M2 research internship proposal

Non-stationary and robust Reinforcement Learning methodologies for drones detection

Internship supervisors:

- *Leïla Gharsalli*, EC IPSA, leila.gharsalli@ipsa.fr .
- *Stefano Fortunati*, EC IPSA/L2S, stefano.fortunati@l2s.centralesupelec.fr .

Guest institution:

- *Laboratoire des signaux et systèmes* (L2S), Bât IBM, Rue Alfred Kastler, 91400 Orsay.

Important dates:

- *Deadline for the submission of candidature*: until mid-January,
- *Duration of the stage*: from 4 to 6 months,
- *Period of the stage*: from February to September 2024,

1 Description

Reinforcement Learning (RL) methodologies are currently adopted in different context requiring sequential decision-making tasks under uncertainty [1]. The RL paradigm is based on the perception-action cycle, characterized by the presence of an agent that senses and explores the unknown environment, tracks the evolution of the system state and intelligently adapts its behavior in order to fulfill a specific mission. This is accomplished through a sequence of actions aiming at optimizing a pre-assigned performance metric (reward).

Despite of their wide applicability, classical RL algorithms are based on a cumbersome assumption: the stationarity of the environment, i.e. the statistical and physical characterization of the scenario, is assumed to be time-invariant. This assumption is clearly violated in surveillance application, where the position and the number of targets, along with the statistical characterization of the disturbance may change over time. To overcome this limitation and include the non-stationarity in the RL framework, both theoretical and application-oriented non-stationary approaches have been proposed recently in the RL literature (e.g. [2,3]). The application of these non-stationary-based line of research to robust radar detection problems has been investigated in [4–6]. Moreover, a PhD project has been also founded by IPSA to develop original non-stationary RL detection algorithm.

The aim of this internship is then to support and complete the ongoing research activity by testing and validating the non-stationary RL algorithms on several realistic scenarios where the radar acts as an agent that continuously senses the unknown environment (i.e., targets and disturbance) and consequently optimizes transmitted waveforms in order to maximize the probability of detection (PD) by focusing the energy in specific range-angle cells. Due to their crucial strategical interest, particular attention will be devoted to scenarios containing drones.

2 Mission

In coordination with the PhD student working on this project, the intern will firstly review the relevant existing literature on stationary and non-stationary RL methodologies as well as on radar multi-target detection in order to better understand both the scientific and application context of this work. Consequently, the intern will propose several realistic scenarios related to target detection surveillance applications to be tested. For these test cases, particular attention will be given to the abstract notions of agent, actions, states and reward that will be specialized in the radar detection framework [5,6]. The non-stationarity of the environment will be characterized by variation of the the number and Signal-to-Noise Ratio (SNR) of targets/sources to be detected along with the statistical characterization of the disturbance. Finally, the validation of the algorithmic solution obtained by the proposed advanced RL methodologies will allow to focus on a much more specific challenging problem, highly non-stationary and physically variable: the detection of drones.

Required profile:

Master 2 or equivalent in machine learning / applied mathematics / statistical signal processing or any related field.

Location:

L2S Laboratory (Laboratoire des signaux et systèmes,UMR8506) - Modeling and Estimation Group (GME) in the Signals and Statistics group.

Duration:

6 months starting from February 2024.

References

[1] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction.* The MIT Press, second ed., 2018.

[2] E. Lecarpentier and E. Rachelson, "Non-stationary markov decision processes, a worst-case approach using model-based reinforcement learning," *Advances in neural information processing systems*, vol. 32, 2019.

[3] S. Padakandla, K. J. Prabuchandran, and S. Bhatnagar, "Reinforcement learning algorithm for non-stationary environments," *Applied Intelligence*, vol. 50, p. 3590?3606, 2020.

[4] S. Fortunati, L. Sanguinetti, F. Gini, M. S. Greco, and B. Himed, "Massive MIMO radar for target detection," *IEEE Transactions on Signal Processing*, vol. 68, pp. 859–871, 2020.

[5] A. M. Ahmed, A. A. Ahmad, S. Fortunati, A. Sezgin, M. S. Greco, and F. Gini, "A reinforcement learning based approach for multitarget detection in massive MIMO radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 5, pp. 2622–2636, 2021.

[6] F. Lisi, S. Fortunati, , M. S. Greco, and F. Gini, "Enhancement of a state-of-the-art RL-based detection algorithm for Massive MIMO radars," *IEEE Transactions on Aerospace and Electronic Systems (accepted)*, 2022.