

Received April 13, 2020, accepted May 5, 2020, date of publication May 12, 2020, date of current version May 26, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2994198

Machine Learning Empowered Spectrum Sharing in Intelligent Unmanned Swarm Communication Systems: Challenges, Requirements and Solutions

XIMING WANG¹, YUHUA XU¹, (Member, IEEE), CHAOHUI CHEN², XIAOQIN YANG¹,
JIAXIN CHEN³, LANG RUAN¹, YIFAN XU¹, AND RUNFENG CHEN¹

¹College of Communications Engineering, Army Engineering University, Nanjing 210007, China

²Guangzhou Haige Communications Group Incorporated Company, Guangzhou 510000, China

³College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

Corresponding author: Yuhua Xu (yuhuaenator@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61771488, Grant 61671473, Grant 61961010, Grant 61631020, and Grant 61971439, and in part by the Natural Science Foundation of Jiangsu Province of China under Grant BK20191329.

ABSTRACT The unmanned swarm system (USS) has been seen as a promising technology, and will play an extremely important role in both the military and civilian fields such as military strikes, disaster relief and transportation business. As the “nerve center” of USS, the unmanned swarm communication system (USCS) provides the necessary information transmission medium so as to ensure the system stability and mission implementation. However, challenges caused by multiple tasks, distributed collaboration, high dynamics, ultra-dense and jamming threat make it hard for USCS to manage limited spectrum resources. To tackle with such problems, the machine learning (ML) empowered intelligent spectrum management technique is introduced in this paper. First, based on the challenges of the spectrum resource management in USCS, the requirement of spectrum sharing is analyzed from the perspective of spectrum collaboration and spectrum confrontation. We found that suitable multi-agent collaborative decision making is promising to realize effective spectrum sharing in both two perspectives. Therefore, a multi-agent learning framework is proposed which contains mobile-computing-assisted and distributed structures. Based on the framework, we provide case studies. Finally, future research directions are discussed.

INDEX TERMS Unmanned swarm system, spectrum sharing, machine learning, multi-agent learning, game theory.

I. INTRODUCTION

With the advance of artificial intelligence (AI), small unmanned robotics and internet of things (IoT), wirelessly connected unmanned swarm systems (USS) are extended from the military and public safety fields to the daily civilian applications [1]–[3]. At present, the intelligent USS has begun to emerge in many military and civilian fields such as collaborative reconnaissance, joint strike and emergency communication [4]–[6].

The unmanned swarm communication system (USCS) provides the necessary information transmission medium for

The associate editor coordinating the review of this manuscript and approving it for publication was Jiankang Zhang¹.

crowd sensing, collaborative control and collective intelligence excitation, which is considered as the “nerve center” of the USS. Spectrum resource is regarded as a key component affecting the communication quality. First, due to the dense deployment of a swarm, neighbor communication nodes need to frequently exchange control information to coordinate behaviours and avoid physical collision. Second, the data collected by agents, such as videos and pictures, rely on high-rate communication techniques to realize fast information fusion and timely backhaul. Last but not the least, when USCS is attacked by malicious radio jammers, a robust anti-jamming communication mechanism is needed. Therefore, the spectrum resource management plays a vital role in the USCS. In summary, an efficient, flexible and adaptable

spectrum resource sharing method is distinctly important, which ensures a reliable network connection and satisfies the great demand for communication resources in a local area. Research in this field is currently a hotspot and has been highly valued by governments, associations, industry and academia.

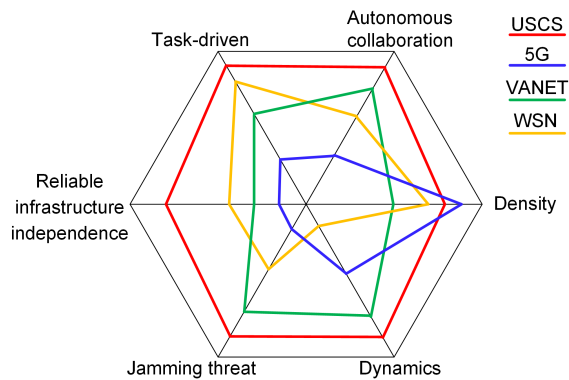


FIGURE 1. Illustration of characteristics of the unmanned swarm communication system. The comparison of the similarities and differences with other communication systems is presented in a radar chart. Note that, the radar chart makes a qualitative comparison which is suitable for most cases. In some extreme cases, there may be discrepancies in the ordering.

However, spectrum resource sharing in USCS is a very complex and challenging problem. In Fig. 1, we illustrate the typical characteristics of USCS and make a qualitative comparison with several communication systems, i.e., the fifth-generation (5G) communication network, vehicle ad hoc network (VANET) and wireless sensor network (WSN). Compared with the 5G communication network [7] and VANET [8], there is no reliable infrastructure such as base station and roadside access point in USCS. Agents in USCS intelligently and autonomously collaborate instead of being controlled by the central station. What's more, when USCS faces the threat of radio jamming attack, agents need to adjust their communication strategy on their own, which cannot be found in the 5G systems. Compared with WSN [9], the USCS is highly dynamic in the three-dimensional space due to the mobile nodes such as unmanned aerial vehicles (UAVs). In addition, the characteristic of “swarm” of USCS also makes the spectrum resource sharing more complex. Besides, the main goal of these agents in the USS is to collaboratively complete tasks. Thus it can be seen that, most of the existing spectrum resource management methods cannot be directly applied to USCS.

Recently, machine learning (ML) empowered spectrum management techniques, for the powerful capabilities, are gaining more and more attention. In this paper, we are motivated to focus on the physical layer spectrum resources of the USCS, and discuss the requirements, challenges and solutions of the system from the perspective of ML. The main contributions of this paper are summarized as follows. First, we summarize five typical challenges within USCS according to the characteristics of USS. To solve them,

we provide two requirements for the communication decision making of agents. Second, we propose two multi-agent learning structures for the intelligent USCS, which conform with the practical application scenarios of USCS. Third, we introduce several case studies of the multi-agent learning structures using existing methods that provide for possible solutions to some of the challenges. In the end, we provide several future research directions.

II. RELATED WORK

The problems of spectrum resource sharing in USCS has drawn more attention due to the rapid development of the wirelessly connected USS. In [10], M. J. Marcus pointed out that while UAV technologies were attracting growing attention, more efficient and effective protocols and methods of spectrum resource management are needed to satisfy the requirements of emerging communication applications. From this perspective, numerous research investigated problems of communication resource optimization in the UAV-involved communication systems [11]–[14]. However, most research considered the cross-system spectrum resource allocation between UAVs and other communication systems such as cellular systems. The intra-system spectrum sharing of UAV swarm communication network was not considered.

Some research investigated distributed spectrum resource sharing in USCS [15]–[18]. All of these works aimed to realize spectrum collaboration for multi-UAV networks in a distributed manner. There are several papers investigating the multi-UAV network in the presence of adversarial jammers [19]. Several survey papers of multi-UAV communication network can be seen in [20]–[22]. In [20], a comprehensive survey of multi-UAV network is provided, including network structures, routing and energy efficiency. H. Wang et. al investigated UAV networks in the cyber-physical-system perspective, and studied the relationship of communication, computation and control in UAV networks [21]. In [22], the spectrum management for UAV swarm networks in millimeter-wave perspective was studied. However, the systematic study on the challenges of intelligent USCS in the spectrum sharing perspective can rarely be seen in these works.

ML-empowered spectrum resource management has drawn a lot of attention recently [23]. It turns out that ML techniques provide wireless communication systems with excellent spectrum resource management capabilities [24]–[28]. However, most of these works studied the spectrum resources management in 5G or IoT communication systems. ML-empowered spectrum resources management methods face new challenges in USCS.

III. CHALLENGES AND REQUIREMENTS OF INTELLIGENT UNMANNED SWARM COMMUNICATION SYSTEMS

A. CHALLENGES

As shown in Fig. 2, compared with the current communication systems, the USCS has new challenges and features.

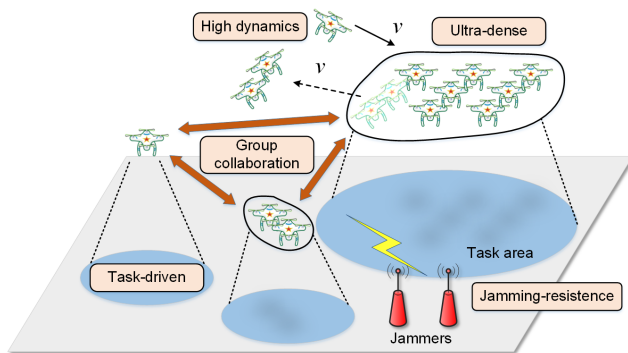


FIGURE 2. Illustration of five challenges in unmanned swarm communication system, where multiple UAVs collaborate to execute the same task in a wide area. Jammers transmit radio jamming to interrupt the communication system.

1) TASK-DRIVEN

The essential and inherent requirement of USCS is task-driven. Different tasks correspond to different task priorities, communication demands, resource utilization priorities and so on. On one hand, tasks determine the USS formation control. To achieve good performance through efficient coordination of multiple agents, accurate maintenance of a task-oriented geometric swarm formation is needed. Hence, periodically exchanging the control information between agents is essential. On the other hand, traffic data transmission, e.g., information dissemination and data collection, also follows the task-driven characteristics such as throughput requirements and delay tolerance. Therefore, compared to other communication systems, different communication demands of USCS are placed.

Different tasks require different utility functions and effectiveness evaluation models. To satisfy the requirements of tasks, USCS has to properly optimize communication resource (such as spectrum, power and relay) according to the effectiveness evaluation models. In the perspective of mathematical formulation, the essence of task-driven is the required constraints (e.g., delay, throughput and packet error rate) when optimizing the spectrum sharing.

2) GROUP COLLABORATION

In order to complete the complex tasks such as post-disaster search and terrain scanning, each agent has to work in a collaborative and cooperative way. For USCS, group collaboration has three aspects. First, collaborative flight control information needs to be frequently exchanged with high priority and robustness. Second, agents are required to disseminate and fuse the task-related traffic information collaboratively such as topographic data, which usually needs high transmission rate. Third, for communication decision making such as spectrum access, transmission power control and routing, information exchange to coordinate actions and avoid conflict are needed. The process of finding effective joint decisions of USCS can be accelerated by information exchange. However, frequent information exchange yields

high communication overhead, which may decrease the efficiency of USCS. The tradeoff between system efficiency and overhead information exchange must be considered.

3) HIGH DYNAMICS

In USS, highly mobile nodes such as drones cause that the network topology changes rapidly, leading to dynamic change of network connection relationship and interference relationship. Besides, the geographical environment decides the surrounding electromagnetic environment. Compared to the conventional low-mobility ad hoc network, the USCS needs more flexible and rapid reconfiguration capabilities.

4) ULTRA-DENSE

In some tasks the density of swarm is large, putting high requirements on the resource management. First, compared to the sparse distribution, ultra-dense agents have to exchange control information more frequently to avoid collision. Second, the ultra-dense scenario will cause the severe conflict of wireless resources and internal interference. How to allocate the limited resources to large number of agents in a local area is challenging.

5) JAMMING-RESISTANCE

If the physical strike is not considering, disabling the “neural system” is the most effective way of destroying USS. One of the most efficient and direct methods is using the physical-layer radio jamming, which can “deafen” every agent in USS. As a result, agents can not collaborate, and the swarm control will fail. More severely, the USS may be destroyed due to the physical collision. Jamming-resistant means that even under jamming attack, the basic requirements for wireless communication can be satisfied. It requires the USCS to not only coordinate the internal communication between agents, but also confront the external jamming attack.

B. REQUIREMENTS

ML techniques can be used in many fields such as pattern recognition, prediction, decision making and so on. In this paper, we mainly study the ML-empowered spectrum sharing methods in the aspect of intelligent decision making. Based on the previous five challenges, we can summarize the requirements for the communication decision making as follows.

1) SPECTRUM COLLABORATION AND CONFRONTATION

From the perspectives of spectrum, we can summarize two requirements for USCS as follows.

- *Spectrum collaboration.* Spectrum collaboration focuses on exploiting spectrum resource collaboratively in order to avoid resource conflict and waste. As described in challenges, agents of the USCS need to coordinate the usage of spectrum to realize high-efficiency communication. Although cognitive radio has promoted the development of dynamic spectrum access and improved

the utilization of spectrum, the current spectrum management of wireless ad hoc network is mainly in a preplanned stage. In order to realize the dynamic spectrum sharing and task-driven resource allocation, more flexible and adaptive spectrum collaboration is needed.

- *Spectrum confrontation.* Spectrum confrontation aims to cope with external threats while exploiting the spectrum. Due to the open nature of electromagnetic spectrum, USCS is under serious threat of adversarial jamming or unintended interference, which is one of the fatal weaknesses of USS. While agents collaborate the spectrum usage with each other, they also need to cope with the external spectrum attack and ambient RF signal. The methods of spectrum confrontation against malicious jammer and background environment should be designed.

Note that, we have proposed the concept of spectrum collaboration and spectrum confrontation in [29]. The work in [29] is from the perspective of anti-jamming communications, where a dynamic spectrum anti-jamming communication framework was proposed. In this paper, we reconsider the spectrum problems in the perspective of USCS.

2) MULTI-AGENT COLLABORATIVE DECISION MAKING

In the perspective of problem formulation, it is intractable to simultaneously optimize spectrum collaboration and confrontation. Since all the agents have individual decision-making ability, any agent’s decision (such as channel accessing, routing or offloading) will affect the spectrum environment and then impact the system performance. The goal of multi-agent decision making in the USS is to find the efficient joint actions. When an agent makes a decision, it should consider others’ decisions in a collaborative way. However, this may cause a huge combinatory action space for the ultra-dense scenario, which will dramatically increase the algorithm complexity. Therefore, suitable multi-agent collaborative decision-making algorithms are needed to realize the spectrum collaboration and confrontation in USCS.

IV. MULTI-AGENT LEARNING FRAMEWORK FOR INTELLIGENT UNMANNED SWARM COMMUNICATION SYSTEMS

In Fig. 3, a multi-agent learning framework for a USCS is presented. Basically, there are two kinds of multi-agent learning structures, i.e., mobile-computing-assisted structure and distributed structure. For the agents in a USCS, minimizing the inter-agent resource conflicts and maximizing the anti-jamming performance are two coupled problems. It is intractable to solve them simultaneously, especially when the jamming is dynamic (e.g., the dynamic power or frequency hopping). For the energy-limited and computing-power-limited mobile agents, mobile-computing-assisted multi-agent learning structure is a good candidate. However, the distributed multi-agent learning structure is needed when the mobile computing structure does not exist.

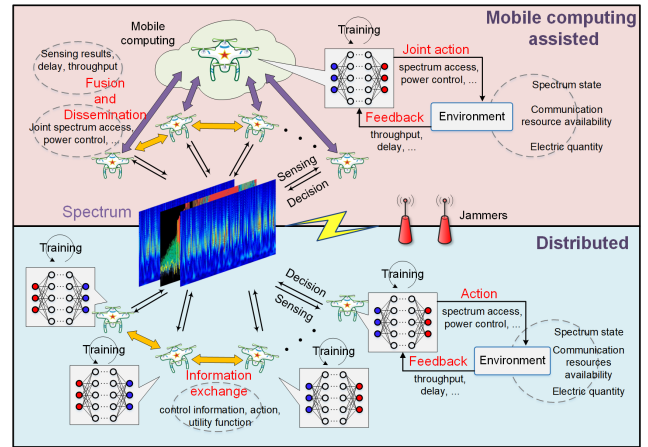


FIGURE 3. Multi-agent learning framework for the intelligent unmanned swarm communication system with mobile-computing-assisted structure and distributed structure. In the mobile-computing-assisted structure, agents are assisted by the mobile edge or mobile cloud to do the sophisticated computation and make joint actions. In the distributed structure, agents do the distributed computation and make distributed decisions to collaborate.

A. MOBILE-COMPUTING-ASSISTED MULTI-AGENT LEARNING

The mobile-computing-assisted structure is considered to be practical for a USCS. In a UAV swarm network, for example, a powerful UAV with more power supply and computation resource can act as the mobile edge server. Here, the server serves as a computing center or a central controller. The advantages are several-fold with the help of the edge server. First, the delay of data processing is decreased. Low delay is suitable for the highly dynamic scenario. Second, the distributed collaboration between agents is facilitated. Third, there can be multiple servers to manage the scarce spectrum resource in the ultra-dense USCS, which forms a hierarchical network structure. Fourth, some advanced but computation-demanding anti-jamming algorithms such as deep reinforcement learning can be used in the server to enhance the anti-jamming performance [29].

Note that, the mobile-computing-assisted structure is different from the centrally controlled structure since agents are intelligent to choose to whether be assisted by the server or make decisions autonomously. However, to take full use of the mobile-computing-assisted structure, two parts should be designed and optimized.

1) FUSION AND DISSEMINATION

The information fusion and dissemination is the primary feature of this structure. In terms of fusion, it can be separated by task aspect and algorithm aspect. Task-related data such as pictures are uploaded to the server for further processing. In the algorithm aspect, agents upload some key information such as spectrum state to feed the spectrum sharing algorithm running in the server. In terms of dissemination, the server disseminates the results of the data fusion and the outputs of multi-agent collaboration algorithm. In the task aspect, agents

can carry forward the tasks based on the fusion results. In the algorithm aspect, agents take actions according to the algorithm outputs to achieve high communication performance. In the fusion and dissemination part, mobile edge computing techniques [30] can be used to optimize the communication cost of uploading.

2) MOBILE-COMPUTING-ASSISTED MULTI-AGENT COLLABORATION

In the mobile-computing-assisted structure, the learning process of the complex multi-agent collaborative algorithms can be handed over to the edge server. As shown in the upper part of Fig. 3, the server may have a deep neural network (DNN) which is used to approximate the multi-agent collaborative decision-making function. The input of the DNN is the algorithm-related information such as spectrum state and communication resource availability. The output can be the joint action of agents. Then, this DNN is trained according to the feedback (i.e., obtained system communication performance such as throughput, delay and so on). With the help of the server, the difficulty of multi-agent collaboration is significantly decreased compared to the no-server case.

B. DISTRIBUTED MULTI-AGENT LEARNING

As shown in the lower part of Fig. 3, agents are in a distributed multi-agent learning structure, each of whom can be equipped with a DNN. In this structure, each agent senses the spectrum state, makes decisions, obtains feedback, and then trains the DNN. In order to realize the multi-agent collaboration, it is important to study and model the multi-agent decision-making relationship. Game theory is a powerful mathematical tool to study and model the interactions of a group of decision makers. It has been widely applied in wireless communications [31]–[34]. Using game theory we can analyse the impact of decision-making interactions of the agents and predict the outcome.

1) GAME MODELING

Generally, a game can be expressed as $\mathcal{G} = \{\mathcal{N}, \mathcal{A}_n, u_n\}$, where $\mathcal{N} = \{1, \dots, N\}$ is the set of participant agents, \mathcal{A}_n is the available action set of agent n (e.g., available channel, transmit power, etc.), and u_n denotes the utility function of agent n . In terms of an agent, the utility function is the evaluation of the decision. The goal of agents is to maximize u_n by adjusting their decision-making policy. The reasons to adopt game theory are twofold:

- *Distributed collaboration*: In conventional wireless networks, users are mostly self-interested and whose goal is to maximize their own utilities. However, selfish actions will result in poor performance in the USCS. To realize spectrum collaboration, the spectrum-involved actions must be restrained by some collaborative rules. For example, if the action of each agent considers not only the payoff it can get but also the punishment of the negative effects on other agents, then the collaboration

of the network can be realized spontaneously. According to this idea, the collaborative rules of the spectrum sharing games can be designed so as to reach the effective equilibrium.

- *Confrontational decision-making*: Game theory can also model the behaviors of jammer and analyse the interactions between the USCS and the jammer. By predicting the equilibrium of the confrontational game, the anti-jamming strategies are obtained.

There are many mature game models that are widely used in wireless communication problems and appropriate for the USCS.

- *Coalition formation games*: The coalition formation game is a kind of cooperative games [35], [36]. The main idea of coalition formation games is to realize the effect of “ $1 + 1 \geq 2$ ”. Generally, if the utility after forming a coalition is larger than the sum of the individual utility of each member, the coalition is formed. The coalition formation principles can be designed by considering task factors. Therefore, the coalition game is an alternative method which enables the agents to achieve task-driven collaboration. In the next section, a case study based on the coalition game is introduced.
- *Evolutionary games*: The evolutionary game is a useful method for its ability to model dynamics in wireless communication as an evolving game, which is widely used in communication resources management [37]. This game model may be helpful for the highly dynamic challenge of USCS.
- *Mean-field games*: The mean-field game is suitable for the large-scale network. In a mean-field game, the effect of other agents’ decisions on an agent is approximated by a mean effect which is assumed to be caused by a virtual agent. In this way, each agent only needs to consider a virtual agent when it is making decisions, thus significantly decreasing the complexity of multi-agent decision making. This game model is promising in the ultra-dense network [38].
- *Confrontational games*: As discussed in [29], Stackelberg games [34] and zero-sum games [39] are two mostly used game types to model the confrontational interactions between legitimate agents and malicious jammers. However, in practice, the accurate information of jammers is unavailable. Bayesian games can be used to estimate the jammers’ strategies and improve the anti-jamming performance.
- *Markov games*: Markov games are the extension of Markov decision process in a multi-agent problem. As shown in [29], confrontational games have to model jammers’ strategies, which are usually unknown to the agents in practice. By treating the jamming as the environment, Markov games enable agents to find the optimal anti-jamming policies in an unknown jamming environment. In the next section, a case study is presented.

Generally, game theory investigates the internal interactions (such as collaboration and competition) and the external confrontation. It theoretically analyses the existence and properties of stable outcomes of a game. But beyond that, algorithms in the framework of game theory are needed to guide agents how to reach the equilibrium of the game.

2) GAME-THEORETIC LEARNING

Game-theoretic learning algorithms are designed to enable the agents to autonomously find the effective solutions of the game. The game-theoretic learning model can be expressed as [40], [41]:

$$a_n(k+1) = F_n[a_n(k), a_{-n}(k), u_n(k), \dots, a_n(1), a_{-n}(1), u_n(1)] \quad (1)$$

where $a_n(k)$ and $a_{-n}(k)$ denote the action of agent n and the action profile of all the rest of agents at the k th decision-making period, utility function $u_n(k)$ is related to $a_n(k-1)$ and $a_{-n}(k-1)$, and $F_n(\cdot)$ is the policy function. As can be seen, the decision-making at the $(k+1)$ th period is adjusted according to the previous action and received reward that has happened. This kind of online learning method can overcome the disadvantages of dynamic and unknown environment. In the framework of game theory, the learning algorithms must at least guarantee the system converge to an equilibrium. More than that, there may exist multiple equilibrium solutions. It is the best for the algorithms to achieve the optimal equilibrium [40], [41].

Note that, the mentioned distributed multi-agent learning methods can also be used in the mobile-computing structure in a manner of centralized learning with decentralized execution. With the advantages of the mobile-computing structure and advanced algorithms, the well-designed multi-agent learning framework is promising to address the challenges of the USCS.

V. CASE STUDY OF MULTI-AGENT LEARNING EMPOWERED SPECTRUM SHARING

In this section, we provide case studies of multi-agent learning empowered spectrum sharing for the USCS.

A. MOBILE-COMPUTING-ASSISTED MULTI-AGENT LEARNING

For a dynamic and unknown environment, the online learning algorithms such as reinforcement learning enable USCS to adapt to the environment in an “action-feedback-adjustment” manner, as shown in Fig. 3. Specifically, the USCS first explores the unknown spectrum environment by repeating the reinforcement manner and accumulates experiences. Then the agents of USCS gradually learn the decision-making policy which can bring high rewards. However, the online learning process is time-consuming. High dynamics characteristic will cause the learning algorithm to fail to converge on the effective decision-making policy.

An “offline-learning online-planning” decision-making framework was proposed in [42], which was applied to order dispatch in a on-demand ride-hailing platform. In order to optimize the long-term accumulative rewards in stead of focusing on the immediate reward, authors proposed to first process the historical data produced by extensive offline experiments to derive the future expected value being in a particular state. Then, the sequential order dispatch decisions with high future accumulative rewards were made in real-time by taking into account both the immediate reward and the future expected value. For more details of the algorithm, the reader is referred to [42]. An obvious advantage of this decision-making framework is that a large amount of time-consuming and power-wasting operations to deal with the complex problem can be completed offline. However, the effectiveness of this framework is based on the premise that the historical data can simulate the data obtained in the actual environment.

Fortunately, particular tasks have particular requirements for USCS. For example, the deployments and formations of a UAV swarm are relatively fixed due to the requirements of tasks, i.e., topologies and interference relationships are relatively fixed. Based on this fact, we can utilize the characteristics to optimize the spectrum-involved decision making with the help of server. Specifically, for those commonly used formations and other requirements, we can solve offline the spectrum resource allocation problems using mathematical methods, such as dynamic programming, game-theoretical learning, and store the obtained schemes in the server. When executing tasks, the USCS can match the optimal resource allocation schemes based on current formation, task and location, which guarantees the effectiveness and timeliness.

B. DISTRIBUTED MULTI-AGENT LEARNING

1) TASK-DRIVEN DYNAMIC SPECTRUM ACCESS

In the task-driven USCS, control channel (CCH) and traffic channel (TCH) are critical to the quality of the task which convey control information (such as formation control commands) and task-related information (such as pictures), respectively. The transmission of control information is required to be periodical and reliable, while that of data information required to be high rate. As discussed in Section III-A, the design of CCH and TCH depends on the task. Most existing approaches only focused on either CCH or TCH, while ignoring the coupling relationship between them. Spectrum collaboration between the CCH and TCH can also improve the spectrum efficiency. Partially overlapping channels (POCs), the channels sharing within their own channel boundaries, can solve the intractable coupling problem [43]. On one hand, the interference can be reduced when two agents work on different channels. On the other hand, the control information exchange between agents can be performed using overlapping area so long as the signal-to-noise ratio (SNR) between them is above a threshold. Therefore, without channel switching, the data transmission under formation keeping can be achieved.

When considering task-driven in the aspect of swarm formation control, the control information exchange link is a two-way choice, i.e., the link can be established when agents reach a consensus. To capture the two-way choice feature, two-way consensus game is a good candidate [44]. In this game, when two agents choose each other at the same time, the communication link between them can exist. It is investigated in [45] that different formation control method corresponds to different communication topology. Hence, joint channel and link selection applying two-way consensus game is a meaningful study in the task-driven USCS.

Consider a USCS under formation keeping operating on POCs. The task time can be divided into slots. In each slot, agents conduct information exchange to keep swarm formation and then transmit traffic data. Two-way consensus game is first proposed in [44], where dynamic spectrum access in UAV communication networks when considering leader-follower formation keeping was investigated. Each agent chooses channel according to the mutual interference as well as information exchange cost. To capture the tradeoff between them, an experiment-dependent tradeoff factor β is assumed. Then, the utility function of the proposed game is defined as the combination of mutual interference in TCH and information exchange cost in CCH.

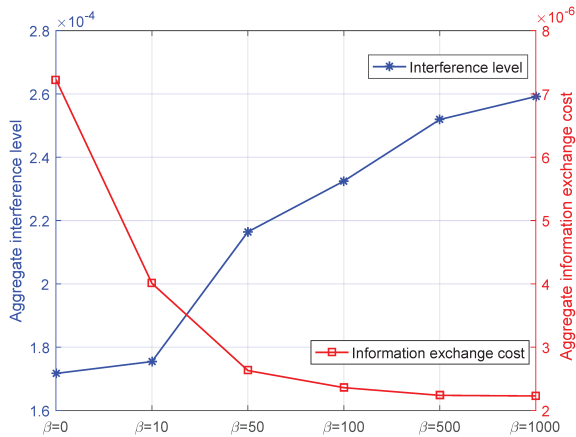


FIGURE 4. The illustration of the tradeoff between interference level and information exchange cost [44].

In [44], a two-way consensus game and a distributed learning based dynamic spectrum access algorithm are proposed. A USCS with 16 UAVs based on four finger squadron formation shape is considered. As can be seen in Fig. 4, different tradeoff factors represent the emphasis of the utility function, i.e., the importance of interference and exchange cost. Hence, it is important to choose suitable factor by practical experiments. Moreover, the performance comparison results are depicted in Fig. 5. Note that best Nash equilibrium (NE) and worst NE are obtained by applying best response algorithm, which can be viewed as the upper and lower bounds of the game. According to Fig. 5, some important results can be observed:

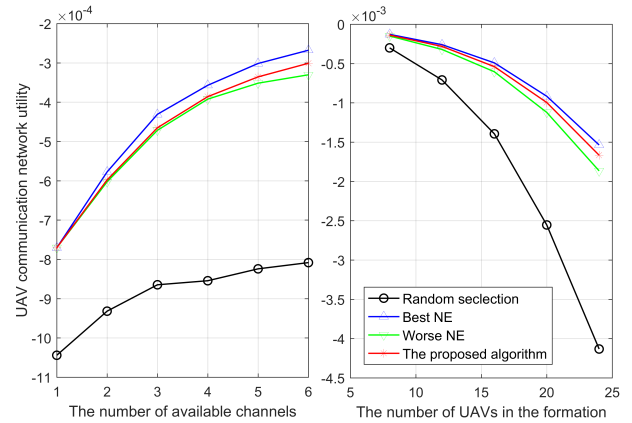


FIGURE 5. UAV communication network utility versus the number of available channels and UAVs, respectively, with different algorithms.

- As the increasing of the number of channels, the aggregate utility shows an upward trend.
- As the formation scale increases, the aggregate utility decreases. The reason is that the large number of UAVs yields not only the serious mutual interference but also the higher information exchange cost.
- The learning algorithm is close to the best NE and far better than the random solution. The effectiveness is validated.

2) TASK-DRIVEN DATA DISSEMINATION

For the sake of a task, agents in a USCS may have overlapping data requirements. Repeatedly transmitting the same data and long-distance transmission will result in high transmission overhead. To solve this problem, a resource allocation optimization method based on distributed data content in a flying ad hoc network (FANET) was proposed in [46]. The throughput maximization problem of resource allocation was constructed as a coalition formation game framework in this work.

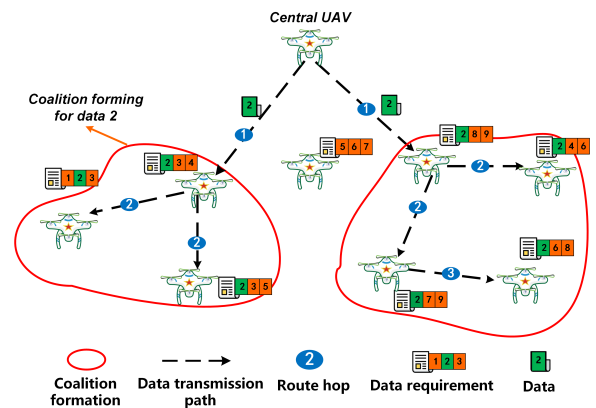


FIGURE 6. Collaborative data multi-hop transmission considering coalition formation in FANET [46].

Based on the formed coalitions, the task will be performed more collaboratively and efficiently. As shown in Fig. 6,

a scenario where a FANET executes reconnaissance and surveillance tasks has been considered in [46], where UAVs share the same spectrum resource. The central UAV disseminates the required data of other UAVs. In order to maximize the utilities, UAVs form coalitions according to their data requirements and locations to increase the throughput and decrease the overhead of transmission. The data transmission and coalition selection problem are modeled as graph game and coalition formation game respectively. Through the design of the utility function, it has been proved that both games have stable solutions. A learning algorithm was proposed to find these solutions. To show the effectiveness, representative simulation results are given in this paper. More results can be found in [46].

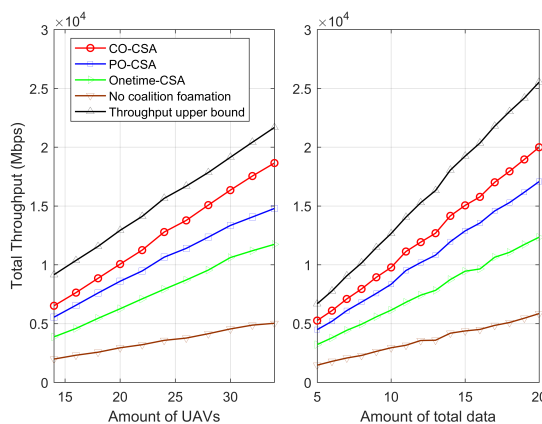


FIGURE 7. Throughput versus amount of UAVs with different algorithms.

Fig. 7 compares the throughput performance with different algorithms [46]. Two criteria of coalition formation are compared, i.e., coalition order/Pareto order based coalition selection algorithm (CO-CSA/PO-CSA). Two benchmark algorithms are compared, i.e., onetime-CSA algorithm which forms coalitions one time without considering the data content, and no coalition formation algorithm in which all data are transmitted directly by the central UAV. As the results show, the coalition game based algorithms enable the UAVs to collaboratively use the shared spectrum resource and achieve high performance.

3) JAMMING-RESISTANCE

One of the most challenging parts in USCS is that agents have to find a way to avoid collisions with others' decisions (spectrum collaboration) and simultaneously adapt to the adversarial spectrum environment (spectrum confrontation). As introduced in Section IV-B, game theory can be used to model the internal conflict relationship between agents and external adversarial relationship between agents and jammers. Based on this, a Markov game based collaborative Q-learning algorithm was proposed in [47] to solve the multi-agent anti-jamming channel access problem, then a followed work proposed by Xu *et al.* [48] applied this algorithm in a multi-UAV communication network. As illustrated

in [47] and [48], the collaborative Q-learning algorithm enables multiple agents to not only coordinate their channel access decisions, but also avoid the dynamic jamming attacks (swept jamming).

However, to apply to the USCS, several issues need to be solved. First, this algorithm is only suitable for small-scale networks. In order to optimize the global objective (e.g., network communication throughput), every agent takes the joint action of other agents into consideration, which results in the problem that the computation complexity and storage space (the dimension of Q-table) increase exponentially with the number of agents, i.e., multi-agent combinatorial explosion. What's more, to achieve the collaborative effect, the algorithm requires every agent to inform others of the algorithmic information (Q-table). Due to the combinatorial explosion, the communication cost is significant and impractical [29]. Second, this algorithm only considers single-link anti-jamming communication. However, multi-hop communication exists in the USCS to realize long-distance transmission. Multi-hop anti-jamming mechanism is needed.

VI. FUTURE RESEARCH DIRECTIONS AND CONCLUSIONS

A. FUTURE RESEARCH DIRECTIONS

1) LIGHTWEIGHT AND CUSTOMIZED ML ALGORITHMS

Many ML algorithms such as deep learning require powerful computation capability and time-consuming training. An "offline-training-online-using" manner is appropriate for USCS. However, due to the high dynamics, agents have to online adapt to the real spectrum environment. Hence, lightweight and customized ML-enabled spectrum sharing algorithms which also enable the agents to online learn the spectrum environment are needed.

2) FAST AND ROBUST SPECTRUM DATA PROCESSING

The cognition of spectrum state is vital for the intelligent spectrum sharing algorithms to obtain the effective policies. However, due to the limited processing power of UAVs and characteristics such as high dynamics, ultra-dense and jamming attack, conventional spectrum sensing and spectrum data processing methods may not satisfy the USCS's requirements for the delay and accuracy.

3) TASK-DRIVEN INFORMATION TRANSMISSION

In Fig. 8, we categorize the information transmitted in USCS into three kinds: network-related, traffic-related and algorithmic. Network-related information is mainly responsible for maintaining the network of USCS, such as formation control information and routing information. Traffic-related information represents the traffic data such as videos and pictures. Algorithmic information is used to guarantee the performance of the spectrum sharing algorithms.

In general, the network-related and traffic-related information transmission are necessary for a USCS. This process adds constraint conditions to the spectrum sharing. But in turn,

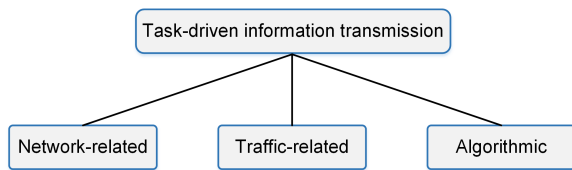


FIGURE 8. Categorization of information transmission in an intelligent unmanned swarm communication system.

these algorithms can optimize the network-related and traffic-related information transmission. In the multi-agent learning framework, algorithmic information exchange can accelerate discovery of successful collaborative spectrum sharing policies.

However, it is noteworthy that optimizing the task-driven information transmission is very challenging. First, these three kinds of information transmission are interrelated. How to allocate limited communication resource needs intensive study. Second, excessive information exchange will significantly increase the communication overhead, which may instead decrease the system performance. A tradeoff needs in-depth investigation.

4) TASK-DRIVEN SPECTRUM SHARING STRUCTURE

In conventional ad hoc networks, the concept of cluster is widely used for the reason that it can facilitate the management of communication resources and improve network efficiency. Most of the clustering-based optimization methods are based on the principle of relative physical locations of communication nodes. However, task-driven factors must be considered in USCS problems if the cluster-like structure is used. As the case study shows, the coalition form game can realize the task-driven spectrum collaboration which take both the physical location and task-related factors into account. The coalition-based spectrum sharing structure in USCS is a promising research direction.

B. CONCLUSIONS

The spectrum sharing problems for USCS were investigated in this paper. From the perspective of spectrum collaboration and spectrum confrontation, five challenges that USCS must overcome were summarized. In the perspective of ML, two requirements for the communication decision making of the agents were proposed. Then, a multi-agent learning framework has been proposed. Based on it, we have introduced four case studies. Finally, future research directions were discussed.

REFERENCES

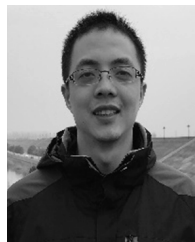
- [1] S. A. R. Naqvi, S. A. Hassan, H. Pervaiz, and Q. Ni, "Drone-aided communication as a key enabler for 5G and resilient public safety networks," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 36–42, Jan. 2018.
- [2] Y. Zeng, J. Lyu, and R. Zhang, "Cellular-connected UAV: Potential, challenges, and promising technologies," *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 120–127, Feb. 2019.
- [3] I. Guvenc, F. Koohifar, S. Singh, M. L. Sichitiu, and D. Matolak, "Detection, tracking, and interdiction for amateur drones," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 75–81, Apr. 2018.

- [4] H. Kim, L. Mokdad, and J. Ben-Othman, "Designing UAV surveillance frameworks for smart city and extensive ocean with differential perspectives," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 98–104, Apr. 2018.
- [5] J. Gu, T. Su, Q. Wang, X. Du, and M. Guizani, "Multiple moving targets surveillance based on a cooperative network for multi-UAV," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 82–89, Apr. 2018.
- [6] J. Wang, C. Jiang, Z. Han, Y. Ren, R. G. Maunder, and L. Hanzo, "Taking drones to the next level: Cooperative distributed unmanned-aerial-vehicular networks for small and mini drones," *IEEE Veh. Technol. Mag.*, vol. 12, no. 3, pp. 73–82, Sep. 2017.
- [7] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1617–1655, 3rd Quart., 2016.
- [8] N. Lu, N. Cheng, N. Zhang, X. Shen, and J. W. Mark, "Connected vehicles: Solutions and challenges," *IEEE Internet Things J.*, vol. 1, no. 4, pp. 289–299, Aug. 2014.
- [9] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.
- [10] M. J. Marcus, "Spectrum policy challenges of UAV/drones [spectrum policy and regulatory issues]," *IEEE Wireless Commun.*, vol. 21, no. 5, pp. 8–9, Oct. 2014.
- [11] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, May 2016.
- [12] H. Wang, J. Wang, G. Ding, J. Chen, Y. Li, and Z. Han, "Spectrum sharing planning for full-duplex UAV relaying systems with underlaid D2D communications," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1986–1999, Sep. 2018.
- [13] L. Wang, H. Yang, J. Long, K. Wu, and J. Chen, "Enabling ultra-dense UAV-aided network with overlapped spectrum sharing: Potential and approaches," *IEEE Netw.*, vol. 32, no. 5, pp. 85–91, Sep. 2018.
- [14] Z. Hu, Z. Zheng, L. Song, T. Wang, and X. Li, "UAV offloading: Spectrum trading contract design for UAV-assisted cellular networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 9, pp. 6093–6107, Sep. 2018.
- [15] J. Chen, Y. Xu, Q. Wu, Y. Zhang, X. Chen, and N. Qi, "Interference-aware online distributed channel selection for multicenter FANET: A potential game approach," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 3792–3804, Apr. 2019.
- [16] K. Yao, J. Wang, Y. Xu, Y. Xu, Y. Yang, Y. Zhang, H. Jiang, and J. Yao, "Self-organizing slot access for neighboring cooperation in UAV swarms," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2800–2812, Apr. 2020.
- [17] D. Liu, Y. Xu, J. Wang, Y. Xu, A. Anpalagan, Q. Wu, H. Wang, and L. Shen, "Self-organizing relay selection in UAV communication networks: A matching game perspective," *IEEE Wireless Commun.*, vol. 26, no. 6, pp. 102–110, Dec. 2019.
- [18] D. Liu, J. Wang, K. Xu, Y. Xu, Y. Yang, Y. Xu, Q. Wu, and A. Anpalagan, "Task-driven relay assignment in distributed UAV communication networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 11003–11017, Nov. 2019.
- [19] J. Peng, Z. Zhang, Q. Wu, and B. Zhang, "Anti-jamming communications in UAV swarms: A reinforcement learning approach," *IEEE Access*, vol. 7, pp. 180532–180543, 2019.
- [20] L. Gupta, R. Jain, and G. Vaszkun, "Survey of important issues in UAV communication networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1123–1152, 2nd Quart., 2016.
- [21] H. Wang, H. Zhao, J. Zhang, D. Ma, J. Li, and J. Wei, "Survey on unmanned aerial vehicle networks: A cyber physical system perspective," *IEEE Commun. Surveys Tuts.*, early access, Dec. 25, 2019, doi: 10.1109/COMST.2019.2962207.
- [22] Z. Feng, L. Ji, Q. Zhang, and W. Li, "Spectrum management for mmWave enabled UAV swarm networks: Challenges and opportunities," *IEEE Commun. Mag.*, vol. 57, no. 1, pp. 146–153, Jan. 2019.
- [23] J. Wang, C. Jiang, H. Zhang, Y. Ren, K.-C. Chen, and L. Hanzo, "Thirty years of machine learning: The road to Pareto-optimal wireless networks," *IEEE Commun. Surveys Tuts.*, early access, Jan. 13, 2020, doi: 10.1109/COMST.2020.2965856.
- [24] D. Wang, B. Song, D. Chen, and X. Du, "Intelligent cognitive radio in 5G: AI-based hierarchical cognitive cellular networks," *IEEE Wireless Commun.*, vol. 26, no. 3, pp. 54–61, Jun. 2019.

- [25] N. Javaid, A. Sher, H. Nasir, and N. Guizani, "Intelligence in IoT-based 5G networks: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 94–100, Oct. 2018.
- [26] X. Liao, J. Shi, Z. Li, L. Zhang, and B. Xia, "A model-driven deep reinforcement learning heuristic algorithm for resource allocation in ultra-dense cellular networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 983–997, Jan. 2020.
- [27] X. Chen, Z. Han, H. Zhang, G. Xue, Y. Xiao, and M. Bennis, "Wireless resource scheduling in virtualized radio access networks using stochastic learning," *IEEE Trans. Mobile Comput.*, vol. 17, no. 4, pp. 961–974, Apr. 2018.
- [28] R. Li, Z. Zhao, X. Zhou, G. Ding, Y. Chen, Z. Wang, and H. Zhang, "Intelligent 5G: When cellular networks meet artificial intelligence," *IEEE Wireless Commun.*, vol. 24, no. 5, pp. 175–183, Oct. 2017.
- [29] X. Wang, J. Wang, Y. Xu, J. Chen, L. Jia, X. Liu, and Y. Yang, "Dynamic spectrum anti-jamming communications: Challenges and opportunities," *IEEE Commun. Mag.*, vol. 58, no. 2, pp. 79–85, Feb. 2020.
- [30] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2322–2358, 4th Quart., 2017.
- [31] Z. Han, D. Niyato, W. Saad, T. Başar, and A. Hjørungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
- [32] Y. Xu, A. Anpalagan, Q. Wu, L. Shen, Z. Gao, and J. Wang, "Decision-theoretic distributed channel selection for opportunistic spectrum access: Strategies, challenges and solutions," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1689–1713, 4th Quart., 2013.
- [33] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y.-D. Yao, "Opportunistic spectrum access in cognitive radio networks: Global optimization using local interaction games," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 2, pp. 180–194, Apr. 2012.
- [34] L. Jia, Y. Xu, Y. Sun, S. Feng, and A. Anpalagan, "Stackelberg game approaches for anti-jamming defence in wireless networks," *IEEE Wireless Commun.*, vol. 25, no. 6, pp. 120–128, Dec. 2018.
- [35] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Basar, "Coalitional game theory for communication networks," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 77–97, Sep. 2009.
- [36] Y. Zhang, Y. Xu, A. Anpalagan, Q. Wu, Y. Xu, Y. Sun, S. Feng, and Y. Luo, "Context-aware group buying in ultra-dense small cell networks: Unity is strength," *IEEE Wireless Commun.*, vol. 26, no. 6, pp. 118–125, Dec. 2019.
- [37] K. Zhu, D. Niyato, P. Wang, and Z. Han, "Dynamic spectrum leasing and service selection in spectrum secondary market of cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1136–1145, Mar. 2012.
- [38] C. Yang, J. Li, P. Semasinghe, E. Hossain, S. M. Perlaza, and Z. Han, "Distributed interference and energy-aware power control for ultra-dense D2D networks: A mean field game," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1205–1217, Feb. 2017.
- [39] C. Chen, M. Song, C. Xin, and J. Backens, "A game-theoretical anti-jamming scheme for cognitive radio networks," *IEEE Netw.*, vol. 27, no. 3, pp. 22–27, May 2013.
- [40] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y.-D. Yao, "Opportunistic spectrum access in unknown dynamic environment: A game-theoretic stochastic learning solution," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1380–1391, Apr. 2012.
- [41] Y. Xu, J. Wang, Q. Wu, J. Zheng, L. Shen, and A. Anpalagan, "Dynamic spectrum access in time-varying environment: Distributed learning beyond expectation optimization," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5305–5318, Dec. 2017.
- [42] Z. Xu, Z. Li, Q. Guan, D. Zhang, Q. Li, J. Nan, C. Liu, W. Bian, and J. Ye, "Large-scale order dispatch in on-demand ride-hailing platforms: A learning and planning approach," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2018, pp. 905–913.
- [43] J.-H. Lim, K. Naito, J.-H. Yun, and M. Gerla, "Revisiting overlapped channels: Efficient broadcast in multi-channel wireless networks," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2015, pp. 1984–1992.
- [44] J. Chen, P. Chen, Y. Xu, N. Qi, C. Dong, and Q. Wu, "Dynamic spectrum access in UAV communication networks: A two-way consensus game," *IEEE Trans. Mobile Comput.*, submitted for publication.
- [45] W. Ren, "Consensus strategies for cooperative control of vehicle formations," *IET Control Theory Appl.*, vol. 1, no. 2, pp. 505–512, Mar. 2007.
- [46] L. Ruan, J. Chen, Q. Guo, X. Zhang, Y. Zhang, and D. Liu, "Group buying-based data transmission in flying ad-hoc networks: A coalition game approach," *Information*, vol. 9, no. 10, p. 253, 2018.
- [47] F. Yao and L. Jia, "A collaborative multi-agent reinforcement learning anti-jamming algorithm in wireless networks," *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1024–1027, Aug. 2019.
- [48] Y. Xu, G. Ren, J. Chen, X. Zhang, L. Jia, and L. Kong, "Interference-aware cooperative anti-jamming distributed channel selection in UAV communication networks," *Appl. Sci.*, vol. 8, no. 10, p. 1911, 2018.



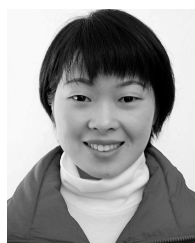
XIMING WANG received the B.E. degree in electronic and information engineering from the Wuhan University of Technology, China, in 2015, and the M.S. degree in communications and information systems from the College of Communication Engineering, Army Engineering University, China, in 2017. He is currently pursuing the Ph.D. degree with Army Engineering University. His current research interests include anti-jamming communications, cognitive radio, and machine learning.



YUHUA XU (Member, IEEE) received the B.S. degree in communications engineering and the Ph.D. degree in communications and information systems from the College of Communications Engineering, PLA University of Science and Technology, in 2006 and 2014, respectively. He is currently an Associate Professor with the College of Communications Engineering, Army Engineering University of PLA. He has published several articles in international conferences and reputed journals in his research area. His research interests include unmanned aerial vehicle communication networks, opportunistic spectrum access, learning theory, and distributed optimization techniques for wireless communications. He received the Certificate of Appreciation as an Exemplary Reviewer for the IEEE COMMUNICATIONS LETTERS, in 2011 and 2012, respectively. He was selected to receive the IEEE Signal Processing Society 2015 Young Author Best Paper Award and the Funds for Distinguished Young Scholars of Jiangsu Province, in 2016.



CHAOHUI CHEN received the B.S. degree from the School of Telecommunications Engineering, Xidian University, in 1989, and the M.S. degree from the School of Electronic Information and Communication, Huazhong University of Science and Technology, in 2005. He is currently the Deputy Chief Engineer with Guangzhou Haige Communications Group Incorporated Company, Guangzhou. His research interests include high-frequency communication, radio detection, and cloud computing.



XIAOQIN YANG received the M.S. degree from the School of Computer Science and Communication Engineering, Jiangsu University, China, in 2010. She is currently a Lecturer with the College of Communication Engineering, Army Engineering University. Her current research interests include anti-jamming communication and cognitive radio.



JIAXIN CHEN received the B.S. degree in communications engineering from the College of Electronic and Information Engineering, Civil Aviation University of China, in 2016. She is currently pursuing the Ph.D. degree in communications and information systems with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics. Her research interests include game theory and resource allocation in unmanned aerial vehicle communication networks.



YIFAN XU received the B.E. degree in electronic and information engineering from the Wuhan University of Technology, China, in 2015, and the M.S. degree in communications and information systems from the College of Communication Engineering, Army Engineering University, China, in 2017. He is currently pursuing the Ph.D. degree with Army Engineering University. His current research interests include anti-jamming communications, cognitive radio, and machine learning.



LANG RUAN received the B.S. degree in information and computing science from Zhejiang University, Hangzhou, China, in 2016, and the M.S. degree from the College of Communications Engineering, Army Engineering University, Nanjing, China, in 2019, where he is currently pursuing the Ph.D. degree. His current research interests include game theory, unmanned aerial vehicle communications networks, and cooperative localization.



RUNFENG CHEN received the B.E. degree from the College of Communication Engineering, Army Engineering University, China, in 2018. He is currently pursuing the Ph.D. degree with Army Engineering University. His current research interests include unmanned aerial vehicle networking, opportunity spectrum access, and collective intelligence.

...