IOP Conf. Series: Materials Science and Engineering **719** (2020) 012073 doi:10.1088/1757-899X/719/1/012073

Trajectory Design and Mode Selection for a Cellular-based UAV Traffic Monitoring System

Kai Wang[2], Fanyi Wu[1], Kaiming Xu[1], Jianjun Wu[1]

[1]School of Electronics Engineering and Computer Science, Peking University, Beijing, China.
[2]Institute of War Studies, Academy of Military Science, Beijing, China.
E-mail: 1801213629@pku.edu.cn

Abstract. In this paper, we propose to enable cellular networks to support a UAV traffic monitoring system, in which multiple UAVs are required to record videos from the target roads, and then transmit their video data back to the ground stations over the cellular spectrum. In this system, the video data can be transmitted directly to the ground stations bypassing the base station (BS), or via the traditional cellular links. Each UAV should select its transmission mode from the direct mode and the cellular mode as it preforms traffic monitoring tasks. Since the mode selections are influenced by the trajectories of UAVs, we study the trajectory design problem for UAVs in consideration of their transmission modes under a reinforcement learning (RL) framework. Then, we propose a trajectory design algorithm to solve this problem. Simulation results show that our proposed algorithm outperforms the single-agent one.

1. Introduction

With high mobility and low operational cost, unmanned aerial vehicle (UAV) has been regarded as an emerging facility to provide real-time monitoring for road traffic [1]. When performing traffic monitoring tasks, UAVs should send the recorded video data to the ground stations effectively. However, in the current unmanned aerial systems for traffic monitoring, the video data is transmitted over the unlicensed spectrum, which is difficult to ensure satisfactory quality-of-service (QoS). To tackle this challenge, the cellular-enabled UAV network is proposed as a promising solution to support UAV communications [2]. In this network, UAVs can send the video data to the base station (BS), after which the BS transmits the received data to the corresponding ground stations. Moreover, when UAVs are close to their ground stations, they can directly transmit their sensory data to the ground stations bypassing the BS, which can further improve the throughput in the system.

In this paper, we investigate a cellular-based UAV traffic monitoring system, in which UAV can transmit their video data through the direct mode and the cellular mode. We consider an orthogonal frequency division multiple access (OFDMA) network consisting of multiple UAVs performing traffic monitoring tasks, where the direct mode work as an overlay to the cellular mode. Since the transmission mode selection of UAVs is influenced by their trajectories of the UAV, we aim to maximize the total utility in the system by designing their trajectories in consideration of their transmission mode.

Nevertheless, since the trajectories of UAVs are coupled with their transmission modes, it is challenging to investigate the trajectory design problem for UAVs. To tackle with this challenge, we adopt a sense-and-send protocol to schedule multiple UAVs, which can be analysed by nested bi-level Markov chains. By this means, the transmission modes of UAVs are determined simultaneously as we design their trajectories. Under this condition, the trajectory design problem can be regarded as a

Markov decision problem (MDP), and thus, we can solve the trajectory design problem using reinforcement learning (RL). Then, we propose a Trajectory design algorithm based on deep Q-network (DQN) to optimize policies for UAVs.

In the literature, several works have investigated the UAV communications over cellular networks. In [3], the authors jointly optimized trajectories, sensing locations and scheduling for UAVs performing sensing tasks in the cellular-enabled UAV network, where the BS was required to collect sensory data from UAVs. Authors in [4] jointly optimized subchannel allocation and flying speed for UAVs in a UAV-to-X cellular network, where UAVs could set up communication links with the BS as well as other UAVs However, as an important practical scenario for traffic monitoring, the direct communications between UAVs and ground stations are lack of consideration. Therefore, different from these works, we consider a cellular-based UAV traffic monitoring system, in which UAV transmissions can be conducted over both the direct mode and the cellular mode.

2. System Model

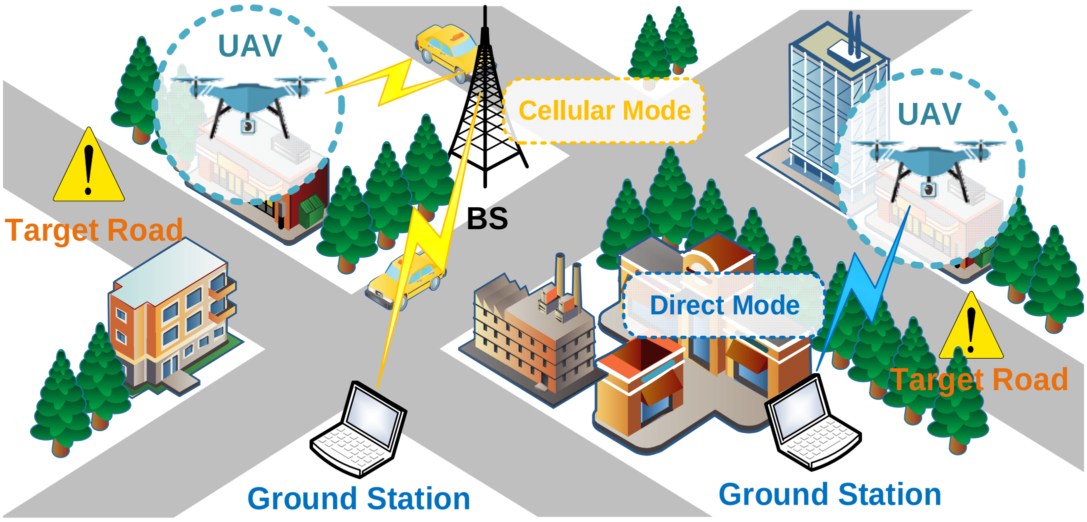

Figure 1. System model.

As illustrated in Fig.1, we consider general city road scenario, which is an OFDMA network with single BS. This network consists of N UAVs performing traffic monitoring tasks. Each UAV is required to record video from its target road and then transmit the video data to the ground stations. The data transmissions can be supported by two modes:

1) Direct Mode: The UAV transmits its video data to its ground station directly;
2) Cellular Mode: The UAV sends the video data to the BS, after which the BS transmits the received data to the ground station.

The time unit for a transmission in both two modes is a frame. To be specific, a UAV in the direct mode takes a whole frame to send the sensory data to its ground station directly, while the two phases of transmissions in the cellular mode take half frame for each.

The network owns C orthogonal subchannels to support the OFDMA transmissions among UAVs and ground station. The BS will allocate orthogonal subchannels to different UAVs, for avoiding the mutual interference among UAVs.

We adopt the probabilistic sensing model in [5] to evaluate UAV's sensing qualities. Specifically, the successful sensing probability for a UAV is an exponential function of the distance between the UAV and its target road. Besides, we adopt the air-to-ground channel model in [6] to evaluate the air-to-ground and ground-to-ground transmissions. Given a threshold, we can express the successful transmission probabilities over the direct mode and the cellular mode from their signal-to-noise ratio (SNR), as referred in [7].

IOP Conf. Series: Materials Science and Engineering **719** (2020) 012073 doi:10.1088/1757-899X/719/1/012073

In our paper, we adopt a sense-and-send protocol to coordinate multiple UAVs performing traffic monitoring tasks simultaneously. This protocol can be formulated by the nested bi-level Markov chains, in which UAV sensing and transmission are formulated as the state transitions in the outer and the inner Markov chains, respectively. More details about the protocol are introduced in [7].

3. Algorithm Design

We introduce a lattice model in [7] to describe UAVs' trajectories, in which the space is divided into a finite set of discrete spatial points. The locations of the BS, UAVs, ground stations and target roads are specified by 3D cartesian coordinates. Thus, the trajectory of a UAV can be characterized by a series of discrete spatial points, which is denoted by a set $\boldsymbol{T}$. We adopt the concept of utility in [7] to evaluate the performance of UAVs, and UAVs aim to maximize their utility by designing their trajectory. To be specific, we define the reward of UAV in the n-th frame is r_n. Besides, we introduce the discounting valuation to evaluate the timeliness of traffic monitoring, in which the discount factor is ρ. Therefore, the trajectory design problem for a UAV can be formulated as follow

$$\max_{\boldsymbol{T}} \quad \sum_{\boldsymbol{n}} \rho^{n-1} r_n$$

Since the utility of a UAV is determined by all UAVs' trajectories, the trajectory design problem is difficult to solve. Note that this problem is a MDP, we can analyze it using RL [8]. Under the RL framework, each UAV is regarded as an agent. The states and the actions of each agent are defined as the locations and the movements of the corresponding UAV, respectively.

Denote the state of UAVs in a frame as $\boldsymbol{s}$, the a UAV taking action a can receive reward r. At the beginning of each frame, a UAV observes state $\boldsymbol{s}$, and then it takes an action a based on its policy, denoted by π. As action $\boldsymbol{a}$ taken by the UAV, it receives a reward r and observes a new state $\boldsymbol{s}'$. Therefore, the trajectory design problem can be solved by maximizing the total expected discounted rewards of all UAVs by optimizing their policies. Under the RL framework, we propose a Trajectory design algorithm based on DQN to optimize policies for UAVs, which is referred in [7].

4. Simulation Result

We present the simulation results on the Trajectory design problem using MATLAB.

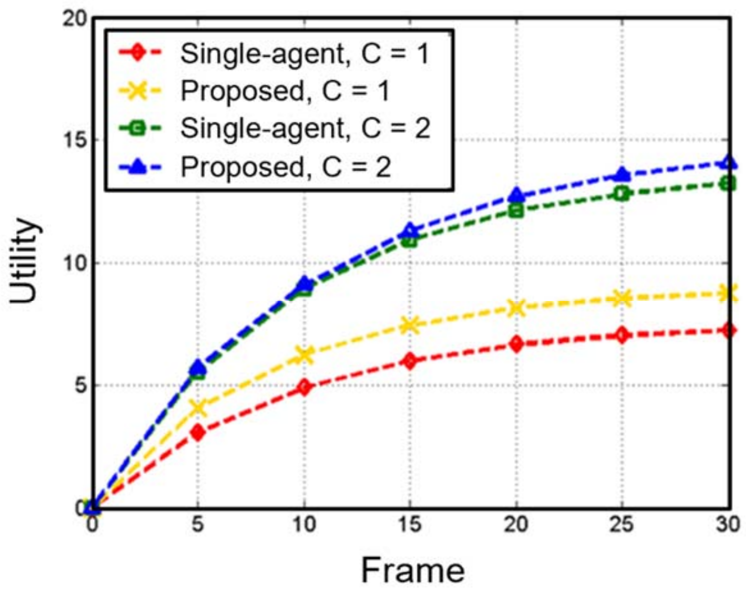

Figure 2. Performance comparison

In Fig. 2, we compare our proposed algorithm with the single-agent Q-learning algorithm [9], in which each UAV updates its policy considering other UAVs as the environment. We can observe that our proposed algorithm achieves a higher utility than the single-agent algorithm. Besides, for either algorithm, the utility increases with the number of subchannels *C*, since more frequency resources are utilized.

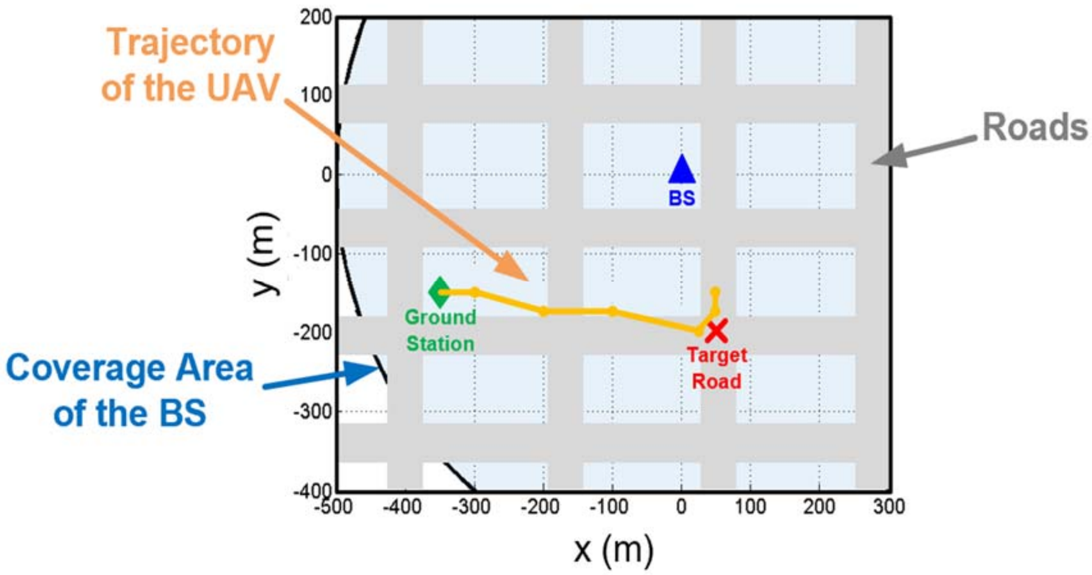

Figure 3. The trajectory of a UAV.

Besides, in Fig. 3, we present the trajectory of a UAV. We can observed that the UAV first flies towards its target road, and then hovers next to its target road rather than on its target road. This is due to the trade-off between UAV sensing and transmission.

5. Conclusion

In this paper, we investigate the trajectory design problem for a cellular-based UAV traffic monitoring system, in which multiple UAVs record videos from the target roads, and then transmit their video data back to the ground stations. We have adopted a sense-and-send protocol to schedule UAV sensing and transmission, and analysed it using nested bi-level Markov chains. As the trajectory design problem is a MDP with large state-action space, we have reformulated this problem using RL, and proposed a DQN-based Trajectory design algorithm to solve it. Simulation results have shown that our proposed algorithm outperforms the single-agent Q-learning method.

6. Acknowledgment

This study was supported by the National Natural Science Foundation of China under Grant No. 61371073.

7. Reference

[1] Y. Zeng, R. Zhang, and T. J. Lim, “Wireless communications with unmanned aerial vehicles: opportunities and challenges,” IEEE Commun. Mag., vol. 54, no. 5, pp. 36-42, May 2016.
[2] H. Zhang, L. Song, Z. Han, and H. V. Poor, “Cooperation techniques for a cellular internet of unmanned aerial vehicles,” IEEE Wireless Commun., to be published.
[3] S. Zhang, H. Zhang, B. Di, and L. Song, “Joint trajectory and power optimization for UAV sensing over cellular networks,” IEEE Commun. Lett., vol. 22, no. 11, pp. 2382-2385, Nov. 2018.

[4] S. Zhang, H. Zhang, B. Di, and L. Song, “Cellular UAV-to-X communications: Design and optimization for multi-UAV networks,” IEEE Trans. Wireless Commun., vol. 18, no. 2, pp. 1346-1359, Feb. 2019.
[5] V. V. Shakhov, and I. Koo, “Experiment design for parameter estimation in probabilistic sensing models,” IEEE Sensors J., vol. 17, no. 24, pp.8431-8437, Oct. 2017.
[6] 3GPP TR 36.777, “Enhanced LTE support for aerial vehicles,” Release 15, Dec. 2017.
[7] F. Wu, H. Zhang, J. Wu, L. Song, ‘‘Trajectory Design for Overlay UAV-to-Device Communications by Deep Reinforcement Learning’’, in Proc. IEEE GLOBECOM, Waikoloa, HI, USA, Dec. 2019.
[8] H. Ye, and G. Y. Li, “Deep reinforcement learning based distributed resource allocation for V2V broadcasting,” in Proc. IEEE IWCMC, Limassol, Cyprus, Jun. 2018
[9] J. Hu, H. Zhang, and L. Song, “Reinforcement learning for decentralized trajectory design in cellular UAV networks with sense-and-send protocol,” IEEE Internet Things J., vol. 6, no. 4, pp. 6177-6189, Aug. 2019.