Dynamic Network Access for Multi-UAV Networks: A Cloud-Assisted Learning Algorithm

Xiaodu Liu[1](✉), Yitao Xu[1], Yanlei Duan[2], Dianxiong Liu[1], and Zhiyong Du[3]

[1] College of Communications Engineering, Army Engineering University of PLA, Nanjing, China
lxdlgdx@163.com
[2] No. 92274 Troops of PLA, Kunming, China
[3] National University of Defense Technology, Changsha, China

Abstract. In this paper, we study the strategy of UAV dynamic network access in the large-scale UAVs swam. We model the master UAV providing communication coverage for the small UAVs which transformed the large-scale UAVs communication problem into the optimization problem. Compared to the traditional ground user network access, the characteristic of UAV's mobility have been considered and each UAV have chance to move to any master UAV for better service. We propose a joint optimization for the throughput and flight loss. Due to the limitation of flight loss, the UAVs can not fly to different networks many times for learning. We set up a load aggregator cloud to help the UAVs simulate the results of each decision. We propose a dynamic network access algorithm based on SLA which is proved to achieve stable solutions with dynamic and incomplete information constraint. The simulation results show that this algorithm can converge to the optimal solution. Also, it is shown that the algorithm has strong robustness and can get good utility than other algorithms regardless of how the environment changing.

Keywords: Dynamic network access · Multi-UAV communication Cloud-assisted · SLA

1 Introduction

The application of the intelligent unmanned aerial vehicles (UAVs) is expanding with the development of the UAV technology [1]. Nowadays, large-scale UAVs and UAV-assisted communication are playing important roles in various fields. In 2017, nearly 300 UAVs flied together to create a dreamlike stage in the USA super bowl. Recently, the company EHang has also achieved the formation of 1000 UAVs. However, the focus of large-scale UAVs is more on the collaborative control [2]. UAV-assisted communication is also only considered as the air base station to assist ground communication [3–7]. However, how to solve the problem of large-scale UAVs' inter-domain communication and how to deal with the

L. Meng and Y. Zhang (Eds.): MLICOM 2018, LNICST 251, pp. 358–369, 2018.
https://doi.org/10.1007/978-3-030-00557-3_36

relationship between the master UAV and the other small UAVs are not having clear answers. There is still relatively little research on the combination of large-scale UAVs and UAV-assisted communication.

In the large-scale UAVs scenario, the communication between UAVs is interactive, resulting in a series of coupling optimization problems. Most of the existing studies have looked at how the UAVs serving ground users. The paper [8] assigns UAV to specific region as the relay, thus enhancing the communication capability of heterogeneous wireless network. However, the work in paper [8] is limited to the uniform distribution of ground users, and does not fully consider the fairness of users' choice in the case of network congestion. In the paper [9], the base station which associated with UAV is determined with the goal of minimizing UAV's transmit power and satisfying the user's rate requirement. The UAV has been used as a mobile base station to serve ground users, and achieves the goal of maximizing the minimum throughput of each ground user by accessing different users in [10]. The authors [11] considered the multi-UAV system and added power control based on the paper [10]. Most of the current researches only consider the UAV as the aerial base station to serve the ground users. The communication problem of the UAV itself is not considered.

In order to solve the problem of large-scale UAVs communication, the relationship between the small UAV and its upper master UAV is analogous to the relationship between the user and the network in the traditional network access scenario. That is, the master UAVs provide communication coverage for the small UAVs. The network connectivity of a UAV-assisted network has been optimized in [12] and [13]. However it does not consider the situation of master UAVs assisting the small UAVs with communication. And there is a lack of research on network access in multi-UAV system. In our paper, the communication problem of the large-scale UAVs is transformed into the optimization problem of the network access. Different from the traditional network access problem, only users in the overlapping areas of the network can choose the access network [14]. Because of the mobility of the UAV, no matter where it is currently located, it can move to the range of any master UAVs to find better service. Such a scenario is more equitable than the original scenario, not just the users in the network's overlapping areas but every one has opportunity to make decisions.

We model the application scenarios as a master UAV has been crashed, and the small UAVs in it are not served. We command the access of small UAVs to other networks through ground control center. The small UAVs of other mater UAVs can also change their location for better service after receiving the impact from outside small UAVs. This brings us more challenges to the study of network access problems. How to define the flight loss of UAV is the first problem to be faced with UAV's mobility. The paper [15–17] considered the UAV energy saving communication but did not consider the flight energy required by the movement. The flight loss have been considered in [18] to solve the problem of energy-efficient UAV communication, but our paper focuses more on the completion of UAV's communication task. We combine the traditional network access throughput optimization with the flight loss of UAV which becomes a joint optimization for the throughput and flight loss. We propose a dynamic network access algorithm based on SLA [19] to find the tradeoff between throughput and UAV flight loss. Due to the limitation of flight loss, the UAVs can not fly to different networks many times for learning. We set up a load aggregator cloud to help the UAVs simulate the results of each decision. The main contributions of our work can be summarized as follows:

- We solve the problem of UAV group communication by clearing the relationship between the master UAV and the small UAV. We model the master UAV providing communication coverage for the small UAVs which transformed the large-scale UAVs communication problem into the optimization problem of the network access. We solve the problem as some master UAVs have been crashed, how the small UAVs to make decisions to guarantee the communication quality.
- We consider the characteristic of UAV's mobility and we make it possible for each UAV to move to any master UAV for better service. We combine the traditional network access throughput optimization with the flight loss of UAV which becomes a joint optimization for the throughput and flight loss.
- A dynamic network access algorithm based on SLA has been proposed to get the Nash equilibrium and find the tradeoff between throughput and UAV flight loss. Due to the limitation of flight loss, the UAVs can not fly to different networks many times for learning. We set up a load aggregator cloud to help the UAVs simulate the results of each decision.

The remainder of this paper is organized as follows. In Sect. 2, we present the system model and problem formation. In Sect. 3, we propose a cloud-assisted learning algorithm based on SLA to solve the problem. Further, we present the simulation results and performance analysis in Sect. 4. Then, we draw the conclusion in Sect. 5.

2 System Model and Problem Formulation

2.1 System Model

We consider a UAV formation consisting of $\mathcal{N} = \{1, 2, ..., N\}$ master UAVs and $\mathcal{M} = \{1, 2, ..., M\}$ small UAVs. The master UAVs provide communication coverage for the small UAVs. The small UAVs are denoted as users which share the master UAVs' resource to send information. Each small UAV has communication tasks to finish, so it must be covered by the master UAV. In this paper, we only consider the case that the number of small UAVs is much larger than the number of master UAVs, so we put three master UAVs N_1, N_2 and N_3 in the system which serve this areas' small UAVs. We denote two kinds of small UAVs in this system. One of them are already in the range of a master UAV, the others do not covered by any master UAVs because their master UAVs have been crushed. In traditional network access scenarios, users are not able to access the network if they are outside of the communication coverage. But this is not a question in the UAV system. Because of the dynamics of UAVs, we can deploy the small UAVs through the control center to move to the coverage of any master UAVs to get service. Due to the addition of external users, the small UAVs which originally in the range of some master UAVs may get less resources for communication. Therefore they also can move to the coverage of different master UAVs to achieve better payoff. Each small UAV has multiple available master UAVs, but the small UAVs can only access one master UAV at any time (Fig. 1).

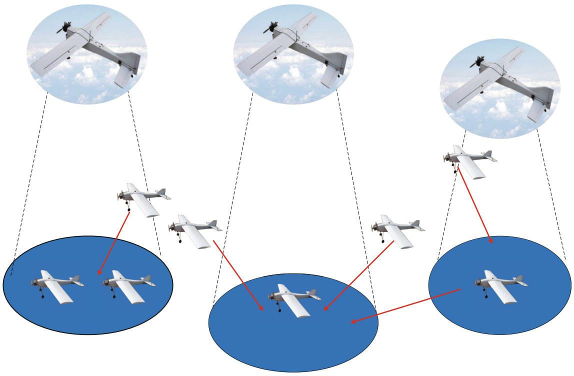

Fig. 1. Dynamic network access system consist of three master UAVs and two kinds of small UAVs. One of the UAVs are already in the range of a master UAV, the others do not covered by any master UAVs because their master UAVs have been crushed.

2.2 Problem Formulation

In our model there are two actions of each small UAV: Firstly, the small UAV is in the coverage of any master UAVs and it does not want to change the master UAV. The throughput of the small UAV accesses the network depends on a number of factors, including the physical layer data transmission rate, the load connect to the network, and the resource allocation strategy adopt by the access network. This paper considers a resource allocation strategy which based on proportional equity. Under this mechanism, the average throughput of the small UAV m to access the master UAV n can be achieved as [14].

$$g_m = \theta_m = \frac{w_m R_{m,n}}{W_n}, \tag{1}$$

where $R_{m,n}$ is the peak data rate between small UAV m and master UAV n, w_m is small UAV m's load, and $W_n = \sum\limits_{i \in \mathrm{M}_n} w_i$ is the total weight of small UAVs that accessed master UAV n. This above model combines many practical considerations. The peek data rate $R_{m,n}$ reflects physical characteristics such as wireless channel quality and modulation encoding. Secondly, the discount factor $\frac{w_m}{W_n}$ reflect the characteristics of the users sharing the network's resources.

The user's weight can distinguish the application types of different small UAVs in the same network. Such that the reconnaissance UAV need to send some photos and videos, so it needs more resources. The attack UAV only needs to receive real-time message, so it has low throughput requirements. In conclusion, the weight of the small UAV depends on the type of the task it needs to do.

Secondly, if the small UAV is not covered by any master UAV so it needs to move to any master UAV to look for the communication service. Some small UAVs which already have master UAVs to access want to move to other master UAVs for better payoff. They all have flight loss caused by movement. The cost E_c can be denoted as [17]

$$E_c = \frac{d_{m,n}}{V}(c_1 V^3 + \frac{c_2}{V}), \tag{2}$$

where $d_{m,n}$ equals to the distance between the small UAV and the communication coverage of the master UAV. That is also what we need to optimize. V is a given UAV speed, c_1, c_2 are constant which related to the weight of the aircraft and the external wind force. The first term in the speed loss is proportional to the third power of the velocity, which is the resistance loss caused by air friction during the flight. The second inverse is the energy loss to overcome the lift. So the utility function of the small UAV m to move to access the master UAV n can be achieved as [8]

$$g_m = \frac{w_m R_{m,n}}{W_n} - \beta \frac{d_{m,n}}{V}(c_1 V^3 + \frac{c_2}{V}), \tag{3}$$

where β is the normalized coefficient. The importance of flight loss can be expressed by changing the size of β. If we improve the value of β means that we do not want the small UAV to change its location. This utility function represents the intentions of each small UAV. We must find the tradeoff between the throughput and flight loss.

We denoted the user-network correlation as M_0. So we defined the utility as the social welfare

$$U_{social}(\mathrm{M}_0) = \sum_{m \in \mathcal{M}} g_m(\theta_m, E_c). \tag{4}$$

The target of the system is to optimize the relationship between the small UAVs and the master UAVs to maximize the net utility which is denoted as

$$(P1) : \max U_{social}(\mathrm{M}_0), \tag{5}$$

3 Dynamic Network Access Algorithm

3.1 SLA: Stochastic Learning Automata

Due to the dynamic and incomplete information constraints, most existing algorithms can not be applied [19]. Based on the SLA (Stochastic learning automata), we propose a new algorithm. Stochastic learning automata is a finite machine that interacts with an unknown environment and tries to learn the best practices provided by the environment [20]. SLA updates the selection probability dynamically through the feedback from each learning and keeps doing the probability update until users reach stable conditions. Due to the limitation of flight loss, the UAVs can not fly to different networks many times for learning. We set up a control center using the SLA algorithm to help the UAVs simulate the results of each decision. When all users converge to Nash equilibrium, the control center deploy the small UAVs to move to the coverage of the specified master UAVs to get service. So as to realize the distributed solution for the original problem.

We extend the dynamic network access game to the form of mixed strategy. We denote that $P = (p_1, ..., p_M)$ is the mixed strategy for all users, where $p_m = (p_{m1}, ...p_{mN})$ is the probability vector when small UAV m access any master UAVs. And p_{mn} is probability of the small UAV m to access the master UAV n. We also denote $h_{nm}(P)$ as the average throughput of small UAV m when the small UAV m access the master UAV n ($a_m = n$) and other small UAVs use the mixed strategy.

$$h_{mn}(P) = u_m(a_1, ..., a_{m-1}, n, a_{m+1}, ..., a_M) \tag{6}$$

According to the number of small UAVs in each master UAV and the location of each small UAV, each small UAV can achieve random return value at the end of each time slot. The small UAV updates its mixed strategy on this basis.

3.2 A Cloud-Assisted Learning Algorithm Based on SLA

Due to the limitation of flight loss, the UAVs can not fly to different networks many times for learning. We set up a load aggregator cloud to help the UAVs simulate the results of each decision. Compared with the existing learning framework, UAVs in the cloud support framework do not need to actually perform frequent network switching, but only report the decision information to the cloud. There is a load aggregator cloud that is responsible for collecting decision information for all UAVs and sending "virtual network load information" to UAVs. Unlike the centralized optimization method adopted in literature [17], the network load aggregation cloud does not make any decision about the allocation of wireless resources. Therefore, the proposed cloud learning framework can also be applied to similar distributed learning algorithms (such as SLA) and improve its operational efficiency. In our paper, UAVs are willing to submit all necessary information to the cloud. Including rate information $R_{m,n}$ and demand information θ_m. After collecting the UAV's information, the network load aggregation cloud represents the benefit of the UAVs, simulating multiple UAVs to run the SLA learning algorithm. The algorithm process is as follows:

The cloud-assisted learning algorithm based on SLA which we proposed has the following characteristics: (i) this algorithm is not a rigid decision, but select the strategy according to a certain probability randomly in the candidate actions; (ii) not blindly choose the optimal utility of action, but improve the access probability of which action has better payoff softly; (iii) the probability of the

Algorithm 1. A Cloud-assisted Learning Algorithm based on SLA

The User Side: .
Step 1: Each user m access to any network and register in the load aggregator cloud.
Step 2: User m reports to the cloud rate information $R_{m,n}$ and demand information θ_m . And wait for the network selection a_m return from the cloud .
The Load Aggregator Cloud Side:
Step 1: Receive all user reports. Maintain a cumulative decision distribution vector for each user.
Step 2: Run the dynamic network access algorithm based on the SLA and update the network access probability vector until the end of a scheduled stop rule.
The SLA part:
Initialize: . Set the number of iterations k=1 and set the initial network access probability as $p_{mn}(k) = 1/N, \forall m \in \mathcal{M}, n \in \{1, ..., N\}$, then generate small UAVs $\mathcal{M} = \{1, 2, ..., M\}$ randomly within or without the range of the master UAV.
Loop for k = 0,1,...
Step 1: At the beginning of the time slot k, firstly the small UAV m which are not in the range of any master UAVs access the master UAV $a_m(k)$ according to its current network access probability vector $p_m(k)$. Secondly the other small UAVs do the same steps.
Step 2: On the basis of the network access in Step1, the UAV do the network awareness and access the master UAV. At the end of the current slot, the small UAV obtains the random return which is calculated by (3) and set the return to the small UAVs.
Step 3: All the small UAVs update the probability of network access according to the following rules:

$$\begin{aligned} p_{mn}(k+1) &= p_{mn}(k) + bg_m(k)(1 - p_{mn}(k)), n = a_m(k) \\ p_{mn}(k+1) &= p_{mn}(k) - bg_m(k)p_{mn}(k), \quad n \neq a_m(k) \end{aligned} \tag{7}$$

where $0 < b < 1$ is the iteration step length, $g_m(k)$ is the normalized throughput.
Step 4: For any small UAV, the corresponding network access probability vector has an element that is close to 1, if greater than 0.99, and the algorithm go back to Step 2. Otherwise, go back to Step 2. Until all the small UAVs' network access probability are close to 1, then the algorithm ends.
Loop end

network access is updated based on the random return value of each slot. The random return is the reinforcement signal of this algorithm. Leaving the small UAVs more exploration space in the dynamic network access system, which can effectively get rid of the local optimal dilemma. According to the real-time change of the network environment, the user can improve the access probability of the current optimal decision at each time slot. So that the probability of optimal access will eventually converge.

4 Simulation Results and Analysis

In this scenario we generate two circular networks as the coverage of master UAVs with a diameter of 500 m. A circular network with a diameter of 300 m has been set to distinguish the different kinds of master UAVs. Then we randomly generate each 10 nodes in first two networks and 5 nodes in the third network as the small UAVs. Twenty nodes have been set as the small UAVs which are not in range of any master UAVs. The number of iterations is set as 500 in the simulation. The location of each small UAV is generated randomly each time. The distance between each small UAV and the distance from the small UAV to the network is the decisive factor of the system. We set up a control center to help the UAVs simulate the results of each decision. We set the speed of UAV as $V = 10dB$, the constant $c_1 = 9.26 \times 10^{-4}$ and $c_2 = 2250$. The learning step is set as 0.5. The link transmission peak rate of the master and the leader UAV is calculating by the Shannon formula $R = B\log_2(1 + P/(B * \sigma^2))$.

The simulation results mainly include the following two parts. The first part is the convergence of the simulation algorithm. In particular, for any selected user, we study the change of the network access probability with the iteration number. In addition, in order to reflect the overall convergence of the system, we also study the change of the number of users in different networks and give the convergence network topology. The second part gives the performance evaluation of the algorithm and compares the utility function of the four methods: (i) The dynamic network access algorithm based on SLA, (ii) The centralized algorithm, (iii) Random access and (iv) The closet network access algorithm. In the centralized algorithm, we assume that each user knows all the message of the system. They know the rewards of access any network and all choose the best one to access. As the users only know the location of the network, the random access algorithm choose the network as the same probability and the closet network access algorithm choose the closet network to access are all feasible.

4.1 Convergence Behavior

Firstly, we study the variation of network access probability with iteration number. Simulation of a dynamic network access system with twenty-five dynamic loads, twenty dynamic users and three networks. And the bandwidth between this three networks is 2:2:1 which represents the size of the communication coverage. The UAV's flight loss is several orders of magnitude larger compared with the throughput. The goal of this paper is to ensure the quality of communication, so we set the normalized coefficient β as 0.015.

The Fig. 2 gives the convergence of network access probability of any selected user. We can see that this user's network access probability vector is approximately running 140 iterations from $\{1/3, 1/3, 1/3\}$ to $\{0, 1, 0\}$. That is to see, this small UAV finally choose the master UAV N_2 to access. The Fig. 4 shows the network topology after the algorithm is convergence. Each users have network to get service. Some users originally in the N_1 may move to N_2 and N_3 to find

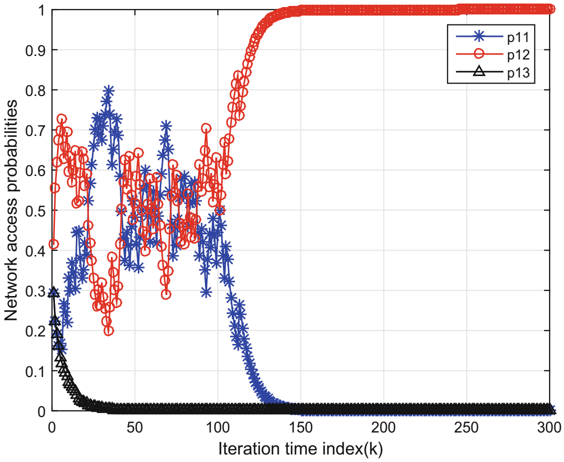

Fig. 2. The convergence process of arbitrary user network access probability.

better service, some users do not change their location. The number of the users in each network is related to the network capacity.

4.2 Performance Analysis

In this section, we study how the different parameters influence the algorithms. Figure 3 represents the relationship between the number of the dynamic users with the social welfare. All these four algorithm, we do 500 independent simulations and then take the average. As can be seen form the figure, the algorithm we proposed is far more efficient than the random access algorithm and the closet network access algorithm no matter how many dynamic users in this system. The algorithm we proposed also get the same social welfare as the centralized algorithm at any time. The reasons are as follows: (i) The proposed algorithm can converge to the optimal solution, and the users will be scattered on different networks. (ii) In the random access algorithm, as the number of dynamic users increases, it is possible to make unreasonable decisions. So this situation may make low throughput and high flight loss. (iii) In the closet network access algorithm, as the number of dynamic users increases, some networks may be accessed by multiple users which make the congestion of the network.

In Fig. 4, we change the normalized coefficient β of the flight loss. This change only has a little effect on our algorithm. The dynamic network access algorithm based on SLA can also get high social welfare like the centralized algorithm. With the low level of the normalized coefficient β, the random access algorithm and the closet network access algorithm may get the social welfare close to the algorithm we proposed. But effected by the improve of the normalized coefficient β, the social welfare get by the random access algorithm and the closet network access algorithm all drop faster. It can be seen, the algorithm we proposed has strong robustness and can get good utility regardless of the environment change.

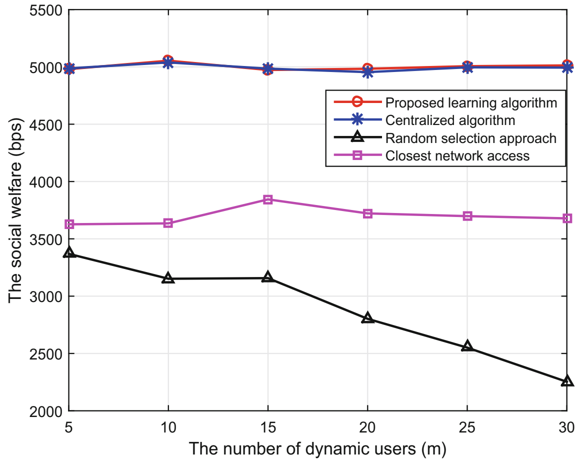

Fig. 3. The utility comparison of the four algorithms with different user numbers.

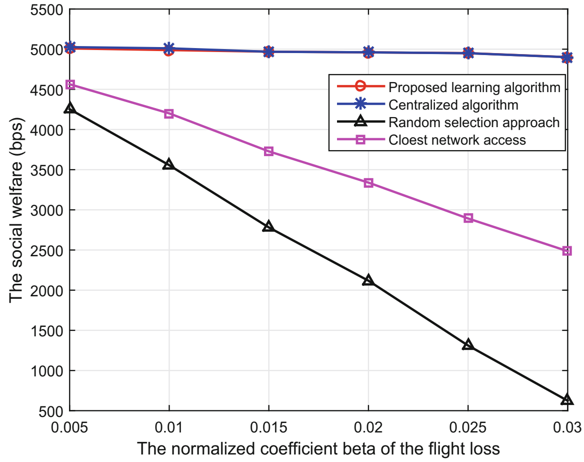

Fig. 4. The utility comparison of the four algorithms with different normalized coefficient β of the flight loss.

5 Conclusion

This paper put forward the network scenario of the master UAVs serving the small UAVs for communication. We transformed the UAV group communication problem into the optimization problem of the network access. We consider the characteristic of UAV's mobility and make it possible for each UAV to move to any master UAV for better service. We combine the traditional network access throughput optimization with the flight loss of UAV which becomes a joint optimization for the throughput and flight loss. We proposed a dynamic network access algorithm based on SLA to get the Nash equilibrium. Due to the limitation of flight loss, the UAVs can not fly to different networks many times for learning. We set up a load aggregator cloud to help the UAVs simulate the results of each decision. The simulation shows that the algorithm we proposed can get good utility than other algorithms regardless of how the environment changes. The algorithm realizes the robust optimization in dynamic unknown environment.

Acknowledgment. This work was supported by the National Science Foundation of China under Grant No. 61771488, No. 61671473, No. 61631020 and No. 61401508, the in part by Natural Science Foundation for Distinguished Young Scholars of Jiangsu Province under Grant No. BK20160034, and in part by the Open Research Foundation of Science and Technology on Communication Networks Laboratory.

References

1. Zeng, Y., Zhang, R., Lim, T.J.: Wireless communications with unmanned aerial vehicles: opportunities and challenges. IEEE Commun. Mag. **54**(5), 36–42 (2016)
2. Kuriki, Y., Namerikawa, T.: Formation control with collision avoidance for a multi-UAV system using decentralized MPC and consensus-based control. In: European Control Conference (ECC), Linz, pp. 3079–3084 (2015)
3. Mozaffari, M., Saad, W., Bennis, M., Debbah, M.: Drone small cells in the clouds: design, deployment and performance analysis. In: IEEE Global Communications Conference (GLOBECOM), San Diego, CA, pp. 1–6 (2015)
4. Mozaffari, M., Saad, W., Bennis, M., Debbah, M.: Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage. IEEE Commun. Lett. **20**(8), 1647–1650 (2016)
5. Al-Hourani, A., Kandeepan, S., Lardner, S.: Optimal LAP altitude for maximum coverage. IEEE Wirel. Commun. Lett. **3**(6), 569–572 (2014)
6. Lyu, J., Zeng, Y., Zhang, R., Lim, T.J.: Placement optimization of UAV-mounted mobile base stations. IEEE Commun. Lett. **21**(3), 604–607 (2017)
7. Bor-Yaliniz, R.I., El-Keyi, A., Yanikomeroglu, H.: Efficient 3-D placement of an aerial base station in next generation cellular networks. In: 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, pp. 1–5 (2016)
8. Sharma, V., Bennis, M., Kumar, R.: UAV-assisted heterogeneous networks for capacity enhancement. IEEE Commun. Lett. **20**(6), 1207–1210 (2016)
9. Mozaffari, M., Saad, W., Bennis, M., Debbah, M.: Optimal transport theory for power-efficient deployment of unmanned aerial vehicles. In: 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, pp. 1–6 (2016)
10. Wu, Q., Zeng, Y., Zhang, R.: Joint trajectory and communication design for UAV-enabled multiple access. In: 2017 IEEE Global Communications Conference, GLOBECOM 2017, Singapore, pp. 1–6 (2017)
11. Wu, Q., Zeng, Y., Zhang, R.: Joint trajectory and communication design for multi-UAV enabled wireless networks. IEEE Trans. Wirel. Commun. **PP**(99), 1
12. Han, Z., Swindlehurst, A.L., Liu, K.J.R.: Optimization of MANET connectivity via smart deployment/movement of unmanned air vehicles. IEEE Trans. Veh. Technol. **58**(7), 3533–3546 (2009)
13. Kim, S., Oh, H., Suk, J., Tsourdos, A.: Coordinated trajectory planning for efficient communication relay using multiple UAVs. Control Eng. Pract. **29**, 42–49 (2014)

14. Du, Z., Wu, Q., Yang, P.: Dynamic user demand driven online network selection. IEEE Commun. Lett. **18**(3), 419–422 (2014)
15. Kandeepan, S., Gomez, K., Reynaud, L., Rasheed, T.: Aerial-terrestrial communications: terrestrial cooperation and energy-efficient transmissions to aerial base stations. IEEE Trans. Aerosp. Electron. Syst. **50**(4), 2715–2735 (2014)
16. Mozaffari, M., Saad, W., Bennis, M., Debbah, M.: Optimal transport theory for power-efficient deployment of unmanned aerial vehicles. In: 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, pp. 1–6 (2016)
17. Li, K., Ni, W., Wang, X., Liu, R.P., Kanhere, S.S., Jha, S.: Energy-efficient cooperative relaying for unmanned aerial vehicles. IEEE Trans. Mobile Comput. **15**(6), 1377–1386 (2016)
18. Zeng, Y., Zhang, R.: Energy-efficient UAV communication with trajectory optimization. IEEE Trans. Wirel. Commun. **16**(6), 3747–3760 (2017)
19. Xu, Y., Xu, Y., Anpalagan, A.: Database-assisted spectrum access in dynamic networks: a distributed learning solution. IEEE Access **3**, 1071–1078 (2015)
20. Kaelbing, L.P., Littman, M.L., Moore, A.W.: Reinforcement learning: a survey. J. Artif. Intell. Res. **4**(1), 237–285 (1996)
21. Xue, P., Gong, P., Park, J.H., Park, D., Kim, D.K.: Radio resource management with proportional rate constraint in the heterogeneous networks. IEEE Trans. Wirel. Commun. **11**(3), 1066–1075 (2012)