Towards Adaptive Traffic Management: A Learning Framework for Spatio-Temporal Signal Optimisation

Date: 11/05/2025
Week: 10
Author: Alvin B. Lin
Student ID: 112652040

Problem 1: The Future Capability of Artificial Intelligence

A major capability that current artificial intelligence systems cannot yet accomplish, but could emerge within the next two decades, is a *fully autonomous and adaptive traffic management system*. This AI would coordinate the entire urban road network in real time, learning how to adjust signal timings, vehicle routing, and emergency responses to optimise citywide mobility.

Unlike today's rule-based or pre-timed systems, such an AI would integrate heterogeneous data sources — live camera feeds, vehicular GPS data, public transport timetables, and crowd-sourced jam detections — to construct a coherent, evolving model of traffic density and flow. When an accident or congestion occurs, it would infer the incident's probability distribution $q(x, t, p)$, predict its downstream effects, and modify intersection timings dynamically to minimise total waiting time and pollution.

Significance. This development would be transformative for several reasons:

- Reduce congestion and emissions in dense cities, improving air quality and fuel efficiency.
- Enhance public safety by detecting incidents early and coordinating rapid rerouting.
- Extend to other infrastructures such as logistics networks or evacuation planning.

Beyond technical progress, such a system represents a shift from reactive control to *predictive governance*: the AI learns not merely to follow instructions, but to anticipate, reason, and balance competing priorities such as efficiency, fairness, and sustainability. It exemplifies a future where AI acts as a cooperative decision-maker embedded within human environments.

Problem 2: Type of Machine Learning Involved

The core of this system would be built upon *reinforcement learning (RL)*, augmented by supervised and unsupervised modules for perception and pattern discovery. Reinforcement learning is appropriate because the system continuously interacts with a dynamic environment — each intersection functions as an agent that observes a state o_t (traffic density, and incident probability), executes an action a_t (signal adjustment), and receives a reward r_t.

Data and Learning Feedback. The primary data sources include live traffic densities, vehicle velocities, and event probabilities inferred from navigation or sensor data. The learning objective is defined through a reward rather than labelled outcomes:

$$r_t = -\sum_i \left(p_i^{\text{wait}}(t) + \lambda_1 q_i(t)\right),$$

which encourages smooth flow and rapid recovery from disruptions.

Hybrid Learning Components. Supervised learning can support vision-based detection of congestion or accidents, while unsupervised learning helps identify latent flow patterns and cluster intersections with similar dynamics. Nonetheless, reinforcement learning remains the principal framework, since it enables continual adaptation through trial, feedback, and policy optimisation.

In essence, the model's success depends on its ability to perceive, act, and learn simultaneously within a highly interactive environment — a hallmark of the next generation of intelligent control systems.

Problem 3: Modelisation

Objective. Design a compact, well-defined *model problem* that serves as the first research step towards the long-term goal described in Problem 1: a fully autonomous, adaptive urban traffic management system. The model must capture the essential decision-making, perception and uncertainty aspects of the final capability while remaining small enough to allow rigorous experimentation and reproducible evaluation.

Model Problem/Statement. Consider a minimal road network in the shape of a **Tian**-grid (a 3×3 node grid). Vehicles are generated according to stochastic origin–destination pairs; the macroscopic vehicle distribution on each edge e is represented by a discrete-time density $p_e(t)$. Each intersection i has a binary signal state $s_i(t) \in \{0, 1\}$ indicating whether east–west (1) or north–south (0) movement is permitted. Accidents/jams occur with a state-dependent probability $q_e(t) = q(p_e(t), x_e, t)$ that is non-zero only on road edges. When an incident occurs on edge e, its effective capacity $C_e(t)$ is reduced for a random duration. The control objective is to choose signal sequences $\{s_i(t)\}$ to minimise expected cumulative delay over a finite horizon T.

How this simplified problem represents the final capability.

- **Perception and uncertainty:** modelling $q_e(t)$ encodes the need to perceive and quantify incident risk from noisy observations (mirroring crowd-sourced jam data in reality).
- **Decision under dynamics:** choosing $\{s_i(t)\}$ to influence $p_e(t)$ is a sequential control task capturing global coordination, trade-offs and delayed effects, which are core to the final system.
- **Scalability and structure:** the grid preserves spatial coupling and routing interactions while remaining small enough to allow exhaustive testing; solutions and methods should generalise to larger graphs.
- **Human-centric objectives:** the cost emphasises societal metrics (delay, fairness, recovery time) rather than purely local performance, aligning with the long-term aim of balancing efficiency and equity.

Testability Define a rigorous evaluation protocol with the following elements:

- **Metrics:** average travel time, average waiting time per vehicle, incident recovery time (time to restore pre-incident throughput), fairness index (variance of waiting times across approaches), and robustness (performance degradation under distributional shift).

- **Baselines:** compare against (i) fixed cycle timings, (ii) heuristic actuated control (e.g. max-pressure or vehicle-count based), and (iii) hybrid optimiser (MPC with short horizon).
- **Scenarios:** train and evaluate across a set of stochastic scenarios: nominal traffic, heavy rush, single-edge incident, multi-edge incident, and sensor noise / missing data. Use cross-validation style splits (train on subset of scenarios, test on withheld scenarios).
- **Success criteria:** a candidate method is successful if it (a) reduces average delay over baselines by a statistically significant margin across test scenarios, (b) restores throughput more quickly after incidents, and (c) maintains reasonable fairness.

Required mathematical and machine learning tools

- **Stochastic process and PDE/ODE tools:** discrete conservation laws (LWR discretisation), queuing models and Markov processes to describe density evolution and incident dynamics.
- **Reinforcement learning:** multi-agent or centralised RL algorithms (e.g. policy gradient methods, actor–critic, PPO or value-based methods) to learn sequential control under uncertainty; training techniques for sample efficiency and stability (experience replay, target networks, reward shaping).
- **Supervised / probabilistic learning:** perception modules (CNN/LSTM) for estimating $q_e(t)$ from synthetic sensor inputs; probabilistic calibration (e.g. Bayesian or ensemble methods) to quantify uncertainty.
- **Graph neural networks (GNNs):** to encode spatial relations between intersections and permit parameter sharing / message passing for scalable coordination.
- **Optimisation and control theory:** model predictive control (MPC) as a hybrid baseline and for warm-starting RL; constrained optimisation to enforce safety rules (minimum green times, pedestrian phases).
- **Simulation and validation tools:** a microscopic or mesoscopic traffic simulator (SUMO, CityFlow) for realistic data generation, and tools for domain randomisation to improve generalisation.

Practical experimental plan (brief).

1. Implement the **Tian**-grid in a simulator with stochastic OD flows and incident model.
2. Implement baselines (fixed, actuated, MPC).
3. Develop a perception head to estimate q from noisy observations (supervised/unsupervised training on simulated traces).

4. Train an RL controller with a GNN state encoder, enforce safety constraints, and evaluate under the protocol above.

5. Analyse failure modes, sensitivity to sensor noise, and transferability to larger grids.

Concluding remark. This model problem preserves the essential elements of the envisioned final capability — perception under uncertainty, sequential decision-making with delayed and spatially coupled effects, and the need for robustness and fairness — while remaining compact enough for rigorous study, reproducible experimentation and incremental extension to real city networks.