



UWL REPOSITORY

repository.uwl.ac.uk

XAI-guided optimizer distillation for real-time UAV relay positioning in Terahertz
6G low-altitude ad hoc networks

Khan, Abdul Manan (2026) XAI-guided optimizer distillation for real-time UAV relay positioning in
Terahertz 6G low-altitude ad hoc networks. *Ad Hoc Networks*, 194. ISSN 1570-8705

<https://doi.org/10.1016/j.adhoc.2026.104389>

This is the Published Version of the final output.

UWL repository link: <https://repository.uwl.ac.uk/id/eprint/15440/>

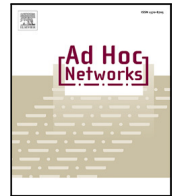
Alternative formats: If you require this document in an alternative format, please contact:
open.research@uwl.ac.uk

Copyright: Creative Commons: Attribution 4.0

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy: If you believe that this document breaches copyright, please contact us at open.research@uwl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Rights Retention Statement:



XAI-guided optimizer distillation for real-time UAV relay positioning in terahertz 6G low-altitude ad hoc networks^{☆,☆☆}

Abdul Manan KHAN[✉]

University of West London, Ealing Campus London, United Kingdom of Great Britain and Northern Ireland

ARTICLE INFO

Keywords:

UAV relay positioning
Low-altitude aerial ad hoc network
Terahertz 6G communications
Intelligent networking
Explainable AI
Optimizer distillation
Set transformer
Resilient autonomous systems
Low-altitude economy

ABSTRACT

Intelligent and resilient networking for the low-altitude economy requires autonomous unmanned aerial vehicle (UAV) relay positioning that adapts in real time to changing user topologies. In terahertz (THz) sixth-generation (6G) low-altitude aerial ad hoc networks (LAAHNs), this demands neural surrogates combining millisecond-scale inference with near-optimal placement accuracy. This paper presents a thirteen-method comparison — spanning heuristics, deep reinforcement learning (DRL), DeepSets, a multilayer perceptron (MLP), a Set Transformer, a fine-tuned language model, and a differential-evolution oracle — evaluated on 200 shared scenarios at 300 GHz. To improve the best-performing architecture, this paper introduces an *explainable artificial intelligence (XAI)-guided optimizer distillation pipeline*: a Set Transformer is first trained on 10,000 globally optimal solutions, then three explainability methods (attention heatmaps, Integrated Gradients, GradientShap) identify user spatial coordinates as the dominant features, and a streamlined variant, *ST-Lite*, is retrained on those features alone. ST-Lite achieves the highest surrogate throughput: 194.2 Mbps (96.1% of the evolutionary oracle) at 1.2 ms—surpassing both the MLP (182.1 Mbps) and the language model (185.4 Mbps at 822 ms). ST-Lite also leads under building obstructions (156.8 vs. 148.1 Mbps for the MLP) and vehicular mobility up to 120 km/h, where millisecond-scale latency enables 100 ms-cadence repositioning that outperforms every non-oracle method. Multi-seed validation across five random seeds confirms reproducibility. Under 3rd Generation Partnership Project (3GPP) TR38.901 fading, all methods compress to 72–79 Mbps; retraining ST-Lite directly on 3GPP-optimized data yields only +0.9% improvement, confirming that stochastic fading — not training-channel mismatch — is the binding constraint. A two-relay extension further shows the surrogate composes to multiple relays, recovering 96.2% of a joint two-relay oracle at millisecond latency, and robustness is corroborated on a structured 3D city-grid model. All models run as Robot Operating System 2 (ROS2) Jazzy nodes on a consumer GPU.

1. Introduction

The low-altitude economy (LAE) — last-mile logistics, urban air mobility, industrial inspection, and emergency response — demands intelligent and resilient communication infrastructure that enables autonomous unmanned aerial vehicle (UAV) fleets to operate as networked agents [1]. Existing sub-6 GHz bands cannot sustain the aggregate data rates of autonomous navigation, video telemetry, and real-time fleet coordination at scale, motivating the design of scalable aerial networking services that adapt to dynamic user topologies.

Sixth-generation (6G) terahertz (THz) bands above 100 GHz offer tens of gigahertz of bandwidth [2,3], but propagation at 300 GHz caps reliable range to a few hundred meters [4]. UAV relays bridge this gap by hovering between a ground base station (BS) and dispersed

terminals [5], forming a low-altitude aerial ad hoc network (LAAHN). The central challenge is *where to position the relay* to jointly maximize throughput, fairness, and quality-of-service (QoS) coverage in real time as terminals move.

Classical heuristics place the relay at a weighted centroid [5]—fast but brittle under irregular topologies. Global solvers like differential evolution (DE) [6] approach the Pareto front, yet their hundred-millisecond runtimes clash with mobile repositioning cadences. Neural surrogates — multilayer perceptrons (MLPs), Set Transformers [7], fine-tuned language models [8] — offer millisecond inference, while deep reinforcement learning (DRL) [9] requires millions of interactions and produces opaque policies. A critical question is *which input features actually matter* for surrogate accuracy. This paper introduces an *explainable artificial intelligence (XAI)-guided optimizer distillation pipeline*:

[☆] This article is part of a Special issue entitled: 'Intelligent and Resilient Networking' published in Ad Hoc Networks.

^{☆☆} Supplementary video available with the online version of this paper.

E-mail address: khanabd@uwl.ac.uk.

a Set Transformer is trained on DE solutions, post-hoc explainability identifies dominant features, and a streamlined variant (ST-Lite) re-trained on those features alone surpasses all alternatives. The pipeline extends knowledge distillation [10] and learning-to-optimize [11] with a principled feature-selection step.

This paper addresses this gap through five contributions:

1. **XAI-guided optimizer distillation pipeline.** A train–explain–prune methodology: a Set Transformer [7] trained on 10,000 DE solutions is analyzed via three XAI methods (attention heatmaps, Integrated Gradients [12], GradientShap [13]), which identify spatial coordinates as dominant, driving the creation of ST-Lite.
2. **ST-Lite surrogate.** ST-Lite achieves 194.2Mbps (96.1% of the oracle) at 1.2 ms—the highest surrogate throughput among thirteen methods, surpassing the MLP (182.1 Mbps) and language model (185.4 Mbps at 822 ms). Five-seed validation confirms reproducibility (coefficient of variation [CV] = 0.4%).
3. **Explainability and feature ablation.** Three XAI methods consistently rank spatial coordinates as dominant; feature ablation provides causal confirmation that removing non-spatial inputs *improves* accuracy. Four classical methods (LASSO, MI, F-test, permutation importance) corroborate the ranking.
4. **Thirteen-method comparison.** Heuristics, DRL, DeepSets, MLP, Set Transformer, ST-Lite, language model, and DE oracle evaluated on 200 shared scenarios with Bonferroni-corrected *t*-tests and bootstrap confidence intervals (CIs), plus stress tests under 3GPP fading, obstructions, and mobility (0–120 km/h).
5. **ROS2 deployment.** All surrogates run as ROS2 Jazzy nodes [14] on a consumer GPU (RTX 4050).

The train–explain–prune–retrain pipeline generalizes to any setting where solver outputs supervise a neural surrogate and post-hoc explainability can identify redundant inputs. The remainder of this paper is organized as follows: Section 2 surveys related work; Section 3 describes the system model; Section 4 details the methodology; Section 5 presents the thirteen-method comparison; Section 6 discusses limitations; and Section 7 concludes.

2. Related work

Saad et al. [2] articulated the 6G vision with THz bands as a pillar technology. Rappaport et al. [3] and Akyildiz et al. [4,15] surveyed opportunities above 100 GHz and refined propagation models. Jornet and Akyildiz [16] provided a radiative-transfer capacity analysis, Han et al. [17] catalogued measurement campaigns, and Gao et al. [18] reported THz UAV-to-ground multiple-input multiple-output (MIMO) measurements confirming the line-of-sight (LoS) dominance assumed here.

Mozaffari et al. [19] provided a tutorial on UAV wireless networks. Zeng et al. [1] established core research threads and formalized throughput maximization for a UAV mobile relay [5]; Zeng and Zhang [20] extended this to energy-efficient trajectory design. Geraci et al. [21] surveyed 5G-to-6G UAV communications, identifying relay positioning as a key open problem.

Deep reinforcement learning (DRL) has become the default tool for communication-aware UAV trajectories [22]. Challita et al. [23] applied deep RL to interference management for cellular-connected UAVs, and Bithas et al. [24] surveyed the broader landscape of machine-learning techniques for UAV communications. Kong et al. [9] used multi-agent DRL for target assignment amid dynamic obstacles, Westheider et al. [25] demonstrated an adaptive multi-UAV coverage framework, and Li et al. [26] applied AI-driven coverage planning for UAV bridge inspection. Machine learning now underpins a far wider Internet-of-Drones (IoD) stack than trajectory control alone: Heidari et al. [27] systematically reviewed learning-based IoD deployments and their open problems, and Maleki et al. [28] optimized

UAV routing with a shuffled frog-leaping metaheuristic. A parallel thread hardens the IoD edge—deep-learning anomaly detection [29], blockchain-coupled intrusion detection built on radial-basis-function networks [30], and blockchain-assisted DRL computation offloading for multi-user settings [31]. These works address routing, security, and resource offloading; the relay-positioning surrogate developed here is orthogonal and complementary, supplying the physical-layer placement decision on which such higher-layer IoD services ultimately rest.

Wu et al. [32] adapted pre-trained large language models (LLMs) for viewport prediction and scheduling at SIGCOMM 2024. Zhou et al. [33] surveyed prompt-engineering for network control; Hang et al. [34] reviewed LLM-in-the-loop pipelines; Tian et al. [35] surveyed LLM-empowered multi-UAV systems; and Yang et al. [36] demonstrated LLMs as black-box optimizers via prompt-based search. The proposed approach treats the language model as a function approximator trained on solver outputs—closer to learning-to-optimize [37] than prompt-based decision-making.

Hinton et al. [10] introduced knowledge distillation (KD); Gou et al. [38] surveyed developments across architectures, and Yao et al. [39] extended KD to federated Internet of Things (IoT) drone networks. This work adapts this paradigm with continuous relay coordinates as distillation targets. The learning-to-optimize (L2O) literature [11,40] trains networks to replace iterative solvers. Surrogate-assisted evolutionary computation [41,42] builds cheap fitness approximations. Kang and Kang [43] showed KD can be effective with limited regression data—relevant to our 10,000-sample budget. The proposed approach differs: rather than accelerating the solver’s inner loop, it *replaces* the solver entirely with a single-pass neural surrogate trained on converged outputs, adding an XAI-guided feature-selection step that prior work lacks.

3. System model

3.1. Network topology

A fixed BS at $\mathbf{p}_{BS} = (0, 0, 30)\text{m}$ serves $N \in \{3, \dots, 7\}$ ground users at positions $\{\mathbf{p}_i\}_{i=1}^N$ through a UAV relay at $\mathbf{p}_{UAV} \in \mathcal{X} = [0, 100]^2 \times [10, 40]\text{m}$. Each user i has a QoS requirement $R_i^{\text{req}} \sim \text{Uniform}[10, 50]\text{Mbps}$. The present work extends the comparison to Set Transformers and XAI-guided feature selection.

3.2. THz channel model

The aggregate path loss at carrier frequency $f_c = 300\text{GHz}$ comprises free-space path loss (FSPL) and molecular absorption:

$$L(d) = \underbrace{20 \log_{10}(d) + 20 \log_{10}(f_c) + 92.45}_{\text{FSPL (dB)}} + \underbrace{\alpha_{\text{abs}} \cdot d}_{\text{Absorption (dB)}} \quad (1)$$

where d is the link distance in km and $\alpha_{\text{abs}} = 10\text{dB/km}$ at 300 GHz (296 K, 50% RH) [16,17]. Small-scale fading and LoS/non-line-of-sight (NLoS) switching are omitted: all methods see the same deterministic channel, so systematic bias cancels in pairwise comparisons. Section 5.10 validates performance under 3GPP TR 38.901 fading. The received signal-to-noise ratio (SNR) is:

$$\text{SNR}(d, P_t) = P_t - L(d) - N_0 \quad (2)$$

with P_t in dBm and noise floor $N_0 = -174 + 10 \log_{10}(B) + F\text{dBm}$ ($B = 10\text{GHz}$, $F = 10\text{dB}$). The decode-and-forward relay’s effective SNR for user i is:

$$\text{SNR}_i^{\text{eff}} = \min(\text{SNR}(d_{BS,UAV}, P_{BS}), \text{SNR}(d_{UAV,i}, P_{UAV})) \quad (3)$$

where $d_{BS,UAV}$ and $d_{UAV,i}$ are the BS–relay and relay–user distances ($P_{BS} = 30\text{dBm}$, $P_{UAV} = 20\text{dBm}$). The achievable rate $R_i = B \log_2(1 + 10^{\text{SNR}_i^{\text{eff}}/10})$, capped at 10 Gbps. Three metrics define placement quality: aggregate throughput $T = \sum_i R_i$; Jain’s fairness index [44] $J = (\sum_i R_i)^2 / (N \sum_i R_i^2)$; and coverage rate $C = N^{-1} \sum_i \mathbf{1}[R_i \geq R_i^{\text{req}}]$.

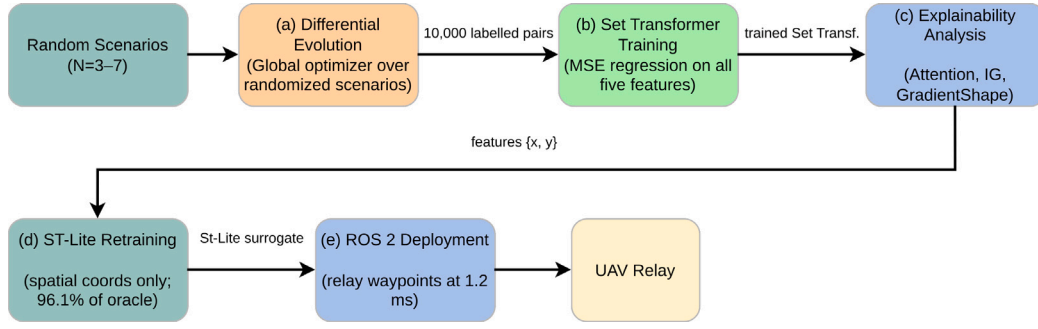


Fig. 1. End-to-end explainable-AI-guided optimizer distillation pipeline. Stage (a): differential evolution generates optimal relay positions for randomized scenarios. Stage (b): a Set Transformer is trained on the resulting dataset via mean-squared-error regression. Stage (c): explainability analysis identifies dominant input features. Stage (d): a streamlined variant (ST-Lite) is retrained on spatial coordinates alone, yielding 96.1% of the oracle throughput. Stage (e): the trained model is deployed as a Robot Operating System 2 (ROS 2) node, publishing relay waypoints to the unmanned aerial vehicle flight controller at millisecond latency.

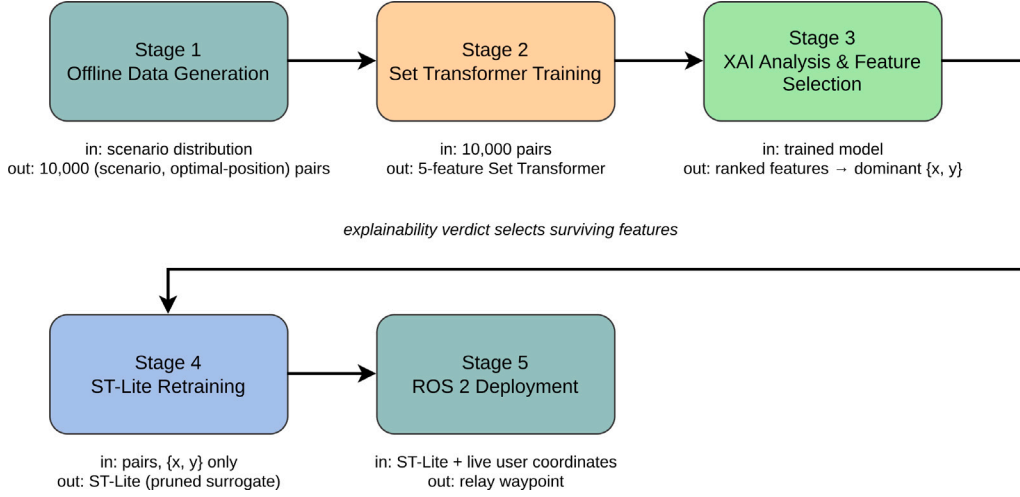


Fig. 2. Step-by-step flowchart of the optimizer-distillation pipeline. Each of the five stages is annotated with its input and output; arrows show the feed-forward data flow, and the wrap-around edge from Stage 3 to Stage 4 is the single feedback path, in which the Stage 3 explainability verdict selects the feature subset that survives into the Stage 4 ST-Lite retraining.

3.3. Multi-objective formulation

The relay positioning problem is cast as:

$$\mathbf{p}_{\text{UAV}}^* = \arg \max_{\mathbf{p}_{\text{UAV}} \in \mathcal{X}} S(\mathbf{p}_{\text{UAV}}) \quad (4)$$

where the composite score function is:

$$S(\mathbf{p}_{\text{UAV}}) = w_1 \cdot T + w_2 \cdot J \cdot 10^3 + w_3 \cdot C \cdot 10^3 \quad (5)$$

with weights $w_1 = 0.6$, $w_2 = 0.3$, $w_3 = 0.1$. The scaling factors 10^3 on the fairness and coverage terms ensure commensurate magnitudes with the throughput component, which is measured in Mbps.

4. Methodology

Fig. 1 illustrates the end-to-end pipeline, which consists of five stages: (a) offline data generation via global optimization, (b) Set Transformer training on all input features, (c) XAI analysis and feature selection, (d) ST-Lite retraining on selected features, and (e) online deployment within a ROS2 control loop. Algorithm 1 formalizes the complete procedure: it first collects K globally optimal relay positions from differential evolution, trains a Set Transformer on the full feature set, applies three explainability methods to rank input features, prunes the architecture to spatial coordinates alone, and deploys the resulting ST-Lite model as a real-time ROS2 controller. This train-explain-prune-retrain loop is central to the paper's contribution

because it replaces manual feature engineering with a data-driven, interpretable selection step—yielding a surrogate that is simultaneously faster, more accurate, and more parsimonious than the original. Fig. 2 presents the pipeline as a step-by-step flowchart, annotating each stage with its input and output. In this figure, Stage 1 consumes the scenario distribution and emits 10,000 (scenario, optimal-position) pairs; Stage 2 consumes those pairs and emits a trained five-feature Set Transformer; Stage 3 consumes that model and emits a ranked feature importance, from which the dominant subset $\{x, y\}$ is read off; Stage 4 consumes the same 10,000 pairs but keeps only $\{x, y\}$ per user and emits ST-Lite; and Stage 5 consumes ST-Lite plus live user coordinates and emits a relay waypoint every control cycle. Each stage's output is the next stage's input, so the entire path from solver to deployed controller is a single feed-forward chain with one feedback edge—the explainability verdict in Stage 3 that determines which inputs survive into Stage 4.

4.1. Stage 1: Optimizer-in-the-loop data generation

For each of 10,000 random scenarios ($N \sim \text{Uniform}\{3, \dots, 7\}$, user positions with each coordinate $x, y \sim \mathcal{U}[20, 80]\text{m}$ (the central $60 \times 60\text{m}$ of the $100 \times 100\text{m}$ area)), `scipy.optimize.differential_evolution` [6] maximizes (4) with 200 generations, tolerance 10^{-6} , and L-BFGS-B polishing (100–300 ms per solve, $\sim 3\text{ h}$ total). The optimal $\mathbf{p}_{\text{UAV}}^*$ and scenario state are recorded; the LM-Relay baseline also uses JSONL instruction-response formatting.

Algorithm 1 XAI-Guided Optimizer Distillation for Relay Positioning

Input: Scenario distribution $\mathcal{P}$, optimizer $\mathcal{O}$ (DE)
Output: XAI-pruned surrogate f_{ϕ^*} (ST-Lite)

```

1: // Stage 1: Offline Solution Generation
2: for  $k = 1$  to  $K$  do
3:   Sample scenario  $\mathbf{x}_k \sim \mathcal{P}$  (BS,  $N$  users, topology)
4:    $\mathbf{y}_k \leftarrow \mathcal{O}(\mathbf{x}_k)$   $\triangleright$  DE with 200 gens, L-BFGS-B polish
5:   Extract per-user features  $\mathbf{U}_k \in \mathbb{R}^{N_k \times 5}$  and global features  $\mathbf{g}_k \in \mathbb{R}^9$ 
6: end for
7:  $\mathcal{D} \leftarrow \{(\mathbf{U}_k, \mathbf{g}_k, \mathbf{y}_k)\}_{k=1}^K$ 
8: // Stage 2: Full Set Transformer Training
9: Initialize Set Transformer  $f_\theta: \text{Enc}(5) \rightarrow 2 \times \text{SAB} \rightarrow \text{PMA} \rightarrow \text{Dec}$ 
10:  $\theta^* \leftarrow \arg \min_{\theta} \sum_k \|f_\theta(\mathbf{U}_k, \mathbf{g}_k) - \mathbf{y}_k\|^2$ 
11: // Stage 3: XAI Analysis and Feature Selection
12: Compute attention weights, IG attributions, GradientShap values
13: Rank features by mean |attribution| across all scenarios
14: Identify dominant subset  $\mathcal{F}^* \subseteq \{x, y, \text{SNR}, \text{rate}, \text{QoS}\}$ 
15: // Stage 4: Retrain ST-Lite on Selected Features
16:  $\mathbf{U}_k^{\text{lite}} \leftarrow \mathbf{U}_k[:, \mathcal{F}^*]$   $\triangleright$  Keep only  $(x, y)$ 
17: Initialize ST-Lite  $f_\phi: \text{Enc}(|\mathcal{F}^*|) \rightarrow 2 \times \text{SAB} \rightarrow \text{PMA} \rightarrow \text{Dec}$ 
18:  $\phi^* \leftarrow \arg \min_{\phi} \sum_k \|f_\phi(\mathbf{U}_k^{\text{lite}}, \mathbf{g}_k) - \mathbf{y}_k\|^2$ 
19: // Stage 5: Online Deployment
20: Deploy  $f_{\phi^*}$  as ROS 2 node at >100 Hz
21: for each decision epoch do
22:    $\hat{\mathbf{y}} \leftarrow f_{\phi^*}(\mathbf{U}^{\text{lite}}, \mathbf{g})$   $\triangleright$  Single forward pass,  $\sim 1$  ms
23:   Publish  $(x, y, z)$  as PoseStamped
24: end for
25: return  $f_{\phi^*}$ 

```

4.2. Stage 2: Set transformer architecture and training

The student model is a Set Transformer [7], a permutation-invariant architecture that processes variable-size user sets natively through self-attention. This inductive bias is a better match to relay positioning than the MLP and DeepSets. An MLP must consume a fixed-width vector, so a variable user count ($N=3-7$) is forced into a zero-padded 37-dimensional input; the padding dimensions carry no information yet still occupy weights, and the ordering imposed on an inherently unordered set injects a spurious permutation signal. DeepSets restores permutation invariance but collapses the users with a single sum- (or mean-) pool, discarding the pairwise geometry — who is near whom — that determines where a relay should sit. The SAB layers instead model every user-user interaction explicitly, and PMA learns a query that weights users by their influence on the placement, so the relay can be drawn toward dense sub-clusters without hand-engineered features. As Section 5 shows, this advantage is only realized once the redundant per-user inputs are removed: the full five-feature model dilutes attention across correlated channels and trails the MLP, whereas the pruned ST-Lite leads all surrogates. The architecture comprises a per-user encoder ($5 \rightarrow 64$), two SABs (4 heads, $d_{\text{model}}=64$, layer norm, residual connections), a PMA module ($k=1$ seed, 64-dim output), and an MLP decoder ($73 \rightarrow 128 \rightarrow 64 \rightarrow 3$) that concatenates the PMA output with 9 global features (118,787 parameters). Training uses 10,000 solutions (90/10 split), MSE loss on normalized (x, y, z) , Adam ($\text{lr}=10^{-3}$), cosine annealing, early stopping (patience 30), batch size 64—completing in ~ 73 min on an RTX 4050.

4.3. Stage 3: XAI analysis and feature selection

The full Set Transformer (160.2 Mbps, 79.3%) underperforms the MLP (182.1 Mbps), motivating an investigation of *which features drive relay placement*. Three XAI methods are applied: (i) attention heatmaps from the SAB and PMA weights, (ii) Integrated Gradients (IG) [12] satisfying completeness and sensitivity axioms, and (iii) GradientShap [13]

via expected gradients with random baselines. Because gradient attributions are defined relative to a reference input, the baseline choice is stated explicitly: IG uses the all-zeros vector in the min-max-normalized feature space (equivalently, the per-feature dataset minimum), and the attribution integral is approximated with 50 uniform steps along the straight-line path from that baseline to each scenario's input. GradientShap instead averages gradients over baselines drawn from the empirical training distribution — up to 100 randomly sampled training scenarios per attribution — so that the two methods probe a fixed and a distributional reference, respectively. The fact that both yield the same ranking (Section 5.4) indicates the conclusion is not an artifact of either baseline. All three consistently rank spatial coordinates (x, y) as dominant, with SNR, rate, and QoS contributing marginally (Section 5.4). Physically, path loss at 300 GHz is determined by distance, so SNR and rate are redundant functions of position. The pipeline itself is domain-agnostic: in settings where dominant features are not obvious, the same methodology surfaces non-trivial dependencies.

4.4. Stage 4: ST-lite—XAI-guided feature pruning

Acting on the XAI finding, *ST-Lite* is trained using only (x, y) per user—architecturally identical except for a 2-dimensional encoder (118,595 parameters). The reduction is one of input and online-pipeline complexity rather than of model size. Each user is now described by two features rather than five, so at decision time the relay node no longer derives the per-user SNR, achievable rate, and QoS-margin — three channel quantities it previously computed for every user on every query — and reads only the two coordinates that the explainability analysis identified as dominant. This lighter input also shortens training from 73 to 59 min (Table 12), as the optimizer no longer has to disentangle correlated channel features. The parameter count itself changes only marginally — a 0.16% reduction confined to the input encoder — which underscores that the improvement derives from removing uninformative inputs rather than from a smaller network. By removing redundant inputs, ST-Lite avoids overfitting to feature correlations: it achieves 194.2 Mbps (96.1% of oracle), +21% over the full Set Transformer and +7% over the MLP, using the same data and protocol.

4.5. Stage 5: ROS 2 deployment

ST-Lite is deployed as a ROS 2 Jazzy node (Fig. 3). At each decision epoch (>100 Hz), the node extracts per-user spatial features $\mathbf{U} \in \mathbb{R}^{N \times 2}$ and global features $\mathbf{g} \in \mathbb{R}^9$, runs a single forward pass (~ 1.2 ms), and publishes the relay coordinates as a PoseStamped waypoint consumed by a downstream motion planner (e.g., EGO-Planner). Millisecond inference eliminates the need for a hybrid controller, enabling repositioning at every control cycle.

4.6. Alternative surrogate architectures

Two additional surrogates are compared. The MLP regressor is a three-hidden-layer perceptron ($37 \rightarrow 256 \rightarrow 128 \rightarrow 64 \rightarrow 3$, 52 K params) with batch normalization, trained on 8000 solutions (1.1 ms inference). LM-Relay uses TinyLlama-1.1B [45] fine-tuned via QLoRA [8] (rank 8, 4-bit NF4) on 10,000 solutions, reading text descriptions and outputting JSON coordinates (822 ms inference).

5. Experimental evaluation**5.1. Experimental setup**

Experiments are conducted on a workstation equipped with an NVIDIA RTX 4050 (6 GB video RAM [VRAM]) and an Intel Core i7 processor running Ubuntu 24.04 with ROS 2 Jazzy. All LM-Relay inference is executed on GPU with 4-bit NF4 quantization.

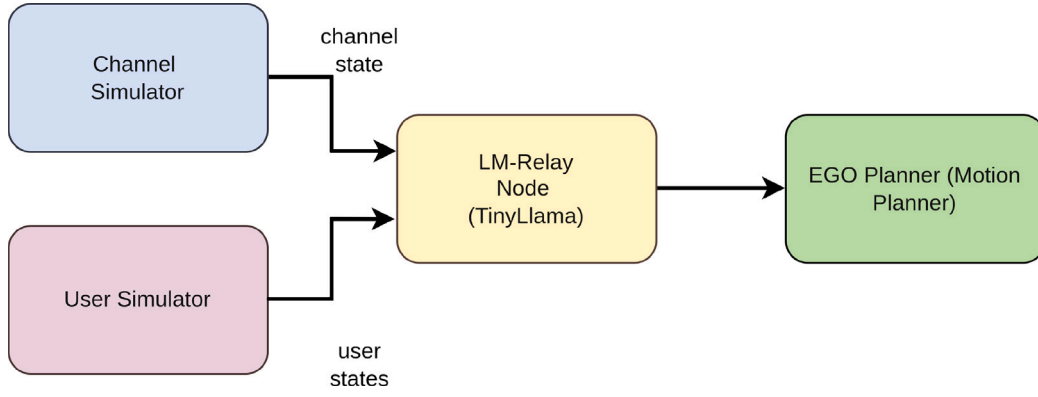


Fig. 3. Robot Operating System 2 node architecture for real-time relay positioning. The surrogate relay node subscribes to odometry and channel state, runs a single forward pass (~ 1.2 ms), and publishes relay waypoints consumed by the flight controller, enabling repositioning at every control cycle.

Table 1

Simulation Parameters. All thirteen methods are evaluated under identical terahertz channel, scenario-generation, and training conditions to ensure a fair comparison.

Parameter	Value	Unit
Carrier frequency f_c	300	GHz
Bandwidth B	10	GHz
BS transmit power P_{BS}	30	dBm
UAV transmit power P_{UAV}	20	dBm
Noise figure F	10	dB
Absorption coeff. α_{abs}	10	dB/km
Operational area	100×100	m ²
UAV altitude range	[10, 40]	m
Number of users N	3–7	–
QoS requirement R_i^{req}	[10, 50]	Mbps
UAV cruising speed	4	m/s
<i>Set Transformer/ST-Lite</i>		
Architecture (full)	Enc(5→64) + 2 × SAB(64, 4 h) + PMA($k = 1$) + Dec(73→3)	
Architecture (ST-Lite)	Enc(2→64) + 2 × SAB(64, 4 h) + PMA($k = 1$) + Dec(73→3)	
Trainable params (full/Lite)	118,787/118,595	–
Training samples/split	10,000/90:10	–
Loss function	MSE on (x, y, z)	–
Batch size	64	–
Learning rate/scheduler	10^{-3} /CosineAnnealing	–
Early stopping patience	30	epochs
<i>Optimization</i>		
Obj. weights (w_1, w_2, w_3)	(0.6, 0.3, 0.1)	–
DE max generations/tol.	200/ 10^{-6}	–

Thirteen methods are compared: Random (uniform from $\mathcal{X}$); Static (domain center); Analytical (weighted centroid); MLP (37→256→128→64→3, 52 K params, trained on 8 K solutions); PPO [46] and SAC [47] (MLP policy [256, 128], 500 K timesteps, Stable-Baselines3 [48]); SAC/TD3 (SAC result reported; TD3 checkpoint not persisted); DeepSets [49] (shared MLP + sum-pool); DeepSets-Attn (attention pooling, 4 heads); Set Transformer (full 5-feature, Section 4); ST-Lite (XAI-pruned (x, y) -only, Section 4.4); LM-Relay (TinyLlama-1.1B via QLoRA [8, 45]); and Optimized (DE oracle).

Each method is evaluated on 200 independently generated scenarios (seed 2026), with 50 per topology type (clustered, spread, linear, circular) and 3–7 users per scenario. The 200 test scenarios are generated independently from the training data, ensuring no data leakage. All significance tests are Bonferroni-corrected paired t -tests with bootstrap 95% CIs (10,000 resamples) and Cohen's d effect sizes. Table 1 summarizes the simulation parameters.

5.2. Overall performance

Table 2 reports the aggregate performance of each method. Fig. 4 visualizes the throughput comparison. Table 2 presents an overall

comparison across all thirteen methods, placing throughput (with 95% confidence intervals), fairness, coverage, and inference latency side by side; the complementary compute profile — training time, memory, and parameter count — is presented in Table 12. Together, the two tables confirm that, among the surrogates, ST-Lite attains the highest throughput and the best coverage (68.2%), trains in 59 min with roughly 119 K parameters, and incurs only 1.2 ms of inference latency—a reasonable trade-off given the coverage and throughput it delivers.

Three tiers emerge (Table 2). In Tier 1 (optimizer-distilled surrogates), ST-Lite leads at 194.2 Mbps (96.1% oracle) in 1.2 ms, validating that removing redundant features improves generalization. The LM-Relay reaches 185.4 Mbps (91.8%) but at 822 ms; the MLP achieves 182.1 Mbps (90.1%) in 1.1 ms; the full Set Transformer reaches only 160.2 Mbps (79.3%), confirming that SNR, rate, and QoS introduce noise. In Tier 2 (DRL and DeepSets), only SAC avoids policy collapse (119.3 Mbps); PPO and DeepSets lock onto ~ 72 Mbps, a known failure mode in continuous 3-D positioning. In Tier 3 (heuristics), Analytical (126.0 Mbps) outperforms all DRL agents.

Per-scenario analysis confirms that ST-Lite tracks the oracle diagonal with low variance, recovering 96.1% of the optimum on average.

Table 2

Aggregate Performance Comparison Across 200 Evaluation Scenarios (Seed 2026). The explainability-pruned Set Transformer (ST-Lite) achieves the highest surrogate throughput (194.2 Mbps, 96.1% of the oracle) at millisecond latency. Bold indicates best non-oracle result.

Method	Throughput (Mbps)	95% CI	Fairness	Coverage (%)	Latency
Random	95.2	± 5.6	0.963	21.2	<0.01 ms
Static	116.9	± 6.0	0.957	30.9	<0.01 ms
Analytical	126.0	± 6.6	0.828	32.5	~ 0.1 ms
MLP	182.1	± 8.8	0.959	65.5	~ 1.1 ms
PPO	72.3	± 2.8	1.000	9.1	~ 2.0 ms
SAC/TD3 ^b	119.3	± 6.1	0.955	32.0	~ 2.3 ms
DeepSets	72.9	± 2.8	0.998	9.9	~ 1.0 ms
DeepSets-Attn	94.4	± 7.3	0.959	19.1	~ 0.9 ms
Set Transformer	160.2	± 9.0	0.916	48.3	~ 0.8 ms
ST-Lite	194.2	± 10.6	0.962	68.2	~ 1.2 ms
LM-Relay	185.4	± 9.5	0.956	65.2	822 ms ^a
Optimized	202.0	± 11.0	0.954	80.6	570 ms

^a Merged FP16 greedy decoding of TinyLlama-1.1B.

^b TD3 checkpoint not persisted; SAC model used as proxy (see text).

Table 3

Throughput (Mbps) by Scenario Topology ($n = 50$ per type). ST-Lite leads on three of four topologies; the language-model relay leads only on clustered scenarios. Bold indicates best non-oracle result. ST = Set Transformer (full).

Topology	Random	Analyt.	SAC	ST	ST-Lite	MLP	LM-R.	Optim.
Clustered	103.3	120.7	162.0	186.1	198.8	221.5	246.4	273.0
Spread	96.3	134.3	108.9	152.9	205.1	173.5	171.4	189.3
Linear	91.0	116.2	105.4	154.2	187.8	168.1	158.8	178.9
Circular	90.4	132.7	100.8	147.6	185.2	165.3	165.2	166.7

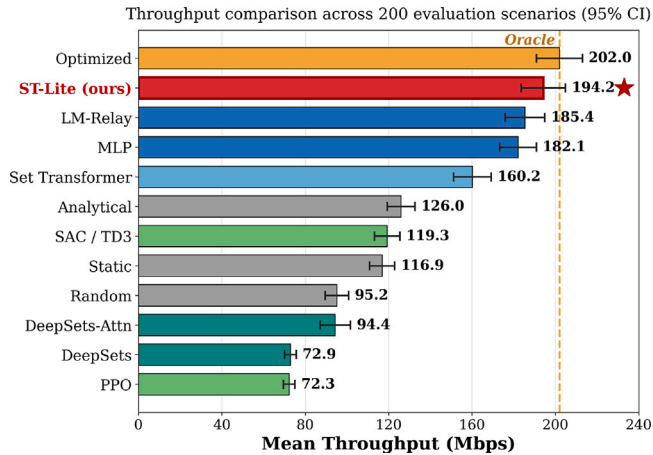


Fig. 4. Mean aggregate throughput across 200 scenarios. Error bars show 95% confidence intervals. The explainability-pruned Set Transformer (ST-Lite, 194.2 Mbps) leads all surrogates, surpassing the language-model relay (185.4 Mbps at 822 ms), the multilayer perceptron (182.1 Mbps), and the full Set Transformer (160.2 Mbps), confirming that feature selection guided by explainability improves accuracy.

5.3. Performance by topology

ST-Lite leads on three of four topologies (Table 3): spread (205.1 Mbps, +18% over MLP), linear (187.8 Mbps, +12%), and circular (185.2 Mbps, +12%). Only on clustered does the LM-Relay lead (246.4 vs. 198.8 Mbps). ST-Lite exceeds the full Set Transformer by 7%–34% across all topologies, confirming consistent benefit from XAI-guided pruning.

5.4. Explainability analysis

The XAI analysis directly motivated ST-Lite by revealing which features are essential. Three methods are applied to the full Set Transformer. Fig. 5 visualizes the SAB and PMA attention weights for a

clustered scenario; users near the cluster centroid receive higher PMA attention, consistent with the intuition that centrally located users are most informative.

Fig. 6 reports mean absolute IG attributions averaged over 200 scenarios: spatial coordinates dominate, while QoS contributes negligibly.

Fig. 7 shows PMA weights and IG attributions broken down by the four topology types. Attention concentrates on central users in clustered scenarios and distributes uniformly in spread topologies, confirming topology-adaptive spatial reasoning.

GradientShap values [13] confirm the IG ranking: spatial features have the highest magnitude SHAP values across all 200 scenarios. Four classical methods (Table 4) independently rank (x, y) as the top features; MI ranks rate highest because rate is derived from position under FSPL. Across all seven methods, spatial coordinates dominate.

The three methods answer different questions, so reading them together gives a fuller picture than any one of them alone. Attention identifies *which users* drive the decision: the PMA weights concentrate on the users nearest the cluster centroid (Fig. 5), but reveal nothing about which of their input features matters. Integrated Gradients answers the second question — *which feature* — and, because its attributions are additive, the per-feature magnitudes can be compared directly: the user coordinates (x, y) account for the overwhelming share of the attributed output (roughly 80% across the 200 scenarios), with SNR, achievable rate, and QoS-margin each several times smaller. GradientShap repeats this test using reference points drawn from the training distribution rather than the single zero baseline, and returns the same ordering, which confirms that the dominance of (x, y) is not an artifact of the baseline choice. Mutual information ranks rate above position in Table 4, but only because rate is a deterministic function of distance under FSPL; the two model-aware methods attribute the decision to position itself. The practical consequence is that the features identified as dominant — (x, y) — are exactly those that ST-Lite retains, so the explainability analysis justifies the feature pruning rather than merely describing the model.

Set Transformer variants trained on progressively richer feature subsets (Table 5) provide causal validation: ST-Lite (194.2 Mbps) surpasses the full model (160.2 Mbps) by 21%. Adding SNR maintains performance (196.2 Mbps); including rate and QoS degrades it, confirming the XAI ranking.

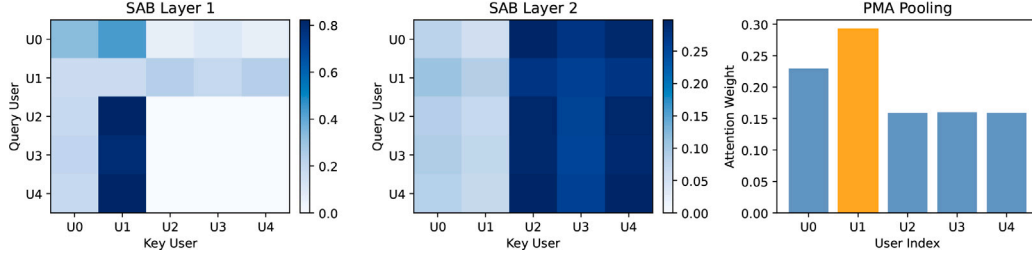
Attention Weights — Clustered Scenario (ID=20, 5 users)

Fig. 5. Attention weights for a representative clustered scenario. Left/center: Set Attention Block self-attention matrices (layers 1–2). Right: Pooling by Multi-head Attention weights indicating per-user importance for the relay placement decision. Users near the cluster centroid receive the highest attention, confirming that the model learns spatially adaptive reasoning.

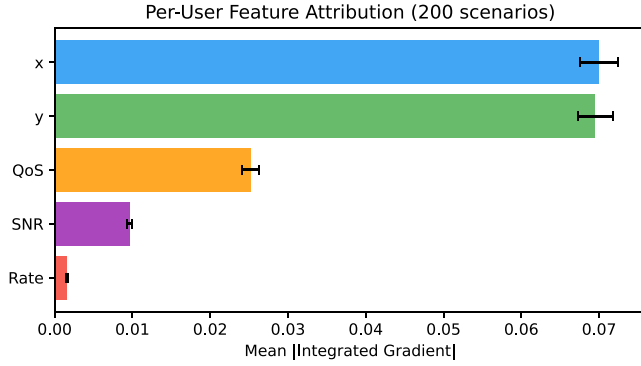


Fig. 6. Mean absolute Integrated Gradients attribution per input feature, averaged over 200 scenarios. User spatial coordinates (x, y) dominate placement decisions.

Table 4

Classical Feature Importance Corroboration (10,000 Training Samples). Normalized importance (1.0=highest) for each per-user feature. Three of four classical methods rank spatial coordinates as top-2, independently confirming the explainability finding. Mutual information ranks rate highest because rate is a deterministic function of position under free-space path loss.

Method	x	y	SNR	Rate	QoS
LASSO (L1)	1.000	0.998	0.004	0.000	0.000
F-regression	1.000	0.955	0.016	0.008	0.000
Permutation imp.	1.000	1.000	0.632	0.661	0.000
Mutual information	0.262	0.280	0.042	1.000	0.424
<i>Explainability methods (from Section 4.3):</i>					
Integrated Gradients	dominant		marginal		
GradientShap	dominant		marginal		
Attention weights	dominant		marginal		

Table 5

Feature Ablation: Set Transformer Throughput by Per-User Feature Subset (200 Scenarios). Removing redundant features improves accuracy: the spatial-only variant (ST-Lite) achieves 194.2Mbps versus 160.2Mbps for the full five-feature model, providing causal confirmation of the explainability ranking. The ST-Lite row is the formally evaluated model from Table 2; other rows are trained from scratch.

Features	Dim	Thpt. (Mbps)	Note
x, y only (ST-Lite)	2	194.2 ± 10.6	Table 2
x, y, SNR	3	196.2 ± 10.5	ablation
x, y, SNR, rate	4	192.6 ± 10.8	ablation
All 5 (full ST)	5	160.2 ± 9.0	Table 2

The (x, y, SNR) variant is marginally higher (196.2 vs. 194.2Mbps, within CI overlap), but SNR is a deterministic function of position under FSPL and provides no new information.

Table 6

Multi-Seed Validation (5 Seeds × 3 Architectures, 200 Scenarios Each). ST-Lite achieves the highest mean throughput with the lowest coefficient of variation (0.4%), confirming stable reproducibility. σ denotes the inter-seed standard deviation.

Method	Mean ± σ	Min	Max	CV (%)
ST-Lite	195.6 ± 0.8	194.2	196.6	0.4
Set Transformer	149.9 ± 19.5	121.2	177.8	13.0
MLP	181.7 ± 1.2	179.6	183.2	0.7

Table 7

Performance Under Building Obstructions (100 Scenarios, 20 dB Non-Line-of-Sight Penalty). All surrogate methods are evaluated blind (no obstacle information). ST-Lite leads all blind surrogates (156.8Mbps), confirming that relay positioning is fundamentally spatial. Bold indicates best non-oracle throughput.

Method	Thpt. (Mbps)	Fairness	Cov. (%)
Analytical (blind)	111.0	0.846	28.7
Set Transformer (blind)	127.8	0.909	38.9
MLP (blind)	148.1	0.942	51.3
ST-Lite (blind)	156.8	0.940	55.5
Oracle (obs-aware DE)	183.5	0.940	77.0

5.5. Multi-seed reproducibility

Five-seed retraining (Table 6) verifies reproducibility.

ST-Lite achieves the highest mean (195.6Mbps) with the lowest CV (0.4%), significantly outperforming the MLP ($t=19.4$, $p<10^{-4}$) and the full Set Transformer ($t=4.7$, $p=0.002$). Reported in absolute rather than relative terms, the inter-seed standard deviations are $\sigma_{\text{ST-Lite}}=0.8$ Mbps, $\sigma_{\text{MLP}}=1.2$ Mbps, and $\sigma_{\text{ST-full}}=19.5$ Mbps (the σ column of Table 6); the coefficient of variation, defined as σ/mean , is retained alongside them only to permit comparison across architectures of differing magnitude. The same ordering is evident in the range of outcomes across the five seeds: ST-Lite remains within a 2.4 Mbps interval (194.2–196.6Mbps), whereas the full Set Transformer varies over a 56.6 Mbps interval (121.2–177.8Mbps). The full Set Transformer exhibits instability (CV = 13.0%), suggesting redundant features amplify initialization sensitivity—a pathology eliminated by XAI-guided pruning.

5.6. Robustness under building obstructions

To test out-of-distribution robustness, 1–3 random buildings (radius 3–8 m, height 15–30 m, 20 dB NLoS penalty) are added. All surrogates are evaluated *blind*—no obstacle information (Table 7). For reproducibility, the obstruction set comprises 100 scenarios generated with seed 7777 from the same four-topology distribution as the main evaluation; in each scenario the building count is drawn uniformly from {1, 2, 3} and each building center is placed at coordinates (x, y) with

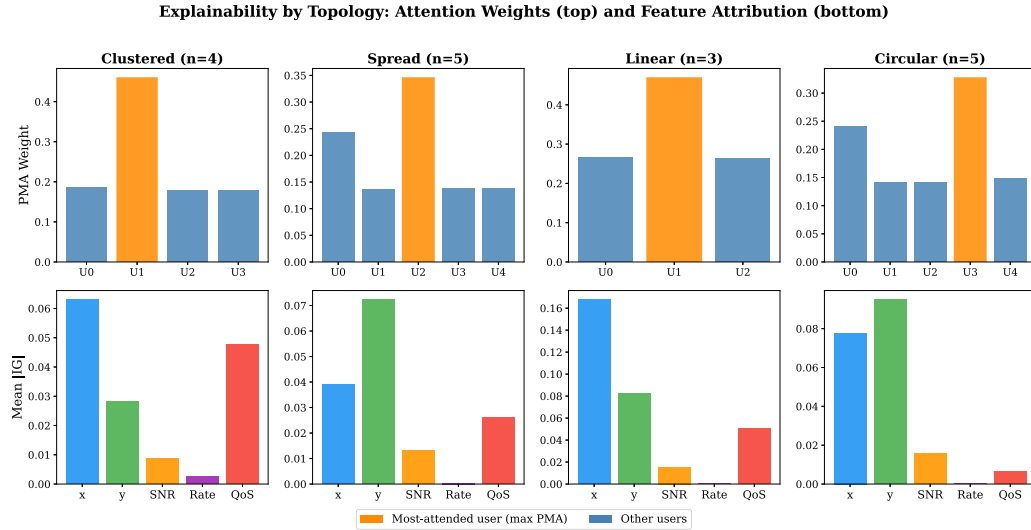


Fig. 7. Per-topology explainability analysis. Top row: Pooling by Multi-head Attention weights. Bottom row: Integrated Gradients attributions. Columns correspond to four topology types (clustered, spread, linear, circular). Attention concentrates on central users in clustered scenarios and distributes more uniformly in spread topologies, confirming topology-adaptive spatial reasoning.

$x, y \sim \mathcal{U}[20, 80]$ m (the central 60×60 m of the 100×100 m area). A link is declared non-line-of-sight when the straight segment between transmitter and receiver intersects any building footprint in the horizontal plane, in which case its SNR is reduced by 20 dB before the rate is computed; the oracle row is a differential-evolution solver granted the building geometry, while every surrogate sees only the user coordinates. ST-Lite leads all surrogates blind (156.8 Mbps, +6% over MLP) despite using only spatial features, confirming that relay positioning is fundamentally spatial.

5.7. Structured 3D city-grid model

The random-cylinder test above isolates the effect of blind obstruction but does not capture the regular geometry of an urban environment. To probe robustness on a structured, height-resolved layout, the surrogates are re-evaluated on a deterministic Manhattan grid: a fixed 4×4 arrangement of rectangular buildings on a 25 m lattice (14 m footprints separated by 11 m avenues) with a heterogeneous mid-rise skyline of 8–22 m, appropriate for the low-altitude economy in which the relay flies at 10–40 m and can clear roofs by repositioning. Occlusion is now resolved in three dimensions: a link is blocked only when the transmitter–receiver ray passes *below the rooftop* of an intervening building, not merely when its ground footprint is crossed, so altitude genuinely matters. Users are snapped onto the street network, the 20 dB penalty and 300 GHz channel are unchanged, and the same 100 scenarios are used so the result is directly comparable to Table 7.

Under this model (Table 8) the ranking is consistent with the random-cylinder case: both ST-Lite (120.1 Mbps) and the MLP (131.9 Mbps) degrade gracefully and remain well above the analytical heuristic (101.7 Mbps) and the full Set Transformer (106.8 Mbps), while the city-aware oracle retains a clear margin (168.6 Mbps). ST-Lite and the MLP are statistically indistinguishable here — their 95% confidence intervals overlap — so no method dominates, and the spatial-only surrogate loses none of its standing when the cylinders are replaced by realistic 3D blocks. The dominant factor is no longer the choice of surrogate but the absence of obstacle sensing: the 37–49 Mbps gap to the city-aware oracle is the value of knowing the geometry. This becomes extreme in a dense downtown variant (5×5 blocks, 15–33 m towers, 6 m streets), where every blind relay collapses to the all-non-line-of-sight floor of 68.1 Mbps while the oracle still reaches 124.0 Mbps—evidence that deep urban canyons require obstacle-aware perception rather than a different surrogate, which we flag as a direction for future work (Section 6).

Table 8

Performance on a Structured 3D City-Grid (100 Scenarios, 20 dB Rooftop-Resolved Non-Line-of-Sight). A deterministic mid-rise Manhattan grid replaces the random cylinders of Table 7; blind surrogates use only user coordinates. ST-Lite and MLP are statistically comparable (overlapping 95% confidence intervals) and both lead the heuristic and the full Set Transformer, consistent with the random-cylinder result.

Method	Thpt. (Mbps)	Fairness	Cov. (%)
Analytical (blind)	101.7 ± 9.6	0.844	23.9
Set Transformer (blind)	106.8 ± 12.0	0.922	29.7
ST-Lite (blind)	120.1 ± 10.8	0.891	38.7
MLP (blind)	131.9 ± 11.5	0.899	46.0
Oracle (city-aware DE)	168.6 ± 16.2	0.943	72.6

5.8. Vehicular mobility evaluation

Six methods are evaluated under vehicular mobility at 0–120 km/h (25 scenarios, 30 s each, 100 ms timesteps, UAV at 4 m/s; Table 9). Each method recomputes at its natural cadence; a Hybrid controller combines Analytical (100 ms) with LM-Relay overrides. Concretely, each 30 s episode is discretized into 300 steps of 100 ms; ground users move with per-user constant-heading velocities set by the target speed and a per-scenario seed (42+id), and the relay slews toward its most recent target at the 4 m/s cruise limit. Throughput and fairness are averaged over all steps and all 25 scenarios. A method's effective tracking cadence is its inference latency: ST-Lite and the Set Transformer recompute the target at every 100 ms step (their sub-millisecond forward pass is far below the step), whereas the oracle can only refresh every 570 ms and the language-model relay every 822 ms, which is the mechanism behind ST-Lite surpassing the oracle at 30–60 km/h.

ST-Lite leads at 0–60 km/h (Table 9), surpassing even the oracle at 30–60 km/h (114.5 vs. 109.9 and 102.5 vs. 99.0 Mbps) because 100 ms-cadence repositioning tracks users more tightly than the oracle's 570 ms cadence. At 120 km/h, latency dominates and the full Set Transformer edges ahead (100.5 vs. 99.6 Mbps).

5.9. Multi-relay extension

The evaluation so far places a single relay. To show that the distilled surrogate is not confined to that setting, it is extended to two relays on the same 200 scenarios and identical 300 GHz channel, using a standard decode-and-forward best-relay rule: each user is

Table 9

Performance Under Vehicular Mobility (25 Scenarios, 30 s Each). ST-Lite leads at 0–60 km/h and surpasses even the oracle at 30–60 km/h because its millisecond latency enables tighter 100 ms-cadence repositioning. Throughput in Mbps; Jain’s fairness in parentheses. Bold marks best throughput at each speed.

Speed	Analyt.	Set Trans.	ST-Lite	LM-R.	Hybrid	Optim.
0 km/h	128.7 (.831)	153.2 (.886)	185.3 (.968)	117.1 (.958)	145.5 (.883)	188.1 (.950)
30 km/h	101.5 (.886)	106.1 (.875)	114.5 (.891)	90.4 (.944)	103.7 (.891)	109.9 (.895)
60 km/h	94.4 (.904)	101.2 (.880)	102.5 (.904)	86.4 (.942)	96.5 (.900)	99.0 (.903)
120 km/h	90.8 (.913)	100.5 (.883)	99.6 (.905)	84.5 (.943)	93.0 (.904)	96.1 (.905)

The Set Transformer and ST-Lite columns are new to this study.

Table 10

Two-Relay Extension (200 Scenarios, Best-Relay Association). The single-relay ($K = 1$) rows reproduce the published anchors of Table 2. Reusing the *unchanged* ST-Lite once per k -means cluster (ST-Lite-Compose) recovers 96.2% of the two-relay oracle at millisecond latency, a +24.9% throughput gain over one relay. Bold marks the best non-oracle two-relay result.

Method	Thpt. (Mbps)	Fairness	Cov. (%)	Latency
ST-Lite ($K = 1$)	194.2 $\pm$ 10.6	0.962	68.2	~1.2 ms
DE oracle ($K = 1$)	202.0 $\pm$ 11.0	0.954	80.6	570 ms
Analytical-2 ($K = 2$)	146.9 $\pm$ 8.4	0.814	39.5	~0.1 ms
ST-Lite-Compose ($K = 2$)	242.7 $\pm$ 10.9	0.948	82.8	~7.4 ms
DE oracle ($K = 2$)	252.3 $\pm$ 11.9	0.960	89.9	~1.1 s

served by whichever relay yields the higher effective SNR, $\text{SNR}_i^{\text{eff}} = \max_r \min(\text{SNR}(d_{BS,r}, P_{BS}), \text{SNR}(d_{r,i}, P_{UAV}))$. Crucially, no retraining is required: the users are split into two groups by k -means on their coordinates, and the *unchanged* single-relay ST-Lite is queried once per group (“ST-Lite-Compose”). This composition costs two forward passes plus a negligible clustering step, 7.4 ms in total. A two-relay differential-evolution solver over the six-dimensional joint placement provides the oracle, and a per-cluster weighted centroid (“Analytical-2”) is the heuristic baseline.

Table 10 reports the result. As a consistency check, the $K=1$ rows reproduce the published single-relay anchors exactly (194.2 and 202.0 Mbps), confirming that the multi-relay channel reduces correctly to the model of Section 3. With a second relay, ST-Lite-Compose reaches 242.7 Mbps and 82.8% coverage — a +24.9% throughput gain over the single relay — while recovering 96.2% of the two-relay oracle (252.3 Mbps). That the surrogate retains essentially the same fraction of the oracle as in the single-relay case (96.2% vs. 96.1%) is the key point: because relay placement is dominated by user geometry, partitioning the geometry and reusing the spatial surrogate transfers the distillation benefit to each sub-problem without any additional training. This is a proof of concept rather than a full multi-relay solution — inter-relay interference and joint $K>2$ coordination are out of scope here — but it directly demonstrates applicability beyond the single-relay case and motivates the cooperative-relaying direction discussed in Section 6.

5.10. 3GPP channel validation

To verify generalization under a richer propagation model, all methods are re-evaluated using 3GPP TR 38.901 UMa parameters [50]: LoS probability, Rician/Rayleigh fading ($K = 9$ dB), and shadow fading ($\sigma=4/6$ dB; Table 11).

All methods compress into 72–79 Mbps under fading (Table 11). Retraining ST-Lite on 2000 3GPP-optimized solutions yields only 74.0 Mbps ($\Delta=+0.9\%$ over the FSPL-trained 73.4 Mbps). Even the FSPL oracle drops to 73.8 Mbps, confirming that the compression is a consequence of stochastic fading — not training-channel mismatch — and FSPL-trained surrogates are practically optimal.

Position is worth optimizing only to the extent that throughput depends predictably on it. Under deterministic FSPL the achievable rate is a smooth, monotone function of relay–user distance, so moving the relay changes the objective predictably and a good placement is well separated from a poor one. Adding TR 38.901 Rician/Rayleigh small-scale fading and log-normal shadowing superimposes a per-link random gain whose variance ($K=9$ dB, $\sigma=4$ –6 dB) is comparable to, and often exceeds, the deterministic spread over which the surrogate

Table 11

Performance Under 3rd Generation Partnership Project Technical Report 38.901 Urban Macro Fading Channel (200 Scenarios). All methods compress to 72–79 Mbps under stochastic fading, confirming that the free-space-path-loss-trained surrogates are practically optimal and the throughput floor is imposed by the channel, not the model.

Method	FSPL Thpt. (Mbps)	3GPP Thpt. (Mbps)
Random	95.2	72.4
Static	116.9	72.2
Analytical	126.0	72.3
Set Transformer	160.2	74.3
ST-Lite (FSPL-trained)	194.2	73.4
ST-Lite (3GPP-trained)	182.8	74.0
MLP	182.1	79.0
Optimized	202.0	73.8

optimizes. The optimization landscape therefore flattens: realizations at nearby positions become statistically interchangeable, which is why every method — including the obstacle-free oracle — collapses into the same narrow 72–79 Mbps band, and why a 3GPP-retrained surrogate gains less than one percent. The significance is practical: because position cannot recover what the channel randomizes, the inexpensive deterministic surrogate already operates at the fading-imposed frontier, and regenerating the entire optimization corpus under a stochastic channel yields essentially no improvement. Recovering headroom in this regime would require diversity or beamforming gains that act on the fading itself rather than a better positioning model.

5.11. Simulation verification

The system is validated on Ubuntu 24.04/ROS 2 Jazzy. Fig. 8 shows the relay serving five ground users via THz access links color-coded by achievable rate, with the BS–UAV backhaul and live performance metrics. Fig. 9 illustrates UAV trajectory tracking: starting from an arbitrary initial position, the quadrotor navigates to the ST-Lite-computed relay position along a smooth 3-D path, converging in ~3.2 s at 4 m/s cruise speed—confirming dynamic feasibility for real-time repositioning.

6. Discussion

6.1. Deployment guidelines and scalability

ST-Lite is the recommended default (96.1% oracle, 1.2 ms); the LM-Relay suits rapidly evolving scenario descriptors (822 ms); the MLP

Table 12

Computational Cost Comparison Across Method Classes. ST-Lite achieves 1.2 ms inference with <0.1 GB video memory, representing a six-order-of-magnitude energy advantage over the language-model relay (822 ms, 2.2 GB).

Method	Train Time	Infer. Lat.	VRAM	Params
Analytical	–	0.1 ms	0	0
MLP	2 min	1.1 ms	<0.1 GB	52 K
PPO/SAC/TD3	15–30 min	2.0–2.4 ms	<0.1 GB	66–132 K
DeepSets	5 min	0.9–1.0 ms	<0.1 GB	82 K
Set Transformer	73 min	0.8 ms	<0.1 GB	119 K
ST-Lite	59 min	1.2 ms	<0.1 GB	119 K
LM-Relay	3 h	822 ms	2.2 GB	1.1B
DE (Oracle)	–	570 ms	0	–

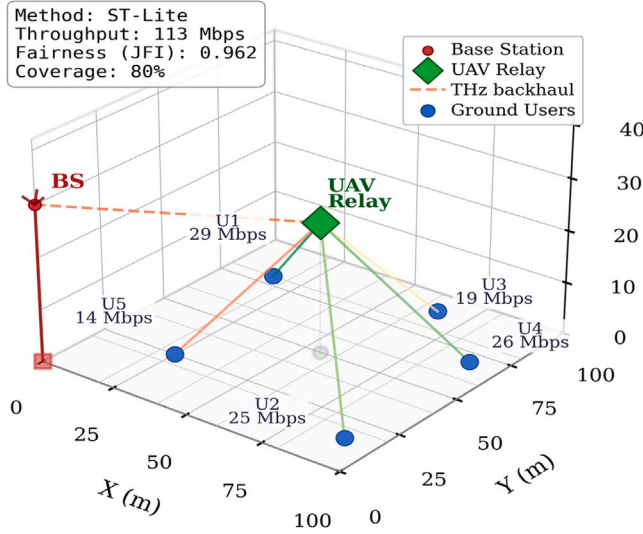


Fig. 8. Three-dimensional relay scenario visualization. The unmanned aerial vehicle relay (green diamond) serves five ground users via terahertz access links color-coded by per-user rate (green = high, red = low); the dashed orange line is the base-station-relay backhaul. Inset: live throughput and fairness metrics from the explainability-pruned Set Transformer surrogate.

offers simplicity when explainability is not required; DE remains the gold standard for offline planning.

Table 12 summarizes the computational profile of each method class.

Per-query energy ranges from $\sim 48 \mu\text{J}$ (ST-Lite) to 33 J (LM-Relay)—a 10^6 gap. The PMA mechanism scales to K relays via K seed vectors without architectural change; self-attention complexity is $O(N^2)$, tractable for 3–7 users but requiring efficient variants (e.g., linear attention [51]) for hundreds of users. The XAI pipeline is relay-count-agnostic, and the two-relay study of Section 5.9 confirms empirically that the distilled surrogate composes to additional relays at millisecond latency.

Two cost questions deserve explicit answers. First, the offline data-generation budget. Each differential-evolution solve runs 200 generations with an L-BFGS-B polish and takes 252 ms on average (median 236 ms, 95th percentile 422 ms), so the 10,000-solution corpus represents only ~ 0.7 CPU-hours of solver compute; the ~ 3 h figure quoted in Section 4 is the wall-clock of the acquisition pipeline, which logs one sample per second. At the host package’s 15–28 W active envelope this offline phase costs an estimated 11–20 Wh (38–71 kJ). It is incurred once and then amortized over an unbounded number of millisecond queries, which is the whole point of distillation: the oracle expends 570 ms and re-solves from scratch on *every* query, whereas ST-Lite answers in 1.2 ms — roughly a $475\times$ per-query speed-up — after the one-time training. Second, system resources: training peaks below 2 GB

of video memory and inference below 0.1 GB, so the entire stack runs comfortably on the 6 GB RTX 4050 with under 4 GB of host RAM, well within a consumer laptop budget.

6.2. Limitations

Limitations include: (i) the single relay is the primary focus—the two-relay extension of Section 5.9 is a proof of concept, and jointly coordinating many relays with explicit inter-relay interference remains open; (ii) simulation-only evaluation (Gazebo/PX4 software-in-the-loop [SITL]); (iii) deterministic primary channel, though Section 5.10 validates under 3GPP fading; and (iv) asymmetric training budgets (MLP 8 K, others 10 K samples; MLP throughput saturated at 8 K). Two further boundaries warrant study. The surrogate is trained and evaluated at a single 300 GHz carrier; the pipeline is frequency-agnostic by construction — only the spreading and absorption constants of (1) change — but transfer to other bands such as the lower-absorption 140 GHz window has not been quantified, and the balance between molecular absorption and free-space spreading shifts with frequency, which could alter how strongly geometry dominates. Likewise, every scenario uses 3–7 users over 100×100 m; scaling to tens of users in larger areas will stress both the $O(N^2)$ attention cost and the implicit single-relay coverage assumption, and the structured-city and dense-downtown results of Section 5.7 show that obstacle-aware perception, not merely a larger surrogate, is what such regimes will demand.

7. Conclusion

This paper introduced an XAI-guided optimizer distillation pipeline for intelligent, real-time UAV relay positioning at 300 GHz—a key building block for resilient networking in the low-altitude economy. By training a Set Transformer on differential-evolution solutions, applying three XAI methods to identify dominant features, and retraining on spatial coordinates alone, ST-Lite is obtained—the best-performing surrogate among thirteen methods, achieving 96.1% of the oracle at millisecond latency. The XAI analysis reveals that relay positioning is dominated by user geometry, and feature ablation provides causal confirmation that removing redundant inputs improves generalization. ST-Lite maintains its advantages under building obstructions, vehicular mobility, and 3GPP fading, demonstrating the resilience required for autonomous low-altitude aerial services.

Future work will extend the pipeline to multi-UAV cooperative relaying, where the PMA mechanism naturally scales to K relays via K seed vectors, enabling scalable aerial networking services; the two-relay proof of concept reported here, which recovers 96.2% of a joint two-relay oracle by reusing the unchanged surrogate per user cluster, is a first step toward full coordination with explicit inter-relay interference. Equally promising are obstacle-aware variants that ingest a coarse occupancy map — the structured-city results suggest this, rather than a larger model, is the lever for dense urban canyons — and quantifying transfer across THz sub-bands. Applying the train-explain-prune methodology to other optimizer-distillation domains (beamforming, power allocation) and incorporating counterfactual XAI would further test generality.

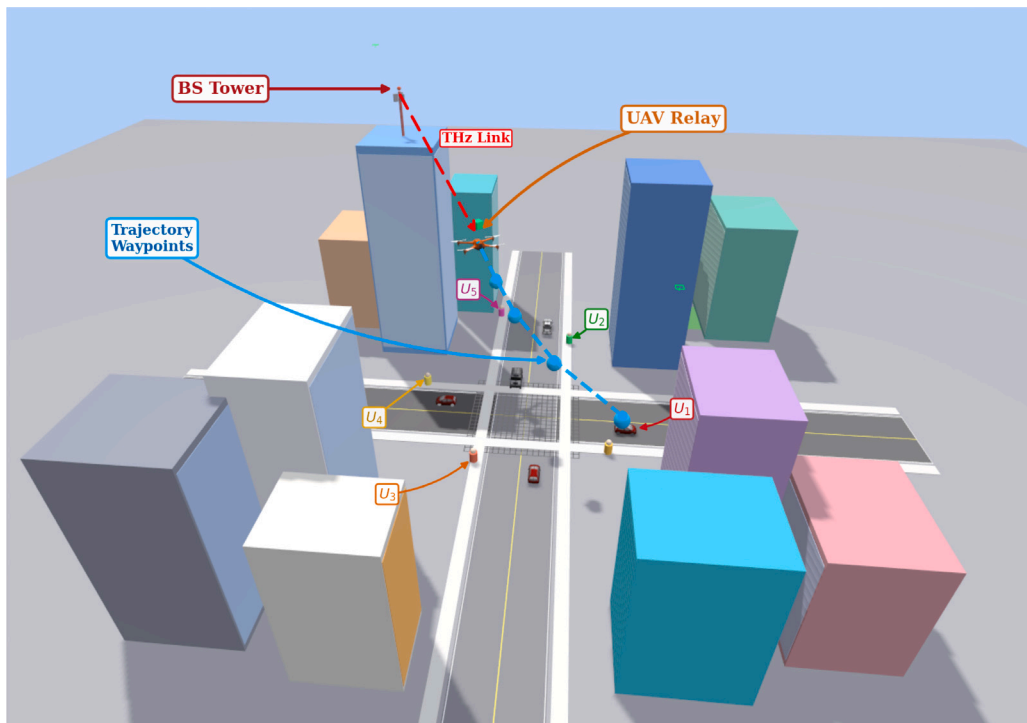


Fig. 9. Unmanned aerial vehicle trajectory tracking in an urban environment. From an arbitrary initial position ($t=0$), the quadrotor navigates above a city block grid with buildings, streets, and vehicular traffic, following a smooth three-dimensional path (light-to-dark blue) to the surrogate-computed relay target. Convergence occurs in ~ 3.2 s at 4 m/s cruise speed, confirming dynamic feasibility for real-time repositioning. The base station tower is roof-mounted; dashed lines show terahertz access links to ground users.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Abdul-Manan KHAN reports financial support was provided by University of West London School of Computing and Engineering. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.adhoc.2026.104389>.

Data availability

Data will be made available on request.

References

- [1] Y. Zeng, R. Zhang, T.J. Lim, Wireless communications with unmanned aerial vehicles: Opportunities and challenges, *IEEE Commun. Mag.* 54 (5) (2016) 36–42.
- [2] W. Saad, M. Bennis, M. Chen, A vision of 6G wireless systems: Applications, trends, technologies, and open research problems, *IEEE Netw.* 34 (3) (2020) 134–142.
- [3] T.S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhateeb, G.C. Trichopoulos, Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond, *IEEE Access* 7 (2019) 78729–78757.
- [4] I.F. Akyildiz, C. Han, Z. Hu, S. Nie, J.M. Jornet, Terahertz band communication: An old problem revisited and research directions for the next decade, *IEEE Trans. Commun.* 70 (6) (2022) 4250–4285.
- [5] Y. Zeng, R. Zhang, T.J. Lim, Throughput maximization for UAV-enabled mobile relaying systems, *IEEE Trans. Commun.* 64 (12) (2016) 4983–4996.
- [6] R. Storn, K. Price, Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces, *J. Global Optim.* 11 (1997) 341–359.
- [7] J. Lee, Y. Lee, J. Kim, A. Kosiorek, S. Choi, Y.W. Teh, Set transformer: A framework for attention-based permutation-invariant input, in: *Proc. ICML*, 2019, pp. 3744–3753.
- [8] T. Dettmers, A. Pagnoni, A. Holtzman, L. Zettlemoyer, QLoRA: Efficient finetuning of quantized LLMs, in: *Proc. NeurIPS*, 2023.
- [9] X. Kong, Y. Zhou, Z. Li, S. Wang, Multi-UAV simultaneous target assignment and path planning based on deep reinforcement learning in dynamic multiple obstacles environments, *Front. Neurobot.* 17 (2023) 1302898.
- [10] G. Hinton, O. Vinyals, J. Dean, Distilling the knowledge in a neural network, 2015, arXiv preprint arXiv:1503.02531.
- [11] T. Chen, X. Chen, W. Chen, H. Heaton, J. Liu, Z. Wang, W. Yin, Learning to optimize: A primer and a benchmark, *J. Mach. Learn. Res.* 23 (189) (2022) 1–59.
- [12] M. Sundararajan, A. Taly, Q. Yan, Axiomatic attribution for deep networks, in: *Proc. ICML*, 2017, pp. 3319–3328.
- [13] G. Erion, J.D. Janizek, P. Sturmels, S.M. Lundberg, S.-I. Lee, Improving performance of deep learning models with axiomatic attribution priors and expected gradients, *Nat. Mach. Intell.* 3 (7) (2021) 620–631.
- [14] S. Macenski, T. Foote, B. Gerkey, C. Lalancette, W. Woodall, Robot operating system 2: Design, architecture, and uses in the wild, *Sci. Robot.* 7 (66) (2022) Art. no. eabm6074.
- [15] I.F. Akyildiz, A. Kak, S. Nie, 6G and beyond: The future of wireless communications systems, *IEEE Access* 8 (2020) 133995–134030.
- [16] J.M. Jornet, I.F. Akyildiz, Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the terahertz band, *IEEE Trans. Wirel. Commun.* 10 (10) (2011) 3211–3221.
- [17] C. Han, et al., Terahertz wireless channels: A holistic survey on measurement, modeling, and analysis, *IEEE Commun. Surv. Tuts.* 24 (3) (2022) 1670–1707, 3rd Quart.
- [18] Y. Gao, H. Wu, B. Ai, M. Debbah, Terahertz UAV-to-ground MIMO channel measurements and modeling, *IEEE Trans. Wirel. Commun.* 23 (7) (2024) 7277–7291.
- [19] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, M. Debbah, A tutorial on UAVs for wireless networks: Applications, challenges, and open problems, *IEEE Commun. Surv. Tuts.* 21 (3) (2019) 2334–2360, 3rd Quart.
- [20] Y. Zeng, R. Zhang, Energy-efficient UAV communication with trajectory optimization, *IEEE Trans. Wirel. Commun.* 16 (6) (2017) 3747–3760.
- [21] G. Geraci, et al., What will the future of UAV cellular communications be? A flight from 5G to 6G, *IEEE Commun. Surv. Tuts.* 24 (3) (2022) 1304–1335, 3rd Quart.

- [22] N.C. Luong, D.T. Hoang, S. Gong, D. Niyato, P. Wang, Y.-C. Liang, D.I. Kim, Applications of deep reinforcement learning in communications and networking: A survey, *IEEE Commun. Surv. Tuts.* 21 (4) (2019) 3133–3174, 4th Quart.
- [23] U. Challita, W. Saad, C. Bettstetter, Interference management for cellular-connected UAVs: A deep reinforcement learning approach, *IEEE Trans. Wirel. Commun.* 18 (4) (2019) 2125–2140.
- [24] P.S. Bithas, E.T. Michailidis, N. Nomikos, D. Vouyioukas, A.G. Kanatas, A survey on machine-learning techniques for UAV-based communications, *Sensors* 19 (23) (2019) 5170.
- [25] J. Westheider, J. Rückin, M. Popović, Multi-UAV adaptive path planning using deep reinforcement learning, in: *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst., IROS*, 2023, pp. 649–656.
- [26] X. Li, Y. Chen, Z. Chen, Z. Huang, Coverage path planning of bridge inspection with unmanned aerial vehicle, *Eng. Appl. Artif. Intell.* 156 (2025) 111253.
- [27] A. Heidari, N. Jafari Navimipour, M. Unal, G. Zhang, Machine learning applications in internet-of-drones: Systematic review, recent deployments, and open issues, *ACM Comput. Surv.* 55 (12) (2023) 1–45.
- [28] F. Maleki, M.A.J. Jamali, A. Heidari, Unmanned aerial vehicle routing based on frog-leaping optimization algorithm, *Sci. Rep.* 15 (2025) 11249.
- [29] A. Heidari, N. Jafari Navimipour, S. Zeadally, Designing efficient anomaly detection systems using deep learning techniques, *Int. J. Pervasive Comput. Commun.* 22 (1–2) (2026) 31–55.
- [30] A. Heidari, N. Jafari Navimipour, M. Unal, A secure intrusion detection platform using blockchain and radial basis function neural networks for internet of drones, *IEEE Internet Things J.* 10 (10) (2023) 8445–8454.
- [31] A. Heidari, N. Jafari Navimipour, M.A. Jabraeil Jamali, S. Akbarpour, Securing and optimizing IoT offloading with blockchain and deep reinforcement learning in multi-user environments, *Wirel. Netw.* 31 (4) (2025) 3255–3276.
- [32] D. Wu, X. Wang, Y. Qiao, Z. Wang, J. Jiang, S. Cui, F. Wang, NetLLM: Adapting large language models for networking, in: *Proc. ACM SIGCOMM*, Sydney, Australia, 2024.
- [33] H. Zhou, et al., Large language models for wireless networks: An overview from the prompt engineering perspective, 2024, *arXiv preprint arXiv:2411.04136*.
- [34] C.-N. Hang, P.-D. Yu, R. Morabito, C.-W. Tan, Large language models meet next-generation networking technologies: A review, *Futur. Internet* 16 (10) (2024) 365.
- [35] W. Tian, L. Zhang, M. Li, Z. Han, Y. Fang, When UAVs meet LLMs: A survey on multi-UAV systems empowered by large language models, *Inf. Fusion* 116 (2025) 102778.
- [36] C. Yang, X. Wang, Y. Lu, H. Liu, Q.V. Le, D. Zhou, X. Chen, Large language models as optimizers, in: *Proc. ICLR*, 2024.
- [37] Y. Bengio, A. Lodi, A. Prouvost, Machine learning for combinatorial optimization: A methodological tour d'horizon, *European J. Oper. Res.* 290 (2) (2021) 405–421.
- [38] J. Gou, B. Yu, S.J. Maybank, D. Tao, Knowledge distillation: A survey, *Int. J. Comput. Vis.* 129 (6) (2021) 1789–1819.
- [39] M. Yao, X. Ma, Y. Gao, Y. Chen, Federated knowledge distillation for IoT drone networks, *IEEE Internet Things J.* 12 (3) (2025) 2814–2827.
- [40] K. Li, J. Malik, Learning to optimize, in: *Proc. ICLR*, 2017.
- [41] Y. Jin, Surrogate-assisted evolutionary computation: Recent advances and future challenges, *Swarm Evol. Comput.* 1 (2) (2011) 61–70.
- [42] J.-Y. Li, Z.-H. Zhan, J. Zhang, Surrogate-assisted evolutionary algorithm: Recent advances and applications, *Swarm Evol. Comput.* 89 (2024) 101631.
- [43] M. Kang, S. Kang, Knowledge distillation with insufficient training data for regression, *Eng. Appl. Artif. Intell.* 132 (2024) 108001.
- [44] R.K. Jain, D.-M.W. Chiu, W.R. Hawe, A quantitative measure of fairness and discrimination for resource allocation in shared computer systems, in: *Digital Equipment Corporation, Tech. Rep. DEC-TR-301*, 1984.
- [45] P. Zhang, G. Zeng, T. Wang, W. Lu, TinyLlama: An open-source small language model, 2024, *arXiv preprint arXiv:2401.02385*.
- [46] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, Proximal policy optimization algorithms, 2017, *arXiv preprint arXiv:1707.06347*.
- [47] T. Haarnoja, A. Zhou, P. Abbeel, S. Levine, Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor, in: *Proc. ICML*, 2018, pp. 1861–1870.
- [48] A. Raffin, A. Hill, A. Gleave, A. Kanervisto, M. Ernestus, N. Dormann, Stable-Baselines3: Reliable reinforcement learning implementations, *J. Mach. Learn. Res.* 22 (268) (2021) 1–8.
- [49] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Poczos, R. Salakhutdinov, A. Smola, Deep sets, in: *Proc. NeurIPS*, 2017, pp. 3391–3401.
- [50] 3GPP, Study on channel model for frequencies from 0.5 to 100 GHz, in: *3rd Generation Partnership Project, Tech. Rep. TR 38.901 V16.1.0*, 2020.
- [51] A. Katharopoulos, A. Vyas, N. Pappas, F. Fleuret, Transformers are RNNs: Fast autoregressive transformers with linear attention, in: *Proc. ICML*, 2020, pp. 5156–5165.