Delay Optimization in Multi-UAV Edge Caching Networks: A Robust Mean Field Game

Lixin Li, *Member, IEEE,* Meng Wang, Kaiyuan Xue, Qianqian Cheng, Dawei Wang, *Member, IEEE,* Wei Chen, *Senior Member, IEEE,* Miao Pan, *Senior Member, IEEE,* and Zhu Han, *Fellow, IEEE*

***Abstract*—The data requirements of a large number of users incur significant delay to base stations/small base stations (BSs/SBSs) in edge caching networks (ECNs). Specially, in certain special scenarios, such as sports events, parades and nature disasters, it is a great challenge to deploy fixed BSs/SBSs. Thus, a kind of flexible replacement of BSs/SBSs, unmanned aerial vehicles (UAVs), are considered as the caching platform to serve users. Hence, this paper proposes a distributed delay optimization algorithm by using the massive UAVs to cache the request contents of users. The purpose of each UAV is to facilitate BSs/SBSs to minimize user delay in downloading content based on the popularity of content and the flight strategy. In addition, we consider the main disturbance caused by the atmospheric turbulence. For the control of large-scale UAVs, we take the robust mean field game theory to model this kind of caching and dynamic flight strategy problem, in which the atmospheric turbulence is described by the disturbance term in the drift function. The simulation results show that the proposed algorithm can reduce the download delay under the limited energy consumption and finally achieve the equilibrium.**

***Index Terms*—edge caching networks, contents popularity, robust mean filed game, ultra-dense networks**

Manuscript received XXX; revised XXX; accepted XXX. Date of publication XXX; date of current version XXX. This work was supported in part by Key R & D plan of Shaanxi Province under Grant 2020GY-034, in part by the Seed Foundation of Innovation and Creation for Graduate Students in Northwestern Polytechnical University under Grant CX2020152, in part by National Natural Science Foundation of China under Grant 61901378, Grant 61901379, Grant 62001387, and in part by Grant NSF EARS-1839818, Grant CNS1717454, Grant CNS-1731424, Grant CNS-1646607, Grant CNS-1801925, Grant CNS-2029569, and Grant CNS-1702850. This paper was presented in part at the 2019 28th Wireless and Optical Communication Conference (WOCC) [1]. The associate editor coordinating the review of this article and approving it for publication was XXX. (Corresponding author: Lixin Li.)

L. Li, M. Wang, K. Xue, Q. Cheng, and D. Wang are with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China (email: lilixin@nwpu.edu.cn; 2019261837@mail.nwpu.edu.cn; xuekaiyuan@nwpu.edu.cn; chengqianqian@mail.nwpu.edu.cn; wangdw@nwpu.edu.cn).

W. Chen is with the Department of Electronic Engineering, Tsinghua University, Beijing, China, 100084, China (e-mail: wchen@tsinghua.edu.cn).

M. Pan is with the Department of Electrical and Computer Engineering, Univeristy of Houston, Houston, 77004, USA (e-mail: mpan2@uh.edu).

Z. Han is with the Department of ECE, University of Houston, Houston, TX 77004 USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul 446-701, South Korea (e-mail: zhan2@uh.edu).

I. INTRODUCTION

THE amount of user data has exploded with the rapid development of 5G technology, which brings a great challenge to the traditional cellular networks [2]. Therefore, the ultra dense network (UDN) technology is proposed in order to solve this problem. Specifically, the UDN core idea is to deploy massive base stations/small base stations (BSs/SBSs) to increase the system capacity and mitigate the pressure on the traditional cellular networks. However, when a large number of BSs/SBSs simultaneously request contents from the core network, serious network congestion can be caused due to the limited backhaul capacity of the BSs/SBSs in the UDN [3]. Meanwhile, most of these requested contents are the repeated downloads of a large number of popular content (for example, popular video and audio), which can be averted by predicting and preloading.

The caching technology is a kind of method to predict and preload the popular contents, which becomes widely adopted as the storage equipment with the capacity increasing and the cost decreasing. Recently, the use of distributed caches at the edge of the network has become an effective solution, which can alleviate network congestion by mitigating backhaul congestion. Specifically, data can be transmited to the requesters by using high-speed and low-cost mobile edge networks during off-peak hours, and then is storaged during the off-peak hours [4]–[10]. With the file stored in advance, the BSs/SBSs can directly send the requested contents to the users (cache-hit user). Therefore, the BSs/SBSs are not required to download and transmit to the users through the backhaul links repeatedly, which reduces the pressure of the backhaul links and the response time to fetch the contents greatly.

In order to achieve edge caching with limited storage capacity, two issues should be considered: (i) what contents the SBS should cache, and (ii) how to cache. Specifically, the first issue means that each cache unit should cache the popular contents to increase the cache hit ratio. This usually requires estimating the short or the long term content requests of the users [11]–[13]. The other issue means the effective caching strategy should be optimized, which can also improve cache efficiency [14], [15]. These two issues are more challenging in 5G ultra-dense networks because of the more intensive information interaction and more complex content distribution. As a potential solution, distributed caching is to cache the content file requested by the users to the BSs/SBSs, which can reduce the load of the backhaul link in the ultra-dense network. By jointly considering the content requirements the users in spatio and temporal as well as the interference management of a large number of BSs/SBSs, works [11]–[15] proposed the distributed cache control strategies under the ultra-dense edge cache networks.

However, in certain special situations (such as disasters, large events, etc.), the number of user devices have increased dramatically or moved constantly, which is a great challenge to deploy BSs/SBSs. Moreover, on one hand, with the temporal and spatial changes of data traffics, the networks are likely to be overloaded with handling bursty geographic traffic. On the other hand, the network traffics are wasted when the traffic is low. To overcome the above challenge, the unmanned aerial vehicles (UAVs) is a promising method to meet the download requirements of matching users by equipping the storage units and the transceivers [23]–[27]. Applying the aerial UAV small cells to the edge caching network (ECN), the performance can be improved by leveraging the low-cost, flexibility and ease of deployment of the UAV. However, the existing UAV caching strategies just model single or several UAVs, which cannot satisfy the scenario of a large number of users and interactions in the ultra-dense edge caching networks. Thus, a new method, mean-field game (MFG) with large-scale agents, is considered to solve this ultra-dense removable scenario [28].

In recent years, the applications of MFG in communication scenarios have been increasing rapidly [29]–[36]. This credit that a huge amount of users in the 5G network. In the ultra-dense network, due to the existence of a large number of agents, it brings the complexity of computing. The MFG is especially suitable for this dense network because the number of players can toward infinity. In this game, the competition between the common agent and all other agents is simplified to the competition with the mass (mean field). Moreover, it is distributed, and each player can make decisions only through the influence from the mass without controls or policies of other players. Therefore, the number of agents in the MFG is independent of the performance, which is suitable for ultra-dense scenarios, especially.

In summary, the flexibility and dynamic flight characteristics of the UAVs make the interaction of inter-UAVs particularly complicated, which is mainly reflected in the intricate interference between UAVs. Furthermore, the low latency and high reliability are the main challenges of UAV communications. Thus, this paper reduces the total delay of the system by adjusting the position of each UAV. There are two methods to decrease the download delay: one is to improve the hit rate of content cache by describing the popular dynamics of contents; and the other approach is to increase the backhaul capacity by controlling the flight of UAVs to reduce the interference among internal UAVs, thereby reducing the delays. Thus, we formulate this flight control problem of the multi-UAVs as a stochastic differential game (SDG), where the popular dynamics of contents are described by Chinese Restaurant Process (CRP) and the Ornstein-Uhlenbeck (OU) process. In this game, each UAV can consider the states of other UAVs to determine the flight state of itself. Meanwhile, we consider the effect of the atmospheric turbulence, which is the vital factor in UAV flight. Then we formulate a robust MFG model to reduce the computational complexity, where the interference of inter-UAVs are simplified. In additon, we give the position distributions of all UAVs at the initial and the final time, which show the variation of the UAV mass. Besides, we study the average download delay with a specific energy consumption under the limited time, and finally obtain the equilibrium.

The main contributions of this paper are summarized as follows:

- We model a dynamic flight network with multiple UAVs in ECN. Meanwhile, considering the popularity of contents and the flight energy consumption, each UAV can dynamically adjust its space position to minimize the download delay of corresponding users.
- The problem of delay optimization is modeled as SDG, which contains the dynamic of UAVs flight and the contents popularity. In the part of contents managements, the CRP and the OU process are used to model the long-term changes and short-term dynamics of contents popularity, respectively.
- In this paper, the flight energy consumption of UAVs and the download delay of the corresponding users are constructed as cost functions. Specifically, the download delay is constituted by two stages: BSs/SBSs-UAVs links and UAVs-users links. And considering the simplification of this problem, the caching and flight strategy of UAVs is remodeled as the MFG framework.
- The robust performance is considered in the proposed framework, in which the atmospheric turbulence, a critical influence factor for UAV flighting, is designed as the disturbance term in the drift function of stochastic differential equation (SDE). Therefore, the MFG framework is redesigned as the robust MFG framework, which contains the novel Hamilton-Jacobi-Bellman (HJB) and Fokker-Planck-Kolmogorov (FPK) equations.

The rest of this paper is organized as follows. Section II introduces the related works of this paper. In Section III, the system models including the basic scenario model, the caching model, the UAV flight dynamic model and the air-to-ground (A2G) channel model are presented. The SDG and problem formulation are provided in Section IV. Section V investigates the robust MFG frameworks and calculates the equilibrium. Section VI shows the numerical results. The conclusions are summarized in Section VII.

II. Related works

The edge caching in ultra-dense networks has been a popular topic recent years [16]–[22]. In [16], the author studied the distributed caching problem of a dense wireless small cellular networks, where the goal of each SBS was to reduce the load on the capacity-limited backhaul link by controlling its own caching strategy. In [17], [18], game theory was used to solve the wireless cache problem. In [19] and [20], the authors proposed a cache algorithm based on random geometry to maximize the local cache gains of users. It is worth mentioning that these works utilized MFG theory to describe the dynamic characteristics of content popularity. Moreover, in [21], the authors investigated prediction of users mobility and the contents popularity in cache-enabled device-to-device (D2D) networks by using two potential recurrent neural network methods. In [22], the authors proposed a algorithm based on the asynchronous advantage actor-critic to minimize total transmission cost from the neighbors in caching networks. Meanwhile, they compared with other classical caching strategies. These policies solved the problem of contents management effectively without users high mobility and great density.

In addition, the UAVs are widely used in the edge caching networks. In [23], the authors proposed an air-ground integrated mobile edge network architecture and introduced the method of UAV-assisted edge caching and computation in the proposed architecture. However, [23] summarized these new conceptions, which were not illustrated in detail. In [24] and [25], the authors proposed a UAV-assisted secure transmission method based on interference alignment for ultra-dense small cell network caches, which can reduce the pressure on the backhaul link and ensure the secure transmission of the information. In [26], the authors investigated the problem of optimizing quality of experience (QoE) of the wireless devices by deploying cache-enabled UAVs in cloud wireless access networks, where the QoE of users is the measure of data transfer rate, delay and device type. In addition, they proposed a machine learning algorithm, which was based on the echo state networks to predict the distribution of each users content request and their mobility. However, this work only adapted several UAVs and did not consider the effects from the fronthaul link caused by a large number of UAVs. In [27], the author proposed an effective method to make UAVs can actively perform content caching in order to solve the problem of weak endurance of UAVs. Moreover, they discussed the trade-off between the file caching cost and the file retrieval cost, which were the two stages of their proposed scheme. Similarly, the number of UAVs cannot satisfy the content demands of ultra-dense networks.

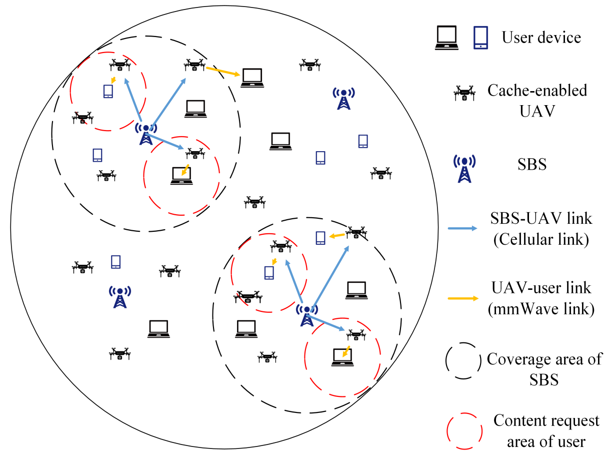


Fig. 1: The multi-UAVs in ECNs

III. System Model

In this section, we present the system models including the basic scenario model, the caching model, the UAV flight dynamic model with disturbance and the A2G channel model. Firstly, we give the summary of the parameter symbols in Table I.

A. Basic scenario model

According to the scenario, we consider an ultra-dense ECN. In the circular area A with a radius r_A, U users are served by N UAVs. Meanwhile, there are K radio assess stations in area A. These radio assess stations can be WiFi or femtocell network access points, even a macrocell base stations. For simplicity, the radio assess stations in our model are SBSs without loss of generality. Moreover, we assume that the user distribution density λ_u and UAV distribution density λ_n are much larger than the distribution density of SBS λ_k, i.e. $\lambda_u \gg \lambda_k$, $\lambda_n \gg \lambda_k$. It is obvious that the number of UAVs and users can be expressed as $N = \lambda_n \pi r_A^2$ and $U = \lambda_u \pi r_A^2$, respectively. Fig. 1 shows the distribution of these U users and N UAVs in the ultra-dense ECN, which follow independent homogeneous Poisson point processes (PPPs) [37].

In Fig. 1, the kth SBS ($k \in \mathcal{K}$) has a coverage area $b_k(c_k, r_k)$ with the coverage radius r_k, where c_k is the coordinate of SBS k. To simplify the model, we assume that the SBSs are uniformly distributed and can cover the entire area. In addition, within the area of content requestion $b_u(c_u, r_u)$, the mathing UAV transmits the contents to the uth users, where r_u and c_u are the radius and the center of circle, respectively. In this area, the request radius represents the average distance of the signal

TABLE I: List of parameters

Parameter	Description	Parameter	Description
N	Number of UAVs	λ_n	Density of UAVs
U	Number of users	λ_u	Density of users
K	Number of SBSs	λ_k	Density of SBSs
b_k	Coverage area of SBS k	b_u	Content request area of user u
b_i	Coverage area of UAV i	σ_1,σ_2	Volatility
M	Number of contents	W_{1t}, W_{2t}	Wiener process
m	Number of contents stored on the UAV	f_c	Carrier frequency
$\mathbb{P}_{i,u}^{(1)\mathrm{LoS}}$, $\mathbb{P}_{i,u}^{(1)\mathrm{NLoS}}$	Probability of LoS/NLoS	c	Speed of light
$\Pr(T)$	Average probability of content request during time T	a_i	Atmospheric turbulence
$U^{(i)}$	Number of the user in b_i	γ	Disturbance attenuation
$X_j^{(i)}$	Probability of requesting content j from UAV i	n_{LoS}, n_{NLoS}	Free space path loss exponent
x_i, v_i	Flight state/control of UAV i	ς_{LoS}, $\varsigma_{\mathrm{NLoS}}$	Gaussian random variables
$d_{i,u}^{(1)}$, $d_{k,i}^{(2)}$	Distance from UAV i to user u/SBS k to UAV i	$\sigma_{i,u}^2$, $\sigma_{k,i}^2$	Noise power
$L_{i,u}^{(1)\mathrm{LoS}}$, $L_{i,u}^{(1)\mathrm{NLoS}}$	LoS/NLoS path loss from the UAV i to user u	C	Size of each content
$L_{k,i}^{(2)\mathrm{LoS}}$, $L_{k,i}^{(2)\mathrm{NLoS}}$	LoS/NLoS path loss from the SBS k to UAV i	P_U, P_S	Transmit power of UAVs/SBSs
$B_{i,u}^{(1)}$, $B_{k,i}^{(2)}$	Bandwidth of two links	E_i	Flight energy consumption
$\mathbb{P}_i^{\mathrm{h}}$, $\mathbb{P}_i^{\mathrm{m}}$	Probability of cache hit/miss	D_i^{h}, D_i^{m}	Delay of cache hit/miss

power larger than the noise power. For ease of exposition, we assume that all M contents are stored in each SBS, where the size of each content is equal to L. Moreover, each UAV can store m contents, and $m < M$.

In addition, we assume that during a period of time T, U users more likely request the same set of $\mathcal{M}$ contents. This assumption is realistic because in practice the contents popularity update period of users is long, usually in days [27]. In general, UAV i downloads the contents for users from the SBS via a backhaul link with limited capacity. User u is associated with any UAV in the content requesting area, which has cached the requested content. If all the UAVs in the content request area do not cache the requested content, user u will connect with the nearest idle UAV, which needs to download the requested content from the matching SBS through the backhaul link.

B. Caching model

In the ECN, the requirements of the users are changed according to the contents popularity. Here, the dynamic of the contents popularity is the spatial and temporal dynamic, because the contents popularity in different geographical regions is different. Therefore, in any region, the contents popularity dynamically changes with time. Besides, we assume that the ith UAV ($i \in \mathcal{N}$) has a coverage area $b_i(c_i, r_i)$, where c_i and r_i represent the coordinate and the radius of UAV i, respectively. Meanwhile, this coverage area is the contents popularity searching area, called community. In these communities, there are $U^{(i)}$ users with dynamic content requests, where $U^{(i)}$ represents the number of users covered by the ith UAV. Specifically, $U^{(i)}$ varies with the number of requesting users, but is regarded as a constant in an optimization period. In each community, we divide the time dynamics of users content request probabilities (i.e., contents popularity) into long-term and short-term dynamics [20], [39]. Specifically, the long-term dynamic describes the long-term variation of the contents popularity, and the short-term dynamic describes the instantaneous variation of the contents popularity.

In this paper, we use the Chinese Restaurant Process (CRP) to describe the long-term changes in contents popularity [20]. In other words, CRP captures the independent random requests of the users at each UAV during each time period T, where the users and the contents can be viewed as the customers and the table in CRP, respectively. Therefore, following [40], we formulate the average probability of user u requests content j from UAV i during time period εT (ε is a positive integer), which can be represented as

$$\Pr_{u,j}^{(i)}(\varepsilon T) = \begin{cases} \frac{U_j^{(i)}(\varepsilon T)}{U_c^{(i)}(\varepsilon T)+\kappa} & \text{for } j \in \mathbb{J}_{i+} \\ \frac{\kappa}{U_c^{(i)}(\varepsilon T)+\kappa} & \text{for } j \in \mathbb{J}_{i-} \end{cases}, u=1,...,U, \forall i, \tag{1}$$

where the $U_j^{(i)}(\varepsilon T)$ represents the number of existing users requesting content j in the current time period τT in b_i. The $U_c^{(i)}(\varepsilon T) = u - 1$ explains the number of users in b_i who have requested the content prior to the requesting of user u. $\mathbb{J}_{i+}$ and $\mathbb{J}_{i-}$ are the set consisting of contents who have and have not requested at least once by users in b_i, respectively. κ is a constant in CRP indicating the skewness of the distribution of the popularity. As it increases, the distribution of contents popularity becomes more diverged. Here we assume that the inter-arrival time of content request is smaller than T and the number of request is sufficiently large. Compared to the classic description model of contents popularity Zipf model, CRP can generate exchangeable UAV assignments [40].

For the short-term dynamics of contents popularity over a time period εT, we use an Ornstein-Uhlenbeck (OU) process $\{X(t) : t > 0\}$ of mean regression model to describe [41]. This process is a mean regression random process that is commonly used to describe contents popularity [19], [20] and [39]. The SDE of this process is as follows:

$$\mathrm{d}X_j^{(i)}(t) = \lambda \left(\mathrm{Pr}_{u,j}^{(i)}(\varepsilon T) - X_j^{(i)}(t) \right) \mathrm{d}t + \sigma_1 \mathrm{d}W_{1t}, \quad (2)$$

where $\lambda > 0$ denotes the mean reversion rate. $X_j^{(i)}(t)$ is UAV i 's request probability of content j at time $t \in T$. $\mathrm{Pr}_{u,j}^{(i)}(\varepsilon T)$ is the long-term average probability shown in (1). σ_1 is the volatility and W_{1t} represents the Wiener process (also known as the Brownian motion, and satisfied $\mathrm{d}W_{1t}/\mathrm{d}t = 0$ and $\mathrm{d}^2 W_{1t} = \mathrm{d}t$ [41]). It is precisely because of the randomness of the Wiener process W_{1t} that the request probability $X_j^{(i)}(t)$ deviates from the long-term average value $\mathrm{Pr}_{u,j}^{(i)}(\varepsilon T)$.

C. UAV flight dynamic model with atmospheric turbulence

The aim of this paper is to minimize the system delay of each UAV by adjusting its 3D position. Moreover, the atmospheric turbulence is one of important factors affecting the smoothness of UAV flight [42]. That means the disturbance from the atmospheric should be well described, instead of simply treated as background noise when designing the flight dynamic. Therefore, in this subsection, we consider the disturbance into the UAV flight dynamic to analyze the effects of atmospheric turbulence. Then, to describe the behavior of the UAVs flight, we formulate the dynamic equation of $x_i(t)$,where $x_i(t)$ is the distance between the UAV i and the who user severed by UAV i. Specifically, in order to describe flight dynamic of the UAVs, the common UAV $(i, i \in \mathcal{N})$ is also adopted. The SDE of UAVs flight dynamic can be represented as follows,

$$\begin{aligned} &\mathrm{d}x_i(t) = (v_i(t) - a_i(t))\,\mathrm{d}t + \sigma_2 \mathrm{d}W_{2t}, \\ &x_i(t+1) = x_i(t) + \mathrm{d}x_i(t), \end{aligned} \quad (3)$$

where $\mathrm{d}x_i(t)$ and $v_i(t)$ are the flight distance and the flight control of UAV i, respectively. $a_i(t)$ represents the atmospheric turbulence, which determines the worst case risk neutral cost function of agent i. σ_2 is the volatility and W_{2t} is the Brownian motion, which represents the influence from the size, weight and the wing area of the UAV. W_{1t} in (2) and the W_{2t} are independent of each other. The dynamic function in (3) describes the flight of general UAV i, where the flight state $x_i(t)$ represents the distance of UAV i and user u who served by UAV i. Flight control $v_i(t)$ is the flight velocity.

D. A2G channel model

The position control of the cache-enabled UAV makes modeling the A2G channel important. In this paper, we mainly investigate the downlink from the SBSs to the UAVs and the downlink from the UAVs to the users. For the ground-to-air (G2A) links from the SBSs to the UAVs, we adapt the line-of-sight (LoS) and the non-line-of-sight (NLoS) links over the licensed band. For the A2G links from the UAVs to the users, we leverage the mmWave frequency spectrum because the high altitude of the UAVs can dramatically reduce the blocking effect of obstacles.

1) UAV-user links: For the closer mmWave propagation channel, the LoS and the NLoS path loss of UAV i which transmits a content to user u at time t can be repersented as (in dB) [26], [43]:

$$L_{i,u}^{(1)\mathrm{LoS}}(t) = L_F(d_0) + 10 n_{\mathrm{LoS}} \log\left(d_{i,u}^{(1)}(t)\right) + \varsigma_{\mathrm{LoS}}, \quad (4a)$$

$$L_{i,u}^{(1)\mathrm{NLoS}}(t) = L_F(d_0) + 10 n_{\mathrm{NLoS}} \log\left(d_{i,u}^{(1)}(t)\right) + \varsigma_{\mathrm{NLoS}}, \quad (4b)$$

where

$$L_F(d_0) = 20 \log\left(4\pi f_c d_0 / c\right), \quad (4c)$$

in which n_{LoS} and n_{NLoS} represent free space path loss exponent of LoS and NLoS, respectively. $d_{i,u}^{(1)}(t)$ is the distance between UAV i and user u. ς_{LoS} and $\varsigma_{\mathrm{NLoS}}$ are the Gaussian random variables, which represent the shadowing random variables. And $(4c)$ denotes the free space path loss, where f_c, d_0, and c represent carrier frequency, the free space reference distance, and the speed of light, respectively. In addition, the probability of LoS is shown in [43]:

$$\mathbb{P}_{i,u}^{(1)\mathrm{LoS}} = \frac{1}{1 + a \exp\left[-b\left(\theta_{i,u}^{(1)} - a\right)\right]}, \quad (5)$$

where a and b indicate the environmental parameter depending on the height and the density of buildings. $\theta_{i,u}^{(1)}$ represents the elevation angle. Similarly, the probability of NLoS is expressed as $\mathbb{P}_{i,u}^{(1)\mathrm{NLoS}} = 1 - \mathbb{P}_{i,u}^{(1)\mathrm{LoS}}$. Consequently, the average path loss can be described as:

$$\begin{aligned} \overline{L}_{i,u}^{(1)}(t) = &L_{i,u}^{(1)\mathrm{LoS}}(t) \times \mathbb{P}_{i,u}^{(1)\mathrm{LoS}}(t) \\ &+ L_{i,u}^{(1)\mathrm{NLoS}}(t) \times \mathbb{P}_{i,u}^{(1)\mathrm{NLoS}}(t). \end{aligned} \quad (6)$$

2) SBS-UAV links: In this model, the distance between SBSs and UAVs is not as close as the link of UAVs and users. Besides, there are more obstructions between the SBSs and the UAVs. Therefore, we adopt the cellular band to satisfy the reliability of the link and reduce the path loss. The LoS and the NLoS path loss of SBS k transmitting a content to UAV i at time $t \in T$ can be respectively expressed as (in dB):

$$\begin{aligned} L_{k,i}^{(2)\mathrm{LoS}}(t) &= d_{k,i}^{(2)}(t)^{-\alpha}, \\ L_{k,i}^{(2)\mathrm{NLoS}}(t) &= \beta d_{k,i}^{(2)}(t)^{-\alpha}, \end{aligned} \quad (7)$$

where $d_{k,i}^{(2)}(t)$ is the distance between SBS k and UAV i. α represents the path loss exponent. β is a positive constant because the NLoS experiences more shadows and diffraction. Meanwhile, the probabilities of LoS and average path loss of SBS-UAV links are identically with those two probabilities in UAV-user links.

IV. Problem Formulation and Robust Stochastic Differential Game

The goal of UAV i is to determine its flight policy, which can be represented as the flight speed vector u_i in order to minimize the cost of the system. Specifically, the cost function contains two main parts, the aviatic energy consumption and the total delay of system. In addition, this minimzing problem is formulated as a dynamic SDG.

A. *Cost function*

The cost of total delay is determined by two parts. One is caused by the probability of the cache hit, and the other is the backhaul capacity. Moreover, we consider the effect of UAVs flight energy consumption.

1) The cost of delay: The cache hit probability $\mathbb{P}_i^h(t)$ ($\mathbb{P}_i^h(t) \in [0,1]$), which belongs to UAV i, is designed as follows in accordance with (2),

$$\mathbb{P}_i^h(t) = \sum_{j=1}^{m} X_j^{(i)}(t), \tag{8}$$

where m represents the storage of each UAV. In this paper, we do not consider the optimization of cache content, which can be investigated in the future works. Thus, the probability of the cache miss can be expressed as $\mathbb{P}_i^{\mathrm{m}}(t) = 1 - \mathbb{P}_i^{\mathrm{h}}(t)$, which relates to the delay of the transmission. Specifically, if the request content j of user u has been cached by UAV i, it can directly transmit content j to user u, which is regarded as cache hit; otherwise, the situation is named as cache miss, where UAV i has to request content j from SBS k and then send it to user u.

Then the delay of cache hit is defined as

$$D_i^h(t) = C \Big/ R_{i,u}^{(1)}(t), \tag{9}$$

where $R_{i,u}^{(1)}(t)$ represents the backhaul capacity among UAV i and user u which is defined as follows:

$$\begin{aligned} R_{i,u}^{(1)}(t) &= B_{i,u}^{(1)} \log_2 \left(1 + \eta_{i,u}^{(1)}(t)\right) \\ &= B_{i,u}^{(1)} \log_2 \left(1 + \frac{P_U - \overline{L}_{i,u}^{(1)}(t)}{I_i(t, \boldsymbol{u_{-i}}) + \sigma_{i,u}^2}\right), \end{aligned} \tag{10}$$

where $B_{i,u}^{(1)}$ and ${\eta_{i,u}}^{(1)}(t)$ are the bandwidth and signal-to-interference-plus-noise ratio (SINR), respectively. P_U is the transmit power with constant value. $\sigma_{i,u}^2$ is the power of the white Gaussian noise. $I_i(t, \boldsymbol{u_{-i}}) = \sum_{l \in N, l \neq i} \left(P_U - \overline{L}_{l,u}^{(1)}(t)\right)$ is the influence from other A2G links. Therefore, the cost of delay is given by:

$$D_i(t) = \mathbb{P}_i^{\mathrm{h}}(t) \times D_i^{\mathrm{h}}(t) + \mathbb{P}_i^{\mathrm{m}}(t) \times D_i^{\mathrm{m}}(t). \tag{11}$$

For simplicity, we ignore the delay of the request from UAV i to SBS k. Therefore, the cache miss $D_i^{\mathrm{m}}(t)$ in (11) is formulated as

$$D_i^m(t) = \frac{C}{R_{i,u}^{(1)}(t)} + \frac{C}{R_{k,i}^{(2)}(t)}, \tag{12}$$

where $R_{k,i}^{(2)}(t)$ represents the backhaul capacity among the SBS k and the UAV i , which can be calculated by refering to the equation (10).

2) The flight energy consumption: In this part, the cost of the flight energy consumption is considered, which is formulated as [45]:

$$E_i(t) = v_i(t) \left(c_1 S^2 + \frac{c_2}{S^2}\right), \tag{13}$$

where c_1 and c_2 are the parameters according to the UAV wing area, weight and air density, etc. S is the speed of each UAV, and we assume it's a constant. $v_i(t)$ denotes the distance that the UAV flight. In (13), these two items represent the air resistance and lift of UAV i, respectively. Then, combine the (11) and (13), the global cost function is constructed as

$$J_i\left(v_i(t), \boldsymbol{v_{-i}}(t)\right) = w_1 D_i(t) + w_2 E_i(t), \tag{14}$$

where $\boldsymbol{v_{-i}}(t)$ is the control vector from other UAVs. w_1 and w_2 are the weights of $D_i(t)$ and $E_i(t)$, respectively. Thus, the average cost of common agent i over time $[0, T]$ can be obtained according to (14), which is formulated as

$$\begin{aligned} \mathcal{J}_i^{\gamma}\left(v_i, \boldsymbol{v_{-i}}, a_i\right) = \mathbb{E}\Bigg[&\int_0^T J_i\left(v_i(t), \boldsymbol{v_{-i}}(t)\right) \mathrm{d}t \\ &+ J_i\left(T\right) - \int_0^T \gamma^2 a_i^2(t) \mathrm{d}t \Bigg], \end{aligned} \tag{15}$$

where $J_i\left(T\right)$ represents the cost at the finial time T. γ is the disturbance attenuation parameter, which satisfied the constrained as [45]:

$$\gamma^2 \geq \frac{\mathcal{J}_i\left(v_i, \boldsymbol{v_{-i}}\right)}{a_i^2 + J_i(0)}, \tag{16}$$

where $\mathcal{J}_i\left(u_i, \boldsymbol{u_{-i}}\right)$ can be expressed as

$$\mathcal{J}_i\left(v_i, \boldsymbol{v_{-i}}\right) = \int_0^T J_i\left(v_i(t), \boldsymbol{v_{-i}}(t)\right) dt + J_i\left(T\right). \tag{17}$$

B. *Robust Stochastic Differential Game and Problem Formulation*

The control strategies for UAVs vary as time according to the dynamic equations in (2) and (3). The goal of each UAV is to find an equilibrium point in minimizing delay and minimizing energy consumption, which means that the UAV needs to reduce delay with as little energy consumption as possible. Thus, the process of minimizing the delay and energy cost in (14) can be modeled as a SDG. So, the SDG is formilated as:

$$\left\{\mathcal{N}, \mathcal{S}_j^{(i)}, \mathcal{U}_i, \mathcal{J}_i^{\gamma}\right\}.$$

The collection consists of four components:

- $\mathcal{N}$ is the number of agents (i.e. UAVs).
- $\mathcal{S}_j^{(i)}$ are the set of states including popularity state and flight state.
- $\mathcal{U}$ represents the control space of UAVi.
- $\mathcal{J}_i^{\gamma}$ is the cost function of UAVi.

Then we define the control problem as **P1**, where the UAV i minimizes the average cost by obtaining its own flight strategy $v_i^*(t)$. The state of UAV i with content j is represented as $\boldsymbol{s_j^{(i)}}(t) = \left\{X_j^{(i)}(t), x_i(t)\right\}$, $\forall i \in \mathcal{N}$, $\forall j \in \mathcal{M}$. **P1** is expressed as:

$$\textbf{P1:} \qquad V_i(t) = \inf_{v_i(t)} J_i^{\gamma}\left(v_i, \boldsymbol{v_{-i}}, a_i\right), t \in [0, T], \quad (18a)$$

subject to:

$$\mathrm{d}X_j^{(i)}(t) = \lambda\left(\mathrm{Pr}_{u,j}^{(i)}(\varepsilon T) - X_j^{(i)}(t)\right)\mathrm{d}t + \sigma_1 \mathrm{d}W_{1t}, \quad (18b)$$

$$\begin{aligned} &\mathrm{d}x_i(t) = (v_i(t) - a_i(t))\,\mathrm{d}t + \sigma_2 \mathrm{d}W_{2t}, \\ &x_i(t+1) = x_i(t) + \mathrm{d}x_i(t), \end{aligned} \quad (18c)$$

where $V_i(t)$ represents the value function.

In order to find out the solution of **P1**, N coupled differential equations with the same form of the following HJB equation, need to be computed.

$$\partial_t V_i(t) + H\left(J_i, v_i, \boldsymbol{s_j^{(i)}}, a_i\right) = 0, \quad (19)$$

where the Hamiltonian $H\left(J_i, v_i, \boldsymbol{s_j^{(i)}}, a_i\right)$ can be defined as

$$\begin{aligned} H\left(J_i, v_i, \boldsymbol{s_j^{(i)}}, a_i\right) &= \inf_{v_i(t)} [J_i(v_i(t), \boldsymbol{v_{-i}}(t)) - \gamma^2 a_i^2(t) \\ &+ \frac{\sigma_1^2}{2}\partial_{XX}^2 V_i(t) + \lambda(\mathrm{Pr}_{u,j}^i - X_j^{(i)}(t))\partial_X V_i(t) \\ &+ \frac{\sigma_2^2}{2}\partial_{xx}^2 V_i(t) + (u_i(t) - a_i(t))\partial_x V_i(t)]. \end{aligned} \quad (20)$$

The detail derivation process can be seen in [49]. Moreover, the existence of the Nash equilibrium (NE) for the differential game can be obtained as follows.

Firstly, all order derivatives of Hamiltonian $H(J_i, v_i, \boldsymbol{s_j^{(i)}}, a_i)$ exist because of the continuity of the cost function $J_i(v_i(t), \boldsymbol{v_{-i}}(t))$. Therefore, Hamiltonian $H(J_i, v_i, \boldsymbol{s_j^{(i)}}, a_i)$ is smooth, which means that exists a solution to the HJB equation.

We can obtain the unique joint solution $V_i^*(t)$ by solving these N equations together, which indicates that **P1** reaches the NE. In practice, when $N \geq 2$, the computational complexity increases exponentially owing to the fact that the control strategy v_i for each agent is affect with the influence of all the others' control strategy $\boldsymbol{v_{-i}}$. In addition, the acquisition of other agent's strategic information requires a large number of information interaction, which is almost impossible in UDNs. Therefore, MFG is considered to deal with this problem in the following section.

V. Robust Mean field game and equilibrium

In this section, we construct the robust SDG as the robust MFG framework and obtain the equilibrium solution and its proof. The MFG describes mass behavior as a mean field term. Therefore, the interactions of individual agent and other agents are expressed as the interaction with mass, which reduces the complexity significantly. Specifically, the agent's interaction with the mean field can be characterized by satisfying the following assumptions.

Assumption 1.

1) The number of agent N is large enough, even $N \to \infty$;

2) The exchangeable of each agent is ensured;

3) There are finite mean filed interactions.

The mean field interactions $I_i(t, \boldsymbol{u_{-i}}) = \sum_{l \in N, l \neq i}\left(P_U - \overline{L}_{l,u}^{(1)}(t)\right)$ in equation (10) should converge to a finite value according to the assumption 3), where P_U is a constant and $\overline{L}_{l,u}^{(1)}(t)$ is the average path loss obtained from equation (6). When $\overline{L}_{l,u}^{(1)}(t)$ is greater than a certain value, the interference $(P_U - \overline{L}_{l,u}^{(1)}(t))$ of the l_{th} agent can be negligible. So, the mean field interactions are finite, although N tends to infinity.

If the agents' control is constant by their index and determined only by their own states, then the agents in the SDG are said to be exchangeable or indistinguishable. It means that the index of replacement agents cannot change their control strategy. With this interchangeability, we can omit the index i and focus on an ordinary UAV.

Then the mean field term $m\left(t, X_j(t), x(t)\right)$ can be defined as the state density of the mass at time t. As a consequence, the empirical distribution of the mass can be given as follows,

$$M\left(t, X_j(t), x(t)\right) = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\delta_{\{x_i(t), X_j(t)\}}, \quad (21)$$

where δ represents the statistic function of the distribution of state. As the number of agents increases, $M\left(t, X_j(t), x(t)\right)$ will converge to $m\left(t, X_j(t), x(t)\right)$. Accordingly, the dynamic of the mass can be described by the FPK equation,

$$\begin{aligned} &\partial_t m(t, X_j(t), x(t)) - \frac{\sigma_1^2}{2}\partial_{XX}^2 m(t, X_j(t), x(t)) - \\ &\frac{\sigma_2^2}{2}\partial_{xx}^2 m(t, X_j(t), x(t)) + \lambda(\mathrm{Pr}_j - X_j(t))\partial_X m(t, X_j(t), x(t)) \\ &+ (v(t) - a(t))\partial_x m(t, X_j(t), x(t)) = 0. \end{aligned} \quad (22)$$

In (22), index i has been omitted because **Assumption 1** is satisfied.

Algorithm 1 Obtaining the MFE

1: Initialization:
2: $m(0)$: initial mean-field distribution;
3: $X(0), x(0)$: initial state;
4: Compute the optimal distribution of the mean field $m^*(t, X, x)$ and the value $V^*(t)$ by solving the equations of HJB (24) and FPK (22)
5: Repeat: Until the system obtain the MFE
6: Calculate the optimal control strategy $v^*(t)$ by using Proposition 1
7: Compute the value of the state $X_j(t)$ and $x(t)$ by using the SDEs (2) and (3)

The optimal mass distribution $m^*(t, X_j(t), x(t))$ can be obtained by solving the FPK equation in (22). Therefore, the interaction between agent i and the mass $I_i(t, \boldsymbol{u_{-i}})$ can be redefined as

$$I(t, m^*(t, X_j(t), x(t))) = \int_X \int_x P_U - m^*(t, X_j(t), x(t))\, \overline{L}_{l,u}^{(1)}(t) \mathrm{d}X \mathrm{d}x. \tag{23}$$

That means we no longer need the policy information of other agents. Hence, each agent only need to joint solve two coupled equations, HJB equation (24) and FPK equation (22), which greatly reduces the complexity of calculations and interactions. HJB equation (19) can be redefined as follows,

$$\partial_t V(t) + \inf_{v(t)} \Big[J(v(t), I(t, m^*(t, X_j(t), x(t)))) - \gamma^2 a^2(t) + \lambda (\mathrm{Pr}_j - X_j(t)) \partial_X V(t) + \frac{\sigma_1^2}{2} \partial_{XX}^2 V(t) + \frac{\sigma_2^2}{2} \partial_{xx}^2 V(t) + (v(t) - a(t)) \partial_x V(t) \Big] = 0. \tag{24}$$

Therefore, we can obtain the $V^*(t)$ and the $m^*(t, X, x)$ by solving HJB equation (24) and FPK equation (22), respectively. And it can reach mean field equilibrium (MFE) when having solved these two partial differential equations. Then we define the MFE of this model as follows.

Definition 1: The MFE with the generic flight policy $v^*(t)$ can be expressed as

$$J(v^*(t), m^*(t, X_j(t), x(t))) \leqslant J(v(t), m^*(t, X_j(t), x(t))). \tag{25}$$

Then the optimal policy $v^*(t)$ can be obtained by utilizing the Karush-Khun-Tucker (KKT) conditions, which is given in **Proposition 1** as follows.

Proposition 1: The optimal flight policy $v^*(t)$ of the generic agent is calculated by getting the critical point as:

$$\frac{\partial}{\partial v(t)} \left[C \Big/ R_t^{(1)} (I(t, m^*(t, X, x))) \right] + \left(c_1 S^2 + \frac{c_2}{S^2} \right) - \partial_x V^*(t) = 0. \tag{26}$$

TABLE II: Parameters in numerical analysis

Parameter	Value	Parameter	Value
λ_n	0.5 UAV/km^2	n_{LoS}, n_{NLoS}	2, 2.4
λ_u	0.4 UAV/km^2	f_c	38 GHz
C	1 MB	c	3×10^8 m/s
$\sigma_{i,u}^2$, $\sigma_{k,i}^2$	-95 dBm	ς_{LoS}, $\varsigma_{\mathrm{NLoS}}$	5.3, 5.27
$B_{i,u}^{(1)}$, $B_{k,i}^{(2)}$	1 MHz	a, b	11.9, 0.13
γ	3	α, β	2, 100
r_A	10 UAV/km^2	P_S, P_U	30dBm, 20dBm
d_0	5 m	h_0	1km

Proof: See Appendix A. ■

Thus, we can obtain the MFE of the flight policy when the initial state $X(0)$, $x(0)$ and initial mean-field distribution $m(0)$ are given. The optimal flight algorithm is shown in **Algorithm 1**. Specifically, the value $V(t)$ can be obtained by calculating the backward equaition, HJB equation in (24) with the mean field $m(t)$, which is obtained by calculate the forward equation, FPK equation in (22). Thus, this iteration terminates as the MFE is acquired, and then the optimal control $v^*(t)$ is obtained by using Proposition 1.

VI. Numerical Results and Discussion

In this section, the numerical results show the superiority of the proposed algorithm. Firstly, we give the mainly parameters settings. Then, the numerical results are shown to verify the effectiveness of the proposed algorithm. In this paper, we assume that users are distributed in a circular area with a radius of $r_A = 10$km, and the area of the circle is recorded as A. In the total area A, the UAVs and the users are distributed in PPP, which means the initial state distribution of UAVs m_0 is known. Considering the user's low speed and short optimization period, all users are regarded as fixed in each optimization period. Moreover, we assume that we have three SBSs normally distributed in the total area A to covering more areas, which is shown in Fig. 2 .

A. Parameter setting

For comparison, the parameter settings are shown in Table II. Besides, we set that the constant transmit power of SBSs and UAVs are $P_S = 30$dBm and $P_U = 20$dBm, respectively. The free-space reference distance d_0 is 5m. The initial altitude h_0 of each UAV is set to 1km. Then we give the numerical results and the analyses as follows.

B. Distribution of the mass and the mean field

Firstly, we show the variation of the distribution of the N UAVs in the total area A. Fig. 2 and Fig. 3 show the initial and the final distribution of U users, N UAVs and K SBSs, respectively, which demonstrates the

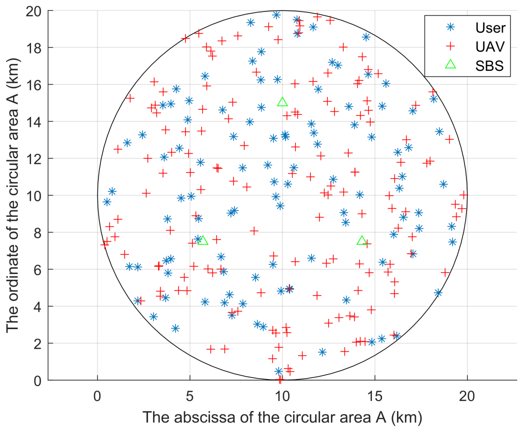


Fig. 2: The distribution of U users, N UAVs and K SBSs at $t = 0$

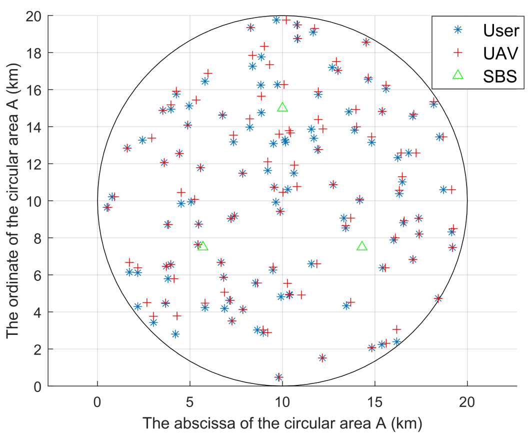


Fig. 3: The distribution of U users, N UAVs and K SBSs at $t = 23$

flight changes of the mass. Specifically, Fig. 2 represents the total area A at time $t = 0$, which is a circular area with the radius as 10km. In this area A, there are three SBSs normally distributed, which are represented as the green triangles. Moreover, the blue starts and the red plus signs represent the U users and the N UAVs, respectively. At the initial time, the masses of UAVs and users are distributed randomly following the independent and homogeneous PPPs. To illustrate the evolution of the mass under the control obtained by **Algorithm 1**, we provide the distribution of the mass DSCs with the users and SBSs fixed at the final time $t = 23$, as shown in Fig. 3. In Fig. 3, each UAV has reached the optimal location to facilitate matching user downloading the contents. The optimality can be verified as follows.

As shown in Fig. 4, we can observe that the distribution of the mean field which is regarded to the time and the distances between the UAVs and matching users. Since

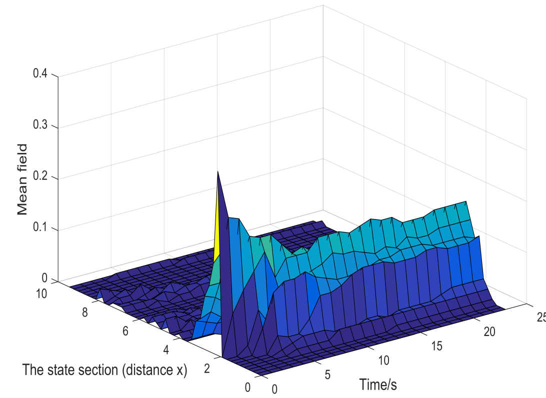


Fig. 4: The distribution of the mean field

the user's content request area is bu, where the users and the UAVs are initially randomly distributed. The distance between the UAVs and the matching users are between 2 km and 3 km. Therefore, a stable trend of the mean field over time can be observed.

C. Performance analysis

In this paper, the optimization objective is the delay of the users downloading with the fewer flight energy consumption. Therefore, Fig. 5 and Fig. 6 demonstrate the average delay of all pairs of the UAVs-users and the average flight energy consumption of all alive UAVs, respectively. Specifically, the average download delay of all users based on proposed algorithm is demonstrated in Fig. 5, where the content downloaded by each UAV depends on the contents popularity of users in the covered area b_i. Meanwhile, the comparison between proposed algorithm and the other two flight manners of UAVs: hovering at initial position, approaching the matching users is implemented, which is shown as Fig. 5. The delay reduction of these two flight manners is obviously smaller than that of the proposed algorithm. In addition, it is unrealistic that UAVs can access users indefinitely. Moreover, we show the delay performance with the perfect information, which can be called the theoretical performance because each agent knows the information of others based on the local information. The red square curve in Fig. 5 shows the delay performance of the perfect information, which is similar to the proposed algorithm.

In Fig. 6, we give the average flight energy consumption of N UAVs over time, where the hovering energy consumption is considered as a constant disregarded in this paper. That means we only foucs on the moving energy consumption of flight. Meanwhile, those three flight strategies and the perfect information which is mentioned above are also demonstrated. In Fig. 6, the flight energy consumption by using the proposed algorithm

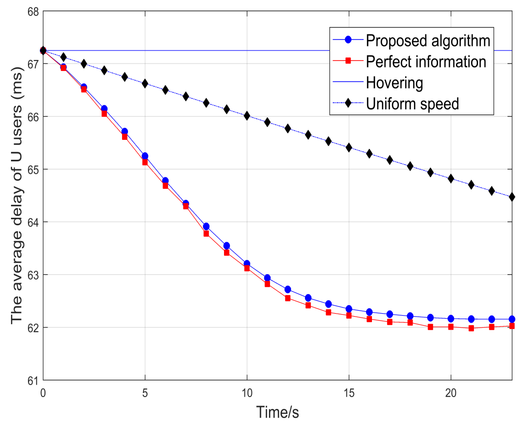


Fig. 5: The average delay of U users as the time varies

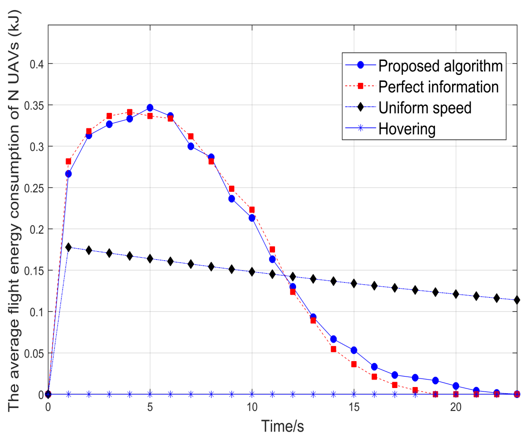


Fig. 6: The average flight energy consumption of N UAVs as the time varies

is represented by the blue circular curve. Moreover, each UAV find the its optimal position at the initial time, which results in the higher flight energy consumption. But after $t = 5$, the flight energy consumption of the proposed algorithm decreases gradually and gets lower than the uniform flight manner after $t = 12$, which is shown as the black diamond curve. Finally, those N UAVs arrive the optimal positions, which makes the average flight energy consumption converges to zero. Moreover, the red square curve show the flight energy consumption with the perfect information, which is similar with the proposed algorithm. In addition, although the lower flight energy consumption of the hovering flight manner, which is shown as the blue start curve, the delay performance is not improved as shown in Fig. 5.

Fig. 7 shows the average flight energy efficiency (EE) of the proposed algorithm and the uniform speed flight manner. Here, the flight EE is define as the delay reduction normalized by the flight energy consumption in an unit time. The blue and red curve in Fig. 7 represent the flight

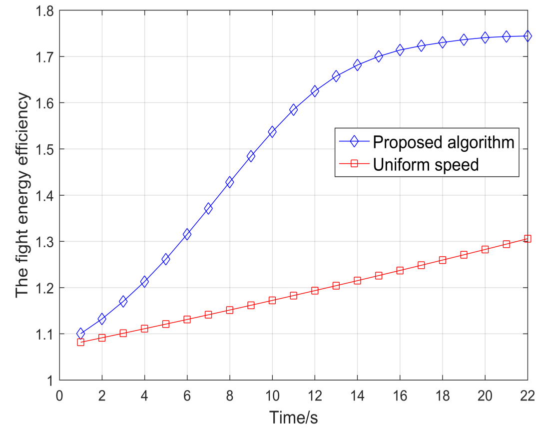


Fig. 7: The flight EE of N UAVs as the time varies

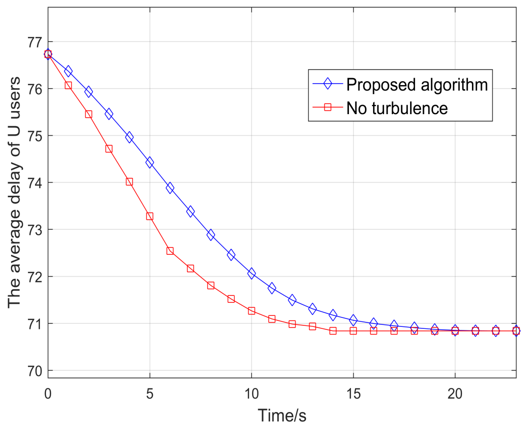
EE by using the proposed algorithm and the uniform speed flight manner, respectively. Clearly, the proposed algorithm has a higher flight EE, which means the proposed algorithm has the larger delay reduction with the similar energy consumption. In this numerical result, the flight EE at the initial time $t = 0$ does not be shown due to the fact that the flight energy consumption at the initial time is zero. In summary, Fig. 5, Fig. 6, and Fig. 7 demonstrate the effectiveness of the proposed algorithm.

D. Robust performance analysis

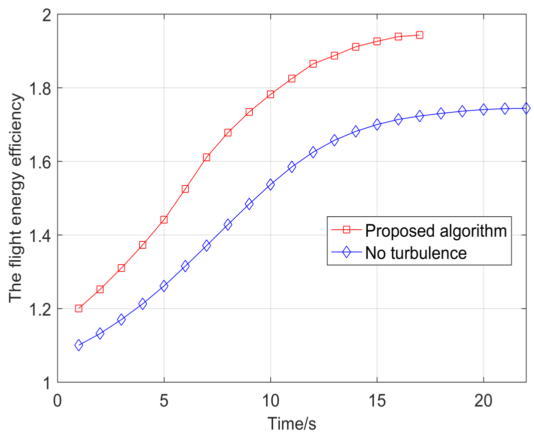
In this subsection, we verify the robust performance of the proposed algorithm by showing the comparison of the delay and the EE variations in the presence of the atmospheric turbulence or not. In this paper, we set the distribution of the atmospheric turbulence following the Gausssian. The numerical results in Figs. 8. (a) and (b) show the average delay and the EE changes, where two red square curves represent the condition without the atmospheric turbulence, and two blue rhombic curves are normal condition. We still can not give the EE at the time $t = 0$ and after $t = 17$ of the red square curve in Fig. 8 (b) because the energy consumption is zero at those time. We can see that the system can reach the MFE earlier without the atmospheric turbulence. However, when the system reaches the MFE with the atmospheric turbulence, the delay and the EE are almost the same as the performance without the atmospheric turbulence, which lie in the robust performance of the proposed algorithm.

VII. Conclusion

This paper builds an ECN with a massive number of UAVs and investigates a delay optimization algorithm. The UAVs are considered as air cache units which can adjust the aerial position according to their own flight energy consumption and the popularity of downloaded content, thereby reducing the download delay of matching users.



(a) The comparison of the average delay



(b) The comparison of the EE

Fig. 8: The comparison of the robust performance

Meanwhile, we consider the effect of the atmospheric turbulence, which is the significant factor in UAV flight. Therefore, we formulate a robust MFG framework to solve this dynamic flight problem, where the atmospheric turbulence is described by the disturbance term in the drift function. The numerical results show the position variations of all UAVs. Furthermore, we show the average download delay of the ground users with the limited flight energy consumption of matching UAVs during $t \in [0, T]$. It can be seen that in a limited time, the flight energy consumption have reached the equilibrium and so do the average download delay. Compared with the other two flight strategies, the proposed algorithm can have a larger delay reduction, while consuming the same energy. In addition, we discuss the robust performance of the proposed method. With the atmospheric turbulence, the proposed method achieves the similar performance to the case without the atmospheric turbulence.

APPENDIX

A. Proof of Proposition 1

We firstly calculate the worst case of the disturbance by getting the critical point as

$$\begin{aligned}&\frac{\partial}{\partial a(t)}\Big[J\left(v(t), I\left(t, m^{*}\left(t, X_{j}(t), x(t)\right)\right)\right)-\gamma^{2} a^{2}(t)\\&+\frac{\sigma_{2}^{2}}{2} \partial_{x x}^{2} v(t)+(x(t)-v(t)+a(t)) \partial_{x} V(t)\\&+\frac{\sigma_{1}^{2}}{2} \partial_{X X}^{2} V(t)+\lambda\left(\operatorname{Pr}_{j}-X_{j}(t)\right) \partial_{X} V(t)\Big]=0.\end{aligned} \tag{27}$$

The strict convexity of the disturbance term $a(t)$ and the smoothness of the drift function in the dynamic equations (18b) and (18c) make the unique solution $a^{*}(t)$ of (27) as follows

$$a^{*}(t)=\frac{\partial_{x} V(t)}{2 \gamma^{2}}, \tag{28}$$

where the value function $V(t)$ satisfies the HJB equation (24). Then the worst case robust HJB (24) is given by

$$\begin{aligned}&\partial_{t} V(t)+\inf_{v(t)}\Big[J\left(v(t), I\left(t, m^{*}\left(t, X_{j}(t), x(t)\right)\right)\right)\\&+\frac{\sigma_{2}^{2}}{2} \partial_{x x}^{2} V(t)+(x(t)-v(t)) \partial_{x} V(t)+\frac{\partial_{x}^{2} V(t)}{4 \gamma^{2}}\\&+\frac{\sigma_{1}^{2}}{2} \partial_{X X}^{2} V(t)+\lambda\left(\operatorname{Pr}_{j}-X_{j}(t)\right) \partial_{X} V(t)\Big]=0.\end{aligned} \tag{29}$$

Therefore, due to the convexity of the HJB equation for control $u(t)$, the unique optimal solution $u^{*}(t)$ can be obtained in (26) by calculating the $\partial \inf(\cdot) / \partial v(t)=0$, $\inf(\cdot)$ represents the part in (29).

REFERENCES

[1] 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," *Wireless and Optical Commun. Conf. (WOCC)*, Beijing, China, May 2019.

[2] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-Dense Networks: A Survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2252-2545, Fourth quarter. 2016.

[3] X. Ge, H. Cheng, M. Guizani, and T. Han, "5G wireless backhaul networks: Challenges and research advances," *IEEE Netw.*, vol. 28, no. 6, pp. 6-11, Nov. 2014.

[4] X. Wang, M. Chen, T. Taleb, A. Ksentini, and V. Leungi, "Cache in the Air: Exploiting Content Caching and Delivery Techniques for 5G Systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 131-139, Feb. 2014.

[5] E. Batu, M. Bennis, M. Kountouris, and M. Debbah, "Cache-enabled small cell networks: Modeling and tradeoffs," *11th International Symposium on Wireless Commun. Systems (ISWCS)*, Barcelona, Spain, Aug. 2014, pp. 649-653.

[6] K. Poularakis, G. Iosifidis, V. Sourlas, and L. Tassiulas, "Exploiting Caching and Multicast for 5G Wireless Networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2995-3007, Apr. 2016.

[7] H. Weng, L. Li, Q. Cheng, W. Chen and Z. Han, "Content Caching Policy Based on GAN and Distributional Reinforcement Learning," *IEEE Int. Conf. Commun. (ICC)*, Dublin, Ireland, Jun. 2020.

[8] J. Qiao, Y. He, and S. Shen, "Proactive caching for mobile video streaming in millimeter wave 5G networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 7187-7198, Oct. 2016.

[9] S. Wang, X. Zhang, Y. Zhang, L. Wang, J. Yang, and W. Wang, "A Survey on Mobile Edge Networks: Convergence of Computing, Caching and Communications," *IEEE Access*, vol. 5, pp. 6757-6779, Mar. 2017.

[10] B. Zhou, Y. Cui, and M. Tao, "Optimal dynamic multicast scheduling for cache-enabled content-centric wireless networks," *IEEE Trans. Commun.*, vol. 65, no. 7, pp. 2956-2970, Jul. 2017.

[11] J. Yin, L. Li, Y. Xu, W. Liang, H. Zhang, and Z. Han, "Joint Content Popularity Prediction and Content Delivery Policy for Cache-Enabled D2D Networks: A Deep Reinforcement Learning Approach," *IEEE Global Conf. on Sig. and Info. Process.(GlobalSIP),* Anaheim, CA, 2018, pp. 609-613.

[12] M. Chen, W. Saad, C. Yin, and M. Debbah, "Echo state networks for proactive caching in cloud-based radio access networks with mobile users," *IEEE Trans. on Wireless Commun.*, vol. 16, no. 6, Mar. 2017.

[13] J. Yin, L. Li, H. Zhang, X. Li, A. Gao, and Z. Han, "A prediction-based coordination caching scheme for content centric networking," *Wireless and Optical Commun. Conf. (WOCC),* Hualien, China, Apr. 2018.

[14] Q. Jia, R. Xie, T. Huang, J. Liu, and Y. Liu, "Efficient caching resource allocation for network slicing in 5G core network," *IET Commun.*, vol. 11, no. 18, pp. 2792-2799, 2017.

[15] S. Wang, X. Zhang, Y. Zhang, L. Wang, J. Yang, and W. Wang, "A Survey on Mobile Edge Networks: Convergence of Computing Caching and Communications," *IEEE Access*, vol. 5, pp. 6757-6779, 2017.

[16] K. Hamidouche, W. Saad, M. Debbah, and H. V. Poor, "Mean-Field Games for Distributed Caching in Ultra-Dense Small Cell Networks," *American Control Conf. (ACC),* Boston, MA, pp. 4699-4704, Jul. 2016.

[17] Z. Xiong, J. Zhao, Z. Yang, D. Niyato and J. Zhang, "Contract Design in Hierarchical Game for Sponsored Content Service Market," *IEEE Trans. on Mobile Comput.*, Apr. 2020. DOI: 10.1109/TMC.2020.2991060.

[18] Z. Xiong, S. Feng, D. Niyato, P. Wang, A. Leshem and Z. Han, "Joint Sponsored and Edge Caching Content Service Market: A Game-Theoretic Approach," IEEE Trans. on Wireless Commun., vol. 18, no. 2, pp. 1166-1181, Feb. 2019.

[19] H. Kim, J. Park, M. Bennis, S.-L. Kim, and M. Debbah, "Ultra-Dense Edge Caching under Spatio-Temporal Demand and Network Dynamics," *IEEE Int. Conf. on Commun. (ICC),* Paris, France, May. 2017.

[20] H. Kim, J. Park, M. Bennis, S. Kim, and M. Debbah, "Mean-Field Game Theoretic Edge Caching in Ultra-Dense Networks," *IEEE Trans. on Veh. Tech.*, vol. 69, no. 1, pp. 935-947, Jan. 2020.

[21] 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 Network," *IEEE Int. of Things,* vol. 7, no. 1, pp. 544 557, Jan. 2020.

[22] Z. Shi, L. Li, Y. Xu, X. Li, W. Chen, and Z. Han, "Content Caching Policy for 5G Network Based on Asynchronous Advantage Actor-Critic Method," *IEEE Global Commun. Conf. (GlobeCom),* Waikoloa, HI, Dec. 2019.

[23] N. Cheng, W. Xu, W. Shi, N. Lu, H. Zhou, and X. Shen, "Air-Ground Integrated Mobile Edge Networks:Architecture, Challenges, and Opportunities," *IEEE Commun. Mag.*, vol. 56, no. 8, pp. 26-32, Aug. 2018.

[24] F. Cheng, G. Gui, N. Zhao, F. R. Yu, Y. Chen, J. Tang and H. Sari, "Caching UAV Assisted Secure Transmission in Small-Cell Networks," *Int. Conf. on Comput., Netw. and Commun. (ICNC),* Maui, Hawaii, pp. 696-701, Mar. 2018.

[25] 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.

[26] 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.

[27] X. Xu, Y. Zeng, Y. L. Guan, and R. Zhang, "Overcoming Endurance Issue: UAV-Enabled Communications With Proactive Caching," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 6, pp. 1231-1244, Jun. 2018.

[28] Y. Sun, L. Li, Q. Cheng, D. Wang. W. Liang, and Z. Han, "Joint Trajectory and Power Optimization in Multi-Type UAVs Network with Mean Field Q-Learning," *IEEE Int. Conf. on Commun.Workshops (ICC Workshops)*, Dublin, Ireland, Jun. 2020.

[29] L. Li, H. Ren, Q. Cheng, K. Xue, M. Debbah, and Z. Han, "Millimeter-Wave Networking in the Sky: A Machine Learning and Mean Field Game Approach for Joint Beamforming and Beam-Steering," *IEEE Trans. on Wireless Commun.*", vol. 19, no. 10, pp. 6393-6408, Oct. 2020.

[30] Q. Cheng, L. Li, Y. Sun, W. Wang, X. Li, and Z. Han,, "Efficient Resource Allocation for NOMA-MEC System in Ultra-Dense Network: A Mean Field Game Approach," *IEEE Int. Conf. on Commun. Workshops (ICC Workshops)*, Dublin, Ireland, Jun. 2020.

[31] 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.

[32] Q. Cheng, L. Li, K. Xue, H. Ren, X. Li, W. Chen, and Z. Han, "Beam-steering Optimization in Multi-UAVs mmWave Networks: A Mean Field Game Approach," *11th Int. Conf. on Wireless Commun. and Sig. Process. (WCSP),* Xi'an, China, Oct. 2019.

[33] L. Li, Z. Zhang, K. Xue, M. Wang, M. Pan, and Z. Han, "AI-Aided Downlink Interference Control in Dense Interference-Aware Drone Small Cells Networks", *IEEE Access*, vol.8, pp. 15110-15122, Jan.2020.

[34] R. A. Banez, L. Li, C. Yang, L. Song, and Z. Han, "A Mean-Field-Type Game Approach to Computation Offloading in Mobile Edge Computing Networks," *IEEE Int. Conf. on Commun. (ICC),* Shanghai, China, May. 2019.

[35] 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," *15th Int. Wireless Commun. & Mobile Comput. Conf. (IWCMC),* Tangier, Morocco, Jun. 2019.

[36] K. Xue, Z. Zhang, L. Li, H. Zhang, X. Li, and Ang Gao, "Adaptive Coverage Solution in Multi-UAVs Emergency Communication System: A Discrete-Time Mean-Field Game," *Int. Wireless Commun. & Mobile Comput. Conf. (IWCMC),* Limassol, Cyprus, Jun. 2018, pp. 1059-1064.

[37] S. Srinivasa and M. Haenggi, "Distance distributions in finite uniformly random networks: Theory and applications," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 940-949, Feb. 2010.

[38] E. Batu, M. Bennis, and M. Debbah, "Living on the edge: The role of proactive caching in 5G wireless networks," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 82-89, Aug. 2014.

[39] T. Qiu, Z. Ge, S. Lee, J. Wang, Q. Zhao, and J. Xu, "Modeling Channel Popularity Dynamics in a Large IPTV System," *ACM SIGMETRICS,* Seattle, WA, pp. 276-286, Jun. 2009.

[40] T. L. Griffiths and Z. Ghahramani, "The Indian Buffet Process: An Introduction and Review," *J. Mach. Learn. Res.*, vol. 12, pp. 11851224, Jul. 2011.

[41] T. Alpcan, H. Boche, M. L. Honig, and H. V. Poor, *Mechanisms and Games for Dynamic Spectrum Allocation-Reacting to the Interference Field. Cambridge, U.K.: Cambridge Univ. Press*, 2014.

[42] J. Lou and K. Zhang, "Design of active disturbance rejection controller for autonomous aerial refueling UAV," *IEEE Int. Conf. of IEEE Region 10 (TENCON),* Xi'an, China, pp. 1-4, Oct. 2013.

[43] 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.

[44] Y. Zeng and R. Zhang, "Energy-Efficient UAV Communication With Trajectory Optimization," *IEEE Trans. on Wireless Commun.*, vol. 16, no. 6, pp. 3747-3760, Jun. 2017.

[45] D. Yang, Q. Wu, Y. Zeng, and R. Zhang, "Energy Tradeoff in Ground-to-UAV Communication via Trajectory Design," *IEEE Trans. on Veh. Tech.*, vol. 67, no. 7, pp. 6721-6726, Jul. 2018.

[46] 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, Jan. 2018.

[47] D. Bauso, H. Tembine, and T. Baar, "Robust mean field games with application to production of an exhaustible resource," *IFAC Proc. Volumes*, vol. 45, no. 13, pp. 454-459, 2012.

[48] J. Moon and T. Baar, "Linear Quadratic Risk-Sensitive and Robust Mean Field Games," *IEEE Trans. on Autom. Control*, vol. 62, no. 3, pp. 1062-1077, Mar. 2017.

[49] M. Burger and J. M. Schulte, *Adjoint methods for Hamilton-Jacobi-Bellman equations, Mnster, Germany: Univ. Mnster*, 2010.

[50] B. Øksendal, *Stochastic Diferential Equations. Springer*, 2003.

Lixin Li (M'12-) 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 Post-Doctoral Fellow with NPU from 2008 to 2010. In 2017, He was a visiting scholar at the University of Houston, Texas. He is currently an Associate Professor in the School of Electronics and Information, NPU. He has authored or coauthored more than 150 peer-reviewed papers in many prestigious journals and conferences, and he holds 12 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.

Meng Wang is currently a master student under the supervision of Prof. Lixin Li with the School of Electronics and Information, Northwestern Polytechnical University, Xian, China. His research interests include unmanned aerial vehicle, game theory and wireless communications.

Kaiyuan Xue was a master student under the supervision of Prof. Lixin Li with the School of Electronics and Information, Northwestern Polytechnical University, Xian, China. His research interests include unmanned aerial vehicle, wireless caching and game theory in wireless communication networks.

Qianqian Cheng is currently a master student under the supervision of Prof. Lixin Li with the School of Electronics and Information, Northwestern Polytechnical University, Xian, China. Her research interests include unmanned aerial vehicle, mobile edge computing and reinforcement learning in wireless communication networks.

Dawei Wang (S14-M18) received the B.S degree from University of Jinan, China, in 2011 and the Ph.D. degree from Xi'an Jiaotong University, China in 2018. From 2016 to 2017, he was a Visiting Student with the School of Engineering, The University of British Columbia. He is currently an Associate Professor with the School of Electronics and Information, Northwestern Polythechnical University, Xi'an, China. He has served as Technical Program Committee (TPC) member for many International conferences, such as IEEE GLOBECOM, IEEE ICC, etc. His research interests include Physical-Layer Security, Cognitive Radio Networks, Cooperative Communication, Energy Harvesting, and Resource Allocation.

Wei Chen (S05-M07-SM13) received the B.S. and Ph.D. degrees (Hons.) from Tsinghua University in 2002 and 2007, respectively. Since 2007, he has been a Faculty Member with Tsinghua University, where he is currently a Tenured Full Professor, the Director of the Degree Office, and a University Council Member. During 2014-2016, he was the Deputy Head of the Department of Electronic Engineering in Tsinghua University. From 2005 to 2007, he was a Visiting Ph.D. Student with the Hong Kong University of Science and Technology. He visited the University of Southampton in 2010, Telecom Paris Tech in 2014, and Princeton University, Princeton, NJ, USA, in 2016. His research interests are in the areas of communication theory and stochastic optimization. He is a Cheung Kong Young Scholar and a member of the National Program for Special Support of Eminent Professionals, also known as 10,000 talent program. He is a standing committee member of All-China Youth Federation as well as, the secretary-general of its education board. He received the IEEE Marconi Prize Paper Award in 2009 and the IEEE Comsoc Asia Pacific Board Best Young Researcher Award in 2011. He is a recipient of the National May 1st Labor Medal and the China Youth May 4th Medal. He has also been supported by the National 973 Youth Project, the NSFC Excellent Young Investigator Project, the New Century Talent Program of the Ministry of Education, and the Beijing Nova Program. He serves as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS. He has served as a TPC Co-Chair for IEEE VTC-Spring in 2011 and a Symposium Co-Chair for IEEE ICC and Globecom.

Miao Pan (S07-M12-SM18) received his BSc degree in Electrical Engineering from Dalian University of Technology, China, in 2004, MASc degree in electrical and computer engineering from Beijing University of Posts and Telecommunications, China, in 2007 and Ph.D. degree in Electrical and Computer Engineering from the University of Florida in 2012, respectively. He is now an Associate Professor in the Department of Electrical and Computer Engineering at University of Houston. He was a recipient of NSF CAREER Award in 2014. His research interests include cybersecurity, big data privacy, deep learning privacy, cyber-physical systems, and cognitive radio networks. His work won IEEE TCGCC (Technical Committee on Green Communications and Computing) Best Conference Paper Awards 2019, and Best Paper Awards in ICC 2019, VTC 2018, Globecom 2017 and Globecom 2015, respectively. Dr. Pan is an Editor for IEEE Open Journal of Vehicular Technology and an Associate Editor for IEEE Internet of Things (IoT) Journal (Area 5: Artificial Intelligence for IoT), and used to be an Associate Editor for IEEE Internet of Things (IoT) Journal (Area 4: Services, Applications, and Other Topics for IoT) from 2015 to 2018. He has also been serving as a Technical Organizing Committee for several conferences such as TPC Co-Chair for Mobiquitous 2019, ACM WUWNet 2019. He is a member of AAAI, a member of ACM, and a senior member of IEEE.

Zhu Han (S01-M04-SM09-F14) 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 an R&D Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an assistant professor at Boise State University, Idaho. Currently, he is a John and Rebecca Moores Professor in the Electrical and Computer Engineering Department as well as in the Computer Science Department at the University of Houston, Texas, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul, South Korea His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. Dr. Han received an 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, IEEE Leonard G. Abraham Prize in the field of Communications Systems (best paper award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. Dr. Han was an IEEE Communications Society Distinguished Lecturer from 2015-2018, AAAS fellow since 2019 and ACM distinguished Member since 2019. Dr. Han is 1% highly cited researcher since 2017 according to Web of Science. Dr. Han is also the winner of 2021 IEEE Kiyo Tomiyasu Award, for outstanding early to mid-career contributions to technologies holding the promise of innovative applications, with the following citation: "for contributions to game theory and distributed management of autonomous communication networks."