value for each ring, the initial value R_z must be ascertained.

3.3.2. Determining the Cluster Radius of the Outermost Ring R_z

To minimize the collective energy consumption of the outermost ring and the internal rings acting as data relays, the cluster radius value of the outermost ring is calculated. The

total energy consumption for the outermost ring comprises the energy expenditures of CHs and CMs, as detailed in Equation (26) and Equation (27), respectively.

$$E_{tcon_CHs(z)} = E_{conCH(z)} \times N_{CH(z)} = \left[3L_1 \times (E_{elec} + E_{fs} \times R_z^2) + 3(N_z - 1)E_{elec} \times L_1 + (N_z - 1)E_{elec} \times L_2 + N_z \times E_{DA} \times L_2 + L_2 \times (E_{elec} + E_{fs} \times d_{CH(z) \rightarrow CH(z-1)}^2) \right] \times \frac{A_z}{S_z} \quad (26)$$

$$E_{tconCMs(z)} = E_{conCM(z)} \times N_{CM(z)} = \left[3L_1 \times (E_{elec} + E_{fs} \times R_z^2) + 2N_z E_{elec} \times L_1 + L_2 \times (E_{elec} + E_{fs} \times d_{CM \rightarrow CH(z)}^2) \right] \times (N_z - 1) \quad (27)$$

where $E_{conCH(z)}$ and $E_{conCM(z)}$ represent the average energy consumption of a single CH and a single CM per round in the outermost ring, respectively; $N_{CH(z)}$ and $N_{CM(z)}$ signify the total count of CHs and CMs in the outermost ring, respectively; and $d_{CM \rightarrow CH(z)}$ indicates the average Euclidean distance from CMs to their corresponding CH in the outermost ring.

The overall energy expenditure of the inner rings, which relay data packets from the outermost ring, is given in Equation (28):

$$E_{add(z)} = \sum_{j=0}^{z-1} \left[\frac{A_z}{S_z} \times L_2 \times (2E_{elec} + E_{fs} \times d_{CH(j) \rightarrow CH(j+1)}^2) \right], \quad (28)$$

where $CH(0)$ denotes the sink; and $d_{CH(0) \rightarrow CH(1)}^2$ represents the average squared distance from the sink to the CHs in the first ring.

In summary, the total energy consumption of the outmost ring and the inner rings for relaying data packets can be expressed as

$$E_{total(z)} = E_{tcon_CHs(z)} + E_{tconCMs(z)} + E_{add(z)} = aR_z^4 + bR_z^2 + c\frac{1}{R_z^2} + e, \quad (29)$$

where a , b , c , and e are positive constants, as delineated in Equations (30)–(33). By taking the first-order derivative of Equation (29) with regard to R_z^2 and setting the derivative to 0, the cluster radius of the outermost ring R_z is ascertained by extracting the square root of the derived result.

$$a = \frac{1}{R_t} \left[E_{fs} \times (3L_1 + L_2) + \frac{2}{R_t} \times L_1 \times E_{elec} \right] \quad (30)$$

$$b = \frac{E_{elec}}{R_t} \times (3L_1 + L_2) - \left[E_{fs} \times \left(3L_1 + \frac{L_2}{2} \right) + \frac{2}{R_t} \times L_1 \times E_{elec} \right] \quad (31)$$

$$c = \left[(2z - 1) \times L_2 \times R_t^2 \right] \times \left[\frac{4}{9}z \times E_{fs} \times R_t^2 + 2(z - 1) \times E_{elec} \right] \quad (32)$$

$$e = (2z - 1) \times R_t \times \left[3L_1 \times (E_{fs} \times R_t + E_{elec}) + L_2 \times (E_{DA} + E_{elec}) \right] - E_{elec}(3L_1 + L_2) \quad (33)$$

3.4. Complexity Analysis of ES-MUCRP

The time complexity of ES-MUCRP can be defined in two ways. The first one defines complexity in terms of control information exchange and data transmission: (1) During the CHs selection and cluster construction sub-stages, nodes in all rings of CRSNs, except the first ring, are required to exchange control messages three times. This exchange includes broadcasting node-specific information, CHs selection weights, and CHs broadcasting CHs announcement messages or normal nodes broadcasting withdrawal messages. The number of exchanged control messages is $3(K - N_{CH(1)})$. Here, K is the total number of CRSN nodes and $N_{CH(1)}$ denotes the number of CHs in ring 1. Additionally, normal nodes send join requests to their respective CHs and form clusters, resulting in an exchange of $K - N_{CH(1)}$

– $N_{CH(2)} - \dots - N_{CH(z-1)} - N_{CH(z)}$ control messages. Therefore, the total number of control messages exchanged during the CHs selection and cluster construction sub-stages is $3(K - N_{CH(1)}) + K - N_{CH(1)} - N_{CH(2)} - \dots - N_{CH(z-1)} - N_{CH(z)} = 4(K - N_{CH(1)}) - N_{CH(2)} - \dots - N_{CH(z-1)} - N_{CH(z)}$. (2) During the route selection sub-stage, first-ring CHs transmit state information to the sink to facilitate the selection of relay nodes by second-ring CHs. The number of exchanged control messages is $N_{CH(1)}$. Due to transmission range constraints, intermediate-ring CHs need to broadcast state information within the R_t range to assist outer-ring CHs in calculating competition values, resulting in an exchange of control messages totaling $N_{CH(2)} + \dots + N_{CH(z-1)}$. Non-first-ring CHs select suitable relay nodes based on the computed competition values and unicast route notification messages, resulting in an exchange of control messages totaling $N_{CH(2)} + \dots + N_{CH(z-1)} + N_{CH(z)}$. Consequently, the total number of control messages exchanged during the route selection sub-stage is $N_{CH(1)} + N_{CH(2)} + \dots + N_{CH(z-1)} + N_{CH(2)} + \dots + N_{CH(z-1)} + N_{CH(z)} = N_{CH(1)} + 2(N_{CH(2)} + \dots + N_{CH(z-1)}) + N_{CH(z)}$. (3) During the data transmission stage, each non-first-ring CH broadcasts TDMA scheduling information within the cluster, instructing its CMs to transmit data on the corresponding channels in their respective time slots. The number of exchanged control messages is $N_{CH(2)} + \dots + N_{CH(z-1)} + N_{CH(z)}$. Therefore, the total complexity of control message exchange for ES-MUCRP is $4(K - N_{CH(1)}) - N_{CH(2)} - \dots - N_{CH(z-1)} - N_{CH(z)} + N_{CH(1)} + 2(N_{CH(2)} + \dots + N_{CH(z-1)}) + N_{CH(z)} + N_{CH(2)} + \dots + N_{CH(z-1)} + N_{CH(z)} = 4(K - N_{CH(1)}) + N_{CH(1)} + 2(N_{CH(2)} + \dots + N_{CH(z-1)}) + N_{CH(z)}$, which is $O(K)$. Each surviving CRSN node j generates one data packet per round, and this packet is forwarded to the sink through a maximum of $2r(j) - 1$ ($r(j)$ is the ring in which it generates) hops. Consequently, the complexity of data transmission remains $O(K)$. The second one measures the complexity of ES-MUCRP based on execution time, which is defined as the maximum total delay introduced by spectrum sensing, CHs selection, cluster construction, route selection, intra-cluster data transmission, and multi-hop inter-cluster data relay. Although we cannot provide theoretical proof, it can be ascertained through simulation. Detailed simulation results are provided in Section 4.

4. Simulation Results and Discussion

This paper utilizes the MATLAB simulation tool to conduct a performance evaluation of ES-MUCRP protocol from the aspects: network lifetime, the balance degree of energy consumption within the network, and network surveillance capability. The effectiveness of ES-MUCRP protocol is validated through comparison with existing clustering routing protocols for CRSNs, such as CogLEACH [15], DSAC [16], NSAC [18], WCM [20], Fuzzy C-means [11], IMOCR [12], and S-MUCRP [27]. Since CogLEACH, DSAC, NSAC, WCM, Fuzzy C-means, and IMOCR are all non-EH clustering routing protocols, a linear EH mechanism is incorporated to ensure a fair comparison, allowing each surviving CRSN node to perform linear EH at the start of each round. CRSN nodes predict the energy that can be harvested from actively communicating PUs based on the signal strength received during the spectrum sensing stage. They then compare this with the RF energy that can be collected from the sink and choose to harvest energy for a duration of t_1 (0.2 s) from the source offering more RF energy, storing the collected energy in their on-board batteries. 450 CRSN nodes are randomly and uniformly deployed in a circular network monitoring area with a radius of 150 m, with the sink located at the center of the network. Given that the maximum transmission range R_t of CRSN nodes is 50 m, the entire network is divided into three rings, indicating that CRSN nodes belong to either ring 1, 2, or 3. The cluster radii for ring 2 and ring 3 are determined using the theoretical derivation method described in Section 3.3. Each surviving node in every round must transmit the monitoring data collected from the environment to the sink. The specific simulation parameters are provided in Table 2.

Table 2. Simulation parameter settings.

Parameters	Values
Network radius (R)	150 m
Total number of CRSN nodes (K)	450
Quantity of PUs (m)	50
Channel occupancy rate of PUs (p_c)	0.8
Number of authorized channels (C)	5
Length of control packets (L_1)	100 bits
Length of data packets (L_2)	1024 bits
Weight factor for adjusting the impact of energy state function (α)	10
Energy expended on data aggregation (E_{DA})	5 nJ/bit/packet
Energy consumption of electronic circuitry (E_{elec})	50 nJ/bit
Energy utilized for channel switching (E_{switch})	10 nJ
Power amplifier coefficient under free-space path loss model (E_{fs})	10 pJ/bit/m ²
Duration of EH (t_1)	0.2 s
Conversion efficiency of linear EH (η)	0.8
Maximum transmission range of CRSN nodes (R_t)	50 m
Interference protection range of PUs	20 m

When a node's residual energy drops to 0, it can no longer perform any operations and loses its network monitoring capability. Therefore, network lifetime is an important metric for assessing CRSNs clustering protocols. In this paper, the number of surviving nodes is used to measure the magnitude of network lifetime. The comparison results of the number of surviving nodes across various protocols are shown in Figure 2, with a detailed comparison shown in Table 3. From Figure 2 and Table 3, it is evident that the first node failure of ES-MUCRP occurs at round 989, which is significantly later than DSAC, WCM, NSAC, and S-MUCRP, but earlier than CogLEACH, Fuzzy C-means, and IMOCR. This indicates that CRSN nodes in ES-MUCRP consume less energy in control information exchange and data transmission. In order to explore the reasons behind this phenomenon, the total control overhead per round, total energy consumption per round, and execution time are recorded in Figures 3–5. A detailed analysis of the results is as follows:

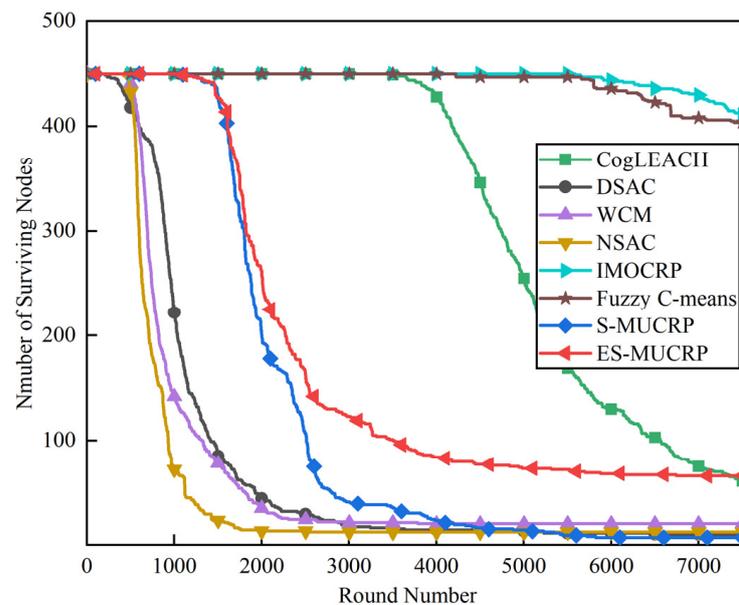
**Figure 2.** Comparison of the number of surviving nodes across various protocols.

Table 3. Specific comparison results of the number of surviving nodes.

Protocols	The Number of Rounds until the First Node Death
CogLEACH	3469
DSAC	203
WCM	478
NSAC	415
IMOCRCP	5588
Fuzzy C-means	4220
S-MUCRP	893
E-MUCRP	989

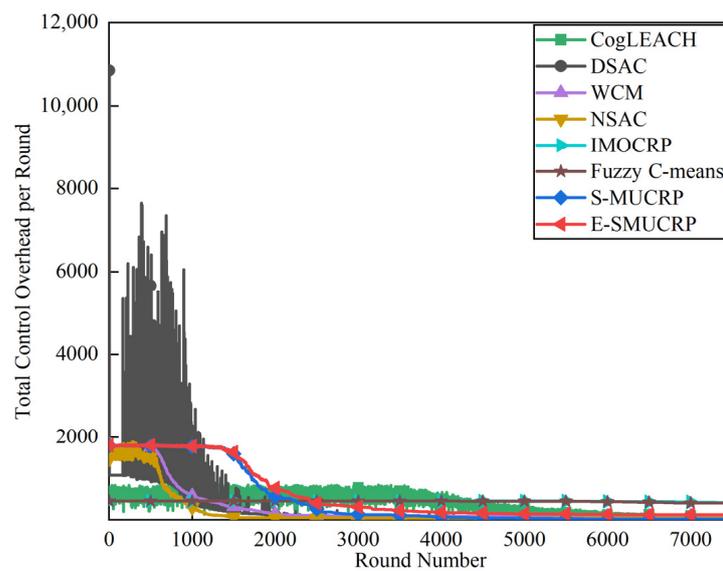


Figure 3. Comparison of the total control overhead per round across various protocols.

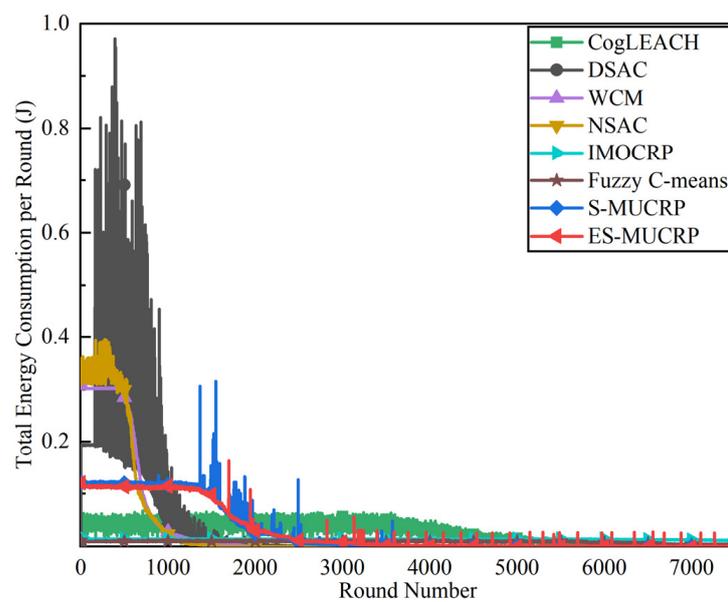


Figure 4. Comparison of the total energy consumption per round across various protocols.

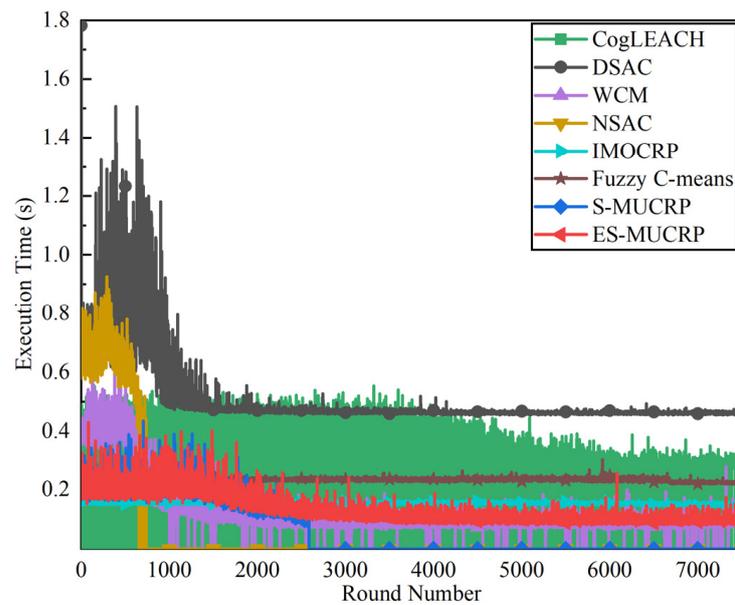


Figure 5. Comparison of the execution time across various protocols.

1. In ES-MUCRP protocol, the overall control overhead incurred during the CHs selection and cluster construction is more than triple but less than four times the count of non-first-ring active nodes. Specifically, CRSN nodes in the first ring become independent CHs directly, eliminating the need for broadcasting CHs selection weights and thus avoiding the associated control overhead. Furthermore, during the CHs selection process, CRSN nodes beyond the first ring broadcast information such as available channels, their locations, and CHs selection weights to their neighbors. Neighboring nodes determine their potential to become CHs based on the received information. Nodes that qualify as CHs then disseminate CHs announcement messages, prompting ordinary nodes that receive these announcements to broadcast their withdrawal from the CHs competition. During cluster formation, nodes not designated as CHs apply to join clusters by sending out join requests.
2. WCM protocol incurs a total control overhead for CHs selection and cluster formation that is roughly four times the count of active nodes: all CRSN nodes broadcast their spectrum sensing results and CHs weights on CCC to select CHs; neighboring nodes that receive this information decide if they are eligible to become CHs, with qualifying nodes broadcasting CHs announcements, and other nodes broadcasting their withdrawal from CHs competition; nodes not designated as CHs apply to join the CH with the greatest weight and shared available channels, after which the CH communicates the cluster details to the sink. In DSAC protocol, each CRSN node starts as a CH and merges with adjacent clusters based on the common available channels and cluster distances until they reach the optimal number derived from theoretical calculations. This involves substantial control information exchanges between CMs and CHs, as well as among neighboring CHs, resulting in considerable energy expenditure. In NSAC protocol, all CRSN nodes calculate their own weight based on the remaining energy and channel quality and continuously update and broadcast their weight information, and the node with the highest weight in the vicinity becomes a CH, with neighboring nodes joining to become CMs. Unclustered nodes repeat this until clustering is completed. This process necessitates extensive control information exchanges among adjacent nodes, leading to significant energy usage. As a result, the first node death occurs earlier in WCM, DSAC, and NSAC protocols than in ES-MUCRP protocol, with a sharp decline in the number of active nodes in subsequent rounds. Fuzzy C-means and IMOCRP represent centralized, single-hop clustering routing protocols for CRSNs. These protocols mandate that each

surviving CRSN node transmits information such as remaining energy to the sink. The sink is responsible for choosing CHs and communicating the clustering results to all CRSN nodes. Consequently, their total control overhead per round is equivalent to the count of surviving nodes. The first node death in CogLEACH happens later than that in ES-MUCRP. This is due to the fact that the control overhead for CHs selection and cluster formation in CogLEACH is roughly twice the number of active nodes, which is relatively modest. In CogLEACH, each CH broadcasts temporary and final CHs announcements, and nodes not designated as CHs send temporary join requests and final confirmation to their CH. Despite lower energy usage under CogLEACH, Fuzzy C-means, and IMOCR, restriction to single-hop communication with the sink substantially restricts the network scalability and monitoring capability.

3. According to the aforementioned analysis, the total control overhead per round of ES-MUCRP is close to that of WCM. However, as shown in Figure 4, its total energy consumption per round is much lower than that of WCM, NSAC, and DSAC. This is because nodes under ES-MUCRP exchange control information within the cluster radius during CHs selection and cluster construction stage, while nodes in other competing protocols exchange information within R_t . Since the cluster radius is smaller than R_t , nodes consume less energy in CHs selection and cluster construction. Moreover, in order to make full use of the direct communication between CHs in ring 1 and the sink, in ES-MUCRP, CHs in ring 1 send their state messages directly to the sink in route selection stage. The sink receives, aggregates, and broadcasts the message, which can reduce the number of control messages received by CHs in ring 2 and the energy consumption of competing for accessing CCC in ring 1. Thus, node energy is saved, and the network lifetime is prolonged.
4. From Figure 5, we can observe that the execution time of DSAC is the longest among all competing protocols. This results from its excessive control information exchange for cluster merging and multi-hop route selection. More competing nodes within R_t increases the time required for successful channel access. ES-MUCRP can reduce the time required for control information exchange by controlling the cluster radius, but more effective data gathering nodes will inevitably increase the execution time of data transmission. This is the price to pay for guaranteeing powerful network surveillance capability. Nonetheless, data packets in ES-MUCRP can still reach the sink within the round time.
5. Additionally, from Figure 2 and Table 3, it is observable that the first node failure in S-MUCRP occurs at round 893. Beyond round 2588, ES-MUCRP protocol maintains a notably higher number of surviving nodes compared to S-MUCRP protocol. Upon calculation, it is found that compared to S-MUCRP protocol, ES-MUCRP protocol exhibits a 37.02% increase in the number of surviving nodes. A detailed analysis of the specific reasons is as follows. By integrating a linear EH mechanism, ES-MUCRP protocol enables CRSN nodes to identify the optimal RF energy source from either the sink or a PU occupying a channel based on the node's location information before cluster formation and route establishment in each round. This RF EH supplements the node's residual energy, delaying node depletion. The energy-intensive activities reduce the residual energy of some nodes to below the dormancy threshold in the network employing ES-MUCRP protocol. The activation of the energy status control mechanism transitions these nodes from an active state to a sleep state, wherein they only engage in EH, thus preventing early node death due to substantial energy depletion from data transfers. Due to the introduction of the inter-cluster SWIPT mechanism, ES-MUCRP protocol allows nodes requiring relay to assess the necessity of energy replenishment through inter-cluster SWIPT, taking into account their own energy situation and that of the relay nodes. After round 2588, some relay nodes experience a sharp decline in energy due to extensive data relaying and forwarding. The activation of the inter-cluster SWIPT mechanism prevents the early death of relay

nodes in critical positions and ensures a more uniform distribution of residual energy among the inter-cluster routing nodes.

To demonstrate the effectiveness of the intra-cluster and inter-cluster SWIPT mechanisms, as well as the energy status management mechanism of ES-MUCRP, this paper introduces the metric of average network energy consumption variance to measure the balance degree of energy consumption among nodes. Simulation results are displayed in Figure 6. Specifically, first, the average network energy consumption $aveE(r)$ for all nodes in the current round r is calculated according to Equation (34).

$$aveE(r) = \frac{\sum_{s=1}^r TotalE(s)}{r}, \quad (34)$$

where $TotalE(s)$ is the sum of the differences between the initial energy and the residual energy at the end of round s for all nodes, which is the cumulative network energy consumption of CRSNs as formulated in Equation (35).

$$TotalE(s) = \sum_{j=1}^K (E_{initial}(j) - E_{res}(j)) \quad (35)$$

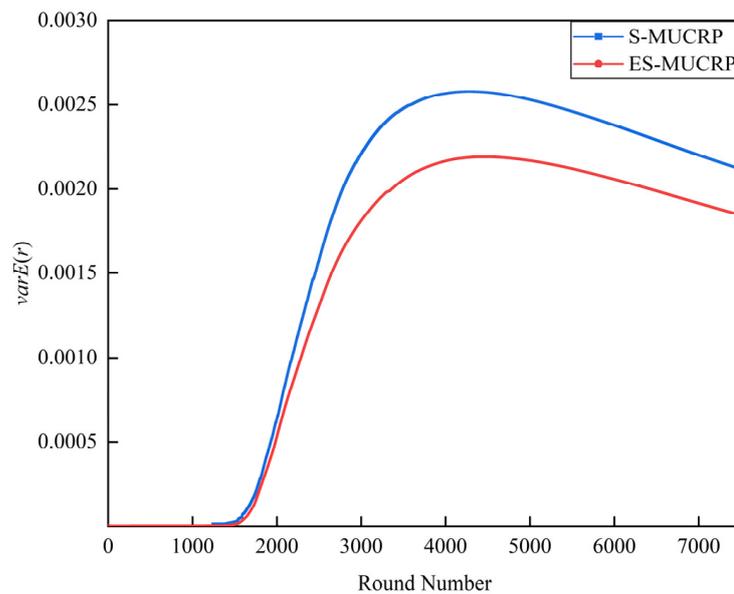


Figure 6. Comparison of average network energy consumption variance for ES-MUCRP and S-MUCRP protocols.

Subsequently, the average network energy consumption variance $varE(r)$ for all nodes in round r is computed, as specified in Equation (36).

$$varE(r) = \frac{\sum_{s=1}^r \sum_{t=1}^s (TotalE(t) - aveE(s))^2}{r} \quad (36)$$

Figure 6 indicates that beyond round 2588, the average network energy consumption variance of ES-MUCRP protocol is notably lower than S-MUCRP protocol, suggesting that ES-MUCRP protocol achieves a more uniform distribution of network energy consumption across nodes. Calculations show that compared to S-MUCRP protocol, the balance degree of energy consumption in ES-MUCRP protocol has improved by 17.69%.

Besides network lifetime, the number of CRSN nodes that effectively gather data serves as a crucial indicator of clustering protocol performance, as it reflects the network

monitoring capabilities of each protocol. This paper records the number of nodes effectively collecting data under each protocol, as illustrated in Figure 7.

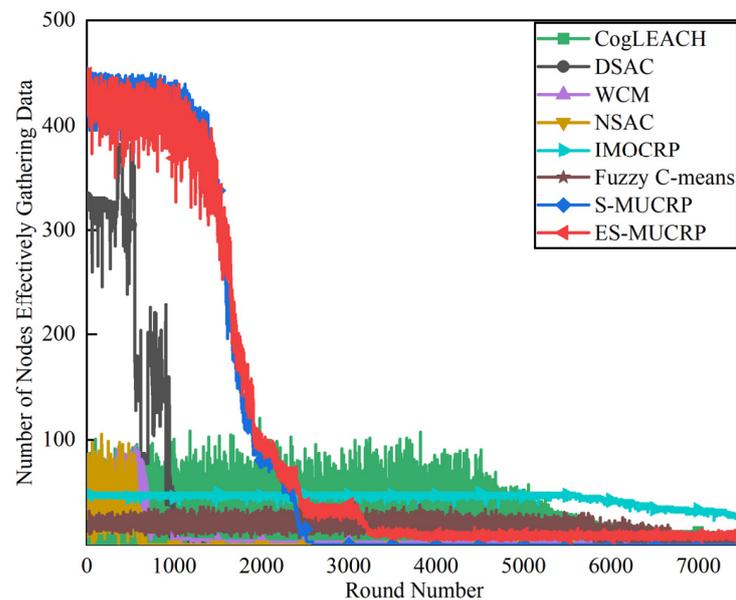


Figure 7. Comparison of the number of nodes effectively gathering data across various protocols.

Figure 7 reveals that the number of nodes effectively collecting data in ES-MUCRP protocol is significantly higher than that of CogLEACH, DSAC, WCM, NSAC, Fuzzy C-means, and IMOCRP protocols, and the reasons are given below. As a multi-hop clustering routing protocol, DSAC initially enables a larger count of nodes to route data to the sink through multiple hops, resulting in a higher number of nodes effectively collecting data. However, substantial control overhead leads to rapid energy depletion of nodes, with a marked decline in the number of surviving nodes and nodes effectively collecting data after round 618. CogLEACH, NSAC, Fuzzy C-means, and IMOCRP protocols function as single-hop CRSNs clustering routing protocols, limiting effective data collection to nodes that can reach the sink in a single hop. Furthermore, the number of nodes effectively collecting data in WCM and NSAC protocols drastically decreases along with the swift reduction in the number of surviving nodes. While building clusters, Fuzzy C-means fails to consider channel availability. Random channel selection may lead to the absence of a common channel between CMs and their CH, severely impacting the successful delivery of data packets. Therefore, the number of nodes effectively collecting data in Fuzzy C-means is low. ES-MUCRP protocol is a multi-hop clustering routing protocol that enables data to be transmitted to the sink through multi-hop routing. It also selects relatively stable channels with the highest count of CMs for intra-cluster and inter-cluster data transmission, leading to infrequent channel reclaim by PUs. Additionally, the number of nodes effectively collecting data in ES-MUCRP protocol is higher than that in S-MUCRP protocol. In an effort to quantitatively assess the surveillance capabilities of the clustering routing protocols, the average number of nodes effectively gathering data for each protocol has been computed, and the results are illustrated in Figure 8. It is evident from Figure 8 that the gap in the average number of nodes effectively collecting data between ES-MUCRP and S-MUCRP protocols widens progressively after round 2588. Calculations reveal a 6.77% increase in the average number of nodes effectively collecting data for ES-MUCRP protocol compared to S-MUCRP protocol. According to Table 4, the duration of effective data collection under ES-MUCRP protocol is significantly extended, with an additional 5281 rounds over S-MUCRP protocol, primarily for the following reasons:

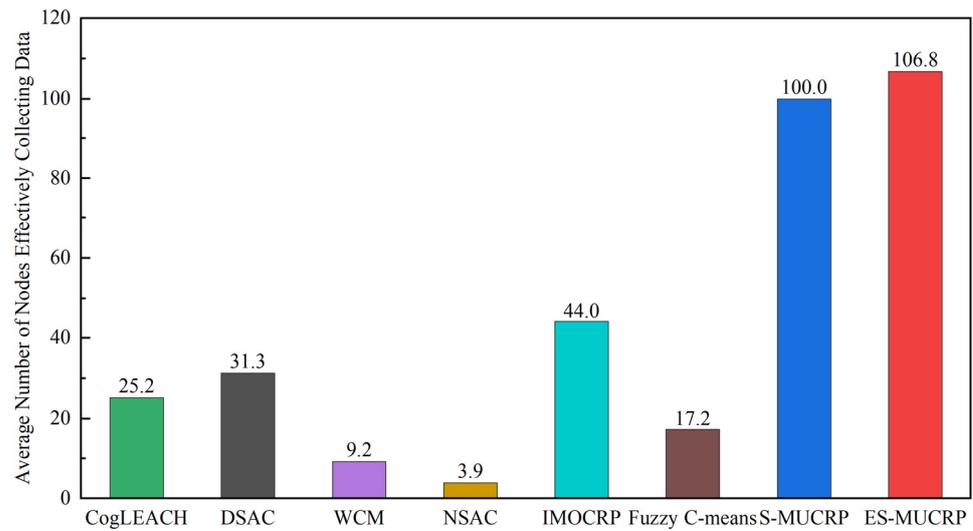


Figure 8. Comparison of the average number of nodes effectively collecting data.

Table 4. Specific comparison results of the duration of effective data collection.

Protocols	The Duration of Effective Data Collection
CogLEACH	7500
DSAC	1097
WCM	7500
NSAC	627
IMOCRIP	7500
Fuzzy C-means	7500
S-MUCRP	2588
ES-MUCRP	7500

1. The EH mechanism effectively replenishes the nodes’ residual energy, decelerating the death rate of the nodes.
2. After round 2588, the residual energy of some nodes and key relay nodes drops significantly. The energy status control mechanism and the inter-cluster SWIPT mechanism within ES-MUCRP protocol are activated, which equalize energy consumption across the network and effectively avoid premature deaths of nodes, thereby notably prolonging the duration of network monitoring.

The proposed ES-MUCRP protocol can be applied in EH-CRSNs to supplement and balance the remaining energy among nodes, thereby enhancing network surveillance capabilities. For instance, it can be used in the industrial IoT mentioned in the Introduction Section, assisting in the collection and transmission of various environmental data to achieve industrial automation and intelligence.

Its limitations can be summarized from the following two aspects:

1. ES-MUCRP protocol is proposed on the assumption of perfect spectrum sensing in this paper, disregarding the potential for sensing errors. Although this assumption simplifies the design of the protocol, it might not align with the actual detection capabilities of CRSN nodes.
2. In EH-CRSNs, communication is restricted to direct links between transmitters and receivers, where substantial path loss due to large Euclidean distance can deplete the limited battery energy of nodes. Although ES-MUCRP can slow down the rate of energy depletion in nodes, it cannot inherently resolve the limitations imposed on network lifespan by finite battery capacity.

5. Conclusions and Future Work

In this study, we have designed a distributed multi-hop clustering routing protocol ES-MUCRP for RF EH-CRSNs. The protocol incorporates downlink RF EH, intra-cluster and inter-cluster SWIPT technologies, and energy status control mechanisms to extend network lifespan, improve the overall balance of network energy consumption among nodes, and enhance effective network monitoring capabilities. Simulation results reveal that ES-MUCRP significantly surpasses existing clustering routing protocols for CRSNs in terms of surviving node count and the number of nodes effectively collecting data. In particular, compared to S-MUCRP protocol, ES-MUCRP protocol improves the balance degree in network energy consumption among nodes by 17.69% and enhances the average number of effective data-collecting nodes by 6.77%.

Focusing on the limitations of ES-MUCRP, we plan to investigate the design of CRSNs clustering routing protocols based on imperfect spectrum sensing in our future work. Additionally, communication is restricted to direct links between transmitters and receivers in EH-CRSNs, where substantial path loss due to large Euclidean distance can deplete the limited battery energy of nodes. Therefore, we intend to introduce an intelligent reflecting surface (IRS) to create additional paths for EH or information transmission, further enhancing the energy-saving effects of clustering routing protocols for EH-CRSNs. This will involve channel modeling, optimal placement of IRS, and the design of detailed clustering protocols. ES-MUCRP protocol addresses the issue of ensuring powerful data collection capabilities over an extended lifetime of CRSN nodes. It provides foundational data for solving predictive maintenance issues. Therefore, we will further study the key aspects of predictive maintenance issues, including effective analysis of the collected data, fault diagnosis, prediction, early warning, and maintenance decision-making, and eventually solve the issues.

Author Contributions: Conceptualization, J.W.; methodology, J.W. and Z.W.; software, Z.W. and L.Z.; validation, Z.W. and L.Z.; writing—original draft preparation, J.W. and Z.W.; supervision, J.W.; funding acquisition, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 61901102.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lilhore, U.K.; Dalal, S.; Simaiya, S. A Cognitive Security Framework for Detecting Intrusions in IoT and 5G Utilizing Deep Learning. *Comput. Secur.* **2024**, *136*, 103560. [[CrossRef](#)]
2. Dalal, S.; Lilhore, U.K.; Faujdar, N.; Simaiya, S.; Ayadi, M.; Almujaally, N.A.; Ksibi, A. Next-generation Cyber Attack Prediction for IoT Systems: Leveraging Multi-class SVM and Optimized CHAID Decision Tree. *J. Cloud Comp.* **2023**, *12*, 137. [[CrossRef](#)]
3. Zheng, M.; Wang, C.Q.; Song, M.; Liang, W.; Yu, H.B. SACR: A Stability-aware Cluster-based Routing Protocol for Cognitive Radio Sensor Networks. *IEEE Sens. J.* **2021**, *21*, 17350–17359. [[CrossRef](#)]
4. Kumar, D.; Singya, P.K.; Choi, K.; Bhatia, V. SWIPT Enabled Cooperative Cognitive Radio Sensor Network with Non-linear Power Amplifier. *IEEE Trans. Cogn. Commun. Netw.* **2023**, *9*, 884–896. [[CrossRef](#)]
5. Talukdar, B.; Kumar, D.; Arif, W. Performance Analysis of a SWIPT Enabled Cognitive Radio Sensor Network Using TS Protocol. In Proceedings of the 2020 Advanced Communication Technologies and Signal Processing (ACTS), Silchar, India, 4–6 December 2020; pp. 1–5. [[CrossRef](#)]
6. Wang, J.H.; Ge, Y.Y. A Radio Frequency Energy Harvesting-based Multihop Clustering Routing Protocol for Cognitive Radio Sensor Networks. *IEEE Sens. J.* **2022**, *22*, 7142–7156. [[CrossRef](#)]
7. Cao, C.F.; Zhang, X.M.; Song, C.Y.; Georgiadis, A.; Goussetis, G. A Highly Integrated Multipolarization Wideband Rectenna for Simultaneous Wireless Information and Power Transfer (SWIPT). *IEEE Trans. Antennas Propag.* **2023**, *71*, 7980–7991. [[CrossRef](#)]

8. Shahraki, A.; Taherkordi, A.; Haugen, O.; Eliassen, F. A Survey and Future Directions on Clustering: From WSNs to IoT and Modern Networking Paradigms. *IEEE Trans. Netw. Serv. Manag.* **2021**, *18*, 2242–2274. [[CrossRef](#)]
9. Afsar, M.M.; Younis, M. A Load-balanced Cross-layer Design for Energy-harvesting Sensor Networks. *J. Netw. Comput. Appl.* **2019**, *145*, 102390. [[CrossRef](#)]
10. Yuan, Y.F.; Liu, M.Y.; Zhuo, X.X.; Wei, Y.; Tu, X.B.; Qu, F.Z. A Q-learning-based Hierarchical Routing Protocol with Unequal Clustering for Underwater Acoustic Sensor Networks. *IEEE Sens. J.* **2023**, *23*, 6312–6325. [[CrossRef](#)]
11. Bhatti, D.M.S.; Saeed, N.; Nam, H. Fuzzy C-means Clustering and Energy Efficient Cluster Head Selection for Cooperative Sensor Network. *Sensors* **2016**, *16*, 1459. [[CrossRef](#)]
12. Wang, J.H.; Li, S.; Ge, Y.Y. Ions Motion Optimization-based Clustering Routing Protocol for Cognitive Radio Sensor Network. *IEEE Access* **2020**, *8*, 187766–187782. [[CrossRef](#)]
13. Latiwesh, A.; Qiu, D.Y. Energy Efficient Spectrum Aware Clustering for Cognitive Sensor Networks: CogLEACH-C. In Proceedings of the 10th International Conference on Communications and Networking in China (ChinaCom), Shanghai, China, 15–17 August 2015; pp. 515–520. [[CrossRef](#)]
14. Kim, S.S.; McLoone, S.; Byeon, J.H.; Lee, S.; Liu, H.B. Cognitively Inspired Artificial Bee Colony Clustering for Cognitive Wireless Sensor Networks. *Cognit. Comput.* **2017**, *9*, 207–224. [[CrossRef](#)]
15. Eletreby, R.M.; Elsayed, H.M.; Khairy, M.M. CogLEACH: A Spectrum Aware Clustering Protocol for Cognitive Radio Sensor Networks. In Proceedings of the 9th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), Oulu, Finland, 2–4 June 2014; pp. 179–184. [[CrossRef](#)]
16. Zhang, H.Z.; Zhang, Z.Y.; Dai, H.Y.; Yin, R.; Chen, X.M. Distributed Spectrum-aware Clustering in Cognitive Radio Sensor Networks. In Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM), Houston, TX, USA, 5–9 December 2011; pp. 1–6. [[CrossRef](#)]
17. Shah, G.A.; Alagoz, F.; Fadel, E.A.; Akan, O.B. A Spectrum-aware Clustering for Efficient Multimedia Routing in Cognitive Radio Sensor Networks. *IEEE Trans. Veh. Technol.* **2014**, *63*, 3369–3380. [[CrossRef](#)]
18. Zheng, M.; Chen, S.; Liang, W.; Song, M. NSAC: A Novel Clustering Protocol in Cognitive Radio Sensor Networks for Internet of Things. *IEEE Internet Things J.* **2019**, *6*, 5864–5865. [[CrossRef](#)]
19. Yadav, R.N.; Misra, R.; Saini, D. Energy Aware Cluster Based Routing Protocol over Distributed Cognitive Radio Sensor Network. *Comput. Commun.* **2018**, *129*, 54–66. [[CrossRef](#)]
20. Wang, T.J.; Guan, X.J.; Wan, X.L.; Shen, H.; Zhu, X.M. A Spectrum-aware Clustering Algorithm Based on Weighted Clustering Metric in Cognitive Radio Sensor Networks. *IEEE Access* **2019**, *7*, 109555–109565. [[CrossRef](#)]
21. Mustapha, I.; Ali, B.M.; Rasid, M.F.A.; Sali, A.; Mohamad, H. An Energy-efficient Spectrum-aware Reinforcement Learning-based Clustering Algorithm for Cognitive Radio Sensor Networks. *Sensors* **2015**, *15*, 19783–19818. [[CrossRef](#)]
22. Pei, E.R.; Han, H.Z.; Sun, Z.H.; Shen, B.; Zhang, T.Q. LEAUCH: Low-energy Adaptive Uneven Clustering Hierarchy for Cognitive Radio Sensor Network. *Eurasip J. Wirel. Commun. Netw.* **2015**, *2015*, 122. [[CrossRef](#)]
23. Stephan, T.; Joseph, K.S. Particle Swarm Optimization-based Energy Efficient Channel Assignment Technique for Clustered Cognitive Radio Sensor Networks. *Comput. J.* **2018**, *61*, 926–936. [[CrossRef](#)]
24. Wang, C.Q.; Wang, S.B. Research on Uneven Clustering APTEEN in CWSN Based on Ant Colony Algorithm. *IEEE Access* **2019**, *7*, 163654–163664. [[CrossRef](#)]
25. Stephan, T.; Al-Turjman, F.; Suresh, J.K.; Balusamy, B. Energy and Spectrum Aware Unequal Clustering with Deep Learning Based Primary User Classification in Cognitive Radio Sensor Networks. *Int. J. Mach. Learn. Cybern.* **2020**, *12*, 3261–3294. [[CrossRef](#)]
26. Liu, W.W.; Zhu, Y.H.; Pan, J. A Grid-based Routing Algorithm with Cross-level Transmission to Prolong Lifetime of Wireless Sensor Networks. *Chin. J. Electron.* **2010**, *19*, 499–502.
27. Wang, J.H.; Zhang, L.D. A Simultaneous Wireless Information and Power Transfer-based Multihop Uneven Clustering Routing Protocol for Cognitive Radio Sensor Networks. *J. Netw. Intell.* **2023**, *8*, 1134–1149.
28. Seo, D.; Varshney, L.R. Information and Energy Transmission with Experimentally Sampled Harvesting Functions. *IEEE Tran. Commun.* **2019**, *67*, 4479–4490. [[CrossRef](#)]
29. Xiao, M.; Zhang, X.D.; Dong, Y.H. An Effective Routing Protocol for Energy Harvesting Wireless Sensor Networks. In Proceedings of the 2013 IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, China, 7–10 April 2013; pp. 2080–2084. [[CrossRef](#)]
30. Yang, L.; Lu, Y.Z.; Zhong, Y.C.; Wu, X.G.; Yang, S.X. A Multi-hop Energy Neutral Clustering Algorithm for Maximizing Network Information Gathering in Energy Harvesting Wireless Sensor Networks. *Sensors* **2016**, *16*, 26. [[CrossRef](#)] [[PubMed](#)]
31. Meng, J.; Zhang, X.D.; Dong, Y.H.; Lin, X.K. Adaptive Energy-harvesting Aware Clustering Routing Protocol for Wireless Sensor Networks. In Proceedings of the 7th International Conference on Communications and Networking in China (ChinaCom), Kunming, China, 8–10 August 2012; pp. 742–747. [[CrossRef](#)]
32. Liu, Z.X.; Zhao, M.Y.; Yuan, Y.Z.; Guan, X.P. Subchannel and Resource Allocation in Cognitive Radio Sensor Network with Wireless Energy Harvesting. *Comput. Netw.* **2020**, *167*, 107028. [[CrossRef](#)]
33. Aslam, S.; Ibnkahla, M. Optimized Node Classification and Channel Pairing Scheme for RF Energy Harvesting Based Cognitive Radio Sensor Networks. In Proceedings of the 2015 IEEE 12th International Multi-Conference on Systems, Signals & Devices (SSD15), Mahdia, Tunisia, 16–19 March 2015; pp. 1–6. [[CrossRef](#)]

34. Zhang, M.C.; Zheng, R.J.; Li, Y.; Wu, Q.T.; Song, L. R-bUCRP: A Novel Reputation-based Uneven Clustering Routing Protocol for Cognitive Wireless Sensor Networks. *J. Sens.* **2016**, *2016*, 5986265. [[CrossRef](#)]
35. Heinzelman, W.B.; Chandrakasan, A.P.; Balakrishnan, H. An Application-specific Protocol Architecture for Wireless Microsensor Networks. *IEEE Trans. Wirel. Commun.* **2002**, *1*, 660–670. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.