Location Prediction and Trajectory Optimization in Multi-UAV Application Missions

Rounak Singh, Chengyi Qu, Alicia Esquivel Morel, and Prasad Calyam

Abstract Unmanned aerial vehicles (a.k.a. drones) have a wide range of applications in e.g., aerial surveillance, mapping, imaging, monitoring, maritime operations, parcel delivery, and disaster response management. Their operations require reliable networking environments and location-based services in air-to-air links with co-operative drones, or air-to-ground links in concert with ground control stations. When equipped with high-resolution video cameras or sensors to gain environmental situation awareness through object detection/tracking, precise location predictions of individual or groups of drones at any instant possible is critical for continuous guidance. The location predictions then can be used in trajectory optimization for achieving efficient operations (i.e., through effective resource utilization in terms of energy or network bandwidth consumption) and safe operations (i.e., through avoidance of obstacles or sudden landing) within application missions. In this chapter, we explain a diverse set of techniques involved in drone location prediction, position and velocity estimation and trajectory optimization involving: (i) Kalman Filtering techniques, and (ii) Machine Learning models such as reinforcement learning and deep-reinforcement learning. These techniques facilitate the drones to follow intelligent paths and establish optimal trajectories while carrying out successful application missions under given resource and network constraints. We detail the techniques using three scenarios. The first scenario involves location prediction based intelligent packet transfer between drones in a disaster response scenario using the various Kalman Filtering techniques along with sensor fusion. The second scenario involves a learning-based trajectory optimization that uses various reinforcement learning models for maintaining high video resolution and effective network performance in

R. Singh · C. Qu · A. Esquivel Morel · P. Calyam (✉)
University of Missouri-Columbia, Columbia, MO, USA
e-mail: calyamp@missouri.edu

R. Singh
e-mail: rsft6@mail.missouri.edu

C. Qu
e-mail: cqy78@mail.missouri.edu

A. Esquivel Morel
e-mail: ace6qv@mail.missouri.edu

a civil application scenario such as aerial monitoring of persons/objects. The third scenario involves salient non-ML-based trajectory optimization techniques that can be adopted within UAV-based applications for public safety networks. We conclude with a list of open challenges and future works for intelligent path planning of drones using location prediction and trajectory optimization techniques.

Keywords Drone swarms · Location-based services · Situational awareness · Deep reinforcement learning

1 Introduction

The use of drones has been increasing at a rapid pace for a diverse range of applications in e.g., aerial surveillance, mapping, imaging, monitoring, maritime operations, parcel delivery, and disaster response management. Many applications involve multi-UAV configurations [1], wherein several drones act as either carrier devices to carry supplies [2], or are used for aerial surveillance for intelligent information gathering [3]. They also are deployed as aerial base stations to provide bandwidth and network coverage for ground users in certain applications [4]. An example of air-to-air links with co-operative drones surveying over a designated area is shown in Fig. 1. These operations require location-aided drone movement and optimal drone paths for reduced energy consumption and efficient resource allocation. We discuss salient challenges in realizing these drone location prediction and trajectory optimization techniques and show their advantages through two scenarios involving: (i) network and video analytics orchestration, and (ii) intelligent packet transfer in a disaster response management scenario. This chapter will illustrate how various predicted location information and intelligent path planning schemes help in achieving efficient performance of application missions.

1.1 How Can Drone's Location Prediction Be Useful in Networking Environments and Application Scenarios?

To explain the significance of drone location prediction in real-time applications, we consider a multi-drone co-ordination and networking system for a critical application mission such as e.g., a disaster response scenario (DRS) [5, 6]. This scenario involves critical tasks such as monitoring the disaster affected area, search and rescue operations, and providing supplies to victims. This system features a Flying Ad-Hoc Network Topology (FANET) [7] to support air-to-air, as well as air-to-ground links. The ground control station (GCS) sends requests to the drones to execute certain tasks and the drones send back situational awareness information to the GCS. Such a scenario, however involves challenges related to drone positioning and path plan-
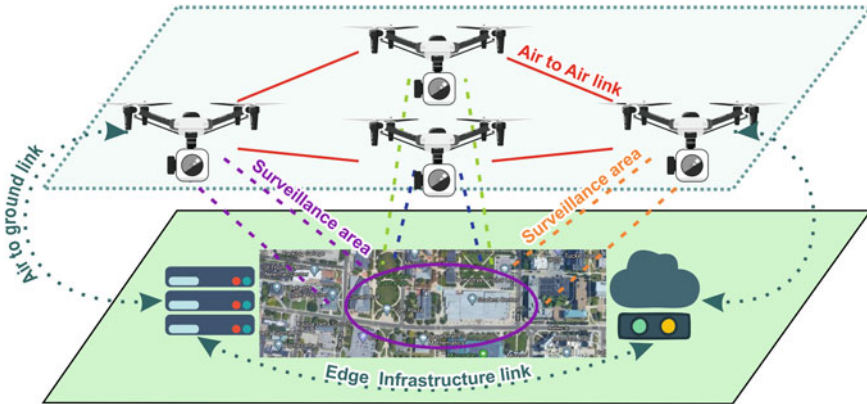
Fig. 1 Overview of multi-drone setup based on air-to-air and air-to-ground links

ning. Particularly, the *location estimation of drones* is necessary for multi-drone co-operation in order to stay on-course and avoid mid-air collisions. Furthermore, *trajectory planning and optimization* is required to efficiently carry out the application mission considering the limitations of energy and resources. To explicitly understand how these two essential methods impact the performance of drones in application missions, we elaborate them in the following:

1. **Location Estimation and Prediction**: Tracking and predicting the locations of drones is important in order to get real-time estimates of drone positions for autonomous control and to improve the accuracy of delivery tasks execution in a specific application scenario. It measures how closely the drones are being monitored and also measures the reliability of the path computation algorithm performance. This can be achieved by using motion models of the drone movements, and by using such models within a tracking algorithm or a recursive filter. To get the near-optimal estimates with the motion model, prior works use the Kalman Filter [8] technique which is widely-used for estimation purposes. The popularity of Kalman Filter is due to the fact that this technique takes in the current values as input data (i.e., measurement) along with noises (i.e., measurement noise and process noise) to produce unbiased estimates of system states [9]. Leveraging this state estimation technique can help achieve predicted positions of drones.

2. **Trajectory Optimization** The path that a drone follows during its operation is crucial for effective communication, computation offloading [10], energy consumption and information transfer. A drone's trajectory design unquestionably plays an important role in the application performance enhancement and effectiveness. During its operation, the drone flies over areas which are prone to network and communication vulnerabilities such as signal-loss, cyber-attacks, coverage and range limitations that could severely impact the drones' performance and put the application mission at risk. Machine learning techniques such as model-free

reinforcement learning [11] and deep reinforcement learning [12] provide effective reliable solutions for tackling these implications. They use trial-and-error path learning techniques for a drone to establish optimal and intelligent trajectory during its overall flight time during an application mission.

1.2   Chapter Organization

This book chapter aims to address the concepts of drone position and trajectory optimization techniques related to intelligent path planning. The chapter will first discuss the challenges related to drone location prediction and trajectory optimization. Next, methods for location prediction will be discussed that involve various Kalman filtering techniques and methods of trajectory optimization using reinforcement and deep reinforcement learning techniques. In this context, we also discuss non-ML-based methods for trajectory optimization. They together provide motivation for localization and intelligent path planning of drones for a given application scenario. Furthermore, we discuss how trajectory optimization of UAVs can aid the operations of public safety networks. These techniques are based on the theoretical and experimental research conducted by the authors in the Virtualization, Multimedia and Networking (VIMAN) Lab at University of Missouri Columbia. Lastly, we discuss the main findings of this chapter and list out the open challenges and future works that can be implemented using our approaches to carry out drone-based application missions effectively and efficiently.

2   Challenges in Drone Location Prediction and Trajectory Optimization

Since drones are classified under unmanned aerial vehicles, it can be presumed that the navigation, operation and controlling is carried out externally by a ground control station or a ground (human) pilot. In most of the applications today, however the drones flight is increasingly becoming autonomous and may require minimal or almost no external (human) guidance. This is possible due to the variety of sensors on-board that constitute the inertial measurement unit (IMU), global positioning system (GPS), inertial navigation system (INS), gyroscope, accelerometer, barometer and high resolution cameras. These sensors facilitate the autonomous drone flights with high accuracy. Nevertheless, these sensors are prone to external noises that can cause inaccuracies malfunctioning. Anther critical elements on which a drone's fight is dependent is the battery that powers the drones flying mechanism, its flight controller and the above-mentioned sensors. Some of the major challenges pertaining to localization and path-planning relating to the above issues are:

Collision avoidance: Real-world application missions are carried out in complex environments and sometimes, civil applications involving drones are conducted in urban areas. The UAVs are only dependent on their on-board sensor capabilities for their traversal through these environments. It is not always feasible to rely on these sensor readings for navigation and the drones may run into obstacles, hit trees, buildings or other drones mid-operation. Many techniques have been proposed for collision avoidance using decentralized control [13, 14]. The drone has to be aware of the location of its neighbor (drone) and itself at any given instant of time. Leveraging this information can help tackle the problem of mid-air collisions. Object detection using computer vision can help in identifying certain objects by training on datasets of images of common environment obstacles [15]. However, drone's system reliance and communication within the network is usually difficult and challenging in large-scale application missions involving complex environments.

System Security: A wide range of drone-based applications are carried out by the military that operate on highly confidential information gathering within classi-fied missions. Also, many civil applications involve sensitive data collection when drones are deployed as aerial base stations or network providers that handle ground user data (e.g., faces and postures of individuals in crowds). Drones are at risk of cyber-attacks and can be hacked, without the drone being physically captured. The information gathered can become vulnerable and exposed to hackers. Mostly, the camera modules are targeted and video captured is received by hackers which may expose the operations that are carried out in the surveillance area. The work in [16] uses Blockchain technology that encrypts the data being transmitted to base stations. An approach for threat analysis of drone based systems is described in [17]. Coun-termeasures to security issues in professional drone based networks are shown in [18].

Energy Limitations: Drones require energy for total flight time including hov-ering over an area for surveillance and data transmission. Additionally, the on-board sensors constantly consume energy to function properly and provide localization of the drones. Energy consumption can also be increased due to attached payloads [19], wind resistance [20] and network issues [21]. The total energy on a drone is limited thus restricting the flight-time of the application mission. The work in [22] provides an energy-aware approach that uses trajectory planning of drones used as mobile anchors to save energy.

Location Awareness and Blockage of Line-of-Sight: In the context of location estimation of drones, blockage of line-of-sight for drones is a very trivial problem that surfaces in the rarest of times [23]. As drones tend to fly long distances based on their application missions, the location awareness becomes essential in order for them to remain in their trajectory and under a predefined network connection for information transfer. It is necessary that they avoid collisions and interference. It becomes a problem if a drone's flight is affected due to external factors and it might become susceptible to unknown attacks. In the worst case scenario, the drone can be thrown off-path and after consuming all its power, it can land or fall in an unknown

territory. Thus, it can render itself and the information collected vulnerable, and any expensive sensors or video camera components are subject to expensive damage or loss. Various types of research is being conducted by many groups to realize location awareness [24] of drones.

3 Methods for Drone Location Estimation and Prediction

In our DRS application the drone environment is considered to be a 2D dynamic and non-linear horizontal plane. As discussed in Sect. 1.1, we assume that all the drones are connected forming a FANET. They communicate the mapping and monitoring information over the same network to the delivery drones in order to carry out a delivery task. Consequently, the network topology of the multi-drone system keeps on changing based on the mobility of the drones. The position estimation of the drones must be performed in very short intervals of time using the new coordinates being updated rapidly within the FANET. Each drone in the FANET is considered to have a GPS module and an IMU to record its current location. This information is broadcast to the FANET so that the other drones in the vicinity are recognized for packet or information transfers when needed. We get the initial measurement data of the drone using GPS and other on-board sensors such as gyroscope, barometer, accelerometer and magnetometer that are all part of the IMU. The drone's rotational movement angles observed and controlled by a gyroscope and rotary movements, for stability are shown in Fig. 2. The accelerations and rotations of the drone can be observed over time to give an estimated position by learning the next measurement values for different time-steps.

The position, velocity, acceleration and heading of a UAV are considered as dynamic states at a given time-step. In order to get the location prediction of an UAV, a state estimator is required to get the true values along with a prediction of these states for the next time-step. Kalman filter can be used to observe state estimates over time along with process noise and measurement noise from sensors to give estimates on which drone position state estimates are closer to true values that cannot be calculated directly [25]. Since the inception of Kalman filter in 1960, it has evolved over time, and the most popular Kalman filters for UAV location esti-

Fig. 2 Motion angles of a drone responsible for movement with six degrees of freedom controlled by the gyroscope and flight controller
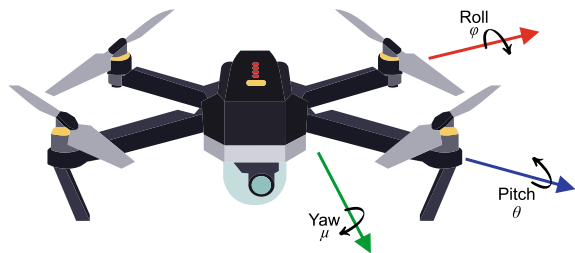
mation are the original Kalman filter, the extended Kalman filter (EKF) [26] and the unscented Kalman filter (UKF) [27].

3.1 State Estimation of Drone Parameters Using Original Kalman Filter

The functionality of Kalman filter relies on consecutive iterations of prediction and filtering i.e., it follows a sequence of prediction and update equations. Along with the inertial navigation system (INS) data, a predefined motion model of the drones' movement is given as input to the Kalman filter. The motion model is basically a state transition matrix having time-periods of the states i.e., x and y coordinate, acceleration and angular velocity. The prediction equations give priori estimates and the update equations give posterior estimates. The update equations take up the previous state's mean and noise covariance and produce the updated mean and noise covariance values for the next state. The filter then combines the predicted states and noisy measurements to produce unbiased estimates of drone system states. In this process, data with process noise and measurement noise from sensors is used as input, and the Kalman filter produces a statistically optimal estimate of the underlying state by recursively acting on the series of observed inputs.

For simplicity, the Kalman Filter can be used to get position and velocity estimates of UAVs but only in a 2D plane, assuming it us flying at a fixed altitude. Other applications of Kalman filter include guidance and navigation systems, tracking of maneuvering targets, dynamic positioning, sensor data fusion and signal processing. An approach for path planning of UAV using a Kalman filter is given in [28].

3.2 Extended Kalman Filter for Non-linear Drone State Estimation

The major limitation of a Kalman filter is that it can only process estimates of linear systems, and it suffers from linearization when operated on nonlinear models. Drone flight operation is generally non-linear and time varying and system parameters with a dynamic motion model cannot be measured directly with on-board sensors because they may be subject to noise and malfunctioning. To overcome this non-linearity issue of drone position estimation, one of the widely used filter for non-linear state estimation, i.e. the extended Kalman filter (EKF) is used. It uses Taylor series expansion and linearizes and approximates the state estimates of a non-linear function around the conditional mean. EKF can be reliable while estimating the drone positions using the drones' dynamic state parameters.

The dynamic motion model is solved by learning the non-linear transition of measurement noise covariance and process noise covariance along with the change

in states to give an optimal estimate of the UAV position. The EKF also follows a series of prediction and update equations. The priori estimates calculated during the prediction process are updated to give the posterior estimates and their covariance. Additionally, Jacobians of dynamic functions are used with respect to system state of the UAV to map its states to observations. Additionally, by recursive operations, the covariance of the estimated error is minimized. Hence, the EKF can be used to get the more accurate positions of the drones through prediction of future positions with insignificant errors, when compared to the original Kalman filter.

The work in [29] shows the non-linear estimation of drone's state along with sensor data for localization and [30] shows an approach for determining the locations of drones using inter-drone distances in 2D co-ordinates.

3.3 Unscented Kalman Filter for Improved Position Estimation and Orientation Tracking of UAVs

The EKF is computationally complex and takes longer to produce estimates, also its accuracy is reliable in real-time but can still be improved. The unscented Kalman filter (UKF) is used for the same applications requiring higher accuracy. It is a deterministic sampling approach involving sampling of distributions using a Gaussian random variable. It employs the unscented transform method to select a set of samples called sigma points around the mean to calculate the mean and covariance of the estimation that eradicates the requirement of using Jacobians, as in the EKF. This preserves the linear update structure of the original Kalman of estimates filter unlike the EKF. Table 1 shows the comparison of various Kalman filtering schemes used for location estimation of UAVs; for a detailed comparison, readers can refer to [31].

In drone localization application, the system dynamics is expanded as the drone's cartesian location i.e., position, velocity and acceleration. These provide a non-liner relationship between the system states and measurements, and thereby the implementation becomes simpler. The orientation tracking of a drone is also carried out using the UKF [32] by considering rigid body dynamics using various types of measurements like acceleration, angular velocity and magnetic field strength. It uses quaternions and UKF, thus proving its computational effectiveness of tracking. Another approach for position estimation using UKF samples images uses a visual target. It uses weights (difference of observed value and estimated value of vision sensor) for observations to prevent divergence in estimated values by UKF [33].

3.4 Sensor Fusion for UAV Localization

Multi-sensor fusion is another technique that shows the importance of using data from distinct sensors to predict the dynamic state estimates of drones for aerial

Table 1 Comparison of Kalman filtering scheme variants for location estimation and prediction of UAVs

Type of filter	Type of system	Accuracy	Model design
Kalman Filter	Linear	Least accurate	Least complex
Linearized Kalman Filter	Non-linear	Moderately accurate	Moderately complex
Extended Kalman Filter (EKF)	Non-linear	Accurate	Most complex
Unscented Kalman Filter (UKF)	Non-linear	Most accurate	Complex

applications. The work in [34] shows how data collected from the GPS, IMU, and INS are fused together for UAV localization using state-dependent Riccati-equation non-linear filter along with a UKF. Drone path planning involves navigating the drone to a desired destination travelling over a predefined path that constitutes obstacles and other environment constraints. The work in [35] shows how the sensor fusion along with real-time kinematic GPS sensors is used to accurately calculate the altitude and position of the drone. They generate a data-set using instantaneous positions of the drone in different directions along with the roll, pitch and yaw angles. Further, they compare this data with the output of the sensor fusion model estimations that are carried out using an EKF to produce position and altitude estimates of drones.

3.5 Location Prediction Based Intelligent Packet Transfer

The location prediction algorithm embedded with the above drone position models along with the position and velocity estimation by Kalman filter and location prediction by EKF, can be run online to make advance decisions by using future location information of the mapping drones, monitoring drones and the delivery drones in the FANET. UKF along with sensor fusion methods can alleviate potential inconsistencies in the dynamic state estimation and can help the algorithm produce accurate results. Thus, the FANET in the DRS scenario can utilize theses location estimation techniques to facilitate efficient packet transfer.

Table 2 summarizes how different methods of location prediction of drones have been proposed in prior works to achieve goals in different application missions. The details of the salient methods used to perform drone location prediction while operating in an application are described in the following:

Table 2 Methods and applications of location estimation of drones

Case study	Method	Solution	Application	Goal
Xiong et al. [25]	Kalman Filter	Linear estimation	State estimation	Autonomous flight
Mao et al. [30]	Extended Kalman Filter	Non-linear estimation	Localization of UAVs	Localization without GPS
Kraft et al. [32]	Unscented Kalman Filter	Linearized estimation	Localization of UAVs	Orientation computation
Abdelfatah et al. [35]	Sensor Fusion	Non-linear estimation	Localization of UAVs	Altitude, position estimation

4 Methods for Drone Trajectory Optimization Using Machine Learning

In context of drone trajectory optimization, we consider an area that is prone to signal-losses, cyber-attacks and potential obstacles like trees, buildings, tall-standing structures which affect the drones' performance and cause hindrance in the application mission. An overview of a drone's trajectory during an application is shown in Fig. 3. To overcome these problems there is a need for intelligent path planning that can enable the drones follow an optimal trajectory, flying in areas free of all the impediments and attacks.

The details of the salient methods used to optimize the drones' trajectories while operating in an application are described in the following:
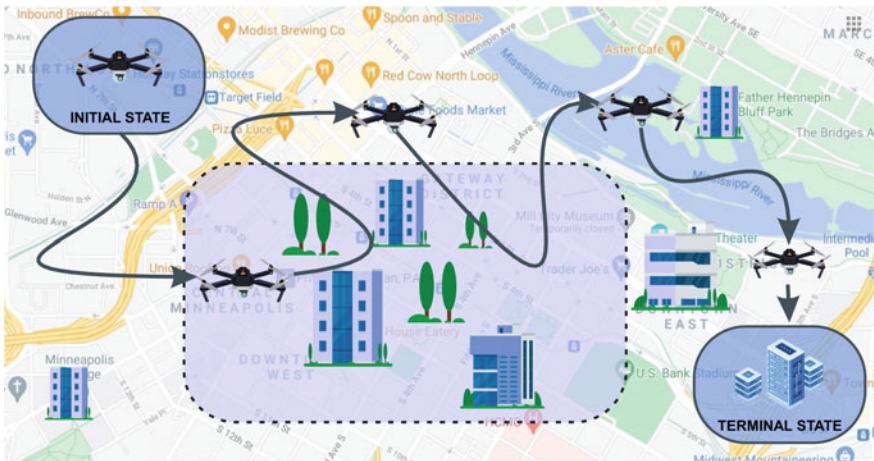

Fig. 3 Overview of drone's trajectory in a learning based environment comprising of potential obstacles

Reinforcement Learning: Path planning of drones is a crucial aspect of research in drone-based applications because the efficiency of missions is dependent on the traversal of the drones in a given area. It correlates with autonomy and has a profound impact on guidance, operation and endurance of the drones. Most drone based application missions are defined in unknown environments. Therefore, Markov Decision Process (MDP) is employed to solve such environments and the Q-Learning algorithm is used that follows the Markov property [36]. It is a model-free reinforcement learning algorithm that puts emphasis on an agent to learn actions under given circumstances to handle problems with stochastic transitions. For any finite MDP, the Q-learning algorithm finds an optimal policy by maximizing the expected value of cumulative rewards over successive actions taken in given states, starting from a current state. There has been a wide usage of reinforcement learning algorithms in varied areas of drone-based application research where drones are allowed to directly and continuously interact with the environment.

Deep Reinforcement learning (DRL): This concept can be considered as a combination of deep learning and reinforcement learning. It employs a deep neural network (DNN) to estimate the Q function $Q(s, a)$ for a given set of state-action pairs. Often reinforcement learning requires the state space and the action space to be fixed and discrete, and the agent learns to make decisions by using a trial and error method. It basically involves employing a Q learning algorithm that maintains a record of values of what actions have been taken in given state spaces and also the rewards associated with the corresponding states and actions in a limited format where the state space is predefined. The DRL method allows the agent to act in an environment that has a continuous and mostly undefined state space. It also uses a set of discrete or continuous actions which are given as a stack of inputs in contrast to the single inputs in case of a simple reinforcement learning. In other words, the DRL makes sure that the agent performs well with extensive input data coming from a large state space to optimize the given objective of any application e.g., it uses pixels as input data in Atari games [37]. The DNN approximates the Q function which estimates the cumulative reward for each state-action pair. A DNN may often suffer with divergence, so it uses a set of experience replay memory and target network to overcome this issue. DQN based RL solutions for drones are necessary because a drone's operation in a given environment is considered as a continuous state space and multi drone scenarios require more robust algorithms such as the multi-agent DQN [38] and the actor-critic [39] networks, which also employ DNNs to generate an optimal policy solution.

4.1 Q-Learning

Q-learning is a type of model-free reinforcement learning as described in [40], which is used to solve MDP based problems with dynamic programming. The Q-learning algorithm creates a table (i.e., Q-table) containing the corresponding values of each

Fig. 4 Overview of
Reinforcement learning
showing the agent's
interaction with the
environment corresponding
to given states and actions to
generate a policy
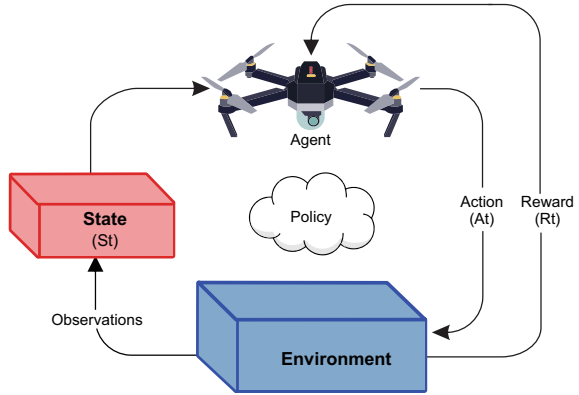
state-action pair and keeps updating them along with the reward values. The scores
obtained in the Q-table are represented as the values of the Q-function $Q(s_t, a_t)$, and
are given by -

$$Q(s_t, a_t) = E\left[\sum_k \gamma^k R_{t+k+1} | (s_t, a_t)\right] \tag{1}$$

where t is the time step and k is the episode. The Q-function is updated for each
episode when the agent performs certain actions in a given state to maximize its
cumulative reward using the Bellman's equation [41], which is given as -

$$Q(s_{t+1}, a_{t+1}) \leftarrow (1 - \alpha)Q(s_t, a_t)$$
$$+ \alpha[R_t + \gamma.max_a Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]; \tag{2}$$

The algorithm converges when maximum reward is reached. The policy encour-
ages the agent to choose optimal actions and receive greater scores in an iterative
fashion, which results in the model rendering high Q-values. The interaction of the
agent with the environment to generate rewards and to establish a policy is shown in
Fig. 4. The output of the Q-learning is the drone trajectory update guidance that is
used to keep the drones as much as possible in the optimal trajectory.

Ensuing the design of the drone's optimal trajectory selection scenario using an
MDP, we can evaluate the overall performance by tuning the values of the discount
factor γ for obtaining the optimal policy $\pi_t^* : S_t \rightarrow A_t$, which maps the state space
with best suitable actions.

4.2   Deep Q-Network

To implement the Q-Learning based algorithm that render optimal trajectories of the drones, we choose a DQN that allows for maximum exploration and exploitation [42] of the learning environment by the agent. The actions in this case are dependent on the weights of the primary DNN, which adds flexibility in the overall learning process i.e., as the weights update, the rewards update accordingly. The intelligent trajectory learning application for DRS scenario renders network performance in terms of throughput and the video quality scores (i.e., rewards) obtained in the process of learning. The DQN is trained using a experience replay, which is memory buffer that stores the sequence of state-action pairs from previous episodes. The process of utilizing replay memory to gain experience by random sampling is called experience replay.

The DQN utilizes the mini-batch from experience replay with the observed state transition samples to update its DNNs after each episode during the training process. Thereby, it breaks any correlation made using sequential state-action pairs in the previous episodes.Sometimes, drones are used as swarms in application missions that are connected via wireless links. For any broken link, the drones have to position themselves to make up the broken link to maintain the same QoS requirements. The work in [43] gives an approach that uses DQN to determine optimal links between drones in swarms and to localize the drones to improve overall network performance of the swarm's wireless network.

4.3   Double Deep Q Network

The Deep Q Network has a single action value function and while updating the primary DNN, same values are used for selection and evaluation of actions. This in turn leads to overestimation that renders over optimistic action value estimates. To avoid this issue, Double Deep Q Learning decouples the selection and evaluation of value function using two separate DNNs (primary and target). It employs two value functions that learn by selecting random experiences that produce two set of weights [44]. It aims to get the most out of Double Q learning with slight increase in computation. For civil and military based application missions, Double Deep Q Network (DDQN) is used for 3 Dimensional path planning of drones using greedy exploitation strategy to improve learning in complex environments [45].

4.4   Dueling Deep Q Network

The Dueling Deep Q Network (Dueling DQN) is another form of a deep reinforcement learning algorithm. It consists of two separate estimators (DNNs) for state value

function and action value function. It is used to overcome the impact caused by similar action values in multiple episodes [46]. Some application missions involve multi-drone connections using cellular networks with each drone acting as a base station. To improve the connectivity over the cellular network, Dueling DQN is used to provide trajectory optimization and coverage-aware navigation for radio mapping [47]. Also in other dynamic environments with unrealized threats, Dueling DQN can provide intelligent path-planning using epsilon greedy policy to render optimal trajectories of the drones [48].

4.5 Actor Critic Networks

Some of the most recent and popular reinforcement learning algorithms are the actor critic networks that aim to achieve optimal policies using low-gradient estimates. The actor network is a DNN that takes in the current environment state and computes continuous actions and the critic judges the performance of the actor network with respect to the input states. It also provides feedback to determine the best possible actions that render higher rewards [49, 50]. An approach to achieve efficient communication and band allocation in the drone network involves determining their 3D trajectory under energy constraints using deep deterministic policy gradient (DDPG) [51] actor-critic networks as shown in [52].

4.6 Orchestration Motivation for Online Learning

The performance in the network links across multi-drone FANETs vary due to certain factors such as, application requirements, weather conditions, obstacles in the path, etc. that cause frequent or intermittent outages in transmission and reception of crucial information inside the FANET. This could also affect the drone's video analytics, when used for civil applications for aerial surveillance. Our proposed orchestration process solves the network links and video analytics disruption by employing an online learning based technique. It analyzes the trajectory during the drone flight, and find ways to optimize the drone's path and even the video quality by selection of pertinent network protocol and video properties during the drone flight.

The Q-learning algorithm forms the basis of the trajectory learning of the drones in different areas and can be applied across all the drones in the FANET. An approach for path planning and obstacle avoidance is shown here [53]. However, Q-learning cannot be used for complex learning environments as it would not allow exploration and exploitation [54] of the total area that the drones are covering during their flights.

To achieve intelligent trajectory learning, we propose a Deep Q-Network based method. The path selection aids the drones to learn and make necessary sequence of decisions under uncertainty in FANET conditions. The learning involved in path selection by the drones can be represented as a Markov Decision Process (MDP) [39]

which forms the basis for the DQN algorithm and is defined as a tuple containing the following-

$$M_=(s_t, a_t, p_t, r_t, s_t^{'}) \tag{3}$$

where s_t is the state space, a_t is the action space, r_t is the reward, p_t is the probability of transition of states and $s_t^{'}$ is the next state. The MDP aims to maximize the cumulative rewards that are received by the drones along their trajectories during the operation over a surveillance area. The drones are assumed to be fully charged before they enter the initial state. The learning environment comprises of all the states and actions.
(1) States: For any MDP, the states used are the current state s_t and the next state $s_t^{'}$.
(2) Actions: These are the actions that the drone chooses to perform during its flight operation.
(3) Reward: It is a feedback parameter, received either as an award or penalty which is a consequence of taking certain actions in the learning environment state-space.
(4) Probability of State Transition: It is defined as the probability distribution of the next state $s_t^{'}$ given the current state s_t and current action a_t.

The video and network analytics of drones can be formulated as states s_t, actions a_t along with corresponding reward functions in a civil application based on requirements. A DQN with pre-defined weights can take state space values (s_t) of drones as input, forward pass the values and generate optimal action value function $Q(s_t, a_t)$, and compare it with optimal action value function $Q_\pi^*(s_t, a_t)$. Through back-propagation, it can perform updates to the weights of the neurons so that in the later iterations the output values come close to the optimal value. The DQN algorithm converges when an optimal value is reached. The DQN model can be further extended to Double DQN, Dueling DQN and Actor-Critic network using the same learning environment based on the requirements for network and video orchestration.

An approach that uses deep reinforcement learning for optimizing UAV trajectories is detailed in [55] and uses flow-Level modeling for UAV base station deployments. A similar approach in [56] uses a deterministic policy gradient (DPG) on a model-free reinforcement learning scenario to obtain intelligent UAV trajectories. Deep reinforcement learning can also be applied to more complex scenarios involving tedious tasks such as real-time resource-allocation in multi-UAV scenarios [57]. We consider a scenario that aims to achieve optimal solution for 'energy harvest time scheduling' in a UAV assisted device-to-device (D2D) communications setup by conceiving a system model that can reflect dynamic positions of UAVs along with unknown channel state information. The system model also uses the deep deterministic policy gradient (DDPG) algorithm to solve the energy efficient optimization game for the D2D communications scenario.

5   Non-ML-Based Trajectory Optimization Techniques for Drones

Although machine learning is gaining traction in solutions for autonomous vehicles, trajectory optimization of UAVs in real-time scenarios is challenging because it is a non-convex optimization problem. There have been advances in drone trajectory planning and optimization techniques for single-UAV, dual-UAV and multi-UAV based applications. A survey for long-distance trajectory optimization of small UAVs is given in [58], and a survey of techniques involving joint trajectory optimization with resource allocation is given in [59]. An approach to perform joint trajectory and communication co-design can be found in [60]. Advances in path-planning techniques feature techniques that are quite different from learning-based methods. To provide high-mobility and flexibility in FANETs, many techniques have been proposed. However, there are several open challenges when it comes to path planning of UAVs. A series of latest works that try to solve the open challenges are as follows.

5.1   Trajectory Optimization Using Quantization Theory-Lagrangian Approach

An approach to provide optimal UAV positions in static networks under spatial user density is described in [61]. This approach uses uniform distribution of ground terminals at zero altitude and determines optimal placement of UAVs in static environments along with ways to reduce power consumption. The optimizations for the static case are done by considering the UAVs at varying altitudes, followed by characterizing optimal UAV deployments in dynamic scenarios. These optimizations are performed by varying ground terminal density in any given dimension for a fixed number of UAVs which are placed at moderate distances from each other. Two two cases are considered: (i) UAVs with no movement, and (ii) UAVs with unlimited movement. This approach aims to achieve lowest possible average power consumption followed by providing a Lagrangian-based descent trajectory optimization technique. The Lagrangian technique is similar to Voronoi based coverage control algorithms and is based on time discretization.

5.2   Joint Optimization of UAV 3D Placement and Path Loss Factor

An approach in [62] aims to fill the gaps of joint aerial base station (ABS) deployments and path loss compensation for ABS placements at certain heights. It puts stress on the power control mechanism needed to establish reliable communication, and on the propagation path-loss that hinders the overall communication performance. The

3D UAV placement procedure involves altitude optimization for maximum coverage along with horizontal position optimization for 2D placement that uses a modified K-means algorithm for aerial base station height with a compensation factor.

5.3 Flexible Path Discretization and Path Compression

This technique considers a piecewise-linear continuous trajectory of a UAV whose path comprises of consecutive line segments connected through a finite number of points in 3D called way-points. It provides a solution to render an optimal path by using a flexible path discretization technique to optimize number of way-points in the path to reduce the complexity in the design of the UAV trajectory [63]. The variables that tend to solve the path-planning are considered in two sets of design-able and non-design-able way-points. The way-points are generated using their sub-path representations that ensure a desired trajectory discretization accuracy. They also help to obtain utility and constraint functions that retain accuracy in e.g., aerial data harvesting using distributed sensors. Following this, a path compression technique is performed that takes the 3D UAV trajectory and decomposes it into a 1D (sub-path) signal to further reduce the path-design complexity.

5.4 Connectivity Constrained Trajectory Optimization

This technique provides a solution to optimize an UAV's trajectory in an energy and connectivity constrained application to reduce the overall mission completion time. It uses graph theory and convex optimizations to achieve high-quality solutions in various scenarios involving: (i) altitude mask constraints, (ii) coordinated multi point (CoMP)-based cellular enabled UAV communications, (iii) QoS requirements based communication using UAVs, and (iv) non-LoS channel model. The degree of freedom of UAV movement is exploited to increase the design flexibility of UAV trajectories with respect to the locations of GCS, and ground users for effective communication. By applying structural properties, effective bounding and approximation techniques, the non-convex trajectory problem is converted into a simple shortest path problem between two vertices and solved using two graph theory based algorithms [64]. A similar technique involving effective trajectory planning under connectivity constraints using graph theory is shown in [65].

5.5 3D Optimal Surveillance and Trajectory Planning

Public safety is another crucial application domain for designing drone based communication systems. Prior works such as [66] have proposed approaches to solve

Table 3 Methods and applications of trajectory optimization of drones

Case study	Objective	Solution	Method	Performance
Koushik et al. [43]	Node Positioning	MHQ-PRP Queuing	Deep Q-Network	Increased throughput of dynamic UAV swarming network
Zhao et al. [45]	3D Path Planning	Greedy Exploration (DRL)	Double Deep Q-Network	Better convergence compared to DQN and DDQN
Yan et al. [48]	Real-time path planning	STAGE scenario	Dueling DDQN	Efficient dynamic environment path planning
Ding et al. [52]	3D Trajectory Planning	DDPG (DRL)	Actor-Critic Network	Increased throughput under fairness conditions
Saxena et al. [55]	Traffic-aware UAV trajectories	Leveraging UAV Base Station Network (UAVBSN)	Deep Reinforcement Learning	Three fold increase in throughput of UAVBSN
Nguyen et al. [57]	UAV Trajectory Optimization	Energy Harvesting Time Scheduling	Deep Deterministic Policy Gradient	Efficient resource allocation under energy and flight-time constraints
Xu et al. [60]	2D Trajectory Planning	Semi-definite Programming	Monotonic Optimization (various)	Significant power saving
Koyuncu et al. [61]	Multi-UAV Trajectory Optimization	Lagrangian Approach (1D)	Quantization Theory	Minimized power consumption
Shakoor et al. [62]	3D Placement and Path-Loss	Placement Compensation Factor	Various Optimization techniques	Improved coverage and performance
Guo et al. [63]	3D Trajectory Design	Flexible Path Discretization and Path Compression	Graph Theory (Shortest path)	Reduced path design complexity
Zhang et al. [64]	3D UAV Trajectory Design	Graph Theory	Optimization - (various)	Improved connectivity
Yang et al. [65]	3D UAV Trajectory Design	Graph Theory, Inequality property	Optimization - (various)	Improved connectivity
Teng et al. [66]	3D Optimal Trajectory Planning	Particle Swarm Optimization	Trajectory Planner	Improved dynamic environment adaptability

challenges for the public safety application domain. Specifically, a swarm optimization based trajectory planner is provided with surveillance-area-importance updating apparatus. The apparatus aims to derive 3D surveillance trajectories of several monitoring drones along with a multi-objective fitness function. The fitness function is used as a metric for various factors of the trajectories generated by the planner such as energy consumption, area priority and flight risk. This approach renders collision-free UAV trajectories with high fitness values and exhibits dynamic environment adaptability and preferential important area selection for multiple drones. Table 3 summarizes how different methods of trajectory estimation and optimization have been proposed to achieve certain goals in various applications.
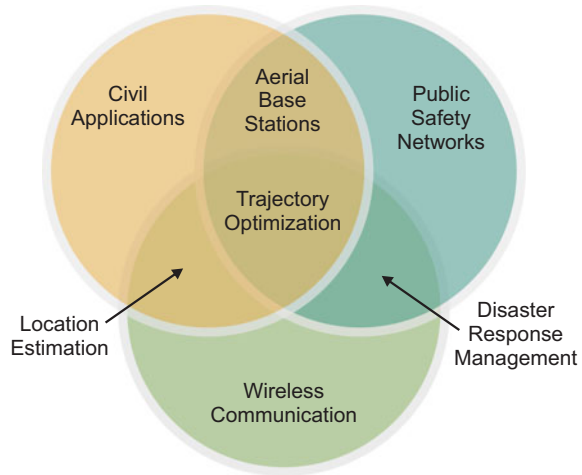
6 How Can Trajectory Optimization Aid UAV-Assisted Public Safety Networks?

Public safety networks (PSNs) are established for public welfare and safety. They are essential means of communication for first responders, security agencies and healthcare facilities. Nowadays, PSNs have been widely relying on wireless technologies such as long range WiFi networks, mobile communication and broadband services that use satellite-aided communication links. In addition, PSNs operate extensively during natural disasters, during times when there is a threat to national security such as terrorist attacks, and any large-scale hazards caused due to human activities. As wireless communication is the backbone of PSNs, advanced and efficient communication technologies such as LTE and 5G-based communications can help establish broadband services that provide improved situational awareness with security and reliability characteristics in the network. In this section, we will discuss how UAVs could be a choice for public safety networks in terms of various use cases, provide case studies on trajectory optimization and localization for UAV-assisted PSNs, and discuss open challenges in UAV-assisted PSNs. Figure 5 provides an overview of multi-UAV operations spanning diverse applications ranging from civil applications to public safety networks.

6.1 UAV-Assisted Public Safety Networks

Since wireless communications play a fundamental role in PSN operations, their effectiveness and responsiveness to emergency situations becomes critical [67]. A few issues that affect the functioning of PSNs include: communication equipment deployment costs, spectrum availability, network coverage and quality of service (QoS). A few of these issues can be solved by improving the ground-based communication systems by fully exploiting the potential of situational awareness and enabling advanced tracking, navigation and localization services [68]. However, to

eradicate these issues of PSNs as a whole, UAVs with enhanced functionalities that can operate as aerial base stations with high-end communication equipment can be used to amplify the effectiveness of communication, improve coverage, reliability, and energy efficiency of wireless networks. In such UAV-assisted PSNS, UAVs are operated by acting as flying mobile terminals within a cellular network or broadband service while monitoring the area, simultaneously. The other advantage on UAV-assisted PSNs is that the UAV base stations are faster and easier to deploy, which provides effectively on cost and can be flexibly reconfigured based on mobility.

6.2 Trajectory Optimization and Path-Planning for UAV-Assisted PSNs

Trajectory Optimization and localization of UAVs can significantly impact the 3D-deployment of the aerial base-stations serving non-stationary users. Optimal path planning can help strengthen the carrier channel transmitting and receiving characteristics. The cellular networks involving aerial base stations can be converted to FANETs, which can help to establish efficient wireless communication in the PSNs. A case study in [69] used path-planning for UAVs in a disaster resilient network. They showed how drones can be used in an wireless infrastructure, allowing a large number of users to establish line-of-sight links for communication. Another approach in [70] uses fast K-means based user cluster model for joint optimization of UAV deployment and resource allocation along with joint optimal power and time transfer allocation for restoring network connectivity during a disaster response scenario. Similarly, research in [71] discusses the role of UAVs in PSNs in terms of energy efficiency and provides a multi-layered architecture that involves UAVs to establish efficient communication by considering the energy consumption considerations.

6.3   Open Challenges in UAV-Assisted PSNs

As we can observe from the previous subsections, UAVs when used as aerial base stations can significantly improve the performance and operation of PSNs. However, there are sill open challenges that need to be resolved. For example, the monitoring of moving objects/target-users becomes an issue after deployment in a disaster scenario. Few challenges such as traffic estimation, frequency allocation and cell association are addressed in [72]. An approach in [73] propose a disaster resilient communications architecture that facilitates edge-computing by providing a UAV cloudlet layer to aid emergency services communication links. Another approach in [74] has a uplink/downlink architecture for a Full-Duplex UAV relay to facilitate ground base stations around the UAVs. The UAVs communicate to distant ground users using non-orthogonal multiple access (NOMA) assisted networks.

Another important concern raised with UAV-based PSNs is security (see Sect. 2). In most cases, These PSNs are handling confidential information and may become vulnerable. They can also be subject to cyber and physical attacks. A variety of security concerns and challenges in drone-assisted PSNs are addressed in [75] such as: WiFi attacks, channel-jams, grey hole attacks, GPS spoofing and other issues relating to interruption, modification, interception and fabrication of information along with procedures to handle them.

7   Conclusion and Future Outlook

In this chapter, we have presented multi-UAV co-operation applications and explained how drone location prediction and trajectory optimization can be performed. We have learnt how location estimation prediction and trajectory optimization of drones can be beneficial in diverse application missions such as disaster response and other civil applications relating e.g., transportation. Various challenges in drone localization, path-planning and trajectory prediction were detailed.

To cope up with the challenges of localization of drones in application scenarios, we studied how techniques such as non-linear dynamic parameter state estimation of drones using distinct Kalman filtering techniques and sensor fusion can solve the drone localization and position prediction problem. We have also seen how Kalman filter can be used for position and velocity estimation of drones followed by location prediction with inter-drone distances and sensory measurements using the Extended Kalman filter. To cope up with sensory malfunctions and other inconsistencies of the filtering techniques, we detailed various machine learning techniques such as reinforcement learning and deep reinforcement learning. Furthermore, to cope up with the challenges of collision avoidance, trajectory optimization and path planning as well as handling of energy constraints, we have seen how a variety of reinforcement and deep reinforcement learning techniques can be used to realize the potential of multi-UAV co-operation.

Further, we presented a scenario corresponding to online orchestration and learning of network and video analytics for civil applications using multi-agent reinforcement learning techniques. These techniques feature prominent mechanisms that can be used for the 2-D and 3-D path-planning of UAVs along with network and resource allocation under bandwidth and energy constraints. Moreover, we discussed non-ML-based trajectory optimization techniques and explained how UAV-based applications can aid public safety networks.

The Road Ahead to More Open Challenges: We conclude this chapter with a list of more open challenges for multi-drone co-ordination in application missions. Addressing these challenges is essential for a variety of multi-drone applications such as aerial surveillance, deployment of UAVs as base stations and aerial mapping and monitoring that are relevant for location estimation and path planning. Few approaches such as [76] shows how joint positioning of UAVs as aerial base stations is done to provide a smart backhaul-fronthaul connectivity network. Other issues are shown in the following-

– **Excessive movements during flight with no hovering:** When the drones are in complex environments or unknown territories with unrealized threats, they tend to fly more rapidly and in different directions in a short span of time. This may be a result of collision avoidance of obstacles in the path or ineffectual attempts to explore the environment to learn threats. This leads to increased energy consumption and affects the battery capacity of drones, thus shortening their overall flight time. To avoid this issue, dynamic programming and scheduling algorithm could be useful if the drone flight plan in the mission is known apriori. The work in [77] provides two cases that show how data services using UAVs is maximized using hover time management for resource allocation, where the optimal hover time can be derived using service load requirements of ground users.
– **Air-resistance due to strong winds:** Severe wind gusts can throw the drones off-course and deviate a drone from following its optimal path. The on-board sensors are subject to vibrations during severe wind conditions and can produce noisy data that may lead to inaccurate estimates of drones parameters. Unexpected wind resistance can also hinder the trajectory learning of the drone using DRL techniques. This hindrance is possible when the drone traversing in optimal path may change course due to the impacts wind. Further research on EKF and UKF based state estimation of gyroscope readings to study the effects of wind could help in developing suitable solutions. The approach in [78] addresses the altitude control problem of UAVs in presence of wind gusts and proposes a control strategy along with stability analysis to solve the issue of air-resistance.
– **Combining LSTMs with Kalman Filters and DQN:** The non-linear state estimation of drone's dynamic parameters is done using individual time-steps of data by on-board sensors and use of the Kalman filter. Also, for the DRL techniques, the drone (agent) takes actions in a given state in independent episodes. Long short term memorys (LSTMs) can be used to utilize the information of previous time-steps of drones instead of just one time step or one episode to make predictions. This way LSTM based Kalman Filtering mechanisms and LSTMs based DRL

mechanisms can use past information of the drone(s) and make much accurate predictions. There are works that show how coupling a Kalman Filter with LSTM network improves performance and provides faster convergence of algorithms for various application purposes [79, 80].

– **Multi-drone Co-ordination under energy constraints:** In missions involving a drone swarm or a fleet of drones, it is difficult to monitor each of the drones' parameters. Factors such as malfunctioning or loss of one drone due to total battery utilization can affect the operation of other drones and compromise the overall application mission. Off-line path planning along with online path-planning can help UAVs find the nearest base stations with recharge units and help alleviate this issue and support multi-drone co-ordination even under available energy limitations. One such approach to solve the issue of multi-drone coordination under energy constraints is detailed in [81].

References

1. Yoo S, Kim K, Jung J, Chung AY, Lee J, Lee SK, ... Kim H (2015) Poster: a multi-drone platform for empowering drones' teamwork. In: Proceedings of the 21st annual international conference on mobile computing and networking, pp 275–277
2. Sorbelli FB, Corò F, Das SK, Pinotti CM (2020) Energy-constrained delivery of goods with drones under varying wind conditions. IEEE Trans Intell Transp Syst
3. Abiodun TF (2020) Usage of drones or unmanned aerial vehicles (UAVs) for effective aerial surveillance, mapping system and intelligence gathering in combating insecurity in Nigeria. Afr J Soc Sci Humanit Res 3(2):29–44
4. Bor-Yaliniz RI, El-Keyi A, Yanikomeroglu H (2016) Efficient 3-D placement of an aerial base station in next generation cellular networks. In: 2016 IEEE international conference on communications (ICC). IEEE, pp 1–5
5. Mayor V, Estepa R, Estepa A, Madinabeitia G (2019) Deploying a reliable UAV-aided communication service in disaster areas. Wirel Commun Mob Comput
6. Mishra B, Garg D, Narang P, Mishra V (2020) Drone-surveillance for search and rescue in natural disaster. Comput Commun 156:1–10
7. Kim GH, Nam JC, Mahmud I, Cho YZ (2016) Multi-drone control and network self-recovery for flying Ad Hoc networks. In: 2016 eighth international conference on ubiquitous and future networks (ICUFN). IEEE, pp 148–150
8. Ribeiro MI (2004) Kalman and extended kalman filters: concept, derivation and properties. Inst Syst Robot 43:46
9. Kalman RE (1960) A new approach to linear filtering and prediction problems
10. Qu C, Calyam P, Yu J, Vandanapu A, Opeoluwa O, Gao K, Palaniappan K (2021) DroneCO-CoNet: learning-based edge computation offloading and control networking for drone video analytics. Future Gener Comput Syst 125:247–262
11. Strehl AL, Li L, Wiewiora E, Langford J, Littman ML (2006) PAC model-free reinforcement learning. In: Proceedings of the 23rd international conference on machine learning, pp 881–888
12. Li Y (2017) Deep reinforcement learning: an overview. arXiv:1701.07274
13. Liu W, Gu W, Sheng W, Meng X, Wu Z, Chen W (2014) Decentralized multi-agent system-based cooperative frequency control for autonomous microgrids with communication constraints. IEEE Trans Sustain Energy 5(2):446–456
14. Abouheaf M, Gueaieb W, Lewis F (2020) Online model-free reinforcement learning for the automatic control of a flexible wing aircraft. IET Control Theory Appl 14(1):73–84

15. Zhu P, Wen L, Bian X, Ling H, Hu Q (2018). Vision meets drones: a challenge. arXiv:1804.07437
16. Rana T, Shankar A, Sultan MK, Patan R, Balusamy B (2019) An intelligent approach for UAV and drone privacy security using blockchain methodology. In: 2019 9th international conference on cloud computing, data science and engineering (confluence). IEEE, pp. 162–167
17. Samland F, Fruth J, Hildebrandt M, Hoppe T, Dittmann J (2012) AR. Drone: security threat analysis and exemplary attack to track persons. In: Intelligent robots and computer vision XXIX: algorithms and techniques, vol 8301. International Society for Optics and Photonics, p 83010G
18. Rodday NM, Schmidt RDO, Pras A (2016) Exploring security vulnerabilities of unmanned aerial vehicles. In: NOMS 2016-2016 IEEE/IFIP network operations and management symposium. IEEE, pp 993–994
19. Di Franco C, Buttazzo G (2015) Energy-aware coverage path planning of UAVs. In: 2015 IEEE international conference on autonomous robot systems and competitions. IEEE, pp 111–117
20. Ware J, Roy N (2016) An analysis of wind field estimation and exploitation for quadrotor flight in the urban canopy layer. In: 2016 IEEE international conference on robotics and automation (ICRA). IEEE, pp 1507–1514
21. Artemenko O, Dominic OJ, Andryeyev O, Mitschele-Thiel A (2016) Energy-aware trajectory planning for the localization of mobile devices using an unmanned aerial vehicle. In: 2016 25th international conference on computer communication and networks (ICCCN). IEEE, pp 1–9
22. Kouroshnezhad S, Peiravi A, Haghighi MS, Jolfaei A (2020) An energy-aware drone trajectory planning scheme for terrestrial sensors localization. Comput Commun 154:542–550
23. Ivancic WD, Kerczewski RJ, Murawski RW, Matheou K, Downey AN (2019) Flying drones beyond visual line of sight using 4g LTE: issues and concerns. In: 2019 integrated communications, navigation and surveillance conference (ICNS). IEEE, pp 1–13
24. Kato N, Kawamoto Y, Aneha A, Yaguchi Y, Miura R, Nakamura H, Kitashima A (2019) Location awareness system for drones flying beyond visual line of sight exploiting the 400 MHz frequency band. IEEE Wirel Commun 26(6):149–155
25. Xiong JJ, Zheng EH (2015) Optimal kalman filter for state estimation of a quadrotor UAV. Optik 126(21):2862–2868
26. Fujii K (2013) Extended kalman filter. Refernce Manual, 14–22
27. Julier SJ, Uhlmann JK (1997) New extension of the Kalman filter to nonlinear systems. In: Signal processing, sensor fusion, and target recognition VI, vol 3068. International Society for Optics and Photonics, pp 182–193
28. Wu Z, Li J, Zuo J, Li S (2018) Path planning of UAVs based on collision probability and Kalman filter. IEEE Access 6:34237–34245
29. Abdelkrim N, Aouf N, Tsourdos A, White B (2008) Robust nonlinear filtering for INS/GPS UAV localization. In: 2008 16th mediterranean conference on control and automation. IEEE, pp 695–702
30. Mao G, Drake S, Anderson BD (2007) Design of an extended kalman filter for uav localization. In: 2007 information, decision and control. IEEE, pp 224–229
31. St-Pierre M, Gingras D (2004) Comparison between the unscented Kalman filter and the extended Kalman filter for the position estimation module of an integrated navigation information system. In: IEEE Intelligent Vehicles Symposium, 2004. IEEE, pp 831–835
32. Kraft E (2003) A quaternion-based unscented Kalman filter for orientation tracking. In: Proceedings of the sixth international conference of information fusion, vol 1, no 1. IEEE Cairns, Queensland, Australia, pp 47–54
33. Tang SH, Kojima T, Namerikawa T, Yeong CF, Su ELM (2015) Unscented Kalman filter for position estimation of UAV by using image information. In: 2015 54th annual conference of the society of instrument and control engineers of Japan (SICE). IEEE, pp 695–700
34. Nemra A, Aouf N (2010) Robust INS/GPS sensor fusion for UAV localization using SDRE nonlinear filtering. IEEE Sens J 10(4):789–798
35. Abdelfatah R, Moawad A, Alshaer N, Ismail T (2021) UAV tracking system using integrated sensor fusion with RTK-GPS. In: 2021 international mobile, intelligent, and ubiquitous computing conference (MIUCC). IEEE, pp 352–356

36. Gurvits L, Ledoux J (2005) Markov property for a function of a Markov chain: a linear algebra approach. Linear Algebra Appl 404:85–117
37. Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, Riedmiller M (2013) Playing atari with deep reinforcement learning. arXiv:1312.5602
38. Tampuu A, Matiisen T, Kodelja D, Kuzovkin I, Korjus K, Aru J, Vicente R (2017) Multiagent cooperation and competition with deep reinforcement learning. PloS One 12(4):e0172395
39. Sutton RS, Barto AG (2018) Reinforcement learning: an introduction. MIT press
40. Roderick M, MacGlashan J, Tellex S (2017) Implementing the deep q-network. arXiv:1711.07478
41. González RLV, Aragone LS (2000) A Bellman's equation for minimizing the maximum cost. Indian J Pure Appl Math 31(12):1621–1632
42. Osband I, Blundell C, Pritzel A, Van Roy B (2016) Deep exploration via bootstrapped DQN. Adv Neural Inf Process Syst 29:4026–4034
43. Koushik AM, Hu F, Kumar S (2019) Deep Q-learning-based node positioning for throughput-optimal communications in dynamic UAV swarm network. IEEE Trans Cogn Commun Netw 5(3):554–566
44. Mnih V, Badia AP, Mirza M, Graves A, Lillicrap T, Harley T, ... Kavukcuoglu K (2016) Asynchronous methods for deep reinforcement learning. In: International conference on machine learning. PMLRpp, 1928–1937
45. Zhao L, Ma Y, Zou J (2020) 3D Path planning for UAV with improved double deep Q-network. In: Chinese intelligent systems conference. Springer, Singapore, pp 374–383
46. Wang Z, Schaul T, Hessel M, Hasselt H, Lanctot M, Freitas N (2016) Dueling network architectures for deep reinforcement learning. In: International conference on machine learning. PMLR, pp 1995–2003
47. Zeng Y, Xu X, Jin S, Zhang R (2021) Simultaneous navigation and radio mapping for cellular-connected UAV with deep reinforcement learning. IEEE Trans Wirel Commun
48. Yan C, Xiang X, Wang C (2020) Towards real-time path planning through deep reinforcement learning for a UAV in dynamic environments. J Intell Robot Syst 98(2):297–309
49. Grondman I, Busoniu L, Lopes GA, Babuska R (2012) A survey of actor-critic reinforcement learning: Standard and natural policy gradients. IEEE Trans Syst Man Cybern Part C (Appl Rev) 42(6):1291–1307
50. Konda VR, Tsitsiklis JN (2000) Actor-critic algorithms. In: Advances in neural information processing systems, pp 1008–1014
51. Hou Y, Liu L, Wei Q, Xu X, Chen C (2017) A novel DDPG method with prioritized experience replay. In: 2017 IEEE international conference on systems, man, and cybernetics (SMC). IEEE, pp 316–321
52. Ding R, Gao F, Shen XS (2020) 3D UAV trajectory design and frequency band allocation for energy-efficient and fair communication: a deep reinforcement learning approach. IEEE Trans Wirel Commun 19(12):7796–7809
53. Zhao YJ, Zheng Z, Zhang XY, Liu Y (2017) Q learning algorithm based UAV path learning and obstacle avoidance approach. In: 2017 36th chinese control conference (CCC). IEEE
54. Coggan M (2004) Exploration and exploitation in reinforcement learning. Research supervised by Prof. Doina Precup, CRA-W DMP Project at McGill University
55. Saxena V, Jaldèn J, Klessig H (2019) Optimal UAV base station trajectories using flow-level models for reinforcement learning. IEEE Trans Cogn Commun Netw 5(4):1101–1112
56. Yin S, Zhao S, Zhao Y, Yu FR (2019) Intelligent trajectory design in UAV-aided communications with reinforcement learning. IEEE Trans Veh Technol 68(8):8227–8231
57. Nguyen KK, Vien NA, Nguyen LD, Le MT, Hanzo L, Duong TQ (2020) Real-time energy harvesting aided scheduling in UAV-assisted D2D networks relying on deep reinforcement learning. IEEE Access 9:3638–3648
58. Langelaan J (2007) Long distance/duration trajectory optimization for small UAVs. In: AIAA guidance, navigation and control conference and exhibit, p 6737
59. Lakew DS, Masood A, Cho S (2020) 3D UAV placement and trajectory optimization in UAV assisted wireless networks. In: 2020 international conference on information networking (ICOIN). IEEE, pp 80–82

60. Xu D, Sun Y, Ng DWK, Schober R (2020) Multiuser MISO UAV communications in uncertain environments with no-fly zones: robust trajectory and resource allocation design. IEEE Trans Commun 68(5):3153–3172
61. Koyuncu E, Shabanighazikelayeh M, Seferoglu H (2018) Deployment and trajectory optimization of UAVs: a quantization theory approach. IEEE Trans Wirel Commun 17(12):8531–8546
62. Shakoor S, Kaleem Z, Do DT, Dobre OA, Jamalipour A (2020) Joint optimization of UAV 3D placement and path loss factor for energy efficient maximal coverage. IEEE Internet Things J
63. Guo Y, You C, Yin C, Zhang R (2021) UAV trajectory and communication co-design: flexible path discretization and path compression. IEEE J Sel Areas Commun
64. Zhang S, Zeng Y, Zhang R (2018) Cellular-enabled UAV communication: a connectivity-constrained trajectory optimization perspective. IEEE Trans Commun 67(3):2580–2604
65. Yang D, Dan Q, Xiao L, Liu C, Cuthbert L (2021) An efficient trajectory planning for cellular-connected UAV under the connectivity constraint. China Commun 18(2):136–151
66. Teng H, Ahmad I, Msm A, Chang K (2020) 3D optimal surveillance trajectory planning for multiple UAVs by using particle swarm optimization with surveillance area priority. IEEE Access 8:86316–86327
67. Fantacci R, Gei F, Marabissi D, Micciullo L (2016) Public safety networks evolution toward broadband: sharing infrastructures and spectrum with commercial systems. IEEE Commun Mag 54(4):24–30
68. Laoudias C, Moreira A, Kim S, Lee S, Wirola L, Fischione C (2018) A survey of enabling technologies for network localization, tracking, and navigation. IEEE Commun Surv Tutor 20(4):3607–3644
69. Naqvi SAR, Hassan SA, Pervaiz H, Ni Q (2018) Drone-aided communication as a key enabler for 5g and resilient public safety networks. IEEE Commun Mag 56(1):36–42
70. Do-Duy T Nguyen LD, Duong TQ, Khosravirad S, Claussen H (2021) Joint optimisation of real-time deployment and resource allocation for UAV-aided disaster emergency communications. IEEE J Sel Areas Commun
71. Shakoor S, Kaleem Z, Baig MI, Chughtai O, Duong TQ, Nguyen LD (2019) Role of UAVs in public safety communications: energy efficiency perspective. IEEE Access 7:140665–140679
72. Mozaffari M, Saad W, Bennis M, Nam YH, Debbah M (2019) A tutorial on UAVs for wireless networks: applications, challenges, and open problems. IEEE Commun Surv Tutor 21(3):2334–2360
73. Kaleem Z, Yousaf M, Qamar A, Ahmad A, Duong TQ, Choi W, Jamalipour A (2019) UAV-empowered disaster-resilient edge architecture for delay-sensitive communication. IEEE Netw 33(6):124–132
74. Do DT, Nguyen TTT, Le CB, Voznak M, Kaleem Z, Rabie KM (2020) UAV relaying enabled NOMA network with hybrid duplexing and multiple antennas. IEEE Access 8:186993–187007
75. He D, Chan S, Guizani M (2017) Drone-assisted public safety networks: the security aspect. IEEE Commun Mag 55(8):218–223
76. Shehzad MK, Ahmad A, Hassan SA, Jung H (2021) Backhaul-aware intelligent positioning of UAVs and association of terrestrial base stations for fronthaul connectivity. IEEE Trans Netw Sci Eng
77. Mozaffari M, Saad W, Bennis M, Debbah M (2017) Wireless communication using unmanned aerial vehicles (UAVs): optimal transport theory for hover time optimization. IEEE Trans Wirel Commun 16(12):8052–8066
78. Shi D, Wu Z, Chou W (2018) Super-twisting extended state observer and sliding mode controller for quadrotor uav attitude system in presence of wind gust and actuator faults. Electronics 7(8):128
79. Pérez-Ortiz JA, Gers FA, Eck D, Schmidhuber J (2003) Kalman filters improve LSTM network performance in problems unsolvable by traditional recurrent nets. Neural Netw 16(2):241–250

80. Coskun H, Achilles F, DiPietro R, Navab N, Tombari F (2017) Long short-term memory kalman filters: recurrent neural estimators for pose regularization. In: Proceedings of the ieee international conference on computer vision, pp 5524–5532
81. Scherer J, Rinner B (2016) Persistent multi-UAV surveillance with energy and communication constraints. In: 2016 IEEE international conference on automation science and engineering (CASE). IEEE, pp 1225–1230