ACCUSTOM: Modeling & Formulations

Bharath Keshavamurthy

January 2024

1 CUSTOM & ACCUSTOM

- **C**ontrolling **U**AV **S**warms for **T**raffic **O**ffloading in **M**IMO (ICC).
- **A**daptive **C**ontrol & **C**oordination of **U**AV **S**warms for **T**raffic **O**ffloading in **M**IMO (JSAC).

2 Deployment model

2.1 Grid tessellation

- Rectangular deployment site of length $x_{\max}$, breadth $y_{\max}$, and height $z_{\max}$, enclosed in a GPS geo-fence.
- This deployment site is tessellated into a grid world, with each rectangular grid voxel of length Δx, breadth Δy, and height Δz.
- Grid tessellation voxel counts:
  - Number of voxels in the X-direction: $N_x{=}x_{\max}/\Delta x$;
  - Number of voxels in the Y-direction: $N_y{=}y_{\max}/\Delta y$;
  - Number of voxels in the Z-direction: $N_z{=}z_{\max}/\Delta z$;
  - Total number of voxels in the grid: $N_{\text{vox}}{=}N_x{\cdot}N_y{\cdot}N_z$.

2.2 UAVs & GNs

- U heterogeneous rotary-wing Unmanned Aerial Vehicles (UAVs).
- G heterogeneous Ground Nodes (GNs) distributed uniformly throughout the site.
- The set of UAVs is defined as $\mathcal{U}{\triangleq}\{1,2,3,\ldots,U\}$.
- The set of GNs is defined as $\mathcal{G}{\triangleq}\{1,2,3,\ldots,G\}$.

2.3 Cartesian coordinate system

- For a UAV $u{\in}\mathcal{U}$, its position at time $t{\in}[0,\infty)$ is defined in a Cartesian coordinate system as $\mathbf{p}'_u(t){\triangleq}[x'_u,y'_u,z'_u]$, where $0{\leq}x'_u{<}x_{\max}$, $0{\leq}y'_u{<}y_{\max}$, and $0{\leq}z'_u{<}z_{\max}$.
- For a GN $g{\in}\mathcal{G}$, its position at time $t{\in}[0,\infty)$ is defined in a Cartesian coordinate system as $\mathbf{p}'_g(t){\triangleq}[x'_g,y'_g,z'_g]$, where $0{\leq}x'_g{<}x_{\max}$, $0{\leq}y'_g{<}y_{\max}$, and $0{\leq}z'_g{<}z_{\max}$.
- The GNs do not necessarily have to be at ground level.
- The GNs are assumed to be stationary, i.e., $\mathbf{p}'_g(t){=}\mathbf{p}'_g, \forall t{\in}[0,\infty)$.

2.4 Voxel-based coordinate system

- Here, we map the Cartesian coordinate system to a voxel-based coordinate system.
- Let $\mathcal{J}{\triangleq}\{1,2,3,\ldots,N_{\text{vox}}\}$ be an index set. Then, the set of voxels in the deployment site is defined as $\mathcal{V}{\triangleq}\{\mathcal{V}_j : j{\in}\mathcal{J}\}$. Here, a voxel $\mathcal{V}_j, \forall j{\in}\mathcal{J}$ is defined as $\mathcal{V}_j{\triangleq}\{[x_j,y_j,z_j] : x_{j,\min}{\leq}x_j{<}x_{j,\max}, y_{j,\min}{\leq}y_j{<}y_{j,\max}, z_{j,\min}{\leq}z_j{<}z_{j,\max}\}$—where $x_{j,\min}$ & $x_{j,\max}$ are the lower- & upper-bounds of voxel $\mathcal{V}_j$ in the X-direction, $y_{j,\min}$ & $y_{j,\max}$ denote the lower- & upper-bounds of voxel $\mathcal{V}_j$ in the Y-direction, and $z_{j,\min}$ & $z_{j,\max}$ are the lower- & upper-bounds of voxel $\mathcal{V}_j$ in the Z-direction.
- At time $t{\in}[0,\infty)$, a UAV $u{\in}\mathcal{U}$ at position $\mathbf{p}'_u(t){=}[x'_u,y'_u,z'_u]$ belongs to a voxel $\mathcal{V}_j,{\in}\mathcal{J}$, if $x_{j,\min}{\leq}x'_u{<}x_{j,\max}, y_{j,\min}{\leq}y'_j{<}y_{j,\max}, z_{j,\min}{\leq}z'_j{<}z_{j,\max}$.
- Thus, mapping $p'_u(t)$ from Cartesian coordinates to our voxel-based coordinate system, we get $\mathbf{p}_u(t){=}\left[\frac{1}{8}\sum_{j'=1}^{8} x_{j'}, \frac{1}{8}\sum_{j'=1}^{8} y_{j'}, \frac{1}{8}\sum_{j'=1}^{8} z_{j'}\right]$, where the centroid of the voxel which UAV u to belongs at time t is used as its position in our voxel-based coordinates.
- We use a similar mapping for the positions of GNs from Cartesian to voxel-based coordinates. A GN $g{\in}\mathcal{G}$ at position $\mathbf{p}'_g(t){=}\mathbf{p}'_g{=}[x'_g,y'_g,z'_g], \forall t{\in}[0,\infty)$ belongs to a voxel $\mathcal{V}_j,{\in}\mathcal{J}$, if $x_{j,\min}{\leq}x'_g{<}x_{j,\max}, y_{j,\min}{\leq}y'_g{<}y_{j,\max}, z_{j,\min}{\leq}z'_g{<}z_{j,\max}$; and subsequently, $\mathbf{p}'_g$ is mapped to voxel coordinates as $\mathbf{p}_g{=}\left[\frac{1}{8}\sum_{j'=1}^{8} x_{j'}, \frac{1}{8}\sum_{j'=1}^{8} y_{j'}, \frac{1}{8}\sum_{j'=1}^{8} z_{j'}\right]$.
- We used this voxel-based coordinate system throughout the rest of our setup.

2.5 Depot

- Let $\mathcal{F}''$ be the contiguous set of coordinates corresponding to the takeoff & landing pads for the UAVs in the fleet (i.e., the UAV depot). In Cartesian coordinates:

$$\mathcal{F}''{=}\{[x'_i,y'_i,z'_i] : 0{\leq}x'_i{<}x_f, 0{\leq}y'_i{<}y_f, 0{\leq}z'_i{<}z_f\},$$

  where x_f, y_f, and z_f are the end coordinates of the depot in the deployment site in X-, Y-, and Z-directions, respectively. Representing $\mathcal{F}''$ in terms of its constituent voxels:

$$\mathcal{F}'{=}\{\mathcal{V}_i : 0{\leq}x_{i,\min}{<}x_{i,\max}{<}x_f, 0{\leq}y_{i,\min}{<}y_{i,\max}{<}y_f, 0{\leq}z_{i,\min}{<}z_{i,\max}{<}z_f\},$$

  where $x_{i,\min}$ & $x_{i,\max}$ are the lower- & upper-bounds of voxel $\mathcal{V}_i$ in the X-direction, $y_{i,\min}$ & $y_{i,\max}$ denote the lower- & upper-bounds of voxel $\mathcal{V}_i$ in the Y-direction, and $z_{i,\min}$ & $z_{i,\max}$ are the lower- & upper-bounds of voxel $\mathcal{V}_i$ in the Z-direction. Finally, representing the depot in terms of the centroids of its constituent voxels, we get

$$\mathcal{F}{=}\left\{\left[\frac{1}{8}\sum_{i'=1}^{8} x_{i'}, \frac{1}{8}\sum_{i'=1}^{8} y_{i'}, \frac{1}{8}\sum_{ji'=1}^{8} z_{i'}\right] : [x_{i'},y_{i'},z_{i'}]{\in}\mathcal{V}_i, \forall\mathcal{V}_i{\in}\mathcal{F}'\right\}.$$

- Therefore, with this voxel-based representation of the depot, we impose the following constraints on the positions of the UAVs:

$$\begin{aligned}
&\mathbf{p}_u(t){\in}\mathcal{F}, \forall t{\in}[0,t_{u,\text{init}}], \forall u{\in}\mathcal{U};\\
&\mathbf{p}_u(t){\in}\mathcal{F}, \forall t{\geq}t_{u,\text{term}}, \forall u{\in}\mathcal{U}; \text{ and}\\
&\mathbf{p}_u(t){\in}\mathcal{F}, \forall t{\in}[t_{u,\text{chrg}}, t_{u,\text{chrg}}{+}\tau_{u,\text{chrg}}], \forall u{\in}\mathcal{U};
\end{aligned}$$

  where $t_{u,\text{init}}$ & $t_{u,\text{term}}$ are the arbitrary mission start and end times of a UAV $u{\in}\mathcal{U}$, respectively—determined dynamically by the Upper Agent (UA) in our Hierarchical Reinforcement Learning (HRL) framework (described later in this document); and $t_{u,\text{chrg}}$ is the time at which a UAV $u{\in}\mathcal{U}$ comes in to the depot $\mathcal{F}$ for recharging, while $\tau_{u,\text{chrg}}$ is the amount of time spent recharging (both these parameters are again determined dynamically by the UA in our HRL framework).

2.6 No-Fly Zones (NFZs) & Unavoidable obstacles

- Let $\mathcal{Z}''$ be the set of Cartesian coordinates corresponding to NFZs and unavoidable obstacles across the deployment site, and let $\mathcal{Z}'$ be its representation in terms of its constituent voxels. Thus, representing NFZs and unavoidable obstacles in terms of the centroids of its constituent voxels, we get

$$\mathcal{Z}{=}\left\{\left[\frac{1}{8}\sum_{i'=1}^{8}x_{i'},\frac{1}{8}\sum_{i'=1}^{8}y_{i'},\frac{1}{8}\sum_{ji'=1}^{8}z_{i'}\right]:[x_{i'},y_{i'},z_{i'}]{\in}\mathcal{V}_i,\forall\mathcal{V}_i{\in}\mathcal{Z}'\right\}.$$

- Therefore, with this voxel-based representation of the NFZs and unavoidable obstacles across the site, we impose the following constraint on the positions of the UAVs:

$$\mathbf{p}_u(t){\notin}\mathcal{Z},\forall u{\in}\mathcal{U},\forall t{\in}[0,\infty).$$

2.7 UAV & GN positioning

- We assume that, during the initial GN deployment (uniformly throughout the site), the GNs do not overlap in the same voxel, i.e., $\mathbf{p}_{g_1}{\neq}\mathbf{p}_{g_2},\forall g_1,g_2{\in}\mathcal{G},g_1{\neq}g_2$.
- In addition to the UAVs possessing LiDARs and other sensors to prevent collisions amongst themselves while operating in the deployment site, we enforce a UAV-UAV collision-avoidance constraint, i.e., $\mathbf{p}_{u_1}(t){\neq}\mathbf{p}_{u_2}(t),\forall u_1,u_2{\in}\mathcal{U},u_1{\neq}u_2,\forall t{\in}[0,\infty)$. In other words, two UAVs cannot ever occupy the same voxel simultaneously.

3 Communication model

3.1 Communication subsystems of the GNs & UAVs

- A GN $g{\in}\mathcal{G}$ has A_g antennas arranged in a rectangular planar array, driven by A_g Tx-Rx chains via the TDD protocol. To enforce heterogeneity in GN design, for two distinct GNs $g_1,g_2{\in}\mathcal{G},g_1{\neq}g_2$, A_{g_1} may or may not be equal to A_{g_2}.
- A UAV $u{\in}\mathcal{U}$ has A_u antennas arranged in a rectangular planar array, again driven by A_u Tx-Rx chains via the TDD protocol. Similarly, to enforce heterogeneity in UAV design, for two distinct UAVs $u_1,u_2{\in}\mathcal{U},u_1{\neq}u_2$, A_{u_1} may or may not be equal to A_{u_2}.

3.2 GN requests

- Each GN $g{\in}\mathcal{G}$ generates a traffic offloading request according to a Poisson process with rate Λ [requests per unit time].
- Each request from a GN $g{\in}\mathcal{G}$ constitutes a header and the data payload. The request header consists of the following fields: the priority value (χ_g), the maximum allowed latency ($\delta_{g,\max}$), the data payload size (ν_g), and the post-deadline discount factor (γ_g) of the request from GN $g{\in}\mathcal{G}$ belonging to a specific traffic class (see table 1).

3.3 GN-UAV uplink service model

- If δ_{gu} is the time taken by GN $g{\in}\mathcal{G}$ to offload its data payload to its serving UAV $u{\in}\mathcal{U}$ (based on GN-UAV positioning, A2G channel conditions, and MIMO beam-forming design), then the reward received by the UAV u is given by $\Omega_{gu}{=}\chi_g\gamma_g^{(\delta_{gu}-\delta_{g,\max})}$.
- A UAV $u{\in}\mathcal{U}$ can serve multiple requests simultaneously, while a specific request from a GN $g{\in}\mathcal{G}$ should only be associated with one UAV. Once a GN request is associated with a UAV, this specific request should be fully served by the UAV.

Traffic Class	Priority χ	Max Latency $\delta_{\max}$ mins	Payload Size ν Mb	Discount Factor γ
Telemetry	100	9.1	256	0.1
File Transfer	24	19.0	536	0.8
Image	72	14.5	512	0.33
Video Stream	84	11.6	1387	0.24

Table 1: Traffic classes, Priorities, Latencies, Payload sizes, and Discount factors.

3.4 Signal model

- For a GN $g \in \mathcal{G}$, at time $t \in [0, \infty)$, the transmitted signal is described by

$$\mathbf{x}_g(t) = \mathbf{\Phi}_g(t)\mathbf{s}_g(t), \tag{1}$$

where $\mathbf{x}_g(t) \in \mathbb{C}^{A_g \times 1}$ is the transmitted signal, $\mathbf{\Phi}_g(t) \in \mathbb{C}^{A_g \times A_g}$ is the linear precoding matrix applied at the GN, and $\mathbf{s}_g(t) \in \mathbb{C}^{A_g \times 1}$ is the symbol vector with $\mathbb{E}[\mathbf{s}_g^H(t)\mathbf{s}_g(t)]=1$.

- For simplicity of notation, omitting the time variable, the transmitted signal from GN $g \in \mathcal{G}$ is $\mathbf{x}_g = \mathbf{\Phi}_g \mathbf{s}_g$, where $\mathbf{x}_g \in \mathbb{C}^{A_g \times 1}$, $\mathbf{\Phi}_g \in \mathbb{C}^{A_g \times A_g}$, and $\mathbf{s}_g \in \mathbb{C}^{A_g \times 1}$ with $\mathbb{E}[\mathbf{s}_g^H \mathbf{s}_g]=1$.
- Consequently, at time $t \in [0, \infty)$, the signal received at UAV $u \in \mathcal{U}$ is described by

$$\mathbf{y}_u = \sum_{g \in \mathcal{G}_u} \mathbf{H}_{gu}\mathbf{x}_g + \mathbf{w}_u, \tag{2}$$

where $\mathbf{y}_u \in \mathbb{C}^{A_u \times 1}$ is the received signal at the UAV, $\mathcal{G}_u$ is the set of GNs associated with the UAV, $\mathbf{H}_{gu} \in \mathbb{C}^{A_u \times A_g}$ is the channel between a GN $g \in \mathcal{G}_u$ and the UAV u obtained via our ray-tracing augmented channel estimation procedure (detailed later in this document), and $\mathbf{w}_u \sim \mathcal{CN}(\mathbf{0}, BN_0\mathbf{I}_{A_u})$ is the AWGN noise vector at the UAV—with B being the bandwidth pre-assigned to the UAV, N_0 is the noise power spectral density, and $\mathbf{I}_{A_u}$ is the identity matrix with dimensions $A_u \times A_u$.

3.5 Channel model

- The channel gain between a GN Tx antenna m and its serving UAV's Rx antenna n is

$$h_{mn} = \beta_{mn} - 10\alpha_{mn}\log_{10} d_{mn} + \xi_{mn} + \lambda_{mn}, \tag{3}$$

where h_{mn} is the channel gain (in dB), β_{mn} is the average channel gain at a reference distance of 1 m (in dB), α_{mn} is the pathloss exponent, d_{mn} is the 3D Euclidean distance between position vectors $\mathbf{p}_m = [x_m, y_m, z_m]$ and $\mathbf{p}_n = [x_n, y_n, z_n]$, $\xi_{mn} \sim \mathcal{N}\left(0, \sigma^2_{\mathrm{SF},mn}\right)$ is a random variable denoting the shadow-fading component with variance $\sigma^2_{\mathrm{SF},mn}$, and $\lambda_{mn} \sim \mathcal{CN}\left(\mu_{\mathrm{SSF},mn}, \sigma^2_{\mathrm{SSF},mn}\right)$ is a random variable denoting the small-scale fading component with mean $\mu_{\mathrm{SSF},mn}$ and variance $\sigma^2_{\mathrm{SSF},mn}$.

- Refer to the ray-tracing augmented channel estimation procedure outlined later in this document to determine these channel metrics $\left(\beta_{mn}, \alpha_{mn}, \sigma^2_{\mathrm{SF},mn}, \mu_{\mathrm{SSF},mn}, \sigma^2_{\mathrm{SSF},mn}\right)$, given the Tx antenna and Rx antenna position vectors $\mathbf{p}_m$ and $\mathbf{p}_n$.

3.6 Additional notes

- The UAVs harvest traffic from the GNs according to a *move-and-receive* protocol. Here, the signal degradation effects due to Doppler shifts brought on by the UAV's motion towards and/or away from the transmitting GNs are compensated via a radio design at the UAVs derived from the PLL-enabled RF-based adaptive Doppler compensation technique proposed in the state-of-the-art.
- The spectrum allocated to this application (i.e., **ACCUSTOM**) is discretized into U channels, with each channel having a preset bandwidth of B. Let the band-edges of this spectrum (each band-edge has a predetermined bandwidth of $B_C << B$) be designated as control channels for UAV-UAV messaging and for coordination messages between the centralized operations hub and the UAVs in the fleet.

- Since the control traffic (UAV-UAV and Hub-UAV) involves very short control frames relative to the large payload frames in the data traffic, we can ignore the latencies from control communication in our system model and subsequent formulations.

4 UAV mobility energy consumption model

The full energy-conscious 3D trajectory of a UAV $u\in\mathcal{U}$ in the fleet (obtained via the Learning Competitive Swarm Optimization algorithm) is defined as

$$\mathcal{Q}_u\left(t_{u,\text{init}},t_{u,\text{term}}\right)\triangleq\Big\{\mathbf{p}_u(\tau),v_u(\tau):\tau\in\left[t_{u,\text{init}},t_{u,\text{term}}\right]\Big\}, \tag{4}$$

where $t_{u,\text{init}}$ & $t_{u,\text{term}}$ denote the trajectory start & end times (determined by HRL), $\mathbf{p}_u(\tau)$ denotes the way-point (in voxel coordinates) at trajectory time $\tau\in\left[t_{u,\text{init}},t_{u,\text{term}}\right]$, and the UAV's velocity in the 3D plane $v_u(\tau)$ is decomposed into its horizontal & vertical components, i.e., $v_{u,\text{horz}}(\tau)=|v_u(\tau)|\cos\angle v_u(\tau)$ & $v_{u,\text{vert}}(\tau)=|v_u(\tau)|\sin\angle v_u(\tau)$, $\forall\tau\in\left[t_{u,\text{init}},t_{u,\text{term}}\right]$.

4.1 Horizontal motion

For a rotary-wing UAV $u\in\mathcal{U}$, its mobility energy consumption upon executing an arbitrary horizontal (2D) trajectory comprising velocities $v_{u,\text{horz}}(\tau)$ from $\tau=t_{u,\text{init}}$ to $\tau=t_{u,\text{term}}$ is

$$\begin{aligned}E_{\text{horz}}\left(\left\{v_{u,\text{horz}}(\tau)\right\}_{\tau=t_{u,\text{init}}}^{\tau=t_{u,\text{term}}}\right)=&\int_{t_{u,\text{init}}}^{t_{u,\text{term}}}C_0\left(1+C_1v_{u,\text{horz}}^2(\tau)\right)d\tau+\\&\int_{t_{u,\text{init}}}^{t_{u,\text{term}}}\kappa(\tau)C_2\left(\sqrt{\kappa^2(\tau)+\frac{v_{u,\text{horz}}^4(\tau)}{C_3^2}}-\frac{v_{u,\text{horz}}^2(\tau)}{C_3}\right)^{\frac{1}{2}}d\tau+\\&\int_{t_{u,\text{init}}}^{t_{u,\text{term}}}C_4v_{u,\text{horz}}^3(\tau)d\tau+\frac{1}{2}m_u\left(v_{u,\text{horz}}^2(t_{u,\text{term}})-v_{u,\text{horz}}^2(t_{u,\text{init}})\right),\end{aligned} \tag{5}$$

where g is the acceleration due to gravity at the deployment site; m_u is the mass of UAV u; C_0, C_1, C_2, C_3, and C_4 are constants that depend on the UAV weight (m_ug), rotor disc area of the UAV, air density, etc.; and the term $\kappa(\tau)$ is defined as follows

$$\kappa(\tau)=\sqrt{1+\frac{\left(\rho\omega_us_u\Gamma_uv_{u,\text{horz}}^2(\tau)+2m_ua_{u,\text{horz}}(\tau)\right)^2}{4m_u^2g^2}}, \tag{6}$$

with ρ being the air density, ω_u being the fuselage drag ratio of the UAV, s_u is the rotor solidity factor of the UAV, Γ_u is the rotor disc area of the UAV, and $a_{u,\text{horz}}(\tau)=\frac{dv_{u,\text{horz}}(\tau)}{d\tau}$ is the horizontal acceleration component of the UAV at trajectory time τ.

4.2 Vertical motion

For a rotary-wing UAV $u\in\mathcal{U}$, its mobility energy consumption upon executing a vertical trajectory comprising velocities $v_{u,\text{vert}}(\tau)$ from $\tau=t_{u,\text{init}}$ to $\tau=t_{u,\text{term}}$ is

$$\begin{aligned}E_{\text{vert}}\left(\left\{v_{u,\text{vert}}(\tau)\right\}_{\tau=t_{u,\text{init}}}^{\tau=t_{u,\text{term}}}\right)=&\int_{t_{u,\text{init}}}^{t_{u,\text{term}}}C_0\left(1+C_1v_{u,\text{vert}}^2(\tau)\right)d\tau\\&\int_{t_{u,\text{init}}}^{t_{u,\text{term}}}\kappa(\tau)C_2\left(\sqrt{\kappa^2(\tau)+\frac{v_{u,\text{vert}}^4(\tau)}{C_3^2}}-\frac{v_{u,\text{vert}}^2(\tau)}{C_3}\right)^{\frac{1}{2}}d\tau,\end{aligned} \tag{7}$$

where C_0, C_1, C_2, C_3, and C_4 are constants that depend on the UAV weight (m_ug), rotor disc area of the UAV, air density, etc.; and the term $\kappa(\tau)$ is defined as follows

$$\kappa(\tau)=\sqrt{1+\frac{\left(\rho\omega_us_u\Gamma_uv_{u,\text{vert}}^2(\tau)+2m_ua_{u,\text{vert}}(\tau)\right)^2}{4m_u^2g^2}}, \tag{8}$$

with ρ being the air density, ω_u being the fuselage drag ratio of the UAV, s_u is the rotor solidity factor of the UAV, Γ_u is the rotor disc area of the UAV, and $a_{u,\text{vert}}(\tau){=}\frac{dv_{u,\text{vert}}(\tau)}{d\tau}$ is the vertical acceleration component of the UAV at trajectory time τ.

4.3 Horizontal + Vertical motion

Under the 3D trajectory setup described earlier in this section, the overall mobility energy consumption of a rotary-wing UAV $u{\in}\mathcal{U}$ is given by

$$E_{\text{3D}}\left(\mathcal{Q}_u\left(t_{u,\text{init}},t_{u,\text{term}}\right)\right){=}E_{\text{3D}}\left(\left\{\mathbf{p}_u(\tau),v_{u,\text{horz}}(\tau),v_{u,\text{vert}}(\tau):\tau{\in}\left[t_{u,\text{init}},t_{u,\text{term}}\right]\right\}\right) \quad (9)$$
$$={E}_{\text{horz}}\left(\left\{v_{u,\text{horz}}(\tau)\right\}_{\tau=t_{u,\text{init}}}^{\tau=t_{u,\text{term}}}\right)+E_{\text{vert}}\left(\left\{v_{u,\text{vert}}(\tau)\right\}_{\tau=t_{u,\text{init}}}^{\tau=t_{u,\text{term}}}\right).$$

4.4 Additional notes

- Each UAV $u{\in}\mathcal{U}$ in the fleet has an on-board energy source (e.g., battery) with maximum capacity $E_{u,\max}$ such that $E_u(0){=}E_{u,\max}$, i.e., each UAV is assumed to be fully-charged at mission start.
- Since the UAVs in the fleet are only receiving traffic from the GNs, and since the UAVs are not themselves involved in any multi-antenna transmissions, we can safely ignore the communication energy consumption of a UAV (used to only operate the receive chains, negligible relative to its mobility energy consumption).

5 Ray-tracing augmented channel estimation

5.1 Problem setup

- We generate a radio map for the entire deployment site using a channel estimation procedure seeded by a small subset of ray-tracing measurements on Wireless InSite.
- The channel gain between a GN Tx antenna m and its serving UAV's Rx antenna n is
$$h_{mn}{=}\beta_{mn}{-}10\alpha_{mn}\log_{10}d_{mn}{+}\xi_{mn}{+}\lambda_{mn},$$
where h_{mn} is the channel gain (in dB), β_{mn} is the average channel gain at a reference distance of 1 m (in dB), α_{mn} is the pathloss exponent, d_{mn} is the 3D Euclidean distance between position vectors $\mathbf{p}_m{=}[x_m,y_m,z_m]$ and $\mathbf{p}_n{=}[x_n,y_n,z_n]$, $\xi_{mn}{\sim}\mathcal{N}\left(0,\sigma^2_{\text{SF},mn}\right)$ is a random variable denoting the shadow-fading component with variance $\sigma^2_{\text{SF},mn}$, and $\lambda_{mn}{\sim}\mathcal{CN}\left(\mu_{\text{SSF},mn},\sigma^2_{\text{SSF},mn}\right)$ is a random variable denoting the small-scale fading component with mean $\mu_{\text{SSF},mn}$ and variance $\sigma^2_{\text{SSF},mn}$.
- Here, the 3D Euclidean distance d_{mn} between the two position vectors $\mathbf{p}_m$ and $\mathbf{p}_n$ is
$$d_{mn}{=}\sqrt{(x_m{-}x_n)^2{+}(y_m{-}y_n)^2{+}(z_m{-}z_n)^2}. \quad (10)$$
- Let $\mathcal{D}{\subseteq}\mathbb{R}^6$ be the set of all possible GN-UAV position pairs. Let R be the number of site partitions for our radio map generation procedure, i.e.,
$$\mathcal{D}{=}\bigcup_{r=1}^{R}\mathcal{D}_r;\ \ D_{r_1}{\cap}D_{r_2}{=}\phi;\ \ \forall r_1,r_2{\in}\{1,2,3,\ldots,R\};\ \ r_1{\neq}r_2. \quad (11)$$
- Thus, with this site partitioning, we can write the channel gain (in dB) between GN Tx antenna position vector $\mathbf{p}_m$ and UAV Rx antenna position vector $\mathbf{p}_n$ as
$$h_{mn}{=}h(\mathbf{p}_m,\mathbf{p}_n){=}\sum_{r=1}^{R}\left(\beta_r{-}10\alpha_r\log_{10}d(\mathbf{p}_m,\mathbf{p}_n){+}\xi_r{+}\lambda_r\right)\mathbb{1}\left\{(\mathbf{p}_m,\mathbf{p}_n){\in}\mathcal{D}_r\right\}, \quad (12)$$

where $d(\mathbf{p}_m, \mathbf{p}_n) = \|\mathbf{p}_m - \mathbf{p}_n\|_2$ is the 3D Euclidean distance between the two position vectors $\mathbf{p}_m = [x_m, y_m, z_m]$ and $\mathbf{p}_n = [x_n, y_n, z_n]$ (as described by Eq. (10). Also, for a site partition $r \in \{1, 2, 3, \ldots, R\}$, β_r is the average channel gain at a reference distance of 1 m (in dB), α_r is the pathloss exponent, $\xi_r \sim \mathcal{N}\left(0, \sigma^2_{\mathrm{SF},r}\right)$ is a random variable denoting the shadow-fading component with variance $\sigma^2_{\mathrm{SF},r}$, and $\lambda_r \sim \mathcal{CN}\left(\mu_{\mathrm{SSF},r}, \sigma^2_{\mathrm{SSF},r}\right)$ is a random variable denoting small-scale fading with mean $\mu_{\mathrm{SSF},r}$ and variance $\sigma^2_{\mathrm{SSF},r}$.

- Let $\mathcal{M} = \left\{\left(\mathbf{p}_m^{(l)}, \mathbf{p}_n^{(l)}, h_{mn}^{(l)}\right) : l = 1, 2, 3, \ldots L\right\}$ be the set of L ray-tracing measurements obtained on Wireless InSite for a subset of GN-UAV position pairs in $\mathcal{D}$, where $\mathbf{p}_m^{(l)}$ is the position vector of the GN Tx antenna, $\mathbf{p}_n^{(l)}$ is the position vector of the UAV Rx antenna, and $h_{mn}^{(l)}$ is the measured channel gain in dB, $\forall l \in \{1, 2, 3, \ldots L\}$.
- Given the channel gain description in Eq. (12) and conditioned on GN-UAV position pair $(\mathbf{p}_m, \mathbf{p}_n) \in \mathcal{D}$ belonging to site partition $\mathcal{D}_r \subset \mathcal{D}, r \in \{1, 2, 3, \ldots, R\}$, i.e., $(\mathbf{p}_m, \mathbf{p}_n) \in \mathcal{D}_r$, we can write the joint probability density function for the GN-UAV position pair and its corresponding channel gain as

$$f_r\left(\mathbf{p}_m, \mathbf{p}_n, h_{mn}\right) = \frac{1}{\sqrt{2\pi\left(\sigma^2_{\mathrm{SF},r} + \sigma^2_{\mathrm{SSF},r}\right)}} \exp\left\{-\frac{\left(h_{mn} - \beta_r + 10\alpha_r \log_{10} d_{mn} - \mu_{\mathrm{SSF},r}\right)^2}{2\left(\sigma^2_{\mathrm{SF},r} + \sigma^2_{\mathrm{SSF},r}\right)}\right\} \tag{13}$$

- Based on this conditional probability density function in Eq. (13), using the *Law of Total Probability*, we can write the general joint probability density function as

$$\begin{aligned} f\left(\mathbf{p}_m, \mathbf{p}_n, h_{mn}, \boldsymbol{\psi}\right) &= \sum_{r=1}^{R} \mathbb{P}\left(\mathbf{p}_m, \mathbf{p}_n, h_{mn} | \psi_r = 1\right) \mathbb{P}\left(\psi_r = 1\right) \\ &= \sum_{r=1}^{R} f_r\left(\mathbf{p}_m, \mathbf{p}_n, h_{mn}\right) \pi_r, \end{aligned} \tag{14}$$

where for a given pair of GN-UAV position vectors $(\mathbf{p}_m, \mathbf{p}_n) \in \mathcal{D}$, $\boldsymbol{\psi} \triangleq [\psi_1, \psi_2, \psi_3, \ldots, \psi_R]$ denotes the partition classification vector—with $\psi_r \in \{0, 1\}$ indicating whether the given GN-UAV position pair belongs to site partition $r \in \{1, 2, 3, \ldots, R\}$ ($\psi_r = 1$) or not ($\psi_r = 0$); and $\pi_r = \mathbb{P}\left(\psi_r = 1\right)$ denotes the marginal probability that the given GN-UAV position pair belongs to site partition $r \in \{1, 2, 3, \ldots, R\}$—with $\sum_{r=1}^{R} \pi_r = \sum_{r=1}^{R} \mathbb{P}\left(\psi_r = 1\right) = 1$. Next, we can rewrite Eq. (14) as

$$f\left(\mathbf{p}_m, \mathbf{p}_n, h_{mn}, \boldsymbol{\psi}\right) = \prod_{r=1}^{R} \left(f_r\left(\mathbf{p}_m, \mathbf{p}_n, h_{mn}\right)\right)^{\psi_r} \cdot \prod_{r=1}^{R} \left(\pi_r\right)^{\psi_r}. \tag{15}$$

- The set of channel parameters to be estimated & associated with the site partitions is

$$\mathcal{B} \triangleq \left\{\left(\alpha_r, \beta_r, \sigma_{\mathrm{SF},r}, \mu_{\mathrm{SSF},r}, \sigma_{\mathrm{SSF},r}, \pi_r\right) : r = 1, 2, 3, \ldots, R\right\}, \tag{16}$$

where for each site partition $r = 1, 2, 3, \ldots, R$, the channel parameters α_r, β_r, $\sigma_{\mathrm{SF},r}$, $\mu_{\mathrm{SSF},r}$, and $\sigma_{\mathrm{SSF},r}$ are defined in Eq. (12), and the marginal probability of a given GN-UAV position pair belonging to the partition r, i.e., π_r, is defined in Eq. (14).

5.2 Maximum Likelihood Estimation (MLE)

- The likelihood function for the set of measurements $\mathcal{M}$ obtained on Wireless InSite is

$$\begin{aligned} g(\mathcal{B}) = \prod_{l=1}^{L} \mathbb{P}\left(\mathbf{p}_m^{(l)}, \mathbf{p}_n^{(l)}, h_{mn}^{(l)}, \boldsymbol{\psi}^{(l)} \middle| \mathcal{B}\right) &= \prod_{l=1}^{L} f\left(\mathbf{p}_m^{(l)}, \mathbf{p}_n^{(l)}, h_{mn}^{(l)}, \boldsymbol{\psi}^{(l)} \middle| \mathcal{B}\right) \\ &= \prod_{l=1}^{L} \left[\prod_{r=1}^{R} \left(f_r\left(\mathbf{p}_m, \mathbf{p}_n, h_{mn}\right)\right)^{\psi_r^{(l)}} \cdot \prod_{r=1}^{R} \left(\pi_r\right)^{\psi_r^{(l)}}\right] \text{ from Eq. (15).} \end{aligned} \tag{17}$$

- With this definition of the likelihood function in Eq. (17), the MLE problem is

$$\begin{aligned}\underset{\mathcal{B},\{\boldsymbol{\psi}^{(l)}:l=1,2,3,\ldots,L\}}{\text{maximize}} \quad & \prod_{l=1}^{L}\left[\prod_{r=1}^{R}\Big(f_r\left(\mathbf{p}_m,\mathbf{p}_n,h_{mn}\right)\Big)^{\psi_r^{(l)}}\cdot\prod_{r=1}^{R}\left(\pi_r\right)^{\psi_r^{(l)}}\right]\\ \text{subject to} \quad & \sum_{r=1}^{R}\pi_r{=}1.\end{aligned} \tag{18}$$

- Reformulating Eq. (18) as MLE in the log-likelihood, we get

$$\begin{aligned}\underset{\mathcal{B},\{\boldsymbol{\psi}^{(l)}:l=1,2,3,\ldots,L\}}{\text{maximize}} \quad & \sum_{l=1}^{L}\sum_{r=1}^{R}\left[\psi_r^{(l)}\Bigg(\log\Big(f_r\left(\mathbf{p}_m,\mathbf{p}_n,h_{mn}\right)\Big)+\log\left(\pi_r\right)\Bigg)\right]\\ \text{subject to} \quad & \sum_{r=1}^{R}\pi_r{=}1.\end{aligned} \tag{19}$$

- Since the site partition labels $\psi_r^{(l)},\forall r{\in}\{1,2,3,\ldots,R\}$ are unknown for the samples in $\mathcal{M}$, i.e., $\left(\mathbf{p}_m^{(l)},\mathbf{p}_n^{(l)},h_{mn}^{(l)}\right),\forall l{=}1,2,3,\ldots L$, the exact expression for this log-likelihood function in Eq. (19) is unattainable. To solve this problem, we can generate a statistical estimate of $\psi_r^{(l)},\forall r{\in}\{1,2,3,\ldots,R\},\forall l{\in}\{1,2,3,\ldots,L\}$ as follows:

$$\bar{\psi}_r^{(l)}(\mathcal{B}){=}\mathbb{E}\left[\psi_r^{(l)}\middle|\mathbf{p}_m^{(l)},\mathbf{p}_n^{(l)},\mathcal{B}\right]. \tag{20}$$

  Since $\psi_r^{(l)}{\in}\{0,1\}$ is a Bernoulli random variable $\forall l{\in}\{1,2,3,\ldots,L\}$, we can write the $\mathbb{E}\left[\psi_r^{(l)}\middle|\mathbf{p}_m^{(l)},\mathbf{p}_n^{(l)},\mathcal{B}\right]$ term in Eq. (20) in terms of the success probability, and simplify it further to get the following:

$$\begin{aligned}\bar{\psi}_r^{(l)}(\mathcal{B}){=}&\mathbb{P}\left(\psi_r^{(l)}{=}1\middle|\mathbf{p}_m^{(l)},\mathbf{p}_n^{(l)},\mathcal{B}\right)\\ =&\frac{\mathbb{P}\left(\mathbf{p}_m^{(l)},\mathbf{p}_n^{(l)},h_{mn}^{(l)}\middle|\mathcal{B},\psi_r^{(l)}{=}1\right)\mathbb{P}\left(\psi_r^{(l)}{=}1\right)}{\sum_{r'=1}^{R}\mathbb{P}\left(\mathbf{p}_m^{(l)},\mathbf{p}_n^{(l)},h_{mn}^{(l)}\middle|\mathcal{B},\psi_{r'}^{(l)}{=}1\right)\mathbb{P}\left(\psi_{r'}^{(l)}{=}1\right)}\\ =&\frac{f_r\left(\mathbf{p}_m^{(l)},\mathbf{p}_n^{(l)},h_{mn}^{(l)}\middle|\mathcal{B}\right)\pi_r}{\sum_{r'=1}^{R}f_{r'}\left(\mathbf{p}_m^{(l)},\mathbf{p}_n^{(l)},h_{mn}^{(l)}\middle|\mathcal{B}\right)\pi_{r'}}.\end{aligned} \tag{21}$$

- Thus, having obtained a statistical estimate of the site partition labels, i.e., $\bar{\psi}_r^{(l)}$, $\forall r{\in}\{1,2,3,\ldots,R\},\forall l{=}1,2,3,\ldots L$, we can rewrite the MLE problem in the log-likelihood (Eq. (19)) as follows:

$$\begin{aligned}\underset{\mathcal{B},\{\boldsymbol{\psi}^{(l)}:l=1,2,3,\ldots,L\}}{\text{maximize}} \quad & \mathbb{E}\left[\sum_{l=1}^{L}\sum_{r=1}^{R}\left\{\psi_r^{(l)}\Bigg(\log\Big(f_r\left(\mathbf{p}_m,\mathbf{p}_n,h_{mn}\right)\Big)+\log\left(\pi_r\right)\Bigg)\right\}\right]\\ \text{subject to} \quad & \sum_{r=1}^{R}\pi_r{=}1;\end{aligned} \tag{22}$$

$$\begin{aligned}\underset{\mathcal{B},\{\boldsymbol{\psi}^{(l)}:l=1,2,3,\ldots,L\}}{\text{maximize}} \quad & \sum_{l=1}^{L}\sum_{r=1}^{R}\left[\bar{\psi}_r^{(l)}(\mathcal{B})\Bigg(\log\Big(f_r\left(\mathbf{p}_m,\mathbf{p}_n,h_{mn}\right)\Big)+\log\left(\pi_r\right)\Bigg)\right]\\ \text{subject to} \quad & \sum_{r=1}^{R}\pi_r{=}1.\end{aligned} \tag{23}$$

- With this reformulation in Eq. (23), we now have a convex problem in α_r, β_r, $\sigma_{\mathrm{SF},r}$, $\mu_{\mathrm{SSF},r}$, $\sigma_{\mathrm{SSF},r}$, and π_r, which can be solved via *Iterative Search* described next.

5.3 Iterative search algorithm

5.3.1 Initialization

- With the channel gains as the clustering variable, use K-means clustering to cluster the set of measurements $\mathcal{M}=\left\{\left(\mathbf{p}_m^{(l)}, \mathbf{p}_n^{(l)}, h_{mn}^{(l)}\right) : l=1,2,3,\ldots,L\right\}$ into R partitions, i.e.,

$$\begin{aligned}&\underset{\{\mathcal{R}_1,\mathcal{R}_2,\mathcal{R}_3,\ldots,\mathcal{R}_R\}}{\operatorname{argmin}} \sum_{r=1}^{R} \sum_{h_{mn}\in\mathcal{R}_r} \left\|h_{mn}-\bar{h}_{r,mn}\right\|_2^2\\&\text{subject to } R<<L \text{ and } |\mathcal{R}_1|+|\mathcal{R}_2|+|\mathcal{R}_3|+\ldots+|\mathcal{R}_R|=L,\\&\text{where } \bar{h}_{r,mn}=\frac{1}{|\mathcal{R}_r|}\sum_{h_{mn}\in\mathcal{R}_r} h_{mn}, \forall r\in\{1,2,3,\ldots,R\}.\end{aligned} \tag{24}$$

- Then, initialize $\bar{\psi}_r^{(l)}\left(\mathcal{B}^{(0)}\right)=1$, if $h_{mn}^{(l)}\in\mathcal{R}_r$, i.e., the measurement sample $\left(\mathbf{p}_m^{(l)}, \mathbf{p}_n^{(l)}, h_{mn}^{(l)}\right)\in\mathcal{M}$ belongs to the partition set $\mathcal{R}_r$, $r\in\{1,2,3,\ldots,R\}$.
- Next, using these initial estimates, compute $\alpha_r^{(0)}$, $\beta_r^{(0)}$, $\sigma_{\mathrm{SF},r}^{(0)}$, $\mu_{\mathrm{SSF},r}^{(0)}$, $\sigma_{\mathrm{SSF},r}^{(0)}$, and $\pi_r^{(0)}$, $\forall r\in\{1,2,3,\ldots,R\}$ using Eqs. (26), (27), (28), (29), and (30) outlined below.
- Starting with these estimates $\bar{\psi}_r^{(l)}\left(\mathcal{B}^{(0)}\right)$, $\forall r\in\{1,2,3,\ldots,R\}, \forall l\in\{1,2,3,\ldots,L\}$, iterate through the steps below (i=1, 2, 3, ...) until the values of the parameters α_r, β_r, $\sigma_{\mathrm{SF},r}$, $\mu_{\mathrm{SSF},r}$, $\sigma_{\mathrm{SSF},r}$, and π_r, $\forall r\in\{1,2,3,\ldots,R\}$ do not change above a pre-defined threshold between successive iterations.

5.3.2 Iteration

At the i^{th} (i=1, 2, 3, ...) iteration, with the value of the channel parameters $\mathcal{B}^{(i-1)}$:

- Compute $\bar{\psi}_r^{(l)}\left(\mathcal{B}^{(i-1)}\right)$ according to Eq. (21), $\forall r\in\{1,2,3,\ldots,R\}$, $\forall l\in\{1,2,3,\ldots,L\}$.
- The objective function of the optimization problem in Eq. (23) is

$$\begin{aligned}&\mathbb{E}\left[\log g\left(\mathcal{B}^{(i-1)}\right)\right]= \qquad (25)\\&\sum_{l=1}^{L}\sum_{r=1}^{R}\left[\bar{\psi}_r^{(l)}\left(\mathcal{B}^{(i-1)}\right)\left(\log\left\{\frac{1}{\sqrt{2\pi\left(\sigma_{\mathrm{SF},r}^2+\sigma_{\mathrm{SSF},r}^2\right)}}\exp\left\{-\frac{\left(h_{mn}^{(l)}-\beta_r+10\alpha_r\log_{10}d_{mn}^{(l)}-\mu_{\mathrm{SSF},r}\right)^2}{2\left(\sigma_{\mathrm{SF},r}^2+\sigma_{\mathrm{SSF},r}^2\right)}\right\}\right\}+\log\left(\pi_r\right)\right)\right].\end{aligned}$$

- To solve for $\left\{(\alpha_r,\beta_r) : r=1,2,3,\ldots,R\right\}$, fix all the other parameters in $\mathcal{B}^{(i-1)}$, i.e., $\sigma_{\mathrm{SF},r}=\sigma_{\mathrm{SF},r}^{(i-1)}$, $\mu_{\mathrm{SSF},r}=\mu_{\mathrm{SSF},r}^{(i-1)}$, $\sigma_{\mathrm{SSF},r}=\sigma_{\mathrm{SSF},r}^{(i-1)}$, and $\pi_r=\pi_r^{(i-1)}$, $\forall r\in\{1,2,3,\ldots,R\}$. Then, solve the following optimization problem:

$$\underset{\left\{(\alpha_r,\beta_r):r=1,2,3,\ldots,R\right\}}{\text{maximize}}\ \mathbb{E}\left[\log g\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}_{\mathrm{SF}}^{(i-1)},\boldsymbol{\mu}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\pi}^{(i-1)}\right)\right], \tag{26}$$

where $\boldsymbol{\alpha}=[\alpha_1,\alpha_2,\alpha_3,\ldots,\alpha_R]$ and $\boldsymbol{\beta}=[\beta_1,\beta_2,\beta_3,\ldots,\beta_R]$ are the optimization variables; while the other variables are fixed in this step of the algorithm $\boldsymbol{\sigma}_{\mathrm{SF}}=\boldsymbol{\sigma}_{\mathrm{SF}}^{(i-1)}=\left[\sigma_{\mathrm{SF},1}^{(i-1)},\sigma_{\mathrm{SF},2}^{(i-1)},\sigma_{\mathrm{SF},3}^{(i-1)},\ldots,\sigma_{\mathrm{SF},R}^{(i-1)}\right]$, $\boldsymbol{\mu}_{\mathrm{SSF}}=\boldsymbol{\mu}_{\mathrm{SSF}}^{(i-1)}=\left[\mu_{\mathrm{SSF},1}^{(i-1)},\mu_{\mathrm{SSF},2}^{(i-1)},\mu_{\mathrm{SSF},3}^{(i-1)},\ldots,\mu_{\mathrm{SSF},R}^{(i-1)}\right]$, $\boldsymbol{\sigma}_{\mathrm{SSF}}=\boldsymbol{\sigma}_{\mathrm{SSF}}^{(i-1)}=\left[\sigma_{\mathrm{SSF},1}^{(i-1)},\sigma_{\mathrm{SSF},2}^{(i-1)},\sigma_{\mathrm{SSF},3}^{(i-1)},\ldots,\sigma_{\mathrm{SSF},R}^{(i-1)}\right]$, and $\boldsymbol{\pi}=\boldsymbol{\pi}^{(i-1)}=\left[\pi_1^{(i-1)},\pi_2^{(i-1)},\pi_3^{(i-1)},\ldots,\pi_R^{(i-1)}\right]$.

To get $\alpha_r^{(i)}, \forall r\in\{1,2,3,\ldots,R\}$: $\dfrac{\partial}{\partial\alpha_r}\mathbb{E}\left[\log g\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}_{\mathrm{SF}}^{(i-1)},\boldsymbol{\mu}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\pi}^{(i-1)}\right)\right]=0;$

To get $\beta_r^{(i)}, \forall r\in\{1,2,3,\ldots,R\}$: $\dfrac{\partial}{\partial\beta_r}\mathbb{E}\left[\log g\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}_{\mathrm{SF}}^{(i-1)},\boldsymbol{\mu}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\pi}^{(i-1)}\right)\right]=0.$

- To solve for $\left\{\sigma_{\mathrm{SF},r}\ :\ r{=}1,2,3,\ldots,R\right\}$, fix all the other parameters in $\mathcal{B}^{(i-1)}$, i.e., $\alpha_r{=}\alpha_r^{(i-1)}$, $\beta_r{=}\beta_r^{(i-1)}$, $\mu_{\mathrm{SSF},r}{=}\mu_{\mathrm{SSF},r}^{(i-1)}$, $\sigma_{\mathrm{SSF},r}{=}\sigma_{\mathrm{SSF},r}^{(i-1)}$, and $\pi_r{=}\pi_r^{(i-1)}$, $\forall r{\in}\{1,2,3,\ldots,R\}$. Then, solve the following optimization problem:

$$\underset{\left\{\sigma_{\mathrm{SF},r}:r=1,2,3,\ldots,R\right\}}{\text{maximize}}\ \mathbb{E}\left[\log g\left(\boldsymbol{\alpha}^{(i-1)},\boldsymbol{\beta}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SF}},\boldsymbol{\mu}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\pi}^{(i-1)}\right)\right], \qquad (27)$$

where $\boldsymbol{\sigma}_{\mathrm{SF}}{=}\left[\sigma_{\mathrm{SF},1},\sigma_{\mathrm{SF},2},\sigma_{\mathrm{SF},3},\ldots,\sigma_{\mathrm{SF},R}\right]$ is the optimization variable; while the other variables are fixed in this step of the algorithm $\boldsymbol{\alpha}{=}\boldsymbol{\alpha}^{(i-1)}{=}\left[\alpha_1^{(i-1)},\alpha_2^{(i-1)},\alpha_3^{(i-1)},\ldots,\alpha_R^{(i-1)}\right]$, $\boldsymbol{\beta}{=}\boldsymbol{\beta}^{(i-1)}{=}\left[\beta_1^{(i-1)},\beta_2^{(i-1)},\beta_3^{(i-1)},\ldots,\beta_R^{(i-1)}\right]$, $\boldsymbol{\mu}_{\mathrm{SSF}}{=}\boldsymbol{\mu}_{\mathrm{SSF}}^{(i-1)}{=}\left[\mu_{\mathrm{SSF},1}^{(i-1)},\mu_{\mathrm{SSF},2}^{(i-1)},\mu_{\mathrm{SSF},3}^{(i-1)},\ldots,\mu_{\mathrm{SSF},R}^{(i-1)}\right]$, $\boldsymbol{\sigma}_{\mathrm{SSF}}{=}\boldsymbol{\sigma}_{\mathrm{SSF}}^{(i-1)}{=}\left[\sigma_{\mathrm{SSF},1}^{(i-1)},\sigma_{\mathrm{SSF},2}^{(i-1)},\sigma_{\mathrm{SSF},3}^{(i-1)},\ldots,\sigma_{\mathrm{SSF},R}^{(i-1)}\right]$, and $\boldsymbol{\pi}{=}\boldsymbol{\pi}^{(i-1)}{=}\left[\pi_1^{(i-1)},\pi_2^{(i-1)},\pi_3^{(i-1)},\ldots,\pi_R^{(i-1)}\right]$.

To get $\sigma_{\mathrm{SF},r}^{(i)}, \forall r{\in}\{1,2,3,\ldots,R\}$: $\dfrac{\partial}{\partial\sigma_{\mathrm{SF},r}}\mathbb{E}\left[\log g\left(\boldsymbol{\alpha}^{(i-1)},\boldsymbol{\beta}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SF}},\boldsymbol{\mu}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\pi}^{(i-1)}\right)\right]{=}0.$

- To solve for $\left\{\mu_{\mathrm{SSF},r}\ :\ r{=}1,2,3,\ldots,R\right\}$, fix all the other parameters in $\mathcal{B}^{(i-1)}$, i.e., $\alpha_r{=}\alpha_r^{(i-1)}$, $\beta_r{=}\beta_r^{(i-1)}$, $\sigma_{\mathrm{SF},r}{=}\sigma_{\mathrm{SF},r}^{(i-1)}$, $\sigma_{\mathrm{SSF},r}{=}\sigma_{\mathrm{SSF},r}^{(i-1)}$, and $\pi_r{=}\pi_r^{(i-1)}$, $\forall r{\in}\{1,2,3,\ldots,R\}$. Then, solve the following optimization problem:

$$\underset{\left\{\mu_{\mathrm{SSF},r}:r=1,2,3,\ldots,R\right\}}{\text{maximize}}\ \mathbb{E}\left[\log g\left(\boldsymbol{\alpha}^{(i-1)},\boldsymbol{\beta}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SF}}^{(i-1)},\boldsymbol{\mu}_{\mathrm{SSF}},\boldsymbol{\sigma}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\pi}^{(i-1)}\right)\right], \qquad (28)$$

where $\boldsymbol{\mu}_{\mathrm{SSF}}{=}\left[\mu_{\mathrm{SSF},1},\mu_{\mathrm{SSF},2},\mu_{\mathrm{SSF},3},\ldots,\mu_{\mathrm{SSF},R}\right]$ is the optimization variable; while the other variables are fixed in this step of the algorithm $\boldsymbol{\alpha}{=}\boldsymbol{\alpha}^{(i-1)}{=}\left[\alpha_1^{(i-1)},\alpha_2^{(i-1)},\alpha_3^{(i-1)},\ldots,\alpha_R^{(i-1)}\right]$, $\boldsymbol{\beta}{=}\boldsymbol{\beta}^{(i-1)}{=}\left[\beta_1^{(i-1)},\beta_2^{(i-1)},\beta_3^{(i-1)},\ldots,\beta_R^{(i-1)}\right]$, $\boldsymbol{\sigma}_{\mathrm{SF}}{=}\boldsymbol{\sigma}_{\mathrm{SF}}^{(i-1)}{=}\left[\sigma_{\mathrm{SF},1}^{(i-1)},\sigma_{\mathrm{SF},2}^{(i-1)},\sigma_{\mathrm{SF},3}^{(i-1)},\ldots,\sigma_{\mathrm{SF},R}^{(i-1)}\right]$, $\boldsymbol{\sigma}_{\mathrm{SSF}}{=}\boldsymbol{\sigma}_{\mathrm{SSF}}^{(i-1)}{=}\left[\sigma_{\mathrm{SSF},1}^{(i-1)},\sigma_{\mathrm{SSF},2}^{(i-1)},\sigma_{\mathrm{SSF},3}^{(i-1)},\ldots,\sigma_{\mathrm{SSF},R}^{(i-1)}\right]$, and $\boldsymbol{\pi}{=}\boldsymbol{\pi}^{(i-1)}{=}\left[\pi_1^{(i-1)},\pi_2^{(i-1)},\pi_3^{(i-1)},\ldots,\pi_R^{(i-1)}\right]$.

To get $\mu_{\mathrm{SSF},r}^{(i)}, \forall r{\in}\{1,2,3,\ldots,R\}$: $\dfrac{\partial}{\partial\mu_{\mathrm{SSF},r}}\mathbb{E}\left[\log g\left(\boldsymbol{\alpha}^{(i-1)},\boldsymbol{\beta}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SF}}^{(i-1)},\boldsymbol{\mu}_{\mathrm{SSF}},\boldsymbol{\sigma}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\pi}^{(i-1)}\right)\right]{=}0.$

- To solve for $\left\{\sigma_{\mathrm{SSF},r}\ :\ r{=}1,2,3,\ldots,R\right\}$, fix all the other parameters in $\mathcal{B}^{(i-1)}$, i.e., $\alpha_r{=}\alpha_r^{(i-1)}$, $\beta_r{=}\beta_r^{(i-1)}$, $\sigma_{\mathrm{SF},r}{=}\sigma_{\mathrm{SF},r}^{(i-1)}$, $\mu_{\mathrm{SSF},r}{=}\mu_{\mathrm{SSF},r}^{(i-1)}$, and $\pi_r{=}\pi_r^{(i-1)}$, $\forall r{\in}\{1,2,3,\ldots,R\}$. Then, solve the following optimization problem:

$$\underset{\left\{\sigma_{\mathrm{SSF},r}:r=1,2,3,\ldots,R\right\}}{\text{maximize}}\ \mathbb{E}\left[\log g\left(\boldsymbol{\alpha}^{(i-1)},\boldsymbol{\beta}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SF}}^{(i-1)},\boldsymbol{\mu}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SSF}},\boldsymbol{\pi}^{(i-1)}\right)\right], \qquad (29)$$

where $\boldsymbol{\sigma}_{\mathrm{SSF}}{=}\left[\sigma_{\mathrm{SSF},1},\sigma_{\mathrm{SSF},2},\sigma_{\mathrm{SSF},3},\ldots,\sigma_{\mathrm{SSF},R}\right]$ is the optimization variable; while the other variables are fixed in this step of the algorithm $\boldsymbol{\alpha}{=}\boldsymbol{\alpha}^{(i-1)}{=}\left[\alpha_1^{(i-1)},\alpha_2^{(i-1)},\alpha_3^{(i-1)},\ldots,\alpha_R^{(i-1)}\right]$, $\boldsymbol{\beta}{=}\boldsymbol{\beta}^{(i-1)}{=}\left[\beta_1^{(i-1)},\beta_2^{(i-1)},\beta_3^{(i-1)},\ldots,\beta_R^{(i-1)}\right]$, $\boldsymbol{\sigma}_{\mathrm{SF}}{=}\boldsymbol{\sigma}_{\mathrm{SF}}^{(i-1)}{=}\left[\sigma_{\mathrm{SF},1}^{(i-1)},\sigma_{\mathrm{SF},2}^{(i-1)},\sigma_{\mathrm{SF},3}^{(i-1)},\ldots,\sigma_{\mathrm{SF},R}^{(i-1)}\right]$, $\boldsymbol{\mu}_{\mathrm{SSF}}{=}\boldsymbol{\mu}_{\mathrm{SSF}}^{(i-1)}{=}\left[\mu_{\mathrm{SSF},1}^{(i-1)},\mu_{\mathrm{SSF},2}^{(i-1)},\mu_{\mathrm{SSF},3}^{(i-1)},\ldots,\mu_{\mathrm{SSF},R}^{(i-1)}\right]$, and $\boldsymbol{\pi}{=}\boldsymbol{\pi}^{(i-1)}{=}\left[\pi_1^{(i-1)},\pi_2^{(i-1)},\pi_3^{(i-1)},\ldots,\pi_R^{(i-1)}\right]$.

To get $\sigma_{\mathrm{SSF},r}^{(i)}, \forall r{\in}\{1,2,3,\ldots,R\}$: $\dfrac{\partial}{\partial\sigma_{\mathrm{SSF},r}}\mathbb{E}\left[\log g\left(\boldsymbol{\alpha}^{(i-1)},\boldsymbol{\beta}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SF}}^{(i-1)},\boldsymbol{\mu}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SSF}},\boldsymbol{\pi}^{(i-1)}\right)\right]{=}0.$

- Finally, to solve for $\left\{\pi_r : r{=}1,2,3,\ldots,R\right\}$, fix all the other parameters in $\mathcal{B}^{(i-1)}$, i.e., $\alpha_r{=}\alpha_r^{(i-1)}$, $\beta_r{=}\beta_r^{(i-1)}$, $\sigma_{\mathrm{SF},r}{=}\sigma_{\mathrm{SF},r}^{(i-1)}$, $\mu_{\mathrm{SSF},r}{=}\mu_{\mathrm{SSF},r}^{(i-1)}$, and $\sigma_{\mathrm{SSF},r}{=}\sigma_{\mathrm{SSF},r}^{(i-1)}$.

Then, solve the following Lagrangian problem:

$$\mathcal{L}\left(\pi,\varrho\middle|\mathcal{B}^{(i-1)}\right)\triangleq\mathbb{E}\left[\log g\left(\boldsymbol{\alpha}^{(i-1)},\boldsymbol{\beta}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SF}}^{(i-1)},\boldsymbol{\mu}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\sigma}_{\mathrm{SSF}}^{(i-1)},\boldsymbol{\pi}\right)\right]+\varrho\left(\sum_{r=1}^{R}\pi_r-1\right). \tag{30}$$

This Lagrangian is solved by computing the solution to the following KKT conditions:

To get $\pi_r^{(i)}, \forall r\in\{1,2,3,\ldots,R\}:\ \dfrac{\partial}{\partial\pi_r}\mathcal{L}\left(\pi,\varrho\middle|\mathcal{B}^{(i-1)}\right)=0,\ \varrho\left(\sum_{r=1}^{R}\pi_r-1\right)=0,$ and $\varrho\geq 0$.

5.3.3 Termination

- Iterate through the steps detailed in the previous subsection (i=1, 2, 3, ...) until the values of the parameters α_r, β_r, $\sigma_{\mathrm{SF},r}$, $\mu_{\mathrm{SSF},r}$, $\sigma_{\mathrm{SSF},r}$, and π_r, $\forall r\in\{1,2,3,\ldots,R\}$ do not change above a pre-defined threshold between successive iterations.
- Upon termination, we have a radio map consisting of our site partition as follows:

$$\mathcal{R}=\{\mathcal{R}_1,\mathcal{R}_2,\mathcal{R}_3,\ldots,\mathcal{R}_R\},\ \text{where} \tag{31}$$

$$\mathcal{R}_r=\left\{\left(\alpha_r,\beta_r,\sigma_{\mathrm{SF},r},\mu_{\mathrm{SSF},r},\sigma_{\mathrm{SSF},r},\pi_r,\left\{\mathbf{p}_m^{(l)},\mathbf{p}_n^{(l)},h_{mn}^{(l)}:\psi_r^{(l)}=1,l=1,2,3,\ldots,L\right\}\right),r=1,2,3,\ldots,R\right\}.$$

- Each site partition $R_r, r\in\{1,2,3,\ldots,R\}$ is represented by its characteristic channel parameters $\left(\alpha_r,\beta_r,\sigma_{\mathrm{SF},r},\mu_{\mathrm{SSF},r},\sigma_{\mathrm{SSF},r}\right)$ and the marginal probability of a GN-UAV position pair belonging to it (i.e., π_r)—along with measurement samples that belong to it, i.e., $\left\{\mathbf{p}_m^{(l)},\mathbf{p}_n^{(l)},h_{mn}^{(l)}:\psi_r^{(l)}=1,l=1,2,3,\ldots,L\right\}$.

5.3.4 Reconstruction

- Upon completing the procedure outlined above, using KNN classification, we need to dynamically classify each *new* GN-UAV positional pair where $(\mathbf{p}_m,\mathbf{p}_n)\in\mathcal{D}$ but $(\mathbf{p}_m,\mathbf{p}_n)\notin\mathcal{M}$, i.e., the GN-UAV positional pair is in the overall deployment site but not among the subset of positional pairs that constitute the ray-tracing measurements.
- KNN classification:
  - We define the index set of S-nearest neighbors of $(\mathbf{p}_m,\mathbf{p}_n)\in\mathcal{D}$ as follows:

$$\mathcal{S}(\mathbf{p}_m,\mathbf{p}_n)=\underset{\tilde{\mathcal{S}}\subseteq\{1,2,3,\ldots,L\}:|\tilde{\mathcal{S}}|=S}{\operatorname{argmin}}\ \sum_{l\in\tilde{\mathcal{S}}}\left[\left\|\mathbf{p}_m-\mathbf{p}_m^{(l)}\right\|_2+\left\|\mathbf{p}_n-\mathbf{p}_n^{(l)}\right\|_2\right]$$
$$\text{subject to } R<<L \text{ and } S<<L. \tag{32}$$

  - Then, the site partition which $(\mathbf{p}_m,\mathbf{p}_n)\in\mathcal{D}$ belongs to is given by

$$\hat{\boldsymbol{\psi}}(\mathbf{p}_m,\mathbf{p}_n)=\mu\sum_{l\in\mathcal{S}(\mathbf{p}_m,\mathbf{p}_n)}\mathcal{K}\left(\mathbf{p}_m,\mathbf{p}_n,\mathbf{p}_m^{(l)},\mathbf{p}_n^{(l)}\right)\bar{\boldsymbol{\psi}}^{(l)},\ \text{where} \tag{33}$$

$$\mathcal{K}\left(\mathbf{p}_m,\mathbf{p}_n,\mathbf{p}_m^{(l)},\mathbf{p}_n^{(l)}\right)\triangleq\exp\left\{-\frac{\left(\left\|\mathbf{p}_m-\mathbf{p}_m^{(l)}\right\|_2+\left\|\mathbf{p}_n-\mathbf{p}_n^{(l)}\right\|_2\right)^2}{\varphi}\right\}\ \text{is the kernel function,}$$

$\varphi>0$ is the kernel parameter, and $\mu>0$ is chosen such that $\sum_{r=1}^{R}\hat{\psi}_r(\mathbf{p}_m,\mathbf{p}_n)=1$.

- Lastly, we get the channel gain (in dB) for this *new* or *unseen* GN-UAV positional pair $(\mathbf{p}_m, \mathbf{p}_n) \in \mathcal{D}$ as follows:

$$\hat{h}_{mn}(\mathbf{p}_m, \mathbf{p}_n) = \sum_{r=1}^{R} \left(\beta_r - 10\alpha_r \log_{10} d_{mn} + \xi_r + \lambda_r\right) \hat{\psi}_r(\mathbf{p}_m, \mathbf{p}_n), \tag{34}$$

where $d_{mn} = \|\mathbf{p}_m - \mathbf{p}_n\|_2$, $\xi_r \sim \mathcal{N}\left(0, \sigma^2_{\mathrm{SF},r}\right)$, and $\lambda_r \sim \mathcal{CN}\left(\mu_{\mathrm{SSF},r}, \sigma_{\mathrm{SSF},r}\right)$.

6 Solution: DPDP solved via Hierarchical RL

6.1 Problem setup

- Let ζ be the fleet orchestration policy that defines the GN request scheduling & UAV association mechanism, the corresponding UAV positioning along with its constituent energy-conscious 3D trajectory, and the re-charging strategy for the UAVs considering long-term network dynamics.
- Here, in this formulation, we assume that the GN requests (which are generated as per a Poisson arrival process with rate Λ [requests per unit time]) are *cached* in a request buffer at the centralized operations hub, and *released* in batches for them to be served by the fleet. This fits into our Hierarchical Reinforcement Learning (HRL) formulation, discussed later in this section.
- First, we define a *policy interval* as the mission time-interval between the UAVs in the *idle phase* waiting for the batch of requests to be released, the UAVs entering the *service phase*, i.e., starting to serve their assigned requests, the UAVs completing the *service phase* after fully servicing their assigned requests, and the UAVs re-entering the *idle phase* waiting for the next batch of requests to be released.
- Let Δ_η be the duration of the η^{th} policy interval, $\eta \in \{1, 2, 3, \ldots\}$. Let $E_{u,\eta}$ be the mobility energy consumption of UAV $u \in \mathcal{U}$ in this policy interval.
- Under the GN request model described earlier in this document, we define the average per-UAV reward under this policy ζ as follows:

$$\bar{\Omega}_\zeta \triangleq \lim_{\eta \to \infty} \mathbb{E}\left[\frac{1}{U(\eta-1)} \sum_{\iota=1}^{\eta-1} \sum_{g \in \mathcal{G}_\iota} \sum_{u=1}^{U} \vartheta_{gu,\iota} \Omega_{gu,\iota}\right], \tag{35}$$

where $\mathcal{G}_\iota$ is the set of GNs whose requests have been *released* in policy interval ι; $\vartheta_{gu,\iota} \in \{0, 1\}$ is a binary variable denoting if the request from $g \in \mathcal{G}_\iota$ is assigned to UAV $u \in \mathcal{U}$; and $\Omega_{gu,\iota} = \chi_g \gamma_g^{\left(\delta_{gu,\iota} - \delta_{g,\max}\right)}$ is the reward obtained by UAV $u \in \mathcal{U}$ for serving a request from GN $g \in \mathcal{G}_\iota$ in $\delta_{gu,\iota}$ time—with the request priority value being χ_g, the maximum latency being $\delta_{g,\max}$, the data payload size being ν_g, and the post-deadline discount factor being γ_g.
- Similarly, under the UAV energy model outlined earlier in this document, we define the average per-UAV energy consumption under this policy ζ as follows:

$$\bar{E}_\zeta \triangleq \lim_{\eta \to \infty} \mathbb{E}\left[\frac{1}{U(\eta-1)} \sum_{\iota=1}^{\eta-1} \sum_{u=1}^{U} \left[E_{u,\iota} - \Xi_{u,\iota} \tau_{u,\mathrm{chrg},\iota} \Upsilon_u\right]\right], \tag{36}$$

where $E_{u,\iota}$ is the energy consumed by UAV $u \in \mathcal{U}$ in policy interval ι; $\Xi_{u,\iota} \in \{0, 1\}$ is a binary variable denoting if the UAV $u \in \mathcal{U}$ is sent into the depot for charging in this policy interval ι; $\tau_{u,\mathrm{chrg},\iota}$ is the amount of time spent charging by UAV $u \in \mathcal{U}$ in policy interval ι; and Υ_u is the charge rate for UAV $u \in \mathcal{U}$, i.e., the rate at which energy is delivered to the UAV's battery by the charging stations at the depot.
- Let the set containing all the information pertaining to the UAV fleet be defined as

$$\mathcal{I}_U \triangleq \left\{\mathcal{U}, \{A_u\}_{u \in \mathcal{U}}, \left\{E_{u,\min}, E_{u,\mathrm{avg}}, E_{u,\max}\right\}_{u \in \mathcal{U}}, \left\{v_{u,\min}, v_{u,\max}, a_{u,\min}, a_{u,\max}\right\}_{u \in \mathcal{U}}, \{\Upsilon_u\}_{u \in \mathcal{U}}\right\}.$$

- Let the set containing all the information pertaining to the GNs be defined as

$$\mathcal{I}_G \triangleq \left\{ \mathcal{G}, \{A_g\}_{g\in\mathcal{G}}, \{\mathbf{p}_g\}_{g\in\mathcal{G}} \right\}.$$

- Let the complete set of channel parameters for the site partitions (and their associated marginal probabilities), obtained via the ray-tracing augmented channel estimation procedure detailed earlier in this document, be defined as follows:

$$\mathcal{B}^* \triangleq \left\{ \left(\alpha_r^*, \beta_r^*, \sigma_{\mathrm{SF},r}^*, \mu_{\mathrm{SSF},r}^*, \sigma_{\mathrm{SSF},r}^*, \pi_r^*\right) : r{=}1,2,3,\ldots,R \right\}.$$

6.2 Optimization problem

(P.0) $\underset{\zeta}{\text{maximize}}\ \bar{\Omega}_\zeta\left(\mathcal{I}_U, \mathcal{I}_G, \mathcal{B}^*\right)$

$$\text{subject to } \bar{E}_\zeta\left(\mathcal{I}_U, \mathcal{I}_G, \mathcal{B}^*\right) \leq \frac{1}{U}\sum_{u=1}^{U} E_{u,\mathrm{avg}} \tag{C.1}$$

$$0{\leq}x{\leq}x_{\max},\ 0{\leq}y{\leq}y_{\max},\ 0{\leq}z{\leq}z_{\max};\ \Delta_x{>}0,\ \Delta_y{>}0,\ \Delta_z{>}0 \tag{C.2}$$

$$|\mathcal{U}|{=}U\ (\text{fixed});\ |\mathcal{G}|{=}G\ (\text{fixed});\ \mathcal{B}^*\ (\text{pre-determined}) \tag{C.3}$$

$$\begin{aligned}&\mathbf{p}_u(0){\in}\mathcal{F}\ (\text{depot});\ \mathbf{p}_{u_1}(t){\neq}\mathbf{p}_{u_2}(t)\ (\text{collision-avoidance});\ \mathbf{p}_u(t){\notin}\mathcal{Z}\ (\text{NFZs});\\ &\forall u, u_1, u_2{\in}\mathcal{U},\ u_1{\neq}u_2,\ \forall t{\in}\left[t_{u,\mathrm{init},\eta}, t_{u,\mathrm{term},\eta}\right],\ \forall \eta{\in}\{1,2,3,\ldots\}\end{aligned} \tag{C.4}$$

$$\mathbf{p}_g(t){=}\mathbf{p}_g\ (\text{fixed}),\ \forall g{\in}\mathcal{G},\ \forall t{\in}\left[t_{u,\mathrm{init},\eta}, t_{u,\mathrm{term},\eta}\right],\ \forall \eta{\in}\{1,2,3,\ldots\} \tag{C.5}$$

$$\begin{aligned}&v_{u,\min}{\leq}v_{u,\mathrm{horz}}(t){\leq}v_{u,\max},\ v_{u,\min}{\leq}v_{u,\mathrm{vert}}(t){\leq}v_{u,\max};\\ &a_{u,\min}{\leq}a_{u,\mathrm{horz}}(t){\leq}a_{u,\max},\ a_{u,\min}{\leq}a_{u,\mathrm{vert}}(t){\leq}a_{u,\max};\\ &\forall u{\in}\mathcal{U},\ \forall t{\in}\left[t_{u,\mathrm{init},\eta}, t_{u,\mathrm{term},\eta}\right],\ \forall \eta{\in}\{1,2,3,\ldots\}\end{aligned} \tag{C.6}$$

$$\vartheta_{gu,\eta}{\in}\{0,1\};\ \sum_{u=1}^{U}\vartheta_{gu,\eta}{=}1;\ \forall u{\in}\mathcal{U},\ \forall g{\in}\mathcal{G}_\eta,\ \mathcal{G}_\eta{\subseteq}\mathcal{G},\ \forall \eta{\in}\{1,2,3,\ldots\} \tag{C.7}$$

$$0{\leq}\sum_{g\in\mathcal{G}_\eta}\vartheta_{gu,\eta}{\leq}|\mathcal{G}_\eta|,\ \forall u{\in}\mathcal{U},\ \mathcal{G}_\eta{\subseteq}\mathcal{G},\ \forall \eta{\in}\{1,2,3,\ldots\} \tag{C.8}$$

$$\begin{aligned}&\int_{t_{u,\mathrm{init},\eta}}^{t_{u,\mathrm{term},\eta}} R_{gu}\Big(\mathbf{p}_u(t), \mathbf{p}_g, \mathcal{B}^*, \mathbf{\Phi}_g(t)\Big)dt{\geq}\nu_g;\\ &\forall u{\in}\mathcal{U},\ \forall g{\in}\mathcal{G}_\eta\ \text{s.t.}\ \vartheta_{gu,\eta}{=}1,\ \mathcal{G}_\eta{\subseteq}\mathcal{G},\ \forall \eta{\in}\{1,2,3,\ldots\}\end{aligned} \tag{C.9}$$

$$\begin{aligned}&\Xi_{u,\eta}{\in}\{0,1\};\ 0{<}\tau_{u,\mathrm{chrg},\eta}{<}\Delta_\eta;\ \Upsilon_{u,\eta}{=}\Upsilon_u\ (\text{fixed});\\ &E_{u,\min}{\leq}E_{u,\mathrm{rem}}(t){\leq}E_{u,\max};\ E_{u,\min}{\leq}E_{u,\eta}{-}\Xi_{u,\eta}\tau_{u,\mathrm{chrg},\eta}\Upsilon_u{\leq}E_{u,\max};\\ &\forall u{\in}\mathcal{U},\ \forall t{\in}\left[t_{u,\mathrm{init},\eta}, t_{u,\mathrm{term},\eta}\right],\ \forall \eta{\in}\{1,2,3,\ldots\}.\end{aligned} \tag{C.10}$$

6.3 DPDP formulation via HRL

- The Lagrangian for problem **(P.0)** is

$$\textbf{(L.0)} \ \underset{\zeta}{\text{maximize}} \ \bar{\Omega}_{\zeta}\left(\mathcal{I}_U, \mathcal{I}_G, \mathcal{B}^*\right) + \epsilon\Big(\bar{E}_{\zeta}\left(\mathcal{I}_U, \mathcal{I}_G, \mathcal{B}^*\right) - \frac{1}{U}\sum_{u=1}^{U} E_{u,\text{avg}}\Big)$$
$$\text{subject to constraints C.2} - \text{C.10 and } \epsilon \geq 0 \text{ (dual variable)}. \quad (37)$$

- This problem can be approached as a Dynamic Pickup and Delivery Problem (DPDP) wherein the GN requests (generated according to a Poisson arrival process) are *cached* and *released* dynamically: this *cache-and-release* is optimized via a Deep Q-Network (DQN). Subsequently, when a batch of GN requests is *released*, the resultant problem is a Static Pickup and Delivery Problem (Static PDP), solved via a Cross-Layer Multiple Traveling Salesman Problem (MTSP) approach.
- This two-stage process can be modeled with a Hierarchical Reinforcement Learning (HRL) construction, i.e., the *cache-and-release* process is optimized by an Upper Agent (UA) involving a DQN, while the resultant Static PDP is solved via a Cross-Layer Multiple Traveling Salesman Problem (MTSP) approach.

6.3.1 Upper Agent (UA)

The Markov Decision Process (MDP) underlying the operations of the Upper Agent (UA) in our DPDP HRL formulation involves the following states, actions, rewards, and solution process.

- State: $s_{\text{UA}} \triangleq \Big[N_{\text{cache}}, N_{\text{UAVs}}, \mathcal{J}_U, \mathcal{J}_G\Big]$, where N_{cache} is the number of requests in the cache (i.e., request buffer), N_{UAVs} is the number of available vehicles, $\mathcal{J}_U$ is the set containing all the information pertaining to the available UAVs (i.e., their positions & trajectories and their remaining energies), and $\mathcal{J}_G$ is the set containing information pertaining to the GN requests cached in the buffer (i.e., the GN positions and the request headers).
- Action: $a_{\text{UA}} \in \{0, 1\}$ is a binary variable indicating whether the agent should continue to cache requests ($a_{\text{UA}}=0$) or if the agent should release all the requests in the current buffer ($a_{\text{UA}}=1$).
- If the requests in the cache are not released, the reward for the upper agent is 0, i.e., $r_{\text{UA}}=0$; else, if all the requests in the current buffer are released, the resultant static PDP in policy interval η is solved by the Lower Agent (LA) via a cross-layer MTSP approach (policy ζ_{LA}) to get a reward

$$r_{\text{UA}} = \bar{\Omega}_{\zeta_{\text{LA}}}\left(\mathcal{I}_U, \mathcal{I}_G, \mathcal{B}^*, \mathcal{J}_{U,\eta}, \mathcal{J}_{G,\eta}\right) + \quad (38)$$
$$\epsilon\left(\bar{E}_{\zeta_{\text{LA}}}\left(\mathcal{I}_U, \mathcal{I}_G, \mathcal{B}^*, \mathcal{J}_{U,\eta}, \mathcal{J}_{G,\eta}\right) - \frac{1}{U}\sum_{u=1}^{U} E_{u,\text{avg}}\right)$$

- This MDP's solution process involves the use of a DQN to optimize the Temporal Difference (TD) cost function:

$$\mathcal{L}_{\text{UA}}\left(\Theta_{\text{UA},i}\right) = \quad (39)$$
$$\underset{(s_{\text{UA}}, a_{\text{UA}}, r_{\text{UA}}, s'_{\text{UA}}) \sim \text{U}(\mathcal{E})}{\mathbb{E}}\left[\left(r_{\text{UA}} + \gamma_{\text{UA}} \max_{a'_{\text{UA}}}\Big\{Q\big(s'_{\text{UA}}, a'_{\text{UA}}; \Theta^{-}_{\text{UA},i}\big) - Q\big(s_{\text{UA}}, a_{\text{UA}}; \Theta_{\text{UA},i}\big)\Big\}\right)^2\right],$$

where $\Theta_{\text{UA},i}$ denotes the Q-network parameters in training iteration i, $\Theta^{-}_{\text{UA},i}$ denotes the target network parameters in training iteration i, $\text{U}(\mathcal{E})$ denotes uniformly sampling experiences from the replay buffer ($\mathcal{E}$), and γ_{UA} is the UA's discount factor for this learning process.

6.3.2 Lower Agent (LA)

Upon releasing all the requests in the cache, in policy interval η, the Lower Agent (LA) solution process involves employing a cross-layer MTSP approach to solve the resultant static PDP, and thus obtain the reward for that policy interval: see our ICC Workshops 2024 paper for more detail on the LA's operations.