

A MARL approach for connectivity-aware search and rescue in urban environments

Andrés Meseguer Valenzuela

Abstract— This work presents a closed-loop experimental framework for connectivity-aware urban search and rescue (SAR) using heterogeneous unmanned ground vehicles (UGVs) and unmanned aerial vehicles (UAVs). The setup couples a physics-based urban digital twin in NVIDIA Isaac Sim with Robot Operating System 2 (ROS2) orchestration, a Proximal Policy Optimization (PPO) multi-agent reinforcement learning (MARL) controller, and a fifth-generation (5G) link evaluation pipeline based on ns-3/5G-LENA key performance indicators (KPIs). Two UGVs execute mission-directed navigation toward a hazard region, while two UAV relays and a gNB-like aerial anchor adapt their positions to sustain end-to-end service under line-of-sight and non-line-of-sight transitions induced by urban occlusions.

Preliminary simulation results validate end-to-end operability and provide quantitative evidence of simultaneous mission progress and network continuity. Across a representative episode, the minimum distance to the hazard-region center decreases from 27.9 m to 1.55 m (final 1.80 m), while latency remains in a low regime (mean 4.88 ms, p95 8.17 ms). Packet loss is bounded (mean 3.5% and 2.2% for the two UGVs), and outages are sparse (101 steps over 9000), even during partial traversal of building-dense areas. The platform enables systematic diagnosis of mobility-connectivity coupling and supports transfer-oriented refinement of relay control and coordination policies.

MOTIVATION AND CONTRIBUTION

Search-and-rescue (SAR) missions increasingly rely on heterogeneous robotic teams under tight time constraints[1]. Unmanned ground vehicles (UGVs) perform inspection and mapping, while unmanned aerial vehicles (UAVs) provide sensing and communication support. In dense urban environments, building occlusions and “urban canyon” propagation trigger frequent transitions between line-of-sight (LoS) and non-line-of-sight (NLoS) conditions[2], causing throughput drops, bursty outages, and latency inflation. Mission execution therefore depends on timely delivery of camera streams, odometry, and map updates, not solely on local control.

This article contributes an end-to-end, closed-loop simulation framework that makes this coupling measurable and actionable. A city-scale simulation runs in NVIDIA Isaac Sim and is orchestrated via Robot Operating System 2 (ROS2), while a Proximal Policy Optimization (PPO) multi-agent reinforcement learning (MARL) policy coordinates two UGV explorers and UAV communication assets (relays and a gNB-like anchor). Network key performance indicators (KPIs) are evaluated alongside mobility, enabling systematic diagnosis of mobility-connectivity failure modes.

Andrés Meseguer-Valenzuela is with Instituto Tecnológico de Informática (ITI), Spain (e-mail: ameseguer@iti.es)

SYSTEM AND SIMULATION OVERVIEW

The experimental setup is conceived as a closed-loop system that couples physics-based urban simulation, middleware orchestration, multi-agent decision making, and radio-performance instrumentation to study mobility-connectivity coupling in SAR-style operations. A city-scale three-dimensional (3D) simulation is executed in NVIDIA Isaac Sim, where two UGVs act as mission executors and three UAVs provide connectivity support through two relay platforms and a gNB-like aerial anchor.

The simulated environment is defined as a stage with structured road corridors and building blocks that induce frequent LoS and NLoS transitions, thereby creating mobility-driven fluctuations in link quality representative of dense urban deployments. The scenario addressed by the system is a target-reaching SAR navigation problem under communication constraints. At the start of each episode, both UGVs are initialized within a designated start region, depicted as a green zone in Fig. 1, and are required to traverse the urban street network toward a predefined target region that represents the mission objective (e.g., a hazard or inspection area). Mission progress is therefore expressed in terms of the UGVs’ convergence to, and eventual entry into, the target zone while operating in the presence of occlusions that intermittently degrade end-to-end connectivity.

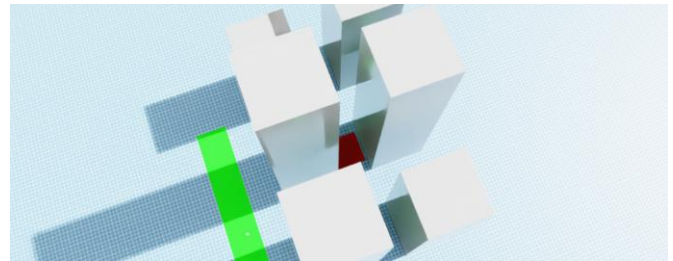


Figure 1. Simulation environment

The UAVs are tasked with supporting this progression by adapting their positions to maintain network service as the UGVs move deeper into occluded areas, effectively shaping a relay-assisted aerial backbone. ROS 2 provides the middleware layer for time-aligned state streaming and command dispatch during live episodes, enabling continuous interaction between simulation, control, and measurement components.

Decision making is implemented via a PPO multi-agent reinforcement learning policy, complemented by a supervisory layer that enforces safety constraints and limits physically implausible behaviors in early-stage trials. Connectivity is assessed through a fifth-generation (5G) link evaluation pipeline based on ns-3 and 5G-LENA simulations, which

returns end-to-end KPIs such as throughput, latency, outage, and packet loss aligned with agent mobility. Fig. 2 depicts the connectivity between the different platforms.



Figure 2. Simulation environment

The operational objective is to preserve connectivity while the UGVs move from the start zone to the target zone, with UAV relay and anchor motion controlled by the PPO policy so that communication performance remains sufficient for timely delivery of mission-relevant telemetry during traversal.

SIMULATION RESULTS

Preliminary experiments provide quantitative evidence that the proposed closed-loop stack operates end to end and supports simultaneous mission progress and connectivity monitoring under urban occlusions. Fig. 3 displays how connectivity behaves, highlighting latency and data loss while distance between the robots increases. Across representative episodes, the integrated pipeline runs continuously, enabling synchronized logging of mobility and KPIs.



Figure 3. Simulation results

Mission progress is reflected by sustained convergence toward the target hazard region. The minimum distance to the hazard-region center decreases from approximately 27.9 m at initialization to 1.55 m, with a final value of 1.80 m, indicating substantial mission-directed motion toward the objective zone. Connectivity remains operational throughout this progression. End-to-end latency stays in a low regime overall, with a global mean of 4.88 ms and a 95th percentile of 8.17 ms; short-lived peaks occur during transient degradations but do not dominate the episode behavior. Packet-loss remains bounded and consistent with these trends, with mean loss of 3.5% for UGV0 and 2.2% for UGV1 (95th percentiles of 20.0% and 13.0%, respectively).

Notably, these results are obtained while the UGVs partially traverse building-dense regions—quantified through inside-obstacle ratios of 5.9% for UGV0 and 31.3% for UGV1—where non-line-of-sight conditions are more likely and radio performance is expected to worsen. Even under these occlusion-prone conditions, connectivity is largely maintained, with 101 outage steps observed over 9000 total steps. This pattern suggests that degradations are predominantly bursty rather than sustained, consistent with short interruptions associated with relay/path reconfiguration rather than prolonged communication collapse.

CONCLUSIONS

The presented closed-loop framework demonstrates the feasibility of jointly evaluating mission progress and end-to-end connectivity for SAR scenarios with heterogeneous UGVs and UAVs. Live experiments confirm that mission-directed motion toward a target hazard region can be achieved while maintaining generally stable communication performance under occlusion-prone conditions, and the collected telemetry provides actionable evidence that remaining degradations are primarily transient and suitable for targeted stabilization. As a result, the current platform constitutes a practical baseline for transfer-oriented refinement of connectivity-aware autonomy in urban environments.

Near-term work focuses on four concrete actions. First, aerial connectivity control will be stabilized by adding hysteresis and minimum dwell-time rules to relay/path switching, together with stronger hold behavior when line-of-sight conditions are already suitable. Second, UGV control will be retuned to increase effective forward motion toward the hazard zone while maintaining obstacle-safe behavior and reducing stagnation. Third, runtime settings will be optimized to improve simulation-time throughput and support longer evaluation horizons per run. Finally, controlled ablation studies will be executed to quantify how supervision settings, pacing mode, and radio-penalty parameters affect outage bursts, latency peaks, and packet-loss levels.

REFERENCES

- [1] C.-L. Lu *et al.*, “A Heterogeneous Unmanned Ground Vehicle and Blimp Robot Team for Search and Rescue using Data-Driven Autonomy and Communication-Aware Navigation,” *Field Robot.*, vol. 2, pp. 557–594, May 2022, doi: 10.55417/fr.2022020.
- [2] A. Saboor, Z. Cui, E. Vinogradov, and S. Pollin, “Air-to-Ground Channel Model for Pedestrian and Vehicle Users in General Urban Environments,” *IEEE Antennas Wirel. Propag. Lett.*, vol. 24, no. 1, pp. 227–231, Jan. 2025, doi: 10.1109/LAWP.2024.3492507.

ACKNOWLEDGEMENTS

Research activities supported by Red de Excelencia Cervera para la Investigación en Tecnologías Avanzadas para la Cognición, simulación Táctica y entrenamiento adaptativo en escenarios críticos a través de Interfaces Extendidas (CER-20251027) funded by Ministerio de Ciencia, Innovación y Universidades a través del Centro para el Desarrollo Tecnológico y la Innovación (CDTI).