Intelligent Resource Allocation in UAV-Enabled Mobile Edge Computing Networks

Meng Wang[1], Shuo Shi[1,2], Shushi Gu[2,3], Ning Zhang[4], Xuemai Gu[1]
[1]School of Electronics and Information Engineering, Harbin Institute of Technology Harbin, Harbin 150001, China
[2]Peng Cheng Laborotary, Shenzhen, 518055, China
[3]School of Electronics and Information Engineering, Harbin Institute of Technology Harbin(Shenzhen), Shenzhen, 518055, China
[4]Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B3P4, Canada
E-mail: wangmeng_hit@163.com, crcss@hit.edu.cn. gushushi@hit.edu.cn, ning.zhang@ieee.org, guxuemai@hit.edu.cn

***Abstract*—Unmanned aerial vehicles (UAVs) have been considered as effective flying base stations (FBSs) to provide on-demand wireless communications. Equipped with computation resource, UAVs are also capable of offering computation offloading opportunities for the mobile users (MUs) in mobile edge computing (MEC) networks. However, due to the small hardware and load capacity, UAVs can only supply limited computation and energy resource. It is thus challenging for UAVs to guarantee the quality of service (QoS) of MUs, while minimizing their total resource consumptions. Toward this end, instead of using all resource for every single task, we propose an intelligent resource allocation algorithm based on reinforcement learning, which enables UAVs to make energy-efficent and computation-efficent allocation decisions intelligently. Then, we take UAVs as learning agents by forming resource allocation decisions as actions and designing a reward function with the aim of minimizing the weighted resource consumptions. Each UAV performs the algorithm only based on its local observations without information exchange among different UAVs. Simulation results show that the proposed reinforcement learning based approach outperforms the benchmark algorithms in terms of weighted consumptions in a whole time period.**

***Index Terms*—UAV communications, intelligent resource allocation, reinforcement learning, mobile edge computing**

I. Introduction

Unmanned aerial vehicles (UAVs) have a wide range of applications in both civil and military fields due to their high flexibility and controllability. In particular, if properly planned and deployed, UAVs can provide reliable and cost-effective wireless communications solutions for a variety of scenarios to deal with the rapid growth of mobile users (MUs) and data traffic [1]. Since it is very costly to deploy more network infrastructure in rural areas (e.g., mountain areas), UAVs with communication transceivers can act as flying base stations (FBSs) to give high-speed and seamless-coverage access especially for MUs in those areas [2,3]. Mobile edge computing (MEC) is a promising technique for MUs with limited computation resource to offload latency-intensive or computation-intensive tasks to network edge servers. Equipped with computation resource, UAVs can play the important role of aerial edge servers with high mobility, namely UAV-enabled MEC communication networks [4]. Many siginificant studies on UAV-enabled MEC communication networks has been extensively investigated, including UAV-ground channel model, offloading tasks model, latency reduction, etc. Nevertheless, resource allocation for UAVs such as transmit power, spectrum and computation resource, as a key communication problem to be solved, is essential to further enhance the energy-efficiency and coverage for UAV-enabled communication networks.

Compared with conventional terristrial edge servers and small cells, UAVs are generally faster and more flexible to deploy. However, due to their limited energy capacity and radius coverage, applying UAVs into practical scenarios is still a challanging mission. In [5], a three-dimensional (3D) deployment algorithm based on circle packing is developed for maximizing the downlink coverage performance. [6] analyzes the coverage performance of UAV-assisted terrestrial cellular networks, where partially energy-harvesting-powered caching UAVs are randomly deployed in the 3-D space with a minimum and maximum altitude, [4] consider the energy-efficient issues where a single UAV performs data collection upon the ground sensors through trajectory design. A UAV-enabled computation offloading problem is detailed discussed in [5] with the goal of optimiziting the uplink transmission bits. A two-layer network architecture is proposed where a single UAV roams around a fixed area acting as the moving edge server, aiming at minimizing the weighted sum of delay and energy consumptions among all MUs in each time period [6]. Due to the versatility and manoeuvrability of UAVs, human intervention becomes restricted for UAVs' flight control. Therefore, machine learning based intelligent control of UAVs is desired for enhancing the performance for UAV-enabled communication networks [10]. [14] proposes an efficent iterative algorithm to minimize the energy consumptions of UAVs by optimizing the hovering time, scheduling and resource allocation of the tasks in a UAV-enabled MEC scenario.

Although there are many exiciting works focus on UAV-enabled communication issues in MEC network, they mainly focus on a single UAV server [7,8,9,14] or multi-UAV scenarios by assuming the availability of complete prior information for each UAV [12]. Moreover, studies in UAV-

This work is supported by Natural Sciences Foundation of China under Grant 61701136 and the project "The Verification Platform of Multi-tier Coverage Communication Network for Oceans (PCL2018KP002)" .

enabled MEC are mainly user-centric. Only a few works are UAV-centric with the aim of optimizing UAVs' performace in MEC scenarios, which is worth studying due to UAVs' limited on-board energy and resource. Motivitated by the aforementioned issues, in this paper, we study the resource allocation issue in multi-UAV-enabled MEC network scenario where UAVs play the role of edge servers. Each UAV executes the offloaded tasks from MUs in each time interval and commnuicate with the ground through downlink transimssion independently by implementing reinforement learning based approach. The objective is to minimize the total resource consumptions, including transmitting energy and computation resource, while guaranteeing the latency constraints of MUs' demand. We first formulate the resource allocation issue as a mixed integer linear programm (MILP) optimization problem, which is difficult to solve with low complexity solutions. Then we propose our multi-agent Q-learning algorithm to solve the problem distributedly and efficiently. Simulation results show the effectiveness of the proposed algorithm compared with other benchmark algorithms.

The remainder of this paper is organized as follows. In Section II, we state the system model and problem formulation. We give the detailed designs of the multi-agent Q-learning algorithm in Section III. Numerical results are presented in Section IV. Finally, the paper is concluded in Section V.

II. System model and problem formulation

A. Network model

Fig. 1 shows a multi-UAV MEC network, where N MUs are distributed randomly in an area with M UAVs hovering in fixed locations. Each UAV is assumed to have communication and computation abilities, enabling MUs to offload the tasks. The objective of this paper is to minimize the energy and computation consumptions of UAVs through resource allocation under the QoS constraints of MUs in terms of latency.

The sets of MUs and UAVs are denoted as $\mathcal{N} = \{1, 2, ..., N\}$ and $\mathcal{M} = \{1, 2, ..., M\}$, respectively. MUs are associated with UAVs using orthogonal frequency-division multiplexing (OFDM). We assume the network operates during a time period K, which consists of T time intervals, denoted as $\mathcal{T} = \{1, 2, ..., T\}$, and the locations of MUs are assumed to be constant in a time interval, denoted as $\mathbf{q}_i(t) = [x_i(t), y_i(t)]$. Furthermore, we assume that UAVs fly horizontally at a constant altitude H with a radius coverage of R and the coordinates of j-th UAV can be expressed as $\mathbf{p}_j = [X_j(t), Y_j(t)]$. Moreover, the k-th task of i-th MU is indicated as $R_i^k = (S_i^k, F_i^k, D)$, where S_i^k,F_j^k,D denotes the input size, the required CPU circles and the maximum tolerant time of k-th task, respectively. Note that, D remains the same for all tasks, which represents the QoS constraints of latency-intensive tasks.

B. Communication model

According to [8], the channel gain from j-th UAV to i-th MU can be modelled as

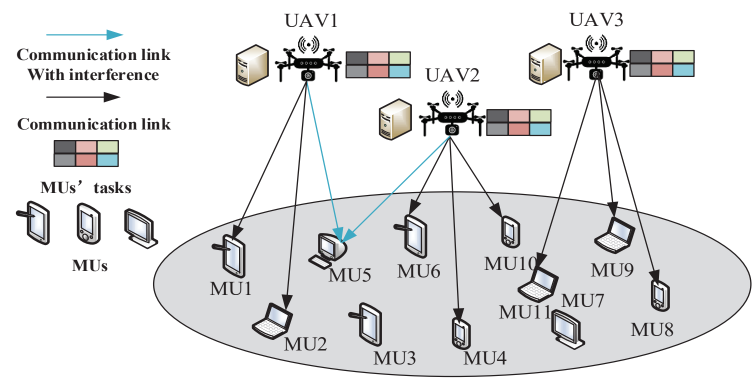

Fig. 1. An illustration of multi-UAV network structure, where each UAV makes resource allocation decisions independently.

$$h_i^j(t) = \beta_0 d^{-2}(t) = \frac{\beta_0}{H^2 + ||\mathbf{q}_i(t) - \mathbf{p}_j(t)||^2}, \tag{1}$$

where β_0 denotes the channel gain at the reference distance d_0 = 1 meter. When UAV j deals with task k from MU i, the data of the task need to be transmitted from j to i, and the throughput can be denoted as

$$r_{i,k}^j(t) = B\log_2\Big(1 + \frac{a_{i,k}^j(t)Ph_i^j(t)}{\sigma^2 + I_i^u(t)}\Big), \tag{2}$$

where $a_{i,k}^j(t) \in [0, 1]$ is the energy allocation indicator, B is the total bandwith can be assigned to MUs, P denotes the maximum transmit power of UAVs. In this paper, we consider the radio resource used by small cells is overlaid, so mutual interference occurs when the same task is transmitted to different UAV servers, with $I_i^u(t) = \sum_{u\in\mathcal{M},u\neq j} p_u(t)h_i^u(t)$. σ^2 denotes the background noise power, and we assume that it is constant on the slow fading channel.

Hence, the transmission time of task k in t-th interval can be derived as

$$L_{i,k}^{j,s}(t) = \frac{S_i^k(t)}{r_{i,k}^j(t)}, \tag{3}$$

where $S_i^k(t)$ denotes the size of the executed task in t-th time interval. The transmit power consumptions of j-th UAV in t-th time interval can be expressed as

$$p_{i,k}^{j,s}(t) = a_{i,k}^j(t)PL_{i,k}^{j,s}(t). \tag{4}$$

C. Computation model

To simplify the analysis, the computation resource of all UAVs are assumed the same, denoted as C circles per second. Thus, the execution time of task k in j-th UAV can be expressed as

$$L_{i,k}^{j,c}(t) = \frac{F_i^k(t)}{b_{i,k}^j(t)C}, \tag{5}$$

where $b_{i,k}^{j}(t) \in [0,1]$ denotes the computation resource allocation decision. The computation consumptions of task k can be given by

$$p_{i,k}^{j,c}(t) = \gamma_0 b_{i,k}^{j}(t) C. \tag{6}$$

where γ_0 is a constant related to the hardware structure.

D. Problem formulation

By jointly considering computation and energy resource allocation, this paper aims to minimize the total resource consumptions of UAVs under the constraint of MUs' QoS in terms of latency. The optimization problem can be formulated as follows:

$$\min_{\mathbf{a},\mathbf{b}} \omega_0 \sum_{j=1}^{M}\sum_{t=1}^{T} l_{i,j,t} p_{i,k}^{j,s}(t) + (1-\omega_0) \sum_{j=1}^{M}\sum_{t=1}^{T} l_{i,j,t} p_{i,k}^{j,c}(t), \tag{7}$$

$$\forall i \in \mathcal{N}, \forall k \in \mathcal{K}$$

$$\begin{aligned}
\text{s.t.}\quad & C1: && a_{i,k}^{j}(t) \in [0,1] \\
& C2: && b_{i,k}^{j}(t) \in [0,1] \\
& C3: && t_{i,k}^{j,c}(t) + t_{i,k}^{j,s}(t) \le D \quad, \forall i \in \mathcal{N}, \forall k \in \mathcal{K} \\
& C4: && ||\mathbf{q}_i(t) - \mathbf{p}_j(t)||^2 \le R^2 \\
& C5: && \sum_{i=1}^{M} l_{i,j,t} \le 1 \quad, \forall j \in \mathcal{M}, \forall t \in \mathcal{T}
\end{aligned}$$

where $l_{i,j,t}$=$\{0,1\}$ denotes whether MU i offloads task to UAV j in time interval t. $C1$ and $C2$ indicate the constraints on radio resource allocation and computation resource allocation, respectively. $C3$ indicates that the total processing time of task k need to satisfy the maximum tolerant time D. Since the radius coverage of UAV is R, $C4$ gurantees the selected MU and UAV in communication range. $C5$ denotes that each UAV can only execute one task in a single time interval.

According to [9], (7) is a mixed integer programm problem (MILP), which can not be solved optimally through optimization schemes. [13] uses Branch and Bound mathematical methods (BB) to give a optimal solution, however, the complexity of BB is equal to exhaustive search, which can not be used in real multi-UAV network. The objective of this paper is to design a intelligent resource allocation algorithm, which enables UAVs to make decisions upon each task arrival distributedly and efficiently.

III. Proposed multi-agent Q-learning Scheme

In this section, we give the detailed description of multi-agent Q-learning algorithm, where all agents conduct a decision algorithm independently but share common Q-learning structure. Fig. 2 shows the structure of the proposed multi-agent Q-learning in multi-UAV network, where each UAV explores and gets reward from the environment distributedly under the same structure. We first formulate the multi-UAV network as a Markov Decision Process (MDP), where the next state of each agent is only related the current state and action. Note that, in this paper, since each UAV makes the resource

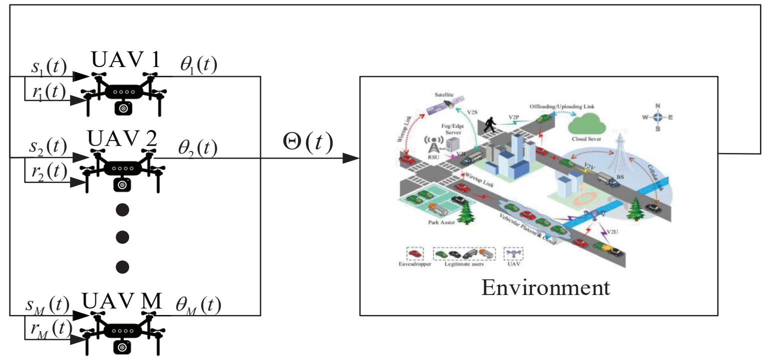

Fig. 2. An illustration of proposed multi-agent Q-learning structure in multi-UAV MEC network.

allocation decision according to its local observations without information exchange between different UAVs, so that the complexity of the proposed multi-agent Q-learning algorithm can be reduced significantly.

Q-learning is a value function based model-free reinforcement learning method, which aims to maximize the long term cumulated average reward of each agent with a specific strategy. The cumulated reward from the current time is defined as return, which can be denoted as

$$R_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1}, \tag{8}$$

where $\gamma \in [0,1]$ represents the discount factor, r_t is the reward value in t-th observation. State value function and action-state value function are defined as the evaluation of a specific strategy, which can be expressed as follows

$$V^{\pi}(s_t) = E_{\pi}\left[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s\right], \tag{9}$$

$$Q^{\pi}(s_t, a_t) = E_{\pi}\left[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s, a_t = a\right], \tag{10}$$

where s_t and a_t denote the state and action in t-th observation, respectively. (10) indicates that a stategy outperforms others if it has the highest state-action value. Furthermore, according to the Bellman equation, the action-state value can be updated as follows

$$Q^{\pi}(s_t, a_t) = E_{\pi}\left[r_{t+1} + \gamma Q^{\pi}(s_{t+1}, a_{t+1}) | s_t = s, a_t = a\right], \tag{11}$$

which indicates that action-state value in current state can be obtained from the action-state value in next state. Q-learning makes decisions based on the Q-table, which consists of state-action values (Q-value) in each time observation. According to [10], we update the Q-table as follows

$$Q_{t+1}(s,a) = Q_t(s,a) + \alpha\Big(r_t + \gamma \max Q_t(s', a') - Q_t(s,a)\Big). \tag{12}$$

In order to apply Q-learning algorithm into our mulit-UAV network, we now define the state, action, reward and the detailed algorithm procedure according to our the optimization objective and constraints.

1) State space: In this paper, we assume that in each time interval t UAV m will execute one task, including downlink transmission and task execution. Since the resource allocation decision is closely related to the property of the task and the SINR $\mathcal{R}$ between UAV and its corresponding MU. We define the state vector of UAV m in time interval t as $\mathbf{s}_j(t) = \Big(X_j(t), Y_j(t), \mathcal{R}_j(t), S_j(t), F_j(t)\Big)$. Each UAV agent only observes its local enviroment and makes actions regardless of the states of other agents, if different UAVs execute a same task simultaneously, each agnet will find that if there are multiple UAVs share the same downlink channel through related media access control (MAC) protocol. Similar as [10], we pre-defined the trajectory of each UAV in each time slot, in order to gurantee the discreteness of state actor, so that Q-learning can be applied to solve the problem.

2) Action space: The action space of UAV j in time interval t can be defined as $\Theta_j(t)$=$(a_j(t), b_j(t))$, where $a_j(t)$, $b_j(t)$ represent the energy and computation resource allocation, respectively. We discretizate the allocation indicators as $a_j(t) \in \{p_j^1(t), p_j^2(t), ..., p_j^T(t)\}$, $b_j(t) \in \{C_j^1(t), C_j^2(t), ..., C_j^T(t)\}$. In each state $\mathbf{s}_t$, each UAV agent chooses an action from action space and obtain a corresponding Q-value. Since Q-learning is a model-free reinforcement learning scheme, there is no transition probability model. After action a_t, the state transition from s_t to s_{t+1}.

3) Reward function: The reward function is an important factor to obtain the optimal Q-table. Moreover, it can modify the action of agents towards a specific objective. In this paper, our aim is to minimize the energy consumptions of UAVs under the constraint of QoS in terms of latency. Based on the optimization objective, the reward fuction can be defined as

$$r_j(t) = \begin{cases} \omega_0(1-a_j(t))p_{i,max}^{j,s}(t) \\ +(1-\omega_0)(1-b_j(t))p_{i,max}^{j,c}(t), \\ \qquad L_{i,k}^{j,s}(t) + L_{i,k}^{j,c}(t) \le D_i^k \\ \\ 0, \qquad L_{i,k}^{j,s}(t) + L_{i,k}^{j,c}(t) > D_i^k \end{cases} \tag{13}$$

where $p_{i,k,max}^{j,s}(t)$, $p_{i,k,max}^{j,c}(t)$ represents the maximum energy consumptions when dealing with task k, respectively. (13) indicates that under the premise of meeting the task delay requirement, the less resource consumption, the higher the reward.

4) Proposed multi-agent Q-learning algorithm: Then we present the proposed multi-agent Q-learning algorithm in Algorithm 1, where each UAV performs the algorithm independently without information exchange. The algorithm consists of two stages, namely learning stage and application stage. In learning stage, we first initialize the state vector $\mathbf{s}_j$, Q-table, α, γ, ϵ. Then we consider a learning period of E episodes. In each time interval of an episode, we first initialize the data size $S_i(k)$ and execution size $F_i(k)$ of task k then we perform the ϵ-greedy policy [10], where i-th MU either chooses a random action with probability ϵ or follows the greedy policy with probability $1-\epsilon$. After the given action, we analyze the total processing time of task k, if the constraint is not satified, $r_j(t)$ is set as 0, otherwise it follows (13). Then we use (12) to update the Q-table until the algorithm converges. Finally we obtain a optimal policy π^* according to the maximum Q-value of each state. In applicaiton stage, each UAV agent j in each state $s_j(t)$ chooses action $\Theta_m(t)$ based on π^*.

Algorithm 1: Proposed multi-agent Q-learning algorithm

1: Initialize $\mathbf{s}_j$, α, γ, ϵ, $Q(s,a)$
2: **for** each episode e = 1,2,...,E **do**
3: **for** each time interval t = 1,2,...,T
4: Initialize $X_j(t)$, $Y_j(t)$, $\mathbf{x}(t)$, $\mathbf{y}(t)$, $h_i^j(t)$, $S_i^k(t)$, $F_i^k(t)$
5: Select an action $a_j(t)$ based on ϵ-policy
6: Compute the total time consumptions on task k
7: Compute the reward $r_j(t)$ according to (13)
8: Update the Q-table, according to (12)
9: Update state $s_j(t) \rightarrow s_j(t+1)$
10: **end for**
11: episode = episode + 1
12: **end for**
13: Output a policy $\boldsymbol{\pi}^*$ for j-th UAV

IV. Simulation Results

In this section, we give numerical results to verify the effectivenss of the proposed multi-agent Q-learning method. We do the simulations through Matlab 2015 in Windows 10 operating system. Specifically, we set the multi-UAV network in a 800m $\times$ 800m field with $N = 100$, $M = 3$, $H = 100$m. For communication related parameters, β_0 is set as 1.42×10^{-4}, the bandwidth is $B = 20$MHz, the noise power is $\sigma^2 =$-90dBm, the maximum transmitt power is $P_{max} = 23$dBm. For computation parameters, The computation capacity of UAVs is $C = 30$GHz, $\gamma_0 = 1 \times 10^{-10}$, time constraints $D = 0.1$s. For reinforcementl learning parameters, the discount factor is $\gamma = 0.9$, the reward factor is $\omega_0 = 0.4$. Fig. 2 show the convergence performance of the proposed multi-agent Q-learning alogrithm with different ϵ values, where $M = 3$, $N = 100$, $\alpha = \frac{1}{(t+c_\alpha)^{\phi\alpha}}$ [10]. ϵ measures the exploration and exploition of the learning algorithm. ϵ is set as 1 in the initial training stage, since the Q-table is empty, agents need to select actions randomly in order to explore the environment. After some training epsidoes the agents tend to choose the maximum Q-value to obtain a higher reward. From Fig. 2, we obtain the proper ϵ value for the convergence of the algorithm, the average maximum Q-Value denotes the maximum Q-Value in each time interval t among different actions.

Fig. 3 gives the comparison of the total weighted resource consumptions with different benchmark algorithms, where $\epsilon = 0.5$, $M = 3$, $N = 100$. The greedy algorithm allocates the whole computation and energy resoure for each offloading task, which promise the latency constraint of tasks to be satisfied, however, with the most resource cosumptions. The random algorithm randomly allocates resource for each task,

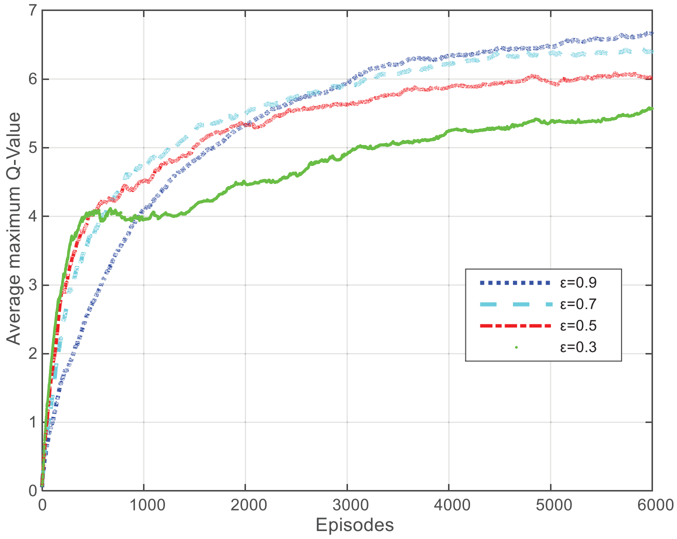

Fig. 3. Comparisons of average Q-Value with different ϵ values, where $M = 3$, $N = 100$, $\alpha = \frac{1}{(t+c_\alpha)^{\phi_\alpha}}$, $\gamma = 0.9, c_\alpha = 0.5, \phi_\alpha = 0.8$

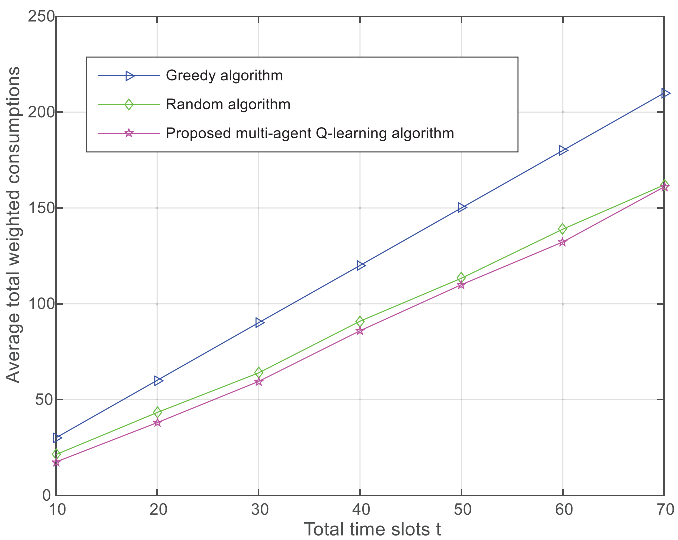

Fig. 4. Comparisons of total weighted resource consumptions with different benchmark algorithms, where $M = 3$, $N = 100$, $\epsilon = 0.5$, $\alpha = \frac{1}{(t+c_\alpha)^{\phi_\alpha}}$, $\gamma = 0.9, c_\alpha = 0.5, \phi_\alpha = 0.8$.

resulting in relatively low consumptions but it is not guaranteed to meet the latency constraint. The proposed algorithm takes actions in each time interval based on the maximum Q-Value while satisfying the latency constraint.

V. Conclusion

In this paper, we study the computation and energy resource allocation in a multi-UAV MEC network. The objective is to enable each UAV to execute tasks offloaded from the MUs with minimized resource consumptions while meeting the latency requirement of MUs. We first give the network, communication and computation model and formulate the resource allocation problem into an MILP problem. Then we propose a mulit-agent Q-learning algorithm to solve the problem with low complexity, which allows each UAV makes decisions without information exchange between different UAVs. Finally, simulations results are given to verify the feasiblity and effectiveness of our Q-learning based algorithm compared with greedy and random algorithms. The processing of continious states space and actions space in the proposed algorithm will be left as our future works.

References

[1] S. Zhang, S. Shi, S. Gu and X. Gu, "Power Control and Trajectory Planning Based Interference Management for UAV-Assisted Wireless Sensor Networks," *IEEE Access*, vol. 8, pp. 3453-3464, January 2020.

[2] N. Zhang, S. Zhang, P. Yang, O. Alhussein, W. Zhuang, and X. Shen, "Software Defined Space-Air-Ground Integrated Vehicular Networks: Challenges and Solutions," *IEEE Communication Magazine*, vol. 55, no. 7, pp. 101-109, July 2017.

[3] Y. Zeng, R. Zhang, and T. Lim, "Wireless Communications with Unmanned Aerial Vehicles: Opportunities and Challenges," *IEEE Commun. Magazine*, vol. 54, no. 5, pp. 36–42, May 2016.

[4] Y. Mao, *et al.*, "A Survey on Mobile Edge Computing: The Communication Perspective," *IEEE Communications Survey and Tutorials*, vol. 19, no. 4, pp. 2322–2358, Fourthquarter 2017.

[5] M. Mozaffari, W. Saad, M. Bennis and M. Debbah "Efficient Deployment of Multiple Unmanned Aerial Vehicles for Optimal Wireless Coverage,", *IEEE Communications Letters*, vol. 20, no. 8 pp. 1647–1650. August 2016.

[6] H. Wu, X. Tao, N. Zhang, and X. Shen, "Cooperative UAV Cluster Assisted Terrestrial Cellular Networks for Ubiquitous Coverage," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 9, pp. 2045-2058, September 2018.

[7] D. Yang, Q. Wu, Y. Zeng, R. Zhang, "Energy Tradeoff in Ground-to-UAV Communication via Trajectory Design," *IEEE Trans. on Vehicular Technology*, vol.67 , no. 7 , pp. 1647–1650. July 2018.

[8] Y. Zeng, R. Zhang, "Energy-Efficient UAV Communication With Trajectory Optimization," *IEEE Trans. on Wireless Communications*, vol. 16, no. 6, pp. 3747–3760, June 2017.

[9] S. Jeong, O. Simeone, J. Kang, "Mobile Edge Computing via a UAV-Mounted Cloudlet: Optimization of Bit Allocation and Path Planning," *IEEE Trans. on Vehicular Technology*, vol. 67, no. 3, pp. 2049–2064, March 2018.

[10] J. Cui, Y. Liu, A. Nallanathan, "Multi-Agent Reinforcement Learning-Based Resource Allocation for UAV Networks,"*IEEE Trans. on Wireless Communications*, vol. 19, no. 2, pp. 729–743, February 2020

[11] C. H. Liu, Z. Chen, Y.Zhan "Energy-Efficient Distributed Mobile Crowd Sensing: A Deep Learning Approach," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 6, pp. 1262–1276, June 2019.

[12] X. Chen, L. Jiao, W. Li, X. Fu "Efficient Multi-User Computation Offloading for Mobile-Edge Cloud Computing," *IEEE/ACM Trans. on Networking*, vol. 24, no. 5, pp. 2795–2809, October 2016.

[13] Tuyen X. Tran, Dario Pompili, "Adaptive Bitrate Video Caching and Processing in Mobile-Edge Computing Networks," *IEEE Trans. on Mobile Computing*, vol. 18, no. 9, pp. 1965–1978, September 2019.

[14] Y. Du, K. Wang, K. Yang and G. Zhang,"Energy-Efficient Resource Allocation in UAV Based MEC System for IoT Devices," *2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 2018, pp. 1-6.