IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received July 6, 2020, accepted July 26, 2020, date of publication August 3, 2020, date of current version August 13, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3013644

A Game Theory Based Clustering Protocol to Support Multicast Routing in Cognitive Radio Mobile Ad Hoc Networks

THONG NHAT TRAN[1], TOAN-VAN NGUYEN[1], (Student Member, IEEE), KYUSUNG SHIM[1], (Student Member, IEEE), AND BEONGKU AN[2], (Member, IEEE)

[1]Department of Electronics and Computer Engineering, Graduate School, Hongik University, Sejong 30016, South Korea
[2]Department of Software and Communications Engineering, Hongik University, Sejong 30016, South Korea

Corresponding author: Beongku An (beongku@hongik.ac.kr)

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2019R1A2C1083996).

ABSTRACT The frequent movement of nodes in cognitive radio mobile ad hoc networks (CRAHNs) causes challenges in scalability, stability, channel sensing and channel access problems that can be solved by using clustering technique. Game theory is a feasible approach to solve such problems by casting clustering problems as distributed optimization problems. The main contributions of this article are as follows. Firstly, we propose a minimum connected weighted inner edge spanning tree (MWIEST) game to find an approximate solution of a MWIEST problem in CRAHNs. In this game, a link-weight function of each link is designed based on a combination of link-stability and link-connectivity ratio functions. Secondly, we prove that the MWIEST game is an exact potential game that exists at least one Nash equilibrium (NE) point which is an approximate solution of the MWIEST problem. Besides, we also prove that best responses (BRs) of the game converge to a NE in finite iterations. Thirdly, based on the MWIEST game, we propose four algorithms including the node information exchange (NIE), the best response selection (BRS), the intermediate nodes selection (INS) and the forming cluster (FC). Specifically, the algorithms NIE, BRS and INS provide a set of intermediate nodes (SetIN) which supports the FC algorithm to form clusters. Finally, we propose the game theory based clustering (GBC) protocol which is combination of the FC algorithm and the proposed cluster maintenance algorithm to construct high stable clusters supporting multicast routing in CRAHNs. Moreover, each obtained cluster includes most members having the same receiving channel which avoids the affected regions of licensed channels. For the performance evaluation, we implement the GBC protocol in OMNET++ platform to demonstrate its performance improvement over the state-of-the-art protocols in terms of network stability and control overheads.

INDEX TERMS Cognitive radio mobile ad hoc network, clustering, multicast routing, game theory, network layer, OMNET++.

I. INTRODUCTION

The rapid growth of wireless communication networks has led to an increasing demand for spectrum, and continue to do so in future. Cognitive radio (CR) technology has been studied to solve overloaded spectrum problems. CR technology allows a wireless device to access spectrum in a dynamic opportunistic way. In CR networks, there are two kinds of users that are primary users (PUs) and secondary users (SUs). PUs can access licensed channels whenever they need to transmit data while SUs can only access licensed channels when these are unoccupied by PUs [1], [2]. A CR network can be classified into a centralized network (including base stations and wireless users) or a distributed network, e.g., mobile ad hoc network (MANET). In a distributed network, each user can directly communicate with the others without any existence of a fixed network infrastructure such as base stations or access points [3], [4]. With a rise of mobile devices, MANETs have been receiving attention with an increasing number of applications in environmental monitoring, health

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

care, commercial, military, private sectors [4], [6]. Moreover, because a MANET does not require an infrastructure network, this provides a convenient decentralized feature and makes networks more flexible and robust. In this article, we focus on the clustering issue in cognitive radio mobile ad hoc networks (CRAHNs). There are some factors which affect to topology management in CRAHNs such as (1) there is no common channel to transmit data throughout the network, (2) topology of a network is constantly changing over time according to user activities, and (3) it takes a long time to spread routing information in whole large size network [4]. Besides, multicast routing is also considered in CRAHNs with many applications and services such as video conferencing, distance learning, cooperative work, replicated database updating and video on-demand, etc [8]. In multicast routing, each multicast group presents a group of users with common characteristics. In clustering process, cluster heads are selected to receive messages from source and forward messages to multicast groups in order to reduce control overheads and packet loss. Hence, clustering is considered as an effective technique to classify nodes into clusters for topology control. In other words, it builds a virtual backbone to ensure network system performance and make network size smaller and easier to control. Moreover, clustering is used to improve network performances such as delay, bandwidth consumption and throughput [9].

In this article, although there exist important issues of lower layers which can be occured at network layer such as throughput, energy efficiency, traffic optimization, packet loss, security, delay, QoS, admission control, mobility, routing and data gathering, we only focus on clustering issue in network layer to obtain high stable clusters. Thus, we assume that these above issues were fully supported by lower layers that are physical and data link layers.

A. RELATED WORKS AND MOTIVATIONS

Clustering algorithms in CRAHNs can be classified into different categories [4], [5], [10] such as (1) Dominating-set-based clustering [11]–[13], (2) Combined-metrics-based clustering [14], [15], (3) Common-control-channel-establishment-based clustering [16], [17], (4) Stability-based clustering [8], [18], (5) Energy-efficient clustering [19], [20] and (6) Spectrum-sense-improvement-based clustering [21]–[23]. Particularly, (1) The idea of dominating-set-based clustering is to find a minimum dominating set of CRAHNs which is used as a minimum routing space to optimize system performance in terms of control overheads, packet delivery ratio, delay, etc. In [11], the authors proposed a connected dominating set (CDS) based distributed algorithm in MANET. In this algorithm, if a node has two neighbor nodes which are not connected, it will become a member of the CDS. Next, some rules were proposed to reduce the CDS size. The obtained CDS was used as an optimal virtual backbone of network to reduce the routing space. In [12], a CDS based routing protocol in MANET was proposed to obtain a stable routing space and prolong the lifetime of network. In this protocol, the authors proposed a status function of a node which integrates three factors such as energy, mobility, and degree to construct a minimum connected dominating set (SoN-MCDS). In [13], the authors proposed an efficient CDS clustering based routing protocol to construct stable clusters and improve system performances such as control overheads, packet delivery ratio and delay. In this protocol, a weighted function and some rules were proposed to construct a minimum CDS which was used for clustering and routing processes. (2) The main goal of combined-metrics-based clustering is to propose a weighted function which is a combination of parameters such as node degree (connectivity of node), residual energy capacity, moving direction, etc. This clustering approach uses a weighted function for cluster head selection process to construct a cluster structure which supports networks to improve the system performances. In [14], a weighted clustering algorithm was studied, called enhancement on weighted clustering algorithm (EWCA). In EWCA, transmission power, transmission range, mobility and battery energy parameters were used to select cluster heads. This achieves a load balancing to enhance the stability of clusters in MANETs. In [15], the authors proposed a method of cluster formation which used two factors that are node degree and bandwidth to select cluster heads and construct clusters. They also proposed a new mechanism of merging two clusters to make clusters more stable and minimize packet loss. (3) Common-control-channel-establishment-based clustering approaches are applied to determine a list of channels in which each channel is used for all members of a cluster to exchange data in cooperative spectrum sensing, broadcasting routing information and spectrum access coordination. In [16], a distributed coordination protocol was considered to build groups according to spectrum heterogeneity in the CR network. This protocol provided large-group size and each group has one common control channel. In [17], the authors proposed two distributed clustering algorithms that are spectrum-opportunity clustering (SOC) and constrained-SOC (C-SOC). In these algorithms, the clustering was formulated as a maximum edge biclique construction problem, the SOC algorithm made a balance between the cluster size and the number of common channels, and the C-SOC algorithm constructs clusters with maximum cluster size. The C-SOC also provides the number of common idle channels in a cluster equal to or greater than a predefined value. (4) In stability-based clustering, the topology of CRAHNs changes rapidly, nodes need to exchange additional messages to repair clusters. This makes a significant increase in control overheads. Thus, main goal of the stability-based clustering is to provide stable clusters to reduce re-affiliation rate and re-clustering. In [8], the authors proposed a mobility-based clustering (MBC) approach that uses combination of both physical and logical partitions of network to support mobility management and multicast routing in MANET. The MBC protocol provided clustered topology which is more stable than other approaches. In [18], the authors presented their studies as follows. Firstly, a realistic mobility model was proposed to describe movement of highly mobile airborne nodes. Secondly, a cluster head selection algorithm was proposed based on node degree level, average number of hops and channel switching from member nodes to cluster heads. Thirdly, two new common control channel (CCC) selection schemes were proposed based on node contraction concept and the discrete particle swarm optimization algorithm. Finally, a routing protocol was proposed with a channel assignment scheme based on node capacity to improve total throughput of control channels, end-to-end delay and packet delivery ratio. (5) In energy-efficient clustering, the aim is to reduce unnecessary energy consumption or balance energy consumption to prolong network lifetime by reducing transmission power, distance between members and cluster heads. In [19], the authors proposed a spectrum-aware cluster-based energy efficient multimedia routing (SCEEM) protocol for cognitive radio sensor network (CRSN) to overcome formidable limitations of energy and spectrum issues. The SCEEM protocol was applied to improve the quality of service (QoS) and energy efficient routing by limiting the participating nodes in route establishment. In [20], EDCRN (Energy-Driven Cognitive Radio Network) protocol was proposed in which the rotation energy threshold was estimated by using cluster head node real time energy load. The authors also showed that the network lifetime of EDCRN is better than of LEACH (Low-Energy Adaptive Clustering Hierarchy) and EDAC. (6) In spectrum sense improvement-based clustering, spectrum sense is a process of sensing activities of primary users to detect idle channels. Because a cognitive user can not make accurate decisions, a cooperative spectrum sensing has been introduced to increase the sensing results. In traditional cooperative spectrum sensing, each user sends its sensing results to the coherence center by using a reporting channel. This causes channel congestion problem when there are a large number of reports. To solve this issue, clustering is used to reduce number of reports to center by reporting through cluster heads [21], [22]. In [21], A cluster-based CSS was introduced to obtain a proper assignment policy and maximize achievable throughput for secondary users. In the obtained policy, all secondary users in each cluster cooperate to sense the same set of PU channels. In [23], the authors proposed a cluster-based cooperative spectrum sensing scheme. The optimal number of clusters was obtained by balancing tradeoff between communication overhead and sensing reliability. This can minimize cooperation overhead without any performance loss of reliability.

From these observations, main advantages of the clustering in CRAHNs can be listed as follows. (i) Channel assignment: adjacent SUs having similar spectrum are grouped in the same cluster and receiving channels of clusters are not affected by licensed channels. (ii) Enhancement of sensing outcomes: SUs in the same cluster cooperate to decide which channels are idle. This can reduce probability of miss-detection and incorrect alarm to network [30]. (iii) Enhances the network stability: a stable clustered structure makes CRAHNs easier to control, reduces the re-affiliation rate and minimizes re-clustering situations [10]. (iv) Network scalability: the flat structure restricts the extension of network because the control overheads is increased. Hence, a clustered structure is necessary to simplify routing and reduce the control overheads because cluster heads and gateways establish a virtual backbone for inter-cluster communication to reduce generation information of routing process [10], [31]. (v) Multicast Routing: Clustering approach selects cluster heads which are used to receive messages from source and forward the messages to multicast groups in order to reduce control overheads and packet loss [8].

Game theory is a collection of mathematical models of strategic interaction between rational players, which has a wide range of useful applications in economics, sociology and psychology, political science, biology social science and computer science. A modern game theory was developed by Morgenstern and von Neumann [24] with the idea of mixed-strategy equilibria in two-person zero-sum games. In the 1950s, game theory was studied by many researchers, at which time, John Nash developed a criterion for mutual consistency of players' strategies known as the Nash equilibrium of the game and it was applied to many different games [25]. Recently, game theory has been applied to network problems because nodes in a distributed network are independent, they have to make decisions by their own interest, and game theory adapts with distributed optimization problems [26]. In [27], the authors modeled clustering problem in MANET as a coalition game framework, and proposed a distributed generic coalition formation algorithm that operated clustering process. This generic algorithm was compatible with unstructured and structured cases. Simulation results indicated that the proposed algorithm outperforms others in terms of cluster size and stability. In [28], an evolutionary game theoretic (EGT) framework was proposed to solve a problem of cluster in-stability in vehicular ad hoc networks (VANETs). The EGT was used to reduce signaling overhead, complexity and increase cluster stability in large scale VANETs. In [29], the authors proposed an energy-efficient clustering algorithm based on game theory. In this algorithm, each node competes to be a cluster head by joining a localized clustering game in order to achieve equilibrium probability. A potential cluster head is selected to be a real cluster head through a properly designed probability method. The authors also indicated that the proposed algorithm was better than low-energy adaptive clustering hierarchy and clustered routing in terms of network lifetime. In [32], the authors proposed an energy-efficient clustering algorithm which was a combination of non-cooperative game and dual-cluster-head (ECGD) to improve energy efficiency and extend the network lifetime in WSNs. In [33], a coalitional game-theoretic clustering (CGTC) algorithm was proposed in WSNs to reduce energy consumption and improve throughput. Particularly, the entire network area was divided into two regions that are an area far from the base station (far area) and a region close to the base station (vicinity area). Sensor nodes in the vicinity area were grouped into small coalitions by using coalitional game to reduce the energy consumption. In the far area, cluster heads were elected based on the set of nodes with the highest residual energy, and clusters are formed by using coalitional game. In [34], the authors investigated a new clustering approach by mixing a non-cooperative game theory technique with a decentralized clustering algorithm in WSNs to maximize the network lifetime. Specifically, the game theory approach was used to limit the number of the forwarding messages as well as to maximize the lifetime of sensors. In [35], the authors designed a clustering protocol based on game theoretic approach for balancing energy consumption in WSNs to improve the network lifetime.

Because nodes in the network randomly move based on random models, some critical problems can be occurred as follows: (i) a member node moves out of its cluster, (ii) a cluster head moves out of its cluster or dies and (iii) multiple clusters move into the same cluster. To solve problems (i)–(iii), local cluster maintenance algorithms were proposed in [36]–[39]. The authors in [36] proposed an improved cluster maintenance scheme (ICMS) in MANETs to improve the stability of clusters and reduce the control overheads. Specifically, the ICMS delays the cluster head change if two cluster heads are in the same cluster area. The new cluster head is elected based on the cluster priority and delay time. In [37], the authors proposed an optimized stable clustering algorithm for MANETs to improve the stability of clusters and reduce the control overheads. This algorithm considered new nodes as backup nodes in the cluster. A backup node acts as a cluster head if the cluster head moves out (or dies) of the cluster. Next, the cluster head chooses a new backup node, which keeps the network available without disturbance. Furthermore, the priority of the cluster heads and the backup nodes is calculated based on the remaining battery life of mobile nodes. In [38], the authors analyzed the performance of several clustering algorithms and discussed about stability and control overheads of network. In [39], a least cluster change (LCC) clustering algorithm was proposed to solve the following problems for MANETs: one is when a non-cluster head moves out of its cluster, and the other is when two cluster heads come within range of each other.

Most recent studies on applying game theory for clustering have restricted to the energy consumption optimization problem in WSNs and MANETs to prolong the network lifetime. However, one of the most important clustering issues in MANET is to construct a highly stable structure of clusters, and game theory is an effective approach to address these clustering issues in CRAHNs. Hence, we apply game theory to propose a novel clustering protocol in order to obtain the high stable clusters supporting multicast routing in CRAHNs.

B. CONTRIBUTION AND ORGANIZATION

The main contributions of this article are summarized as follows:

- Firstly, we propose a minimum connected weighted inner edge spanning tree (MWIEST) game to solve the MWIEST problem in CRAHNs. The solution of this game is used to find a set of intermediate nodes (SetIN) which supports the clustering process to choose cluster heads and gateways from the SetIN. In this game, a link-weight function of each edge is designed based on a combination of link-stability and link-connectivity ratio functions.
- Secondly, we prove that the MWIEST game is an exact potential game and there exists at least one Nash equilibrium (NE) point which is an approximate solution of the MWIEST problem (Theorem 1). Besides, we also prove that best responses (BRs) of the game converges to a NE in $6^4 \times N$ iterations at most, where N is the total number of nodes (Theorem 2).
- Thirdly, based on the MWIEST game, we propose four algorithms that are the node information exchange (NIE), the best response selection (BRS), the intermediate nodes selection (INS) and the forming cluster (FC). Specifically, The NIE algorithm is used for nodes to exchange node's information and neighbors' information list with their neighbors. The BRS algorithm is built based on the MWIEST game to achieve a NE point. The INS algorithm uses the NE point to obtain the SetIN which supports the FC algorithm to form clusters.
- Finally, we propose the game theory based clustering (GBC) protocol which is combination of the FC algorithm and the proposed cluster maintenance algorithm to construct high stable clusters supporting multicast routing in CRAHNs. Moreover, each obtained cluster includes most members having the same receiving channel and avoids affected regions of licensed channels.
- For the performance evaluation of the proposed GBC protocol, we implement the GBC protocol in OMNET++ platform to demonstrate that it provides better performance than the state-of-the-art protocols in terms of network stability and control overheads.

The rest of this article is organized as follows. Section II describes system model and the basic concept of the proposed game based clustering protocol. Section III includes necessary functions and definitions, the MWIEST problem and the proposed MWIEST game. Section IV presents necessary algorithms that are NIE, BRS, INS and FC to support the proposed GBC protocol. Section V shows the performance evaluation of the proposed GBC protocol. Finally, Section VI concludes the paper.

II. SYSTEM MODEL

In this section, we present the system model and the basic concept of the proposed GBC protocol, as shown in Figure 1. We consider a CRAHN consisting of multiple SUs and PUs, called nodes that have some characteristics as follows:

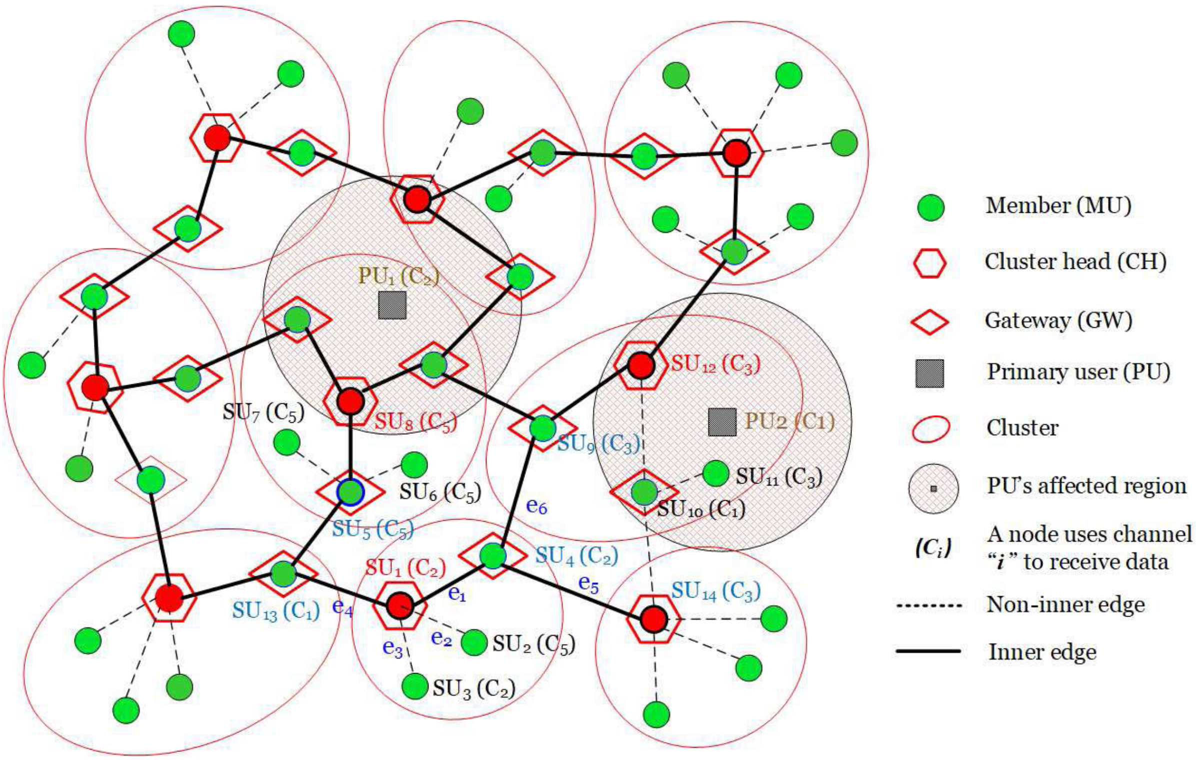


FIGURE 1. Basic concept of the proposed game based clustering protocol.

- All nodes can move randomly in two-dimension space based on the random waypoint (RWP) model.
- Each node can know its own location through the global positioning system (GPS) [40], [41].
- The transmission range of nodes is fixed.
- Each node can exchange control packets by using a given channel (control channel) which is different from data channels. Moreover, the control channel does not affect the licensed channels of PUs.
- All nodes are classified into groups, called clusters. Each cluster includes one cluster head (CH), gateways (GWs) and member users (MUs), where a CH is a representation of a cluster and its tasks are aggregating and transmitting messages of rest nodes in its cluster within a single-hop. Each CH can directly communicate with GWs. A GW is an intermediate node which can directly communicate with its CH or CHs/GWs of other clusters within a single-hop. If a node is not a CH or a GW, it is considered as a member user (MU) which can only directly communicate with its CH [42].

The basic concept of the GBC protocol in Figure 1 can be explained as follows:

- Step 1. Each node exchanges node's information and neighbors' information list with their neighbors by using the NIE algorithm.
- Step 2. Each node uses the BRS algorithm to obtain a NE point of the MWIEST game. Particularly, each node v chooses its BR (best response) based on the MWIEST game. A BR is a set of links (edges) which implies a weighted spanning tree with minimum possible total weight of inner edges in graph $\mathsf{Gr}_v = (\delta[v], \mathsf{E}_{\delta[v]})$, where E is the set of all edges of the network, $\delta[v]$ is a set of neighbors of node v including v, $\mathsf{E}_{\delta[v]}$ is a subset of E such that all edges $e \in \mathsf{E}_{\delta[v]}$ have their endpoints in $\delta[v]$ and an edge is an inner edge if both its endpoints are non-leaf nodes. The BRS algorithm repeats the BR selection until a NE point is reached.
- Step 3. Each node uses the INS algorithm to select a set of intermediate nodes based on the NE point which is obtained in Step 2. A node is an intermediate node if its degree is strictly greater than 1 in graph $G_v = (\delta[v], \mathsf{E}(\mathsf{BR}))$, where $\mathsf{E}(\mathsf{BR})$ is the set of edges which is selected by BR.
- Step 4. Each node uses the GBC protocol to select CHs and GWs from the set of intermediate nodes which is obtained in Step 3. Based on the CHs and GWs, the GBC protocol classifies nodes into high stable clusters.

Based on the basic concept of the proposed GBC protocol, the example in Figure 1 can be explained as follows:

- Step 1. Each node exchanges node's information and neighbors' information list with their neighbors by using the NIE algorithm.
- Step 2. Each node uses the BRS algorithm to obtain a NE point which is a combination of all BR_i, where BR_i is the best response of SU_i. We have $\mathsf{BR}_1 = \{e_1, e_2, e_3, e_4\}$ and $\mathsf{BR}_4 = \{e_1, e_5, e_6\}$.

- Step 3. Because the edges e_1, e_4, e_5 and e_6 are inner edges and endpoints of these edges are $\mathsf{SU}_1, \mathsf{SU}_4, \mathsf{SU}_9, \mathsf{SU}_{13}$ and SU_{14}, these endpoints will be selected to become CHs or GWs.
- Step 4. Clustering process:
  - Node SU_3 (channel C_2) elects intermediate node SU_1 (channel C_2) as a CH. Node SU_3 becomes a MU.
  - Because node SU_2 (channel C_5) has no adjacent intermediate node (channel C_5), it chooses the CH (SU_1) as its CH based on node-weight function. Node SU_2 becomes a MU.
  - Because the intermediate node SU_4 (channel C_2) is not elected by any non-intermediate node, it chooses the CH (SU_1) (channel C_2) as its CH. Node SU_4 becomes a GW.
  - The group of nodes $\{\mathsf{SU}_1, \mathsf{SU}_2, \mathsf{SU}_3, \mathsf{SU}_4\}$ becomes a cluster with maximum number of nodes having the same receiving channel C_2.
  - Non-intermediate node SU_{11} elects intermediate node SU_{12} as a CH because SU_{12} has the receiving channel C_3 which is different from PU_2's channel C_1. Node SU_{11} becomes a MU.
  - Because the intermediate nodes SU_9 and SU_{10} are not elected by any non-intermediate node, the intermediate node SU_9 (channel C_3) chooses the CH (SU_{12}) with channel C_3 as its CH, and intermediate node SU_{10} chooses CH (SU_{12}) as its CH based on node-weight function. Nodes SU_9 and SU_{10} becomes GWs.
  - The group of nodes $\{\mathsf{SU}_9, \mathsf{SU}_{10}, \mathsf{SU}_{11}, \mathsf{SU}_{12}\}$ becomes a cluster with maximum nodes having the same receiving channel C_3, node SU_{11} becomes a MU and the cluster's receiving channel C_3 is not affected by PU_2's channel C_1.

The rest of SUs are operated to become CHs, GWs or MUs by the same way as the above SUs.

III. PRELIMINARIES

We consider the CRAHN as an un-directed connected graph Gr. We assume that each edge e in Gr has a weight $\mathsf{W}(e)$ which measures the stability of this edge and the smaller $\mathsf{W}(e)$ is, the higher the stability of edge e is. We have that a spanning tree of Gr with minimum possible total weight of inner edges provides endpoints of these inner edges which keep connection together in the high stable way. In this section, we present some necessary definitions and notations in graph theory and a link-weight function to propose an optimization problem, called MWIEST problem. The solution of the MWIEST problem is a spanning tree with minimum possible total weight of inner edges of the network which supports the proposed GBC protocol to obtain high stable clusters. Because the MWIEST problem is a NP-complete proven in Remark 4, we propose a MWIEST game to find a approximate solution of the MWIEST problem.

A. DEFINITIONS AND NOTATIONS

A CRAHN is considered as an un-directed connected graph $\mathsf{Gr} = (\mathsf{V}, \mathsf{E})$, where $\mathsf{V} = \{v_1, v_2, \ldots, v_N\}$ is a set of vertices (nodes) and $\mathsf{E} = \{e_1, e_2, \ldots, e_M\}$ is a set of edges (un-directed link). An edge between two nodes u and v indicates that both nodes v and u are in their wireless transmitter ranges. We assume that the transmission range of nodes is the same. Thus, if there is an edge $e = (v, u)$ in E, u is within v's range and v is within u's range.

Edge selection (ES) vector: A vector $\mathsf{s} = (\mathsf{s}_1, \mathsf{s}_2, \ldots, \mathsf{s}_M) \in \{0, 1\}^M$ is called an ES vector. It implies a subset of E which is denoted by

$$\mathsf{E}(\mathsf{s}) = \{e_i \in \mathsf{E} \mid \mathsf{s}_i = 1\}. \quad (1)$$

Spanning tree (ST): A ST of Gr is an un-directed connected sub-graph which includes all of vertices of Gr, without any cycles, with minimum possible number of edges.

Weighted Edge Spanning Tree (WST): A WST is a ST where edges have weights or values.

Minimum Weighted Inner Edge Spanning Tree: A MWIEST is a WST with minimum possible total weight of inner edges, where an edge is an inner edge if both its endpoints have degree ≥ 2 in the WST.

For ease of presentation, we give some notations and definitions as shown in Table 1.

TABLE 1. Notations and definitions.

Symbol	Definition
$\vert.\vert$	The cardinality of a set. $\vert\mathsf{V}\vert = N$ and $\vert\mathsf{E}\vert = M$.
$\delta(i)$	The set of neighbors of node i.
$\delta[i]$	$\delta[i] = \delta(i) \cup \{i\}$.
$\mathsf{deg}(i)$	The degree of node i. It is the number of edges that are incident to node i.
$\mathsf{E_A}$	A is a subset of V. $\mathsf{E_A}$ is a subset including all edges $e \in \mathsf{E}$ which have their endpoints in A.
$\mathsf{Gr[B]}$	Given $\mathsf{B} \subset \mathsf{E}$, Gr[B] is a sub-graph of Gr induced by B, i.e., Gr[B] consists B and endpoints of all edge in B.
$\mathsf{s}(e)$	Let $\mathsf{s} = (\mathsf{s}_1, \mathsf{s}_2, \ldots, \mathsf{s}_M)$ be an ES vector and $e \in \mathsf{E}$, $\mathsf{s}(e) = \mathsf{s}_i$, with $e_i = e \in \mathsf{E}$.
$\mathsf{sv(B)}$	Let B be a subset of E, $\mathsf{sv(B)} = (\mathsf{s}_1, \mathsf{s}_2, \ldots, \mathsf{s}_M)$ is an ES vector, with $\mathsf{s}_i = 1$ if edge e_i is in B and $\mathsf{s}_i = 0$ if otherwise.
leaf node	In a given graph, a node i is called a leaf node if $\mathsf{deg}(i) = 1$.
inner edge	In a given graph, an edge e is called an inner edge if both its endpoints are non-leaf nodes.
$\mathsf{inner}(e)$	In a given graph, $\mathsf{inner}(e) = 1$ if e is an inner edge, and $\mathsf{inner}(e) = 0$ if otherwise.

1) THE PROPOSED LINK-STABILITY FUNCTION: LS FUNCTION

Each edge e in Gr has two endpoints i and j which have movement directions $(d_x^{(i)}, d_y^{(i)})$ with speed v_i and angle $\alpha_i = \arctan(d_y^{(i)}/d_x^{(i)})$, and $(d_x^{(j)}, d_y^{(j)})$ with speed v_j and angle $\alpha_j = \arctan(d_y^{(j)}/d_x^{(j)})$, respectively.

At time t_0, position of nodes i and j are $(x_0^{(i)}, y_0^{(i)})$ and $(x_0^{(j)}, y_0^{(j)})$, respectively. At time t_1, position of nodes i and j are $(x_1^{(i)}, y_1^{(i)})$ and $(x_1^{(j)}, y_1^{(j)})$, respectively, where positions at time t_1 can be calculated as

$$x_1^{(i)} = x_0^{(i)} + v_i(t_1 - t_0)\cos(\alpha_i),$$

$$y_1^{(i)} = x_0^{(i)} + v_i(t_1 - t_0)\sin(\alpha_i),$$
$$x_1^{(j)} = x_0^{(j)} + v_j(t_1 - t_0)\cos(\alpha_j),$$
$$y_1^{(j)} = x_0^{(j)} + v_j(t_1 - t_0)\sin(\alpha_j). \qquad (2)$$

We denote $\mathsf{D}_{t_0}(i,j)$ and $\mathsf{D}_{t_1}(i,j)$ as distances between i and j at time t_0 and t_1, respectively. By using Eq. (2), $\mathsf{D}_{t_0}(i,j)$ and $\mathsf{D}_{t_0+\Delta t}(i,j)$ can be calculated as

$$\mathsf{D}_{t_0}(i,j) = \sqrt{(x_{t_0}^{(i)} - x_{t_0}^{(j)})^2 + (y_{t_0}^{(i)} - y_{t_0}^{(j)})^2},$$
$$\mathsf{D}_{t_0+\Delta t}(i,j) = \sqrt{(x_{t_0+\Delta t}^{(i)} - x_{t_0+\Delta t}^{(j)})^2 + (y_{t_0+\Delta t}^{(i)} - y_{t_0+\Delta t}^{(j)})^2}. \qquad (3)$$

The link-stability function $\mathsf{LS}_{\Delta t}$ of link e over interval time Δt is defined as

$$\mathsf{LS}_{\Delta t}(e) = \frac{|\mathsf{D}_{t_0}(i,j) - \mathsf{D}_{t_0+\Delta t}(i,j)|}{2S_{\max}\Delta t}, \qquad (4)$$

where $S_{\max}$ is the maximum speed of nodes.

Remark 1: The value of $\mathsf{LS}(e)$ indicates that the smaller the $\mathsf{LS}(e)$ is, the higher the stability of the distance between two endpoints of edge e is.

2) THE PROPOSED LINK-CONNECTIVITY FUNCTION: LC FUNCTION

We denote the function LC as a link-connectivity function of edge e which can be express as

$$\mathsf{LC}(e) = \frac{\mathsf{deg}(u) + \mathsf{deg}(v)}{2\mathsf{deg}_{\max}}, \qquad (5)$$

where u and v are two endpoints of edge e and $\mathsf{deg}_{\max}$ is the maximum degree of each node.

Remark 2: The value of $\mathsf{LC}(e)$ indicates that the higher the $\mathsf{LC}(e)$ is, the higher the total of nodes adjacent to the endpoints of edge e is. Thus, the value of $1 - \mathsf{LC}(e)$ indicates that the smaller the $1 - \mathsf{LC}(e)$ is, the higher the total of nodes adjacent to the endpoints of edge e is.

3) THE PROPOSED LINK-WEIGHT FUNCTION: LW FUNCTION

The link-weight function W of edge e is defined as a combination of the functions LS and LC in Eq. (4) and Eq. (5), respectively. It can be expressed as

$$\mathsf{W}(e) = w_1\mathsf{LS}_{\Delta t}(e) + w_2(1 - \mathsf{LC}(e)), \qquad (6)$$

where Δt is a time constant, $w_1, w_2 \in \mathbb{R}$ and $w_1 + w_2 = 1$.

Remark 3: The value of $\mathsf{W}(e)$ indicates that the smaller the $\mathsf{W}(e)$ is, the higher the opportunity of edge e becoming an edge of a MWIEST is.

B. THE MWIEST PROBLEM

The problem of finding a MWIEST of a graph Gr, called MWIEST problem, can be described as

$$\min_{\mathsf{s}} F(\mathsf{s}) = \sum_{i=1}^{M} \mathsf{s}_i \mathsf{inner}(e_i)\mathsf{W}(e_i)$$
$$\text{subject to } \mathsf{s} = (\mathsf{s}_1, \mathsf{s}_2, \ldots, \mathsf{s}_M) \in \{0,1\}^M$$
$$\text{sub-graph } (V, \mathsf{E}(s)) \text{ of } \mathsf{Gr} \text{ is a WST},$$
$$\mathsf{inner}(e_i) \text{ is calculated in } (V, \mathsf{E}(s)), \qquad (7)$$

where each e_i is an edge in E, $\mathsf{s}_i = \mathsf{s}(e_i)$ and $\mathsf{inner}(e_i)$ are defined in Table 1, $\mathsf{E}(s)$ is a set of edges in Eq. (1) and $\mathsf{W}(e_i)$ in Eq. (6) is the link-weight function of edge e.

Remark 4: The MWIEST problem is NP-complete.

Proof: If we set the weight of all edges equal to 1, the MWIEST problem is equivalent to the maximum leaf spanning tree (MLST) problem. According to [46], [47], the MLST problem is NP-complete. Thus the MWIEST problem is also NP-complete. ■

Because the MWIEST problem is NP-complete, we will propose a MWIEST game in section III-C2 to find a NE point which is a combination of all node's BRs. Each BR of a node is a set of edges which implies a local weight edge spanning tree with possible total weight of inner edges. The obtained NE point can be considered as an approximate solution of the MWIEST game.

C. THE PROPOSED MINIMUM WEIGHTED INNER EDGE SPANNING TREE GAME: MWIEST GAME

1) POTENTIAL GAME

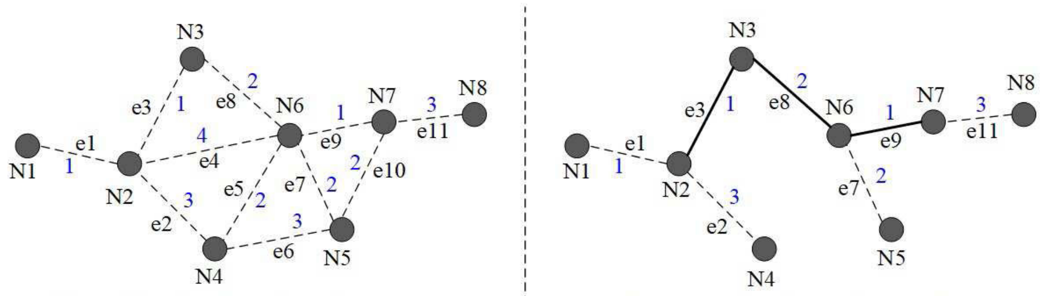
A standard representation of a game is modeled as a normal form game, or a game in strategic form:

- The set of players is $\mathsf{N} = \{1, \ldots, N\}$.
- Player i has an available set of strategies (actions) S_i. This set might be finite or infinite.
- Let $\mathsf{S} = \mathsf{S}_1 \times \cdots \times \mathsf{S}_N$ be a N dimension space of all S_i, $i = 1, \ldots, n$. An element $\mathsf{s} \in \mathsf{S}$ is denoted by $\mathsf{s} = (\mathsf{s}_1, \ldots, \mathsf{s}_N)$, where s_i is a strategy of S_i.
- Player i's payoff is a function of vector s which $u_i\colon \mathsf{S} \to \mathbb{R}$, $\mathsf{s} \mapsto u_i(\mathsf{s})$, where $\mathsf{s} \in \mathsf{S}$ is chosen by all players in the game.

A game is called a potential game if the strategies of all players can be expressed by using a single global function called the potential function [43]. The potential function is used to analyze equilibrium properties of the game because the strategies of all players are mapped into one potential function, and the Nash equilibrium points can be found by the local optima selection process of the potential function. There are several types of potential games which are presented as follows:

Let $\mathsf{Gm} = (\mathsf{N}; \mathsf{S} = (\mathsf{S}_1 \times \cdots \times \mathsf{S}_N); u_i\colon \mathsf{S} \to \mathbb{R}, i = 1, \ldots, N)$ be a game. We denote s_i as a strategy in S_i, s_{-i} as the strategies of all players except player i, S_{-i} as the set of all s_{-i}, and $(\mathsf{s}_i, \mathsf{s}_{-i}) = (\mathsf{s}_1, \ldots, \mathsf{s}_i, \ldots, \mathsf{s}_N)$ as a vector in S. The types of potential game are defined as follows:

- The game Gm is an exact potential game if there is a function $\Phi\colon \mathsf{S} \to \mathbb{R}$ such that $\forall \mathsf{s}_{-i} \in \mathsf{S}_{-i}, \forall \mathsf{s}'_i, \mathsf{s}''_i \in \mathsf{S}_i$,
$$\Phi(\mathsf{s}'_i, \mathsf{s}_{-i}) - \Phi(\mathsf{s}''_i, \mathsf{s}_{-i}) = u_i(\mathsf{s}'_i, \mathsf{s}_{-i}) - u_i(\mathsf{s}''_i, \mathsf{s}_{-i}). (8)$$
- The game Gm is a weighted potential game if there is a function $\Phi\colon \mathsf{S} \to \mathbb{R}$ such that $\forall \mathsf{s}_{-i} \in \mathsf{S}_{-i}, \forall \mathsf{s}'_i, \mathsf{s}''_i \in \mathsf{S}_i$ and a vector $w \in \mathbb{R}^N_{++}$ such that
$$\Phi(\mathsf{s}'_i, \mathsf{s}_{-i}) - \Phi(\mathsf{s}''_i, \mathsf{s}_{-i}) = w_i(u_i(\mathsf{s}'_i, \mathsf{s}_{-i}) - u_i(\mathsf{s}''_i, \mathsf{s}_{-i})), \qquad (9)$$
where $\mathbb{R}_{++}$ is a set of positive real numbers.



The weighted undirected graph Gr

A minimum weighted inner edge spanning tree of the graph Gr with total weight of all inner edges 4

FIGURE 2. **Example of MWIEST problem.**

- The game Gm is an ordinal potential game if there is a function $\Phi\colon \mathsf{S} \to \mathbb{R}$ such that $\forall \mathsf{s}_{-i} \in \mathsf{S}_{-i}, \forall \mathsf{s}'_i, \mathsf{s}''_i \in \mathsf{S}_i$,

$$\Phi(\mathsf{s}'_i, \mathsf{s}_{-i}) - \Phi(\mathsf{s}''_i, \mathsf{s}_{-i}) > 0 \Leftrightarrow u_i(\mathsf{s}'_i, \mathsf{s}_{-i}) - u_i(\mathsf{s}''_i, \mathsf{s}_{-i}) > 0. \quad (10)$$

- The game Gm is a best response potential game if there is a function if there is a function $\Phi\colon \mathsf{S} \to \mathbb{R}$ such that $\forall i \in N, \forall \mathsf{s}_{-i} \in \mathsf{S}_{-i}$,

$$b_i(\mathsf{s}_{-i}) = \arg\max_{\mathsf{s}_i \in \mathsf{S}_i} \Phi(\mathsf{s}_i, \mathsf{s}_{-i}), \quad (11)$$

where $b_i(\mathsf{s}_{-i})$ is the best strategy for player i given s_{-i}.

2) THE MWIEST GAME

We propose a MWIEST game to model the optimization problem (7) as an exact potential game. Figure 2 presents an example of the MWIEST problem while Figures 3 and 4 explain the operation of all nodes and Figure 5 shows the solution of the MWIEST game to compare with the solution of the MWIEST problem. The MWIEST game can be described as follows:

- **Player:** Each node is considered as a player in the game.
- **Strategy:** An ES vector $\mathsf{s}_i = (\mathsf{s}_{i1}, \mathsf{s}_{i2}, \ldots, \mathsf{s}_{iM}) \in \{0, 1\}^M$ is called a strategy of node i if $\mathsf{s}_i(e) = 0, \forall e \in \mathsf{V} \setminus \delta[i]$ and s_i implies a local WST $\mathcal{T}_i(\mathsf{s}_i) = (\delta[i], \mathsf{E}(\mathsf{s}_i))$ of sub-graph $\mathsf{Gr}_i = (\delta[i], \mathsf{E}_{\delta[i]})$.
- **Strategy Selection (SS) Rule:** The game is divided into stages, where each stage is a period of time. At stage τ, all nodes have to choose their respective strategies based on rules as follows:
  - Each node chooses its strategy one by one such that a node with a smaller ID will choose its strategy before a node with a larger ID. To do this, a node i is assigned a time slot t_i such that if $\forall i, j \in \{1, \ldots, N\}$ and $i < j$, then $t_i < t_j$.
  - Node i can be only allowed to choose its strategies, if time slot t_i is active.
  - Node i can only choose strategy $\mathsf{s}_i^{(\tau,t_i)}$ which don't affect any strategy $\mathsf{s}_j^{(\tau,t_i)}$, $\forall j \in \delta(i)$, $j < i$, i.e., $\mathsf{E}_{\delta[i]\cap\delta[j]}(\mathsf{s}_i^{(\tau,t_i)}) = \mathsf{E}_{\delta[i]\cap\delta[j]}(\mathsf{s}_j^{(\tau,t_i)})$, $\forall j \in \delta(i)$, $j < i$.
  - If node N finishes the strategy selection, the game will move to a new stage and time slots of all nodes are updated in the new stage.
  - In the new stage, each node will repeat its work similar to the previous stage.
- **Payoff function:** At stage τ and time slot t_i, if s_i does not satisfy the SS rule, the payoff of strategies $(\mathsf{s}_i^{(\tau,t_i)}, \mathsf{s}_{-i}^{(\tau,t_i)})$ is assigned to $-\infty$. Otherwise, the payoff function of strategies $(\mathsf{s}_i^{(\tau,t_i)}, \mathsf{s}_{-i}^{(\tau,t_i)})$ is defined as follows:

$$u_i^{(\tau,t_i)}(\mathsf{s}_i^{(\tau,t_i)}, \mathsf{s}_{-i}^{(\tau,t_i)}) = -\sum_{j,k\in\delta[i]} \mathsf{s}_i^{(\tau,t_i)}(e_{jk})\mathsf{inner}_i^{(\tau,t_i)}(e_{jk})\mathsf{W}(e_{jk}), \quad (12)$$

where $\mathsf{inner}_i^{(\tau,t_i)}(e)$ is a function of edge e with strategies $(\mathsf{s}_i^{(\tau,t_i)}, \mathsf{s}_{-i}^{(\tau,t_i)})$ to determine whether an edge e is inner edge or not. To do this, we need to compute degree of endpoints of edge e. The degree of a node using for node i's payoff at stage τ and time slot t_i, denoted by $\mathsf{deg}_i^{(\tau,t_i)}(*)$, is defined as

  - If $j \leq i$, the $\mathsf{deg}_i^{(\tau,i)}(j)$ is total number of edges $e \in \mathsf{E}(\mathsf{s}_j^{(\tau,t_i)})$ which are incident to node j.
  - If $k > i$, the $\mathsf{deg}_i^{(\tau,t_i)}(k)$ is total number of edges $e \in \mathsf{E}(\mathsf{s}_i^{(\tau,t_i)}) \cup \mathsf{E}_{\{k\}\cup(\mathsf{V}\setminus\delta[\leq i])}(\mathsf{s}_k^{(\tau,t_i)})$, where $\bigcup_{l\leq i} \delta[l]$ is denoted by $\delta[\leq i]$.

This payoff function (12) is a local function which can be calculated at node i based on strategy information of its neighbors. Moreover, the minus of payoff function of node i gives the total weight of inner edges of a local WST $\mathcal{T}_i(\mathsf{s}_i^{(\tau,t_i)})$.

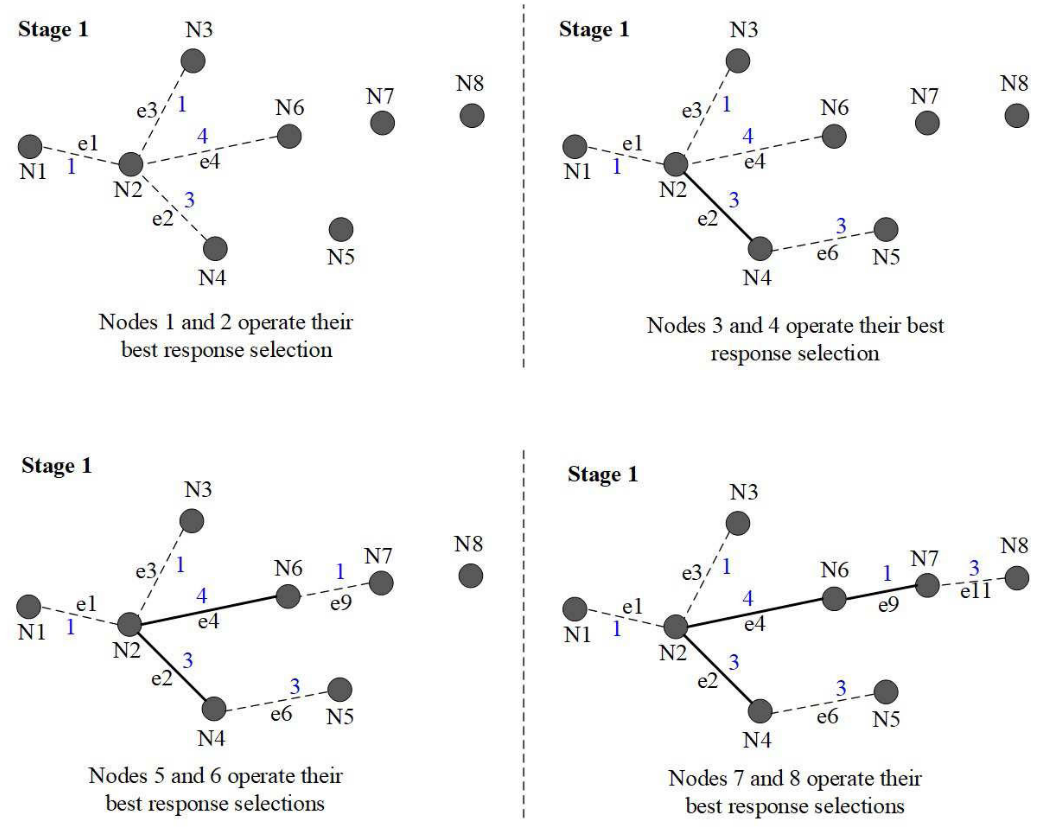


FIGURE 3. Operation of MWIEST game (a).

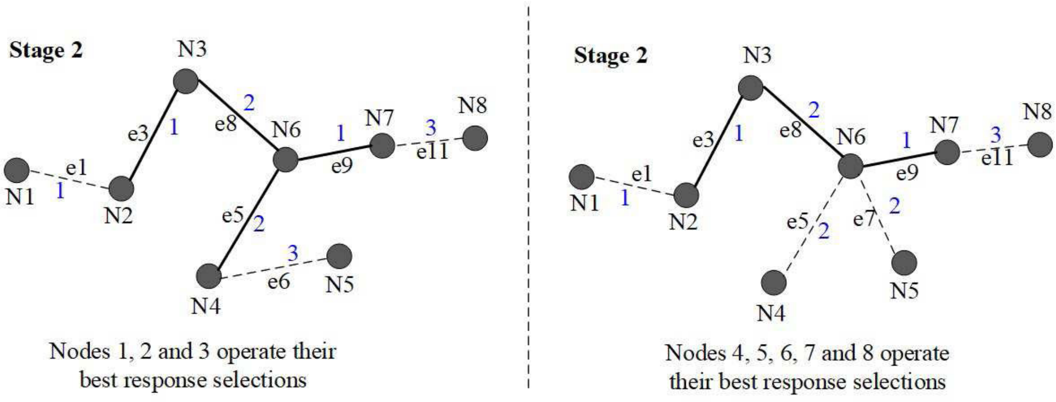


FIGURE 4. Operation of MWIEST game (b).

- **Global Strategy:** At stage τ and time slot t_i, a global strategy vector $\mathsf{gs}_i^{(\tau,t_i)}$ of $(\mathsf{s}_i^{(\tau,t_i)}, \mathsf{s}_{-i}^{(\tau,t_i)})$ is defined as

$$\mathsf{gs}_i^{(\tau,t_i)} = \mathsf{sv}\left(\bigcup_{j\leq i} \mathsf{E}(\mathsf{s}_j^{(\tau,t_i)}) \cup \bigcup_{k>i} \mathsf{E}_{\{k\}\cup(\mathsf{V}\setminus\delta[\leq i])}(\mathsf{s}_k^{(\tau,t_i)})\right). \tag{13}$$

- **Best Response** (BR): The BR of node i at stage τ and time slot t_i can be expressed as

$$\mathsf{s}_i^{*(\tau,t_i)} = \arg\max_{\mathsf{s}^{(\tau,t_i)}\in\mathsf{S}_i} u_i^{(\tau,t_i)}(\mathsf{s}^{(\tau,t_i)}, \mathsf{s}_{-i}^{(\tau,t_i)}), \tag{14}$$

where S_i is the set of strategies of node i such that each strategy in S_i satisfies the SS rule. In order words, a BR of node i gives a local MWIEST $\mathcal{T}_i(\mathsf{s}_i^{*(\tau,t_i)})$.

- **Objective function:** At stage τ and time slot t_i, the objective function Φ of the MWIEST game is defined as

$$\Phi((\mathsf{s}_i^{(\tau,t_i)}, \mathsf{s}_{-i}^{(\tau,t_i)})) = -\sum_{e\in \mathsf{E}} \mathsf{gs}_i^{(\tau,t_i)}(e)\mathsf{inner}(e)\mathsf{W}(e), \tag{15}$$

where $\mathsf{inner}(e)$ is calculated in graph $(\mathsf{V}, \mathsf{E}(\mathsf{gs}_i^{(\tau,t_i)}))$.

Remark 5: The solution of the MWIEST-GAME is an approximate solution of the MWIEST problem (7). This will be proven in Theorem 1.

Theorem 1: The MWIEST game is an exact potential game, i.e., at stage τ and time slot t_i,

$$\begin{aligned} u_i(\mathsf{s}_i^{\prime\prime(\tau,t_i)}, \mathsf{s}_{-i}^{\prime\prime(\tau,t_i)}) - u_i(\mathsf{s}_i^{\prime(\tau,t_i)}, \mathsf{s}_{-i}^{\prime(\tau,t_i)}) \\ = \Phi(\mathsf{s}_i^{\prime\prime(\tau,t_i)}, \mathsf{s}_{-i}^{\prime\prime(\tau,t_i)}) - \Phi(\mathsf{s}_{t,i}^{\prime(\tau,t_i)}, \mathsf{s}_{t,-i}^{\prime(\tau,t_i)}). \end{aligned} \tag{16}$$

In addition, there exists at least one Nash equilibrium point which is an approximate solution of the MWIEST problem.

Proof: In this proof, for convenience, we remove the notation "(τ, t_i)" which presents stage τ and time slot t_i. The set of edges E can be presented as $\mathsf{E} = (\mathsf{E} \setminus E_{\delta[i]}) \cup E_{\delta[i]}$. An edge e has two endpoints v_s (s-th node) and v_d (d-th node), and two status inner edge and non-inner edge. Let s_i is an arbitrary strategy of node i. We can see that an edge e in $\mathsf{E} \setminus E_{\delta[i]}$ falls into one of the following cases

(i) Both v_s and v_d are not in $\delta[i]$: Based on the SS rule, it is easily to see that edge e is not affected by s_i, i.e., the status of edge e is not changed by s_i.

(ii) $v_s \notin \delta[i]$ and $v_d \in \delta[i], d < i$: Due to the SS rule, because s_i does not affect to the strategy s_d, the $\mathsf{deg}(v_d)$ is not changed by s_i. Moreover, v_s is not in $\delta[i]$, the $\mathsf{deg}(v_s)$ is also not changed by s_i. Thus, edge e is not affected by s_i.

(iii) $v_s \notin \delta[i]$ and $v_d \in \delta[i], d \geq i$: According to the SS rule, because s_i implies a local WST, there is at least one edge $e \in \mathsf{E}(\mathsf{s}_i)$ which has one endpoint v_d and one endpoint $v_k \in \delta[i]$. Moreover, v_d is also an endpoint of edge $e \notin \mathsf{E}(\mathsf{s}_i)$. Thus, degree of node v_d is always greater than 2. We have that edge e is not changed by s_i.

(iv) Both v_s and v_d are not in $\delta[i]$: In this case, the status of edge e only depends on s_i.

According to above results, the objective function Φ can be expressed in Eq. (17), as shown at the bottom of the page, of which we have

$$\begin{aligned} u_i(\mathsf{s}_i^{\prime\prime(\tau,t_i)}, \mathsf{s}_{-i}^{\prime\prime(\tau,t_i)}) - u_i(\mathsf{s}_i^{\prime(\tau,t_i)}, \mathsf{s}_{-i}^{\prime(\tau,t_i)}) \\ = \Phi(\mathsf{s}_i^{\prime\prime(\tau,t_i)}, \mathsf{s}_{-i}^{\prime\prime(\tau,t_i)}) - \Phi(\mathsf{s}_{t,i}^{\prime(\tau,t_i)}, \mathsf{s}_{t,-i}^{\prime(\tau,t_i)}). \end{aligned} \tag{18}$$

First, according to Eq. (18), we have that the MWIEST game is an exact potential game.

Second, at the stage t_1, if there exits a node which has not selected its BR, i.e., the global strategy does not include all BRs, the value of Φ can decrease until all nodes finish their BR selection processes. When the stage t_1 is end, the global strategy includes all BRs. Hence, at the stage t_i, $i > 1$, the Eq. (16) indicates that the value of Φ will increase after each node chooses its BR. Moreover, because the value of Φ is less than 0 and each node has finite strategies, there exists at least one NE point and the BRs of the game that will converge to the NE point in finite iterations.

Finally, according to the SS rule, we have that a global strategy (gs) is a combination of all BRs and graph $(\mathsf{V}, \mathsf{E}(\mathsf{gs}))$ is a combination of all local MWIESTs. Thus, the solution of the MWIEST game is an approximate solution of the MWIEST problem. ■

Theorem 2: The best responses (BRs) of the MWIEST game will converge to a NE point in $6^4 \times |\mathsf{V}|$ iterations at most, where V is the set of all nodes.

Proof: In this game, to reduce the computation complexity in clustering process, each node i only considers around 5 neighbors for best response selection, i.e., $|\delta[i]| \leq 6, \forall i \in \mathsf{V}$. The reason is that the number of nodes in our simulation environment is 50 node, and we expect the number of cluster to be 9. Thus, the average number of members of each cluster is expected to be 6. Any spanning tree $\mathcal{T}_i$ of $\mathsf{Gr} = (\delta[i], \mathsf{E}_{\delta[i]})$ has 5 edges at most. Based on [48], the number of spanning trees in the complete graph is $N^{(N-2)}$. Thus, each node i has 6^4 strategies to choose. In each stage, because each node chooses its BR one-time, we have maximum $6^4 \times |\mathsf{V}|$ iterations for the strategy selection process. Hence, the BRs of the MWIEST game will converges to a Nash Equilibrium and the number of iterations to converge is less than $6^4 \times |\mathsf{V}|$. ■

$$\begin{aligned} \Phi((\mathsf{s}_i, \mathsf{s}_{-i})) &= -\sum_{e\in\mathsf{E}} \mathsf{gs}_i(e)\mathsf{inner}(e)\mathsf{W}(e) = -\sum_{e\in\mathsf{E}(\mathsf{gs}_i)} \mathsf{inner}(e)\mathsf{W}(e) \\ &= -\sum_{e\in\mathsf{E}(\mathsf{gs}_i)\setminus E(\mathsf{s}_i)} \mathsf{inner}(e)\mathsf{W}(e) - \sum_{e\in\mathsf{E}(\mathsf{s}_i)} \mathsf{inner}(e)\mathsf{W}(e) \\ &= -\sum_{e\in\mathsf{E}(\mathsf{gs}_i)\setminus E(\mathsf{s}_i)} \mathsf{inner}(e)\mathsf{W}(e) - \sum_{j,k\in\delta[i]} \mathsf{s}_i(e_{jk})\mathsf{inner}_i(e_{jk})\mathsf{W}(e_{jk}) \\ &= -\sum_{e\in\mathsf{E}(\mathsf{gs}_i)\setminus E(\mathsf{s}_i)} \mathsf{inner}(e)\mathsf{W}(e) - u_i(\mathsf{s}_i, \mathsf{s}_{-i}). \end{aligned} \tag{17}$$

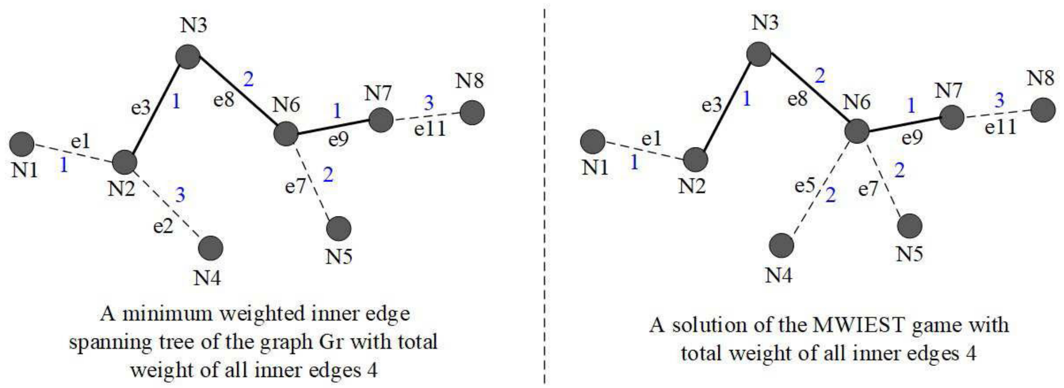


FIGURE 5. Solutions of MWIEST problem and MWIEST game.

TABLE 2. The MWIEST game example.

L.I.E.: the set of inner edges of the local WST $\mathcal{T}_i(\mathsf{s}_i^*)$.
G.I.E.: the set of inner edges of the global BR.
$\Phi(*)$: the value of the objective function of the game.

Iteration	Node	BR_i	L.I.E.	Max payoff	G.I.E	$\Phi(*)$
1	1	e_1	$\emptyset$	0	$\emptyset$	0
2	2	e_1, e_2, e_3, e_4	$\emptyset$	0	$\emptyset$	0
3	3	e_3, e_4	$\emptyset$	0	$\emptyset$	0
4	4	e_2, e_4, e_6	e_2	-3	e_2	-3
5	5	e_6, e_9	$\emptyset$	0	e_2, e_4	-7
6	6	e_2, e_3, e_4, e_6, e_9	e_2, e_4	-7	e_2, e_4	-7
7	7	e_9, e_{11}	e_9	-1	e_2, e_4, e_9	-8
8	8	e_{11}	$\emptyset$	0	e_2, e_4, e_9	-8
9	1	e_1	$\emptyset$	0	e_2, e_4, e_9	-8
10	2	e_1, e_3, e_8, e_5	e_3, e_5, e_8	-5	e_3, e_5, e_8, e_9	-6
11	3	e_3, e_8	e_3, e_8	-3	e_3, e_5, e_8, e_9	-6
12	4	e_5, e_7	$\emptyset$	0	e_3, e_8, e_9	-4
13	5	e_5, e_7, e_9	e_9	-1	e_3, e_8, e_9	-4
14	6	e_5, e_7, e_8, e_9	e_8, e_9	-3	e_3, e_8, e_9	-4
15	7	e_9, e_{11}	e_9	-1	e_3, e_8, e_9	-4
16	8	e_{11}	$\emptyset$	0	e_3, e_8, e_9	-4

3) AN EXAMPLE OF THE MWIEST GAME

In this part, we give an example of the MWIEST game to explain the basic concept of the MWIEST game. The operation of this game is divided into stages. At a stage τ, node i has a time slot t_i to active its BR selection process, where the stage τ and time slot t_i have been already defined in the SS rule. The operation of node i at a stage t and time slot t_i can be expressed as follows:

- Step 1. Based on SS rule, node i considers all strategy and chooses the best strategy with maximum its payoff (a new BR). In order words, node i chooses a local MWIEST which satisfies the SS rule.
- Step 2. If the new BR of node i is not different from the current BR of node i, the game is end. Otherwise, node i will wait for the next stage $t + 1$ and return to Step 1.

When the game is end, the global strategy of all BRs is a NE point which is considered as a solution of the MWIEST game. The obtained NE is an approximate solution of the MWIEST problem (7).

In Figure 2, we present an example of the MWIEST problem and a solution of this problem. Figures 3 and 4 present the operation of all nodes in the MWIEST game and Figure 5 shows the solution of the MWIEST game to compare with the solution of the MWIEST problem. We also give Table 2 to show the calculation of the MWIEST game. The example can be explained as follows:

STAGE τ_1

- Time slot t_1: Node 1 chooses the set $\{e_1\}$ as its BR with the set of local inner edges $\emptyset$, maximum payoff 0, the set of global inner edges $\emptyset$, and the value of objective function $\Phi = 0$.
- Time slot t_2: Node 2 chooses the set $\{e_1, e_2, e_3, e_4\}$ as its BR with the set of local inner edges $\emptyset$, maximum payoff 0, the set of global inner edges $\emptyset$, and the value of objective function $\Phi = 0$.
- Time slot t_3: Node 3 chooses the set $\{e_3, e_4\}$ as its BR with the set of local inner edges $\emptyset$, maximum payoff 0, the set of global inner edges $\emptyset$, and the value of objective function $\Phi = 0$.

- Time slot t_4: Node 4 chooses the set $\{e_2, e_4, e_6\}$ as its BR with the set of local inner edges $\{e_2\}$, maximum payoff -3, the set of global inner edges $\{e_2\}$, and the value of objective function $\Phi = -3$.
- Time slot t_5: Node 5 chooses the set $\{e_6, e_9\}$ as its BR with the set of local inner edges $\emptyset$, maximum payoff -7, the set of global inner edges $\{e_2, e_4\}$, and the value of objective function $\Phi = -7$.
- Time slot t_6: Node 6 chooses the set $\{e_2, e_3, e_4, e_6, e_9\}$ as its BR with the set of local inner edges $\{e_2, e_4\}$, maximum payoff -7, the set of global inner edges $\{e_2, e_4\}$, and the value of objective function $\Phi = -7$.
- Time slot t_7: Node 7 chooses the set $\{e_9, e_{11}\}$ as its BR with the set of local inner edges $\{e_9\}$, maximum payoff -1, the set of global inner edges $\{e_2, e_4, e_9\}$, and the value of objective function $\Phi = -8$.
- Time slot t_8: Node 8 chooses the set $\{e_{11}\}$ as its BR with the set of local inner edges $\emptyset$, maximum payoff 0, the set of global inner edges $\{e_2, e_4, e_9\}$, and the value of objective function $\Phi = -8$.

According to the proof of Theorem 1, we note that at the stage t_1, the value of Φ decreases until all nodes finish their BR selection processes. At the stage t_i, $i > 1$, the value of Φ will increase after each node chooses its BR.

STAGE τ_2

- Time slot t_1: Node 1 chooses the set $\{e_1\}$ as its BR with the set of local inner edges $\emptyset$, maximum payoff 0, the set of global inner edges $\{e_2, e_4, e_9\}$, and the value of objective function $\Phi = -8$. Node 1 does not change its BR.
- Time slot t_2: Node 2 chooses the set $\{e_1, e_3, e_8, e_5\}$ as its BR with the set of local inner edges $\{e_3, e_5, e_8\}$, maximum payoff -5, the set of global inner edges $\{e_3, e_5, e_8, e_9\}$, and the value of objective function $\Phi = -6$. Node 2 changes its BR.
- Time slot t_3: Node 3 chooses the set $\{e_3, e_8\}$ as its BR with the set of local inner edges $\{e_3, e_8\}$, maximum payoff -3, the set of global inner edges $\{e_3, e_5, e_8, e_9\}$, and the value of objective function $\Phi = -6$. Node 3 changes its BR.
- Time slot t_4: Node 4 chooses the set $\{e_5, e_7\}$ as its BR with the set of local inner edges $\emptyset$, maximum payoff 0, the set of global inner edges $\{e_3, e_8, e_9\}$, and the value of objective function $\Phi = -4$. Node 4 changes its BR.
- Time slot t_5: Node 5 chooses the set $\{e_5, e_7, e_9\}$ as its BR with the set of local inner edges $\{e_9\}$, maximum payoff -1, the set of global inner edges $\{e_3, e_8, e_9\}$, and the value of objective function $\Phi = -4$. Node 5 changes its BR.
- Time slot t_6: Node 6 chooses the set $\{e_5, e_7, e_8, e_9\}$ as its BR with the set of local inner edges $\{e_8, e_9\}$, maximum payoff -3, the set of global inner edges $\{e_3, e_8, e_9\}$, and the value of objective function $\Phi = -4$. Node 6 changes its BR.
- Time slot t_7: Node 7 chooses the set $\{e_9, e_{11}\}$ as its BR with the set of local inner edges $\{e_9\}$, maximum payoff -1, the set of global inner edges $\{e_3, e_8, e_9\}$, and the value of objective function $\Phi = -4$. Node 7 does not change its BR. Node 7 does not change its BR.
- Time slot t_8: Node 8 chooses the set $\{e_{11}\}$ as its BR with the set of local inner edges $\emptyset$, maximum payoff 0, the set of global inner edges $\{e_3, e_8, e_9\}$, and the value of objective function $\Phi = -4$. Node 8 does not change its BR. Node 8 does not change its BR.

STAGE τ_3

In stage τ_3, all nodes do not change their BRs. Thus, the NE point is a global strategy of all BRs in stage τ_2, as shown in Figure 5. The MWIEST game stops at this stage.

IV. THE GAME THEORY BASED CLUSTERING PROTOCOL: GBC PROTOCOL

In this section, we describe the algorithms NIE, BRS, INS which support the FC algorithm to form clusters. After that, the GBC protocol is described based on the FC algorithm and the proposed cluster maintenance algorithm to construct high stable clusters supporting multicast routing in CRAHNs.

A. THE PROPOSED NODE INFORMATION EXCHANGE ALGORITHM: NIE ALGORITHM

A node uses the proposed NIE algorithm to exchange node's information and neighbors' information list with its neighbors. At each node v, the NIE algorithm, as shown in Figure 6, can be expressed as follows:

- Process 1: Node v exchanges its node's information with its neighbors.
  - Step 1. Node v broadcasts an INFO-PACKET to all neighbors by using control channel. Go to Step 2. The INFO-PACKET contains the following fields: ⟨ Packet_Type, Sender_ID, Receiver_ID, Location, Speed, Moving_Direction, Role, Receiving_Channel ⟩, as shown in Figure 6.
  - Step 2. If node v receives an INFO-PACKET, it will record the sender's information (Sender_ID, Location, Speed, Moving_Direction, Role, Receiving_Channel) into the neighbors' information table. Go to Step 3.
- Process 2: Node v exchanges its neighbors' information list with its neighbors.
  - Step 3. Node v broadcasts a NBI-PACKET to all neighbors by using control channel. Go to Step 4. The NBI-PACKET contains the following fields: ⟨ Packet_Type, Sender_ID, Receiver_ID, Neighbor_ID_List, Licensed_Channel ⟩, as shown in Figure 6.
  - Step 4. If node v receives a NBI-PACKET, it will record the neighbors' information list of sender (Neighbors_ID_List, Lisenced_Channel) into the

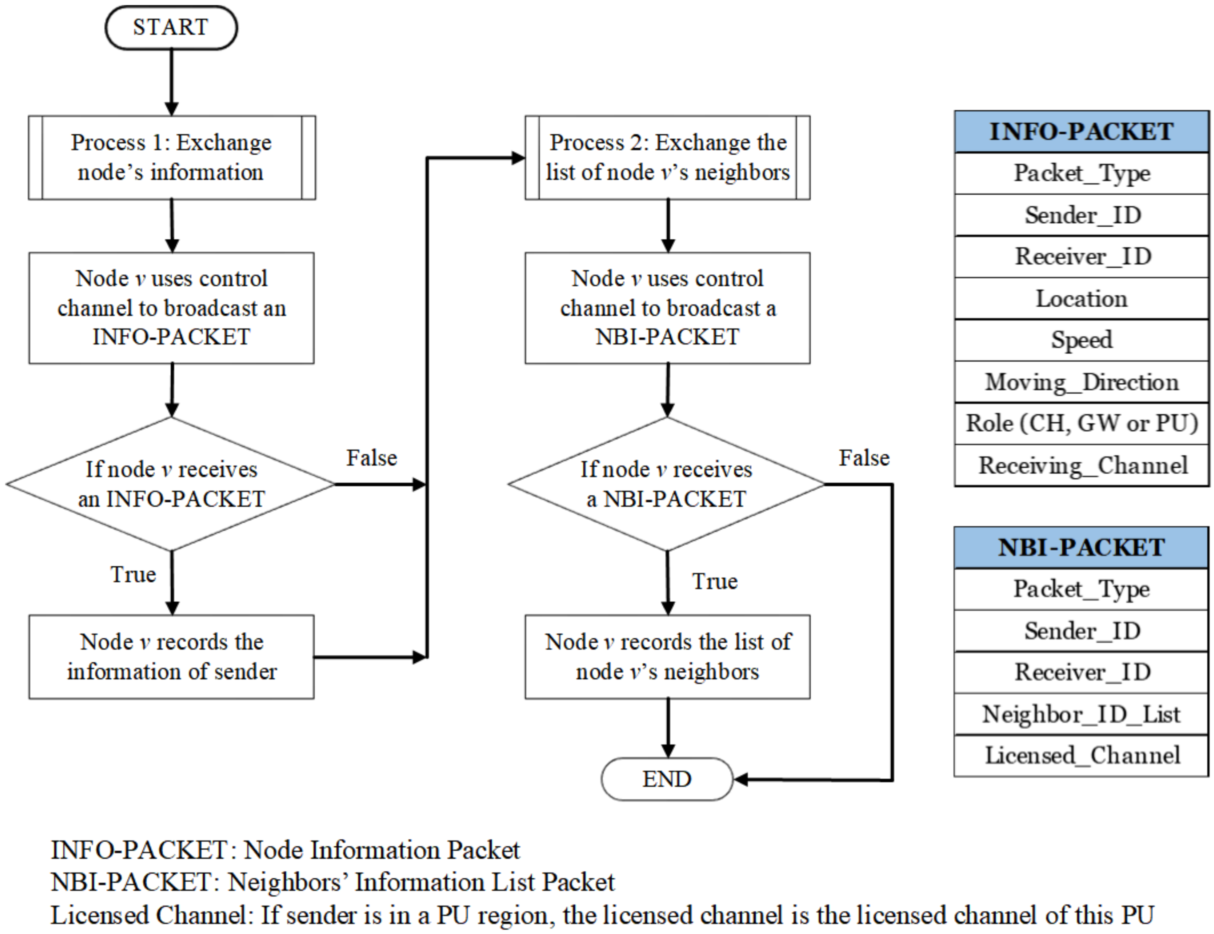


FIGURE 6. **The proposed node information exchange algorithm: NIE algorithm (at Node *v*).**

neighbor' information table based on Sender_ID. The NIE algorithm is ended.

B. THE PROPOSED BEST RESPONSE SELECTION ALGORITHM: BRS ALGORITHM

All nodes in a certain group of nodes use the proposed BRS algorithm to choose a NE point which is an approximate solution of the MWIEST problem. In this algorithm, these nodes have to choose their BRs one by one such that a node with smaller ID will choose its BR before a node with larger ID by using different time slots. This process is looped until each node can't find a new BR different from its current BR. A combination of all final BRs is a NE point which is used for the INS algorithm. In the BRS algorithm, each node needs to determine the set of all strategies which can be found by considering all local WSTs. To reduce the complexity of the BRS algorithm, each node v only considers 5 nearest neighbors at most. We have $|\delta[v]| \leq 6, \forall i \in \mathsf{V}$. At node i-th, the BRS algorithm, as shown in Figure 7, can be explained as follows:

- Step 1. If time slot of node i-th is active, go to Step 2. Otherwise, go to Step 7.
- Step 2. All strategies of node i are assigned as unmarked strategies. The initial value of the max_payoff is assigned to $-\infty$. The new BR is assigned to the current BR. Go to Step 3.
- Step 3. If there exists at least one unmarked strategy, go to Step 4. Otherwise, go to Step 6.
- Step 4. If s is an unmarked strategy and satisfies the SS rule, go to Step 5. Otherwise, go back to Step 3.
- Step 5. The algorithm calculates payoff(s) in Eq. (12) and marks s as a marked strategy. If the payoff(s) greater than max_payoff, the value of max_payoff is assigned to payoff(s) and the new BR is assigned to s. Go back to Step 3.
- Step 6. If the new BR is different from the current BR, the current BR is assigned to new BR, the max_payoff is assigned to payoff(s) and node i-th broadcasts a BR-PACKET to all neighbors by using control channel and go to Step 7. Otherwise, the BRS algorithm is end. The BR-PACKET contains the following fields: ⟨ Packet_Type, Sender_ID, Receiver_ID, BR_Information ⟩, as shown in Figure 7.
- Step 7. If time slot of node N-th is over, node i updates its new time slot go back to Step 1. Otherwise,

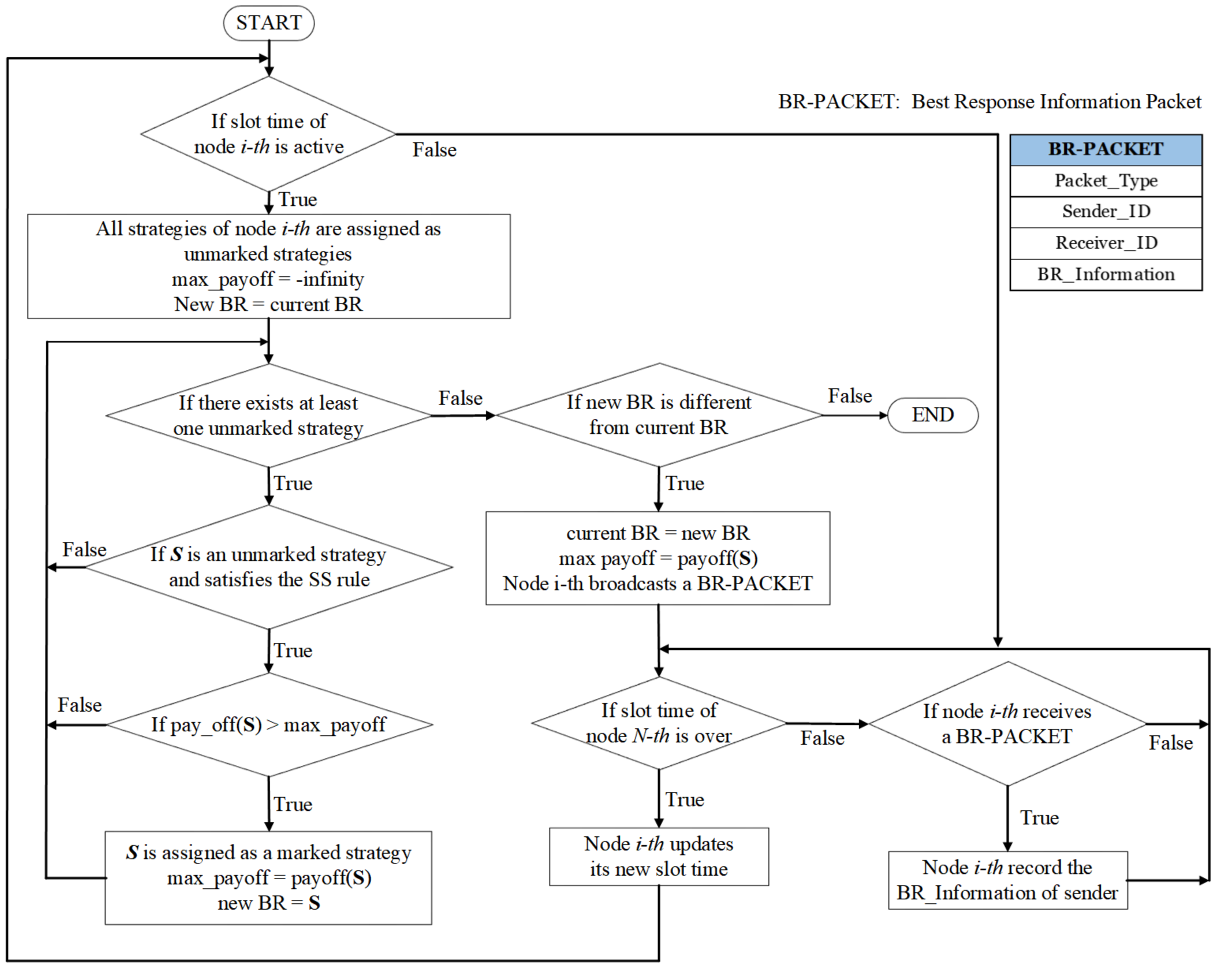


FIGURE 7. The proposed best response selection algorithm: BRS algorithm (at Node i-th).

if node i-th receives a BR-PACKET, it will record the BR_Information of sender.

C. THE PROPOSED INTERMEDIATE NODE SELECTION ALGORITHM: INS ALGORITHM

After all nodes in a certain group of nodes finish the BRS algorithm, these nodes use the proposed INS algorithm, as shown in Figure 8, to determine intermediate nodes. At node v, the INS algorithm can be explained easily as follows. Node v uses the information of its BR to calculate its degree (the number of edges that are incident to node v) in graph $G_v = (\delta[v], \mathsf{E}(\mathsf{BR}))$. If degree of node v is strictly greater than 1, it will become as an intermediate node.

D. THE BEST CHCANDIDATE

1) THE NODE-WEIGHT FUNCTION

To support the cluster head selection process, we define a node weight function NW of node v as

$$\mathsf{NW}(v) = \frac{\sum_{u \in \delta(v)} \mathsf{W}(v, u)}{|\delta(v)|}. \tag{19}$$

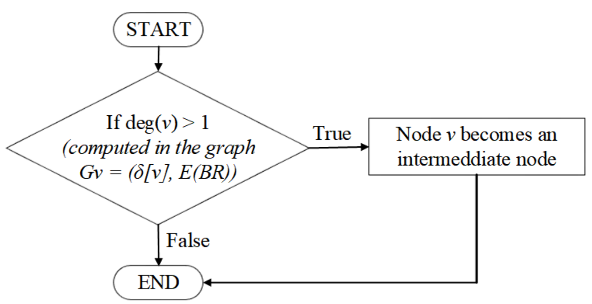


FIGURE 8. The proposed intermediate node selection algorithm: INS algorithm (at Node v).

2) THE BEST CH CANDIDATE

Let v be a node and the set $\mathtt{U}$ be a set of intermediate nodes (not including v) in neighbors list of node v. We denote c_v as the receiving channel of node v, and $\mathtt{U}^*$ is a set of intermediate nodes in $\mathtt{U}$ which are not affected by the licensed channel. An intermediate node u in the SetIN is the best CH candidate of node v if it satisfies one of following conditions

1) The set $\mathtt{U}^*$ is not empty, and node $u \in \mathtt{U}^*$ has $c_u = c_v$, and $\mathsf{NW}(u) = \min\{\mathsf{NW}(u') \mid (u' \in \mathtt{U}^*) \wedge (c_{u'} = c_v)\}$

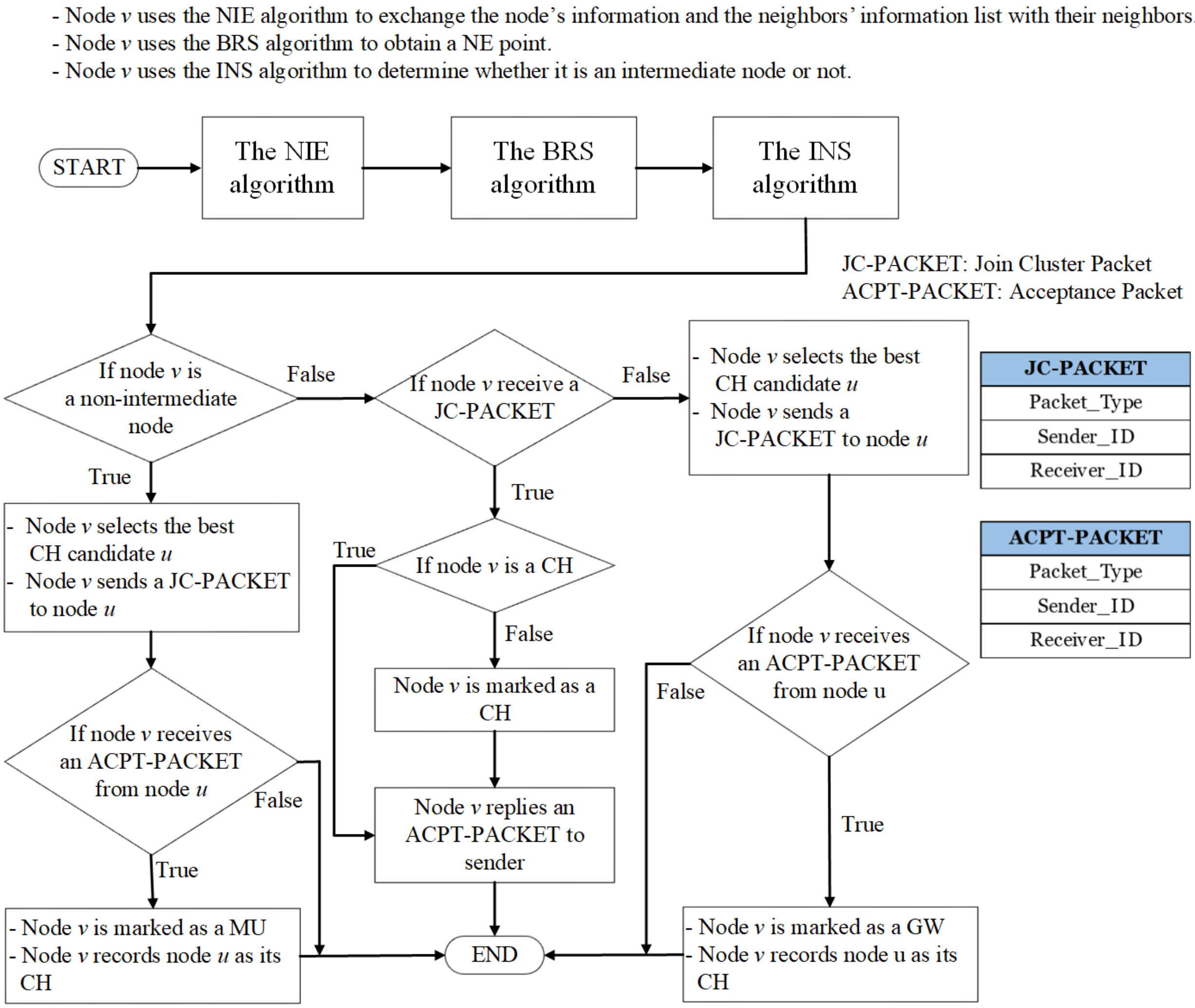


FIGURE 9. The proposed forming cluster algorithm: FC algorithm (at Node v).

2) The set U^* is not empty, and node $u \in U^*$ has $c_u \neq c_v$, and $c_{u'} \neq c_v, \forall u' \in U^*$ and $NW(u) = \min\{NW(u') \mid u' \in U^*\}$
3) The set U^* is empty, and node $u \in U$ has $c_u = c_v$, and $NW(u) = \min\{NW(u') \mid (u' \in U) \wedge (c_{u'} = c_v)\}$
4) The set U^* is empty, and node $u \in U$ has $c_u \neq c_v$, and $c_{u'} \neq c_v, \forall u' \in U$ and $NW(u) = \min\{NW(u') \mid u' \in U\}$

E. THE PROPOSED FORMING CLUSTER ALGORITHM: FC ALGORITHM

After all nodes in a certain group of nodes finish the INS algorithm. These nodes use the FC algorithm to form clusters. At a node v, the proposed FC algorithm, as shown in Figure 9, can be presented as follows:

- Step 1. Node v uses the NIE algorithm to exchange the node's information and the neighbors' information list with their neighbors. Go to Step 2.
- Step 2. Node v uses the BRS algorithm to obtain a NE point which is an approximate solution of the MWIEST problem. Go to Step 3.
- Step 3. Node v uses the INS algorithm to determine whether it is an intermediate node or not. Go to Step 4.
- Step 4. If node v is a non-intermediate node, go to Step 5. Otherwise, go to Step 7.
- Step 5. (Determining a MU from a non-intermediate node)
  Node v chooses a best CH candidate u and sends a JC-PACKET by using control channel to node u. Go to Step 6. The JC-PACKET contains the following fields: ⟨ Packet_Type, Sender_ID, Receiver_ID ⟩, as shown in Figure 9.
- Step 6. If node v receives an ACPT-PACKET from node u, node v will be marked as a MU, records node u as its CH. The FC algorithm is end. The ACPT-PACKET contains the following fields: ⟨ Packet_Type, Sender_ID, Receiver_ID ⟩, as shown in Figure 9.
- Step 7. (Determining a GW or a CH from an intermediate node)
  If node v receives a JC-PACKET, go to Step 8. Otherwise, go to Step 9.

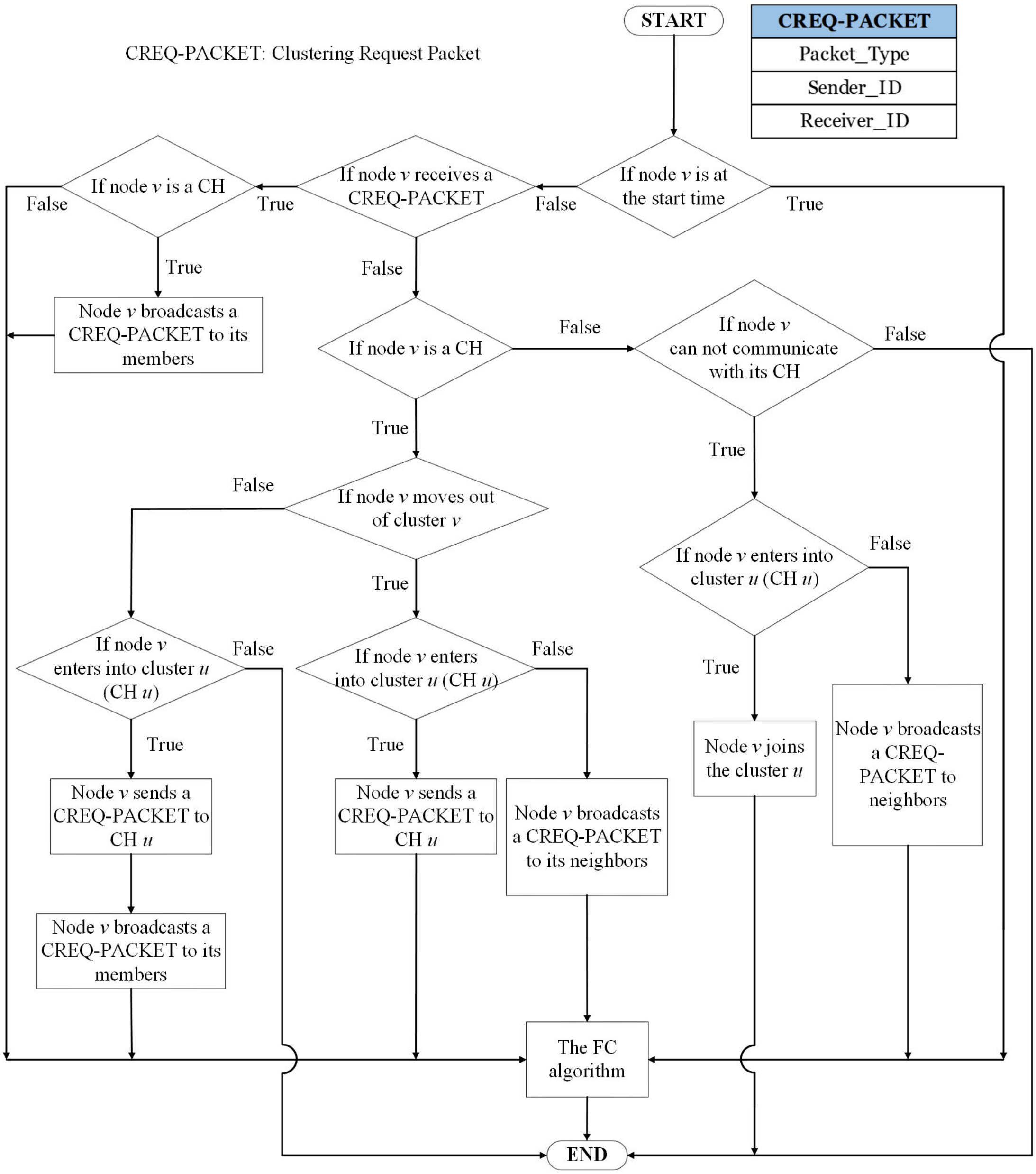


FIGURE 10. The proposed game based clustering protocol: GBC protocol (at Node *v*).

- Step 8. If node v is not a CH, node v is marked as a CH. Node v replies an ACPT-PACKET by using control channel to sender. The FC algorithm is end.
- Step 9. Node v chooses a best CH candidate u and sends a JC-PACKET to node u by using control channel. Go to Step 10.
- Step 10. If node v receives an ACPT-PACKET from node u, node v will be marked as a GW, records node u as its CH. The FC algorithm is end.

F. THE PROPOSED GAME BASED CLUSTERING PROTOCOL: GBC PROTOCOL

In summary, we propose the GBC protocol, as shown in Figure 10, by using the proposed FC algorithm which is supported by NIE, BRS and INS algorithms and the proposed cluster maintenance algorithm. Particularly, at the beginning, all nodes use the FC algorithm to form clusters for the whole network. After that, the cluster maintenance algorithm which is explained in details in subsection IV-G is designed to maintain clusters by using the FC algorithm locally. At a node v, the proposed GBC protocol can be presented as follows:

- Step 1. If node v is at the beginning time, go to Step 10. Otherwise, go to Step 2.
- Step 2. If node v receives a CREQ-PACKET (clustering request packet), go to Step 3. Otherwise, go to Step 4. The CREQ-PACKET contains the following fields: ⟨ Packet_Type, Sender_ID, Receiver_ID ⟩, as shown in Figure 10.

- Step 3. If node v is a cluster head, the CH_v broadcasts a CREQ-PACKET to all its members and go to Step 10. Otherwise go to Step 10.
- Step 4. If node v is a CH, go to Step 5. Otherwise, go to Step 8.
- Step 5. If the CH_v moves out of cluster v, i.e., the CH_v can not communicate with all its members, go to Step 6. Otherwise, go to Step 7.
- Step 6. If the CH_v enters into cluster u (CH u), the CH_v sends a CREQ-PACKET to CH u and go to Step 10. Otherwise, node v broadcasts a CREQ-PACKET to its neighbors and go to Step 10.
- Step 7. If the CH_v enters into cluster u (CH u), the CH_v sends a CREQ-PACKET to CH u and broadcasts a CREQ-PACKET to all its members and go to Step 10. Otherwise, go to Step 11.
- Step 8. If node v can not communicate with its CH, go to Step 9. Otherwise, go to Step 11.
- Step 9. If node v enters into cluster u (CH u), node v joins the cluster u and go to Step 11. Otherwise, node v broadcasts a CREQ-PACKET to its neighbors and go to Step 10.
- Step 10. Node v uses the proposed FC algorithm to form clusters and go to Step 11.
- Step 11. The GBC protocol is end.

G. THE EXPLANATION OF THE CLUSTER MAINTENANCE ALGORITHM

Because nodes in the network randomly move, some critical problems can be occurred as follows: (i) a member node moves out of its cluster or it can not communicate with its cluster head, (ii) a cluster head moves out of its cluster or dies and (iii) multiple clusters move into the same cluster. These critical problems are addressed by the proposed cluster maintenance algorithm as follows:

(i) If a member node v moves out of its cluster or it can not communicate with its cluster head, the following process will be operated:

(i.1) If node v does not enter into any existing cluster, node v broadcasts a CREQ-PACKET to all its neighbors to form a new cluster. The re-clustering process is locally operated by using the FC algorithm at node v and its neighbors. This problem is shown in Steps 2, 3, 9 and 10 of the GBC protocol.

(i.2) If node v enters into a cluster u, node v joins the cluster u. This problem is shown in Steps 8 and 9 of the GBC protocol.

(ii) If a cluster head v moves out of its cluster or dies, the following processes will be operated:

(ii.1) If the CH_v enters into cluster u (CH_u), the CH_v sends a CREQ-PACKET to CH_u. When CH_u receives a CREQ-PACKET, it broadcasts a CREQ-PACKET to its members to form a new cluster. The re-clustering process is locally operated by using the FC algorithm at CH_v, CH_u and cluster u's members. This problem is shown in Steps 2, 3, 6 and 10 of the GBC protocol.

(ii.2) If CH_v does not enter into any existing cluster, the CH_v broadcasts a CREQ-PACKET to its neighbors to form a new cluster. The re-clustering process is locally operated by using the FC algorithm at CH_v and CH_v's neighbors. This problem is shown in Steps 2, 3, 6 and 10.

(ii.3) When CH_v moves out of its cluster or dies, all cluster v's members can not communicate with CH_v. These members will be operated as problem (i).

(iii) If multiple clusters $1, 2, \ldots$ and K are in the same cluster, each CH_i, $i \in \{1, \ldots K\}$ broadcasts a CREQ-PACKET to other cluster heads. If each CH_i, $i \in \{1, \ldots K\}$ receives a CREQ-PACKET, it will broadcast a CREQ-PACKET its members to form a new cluster. The re-clustering process is locally operated by using the FC algorithm at CH_i, $i \in \{1, \ldots K\}$ and their members. This problem is shown in Steps 2, 3, 5, 7 and 10.

V. PERFORMANCE EVALUATION

In this section, we evaluate the performance of the proposed GBC protocol under different *max-speeds* of mobile nodes (20 km/h, 40 km/h, 60 km/h and 80 km/h). The simulation environments and parameters are presented in Table 3. The simulation of the proposed GBC protocol is implemented in OMNET++ platform. To balance the effects of Link-Stability function (LS) and Link-Connectivity function (LC) in the intermediate node selection algorithm, the parameters w_1, w_2 of the link-weight function (6) are chosen so that their value are nearly equal 0.5.

TABLE 3. Simulation environments and parameters.

Network size	1×1 km^2
Number of nodes	50
Number of PUs	2
Session length	5 s
Transmission range of nodes	250 m
Coverage range of PUs	250 m
Number of data channels	5
Max speed of nodes	20, 40, 60 and 80 km/h
Mobility model	Random waypoint
The time of the simulation	1000 seconds

To measure and valuate the stability of obtained clusters, we propose some metrics which are used in this article as follows:

- The number of cluster heads: It presents a measure of the average number of cluster heads in each session. It is also the number of obtained average clusters in the network.
- The number of new cluster heads: It presents a measure of the average number of new cluster heads between two consecutive sessions. It is used to measure the stability of clusters through several sessions.
- The number of intermediate nodes: It presents a measure of the average number of intermediate nodes which

includes the number of cluster heads and gateways in each session.

- The number of new intermediate nodes: It presents a measure of the average number of new intermediate nodes between two consecutive sessions. It is used to measure the stability of cluster heads and gateways which is provided by a solution of the MWIEST game.
- The number of broken member links: It presents the average number of broken links between member users and their cluster heads in each session. It is used to measure the stability of links between member users and their cluster heads.
- The number of broken links between intermediate nodes: It presents the average number of links between gateways and cluster heads which is broken in each session. It is used to measure the stability of links between gateways and cluster heads which is provided by a solution of the MWIEST game.
- The number of control packets, called the control overheads: It presents the average number of control packets which is used for clustering process in each session. It shows how many the average number of control packets are needed for the clustering process used to obtain clusters in each session.

To show the advantage of the proposed GBC protocol, we compare it with two protocols that are lowest ID based clustering and highest connectivity based clustering protocols [8], [49] in the same environment and parameters as GBC protocol. The simulation results also indicate that the proposed GBC protocol always obtains cluster heads that avoid the affected region of licensed channels based on the choosing best CH process.

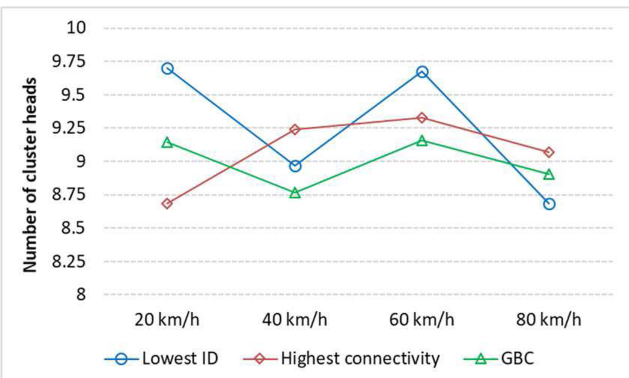


FIGURE 11. The number of cluster heads.

Figure 11 shows the number of cluster heads in each session as a function of node mobility. In GBC protocol, the number of clusters ranges from 8.76 to 9.15, i.e., the number of members in each cluster is around 5.5. The reason is that the BRS algorithm uses the proposed MIWEST game to obtain a reasonable sub-optimal solution. Thus, the INS algorithm provides the reasonable set of intermediate nodes which is used for the CH and GW selection processes in the FC algorithm. Figure 11 also shows that the number of clusters of the lowest ID and highest connectivity based clustering protocols are around 9. Moreover, in the proposed GBC protocol the number of clusters does not depend on the speed of nodes but it depends on the topology of network [44], [45].

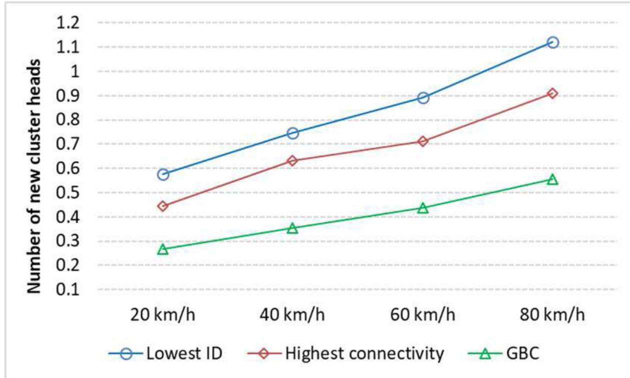


FIGURE 12. The number of new cluster heads.

Figure 12 indicates number of new cluster heads between two consecutive sessions as a function of node mobility. In GBC protocol, this number is increased from 0.26 to 0.55 corresponding to the max speed from 20 km/h to 80 km/h. We can also observe that the number of new cluster heads of the lowest ID and highest connectivity based clustering protocol are higher than that of the GBC protocol at each value of max speed. The reason is that the MWIEST game and link-weight function support BRS and INS algorithms to obtain the high stability set of intermediate nodes which is used for the FC algorithm to select the best CHs. These CHs can be unchanged through multiple sessions.

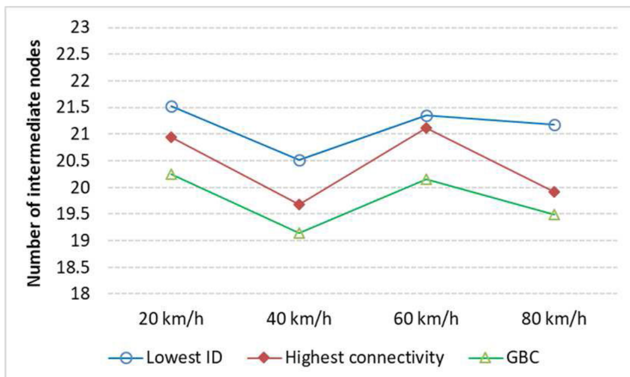


FIGURE 13. The number of intermediate nodes.

Figure 13 illustrates the number of intermediate nodes in each session as a function of node mobility. In GBC protocol, this number implies the total number of CHs and GWs that ranges from 19.1 to 20.2. Based on approximately 9 CHs in Figure 11, we have that the number of GWs is around 11. It means that on average, each cluster has 1.22 GWs. As can be observed in Figure 13, the lowest ID and highest connectivity based clustering protocols also provide the number of about

19.6 to 21.5 intermediate nodes, respectively. By the same reason as the number of CHs in Figure 11, the FC algorithm provides a reasonable number of intermediate nodes in each session. We can also see that in the proposed GBC protocol the number of intermediate nodes does not depend on the speed of nodes but it depends on the topology of network.

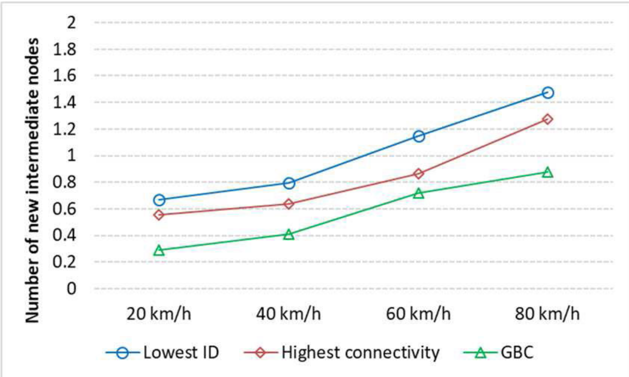


FIGURE 14. The number of new intermediate nodes.

Figure 14 presents the number of new intermediate nodes between two consecutive sessions as a function of node mobility. In GBC protocol, this number is changed from 0.29 to 0.87 corresponding to the speed from 20 km/h to 80 km/h. Figure 14 also presents that the numbers of intermediate node changes of the lowest ID and highest connectivity based clustering protocols are higher than that of the GBC prtocol at each value of max speed. This result also has the same reason as the number of cluster heads in Figure 12. The FC algorithm provides the stability set of intermediate nodes which changes a little bit over multiple sessions.

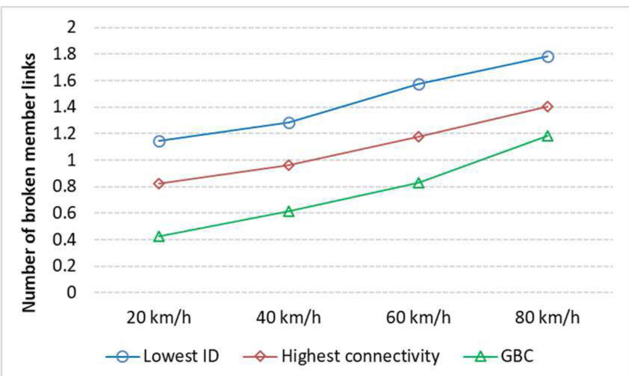


FIGURE 15. The number of broken member links.

Figure 15 shows the number of broken member links in each session. In GBC protocol, this number ranges from 0.42 to 1.18 corresponds with the speed from 20 km/h to 80 km/h. This number is a small number, i.e., the links between member users and cluster heads is stable through sessions. Besides, the numbers of broken member links of the lowest ID and highest connectivity based clustering protocols are higher than that of the GBC prtocol at each value of max speed. The reason is that the MWIEST game helps each node to choose a reasonable BR (a local minimum weighted inner edge spanning tree) which contributes to find stable clusters in GBC protocol. The obtained clusters can keep the connection between MUs and CHs in a stable way.

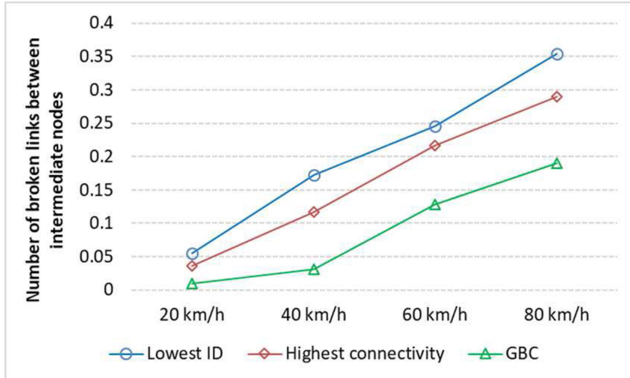


FIGURE 16. The number of broken links between intermediate nodes.

Figure 16 indicates the number of broken links between intermediate nodes in each session. In GBC protocol, this number increases from 0.01 to 0.19 corresponding to the speed from 20 km/h to 80 km/h. It means that the number of broken virtual backbone links in each session is too small, i.e., the set of intermediate nodes is stable through sessions. Moreover, the numbers of broken virtual backbone links of the lowest ID and highest connectivity based clustering protocols are higher than that of the GBC protocol at each value of max speed. The reason is that the link-weight function helps the BRS and INS algorithms to construct a high stable set of intermediate nodes. Thus, the GBC protocol can provide high stable clusters which can keep connection together in a high stable way.

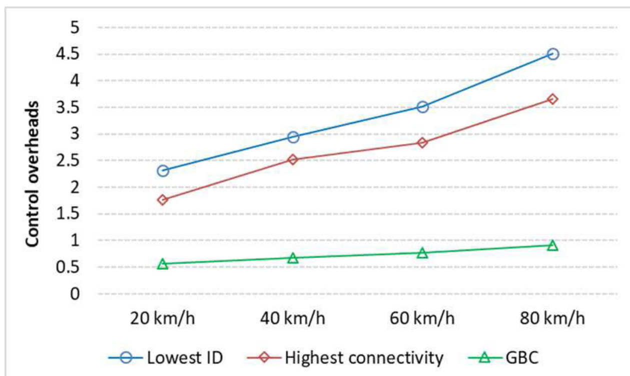


FIGURE 17. The control overheads.

Figure 17 presents the control overheads as a function of node mobility. In GBC protocol, the control overheads change from 0.55 to 0.9. At the first time, the FC algorithm spends around 250 control packets to form clusters for the whole network (50 nodes). Next time, because the GBC protocol locally implements the re-clustering process by using the proposed cluster maintenance algorithm, it spends around 60 control packets for 10 nodes. At 20 km/h, because the obtained clusters of GBC protocol are used through average 3.84 sessions, the number of control packets is around $60/3.85 = 15.58$ packets for each session. Thus, the GBC protocol spends around $(250/200 + 15.58)/(50 \text{ nodes}) = 0.56$ control packets per 1 node for each session. At 80 km/h, because the obtained clusters of GBC protocol are used through average 1.8 sessions, the number of control packets is around $60/1.8 = 33.33$ packets for each session. Thus, the GBC protocol spends around $(250/200 + 33.33)/(50 \text{ nodes}) = 0.91$ control packets per 1 node for each session. The reason is that the GBC protocol helps to obtain high stable clusters. These clusters can be reused through multiple sessions resulting in the reduction of control packets in GBC protocol. Moreover, the cluster maintenance algorithm is locally implemented to significantly reduce the control packets for re-clustering process. Finally, Figure 17 also indicates that the control overheads of the lowest ID and highest connectivity based clustering protocols are higher than that of GBC protocol at each value of max speed.

VI. CONCLUSION

In this article, we proposed a GBC protocol to achieve high stable clusters which supports multicast routing in CRAHNs. Firstly, we propose the MWIEST game to model the minimum connected weighted inner edge spanning tree problem as a game. In this game, the weight of each edge is proposed as a link-weight function which is a combination of the link-stability function and the link-connectivity ratio function. Secondly, we prove that the MWIEST game is an exact potential game and there exists at least one Nash equilibrium (NE) point which is an approximate solution of the MWIEST problem (Theorem 1). Besides, we also prove that best responses (BRs) of the game converges to a NE in $6^4 \times N$ iterations at most, where N is the total number of nodes (Theorem 2). Thirdly, based on the MWIEST game, we propose four algorithms including the node information exchange (NIE), the best response selection (BRS), the intermediate nodes selection (INS) and the forming cluster (FC). Specifically, the algorithms NIE, BRS and INS provide a set of intermediate nodes (SetIN) which supports the FC algorithm to form clusters. Finally, we propose the GBC protocol which is combination of the FC algorithm and the proposed cluster maintenance algorithm to construct high stable clusters supporting multicast routing in CRAHNs. Moreover, each obtained cluster includes most members having the same receiving channel which avoids the affected regions of licensed channels. For performance evaluation, we implement the GBC protocol in OMNET++ platform to show that it provides better performance than the lowest ID and highest connectivity protocols in terms of network stability and control overheads.

In future works, we will develop deep learning frameworks to solve clustering and routing problems in CRAHNs. Particularly, we will implement the cross-layer design to optimize system parameters in physical, data link, and network layers under the developed deep learning frameworks to improve the PDR, control overhead, and routing delay in the future Internet-of-Things CRAHNs.

REFERENCES

[1] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[2] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Comput. Netw.*, vol. 50, no. 13, pp. 2127–2159, Sep. 2006.

[3] I. F. Akyildiz, W.-Y. Lee, and K. R. Chowdhury, "CRAHNs: Cognitive radio ad hoc networks," *Ad Hoc Netw.*, vol. 7, no. 5, pp. 810–836, Jul. 2009.

[4] M. Osman, S. K. Syed-Yusof, N. Malik, and S. Zubair, "A survey of clustering algorithms for cognitive radio ad hoc networks," *Wireless Netw.*, vol. 24, no. 5, pp. 1451–1475, Nov. 2016.

[5] M. Ahmad, A. Hameed, A. A. Ikram, and I. Wahid, "State-of-the-art clustering schemes in mobile ad hoc networks: Objectives, challenges, and future directions," *IEEE Access*, vol. 7, pp. 17067–17081, 2019.

[6] L. Raja and S. S. Baboo, "An overview of MANET: Applications, attacks and challenges," *Int. J. Comp. Sci. Mobile Comput.*, vol. 3, no. 1, pp. 408–417, 2014.

[7] M. Y. Arafat and S. Moh, "A survey on cluster-based routing protocols for unmanned aerial vehicle networks," *IEEE Access*, vol. 7, pp. 498–516, 2019.

[8] B. An and S. Papavassiliou, "A mobility-based clustering approach to support mobility management and multicast routing in mobile ad-hoc wireless networks," *Int. J. Netw. Manage.*, vol. 11, no. 6, pp. 387–395, Nov./Dec. 2001.

[9] S. Chinara and S. K. Rath, "A survey on one-hop clustering algorithms in mobile ad hoc networks," *J. Netw. Syst. Manage.*, vol. 17, nos. 1–2, pp. 183–207, Jun. 2009.

[10] J. Y. Yu and P. H. J. Chong, "A survey of clustering schemes for mobile ad hoc networks," *IEEE Commun. Surveys Tuts.*, vol. 7, no. 1, pp. 32–48, 1st Quart., 2005.

[11] J. Wu and H. Li, "A dominating-set-based routing scheme in ad hoc wireless networks," *Telecomm. Syst. J.*, vol. 3, pp. 63–84, Sep. 2001.

[12] Z. Yang, D. Fu, L. Han, and S. T. Jhang, "Improved route discovery based on constructing connected dominating set in MANET," *Int. J. Distrib. Sensor Netw.*, vol. 11, no. 8, Aug. 2015, Art. no. 612102.

[13] T. N. Tran, T.-V. Nguyen, and B. An, "An efficient connected dominating set clustering based routing protocol with dynamic channel selection in cognitive mobile ad hoc networks," *Electronics*, vol. 8, no. 11, p. 1332, Nov. 2019.

[14] W. Bednarczyk and P. Gajewski, "An enhanced algorithm for MANET clustering based on weighted parameters," *Universal J. Commun. Netw.*, vol. 1, no. 3, pp. 88–94, 2013.

[15] S. Pathak and S. Jain, "A novel weight based clustering algorithm for routing in MANET," *Wireless Netw.*, vol. 22, no. 8, pp. 2695–2704, Nov. 2016.

[16] L. Lazos, S. Liu, and M. Krunz, "Spectrum opportunity-based control channel assignment in cognitive radio networks," in *Proc. IEEE SECON*, Jun. 2009, pp. 135–143.

[17] S. Liu, L. Lazos, and M. Krunz, "Cluster-based control channel allocation in opportunistic cognitive radio networks," *IEEE Trans. Mobile Comput.*, vol. 11, no. 10, pp. 1436–1449, Oct. 2012.

[18] X.-L. Huang, G. Wang, F. Hu, and S. Kumar, "Stability-capacity-adaptive routing for high-mobility multihop cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2714–2729, Jul. 2011.

[19] G. A. Shah, F. Alagoz, E. A. Fadel, and O. B. Akan, "A spectrum-aware clustering for efficient multimedia routing in cognitive radio sensor networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 7, pp. 3369–3380, Sep. 2014.

[20] G. Xu, X. Tan, S. Wei, S. Guo, and B. Wang, "An energy-driven adaptive cluster-head rotation algorithm for cognitive radio network," in *Proc. 1st Int. Conf. Pervas. Comput., Signal Process. Appl.*, Sep. 2010, pp. 138–141.

[21] W. Zhang, Y. Yang, and C. K. Yeo, "Cluster-based cooperative spectrum sensing assignment strategy for heterogeneous cognitive radio network," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2637–2647, Jun. 2015.

[22] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Proc. Conf. Rec. 38th Asilomar Conf. Signals, Syst. Comput.*, vol. 1, 2004, pp. 772–776.

[23] C. Guo, T. Peng, S. Xu, H. Wang, and W. Wang, "Cooperative spectrum sensing with cluster-based architecture in cognitive radio networks," in *Proc. VTC Spring*, Apr. 2009, pp. 1–5.
[24] O. Morgenstern and J. V. Neumann, *The Theory of Games and Economic Behavior*. Princeton, NJ, USA: Princeton Univ. Press, 1947.
[25] J. F. Nash, "Equilibrium points in N-person games," *Proc. Nat. Acad. Sci. USA*, vol. 36, no. 1, pp. 48–49, Jan. 1950.
[26] D. Yang, X. Fang, and G. Xue, "Game theory in cooperative communications," *IEEE Wireless Commun.*, vol. 19, no. 2, pp. 44–49, Apr. 2012.
[27] R. Massin, C. J. Le Martret, and P. Ciblat, "A coalition formation game for distributed node clustering in mobile ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3940–3952, Jun. 2017.
[28] A. A. Khan, M. Abolhasan, and W. Ni, "An evolutionary game theoretic approach for stable and optimized clustering in VANETs," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4501–4513, May 2018.
[29] Q. Liu and M. Liu, "Energy-efficient clustering algorithm based on game theory for wireless sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 13, no. 11, Nov. 2017, Art. no. 155014771774370.
[30] C. Sun, W. Zhang, and K. B. Letaief, "Cluster-based cooperative spectrum sensing in cognitive radio systems," in *Proc. IEEE ICC*, Jun. 2007, pp. 2511–2515.
[31] U. C. Kozat, G. Kondylis, B. Ryu, and M. K. Marina, "Virtual dynamic backbone for mobile ad hoc networks," in *Proc. IEEE ICC*, vol. 1, Jun. 2001, pp. 250–255.
[32] D. Lin and Q. Wang, "An energy-efficient clustering algorithm combined game theory and dual-cluster-head mechanism for WSNs," *IEEE Access*, vol. 7, pp. 49894–49905, Apr. 2019.
[33] M. M. Afsar, R. T. Crump, and B. H. Far, "Energy-efficient coalition formation in sensor networks: A game-theoretic approach," in *Proc. IEEE CCECE Conf.*, Oct. 2019, pp. 1–6.
[34] S. Kassan, J. Gaber, and P. Lorenz, "Game theory based distributed clustering approach to maximize wireless sensors network lifetime," *J. Netw. Comput. Appl.*, vol. 123, pp. 80–88, Dec. 2018.
[35] L. Yang, Y. Lu, L. Xiong, Y. Tao, and Y. Zhong, "A game theoretic approach for balancing energy consumption in clustered wireless sensor networks," *Sensors*, vol. 17, no. 11, p. 2654, Nov. 2017.
[36] S. Pathak, N. Dutta, and S. Jain, "An improved cluster maintenance scheme for mobile AdHoc networks," in *Proc. Conf. ICACCI*, Dec. 2014, pp. 2117–2121.
[37] S. Pathak and S. Jain, "An optimized stable clustering algorithm for mobile ad hoc networks," *EURASIP J. Wireless Commun. Netw.*, vol. 2017, no. 1, pp. 1–11, Mar. 2017.
[38] S. Pathak and S. Jain, "Comparative study of clustering algorithms for MANETs," *J. Statist. Manage. Syst.*, vol. 22, no. 4, pp. 653–664, Jun. 2019.
[39] C.-C. Chiang, H.-K. Wu, W. Liu, and M. Gerla, "Routing in clustered multihop, mobile wireless networks with fading channel," in *Proc. IEEE SICON*, Apr. 1997, pp. 197–211.
[40] P. Enge and P. Misra, "Special issue on global positioning system," *Proc. IEEE*, vol. 87, no. 1, pp. 3–15, Jan. 1999.
[41] E. D. Kaplan and C. J. Hegarty, Eds. *Understanding the GPS: Principles and Applications*. Boston, MA, USA: Artech House, 1996.
[42] E. Hossain, R. Palit, and P. Thulasiraman, "Clustering in mobile wireless ad hoc networks: Issues and approaches," in *Wireless Communications Systems and Networks*, M. Guizani, Ed. Boston, MA, USA: Kluwer, 2004, pp. 383–424.
[43] D. Monderer and L. S. Shapley, "Potential games," *Games Econ. Behav.*, vol. 14, no. 1, pp. 124–143, 1996.
[44] M. E. Elhdhili, L. B. Azzouz, and F. Kamoun, "CASAN: Clustering algorithm for security in ad hoc networks," *Comput. Commun.*, vol. 31, no. 13, pp. 2972–2980, Aug. 2008.
[45] M. Rajkumar, S. Subramanian, and S. Shankar, "Internal and external factors based clustering algorithm for MANET," *Int. J. Eng. Technol.*, vol. 6, no. 4, 1778–1785, 2014.
[46] T. Fujie, "An exact algorithm for the maximum leaf spanning tree problem," *Comput. Oper. Res.*, vol. 30, no. 13, pp. 1931–1944, Nov. 2003.
[47] H. Fernau, J. Kneis, D. Kratsch, A. Langer, M. Liedloff, D. Raible, and P. Rossmanith, "An exact algorithm for the maximum leaf spanning tree problem," *Theor. Comput. Sci.*, vol. 412, no. 45, pp. 6290–6302, Oct. 2011.
[48] *Collected Mathematical Papers*, vol. 13. Cambridge, U.K.: Cambridge Univ. Press, 1897, pp. 26–28, doi: 10.1017/CBO9780511703799.010.
[49] K.-L.-A. Yau, N. Ramli, W. Hashim, and H. Mohamad, "Clustering algorithms for cognitive radio networks: A survey," *J. Netw. Comput. Appl.*, vol. 45, pp. 79–95, Oct. 2014.

THONG NHAT TRAN received the B.S. degree in mathematics-informatics and the M.S. degree in mathematics from Da Lat University, Vietnam, in 2002 and 2006, respectively. He is currently pursuing the Ph.D. degree in electronics and computer engineering with Hongik University, South Korea. His current research interests include ad hoc networks, wireless sensor networks, wireless cognitive radio networks, clustering, and routing.

TOAN-VAN NGUYEN (Student Member, IEEE) received the B.S. degree in electronics and telecommunications engineering and the M.S. degree in electronics engineering from the Ho Chi Minh City University of Technology and Education, Vietnam, in 2011 and 2014, respectively. He is currently pursuing the Ph.D. degree with the Department of Electronics and Computer Engineering, Graduate School, Hongik University, South Korea. His current research interests include mathematical modeling of 5G networks and machine learning for wireless communications.

KYUSUNG SHIM (Student Member, IEEE) received the B.S. degree in computer and information communications engineering and the M.S. degree in information system from Hongik University, South Korea, in 2012 and 2017, respectively, where he is currently pursuing the Ph.D. degree in electronics and computer engineering. His main research interests include wireless communications, physical layer security, and cooperative communications.

BEONGKU AN (Member, IEEE) received the B.S. degree in electronic engineering from Kyungpook National University, South Korea, in 1988, the M.S. degree in electrical engineering from New York University (Polytechnic), New York City, NY, USA, in 1996, and the Ph.D. degree from the New Jersey Institute of Technology (NJIT), Newark, NJ, USA, in 2002. He joined the Faculty of the Department of Software and Communications Engineering, Hongik University, South Korea, where he is currently a Professor. From 1989 to 1993, he was a Senior Researcher with RIST, Pohang, South Korea. He was also a Lecturer and a RA in NJIT, from 1997 to 2002. His current research interests include mobile wireless networks and communications, such as ad-hoc networks, sensor networks, wireless cognitive radio networks, cellular networks, cooperative communication, multicast routing, QoS routing, energy harvesting, physical layer security, M2M/D2D, the IoT, visible light communication (VLC), cross-layer technology, 5G/beyond 5G, NOMA, SWIPT, machine learning and block chain, and mobile cloud computing. He was the President of the IEIE Computer Society (The Institute of Electronics and Information Engineers, Computer Society), in 2012. Since 2013, he has been the General Chair of the International Conference on Green and Human Information Technology (ICGHIT). He was listed in the Marquis Who's Who in Science and Engineering and the Marquis Who's Who in the World.