ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/335895555

Task-Driven Relay Assignment in Distributed UAV Communication Networks

Article *in* IEEE Transactions on Vehicular Technology · September 2019

DOI: 10.1109/TVT.2019.2942095

CITATIONS
51

READS
394

8 authors, including:

Dianxiong Liu
PLA University of Science and Technology
47 PUBLICATIONS **468** CITATIONS

SEE PROFILE

Kun Xu
84 PUBLICATIONS **377** CITATIONS

SEE PROFILE

Yuhua Xu
PLA University of Science and Technology
199 PUBLICATIONS **4,984** CITATIONS

SEE PROFILE

Qihui Wu
PLA University of Science and Technology
276 PUBLICATIONS **6,911** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Lyric: Local recall-based Dynamic Double Spectrums Auction Mechanism with Heterogeneous-demand Secondary Users View project

Project Hierarchical Game Optimization for Resource Allocation in Cognitive Smallcell Network View project

All content following this page was uploaded by Dianxiong Liu on 22 September 2019.

The user has requested enhancement of the downloaded file.

Task-Driven Relay Assignment in Distributed UAV Communication Networks

Dianxiong Liu, *Student Member, IEEE*, Jinlong Wang, *Senior Member, IEEE*, Kun Xu, Yuhua Xu, *Member, IEEE*, Yang Yang, Yitao Xu, Qihui Wu, *Senior Member, IEEE*, and Alagan Anpalagan, *Senior Member, IEEE*

***Abstract*—In this paper, we study the distributed relay assignment problem in multi-channel multi-radio unmanned aerial vehicle (UAV) communication networks. Multi-UAVs are driven by the overall task and fly in certain formation, where UAVs with different tasks have various transmission requirements. Source UAVs equipped with multi-radio can select more than one relay radios to achieve high data rate, and each relay radio can be shared by multiple source UAVs. We construct distributed game models to promote the global transmission performance by self-organizing coordination among UAVs. Specifically, the channel competition relationship between relay UAVs is modeled as a congestion game model, while the task-driven relay selection among UAVs is modeled as a many-to-many matching market without substitutability. With the proposed game models, the optimizing of local optimized process will lead to the improvement of global transmission results. After that, we design algorithms for the stable and changeable topology structures, respectively. Based on the given formation shape of UAVs, a learning matching algorithm is proposed to reach the optimum result with a large probability. A fast potential matching algorithm is propose to deal with the topological change of UAV networks. We prove that two proposed algorithms can both achieve the stable matching results. Simulation results show that the proposed relay assignment approaches yield good performances in terms of the global transmission satisfaction and fairness. Particularly, the result of the learning algorithm is close to the global optimum and the fast potential matching approach is robust to the perturbation of UAV networks.**

***Index Terms*—UAV communication networks, distributed multi-channel multi-radio, transmission requirement, game models, matching without substitutability.**

I. Introduction

Unmanned aerial vehicle (UAV) communication has become a hot area of research in wireless networks [1]–[5]. In addition to being used in urban systems such as cellular-connected UAVs and UAV-assisted terrestrial communications [1], [2], drones are also expected to perform many tasks in the wild or dangerous places [3]–[5]. To play broader applications such as a wider range of search, combat and patrol tasks [5]–[7], multiple UAVs are expected to be deployed. Cooperation of UAVs is considered as an important research branch, where the cluster based communication structure is regarded as a useful model for the task driven UAVs [3]–[5].

The efficient information interaction in the UAV cluster is important to maintain the flight and mission implementation. Each UAV cluster head is responsible to collect and deliver the information from cluster members to the ground controller [3]–[5]. However, it is difficult to achieve entire network coverage by one-hop transmission due to the limited transmission power of UAVs. To tackle this problem, one solution is to perform relay transmission, while the other is to change the position and optimize UAV trajectories. Carrying distant mission, UAVs are usually in formation flight to ensure the coverage of the ground and the stability of the swarm [8]. The relative movement of drones is likely to cause a collision risk. Frequent trajectory changes also consume much propulsion energy, which is not conducive to the endurance of swarms. By contrast, the relay communication can improve the transmission rate and coverage of networks without destroying the structure of the formation, which is promising in UAV communication networks [4].

For the UAV swarms with multiple source and relay UAVs, only efficient relay assignment can lead to improved performances, while inappropriate methods may cause negatively influence [9]. Therefore, the assignment problem of relay resource is considered to promote the transmission performance of task-driven UAVs. The research on the relay assignment in UAV communication networks is still in the infancy [10]–[13]. In a few articles, Ho et al. [10] studied the single path selection problem with the goal of minimizing the time and energy consumption. Similarly, Zhang et al. [11] learned about the problem of route selection in a three-dimensional UAV swarm. Mukherjee et al. [12] tackled the problem of data resource scheduling through path selection. Sbeiti et al. [13] proposed a secure routing protocol to face attacks in airborne mesh networks. The existing work confirmed the importance of relay communication within UAVs. However, these articles mainly studied the single-link path selection problem. The limitations of channel resources and the task-driven features of UAV networks were neglected.

Compared to UAV swarms, the relay selection problem

Part of this work has been presented in IEEE/CIC International Conference on Communications in China (ICCC'2017). This work was supported by the National Science Foundation of China under Grant No. 61771488, No. 61901520, No. 61801492, No. 61931011, No. 61971439 and No. 61701533, in part by the Natural Science Foundation of Jiangsu Province of China (No. SBK2019020930), Postdoctoral Science Fund of China (No. 2018M633684) and by research fund of National University of Defense Technology with contract No. ZK18-03-20.

Dianxiong Liu, Jinlong Wang, Yuhua Xu, Yang Yang and Yitao Xu are at the College of Communications Engineering, Army Engineering University, Nanjing, China. (e-mails: dianxiongliu@163.com, wjl543@sina.com, yuhuaenator@gmail.com, sheep_1009@163.com, yitaoxu@126.com). Yuhua Xu is the corresponding author.

Kun Xu is with the College of Information and Communication, National University of Defense Technology, Wuhan, China (e-mail: xukunown@163.com).

Qihui Wu is with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China (e-mail: wuqihui2014@sina.com).

Alagan Anpalagan is with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, Canada (e-mail: alagan@ee.ryerson.ca).

has been well studied in terrestrial communications. In [14], [15], the authors optimized the single-link relay selection with goals of maximizing the global transmission rate and the minimum transmission rate, respectively. After that, many-to-one relay models were studied, where one relay node can assist multiple source nodes. In [16], Xu et al. studied the max-min and max total throughput problem for many-to-one relay models by centralized algorithms. Chen et al. [17] proposed a decentralized learning to optimize the fairness performance of spectrum sharing networks. Moreover, in buffer-aid relay model [18], [19] or downlink transmission of non-orthogonal multiple-access systems [20]–[22], one source node can select multiple relay nodes (one-to-many) to improve the transmission performance. The relay selection problem also can be considered joint the other optimizing objectives such as channel assignment [23], security [24] and so on. However, some differences and challenges should be concerned when extending relay technologies from terrestrial networks to UAV networks.

Firstly, with the help of the centralized controller, the centralized approach [25], [26] can be used to optimize the resource allocation in terrestrial networks. However, there may not be a powerful central controller flying in the air [5], especially in distant missions. The optimization problems for UAV communications are suitable to be solved in distributed manners [4].

Secondly, in terrestrial networks, the data transmission is mainly generated by users selfishly in order to meet its own needs. In contrast, the UAV swarm is mainly driven by overall tasks, and the data transmission is mainly to ensure the task satisfaction performance of global networks. With limited information exchange, UAVs will easily interfere the performance of other UAVs when maximizing its own transmission.

Thirdly, terrestrial relay networks are mainly interested in one-to-one [14], [15] or many-to-one models [16]–[22]. In distributed UAV networks, channel resource will be divided and utilized in a shared manner. The single-link mode may be difficult to meet the needs of UAVs when the transmission services of UAVs are heterogeneous. With the many-to-many relay model, drones can configure the number of radios and channel selection strategies flexibly [27]. However, the many-to-many transmission will also lead to more competition among UAVs. Without suitable scheme of interference avoidance and resource allocation, the performance of the whole network is more difficult to guarantee.

Therefore, from the perspective of global transmission tasks, we study the self-organizing resource assignment problem in multi-channel multi-radio UAV relay communications. A distributed game model is constructed. Due to the sharing characteristics of channel resources, the channel access problem of UAVs is modeled as a congestion game model with player-specific payoff functions [28], while the selection relationship between UAVs is modeled as a many-to-many matching market [29]. We take the satisfaction experience of global task completion as the performance metric. The contributions of this paper are as follows,

- We study the multiple access relay selection system with limited channel resource, where the transmission demands among UAVs are considered. To the best of our knowledge, it is the first work to optimize the relay assignment in many-to-many UAV networks.
- The features of resource sharing and peer effects [30] among UAVs make the classic matching model no longer applicable. We model the relay selection model as a many-to-many matching market without substitutability [4], which is rarely studied in wireless networks.
- Combining with the potential game [31], we design a potential matching framework to achieve the stable matching. In this framework, the global result can be improved by optimizing the local matching performance and thus the equilibrium of the non-substitution model can be achieved.
- Based on the given shape of UAV formation, we propose a many-to-many learning matching approach (MLMA) to achieve the optimal performance with an arbitrarily high probability. Aimed at the changeable topology, a fast potential matching approach (FPMA) is proposed, which spends a short time to converge and is robust to the dynamic network. Both two algorithms are proved to achieve stable matching results.

This paper is an extension of the paper [32]. Compared with the preliminary work, the multi-channel multi-radio competition problem is further modeled, and two algorithms are proposed to achieve different objectives of optimizing requirements in UAV networks. The discussion of the matching without substitutability and the theory of the potential-matching framework are systematically given.

The rest of this paper is organized as follows. In Section II, the motivation of developing the potential matching model is given. In Section III, the system model is provided and the problem is formulated. In Section IV, the congestion model is constructed for the channel access among relay UAVs. In Section V, the many-to-many matching game for the task-driven relay assignment is analyzed, and the matching approaches without substitutability are studied. In Section VI, a joint algorithm of channel access and relay selection is proposed. In Section VII, simulation results are shown. The conclusion is drawn in Section VIII.

II. The Motivation of The Matching Game Model

Matching markets [29] are powerful distributed models to address assignment issues in wireless networks. Compared with other game models and other useful approaches for the distributed relay selection optimization of UAV communication networks, the advantages of matching game models can be summarized as follows, i) Players in classical game-theoretical algorithms require the knowledge of other players' actions [31]. However, it is difficult for source UAVs (SVs) to obtain all strategy information of other drones. Matching game with independent decision-making will enhance the distributed implementation; ii) Driven by diverse communication tasks, SVs and relay UAVs (RVs) have different communication demands. Diversified transmission needs make the optimization method more flexible; iii) Matching game can define individual utilities for SVs and RVs. The available algorithmic implementations can solve the resource allocation problem without obtaining global information. Therefore, the UAV relay selection system is modeled as a matching market, in which SVs and RVs have decision-making capabilities to achieve their own transmission needs.

The existing work of matching game mainly used one-to-one or many-to-one markets [33]–[35] to solve the resource assignment problem of wireless networks. With the development of multiple access techniques, researchers began to focus on the many-to-many matching market [36]–[38]. For example, Di et al. [36] studied codebook-based resource allocation problem by the matching game for an uplink sparse code multiple access network. Similarly, Zhang et. al [37] studied relay networks with non-orthogonal multiple access. It is noted that all of existing work above referred to the matching market with substitutability [35]–[38], where the relationship between players is solely competitive. Players replace other competitors when they are successfully accepted by players in the opposing set. Such work does not mention the limited resource sharing. In distributed UAV communication models, multiple UAVs have the same opportunity to share resources equally, and they aim at finishing the cooperative work. While improving own transmission rate, UAVs should also consider the performance of other drones. The model of resource assignment is much different from the matching with substitutability.

Matching market without substitutability [4], [39] can be used to model distributed networks with resource sharing. In this matching game, the relationship between applicants is neither bad or good, and the priority of players is not the criterion for filtering. After received matching requests from players in the opposite side, players do not accept or reject a matching connection according to the unilateral performance of applicants. The performance of the whole network and the applicants of other connections will affect the matching process. This class is suitable for the fully sharing networks, such as spectrum sharing equally with unfixed quotas [4]. To solve the selection problem in this kind of UAV networks, we combine the matching market with the potential game according to the optimization objective and the characteristic of matching game without substitutability.

III. System Model and Problem Formulation

A. The Relay Model of UAV Communication Networks

Consider a UAV swarm maintaining a certain formation shape to execute determined tasks. In the cluster-based system shown as Fig. 1, cluster members perform various tasks such as surveillance monitoring and patrol tracking, and then summarize the collected data to the cluster head [4]. SVs at cluster edge with limited transmission power have opportunities to be assisted by RVs near the cluster head. The set of SVs, RVs and UAV destinations are denoted by $\mathcal{S} = \{1, ..., s, ..., S\}$, $\mathcal{R} = \{1, ..., r, ..., R\}$ and d, respectively. Each UAV is equipped with one or more radios so as to work on multiple channels denoted by $\mathcal{L} = \{1, ..., l, ..., L\}$ at the same time. Without loss of generality, we assume that $L \le R \le S$. Each radio of UAVs can work on one channel at most, so each SV s can connect with the number of channels

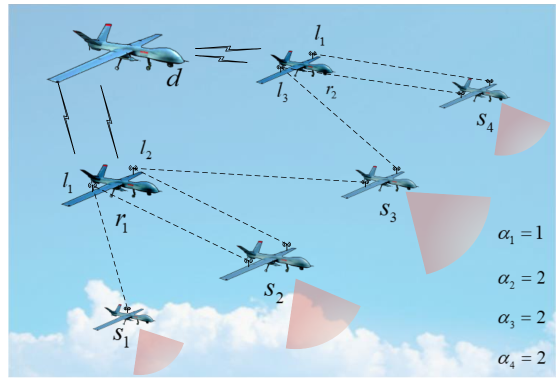


Fig. 1. A multi-channel multi-radio UAV relay selection model.

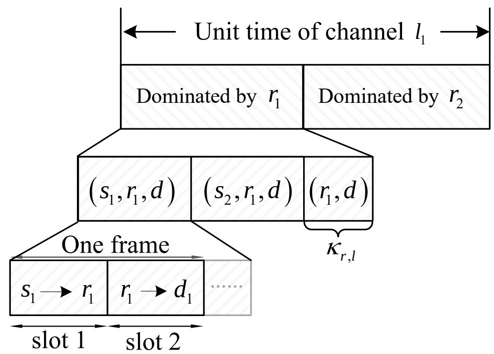


Fig. 2. The transmission process of the UAV relay model.

that do not exceed the number of radios it equipped with. That is,

$$\sum\nolimits_{l \in \mathcal{L}} \delta_{sl} \le \alpha_s, \tag{1}$$

where the number of radios of s is denoted by α_s. δ_{sl} represents the number of radios that SV s works on channel l, and $0 \le \delta_{sl} \le 1$ because there is no capacity benefit if two different radios of one drone work in the same channel. Similarly, there are $\sum\nolimits_{l \in \mathcal{L}} \delta_{rl} \le \alpha_r$, $\delta_{rl} \le 1$, $\sum\nolimits_{r \in \mathcal{R}} \delta_{sr} \le \alpha_s$ and $\delta_{sr} \le \min\{\alpha_s, \alpha_r\}$, where α_r represents the number of radios of r. δ_{sr} is the number of links that RV r assists s. A RV can provide multi-channel services for one SV by multiple radios.

The transmission model of UAV communication networks is shown in Fig. 2. To report own service data, RVs access and share channel resources in competitive mechanisms [40]. The transmission rate that r can achieve on l is:

$$u_r(l, d) = \frac{W}{A_{rl}} \cdot \log_2(1 + \gamma_{rd}). \tag{2}$$

where the channel bandwidth is denoted by W. For ease of expression, the signal-to-noise-ratio (SNR) at the d coming from RV r is denoted by γ_{rd} and $\gamma_{rd} = P_r |h_{r,d}|^2 / \sigma^2$, where P_r is the transmission power of the RV and the parameter σ^2 is the variance of additive white gaussian noise. $h_{r,d}$ represents the path gain between r and d, where Doppler effects caused by the mobility of UAVs are assumed to be compensated perfectly [26]. RVs share channel resources equally, and the number of RVs choosing the channel l is denoted by A_{rl}.

RVs will occupy a part of channel resource to transmit their own data, which is denoted by $\kappa(r,l)$, $0 \le \kappa(r,l) \le 1$. After that, RVs sense and retransmit the data from SVs. For the relay model, each radio works in a half-duplex mode and the communication pattern is frame-by-frame, where each frame is divided into two time slots. The first time slot is used for the link of the SV to the RV, while the second time slot is used for the link of the RV to the destination. In this paper, the amplify-and-forward (AF) relay mode is applied, and the proposed scheme can be extended to other modes such as decode-and-forward (DF) or hybrid modes [9]. Under the AF mode, RVs receive signals from SVs and then amplify and transmit them to the destination.

In order to represent the performance of relay assignment among SVs, we assume that the direct transmission diversity between the SVs and destination would not exist. Therefore, if one of radio of s is assisted by r, the ideal capacity can be expressed as [9],

$$C_{AF}(s,r,d) = \frac{W}{2} \cdot \log_2\left(1 + \frac{\gamma_{sr}\gamma_{rd}}{1+\gamma_{sr}+\gamma_{rd}}\right), \tag{3}$$

where the SNR from s to r and s to d are denoted by γ_{sr} and γ_{sd}, respectively. Due to the limited number of channels and RVs, SVs also need to compete for resources with other UAVs. Considering factors above, the available data rate that s assisted by r on channel l can be given,

$$u_s(r,l,d) = \frac{C_{AF}(s,r,d)(1-\kappa_{r,l})}{A_{rl}A_{srl}}. \tag{4}$$

The data rate of one pair is related to the number of transmission pairs choosing the same radio of r [17], which is denoted by A_{srl}. With the help of multiple relay radios, the data rate of communication pair (s,d) denoted by $u_{s,d}$ can be given,

$$u_{s,d} = \sum_{r\in\mathcal{R}}\sum_{l\in\mathcal{L}} u_s(r,l,d)\,\delta_{srl}, \tag{5}$$

$$\text{s.t.} \quad \sum\nolimits_{l\in\mathcal{L}} \delta_{sl} \le \alpha_s, \quad \forall\delta_{sl}\in\{0,1\}, \tag{6}$$

$$\sum\nolimits_{l\in\mathcal{L}} \delta_{rl} \le \alpha_r, \quad \forall\delta_{rl}\in\{0,1\}, \tag{7}$$

$$\sum\nolimits_{r\in\mathcal{R}} \delta_{sr} \le \alpha_s, \quad \delta_{sr} \le \min\{\alpha_s,\alpha_r\}, \tag{8}$$

where δ_{srl} is an indicator function, representing that s is assisted by r on the channel l if $\delta_{srl}=1$, otherwise $\delta_{srl}=0$.

Driven by various tasks of data transmission, pairs have different type of services, data rate requirements and tolerance levels of service quality. SVs aim at finding the set of relay radios as well as channels to meet the transmission demand. Let $f(s,d)$ indicate the satisfaction degree of the pair between s and d for transmission rate, where the range of $f(s,d)$ is [0, 1]. We assume that the form of the satisfaction function is a universal sigmoid function, which can describe diverse communication services of SVs [41],

$$f(s,d) = \frac{1}{1+\exp\left[-\lambda_{s,d}\left(u_{s,d} - u'_{s,d} + \frac{\nu}{\lambda_{s,d}}\right)\right]}, \tag{9}$$

where $u'_{s,d}$ denotes the communication task requirement of s, and we set $\nu > 7$ so that $f(s,d) \ge \frac{1}{1+\exp(-7)} \approx 1$ when the obtained data rate is higher than that of required, i.e., $u_{s,d} \ge u'_{s,d}$. The satisfaction function is shown as a slightly S-shaped curve, which is suitable to express different types of demands. The trend-changing speed denoted by $\lambda_{s,d}$ reflects the urgency of transmitting tasks. Lower values of $\lambda_{s,d}$ lead to lower slopes of the utility curve which means the transmission task is not urgent. Higher values of $\lambda_{s,d}$ lead to steeper curve around $u'_{s,d}$, which indicates the urgent task and strong requirement of resource.

B. Problem Formulation

In the distributed UAV networks with limited resource, UAVs should compete for relay and channel resources to achieve the transmission requirements. However, the decision of one UAVs will affect others, and even affect the performance of global networks. Different from the terrestrial network, the task-driven UAV network needs to ensure the task completion of global network more than the individual respective transmission needs. We study the problem of multi-channel multi-radio relay assignment to improve the global satisfaction experience. The optimization problem can be defined,

$$\text{maxmize} \quad \Lambda = \sum_{s\in\mathcal{S}} f(s,d), \tag{10}$$

$$\text{s.t.} \quad (6),\ (7),\ (8),$$

$$u_r(l,d)\,\kappa_{r,l} \ge u'_{r,l,d}, \tag{11}$$

where the global transmission satisfaction is denoted by Λ. The problem in (10) aims to maximize the aggregate satisfaction of all SVs by optimizing their selection strategies. The constraint in (11) represents that RVs optimize the transmission of SVs under the premise of ensuring its own transmission, where the service demand of r on l is denoted by $u'_{r,l,d}$.

The objective function given by (10) is non-convex/non-concave, which is difficult to be solved in distributed systems with heterogeneous demand and multi-link competition. Both SVs and RVs have capabilities of decision making. SVs need to select appropriate RVs and channels for transmission, while RVs need to access channels to transmit their own data and assist SVs. We analyze and model the strategy of SVs and RVs separately, and a joint algorithm is proposed to promote the global transmission performance. The channel access strategy of RV is identified as ζ_r, $\zeta_r \subseteq \mathcal{L}$. The relay selection strategy of SVs is denoted by ψ_s, $\psi_s \subseteq \mathcal{R}(\mathcal{L})$, which means the radios of $\mathcal{R}$ working on $\mathcal{L}$.

IV. The Congestion Game Model for Distributed Channel Access of RVs

RVs access channels to finish transmission tasks, and help the transmission of SVs on this basis. Due to the sharing characteristics of channel resources, the competition model of RVs is modeled as a distributed congestion game model with player-specific payoff functions [28].

Definition 1. *The channel competition among RVs can be defined as a tuple $\mathcal{G}(\mathcal{R},\mathcal{L},\zeta_r,u_r(\zeta_r,\zeta_{-r}))$. This model will* *reach a state of Nash equilibrium (NE)* $\{\zeta_1, ..., \zeta_r, ..., \zeta_R\}$ *if there is* $\nexists \{\zeta_1, ..., \zeta_r^*, ..., \zeta_R\}$ *satisfying:*

$$u_r\left(\zeta_r^*, \zeta_{-r}\right) > u_r\left(\zeta_r, \zeta_{-r}\right), \tag{12}$$

where the channel access strategies of RVs except r is denoted by $\zeta_{-r} = \{\zeta_1, ..., \zeta_{r-1}, \zeta_{r+1}, , ..., \zeta_R\}$, and $u_r\left(\zeta_r^*, \zeta_{-r}\right)$ represents the utility of r when it changes strategies from ζ_r to ζ_r^*. (12) means that no RV can improve the utility function by changing strategies unilaterally. According to (2), the utility of channel access can be expressed as

$$u_r\left(\zeta_r, \zeta_{-r}\right) = \sum_{l \in \mathcal{L}} \frac{W \log_2\left(1 + \gamma_{rd}\right)}{A_l\left(\zeta_r\right)}, \tag{13}$$

where the number of RVs choosing the same strategy with r on channel l is denoted by $A_l\left(\zeta_r\right)$. This type of model is proved to have a stable solution with at least one Nash equilibrium (NE) as follow. Based on Rosenthal' s potential function [28]

$$\phi_r\left(\zeta_r, \zeta_{-r}\right) = \sum_{l \in \mathcal{L}} \sum_{i=1}^{A_l(\zeta_r)} \frac{W \log_2\left(1 + \gamma_{rd}\right)}{i}, \tag{14}$$

when one of radio of r changes the channel access strategy from l to l^*, the effect of the unilateral change can be given:

$$\begin{aligned} & u_r\left(\zeta_r^*, \zeta_{-r}\right) - u_r\left(\zeta_r, \zeta_{-r}\right) \\ & \quad = \frac{W \log_2\left(1 + \gamma_{rd}^*\right)}{A_{l^*}\left(\zeta_r^*\right)} - \frac{W \log_2\left(1 + \gamma_{rd}\right)}{A_l\left(\zeta_r\right)}. \end{aligned} \tag{15}$$

With the strategies change of RV r, we have $A_{l^*}\left(\zeta_r^*\right) = A_{l^*}\left(\zeta_r\right) + 1$ and $A_l\left(\zeta_r^*\right) = A_l\left(\zeta_r\right) - 1$. The deviating value of the potential function is

$$\begin{aligned} & \phi_r\left(\zeta_r^*, \zeta_{-r}\right) - \phi_r\left(\zeta_r, \zeta_{-r}\right) \\ & = \sum_{i=1}^{A_{l^*}(\zeta_r)+1} \frac{W \log_2\left(1 + \gamma_{rd}^*\right)}{i} + \sum_{i=1}^{A_l(\zeta_r^*)} \frac{W \log_2\left(1 + \gamma_{rd}\right)}{i} \\ & \quad - \sum_{i=1}^{A_{l^*}(\zeta_r)} \frac{W \log_2\left(1 + \gamma_{rd}^*\right)}{i} - \sum_{i=1}^{A_l(\zeta_r)} \frac{W \log_2\left(1 + \gamma_{rd}\right)}{i} \\ & = \frac{W \log_2\left(1 + \gamma_{rd}^*\right)}{A_{l^*}\left(\zeta_r\right) + 1} - \frac{W \log_2\left(1 + \gamma_{rd}\right)}{A_l\left(\zeta_r\right)} \\ & = \frac{W \log_2\left(1 + \gamma_{rd}^*\right)}{A_{l^*}\left(\zeta_r^*\right)} - \frac{W \log_2\left(1 + \gamma_{rd}\right)}{A_l\left(\zeta_r\right)}. \end{aligned} \tag{16}$$

Combining with (15), there is

$$\phi_r\left(\zeta_r^*, \zeta_{-r}\right) - \phi_r\left(\zeta_r, \zeta_{-r}\right) = u_r\left(\zeta_r^*, \zeta_{-r}\right) - u_r\left(\zeta_r, \zeta_{-r}\right), \tag{17}$$

which always holds with $\forall r \in \mathcal{R}, \forall \zeta_r, \zeta_{-r} \subseteq \mathcal{L}$. Based on [28], the proposed multi-channel access model is an exact potential game (EPG) [31] with (14). The most important properties of the EPG can be given, 1) a Potential game has at least one pure strategy NE; 2) global or local maximum potential function constitutes a pure strategy NE.

As shown in Algorithm 1, we propose a distributed channel access algorithm based on the best response (BR) to achieve the NE point. In the initialization, RVs access available channels randomly, which is subject to (7). After that, the loop of channel competition among RVs starts with (18) in Stage II,

Algorithm 1*: Distributed channel competition algorithm*

Stage I: Initialization
- Every radio of RVs accesses one channel randomly.

Stage II: Evolution of channel competition
Repeat in location interaction N
- RVs detect and calculate the expected rewards of the channel access.

For $t = 1, 2, 3...$
- RVs sense and listen to the channel state and competes for opportunity of the update strategy on the selected channel.
- Pick one drone r to select channels with maximal utility function as the strategy $\zeta_r\left(t+1\right)$ which satisfies:

$$\zeta_r\left(t+1\right) = \arg \max_{\zeta_r \subseteq \mathcal{L}} u_r\left(\zeta_r, \zeta_{-r} | t\right), \tag{18}$$

while the other RVs repeat previous strategies of channel access, that is $\zeta_{-r}\left(t+1\right) = \zeta_{-r}\left(t\right)$.
end
Until all of RVs maintain the strategies of channel selection.

where the decision update iteration of RVs is denoted by t. RVs sense to the channel state and compete for update opportunity on the selected channel at each moment. One RV r can change ζ_r to ζ_r^* at time $t+1$, and the other RVs remain the original strategies $\zeta_{-r}\left(t+1\right) = \zeta_{-r}\left(t\right)$. Finally, the stable strategies of channel access can be determined with the evolution of the Stage II. According to [31], the optimal or sub-optimal result of channel access can be achieved by the proposed distributed algorithm in finite iterations among RVs.

V. A Many-to-Many Matching Game for Global Satisfaction-Aware Relay Assignment

A. Many-to-Many Matching Game Model

RVs will assist the transmission of SVs after occupying a part of channel resource. In the multi-UAV communication network, each SV $s \in \mathcal{S}$ will be assigned to one or multiple radios of RVs with occupied certain channels $r\left(l\right) \in \mathcal{R}\left(\mathcal{L}\right)$. Each relay radio can also be shared by multiple SVs. As discussed in Section II, the many-to-many matching game is suitable to model the selection problem, and the definition [38] is given as follows,

Definition 2. *A many-to-many matching* μ *is a mapping by two sets of players* $(\mathcal{S}, \mathcal{R}(\mathcal{L}))$ *and two preference relations* $\succ_s, \succ_{r(l)}$. *Each player* $s \in \mathcal{S}$ *and* $r(l) \in \mathcal{R}(\mathcal{L})$ *constructs preference lists over one another, ranking players in* $\mathcal{S}$ *and* $\mathcal{R}(\mathcal{L})$*, respectively. The matching process is constrained to,*

- $\mu\left(s\right) \subseteq \mathcal{R}(\mathcal{L})$ and $\mu\left(r(l)\right) \subseteq \mathcal{S}$, respectively;
- $|\mu\left(s\right)| \leq q_s$ for $\forall s \in \mathcal{S}$;
- s in $\mu\left(r(l)\right)$ if and only if $r(l)$ in $\mu\left(s\right)$,

where the preference relation $\succ$ is defined as a complete and reflexive binary relation between players in $\mathcal{S}$ and $\mathcal{R}(\mathcal{L})$. $\mu\left(s\right)$ means the subset of relay radios which is connected by SV s, and $\mu\left(r(l)\right)$ is the subset of SVs connecting to relay r with channel l. The quota is the maximal number that each player can match, denoted by q. The second constraint implies that quotas of SVs are fixed while quotas of RVs are dynamic. The third constraint ensures that the matching between two sets of players is by mutual consent. Therefore, the matching game can be expressed as a tuple $\mathcal{G}\left(\mathcal{S}, \mathcal{R}(\mathcal{L}), \succ_s, \succ_{r(l)}, q_s\right)$.

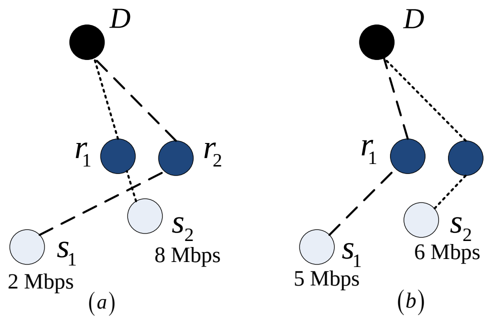


Fig. 3. The shortcoming of the classic matching game model and the comparison with optimum result.

Remark 1. *The unfixed quotas of RVs lead to peer effects among SVs, which perplexes the problem of matching game.*

In classic matching games, quotas of players in two sets are both fixed. However, quotas of RVs in the UAV relay model are unfixed. One relay radio can serve a dynamic number of SVs simultaneously, where SVs connecting with the RV will share the time resource equally. Therefore, the strategy decision of one SV depends on not only the communication quality of the selected relay radios but also the number of SVs having the same strategies. Such problems in the matching game are defined as peer effects [30], these situations cause SVs' decision-making process becoming more complicated.

Remark 2. *The matching game with substitutability developed in the existing work is unsuitable for the sharing multi-access UAV communication networks.*

Existing researches of the matching game in wireless networks mainly developed the matching with substitutability. One player will pursue the best targets within its range of the quota and the other candidates will be rejected. The classic matching game is inapplicable to the resource sharing system mentioned in this paper. Due to unfixed quotas of the RVs, it is reasonable to model the proposed UAV communication networks as a matching market without substitutability.

Moreover, the classic matching game mainly focuses on the individual performance and stable state, but not the global performance. The traditional approach encourages each player to choose and accept the best matching item. For example shown in Fig. 3, even if the global performance obtained by Fig. 3 (b) will be better than that by Fig. 3 (a), the connection result of the classic matching approach will become Fig. 3 (a) because the higher priority of s_2 takes more initiative.

The peer effects, dynamic quotas and the matching without substitutability motivate us to develop new schemes that significantly differ from existing applications of matching game theory in wireless networks such as [33]–[37]. Therefore, the many-to-many matching without substitutability is proposed.

B. Proposed Model of The Matching Game Without Substitutability

In the proposed game, SVs and radios of RVs select elements in the opposing set according to individual matching criteria.

1) For SVs preferences: With heterogeneous types of service, each SV s seeks to find relay radios with high data rate so as to achieve its demands.

$$\begin{aligned} f_s\left(\psi_s, \psi_{-s}\right) \rightarrow 1 \Rightarrow u_s\left(\psi_s, \psi_{-s}\right) \geq {u'}_{s,d} \\ \Rightarrow \text{find } \mu\left(s\right) = \left\{r\left(l\right) \middle| r\left(l\right) \in \mathcal{R}\left(\mathcal{L}\right)\right\}, \end{aligned} \quad (19)$$

$$\text{s.t. } (6), \ (8), \quad (20)$$

where the relay selection strategy of s is denoted by ψ_s, the strategy of the other SVs is denoted by ψ_{-s}, and $\psi_s, \psi_{-s} \subseteq \mathcal{R}\left(\mathcal{L}\right)$. $u_s\left(\psi_s, \psi_{-s}\right)$ represents the utility of s according to (5) with strategies $\left(\psi_s, \psi_{-s}\right)$. $f_s\left(\psi_s, \psi_{-s}\right) \rightarrow 1$ means that s tries to the desired transmission rate ${u'}_{s,d}$, and s will search for suitable relay radio with occupied channel rather than the blind pursuit of high throughput. According to data rate requirements, the preference of s can be expressed as,

$$\left(s, r_n(l_i)\right) \succ_s \left(s, r_m(l_j)\right) \Leftrightarrow u_s\left(r_n, l_i, d\right) > u_s\left(r_m, l_j, d\right), \quad (21)$$

where $r_n, r_m \in \mathcal{R}$, $l_i, l_j \in \mathcal{L}$, $r_n(l_i) \neq r_m(l_j)$. $u_s\left(r_n, l_i, d\right)$ and $u_s\left(r_m, l_j, d\right)$ represent the estimated data rates of s assisted by $r_n(l_i)$ and $r_m(l_j)$, respectively. Therefore, relay radios which can provide better services will obtain higher priorities. The available resource of s is dynamically influenced by the other SVs' choices. Because strategies of the other SVs are uncertain, the preference ordering of one SV is dynamic in each matching iteration according to the obtained data rate.

2) Preferences of radios of RVs: RVs assist the transmission of SVs to improve the global transmission performance. In the matching process, RVs decide whether or not to accept the request of SVs according to the preference criterion. Because the selection strategy of one SV may disturb the other SVs that have the same strategy sets, we define the utility function of relay radios driven by s as,

$$\begin{aligned} U_{r(l)}\left(s, \psi_s\right) = f_s\left(\psi_s, \psi_{\mathcal{J}_s}\right) \\ + \sum_{s_k \in \mathcal{J}_s} \left[f_k\left(\psi_k, \psi_{\mathcal{J}_k}\right) - f_k\left(\psi_k, \psi_{\mathcal{J}_k \setminus s}\right)\right], \end{aligned} \quad (22)$$

where $s_k \in \mathcal{J}_s$ denotes SVs which may be impacted by the decision of SV s. The preference lists of these SVs have partly same elements with s. $f_k\left(\psi_k, \psi_{\mathcal{J}_k}\right)$ is the satisfaction result of the other SVs s_k, and $f_k\left(\psi_k, \psi_{\mathcal{J}_k \setminus s}\right)$ is the satisfaction of s_k if s gives up the selecting radio, which indicates the marginal contribution utility [41].

It should be noted that not all the SVs in $\mathcal{J}_s$ would exactly choose the same radios even if their preference lists have overlapping parts. If s_n exchanges its strategies of selection from ψ_s to $\bar{\psi}_s$, it will affect the SVs which have the same choices of radios. Therefore, the SVs influenced by s can be represented as

$$\mathcal{I}_s = \left\{\forall s_k \middle| \delta_{krl} = 1\right\}, \text{ s.t. } r\left(l\right) \in \left\{\bar{\psi}_s \cup \psi_s\right\}, \quad (23)$$

where $\bar{\psi}_s$ and ψ_s mean the prepared and the current selection strategy of s, respectively. With the matching swap proposed by s, the RVs in both current and prepared strategies will consider the results of the swap. The information exchanges will happen among the radios of RVs in $r\left(l\right) \in \left\{\bar{\psi}_s \bigcup \psi_s\right\}$.

Only SVs in $\mathcal{I}_s$ can influence the transmission performance of s. Thus, it can be found that,

$$f_s\left(\psi_s, \psi_{\mathcal{J}_s}\right)=f_s\left(\psi_s, \psi_{\mathcal{I}_s}\right). \tag{24}$$

In this case, the satisfaction of SVs in $\mathcal{J}_s \backslash \mathcal{I}_s$ will not be influenced no matter which relay radios are chosen by s. Thus, it can also be found that

$$f_k\left(\psi_k, \psi_{\mathcal{J}_k}\right) = f_k\left(\psi_k, \psi_{\mathcal{J}_k \backslash s}\right), k \in \mathcal{J}_s \backslash \mathcal{I}_s. \tag{25}$$

Based on (24)-(25), $U_{r(l)}\left(s, \psi_s\right)$ in (22) can be simplified as

$$\begin{aligned} U_{r(l)}\left(s, \psi_s\right) &= f_s\left(\psi_s, \psi_{\mathcal{I}_s}\right) \\ &+ \sum_{s_k \in \mathcal{I}_s} \left[f_k\left(\psi_k, \psi_{\mathcal{I}_k}\right) - f_k\left(\psi_k, \psi_{\mathcal{I}_k \backslash s}\right)\right], \end{aligned} \tag{26}$$

where the strategies of s will influence SVs $s_k \in \mathcal{I}_s$ and we denote the expected result of $\bar{\psi}_s$ changing from ψ_s as $U_{r(l)}\left(s, \bar{\psi}_s\right)$, that is,

$$\begin{aligned} U_{r(l)}\left(s, \bar{\psi}_s\right) &= f_s\left(\bar{\psi}_s, \psi_{\mathcal{I}_s}\right) \\ &+ \sum_{s_k \in \mathcal{I}_s} \left[f_k\left(\psi_k, \bar{\psi}_{\mathcal{I}_k}\right) - f_k\left(\psi_k, \bar{\psi}_{\mathcal{I}_k \backslash s}\right)\right], \end{aligned} \tag{27}$$

where $\mathcal{I}_s$ is constrained to (23). $f_k\left(\psi_k, \bar{\psi}_{\mathcal{I}_k}\right)$ denotes the satisfaction result of s_k $(k \in \mathcal{I}_s)$ when s changes its selection strategy, and $f_k\left(\psi_k, \bar{\psi}_{\mathcal{I}_k \backslash s}\right)$ is the marginal contribution utility of the strategy $\bar{\psi}_s$ of s. Relay radios determine whether or not to accept the strategy requests of s according to the comparison of (26) and (27). If the utility can be increased, that is,

$$U_{r(l)}\left(s, \bar{\psi}_s\right) > U_{r(l)}\left(s, \psi_s\right), \tag{28}$$

the requests will be accepted. Otherwise, the request will be rejected. Different from the classic matching, matching processes in the proposed matching game without substitutability are not substituted. Once one SV is accepted by the RV, it will share resources of relay radios with other SVs rather than replacing them. The stable matching of the matching game without substitutability is defined.

Definition 3. *A matching result is a global stable matching if and only if no player can improve its matching utility by deviating unilaterally, i.e.,*

$$\nexists \Lambda\left(\bar{\psi}_s, \psi_{-s}\right) > \Lambda\left(\psi_s, \psi_{-s}\right), \tag{29}$$

$$\text{s.t.} \quad \forall s \in \mathcal{S}, \psi_s \subseteq \mathcal{R}\left(\mathcal{L}\right), \left|\psi_s\right| \leq \alpha_s. \tag{30}$$

Inspired by the potential function of potential games in [31], we prove that the optimal solution of the distributed matching problem in (19) and (26) will lead to the global satisfaction result optimizing.

Theorem 1. *Optimizing the local matching process, the global network will achieve a stable matching result.*

Proof: According to the system model and the optimization objective, we construct the potential function [31] as,

$$\Phi\left(\psi_1, \psi_2, ...\psi_S\right) = \sum\nolimits_{s \in \mathcal{S}} f\left(\psi_s, \psi_{\mathcal{I}_s}\right). \tag{31}$$

It can be noted that $\Phi\left(\psi_1, \psi_2, ...\psi_S\right) = \Lambda$, i.e. the defined potential function is equal to the global network satisfaction. Based on (26), the change of matching results caused by the SV's unilateral change can be given,

$$\begin{aligned} U_{r(l)}\left(s, \bar{\psi}_s\right) - U_{r(l)}\left(s, \psi_s\right) &= f_s\left(\bar{\psi}_s, \psi_{\mathcal{I}_s}\right) - f_s\left(\psi_s, \psi_{\mathcal{I}_s}\right) \\ &+ \sum_{s_k \in \mathcal{I}_s} \left[f_k\left(\psi_k, \bar{\psi}_{\mathcal{I}_k}\right) - f_k\left(\psi_k, \psi_{\mathcal{I}_k}\right)\right], \end{aligned} \tag{32}$$

where we suppose that s unilaterally changes its selection from ψ_s to $\bar{\psi}_s$. The $f_k\left(\psi_k, \bar{\psi}_{\mathcal{I}_k}\right)$ is the satisfaction experience of the other SVs when the SV s changes its choice strategies. It can be noted that the $f_k\left(\psi_k, \psi_{\mathcal{I}_k \backslash s}\right)$ is eliminated in the subtraction operation because the strategies of $\psi_{\mathcal{I}_k \backslash s}$ will not be influenced by s.

Similarly, the change of the potential function caused by the unilateral change of s can be given,

$$\begin{aligned} &\Phi\left(\bar{\psi}_s, \psi_{-s}\right) - \Phi\left(\psi_s, \psi_{-s}\right) \\ &= \Big\{ f_s\left(\bar{\psi}_s, \psi_{\mathcal{I}_s}\right) - f_s\left(\psi_s, \psi_{\mathcal{I}_s}\right) \\ &\quad + \sum_{s_k \in \mathcal{I}_s} \left[f_k\left(\psi_k, \bar{\psi}_{\mathcal{I}_k}\right) - f_k\left(\psi_k, \psi_{\mathcal{I}_k}\right)\right] \\ &\quad + \sum_{s_k \in \{\mathcal{J}_s \backslash \mathcal{I}_s\}} \left[f_k\left(\psi_k, \bar{\psi}_{\mathcal{I}_k}\right) - f_k\left(\psi_k, \psi_{\mathcal{I}_k}\right)\right] \\ &\quad + \sum_{s_k \in S \backslash \mathcal{J}_s, s_k \neq s} \left[f_k\left(\psi_k, \bar{\psi}_{\mathcal{I}_k}\right) - f_k\left(\psi_k, \psi_{\mathcal{I}_k}\right)\right] \Big\} \end{aligned} \tag{33}$$

According to the mutual influence relation, the following equation can be given,

$$\sum_{s_k \in \mathcal{S} \backslash \mathcal{J}_s, s_k \neq s} \left[f_k\left(\psi_k, \bar{\psi}_{\mathcal{I}_k}\right) - f_k\left(\psi_k, \psi_{\mathcal{I}_k}\right)\right] = 0. \tag{34}$$

where the preference list of $s_k \in \mathcal{S} \backslash \mathcal{J}_s, s_k \neq s$ has no common elements with s and the strategies changes of s will not influence s_k. Moreover, since the action of s only affects partial SVs whose choice of radio sets contain at least one of the selecting radios of s, the following equation is known:

$$\sum_{s_k \in \{\mathcal{J}_s \backslash \mathcal{I}_s\}} \left[f_k\left(\psi_k, \bar{\psi}_{\mathcal{I}_k}\right) - f_k\left(\psi_k, \psi_{\mathcal{I}_k}\right)\right] = 0. \tag{35}$$

Therefore, (33) can be simplified as,

$$\begin{aligned} &\Phi\left(\bar{\psi}_s, \psi_{-s}\right) - \Phi\left(\psi_s, \psi_{-s}\right) \\ &= f_s\left(\bar{\psi}_s, \psi_{\mathcal{I}_s}\right) - f_s\left(\psi_s, \psi_{\mathcal{I}_s}\right) + \\ &\quad \sum_{s_k \in \mathcal{I}_s} \left[f_k\left(\psi_k, \bar{\psi}_{\mathcal{I}_k}\right) - f_k\left(\psi_k, \psi_{\mathcal{I}_k}\right)\right]. \end{aligned} \tag{36}$$

When s takes a unilateral matching request, the change of the utility is the same with that of the potential function, that is $U_{r(l)}\left(s, \bar{\psi}_s\right) - U_{r(l)}\left(s, \psi_s\right) = \Phi\left(\bar{\psi}_s, \psi_{-s}\right) - \Phi\left(\psi_s, \psi_{-s}\right)$. Thus, the result of each matching process will lead to the same result as the global network performance. The strategies of RVs are based on the properties of EPG, where the definition of NE [31] is similar with the concept of stable matching. Thus, the stable results can be achieved when we optimize the local matching process. Theorem 1 is proved. ■

The game model reduces the information exchange of drones without affecting the quality of the communication. SVs needn't exchange information with the other SVs. RVs

Algorithm 2: Distributed many-to-many learning matching approach (MLMA)

Initialization: Each SV estimates all available relay radio according to the location information, then it forms the preference list of RVs based on (21).
Loop in location interaction N;
Stage I: Relay radios estimate and utilities computation
SVs reform the preference list of RVs based on the CSI in each matching iteration.
Satisfied SVs maintain their strategies of radios selection.
Each unsatisfied SV calculates utility function (9) over its all available relay radio based on the CSI, then it chooses a set of available relay radios based on the mixed strategy, where the component denotes the probability of radio selection $\bar{\psi}_s$ in the mixed strategy is given as

$$p\left(\bar{\psi}_s\right)=\frac{u_s\left(\bar{\psi}_s,\psi_{-s}\right)}{\sum_{\bar{\psi}_s\subseteq\mathcal{R}(\mathcal{L})}u_s\left(\bar{\psi}_s,\psi_{-s}\right)}. \tag{37}$$

Stage II: Matching evaluation
Based on (26) and (27), RVs exchange information with relative RVs and decide whether accept or not with a mixed strategy, where the component contains "accept" $\left(\bar{\psi}_s\right)$ or "reject" $\left(\psi_s\right)$.

$$P=\frac{\exp\left\{\beta\cdot U_{r(l)}\left(\bar{\psi}_s\right),\beta\cdot U_{r(l)}\left(\psi_s\right)\right\}}{\exp\left\{\beta\cdot U_{r(l)}\left(\bar{\psi}_s\right)\right\}+\exp\left\{\beta\cdot U_{r(l)}\left(\psi_s\right)\right\}}, \tag{38}$$

for some learning parameter $\beta>0$.
end Loop until $\nexists\mu$ can improve the satisfaction results.
Output: Convergence to a stable matching result.

only need to exchange information with a portion of the other RVs in the same strategy sets of SVs.

Based on the potential matching game, we design specific implementation methods. We propose different selection schemes for resource allocation in two various cases. For the UAV cluster maintains certain formation shape during the flight, we propose a learning algorithm to achieve the optimal relay allocation result of the whole network with a large probability. Considering the perturbation of UAV networks caused by the deployment changes, a fast matching optimization scheme is proposed, which spends a short time to converge and is robust to the dynamic network.

C. Many-to-Many Learning Matching Approach (MLMA)

We study the problem of relay selection under the scenario with partial information, in which each SV knows the channel state information (CSI) so as to reorganize the preference ordering of relay radios.

The idea of learning algorithms achieving NE point can be seen as promising schemes in the proposed matching game without substitutability. Based on the Theorem 1, the optimum of global satisfaction can be obtained if the optimal stable matching is achieved. Although there is existing work such as [43] studied the optimal solution by spatial adaptive play (SAP) to achieve optimal NE with a large probability. SAP is not suitable for the relay models since it implicitly needs a global controller and its convergence speed is slow. Even in [44] we proposed a C-SAP algorithm, it needs global available strategies information, which may exchange with all of players in the network. SVs with limited ability can not exchange information with the other SVs during every matching process. In order to study the relay selection with limited information exchanges, the distributed learning matching approach is proposed in *Algorithm 2*.

Here, *Algorithm 2* consists of the initialization and two stages in loops: the utility computation of relay radio and the matching evaluation. In the initialization, based on the location information of relay systems, each SV forms its own preference list by estimating payoffs. In Stage I, unsatisfied SVs choose a set of available relay radios based on the mixed strategy (37). In contrast, satisfied SVs maintain their strategies of radios selection. In Stage II, receiving requests from SVs, RVs exchange information with relative RVs and decide to accept or reject the exchange proposal according to (38), where the parameter β balances the tradeoff between exploration and exploitation. The solution is referred to the Boltzmann exploration strategy [45].

Theorem 2. *With a sufficiently large β, the proposed algorithm achieves the global optimum of distributed relay channel matching with an arbitrarily high probability [44].*

Proof: Inspired by [44], we give the proof that the proposed matching approach achieves the maximum potential function values with an arbitrarily high probability.

Intuitively, SV changes the strategy from ψ_s to $\bar{\psi}_s$, with a lightly guidance of performance estimation (37). RVs with higher predicted data rate would achieve the demanded data rate using shorter time resource. Connecting to the prior RVs, SVs have relatively high probability to increase data rate. RVs receive the connection proposals from SVs, judge the rationality of the request and decide whether or not to accept the request according to the (38). If the utility $U_{r(l)}\left(\bar{\psi}_s\right)$ is better than $U_{r(l)}\left(\psi_s\right)$, with a large β, the following inequality holds:

$$\exp\left(\beta\cdot\Phi\left(\bar{\psi}\right)\right)\gg\exp\left(\beta\cdot\Phi\left(\psi\right)\right), \tag{39}$$

and it means that the better strategy can be chosen with an arbitrarily high probability. Similarly, the optimal result of matching process is determined by $\Phi_{\max}=\Phi\left(\psi_{opt}\right)$. When β becomes sufficiently large, i,e, $\beta\to\infty$, the comparison between optimal results and the other results can be given:

$$\exp\left(\beta\cdot\Phi\left(\psi_{opt}\right)\right)\gg\exp\left(\beta\cdot\Phi\left(\psi'\right)\right),\forall\psi'\subseteq\mathcal{R}(\mathcal{L})\backslash\psi_{opt}. \tag{40}$$

In addition, it is assumed that there are one or more optimal selection profiles indexed by $\psi_{opt_1},\psi_{opt_2},...,\psi_{opt_K},K\geq 1$, based on (38) and (39), the optimal results have:

$$\lim_{\beta\to\infty}\sum_{k=1}^{K}\mu\left(\psi_{opt_k}\right)=1, \tag{41}$$

and

$$\lim_{\beta\to\infty}\mu\left(\psi_{opt_k}\right)=\frac{1}{K},\forall k=1,...,K. \tag{42}$$

Therefore, when β becomes large enough, the stationary distribution over all the optimal relay selection profiles, and one or more optimal selection strategies have equal probability to be achieved with limited information exchange of RVs. ■

D. Fast Potential Matching Without Substitutability (FPMA)

To achieve the global optimum of relay selection, the learning algorithm requires a period of iteration for convergence, which can be used in UAVs with stable topology structure.

Algorithm 3*: Fast potential matching approach (FPMA)*

Initialization: Each SV estimates all available relay radio according to the location information, then it forms the preference list of RVs based on (21).
Loop in location interaction N;
Stage I: Relay radios estimate and utilities computation
SVs reform the preference list of RVs based on the CSI in each matching iteration.
Unsatisfied SVs choose a set of available relay radios which they preferred according to the preference list (21).
Satisfied SVs try to reduce one or more connections to save the transmission power. If one SV reduces one radio connection, the satisfaction performance (9) does not become worse, it will quit the connection.
Stage II: Matching evaluation
Based on (28), RVs exchange information with relative RVs and decide to accept the requsts from SVs or not.
RVs equally assign their time resource to SVs which connect to them.
end Loop until $\nexists\mu$ can improve the satisfaction results.
Output: Convergence to a stable matching result.

When the network topology is disturbed by such as position changing, a faster matching algorithm is needed to adapt to the dynamically changing network environment.

We discuss the fast matching method without substitutability in this section, shown in *Algorithm 3*. In the initialization, each SV forms its own preference list according to the estimation of data rate. Although SVs can obtain the CSI information, they can't understand either the strategies of the other SVs or what influence would cause if they make a strategy. Therefore, each SV constructs the preference list according to the estimated value. After the initialization, each loop in the iteration N is divided into two steps:

In Stage I, SVs reform the preference list of RVs based on the CSI in each iteration. Each unsatisfied SV requests to its prior RVs with the number smaller than that of its radios. In contrast, in order to save the transmission power, SVs which satisfy their data rates will try to reduce the number of connections. In Stage II, receiving requests from SVs, RVs exchange information with relative RVs and determine whether or not to accept matching requests according to (28). It can be noted that if the change of the matching results improves the local performance of RVs, the requests of SVs would be accepted by RVs.

Theorem 3. *After a finite number of matching steps, SVs and RVs become certain matching results, and thus the proposed FPMA can converge to a stable matching result.*

Proof: In the UAV communication system, the numbers of both SVs and RVs are finite, thus the length of the preferences of SVs is limited. For SVs s feeling satisfied, that is $f_s(\psi_s, \psi_{-s}) = 1$, there is no need to select other relay radios.

For SVs which are not satisfied with obtained data rates, they will reform their preference lists before the strategies making and then select the preferred RVs. RVs will filter the requirement according to (28). According to *Theorem 1*, the acceptance decision guarantees that,

$$\Phi\left(\bar{\psi_s}, \psi_{-s}\right) > \Phi\left(\psi_s, \psi_{-s}\right), \tag{43}$$

It can be noted that proposals from SVs will be accepted only if the global performance can be promoted. Due to the upper limit of global optimum, SVs will not be able to achieve results better than the optimum, thus it will converge to a stable solution. ■

Algorithm 4*: Joint algorithm of channel access and relay selection*

Phase I: Initialization
- Perform the Stage I of Algorithm 1, where all RVs randomly access channels.

Phase II: Double layer evolution
Repeat in location interaction N
- Radios of RVs occupy partial channel resources so that $u_r(l,d)\,\kappa_{r,l} \geq u'_{r,l,d}$, and remainder resources are used to assist SVs.
- Perform Algorithm 2 or Algorithm 3 according to the environment and the requirement of optimization.
- Perform Stage II of Algorithm 1, where utilities of RVs are based on (13).

Until no selection strategy is changed at the previous iteration.

Although there are a large number of selection strategies for every SV, the reasonable selection is the RVs located closely. With the filtration by the preference ordering, prior radios can be quickly found. Therefore, in the model with finite number of UAVs, the FPMA can fast converge to the stable matching.

VI. Joint Algorithm of Channel Access and Relay Assignment

In the previous two sections, we analyzed the models and methods of channel access and relay selection, respectively. In this section, the two models are combined together and a joint algorithm of channel access and relay selection is proposed to optimize both the transmission of RVs and SVs. The proposed algorithm is shown in Algorithm 4.

The overall joint algorithm begins with random channel access for RVs. RVs intercept part of the time resources $\kappa_{r,l}$ according to its own transmission requirements (11) and uses the remaining resources to assist SVs. The relay allocation can be divided into two ways according to the network environment. When the UAV network is in a stable flight state, the optimal result of the whole network is realized by Algorithm 2. When the drone is disturbed, the Algorithm 3 is used to achieve rapid response of the dynamic network. RVs and SVs iterate channel access and relay selection at each moment. The relay selection of SVs will affect the channel access of the next stage.

Due to the sharing characteristics of channel resources, the competition among RVs will result in an average utilization of channel resources. In the UAV relay selection model, the numbers of both SVs and RVs are finite, thus the length of the preferences of SVs is limited. Based on the Theorem 1, proposals from SVs will be accepted only if the global performance can be promoted. Due to the upper limit of global optimum, SVs will not be able to achieve results better than the optimum. When channel access is in an average state, the relay selection will converge to the stable results according to Theorem 2 and Theorem 3.

Remark 3. *The complexity costs and the distributed implementation of proposed algorithms are discussed.*

The convergence iterations of the MLMA and the FPMA are denoted by N_1 and N_2, respectively. The complexity of SVs and RVs estimating the received payoff can be expressed as $\mathcal{O}(C_1)$, where C_1 is a constant decided by the time of the estimation period [46]. For the MLMA, the procedure of updating the relay selection involves finite sums and 1 scalar-vector product, and the complexity can be expressed as $\mathcal{O}(C_2)$, where C_2 is a small constant. The procedure of exchanging information among RVs involves 3 exponents and 1 scalar-vector product, and we denoted the complexity as $\mathcal{O}(C_3)$. The total complexity of the Algorithm 4 with MLMA can be given by,

$$N_1\left[\left(S+\sum_{r\in\mathcal{R}}\alpha_r\right)\left(\mathcal{O}\left(C_1\right)+\mathcal{O}\left(C_2\right)\right)+\sum_{r\in\mathcal{R}}\alpha_r\mathcal{O}\left(C_3\right)\right], \tag{44}$$

where the complexity of channel access in each iteration is $\sum_{r\in\mathcal{R}}\alpha_r(\mathcal{O}\left(C_1\right)+\mathcal{O}\left(C_2\right))$.

For the FPMA, satisfied SVs need not exchange information with other drones. To calculate maximum complexity, we assume that all SVs are unsatisfied during iterations. The procedure of exchanging information among RVs involves operations of no more than L sums and 1 comparison, so its maximum complexity can be expressed as $\mathcal{O}(C_5)$, where C_5 is a small constant. Thus, the total complexity of the Algorithm 4 with FPMA can be given by,

$$N_2\left[\left(S+\sum_{r\in\mathcal{R}}\alpha_r\right)\mathcal{O}\left(C_1\right)+\sum_{r\in\mathcal{R}}\alpha_r\left(\mathcal{O}\left(C_2\right)+\mathcal{O}\left(C_5\right)\right)\right]. \tag{45}$$

The complexity costs of both proposed algorithm have the linear relationship with number of drones. If the centralized method is applied, the complexity will be greatly increased caused by

$$\prod\nolimits_{s=1}^{S}\left(\begin{array}{c}\alpha_s\\ \sum_{r\in\mathcal{R}}\alpha_r\end{array}\right)\cdot\prod\nolimits_{r=1}^{R}\left(\begin{array}{c}\alpha_r\\ L\end{array}\right), \tag{46}$$

where α_s and α_r mean the number of radios of s and r, respectively. Unlike the centralized control, the computational pressure is dispersed to individual UAVs by the proposed methods, which reduces the load on the UAV cluster head. Therefore, the proposed distributed algorithms are more suitable for self-organizing UAV communication networks.

VII. Simulation Results and Discussions

In a 2000 cubic meters topology, as shown in Fig. 4, there is a destination located at the center of this topology. Several RVs (represented by pentagrams) are close to the UAV destination ($50m < D_{rd} < 650m$) and cell edged SVs (represented by squares) are randomly distributed ($700m < D_{sd}$). Each UAV is equipped with two radios. In order to compare the proposed algorithms with the existing work [17], [23] of relay models, the experimental parameters follow the simulation methodologies of 3GPP specifications [47]. It can be noted that the numerical results can be extended to the other communication environments such as in [25].

The maximum transmission powers of SVs and RVs are set to be 20 dBm and 30 dBm, respectively. The transmission requirement $u'_{r,l,d}$ of RVs on channels l are randomly ranging from 0-5 Mbps, and we optimize the transmission of SVs while ensuring the requirements of RVs. SVs have individual

TABLE I
Parameter settings in simulations.

Transmission bandwidth	10 MHz
Transmission power of SVs	20 dBm
Transmission power of RVs	30 dBm
Noise density	-174 dBm/Hz
Carrier Frequency [47]	2 GHz
Channel gain of direct links [47]	131.1+42.8lg$D$(km)
Channel gain of relay links [47]	103.8+20.9lg$D$(km)
Requirements of source drones	5 Mbps - 25 Mbps
Trend-changing speed in (9)	$0.5 < \lambda_{s,d} \leq 3$ randomly

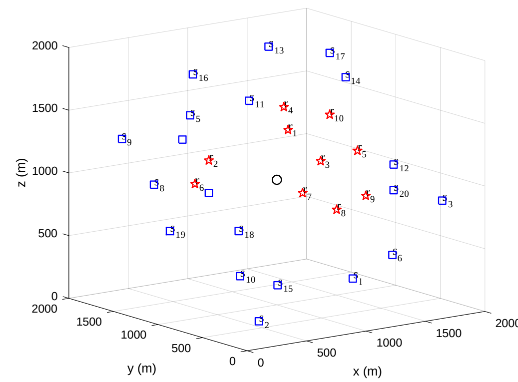


Fig. 4. The UAV communication network topology of the simulation.

data rate requirements ranging from 5-25 Mbps randomly and the $\lambda_{s,d}$ of each SV s is also random ($0.5 < \lambda_{s,d} \leq 3$). All results are obtained by simulating 800 topologies independently and taken the expected values. To balance the tradeoff between exploration and exploitation, $\beta = 1 + b \times N$ is chosen in our simulation, where N means the the iteration step and b is a constant which equals to 0.5 in the simulation.

A. Satisfaction Performance

We consider a UAV communication network consisting of 20 SVs and 12 RVs with 10 available channels. The average convergence performance of proposed distributed potential

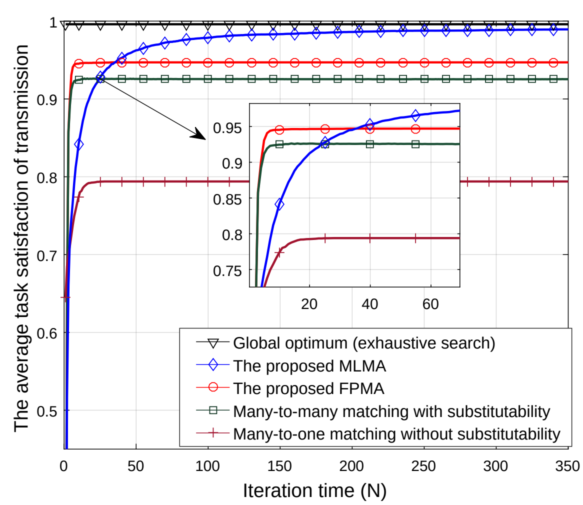


Fig. 5. Evolution results of satisfaction performance with 20 SVs, 12 RVs and 10 channels.

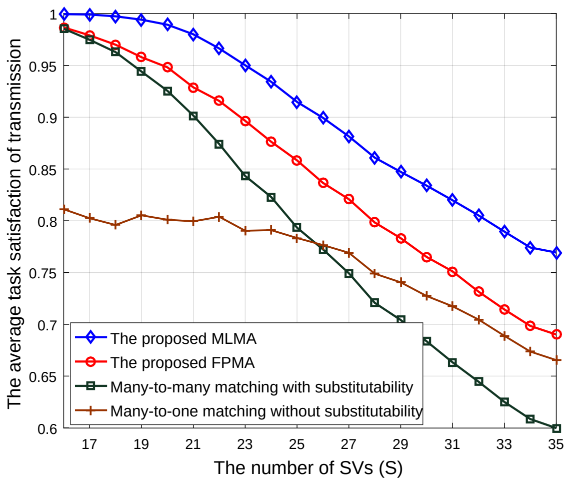


Fig. 6. The comparison of average satisfaction among SVs with heterogeneous requirements.

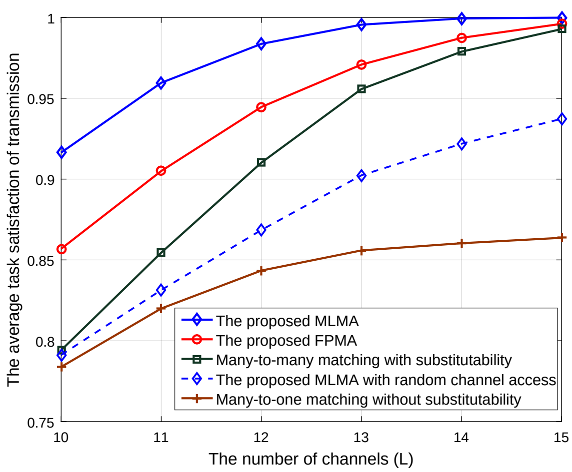


Fig. 7. The comparison of average satisfaction among 25 SVs with 12 RVs and varying number of channels.

matching approaches are shown in Fig. 5, where the global optimum result is obtained by the exhaustive searching solution. Combined with Algorithm 1, both the FPMA and the MLMA are applied. It is noted from Fig. 5 that the MLMA catches up with the global optimum, while the FPMA achieves a stable matching result by fewer iterations. The performance comparison of the many-to-many matching with substitutability [42] and many-to-one matching without substitutability [39] are given. It can be seen from the figure that the proposed algorithms are more effective than the matching with substitutability model and the many-to-one matching model.

To demonstrate the performance of proposed algorithms under different network sizes, the average satisfaction results for different algorithms with varying numbers of SVs are compared. We consider the UAV networks with 12 RVs and varies the number of SVs working on 10 channels. Fig. 6 shows the average global satisfaction values of various algorithms. It is illustrated in Fig. 6 that, two proposed many-to-many relay selection algorithms have significant advantage in the term of satisfaction at all network sizes. Particularly, the global satisfaction value of MLMA is higher than 0.95 when the number of source-destination pairs is no more than 26. It can be noted that, with more time to learn the selection strategies, the MLMA performs better than the FPMA.

In the simulation, the classic algorithm in the many-to-many model and a many-to-one matching with substitutability are given, respectively. When the number of SVs is less than 26, the performance of the alternative many-to-one model is significantly lower than other algorithms. The many-to-many model with substitutability declined significantly. When the number of source nodes is greater than 26, the many-to-many matching with substitutability has the lowest performance. The reason are discuss as follow.

Although the many-to-many optimization model is conducive to flexible resource utilization, the unreasonable allocation method may intensify the competition between drones, resulting in a drastic decline in performance. Reasonable many-to-many optimization methods, such as the proposed two algorithms, can ensure the improvement of transmission performance. The optimization performance of FPMA tends to be stable with the increase SVs, which indicates that the algorithm avoids resource conflicts caused by many-to-many models under the premise of fast convergence. Compared with the FPMA, the MLMA is more robust to the dense networks.

Fig. 7 shows the performance variation of UAV network with 25 SVs and 12 RVs as the number of channels increases from 10 to 15. The figure verifies the effective performance of the proposed algorithms under different channel resources. It is noted that the flexibility of the many-to-many optimization model has an obvious advantage on the improvement of network performance with the increase of resources. Furthermore, the result of the MLMA with random channel access shows that unreasonable channel allocation scheme will result in reduced network performance. The result verifies the satisfactory performance of the proposed algorithm of channel access.

B. Fairness Performance

Here, the global fairness performance among various algorithms is compared based on the Jains fairness index (JFI) [48], which is given as:

$$J = \frac{\left(\sum_{s\in\mathcal{S}} f\left(s,d\right)\right)^2}{S\sum_{s\in\mathcal{S}} f^2\left(s,d\right)}, \tag{47}$$

where $f\left(s,d\right)$ is the satisfaction index in (9) and S is the number of SVs. High satisfaction fairness of SVs will bring high value of J which will arrive at the largest value 1 if all SVs have a same level of transmission task completion.

As shown in Fig. 8, JFI of satisfaction performances of various algorithms are compared in UAV networks consisting of 12 RVs, 10 channels and varying SVs. Similar to the satisfaction performance, two proposed algorithms perform very good results in fairness performance and MLMA has a better performance than the FPMA. The proposed algorithms have significant advantage in term of the JFI of satisfaction. These results demonstrate that the proposed matching game approaches can achieve good fairness performance among SVs under the premise of satisfaction optimized. Due to the limited

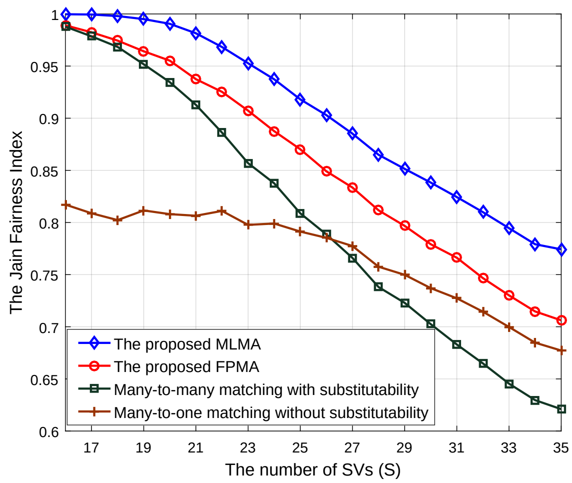


Fig. 8. The comparison of fairness performance of satisfaction among SVs with heterogeneous requirements.

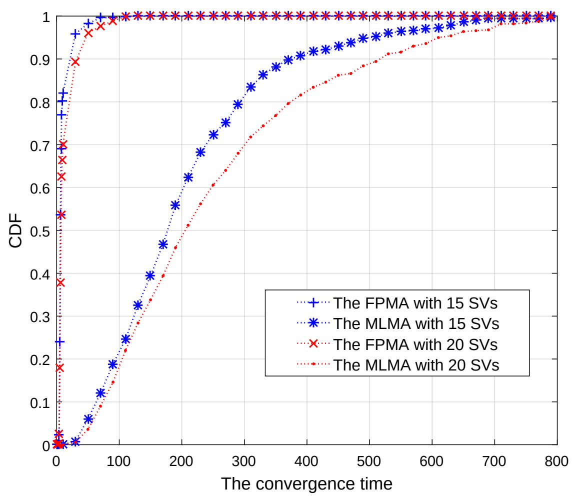
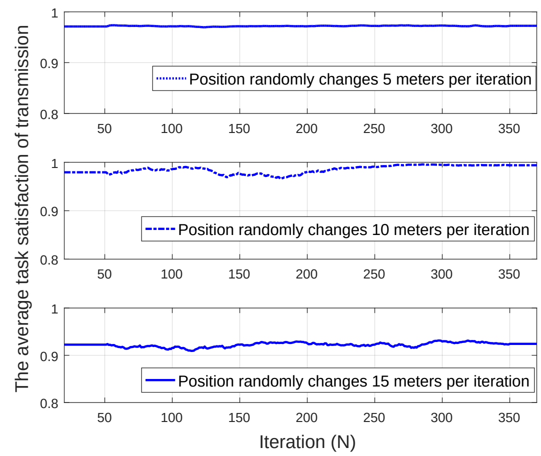


Fig. 9. The convergence performance of FPMA and MLMA with various SVs in UAV networks.

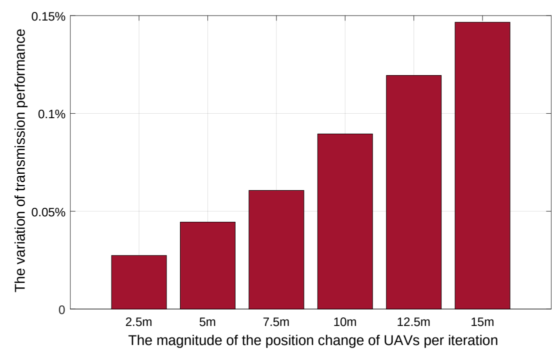
number of relay radios, the JFI performance of all the schemes decreases as the number of SVs increases.

C. Convergence Performance

Fig. 9 shows the cumulative distribution function (CDF) for the convergence time of two proposed algorithms. The simulation results are calculated by 500 independent experiments. It is seen that the convergence time of the FPMA does not exceed 100 iterations for the network with 25 SVs and 12 RVs. More than 90 percent of the probability that the FPMA achieves the stable matching within 30 iterations. The MLMA needs more time than the FPMA to achieve the stable matching so as to achieve the optimum. We can see that, the average number of iteration increases due to the increase of the number of players in the UAV network. With the number of drones increases, the collisions among SVs also increase. However, the collisions have a final upper boundary because the UAV network becomes a saturated state. Fig. 9 demonstrates that the MLMA has a reasonable convergence time that there is more than 80 percent of the probability to achieve the stable results within 400 iterations. These results validate the convergence performance of the proposed distributed algorithms.



(a) Jitter amplitude of global transmission performance with different varying position change



(b) The statistical result

Fig. 10. The effect of SVs position change on the transmission stability.

We verify the robustness of the FPMA in dynamic environment by assuming that the positions of UAVs are changing during the flight process. Assume that the update interval for location interaction of the UAV network is 100ms, and the position change range of each UAV in one iteration time is lower than 15 meters. In the simulation, during the iteration time between 50 and 350, the relative positions of all SVs vary randomly and continually. The three subgraphs in Fig. 10 (a) show the effect on the transmission performance when the upper limit of the position change is 5, 10 and 15 meters respectively. As the magnitude of the change increases, the jitter of the transmission performance is more severe. The statistical results are shown in Fig. 10 (b). When the movement range of UAVs $\leq$ 15 meters in each iteration, the jitter of the transmission performance of the global network remains within a stable range, which is below 0.15%. Therefore, the proposed FPMA can better cope with the dynamic changes of UAV network topology, so that the network maintains stable performance in a dynamic environment.

The robustness of the self-organizing FPMA algorithm is further investigated. Based on the topology in Fig. 4, we

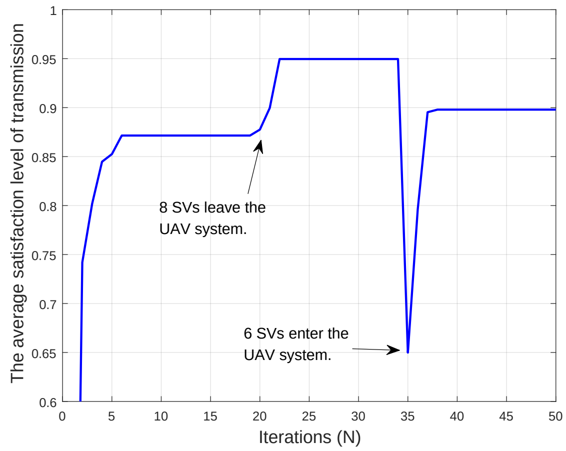


Fig. 11. Dynamics of satisfaction performance with perturbations. At iteration N = 20 and 35, 8 SVs leave the system and 6 new SVs enter the system, respectively.

consider $S = 20$ SVs sharing 12 RVs with 10 channels. At iteration $N = 20$ and 35, we let 8 SVs leave the system and 6 new SVs enter the system, respectively. The results in Fig. 11 show that the proposed FPMA quickly converges to a stable matching result after the perturbations occur. It verifies that the FPMA is robust to the dynamics of communication tasks in the UAV networks. In our future work, the dynamic situation of drones will be considered in the real-time strategy making.

VIII. Conclusion

This paper modeled the multiple access relay assignment system as a game model, where the channel access is constructed as a congestion game for the shared competition characteristics, and the selection relationship between UAVs is modeled as a many-to-many matching market. To optimize the global satisfaction performance among task-driven UAVs, we designed a potential matching framework, in which the global satisfaction performance can be improved by optimizing the local matching. We proposed two distributed algorithms for different cases, namely, the MLMA and FPMA algorithms, respectively. Both two algorithms were proved to achieve the stable matching of global networks. Numerical results shown that the proposed algorithm improves both of the global satisfaction and fairness performances. Specifically, the MLMA performs excellent satisfaction results, while the FPMA has a remarkable convergence performance which is robust to dynamic UAV communication networks.

References

[1] M. Mozaffari, W. Saad, M. Bennis, et al., "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2334-2360, Thirdquarter 2019.

[2] Y. Zeng, J. Lyu and R. Zhang, "Cellular-connected UAV: Potential, challenges, and promising technologies," *IEEE Wireless Communications*, vol. 26, no. 1, pp. 120-127, February 2019.

[3] L. Gupta, R. Jain and G. Vaszkun, "Survey of important issues in UAV communication networks," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1123-1152, Secondquarter 2016.

[4] D. Liu, Y. Xu, J. Wang, et al., "Self-organizing relay selection in UAV communication networks: A matching game perspective," *IEEE Wireless Communications*, early access. (DOI: 10.1109/MWC.001.1700434)

[5] I. Bekmezci, O. K. Sahingoz, and S. Temel, "Flying ad-hoc networks (FANETs): A survey," *Ad Hoc Netw.*, vol. 11, no. 3, pp. 1254-1270, 2013.

[6] Joel G. Manathara, P. B. Sujit and Randal W. Beard, "Multiple UAV coalitions for a search and prosecute mission," *Intell. Robotics Syst.*, vol. 62, no. 1, pp. 125-158, April 2011.

[7] S. Mousavi, F. Afghah, J. D. Ashdown and K. Turck, "Leader-follower based coalition formation in large-scale UAV networks, a quantum evolutionary approach," in *Proc. INFOCOM WKSHPS*, Honolulu, HI, pp. 882-887, 2018.

[8] B. D. O. Anderson, B. Fidan, Yu. C , et al., "UAV formation control: Theory and application," *Lecture Notes in Control & Information Sciences*, vol. 371, no. 3, pp. 15-33, 2008.

[9] J. N. Laneman, D. N. C. Tse and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062-3080, Dec. 2004.

[10] D. T. Ho, E. I. Grøtli, S. Shimamoto, et al., "Optimal relay path selection and cooperative communication protocol for a swarm of UAVs," in *Proc. IEEE Globecom Workshops*, Anaheim, CA, pp. 1585-1590, 2012.

[11] P. Zhang, Q. Zhang, M. Jiang and Z. Feng, "Cube based space region partition routing algorithm in UAV networks," in *Proc. APCC*, Perth, WA, pp. 1-6, 2017.

[12] A. Mukherjee, S. Misra, V. S. P. Chandra and M. S. Obaidat, "Resource-optimized multiarmed bandit-based offload path selection in edge UAV swarms," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4889-4896, June 2019.

[13] M. Sbeiti, N. Goddemeier, D. Behnke, et al., "PASER: Secure and efficient routing approach for airborne mesh networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 1950-1964, Mar., 2016.

[14] D. Yang, X. Fang, and G. Xue, "OPRA: Optimal relay assignment for capacity maximization in cooperative networks," in *Proc. IEEE Int. Conf. Commun.* (ICC), pp. 1-6, Kyoto, Japan, Jun. 2011.

[15] S. Sharma, Y. Shi, Y. T. Hou, and S. Kompella, "An optimal algorithm for relay node assignment in cooperative ad hoc networks," *IEEE/ACM Trans. Netw.*, vol. 19, no. 3, pp. 879-892, Jun. 2011.

[16] H. Xu, L. Huang, C. Qiao, et al., "Shared relay assignment (SRA) for many-to-one traffic in cooperative networks," *IEEE Transactions on Mobile Computing*, vol. 15, no. 6, pp. 1499-1513, 2016.

[17] Z. Chen, T. Lin and C. Wu, "Decentralized learning-based relay assignment for cooperative communications," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 2, pp. 813-826, Feb. 2016.

[18] N. Nomikos et al., "A survey on buffer-aided relay selection," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1073-1097, Secondquarter 2016.

[19] N. Nomikos, T. Charalambous, D. Vouyioukas, et al., "Hybrid NOMA/OMA with buffer-aided relay selection in cooperative networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 3, pp. 524-537, June 2019.

[20] Z. Ding, H. Dai and H. V. Poor, "Relay selection for cooperative NOMA," *IEEE Wireless Communications Letters*, vol. 5, no. 4, pp. 416-419, Aug. 2016.

[21] Zheng Yang, Zhiguo Ding, Yi Wu, et al., "Novel relay selection strategies for cooperative NOMA," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10114-10123, 2017.

[22] Long Yang, Qiang Ni, Lu Lv, et al., "Cooperative non-orthogonal layered multicast multiple access for heterogeneous networks," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1148-1165, 2019.

[23] J. Jing, D. Liu, Y. Xu, et al., "Distributed channel access, relay selection and time assignment for QoE-aware relay networks," *IEEE Access*, vol. 6, pp. 28790-28800, 2018.

[24] A. Kuhestani, A. Mohammadi and M. Mohammadi, "Joint relay selection and power allocation in large-scale MIMO systems with untrusted relays and passive eavesdroppers," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 2, pp. 341-355, Feb. 2018.

[25] M. Mozaffari, W. Saad, M. Bennis, et al., "Wireless communication using unmanned aerial vehicles (UAVs): Optimal transport theory for hover time optimization," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 8052-8066, Dec. 2017.

[26] S. Zhang, Y. Zeng and R. Zhang, "Cellular-enabled UAV communication: A connectivity-constrained trajectory optimization perspective," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2580-2604, March 2019.

[27] J. Zhao, F. Gao, L. Kuang, et al., "Channel tracking with flight control system for UAV mmWave MIMO communications," *IEEE Communications Letters*, vol. 22, no. 6, pp. 1224-1227, June 2018.

[28] B. Vöking and R. Aachen, "Congestion games: Optimization in competition," in *Proc. Algorithms Complex. Durham Workshop*, pp. 9-20, 2006.

[29] Y. Gu, W. Saad, M. Bennis, et al., "Matching theory for future wireless networks: Fundamentals and applications," *IEEE Communications Magazine*, vol. 53, no.5, pp. 52-59, 2015.

[30] E. Bodine-Baron, C. Lee, A. Chong, et al., "Peer effects and stability in matching markets," in *Proc. 4th Symp. Algorithmic Game Theory*, pp. 117-129, 2011.

[31] D. Monderer and L. S. Shapley, "Potential games," *Games Econ. Behavior*, vol. 14, no. 1, pp. 124-143, 1996.

[32] D. Liu, Y. Xu, Y. Xu, et al., "Distributed relay selection for heterogeneous UAV communication networks using a many-to-many matching game without substitutability", in *Proc. IEEE/CIC ICCC*, pp. 1-6, Oct. 2017.

[33] O. Semiari, W. Saad, M. Bennis, et al., "Caching meets millimeter wave communications for enhanced mobility management in 5G networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 779-793, Feb. 2018.

[34] S. Bayat, Y. Li, L. Song, et al., "Matching theory: applications in wireless communications," *IEEE Signal Processing Magazine*, vol. 33, no. 6, pp. 103-122, Nov. 2016.

[35] M. Simsek, T. Hößler, E. Jorswieck, et al., "Multiconnectivity in multicellular, multiuser systems: A matching-based approach," *Proceedings of the IEEE*, vol. 107, no. 2, pp. 394-413, Feb. 2019.

[36] B. Di, L. Song and Y. Li, "Radio resource allocation for uplink sparse code multiple access (SCMA) networks using matching game," in *Proc. IEEE ICC*, Kuala Lumpur, pp. 1-6, 2016.

[37] S. Zhang, B. Di, L. Song et al., "Radio resource allocation for non-orthogonal multiple access (NOMA) relay network using matching game," in *Proc. IEEE ICC*, Kuala Lumpur, pp. 1-6, 2016.

[38] K. Hamidouche, W. Saad and M. Debbah, "Many-to-many matching games for proactive social-caching in wireless small cell networks," in Proc. *Int. Symp. Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pp. 569-574, Hammamet, Tunisia, May 2014.

[39] M. Pycia, "Many-to-one matching without substitutability," *Mit Industrial Performance Center Working Paper*, 2005.

[40] Y. Xu, J. Wang, Q. Wu, et al., "Dynamic spectrum access in time-varying environment: Distributed learning beyond expectation optimization," *IEEE Transactions on Communications*, vol. 65, no. 12, pp. 5305-5318, Dec. 2017.

[41] Q. Wu, D. Wu, Y. Xu, et al., "Demand-aware multichannel opportunistic spectrum access: A local interaction game approach with reduced information exchange," *IEEE Transactions on Vehicular Technology*, vol. 64, no.10, pp. 4899-4904, 2015.

[42] D. Gale and L. Shapley, "College admissions and the stability of marriage," *Amer. Math. Monthly*, vol. 69, pp. 9-15, Jan. 1962.

[43] J. Marden, G. Arslan, and J. Shamma, "Cooperative control and potential games," *IEEE Trans. Syst.*, Man, Cybern. B, vol. 39, no. 6, pp. 1393-1407, Dec. 2009.

[44] Y. Xu, J. Wang, Q. Wu, et al., "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, April 2012.

[45] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction.* Cambridge, MA: MIT Press, 1998.

[46] Y. Xu, Q. Wu, L. Shen, et al., "Opportunistic spectrum access with spatial reuse: Graphical game and uncoupled learning solutions," *IEEE Transactions on Wireless Communications*, vol. 12, no. 10, pp. 4814-4826, Oct. 2013.

[47] 3GPP TR 36.814, "Further advancements for E-Utra: Physical layer aspects," *Technical Specification Group Radio Access Network*, June 2009.

[48] R. Jain, D. Chiu, and W. Haws, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system," *Technical Report*, 1984.

View publication stats