An Adaptive Transport Management Approach using Imitation Learning

Mark Bugeja
University of Malta
Msida, Malta
mark.bugeja@um.edu.mt

Alexiei Dingli
University of Malta
Msida, Malta
alexiei.dingli@um.edu.mt

Maria Attard
University of Malta
Msida, Malta
maria.attard@um.edu.mt

Dylan Seychell
University of Malta
Msida, Malta
dylan.seychell@um.edu.mt

ABSTRACT

The area of Intelligent Transport Systems has been critical in traffic management and intelligent systems for the past decades. In this paper, we introduce a novel approach to traffic management. We develop a process that that starts with the development of a "game" based upon different road networks that are used to gather data based upon user actions. The user's decision directly affect how traffic light states change. This data is then passed to an Imitation Learning model that can observe actions and imitate the same decisions on a similar road network.

CCS CONCEPTS

• **Computing methodologies → Intelligent agents**; • **Applied computing → Transportation**; • **Theory of computation →** *Inverse reinforcement learning*.

KEYWORDS

datasets, neural networks, gaze detection, text tagging

ACM Reference Format:
Mark Bugeja, Alexiei Dingli, Maria Attard, and Dylan Seychell. 2019. An Adaptive Transport Management Approach using Imitation Learning. In *3rd ACM CSCS '19, October 18th, 2019, Kaiserslautern, Germany*. ACM, New York, NY, USA, 3 pages.

1 INTRODUCTION

In this paper we introduce a framework for an Adaptive Traffic Management System. Reinforcement Learning (RL) is considered a state-of-the-art approach in adaptive traffic light systems. Depending on the traffic model used, this area of research offers a variety of techniques that have been applied with various degrees of success. For example, Dynamic Programming techniques are mostly used when modelling micro models [12]. That is modelling one single junction or roundabout. The problem with such an approach is that the optimisation of one junction might have adverse effects on the rest of the road network. On the other hand, more advanced techniques such as Deep-Q-networks are ideally suited to solve macro problems that are more complex problems where the number of states and dimension are significantly higher [11]. The drawback in these cases is usually the limitation in the amount of training needed. In addition, these models have a tendency to overfit for the specific environment they were trained upon. Thus, limiting the model to solving one road network problem. Another approach which has not been used in traffic management is a technique developed by Abbeel et al [2]. This machine learning model is based upon RL. This technique referred to as Imitation Learning (IL) teaches machines to solve tasks not by "doing" but by "observing". The observed is usually a human who is performing some task. This technique has already been successfully used to teach an agent to pilot a helicopter [1]. Our research was inspired by the idea of having one person watching a live traffic camera and being able to adapt traffic light signals to the traffic information he can observe in the live feed. This motivation might also imply that a technique such as IL might also provide a solution which automates the same process. Moreover, we intend on using data which simulates both human drivers and autonomous vehicles. An adaptive traffic management system should not only provide data to intelligent control structures but should also extend information to self driving cars. This would optimise vehicle routing as well as provide predictive data to the system to further optimise the decision making process.

2 REINFORCEMENT LEARNING CHALLENGES

The coordination of data between agents is one of the most significant challenges present in RL applied to transport [3],[10]. The actions undertaken by an agent has a significant effect on the environment and is influenced by the actions taken by other agents working in the same ecosystem [4]. This is mostly because individual agents are interested in solving only the problem for one and optimise the traffic for one junction or traffic light. Consequently, this will result in maximisation of traffic flow in one particular section of the road network but will have a detrimental effect on other areas of the network which will reduce the effectiveness of other agents in optimising traffic flow across the whole traffic network. Thus, an ideal traffic network managed by a multi-agent RL

3rd ACM CSCS '19, October 18th, 2019, Kaiserslautern, Germany

model is not based upon greedy agents that only optimise for their portion of the ecosystem, but, other mechanisms are put in place to orchestrate the different agents and optimise for the whole network. This orchestration of agents will result in some cases where, the agent will not act entirely for its self-interest, and the traffic flow might not be completely optimised. This aspect is an important part of this study as we are not dealing with theoretical frameworks but developing solutions for real-world problems.Another challenge in RL for Transport Orchestration is finding the optimal traffic control policy given the vast scale of the problem. A term often used in RL is the "curse of dimensionality" which describes the problem with managing the substantial amount of data generated by the road networks [14], [16], [5]. This difficulty is due to the number of agents needed to control different sections of the system and usually scale proportionally to the size of the road network. Moreover, by orchestrating a multi-agent environment, new variables are introduced that increase this dimensionality problem [7].

3 IMITATION LEARNING

RL techniques learn through exploration, the experience they gain when interacting with the environment enable these models to find optimal decision paths. While this technique has shown to be particularly useful in a number of tasks related to transportation. Nonetheless, the models suffer from limitations in terms of data availability as well as the initial difficult task of accurately representing the real state of the environment. IL, on the other hand, is a technique that learns by imitating some behaviour, it gains skills in observing demonstrations. This ability of the model to learn through "seeing" is particularly useful when trying to solve complex tasks [13]. The advantage of using this technique is that it can benefit from watching an expert perform some task. Humans, learn through a number of different cognitive abilities. We can make split-second decisions based on intuition and past experiences. In the case of adaptive traffic signal control, human experts define the policies needed for Reinforcement Models, as discussed above, problems arise when optimising for multiple traffic lights systems in a road network. On the other hand, imitation through IL different 'agents' can be orchestrated by building an adaptive policy that learns by watching human experts perform orchestration tasks. IL is divided into two different categories. The first type, called behavioural cloning, performs supervised learning from observations to actions [9]. The second type uses inverse RL. In the latter type, the model learns by estimating reward functions that describe a demonstration as (near) optimal behaviour [8]. Although a lot of work already exists using IL up to the time of writing, only some work related to self-driving cars [15] has been explored in relation to transport.

4 PROPOSED SOLUTION

From the literature, we established that RL techniques are excellent at optimising the decision making processes given some finite goal. On the other hand, they tend to not do so well in the context of adaptive traffic signal control when deployed on large road networks. This lack in optimality is usually due to the lack of communication between the 'agents'. Multi-agent systems have shown promise in recent research [6], but, they have not managed to solve the problem completely. Our proposed approached aims to solve this problem by developing by using IL. Figure 1 illustrates the proposed solution. A game environment is initially developed based upon real road network data. This "game" will ask users to switch different traffic light states to optimise traffic flow. The better and more optimised the traffic flow, the higher the number of points they receive. The actions undertaken by the users are stored as observations. The score is used as part of the filtering process, only users actions which achieved an overall mean score which is greater then the average score is stored in the database (Module 1.1). The IL module learns from the filtered observation data and affects a series of decisions on the Traffic Management Module. Finally, feedback is sent back to the IL Module, describing the effectiveness of the latest decision as well as the new state of the world.

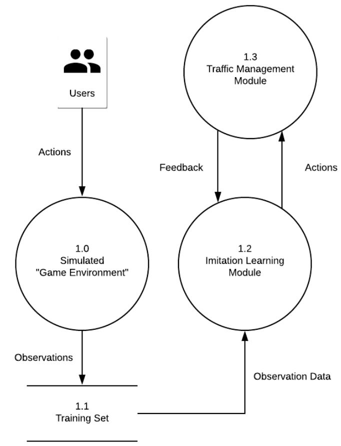

Figure 1: Proposed DFD for the Adaptive Traffic Light Management System

5 PROPOSED EVALUATION

The evaluation methods used for this research include an extensive evaluation modelling different road networks. Each roadwork is evaluated against different level of observations taking into consideration the best 10% to 40% of the scores. In addition, traditional multi-agent RL techniques will also be trained using the same environment to compare the current state-of-the-art implementations with our proposed method.

6 CONCLUSION

In this paper we explore Rl techniques applied to transport management systems. We also introduce a process that uses IL to develop an agent that can manage traffic control structures. Finally, we present a theoretical framework that leverages the idea of a "game" to build the training set needed to develop an IL traffic management model together with a proposed evaluation strategy.

REFERENCES

[1] Pieter Abbeel, Adam Coates, Morgan Quigley, and Andrew Y Ng. 2007. An application of reinforcement learning to aerobatic helicopter flight. In *Advances in neural information processing systems*. 1–8.

[2] Pieter Abbeel and Andrew Y Ng. 2004. Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the twenty-first international conference on Machine learning*. ACM, 1.

[3] Ana LC Bazzan. 2009. Opportunities for multiagent systems and multiagent reinforcement learning in traffic control. *Autonomous Agents and Multi-Agent Systems* 18, 3 (2009), 342.

[4] Lucian Buşoniu, Robert Babuška, and Bart De Schutter. 2010. Multi-agent reinforcement learning: An overview. In *Innovations in multi-agent systems and applications-1*. Springer, 183–221.

[5] Mayank Daswani, Peter Sunehag, and Marcus Hutter. 2014. Feature reinforcement learning: state of the art. In *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*.

[6] Samah El-Tantawy and Baher Abdulhai. 2017. Multi-agent reinforcement learning for integrated and networked adaptive traffic signal control. US Patent 9,818,297.

[7] Samah El-Tantawy, Baher Abdulhai, and Hossam Abdelgawad. 2013. Multiagent reinforcement learning for integrated network of adaptive traffic signal controllers (MARLIN-ATSC): methodology and large-scale application on downtown Toronto. *IEEE Transactions on Intelligent Transportation Systems* 14, 3 (2013), 1140–1150.

[8] Jonathan Ho and Stefano Ermon. 2016. Generative adversarial imitation learning. In *Advances in neural information processing systems*. 4565–4573.

[9] Ahmed Hussein, Mohamed Medhat Gaber, Eyad Elyan, and Chrisina Jayne. 2017. Imitation learning: A survey of learning methods. *ACM Computing Surveys (CSUR)* 50, 2 (2017), 21.

[10] Lior Kuyer, Shimon Whiteson, Bram Bakker, and Nikos Vlassis. 2008. Multiagent reinforcement learning for urban traffic control using coordination graphs. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 656–671.

[11] Li Li, Yisheng Lv, and Fei-Yue Wang. 2016. Traffic signal timing via deep reinforcement learning. *IEEE/CAA Journal of Automatica Sinica* 3, 3 (2016), 247–254.

[12] Tao Li, Dongbin Zhao, and Jianqiang Yi. 2008. Adaptive dynamic programming for multi-intersections traffic signal intelligent control. In *2008 11th International IEEE Conference on Intelligent Transportation Systems*. IEEE, 286–291.

[13] Takayuki Osa, Joni Pajarinen, Gerhard Neumann, J Andrew Bagnell, Pieter Abbeel, Jan Peters, et al. 2018. An algorithmic perspective on imitation learning. *Foundations and Trends® in Robotics* 7, 1-2 (2018), 1–179.

[14] LA Prashanth and Shalabh Bhatnagar. 2011. Reinforcement learning with average cost for adaptive control of traffic lights at intersections. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 1640–1645.

[15] Wilko Schwarting, Javier Alonso-Mora, and Daniela Rus. 2018. Planning and decision-making for autonomous vehicles. *Annual Review of Control, Robotics, and Autonomous Systems* (2018).

[16] Richard S Sutton, Andrew G Barto, et al. 1998. *Introduction to reinforcement learning*. Vol. 2. MIT press Cambridge.