Auto-Regulated Traffic Signal Control in Multi-Modal Urban Networks Using Graph-Based Deep Reinforcement Learning

Context and Motivation

Traffic signal control (TSC) is a cornerstone of urban traffic management, directly impacting traffic efficiency, network stability, and environmental performance [1]. Over the past decade, adaptive and intelligent TSC approaches have become essential tools to mitigate congestion. These methods adjust signal timings based on real-time traffic conditions, helping to reduce delays and improve throughput. Among these approaches, reinforcement learning (RL), particularly deep RL (DRL), has emerged as a promising paradigm capable of capturing complex traffic dynamics through interaction with the environment [2].

In real-world traffic networks, intersections are interdependent: the conditions at one intersection are influenced by upstream inflows and downstream congestion, forming tightly coupled spatial dependencies. This complexity becomes more pronounced when multiple intersections share major traffic flows or transit routes. As such, isolated signal optimization is insufficient. Recent work has explored multi-agent reinforcement learning (MARL) to coordinate control across multiple intersections through distributed agents. These decentralized methods offer scalability and robustness but require careful coordination strategies to avoid myopic or conflicting decisions [3].

An important open issue concerns: (i) how intersections can effectively exchange and process relevant information, then (ii) to what extend an intersection is interlinked with other intersections [4]. In most practical deployments, controllers use data from signalized intersections only, without accounting for the impact of non-signalized nodes (such as roundabouts or priority-to-the-right) within the same area. These configurations are common in urban networks and can significantly affect the dynamics at nearby controlled intersections. Typically, this problem might be interpreted as a Partial Observability issue, similar to the ones encountered when deploying agents in real world scenarios. There is thus a need to develop models that can capture heterogeneous neighborhood effects by identifying which nearby nodes influence a given intersection and integrating only the relevant neighborhood information into decision-making [4, 5].
Additionnally, when the surrounding is incorporated into the agent controlling an intersection [4], the process is usually mono-directionnal, i.e. the information goes only one-direction to feed the decision process of the agent, without introducing a truly mutual relationship toward cooperativeness. Consequently, an agent controlling an intersection operate thanks to an augmented perception enabled by its surrounding, but usually maintain selfish decisions with few consideration for its impact on its surrounding.

Furthermore, even with sophisticated multi-agent control, oversaturated conditions (e.g., during rush hours or large public events) can lead to gridlock and system collapse due to spillback effects. To address this, the concept of perimeter control has been proposed in the literature [6], which involves restricting vehicle inflows into a high-demand area to preserve its internal flow. However, most existing approaches rely on static boundaries and centralized coordination, limiting their scalability, transferability and adaptability to real-time changes. There is a clear need for adaptive, agent-driven perimeter protection capable of dynamically identifying and regulating protected zones based on local observations and decentralized operations [7].

Achieving independent agents equipped with skills restricted to local perception and control, but able to exchange information with agents in its surroundings towards cooperativeness remains an opened question. This is a key stage toward the emergence of self-organised and pro-active urban traffic management strategies with the perspective of spatially dynamic protected networks.

Lastly, managing multi-objectivity and multi-modality of traffic is becoming increasingly essential. Urban intersections handle a diverse mix of users, including private vehicles, freight, bicycles, pedestrians, and public transit. Buses, in particular, are sensitive to signal timing and congestion and require headway regularity to avoid bunching and deliver reliable service. Yet, integrating bus headway objectives into RL-based control and rewards remains a challenge. Although a few recent studies [8] have begun to address multi-modality in learning-based TSC, most approaches still fail to model real-world bus dynamics, such as open-loop operations and heterogeneous passenger demand. Beyond the incorporation of multi-modality, the multi-objectivity in TSC comes into play through the consideration of a large set of efficiency Key Performance Indicators encompassing traffic, safety and pollutant emissions. Among others, the gas emissions are of paramount importance for most of the urban areas and cities. Consequently, we suggest applying a similar framework to explore selective access to limited zones by creating dynamic protected areas based on pollution levels. This would involve permission mechanisms for restricted vehicle types depending on real-time emissions data.

No existing approach fully combines these aspects: multi-modal control, dynamic perimeter flow protection, graph-based coordination, and resilience to incomplete data or heterogeneous infrastructures in one coherent framework. This PhD thesis proposal aims to fill that gap by developing an advanced traffic signal control framework that leverages learning-based, graph-aware multi-agent architectures to optimize performance in multi-modal, oversaturated, and heterogeneous urban networks. The ultimate goal is a holistic, scalable urban traffic control approach that improves both network efficiency and transit service reliability, pushing beyond the current state of the art in smart mobility management.

Objectives and scientific challenges

To address the above challenges, this thesis proposes to pursue the following objectives:

1. **Multi-Objective, Multi-Modal TSC Optimization:** Design a reinforcement learning framework where each intersection agent optimizes for multiple goals: reducing overall vehicle delay/queuing, enhancing bus headway regularity and controlling pollutant emissions. The controller will explicitly include gas emission objectives and bus transit performance (e.g. headway deviation or punctuality) in its reward function, alongside traditional traffic efficiency metrics, to ensure a balanced improvement of both general traffic flow and bus service reliability

   Challenges - *Toward Multi-Objective Reward Design:* Balancing the comprehensive multi-modal demand (including car traffic efficiency), the gas emission according to transportation means and the public transit reliability within a unified reward framework.

2. **Efficient Graph-Based Communication for Local Coordination:** Ensure that the MARL system can cope with Partial Observability of its environment. The challenge lies in achieving a proper recognition of oversaturation, despite the restricted perception. Then, the

agent has to act cooperatively to prevent gridlock. Rather than hard-coding a gating logic, we aim for emergent perimeter protection: the agents should learn to occasionally hold traffic at boundary intersections (creating metering) when the downstream network is saturated, thereby mimicking perimeter control.

Challenges - *State Representation and Neighborhood Modeling:* Integrating heterogeneous neighborhood information using graph structures and attention mechanisms. Selecting the optimal spatio-temporal graph embedding and adjacency matrix representation to obtain representative local and global intersection observations within a specific latent space.

3. **Dynamic Perimeter Protection via Emergent Cooperation:** Enable agents to detect local congestion and collaboratively form dynamic protected zones. These zones will be delimited by specific intersections that collectively agree on the need to establish a protected area around them. Neighboring intersections located at the periphery of the zone will take over to help relieve internal pressure. The creation and management of such zones will be carried out in a fully dynamic and autonomous manner, coordinated by and between the intersection agents.

   Challenges - *Emergence of Protected Zones:* Developing agent policies that can trigger, manage, and dissolve protection zones in response to real-time congestion signals.

The thesis aims to develop a scalable and decentralized traffic signal control system that enhances both traffic flow efficiency and public transit reliability in urban areas. This system will be designed to function in realistic, infrastructure-supported environments and to provide effective protection of urban zones in situations of network saturation. To support these practical objectives, the thesis also seeks to advance the theoretical foundations of multi-agent reinforcement learning by addressing multi-objective coordination under partial observability, using graph neural networks to enable efficient and cooperative control. By jointly considering traffic efficiency and public transit reliability, and by enabling decentralized, context-aware coordination among intersections, the thesis aims to develop urban traffic networks more resilient, fair, and adaptive.

Beyond its core objectives, this work also opens the door to optimizing sensor deployment. The use of attention mechanisms offers a basis for identifying the most informative intersections, paving the way for cost-effective and efficient sensor placement strategies.

Candidate Profile and Application

We are looking for a motivated PhD candidate with the following background:

- **Academic Background:** Master's degree or engineering diploma in Computer Science, Artificial Intelligence, Applied Mathematics, or related fields.
- **AI and Modeling Skills:** Solid knowledge in deep learning (PyTorch or TensorFlow), with interest or initial experience in reinforcement learning; familiarity with graph neural networks (e.g., GCNs) is appreciated.
- **Traffic Simulation Tools:** Experience with traffic simulators (e.g., SUMO, CityFlow) is a plus.

- **Research Skills:** Strong analytical mindset, autonomy, scientific curiosity, and good communication skills in English.

To apply: Please submit your application via the following link. `https://docs.google.com/forms/d/e/1FAIpQLSfL8J9zItyc-lQuXOcTQOiN86QCZfA9JAy9wJUDS32NatIaBw/viewform?usp=dialog`

Working Environment

The PhD will be conducted within the **LICIT-ECO7 laboratory**, a joint research unit between the ENTPE engineering school and Université Gustave Eiffel, located in the Lyon metropolitan area. The lab is renowned for its work on intelligent transportation systems, traffic management, and mobility modeling, combining expertise from both civil engineering and artificial intelligence. More information about the lab is available at: `https://licit-eco7.fr`

References

[1] Wei Miao, Long Li, and Zhiwen Wang. A survey on deep reinforcement learning for traffic signal control. In *2021 33rd Chinese Control and Decision Conference (CCDC)*, pages 1092–1097. IEEE, 2021.

[2] Haiyan Zhao, Chengcheng Dong, Jian Cao, and Qingkui Chen. A survey on deep reinforcement learning approaches for traffic signal control. *Engineering Applications of Artificial Intelligence*, 133:108100, 2024.

[3] Chengdong Ma, Aming Li, Yali Du, Hao Dong, and Yaodong Yang. Efficient and scalable reinforcement learning for large-scale network control. *Nature Machine Intelligence*, 6(9):1006–1020, 2024.

[4] Jing Shang, Shunmei Meng, Jun Hou, Xiaoran Zhao, Xiaokang Zhou, Rong Jiang, Lianyong Qi, and Qianmu Li. Graph-based cooperation multi-agent reinforcement learning for intelligent traffic signal control. *IEEE Internet of Things Journal*, 2025.

[5] Yiming Bie, Yuting Ji, and Dongfang Ma. Multi-agent deep reinforcement learning collaborative traffic signal control method considering intersection heterogeneity. *Transportation Research Part C: Emerging Technologies*, 164:104663, 2024.

[6] Xinfeng Ru, Weiguo Xia, Xiang Fan, and Tao Sun. Nearly optimal perimeter tracking control for two urban regions with unknown dynamics. *IEEE Control Systems Letters*, 2024.

[7] Jiajie Yu, Pierre-Antoine Laharotte, Yu Han, Wei Ma, and Ludovic Leclercq. Perimeter control with heterogeneous metering rates for cordon signals: A physics-regularized multi-agent reinforcement learning approach. *Transportation Research Part C: Emerging Technologies*, 171:104944, 2025.

[8] Jiajie Yu, Pierre-Antoine Laharotte, Yu Han, and Ludovic Leclercq. Decentralized signal control for multi-modal traffic network: A deep reinforcement learning approach. *Transportation Research Part C: Emerging Technologies*, 154:104281, 2023.