IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 27, 2020, accepted October 4, 2020, date of publication October 13, 2020, date of current version October 26, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3030808

Ions Motion Optimization-Based Clustering Routing Protocol for Cognitive Radio Sensor Network

JIHONG WANG, (Member, IEEE), SHUO LI, AND YIYANG GE
School of Electrical Engineering, Northeast Electric Power University, Jilin City 132012, China

Corresponding author: Jihong Wang (wangjihong07@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61901102.

ABSTRACT Cognitive radio sensor network (CRSN) is an intelligent and reasonable combination of cognitive radio and wireless sensor networks. Clustering is an effective method to manage the network topology of CRSN. In order to solve the optimal cluster head selection problem in clustering which is proved to be NP-hard, inspired by ions motion optimization (IMO) algorithm, a novel centralized clustering routing protocol, i.e., IMO-based clustering routing protocol (IMOCRP) which can adapt to the dynamic characteristics of CRSN is presented in this paper. Optimal number of clusters and identity of cluster heads are automatically determined by combining the excellent characteristics of IMO algorithm and the rich resource, comprehensive information at the sink together. Thanks to its centralized structure, IMOCRP proposed in this paper avoids excessive overhead in cluster head selection and cluster formation, which in turn helps conserve energy. Available channel list at each living node is taken into consideration to perform reasonable channel allocation for clusters, and the probability of collisions with primary users can be reduced, which promotes more successful information delivery. Simulation results have shown that IMOCRP can balance network lifetime and effective information collection capability, and it is superior to other competing protocols.

INDEX TERMS Cognitive radio sensor network, clustering algorithms, ions motion optimization, computational complexity.

I. INTRODUCTION

Cognitive radio sensor network (CRSN) is an intelligent and reasonable combination of cognitive radio (CR) and wireless sensor networks (WSNs) [1]. With the capability of spectrum sensing, CRSN nodes can opportunistically access the licensed spectrum bands when they are not occupied by primary users (PUs), thereby alleviating the spectrum scarcity faced by traditional WSNs and enhancing network performance [2].

Clustering routing protocol is a type of energy-efficient routing protocol in CRSN. In clustering, data aggregation at cluster head (CH) reduces the number of data transmissions throughout the whole network, and fewer data transmissions mean lower energy consumption and smaller probability of collisions with PUs. Additionally, cooperative spectrum sensing among nodes within a cluster can also help decrease the probability of collisions with PUs due to miss detection in spectrum sensing [3], [4]. In view of its importance in improving network performance, clustering routing protocol design for CRSN has become a hot research area which is wildly concerned by scholars [5]. CH selection, optimal number of clusters and reasonable channel allocation are three important aspects in clustering routing protocol design for CRSN.

- CH selection: CHs are responsible for data collection, data aggregation, data transmission within the cluster, and even data relay for other clusters. Therefore, its energy consumption is usually higher than normal nodes. Residual energy and number of available channels should be considered when selecting CHs [6].
- Optimal number of clusters: Factors such as node deaths (caused by running out of energy) or transitions in PUs' channel occupancy state will lead to network topology changes in CRSN, therefore, the optimal number of

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

clusters should adapt to the dynamic characteristics of CRSN [7], 8].

- Reasonable channel allocation: Channel allocation should satisfy intra-cluster communication requirements and minimize the potential impact on PUs [9].

Inspired by ions motion optimization (IMO) algorithm [10], a novel centralized clustering routing protocol, i.e., IMO-based clustering routing protocol (IMOCRP) for CRSN is proposed in this paper. IMO is chosen as the basis for clustering routing protocol design due to the following reasons: IMO has fewer input parameters which makes it simple to implement, and it has powerful exploring and exploiting capability than many other swarm intelligence-inspired optimization algorithms [11]; The automatic adjustment of ion positions by gravitation is consistent with the properties of IMOCRP, which makes it very suitable for solving clustering problem in CRSN [12]. IMOCRP makes full use of the advantages of IMO, and the optimal number of clusters and CHs are automatically determined to adapt to the dynamic characteristic of CRSN. Simulation experiments show that IMOCRP can extend the lifetime of CRSN and guarantee the quality of network monitoring simultaneously.

The technical contributions and novelty of our work are summarized as follows:

- By combining the excellent characteristics of IMO algorithm and the rich resource, comprehensive information at the sink together, IMOCRP can automatically determine the optimal number of CHs and corresponding clustering results while minimizing total control overhead incurred during clustering process. Therefore, IMOCRP is energy efficient and adaptable to dynamic topology changes.
- Available channel list at each living node is taken into consideration to perform reasonable channel allocation for clusters, and the probability of collisions with PUs can be reduced, which promotes more successful information delivery. Therefore, the effective monitoring capability required by CRSN applications can be guaranteed.
- The computational complexity of IMOCRP is analyzed and its performance is confirmed through extensive simulations. The simulation results have shown that IMOCRP can balance network lifetime and effective information collection capability, and it is superior to other competing protocols.

The rest of this paper is organized as follows. Section II summarizes existing works related to clustering routing protocols for CRSN and swarm intelligence-inspired algorithms for clustering in WSNs. The network model and assumptions are given in Section III. Section IV briefly explains the principles of original IMO algorithm which are the basis for IMOCRP design. We explain the proposed IMOCRP in details in terms of its frame structure, optimization target, the mapping relationship between ions and clustering results, the liquid and crystal iteration process and typical examples in Section V. Section VI is the performance evaluation and time complexity analysis of IMOCRP. Section VII concludes the paper and indicates our future work.

II. RELATED WORK

This paper focuses on developing new clustering routing protocol IMOCRP for CRSN, and it utilizes swarm intelligence-inspired algorithm IMO as its basis. Therefore, in this section, we will survey related works from the following two aspects: swarm intelligence-inspired algorithms for clustering in WSNs and current clustering routing protocols for CRSN.

A. SWARM INTELLIGENCE-INSPIRED ALGORITHMS FOR CLUSTERING IN WSNs

Clustering has been a hot research topic in WSNs, and many swarm intelligence-inspired algorithms are applied to solve this problem. Here we only list some recent works in this area. In [13], a new version of gravitational search algorithm (GSA) is utilized to solve the energy-efficient clustering problem in WSNs. Fitness function which is composed of efficient energy consumption and link quality is proposed to evaluate the clustering results produced. By utilizing the power distance sums scaling method to calculate the mass values and fuzzy logic controller to identify corresponding parameters, GSA finds the optimal number of clusters and properly organizes these clusters. EECP based on genetic algorithm [14] is proposed to perform energy-efficient CH selection and balance energy consumption load among routes. CHs are picked based on an enhanced search equation to improve exploitation competences and convergence rate. A quantum particle swarm optimization (PSO) algorithm [15] is proposed to improve network lifetime of WSNs. It is composed of 3 steps including position updating, CH selection and cluster formation. After position updating, the fitness function which is dependent on average distance among clusters, average distance between CH and the sink, and total energy of all selected CHs is used to find the best particle. CH selection is done according to the best particle, and then clusters are formed. Multi-objective Taylor crow optimization algorithm is proposed to perform the optimal CH selection in [16]. The optimal CHs are determined by minimizing its objective function which is the weighted combination of factors such as distance between nodes in a cluster, energy of nodes, traffic density of cluster and delay in transmitting data packets. In [17], a hybrid method based on firefly algorithm and PSO is proposed to find the optimal CH selection in LEACH-C, in which PSO is exploited to help improve the global search behavior of fireflies. In [18], a distributed swarm artificial bee colony algorithm is proposed to optimize the dynamics of CHs and normal nodes in WSNs. It can minimize the energy dissipation of nodes and balance the interference-aware energy consumption of the network. However, above works are specially designed for WSNs, and they consider nothing about CR and available channels, therefore, they cannot be applied to CRSN without any modification.

B. CURRENT CLUSTERING ROUTING PROTOCOLS FOR CRSN

CogLEACH [19] is a distributed and spectrum-aware extension of legacy LEACH protocol. It regards number of available channels as the weight function to be CH. Nodes with larger number of idle channels have higher probability to become CH. Ordinary nodes select the nearest CH which shares common available channels with them to form clusters. However, the number of CHs selected in CogLEACH may be larger than the optimal value, and this will result in unnecessary energy waste. In order to guarantee the optimal number of CHs, centralized CogLEACH-C [20] is proposed on the basis of CogLEACH. The sink selects the optimal CHs from all living nodes according to the number of vacant channels, residual energy and geographical location of nodes. The cluster formation of CogLEACH-C is the same as CogLEACH. Fuzzy C-means algorithm [21] divides the whole network into K clusters by minimizing the sum of the square of distances from each cluster member (CM) to the cluster center. In each cluster, CH is selected based on four parameters, that is, location within each cluster, location with respect to the sink, signal-to-noise ratio of channels and residual energy. Nodes acquire spectrum information through cooperative sensing to avoid the influence of shadowing and fading on sensing accuracy. In case of multihop communications, CHs are not only responsible for gathering intra-cluster data, but also relaying data for other clusters. Therefore, compared with CHs far away from the sink, CHs near the sink consume more energy. LEAUCH [22] employs uneven clustering method to reduce the size of clusters close to the sink, thereby balancing energy consumption among CHs. Among all candidate CHs within competition radius, the one with the highest residual energy becomes final CH. Here, competition radius concept is introduced to form uneven clusters, and it becomes smaller as distance to the sink decreases. IACUCAPTEEN [23] is also an uneven clustering routing protocol which uses probability influence factors such as remaining energy of nodes and number of available channels for candidate CHs selection. Ant colony algorithm is employed to search inter-cluster paths with the purpose of extending node usage time. In ESAUC [24], CRSN nodes with higher number of available channels are more likely to become candidate CHs. Among them, those with more residual energy and common channels within competition radius become final CHs. Here, competition radius is introduced to implement unequal clustering and its definition is different from LEAUCH by taking neighbor count and probability of idle channels into consideration further. ESAUC uses a CogAODV based routing mechanism to perform inter-cluster forwarding, and expected path reliability is used to select the optimal routing path. WCM-based SAC protocol [25] is a spectrum-aware clustering routing protocol and it obtains the optimal clustering results by solving optimization model heuristically. It selects CHs according to temporal-spatial correlation, sensing confidence and residual energy. After clustering, CHs perform spectrum sensing instead of their member nodes, thereby saving the energy consumption of clusters. NSAC [26] integrates spectrum dynamics and energy consumption into clustering protocol design. It sets node weight according to remaining energy and quality of available channels, and selects node with the largest weight in neighborhood as CH. Other nodes in its maximum edge biclique join the cluster as CMs, and the clustering process continues until all nodes are clustered. HLEACH [27] is applicable to heterogeneous CRSN which is composed of ordinary nodes and cognitive nodes. HLEACH determines cluster number with the goal of minimizing node energy consumption. Cognitive nodes calculate competition radiuses through the global information broadcasted by the sink and confirm whether they can be CHs. HLEACH has superior performance in terms of channel detection probability, network lifetime and energy consumption among CHs. PROP [28] is a clustering routing protocol for CRSN with energy heterogeneity, and it divides the whole network into different regions according to the location of the sink. Weighted election probability to be CH for each type of nodes is calculated based on optimal probability value and average distance to the sink. Multihop path is utilized to transmit data from CHs far away from the sink. EACRP [29] is a distributed event-driven clustering routing protocol for CRSN, and it selects CH based on remaining energy, number of available channels, number of neighbors and distance to the sink. For inter-cluster communication, gateway nodes with more residual energy, more common channels with neighbors and closer to the sink are chosen to help route event data to the sink. ESUCR [30] is another event-driven clustering routing protocol for CRSN. A channel ranking algorithm is proposed to calculate channel stability according to idle probability and total number of transitions observed over a channel, and the most stable one is used for communication. CHs are selected and rotated based on residual energy, distance to the sink, neighbor connectivity with other clusters and intra-cluster channel stability. Neighboring clusters are merged according to the number of commonly unoccupied channels and measured distance between them until the optimal number of clusters is achieved. Primary and secondary gateways are selected hop by hop for packet forwarding towards the sink. ABCC [31] is a swarm intelligence-inspired clustering routing protocol for CRSN, and it takes full advantages of the cognitive behaviors of artificial bees which match perfectly with the dynamics of CRSN. It automatically determines the optimal number of clusters and CHs, and achieves the goal of extending network usage time.

Through above literature review, we find that most clustering routing protocols proposed for CRSN are heuristic algorithms. They usually calculate the optimal number of clusters based on assumptions such as CHs are at cluster centers and uniform distribution of nodes. However, when topology changes occur in CRSN due to node deaths, above assumptions cannot hold, and it is difficult to determine the new number of optimal clusters. Other heuristic clustering algorithms may rely on local information exchange in neighborhood to determine which node should become CH and which should be normal CMs. Decision made by utilizing limited local information cannot guarantee good performance from the whole network point of view. Therefore, heuristic algorithms exhibit poor adaptability to dynamic network changes. Swarm intelligence-based algorithms can automatically determine the optimal number of clusters based on network parameters and topology, which is consistent with the dynamic characteristics of CRSN [32]. However, existing swarm intelligence-based clustering protocols for CRSN such as ABCC are limited by high computational complexity and slow convergence speed. If algorithm complexity can be reduced, their performance can be effectively improved. Additionally, the sink node in CRSN usually has rich hardware resources and is able to acquire comprehensive network information. Compared with distributed protocols, centralized protocols have more advantages in small or medium-sized CRSN. Therefore, IMOCRP which adopts a centralized structure is proposed in this paper. The sink automatically determines CHs and the optimal number of clusters according to IMO algorithm, which can adapt to dynamic changes of network topology.

III. SYSTEM MODEL AND ASSUMPTIONS

N homogeneous CRSN nodes are uniformly distributed in the monitoring area, and one sink is located at the center or (0, 0) position in the network. These CRSN nodes act as secondary users (SUs) to P randomly distributed PUs in the same region. PUs send data on licensed spectrum bands intermittently, and Semi-Markov process is adopted to model PUs' behaviors. In this model, ON (Busy) and OFF (Idle) states of licensed spectrum are assumed to be independent of each other, and they both obey exponential distribution with parameters q_i and p_i, respectively [2]. In ON state, CRSN nodes within the interference protection range of PUs (a circle centered at PU and with radius 20m) cannot use the spectrum; in OFF state, nodes can use the spectrum for communication opportunistically. As a result, spectrum scarcity faced by traditional WSNs can be alleviated. As the spectrum usage of PUs is intermittent, channel occupancy state switches between ON and OFF states constantly as shown in Fig.1.

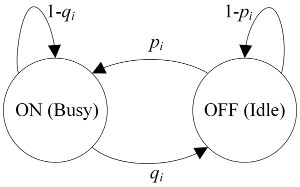

FIGURE 1. Semi-Markov ON/OFF process for modeling PUs' behaviors.

For the development of our protocol, we make the following basic assumptions in this paper:

- Each CRSN node has unique ID to distinguish from each other, and its position (x_i, y_i) is kept unchanged once deployed, unless it runs out of energy.
- Each CRSN node is configured with only one transceiver. Due to hardware constraint, it cannot perform spectrum sensing and data transmission at the same time.
- Each CRSN node can learn its geographical location information and remaining energy in each round, which is not in the scope of this paper. It can adjust its transmission power according to the distance from the destination.
- Perfect spectrum sensing is assumed, that is, each CRSN node can accurately detect the spectrum occupancy of PUs at its own location and determine its available channel list. Sensing errors such as miss detection and false alarm are ignored. Here, energy detection method is adopted due to its simplicity [33].
- Network-wide common control channel (CCC) is available to exchange control information for CH selection, cluster formation and network schedule [34].

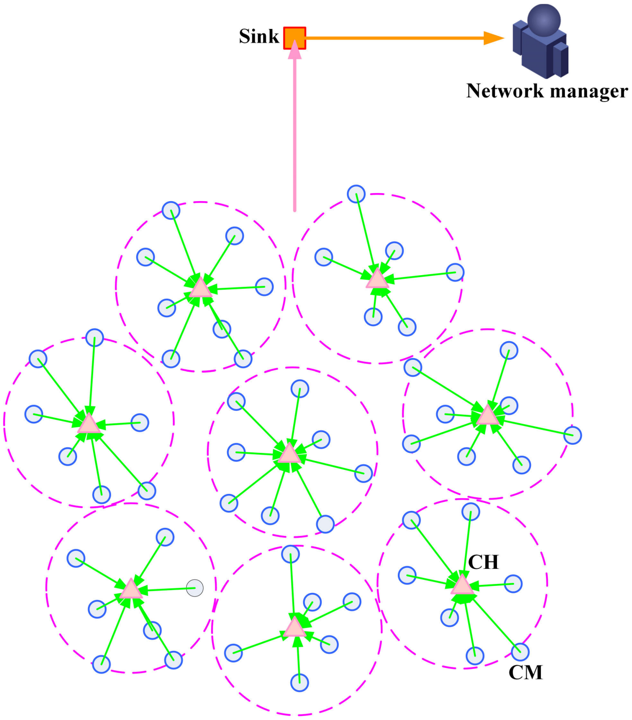

FIGURE 2. The architecture of cluster-based CRSN.

- The architecture of our cluster-based CRSN is presented in Fig.2. Each living CRSN node transmits its monitoring data to corresponding CH periodically, one packet per round. At CH, perfect aggregation [35] is done to compress all data from its CMs into single outgoing packet which is transmitted directly to the sink and onto the network manager, i.e., the aggregation coefficient α is the reciprocal of the number of CMs from which the CH can receive data report. Here, network manager can retrieve the desired information from the sink or it can request CRSN nodes for data collection.

The energy consumption model of CM is related to the Euclidean distance to its corresponding CH. Distance threshold d_0 determines the loss model used to model the radio wave propagation between the transmitter and receiver. Taking node *i* and its corresponding CH (denoted by CH$_j$) as an example, when their Euclidean distance d_{ij} is shorter than or equal to threshold d_0, free space loss model is applied. Otherwise, multipath fading loss model is used [36]. The energy consumption models of CM and CH are shown as below, respectively.

$$E\,(CM)_{dissipate} = \begin{cases} l \times E_{elec} + l \times E_{fs} \times d_{ij}^2, & if\ d_{ij} \le d_0 \\ l \times E_{elec} + l \times E_{mp} \times d_{ij}^4, & if\ d_{ij} > d_0 \end{cases} \tag{1}$$

$$E(CH)_{dissipate} = \begin{cases} l \times E_{elec} \times C_{num} + l \times E_{DA} \\ \quad \times (C_{num}+1) + l \times E_{elec} + l \\ \quad \times E_{fs} \times d_{tosink}^2, \quad if\ d_{tosink} \le d_0 \\ l \times E_{elec} \times C_{num} + l \times E_{DA} \\ \quad \times\ (C_{num}+1) + l \times E_{elec} + l \times E_{mp} \\ \quad \times d_{tosink}^4, \quad if\ d_{tosink} > d_0 \end{cases} \tag{2}$$

where l represents data packet size. E_{elec} is the electronics energy. E_{fs} is the amplifier energy consumption in free space model, and E_{mp} is the amplifier energy consumption in multipath fading model. Distance threshold $d_0 = (E_{fs}/E_{mp})^{1/2}$ [37]. C_{num} is the number of CMs in the cluster. E_{DA} is the energy required for data aggregation, and d_{tosink} is the Euclidean distance between the CH and the sink node. Since CHs need to receive and aggregate data from CMs in addition to sending information, their energy consumption is usually higher than that of CMs. Table 1 summarizes the notations used in this paper.

TABLE 1. Notation description.

Notation	Meaning
P	Number of PUs
N	Total number of CRSN nodes in the network
N_t	The number of nodes alive at the beginning of round *t*
α	Weight parameter of *OBJ*(*E*)
E_J	Initial energy of CRSN node when it is fully charged
p_i	Transition probability from OFF state to ON state
q_i	Transition probability from ON state to OFF state
E_{elec}	Electronics energy
E_{fs}	Energy consumption of amplifier in free-space loss model
E_{mp}	Energy consumption of amplifier in multipath fading loss model
E_{DA}	Energy consumption of fusing 1 bit data per packet
C_{num}	Number of CMs in a cluster
l	Data packet size
l_1	Control packet size
d_0	Distance threshold which determines the type of loss model used
$d_{i,j}$	Euclidean distance between node *i* and CH$_j$
d_{tosink}	Euclidean distance between CRSN node and the sink
$Anion_{i,j}$	j^{th} dimension of anion *i*
$Cation_{i,j}$	j^{th} dimension of cation *i*
$AD_{i,j}$	Euclidean distance between the j^{th} dimension of anion *i* and the best cation
$CD_{i,j}$	Euclidean distance between the j^{th} dimension of cation *i* and the best anion
$AF_{i,j}$	Gravitational attraction of the best cation to the j^{th} dimension of anion *i*
$CF_{i,j}$	Gravitational attraction of the best anion to the j^{th} dimension of cation *i*
$AVG(E_{dissipate})$	Average energy consumption of all living nodes
$STD(E_{residual})$	Standard deviation of remaining energy of all living nodes
t	Round number
$E(n)_{dissipate\_t}$	Energy consumption of node *n* in round *t*
N_t	Number of nodes alive at the beginning of round *t*
$E(n)_{residual\_t}$	Residual energy of node *n* at the end of round *t*
Q	Total number of CHs in Algorithm2
IN	Population of ions
$iter$	Iteration number
ϕ_a,ϕ_c,ϕ_{ac}	Random number belonging to range [0,1]
ϕ_1,ϕ_2	Random number belonging to range [–1,1]
T_1,T_2,T_3	Time duration of each phase in IMOCRP operation
t_1,t_2,t_3	Time slot length in each phase of IMOCRP

IV. ORIGINAL IONS MOTION OPTIMIZATION ALGORITHM

For a given optimization problem, IMO divides ion population equally into two groups, namely the anion group and the cation group. Due to attractive force between heterogeneous ions, ions move continuously in the search space, and their moving states can be divided into liquid phase and crystal phase. These two phases can transform to each other, and ions evolve continuously during the transformation process until termination conditions are satisfied.

In liquid phase, anions move towards the direction of the best cation ($Cation_{best}$), and cations move towards the direction of the best anion ($Anion_{best}$). Their position updating formulas are shown in (3) and (4), respectively.

$$Anion_{i,j} = Anion_{i,j} + AF_{i,j} \times AD_{i,j} \tag{3}$$

$$Cation_{i,j} = Cation_{i,j} + CF_{i,j} \times CD_{i,j} \tag{4}$$

where $Anion_{i,j}$ is the j^{th} dimension of anion *i* and $Cation_{i,j}$ is the j^{th} dimension of cation *i*. $AD_{i,j}$ is the Euclidean distance between $Anion_{i,j}$ and the best cation, therefore $AD_{i,j} = |Anion_{i,j}–Cation_{best,j}|$. $CD_{i,j}$ is the Euclidean distance between $Cation_{i,j}$ and the best anion, and we have $CD_{i,j} = |Cation_{i,j}–Anion_{best,j}|$. $AF_{i,j}$ and $CF_{i,j}$ represent the gravitational attraction of the best cation to $Anion_{i,j}$ and the gravitational attraction of the best anion to $Cation_{i,j}$, respectively. Their expressions are shown in (5) and (6), respectively.

$$AF_{i,j} = \frac{1}{1 + e^{\frac{-0.1}{AD_{i,j}}}} \tag{5}$$

$$CF_{i,j} = \frac{1}{1 + e^{\frac{-0.1}{CD_{i,j}}}} \tag{6}$$

In crystal phase, ions have two small-scale search methods and one initialization method. Ions randomly choose one of them to update their positions and avoid falling into local optimum [38]. The two small-scale search methods are distinguished by a random number ($\epsilon[0, 1]$) which is generated for each ion. Taking anion $Anion_i$ as an example, if the random number ϕ_a is larger than 0.5, $Anion_i$ is updated according to the first part of (7), otherwise, $Anion_i$ is updated according to the second part of (7) [10].

$$Anion_i = \begin{cases} Anion_i + \phi_1 \times (Cation_{best} - 1) & if\ \phi_a > 0.5 \\ Anion_i + \phi_1 \times Cation_{best} & else \end{cases} \tag{7}$$

where ϕ_1 is a random number which belongs to $[-1, 1]$. $Cation_{best}$ is the best cation in terms of fitness function value. Similarly, cations are updated according to the randomly generated number ϕ_c as below:

$$Cation_i = \begin{cases} Cation_i + \phi_2 \times (Anion_{best} - 1) & if\ \phi_c > 0.5 \\ Cation_i + \phi_2 \times Anion_{best} & else \end{cases} \tag{8}$$

where ϕ_2 is a random number which belongs to $[-1, 1]$. $Anion_{best}$ is the best anion in terms of fitness function value.

After position updating, a random number ϕ_{ac} is generated for each ion again by IMO, and if $\phi_{ac} \leq 0.05$, corresponding ion will be initialized. Therefore, the original IMO algorithm applies 3 update methods to each ion in crystal phase, here taking $Anion_i$ as an example, the three update methods are summarized below:

$$Anion_i = \begin{cases} Anion_i + \phi_1 \times (Cation_{best} - 1) & if\ \phi_a > 0.5 \\ & \cap\ \phi_{ac} > 0.05 \\ Anion_i + \phi_1 \times Cation_{best} & if\ \phi_a \leq 0.5 \\ & \cap\ \phi_{ac} > 0.05 \\ initialized & else \end{cases} \tag{9}$$

In order to be consistent with the original IMO algorithm, our IMOCRP also adopts 3 update methods in crystal phase.

V. PROPOSED IMO-BASED CLUSTERING ROUTING PROTOCOL

A. FRAME STRUCTURE OF IMOCRP

Fig.3 shows the frame structure of proposed IMOCRP. The frame is based on time division multiple access (TDMA) scheduling to ensure conflict-free transmissions. The whole frame can be divided into 3 phases, that is, spectrum sensing phase (Phase 1), clustering phase (Phase 2), and data transmission phase (Phase 3), whose duration is T_1, T_2 and T_3, respectively.

- Phase 1 Spectrum Sensing: each node performs spectrum sensing individually. Sensing errors are not taken into consideration, i.e., perfect sensing is assumed. There are C licensed channels in total, therefore $T_1 = C \times t_1$. Here t_1 is the spectrum sensing time per channel.
- Phase 2 Clustering: each living CRSN node sends its available channel list obtained in Phase 1, its geographical location and its remaining energy to the sink on CCC. Each node is assigned with a time slot whose length is t_2 according to the ID order to avoid interference among these control message transmissions. The sink receives the information in each slot. The final clustering results are produced by the sink through running IMOCRP (which is explained in detail in Algorithm1), as it is assumed that the sink is rich in hardware resources, the time of running IMOCRP can be neglected. At last, the sink broadcasts final clustering results including the role of each node, the cluster each node belonging to, cluster channel and schedule time of each node in a time slot, and each living CRSN node receives this message immediately. Therefore, the time length of T_2 is $(N_r + 1) \times t_2$. Here, N_r is the number of living CRSN nodes at the beginning of round r.

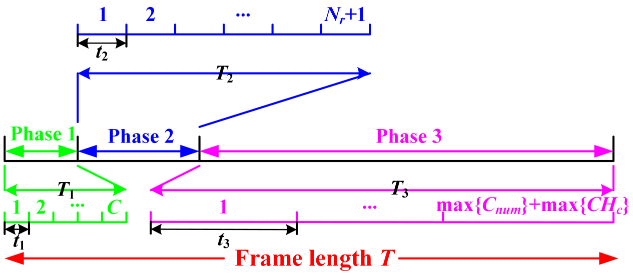

FIGURE 3. Frame structure of IMOCRP.

- Phase 3 Data Transmission: each CM sends its monitoring data to its CH in its assigned time slot whose length is t_3, and the CH receives data packet accordingly. The time spent in data aggregation is neglected. Hence, the time interval for intra-cluster data collection is $\max\{C_{num}\} \times t_3$. Here, C_{num} is the number of CMs in one cluster, and $\max\{C_{num}\}$ is the largest C_{num} in all clusters. After that, each CH sends the aggregated data to the sink on its cluster channel in its exclusive time slot determined by the sink. The sink can receive packets on C licensed channels at the same time through its multiple transceivers. CHs using the same channel cannot activate at the same time, and they should occupy the channel in sequence. However, CHs on different channels can activate simultaneously as they do not interfere with each other. Therefore, inter-cluster data transmission takes $\max\{CH_c\} \times t_3$ length of time. Here, CH_c is the number of CHs using channel c, and $\max\{CH_c\}$ is the largest CH_c on all licensed channels. Therefore, we have $T_3 = (\max\{C_{num}\} + \max\{CH_c\}) \times t_3$. Of course, it is assumed that the available channels in Phase 1 are still available in Phase 3, and nodes will not sense the spectrum again at the beginning of Phase 3. If the available channels are reclaimed by PUs during the rest time of this round, collisions occur. In this case, some CRSN nodes cannot transmit their data to the sink successfully.

B. OPTIMIZATION TARGET OF IMOCRP

IMOCRP aims at prolonging network lifetime and improving data collection capability of CRSN. On the one hand, CRSN nodes usually use micro-batteries with very limited capacity for power supply, and due to environmental factors, it is difficult to replace batteries after node deployment. Therefore, energy consumption has become a key factor which restricts the lifespan of CRSN. Reducing average energy consumption of nodes $AVG(E_{dissipate})$ is beneficial for network lifetime expansion. On the other hand, due to their limited coverage range, premature death of CRSN nodes will lead to significant reduction in monitoring scope and information collection capabilities. Reducing the standard deviation of node residual energy $STD(E_{residual})$ can delay deaths of high energy-consuming nodes and improve information collection capability of CRSN. Taking above two factors both into consideration, the optimization target of IMOCRP is to maximize the following function:

$$OBJ\ (E) = \frac{1}{AVG\left(E_{dissipate\_t}\right) + \alpha \times STD\left(E_{residual\_t}\right)} \tag{10}$$

where $AVG(E_{dissipate\_t})$ is node average energy consumption in round t. $STD(E_{residual\_t})$ is the standard deviation of remaining energy in each node in round t. α is a tunable weight assigned to $STD(E_{residual\_t})$ to quantify its relative effect on the whole function. The higher the value of α, the more important $STD(E_{residual\_t})$ in the entire objective is.

$$AVG\left(E_{dissipate\_t}\right) = \left(\sum_{n=1}^{N_t} E\ (n)_{dissipate\_t}\right) / N_t \tag{11}$$

$$STD\left(E_{residual\_t}\right) = \sqrt{\frac{1}{N_t}\sum_{n=1}^{N_t}\left(E\ (n)_{residual\_t} - AVG\left(E_{residual\_t}\right)\right)^2} \tag{12}$$

where $E(n)_{dissipate\_t}$ is the energy consumption of node n in round t, which can be calculated by (1) or (2). $N_t(\neq 0)$ is the number of nodes alive at the beginning of round t. $E(n)_{residual\_t}$ is the residual energy of node n at the end of round t, and its definition is shown in (13).

$$E\ (n)_{residual\_t} = E\ (n)_{residual\_t-1} - E\ (n)_{dissipate\_t} \tag{13}$$

If $t = 1$, $E(n)_{residual\_t-1}$ is defined as node initial energy E_J. $AVG(E_{residual\_t})$ is the average remaining energy of all living nodes at the end of round t, and it is expressed as:

$$AVG\left(E_{residual\_t}\right) = \frac{1}{N_t}\sum_{n=1}^{N_t} E\ (n)_{residual\_t} \tag{14}$$

C. MAPPING RALATIONSHIP BETWEEN IONS AND CLUSTERING RESULTS IN IMOCRP

IMOCRP uses each anion or cation as possible clustering result. Each anion or cation is represented by a binary vector with N_t dimensions, and each element in it represents the identity of a living CRSN node, that is, value 1 indicates CH role, and 0 means CM. Taking CRSN A in 10 m × 10 m area as an example, 10 CRSN nodes and 3 PUs are randomly distributed in it. Positions of CRSN nodes and PUs are shown in Table 2. $Anion_1$ and $Cation_1$ are randomly generated for network A, as shown in Table 3.

TABLE 2. Locations of CRSN nodes and PUs.

node	1	2	3	4	5	6	7	8	9	10
x(m)	8	1	1	3	1.5	5	8	2	9.5	7
y(m)	9	9	6	5.5	8	6	1	1	3.5	2.5
PU	PU_1	PU_2	PU_3							
x(m)	3	6	2							
y(m)	4	6	9							

TABLE 3. Values of $Anion_1$ and $Cation_1$.

dimensions	1	2	3	4	5	6	7	8	9	10
$Anion_1$	0	0	1	0	0	0	0	0	1	0
$Cation_1$	0	0	0	0	1	1	0	0	0	0

When clustering the network according to $Anion_1$, it can be known from $Anion_{1,3} = 1$, $Anion_{1,9} = 1$ that the third and ninth nodes are CHs, and remaining nodes are CMs. Similarly, when clustering the network with the result shown by $Cation_1$, it can be known from $Cation_{1,5} = 1$ and $Cation_{1,6} = 1$ that the fifth and sixth nodes are CHs, and remaining nodes are CMs.

D. OPERATION PROCESS OF IMOCRP

IMOCRP aims at maximizing $OBJ(E)$ in (10) and its pseudo code is shown in Fig.4.

Network parameters are initialized in the first line. These parameters include node initial energy E_J, electronic circuit energy consumption E_{elec}, node amplifier energy consumption E_{fs} (in free space loss model) and E_{mp} (in multipath fading loss model), data aggregation energy E_{DA}, data packet size l, control packet size l_1, population number of ions IN.

Spectrum sensing phase is shown in lines 2-3. Each node senses spectrum through energy detection method, and then sends its remaining energy, geographical location, and current spectrum detection results to the sink on CCC.

In clustering phase (lines 4-16), the sink randomly generates $IN/2$ anions $\{Anion_1, Anion_2 \ldots Anion_{IN/2}\}$ and $IN/2$ cations $\{Cation_1, Cation_2 \ldots Cation_{IN/2}\}$. CRSN is clustered by these anions and cations and the sink evaluates their fitness values, i.e., the $OBJ(E)$ defined by (10). The higher the value of $OBJ(E)$ goes, the better the clustering result of ion is. In order to obtain the fitness value of each ion, the connection relationship between each normal node and the CHs should be determined first. In other words, we need to find proper CMs for each CH to form clusters according to adjacency relationship and common available channels. This process is done according to Algorithm2 which is shown in Fig.5. Only in this case, we can know the Euclidean distance between any CM and its corresponding CH. Based on the Euclidean distance obtained, we can calculate energy consumption of each node $E(n)_{dissipate\_t}$ and its residual energy

```
Algorithm1: IMO-based Clustering Routing Protocol
1. Use an $N$ nodes and $P$ PUs network with a sink node.
   Set the energy parameter $E_J$, $E_{elec}$, $E_{mp}$, $E_{fs}$, $E_{DA}$, $l$, $l_1$ and the population
   of ions $IN$.
Phase1:Spectrum Sensing
2. CRSN nodes sense available channels.
3. Each node sends its residual energy, location and available channel list
   to the sink.
Phase2:Clustering
4. The sink randomly generates $IN/2$ Anions and $IN/2$ Cations.
5. Anions and Cations cluster CRSN according to Algorithm2.
6. Evaluate Ions according to (10) and elect $Anion_{best}(0)$, $Cation_{best}(0)$ and
   $Ion_{best}(0)$.
7. $iter$=0, $indicator$=1.
8. while $indicator$=1
9.    $iter$=$iter$+1
10.   Liquid Phase: Anions($iter$–1) move close to $Cation_{best}(iter–1)$ accord-
      ing to (15). Cations($iter$–1) move close to $Anion_{best}(iter–1)$ according
      to (16).
11.   Crystal Phase: Divide Ions evenly into 3 groups. In $group_1$, $Anion=G_1(Anion)$,
      $Cation=G_1(Cation)$. In $group_2$, $Anion=G_2(Anion)$,
      $Cation=G_2(Cation)$. In $group_3$, $Anion=G_3(Anion)$, $Cation=G_3(Cation)$.
      $G_1$, $G_2$, $G_3$ are different indicator functions.
12.   Evaluate Ions according to (10) and elect $Anion_{prebest}(iter)$ and
      $Cation_{prebest}(iter)$.
13.   Elect $Anion_{best}(iter)$ from $Anion_{prebest}(iter)$ and $Anion_{best}(iter–1)$.
      Elect $Cation_{best}(iter)$ from $Cation_{prebest}(iter)$ and $Cation_{best}(iter–1)$.
      Elect $Ion_{best}(iter)$ from $Anion_{best}(iter)$ and $Cation_{best}(iter)$.
14.   if $iter<3N/IN$
         $indicator$=1.
      elseif $3N/IN \le iter < 2N$
               if $Ion_{best}(iter) \ne Ion_{best}(iter–3N/IN)$
                  $indicator$=1.
               else
                  $indicator$=0.
               end
         else
            $indicator$=0.
         end
      end
15.end
16.CRSN is clustered by $Ion_{best}(iter)$.
Phase3:Data Transmission
17.Each CM sends data to its CH in corresponding time slot.
18.CH aggregates data and sends it to the sink.
Reiterate
19.After round time, clustering is re-triggered and Phase1-3 are reiterated
   until all nodes die.
```

FIGURE 4. Pseudo-code of IMOCRP.

```
Algorithm2: Cluster Formation
Step1: Tentative Cluster Formation
1. for $i$=1:$N_r$–$Q$
2.   for $j$=1:$Q$ ($d_{i,1} \le d_{i,2} \le \ldots\ldots \le d_{i,Q}$)
3.     if node$_i$(channel)$\cap$CH$_j$(channel)$\ne\Phi$
4.        node$_i$ becomes tentative CM of Cluster(CH$_j$).
5.        break
6.     end
7.   end
8. end
Step2: Cluster Channels Selection
9. for $j$=1:$Q$
10.  the sink selects Channel(CH_$j$) as cluster channel for Cluster(CH$_j$).
11.end
Step3: Final Cluster Formation
12.for $i$=1:$N_r$–$Q$
13.  for $j$=1:$Q$ ($d_{i,1} \le d_{i,2} \le \ldots\ldots \le d_{i,Q}$)
14.     if Channel(CH_$j$)$\in$node$_i$(channel)
15.        node$_i$ becomes CM of Cluster(CH$_j$).
16.        break
17.     end
18.  end
19.  if node$_i$ $\notin$\{Cluster(CH$_1$)$\cup$Cluster(CH$_2$)$\cup$……$\cup$Cluster(CH$_Q$)\}
20.     node$_i$ becomes separate CH.
21.  end
22.end
```

FIGURE 5. Sub-algorithm for cluster formation in IMOCRP.

TABLE 4. Node channel information for CRSN *A*.

node	1	2	3	4	5	6	7	8	9	10
channel 1	√	√	√		√	√	√	√	√	√
channel 2	√	√	√	√	√		√	√	√	√
channel 3	√		√	√		√	√	√	√	√
PU	PU_1	PU_2	PU_3							
channel 1	√									
channel 2		√								
channel 3			√							

$E(n)_{residual\_t}$, and then $AVG(E_{dissipate\_t})$, $STD(E_{residual\_t})$ and fitness value at last. Of course, these connection relationships are only temporary and used for ions evolution. The connection relationships in the final best ion of this round $Ion_{best}(iter)$ are used to form actual clusters.

Algorithm2 is composed of three execution steps. In step 1, as shown in lines 1-8, assuming that there are Q CHs in the network, each ordinary node becomes tentative CM of the closest CH with common channels. Then, the sink records the available channels in each cluster. In step 2, as shown in lines 9-11, IMOCRP selects the channel with the largest number of covered nodes from the CH's available channel list as cluster channel. Randomly selection is used to break the tie. In step 3, as shown in lines 12-22, for normal node, say i, the CH which is closest to it and whose cluster channel is contained in the available channel list of i is chosen as its final CH. If the CH that meets above conditions does not exist, the node itself becomes an isolated CH and communicates with the sink directly.

Taking CRSN *A* as an example again, it is assumed that interference protection range of each PU is 2 m and PUs are in ON state. The corresponding channel information is listed in Table 4 in which "√" indicates that the corresponding channel is currently available to the node. The node distribution and available channels of CRSN *A* are shown in Fig.6(a).

Suppose $Anion_1$ is used to cluster CRSN *A*. In tentative cluster formation step, nodes 2, 4, 5, 6, and 8 become tentative CMs of CH node 3, and nodes 1, 7, and 10 become tentative CMs of CH node 9. In cluster channel selection step, for CH node 3, its available channel 1 covers four nodes 2, 5, 6, 8, and channel 2 covers four nodes 2, 4, 5, 8, while channel 3 covers three nodes 4, 6, 8. As equal number of nodes operate on channel 1 and channel 2, CH node 3 randomly selects one from them, say channel 2, as final cluster channel. Similarly, suppose CH node 9 selects channel 1. In final cluster formation step, channel 2 is included in the available channel lists of nodes 2, 4, 5, 8, and they are closest to CH node 3, so they form cluster with CH node 3. In the same way, nodes 1, 6, 7, 10 are clustered with CH node 9. The final clustering result of CRSN *A* is shown in Fig.6(b).

The original best anion $Anion_{best}(0)$ and the original best cation $Cation_{best}(0)$ are selected from all anions and cations

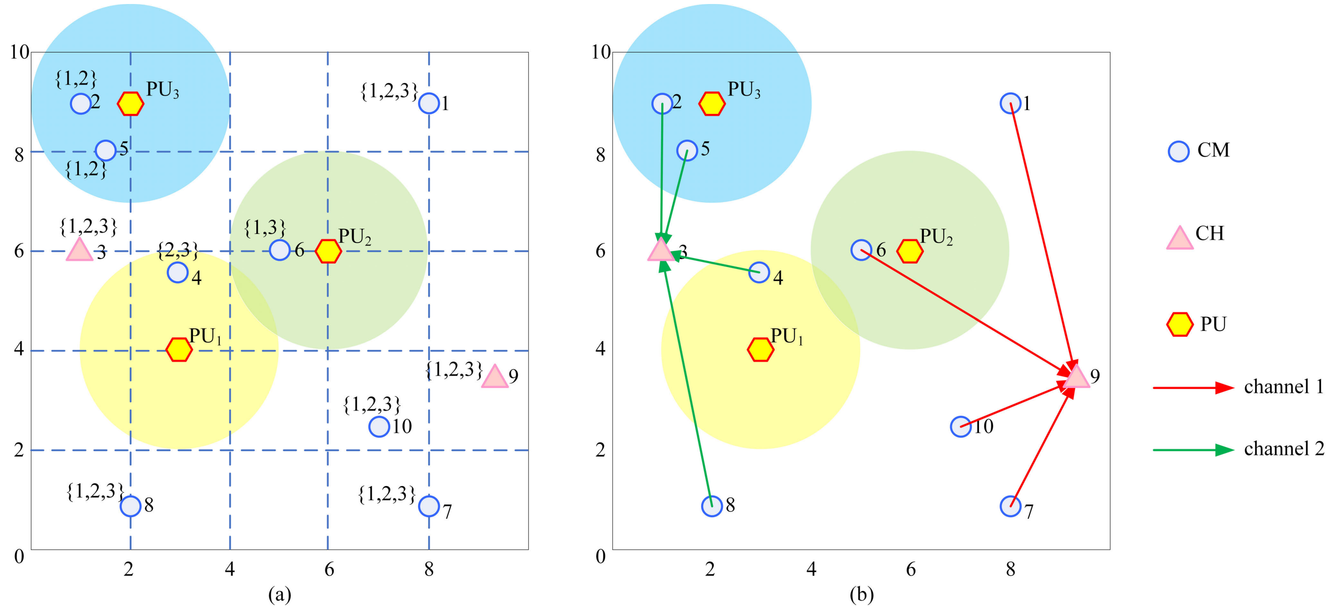

FIGURE 6. Node distribution and final clustering results of CRSN *A*.

generated above, and the one with larger fitness value among them is selected as the original best ion $Ion_{best}(0)$. Line 10 shows the liquid phase, and anions and cations in it are updated as follows:

$$Anion_{i,j}\,(iter) = \begin{cases} Anion_{i,j}\,(iter-1) & if\ \ i = best \\ \operatorname{mod}\left(Anion_{i,j}\,(iter-1) + f\left(AF_i \times \left|Cation_{best,j}\,(iter-1) - Anion_{i,j}\,(iter-1)\right|\right),2\right) & else \end{cases} \tag{15}$$

$$Cation_{i,j}\,(iter) = \begin{cases} Cation_{i,j}\,(iter-1) & if\ \ i = best \\ \operatorname{mod}\left(Cation_{i,j}\,(iter-1) + f\left(CF_i \times \left|Anion_{best,j}\,(iter-1) - Cation_{i,j}\,(iter-1)\right|\right),2\right) & else \end{cases} \tag{16}$$

where mod(y, 2) takes the remainder when y is divided by 2. $f(x)$ means a randomly generated integer 0 or 1, where x is the probability of 1. $|Cation_{best,j}(iter–1)–Anion_{i,j}(iter–1)|$ represents the Euclidean distance between the j^{th} dimension of $Anion_i$ and $Cation_{best}$ in iteration $iter$–1. For example, when $Anion_{i,j}(iter–1) = 1$, $Cation_{best,j}(iter–1) = 0$, $|Cation_{best,j}(iter–1)–Anion_{i,j}(iter–1)| = 1$. AF_i is the gravity between $Cation_{best}(iter–1)$ and $Anion_i(iter–1)$, which is calculated by:

$$AF_i = \frac{\sum_{j=1}^{N_t}\left(Cation_{best,j}\,(iter-1) - Anion_{i,j}\,(iter-1)\right)}{N_t} \tag{17}$$

Similarly, $|Anion_{best,j}(iter–1)–Cation_{i,j}(iter–1)|$ represents the Euclidean distance between the j^{th} dimension of $Cation_{\ i}$ and $Anion_{best}$ in iteration $iter$–1. CF_i is the gravity between $Anion_{best}(iter–1)$ and $Cation_i(iter–1)$.

$$CF_i = \frac{\sum_{j=1}^{N_t}\left(Anion_{best,j}\,(iter-1) - Cation_{i,j}\,(iter-1)\right)}{N_t} \tag{18}$$

The crystal phase (line 11) divides anions and cations into 3 groups, namely $group_1$, $group_2$, and $group_3$. In $group_1$, ions randomly select dimension j with value 1 and change its value to 0. The expressions are $Anion = G_1(Anion)$ and $Cation = G_1(Cation)$. In $group_2$, ions randomly select dimension k with value 0 and change its value to 1. The expressions are $Anion = G_2(Anion)$ and $Cation = G_2(Cation)$. In $group_3$, ions randomly select dimension j with value 1 and dimension k with value 0 and exchange them. The expressions are $Anion = G_3(Anion)$ and $Cation = G_3(Cation)$. Taking $Anion_1$ (in Table 3) as an example, the values of different groups after updating are shown in Table 5.

TABLE 5. Update table of $Anion_1$ in crystal phase.

dimensions	1	2	3	4	5	6	7	8	9	10
$G_1(Anion_1)$	0	0	0	0	0	0	0	0	1	0
$G_2(Anion_1)$	0	0	1	1	0	0	0	0	1	0
$G_3(Anion_1)$	0	0	1	0	0	0	0	1	0	0

After crystal phase, candidate best anion $Anion_{prebest}(iter)$ and candidate best cation $Cation_{prebest}(iter)$ for this iteration are selected according to (10). Then, the best anion $Anion_{best}(iter)$ is selected from $Anion_{prebest}(iter)$ and $Anion_{best}(iter–1)$, and the best cation $Cation_{best}(iter)$ is selected from $Cation_{prebest}(iter)$ and $Cation_{best}(iter–1)$.

Finally, the best ion $Ion_{best}(iter)$ of this iteration is picked out from $Anion_{best}(iter)$ and $Cation_{best}(iter)$. The iterative cycle is performed at least $3N/IN$ times and at most $2N$ times. When the number of cycles is greater than or equal to $3N/IN$, if the best ion Ion_{best} does not change in consecutive $3N/IN$ iterations, the liquid to crystal cycle ends ahead of time, and $Ion_{best}(iter)$ is the final clustering result of this round. Here, $3N/IN$ is set to ensure sufficient local search and avoid falling into local optima, and $2N$ is set to strike a compromise between network performance and computational complexity according to extensive simulation results.

In data transmission phase (lines 17-19), CMs send data to their CH in corresponding time slot on cluster channel. The CH aggregates all received data and sends it to the sink directly. The completion of all data transmissions marks the end of this round. Spectrum sensing, clustering, and data transmission phases of IMOCRP are repeated until the end of network lifetime.

VI. SIMULATION RESULTS AND ANALYSIS

In this section, our proposed protocol is studied and analyzed by using MATLAB tool. Performance comparisons with existing protocols, such as CogLEACH [19], CogLEACH-C [20], LEAUCH [22], Fuzzy C-means [21], NSAC [26], WCM-based SAC [25], and ABCC [31], verify the advantages of IMOCRP in energy efficiency and information collection capability. Among them, Fuzzy C-means and ABCC leave channel occupancy out of consideration, while others all consider about channel availability at each node and the PUs activity. In order to further substantiate the superiority of our proposed IMOCRP, we improved Fuzzy C-means and ABCC by performing spectrum sensing at the beginning of each round and taking channel availability into account while forming clusters and selecting cluster channel. The improved ones are denoted as i-Fuzzy C-means and i-ABCC, respectively. Simulations are conducted in a square network of 100 m × 100 m with the sink located at the center and (0, 0), respectively. Simulation parameters are shown in Table 6. For fair comparisons, the input parameters of ABCC are set according to [31], i.e., $SN = 20$, $MaxCycle = 1500$.

A. THE OPTIMAL VALUE OF WEIGHT α

In the optimization target of IMOCRP (see (10)), weight α is assigned to the component $STD(E_{residual\_t})$ to quantify its relative effect on the whole function. $AVG(E_{dissipate\_t})$ and $STD(E_{residual\_t})$ both change round by round, and it is difficult to provide theoretical analysis for determining the optimal value of α. Therefore, we conduct extensive simulations and observe the performance of IMOCRP while changing α from small value (such as 0.2) to large value (such as 100). We also consider about minimizing total energy consumption (denoted by Obj_1, equivalent to $\alpha \rightarrow 0$) and minimizing standard deviation of node remaining energy (denoted by Obj_2, equivalent to $\alpha \rightarrow \infty$) separately, and the simulation results are shown in Fig.7 below.

TABLE 6. Simulation parameters.

Parameter	Value
initial energy of CRSN node E_J	0.5J
electronics energy E_{elec}	50nJ/bit
energy consumption of amplifier in free-space loss model E_{fs}	10pJ/bit/m^2
energy consumption of amplifier in multipath fading loss model E_{mp}	0.0013pJ/bit/m^4
energy consumption of fusing 1 bit data E_{DA}	5nJ/bit/packet
length of data packet l	1024bit
length of control packet l_1	20bit
population of ions IN	30
number of CRSN nodes N	100
number of PUs P	5
probability of PU in ON state p_i	0.7
distance threshold d_0	87.7m
spectrum sensing time per channel t_1	5ms
channel bandwidth (CCC)	256kbps
channel bandwidth (data channel)	1Mbps

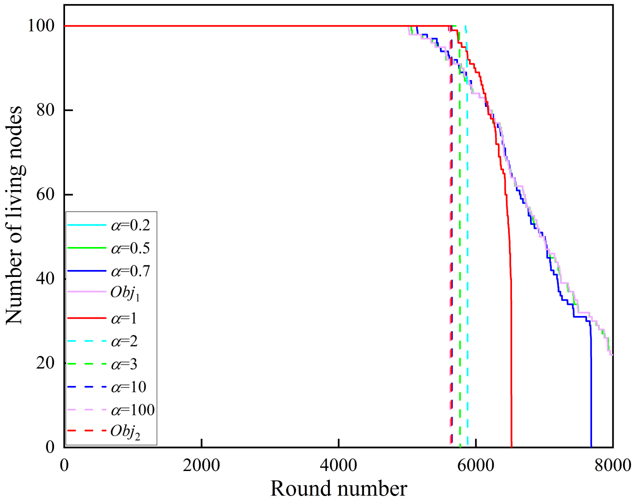

FIGURE 7. Effect of α value on the number of living nodes for IMOCRP.

Through simulations, we can observe that for low-power CRSN nodes, usually $AVG(E_{dissipate\_t})$ is not more than $STD(E_{residual\_t})$. Therefore, if $\alpha > 1$, $STD(E_{residual\_t})$ dominates the objective function and all nodes almost die in the same round. While $\alpha < 1$, $AVG(E_{dissipate\_t})$ dominates the objective function and the round difference between the first-dead node and the last-dead node becomes larger. In other words, in this case IMOCRP cares more about $AVG(E_{dissipate\_t})$ and less about the energy consumption balance among different nodes. Some nodes may die earlier due to faster energy consumption than others. This will negatively affect the network lifetime. According to the above simulation results and analysis, we set $\alpha = 1$ to take average energy consumption of nodes and their energy balance comprehensively into consideration, and this configuration is applied to all the simulations in this paper below.

B. NETWORK LIFETIME

In this section, we show the relationship between the number of surviving nodes with the running time (represented by rounds) when the sink is located at the center and location (0,0), respectively. The results are given in Fig.8 and Fig.9.

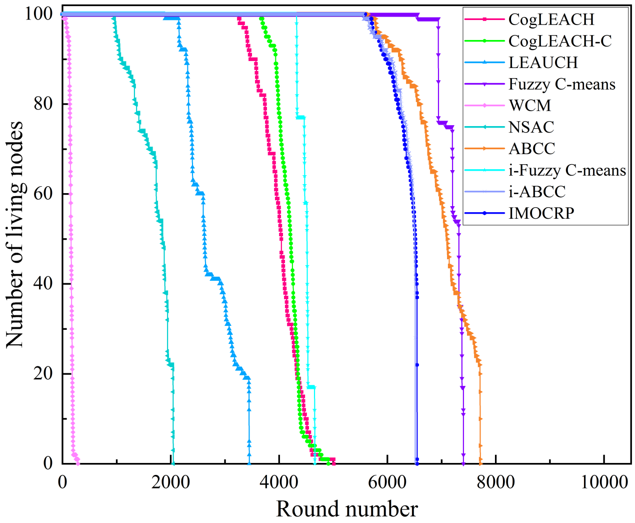

FIGURE 8. Number of living nodes in each round (with the sink located at the center).

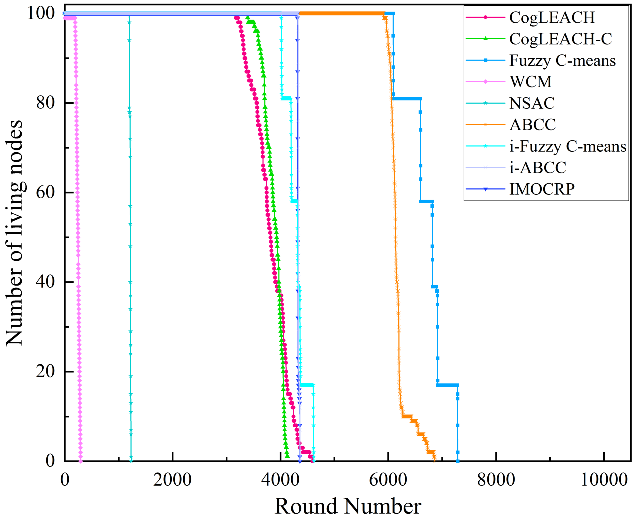

FIGURE 9. Number of living nodes in each round (with the sink located at (0, 0)).

We can know from Fig.8 and Fig.9 that: Firstly, rounds in which the first node dies under IMOCRP are 5594 and 4315 respectively, which are significantly larger than majority protocols except ABCC and Fuzzy C-means. It means that IMOCRP can reduce energy consumption of nodes and postpone first node death time. The reduction of node energy consumption is mainly achieved by reducing the overhead of control information exchange during CH selection and cluster construction. In CogLEACH, CHs need to broadcast temporary CH announcements and final CH announcement messages in cluster construction; CMs send temporary join requests and final confirmation requests to corresponding CHs. Therefore, the total overhead of CogLEACH protocol in CH selection and cluster construction is twice the number of living nodes. CogLEACH-C requires all nodes to send node information (including their energy level, geographic location and number of available channels) to the sink when selecting CHs. The sink determines and broadcasts node identity according to the node information. CHs broadcast cluster channel information; CMs send temporary joining requests and final confirmation requests to corresponding CHs. The CH selection and cluster construction cost of CogLEACH-C protocol is the sum of two times the number of living nodes and the number of CMs. In LEAUCH, candidate CHs compete for final CHs by comparing remaining energy within competition radius and exchanging information. The cluster construction method of LEAUCH is the same as that of CogLEACH protocol, so the total overhead of LEAUCH in CH selection and cluster construction is twice the number of candidate CHs plus twice the number of living nodes. WCM-based SAC protocol requires all CRSN nodes to exchange spectrum sensing information with neighboring nodes and calculate CH weight. Nodes exchange weight information to select CHs and build initial clusters. The optimal number of clusters is reached by continuously merging clusters with high temporal-spatial correlation. Therefore, the total cost of WCM-based SAC in CH selection and cluster construction is about 4 times the number of living nodes. In NSAC protocol, all nodes that have not been clustered continuously update and broadcast node weight information, select CHs through weight comparison in neighborhood, and build clusters until all nodes are clustered. Both WCM-based SAC and NSAC protocol require a large amount of control information exchange between nodes, which increases energy consumption of nodes. IMOCRP only requires each living CRSN node to send node information to the sink once during cluster formation phase. If the rich hardware resource at the sink is taken into consideration, the total cost in CH selection and cluster construction is approximately the number of living nodes in the network, which is significantly lower than that of CogLEACH, CogLEACH-C, LEAUCH, WCM-based SAC, etc. Secondly, when comparing with original clustering algorithms, the average packet delivery ratio of improved ones, i.e., i-Fuzzy C-means and i-ABCC, is greatly improved. However, high packet delivery ratio comes at the cost of high energy consumption. If packets can be received at the CH on the same available channel, the CH should consume energy for receiving and aggregating these packets, and forward the aggregated packet further to the sink. Therefore, the network lifetime of i-ABCC and i-Fuzzy C-means is much shorter than the original algorithms. In addition, the performance of IMOCRP is better than that of i-Fuzzy C-means and better than or at least equal to that of i-ABCC. However, through the computational complexity analysis in the *E* part, we know that IMOCRP gains advantages over i-ABCC. Thirdly, in the two topologies, difference between the round in which the first node and the last node die in IMOCRP is 956 and 46, respectively, which is smaller than

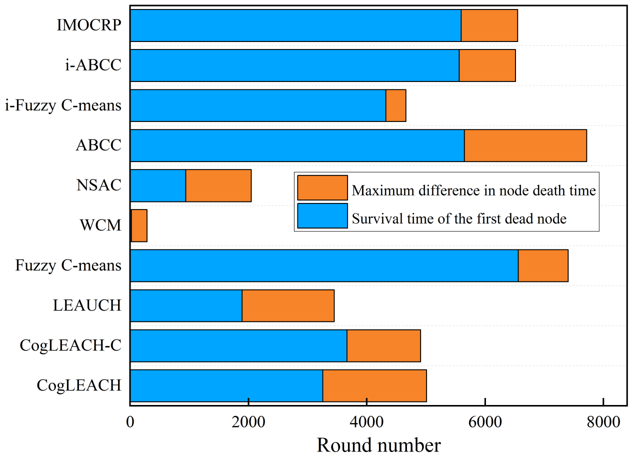

FIGURE 10. Comparison of node death time (with the sink at the center).

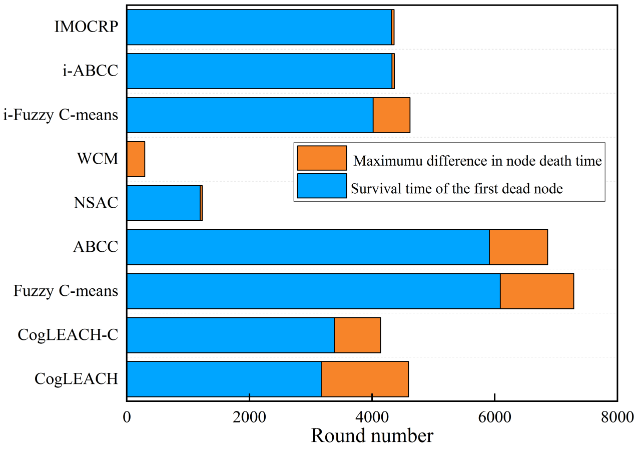

FIGURE 11. Comparison of node death time (with the sink located at (0,0)).

most of the comparison protocols. The maximum differences between the first node death (expressed in rounds) and the last node death under different protocols are shown in Fig. 10 and Fig. 11.

The smaller the difference in the number of rounds, the closer the death time of CRSN nodes, that is, the more balanced the energy consumption between nodes. Above results illustrate that IMOCRP can simultaneously achieve the goals of reducing average node energy consumption and average node residual energy standard deviation, thereby extending network lifetime and improving network performance.

In order to verify the advantages of IMOCRP further, we compare the number of CHs selected by each protocol in 4 randomly selected rounds, that is, round 1000, 2000, 3000, and 4000. The results are shown in Fig.12.

Fig.12 shows that the number of CHs of IMOCRP is significantly higher than other protocols. When the sink is located at the center, most nodes are CH nodes without CMs, as shown in Fig.13(a); whereas CogLEACH, CogLEACH-C, NSAC, and other protocols have relatively low number of CHs, as shown in Fig.13(b).

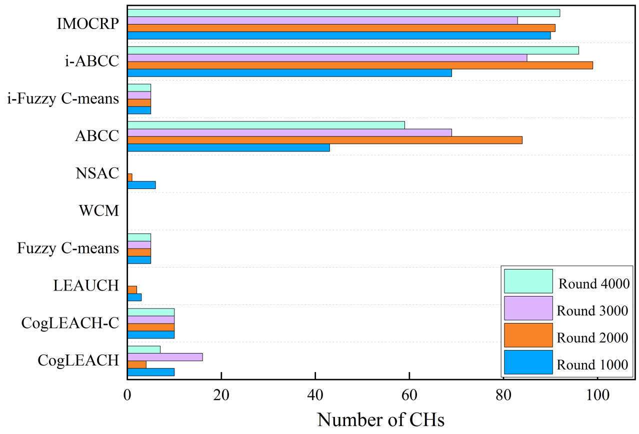

FIGURE 12. Number of CHs statistics (with the sink at the center).

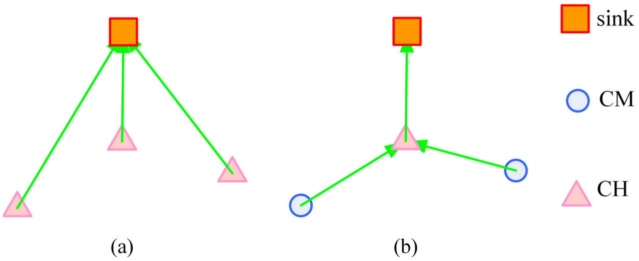

FIGURE 13. Schematic diagram of clustering.

Assuming that there are n nodes in the network, in Fig.13(a), the number of CMs in each cluster is zero. When the sink is at the center, since $d_{tosink} < d_0$, according to (2), the total energy consumption of all CHs is:

$$E\left(a\right)_{dissipate} = n\times l\times E_{elec} + l\times E_{fs}\times\sum_{i=1}^{n} d_{tosink}^{2}\left(n_i\right) \tag{19}$$

In Fig.13(b), CMs are closer to the CH than to the sink. Assuming $d_{toCH} < d_0$, according to (1) and (2), the total energy consumption of all nodes is:

$$\begin{aligned} E\left(b\right)_{dissipate} &= E\left(CH\right)_{dissipate} + E\left(CM\right)_{dissipate} \\ &= (2n-1)\times l\times E_{elec} + n\times l\times E_{DA} \\ &\quad + l\times E_{fs}\times d_{tosink}^{2} + l\times E_{fs}\times\sum_{i=1}^{n-1} d_{toCH}^{2}\left(n_i\right) \end{aligned} \tag{20}$$

The energy consumption difference between the two cases is:

$$\begin{aligned} E_{DIF} &= E\left(a\right)_{dissipate} - E\left(b\right)_{dissipate} \\ &= l\times E_{fs}\times\left(\sum_{i=1}^{n-1} d_{tosink}^{2}\left(n_i\right) - \sum_{i=1}^{n-1} d_{toCH}^{2}\left(n_i\right)\right) \\ &\quad - (n-1)\times l\times E_{elec} - n\times l\times E_{DA} \end{aligned} \tag{21}$$

From (21), we can know that Fig.13(b) is more energy-efficient than Fig.13(a) when the distance between

CMs and the sink is longer and the distance to the CH is shorter. However, when the sink is located at the center, CRSN nodes and the sink are very close. Under this circumstance, nodes are clustered separately to save energy, that is, the energy saving effect is better when number of CHs is high. According to the network topology, the sink automatically determines the optimal number of clusters, which makes IMOCRP achieve high number of CHs, thereby achieving the goal of reducing average energy consumption of nodes.

C. PACKET DELIVERY RATIO

In addition to network lifetime, information collection capability is also an important performance indicator of CRSN clustering routing protocol. We use packet delivery ratio in each round and average packet delivery ratio to measure the information collection capability of the network. Here, packet delivery ratio in each round is defined as the ratio between the number of packets received successfully at the sink in each round and total number of packets supposed to be delivered in corresponding round. We count the number of packets received at the sink in each round and sum them up, then we obtain the total number of packets received at the sink *TotalReceived*. We also count the number of packets supposed to be sent by all living nodes in each round and sum them up, and we obtain the total number of packets supposed to be transmitted *TotalTrans*. Then the average packet delivery ratio is defined as the ratio between total number of packets received at the sink and total number of packets supposed to be sent, as shown below:

$$APDR = \frac{TotalReceived}{TotalTrans} \tag{22}$$

Packet delivery ratio in each round shows the variation of packet delivery ratio as the network operation moves forward, and average packet delivery ratio shows the comparing results more intuitively. The results of different protocols with the sink located at the center and (0, 0) are shown in Fig.14 to Fig.17, respectively.

As can be seen from Fig.14 - Fig.17, the packet delivery ratio of IMPORP in each round is usually above 0.8, and the average packet delivery ratio is above 0.96, which are significantly higher than other protocols. The reasons are as follows:

Firstly, unlike random channel selection methods used by Fuzzy C-means and ABCC, IMOCRP considers available channel conditions of nodes during cluster construction. This can help avoid transmission failures due to cluster channel being occupied by PUs at the beginning of data transmission. When data transmission fails, the node energy consumption is less than normal communication, which are shown by the first part and second part of (23), respectively.

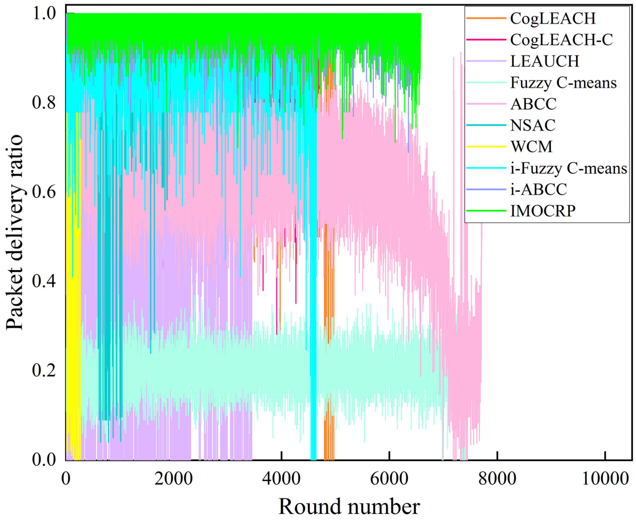

FIGURE 14. Statistics of packet delivery ratio in each round (with the sink at the center).

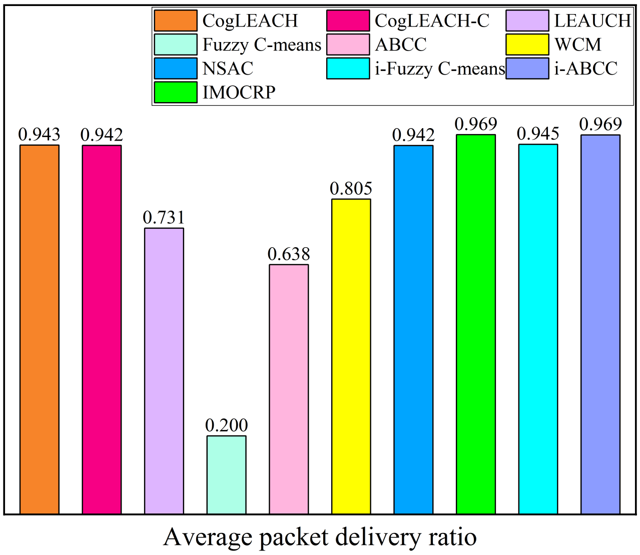

FIGURE 15. Average packet delivery ratio (with the sink at the center).

$$E_{rxag} = \begin{cases} 0 & if\ \text{CM(channel)} \\ & \cap\ \text{CH(channel)} = \emptyset \\ (E_{elec} + E_{DA}) \times l & if\ \text{CM(channel)} \\ & \cap\ \text{CH(channel)} \neq \emptyset \\ & \cap\ no\ collision\ occurs \end{cases} \tag{23}$$

So, the information collection capability of Fuzzy C-means and ABCC is weak, which makes nodes survive longer. This explains why the round of first and last node death under Fuzzy C-means and ABCC in Fig. 8 and Fig. 9 are higher than IMOCRP. However, CRSN should possess the network information collection capabilities required by applications,

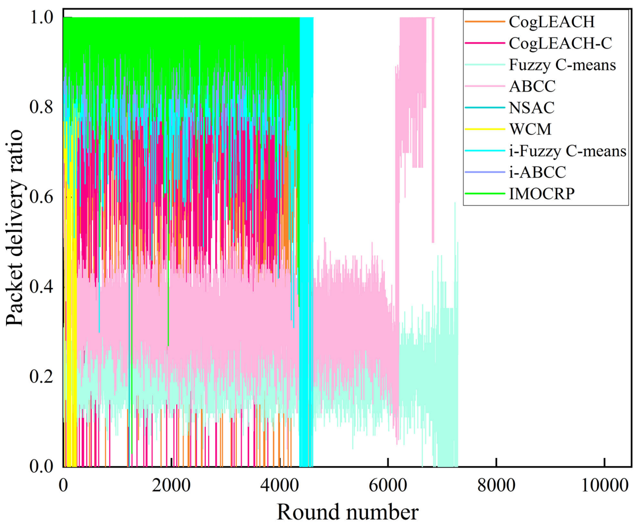

FIGURE 16. Statistics of packet delivery ratio in each round (with the sink located at (0, 0)).

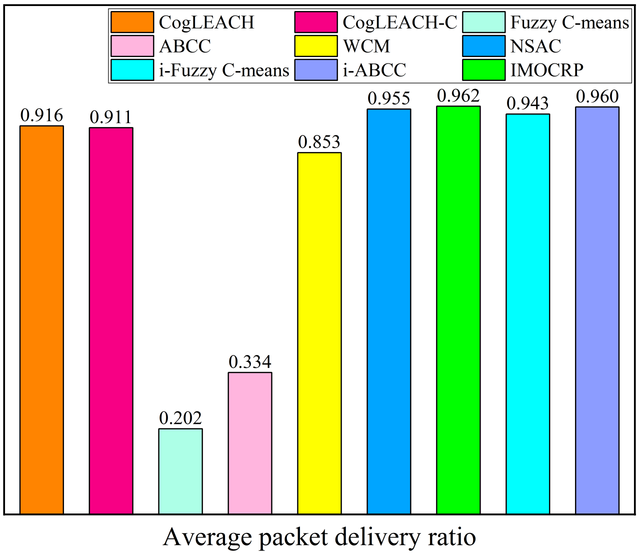

FIGURE 17. Average packet delivery ratio (with the sink located at (0, 0)).

so IMOCRP which takes network lifetime and effective information collection into account is more suitable for CRSN.

Secondly, in 100m × 100m small-scale network, compared with CogLEACH, CogLEACH-C, and other protocols which also consider about the channel allocation problem, IMOCRP builds many small clusters, and the probability of nodes in the cluster being outside interference protection ranges of PUs is higher. This further reduces data transmission failures due to channel reclaim of PUs during data transmissions, so the packet delivery ratio of IMOCRP is higher.

D. INFLUENCE FROM PUs AND IMPACT ON PUs

In order to test the capability of IMOCRP to adapt to dynamic spectrum conditions, we changed the probability of PUs in ON state p_i from 0 to 0.9 in CRSN with the sink located at the center, and the simulation results are shown below:

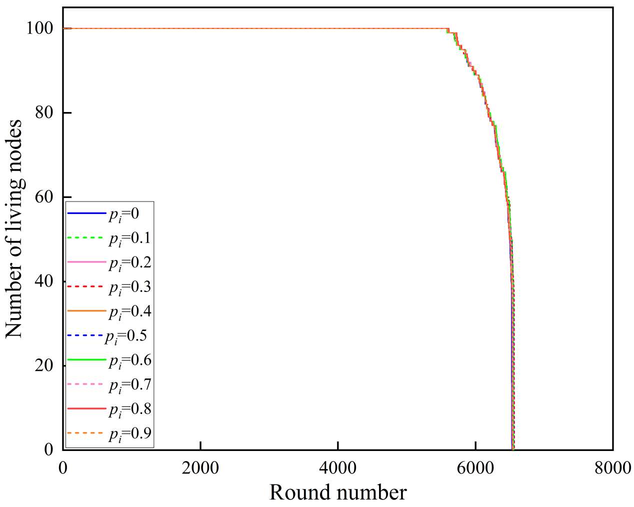

FIGURE 18. Number of living nodes under different ON states of PUs p_i (with the sink located at the center).

From Fig.18, we can see that under different ON states of PUs, IMOCRP can achieve nearly equal performance in terms of number of living nodes, i.e., the first dead node appears in approximately round 5600 and the last node dies after about 1000 rounds. It demonstrates that IMOCRP can balance the average energy consumption and deviation of residual energy among nodes, which reduces the duration of node deaths. Apart from network lifetime, we also provide the simulation results in terms of average packet delivery ratio, as shown in Fig.19.

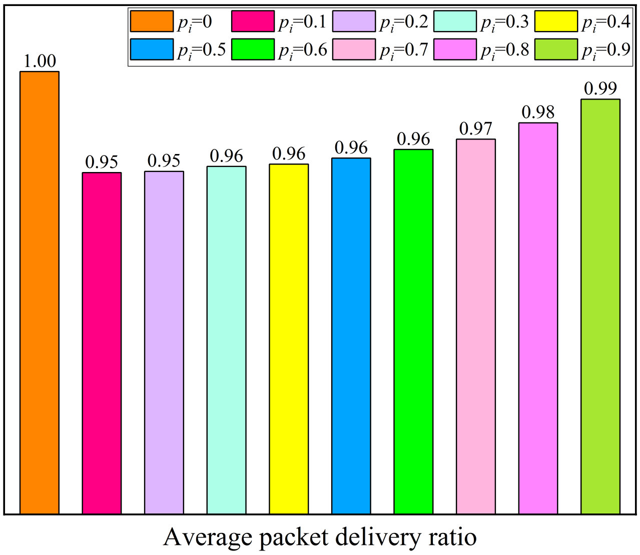

FIGURE 19. Average packet delivery ratio under different ON states of PUs p_i (with the sink located at the center).

Fig.19 shows the variations of average packet delivery ratio under different ON states of PUs. From the figure, we can see that all packets can be transmitted successfully if ON state of PUs is 0. The reason is: when $p_i = 0$, CRSN nodes can freely utilize all licensed channels, and there will be none collisions between CRSN nodes and PUs. Therefore, IMOCRP can achieve 100 percent average packet delivery ratio. In IMOCRP, the sink will select the licensed channel with the largest number of covered CMs as cluster channel. If multiple channels have the same largest number of covered CMs, the sink will randomly select one from them to break the tie. From the clustering results of IMOCRP, we can see that many CRSN nodes become separate CHs without any CM, especially when the sink is located at the center. In this case, the sink will randomly select a licensed channel from the available channel list of the CH. Higher p_i means more spectrum occupancy, which can be sensed by CRSN nodes in spectrum sensing phase (i.e., Phase 1). If a licensed channel is sensed as busy in Phase 1, it is thought to be busy in Phase 3 by IMOCRP. Therefore, higher p_i can help CHs select the channel which will not be occupied by any PU with higher probability. From the average point of view, this decreases the number of collisions to PUs. This is why the average packet delivery ratio of IMOCRP increases when p_i varies from 0.1 to 0.9. The above simulation results demonstrate that IMOCRP is capable of adapting to the dynamic spectrum environment.

As stated above, when channel sensed as idle at sensing moment is used by CRSN nodes to transmit data to CHs or to the sink, PUs may reappear on this channel. CRSN node cannot discover the change in channel occupancy state immediately because it cannot sense the spectrum simultaneously when it is performing data transmission. Therefore, it will not vacate the channel immediately, and in this case, collision occurs. In order to test the impact of individual clustering protocol on PUs, we count the number of collisions occurring between all living CRSN nodes and all PUs in each round, and one collision means that the packet delivery of PUs during SUs occupancy is affected. If they appear on the same channel simultaneously for communication, the number of collisions increases by 1. The simulation results for different sink locations are shown in Fig.20 and Fig.21, respectively.

From these figures, we can see that the number of collisions with PUs in IMOCRP is usually lower than other competing protocols. Actually, the normalized number of collisions with PUs is complementary with packet delivery ratio of CRSN nodes. More collisions with PUs mean more transmission failures and the packet delivery ratio of CRSN nodes naturally reduces. However, the number of collisions with PUs in Fuzzy C-means and ABCC is also low in the two simulation scenarios, and the reason is that channel selection and usage is not taken into consideration in Fuzzy C-means and ABCC. Therefore, CRSN node randomly selects one channel as its available channel, and this randomness together with the random channel occupancy of PUs produces the above results.

E. TIME COMPLEXITY OF IMOCRP

Both IMOCRP and ABCC are swarm intelligence-based algorithms. In addition to comparing network lifetime and information collection capability, we also compare their time complexity. The time complexity analysis of IMOCRP and ABCC in single iteration is summarized in Table 7.

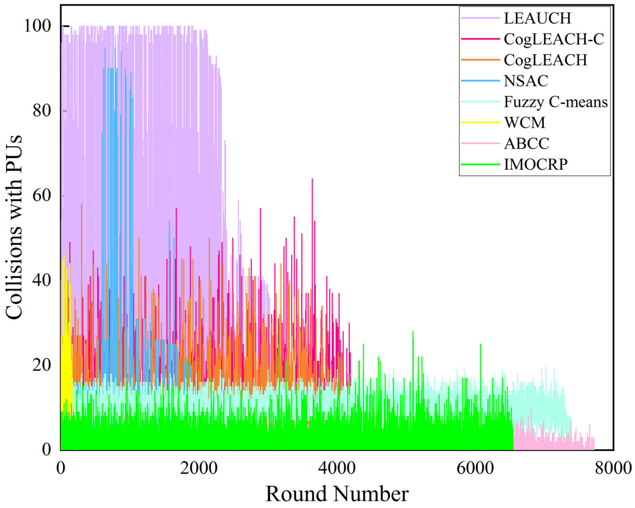

FIGURE 20. Number of collisions with PUs for each clustering protocol (with the sink located at the center).

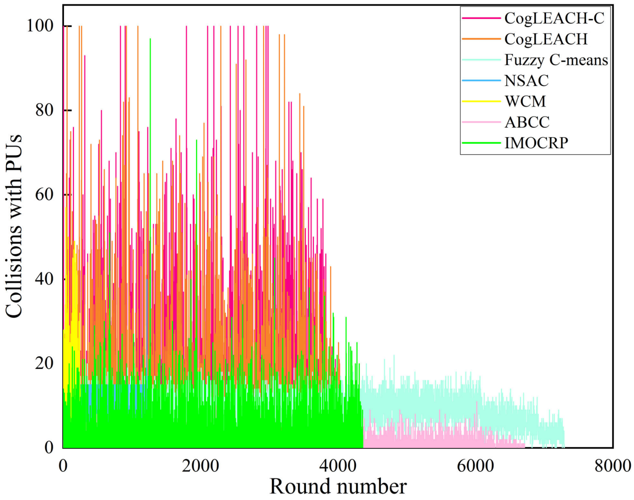

FIGURE 21. Number of collisions with PUs for each clustering protocol (with the sink located at (0, 0)).

TABLE 7. Time complexity analysis.

Protocol	Steps	Time complexity
IMOCRP	Position update in liquid phase	$O(IN \times N)$
	Position update in crystal phase	$O(IN \times N)$
	Fitness function evaluation	$O(N^2)$
	Single liquid and crystal iteration	$O(N^2)$
ABCC	Employed bee phase	$O(SN \times N^2)$
	Onlooker bee phase	$O(SN \times N^2)$
	Single iteration	$O(SN \times N^2)$

From above table, we can see that the time complexity in single liquid and crystal iteration of IMOCRP is $O(N^2)$, and the time complexity in single iteration of ABCC is

$O(SN \times N^2)$. In other words, the time complexity of single iteration of IMOCRP is lower than that of ABCC. The maximum number of iterations set by ABCC is *MaxCycle*, and *SN* employed bees and *SN* onlooker bees search around food sources (*SN* is the number of food sources) in each iteration. Additionally, there is also extra time complexity if there is no improvement in consecutive *Limit* trials. For IMOCRP, in the worst case, the liquid to crystal iteration is executed $2N$ times, and our simulation experiments show that in most cases the number of liquid to crystal iterations is far smaller than $2N$. Therefore, the total time complexity of ABCC is much higher than that of IMOCRP.

In summary, IMOCRP reduces the energy consumption during CH selection and cluster construction, and it can automatically determine the optimal number of clusters in the network, thereby extending network lifetime. For information collection capability, IMOCRP increases the package delivery ratio through reasonable allocation of channels. At the same time, IMOCRP is less affected by the status of PUs and has a lower probability of conflict with PUs than other competing protocols. Compared with ABCC, IMOCRP has lower total time complexity which makes it easier to implement in practical CRSN.

VII. CONCLUSION AND FUTURE WORK

In this paper, we design a new centralized clustering routing protocol IMOCRP for CRSN. It aims at minimizing the weighted sum of average energy consumption of nodes and standard deviation of node remaining energy. If the time till the first node death is defined as network lifetime, when the sink is located at the center, IMOCRP can achieve 72 percent and 53 percent longer lifetime than CogLEACH and CogLEACH-C, respectively. By reasonably allocating available channels for clusters, IMOCRP reduces the probability of collisions with PUs and promotes successful information transmissions. Therefore, its packet delivery ratio is usually above 0.8, while Fuzzy C-means can only achieve about 0.2. We also change the probability of PUs in ON states to test the adaptability of IMOCRP to dynamic spectrum environment. With current network configurations, we can observe through simulations that the network lifetime and packet delivery ratio of IMOCRP are basically kept unchanged even under very high PUs activity, such as $p_i = 0.9$. Additionally, the impact on PUs is also evaluated by number of collisions with PUs. We can observe that IMOCRP can achieve fewer collisions with PUs thanks to its reasonable channel allocation and small cluster size. In order to simplify protocol design, this paper assumes that each CRSN node sends the remaining energy, available channel list, and other information to the sink through single-hop communication. In future work, we will further study multihop information transmissions between CRSN nodes and the sink, thereby further expanding the application scope of our IMOCRP. Additionally, the current version of IMOCRP is suitable for time-triggered applications, therefore, we will improve it to make it work well in both time-triggered applications and event-driven applications in our future work.

REFERENCES

[1] M. Zheng, C. Wang, M. Du, L. Chen, W. Liang, and H. Yu, "A short preamble cognitive MAC protocol in cognitive radio sensor networks," *IEEE Sensors J.*, vol. 19, no. 15, pp. 6530–6538, Aug. 2019.

[2] S. H. R. Bukhari, M. H. Rehmani, and S. Siraj, "A survey of channel banding for wireless networks and guidelines of channel bonding for futuristic cognitive radio sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 924–948, 2nd Quart., 2016.

[3] M. M. A. Osman, S. K. Syed-Yusof, N. N. N. A. Malik, and S. Zubair, "A survey of clustering algorithms for cognitive radio ad hoc networks," *Wireless Netw.*, vol. 24, no. 5, pp. 1451–1475, Jul. 2018.

[4] N. U. Hasan, H. S. Kim, W. Ejaz, and S. Lee, "Knapsack-based energy-efficient node selection scheme for cooperative spectrum sensing in cognitive radio sensor networks," *IET Commun.*, vol. 6, no. 17, pp. 2998–3005, Nov. 2012.

[5] G. P. Joshi and S. W. Kim, "A survey on node clustering in cognitive radio wireless sensor networks," *Sensors*, vol. 16, no. 9, pp. 1–19, Jun. 2016.

[6] E. Helal, A. Khattab, and Y. A. Fahmy, "Energy-efficient cluster head selection for cognitive radio sensor networks," in *Proc. 28th Int. Conf. Microelectron. (ICM)*, Giza, Egypt, Dec. 2016, pp. 205–208.

[7] H. Zhang, Z. Zhang, H. Dai, R. Yin, and X. Chen, "Distributed spectrum-aware clustering in cognitive radio sensor networks," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Houston, MA, USA, Dec. 2011, pp. 1–6.

[8] I. Mustapha, B. Ali, M. Rasid, A. Sali, and H. Mohamad, "An energy-efficient spectrum-aware reinforcement learning-based clustering algorithm for cognitive radio sensor networks," *Sensors*, vol. 15, no. 8, pp. 19783–19818, Aug. 2015.

[9] J. Ren, Y. Zhang, N. Zhang, D. Zhang, and X. Shen, "Dynamic channel access to improve energy efficiency in cognitive radio sensor networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3143–3156, May 2016.

[10] B. Javidy, A. Hatamlou, and S. Mirjalili, "Ions motion algorithm for solving optimization problems," *Appl. Soft Comput.*, vol. 32, pp. 72–79, Jul. 2015.

[11] C.-H. Yang, K.-C. Wu, Y.-S. Lin, L.-Y. Chuang, and H.-W. Chang, "Protein folding prediction in the HP model using ions motion optimization with a greedy algorithm," *BioData Mining*, vol. 11, no. 1, pp. 1–14, Aug. 2018.

[12] L. Özbakır and F. Turna, "Clustering performance comparison of new generation meta-heuristic algorithms," *Knowl.-Based Syst.*, vol. 130, pp. 1–16, Aug. 2017.

[13] S. Ebrahimi Mood and M. M. Javidi, "Energy-efficient clustering method for wireless sensor networks using modified gravitational search algorithm," *Evolving Syst.*, vol. 11, no. 4, pp. 575–587, Dec. 2020, doi: 10.1007/s12530-019-09264-x.

[14] N. Srikanth and M. S. G. Prasad, "Efficient energy clustering protocol using genetic algorithm in wireless sensor networks," *J. Eng. Sci. Technol. Rev.*, vol. 11, no. 6, pp. 85–93, Dec. 2018.

[15] P. Kanchan and D. P. Shetty, "Quantum PSO algorithm for clustering in wireless sensor networks to improve network lifetime," *Adv. Intell. Syst. Comput.*, vol. 814, pp. 699–713, 2019.

[16] J. John and P. Rodrigues, "MOTCO: Multi-objective Taylor crow optimization algorithm for cluster head selection in energy aware wireless sensor network," *Mobile Netw. Appl.*, vol. 24, no. 5, pp. 1509–1525, Oct. 2019.

[17] B. Pitchaimanickam and G. Murugaboopathi, "A hybrid firefly algorithm with particle swarm optimization for energy efficient optimal cluster head selection in wireless sensor networks," *Neural Comput. Appl.*, vol. 32, no. 12, pp. 7709–7723, Jun. 2020.

[18] S. Panda, "Performance improvement of clustered wireless sensor networks using swarm based algorithm," *Wireless Pers. Commun.*, vol. 103, no. 3, pp. 2657–2678, Dec. 2018.

[19] R. M. Eletreby, H. M. Elsayed, and M. M. Khairy, "CogLEACH: A spectrum aware clustering protocol for cognitive radio sensor networks," in *Proc. 2014 9th Int. Conf. Cogn. Radio Oriented Wireless Netw. Commun. (CROWNCOM)*, Oulu, Finland, Jun. 2014, pp. 179–184.

[20] A. Latiwesh and D. Qiu, "Energy efficient spectrum aware clustering for cognitive sensor networks: CogLEACH-C," in *Proc. 10th Int. Conf. Commun. Netw. China (ChinaCom)*, Shanghai, China, Aug. 2015, pp. 515–520.

[21] D. M. S. Bhatti, N. Saeed, and H. Nam, "Fuzzy C-means clustering and energy efficient cluster head selection for cooperative sensor network," *Sensors*, vol. 16, no. 9, pp. 1–17, Sep. 2016.

[22] E. Pei, H. Han, Z. Sun, B. Shen, and T. Zhang, "LEAUCH: Low-energy adaptive uneven clustering hierarchy for cognitive radio sensor network," *EURASIP J. Wireless Commun. Netw.*, vol. 2015, no. 1, pp. 1–8, Apr. 2015.

[23] C. Q. Wang and S. B. Wang, "Research on uneven clustering APTEEN in CWSN based on ant colony algorithm," *IEEE Access*, vol. 7, pp. 163654–163664, Nov. 2019.

[24] T. Stephan, F. Al-Turjman, S. Joseph K., and B. Balusamy, "Energy and spectrum aware unequal clustering with deep learning based primary user classification in cognitive radio sensor networks," *Int. J. Mach. Learn. Cybern.*, early access, Jun. 2020, doi: 10.1007/s13042-020-01154-y.

[25] T. Wang, X. Guan, X. Wan, H. Shen, and X. Zhu, "A spectrum-aware clustering algorithm based on weighted clustering metric in cognitive radio sensor networks," *IEEE Access*, vol. 7, pp. 109555–109565, Aug. 2019.

[26] M. Zheng, S. Chen, W. Liang, and M. Song, "NSAC: A novel clustering protocol in cognitive radio sensor networks for Internet of Things," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5864–5865, Jun. 2019.

[27] E. Pei, J. Pei, S. Liu, W. Cheng, Y. Li, and Z. Zhang, "A heterogeneous nodes-based low energy adaptive clustering hierarchy in cognitive radio sensor network," *IEEE Access*, vol. 7, pp. 132010–132026, Sep. 2019.

[28] V. Srividhya and T. Shankar, "Energy proficient clustering technique for lifetime enhancement of cognitive radio-based heterogeneous wireless sensor network," *Int. J. Distrib. Sensor Netw.*, vol. 14, no. 3, pp. 1–18, Mar. 2018.

[29] R. N. Yadav, R. Misra, and D. Saini, "Energy aware cluster based routing protocol over distributed cognitive radio sensor network," *Comput. Commun.*, vol. 129, pp. 54–66, Sep. 2018.

[30] S. Thompson, A. T. Fadi, J. K. Suresh, B. Balamurugan, and S. Srivastava, "Artificial intelligence inspired energy and spectrum aware cluster based routing protocol for cognitive radio sensor networks," *J. Parallel Distrib. Comput.*, vol. 142, pp. 90–105, Aug. 2020.

[31] S.-S. Kim, S. McLoone, J.-H. Byeon, S. Lee, and H. Liu, "Cognitively inspired artificial bee colony clustering for cognitive wireless sensor networks," *Cognit. Comput.*, vol. 9, no. 2, pp. 207–224, Apr. 2017.

[32] J. H. Wang and W. X. Shi, "Survey on cluster based routing protocols for cognitive radio sensor networks," *J. Commun.*, vol. 39, no. 11, pp. 156–169, Nov. 2018.

[33] O. M. Al-Kofahi, H. M. Almasaeid, and H. Al-Mefleh, "Efficient on-demand spectrum sensing in sensor-aided cognitive radio networks," *Comput. Commun.*, vol. 156, pp. 11–24, Apr. 2020.

[34] R. N. Yadav and R. Misra, "Approximating common control channel problem in cognitive radio networks," *IEEE Syst. J.*, vol. 13, no. 1, pp. 301–312, Mar. 2019.

[35] S. Soro and W. B. Heinzelman, "Prolonging the lifetime of wireless sensor networks via unequal clustering," in *Proc. 19th IEEE Int. Parallel Distrib. Process. Symp.*, Denver, CO, USA, Apr. 2005, pp. 1–8.

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

[37] G. Smaragdakis, I. Matta, and A. Bestavros, "SEP: A stable election protocol for clustered heterogeneous wireless sensor networks," in *Proc. 2nd Int. Workshop Sensor Actor Netw. Protocols Appl. (SANPA)*, Boston, MA, USA, Apr. 2004, pp. 1–11.

[38] Y. J. Wang and C. L. Ma, "Dynamic ion motion optimization algorithm based on memory strategy," *J. Jilin Univ., Eng. Technol. Ed.*, vol. 50, no. 3, pp. 1047–1060, May 2020.

JIHONG WANG (Member, IEEE) received the B.S. degree in communication engineering from Jilin University, Changchun, Jilin, in 2011, and the Ph.D. degree in communication and information systems from the School of Communication Engineering, Jilin University, in 2016. She is currently an Associate Professor and a Master Supervisor with Northeast Electric Power University. Her research interests include cognitive radio sensor networks, routing and resource allocation in wireless networks, energy harvesting, and so on.

SHUO LI was born in Changchun, Jilin, China, in 1995. He received the bachelor's degree in communication engineering from Northeast Electric Power University, China, in 2018, and the master's degree from Northeast Electric Power University. His research interests include clustering and routing in cognitive radio sensor networks.

YIYANG GE was born in Changchun, Jilin, China, in 1996. She received the B.S. degree in electronic information science and technology from Northeast Electric Power University, China, in 2019. Her research interests include cluster routing protocol design and resource allocation in cognitive radio wireless sensor networks with energy harvesting.