WS 2025-2026, SMT-ROB, DISAL-MP60 Project Summary

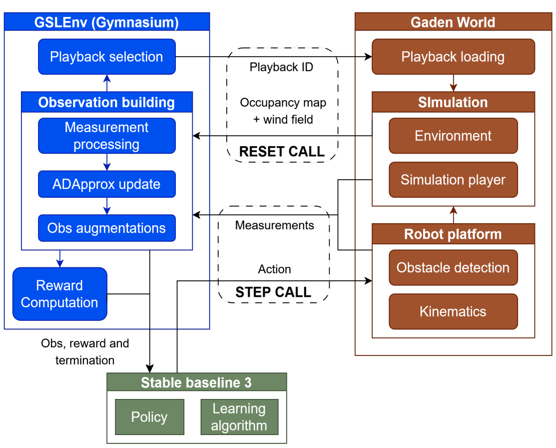
Reinforcement Learning for Gas Source Localization

Badil Mujovi

Professor : Alcherio Martinoli
Assistant(s) : Nicolaj-Bösel Schmid

Gas source localization (GSL) is the problem of identifying the position of a gas release using sparse concentration measurements collected by a mobile robot. In indoor environments, this task is particularly challenging due to turbulent airflow, complex obstacle layouts, and the indirect relationship between measurements and the true source location. While physics-based gas distribution mapping algorithms can provide accurate source estimates, they do not specify how a robot should actively navigate in order to collect informative measurements.

In this thesis, reinforcement learning is used to learn navigation strategies that actively exploit a physics-based gas distribution mapping algorithm, ADApprox. Instead of replacing the mapping algorithm, the proposed approach embeds reinforcement learning in a closed loop with ADApprox: the learning agent selects actions that guide exploration and plume tracking, while ADApprox continuously estimates the gas source location from the collected measurements. This allows the agent to learn how to efficiently move through obstacle-dense environments while progressively refining the source estimate.

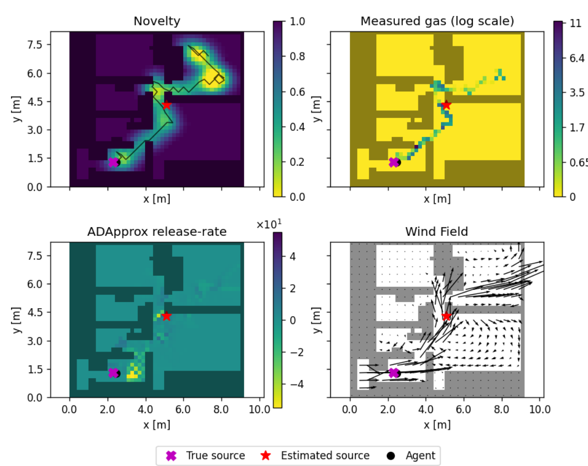


Training pipeline structure

A complete training pipeline was developed using a realistic simulation framework. Gas dispersion is simulated using GADEN-RT, a filament-based gas dispersion simulator with real-time performance, combined with pre-computed CFD wind fields. The learning agent is trained using the Soft Actor-Critic algorithm with a multi-input neural network policy that processes spatial map-based information and local scalar observations. The reward formulation explicitly separates exploration and plume-search phases, encouraging the agent to first discover the gas plume and then follow it upwind toward the source.

The proposed method was evaluated across multiple indoor environments with varying obstacle layouts and wind configurations. The trained policy achieves high localization success rates and demonstrates stable performance across multiple random seeds. Performance remains consistent across different dataset splits, indicating that the learned behavior generalizes across environments rather than overfitting to specific wind or source configurations.

An ablation study analyzes the contribution of individual reward components and observation inputs, showing that several dense rewards and map-based observations substantially facilitate learning, while the removal of specific elements leads to measurable performance degradation. Sensitivity analyses further show that the method is robust to changes in parameters. These results confirm that the learned behavior is not tied to a narrow parameter regime.



Example of an episode

Qualitative analysis complements the quantitative results by revealing structured and interpretable navigation strategies. The agent exhibits efficient exploration of previously unvisited regions during early stages of an episode, followed by upwind plume-following once gas is detected. In cluttered environments, the policy adapts its motion to navigate narrow passages, including temporary downwind movements when required, while maintaining informative sampling behavior. This illustrates how reinforcement learning and physics-based estimation jointly shape effective gas source localization strategies.

Overall, the results demonstrate that reinforcement learning can effectively complement physics-based gas distribution mapping by learning active navigation strategies for gas source localization in realistic indoor settings.