RAPID: Reinforcement Learning-Aided Femtocell Placement for Indoor Drone Localization

Alireza Famili*, Amin Tabrizian†, Tolga Atalay‡, Angelos Stavrou*‡, Peng Wei†
*WayWave Inc, Virginia, USA
†Department of Computer Science, George Washington University, Washington, D.C., USA
‡Department of Electrical & Computer Engineering, Virginia Tech, Virginia, USA
afamili@waywave.com, amin_tabrizian@gwu.edu, tolgaoa@vt.edu, angelos@vt.edu, pwei@gwu.edu

***Abstract*—Mobile networks are swiftly advancing to accommodate the burgeoning spectrum of applications. The architecture of 5G networks integrates the principle of network slices, logically isolated end-to-end segments tailored to offer specific services. In this architectural schema, drones have emerged as a significant service category. Achieving the successful deployment of drone networks is heavily contingent upon the ability to accurately localize them in a three-dimensional (3D) setting, beyond the critical requirement for tight latency control. Transitioning from 4G to 5G, these networks are characterized by their operation at elevated frequency spectrums and more densely packed deployment configurations. Within such environments, the task of ensuring precise indoor localization poses a significant challenge, primarily due to the distinctive signal behavior at higher frequencies. To achieve this goal, we propose the *RAPID* framework, utilizing foundational principles from the third-generation partnership project (3GPP) to design a radio access network (RAN) that includes 5G femtocells. This architecture aims to shift positioning responsibilities from outdoor base stations (BSs) to improve indoor localization performance. Our study's principal contribution is the demonstration of how the spatial distribution of 5G femtocells significantly influences the accuracy of drone positioning. To address the challenges inherent in femtocell deployment, we develop an innovative optimization framework coupled with a deep reinforcement learning (DRL) strategy, aimed at solving the NP-hard problem. Our findings reveal that adopting our DRL-based placement strategy significantly improves positioning accuracy compared to regular arbitrary deployment approaches.**

***Index Terms*—indoor positioning, autonomous vehicles, drone, 5G femtocell, GDOP, optimal placement, RL**

I. Introduction

The next generation of mobile networks is set to provide services to a wide array of sectors, each with unique quality of service (QoS) demands. To meet these demands, these next generation networks will employ a heterogeneous radio access network (RAN) architecture. This RAN integrates multiple radio access technologies (RATs) defined by the 3rd Generation Partnership Project (3GPP), facilitating wireless connectivity across diverse environments.

In this ecosystem, the communication between autonomous vehicles (AVs) assumes critical importance. A prime example of this is the interaction between unmanned aerial vehicles (UAVs), commonly known as drones, and their ground controllers, which has surfaced as a significant use case drawing substantial interest. Leading corporations, such as Amazon and Google, are seeking to create a navigation network that encompasses both indoor and outdoor settings, alongside urban and rural territories. Within this framework, the attainment of precise localization emerges as a fundamental necessity [1].

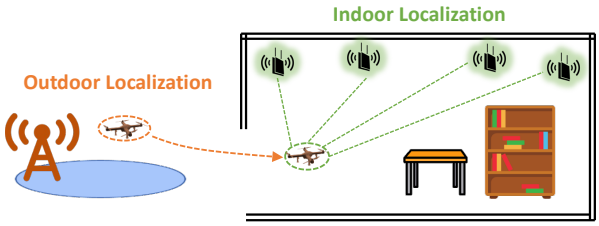


Figure 1: Overview of the *RAPID* framework, showing the shift of localization responsibilities from the 5G base stations to the 5G femtocells, facilitating a seamless transition from outdoor to indoor environments

The Global Positioning System (GPS) is widely acknowledged for providing accurate location information in outdoor environments. Yet, its effectiveness diminishes significantly within indoor spaces, often resulting in service disruptions. In traditional Long-Term Evolution (LTE) networks, the positioning accuracy reported falls substantially short of that offered by GPS. Nonetheless, both cellular and GPS technologies face obstacles in maintaining coverage and accuracy indoors.

The widespread implementation of diverse 3GPP RAN architectures in 5G has enabled uninterrupted communication for 5G-compatible UAVs, ensuring reliable coverage and remarkable data transfer rates in outdoor scenarios. However, achieving a comprehensive solution that offers precise and dependable localization for drones in both indoor and outdoor settings continues to be a formidable challenge. To address this and enhance indoor coverage, we advocate for the adoption of 5G femtocells for indoor localization, as shown in Figure 1. This strategy ensures that drones experience consistent and reliable positioning throughout their interaction.

The accuracy of indoor positioning is influenced by both the ranging error and errors caused by relative spatial geometric configurations. For example, when the ranging error is kept to 10 cm or less, a target achievable with the extensive bandwidth of 5G mm-Wave technology, the geometric dilution of precision (GDOP) at a given location is found to be 20. This results in an ultimate localization error of approximately 2 m, calculated as $10\ cm \times 20 = 200\ cm$. This twentyfold increase in error, from

10 cm to 2 m, is deemed unsuitable for applications requiring high precision in positioning, such as the intricate localization needs of drones.

Our goal is to examine the impact of GDOP and to devise optimal configurations for positioning beacons, which are essentially 5G femtocells. Whereas existing studies [2]–[4] have largely focused on mitigating ranging errors to improve localization accuracy, our methodology includes an analysis of the spatial geometric configuration of the system, a crucial element overlooked in previous studies.

To that end, we propose ***RAPID:*** ***R****einforcement* ***L****earning-****A****ided Femtocell* ***P****lacement for* ***I****ndoor* ***D****rone Localization*, a framework for high-accuracy localization of indoor drones leveraging 5G femtocells that are configured via deep reinforcement learning (DRL) approaches. We compute the Cramer-Rao lower bound (CRLB) for the position estimator. This calculation reveals that the localization error in round-trip time (RTT) trilateration methods stems from both ranging errors and the spatial location of the drone relative to the positioning nodes. Our primary objective is to offer femtocell configurations for any indoor setting that minimize the latter source of error to facilitate precise localization. Towards this goal, we introduce an innovative optimization formulation and employ a DRL strategy to identify desired 5G femtocell configurations.

DRL has played a significant role in addressing various sequential decision-making and control challenges, such as gaming, robotic manipulation, optimizing chemical reactions, and aviation [5], [6]. The majority of conventional DRL research focuses on creating agents capable of learning to tackle sequential decision problems arising from the intrinsic dynamics of a task, like the differential equations governing the cart-pole task in the classic control suite [6].

In the case of combinatorial optimization problems, numerous conventional approaches rely on employing manually devised heuristics, which systematically build a solution. These heuristics are typically crafted by domain experts but can frequently be suboptimal due to the inherent complexity of the problems. The majority of the current literature for the optimal placement of positioning nodes is based on heuristic approaches [7]. On the other hand, DRL offers a promising alternative for automating the exploration of these heuristics by training an agent through supervised or self-supervised methods [8]. To that end, *RAPID* proposes a novel approach to find the optimal placement of 5G femtocells using an innovative DRL-based method, outperforming the existing literature.

Our experimental validation indicates that, even under conditions of minimal ranging errors, suboptimal GDOP can markedly compromise the final localization accuracy of drones in a three-dimensional (3D) environment. By utilizing coverage heatmaps, we demonstrate that GDOP arises from both vertical and horizontal dilutions of precision, known as VDOP and HDOP, respectively; with vertical dilution having a more pronounced impact on the overall reduction of estimation accuracy than horizontal dilution.

Our contributions can be succinctly outlined as follows:

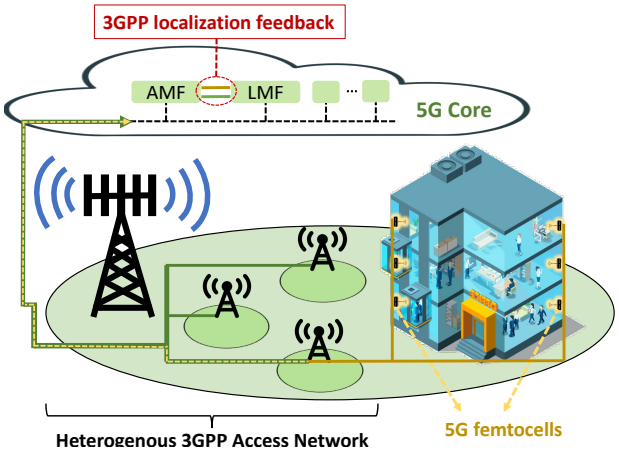


Figure 2: An overview of a combined RAT 5G network offering both indoor and outdoor coverage with a unified core network

- The introduction of an integrated localization system that harmonizes outdoor 5G base stations (BSs) with indoor 5G femtocells, all cohesively linked to a central localization unit within the 5G core network.
- The derivation of the positioning error bound (PEB) for an RTT-based trilateration system, attributing the final error to both ranging inaccuracies and geometry-induced errors.
- The investigation into the effects of VDOP and HDOP, revealing that VDOP poses a greater source of error compared to HDOP on the overall localization accuracy.
- The proposal of a principled DRL-based learning approach to solve the NP-hard problem of 5G femtocell placement in indoor settings.
- The comparison of the proposed DRL-based approach with random placement patterns to demonstrate the superior GDOP of *RAPID*. According to our results, *RAPID* yields a significant decrease in the GDOP error, thereby significantly improving the overall localization estimation.

II. *RAPID* Core Network Model

Within a 5G network, maintaining continuous tracking of localization data is crucial, especially in situations where User Equipment (UE) has high mobility. To address this, the model in Figure 2 aims to create a centralized core network repository designed to keep accurate tabs on the location information of specific UEs across their entire connectivity span within a diverse 5G infrastructure, encompassing both outdoor and indoor nodes. This strategy ensures that regardless of the UE's mobility or the complexity of the network environment, its location can be accurately and consistently monitored, facilitating seamless service delivery and enhanced user experiences.

The scenario in Figure 2 illustrates an integrated setup encompassing both macro and micro BSs, specifically tailored for outdoor coverage using the 5G New Radio (NR) technology, as standardized by the 3GPP. In parallel, indoor connectivity is secured through the deployment of 5G femtocells. A pivotal element of this system is the centralized Location Management Function (LMF) within the core network, which is tasked with

the integration and processing of localization data originating from both the outdoor 5G NR infrastructure and the indoor femtocells. This unified approach to location management ensures that there is a cohesive and accurate tracking mechanism in place, capable of localizing a drone as it moves between the indoor and outdoor RAN. This architecture not only facilitates seamless connectivity across different environments but also significantly enhances the efficiency and reliability of services that depend on precise location information, such as navigation and logistics applications.

In this setup, a drone moves between outdoor and indoor coverage, keeping its location accurate indoors with femtocells and outdoors with BSs. This approach aligns with ongoing standards, allowing the Location Management Function (LMF) API to remain unchanged for streamlined operation.

In the following sections, we explore the challenges related to indoor localization accuracy. We base our analysis on the premise of a consistent outdoor cellular environment, predominantly utilizing 5G NR nodes for positioning purposes.

III. *RAPID* Positioning

This section begins by outlining the essential prerequisites for localization. It then proceeds to describe how *RAPID* facilitates real-time, precise location tracking for drones via RTT-based trilateration. The final part provides the derivation of the positioning error bound (PEB) for *RAPID* and highlights the influence of GDOP on the overall localization error.

A. Foundational Prerequisites for Localization

Recent studies have explored the anchor placement optimization to enhance indoor localization [9]–[13]. This research primarily focuses on determining the ideal number of anchors needed for comprehensive indoor coverage and optimizing their positions to minimize errors due to spatial relationships. The choice of sensors significantly affects coverage due to their varying capabilities. For instance, systems employing low-power Bluetooth sensors offer omnidirectional transmission limited by range and physical barriers. In contrast, systems utilizing ultrasound sensors face restrictions based on the sensor's beam angle, influencing the sensor count and placement strategy [14]. Once the requisite number of anchors is established, an optimization algorithm can be applied to ascertain their optimal locations, thereby reducing localization errors stemming from geometric considerations.

Our main focus is on localization methods that utilize ranging-based strategies [15]. These techniques involve using transmit/receive signals between a target and one or more positioning nodes to frame the localization issue.

The signals employed in these localization techniques may either be acoustic [4] or Radio Frequency (RF) [16], contingent upon the type of transceivers used. Acoustic (or ultrasound) signals, due to their slower propagation speed, are capable of delivering high-precision location estimates [12]. While they excel in closely-spaced, high-accuracy gesture tracking scenarios [15], their effectiveness diminishes over longer distances due to limited coverage, rendering them less appropriate for extensive range indoor/outdoor localization tasks.

Beyond signal type, the efficacy of a localization system hinges on measurement techniques such as Received Signal Strength (RSS), Time of Arrival (TOA), or Angle of Arrival (AOA). RSS, commonly paired with environmental fingerprinting, offers a cost-effective and straightforward implementation. However, localization strategies relying on RSS fingerprinting are notably susceptible to dynamic environmental variations, compromising their reliability.

AOA estimation, employing angulation for localization, necessitates specialized antenna arrays [17]. This approach demands significant computational resources, requiring additional processing power allocation. The complexity arises from the use of sophisticated algorithms, such as Multiple Signal Classification (MUSIC), essential for accurate angle estimation.

In contrast to RSS and AOA, Time of Arrival (TOA) offers a simpler and more dependable measurement alternative, requiring neither high computational resources nor specialized antenna setups [18]. TOA estimates are derived by converting time measurements into distances using $d = c \times t$, where d represents the distance between the transmitter and receiver, c is the propagation speed of the signal (the speed of light for RF signals), and t is the duration RF waves take to reach the receiver from the transmitter. Utilizing distance data from several anchor nodes to the target allows for the application of trilateration methods to determine the precise location.

B. RTT-based Trilateration

In 5G localization, two primary techniques are commonly employed: Time Difference of Arrival (TDOA) and RTT-based methods. While both techniques alleviate the stringent synchronization requirements between the UE and positioning nodes, research presented in [19] demonstrates that RTT-based methods offer superior ranging accuracy, as validated through CRLB derivations. In our setup, we utilize TOA measurements derived from RTT for acquiring distance measurements between 5G femtocells and drones, enabling 3D localization. Trilateration necessitates a minimum of three anchors for two-dimensional (2D) localization and at least four for 3D localization.

Consider the distance between the drone and the i-th 5G femtocell, denoted as d_i. Let the drone's position be represented by $[x\ y\ z]^T$, and the position of the i-th 5G femtocell be denoted as $[x_i\ y_i\ z_i]^T$. The trilateration problem can be formulated as $(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2 = d_i^2$, where $i \in \{1, \cdots, n\}$ for a setup with n number of 5G femtocells. With further mathematical manipulation, we can rewrite the problem as $\boldsymbol{\Phi\theta} = \boldsymbol{\delta}$, where $\boldsymbol{\Phi}$ and $\boldsymbol{\delta}$ are equal to:

$$\boldsymbol{\Phi} = \begin{bmatrix} 2(x_n - x_1) & 2(y_n - y_1) & 2(z_n - z_1) \\ \vdots & \vdots & \vdots \\ 2(x_n - x_{n-1}) & 2(y_n - y_{n-1}) & 2(z_n - z_{n-1}) \end{bmatrix},$$

$$\boldsymbol{\delta} = \begin{bmatrix} d_1^2 - d_n^2 - x_1^2 - y_1^2 - z_1^2 + x_n^2 + y_n^2 + z_n^2 \\ \vdots \\ d_{n-1}^2 - d_n^2 - x_{n-1}^2 - y_{n-1}^2 - z_{n-1}^2 + x_n^2 + y_n^2 + z_n^2 \end{bmatrix}.$$

The vector $\boldsymbol{\theta} = [x\ y\ z]^T$, containing the drone's coordinates, is expressed as: $\boldsymbol{\theta} = (\boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^T\boldsymbol{\delta}$.

C. Positioning Error Bound

A valuable metric for evaluating localization accuracy is the CRLB, representing the minimum achievable variance in location estimation using an unbiased estimator. Assuming independent range measurements with zero-mean additive Gaussian noise and constant variance σ_r^2 [9], we can demonstrate that for a 2D trilateration system with an unbiased estimator, the CRLB variance of the positional error $\sigma^2(r) = \sigma_x^2(r) + \sigma_y^2(r)$ at position r, is given by:

$$\sigma(r) = \sigma_r \times \sqrt{\frac{N_F}{\sum_{i=1}^{N_F-1}\sum_{j=i+1}^{N_F} F_{ij}}},$$

where N_F is the number of 5G femtocells, θ_i is the angle between F_i and r, F_i is the i-th 5G femtocell, and $F_{ij} = |\sin(\theta_i - \theta_j)|$.

Hence, we can infer that the localization error is proportional to the product of the ranging measurement inaccuracies and a function involving the number of 5G femtocells and the angle between them and the drone. This function is known as GDOP, defined as $GDOP = \frac{\sigma(r)}{\sigma_r}$.

As CRLB is directly linked to GDOP, we can regard GDOP as a reliable metric for assessing localization accuracy [13], [20].

For the 3D localization of an object at coordinates (x, y, z) utilizing 5G femtocells, the GDOP equation is given by:

$$GDOP = \frac{\sqrt{Var(x) + Var(y) + Var(z) + Var(c\tau)}}{\sigma_r},$$

here, c represents the speed of light, and τ denotes the clock offset of the receiver. We make the assumption that both the transmitter and receiver share the same clock, thereby setting the timing offset to zero. This simplification yields:

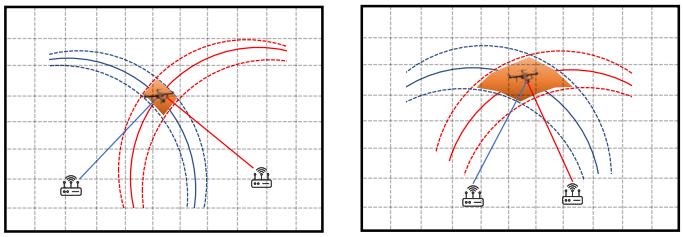
$$GDOP = \sqrt{\frac{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}{\sigma_r^2}}. \tag{1}$$

Figure 3 illustrates the visual representation of GDOP for two distinct scenarios. In the first case, depicted in Figure 3a, the beacons are positioned more optimally compared to the configuration shown in Figure 3b. Consequently, this disparity in placement leads to varying location estimation errors, as indicated by the respective shaded regions.

Let the position of the drone be denoted as (x, y, z) and the position of each 5G femtocell as (x_i, y_i, z_i). The distance between them is given by:

$$r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}. \tag{2}$$

Due to the ranging measurement error, the precise value of r_i is unknown, leading to inaccuracies when solving for (x, y, z) in Eq.2. To establish a relationship between the solution error and the actual ranging errors between the drone and the anchors, we differentiate Eq.2 with respect to the variables, neglecting



(a) Beacon deployment with low localization error

(b) Beacon deployment with high localization error

Figure 3: Assessment of GDOP's impact on localization error resulting from varying beacon deployments

terms beyond the first order, as Massatt has done [21]. This yields:

$$\begin{aligned}\Delta r_i &= \frac{\Delta x(x - x_i) + \Delta y(y - y_i) + \Delta z(z - z_i)}{\sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}} \\ &= \Delta x \cos\alpha_i + \Delta y \cos\beta_i + \Delta z \cos\gamma_i\end{aligned}$$

where $[\cos\alpha_i\ \cos\beta_i\ \cos\gamma_i]^T$ represents the unit vector pointing from the drone to the i-th 5G femtocell.

Let's denote $\boldsymbol{\Theta} = [\Delta x\ \Delta y\ \Delta z]^T$ as the position error vector and $\boldsymbol{\Psi} = [\Delta r_1 \cdots \Delta r_n]^T$ as the target range error vector. Then, we can express the matrix $\boldsymbol{\Upsilon}$ as follows:

$$\boldsymbol{\Upsilon} = \begin{bmatrix} \upsilon_1^1 & \upsilon_2^1 & \upsilon_3^1 \\ \vdots & \vdots & \vdots \\ \upsilon_1^n & \upsilon_2^n & \upsilon_3^n \end{bmatrix}$$

where $[\upsilon_1^i\ \upsilon_2^i\ \upsilon_3^i] = [\cos\alpha_i \cos\beta_i \cos\gamma_i]$. With this, we can express $\boldsymbol{\Psi} = \boldsymbol{\Upsilon}\boldsymbol{\Theta}$, leading to $\boldsymbol{\Theta} = (\boldsymbol{\Upsilon}^T\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}^T\boldsymbol{\Psi}$. Given that:

$$\mathbf{Cov}(\boldsymbol{\Theta}) = \mathbb{E}(\boldsymbol{\Theta}\boldsymbol{\Theta}^T) = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_z^2 \end{bmatrix}. \tag{3}$$

We assume Var(r_i) = σ_r^2 and that the errors Δr_i are uncorrelated, then:

$$\begin{aligned}\mathbb{E}(\boldsymbol{\Theta}\boldsymbol{\Theta}^T) &= \mathbb{E}(((\boldsymbol{\Upsilon}^T\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}^T\boldsymbol{\Psi})((\boldsymbol{\Upsilon}^T\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}^T\boldsymbol{\Psi})^T) \\ &= (\boldsymbol{\Upsilon}^T\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}^T\mathbb{E}(\boldsymbol{\Psi}\boldsymbol{\Psi}^T)((\boldsymbol{\Upsilon}^T\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}^T)^T \\ &= (\boldsymbol{\Upsilon}^T\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}^T\boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}\boldsymbol{\Upsilon}^T)^{-1}\sigma_r^2 = (\boldsymbol{\Upsilon}^T\boldsymbol{\Upsilon})^{-1}\sigma_r^2\end{aligned}$$

By combining this outcome with Eq. 1 and Eq. 3, we demonstrate that the diagonal elements of $(\boldsymbol{\Upsilon}^T\boldsymbol{\Upsilon})^{-1}$ are utilized in computing the GDOP, given by $GDOP = \sqrt{HDOP^2 + VDOP^2}$. Here, HDOP characterizes the impact of relative geometry on the $X - Y$ plane, while VDOP represents the influence of geometry on the estimation along the Z-axis. Table I provides an assessment of the GDOP values.

It is crucial to emphasize our aim not only for a favorable GDOP but also for maintaining accurate HDOP and VDOP values concurrently. A scenario where HDOP is good while VDOP is poor may yield precise estimation in the $X-Y$ plane (i.e., effective 2D localization) but may lead to inaccurate estimations along the Z-axis (i.e., inadequate height estimation).

Table I: Evaluation of GDOP Values

GDOP Values	Evaluation of Femtocells Placement
1	Ideal
$1-2$	Very Good
$2-5$	Good
$5-10$	Medium
$10-20$	Sufficient
>20	Bad

Much of the existing literature predominantly concentrates on 2D localization [9], [13], often overlooking the importance of geometry and optimal anchor node placement. However, when dealing with drones in a three-dimensional setting, ensuring accurate estimations along the Z-axis becomes equally crucial as those in the $X-Y$ plane. Thus, we have underscored the localization error sources in this section and introduced concepts such as GDOP, HDOP, and VDOP to address these considerations.

IV. *RAPID* Problem Formulation for Femtocell Deployment

Previously, localization techniques were primarily designed for static targets in 2D scenarios, overlooking the real-world geometric relationship between the target and anchors in three dimensions. Consequently, determining the optimal placement for a mobile target in a 3D space remains an unresolved challenge [13]. Additionally, identifying the anchor placement configuration for indoor localization to minimize the relative geometric error between transmitters and the user at any position is a recognized NP-hard problem [9]–[11], [13].

In this section, we introduce an optimization problem formulation aimed at determining the optimal placement of 5G femtocells to mitigate localization errors stemming from unfavorable relative geometry. Subsequently, in the next section, we introduce *RAPID's* DRL-based technique and elucidate the mechanism behind our framework to solve the problem.

The proposed algorithm must ensure high estimation accuracy for both the $X-Y$ plane (horizontal) and the Z-axis (vertical). It is worth noting that the Z-axis is particularly susceptible to geometry-induced errors. Even if the overall GDOP for a specific point is satisfactory, there is no assurance that the VDOP matches the quality of the HDOP for that point. This often leads to highly accurate estimations in the $X-Y$ plane but significantly erroneous estimations along the Z-axis. To the best of our knowledge, we are the first to propose a DRL-based algorithm aimed at enhancing estimation accuracy in both the $X-Y$ plane and the Z-axis.

The drone domain, denoted as set $\mathbb{D}$, represents a subspace within the indoor environment encompassing all potential trajectories of the drone. Optimization computations are conducted across all points within this domain to determine VDOP and HDOP values for all the points. The 5G femtocells domain, designated as set $\mathbb{F}$, constitutes the permissible locations for deploying the 5G femtocells, encompassing the entire ceiling. A tolerance (h_{tol}) for HDOP is enforced as a constraint to ensure that HDOP remains below a certain threshold for every single point in set $\mathbb{D}$, while a tolerance (v_{tol}) for VDOP is imposed to mandate that VDOP is also smaller than its required threshold.

Based on our findings in Sec. III-C, if each measurement is uncorrelated and possesses identical uncertainty with zero mean and unit variance, then HDOP and VDOP can be inferred from the diagonal elements of the matrix $\mathbf{\Omega}$ as illustrated below:

$$\mathbf{\Omega} = (\mathbf{\Upsilon}^T\mathbf{\Upsilon})^{-1} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_z^2 \end{bmatrix},$$

where $VDOP = \sqrt{\sigma_z^2}$ and $HDOP = \sqrt{\sigma_x^2 + \sigma_y^2}$.

$$\mathbf{\Upsilon} = \begin{bmatrix} \frac{x_1-x}{r_1} & \frac{y_1-y}{r_1} & \frac{z_1-z}{r_1} \\ \frac{x_2-x}{r_2} & \frac{y_2-y}{r_2} & \frac{z_2-z}{r_2} \\ \frac{x_3-x}{r_3} & \frac{y_3-y}{r_3} & \frac{z_3-z}{r_3} \\ \frac{x_4-x}{r_4} & \frac{y_4-y}{r_4} & \frac{z_4-z}{r_4} \end{bmatrix}.$$

The coordinates of the drone and the i-th 5G femtocell are represented by (x, y, z) and (x_i, y_i, z_i), respectively, where r_i denotes the distance between them.

Our primary objective is to determine the optimal placement for a group of four 5G femtocells, aiming to minimize GDOP for each point in set $\mathbb{D}$ while ensuring that HDOP and VDOP associated with that point remain below specified thresholds.

The optimization formulation can be expressed as follows:

$$\begin{aligned} \min \quad & \textstyle\sum_{\mathbb{D}} \textbf{Trace}\{(\mathbf{\Upsilon}^T\mathbf{\Upsilon})^{-1}\} \\ s.t. \quad & \forall (x,y,z) \in \mathbb{D}: \ HDOP(x,y,z) < h_{tol}; \\ s.t. \quad & \forall (x,y,z) \in \mathbb{D}: \ VDOP(x,y,z) < v_{tol}. \end{aligned} \tag{4}$$

The objective is to minimize $GDOP(x, y, z)$ for any point in the drone domain $(\forall (x, y, z) \in \mathbb{D})$ to optimize the placement of 5G femtocells, thereby enhancing localization accuracy. Simultaneously, to uphold low errors in 2D localization and Z-axis estimation, we incorporate constraints for $HDOP(x, y, z)$ and $VDOP(x, y, z)$, respectively.

V. *RAPID* Femtocell Deployment Strategy

RAPID proposes a DRL-based technique to reduce the negative impact of the geometry on overall positioning error. In this section, we describe *RAPID's* strategy for determining the 5G femtocell placement. First in Sec. V-A, we briefly overview the RL. Next, in Sec. V-B, we introduce a novel DRL formulation designed to address the placement problem. Our goal is to minimize the GDOP, while also ensuring that both the HDOP and VDOP remain low. Lastly, in Sec. V-C we outline the *RAPID* network architecture.

A. RL Foundamentals

1) Markov Decision Processes: In RL, an agent learns to make decisions by interacting with an environment. This interaction is typically modeled as a Markov decision process (MDP), defined by a tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$, where:

- $\mathcal{S}$ is the set of states representing the environment's possible configurations.

- $\mathcal{A}$ is the set of actions the agent can take.
- $\mathcal{P}(s', r|s, a)$ is the transition probability function, giving the probability of transitioning to state s' and receiving reward r when taking action a in state s.
- $\mathcal{R}(s, a, s')$ is the reward function, defining the immediate reward received upon transitioning from state s to state s' by taking action a.
- $\gamma \in [0, 1]$ is the discount factor, determining the importance of future rewards relative to immediate rewards.

The agent aims to learn a policy $\pi(a|s)$ that maps states to actions in a way that maximizes cumulative rewards over time [22].

$$\pi^* = \underset{\pi(a|s)}{\operatorname{argmax}} \mathbb{E}_\pi[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1}|s_t = s], \forall s \in \mathcal{S} \tag{5}$$

This is typically achieved through value-based methods, policy-based methods, or a combination of both.

2) Deep Reinforcement Learning: DRL extends RL by utilizing deep neural networks to approximate complex policies or value functions. DRL algorithms leverage the representational power of neural networks to handle high-dimensional state spaces and complex decision-making processes.

3) Policy Gradient Methods: Policy gradient methods directly parameterize the policy $\pi(a|s)$ and update its parameters to maximize expected cumulative rewards. The objective function for policy optimization is given by:

$$J(\theta) = \mathbb{E}_{\tau \sim \pi_\theta} \left[\sum_{t=0}^{T} \gamma^t r_t \right], \tag{6}$$

where $\tau = (s_0, a_0, r_0, \ldots, s_T, a_T, r_T)$ is a trajectory sampled from the policy π_θ.

One of the well-established policy gradient methods is proximal policy optimization (PPO), a state-of-the-art policy gradient method seeking to optimize the policy while ensuring stable and efficient learning [23]. PPO maximizes an objective function that penalizes large policy changes. The objective is given by:

$$\max_\theta \mathbb{E}_t \left[\min \left(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t \right) \right], \tag{7}$$

where $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}$ is the ratio of new policy to old policy, $\hat{A}_t$ is the advantage estimate, and ϵ is a hyperparameter controlling the extent of policy updates.

B. RAPID MDP Formulation

The 5G femtocell placement is formulated as an MDP problem. As illustrated in Figure 4, the problem begins with an initial room dimension and the PPO agent will propose a strategy for placing femtocells. Then the overall GDOP value for the proposed configuration will be computed and the environment will evolve to the next room dimension. The goal is to find a configuration that has the minimum GDOP among all possible points within a particular room dimension.

1) *State*: The environment state is the room dimension. In this study, it is assumed that the height of the room is fixed and its length and width will vary between a certain range (a_r, b_r).
$$s = (l, w, h) \tag{8}$$
2) *Action*: The action space is a vector of length 12 that represents the location of all femtocells.
$$a \in [p_1, p_2, p_3, p_4], \tag{9}$$
where $p_i = [x_i, y_i, z_i]$ is the location of the i-th femtocell. It is worth mentioning that all of the locations are zero-centered to enhance training effectiveness.
3) *Transition Model*: For training an agent capable of finding the optimal femtocell placement for different room dimensions, the room dimensions should vary during training. Therefore, after an action is taken, the environment will transit to another room dimension sampled from a uniform distribution between a_r and b_r. The height of all rooms is considered the fixed value of c_r;
$$L, W \sim U(a_r, b_r). \tag{10}$$
It should be emphasized that this definition is unlike traditional RL problems' transition models as the state transition is not dependent on the previous state or the action taken.
4) *Reward Function*: To ensure having a low average GDOP while keeping its value low in corner cases, the reward function is defined as follows.
$$r = -\sum_{i,j,k} \sqrt{\alpha VDOP^2(a, p_{ijk}) + \beta HDOP^2(a, p_{ijk})} - \zeta \max_{p_{ijk}} VDOP(a, p_{ijk}) - \delta \max_{p_{ijk}} HDOP(a, p_{ijk}), \tag{11}$$
where α, β, ζ, and δ are hyperparameters and p_{ijk} is the point we are calculating its GDOP value. By this reward function, we are penalizing the agent for creating strategies that have either a high overall GDOP or a high maximum value for HDOP and VDOP in the corner cases. Considering the continual mobility of a drone, it is inadequate to calculate the GDOP, VDOP, and HDOP for just one position. Thus, we assessed all potential locations within the indoor space that the drone might traverse and computed the cumulative rewards across these locations. This enables us to derive the reward values for the entire 3D area of interest.

C. RAPID Network Architecture

Here, we present the configuration and architecture details of the PPO network of the *RAPID* utilized for femtocell placement. Both the actor and critic networks have a similar architecture consisting of three fully connected layers with 64 and 128 hidden units, respectively. The input dimension matches the state space dimension. The output dimension in actor network outputs the probability distribution over the action space, while the critic network estimates the state-value function. The hyperparameters of the PPO network are outlined in Table II.

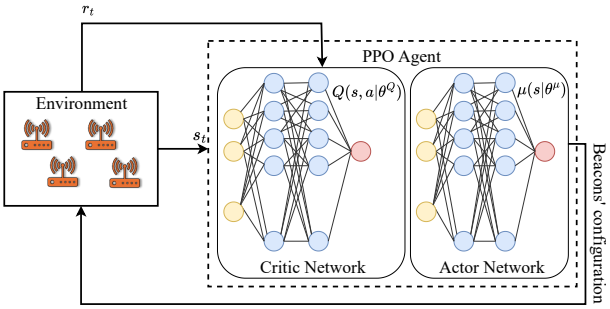


Figure 4: An overview of the *RAPID's* PPO learning framework

Table II: Hyperparameters of *RAPID's* PPO network

Hyperparameter	Value
Training Timesteps	40,000
Max Timesteps per Episode	20
Starting Std of Action Distribution	0.6
Decay Rate of Std	0.05
Update Frequency	80
Number of Policy Updates	80
ϵ (Epsilon Clip)	0.2
γ (Discount Factor)	0.99
Actor Learning Rate	0.0006
Critic Learning Rate	0.0005

VI. Results & Evaluations

In this section, we first present the experimental DRL setup used for training *RAPID* and the simulation environment used for evaluating the performance of the femtocell placement patterns. Next, we utilize the optimal pattern obtained with *RAPID* from our DRL setup and show how it improves localization accuracy over random placement patterns.

RAPID's DRL-based placement approach is trained using an *NVIDIA GeForce RTX 3090* GPU for 40,000 timesteps. Each episode consists of 20 timesteps, with room dimensions randomly sampled from a uniform distribution $U(5, 15)$. The room height is fixed at 4 m, and all the femtocells hung from the ceiling have a height range of 3.5 m to 4 m. The training average return of *RAPID* is depicted in Figure 5. As observed, *RAPID* adeptly learns the underlying dynamics of GDOP across diverse room dimensions and converges to a stable average return within the allotted training timesteps.

After training, *RAPID* can now be tested with any given indoor setup using custom room dimensions. To test the performance of *RAPID*, we design a simulation setup using *MATLAB 2022a* running on a *Dell Optiplex 7080* computer. Simulations are based on the deployment of four 5G femtocells. The primary goal of these simulations is to show that after *RAPID* is done with its training algorithm, it can provide the optimal placement of 5G femtocells for any given indoor setup in real time. Since this problem is inherently dependent on the dimensions of the room, we show the performance results on different dimensions to ensure that the proposed algorithm works on any given setup.

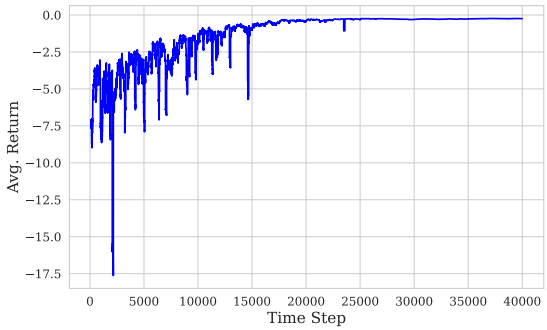


Figure 5: Training average return of *RAPID's* PPO network

Table III: 5G femtocells deployment for different room sizes with a random solution as a benchmark

Room Dimensions	5G-FC #1	5G-FC #2	5G-FC #3	5G-FC #4
Office Room (5m × 5m × 4m)	(4,1,3.9)	(5,3,3.9)	(3,1,3.6)	(2,4,3.9)
Conference Room (10m × 10m × 4m)	(10,4,3.5)	(9,6,3.7)	(10,2,3.8,)	(6,4,3.6)
Large Hall (15m × 15m × 4m)	(3,12,3.9)	(14,2,3.6)	(1,7,3.6)	(13,1,3.5)

To assess *RAPID's* effectiveness and adaptability, we test it on three distinct floor plans, each representing different room dimensions. The first scenario involves a typical office room with dimensions of $5\ m\ \times\ 5\ m\ \times\ 4\ m$, considered as a small space. The second scenario is that of a larger space, (e.g., conference room) measuring $10\ m\ \times\ 10\ m\ \times\ 4\ m$, considered as a medium-sized space. Lastly, the algorithm is tested for an edge case, (e.g., large hall) spanning $15\ m\ \times\ 15\ m\ \times\ 4\ m$. These diverse tests showcase the algorithm's versatility across room dimensions ranging from moderate to exceedingly spacious. We start with small to large to ensure that the *RAPID* can deliver the same performance across all of them. Moreover, in instances where the dimensions are extremely larger, such as a storage venue measuring $25\ m\ \times\ 25\ m\ \times\ 4\ m$, where the coverage range of 5G femtocells may fall short of spanning the entire floor plan, the sole adjustment needed is the introduction of additional 5G femtocells to meet the coverage demands, and apply *RAPID* to find a location for installing the newly added femtocell.

The non-optimized benchmark and the *RAPID* deployments are listed in Table III and Table IV, respectively. Measurements in both tables are given in meters.

Our goal is to demonstrate the xDOP values, including HDOP, VDOP, and GDOP, in 3D spaces. We calculate the average xDOP values for each (x, y) point across all z planes and depict them with heat maps. This method simplifies the spatial representation and enables visualization across all z planes, not just a select few, yielding more thorough results.

To understand the specific effects on localization accuracy between the $X - Y$ plane and the Z-axis, we individually present HDOP, VDOP, and GDOP. This approach allows us to identify situations where a certain setup offers better hori-

Table IV: 5G femtocells deployment for different room sizes with *RAPID* solution

Room Dimensions	5G-FC #1	5G-FC #2	5G-FC #3	5G-FC #4
Office Room (5m × 5m × 4m)	(0,5,3.5)	(2,0,3.6)	(2,3,4)	(5,5,3.5)
Conference Room (10m × 10m × 4m)	(0,9,3.5)	(5,0,3.6)	(4,6,4)	(10,10,3.5)
Large Hall (15m × 15m × 4m)	(0,13,3.5)	(7,0,3.6)	(7,9,4)	(15,14,3.5)

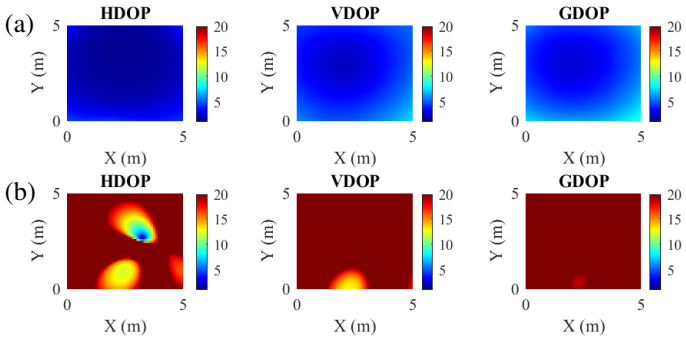


Figure 6: xDOP representation for an office room and comparison between (a) *RAPID's* optimal deployment as shown in Table IV vs. (b) a random placement as shown in Table III

zontal accuracy but poorer vertical estimations, offering a more detailed perspective on performance.

Figure 6 presents a comparison of xDOP values, analyzing optimal solutions from *RAPID* against random placements, with a focus on a smaller space, such as an office room. Similar comparisons for larger spaces, including a conference room and a large hall, are illustrated in Figure 7 and Figure 8, respectively. These comparisons aim to assess *RAPID's* efficacy across various room sizes. As depicted in these figures, *RAPID* demonstrates strong performance in all scenarios.

Additionally, these figures underscore the importance of utilizing the *RAPID* framework for deploying 5G femtocells, rather than relying on random placement strategies. This preference is justified by the substantial impact of GDOP values on the system's overall accuracy. To reduce the multiplicative factor in the equation $\sigma_T^2(x,y,z) = GDOP \cdot \sigma_r^2$, achieving lower GDOP values is essential. By employing *RAPID*, we can

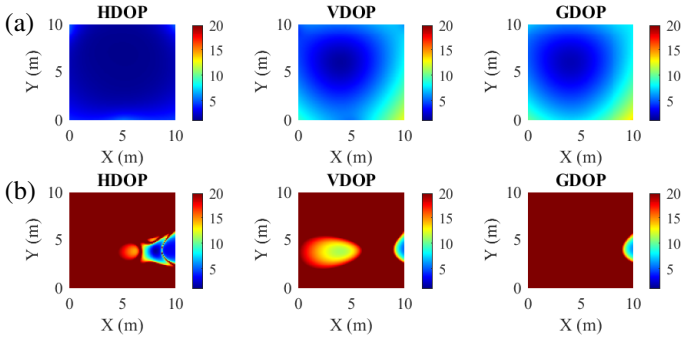


Figure 7: xDOPP representation for a conference room and comparison between (a) *RAPID's* optimal deployment as shown in Table IV vs. (b) a random placement as shown in Table III

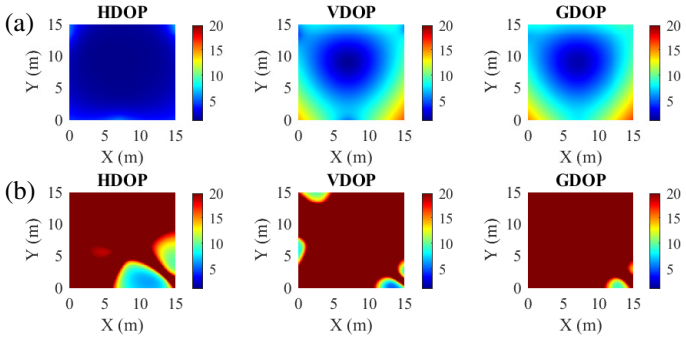


Figure 8: xDOPP representation for a large hall and comparison between (a) *RAPID's* optimal deployment as shown in Table IV vs. (b) a random placement as shown in Table III

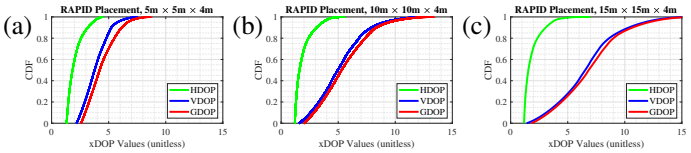


Figure 9: CDF plots for xDOP values for various room dimensions with the *RAPID* configurations shown in Table IV

successfully reach this objective.

Figure 9 illustrates the cumulative distribution function (CDF) of xDOP values for optimal placements across three different setups. This analysis aims to highlight the points within the drone domain (set $\mathbb{D}$) that exhibit xDOP values beneath a specific threshold. The primary objective of the optimization was to ensure that the majority of points maintained xDOP values under 20, a goal that was successfully met in all examined scenarios. Moreover, the figure reveals that most points actually display xDOP values below 15, surpassing initial expectations. This indicates that, in the final 3D localization process, the spatial configuration of the femtocells has a minimal detrimental impact on overall accuracy.

For example, in the office room, GDOP values fall below 8, while in the conference room, it's 13, and in a large hall, it's 15. All these values, being under 20, underscore the efficiency of the *RAPID* framework.

To set a benchmark and underline the critical role of optimal femtocell placement, we contrast the outcomes of random placement with those of strategic deployment. Figure 10 showcases plots similar to those in Figure 9, but for randomly placed femtocells. In these scenarios, the xDOP values escalate to the thousands, a stark contrast to their sub-20 levels achieved through optimal deployment. This significant increase suggests

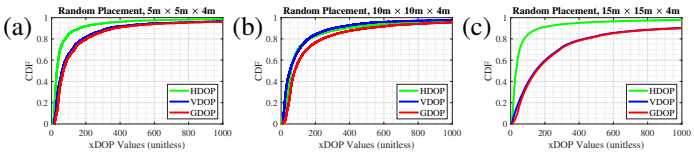


Figure 10: CDF plots for xDOP values for various room dimensions with random configurations shown in Table III

that 3D localization accuracy becomes highly unreliable without a placement strategy, as evidenced by xDOP values soaring.

For instance, consider a case where the same high-accuracy ranging, with a σ_r^2 under $1\ cm$–achieved through substantial effort and expense–is employed. Based on Figure 10, a random arrangement could result in a final 3D localization accuracy with a σ_T^2 around $10\ m$, rendering the precision efforts futile and leading to an unacceptable localization accuracy far exceeding the desirable sub-meter precision. Conversely, based on results seen in Figure 9, employing the *RAPID* framework for femtocell deployment ensures the σ_T^2 for the overall 3D localization accuracy remains around $10\ cm$. This represents almost a thousandfold improvement in accuracy over random placement. Such a comparison underscores the pivotal impact of spatial geometry between the user and positioning nodes on localization accuracy and the critical need for an algorithm like *RAPID* that minimizes this error effectively.

VII. Conclusion and Future Work

Conclusion. In this research, we explored the impact of the geometric relationship between 5G-enabled drones and 5G femtocells on the accuracy of indoor localization. To address this, we introduced our framework, *RAPID*. We discovered that localization errors stem from both range measurements and the GDOP, with the latter having a more significant negative effect on the Z-axis than on the $X-Y$ plane. To enhance positioning accuracy, we devised a DRL-aided framework to determine the optimal placement of 5G femtocells. Our assessments showed that our DRL methodology substantially improves estimation along the Z-axis and overall GDOP.

Future Work. Using a small-scale 5G testbed, we have conducted comprehensive measurements measurements for the over-the-air performance of UEs [24] to establish a benchmark for our real-life testing. In the future, we seek to integrate *RAPID* into our existing testbed to create a proof-of-concept that can demonstrate how the localization accuracy is improved for a real-life UE.

References

[1] H. Y. Jeong, B. D. Song, and S. Lee, "Optimal scheduling and quantitative analysis for multi-flying warehouse scheduling problem: Amazon airborne fulfillment center," *Transportation Research Part C: Emerging Technologies*, vol. 143, p. 103831, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X22002522

[2] J. V.-V. Gerwen, K. Geebelen, J. Wan, W. Joseph, J. Hoebeke, and E. De Poorter, "Indoor drone positioning: Accuracy and cost trade-off for sensor fusion," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 1, pp. 961–974, 2022.

[3] G. Chi, Z. Yang, J. Xu, C. Wu, J. Zhang, J. Liang, and Y. Liu, "Wi-Drone: Wi-Fi-based 6-DoF tracking for indoor drone flight control," in *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*, ser. MobiSys '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 56–68. [Online]. Available: https://doi.org/10.1145/3498361.3538936

[4] A. Famili, A. Stavrou, H. Wang, and J.-M. J. Park, "PILOT: high-precision indoor localization for autonomous drones," *IEEE Transactions on Vehicular Technology*, pp. 1–15, 2022.

[5] P. Razzaghi, A. Tabrizian, W. Guo, S. Chen, A. Taye, E. Thompson, A. Bregeon, A. Baheri, and P. Wei, "A survey on reinforcement learning in aviation applications," *arXiv preprint arXiv:2211.02147*, 2022.

[6] A. Mohan, A. Zhang, and M. Lindauer, "Structure in reinforcement learning: A survey and open problems," 2023.

[7] A. Famili, T. O. Atalay, and A. Stavrou, "5GPS: 5G femtocell placement strategies for ultra-precise indoor localization in the Metaverse," *International Conference on Computing, Networking and Communications (ICNC 2024)*, 2024.

[8] N. Mazyavkina, S. Sviridov, S. Ivanov, and E. Burnaev, "Reinforcement learning for combinatorial optimization: A survey," *Computers & Operations Research*, vol. 134, p. 105400, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0305054821001660

[9] H. Wang, N. Rajagopal, A. Rowe, B. Sinopoli, and J. Gao, "Efficient beacon placement algorithms for time-of-flight indoor localization," in *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ser. SIGSPATIAL '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 119–128. [Online]. Available: https://doi.org/10.1145/3347146.3359344

[10] R. Sharma and V. Badarla, "Analysis of a novel beacon placement strategy 3D localization in indoor spaces," in *2019 11th International Conference on Communication Systems Networks (COMSNETS)*, Jan 2019, pp. 320–327.

[11] J. Schmalenstroeer and R. Haeb-Umbach, "Investigations into bluetooth low energy localization precision limits," in *2016 24th European Signal Processing Conference (EUSIPCO)*, Aug 2016, pp. 652–656.

[12] A. Famili, A. Stavrou, H. Wang, and J.-M. Park, "iDROP: robust localization for indoor navigation of drones with optimized beacon placement," *IEEE Internet of Things Journal*, vol. 10, no. 16, pp. 14 226–14 238, 2023.

[13] N. Rajagopal, S. Chayapathy, B. Sinopoli, and A. Rowe, "Beacon placement for range-based indoor localization," in *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Oct 2016, pp. 1–8.

[14] A. Famili, A. Stavrou, H. Wang, and J.-M. Park, "SPIN: sensor placement for indoor navigation of drones," in *IEEE LATINCOM 2022*, nov 2022.

[15] W. Wang, A. X. Liu, and K. Sun, "Device-free gesture tracking using acoustic signals," in *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 82–94. [Online]. Available: https://doi.org/10.1145/2973750.2973764

[16] R. C. Luo and T.-J. Hsiao, "Indoor localization system based on hybrid Wi-Fi/BLE and hierarchical topological fingerprinting approach," *IEEE Transactions on Vehicular Technology*, pp. 10 791–10 806, 2019.

[17] A. B. Pizarro, J. P. Beltrán, M. Cominelli, F. Gringoli, and J. Widmer, "Accurate ubiquitous localization with off-the-shelf IEEE 802.11ac devices," in *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 241–254. [Online]. Available: https://doi.org/10.1145/3458864.3468850

[18] G. Himona, A. Famili, A. Stavrou, V. Kovanis, and Y. Kominis, "Isochrons in tunable photonic oscillators and applications in precise positioning," in *Physics and Simulation of Optoelectronic Devices XXXI*, vol. 12415. SPIE, 2023, pp. 82–86.

[19] A. Famili, T. O. Atalay, A. Stavrou, H. Wang, and J.-M. Park, "OFDRA: Optimal femtocell deployment for accurate indoor positioning of RIS-mounted AVs," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 12, pp. 3783–3798, 2023.

[20] M. A. Spirito, "On the accuracy of cellular mobile station location estimation," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 3, pp. 674–685, May 2001.

[21] P. Massatt and K. Rudnick, "Geometric formulas for dilution of precision calculations," *NAVIGATION, Journal of The Institute of Navigation*, vol. 37, no. 4, pp. 379–392, Winter 1990-1991.

[22] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. The MIT Press, 2018. [Online]. Available: http://incompleteideas.net/book/the-book-2nd.html

[23] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *CoRR*, vol. abs/1707.06347, 2017. [Online]. Available: http://arxiv.org/abs/1707.06347

[24] T. O. Atalay, A. Famili, D. Stojadinovic, and A. Stavrou, "Demystifying 5g traffic patterns with an indoor ran measurement campaign," in *GLOBECOM 2023-2023 IEEE Global Communications Conference*. IEEE, 2023, pp. 1185–1190.