



UAV-assisted wireless relay networks for mobile offloading and trajectory optimization

Guangsheng Feng¹ · Chengbo Wang¹ · Bingyang Li¹ · Haibin Lv¹ · Xiaoxiao Zhuang¹ · Hongwu Lv¹ · Huiqiang Wang¹ · Xianlang Hu²

Received: 30 December 2018 / Accepted: 17 July 2019 / Published online: 14 August 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

In some remote areas under extremely scarce computation and communication resources, unmanned aerial vehicles (UAVs) assisted wireless communications are quite attractive for the more widely communication coverage and powerful computation capacity in contrast to user mobile devices (MDs). In addition to providing relay communication capabilities, the UAV can not only perform user tasks locally, but also offload them to base station or cloudlet for computing, provided that the user tasks require more powerful computation capacities. Considering the fact that the mobility trajectory of UAV could exert a negative impact on mobile offloading, we investigate the mobile offloading problem with a comparative consideration of mobility trajectory, communication, and computation resource allocation, in which the UAV-assisted mobile offloading and trajectory optimization problem (UOTO) is formulated with the objective of maximizing minimum user utility. Unfortunately, the UOTO provided is proven to be non-convex and there is no effective way to solve it. To overcome the difficulties, we develop a near-optimal algorithm by integrating the nonlinear fractional programming (NFP) and successive convex approximation (SCA). The extensive experiments are conducted to illustrate the performance of the proposed scheme.

Keywords UAV · Relay · Traffic offloading · Computation offloading · Fairness

1 Introduction

With the enduring prosperity of wireless communication techniques, pervasive Wireless Wide Area Networks (WWANs), e.g., the LTE-A networks, the next generation

networks (5G) and beyond, have triggered the dramatic shift in people lifestyle and bring plentiful benefits to contemporaries. With WWANs, users can easily access remote Internet resources and enjoy services via Base Stations (BSs) anytime and anywhere, e.g., online shopping on the way to work, watching live broadcast videos during heavy traffic jam, playing interactive games on a high-speed train, to name a few. An irresistible tendency is that the contemporaries are heavily relying on WWANs to fulfill their transmission and computation requirements. Unfortunately, such a lifestyle shift could suffer from severe service degradation under massive transmission and computation scenarios, where huge volumes of mobile traffic are generated and aggregated at the side of network edge, exerting significant pressure on cellular BSs. Even more seriously, users may not be served properly when the BSs are heavily jammed, or even some of them are damaged by natural disasters, e.g., earthquakes or hurricanes.

To address this issue, mobile offloading [2], including traffic offloading and computation offloading, is considered to be of great significance, of which one traditional scheme

This article is part of the Topical Collection: *Special Issue on Networked Cyber-Physical Systems*

Guest Editors: Heng Zhang, Mohammed Chadli, Zhiguo Shi, Yanzheng Zhu, and Zhaojian Li

A preliminary version of this work has been published in the 2018 IEEE 15th International Conference on Mobile Ad Hoc and Sensor Systems(MASS) [1].

✉ Bingyang Li
jensen.hrbeu@gmail.com

Guangsheng Feng
fengguangsheng@hrbeu.edu.cn

¹ Harbin Engineering University, Harbin, 150001, China

² Jiangsu Institute of Automation, Lianyungang, 222000, China

is that the overloaded networks offload partial tasks to small-cell networks, e.g., WiFi access points and femtocells. For example, Losifidis et al. [3] assumed that the cellular networks can competitively lease residential users as access points for data offloading, while Al-Kanj et al. [4] proposed that the mobile device holders can actively form device-to-device (D2D) networks to offload partial user traffic for the overloaded cellular networks. In addition, Wang et al. [5] proposed a novel architecture that reduced resource consumption and ensured the extensibility of trust evaluation mechanism. The above existing work can be of feasibility and practicality only when the cooperative partners are densely deployed. Unfortunately, such assumptions cannot always hold, especially when partial cellular networks in populated urban areas are severely jammed by overloaded user traffic requirements or damaged by natural disasters. On the other hand, in remote areas out scope of the BS transmission coverage, users' demands can hardly be fulfilled without the direct connections between users and BS. Considering this fact, leveraging Unmanned Aerial Vehicles (UAVs) to assist traffic or computation offloading in cellular networks is proposed [6–9], in which the UAVs serve as moving cloudlets or relays to accommodate user demands thanks to the convenient mobility in 3-D spaces [10–15]. With this offloading paradigm, the UAV can not only provide a much larger transmission coverage during flight, but also maintain a reliable connection with the cellular networks. As such, the user requirements can be satisfied to a certain extent. The existing work can achieve a better performance when the user demands are homogeneous, i.e., all users require only the same one of services, namely, transmission offloading or computation offloading, rather than both of them.

In this paper, we consider a more complicated and practical scenario, where users are heterogeneous, and each user might have different offloading requirement. Different from the data and computation co-offloading in [16], we use the UAV as a moving cloudlet to provide concurrent serve and relay when the heterogeneous users are not covered by the cellular BS. It is a necessity that the UAV has to fulfill its offloading tasks before a hard deadline during flying to the destination, and new challenges to the UAV-assisted traffic and computation offloading (UOTO) problem are brought in three folds. Firstly, the resource competition in the UOTO problem, including bandwidth and energy, is unavoidable. In comparison to the existing offloading optimization [17, 18], non-linear couple between UAV's transmit power and bandwidth allocation is incurred, which leads to the UOTO problem difficult to solve optimally. Secondly, different from the assumption of sufficient UAV computation capacity in existing work [19–23], the UAV

usually contributes to the computation offloading task with a constrained on-board computation resource during serving as a moving cloudlet. In other words, the user computation tasks could be performed at the locality, the UAV, or the BS server. Therefore, the UAV has to make an instant decision when multiple user tasks come. In addition, the UAV's trajectory has to be considered since it can affect the resource allocation manner including energy and computation resources. Compared with the existing work [24–26], we not only consider the power allocation and trajectory optimization, also consider the heterogeneity of users need. The above three folds motivate us to optimize the UOTO problem by jointly considering the bandwidth allocation, partial offloading decision and UAV's trajectory planning. The UOTO problem considered in this paper is more practical since most of users might require traffic offloading and computation offloading concurrently, rather than only one of them. The optimal solution not only relies on the competition in resource competition between traffic offloading and computation offloading but also on the UAV's trajectory planning. In consideration of the fairness among users, the objective of the UOTO problem is to maximize the minimum user utility. In the formulated problem, there exists non-linear couples among variables and it is an NP-hard problem. To deal with it, we propose a three-stage alternative optimization (TSAO) algorithm to solve it. The main contributions of this paper are summarized as follows.

- Considering the heterogeneous demands for mobile offloading, the UAV-assisted mobile offloading and trajectory optimization problem (UOTO) is formulated with the objective of maximizing minimum user utility. And it is a non-convex optimization problem.
- We carefully investigate the structure of the UOTO problem, which is decomposed three sub-problems, namely, bandwidth resource allocation, partial offloading decision, and UAV trajectory optimization. The closed form of optimal solution for bandwidth resource allocation is obtained by its dual multipliers.
- We develop a three-stage alternative optimization algorithm to solve it iteratively by leveraging the nonlinear fractional programming (NFP) and successive convex approximation (SCA) methods.

The rest of this work is as follows. Section II introduces the system model and presents the formulation of UOTO problem. In Section III, we analyze the property of the problem and propose a three-stage alternative optimization algorithm to solve it. Then, numerical results are given in Section IV. Finally, in Section V, the conclusions are drawn.

2 System model

Considering a mobile offloading scenario with a three-dimensional (3D) Euclidean coordinate, there are N heterogeneous users with fixed locations on the ground. The users might have traffic demand, computation demand, or both of them. In the scenario, although the cellular BS server can provide a more powerful transmission and computing capability, the users cannot connect to the cellular networks directly due to the existing obstacles blocking the transmission between the users and cellular BS, or the users are not in the coverage of the cellular BS. One UAV, equipped with on-board communication circuits and computing processors while constrained energy and computation resources, can provide transmission and computation services for the covered users during floating from one position to another. Assume that there is a LOS transmission condition between the UAV and BS, i.e., the UAV can serve as an aerial platform for the covered users and meanwhile maintain a reliable connection with the BS. The considered system is time slotted with a duration T , which is divided into M identical slots.

In the system model, three types of users, i.e., *users with computation offloading demand*, *users with traffic offloading demand*, and *users with both offloading commands* are considered. Different from our previous work [1], when the UAV faces some resource-hungry computation tasks, the UAV might transmit partial tasks to the BS server for the sake of saving resources. For the first

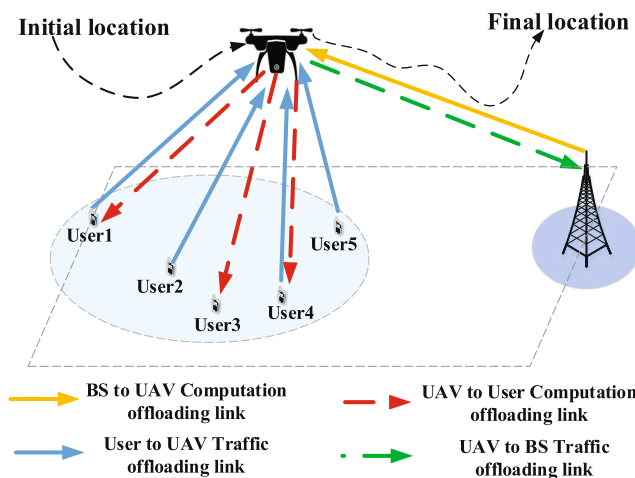


Fig. 1 A UAV-assisted wireless relay networks for mobile offloading scenario for heterogeneous users, where User 2, User 3 and User 5 require only traffic or computation offloading, while others have both demands. For the users with traffic demand, the UAV will request the relative contents from the BS and then transmit to the users. For the users with computation demand, the UAV and BS server will jointly process the tasks. UAV has a fixed initial and final location [27]

user type, the UAV functions as a mobile cloudlet to provide computation services for this type of users. For the second user type, the UAV is employed as a relay to download the traffic from BS and then transmit to the kind of users. For the third user type, the UAV will perform traffic and computation offloading for users concurrently. The example of scenario is illustrated in Fig. 1.

The computation tasks of users are partitioned into two parts, one is completed by the UAV and the other is transmitted to the BS server to perform, which is also called *partial computation offloading mode*.

The notations used in this paper are summarized in Table 1. Denote the set of UEs as \mathcal{N} and $\mathbf{L}_i = [x_i^{(1)}, y_i^{(1)}, z_i^{(1)}]$ as the coordinate of user i ($1 \leq i \leq N$), which can be obtained by the built-in GPS and other sensors [28]. We assume that all the users stand on a flat ground (i.e., $z_i^{(1)}$ is a constant) and the UAV flies at a fixed height H , which is the minimum altitude that is appropriate to the terrain. The wireless communication channels between MDs and UAV are dominated by line-of-sight links. Denote \mathbf{q}_{st} and \mathbf{q}_{end} as UAV's the initial location and destination location, respectively. Considering the energy consumption, the UAV's flight speed has an upper limit, denoted by v_{max} [28]. According to the slotted system model, M can be sufficiently large, such that the slot duration to be sufficiently small. Therefore, the location of UAV within each slot can be approximately constant [23, 27]. Noted that there are M sets of variables, which means the variables with different slots are irrelevant. Denote $\mathbf{q}(t) = [x^{(c)}(t), y^{(c)}(t), H]$ as UAV's location within slot t , where $x^{(c)}(t)$ and $y^{(c)}(t)$ represent the horizontal and vertical coordinates, respectively. The location of BS is denoted as $\mathbf{S} = [x^{(s)}, y^{(s)}, H^{(s)}]$, where $H^{(s)}$ denotes the altitude of BS.

Assume the bandwidth of downlink and uplink are both B MHz. And time division multiple access (TDMA) scheme is used in downlink and uplink transmission, respectively. Let $\alpha_i^{(d)}(t)$ and $\alpha_i^{(u)}(t)$ represent the percentage of downlink and uplink communication duration allocated for user i , respectively. Obviously, for the users who have upload demands, the sum of their allocated upload proportion cannot exceed 1. The users who have download demands are similar. The constraints are given as

$$\sum_{i=0}^N \alpha_i^{(u)}(t) \theta_i^{(u)} \leq 1, \quad (1)$$

and

$$\sum_{i=0}^N \alpha_i^{(d)}(t) \theta_i^{(d)} \leq 1, \quad (2)$$

Table 1 Parameter definitions

| Parameters | Definitions |
|--|---|
| $x_i^{(l)}, y_i^{(l)}$ | The horizontal and vertical coordinates of user i , respectively |
| $x^{(c)}(t), y^{(c)}(t)$ | The horizontal and vertical coordinates of UAV during slot t , respectively |
| $x^{(s)}, y^{(s)}$ | The horizontal and vertical coordinates of BS, respectively |
| $H, H^{(s)}$ | The fixed altitude of UAV and BS, respectively |
| \mathbf{S} | The location of BS |
| $\mathbf{q}_{st}, \mathbf{q}_{end}$ | The initial and destination location of UAV, respectively |
| $\mathbf{q}(t)$ | The location of UAV in slot t |
| $\theta_i^{(u)}, \theta_i^{(d)}$ | Whether user i has upload and download demands or not, respectively |
| $\alpha_i^{(u)}(t), \alpha_i^{(d)}(t)$ | The proportion of upload and download duration allocated to user i in slot t , respectively |
| δ | The channel power gain at a reference distance equals to 1m |
| $R_i^{(u)}(t), R_i^{(d)}(t)$ | The upload and download rate of use i during slot t , respectively |
| $P_i^u(t)$ | The transmit power of use i during slot t |
| $P_i^d(t)$ | The transmit power that UAV allocates to user i during slot t |
| $D_i^{(d)}(t)$ | The download traffic size of user i during slot t |
| $R_i^{(s)}$ | The upload rate of UAV to BS during slot t |
| $v(t)$ | The flight speed of UAV during slot t |
| $v_{max}^{(uav)}$ | The maximum flight speed of UAV |
| $P_{max}^{(uav)}, P_{i,max}^u$ | UAV’s and user i ’s maximum transmit power, respectively |
| B_s | The bandwidth between UAV and BS |
| p^b | The transmit power of BS |
| $R_{i,min}^d(t)$ | The minimum downlink capacity requirement of user i |
| γ_i | Number of CPU cycles per input bit of user i needed for computing |
| f_i^c | the CPU frequency that UAV allocates for user i |
| $\eta_i(t)$ | The proportion of computation task that user i allocates to the UAV |
| $T_i^{(uu)}(t), T_i^{(uc)}(t)$ | The computation delay of user i on the UAV and central cloud within slot t , respectively |
| $T_i^{(us)}(t)$ | The delay that UAV offloads user i ’s partial task to the BS within slot t |
| $T_i^{(ucd)}(t)$ | The computation delay of user i within slot t |
| $T_i^{(ud)}(t)$ | The delay of user i within slot t |
| U_i | The utility of user i |
| $\tau_i^{(d)}, \tau_i^{(t)}$ | The weights of computation size and delay of user i |

where $\theta_i^{(d)}$ and $\theta_i^{(u)}$ represent user i ’s traffic and computation offloading demands, respectively. And user i has offloading demands if it equals to 1, otherwise not.

The distance between user i and UAV and the distance between BS and UAV within slot t are respectively expressed as

$$d_i(t) = \sqrt{H^2 + \|\mathbf{q}(t) - \mathbf{l}(t)\|^2}, \tag{3}$$

$$d_s(t) = \sqrt{(H^{(s)})^2 + \|\mathbf{q}(t) - \mathbf{S}\|^2}. \tag{4}$$

According to [27, 29], the time-varying communication channel between user i and UAV and the time-varying communication channel between BS and UAV follow the free-space path loss model

$$h_i(t) = \delta d_i^{-2}(t), \tag{5}$$

$$h_s(t) = \delta d_s^{-2}(t), \tag{6}$$

where δ denotes the communication channel gain at the reference distance equals to 1 m.

According to the standard information theory [30], user i ’s transmission rate in uplink within slot t is express as

$$R_i^{(u)}(t) = \theta_i^{(u)} B \log_2(1 + \frac{P_i^{(u)}(t)h_i(t)}{N_0}), \tag{7}$$

where $p_i^{(u)}(t)$ represents user i ’s transmit power and N_0 represents the gaussian noise. Therefore, the traffic volume that user i uploaded is given by

$$D_i^{(u)}(t) = \frac{T}{M} \alpha_i^{(u)}(t) R_i^{(u)}(t). \tag{8}$$

Similarly, user i ’s transmission rate in downlink can be expressed as

$$R_i^{(d)}(t) = \alpha_i^{(d)}(t) \theta_i^{(d)} B \log_2(1 + \frac{p_i^{(d)}(t)h_i(t)}{N_0}), \tag{9}$$

where $p_i^{(d)}(t)$ represents the transmit power that UAV allocates for user i . Therefore, traffic volume that user i downloaded within slot t is given by

$$D_i^{(d)}(t) = \frac{T}{M} \alpha_i^{(d)}(t) R_i^{(d)}(t). \tag{10}$$

Similarly, the transmit rate of UAV to BS within slot t is given by

$$R_i^{(s)} = B_s \log_2 \left(1 + \frac{p_s(t) h_s(t)}{N_0} \right). \quad (11)$$

Similar to [31, 32], the computing time of UAV and the downloading time of computation result are neglected, which is because UAV's computing capability is significantly bigger than MDs' and the returned traffic volume is relatively small. Then, the flight speed of UAV within slot t can be expressed as

$$v(t) = \frac{\|\mathbf{q}(t) - \mathbf{q}(t-1)\|}{T/M}. \quad (12)$$

Because the flight speed of UAV is upper bounded to its maximum flight speed v_{\max} , one has

$$v(t) \leq v_{\max}. \quad (13)$$

Similarly, for UAV's and users' transmit power, one has

$$p_s + \sum_{i=1}^N \theta_i^{(d)} p_i^{(d)}(t) \leq p_{\max}^{(uav)}, \quad (14)$$

and

$$p_i^{(u)}(t) \leq p_{i,\max}^{(u)}, \quad (15)$$

where $p_{i,\max}^{(u)}$, $p_{\max}^{(uav)}$ and p_s represent the maximum transmit power of user i , UAV and the transmit power that UAV allocates for BS, respectively.

Due to users' and UAV's transmit power should be positive or equal to zero, one has

$$p_i^{(u)}(t) \geq 0, \quad (16)$$

$$p_s(t) \geq 0, \quad (17)$$

$$p_i^{(d)}(t) \geq 0. \quad (18)$$

Besides, the sum of the transmission rates allocated by the UAV to all users is less than or equal to the bandwidth allocated by the UAV to the BS, one has

$$\sum_{i=1}^N R_i^{(d)}(t) \leq B_s \log_2 \left(1 + \frac{p^{(b)} h_s(t)}{N_0} \right), \quad (19)$$

where B_s and $p^{(b)}$ denote the bandwidth that the UAV allocates to the BS and the transmit power of BS, respectively.

Each user has a downlink capacity requirement

$$R_i^{(d)}(t) \geq R_{i,\min}^{(d)}(t), \quad (20)$$

where $R_{i,\min}^{(d)}(t)$ denotes the minimum downlink capacity requirement of user i .

According to [19, 33], the computation time of user i on the UAV within slot t is given by

$$T_i^{(uu)}(t) = \frac{\eta_i(t) D_i^{(u)}(t) \gamma_i}{f_i^{(c)}}, \quad (21)$$

where γ_i and $f_i^{(c)}$ represent the number of CPU cycles per input bit needed for computing and the CPU frequency allocated for user i , respectively. $\eta_i(t)$ denotes the proportion of computation task that user i computed on the UAV within slot t .

The computation time of user i on the central cloud within slot t is expressed as

$$T_i^{(uc)}(t) = \frac{(1 - \eta_i(t)) D_i^{(u)}(t) \gamma_i}{f_i^{(s)}}. \quad (22)$$

The time that UAV offload user i 's partial task to the BS within slot t is given by

$$T_i^{(us)}(t) = \frac{(1 - \eta_i(t)) D_i^{(u)}(t)}{R_i^{(s)}}. \quad (23)$$

According to Eqs. 21–23, the computation delay of user i within slot t is expressed as

$$T_i^{(ucd)}(t) = \max\{T_i^{(uu)}(t), T_i^{(us)}(t) + T_i^{(uc)}(t)\}. \quad (24)$$

Therefore, the delay of user i within slot t is given by

$$T_i^{(ud)}(t) = \alpha_i^{(u)} \frac{T}{M} + T_i^{(ucd)}(t). \quad (25)$$

The objective is to maximize user utility [34, 35], which is indicated by the computation size minus the delay, and it is expressed as

$$U_i = \sum_{t=1}^M \left(\tau_i^{(d)} D_i^{(u)}(t) - \tau_i^{(t)} T_i^{(ud)}(t) \right), \quad (26)$$

where $\tau_i^{(d)}$ and $\tau_i^{(t)}$ denote the weights of computation size and delay of user i , respectively. It should be noted that $0 \leq \tau_i^{(d)}, \tau_i^{(t)} \leq 1$. Users can set different values of $\tau_i^{(d)}$ and $\tau_i^{(t)}$ for their demands. For example, when a user is sensitive to the delay (e.g., online game), the user can set $\tau_i^{(t)} = 1$ and $\tau_i^{(d)} = 0$. When user concerns more about the computation size, the value of $\tau_i^{(d)}$ can be assigned larger. In practice the proper weights can be calculated by the multi-attribute utility approach [36].

We take the Max-Min fairness scheme [37, 38] for users mobile offloading. We introduce a variable $\xi = \min_{i \in \mathcal{N}} U_i$ and add following constraint:

$$U_i \geq \xi. \quad (27)$$

Then, the objective is to maximize ξ . Therefore, we can obtain the UOTO optimization problem as **P0**.

$$\mathbf{P0} : \max_{\alpha^{(u)}, \alpha^{(d)}, \mathbf{p}^{(d)}, \mathbf{q}} \xi \quad (28a)$$

$$s.t. (1), (2), (13) - (20) \text{ and } (27) \quad (28b)$$

In problem **P0**, the objective function represents user utility. Equations 1 and 2 denote the uplink and downlink bandwidth allocation constraints, respectively. And Eqs. 13–14 denote the constraint of UAV's flight speed and transmit power. Equations 15–20 denote the download rate and transmit power constraints of users. Besides, Eq. 27 denotes the fairness constraint among users.

3 Problem analysis and algorithm design

Due to there exists nonlinear coupling in constraints (20) and (27), the problem **P0** is therefore non-convex. At present, there are no effective solutions to obtain the optimal solution. Therefore, instead of resorting to expensive global optimization methods, we propose the TSAO algorithm to decouple the nonlinear coupling variables and solve them iteratively. In detail, we separate the variables into several subsets of non-overlapping variables and then get three sub-problems. For each sub-problem, we replace the non-convex function with suitable convex approximation by leveraging SCA and NFP [39–42].

Stage I: For the given trajectory and variables $\eta_i(t)$, $f_i^{(c)}$, $p^{(s)}$, we have following bandwidth allocation and power optimization problem **P1** that related to users and UAV. However, due to the constraint (27) is non-convex, **P1** is a non-convex problem.

Let $\phi_i^{(u)}(t) = \alpha_i^{(u)}(t)p_i^{(u)}(t)$, $\phi_i^{(d)}(t) = \alpha_i^{(d)}(t)p_i^{(d)}(t)$, $D_i^{(u)}(t)$ is then transformed into $D_i^{(u1)}(t)$, which is given by

$$D_i^{(u1)}(t) = \frac{T}{M} \alpha_i^{(u)}(t) \theta_i^{(u)} B \log_2 \left(1 + \frac{\phi_i^{(u)}(t) h_s(t)}{\alpha_i^{(u)}(t) N_0} \right) \quad (29)$$

Besides, constraint (27) can be rewritten as

$$\sum_{t=1}^M (\tau_i^{(d)} D_i^{(u1)}(t) - \tau_i^{(t)} \left(\frac{T}{M} + \max \left\{ \frac{\varphi \eta_i(t) D_i^{(u1)}(t) \gamma_i}{f_i^{(c)}(t)}, \frac{(1-\eta_i(t)) D_i^{(u1)}(t)}{B_s \log_2(1+p_s(t)h_s(t)/N_0)} \right\} \right)) \geq \xi \quad (30)$$

Thereafter, we can identically transform (30) into following constraints:

$$\sum_{t=1}^M (\tau_i^{(d)} D_i^{(u1)}(t) - \tau_i^{(t)} \left(\frac{T}{M} + \frac{\varphi \eta_i(t) D_i^{(u1)}(t) \gamma_i}{f_i^{(c)}(t)} \right)) \geq \xi, \quad (31)$$

$$\sum_{t=1}^M (\tau_i^{(d)} D_i^{(u1)}(t) - \tau_i^{(t)} \left(\frac{(1-\eta_i(t)) D_i^{(u1)}(t)}{B_s \log_2(1+p_s(t)h_s(t)/N_0)} + \frac{T}{M} \right)) \geq \xi. \quad (32)$$

Therefore, **P0** is reformed as **P1**.

$$\mathbf{P1} : \max_{\alpha_i^{(u)}(t), \alpha_i^{(d)}(t), p_i^{(u)}(t), p_i^{(d)}(t)} \xi \quad (33a)$$

$$s.t. (1), (2), (14), (15), (16), (18), (31) \text{ and } (32) \quad (33b)$$

According to [43, 44], it can be proved that **P1** is a convex optimization problem. We can tackle it by adopting the classic algorithm (e.g. Gradient descent) [45–48], and the optimal solutions of $\alpha_{i,\text{opt}}^{(u)}(t)$, $\alpha_{i,\text{opt}}^{(d)}(t)$, $p_{i,\text{opt}}^{(d)}(t)$, $p_{i,\text{opt}}^{(u)}(t)$ will be obtained. Moreover, we have the following Theorem 1.

Theorem 1 For a given trajectory, the closed-form of the optimal solution can be directly obtained by the dual multipliers that associated with constraints in problem **P1**.

Proof Let $\lambda^{(uu)}, \lambda^{(ud)}, \lambda^{(v)}, \lambda^{(pdm)}, \lambda^{(pum)}, \lambda^{(sum)}, \lambda_i^{(u)}, \lambda^{(ps)}, \lambda_i^{(pd)}, \lambda^{(r)}, \lambda_i^{(rm)}, \lambda^{(e)}, \lambda_i^{(f1)}, \lambda_i^{(f2)}$ denote the dual multipliers w.r.t. constraints (13)–(20), (31) and (32). Thus, the general Lagrangian function of problem **P1** can be expressed as (34) (see next page).

$$\begin{aligned} L = & \xi + \sum_{t=1}^M \lambda^{(uu)}(t) \sum_{i=0}^N (\theta_i^{(u)} \alpha_i^{(u)}(t) - 1) + \sum_{i=0}^N \sum_{t=1}^M (\lambda_i^{(pd)} p_i^{(d)}(t)) + \sum_{i=0}^N \sum_{t=0}^M (\lambda_i^{(pum)}(t) (p_i^{(u)}(t) - p_{i,\text{max}}^{(u)})) \\ & + \sum_{i=0}^N (\lambda^{(sum)} (p_s - p_{i,\text{max}}^{(u)})) + \sum_{t=1}^M \lambda^{(ud)}(t) \sum_{i=0}^N (\theta_i^{(d)} \alpha_i^{(d)}(t) - 1) + \sum_{t=1}^M (\lambda^{(v)}(t) (v(t) - v_{\text{max}})) + \lambda^{(ps)} p_s(t) \\ & + \sum_{t=1}^M (\lambda^{(pdm)} (p_s(t) + \sum_{i=1}^N p_i^{(d)}(t) - p_{\text{max}}^{(d)})) + \sum_{i=0}^N \sum_{t=1}^M (\lambda_i^{(u)} p_i^{(u)}(t)) + \sum_{i=0}^N \sum_{t=1}^M (\lambda^{(r)} (R_i^{(d)}(t) - B_s \log_2(1 + \frac{p^{(b)} h_s(t)}{N_0}))) \\ & + \sum_{i=0}^N \sum_{t=1}^M (\lambda_i^{(f1)} (\tau_i^{(d)} D_i^{(u)}(t) - \tau_i^{(t)} (\frac{T}{M} + \frac{\eta_i(t) D_i^{(u)}(t) \gamma_i}{f_i^{(c)}(t)} - \xi)) + \sum_{i=0}^N \sum_{t=1}^M (\lambda_i^{(rm)} (R_i^{(d)}(t) - R_{i,\text{min}}^{(d)}(t))) \\ & + \sum_{i=0}^N \sum_{t=1}^M (\lambda_i^{(f2)} (\tau_i^{(d)} D_i^{(u)}(t) - \tau_i^{(t)} (\frac{T}{M} + \frac{(1-\eta_i(t)) D_i^{(u)}(t)}{B_s \log_2(1+p_s(t)h_s(t)/N_0)} - \xi)). \end{aligned} \quad (34)$$

Therefore, the general Lagrangian dual function of problem **P1** is given by

$$\kappa(\lambda^{(uu)}, \lambda^{(ud)}, \lambda^{(v)}, \lambda^{(pdm)}, \lambda_i^{(pum)}, \lambda^{(sum)}, \lambda_i^{(u)}, \lambda^{(ps)}, \lambda_i^{(pd)}, \lambda^{(r)}, \lambda_i^{(rm)}, \lambda^{(e)}, \lambda_i^{(f1)}, \lambda_i^{(f2)}) = \max L. \tag{35}$$

Then, the solution of **P1** can be obtained by solving its dual problem, which is expressed as

$$\min \kappa(\lambda^{(uu)}, \lambda^{(ud)}, \lambda^{(v)}, \lambda^{(pdm)}, \lambda_i^{(pum)}, \lambda^{(sum)}, \lambda_i^{(u)}, \lambda^{(ps)}, \lambda_i^{(pd)}, \lambda^{(r)}, \lambda_i^{(rm)}, \lambda_i^{(dm)}, \lambda^{(e)}, \lambda_i^{(f1)}, \lambda_i^{(f2)}) \tag{36}$$

It can be seen that the dual problem of **P1** can be divided into N independent convex optimization problems. Thus, by getting the derivation of each sub-problem with respect to variables $\alpha_i^{(u)}(t), \alpha_i^{(d)}(t), p_i^{(u)}(t), p_i^{(d)}(t)$, we can obtain the closed-form of the optimal solution. □

Stage II: After **P1** is solved, the optimal values of $p_{i,opt}^{(d)}(t), p_{i,opt}^{(u)}(t), \alpha_{i,opt}^{(u)}(t), \alpha_{i,opt}^{(d)}(t)$ under some given variables can be obtained. Then, the problem related to computation can be obtained as **P2** for given variables $\alpha_{i,opt}^{(u)}(t), \alpha_{i,opt}^{(d)}(t), p_{i,opt}^{(d)}(t), p_{i,opt}^{(u)}(t)$ and $q(t)$. **P2** aims at solving the proportion of partial offloading and power optimization related to UAV and BS, which is given as

$$\mathbf{P2} : \max_{\eta_i(t), f_i^{(c)}(t), p_s(t)} \xi \tag{37a}$$

$$s.t. (14), (17), (31) \text{ and } (32) \tag{37b}$$

Due to the fraction in Eq. 32, problem **P2** is thus a nonlinear fraction optimization problem, which is hard to solve. We have following lemma by the reformulation of **P2**.

Lemma 1 *problem P2 can be transformed into an equivalent convex optimization.*

Proof By introducing some additional variables to replace the fraction expressions, we have the following equations:

$$\mu_i^{(c)}(t) = \frac{\eta_i(t)}{f_i^{(c)}(t)}, \tag{38}$$

$$\mu_i^{(s)}(t) = \frac{\eta_i(t)}{\log_2(1 + p_s(t)h_s(t)/N_0)}, \tag{39}$$

$$v_s(t) = \frac{1}{\log_2(1 + p_s(t)h_s(t)/N_0)}. \tag{40}$$

Then, constraints (31) and (32) are transformed as follows:

$$\sum_{i=1}^M (\tau_i^{(d)} D_i^{(u)}(t) - \tau_i^{(t)} (\frac{T}{M} + \varphi D_i^{(u)}(t) \gamma_i \mu_i^{(c)}(t))) \geq \xi, \tag{41}$$

$$\sum_{i=1}^M (\tau_i^{(d)} D_i^{(u)}(t) - \tau_i^{(t)} (\frac{T}{M} + (v_s - \mu_i^{(s)}(t)) D_i^{(u)}(t) B_s)) \geq \xi. \tag{42}$$

Besides, in order to transform **P2** into a convex optimization problem, two following constraints should be satisfied [49]:

$$\mu_i^{(c)}(t) \leq \frac{1}{f_i^{(c)}(t)}, \tag{43}$$

$$v_s h_s(t) \log_2 \left(1 + \frac{\mu_i^{(s)}(t)}{v_s h_s(t) N_0} \right) \leq 1. \tag{44}$$

Therefore, we have following convex problem **P3**:

$$\mathbf{P3} : \max_{\mu_i^{(c)}(t), \mu_i^{(s)}(t), v_s(t)} \xi \tag{45a}$$

$$s.t. (14), (17), (41) - (44) \tag{45b}$$

□

Therefore, problem **P2** is transformed into the convex problem **P3**, which can be tackled by leveraging the classic optimization algorithm, and the corresponding optimal solutions can be then obtained.

Stage III: When **P1** and **P3** are solved, the optimal variables expect UAV’s trajectory are obtained. Then, UAV’s trajectory can be obtained by following optimization problem **P4**.

$$\mathbf{P4} : \max_{\mathbf{q}(t)} \xi \tag{46a}$$

$$s.t. (19), (13), (31) \text{ and } (32) \tag{46b}$$

Because (32) is non-convex, **P4** is a non-convex problem. In Eq. 32, variables $p_i^{(u)}$ and $q(t)$ are nonlinear coupled in the form of bilinear expression [50, 51]. We leverage the SCA method to tackle it, which decouples the variables and approximate it to a convex problem. Thereafter, by relaxing the non-convex expressions with Taylor formula, Theorem 2 can be obtained.

Theorem 2 *For any given feasible UAV trajectory $\mathbf{q}_0(t)$, following inequalities can hold:*

$$\begin{aligned} & \log_2 \left(1 + \frac{\delta p_i^{(u)}(t)}{N_0(H^2 + \|\mathbf{q}(t) - \mathbf{I}(t)\|^2)} \right) \\ & \geq R^{(u)}(t) = \log_2 \left(1 + \frac{\delta p_i^{(u)}(t)}{N_0(H^2 + \|\mathbf{q}_0(t) - \mathbf{I}(t)\|^2)} \right) - \\ & \left(\frac{\log_2(e) p_i^{(u)}(t) (\|\mathbf{q}(t)\|^2 - \|\mathbf{q}_0(t)\|^2)}{(N_0 H^2 + \delta p_i^{(u)}(t) + N_0 \|\mathbf{q}_0(t)\|^2)(H^2 + \|\mathbf{q}_0(t)\|^2)} \right), \end{aligned} \tag{47}$$

$$\begin{aligned} & \log_2 \left(1 + \frac{\delta p_i^{(d)}(t)}{N_0(H^2 + \|\mathbf{q}(t) - \mathbf{I}(t)\|^2)} \right) \\ & \geq R^{(d)}(t) = \log_2 \left(1 + \frac{\delta p_i^{(d)}(t)}{N_0(H^2 + \|\mathbf{q}_0(t) - \mathbf{I}(t)\|^2)} \right) - \\ & \left(\frac{\log_2(e) p_i^{(d)}(t) (\|\mathbf{q}(t)\|^2 - \|\mathbf{q}_0(t)\|^2)}{(N_0 H^2 + \delta p_i^{(d)}(t) + N_0 \|\mathbf{q}_0(t)\|^2)(H^2 + \|\mathbf{q}_0(t)\|^2)} \right), \end{aligned} \tag{48}$$

$$\begin{aligned} & \log_2 \left(1 + \frac{\delta p_r}{N_0((H^s)^2 + \|\mathbf{q}(t) - \mathbf{s}(t)\|^2)} \right) \\ & \geq R^{(s)}(t) = \log_2 \left(1 + \frac{\delta p_r}{N_0((H^s)^2 + \|\mathbf{q}_0(t) - \mathbf{s}(t)\|^2)} \right) - \\ & \left(\frac{\log_2(e) p_r (\|\mathbf{q}(t)\|^2 - \|\mathbf{q}_0(t)\|^2)}{(N_0 (H^s)^2 + \delta p_r + N_0 \|\mathbf{q}_0(t)\|^2)((H^s)^2 + \|\mathbf{q}_0(t)\|^2)} \right), \end{aligned} \tag{49}$$

where the equalities hold when $q(t) = q_0(t)$.

Proof Let $\delta(t) = \|\mathbf{q}(t) - \mathbf{I}(t)\|^2$, and then $\log_2(1 + \frac{p_i^{(u)}(t)}{N_0(H^2 + \|\mathbf{q}(t) - \mathbf{I}(t)\|^2)})$ is converted into $\log_2(1 + \frac{p_i^{(u)}(t)}{N_0(H^2 + \delta(t))})$.

Let $f(\delta(t)) = \log_2(1 + \frac{p_i^{(u)}(t)}{N_0(H^2 + \delta(t))})$, in which $\delta(t)$ is the variable of function $f(\delta(t))$. Therefore, $f(\delta(t))$ is a convex function. According to Taylor expansion, the following inequality can be obtained.

$$f(\delta(t)) \geq f(\delta_0(t)) - f'(\delta_0(t))(\delta(t) - \delta_0(t)). \quad (50)$$

By substituting $\delta(t) = \|\mathbf{q}(t) - \mathbf{I}(t)\|^2$ into (50), inequality (47) is obtained. Similarly, the inequality (48) and (49) are obtained. \square

Therefore, the upload traffic size of user i within slot t is transformed into $D_i^{(u)}(t)$, which is given by

$$D_i^{(u)}(t) = \frac{T\theta_i^{(u)} B\alpha_i^{(u)}(t)}{M} \log_2\left(1 + \frac{\delta p_i^{(u)}(t)/N_0}{(H^2 + \|\mathbf{q}(t) - \mathbf{I}(t)\|^2)}\right), \quad (51)$$

$$\sum_{t=1}^M (\tau_i^{(d)} D_i^{(u)}(t) - \tau_i^{(t)} (\frac{T}{M} + \frac{\varphi\eta_i(t)D_i^{(u)}(t)\gamma_i}{f_i^{(c)}(t)})) \geq \xi, \quad (52)$$

$$\sum_{t=1}^M (\tau_i^{(d)} D_i^{(u)}(t) - \tau_i^{(t)} (\frac{(1 - \eta_i(t))D_i^{(u)}(t)}{B_s \log_2(1 + \frac{p_s \delta}{N_0((H^{(s)})^2 + \|\mathbf{q}(t) - \mathbf{S}\|^2)})}) \quad (53)$$

$$+ \frac{T}{M})) \geq \xi. \quad (54)$$

Substituting (51), (52) and (53) into **P5**, we have problem **P5**.

$$\mathbf{P5} : \max_{\mathbf{q}(t)} \xi \quad (54a)$$

$$s.t. (19), (13), (52) \text{ and } (53) \quad (54b)$$

According to [19, 27, 52], it can be proven that problem **P5** is convex. Thus, we can tackle it by leveraging classic algorithm (e.g., Gradient descent). Thereafter, substituting the optimal trajectory $\underline{\text{old}}q_{\text{opt}}(t)$ into problem **P1**. Then, solve **P1**, **P3** and **P5** iteratively until the growth of objective function is less than the threshold ρ_1 . The details are depicted in Algorithm 1.

Algorithm 1 TSAO algorithm based on NFP and SCA for **P0**.

- 1: **Input:** the iterative number $k=1$, the initial user utility $U_{\text{total}}^0 = -\infty$ and UAV trajectory $q_0(t)$, the tolerance errors $\rho_1 = 0.0001$ and $\rho_2 = 0.0001$. And other related parameters listed in Table 2.
- 2: **Output:** the upload duration $\alpha^{(u)}$, the download duration $\alpha^{(d)}$, the transmit power allocated to users $\mathbf{p}^{(d)}$ and the location of UAV \mathbf{q} .
- 3: **while** $U_{\text{total}}^k - U_{\text{total}}^{k-1} > \rho_1$ **do**
- 4: Solve **P1** by leveraging gradient descent method and the optimal values of $\alpha_{i,\text{opt}}^{(u)}(t)$, $\alpha_{i,\text{opt}}^{(d)}(t)$, $p_{i,\text{opt}}^{(d)}(t)$, $p_{i,\text{opt}}^{(u)}(t)$ are obtained ;
- 5: Substituting the variables obtained by solving **P1** into **P3**, and solve **P3** by using gradient descent method. The optimal values of variables $\mu_i^{(c)}(t)$, $\mu_i^{(s)}(t)$, $v_s(t)$ are obtained.
- 6: **while** $\sum_{j=1}^M \|\mathbf{q}_j - \mathbf{q}_{\text{opt}}\| > \rho_2$ **do**
- 7: Solve **P5** by leveraging gradient descent method and the optimal trajectory \mathbf{q}_{opt} is obtained;
- 8: $j = j + 1$;
- 9: $\mathbf{q}_{\text{opt}} = \mathbf{q}_j$
- 10: **end while**
- 11: $k = k + 1$
- 12: $U_{\text{total}}^k = U_{\text{opt}}$
- 13: $\mathbf{q}^k = \mathbf{q}_{\text{opt}}$
- 14: **end while**

The convergence of Algorithm 1 is discussed in [53]. Besides, the complexity of Algorithm 1 is composed of four aspects. The first is the computation of problem **P1** by using CVX [54, 55]. The second is the computation of problem **P3** by using CVX. Denote A_1 , B_1 as the number of required iterations for the outer loop and inner loop in Algorithm 1, respectively. Thus, according to [56], the complexity of Algorithm 1 is $\mathcal{O}(A_1(2M^3 + B_1M^3))$.

4 Numerical results

In this section, numerical results are provided to evaluate the performance of proposed scheme. The simulation settings are based on some existing work [27, 57]. There are $N = 4$ users and their positions are $\mathbf{L}_1 = [0,0]$, $\mathbf{L}_2 = [10,0]$, $\mathbf{L}_3 = [0,10]$ and $\mathbf{L}_4 = [10,10]$, respectively [43]. The location of BS is set as $\mathbf{S} = [70,70,0]$. Besides, UAV's initial position and destination position are set as $\mathbf{q}_{\text{st}} = [0,0]$ and $\mathbf{q}_{\text{st}} =$

Table 2 Simulation parameters

| Parameter | Value | Parameter | Value |
|------------|--------|-------------|-------------|
| H | 10 m | B_s | 40 MHz |
| γ_i | 10^3 | $f_i^{(c)}$ | 1 GHz |
| δ | -50 dB | N_0 | -174 dBm/Hz |

[0,10], respectively. Assume users are static within each slot. Users’ traffic and computation offloading demands are set as $\mathbf{a}^{(d)} = [1,1,1,0]$ and $\mathbf{a}^{(u)} = [1,0,1,1]$, respectively. Besides, the duration time is set as $T = 2s$, which is decomposed into $M = 50$ slots. UAV’s maximum flight speed is set as 20m/s. Both uplink and downlink bandwidths are set as 40MHz. $p_{max}^{(d)}$ and $p_{i,max}^{(u)}$ are set as 4W and 4W, respectively. $\tau_i^{(d)}$ and $\tau_i^{(t)}$ are set as 0.5 and 0.5, respectively. The remaining parameters used in the simulations, unless specified otherwise, are summarized in Table 2. The altitude of UAV is set as $H = 10m$, the channel power gain is set as $\delta = -50dB$ and the noise power is set as $N_0 = -174dBm/Hz$. Besides, it is assumed the required CPU cycles per bit of user i is set as $\gamma_i = 10^3$ and the computing capacity that cloud allocated to user i is set as $f_i^{(c)} = 1GHz$.

Figure 2 shows the dynamics of the minimum user utility during iteration. Similar to [27, 43, 58, 59], the iteration number in Fig. 2 is defined as the number that the three-stage has been looped. That is, we only calculate the number of outer loop in Algorithm 1, and the iterations of solving problem P1, P3 and P5 is not counted. It can be observed that the value of utility will decrease with the iteration. The transmit power of UAV larger, the minimum user

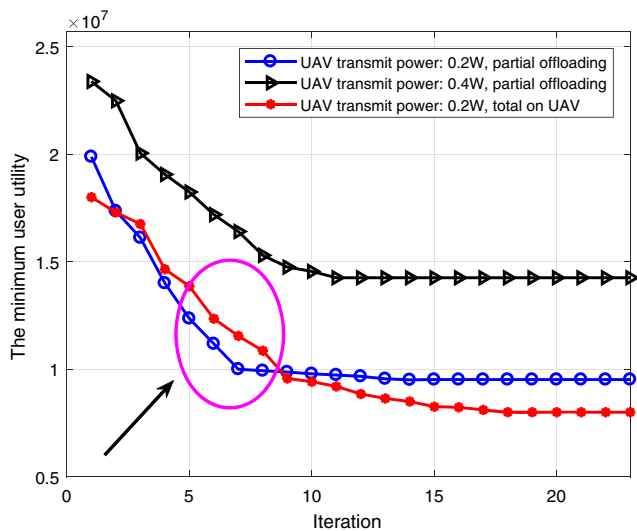


Fig. 2 The dynamics of the minimum user utility

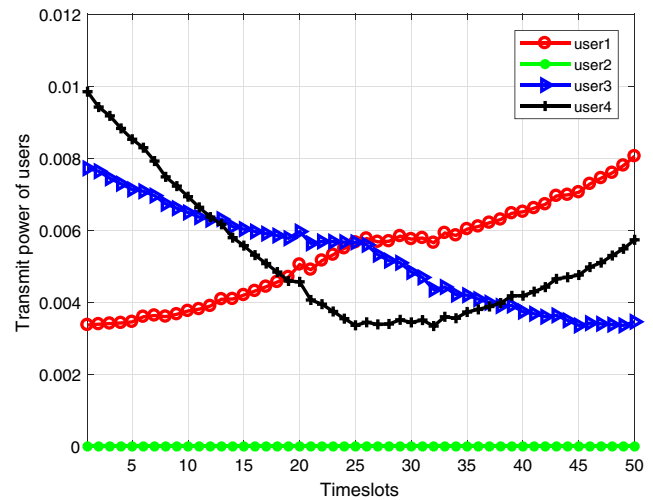


Fig. 3 a Users’ transmit power with different timeslots

utility larger, which is because the transmit rate between UAV and BS increases. More clearly, the transmit power larger, the more data can be processed on UAV and the less delay it takes. And the proposed partial offloading scheme outperforms the scheme that only computing on UAV. The reason is that the proposed partial offloading scheme can take use of BS for computing, which augments the computation capacity of UAV. However, the minimum user utility of proposed partial is less than that of total on UAV during iteration. It’s because that in the alternative process of solving the three-stage problem, the obtained value of some variables are not global optimal. Besides, the proposed algorithm only need several iterations to converge, which indicates it is computationally efficient and has a fast convergence rate.

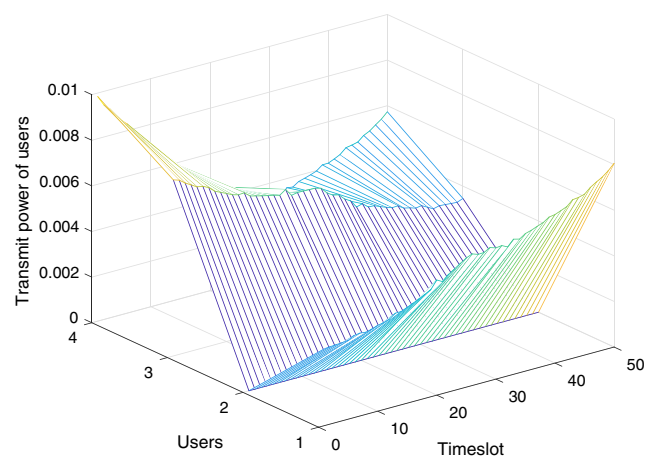


Fig. 4 b Users’ transmit power with different timeslots

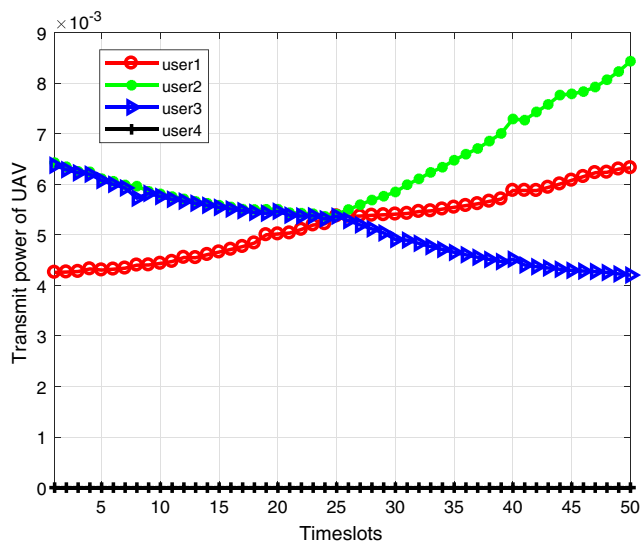


Fig. 5 a UAV’s transmit power to each user with different timeslots

Figures 3 and 4 show the dynamics of users’ transmit power with different time slots under a given semicircle trajectory. It can be observed that user 1’s transmit power is gradually increasing, which is due to the distance between UAV and user 1 is gradually increasing. Due to the requirements of fairness among users, users will increase transmit power to keep communication rate steady when the distance increases. Besides, it can be observed that the trends of users’ transmit power is opposite to that of the distance between users and UAV. Since user 2 has no computation offloading demand, the transmit power of user 2 equals to zero. Additional, the transmit power of user 4 has an increasing trend after decreasing, which is due to the distance between user 4 and UAV decreases first and increases after slot 25. Thus, it can be inferred that users’ transmit power is related to the distance between UAV and users. In addition, it can be observed that the value of users’

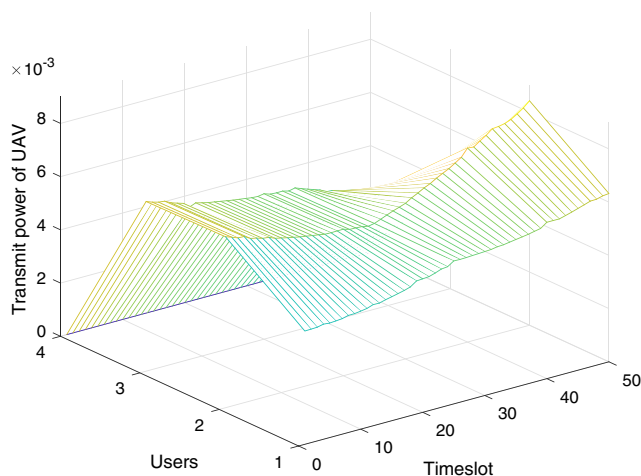


Fig. 6 b UAV’s transmit power to each user with different timeslots

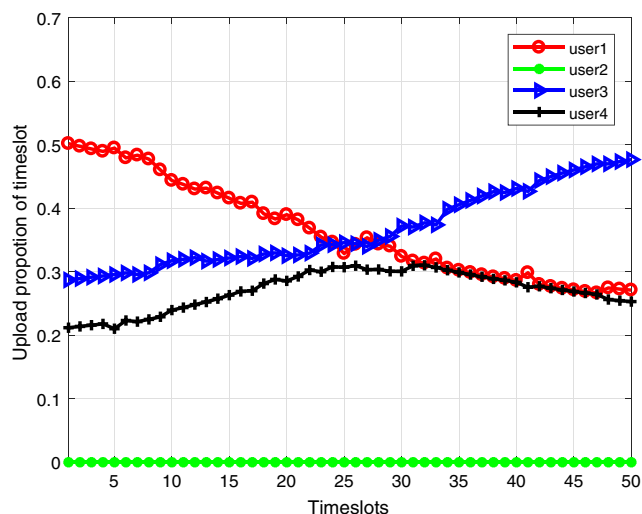


Fig. 7 a Proportion of upload time about users within each slot

average transmit power are close, and it will lead to the traffic size that users uploaded are close, which denotes the fairness among users.

Similar to Figs. 3 and 4, 5 and 6 show UAV’s transmit power that allocated to each user with different time slots under a given semicircle trajectory. It can be seen that the result is similar to Figs. 3 and 4, which has been described and analyzed above. Moreover, to meet users’ minimum download rate requirement, the minimum transmit power that UAV allocates to users is still relatively high. It implies that the proposed scheme not only guarantees the computation offloading fairness, but also traffic offloading fairness among users.

The proportion of users’ upload duration within each slot is shown in Figs. 7 and 8. It can be seen that the dynamic changes trends in Figs. 7 and 8 are opposite to that in Figs. 3 and 4. That is, the transmit power of users smaller, the proportion of users upload time larger. And it denotes that

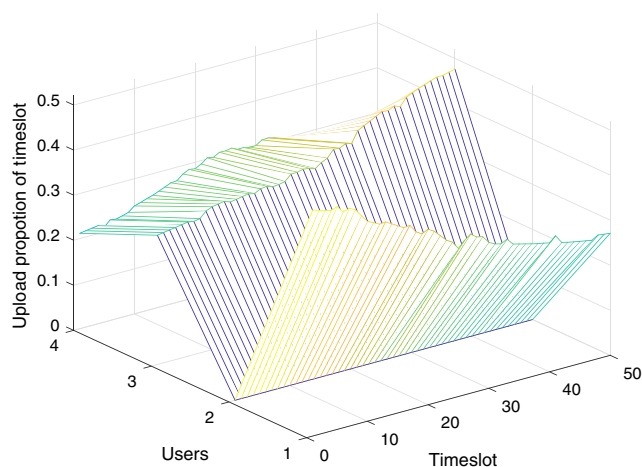


Fig. 8 b Proportion of upload time about users within each slot

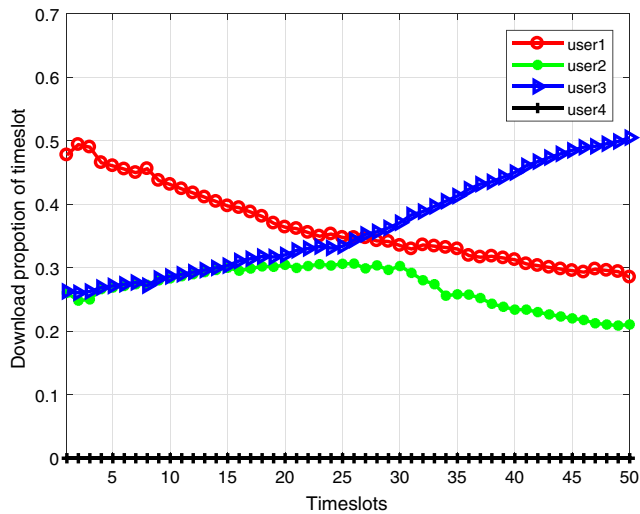


Fig. 9 a Proportion of download time about users during each slot

more traffic of users will be uploaded when the transmit power in low state. In addition, it can be observed that the distance between UAV and users smaller, the upload time of users larger, which guarantees the high computation rate of users.

The proportion of users' download duration within each slot is shown in Figs. 9 and 10. It can be observed that the dynamic changes trends in Figs. 9 and 10 are opposite to that in Figs. 5 and 6. The transmit power that UAV allocates to users smaller, the proportion of download duration larger, which is due to the minimum download rate requirements of users. That is, in order to satisfy the requirements, the download time of users will increase when in low transmit power range. From another aspect of view, UAV offloads more traffic for users in the low transmit power state, which guarantees UAV can assist more users.

Figure 11 shows the minimum user utility with different user's and UAV's maximum transmit power and offloading

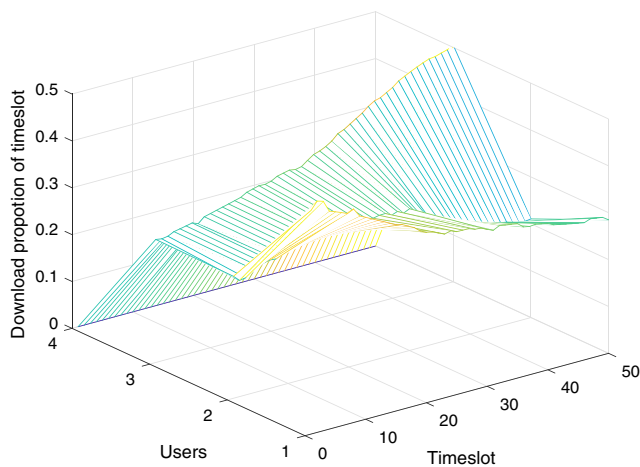


Fig. 10 b Proportion of download time about users during each slot

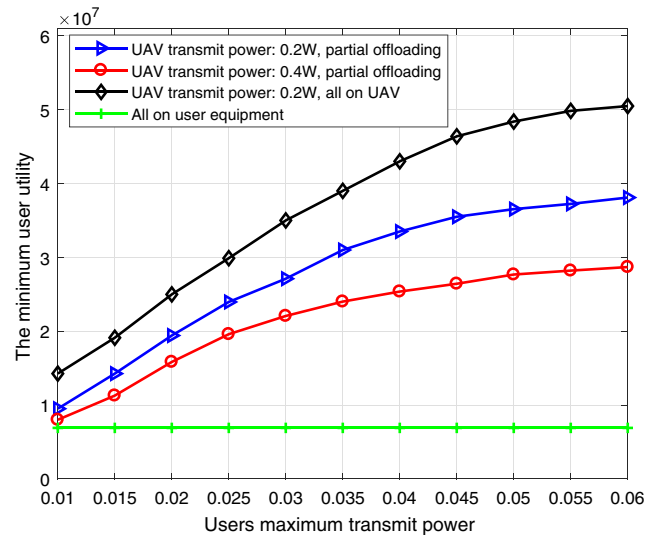


Fig. 11 The minimum user utility with different user's maximum transmit power

schemes [27, 43]. It can be observed that with the increasing of users' maximum transmit power, the minimum user utility is also increasing. The reason is that the transmit power of users higher, the size of data uploaded larger, which is contribute to the user utility. However, the minimum user utility increasing gradually slowly, which is because the transmit rate between UAV and BS is limited. With the upload data increasing, UAV has no enough computation capacity, which leads to the increasing of delay. Therefore, the minimum user utility increases gradually slowly. In addition, it can be observed that minimum user utility under the schemes that only perform computation on UAV and local computation is relatively

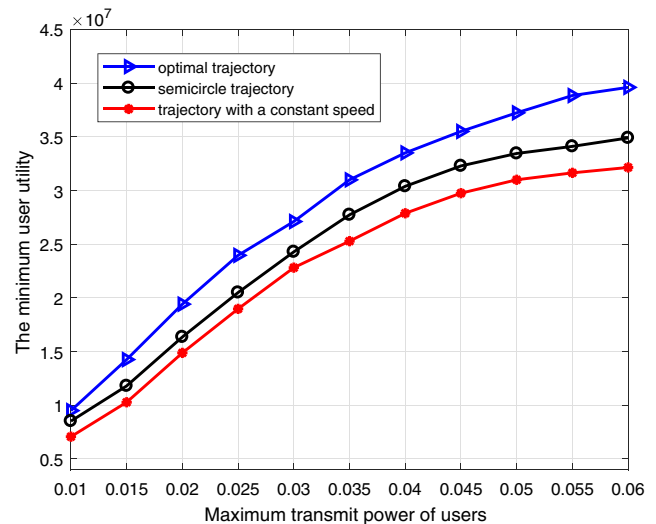


Fig. 12 The minimum user utility under different UAV trajectories

smaller than the proposed schemes, which indicates the efficiency of proposed algorithm.

Figure 12 show the minimum user utility versus the maximum transmit power of users under several different UAV trajectories. As shown in Fig. 12, the minimum user utility achieved by using proposed scheme is larger than that of semi-circle trajectory mode and trajectory with a constant speed mode. The reason is that users who are far from UAV have low transmit rate, and it leads to the low computation rate and high delay. Besides, the result indicates that the optimization of the trajectory of the UAV can improve the minimum user utility. It also verifies that our proposed scheme outperforms the disjoint optimization scheme. Besides, the relation between the transmit power of users and minimum user utility is consistent with Fig. 11.

Figure 13 show the minimum user utility versus the number of users under different maximum transmit power of UAV. The maximum transmit power of UAV is set as $p_{\max}^{(uav)} = 0.2W$, $p_{\max}^{(uav)} = 0.3W$ or $p_{\max}^{(uav)} = 0.4W$. As shown in Fig. 13, the minimum user utility increases with the number of users. The reason is that the number of users larger, the computation rate larger. However, the delay doesn't increase too much, which is because the UAV offloading some tasks to the BS for computing. It indicates the proposed schemes is adaptive for the large number of users. Besides, it can be observed that the minimum use utility under $p_{\max}^{(uav)} = 0.4W$ is larger than that of utility under $p_{\max}^{(uav)} = 0.3W$ and $p_{\max}^{(uav)} = 0.2W$. It's because the maximum transmit power of the UAV larger, the transmit rate between UAV and BS larger, which can reduce the delay and increase the computation rate.

As shown in Fig. 14, the utilities of users are almost equally. With the increasing of user transmit power, the results are similar. It indicates fairness among users.

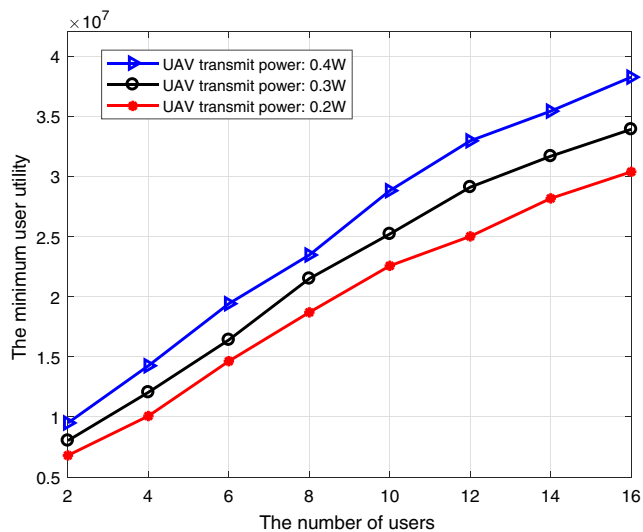


Fig. 13 The minimum user utility with different number of users

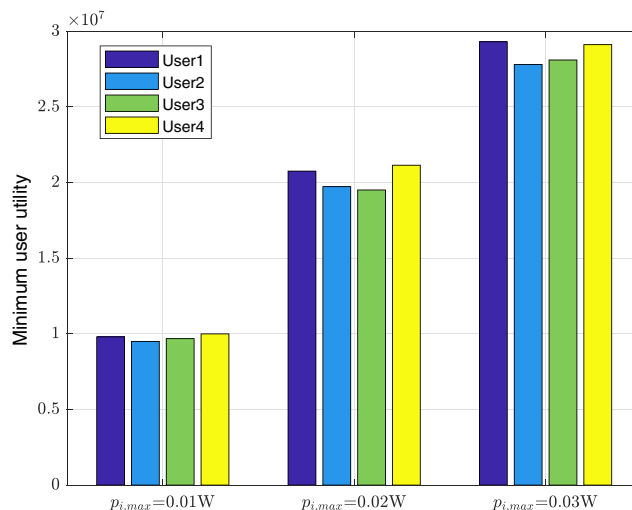


Fig. 14 Users utilities with transmit power

In conclusion, we evaluate the performance of proposed scheme from several different aspects. First, Fig. 2 shows the proposed algorithm has a fast convergence rate. Then, Figs. 3-6 study the dynamics of users' and UAV's transmit power with different time slots, and Fig. 7-10 study the dynamics of download and upload duration allocation. The results show that the resources allocated to each user are almost equally, which demonstrates the fairness among users. Last, the minimum user utility with different parameters and other disjoint optimization schemes are studied in Figs. 11-14. The results show that the proposed scheme can achieve larger user utility under different parameters, which implies the proposed scheme has a better performance than other disjoint optimization schemes.

5 Conclusion

In this paper, we study the UOTO optimization problem. The minimum user utility is maximized under the partial offloading schemes. Since the problem is formulated as a non-convex problem, a three-stage alternative optimization algorithm based on SCA and Nonlinear fractional programming is proposed to jointly optimize the power allocation, bandwidth allocation, offloading decision and UAV trajectory. Simulation results show that the proposed schemes outperform other disjoint optimization schemes. Besides, it is shown that the convergence speed of proposed algorithm is fast.

For the future work, we are going to study the more general case where multiple UAVs assisted mobile offloading, and the collision avoidance among UAVs should be carefully investigated. Another research direction is to consider UAV flight with a variable altitude, which would be very interesting and challenging.

Acknowledgments This work is supported by the Natural Science Foundation of China (No. 61872104), the Natural Science Foundation of Heilongjiang Province in China (No. F201 6009), the Fundamental Research Fund for the Central Universities in China (No. HEUCF180602) and the National Science and Technology Major Project (No. 2016 ZX03001023-005).

References

- Hu X, Zhuang X, Feng G, Lv H, Wang H, Lin J (2018) Joint optimization of traffic and computation offloading in uav-assisted wireless networks. In: 2018 IEEE 15th international conference on mobile Ad Hoc and sensor systems (MASS). IEEE, pp 475–480
- Wang Y, Wang K, Huang H, Miyazaki T, Guo S (2019) Traffic and computation co-offloading with reinforcement learning in fog computing for industrial applications. *IEEE Trans Indust Inf* 15(2):976–986
- Iosifidis G, Gao L, Huang J, Tassiulas L (2015) A double-auction mechanism for mobile data-offloading markets. *IEEE/ACM Trans Netw* 23(5):1634–1647
- Al-Kanj L, Vincent Poor H, Dawy Z (2014) Optimal cellular offloading via device-to-device communication networks with fairness constraints. *IEEE Trans Wirel Commun* 13(8):4628–4643
- Wang T, Zhang G, Liu A, Bhuiyan MZA, Jin Q (2019) A secure iot service architecture with an efficient balance dynamics based on cloud and edge computing. *IEEE Internet of Things Journal* 6(3):4831–4843
- Zeng Y, Zhang R, Teng JL (2016) Wireless communications with unmanned aerial vehicles: opportunities and challenges. *IEEE Commun Mag* 54(5):36–42
- Wang K, Zhou Q, Guo S, Luo J (2018) Cluster frameworks for efficient scheduling and resource allocation in data center networks: A survey. *IEEE Commun Surv Tutor* 20(4):3560–3580
- Ministry of Internal Affairs Economic Research Office and Japan Communications (MIC). Information and communications in japan. <http://www.soumu.go.jp/johotsusintokei/whitepaper/eng/WP2011/2011-index.html>
- Qiantori A, Sutiono AB, Hariyanto H, Suwa H, Ohta T (2012) An emergency medical communications system by low altitude platform at the early stages of a natural disaster in indonesia. *J Med Syst* 36(1):41–52
- Cheng F, Zhang S, Li Z, Chen Y, Zhao N, Yu R, Leung VCM (2018) Uav trajectory optimization for data offloading at the edge of multiple cells. *IEEE Trans Veh Technol* 67(7):6732–6736
- Li Y, Feng G, Ghasemahmadi M, Cai L (2019) Power allocation and 3-d placement for floating relay supporting indoor communications. *IEEE Trans Mob Comput* 18(3):618–631
- Zhan P, Yu K, Swindlehurst AL (2011) Wireless relay communications with unmanned aerial vehicles: Performance and optimization. *IEEE Trans Aerosp Electron Syst* 47(3):2068–2085
- Zhang J, Zeng Y, Zhang R (2018) UAV-Enabled Radio Access Network: Multi-Mode Communication and Trajectory Design. ArXiv e-prints
- Zhang J, Zeng Y, Zhang R (2017) Spectrum and energy efficiency maximization in uav-enabled mobile relaying. In: IEEE international conference on communications, pp 1–6
- Zhao W (2004) A message ferrying approach for data delivery in sparse mobile ad hoc networks. In: Proceedings of ACM Int'l symposium on mobile Ad Hoc networking and computing, pp 187–198
- Wang Y, Wang K, Huang H, Miyazaki T, Guo S (2019) Traffic and computation co-offloading with reinforcement learning in fog computing for industrial applications. *IEEE Trans Indust Inf* 15(2):976–986
- Wang T, Zhou J, Liu A, Bhuiyan MZA, Wang G, Jia W (2019) Fog-based computing and storage offloading for data synchronization in iot. *IEEE Internet of Things Journal* 6(3):4272–4282
- Wang T, Zeng J, Lai Y, Cai Y, Tian H, Chen Y, Wang B Data collection from wsns to the cloud based on mobile fog elements. *Future Generation Computer Systems*, <https://doi.org/10.1016/j.future.2017.07.031>
- Jeong S, Simeone O, Kang J (2017) Mobile cloud computing with a uav-mounted cloudlet: optimal bit allocation for communication and computation. *Iet Commun* 11(7):969–974
- He X, Wang K, Huang H, Miyazaki T, Wang Y, Guo S (2018) Green resource allocation based on deep reinforcement learning in content-centric iot. *IEEE Transactions on Emerging Topics in Computing*, pp 1–1. <https://doi.org/10.1109/TETC.2018.2805718>
- Zeng Y, Zhang Rx (2017) Energy-efficient uav communication with trajectory optimization. *IEEE Trans Wirel Commun* 16(6):3747–3760
- Hua M, Wang Y, Wu Q, Li C, Huang Y, Yang L (June 2018) Energy Optimization for Cellular-Connected multi-UAV Mobile Edge Computing Systems with Multi-Access Schemes. ArXiv e-prints
- Zhang J, Zeng Y, Zhang R (2018) Uav-enabled radio access network: Multi-mode communication and trajectory design
- Huang Y, Xu J, Qiu L, Zhang R (2018) Cognitive UAV Communication via Joint Trajectory and Power Control. arXiv:1802.05090
- Li P, Guo S, Miyazaki T, Liao X, Jin H, Zomaya AY, Wang K (2017) Traffic-aware geo-distributed big data analytics with predictable job completion time. *IEEE Trans Parallel Distrib Syst* 28(6):1785–1796
- Xu J, Zeng Y, Zhang R (2018) Uav-enabled wireless power transfer: Trajectory design and energy optimization. *IEEE Transactions on Wireless Communications* 17(8):5092–5106
- Jeong S, Simeone O, Kang J (2018) Mobile edge computing via a uav-mounted cloudlet: Optimization of bit allocation and path planning. *IEEE Trans Veh Technol* 67(3):2049–2063
- Motlagh NH, Taleb T, Arouk O (2016) Low-altitude unmanned aerial vehicles-based internet of things services Comprehensive survey and future perspectives. *IEEE Internet J* 3(6):899–922
- Zhang H, Zheng WX (2018) Denial-of-service power dispatch against linear quadratic control via a fading channel. *IEEE Trans Autom Control* 63(9):3032–3039
- Cover TM, Thomas JA (2006) Elements of information theory (wiley series in telecommunications and signal processing). Wiley-Interscience, New York
- Wang F, Xu J, Wang X, Cui S (2017) Joint offloading and computing optimization in wireless powered mobile-edge computing systems. *IEEE Trans Wirel Commun* 17(3):1784–1797
- Zhang H, Meng W, Qi J, Wang X, Zheng WX (2019) Distributed load sharing under false data injection attack in an inverter-based microgrid. *IEEE Trans Indust Electron* 66(2):1543–1551
- Yang G, He S, Shi Z (2017) Leveraging crowdsourcing for efficient malicious users detection in large-scale social networks. *IEEE Internet J* 4(2):330–339
- Xu C, Jiao L, Li W, Fu X (2015) Efficient multi-user computation offloading for mobile-edge cloud computing. *IEEE/ACM Trans Netw* 24(5):2795–2808
- Thengvall BG, Bard JF, Yu G (2003) A bundle algorithm approach for the aircraft schedule recovery problem during hub closures. *Transp Sci* 37(4):392–407
- Wallenius J, Dyer JS, Fishburn PC, Steuer RE, Zionts S, Deb K (2008) Multiple criteria decision making, multiattribute utility

- theory: Recent accomplishments and what lies ahead. *Manage Sci* 54(7):1336–1349
37. Sun R, Hong M, Luo ZQ (2014) Joint downlink base station association and power control for max-min fairness: Computation and complexity. *IEEE J Sel Areas Commun* 33(6):1040–1054
 38. Li Y, Sheng M, Tan CW, Zhang Y, Sun Y, Wang X, Shi Y, Li J (2015) Energy-efficient subcarrier assignment and power allocation in ofdma systems with max-min fairness guarantees. *IEEE Trans Commun* 63(9):3183–3195
 39. Fang B, Qian Z, Zhong W, Shao Y (2015) An-aided secrecy precoding for swipt in cognitive mimo broadcast channels. *IEEE Commun Lett* 19(9):1632–1635
 40. Mokhtari A, Koppel A, Scutari G, Ribeiro A (2017) Large-scale nonconvex stochastic optimization by doubly stochastic successive convex approximation. In: *IEEE international conference on acoustics, speech and signal processing*, pp 4701–4705
 41. Tian X, Zhang W, Yang Y, Wu X, Peng Y, Wang X (2018) Toward a quality-aware online pricing mechanism for crowdsensed wireless fingerprints. *IEEE Trans Veh Technol* 67(7):5953–5964
 42. Yang C, Shi Z, Han K, Zhang JJ, Gu Y, Qin Z (2018) Optimization of particle cbmember filters for hardware implementation. *IEEE Trans Veh Technol* 67(9):9027–9031
 43. Zhou F, Wu Y, Qingyang Hu R, Qian Y (June 2018) Computation Rate Maximization in UAV-enabled Wireless Powered Mobile-Edge Computing Systems. *ArXiv e-prints*
 44. Yang G, He S, Shi Z, Chen J (2017) Promoting cooperation by the social incentive mechanism in mobile crowdsensing. *IEEE Commun Mag* 55(3):86–92
 45. Faybusovich BV (2006) Convex optimization. *IEEE Trans Autom Control* 51(11):1859–1859
 46. Bai Y, Zhang L (2017) A full-newton step interior-point algorithm for symmetric cone convex quadratic optimization. *J Indust Manag Optim* 7(4):891–906
 47. Tian X, Li W, Yang Y, Zhang Z, Wang X (2018) Optimization of fingerprints reporting strategy for WLAN, indoor localization. *IEEE Trans Mob Comput* 17(2):390–403
 48. Wang Y, Lopez JA, Sznaiier M (2018) Convex optimization approaches to information structured decentralized control. *IEEE Trans Autom Control* 63(10):3393–3403
 49. Shen K, Yu W (2018) Fractional programming for communication systems-part I: power control and beamforming. *CoRR*, arXiv:1802.10192
 50. Serali HD, Alameddine A (1992) A new reformulation-linearization technique for bilinear programming problems. *J Glob Optim* 2(4):379–410
 51. Zhu Y, Zhong Z, Basin MV, Zhou D (2018) A descriptor system approach to stability and stabilization of discrete-time switched pwa systems. *IEEE Trans Autom Control* 63(10):3456–3463
 52. Zhu Y, Zhong Z, Zheng WX, Zhou D (2018) Hmm-based h-infinity filtering for discrete-time markov jump LPV systems over unreliable communication channels. *IEEE Trans Syst Man Cybern Syst* 48(12):2035–2046
 53. Scutari G, Facchinei F, Lampariello L, Song P (2014) Parallel and distributed methods for nonconvex Optimization-Part I: Theory. *ArXiv e-prints*
 54. Grant M, Stephen B (2014) CVX: Matlab software for disciplined convex programming, version 2.1. <http://cvxr.com/cvx>
 55. Chen J, Hu K, Wang Q, Sun Y, Shi Z, He S (2017) Narrowband internet of things Implementations and applications. *IEEE Internet J* 4(6):2309–2314
 56. Bubeck S (May 2014) Convex Optimization: Algorithms and Complexity. *ArXiv e-prints*
 57. Cao X, Xu J, Zhang R (2018) Mobile edge computing for cellular-connected uav: Computation offloading and trajectory optimization. In: *2018 IEEE 19th international workshop on*

signal processing advances in wireless communications (SPAWC), <https://doi.org/10.1109/SPAWC.2018.8445936>

58. Zhou F, Wu Y, Sun H, Chu Z (2018) Uav-enabled mobile edge computing: Offloading optimization and trajectory design. In: *2018 IEEE international conference on communications, ICC 2018, Kansas City, MO*
59. Hua M, Huang Y, Yi W, Wu Q, Dai H, Yang L (2018) Energy optimization for cellular-connected multi-uav mobile edge computing systems with multi-access schemes. *J Comm Inform Netw* 3(4):33–44

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Guangsheng Feng (M'18) received his B.E., M.E, and Ph.D. degrees from Harbin Engineering University (HEU) in 2003, from Harbin Institute Technology (HIT) in 2005, and from HEU in 2009, respectively. Now, he is an associate professor in HEU and researching at mobile data offloading in cellular networks and optimization. His research interests involve edge computing, LTE (LTE-A) networks and wireless channel access control.



Chengbo Wang received his B.E. degree from Shandong University of Finance and Economics in 2017. Now, he is a postgraduate student in Harbin Engineering University in China and majoring in computer science and technology. His research interests involve computation migration and optimization.



Bingyang Li received his B.E., M.E, and Ph.D. degrees from Harbin University of Science and Technology (HUST) in 2001, from Harbin Engineering University (HEU) in 2004, and from HEU in 2009, respectively. Now, he is an associate professor in HEU and researching at mobile ad hoc networks and security. He is a full staff at HEU. His research interests involve cloud computing and wireless network security.



Haibin Lv received his B.E. degree from Tianjin Normal University in China in 2016. Now, he is a postgraduate student in Harbin Engineering University and majoring in computer science and technology. His research interests involve mobile data offloading and optimization.



Huiqiang Wang received his M.E. and Ph.D. degrees from HEU in 1985 and 2005, respectively. From 2001 to 2002, he was at Queens University, Ontario, Canada, as a senior visiting scholar. Now, he is engaged in teaching and researching as a professor and a doctoral advisor at HEU. Up to now, he holds ten Chinese patents. His research interests involve network security, cognitive networks, and autonomic computing.



Xiaoxiao Zhuang received his B.E. degree from Henan University of Science and Technology in China in 2016. Now, she is a postgraduate student in Harbin Engineering University and majoring in computer science and technology. Her research interests involve mobile data offloading and UAV's trajectory optimization.



Xianlang Hu received his B.E. degrees from Harbin Engineering University (HEU) in 2003. Now, he is a full staff in Jiangsu Automation Research Institute and researching at mobile data offloading in industrial internet of things. His research interests involve embedded Systems and edge computing.



Hongwu Lv received his M.S. and Ph.D. degrees from HEU in 2009 and 2011, respectively. Now, he is engaged in teaching and researching as a lecturer at HEU. His research interests involve cognitive computing and dependability analysis.