IEEE ComSoc IEEE Open Journal of the Communications Society

Received XX Month, XXXX; revised XX Month, XXXX; accepted XX Month, XXXX; Date of publication XX Month, XXXX; date of current version 27 June, 2024.

Digital Object Identifier 10.1109/OJVT.2024.0627000

DRL-based UAV Path Planning for Coverage Hole Avoidance: Energy Consumption and Outage Time Minimization Trade-offs

Bahareh Jafari[1], Mazen Hasna[1] *(Senior Member, IEEE)*, Hossein Pishro-Nik[2] *(Senior Member, IEEE)*, Nizar Zorba[1] *(Senior Member, IEEE)*, Tamer Khattab[1] *(Senior Member, IEEE)*, and Hamid Saeedi[3] *(Member, IEEE)*

[1]Qatar University, Doha, Qatar
[2]University of Massachusetts, Amherst, MA, 01003, USA
[3]University of Doha for Science and Technology, Doha, Qatar

CORRESPONDING AUTHOR: Nizar Zorba (e-mail: nizarz@qu.edu.qa).

This work was supported by Qatar National Research Fund (QNRF) under grants NPRP13S-0130-200200 and ARG01-0511-230129. The work of the third author was supported by National Science Foundation under grant CNS-2150832.
A preliminary version of this paper has been presented in IEEE Global Communications Conference, Cape Town, South Africa, Dec. 2024.

ABSTRACT Coverage holes pose critical challenges to reliability of wireless networks and their quality of service (QoS) and therefore should be avoided in the coverage design. In this paper, we address this issue through the deployment of unmanned aerial vehicles (UAVs) as mobile base stations, and we propose specific UAV path planning. A point is said to be in a coverage hole if the coverage probability for that point is below a certain threshold, e.g., 90%. This definition is more suitable for applications such as surveillance or sensor networks. In this paper, we target applications such as wireless communications for which QoS requirement allow only for short time disconnections, i.e., minimal outage time. As such, in addition to avoiding coverage holes, we should also make the outage time as small as possible. By deploying a deep reinforcement learning algorithm, we find optimal UAV paths based on the two families of trajectories: spiral and oval curves, to tackle different design considerations and constraints, in terms of QoS, energy consumption and coverage hole avoidance. We show that for a typical point on the cell, there is a trade-off between minimizing the maximum outage time length and consumed mechanical energy. Our observations indicate that such a trade-off is more pronounced for spiral trajectories compared to oval trajectories, but both of them are useful depending on the QoS and energy constraints imposed by the system.

INDEX TERMS Coverage hole, QoS, path planning, UAV, mechanical energy, deep reinforcement learning (DRL).

I. INTRODUCTION

A. Background and Literature Review

THE next-generation of wireless networks impose great demands on network quality of service (QoS) and reliability [1], [2]. One of the critical challenges which affects the reliability is the existence of coverage holes within a cell. Coverage holes are defined from a statistical point of view, that is, a location on the cell for which the coverage probability is less than a certain threshold is said to be a coverage hole. Coverage holes are caused by shadowing, fading effects and radio link failure (RLF). Despite robust network-coverage planning, coverage holes may still occur due to the heterogeneity in the network and the randomness of the wireless channel, among others [3]–[5]. Consequently, to ensure reliability, it is important to effectively deal with the coverage holes. Most of the works in the literature assume a terrestrial network system model in which solutions are sought to eliminate or reduce the

coverage holes' adverse impact on the reliability of the networks. The conventional method for identifying coverage gaps in cellular networks involves conducting costly drive tests paired with mathematical propagation models to create a radio map [6]. However, this technique is unreliable. In order to address these challenges, 3GPP has introduced the minimization of drive tests (MDT) technique [7]. The MDT mechanism enables the serving base stations (BS) to utilize user equipment (UE) measurement reports as well as RLF report to create coverage maps, aiding in the identification of coverage holes [8]. Needless to say, MDT may still suffer from positioning error on coverage estimation.

Due to the unpredictable nature of the operational environment, recent studies have explored an alternative approach, i.e., utilizing machine learning (ML) algorithms, to detect the coverage holes. The work of [9] proposes an ML-based approach for detecting coverage holes without relying on location information or measurement reports. In fact, having access to the channel estimation database at the BS can facilitate data-driven ML methods to uncover hidden spatial patterns and detect anomalies like coverage gaps. However, [9] does not propose any solutions to deal with the coverage holes. The study by [10] pioneers the integration of artificial intelligence (AI) and mobile edge computing (MEC) within 6G-enabled IoT frameworks, presenting an advanced solution to the challenge of coverage holes. In such a framework, inspired by disc model and the confident information coverage (CIC) model for sensor coverage [11], [12], a reinforcement learning (RL) algorithm is developed based on the movement of mobile edge nodes, enabling more precise and energy-efficient recovery of coverage holes.

Recent advancements in unmanned aerial vehicles (UAVs) technology have leveraged their intrinsic mobility and operational flexibility, facilitating their deployment across surveillance and communication domains [13]–[15]. For the latter application, they are also referred to as aerial base stations (ABSs). An important issue in deploying the UAVs is their mechanical energy consumption as it directly dictates the total time the UAV can be on air before returning to its base for a recharge. It has been shown in literature [16]–[18] that the energy consumption is profoundly influenced by both the trajectory of the traveled path and the UAV's speed and acceleration profile along that path. In telecommunication applications, the energy consumption of the UAV encompasses both propulsion-related mechanical power for aerial maneuvers and RF transmission power for communication functionalities. In practice, the mechanical power has a considerably higher share of the total energy. As such, an energy-efficient path planning of moving UAVs should focus more on optimizing the mechanical energy rather than the energy used for RF communications. The analytical framework presented in [19] establishes a rigorous closed-form expressions for fixed-wing UAV energy consumption for 2-D movement.

UAV platforms offer strategic solutions for addressing coverage deficiencies through boosting the coverage of existing terrestrial cellular networks, while simultaneously providing emergency capacity enhancement for base stations to accommodate the demand for ever-increasing data transmission rates [20]–[23]. As such, for the UAV-assisted communication networks, designing an optimal path or optimal placement while addressing the coverage holes is of utmost significance. The study by [24] explores the use of UAV-based base stations for detecting coverage gaps and delivering temporary on-demand coverage. They employ Q-learning as a trajectory planner to autonomously identify coverage holes within a specified area. Inspired by stochastic geometry analysis, the work of [25] proposes an RL framework for concurrent optimization of UAV placement and antenna beam pattern configuration, with the explicit objective of minimizing maximum outage probability metrics.

B. Contributions

In this paper, in contrast to the existing literature, we use a totally different approach to deal with coverage holes. Our aim is to design a network of moving UAVs such that for any typical point on the cell, coverage holes are avoided at any time. In other words, we are not going to first detect the holes and then try to position the UAV on a particular location to avoid that hole, as our UAVs will be continuously moving. This is beneficial as we need not to be concerned with the dynamic nature of coverage holes any more. To attain this goal, we utilize the framework from [26], which suggests two types of trajectories and speed profiles: spiral paths and oval paths. When UAVs follow these paths, they ensure relatively uniform coverage for users at any location within the cell.

We first design UAV trajectories which avoid coverage holes with minimal energy consumption. Assume that we consider a point to be a coverage hole if it has an outage probability [27] higher than, e.g., 10%. Based on the definition, if we can maintain coverage probability of, e.g., 90%, over all points of the cell, we have fully addressed the issue of coverage holes. We, in fact, show in this paper that this is possible through proper UAV path planning, where oval paths require slightly higher energy compared to spiral paths.

Despite the above achievement, for certain QoS requirements, such as in wireless communications, the above scenario may not be appealing. In the aforementioned scenario, we guarantee that any typical point of the cell will not be uncovered more than 10% of the time, i.e., outage probability is less than 10%. Consider a time span of τ seconds, e.g., 1000 s. This means that we may encounter a worst case scenario with $\tau/10$ s, e.g., 100 s, of continuous outage for a typical point which may not be satisfying the required QoS. As an example, the acceptable packet delay budget for most existing services is anywhere between a fraction of a millisecond to a few hundreds of milliseconds [28], Table 6.1.7-A, [29], Table 5.7.4-1, and [30], P. 16. As such, another

important metric to consider is the maximum outage time for any point on the cell, denoted by T_{max}.

Accordingly, we aim to find curves that provide the minimum value of T_{max} for all points on the cell. We show that we can achieve $T_{max} = 0$, i.e., no outage for any point of the cell, but at the expense of mechanical energy getting unacceptably large when we use spiral paths. Using oval paths, we can make a compromise, i.e., we can lower energy consumption if we can tolerate larger than zero values for T_{max}.

Alternatively, we may want to keep T_{max} for any point of the cell below a given value and try to find a path to guarantee such requirement, with minimum energy. This will provide a controlled compromise between T_{max} and the consumed energy. Accordingly, for different values of T_{max} and starting from 0, we show there is a well-established trade-off between the energy and maximum outage time length.

To find optimal paths in all above-mentioned scenarios, we use deep reinforcement learning (DRL) through the implementation of an actor-critic DRL framework, called deep deterministic policy gradient algorithm (DDPG) [31]. To get closer to the optimal result, we use an improved version of this algorithm called twin delayed deep deterministic policy gradient (TD3) [32], which deploys two critic networks instead of one at the expense of more computational complexity.

This paper is structured as follows: section II outlines the system model and introduces essential preliminary formulations. In section III, the optimization problems are proposed to be solved by the TD3 algorithm, detailed in Section IV. Section V presents the numerical results and conclusions are drawn in Section V.

II. PRELIMINARIES AND SYSTEM MODEL

In this section, we explain the system model and review the concept of the deployed trajectories, i.e., spiral, and oval curves. Then we review the formulations for propulsion energy consumption of fixed-wing UAVs. Finally, we clearly define coverage holes and outage time. In this paper, we use capital bold letters to denote vectors and small regular letters to denote scalars.

A. Spiral Trajectories

The family of general curves below represent the trajectory of a spiral family over the cell:

$$\boldsymbol{Q}_s(s) = \left[\rho s^k \cos(\zeta s), \rho s^k \sin(\zeta s)\right], \qquad s \in [0,1], \quad (1)$$

where ρ is the radius of the cell, s is an arbitrary constant i.e., $0 \leq s \leq 1$, and k and ζ are constants that determine the shape of the curve. Specifically, by setting $\zeta = 0$ and $k = 1$, a series of curves are generated, each corresponding a radius of the cell. This configuration forms the *radial* trajectory, representing the most straightforward trajectory in this family.

Each UAV begins its flight from the cell centroid, moving towards the cell edge along the path $\boldsymbol{Q}_s(s)$ within τ seconds. Upon reaching the cell edge, the UAV returns to the origin along the same path, subsequently following the trajectory $-\boldsymbol{Q}_s(s)$ to reach the opposite side of cell edge and this process repeats continuously.

For the i'th UAV flying on the spiral trajectory, its instantaneous location can be determined by setting $s = \sqrt[2k]{\frac{t-k\tau-T_i}{\tau}}$ in (1). In this setup, N UAVs start their flights at times $T_1, T_2, ..., T_N$ where $T_i = i\tau/(N+1)$, $i \in \{1,2,...,N\}$, and use rotational phase $\theta_i = 2\pi(i-1)/N, i \in \{1,2,...,N\}$. In other words, for the ith UAV we have

$$\begin{aligned}\boldsymbol{Q}_{s_i}(t) =& [x_i(t), y_i(t)] \qquad (2)\\ =& \left[\rho\sqrt{\frac{t-k\tau-T_i}{\tau}}\cos\left(\zeta\sqrt[2k]{\frac{t-k\tau-T_i}{\tau}}+\theta_i\right),\dots\right.\\ & \left.\rho\sqrt{\frac{t-k\tau-T_i}{\tau}}\sin\left(\zeta\sqrt[2k]{\frac{t-k\tau-T_i}{\tau}}+\theta_i\right)\right].\end{aligned}$$

It is shown in [26] that if the above trajectories are followed, uniform coverage is guaranteed over the cell and the coverage probability tends to 1 as N increases.

B. Oval Trajectories

The family of general curves below represent the oval family trajectory over the cell:

$$\boldsymbol{Q}_o(s) = \left[q\cos\left(\frac{\pi}{2}s\right), q\sin\left(\frac{\pi}{2}s\right)\right], \qquad s \in [0,1], \quad (3)$$

where $q = a + (b-a)s$, for any given $a, b \in \mathbb{R}^+$, $0 \leq a \leq b \leq \rho$. The two constants a and b determine the shape of oval curves and its overall characteristics. There are two special cases of oval trajectories, namely *ring* and *ellipse* trajectories. By setting $a = b = \rho$, we achieve a set of curves, each being a circle with radius ρ. This path is called the ring trajectory, recognized as the simplest and most intuitive option within this family of trajectories. Another specific case can be obtained by assigning $b = \rho$ and $0 \leq a \leq \rho$ which is called *ellipse* trajectory.

The instantaneous location of the i'th UAV on the oval trajectory can be obtained by setting $s = (t - T_i)/(\frac{\tau}{4})$ in (3). In this setup, N UAVs start their flights at times $T_1, T_2, ..., T_N$ where $T_i = i\tau/N, i \in \{1,2,...,N\}$, with rotational phase θ_i where $\theta_i = 2\pi i/N, i \in \{1,2,...,N\}$.[1] In other words, for the i'th UAV we have

$$\begin{aligned}&\boldsymbol{Q}_{o_i}(t) = [x_i(t), y_i(t)] = \qquad (4)\\ &\left[q\cos\left(\frac{\pi}{2}\left(\frac{t-T_i}{\frac{\tau}{4}}\right)+\theta_i\right), q\sin\left(\frac{\pi}{2}\left(\frac{t-T_i}{\frac{\tau}{4}}\right)+\theta_i\right)\right],\end{aligned}$$

where $q = a + (b-a)(t-T_i)/\left(\frac{\tau}{4}\right)$, $0 \leq t - T_i \leq \frac{\tau}{4}$, and $0 \leq a \leq b \leq \rho$.

[1]The UAV's can start from anywhere on the curve as long as they maintain the rotational phase difference with respect to each other.

The extended oval trajectory $\boldsymbol{Z}(t)$ over $0 \leq t \leq \tau$ can be obtained by the reflection of $\boldsymbol{Q}_o(t)$ over x-axis and y-axis which is represented as follows:

$$\boldsymbol{Z}(t) = \begin{cases} [x(t), y(t)] & 1^{st} \text{ quadrant} \\ [-x(t), y(t)] & 2^{nd} \text{ quadrant} \\ [-x(t), -y(t)] & 3^{rd} \text{ quadrant} \\ [x(t), -y(t)] & 4^{th} \text{ quadrant} \end{cases} \tag{5}$$

C. Energy Consumption Model

For the rest of the paper, if a formulation applies to both spiral and oval paths, instead of $\boldsymbol{Q}_s(t)$ or $\boldsymbol{Q}_o(t)$, we use $\boldsymbol{Q}(t)$ to represent the path. The propulsion energy consumption of fixed-wing UAV moving on a 2-dimensional plane is given by [19]:

$$E = \int_0^\tau c_1 \|\boldsymbol{V}(t)\|^3 + \frac{c_2}{\|\boldsymbol{V}(t)\|} \left(1 + \frac{\|\boldsymbol{A}(t)\|^2 - \frac{(\boldsymbol{A}(t).\boldsymbol{V}^T(t))^2}{\|\boldsymbol{V}(t)\|^2}}{g^2} \right) dt + \int_0^\tau m\boldsymbol{A}(t).\boldsymbol{V}^T(t) dt, \tag{6}$$

where $\boldsymbol{V}(t) = \dot{\boldsymbol{Q}}(t) = [\dot{x}(t), \dot{y}(t)]$ and $\boldsymbol{A}(t) = \ddot{\boldsymbol{Q}}(t) = [\ddot{x}(t), \ddot{y}(t)]$ indicate the instantaneous speed and acceleration vectors respectively, and c_1 and c_2 are two constants defined as

$$c_1 \triangleq \frac{1}{2}\rho_a C_{D_0} S \ , \ c_2 \triangleq \frac{2W^2}{(\pi e_0 \mathcal{A}_{\mathcal{R}})\rho_a S}. \tag{7}$$

In the above equations, $W = mg$ represents the gravitational force, with m denoting the UAVs mass including its payload, and g is the gravitational acceleration. Furthermore, ρ_a refers to the air density measured in kg/m^3, C_{D_0} denotes the zero-lift drag coefficient, S indicates a reference area (e.g., the wing area), e_0 is the Oswald efficiency which accounts for the aerodynamic efficiency of the wing, and $\mathcal{A}_{\mathcal{R}}$ denotes the aspect ratio of the wing which influences the lift and drag characteristics.

This provides a measure of the average power required for the UAV's movement over a given period. Moreover, the sum of the 2 integrands is the instantaneous power, P_{inst}, which can be written as

$$P_{inst}(t) = c_1 \|\boldsymbol{V}(t)\|^3 + \frac{c_2}{\|\boldsymbol{V}(t)\|} \left(1 + \frac{\|\boldsymbol{A}(t)\|^2 - \frac{(\boldsymbol{A}(t).\boldsymbol{V}^T(t))^2}{\|\boldsymbol{V}(t)\|^2}}{g^2} \right) + m\boldsymbol{A}(t).\boldsymbol{V}^T(t). \tag{8}$$

D. Coverage Hole

For any arbitrary user location within the cellular domain, the coverage probability P_c is defined as the probability that the received signal-to-noise ratio (SNR) exceeds a specific threshold γ. Mathematically, this can be expressed as:

$$P_c = P\left(SNR_{Rx} \geq \gamma\right). \tag{9}$$

Consequently, the outage probability is obtained as $1 - P_c$. The threshold γ is determined by the system specifications, particularly the receiver's capability to successfully recover data at lower SNR levels.

For a UAV flying horizontally on the trajectory $\boldsymbol{Q}(t) = [x(t), y(t)]^T \in \mathbb{R}^{2\times 1}$ at a constant altitude H, and assuming line of sight (LoS) link [2], the received SNR can be expressed as:

$$SNR_{Rx} = \frac{\gamma_0}{H^2 + ||\boldsymbol{Q}(t)||^2}, \tag{10}$$

where we either consider $\boldsymbol{Q}(t) = \boldsymbol{Q}_s(t)$ for spiral paths or $\boldsymbol{Q}(t) = \boldsymbol{Q}_o(t)$ for oval paths. Moreover, $\gamma_0 = \beta_0 P_T/\sigma^2$ is the reference received signal-to-noise ratio (SNR) in which P_T denotes the transmission power, $\beta_0 P_T$ is the received power at the reference distance $d_0 = 1$, and σ^2 is the white Gaussian noise power. In this work, we consider a scenario in which the UAV frequencies are orthogonal to avoid interference. At each moment, the user is served by the closest UAV with a given frequency. Once the user is out of the coverage area of that UAV, it is handed-off to another UAV with different frequency. Exploring more complex scenarios that involve interference, e.g.., assuming all UAVs reuse the same frequency, could be a valuable extension of this study.

Given the above definition for coverage, a point is referred to as a coverage hole in the literature if P_c is less than a certain value, e.g., less than 0.9, or alternatively, the outage probability is greater than 0.1. Even if a point is not a coverage hole, e.g., if $P_c > 0.9$, for a time span of τ, the outage time can be as high as $\tau(1 - P_c)$ which may not be acceptable in practice [28], [29], [30]. For any point of the cell, we denote the maximum outage time as T_{max} and we either try to directly minimize it or make sure it is less than a given threshold.

III. COVERAGE HOLE AVOIDANCE THROUGH UAV PATH PLANNING

A. Preliminaries

As mentioned before, in the literature, a point is deemed to be in coverage hole if the coverage probability for that point is less than a certain threshold. In the next subsection B, we find energy-optimized spiral and oval paths that can provide coverage hole-free cells as will be shown in the corresponding simulation results. However, we observe that, the maximum outage time, T_{max}, for some points might be too high. As such, in Subsection C, using both oval and spiral trajectories, we obtain paths with min. value for T_{max}. We will see in the simulation results that for spiral trajectories, we can achieve $T_{max} = 0$, for any point at the expense of considerably higher mechanical energy compared to Subsection B where we minimized the energy. We also show that using oval trajectories, we can have

[2]It is important to note that although as we mentioned, coverage holes are caused due to shadowing and fading effects in terrestrial networks, when UAV's are deployed to deal with them, we can assume LoS channels due to their higher altitude with respect to ground users.

smaller value of consumed energy at the expense of an increase in minimum value of T_{max}. This shows that there might be a trade-off between the min. value of T_{max} and the consumed mechanical energy. As such in Subsection D, we propose a framework in which we try to minimize energy while guaranteeing a fixed value of T_{max} for any point on the cell. In simulation results, we see that there is in fact a well-established trade-off between these 2 parameters.

B. Path Planning to Minimize Mechanical Energy Consumption while Addressing the Coverage Probability

For the UAVs moving on the spiral trajectories with a certain period τ, the optimal path planning problem can be formulated as:

$$\min_{\boldsymbol{Q}_s(t)} \; E, \tag{11a}$$

$$s.t. \; C_q : \boldsymbol{Q}_s(t) = \left[\rho\sqrt{\frac{t}{\tau}}\cos(\zeta \sqrt[2k]{\frac{t}{\tau}}), \rho\sqrt{\frac{t}{\tau}}\sin(\zeta \sqrt[2k]{\frac{t}{\tau}})\right]. \tag{11b}$$

In this problem, C_q limits the path to a spiral path with parameters k, ζ, with cell radius ρ and travel time τ to provide the uniform coverage on the cell. By applying C_q into (11a), we end up with the following optimization problem:

$$\mathbf{P1} : \min_{\tau,k,\zeta} E(k,\zeta) = \int_0^{\tau} P_{inst}(t,\tau,k,\zeta)dt, \tag{12}$$

where $P_{inst}(t,\tau,k,\zeta)$ is the instantaneous power of the UAVs moving on the spiral trajectories. We derived the closed-form formulation for $P_{inst}(t,\tau,k,\zeta)$ by replacing $\boldsymbol{V}(t) = \dot{\boldsymbol{Q}}_s(t)$ and $\boldsymbol{A}(t) = \ddot{\boldsymbol{Q}}_s(t)$ in (8) as follows:

$$P_{inst}(t,\tau,k,\zeta) = \frac{1}{8}c_1\left[\frac{\rho^2(k^2+\zeta^2(\frac{t}{\tau})^{\frac{1}{k}}}{k^2t\tau}\right]^{\frac{3}{2}} + 2c_2\left[1+\frac{\rho^2\zeta^2(\frac{t}{\tau})^{\frac{1}{k}}(k+k^2+\zeta^2(\frac{t}{\tau})^{\frac{1}{k}})^2}{16k^4t^3g^2\tau(k^2+\zeta^2(\frac{t}{\tau})^{\frac{1}{k}})}\right]\times \frac{1}{\sqrt{\frac{\rho^2(k^2+\zeta^2(\frac{t}{\tau})^{\frac{1}{k}})}{k^2t\tau}}} + \frac{m\rho^2\left(-k^3-(-1+k)\zeta^2(\frac{t}{\tau})^{\frac{1}{k}}\right)}{8k^3t^2\tau}. \tag{13}$$

Through the solution of this optimization problem, we aim to derive the optimal parameters k, ζ, and τ that will yield a trajectory minimizing the UAV's energy consumption. For the optimized path, we measure T_{max} for each point within a cell.

In the case of UAVs moving on oval trajectories, we must solve the following path planning problem:

$$\min_{\boldsymbol{Q}_o(t)} \; E, \tag{14a}$$

$$s.t. \; C_q : \boldsymbol{Q}_o(t) = \left[q\cos\left(\frac{\pi}{2}\left(\frac{t}{\frac{\tau}{4}}\right)\right), q\sin\left(\frac{\pi}{2}\left(\frac{t}{\frac{\tau}{4}}\right)\right)\right]. \tag{14b}$$

In this problem, C_q limits the path to an oval path with parameters a, b, with cell radius ρ and travel time τ. By applying C_q in (14a), we end up with the following optimization problem where $P_{inst}(t,\tau,a,b)$ is the instantaneous power of the UAVs moving on the oval trajectories, and we derived it by substituting $\boldsymbol{V}(t) = \dot{\boldsymbol{Q}}_o(t)$ and $\boldsymbol{A}(t) = \ddot{\boldsymbol{Q}}_o(t)$ into (8) as follows:

$$\mathbf{P2} : \min_{\tau,a,b} E(a,b) = \int_0^{\tau} P_{inst}(t,\tau,a,b)dt, \tag{15}$$

$$P_{inst}(t,\tau,a,b) = \frac{\tau}{2\pi q}\left(c_2 + \frac{16\pi^4q^2(c_2+c_1g^2q^2)}{g^2\tau^4}\right), \tag{16}$$

In the above equation, $q = a + (b-a)\left(\frac{t}{\tau/4}\right)$, and $0 \le a \le b \le \rho$.

Through the solution of this optimization problem, we obtain the optimal values for a, b, and τ to achieve an optimal path with minimal energy consumption. Then we measure T_{max} for each point within a cell.

C. Path Planning to Minimize Maximum Outage Time

In this subsection, we try to obtain paths with minimum T_{max} according to the following problems:

$$\mathbf{P3.a} : \min_{\boldsymbol{Q}_s(t)} T_{max}, \tag{17a}$$

$$s.t. \; C_{\boldsymbol{Q}_s} : (2), \tag{17b}$$

$$\mathbf{P3.b} : \min_{\boldsymbol{Q}_o(t)} T_{max}, \tag{18a}$$

$$s.t. \; C_{\boldsymbol{Q}_o} : (4). \tag{18b}$$

In $P3.a$, $C_{\boldsymbol{Q}_s}$ limits the path to a spiral path with parameters k, ζ, with cell radius ρ and travel time τ. Similarly, in $P3.b$, $C_{\boldsymbol{Q}_o}$ limits the path to an oval path with parameters a, b, with cell radius ρ and travel time τ.

D. Path Planning for Minimum Energy Consumption with Guaranteed Maximum Outage Time

In this part, we propose a new scenario to establish a controlled compromise between T_{max} and the consumed energy. To this end, focusing on spiral paths, we propose the optimization problem P4 in which we minimize the mechanical energy while guaranteeing T_{max} to be lower than a given threshold T_{th} for all point on the cell.

$$\mathbf{P4} : \min_{\boldsymbol{Q}_s(t)} E, \tag{19a}$$

$$s.t. \; C_{\boldsymbol{Q}} : (2), \tag{19b}$$

$$C_t : T_{max} \le T_{th}. \tag{19c}$$

Through this optimization problem, we obtain an optimal trajectory with optimized parameters k and ζ leading to a robust coverage planning with the minimum energy consumption.

It is worth noting that according to simulation results for problem **P3.b** in which we minimize T_{max} for oval path, the minimum value is bounded away form zero in contrast to the case with spiral path. As such, we can not demonstrate the trade-off between T_{max} and consumed energy as efficient as spiral paths and thus, we only consider spiral paths for this problem.

IV. DEEP REINFORCEMENT LEARNING BASED SOLUTION

The trajectory optimization in all the proposed optimization problems could be complicated to solve directly due to their non-convex and non-linear objective functions and constraints. As such, we propose a model-free DRL approach to derive optimal paths by reformulating the problem within the framework of a Markov decision process (MDP). MDP can be defined as a tuple $(\mathcal{S}, \mathcal{A}, n, \mathcal{P}_{ss'}, \mathcal{R})$, where $\mathcal{S}$ and $\mathcal{A}$ represent the state and action spaces, respectively. Specifically, $\mathbf{s}_n \in \mathcal{S}$ is the state and $\mathbf{a}_n \in \mathcal{A}$ is the action at time step $n \in \mathcal{N} = \{1, 2, ...\}$. In addition, $\mathcal{P}_{ss'}$ denotes the transition probability where $\mathcal{P}_{ss'} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow [0, 1]$ represents the probability of transitioning to the next state $\mathbf{s}' = \mathbf{s}_{n+1}$ upon taking action $\mathbf{a}_n$ in state $\mathbf{s}_n$. The reward function $\mathcal{R} : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ assigns an evaluation metric to each state-action pair $(\mathbf{s}_n, \mathbf{a}_n)$, quantifying the optimality of the agent's decisions.

A. MDP reformulation

We formulate the trajectory of UAV-assisted network as follows:

- *Agent*: The UAV operates as an intelligent agent, continuously interacting with its environment to derive an optimal policy π_θ, with parameter θ, that maps state space to action space while maximizing the reward.
- *State space*: The state space $\mathcal{S}$ consists of the trajectory parameters that define the UAV's path. Specifically, for a fixed τ, the UAV state flying along the spiral trajectory with two parameters k and ζ at time step n will be $\mathbf{s}_n = [k_n, \zeta_n]$ and the UAV state flying along the oval trajectory with two parameters a and b at time step n will be $\mathbf{s}_n = [a_n, b_n]$.
- *Action space*: The action space $\mathcal{A}$ consists of 2-dimensional vector $\mathbf{a}_n = [\Delta k_n, \Delta \zeta_n]$ for the spiral trajectory and $\mathbf{a}_n = [\Delta a_n, \Delta b_n]$[3] for the oval trajectory, where each component is bounded within $[-1, 1]$. These values represent normalized parameter adjustments determined by the agent at each decision epoch. The UAV at state $\mathbf{s}_n$ chooses the action $\mathbf{a}_n$ of action space which leads to next state $\mathbf{s}_{n+1}$.
- *Reward*: When the UAV takes an action $\mathbf{a}_n$ in state $\mathbf{s}_n$, it achieves its reward. Specifically, the reward function associated with each optimization problem **P1** and **P2** can be characterized by energy reward, which is formulated as the negative of the UAV's energy consumption, providing an incentive to minimize energy expenditure. The corresponding formulations are shown in (20a) and (20b) for **P1** and **P2**, respectively:

$$R_{s,E}(n) = -E(n, k, \zeta), \tag{20a}$$
$$R_{o,E}(n) = -E(n, a, b). \tag{20b}$$

The reward functions for the optimization problems **P3.a** and **P3.b** can be characterized by the T_{max} reward as shown in (21a) and (21b), respectively:

$$R_{s,T}(n) = -T_{max}(n, k, \zeta), \tag{21a}$$
$$R_{o,T}(n) = -T_{max}(n, a, b). \tag{21b}$$

This setting compels UAVs to fly along optimal trajectories leading to the minimum T_{max}. Finally, the reward function regarding the optimization problem **P4** can be expressed as follows:

$$R_{s,ET}(n) = -E(n, k, \zeta) - R_c(n), \tag{22}$$

where the penalty term $R_c(n)$ is used to impose the required cap on T_{max}, i.e.,

$$R_c(n) = \lambda_c \sum_{i,j} \mathbb{I}\left(T_{max}(n, k, \zeta) > T_{th}\right), \tag{23}$$

where $\mathbb{I}(\cdot)$ denotes an indicator function, which returns 1 if the condition within its argument is satisfied, and 0 otherwise. The summation is taken over T_{max} corresponding to all grid locations (i, j) when deploying a spiral trajectory with parameter k and ζ at time step n. λ_c is a coefficient to adjust the penalty impact.

B. The DRL Algorithm

The continuity of the UAV's trajectory can be effectively handled through the implementation of an actor-critic DRL framework, specifically the DDPG, which leverages two distinct deep neural networks (DNNs) for the approximation of the policy and value functions independently. Let θ represent the parameter of the DNN that defines the UAV's policy function, i.e., the actor network. Operating within a deterministic policy paradigm, this parameterized actor network maps each observed state $\mathbf{s}_n$ to a unique, deterministic action $\mathbf{a}_n = \pi(\mathbf{s}_n|\theta)$, with the objective of optimizing the expected cumulative discounted reward trajectory. The objective function can be written as:

$$J(\theta) = \mathbb{E}_{\mathbf{s}\sim d^\pi}[Q^\pi(\mathbf{s}, \pi(\mathbf{s}|\theta))] \approx \mathbb{E}_{\mathbf{s}_n\sim\mathcal{B}}[Q^\pi(\mathbf{s}_n, \pi(\mathbf{s}_n|\theta))], \tag{24}$$

where $d^\pi(\mathbf{s})$ represents the stationary state distribution corresponding to the deterministic policy $\pi(\mathbf{s}_n|\theta)$, $Q^\pi(\mathbf{s}, \mathbf{a})$ denotes the value (Q-value) function, defined as $Q^\pi(\mathbf{s}, \mathbf{a}) = \mathbb{E}_\pi\left[\sum_{k=0}^{\infty} \gamma^k r_{n+k} | \mathbf{s}_n = \mathbf{s}, \mathbf{a}_n = \mathbf{a}\right]$ where γ is the discount factor and r_n denotes the immediate reward at time step n, and $\mathcal{B}$ represents the replay buffer. The Q-value $Q^\pi(\mathbf{s}_n, \pi(\mathbf{s}_n|\theta))$ serves as an evaluation metric for the

[3]The action vector "$\mathbf{a}_n$" (boldface) shall not be confused with "a_n" that denotes the oval curve parameter.

effectiveness of the policy $\pi(\mathbf{s}_n|\theta)$, i.e, the higher the Q-value, the more preferred is the action $\mathbf{a}_n$, leading to finding a policy that maximizes the cumulative discounted reward. To estimate the Q-value function, DDPG utilizes a critic network with parameter ω, denoted as $Q^\pi(\mathbf{s}_n, \mathbf{a}_n|\omega)$, i.e., $Q^\pi(\mathbf{s}_n, \mathbf{a}_n|\omega) \approx Q^\pi(\mathbf{s}_n, \pi(\mathbf{s}_n|\theta))$. As such $J(\theta)$ depends on both θ and ω of the DNN parameters applied to actor and critic networks, respectively. Additionally, to improve training stability, DDPG utilizes target networks for both the actor and critic, with parameters θ' and ω', respectively.

To find the optimal policy, we leverage the deterministic policy gradient (DPG) theorem [31] to optimize the objective function $J(\theta)$ with respect to the policy parameter θ. As such, the policy gradient can be estimated as follows:

$$\nabla_\theta J(\theta) = \mathbb{E}_{\mathbf{s}_n \sim \mathcal{B}}[\nabla_a Q^\pi(\mathbf{s}_n, \mathbf{a}|\omega)|_{\mathbf{a}=\pi(\mathbf{s}_n)} \nabla_\theta \pi(\mathbf{s}_n|\theta)]. \tag{25}$$

The critic network's parameters can be optimized through temporal-difference (TD) error between the current critic network's Q-value estimation $Q^\pi(\mathbf{s}_n, \mathbf{a}_n|\omega)$ and its target, using Bellman equation: $y_n = r_n + \gamma Q^\pi(\mathbf{s}_n, \mathbf{a}_n|\omega')$, where $\mathbf{a}_n = \pi(\mathbf{s}_n|\theta')$, r_n represents the instantaneous reward and ω' denotes the parameter of the target critic network. Then, the optimization of critic network parameters is achieved through minimization of the TD error, expressed as:

$$L(\omega) = \mathbb{E}_{(\mathbf{s}_n, \mathbf{a}_n, r_n, \mathbf{s}_{n+1}) \sim \mathcal{B}}[|y_n - Q^\pi(\mathbf{s}_n, \mathbf{a}_n|\omega)|^2]. \tag{26}$$

Through temporal difference learning, both actor and critic networks are updated. Also, the target networks are updated using soft updates as follows:

$$\theta' \leftarrow \xi\theta + (1-\xi)\theta', \tag{27}$$

$$\omega' \leftarrow \xi\omega + (1-\xi)\omega', \tag{28}$$

where ξ is the soft update parameter and $\xi \ll 1$.

In the aforementioned off-policy and model-free RL approach, i.e., the DDPG algorithm, substituting a function approximator $Q^\pi(\mathbf{s}_n, \mathbf{a}_n|\omega)$ for the true Q-value function $Q^\pi(\mathbf{s}_n, \mathbf{a}_n|\theta)$ may introduce bias. This overestimation bias in Q-value estimation can lead to sub-optimal policies, which could eventually lead to a local optimum. This was in fact the case when we applied it to our **P4** problem, when we obtained the optimal result through exhaustive search.

To address this issue, twin delayed deep deterministic policy gradient (TD3) is proposed. TD3 deploys two critic networks with parameters ω_1 and ω_2 and their corresponding target networks with parameters ω_1' and ω_2'. Then their minimum is applied to approximate the target Q-value function, i.e., $y_n = r_n + \gamma \min_{i=1,2} Q^\pi(\mathbf{s}_{n+1}, \tilde{\mathbf{a}}_{n+1}|\omega_i')$, where $\tilde{\mathbf{a}}_{n+1} = \text{clip}[\pi(\mathbf{s}_{n+1}|\theta') + \epsilon, -c, c]$, with $\epsilon \sim N(0, \sigma)$ being the target policy smoothing noise and c is the noise clipping limit. The *clip* function limits its first input argument between second and third inputs. Through temporal difference learning, two critic networks will be trained by minimizing:

$$\begin{cases} L(\omega_1) = \mathbb{E}_{(\mathbf{s}_n, \mathbf{a}_n, r_n, \mathbf{s}_{n+1}) \sim \mathcal{B}}[|y_n - Q^\pi(\mathbf{s}_n, \mathbf{a}_n|\omega_1)|^2], \\ L(\omega_2) = \mathbb{E}_{(\mathbf{s}_n, \mathbf{a}_n, r_n, \mathbf{s}_{n+1}) \sim \mathcal{B}}[|y_n - Q^\pi(\mathbf{s}_n, \mathbf{a}_n|\omega_2)|^2]. \end{cases} \tag{29}$$

Similar to DDPG, TD3 uses soft updates for the target networks: $\omega_i' \leftarrow \xi\omega_i + (1-\xi)\omega_i'$, $\theta' \leftarrow \xi\theta + (1-\xi)\theta'$ where $\xi \ll 1$ is the soft update parameter. Although DDPG works pretty well for other problems, i.e., **P1**, **P2**, and **P3**, we used TD3 for them as well. The downside is indeed more implementation complexity.

Algorithm 1 TD3 algorithm for trajectory optimization in UAV-aided networks

Initialization:
  Initialize the network: (ρ, τ, H, γ_0, γ, T_{th})
  Initialize the UAV parameters (c_1, c_2, m, g)
procedure CALCULATE ENERGY AND T_{max}
  Create trajectories, i.e., (2) and (4); $(x_s, y_s) \leftarrow \boldsymbol{Q}_s(k, \zeta, t)$, $(x_o, y_o) \leftarrow \boldsymbol{Q}_o(a, b, t)$
  Create the energy consumption model for each trajectory, i.e., $E(k, \zeta) = \int_0^\tau$ Eq. (13) dt, $E(a, b) = \int_0^\tau$ Eq. (16) dt.
  Compute coverage probability and T_{max} for each trajectory.
  return $E(k, \zeta), E(a, b), T_{max}(k, \zeta), T_{max}(a, b)$.
end procedure
Initialize environment
  Initialize critic networks with parameters ω_1, ω_2, i.e., $Q(\mathbf{s}_n, \mathbf{a}_n, \omega_1), Q(\mathbf{s}_n, \mathbf{a}_n, \omega_2)$
  Initialize actor network $\pi(\mathbf{s}_n, \theta)$ with parameter θ
  Initialize target networks: $\omega_1' \leftarrow \omega_1, \omega_2' \leftarrow \omega_2, \theta' \leftarrow \theta$
  Initialize replay buffer $\mathcal{B}$, exploration noise $N(0, \sigma)$
  Select reward function from $(R_{s,E}(n), R_{o,E}(n), R_{s,T}(n), R_{o,T}(n), R_{s,ET}(n))$.
for episode $= 1$ to N_{episodes} **do**
    Receive initial states s_0
  for $n = 1$ to N_{steps} **do**
      Observe state $\mathbf{s}_n$
      Obtain the trajectory $\boldsymbol{Q}(t)$
      Convert $\boldsymbol{Q}(t)$ into the UAV's action
      Select action with noise ϵ: $\mathbf{a}_n = \pi_\theta(\mathbf{s}_n|\theta) + \epsilon$, $\epsilon \sim N(0, \sigma)$
      Execute action and estimate the reward r_n
      Update state observation: $\mathbf{s}_n \leftarrow \mathbf{s}_{n+1}$
      Record tuple $(\mathbf{s}_n, \mathbf{a}_n, r_n, \mathbf{s}_{n+1})$ into $\mathcal{B}$
      Randomly extract training batch of K transition $(\mathbf{s}_k, \mathbf{a}_k, r_k, \mathbf{s}_{k+1})$ from $\mathcal{B}$
      $\tilde{\mathbf{a}} \leftarrow \pi(\mathbf{s}_{k+1}|\theta') + \epsilon$, $\epsilon = \text{clip}(N(0, \sigma), [-c, c])$
      $y_k \leftarrow r_k + \gamma \cdot \min\{Q(\mathbf{s}_{k+1}, \tilde{\mathbf{a}}|\omega_1'), Q(\mathbf{s}_{k+1}, \tilde{\mathbf{a}}|\omega_2')\}$
      Update the actor and critic network of the UAV
      Update target networks:
      $\omega_i' \leftarrow \xi\omega_i + (1-\xi)\omega_i'$,
      $\theta' \leftarrow \xi\theta + (1-\xi)\theta'$
  end for
end for
return Optimal trajectory parameters (k^*, ζ^*), or (a^*, b^*)

TABLE 1. SIMULATION PARAMETERS

Name of variables	*Value*
Parameters of Network	
τ : Total Flight Time	1000 s
ρ : Cell Radius	4 km
H: UAV altitude	100 m
P_T: UAV Transmission Power	10 dBm
B: Bandwidth	1 MHz
β_0: Reference channel power	-50 dB
σ^2: Power Spectrum Density of the Noise	-170 dBm
γ_0: Reference Received SNR	70 dB
γ: Received SNR Threshold	0.1 dB
Parameters of UAV	
m: The UAV total mass	10 kg
c_1: Constant Coefficient	0.00092
c_2: Constant Coefficient	2250
Parameters of TD3 Algorithm	
N_{episodes}: Number of episodes	3000
$\mathcal{B}$: Replay buffer	10^6
K: Batch size	256
α: Actor's learning rate	0.0003
β: Critic's learning rate	0.0003
σ: Standard deviation of noise	0.1
$\mathcal{T}$: Target network update period	0.005
γ: Discount factor	0.99

V. SIMULATION RESULTS

In this part, we propose the simulation results. In Table 1, we have listed the simulation parameters that apply to all considered scenarios. We consider $N = 5$ UAVs with similar specs and they are scheduled to move on the proposed paths. An important advantage of the proposed paths is that despite using multiple UAV's, it can be mathematically proven that the chance of collision is zero, i.e., the UAV's do not cross-path a point at the same time. Please note that throughout this section, instead of mechanical energy, we focus on average power which is equal to the total energy divided by the flying period.

It is important to note that while the proposed system model is quite general, the reported results here are specific to the considered system parameters in Table 1. As an example, by changing the received SNR threshold to a higher value, we may need to increase the number of UAVs, N, to achieve the same coverage results. This also holds true for the T_{max} threshold. To satisfy lower values of this threshold, the required power may become extremely large.

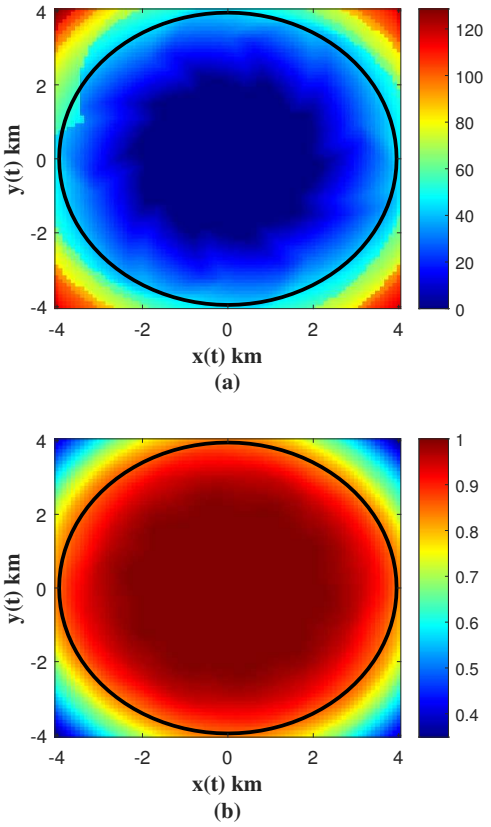

FIGURE 1. **(a) T_{max} heat map, and (b) coverage probability heat map for the spiral path with minimum energy consumption (Problem P1).**

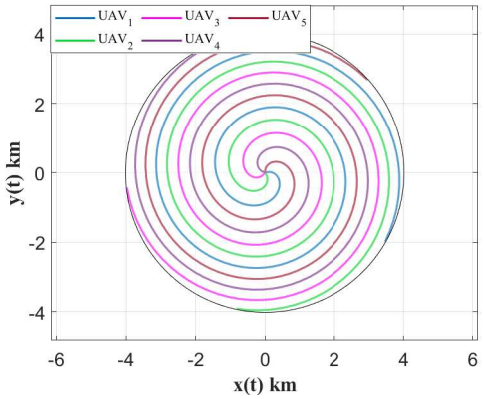

FIGURE 2. **The optimal spiral path with minimum energy consumption (Problem P1).**

A. Performance Comparison: Energy Minimization vs. Maximum Outage Time Minimization

In this part, we propose the results corresponding to problems P1, P2, P3.a and P3.b. Table 2 shows the optimized values corresponding to P1 whereas Fig. 1 shows the heat map for T_{max} and coverage. Fig. 2 shows the corresponding path for 5 UAVs. As can be seen, we have been able to provide coverage probability above 0.9 for all points on the cell, i.e., we have practically avoided any coverage holes.

This article has been accepted for publication in IEEE Open Journal of the Communications Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCOMS.2025.3564837

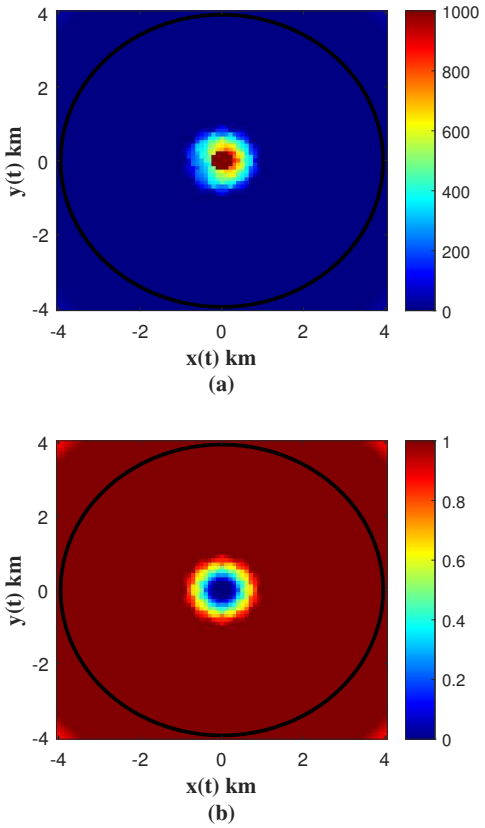

FIGURE 3. (a) T_{max} heat map, and (b) coverage probability heat map for the oval path with minimum energy consumption, (Problem P2) - Curve 1.

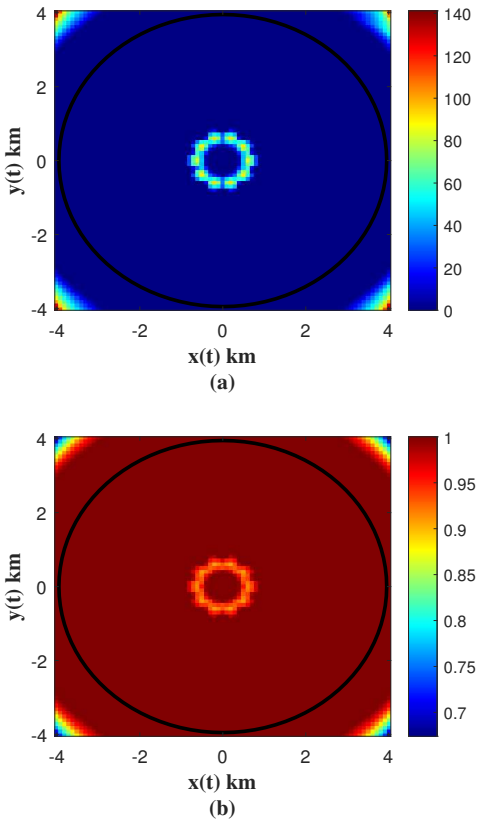

FIGURE 5. (a) T_{max} heat map, and (b) coverage probability heat map for the oval path with minimum energy consumption, (Problem P2) - Curve 2

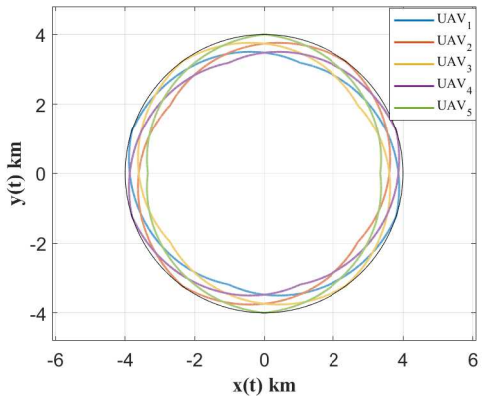

FIGURE 4. The oval path with minimum energy consumption (Problem P2) - Curve1.

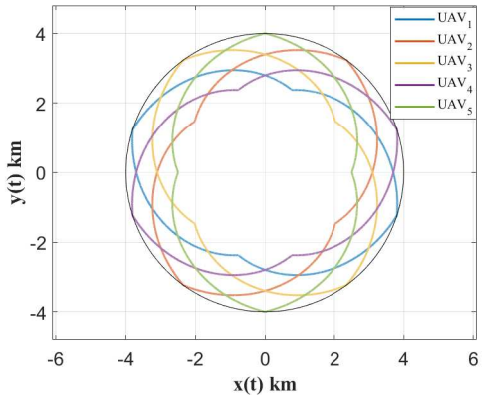

FIGURE 6. The oval path with minimum energy consumption (Problem P2) - Curve 2.

However as can be seen in the table, the maximum value for T_{max} is 50 seconds which is not acceptable.

For the oval curves, we have shown the results for 2 cases in Table 3. The 2nd row of the table shows the optimized values that results in minimum energy where the value is very close to that of the spiral case. However, this trajectory performs very poorly in terms of coverage as seen in Fig. 3 in the cell center. The reason is obvious when we look the curves in Fig. 4. By accepting a slightly larger energy, we can address this issue as can be seen in the 3rd row of Table 3 as well as in Figs 5 and 6. where the average power is increased from 104 to 108 but the coverage is comparable to the spiral case. Although the coverage is significantly improved, the maximum value of T_{max} is 93s which is not acceptable.

To address the issue of higher values of T_{max}, we focus on P3.a and P3.b for both oval and spiral curves, respectively. In Table 4.a, we have shown the results corresponding to P3.a when spiral curves are considered. As can be seen, we have been able to reach $T_{max} = 0$ for all points on the cell. This

TABLE 2. Average consumed power and T_{max} (in seconds) for optimal spiral path in problem P1.

Curve	$\overline{P}$ (Watts)	T_{max}	k	ζ	τ
Spiral	98.3	50	0.77	4.2705π	1000

TABLE 3. Average consumed power and T_{max} (in seconds) for optimal oval path in problem P2.

Curve	$\overline{P}$ (Watts)	T_{max}	a	b	τ
Oval 1	104	1000	3351	4000	1000
Oval 2	108	91	2510	4000	1000

is a great achievement but as can be seen, the corresponding mechanical energy has increased more than 60 folds. To get a better insight, we have plotted maximum outage time and coverage probability heat maps for all points of the cell in Fig. 7. In Fig. 8 we have plotted the corresponding spiral curve. Note that we use 5 UAVs here and we should plot 5 spiral curves, but we did so for only 1 for better clarity. The other 4 curves are exactly the same with rotational phases of 72 degrees with respect to each other. In Table 4.b, we have shown the results corresponding to P3.b when oval curves are considered. As can be seen, the value for energy has been greatly improved. However, this is at the expense of increasing T_{max} to 4.3 seconds. We have also plotted the outage time length and coverage probability heat map in Fig. 9 and the corresponding oval curves in Fig. 10.

The results suggest that there is a kind of trade-off between energy minimization and maximum outage time minimization and the oval curves can suggest a balanced compromise in this case. However, we may need a lower value for T_{max} than the one provided by the oval curves. In the next part, we will address this issue.

B. The Trade-off between Energy and T_{max} Minimization

In the previous part, we observed that the oval curves can not provide T_{max} that are arbitrarily close to 0 seconds. As such in this subsection, we focus on P4 and using spiral curves, we try to minimize energy while guaranteeing that T_{max} remains below a given value which is practically appealing. In this work we set 1 second as the acceptable outage time. The results have been reported in Table 5. As can be seen, the power has significantly improved compared to P3.a but still more than the case of oval curves.

TABLE 4. Optimized values of T_{max} (in seconds) in problems P3.a and P3.b.

(a) P3.a

Curve	$\overline{P}$ (Watts)	T_{max}	k	ζ	τ
Spiral	6.49×10^3	0	0.49962	20.912π	1000

(b) P3.b

Curve	$\overline{P}$ (Watts)	T_{max}	a	b	τ
oval	153.1	4.3	689.45	3000.15	1000

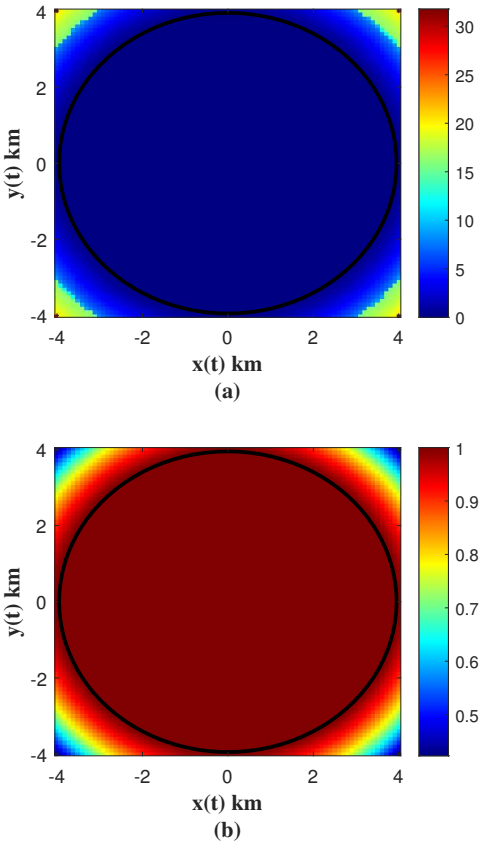

FIGURE 7. (a) T_{max} heat map for the optimal spiral path, and (b) coverage probability heat map for the optimal spiral path, (Problem P3.a).

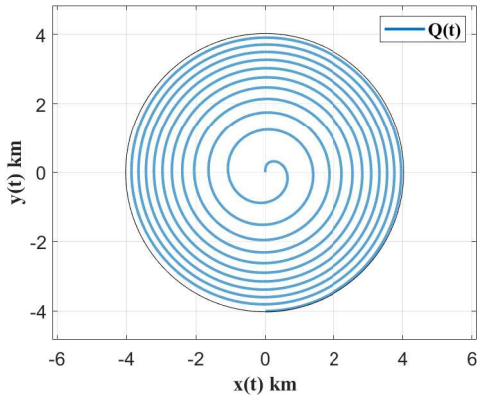

FIGURE 8. The optimal spiral path with minimum T_{max} (Problem P3.a).

The heat map for this case is quite similar to Fig. 7 and is not reported here.[4] The spiral path for this case has been reported in Fig. 11. When we compare Figs 2, 8, and 11, we can see that to minimize energy, a spiral curve with less rotations is optimal. To provide the best coverage, the spiral curve should have so many rotations which causes large

[4]Since at the best case, the oval path cannot provide T_{max} smaller than 1 second, we did not consider a similar scenario to P4 for oval case.

This article has been accepted for publication in IEEE Open Journal of the Communications Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCOMS.2025.3564837

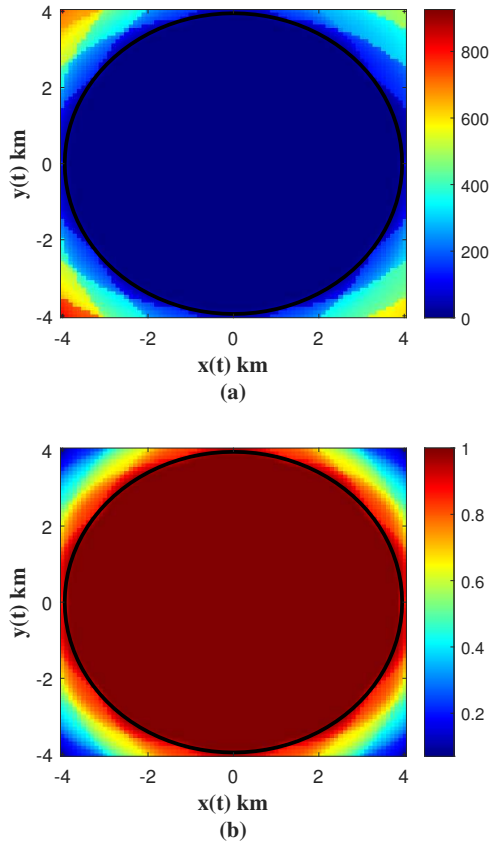

FIGURE 9. **(a) T_{max} heat map for the optimal oval path, and (b) coverage probability heat map for the optimal oval path, (Problem P3.b).**

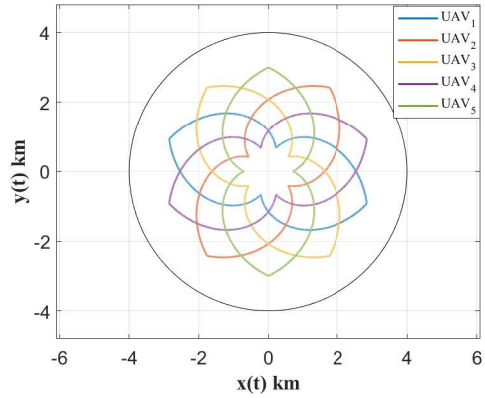

FIGURE 10. **The optimal oval path with minimum T_{max}, (Problem P3.b).**

energy consumption. The number of turns in Fig. 11 is less than Fig. 8 but more than that of Fig. 2.

To better picture the trade-off between energy and T_{max}, we solve P4 for different values of T_{max} and obtain the minimum energy. The results have been shown in Fig. 12. As can be seen, for larger values of T_{max}, we can get average power as low as the one for P1. On the other hand, as T_{max} tends to zero, the average power gets drastically large, and we tend to the value already obtained in P2.

TABLE 5. **Optimized value of consumed power for the optimal spiral path in problem P4 when we set $T_{max} < 1s$.**

Curve	$\overline{P}$ (Watts)	T_{max}	k	ζ	τ
Spiral	781.76	1	0.54	10.90π	1000

This shows a clear trade-off between T_{max} and average consumed power. Using this plot, we can have a control on the amount of T_{max} over the cell and choose the scenario of interest while adhering to the QoS and energy consumption requirements imposed by the system. Moreover, for T_{max} around 4 seconds, spiral path need considerably more power compared to oval paths which suggests that oval paths might be always preferred. However, we cannot always rely on oval paths if we need extremely low values for T_{max}.

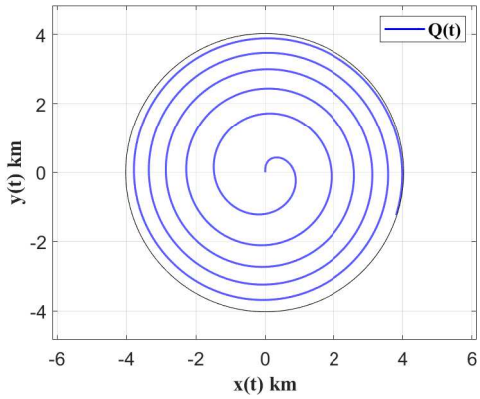

FIGURE 11. **The optimal spiral path with minimum energy and $T_{max} < 1s$, (Problem P4).**

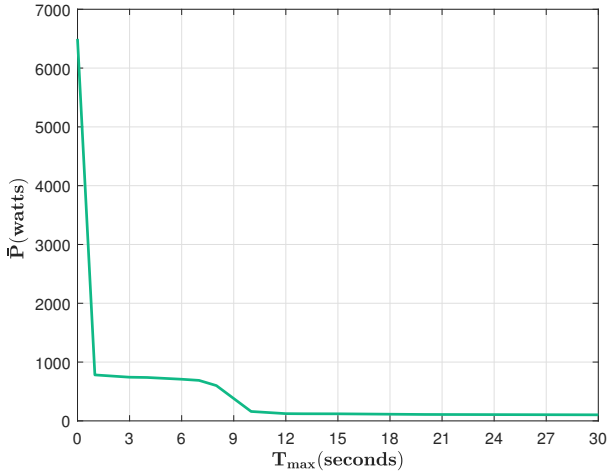

FIGURE 12. **Trade-off between T_{max} and average consumed power.**

VI. CONCLUSIONS

In this paper we addressed the issue of coverage holes and outage time in cellular communications and proposed optimal UAV paths to minimize them, taking into consideration the amount of energy consumption. We first focused on the

coverage holes and showed that we can deploy both spiral and oval trajectories to minimize such holes by guaranteeing that all points on the cell have outage probability less than 10% with average mechanical power consumption as low as 100 watts. We then focused on outage time and aimed to guarantee, for any point on the cell, a maximum time length being in outage. We established a trade-off between energy consumption and outage time minimization. Outage time length of '0' second using spiral paths is at hand at the expense of very high average power consumption. As for the oval curves, we showed that the required power is much less if we can live with non-zero outage time. Finally, for spiral paths, we provided a very informative plot in which we consider different scenarios of average power and T_{max} requirements. It can be seen that for the same value of allowed T_{max} that is feasible for both types of trajectories, oval paths outperform spiral paths in terms of consumed energy. On the other hand, spiral paths should still be deployed if extremely low values for T_{max} are needed, at the expense of more energy consumption. Spectrum allocation strategies can be considered as the extension of this work. For instance, in configurations where multiple UAVs provide overlapping coverage of a target area, an adaptive frequency allocation scheme could be developed based on proximity metrics. When two or more UAVs service the same region with different distances to ground users, the system could dynamically reuse frequency resources to the UAV with superior channel conditions (typically the closer one), thereby maximizing achievable data rates.

REFERENCES

[1] M. Wang, C. Zhang, X. Chen, and S. Tang, "Performance analysis of millimeter wave wireless power transfer with imperfect beam alignment," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 3, pp. 2605–2618, 2021.

[2] Y. Liu, C. She, Y. Zhong, W. Hardjawana, F.-C. Zheng, and B. Vucetic, "Graph neural networks for quality of service improvement in interference-limited ultra-reliable and low-latency communications," *IEEE Transactions on Vehicular Technology*, 2023.

[3] X. Deng, Z. Tang, L. Yi, and L. T. Yang, "Healing multimodal confident information coverage holes in nb-IoT-enabled networks," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1463–1473, 2017.

[4] C. K. Anjinappa, F. Erden, and I. Güvenç, "Base station and passive reflectors placement for urban mmwave networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3525–3539, 2021.

[5] X. Deng, M. Xu, L. T. Yang, M. Lin, L. Yi, and M. Wang, "Energy balanced dispatch of mobile edge nodes for confident information coverage hole repairing in IoT," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4782–4790, 2018.

[6] H.-W. Liang, C.-H. Ho, L.-S. Chen, W.-H. Chung, S.-Y. Yuan, and S.-Y. Kuo, "Coverage hole detection in cellular networks with deterministic propagation model," in *2016 2nd International Conference on Intelligent Green Building and Smart Grid (IGBSG)*. IEEE, 2016, pp. 1–6.

[7] W. A. Hapsari, A. Umesh, M. Iwamura, M. Tomala, B. Gyula, and B. Sebire, "Minimization of drive tests solution in 3gpp," *IEEE Communications Magazine*, vol. 50, no. 6, pp. 28–36, 2012.

[8] I. Akbari, O. Onireti, A. Imran, M. A. Imran, and R. Tafazolli, "How reliable is mdt-based autonomous coverage estimation in the presence of user and bs positioning error?" *IEEE Wireless Communications Letters*, vol. 5, no. 2, pp. 196–199, 2016.

[9] C. K. Anjinappa and İ. Güvenç, "Coverage hole detection for mmwave networks: An unsupervised learning approach," *IEEE Communications Letters*, vol. 25, no. 11, pp. 3580–3584, 2021.

[10] Y. Xia, X. Deng, L. Yi, L. T. Yang, X. Tang, C. Zhu, and Z. Tian, "Ai-driven and mec-empowered confident information coverage hole recovery in 6g-enabled IoT," *IEEE Transactions on Network Science and Engineering*, vol. 10, no. 3, pp. 1256–1269, 2022.

[11] X. Deng, B. Wang, W. Liu, and L. T. Yang, "Sensor scheduling for multi-modal confident information coverage in sensor networks," *IEEE Transactions on parallel and distributed systems*, vol. 26, no. 3, pp. 902–913, 2014.

[12] X. Deng, Y. Jiang, L. T. Yang, L. Yi, J. Chen, Y. Liu, and X. Li, "Learning-automata-based confident information coverage barriers for smart ocean internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9919–9929, 2020.

[13] B. Jafari, H. Saeedi, and H. Pishro-Nik, "UAV path planning for surveillance applications: Rotary-wing vs. fixed-wing UAVs," in *2024 IEEE 98th Vehicular Technology Conference (VTC2024-Spring)*. IEEE, 2024, pp. 1–5.

[14] Y. Pan, R. Li, X. Da, H. Hu, M. Zhang, D. Zhai, K. Cumanan, and O. A. Dobre, "Cooperative trajectory planning and resource allocation for uav-enabled integrated sensing and communication systems," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 5, pp. 6502–6516, 2024.

[15] A. Shamsoshoara, F. Afghah, E. Blasch, J. Ashdown, and M. Bennis, "Uav-assisted communication in remote disaster areas using imitation learning," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 738–753, 2021.

[16] B. Jafari, H. Saeedi, S. Enayati, and H. Pishro-Nik, "Energy-optimized path planning for moving aerial base stations: A non user-oriented framework," *IEEE Communications Letters*, vol. 26, no. 3, pp. 672–676, 2021.

[17] ——, "Path planning for unmanned aerial vehicles: Peak power minimization," in *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*. IEEE, 2023, pp. 1–5.

[18] N. Gupta, S. Agarwal, D. Mishra, and B. Kumbhani, "Trajectory and resource allocation for uav replacement to provide uninterrupted service," *IEEE Transactions on Communications*, 2023.

[19] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 3747–3760, 2017.

[20] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE communications surveys & tutorials*, vol. 21, no. 3, pp. 2334–2360, 2019.

[21] S. Sai, A. Garg, K. Jhawar, V. Chamola, and B. Sikdar, "A comprehensive survey on artificial intelligence for unmanned aerial vehicles," *IEEE Open Journal of Vehicular Technology*, vol. 4, pp. 713–738, 2023.

[22] Y. Sun, Z. Ding, and X. Dai, "A user-centric cooperative scheme for UAV-assisted wireless networks in malfunction areas," *IEEE Transactions on Communications*, vol. 67, no. 12, pp. 8786–8800, 2019.

[23] X. Guo, C. Zhang, F. Yu, and H. Chen, "Coverage analysis for UAV-assisted mmwave cellular networks using poisson hole process," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 3, pp. 3171–3186, 2021.

[24] S. A. Al-Ahmed, M. Z. Shakir, and S. A. R. Zaidi, "Optimal 3d UAV base station placement by considering autonomous coverage hole detection, wireless backhaul and user demand," *Journal of Communications and Networks*, vol. 22, no. 6, pp. 467–475, 2020.

[25] M. T. Dabiri, M. Hasna, N. Zorba, and T. Khattab, "Optimal trajectory and positioning of UAVs for small cell hetnets: Geometrical analysis and reinforcement learning approach," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 2667–2683, 2023.

[26] S. Enayati, H. Saeedi, H. Pishro-Nik, and H. Yanikomeroglu, "Moving aerial base station networks: A stochastic geometry analysis and design perspective," *IEEE Transactions on Wireless Communications*, vol. 18, no. 6, pp. 2977–2988, June 2019.

[27] R. Imran, M. Odeh, N. Zorba, and C. Verikoukis, "Quality of experience for spatial cognitive systems within multiple antenna scenarios," *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 4153–4161, 2013.

[28] 3GPP, "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Policy and charging control architecture ," 3rd Generation Partnership Project

(3GPP), Technical Specification (TS) 23.203, 07 2018, version 15.3.0 Release 15.

[29] ——, "5G; System architecture for the 5G System (5GS)," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 23.501, 10 2020, version 16.6.0 Release 16.

[30] International Telecommunication Union (ITU), "Framework and overall objectives of the future development of IMT for 2030 and beyond," International Telecommunication Union, Tech. Rep. Recommendation ITU-R M.2160, 2023. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2160-0-202311-I%21%21PDF-E.pdf

[31] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," in *International conference on machine learning*. Pmlr, 2014, pp. 387–395.

[32] S. Fujimoto, H. Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *International conference on machine learning*. PMLR, 2018, pp. 1587–1596.