Reinforcement Learning for Gas Source Localization

Emilien Coudurier

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

Gas Source Localization (GSL) is an important field of mobile robotics, notably for environmental monitoring and industrial safety purposes. While several types of strategies emerged in the last decade, this project focuses on the extension of the ADApprox [1] algorithm, originally a Gas Distribution Mapping (GDM) algorithm applied to GSL which was shown to outperform state-of-the-art methods. Its lack of proper path planning and testing in cluttered environments led to a follow-up master thesis [2] which implemented a Deep Reinforcement Learning (DRL) pipeline with ADApprox as input. This work, however, assumes ground truth accessibility by the agent, inducing a simulation-to-reality gap. This semester project aims at addressing this issue, specifically by enhancing the pipeline with sensor models, noise and domain randomization (see Figure 1) to obtain a policy robust to model mismatch.

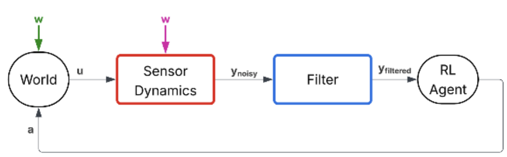

Figure 1: Proposed robust pipeline.

The gas sensor is modelled using a switching first-order system, mimicking the typical nonlinear response found in Metal-Oxide (MOX) sensors. With the intention of providing realistic readings, the sensor output is processed with a multiplicative gaussian noise, and the model is fitted on the logarithmic version of the gas power law. The anemometer is also modelled as a first-order system, with decoupled x and y components. A bounded random-walk noise, depicted by an Orstein-Uhlenbeck process is added on top of its response. The gas filament compiled by GADEN has randomized parameters, including among others temperature, pressure, filament growth and emission rate. The whole procedure is made modular and easily configurable, for convenient use in the future.

Evaluation is assessed over a statistically significant number of episodes, and performances are discussed based on their 95% confidence interval. A double-sided Fisher's test is used to quantitatively validate differences. Parameters randomized between nominal bounds and out-of-bounds are used separately to assess generalization to unknown contexts. Results show an overall low drop in successful gas source position identification (maintained above 80%) when adding them into the pipeline, while having a significant increase in robustness in randomized, unknown environments (up to +20% compared to the original baseline). A proper ablation study on each process, including an added Exponential Moving Average (EMA) filter is conducted, indicating that a randomized environment during training, as well as sensor models, participate significantly to improving the overall robustness. The use of the EMA filter is shown non-necessary, as the agent seems to be able to process raw data and obtain similar performances. Sensor noise is not highlighted as significant, inducing that the amount of noise introduced by other processes is sufficient to obtain a robust policy.

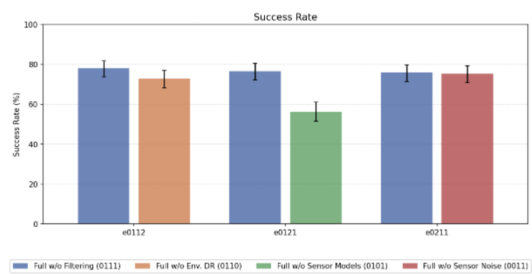

Figure 2: Success rates for out-of-bounds evaluations. In blue is the proposed full, robust pipeline, in orange the agent trained in a fixed environment, in green the agent trained without sensor models and in red the agent trained without sensor noise.

This work concludes with the proposed new robust baseline. It provides a realistic simulation context regarding sensor dynamics, advancing towards physical deployment of the trained policy. Its flexible code structure offers an easy fine-tuning process.