IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON ARTIFICIAL INTELLIGENCE FOR PHYSICAL-LAYER WIRELESS COMMUNICATIONS

Received December 31, 2019, accepted January 11, 2020, date of publication January 15, 2020, date of current version January 27, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2966740

AI-Aided Downlink Interference Control in Dense Interference-Aware Drone Small Cells Networks

LIXIN LI[1], (Member, IEEE), ZIHE ZHANG[1,2], KAIYUAN XUE[1], MENG WANG[1], MIAO PAN[3], (Senior Member, IEEE), AND ZHU HAN[3,4], (Fellow, IEEE)

[1]School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China
[2]54th Research Institute of China Electronic Science and Technology Group Corporation, Shijiazhuang 050081, China
[3]Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004, USA
[4]Department of Computer Science and Engineering, Kyung Hee University, Seoul 446-701, South Korea

Corresponding author: Lixin Li (lilixin@nwpu.edu.cn)

This work was supported in part by the Aerospace Science and Technology Innovation Fund of China Aerospace Science and Technology Corporation, in part by the Shanghai Aerospace Science and Technology Innovation Fund under Grant SAST2018045, Grant SAST2016034, and Grant SAST2017049, in part by the China Fundamental Research Fund for the Central Universities under Grant 3102018QD096, in part by the National Natural Science Foundation of China under Grant 61901379, in part by the Natural Science Basic Research Plan in Shaanxi Province of China under Grant 2019JQ-253, in part by the Seed Foundation of Innovation and Creation for Graduate Students in Northwestern Polytechnical University under Grant ZZ2019024, and in part by U.S. MURI AFOSR MURI under Grant 18RT0073, Grant NSF EARS-1839818, Grant CNS1717454, Grant CNS1731424, Grant CNS-1702850, Grant CNS-1646607, Grant CNS-1350230 (CAREER), and Grant CNS-1801925.

ABSTRACT Recently, drone small cells (DSCs) has been brought into significant focus, which is the one key enabler for potentially facilitating terrestrial wireless communication systems. Meanwhile, in ultra-dense unmanned networks, artificial intelligence (AI) has been a useful and efficient tool for control and management of the multi-agents. This paper investigates a downlink interference control problem in ultra-dense unmanned networks with AI-aided approach, that each DSC can adjust its altitude to increase the data-rate. This problem is formulated as a mean field game (MFG) framework, an AI-aided method to make decisions. In this framework, each DSC controls its velocity to minimize the cost over a period, where the cost function is composed by the data-rate and height adjusting consumption. Meanwhile, in this model, we adopt the mean-field approximation (MFA) approach to derive the interference introduced from a large number of DSCs. Besides, the control strategy is described and explained by using the related Hamilton-Jacobi-Bellman (HJB) and Fokker-Planck-Kolmogorov (FPK) equations, respectively. Thus, a finite difference algorithm is proposed to solve the coupled partial differential equations, which can obtain the optimal altitude control strategy. The algorithm outputs show the optimal behaviors of DSCs in different environment scenarios. In additon, the simulation results verify that the proposed control strategy has better average signal to interference plus noise ratio (SINR) compared with the baseline method.

INDEX TERMS Artificial intelligence, drone small cell, downlink interference control, mean field game, ultra-dense unmanned networks.

I. INTRODUCTION

Recently, unmanned aerial vehicles (UAVs), which have many advantages such as flight flexibly as well as excellent volume and mass have became an available emerging choice for air-to-ground wireless links. Specially, it plays an essential role in extensive application scenarios such as military damage assessments, environmental monitoring, and so far [2]–[6]. Although UAVs have been traditionally developed by the government and military to carry on mission-critical supervision and monitoring works, they are also getting more applicable to many civilian applications. A representative UAV, equipped with wireless transmitters, has the abilities to communicate with other aerial or ground wireless equipments (referred to ''users'' in this paper). The authors in [7] proposed an aerial-ground cooperative vehicular networking architecture, consisting of an aerial subnetwork and a ground vehicular subnetwork. And the authors in [8] also considered a multi-UAV enabled wireless communication system. In [9], the drone small cells (DSCs) are proposed to be applied as the aerial base stations (BSs) to provide the air-to-ground cellular link services in some crucial

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

environments or marginal demands. The primary advantage of employing DSCs is that they can operate without pilots, and hence they are able to be autonomously deployed in any scenarios for different applications such as monitoring, rescue, and communication.

What's more, DSCs can be relocated easily and rapidly based on demand, and the adjustable height makes the drones can meet different requirements based on user densities, desired data rate and interference effects. Meanwhile, DSCs have the potential to complement even substitute the terrestrial cellular networks by serving users in severe shadowing conditions, serving the overloaded or damaged terrestrial BSs, serving users around overloaded BSs (such as mobile data offloading) in ultra-dense networks. Captive UAVs have been employed in these drones-aided cellular networks recently. Due to the highly flexible and energy-constrained coexisting in DSCs, there are many issues on them such as their management and charging. Particularly, the placement as well as deployment of DSCs has gained significant interests [10]–[13]. For example, a DSCs 3-D placement problem was modeled in [10] with the target of maximizing the utility of the network. Therefore, as a kind of stationary captive platform, tethered buoyant DSCs, powered by ground vehicles with cable, can maintain a long time operation, which quite solve the energy constraints on DSCs. Therefore, this kind of DSCs have drawn increasing attentions in recent years.

In order to adapt the evolving cellular networks which are ultra-dense and heterogeneous, the number of DSCs have to increase incessantly. Meanwhile, the distributed and mass deployments of DSCs cause the decisions of DSCs' resource allocation anticipated to be taken individually (i.e., distributed static policies or dynamic strategies). Some of the existing strategies or policies on cellular networks are not suitable in the context of DSCs as they mostly depend on central managers to make strategic decisions. For instance, the placement algorithm of DSCs in [13] could calculate the optimal altitudes and transmit power for each DSC, while it worked with the given global information. What's more, for the dense networks, shortening the information exchanges among DSCs is preferred as a result of the limited energy of DSCs. And there are some works that focus on the resource allocation via distributed algorithms in the context of cellular networks. The authors in [14] described a cell coloring based distributed frequency allocation approach for all kinds of cellular networks. And in [15], the distributed resource allocation for D2D communications underlying cellular networks was considered. Different from the existing centralized algorithms for coordinated multi-cell interference alignment (IA), [16] proposed a low-complexity distributed algorithm that was easy to implement on large networks. For the existing dense DSCs networks, it tends that the DSCs mostly depend on the central managers to make strategic decisions. Therefore, investigating the effective strategies or policies schemes for dense DSCs networks is still a challenging and open problem.

In addition, artificial intelligence (AI), such as the machine learning and deep learning, has been extensive used in wireless communications [17]–[23]. These works have used deep learning to realize automatic modulation recognition, hybrid precoding in millimeter-wave massive MIMO, and predication of channel state information, etc. However, it is still difficult to deal with the issue brought from the multi-agents, especially in ultra-dense unmanned networks. In these cases, game theory has been proven to be an efficient tool to obtain the effective distributed strategies as well as optimal control policies [24], [25]. Since classical games need to model the interactions of the considered agent with each other agents, the analyses of those systems with a large amount of agents is tedious and hard to tackle. Therefore, when faced with a network of densely deployed DSCs, classical game theory is too difficult to solve the placement control problem because of the huge number of players. For this condition, the mean field game (MFG) [27], [28] become an ideal tool for this sort of dense networks because MFG can be used to model the interactions between a subjective player and the average effect of other players [4], [29]–[35], [37]–[42]. For instance, in [30], the authors have designed a theoretical framework for the MFG by using mean field approximation (MFA) to deal with the interference among large number of players. The authors of [31] have obtained the distributed power control strategy for the dense deployment small BSs through the MFG method. The authors of [4] have investigated mean field models for cognitive radio networks with energy and mobility constraints. The authors in [37] have considered the last-level cache sharing problems in large-scale cloud networks, where they provided the closed-form expression of the optimal pricing that gives an efficient resource-sharing policy. Here, the behavior of mass can be expressed as the mean field. The Hamilton-Jacobi-Bellman (HJB) equation in MFGs is formulated to model the individual player's interactions with the mean field. While the evolution of the mean field in terms of the actions of players is described by a Fokker-Planck-Kolmogorov (FPK) equation. Therefore, the solution of MFG can be obtained by solving these two coupled equations. To use the MFGs, one of the advantages is that the solutions of the MFGs are able to be obtained separately if some of the boundary conditions are given. And in the MFGs, all the behaviors of players can be traced by one policy. What's more, MFGs are able to consider the stochastic of system. All the properties mentioned above make MFGs suitable for solving the DSCs placement control problem for dense DSCs downlink networks.

In this paper, we consider an interference management problem with the dense DSCs downlink networks, which is modeled as a MFG. We regard all DSCs which are utilizing the same downlink channel as the players of the MFG, that each DSC compete with all the others. As the deployment of DSCs is a critical part in air-to-ground downlinks, each DSC can improve its downlink's performance by adjusting its position. Thus, we have to considered the competition resulting from the cross-tier interference, i.e., while each

DSC flies to more suitable position to improve the quality of its own downlink, it probably cause higher interference to all other users served by other DSCs. We assume that there are no distinct horizontal displacements for all DSCs in the game during the control period due to the employed captive UAVs. Hence, DSCs can control their vertical velocity to decrease the cost during the preset control interval. The cost function of a DSC is derived by using the MFA approach, which contains the signal-to-interference-plus-noise ratio (SINR) at its user and the cost for rising and falling. Moreover, we develop a finite difference method to derive the HJB and FPK equations for the formulated MFG framework. There are two main advantages of the proposed algorithm. One is that it don't need real-time information from others, which can be performed offline. Each DSC can attain an optimal velocity control strategy with the proposed algorithm, which is influenced by the initial altitude distribution among the DSCs. Then the velocity control strategy can be used to implement downlink communication within a pre-defined interval of time. The other is that it would minimize the cost during a certain period of time rather than making decisions only on the basis of cost at a moment. The contributions of this work can be summarized as follows:

- We consider the co-channel interference circumstance in downlink dense DSCs networks with AI-aided approach. In this network, each DSC's altitude is related to the distance and the elevation angle between it to user, which can directly effect the transmission quality of downlink signals. Therefore, each DSC can control its vertical flight velocity to achieve better quality of communication by considering the influence from others.
- The altitude adjusting problem is modeled as stochastic differential game. Meanwhile, the existence of the Nash Equilibrium (NE) is proved. Faced with the difficulties in solving the differential game when the number of DSCs becomes large, we formulate a MFG framework, which can simplify the mutual interference among a large number of DSCs. In this framework, the altitude of DSCs is regard as the state, and cost function is designed to be the combination of energy consumption and SINR.
- The interference caused by the mass of DSCs (the mean field term) is describe by the MFA method. Moreover, we derive the HJB and FPK equations with regard to the proposed MFG framework, where these two equations represent the forward and backward equations, respectively. In addition, we propose a finite difference method to solve these two differential equations to obtain the solution of MFG.
- Simulation results show the behavior characteristics of DSCs in different DSCs densities and environment scenarios. Meanwhile, it also validate the SINR performance of proposed algorithm. Particularly, this algorithm can perform offline, which is valuable in practice.

Compared to the existing works as [30]–[35], which concentrated on power control with linear remaining energy state into consideration, our paper proposes the state dynamics about altitudes control and applies MFG to obtain the optimal altitude control strategies for distributed DSCs. Then, compared to MFG in existing works [30]–[35], it is intractable for distributed players to obtain the instantaneous interactions from other players. We propose a MFA method for DSCs to obtain the interactions with the mean field. And compared to the existing works [30]–[35], which mostly used a finite difference method with Lax-Friedrichs scheme and Lagrange relaxation, we propose a finite difference method with upwind scheme, and updating the optimal control by applying the first order necessary condition on the Hamiltonian.

The rest of this paper is organized as follows. In Section II, we introduce the system model containing spectrum sharing dense DSCs and assumptions. In Section III, the altitude adjusting problem is converted into a differential game. Then the theoretical analysis on the optimal control of DSCs at the NE are given. We then extend the differential game to the MFG. Also two couple of different equations of proposed MFG framework is deduced in Section IV. In Section V, we adopt a finite differential method to obtain the solutions. Numerical results are analyzed in Section VI. Finally, conclusions are drawn in Section VII.

II. SYSTEM MODEL

In this paper, we investigate a large number of DSCs acting as the flight BSs providing the air-to-ground wireless communication in a given area, where we adopt the tethered buoyant platforms powered by ground vehicles with cables. Therefore, these DSCs have able to maintain flight for a long time, and we have the assumption that we do not consider the distinct horizontal displacements during transient operation interval. The horizontal spatial distributions of DSCs are decided by some central controllers considering some security issues and coverage, which are not considered within this paper.

The interference problem in downlink communications is considered in this paper. As the different channels having been used by the large number of DSCs in this dense networks, we consider the interference among the DSCs, which transmit on the same channel. Meanwhile, we have the assumption that DSCs apply orthogonal multiple access (OMA) technology, so that for a DSC there is only one user can be served by one channel. In this paper, the dense networks contain N DSC-to-user downlinks sharing the same channel, as shown in Fig. II. For example, there may be N users occupying the k-th sub-channel to communicate with their corresponding DSCs during the same period. We regard them as the agents of the game. In comparison to common BS, flexible DSCs can search the optimal altitude to realize the optimal communication with users. However, this behaviour probably has the increase interference to other DSCs. Therefore, other DSCs wil intuitively change their altitude. To maintain a stable downlink, for each DSC in the network, an altitude control policy is decided aiming at

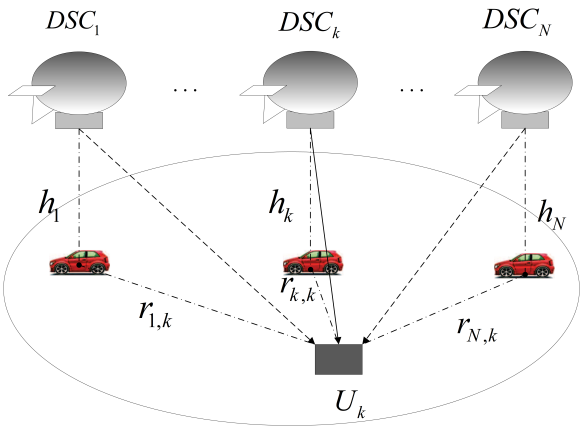


FIGURE 1. System model of N spectrum sharing DSCs.

minimizing the average cost of it during a preset period T. The control period T can be adjusted according to the actual conditions. Even though each DSC has its individual initial altitude at the beginning, the same flight space are deployed on all DSCs, which means all DSCs may have the arbitrary altitude in $[h_{\min}, h_{\max}]$. Different initial altitudes of DSCs is considered in this case.

Different channel models will have a great impact on the interference mitigation techniques. In our considered dense DSCs networks, the distance among DSCs are long enough that the interference from one DSC to another DSC is nearly negligible, which is in the line with the MFG set. As the distances are long, the major impact factor of the channel gain is the None Line of Sight (NLoS) link. Specifically, there are three kinds of signals can be received by users including Line of Sight (LoS) links, NLoS links and multipath signals [9]. Thus, we describe these three signals by using different probabilities of occurrencec as proposed in [42] and [43]. We assume that the received signal is categorized in only one of those groups. Each group has its own probability of occurrence, which is a function of environment, user density, altitude of buildings and the elevation angle. Meanwhile, the probability of occurrence of multipath fading, which is cased by the multipath signals can be ignored because its lower than other two links.Therefore, the typical method is used to model the path loss of the LoS and NLoS, where the received power at k-th user from i-th DSC is given by

$$P_{r,i,k} = \begin{cases} P_i \left|X_{i,k}\right|^{-\alpha} & \text{LoS link}, \\ \eta P_i \left|X_{i,k}\right|^{-\alpha} & \text{NLoS link}, \end{cases} \tag{1}$$

where P_i is the transmit power of i-th DSC. In our considered DSCs network, we assume that the P_i is a constant because the same transmit power is used by these N DSCs. $\left|X_{i,k}\right|$ represents the distance from i-th DSC to k-th user. α and η are the path loss exponent of the air-to-ground links and the attenuation coefficient over the NLoS links, respectively. Then the probability of the LoS links is expressed as follows,

$$P_{\text{LoS}} = \left(1 + c\exp(-b[\theta - c])\right)^{-1}, \tag{2}$$

where c and b are parameter related to the environment such as suburb, urban or dense urban. The θ represents the elevation angle, which can be expressed as,

$$\theta = \frac{180}{\pi} \times \sin^{-1}\left(\frac{h}{|X_i|}\right), \tag{3}$$

where $|X_i| = \sqrt{h_i(t)^2 + d_{i,k}^2}$ and $d_{i,k}$ is the horizontal distance between DSC i and the user it serves. Therefore, the received power of k-th user transmitted from i-th DSC is defined as

$$P_{r,i,k} = P_{\text{LoS}} P_i(t) \left|X_{i,k}\right|^{-\alpha} + P_{\text{NLoS}} \eta P_i(t) \left|X_{i,k}\right|^{-\alpha}. \tag{4}$$

Hence, the achieved SINR with the considered interference channel at k-th user is given by

$$\lambda_k(t) = \frac{P_{r,i,k}}{I_{u,k}(t) + N_0}, \tag{5}$$

where N_0 denotes the noise power. $I_{u,k}(t) = \sum_{j \in \mathbf{N}, j \neq i} P_{r,j,k}$ represents the influence from other DSCs. Then, considering the $I_{u,k}(t)$, i-th DSC determine its optimal velocity $v_i^*(t)$. Accordingly, a differential game is suitable to model the velocity control problem, where the detail will be introduced as follows.

III. DIFFERENTIAL GAME MODEL FOR ALTITUDE CONTROL OF A COMMON DSC

In this section, the differential game model is formulated to solve the velocity control problem in the dense DSC networks, where each DSC decides its flight policy during the control interval. In a differential game, players should determine determine the optimal velocity to minimize the proposed cost function.

A. STATE, ACTION AND STATE EQUATION OF THE COMMON DSC

In this game, we define that the $h_i(t)$, which represents the altitude of i-th DSC at time t, is regard as the state. For the convenience of description, in this paper, the i-th DSC is used to describe the state evolution, the cost function and so on named common DSC because it is any and replaceable of the mass. Here, as the consideration of flight safety and DSC's restrictions, the flight altitude of each DSC $h(t)$ belongs to $[h_{\min}, h_{\max}]$. Therefore, the evolution of state is defined as follows:

Definition 1: The dynamic of state $h_i(t) \in [h_{\min}, h_{\max}]$ can be expressed by a differential equation as,

$$\mathrm{d}h_i(t) = v_i(t)\mathrm{d}t + \sigma_t \mathrm{d}W_i(t), \quad 0 \leq t \leq T, \tag{6}$$

where the velocity $v_k(t)$ represents the drift function of the different equation (6). The $W_k(t)$ is the Wiener process with the volatility σ_t. $\sigma_t \mathrm{d}W_i(t)$ stands for the disturbance, and its detail can be found in [32].

The control policy v_i can be thought of the mapping from a state to an action. Then, each DSC seeks for the optimal control policy v_i^* to minimize its average cost a_i^* during the

time interval $[0, T]$. Thus, the long time average cost of common DSC is given by,

$$a_i^* = \arg \min_{v_i} \mathbb{E}\left[\int_0^T c_i(t)\mathrm{d}t + c_i(T)\right], \quad (7)$$

where $c_i(T)$ is the terminal cost. $\mathbb{E}[\cdot]$ denotes the expectation operator.

B. COST FUNCTION OF THE COMMON DSC

In this paper, the goal of each agent (DSC) is the improvement of transmission performance, which is determined by the SINR. As we know, the SINR directly reflects the rate of which information is successfully transmitted, while the larger SINR induces the DSC to decrease its altitude, which signifies the larger interference to other DSCs-to-users links. This issue can be restricted by the energy constraint, which not only contains the communication energy, but the energy used for adjusting the altitude of DSCs. To simplify the model, the flight energy consumption is proportional to the quadratic form of the flight velocity $v_i(t)$. Therefore, the cost function of common DSC i constituted by two parts including SINR and flight energy consumption is designed as follows:

$$c_i(t) = -\omega_1 \frac{P_{r,i,k}}{I_{u,k}(t) + N_0} + \omega_2 {v_i}^2(t), \quad (8)$$

where ω_1, ω_2 are the weight of SINR and flight energy consumption.

C. DIFFERENTIAL GAME FORMULATION

On the basis of Bellman's principle of optimality [44], one attribute of optimal control strategy is that regardless of initial state and initial decision, the residual decisions constitute the optimal strategy for the state of the first decision necessarily. Therefore, we define the optimal velocity control strategy according to the value function, which can be expressed as

$$u_i(t) = \min_{v_i} \mathbb{E}\left[\int_t^T c_i(\tau)\mathrm{d}\tau + c_i(T)\right], \quad t \in [0, T], \quad (9)$$

where the expectation operator present because the independent Brownian motion (Wiener process) in (6). The optimal accelerated velocity control policy for i-th DSC satisfies for any $t \in [0, T]$. Therefore, we can give the definition as follows:

Definition 2: The velocity profile v_i^* $(0 \to T)$ is the optimal altitude control policy of i-th DSC during the period $t \in [0, T]$, the value function can be defined as

$$\mathbb{E}\left[\int_t^T c_i(v_i^*(\tau))\mathrm{d}\tau + c_i(T)\right] = u_i(t). \quad (10)$$

With the state equation and the value function, we can form the HJB equation, a partial differential equations (PDEs) [45], for this optimal control problem of a generic DSC as

$$\frac{\partial u_i(t)}{\partial t} + \min_{v_i(t)}(c_i(v_i(t)) - v_i(t)\frac{\partial u_i(t)}{\partial h}) + \frac{\sigma^2}{2}\frac{\partial^2 u_i(t)}{\partial^2 h} = 0, \quad (11)$$

where

$$H(h_i(t), \frac{\partial u_i(t)}{\partial h}) = \min_{v_i(t)}\left(c_i(t) - v_i(t)\frac{\partial u_i(t)}{\partial h}\right), \quad (12)$$

which is named the Hamiltonian. As the disturbance term $W_i(t)$ follow the rules of Ito calculus. Hence, the expectation operator disappeared after the deduce for the HJB function.

Then, to achieve the NE of this game, we consider that each DSC can not decrease the cost via unilaterally deviating from current velocity control strategies. Thus, with the HJB equation of this differential game, the NE can be obtained depending the definition 3.

Definition 3: The velocity profile $v^* = [v_1^*, v_2^*, \ldots v_i^*, \ldots v_N^*]$ during the time interval $t \in [0, T]$ represents the NE of the proposed differential game when

$$v_i^* = \arg \min_{v_i} \mathbb{E}\left[\int_0^T c_i(v_i(t), v_{others}^*)\mathrm{d}t + c_i(T)\right]. \quad (13)$$

Here, v_{others}^* denotes the velocity vector of other DSCs. Noticing that states of all DSCs evolve with time based on 6.

Theorem 1: At least one NE exists in the differential game.

Proof 1: Proof of Theorem 1 is given in Appendix A.

Then, in this proposed game, the NE can be obtained by solving at least N PDEs as shown in 11 because there are N DSCs existing in this model [46]. It is unrealistic to solve such large number of PDEs because the complexity grows exponentially named the curse of dimensionality as the number of PDEs is more than three, which means the differential game used in dense networks can only obtain the theoretic solution. Therefore, the MFG is used in this model on account of this game converts the large number of individual behaviors into the mass behavior, which can be simplified by two coupling equations. The details can be found in the next section.

IV. MFG APPROACH FOR DENSE DSCs

A. MEAN FIELD AND MFA

As a special kind of differential game, MFG has a main feature named the similarity hypothesis. Specifically, all agents are similar as well as interchangeable, and each of them follows the same strategy [33]. In other words, all agents are different merely in states, which leads to the affect of the individual almost been neglected as the number of agents is large enough (even tending to infinite). Therefore, in this paper, when the number of DSC is large enough, the MFG approach is considered to solve the velocity control problem. Then, the new velocity control MFG for DSC downlink networks can be formulated depending on the following properties each DSC has:

- Rationality: Each DSC can make the rational velocity control decision by itself to minimize its long time average cost over the presetting control period.
- The existence of the DSCs' continuum: It can ensure the continuity of the mean field. In our considered scenario, the presence of a huge amount of DSCs guarantees the existence of the DSCs continuum.

- Interchangeability of the states of all DSCs: It should be insured that permutation of the states would not influence the outcome. As we derive the cost function by a MFA approach, the states among DSCs can maintain the interchangeability.
- Interaction between a generic player with mean field: It is the core idea of MFG. Each DSC only consider the interaction with the mean field rather than that with all the rest.

For our considered ultra-dense DSCs network, the first condition can be meet [30]–[34]. And in our system, we assume that all DSCs have the same spaces of the state and the control, which is rational in practical as the DSCs should obey the same flight rules. Hence their vector of states can be exchangeable.

The mass is described by the mean field term $m_t(h)$, which is composed by the states distribution of all DSCs defined as follows:

$$m_t(h) = \lim_{N\to\infty} \frac{1}{N} \sum_{\forall i \in N} \mathbb{1}_{\{h_i(t)=h\}}, \tag{14}$$

where $\mathbb{1}_{\{.\}}$ is the indicator function which will return 1 when the condition is true. Otherwise, it will return 0. At each given time, the mean field term $m_t(h)$ is the states probability distribution of all agents. Here, the lager the number of agents N is, the smoother the mean field term $m_t(h)$ becomes, until $N \to \infty$ causes $m_t(h)$ turning into a continuous distribution function.

Similarly, in MFG framework, we adopt the common DSC (agent) i to describe the velocity control problem, and the altitude $h_i(t)$ still is the state of DSC i. The different is that each DSC adjust its flight altitude influenced by the mass rather than each individual to achieve its goal. Therefore, all DSC make decision with the same grope of constraints and equations, which means the complicated interaction can be simplified as the interacting between the common DSC and the mass. Specifically, the process of flight control policy acquiring is reduced to two couple of PDEs, describing the dynamic decision and the mass evolution. In addition, the interference of the mass is simplified as the function of the mean field term, which cases the SINR in cost is the function of $m_t(h)$. In other words, the interference $I_{u,k}(t)$ of other DSCs can convert into the average influence $\bar{I}_{u,k}(t)$ from the mass. Besides, the average distance between other DSCs and the user served by DSC i is necessary. Thus, this distance $\bar{d}_{j,k}$, $j \neq i$ can be obtained by using the MFA [30], where the process is shown as follows:

$$\begin{aligned}\bar{I}_{u,k}(t) &= \sum_{j=1,j\neq i}^{N} P_{r,j,k} \\ &= \sum_{j=1,j\neq i}^{N} P_j(t) f\left(h_j(t), d_{j,k}\right) \\ &\approx N P_{test}(t) \bar{f}\left(h_j(t), d_{j,k}\right). \end{aligned} \tag{15}$$

In equation 15, the $f(.) = P_{\mathrm{LoS}}\left|X_{i,k}\right|^{-\alpha} + P_{\mathrm{NLoS}}\eta\left|X_{i,k}\right|^{-\alpha}$ represents the channel conditions, which is the function of $h_j(t)$, $j \in N, j \neq i$ and $\bar{d}_{j,k}$. The $P_{test}(t)$ is a test transmit power, a preset constant, which is same for each DSC. Meanwhile, we assume that each DSC has the same initial state, i.e. $h_i(0) = h_j(0)$. Then, the previous part of the cost function $\hat{c}_i^1(t)$ for a generic DSC in this MFG can be expressed as

$$\begin{aligned}\hat{c}_i^1(t) &= \frac{P_{r,i,k}}{\bar{I}_{u,k}(t) + N_0} \\ &= \frac{P_i(t) f\left(h_i(t), d_{i,k}\right)}{N P_{test}(t) \int\limits_h f\left(h, \bar{d}_{j,k}\right) m_t(h)\mathrm{d}h + N_0}, \end{aligned} \tag{16}$$

where

$$f(h, \bar{d}_{j,k}) = P_{\mathrm{LoS}}\left|X_{j,k}\right|^{-\alpha}(1-\eta) + \eta\left|X_{j,k}\right|^{-\alpha}. \tag{17}$$

Accordingly, the mean distance $\bar{d}_{j,k}$ can be estimated by the above method in practice. Noticing that at the numerator, $X_{j,k} = \sqrt{h_j(t)^2 + d_{j,k}^2}$ and

$$P_{\mathrm{LoS}} = \left(1 + c\exp\left(-b\left[\sin^{-1}\left(\frac{h_j(t)}{d_{j,k}}\right) - c\right]\right)\right)^{-1}, \tag{18}$$

are functions of current altitude. According the altitude of the common DSC i and the backward evolving of HJB equation, the control of each DSC is only related to its state. Thus, the cost expressed in equation (8) can be redefined as

$$\hat{c}_i(t,h) = \frac{-\omega_1 P_i(t) f\left(h_i(t), d_{i,k}\right)}{N P_{test}(t) \int\limits_h f\left(h, \bar{d}_{j,k}\right) m_t(h)\mathrm{d}h + N_0} + \omega_2 v_i^2(t,h). \tag{19}$$

The same like the cost function $c_k(t)$ proposed in section III, the first portion of $\hat{c}_k(t,h)$ represent the SINR on the user served by the common DSC, and the latter portion denotes the cost caused by rising and falling.

B. FORWARD-BACKWARD EQUATIONS OF MFG

As shown in equation (19), the cost only is the function of two parameters, the mean field term and the control, which means, in MFG frameworks, all agents (DSCs) have the similar control problem and different initial states. Thus, the differential game's HJB in (11) is modified as follows:

$$\begin{aligned}&\frac{\partial u_t(h)}{\partial t} + \frac{\sigma^2}{2}\frac{\partial^2 u_t(h)}{\partial^2 h} \\ &\quad + \min_{v_t(h)}\left(c(v_t(h), m_t(h)) - v_t(h)\cdot\frac{\partial u_t(h)}{\partial h}\right) = 0. \end{aligned} \tag{20}$$

Then, we formulate the forward (FPK) equation to describe the evolving of the mass. The following theorem gives the derivation of the FPK equation.

Theorem 2: The FPK equation in this MFG framework is given by

$$\frac{\partial m_t(h)}{\partial t} + \frac{\partial}{\partial h}\left(m_t(h) v_t(h)\right) + \frac{\sigma^2}{2}\frac{\partial^2 m_t(h)}{\partial^2 h} = 0. \tag{21}$$

Proof 2: Theorem 2 is proved in Appendix B.

Since the HJB equation in (20) iterating backward over time determines the optimal path of the derivation for each agent. Meanwhile, the FPK equation in (21) iterating forward over time determines the evolving of the mean field term. Therefore, these two coupled PEDs are solved necessarily to achieve the NE of this MFG. During the period $t \in [0, T]$, the value function $u_t(h)$ can be expressed as

$$u_t(h) = \mathbb{E}\int_t^T \left(-\omega_1 \frac{P_{\mathrm{LoS}} P_i \left|X_{i,k}\right|^{-\alpha} + P_{\mathrm{NLoS}} \eta P_i \left|X_{i,k}\right|^{-\alpha}}{N P_{test} \int\limits_H f\left(m(\tau, h), \bar{d}_{j,k}\right) + N_0} + \omega_2 v_i^2(\tau) \right) \mathrm{d}\tau. \quad (22)$$

Thus, the solution of MFG can be regarded as the set of $[m, u]$.

Lemma 1: The average velocity $\bar{v}$ of DSC i is the derivative of average height versus time, which is derived by

$$\bar{v} = \frac{\mathrm{d}}{\mathrm{d}t}\int_{h_{\min}}^{h_{\max}} m_t(h)\mathrm{d}h. \quad (23)$$

Proof 3: Lemma 1 is proved in Appendix C.

Lemma 2: If (U_1, m_1) and (U_2, m_2) are two solutions of the forward-backward equations above and $m_1 = m_2$, then $U_1 = U_2$.

Proof 4: With two value functions u_1 and u_2, we have two FPK equations as

$$\frac{\partial m_1}{\partial t} + \frac{\partial}{\partial h}(m_1 v_1) + \frac{\sigma^2}{2}\frac{\partial^2 m_1}{\partial^2 h} = 0, \quad (24)$$

$$\frac{\partial m_2}{\partial t} + \frac{\partial}{\partial h}(m_2 v_2) + \frac{\sigma^2}{2}\frac{\partial^2 m_2}{\partial^2 h} = 0. \quad (25)$$

Since $m_1 = m_2 = m$, we subtract (24) from (25) to obtain $\frac{\partial}{\partial h}((v_1 - v_2)\, m) = 0$. From Lemma 1, if $m_1 = m_2$ is set up, then $\bar{v}_1$ equals $\bar{v}_2$. Since $\bar{v} = \int_{h_{\min}}^{h_{\max}} v m \mathrm{d}h$, we have

$$\begin{aligned} \int_{h_{\min}}^{h_{\max}} v_1 m dh &= \int_{h_{\min}}^{h_{\max}} v_2 m dh \\ &\Rightarrow \int_{h_{\min}}^{h_{\max}} (v_1 - v_2)\, m dh \\ &= 0, \quad \forall t. \end{aligned} \quad (26)$$

This means $v_1 = v_2$. Meanwhile, the $u_t(h)$ is related to $\bar{v}$ and v. Hence, the equivalences in $\bar{v}$ and v lead to $u_1 = u_2$. Therefore, the lemma 2 verifies that the control of each agent only depends on the mean field term. In other worlds, the similar mean field term accompanies the similar behavior.

V. SOLUTION OF THE MFG BASED ON THE FINITE DIFFERENCE METHOD

The goal of this MFG framework is to acquire the optimal strategy for each agent. Two couple of equations, HJB and FPK, can obtain the solution of the MFG mentioned above. Thus, in this section, we adopt a finite different method, which is an effective tool to solve the PDEs. By using this method, the derivatives in PDEs are approximated as the form of finite differences, where the HJB calculates the optimal control strategy backwardly with the minimum cost, and the FPK describes the behavior of the mass by using the strategy obtained from HJB solving. Finally, the iteration in these two equations can derive the solution (NE) of the MFG.

Algorithm 1 NE of the MFG

Initialize:
$T \times (1 + h_{max} - h_{min})$: Initial state space.
M_0^i: Initialization of the mean field.
V_t^h: Initialization of the control space randomly.
$U_{t_{max}}^i = 0$
Repeat:
for $t = 1 : t_{max}$ & $h \in [h_{min}, h_{max}]$ **do**
  M can be obtained by solving 30.
end for
for $t = t_{max} : 2$ & $h \in [h_{min}, h_{max}]$ **do**
  Iterating U by solving 32.
end for
V^* can be obtained by solving 31.
$V = \eta_1 V + \eta_2 V^*$, $\eta_1 + \eta_1 = 1$.
End

In this method, the state and solution are preprocessed by discretization of their space. Specifically, the state space can be discretized as the $[h_{min}, \ldots, h, h + dh, \ldots, h_{max}]$, where dh represents the step size, and h belongs to the $[h_{min}, h_{max}]$. Meanwhile, the time period $[0, T]$ of the control is discretized as $[0, \ldots t, t + dt, \ldots, T]$. Similarly, dt is the time interval. Therefore, the time and state in (20) and (21) can be defined in the space of $T \times (1 + h_{max} - h_{min})$. Then, according to the Upwind scheme, which is a kind of…., the differential in (20) and (21), which is the continuous can be formulated as follows:

$$\frac{\partial u_t^h}{\partial t} = \frac{u_{t+1}^h - u_t^h}{dt}, \quad (27)$$

$$\frac{\partial u_t^h}{\partial h} = \frac{u_t^h - u_t^{h-1}}{dh}, \quad (28)$$

$$\frac{\partial^2 u_t^h}{\partial^2 h} = \frac{u_t^{h+1} - 2u_t^h + u_t^{h-1}}{(dh)^2}. \quad (29)$$

Thus, the FPK in equation (21) is described as

$$\begin{aligned} M_{t+1}^h = M_t^h &+ \frac{dt}{dh}\left[M_t^{h-1} V_t^{h-1} - M_t^h V_t^h\right] \\ &+ \frac{\sigma_t^2 dt}{2(dh)^2}\left[2M_t^h - M_t^{h+1} - M_t^{h-1}\right]. \end{aligned} \quad (30)$$

Since the optimal value function can be obtained by solving the HJB, we can derive the optimal control as

$$v_t^{h*} = \frac{\partial u_t^h}{2\omega_2 \partial h}. \quad (31)$$

TABLE 1. Parameters in simulations.

Descriptions	Parameters	Values
the time interval	$[0, T]$	[0s,10s]
the altitude interval	$[h_{\min}, h_{\max}]$	[1000m,1100m]
the maximal velocity	$\lvert v_{\max}\rvert$	10m/s
the transmit power	P	42dBm
the parameters in air-to-ground channel model	α, η	0.2,0.05
the considered environment parameters	C, B	10,0.03

Substituting optimal control v^* into the HJB equation, and after some deductions, the expression of the value equation can be expressed as

$$U_{t-1}^h = U_t^h + \frac{\frac{dh}{dt}\left(U_{t+1}^h - U_t^h\right) + dhC\left(V_t^{h*}, M_t^h\right)}{V_t^{h*}} + \frac{\sigma_t^2 dt}{2(dh)^2}\left[2U_t^h - U_t^{h+1} - U_t^{h-1}\right]. \quad (32)$$

Here, we consider some boundary conditions during the iterations. For instance, the value U_t^{h+1} is inexistence when the h equals to h_{max}. Thus, the $U_t^{h_{max}}$ is used instead of the $U_t^{h_{max}+1}$. And considering the altitude and velocity constrains, the velocity of each DSC at the maximum and minimum altitude can not be positive and negative, respectively.

The convergent solution can be obtained by solving the (30), (31) and (32), iteratively, which is detailed in Algorithm 1.For our proposed algorithm, the space size $s = X \times Y$, where X, Y are the space size of time interval and altitude interval, respectively. In each iteration, we go through all the space two times, and there are finite computing times in each traversal. So the computation complexity of every iteration is $O(s)$, which means the iteration is also finite. Therefore the computational complexity is acceptable in practical application. In Algorithm 1, all DSCs have the distributions of initial mean field. Meanwhile, the MFA is used to obtain the $\bar{r}_{j,k}$ at the initial time. The iteration ends when the iterative times exceed the threshold.

VI. NUMERICAL RESULTS AND ANALYZATION

In this section, we show the distributions of the mean field as the time and state vary in various environments. The parameters in simulation are shown in Table 1.

The simulation gives the mean field distributions in Fig. 2. At the initial time, we can see the same number of DSCs at each altitude, because the initial altitude is an uniform distribution. Then, as time varying, the DSCs in higher altitude decrease until disappear at $t = 10s$. Most of them choose the lower altitude because the lower altitude causes the weaker path loss, which has the similar trend mentioned above. However, there are still a part of DSCs hovering in higher altitude because the effectiveness of control strategy.

For the convenience of analysis, we show the dynamic of the mean field at some specific altitudes as time varies, which is shown in Fig. 3. The DSCs at 1100m reduce sharply at

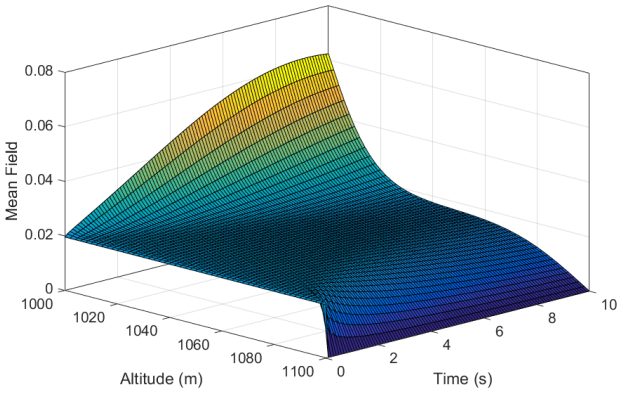


FIGURE 2. The distribution of mean field.

the initial stage. Meanwhile, the distribution of the DSCs at 1080m also decreases after a stable stage because the DSCs at higher altitude become scarce. Conversely, the DSCs' distributions at 1020m and 1000m increase in varying degrees. Clearly, during the control period, most DSCs stay in lower altitude instead of the higher environment because the better communication conditions establish as the closer distance between the DSCs and users. Besides, these curves flatten out as the time varies. The reason of this is that the more DSCs assemble at the same altitude, the lager interference received by users. Therefore, the DSCs in this condition should to increase their flight velocities to achieve their goals (the SINR) reflected in the cost function, which corresponds to more flight energy consumptions.

Fig. 4 proves the convergence of the mean field represented by four specified altitudes in Fig. 3. All mean field are at time T. Simulation results illustrate that the equilibrium convergence with a low iterations, which is acceptable in practical application.

Fig. 5 shows the velocity control strategies of all DSCs. The dynamical system of each DSC can be adjusted to catch the velocity calculated at the beginning of every distributed control time interval based on its current altitude at each time instant. As shown in Fig. 5, the initial velocities at different altitude assign randomly. The all of them decrease over time until $v = 0$, which donates the DSCs hovering in their optimal positions. The reason of this phenomenon is that the increase interference at lower altitude causes the slowdown in a trend of continued decline.

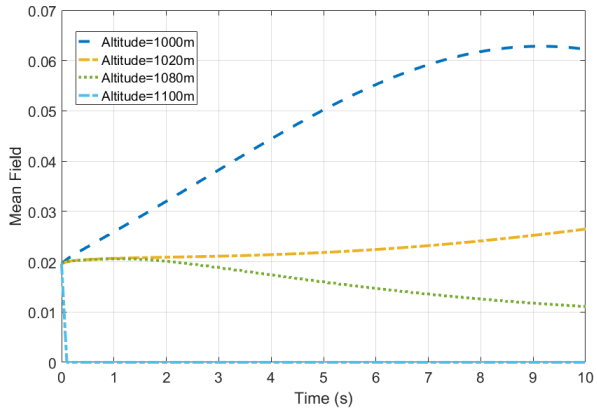


FIGURE 3. Dynamic of the mean field at different altitudes.

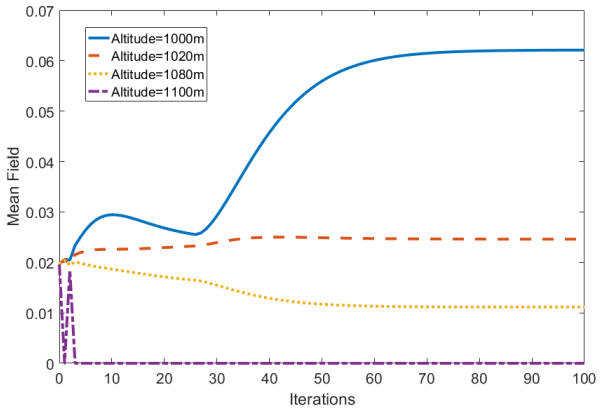


FIGURE 4. Convergence of the value iteration.

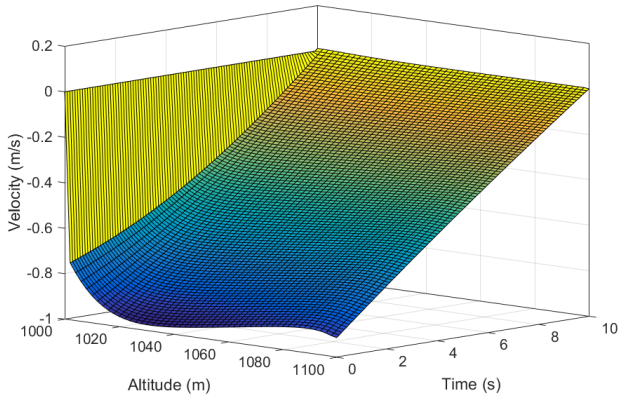


FIGURE 5. The distribution of optimal control (velocity) strategies.

As the air-to-ground channels are affected by variational factors, which are decided by different environment and the number of the DSCs, we are interested in the distribution of DSCs in those scenarios. Fig. 6 shows the mean fields with different densities of DSCs. We can see that with the increasing density of DSCs, the displacement distances of them decrease, which lead to most DSCs converge to the middle altitude of the range.

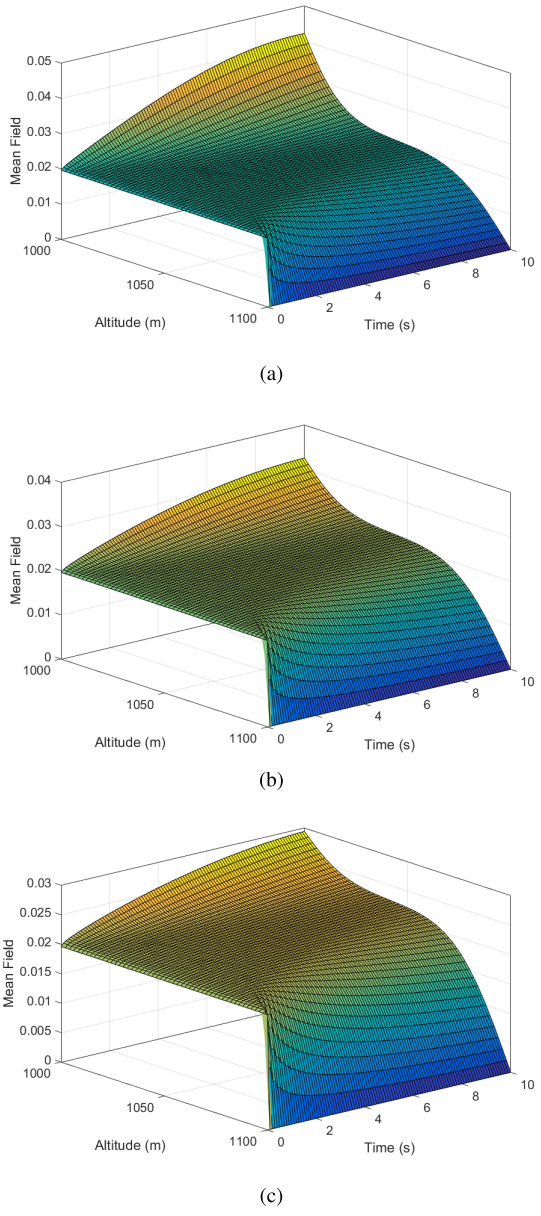


FIGURE 6. The distribution of mean field with (*a*) $N = 100$; (*b*) $N = 200$; (*c*) $N = 300$, respectively.

And we illustrate some specific altitude distribution with the DSCs density in Fig. 7. It also demonstrates the results that, the distributions at different altitudes decrease with the number of the players in this game. The reason is that the increasing density of DSCs causes larger interference to each DSC. Compared with improving the channel by dropping, DSCs are inclined to save power used in dynamic.

In order to visualize the effect of various urban environment on the mean field distributions of DSCs, we plot the Fig. 8 with the following parameters $(B,\ C)$ pairs $(0.06, 11.25)$, $(0.05, 10.39)$, $(0.04, 8.96)$ and $(0.03, 7.37)$ corresponding to various density of the urban environment [10]. As discussed above, we focus on the distributions when the altitude is lowest, which can demonstrate the

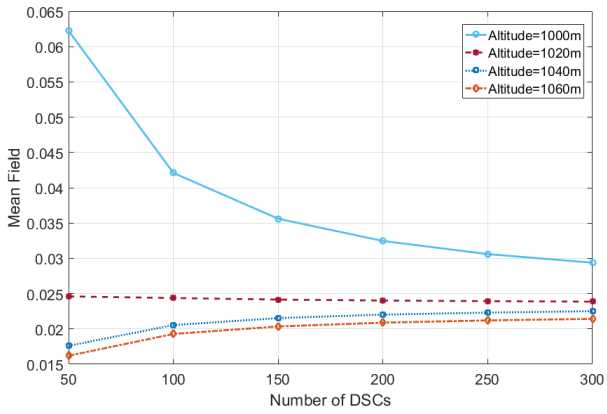


FIGURE 7. Some specific altitudes distributions with the DSCs density.

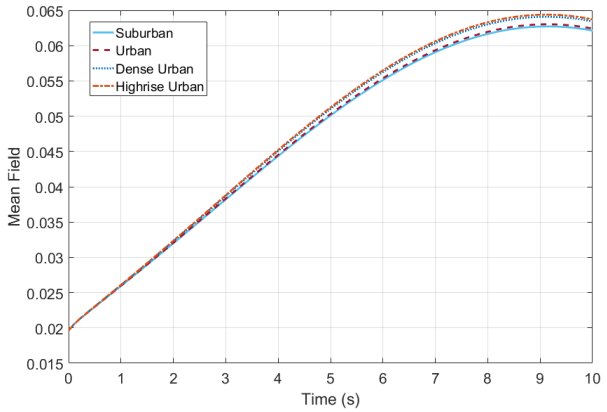


FIGURE 8. The mean field distributions when Altitude=1000m with different urban environment.

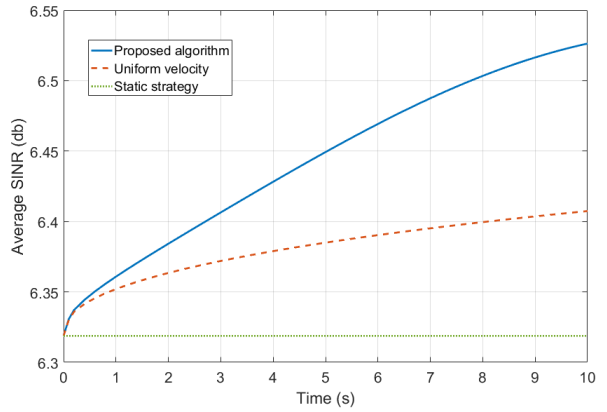


FIGURE 9. Average SINR of all users over time.

downtrend of DSCs. It can be seen that more DSCs converge to the lower altitude when the urban environment gets complex.

Finally, we discuss the performence of the algorithm, the spectral efficiency, which is affected by the SINR. Therefore, we show the average SINR of all users as the time varies in Fig. 9. Meanwhile, for comparison, the uniform decline scheme and hovering scheme are considered as the baseline shown as the red and green curve, respectively. Clearly, the average SINR of proposed method has the better average SINR than these two scheme, which can prove the effectiveness of algorithm.

VII. CONCLUSION

In this paper, we investigate an AI-aided interference management scheme by finding the optimal altitude strategy of the DSC in ultra-dense DSCs downlink networks. In this network, there is a large number of cross-tier interference because the high-density deployed DSCs share same channel. Each DSC reduces interference by adjusting its altitude to improve communication quality, the average SINR and the spectral efficiency. To obtain the optimal altitude control policy, firstly, we model the altitude control problem as a MFG theoretical framework, which is a method to assist AI for making decisions. In this framework, the MFA method is used to derive the large number of interference caused by other DSCs. Meanwhile, the finite difference algorithm based on the upwind scheme is proposed to solve the HJB and FPK equations aiming to obtain the solution (NE) of MFG. Simulation results show that the proposed algorithm can obtain better SINR by learning the optimal altitude policy comparing with the benchmark algorithm. Finally, we believe that there is a optimal joint strategy of power and velocity for those DSCs, but is intractable based on the existing development of MFG, which we will study in future.

APPENDIX

A. PROOF OF THEOREM 1

Since the HJB equation gives the optimal velocity choices for the DSCs in the time space, which cause the evolution of the DSCs' altitudes. In this MFG, the number of the DSCs is very large. Thus, we consider the normalized density function, the $m_t(h)$ of the DSCs. With given initial distribution $m_0(h)$, a FPK equation can be deduced to evolve the mean field forward in time.

In this game, we have the continuum limit $N \to \infty$. To describe the evolution of the continuum, we introduce a smooth and compactly supported function, $\phi(h)$. Therefore, $\int_{h_{\min}}^{h_{\max}} m_t(h)\phi(h)dh$ can be approximated as the continuum limit of the $\frac{1}{N}\sum_{i=1}^{N}\phi(h_i)$, i.e.,

$$\int_{h_{\min}}^{h_{\max}} m_t(h)\phi(h)\mathrm{d}h \approx \frac{1}{N}\sum_{i=1}^{N}\phi(h_i). \tag{33}$$

Our goal is to find how the mean field goes in time, differentiate both sides with time we have

$$\int_{h_{\min}}^{h_{\max}} \frac{\partial m_t(h)}{\partial t}\phi(h)\mathrm{d}h \approx \frac{1}{N}\sum_{i=1}^{N} v(h_i)\frac{\partial \phi(h_i)}{\partial h}. \tag{34}$$

In the continuum limit $N \to \infty$, we have

$$\int_{h_{\min}}^{h_{\max}} \frac{\partial m_t(h)}{\partial t}\phi(h)\mathrm{d}h = \int_{h_{\min}}^{h_{\max}} v(h_i)m_t(h)\frac{\partial \phi(h_i)}{\partial h}\mathrm{d}h. \tag{35}$$

Integration by parts leads to

$$\int_{h_{\min}}^{h_{\max}} \frac{\partial m_t(h)}{\partial t} + \mathrm{div}(v_t(h)m_t(h))\phi(h)\mathrm{d}h = 0. \quad (36)$$

This is valid for every test function $\phi(h)$, so we have derived the advection equation

$$\frac{\partial m_t(h)}{\partial t} + \frac{\partial}{\partial h}\left(m_t(h)v_t(h)\right) = 0. \quad (37)$$

B. PROOF OF THEOREM 2

To assure the existing of the NE under this differential game, we should prove the solution of HJB equation existing. Based on [38], the solution of HJB exists if the Hamiltonian is proven to be smooth. As the Hamiltonian we derived, i.e.,

$$\begin{aligned} &H(h_i(t), \frac{\partial u_i(t)}{\partial h}) \\ &= \min_{v_i(t)} \left(c_i(t) - v_i(t)\frac{\partial u_i(t)}{\partial h} \right) \\ &= \min_{v_i(t)} \left(-\omega_1 \frac{p_{r,i,k}}{I_{u,k}(t) + N_0} + \omega_2 v_i^2(t) - v_i(t)\frac{\partial u_i(t)}{\partial h} \right). \end{aligned} \quad (38)$$

It's evident that the derivatives of the Hamiltonian,

$$\frac{\partial H}{\partial v(t)} = 2\omega_2 v(t) - \frac{\partial u_i(t)}{\partial h}, \quad (39)$$

$$\frac{\partial^2 H}{\partial v(t)} = 2\omega_2, \quad (40)$$

$$\frac{\partial^3 H}{\partial v(t)} = 0. \quad (41)$$

Hence, the function is smooth because it has derivatives of all orders. Thus, at least one NE existing in this differential game can be proven.

C. PROOF OF LEMMA 1

We can derive the integral form of the state equation (6) as

$$\begin{aligned} h(t+dt) - h(t) &= \int_t^{t+dt} v_t(h)\mathrm{d}t + \int_t^{t+dt} \sigma \mathrm{d}W_t \\ &= v_{\tilde{t}}(h)dt + \sigma\left(W_{t+dt} - W_t\right), \end{aligned} \quad (42)$$

where $\tilde{t} \in (t, t+dt)$. It can be proved through the mean value theorem for integrals. Since (41) is true for all DSCs, taking expectation of this equality above for all DSCs and all current height of DSCs, we have

$$\mathbb{E}\left[h(t+dt)\right] - \mathbb{E}\left[h(t)\right] = \mathbb{E}\left[v(\tilde{t}, h)\right]dt + \sigma\mathbb{E}\left[W_{t+dt} - W_t\right], \quad (43)$$

$$\int_{h_{\min}}^{h_{\max}} hm(t+dt, h)\mathrm{d}h - \int_{h_{\min}}^{h_{\max}} hm_t(h)\mathrm{d}h = dt\mathbb{E}\left[v(\tilde{t}, h)\right], \quad (44)$$

where $\mathbb{E}\left[W_{t+dt} - W_t\right] = 0$ satisfies because the Wiener process $W_{t+dt} - W_t$ follow a normal distribution with mean zero.

We assume dt is very small then $\tilde{t} \to t$. Hence, we have

$$\frac{\mathrm{d}\int_{h_{\min}}^{h_{\max}} hm_t(h)\mathrm{d}h}{\mathrm{d}t} = \int_{h_{\min}}^{h_{\max}} v_t(h)m_t(h)\mathrm{d}h = \bar{v}(t). \quad (45)$$

The proof is completed.

ACKNOWLEDGMENT

This article was presented at the 2018 10th International Conference on Wireless Communications and Signal Processing (WCSP).

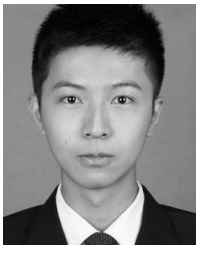
REFERENCES

[1] Z. Zhang, L. Li, W. Liang, X. Li, A. Gao, W. Chen, and Z. Han, "Downlink interference management in dense drone small cells networks using mean-field game theory," in *Proc. 10th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Hangzhou, China, Oct. 2018.
[2] A. Kaul, K. Obraczka, M. A. S. Santos, C. E. Rothenberg, and T. Turletti, "Dynamically distributed network control for message dissemination in ITS," in *Proc. IEEE/ACM 21st Int. Symp. Distrib. Simul. Real Time Appl. (DS-RT)*, Rome, Italy, Oct. 2017.
[3] N. Zhao, F. Cheng, F. R. Yu, J. Tang, Y. Chen, G. Gui, and H. Sari, "Caching UAV assisted secure transmission in hyper-dense networks based on interference alignment," *IEEE Trans. Commun.*, vol. 66, no. 5, pp. 2281–2294, May 2018.
[4] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 3949–3963, Jun. 2016.
[5] P. Yang, X. Cao, C. Yin, Z. Xiao, X. Xi, and D. Wu, "Proactive drone-cell deployment: Overload relief for a cellular network under flash crowd traffic," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 10, pp. 2877–2892, Oct. 2017.
[6] M. Chen, M. Mozaffari, W. Saad, C. Yin, M. Debbah, and C. S. Hong, "Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1046–1061, May 2017.
[7] Y. Zhou, N. Cheng, N. Lu, and X. S. Shen, "Multi-UAV-aided networks: Aerial-ground cooperative vehicular networking architecture," *IEEE Veh. Technol. Mag.*, vol. 10, no. 4, pp. 36–44, Dec. 2015.
[8] Q. Wu, Y. Zeng, and R. Zhang, "Joint trajectory and communication design for multi-UAV enabled wireless networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 2109–2121, Mar. 2018.
[9] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Drone small cells in the clouds: Design, deployment and performance analysis," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Diego, CA, USA, Dec. 2014.
[10] R. I. Bor-Yaliniz, A. El-Keyi, and H. Yanikomeroglu, "Efficient 3-D placement of an aerial base station in next generation cellular networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, May 2016.
[11] J. Lyu, Y. Zeng, R. Zhang, and T. J. Lim, "Placement optimization of UAV-mounted mobile base stations," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 604–607, Mar. 2017.
[12] M. Alzenad, A. El-Keyi, F. Lagum, and H. Yanikomeroglu, "3-D placement of an unmanned aerial vehicle base station (UAV-BS) for energy-efficient maximal coverage," *IEEE Wireless Commun. Lett.*, vol. 6, no. 4, pp. 434–437, Aug. 2017.
[13] E. Kalantari, M. Z. Shakir, H. Yanikomeroglu, and A. Yongacoglu, "Backhaul-aware robust 3D drone placement in 5G+ wireless networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Paris, France, May 2017.
[14] H. Mei, J. Bigham, P. Jiang, and E. Bodanese, "Distributed dynamic frequency allocation in fractional frequency reused relay based cellular networks," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1327–1336, Apr. 2013.
[15] N. Liu, T. Liu, W. Zhang, S. Han, and Q. Wang, "Distributed power allocation for device-to-device communications underlaying cellular networks," in *Proc. 13th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Valencia, Spain, Jun. 2017.
[16] S. Kumar and K. Rajawat, "Distributed interference alignment for MIMO cellular network via consensus ADMM," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Washington, DC, USA, Dec. 2016.

[17] J. Wang, Y. Ding, S. Bian, Y. Peng, M. Liu, and G. Gui, "UL-CSI data driven deep learning for predicting DL-CSI in cellular FDD systems," *IEEE Access*, vol. 7, pp. 96105–96112, 2019.
[18] G. Gui, H. Huang, Y. Song, and H. Sari, "Deep learning for an effective nonorthogonal multiple access scheme," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8440–8450, Sep. 2018.
[19] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4074–4077, Apr. 2019.
[20] H. Huang, Y. Song, J. Yang, G. Gui, and F. Adachi, "Deep-learning-based millimeter-wave massive MIMO for hybrid precoding," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3027–3032, Mar. 2019.
[21] L. Li, H. Ren, X. Li, W. Chen, and Z. Han, "Machine learning-based spectrum efficiency hybrid precoding with lens array and low-resolution ADCs," *IEEE Access*, vol. 7, pp. 117986–117996, 2019.
[22] L. Li, Y. Xu, J. Yin, W. Liang, X. Li, W. Chen, and Z. Han, "Deep reinforcement learning approaches for content caching in cache-enabled D2D networks," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 544–557, Jan. 2020.
[23] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3039–3071, 4th Quart., 2019, doi: 10.1109/comst.2019.2926625.
[24] Z. Han, D. Niyato, T. Basar, W. Saad, and A. Hjorungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge, U.K: Cambridge Univ. Press, 2012.
[25] S. M. Yu and S.-L. Kim, "Game-theoretic understanding of price dynamics in mobile communication services," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 5120–5131, Sep. 2014.
[26] M. N. Soorki, W. Saad, M. H. Manshaei, and H. Saidi, "Stochastic coalitional games for cooperative random access in M2M communications," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6179–6192, Sep. 2017.
[27] J.-M. Lasry and P.-L. Lions, "Mean field games," *Jpn. J. Math.*, vol. 2, no. 1, pp. 229–260, Sep. 2007.
[28] M. Huang, P. E. Caines, and R. P. Malhame, "Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε-Nash equilibria," *IEEE Trans. Autom. Control.*, vol. 52, no. 9, pp. 1560–1571, Sep. 2007.
[29] O. Gueant, J. M. Lasry, and P. L. Lions, "Mean field games and applications," in *Paris-Princeton Lectures on Mathematical Finance*. New York, NY, Springer, 2011, pp. 205–266.
[30] 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.
[31] T. K. Thuc, E. Hossain, and H. Tabassum, "Downlink power control in two-tier cellular networks with energy-harvesting small cells as stochastic games," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 5267–5282, Dec. 2015.
[32] L. Li, Y. Xu, Z. Zhang, J. Yin, W. Chen, and Z. Han, "A prediction-based charging policy and interference mitigation approach in the wireless powered Internet of Things," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 2, pp. 439–451, Feb. 2019.
[33] C. Yang, H. Dai, J. Li, Y. Zhang, and Z. Han, "Distributed interference-aware power control in ultra-dense small cell networks: A robust mean field game," *IEEE Access*, vol. 6, pp. 12608–12619, 2018.
[34] K. Xue, L. Li, F. Yang, H. Zhang, X. Li, and Z. Han, "Multi-UAV delay optimization in edge caching networks: A mean field game approach," in *Proc. 28th Wireless Opt. Commun. Conf. (WOCC)*, Beijing, China, May 2019.
[35] P. Semasinghe and E. Hossain, "Downlink power control in self-organizing dense small cells underlaying macrocells: A mean field game," *IEEE Trans. Mobile Comput.*, vol. 15, no. 2, pp. 350–363, Feb. 2016.
[36] H. Tembine, R. Tempone, and P. Vilanova, "Mean field games for cognitive radio networks," in *Proc. Amer. Control Conf. (ACC)*, Montreal, QC, Canada, Jun. 2012.
[37] A. F. Hanif, H. Tembine, M. Assaad, and D. Zeghlache, "Mean-field games for resource sharing in cloud-based networks," *IEEE/ACM Trans. Netw.*, vol. 24, no. 1, pp. 624–637, Feb. 2016.
[38] R. Couillet, S. M. Perlaza, H. Tembine, and M. Debbah, "Electrical vehicles in the smart grid: A mean field game analysis," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 6, pp. 1086–1096, Jul. 2012.
[39] Z. Zhang, L. Li, X. Liu, W. Liang, and Z. Han, "Matching-based resource allocation and distributed power control using mean field game in the NOMA-based UAV networks," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Honolulu, HI, USA, Nov. 2018, pp. 420–426..
[40] Y. Sun, L. Li, K. Xue, X. Li, W. Liang, and Z. Han, "Inhomogeneous multi-UAV aerial base stations deployment: A mean-field-type game approach," in *Proc. 15th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Tangier, Morocco, Jun. 2019.
[41] Y. Xu, L. Li, Z. Zhang, K. Xue, and Z. Han, "A discrete-time mean field game in multi-UAV wireless communication systems," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Beijing, China, Aug. 2018, pp. 714–718.
[42] K. Xue, Z. Zhang, L. Li, H. Zhang, X. Li, and A. Gao, "Adaptive coverage solution in multi-UAVs emergency communication system: A discrete-time mean-field game," in *Proc. 14th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Limassol, Cyprus, Jun. 2018, pp. 1059–1064.
[43] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *Proc. IEEE Global Commun. Conf.*, Austin, TX, USA, Dec. 2014, pp. 2898–2904.
[44] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 569–572, Dec. 2014.
[45] R. Bellman, "Dynamic programming and lagrange multipliers," *Proc. Nat. Acad. Sci. USA*, vol. 42, no. 10, pp. 767–769, Oct. 1956.
[46] M. Bardi and I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. Philadelphia, PA, USA: SIAM, 2008.

LIXIN LI (Member, IEEE) received the B.Sc. and M.Sc. degrees in communication engineering, and the Ph.D. degree in control theory and its applications from Northwestern Polytechnical University (NPU), Xi'an, China, in 2001, 2004, and 2008, respectively. He was a Postdoctoral Fellow of NPU, from 2008 to 2010. In 2017, he was a Visiting Scholar with the University of Houston, TX, USA. He is currently an Associate Professor with the School of Electronics and Information, NPU. He has authored or coauthored over 100 technical articles in journals and international conferences. He holds 10 patents. His current research interests include wireless communications, game theory, and machine learning. He received the 2016 NPU Outstanding Young Teacher Award, which is the highest research and education honors for young faculties in NPU.

ZIHE ZHANG received the B.Sc. degree in communication engineering and the M.Sc. degree in communication and information systems under the supervision of Prof. Lixin Li from Northwestern Polytechnical University, Xi'an, China, in 2016 and 2019, respectively. He is currently a Technical Engineer with the 54th Research Institute of China Electronic Science and Technology Group Corporation. His research interests include mean field game, UAVs communications, and simultaneous wireless information, and power transfer.



KAIYUAN XUE is currently pursuing the master's degree with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China, under the supervision of Prof. Lixin Li. His research interests include unmanned aerial vehicles, wireless caching, and game theory in wireless communication networks.

MENG WANG is currently pursuing the master's degree with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China, under the supervision of Prof. Lixin Li. His research interest includes game theory in wireless communication networks, and reinforcement learning.

MIAO PAN (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from the Dalian University of Technology, China, in 2004, the M.A.Sc. degree in electrical and computer engineering from the Beijing University of Posts and Telecommunications, China, in 2007, and the Ph.D. degree in electrical and computer engineering from the University of Florida, in 2012. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, University of Houston. His research interests include big data privacy, cybersecurity, cyber-physical systems, and cognitive radio networking. He is a member of the ACM. His work won the Best Paper Award from VTC 2018, and GLOBECOM 2015 and 2017. He was an Associate Editor of the IEEE Internet of Things (IoT) Journal, from 2015 to 2018. He was a recipient of the NSF CAREER Award, in 2014.

ZHU HAN (Fellow, IEEE) received the B.S. degree in electronic engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was a Research and Development Engineer of JDSU, Germantown, MD, USA. From 2003 to 2006, he was a Research Associate with the University of Maryland. From 2006 to 2008, he was an Assistant Professor with Boise State University, ID, USA. He is currently a John and Rebecca Moores Professor with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX, and the Department of Computer Science and Engineering, Kyung Hee University, Seoul, South Korea. He is also a Chair Professor with National Chiao Tung University, China. He has been a 1% highly-cited researcher since 2017 according to the Web of Science. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grids. He was an IEEE Communications Society Distinguished Lecturer, from 2015 to 2018. He received the NSF Career Award, in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society, in 2011, the EURASIP Best Paper Award for the *Journal on Advances in Signal Processing*, in 2015, the IEEE Leonard G. Abraham Prize in the field of communications systems (Best Paper Award of the IEEE JSAC) in 2016, and several best paper awards from the IEEE conferences. He has been a Fellow of the AAAS and ACM Distinguished Member, since 2019.