Discovering Effective Traffic Policies via Reinforcement Learning for Mixed Traffic Scenarios*

A study of Philippine Cities†

Mark Jethro Cortejo
De La Salle University
Manila, Philippines
mark_jethro_cortejo@dlsu.edu.ph

Dr. Briane Paul Samson
De La Salle University
Manila, Philippines
briane.samson@dlsu.edu.ph

ABSTRACT

Many cities in the world evolve rapidly with the inclusion of modern transportation methods as well as the growing population of people. Modern policies were supposed to be equipped to handle any volume of commuters. However, eventually policies give out and the newly encountered traffic behaviors starts to emerge. This paper is a study of the use of Reinforcement Learning in discovering effective policies for high-traffic junctions studied in the Philippines.

CCS CONCEPTS

• **Computing methodologies** → **Agent / discrete models**; **Reinforcement learning**;

KEYWORDS

ACM proceedings, LaTeX, text tagging

1 INTRODUCTION

1.1 Overview

The rapid increase in vehicle volume in the Philippines has exacerbated traffic congestion to unprecedented levels. In 2023, the country recorded 1.4 million registered vehicles, reflecting a significant 7.63% increase from the previous year. Over the years, the Metropolitan Manila Development Authority (MMDA) has implemented various traffic management policies to address these challenges. One of their most notable initiatives is the EDSA Carousel program, which allocates dedicated bus lanes separated from other lanes by barriers. This initiative has successfully reduced commute times for bus riders and encouraged more commuters to shift to public transportation.

However, despite its benefits, the program faces significant challenges. Accessing the carousel requires commuters to cross multiple lanes to reach the central bus stops, leaving them exposed to environmental elements and safety hazards. Additionally, unauthorized entry of private vehicles into the bus lanes further complicates traffic dynamics. While the MMDA continues to propose incremental solutions to alleviate congestion, these efforts often lack comprehensive studies and simulations to evaluate their long-term effectiveness.

Research shows that dynamic approaches to traffic management, supported by advanced simulations, can deliver superior results compared to static rule-based systems. This paper explores the potential of such methods in creating more efficient and effective traffic policies tailored to the unique challenges of the Philippines.

1.2 Objectives

1.3 Scope and Limitations

2 RELATED LITERATURE AND STUDIES

To optimize the flow of traffic in intersections, traffic light systems must be able to optimize their scheduling or control signals based on multiple factors. In this section, describes the methodologies applied to dynamically manage traffic light signals, ranging from traditional to machine learning approaches.

2.1 Traffic Light Control Systems

A recent study utilized reinforcement learning, particularly deep-policy gradient, and value-function-based reinforcement learning, to train the agent to select the most optimal control action. They were able to prove that their methods work in a traffic network simulated in the simulation of an urban mobility traffic simulator, without suffering from instability issues during the training process [3].

Another method via a value-based- deep reinforcement learning model, which is modified from the deep Q-Learning Network (DQN)algorithm, was implemented on adaptive signal control [6]. The agent was trained in a micro-simulator called VISSIM at an isolated intersection. This method generates an action that leads to an optimal estimated value within a finite set as the agent's policy. The study shows that the agent outperforms a fixed timing plan in all testing cases by reducing the system's total delay by 20%.

Graph Convolutional Neural Networks (GCNNs) can be applied to handle and oversee multiple traffic light controls. Nishi et al. conducted research where they utilized GCNNs in a multi-intersection network to automatically extract features and also take into account the traffic characteristics between distant roads by layering multiple neural network layers.[4]

A new approach was proposed that used semi-supervised double dueling broad reinforcement learning now in use in smart cities [5]. With their method, they significantly reduced waiting time by about 11.7% of the average waiting time to the comparative approach. The agent is rewarded for reducing its total cumulative waiting time of consecutive time steps.

Chu et al. proposed an algorithm called Multi-Agent Advantage Actor-Critic (MA2C), which is a multi-agent variant of the classical independent Advantage Actor-Critic (A2C). This study proposed dividing the overall traffic controls into multiple local RL agents. The

---

significant contribution of this proposal is to create an optimal and robust model for each agent by avoiding the large dimensionality of the environment. [1]

Recently, the attention mechanism from transformers was also introduced as a new Reinforcement Learning-based strategy for Large-Scale Adaptive Traffic Signal Controls. This approach integrates the attention mechanism into a multi-agent proximal policy optimization (MAPPO) RL model to enable more effective, scalable, and stable learning in a complex ATSC environment. The results by the author claim that the approach was able to learn stable and sustainable policies to achieve lower congestion and claimed to outperform other state-of-the-art approaches [2].

2.2 Traffic Simulators

Traffic simulators are essential tools for evaluating and optimizing traffic management strategies. Two widely used traffic simulators are Simulation of Urban MObility (SUMO) and Verkehr In Städten - SIMulationsmodell (VISSIM). SUMO is an open-source, highly portable, microscopic, and continuous road traffic simulation package designed to handle large road networks. It allows for the simulation of various traffic scenarios and supports the integration of external control algorithms. VISSIM, on the other hand, is a proprietary microscopic multi-modal traffic flow simulation software. It provides a detailed and realistic representation of traffic operations and is often used for traffic engineering and planning purposes. Both simulators are instrumental in testing and validating traffic control systems, including those based on reinforcement learning and other advanced methodologies.

2.3 Reinforcement Learning

3 CONCEPTUAL FRAMEWORK

3.1 Complex Systems

According to Newman (2011), a complex system is fundamentally defined as one comprised of many interacting parts often referred to as 'agents', that exhibit collective behavior distinct from the actions of their individual components.

Consider a flock of birds, their coordinated movements are a prime example of emergent behavior. Each individual bird is simply following a few basic rules: maintain a certain distance from its neighbors and align its direction with them. However, when many birds follow these simple rules simultaneously, a stunningly complex and beautiful pattern emerges—a swirling, shifting flock that seems to move as one entity. This collective motion isn't dictated by any central leader; it arises spontaneously from the interactions between individual birds, demonstrating how complex patterns can emerge from relatively simple local rules

Ant colonies provide a compelling example of emergent behavior. Individual ants operate with relatively simple rules: following pheromone trails, carrying food, and responding to local interactions. However, when thousands of these individuals interact, incredibly complex collective behaviors arise, such as efficient foraging routes, nest construction, and coordinated defense strategies

Traffic patterns offer a compelling illustration of complex system dynamics. Individual drivers and pedestrians, each pursuing their own route to a destination, interact within a shared space governed by simple rules: maintaining distance, aligning direction; yet collectively produce intricate behaviors like traffic jams and wave propagation, which are disturbances like a sudden braking event or an accident spread through a network of vehicles.

3.2 Agent-Based Modelling

Agent-based modeling is presented as a key method within complex systems research, particularly developed by complex systems scientists. As highlighted in Page and Blume (2008), it involves simulating individual agents and their interactions to observe emergent system behaviors

3.3 Traffic Dynamics

Traffic dynamics refer to the complex interactions and behaviors of various elements within a traffic system, including vehicles, pedestrians, and traffic control devices. Understanding these dynamics is crucial for developing effective traffic management strategies. This section delves into the fundamental principles of traffic dynamics, exploring how individual behaviors and interactions contribute to overall traffic flow and congestion patterns. By examining these principles, we can better understand the challenges and opportunities in optimizing traffic systems.

4 METHODOLOGY

4.1 Environment Design

In this study, we utilized the SUMO simulator for our enviroment. The reason we chose this is because it was the most common platform to train and implement agent based modelling via the TraCI API. TraCI API is an interface used to connect the environment to a programming language such as Python.

We based off our initial environment mirroring the intersection around Uptown Mall in BGC, Philippines. The reason we chose this is the proximity and ease of access from the location of the author. We also believed this is the perfect area to observe different traffic patterns, and driver behaviors to model in the environment as well as the varying transport modes found in this area.

The environment itself consists of 4 legs, each with 3 lanes: 2 regular lanes, and a bike lane, which is very common in BGC. We then implemented sub-lanes to simulate the lane filtering behavior as well as vehicles veering off in different lanes aggressively. This implementation also simulates the lane filtering behaviors commonly found in motorcycles. Traffic light phase design is a rotating mechanism, which means that for every phase only 1 leg at a time allows vehicles to go through. Right turns are always available for all vehicles as per the design of the junctions in BGC.

4.2 Modelling Agents

The agent behaviors are initially based off on the default models by SUMO simulator. We then customized the attributes of each of the vehicles by dimensions, and lane changing behaviors.

We have currently modelled 3 different transport vehicles: Cars, Motorcycles, and Bicycles. Pedestrians are almost modelled using the default jupedsim model by SUMO. This is to simulate the impact of traffic patterns when pedestrians are included in the equation.

For pedestrians, there are crossings in all 4 legs. The adjacent crossing of each leg also turns green when that particular leg for vehicles also turns green.

4.3 Traffic Light Management Agent

In order to train our Traffic Management agent, we plan to utilize Reinforcement Learning in order to find the best traffic light policy for a given condition. We utilized stable baselines reinforcement learning policies for our initial POC.

4.3.1 Observation Spaces. For the observation spaces, we implemented the following features to provide the agent with a comprehensive understanding of the traffic conditions:

1. **Current Phase Lane**: Represented as a one-hot encoded vector, this feature indicates the current active phase of the traffic light, allowing the agent to understand which lanes are currently allowed to move.

2. **Minimum Green Duration**: A binary indicator that shows whether the current phase duration has met the minimum green time requirement. This helps the agent ensure that each lane receives a fair amount of green light time.

3. **Lane Density**: This feature measures the density of each lane by calculating the ratio of the number of incoming vehicles to the total capacity of the lane. It provides the agent with information about how congested each lane is.

4. **Lane Queue Length**: Indicates the number of vehicles in each lane that are moving at speeds below 0.1 miles per hour, effectively capturing the length of the queue and helping the agent identify lanes with significant traffic buildup.

4.3.2 Action Spaces. For the action spaces, the agent is free to pick how long the next phase would be.

4.3.3 Reward Functions. For the reward functions, we implemented the following: 1. Diff Waiting time - produces a reward if the waiting time is improved from the last measure. 2. Average Speed - produces a reward the higher the average speed of the network is. 3. Queue Time - Produces a reward when the number of incoming vehicles is less than the capacity of each lane. 4. Lane pressure - produces as reward by computing the difference between the vehicles outgoing vs incoming in each lane. 5. Pedestrian Waiting Time - produces a reward by checking if the pedestrian waiting time is improved from the last measure.

5 PRELIMINARY RESULTS

Based on our initial observations, we observed that the conditions improve when we implement Reinforcement learning. We observed that the average speed of the network has been increased while at the same time capable of also improving the conditions for the pedestrians.

6 FUTURE WORK

While the study itself is still in progress. We wish to expand the study into a much bigger area, perferrably the entirety of BGC. However, this would require gathering even more enormous amounts of data as we expect different traffic behaviors occuring at different junctions in BGC. We also expect to change our methodology to using a multi-agent reinforcement learning (MARL) approach, as the common practice for handling multiple traffic lights is through MARL.

MARL also requires a heavy duty hardware capable of training multiple agents within reasonable time, so we expect performing our studies in a multi-GPU and multi-CPU hardware, which requires further investment.

7 CONCLUSIONS

This study highlights the potential of leveraging reinforcement learning to address traffic congestion in urban environments, specifically within the unique context of the Philippines. Preliminary results indicate that dynamic traffic management approaches not only improve the average speed of the network but also enhance conditions for pedestrians. These findings underscore the effectiveness of advanced simulation-driven policies over static rule-based solutions. While the study's scope is currently limited, the promising outcomes pave the way for expanding to larger and more complex areas, using multi-agent reinforcement learning to address the challenges of scaling. Future research will focus on enhancing hardware capabilities and gathering extensive data to optimize these models further.

REFERENCES

[1] Tianshu Chu, Jie Wang, Lara Codecà, and Zhaojian Li. 2020. Multi-Agent Deep Reinforcement Learning for Large-Scale Traffic Signal Control. *IEEE Transactions on Intelligent Transportation Systems* 21, 3 (March 2020), 1086–1095. doi:10.1109/TITS.2019.2901791 Conference Name: IEEE Transactions on Intelligent Transportation Systems.

[2] Gengyue Han, Xiaohan Liu, Hao Wang, Changyin Dong, and Yu Han. 2024. An Attention Reinforcement Learning–Based Strategy for Large-Scale Adaptive Traffic Signal Control System. *Journal of Transportation Engineering, Part A: Systems* 150, 3 (March 2024), 04024001. doi:10.1061/JTEPBS.TEENG-8261

[3] Seyed Sajad Mousavi, Michael Schukat, and Enda Howley. 2017. Traffic light control using deep policy-gradient and value-function-based reinforcement learning. *IET Intelligent Transport Systems* 11, 7 (2017), 417–423.

[4] Tomoki Nishi, Keisuke Otaki, Keiichiro Hayakawa, and Takayoshi Yoshimura. 2018. Traffic Signal Control Based on Reinforcement Learning with Graph Convolutional Neural Nets. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. 877–883. doi:10.1109/ITSC.2018.8569301 ISSN: 2153-0017.

[5] Jing Tang, Xin Wei, Jialin Zhao, and Yun Gao. 2020. Semi-supervised double duelling broad reinforcement learning in support of traffic service in smart cities. *IET Intelligent Transport Systems* 14, 10 (2020), 1278–1285.

[6] Chia-Hao Wan and Ming-Chorng Hwang. 2018. Value-based deep reinforcement learning for adaptive isolated intersection signal control. *IET Intelligent Transport Systems* 12, 9 (2018), 1005–1010.