Power Grid Control with Graph-Based Distributed Reinforcement Learning

Carlo Fabrizio*, Gianvito Losapio*, Marco Mussi, Alberto Maria Metelli, and Marcello Restelli
{carlo.fabrizio, gianvito.losapio, marco.mussi, albertomaria.metelli,marcello.restelli}@polimi.it

Objective

Design a **distributed reinforcement learning controller** that **dynamically adjusts loads** over power lines

Motivations

- **Unmanageably large action space** due to:
  - Proliferation of renewable energy sources;
  - Increasing size of the controlled power grid.
- **Now there are obsolete traditional controllers** which strongly rely on:
  - Human-in-the-loop interventions;
  - Computationally intensive optimization algorithms.

Requirements

- **Scalability**: Enable RL to control *large power grids* by decomposing the problem over multiple agents.
- **Performance**: Achieve longer grid survival time compared to standard baselines.
- **Computational Efficiency**: Guarantee lightweight inference and training for real-time deployment.

Contributions

- Novel **Graph-based Distributed Deep Reinforcement Learning** (G2DRL) algorithm **splitting both the action and state spaces** among several agents:
  - **Innovative framework** for **graph-like data creation** from a power grid snapshot;
  - **Shared Graph Neural Network** (GNN) to process graph-like data, promoting **information sharing** among the agents.
- **Practical techniques to speed-up the convergence** in this graph-based multi-agent scenario:
  - **Imitation Learning – Deep Q-Learning from Demonstrations** (DQfD);
  - **Potential-based reward shaping – bootstrapped reward shaping**.

Algorithm: Graph-based Distributed Deep RL (G2DRL)

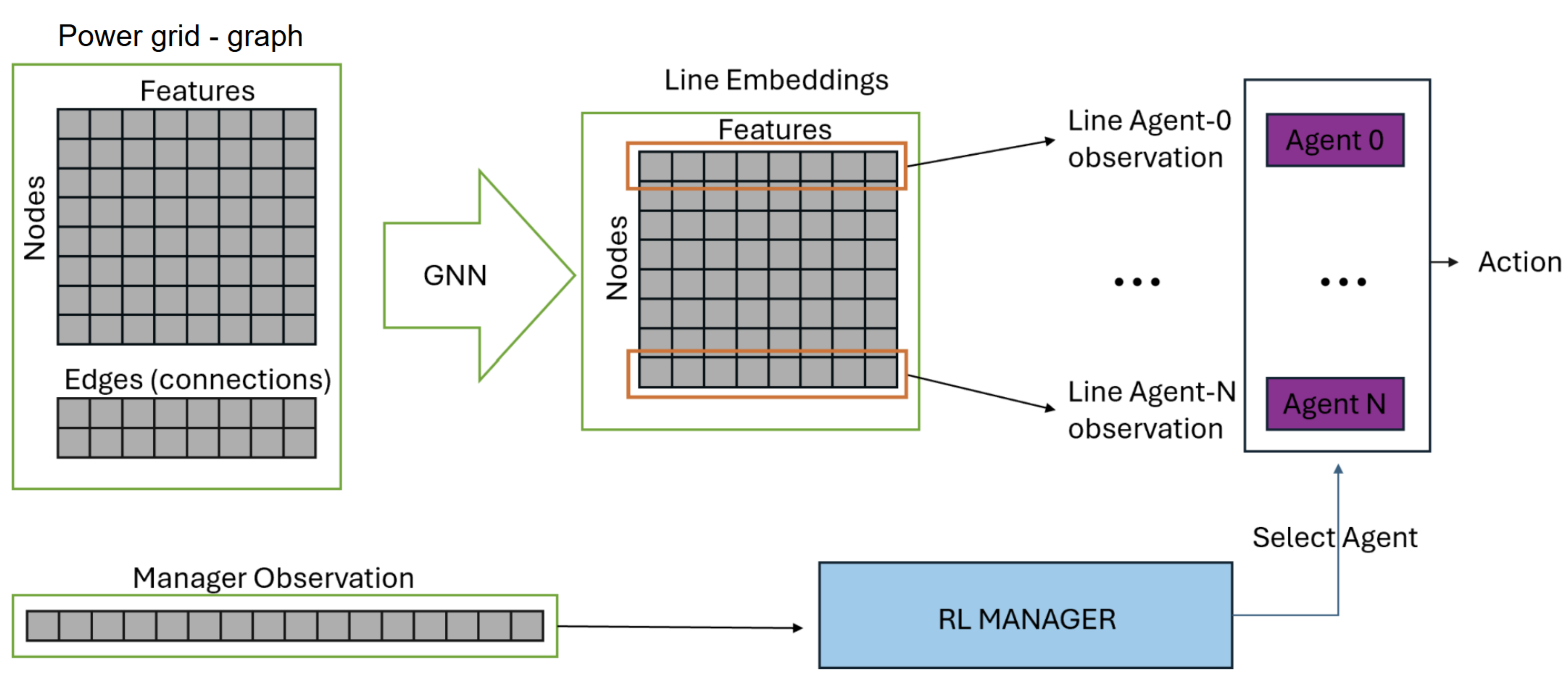
Model's components:

- Agents & Manager → **Deep Double Dueling Q-Learning** agents with **Prioritized Experience Replay Buffer**.
- GNN → **Graph Attention Network v2**.

Algorithm 1: G2DRL computational flow

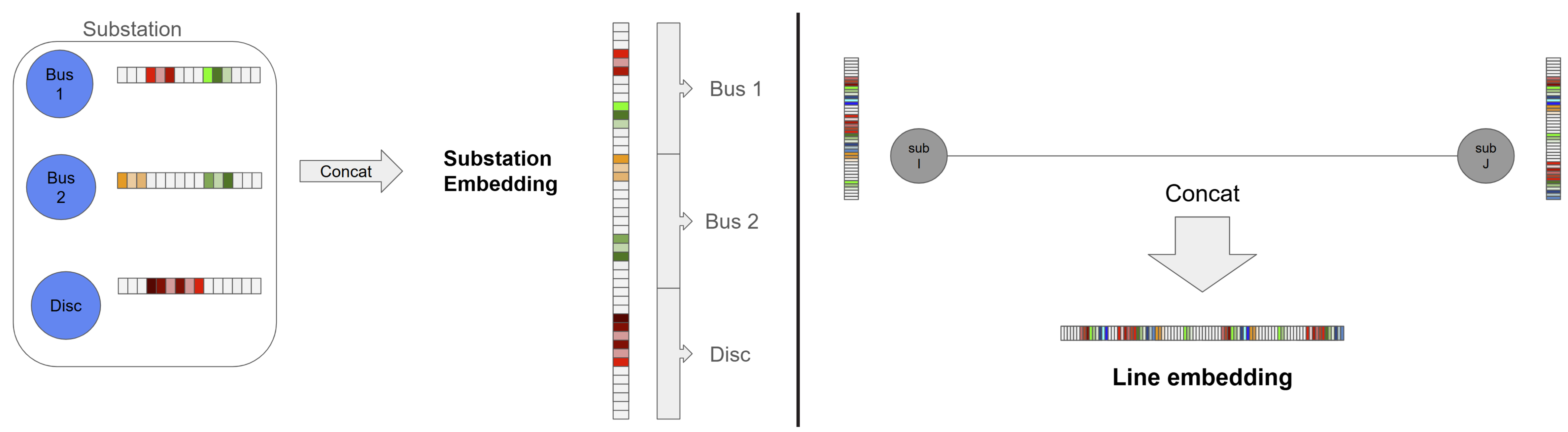
if *Danger situation* **then**
  $i \leftarrow$ Manager.select_agent$(s_t|\mu^t)$
  $g_t \leftarrow$ convert_graph(s_t)
  $o_t \leftarrow$ GNN$(g_t \mid \phi^t)[i]$
  $a_t \leftarrow$ Agent$_i$.policy_play$(o_t \mid \theta_i^t, \alpha_i^t, \beta_i^t)$
  return a_t

High-Level Overview



Data Construction: Power Grid as a Graph

Homogeneous graph representation of a power grid



Experiments

Setting

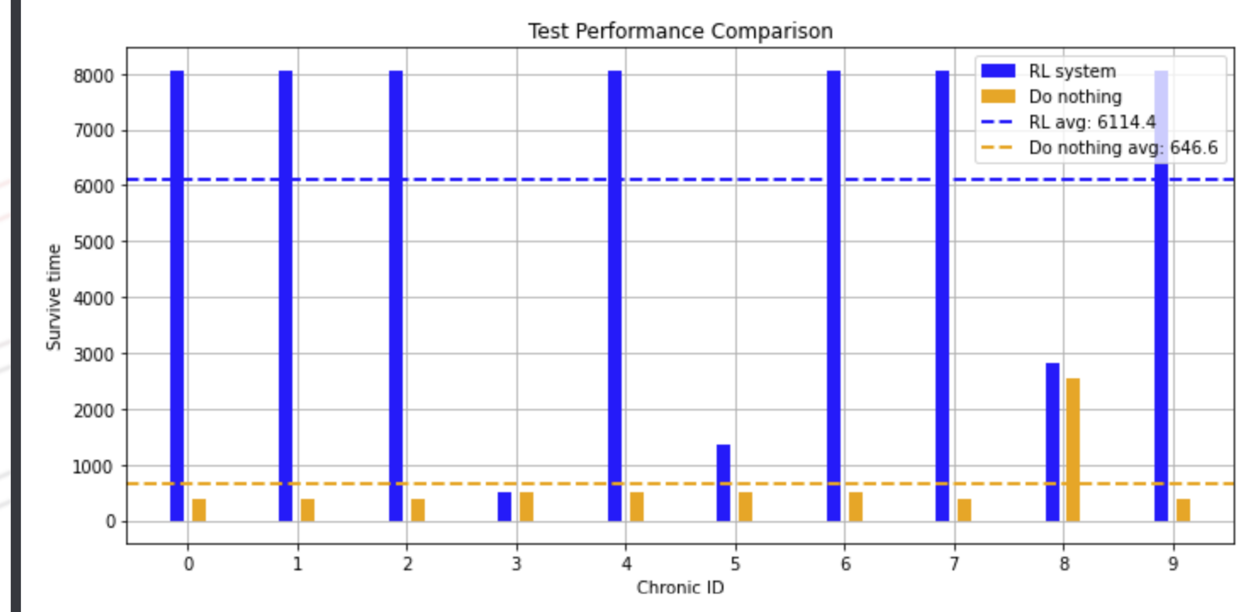
Grid2Op Simulator on the *"l2rpn_case14_sandbox"* environment (12 stations and 20 links) using:

- 984 training episodes
- 10 validation episodes
- 10 test episodes

Baseline: *Do-nothing* agent

Results

Test-Set Performance



Ablation Study

Configuration	Test Performance
Basline	646.6
No DQfD	1877.9
No GNN	785.2
No Reward Shaping	5324.3
Complete System	**6114.4**

Computational Efficiency

Model	Inference Time (avg ± std) [s]
Proposed Model	0.187 ± 0.145
Expert Agent	2.56 ± 0.223

References

Matthijs de Jong, Jan Viebahn, and Yuliya Shapovalova. Generalizable graph neural networks for robust power grid topology control. *arXiv preprint arXiv:2501.07186*, 2025.

Todd Hester, Matej Vecerik, Olivier Pietquin, Marc Lanctot, Tom Schaul, Bilal Piot, Dan Horgan, John Quan, Andrew Sendonaris, Gabriel Dulac-Arnold, Ian Osband, John Agapiou, Joel Z. Leibo, and Audrunas Gruslys. Deep q-learning from demonstrations. In *AAAI conference on artificial intelligence*, 2018.

Antoine Marot, Benjamin Donnot, Gabriel Dulac-Arnold, Adrian Kelly, Aidan O'Sullivan, Jan Viebahn, Mariette Awad, Isabelle Guyon, Patrick Panciatici, and Camilo Romero. Learning to run a power network challenge: a retrospective analysis. In *NeurIPS Competition and Demonstration Track*, 2020a.

Antoine Marot, Benjamin Donnot, Camilo Romero, Balthazar Donon, Marvin Lerousseau, Luca Veyrin-Forrer, and Isabelle Guyon. Learning to run a power network challenge for training topology controllers. *Electric Power Systems Research*, 2020b.