A comparison of algorithms used in traffic control systems

December 22, 2021

Philip Salqvist

phisal@kth.se
1992-04-17

1 Summary

It is in large part a functioning infrastructure that enables cooperation between nations, exchange of goods, and the ability to cast our votes as citizens. [13] The importance of both development and maintenance of infrastructure, has recently been widely debated since the Biden administration decided to sign a 1 trillion $ bill into law [4]. This has raised an important question, namely how to spend this money effectively? Richard Geddes, a scholar on infrastructure policy from Cornell University, recently argued that not enough resources are devoted to the development of new technology in this field, such as intelligent traffic control systems [1]. One can easily imagine the logistic, economic and environmental damages caused by the present relatively ineffective systems. Thus, technological and scientific development in this area can have greater consequences than one intuitively would expect. This bachelor thesis attempts to address this issue, by comparing different traffic control systems, and more precisely the algorithms that control them.

One could ask, is there is a need for this comparison? Isn't there always an optimal solution to the traffic flow problem? And why can't we simply always just deliver this solution? True, there is always an optimal solution. However, delivering this optimal solution in real-time at an intersection would take too much time. Rather, we want to use the algorithm, that is best at approximating the optimal solution reasonably fast and consistently across different traffic scenarios. In this bachelor thesis, the authors compare three algorithms that attempt to do this [5].

One algorithm is pre-timed, meaning that it allows traffic to flow on one road during a fixed amount of time, to then allow traffic on the other for the same amount of time. The deterministic algorithm, on the other hand, collects data through sensors, and can thus make an informed decision. Finally, the reinforcement learning algorithm works similarly to a parent that raises their child to make better decisions. The parent will reward the child to reinforce good behavior, like giving them a toy for cleaning their room, but will perhaps take away the toy for behaving badly. In a similar manner, the reinforcement algorithm gets rewarded when making good decisions, and punished for making bad ones. The algorithm will then strive to optimize the reward and learn as a consequence of this, similar to a child learning to clean their room to gain the benefits associated with it.

The algorithms were tested through simulation, using a four-way intersection and varying traffic demands. The results show that the deterministic algorithm outperforms the others across all different traffic demands. While the reinforcement algorithm performed better than the pre-timed for lower demands, and worse during higher demands. [5]

2 Scientific considerations

One central core of science is the concept of reproducibility. There is, however, not much incentive to replicate studies, since replication rarely gets published and since publications often are necessary to receive grants. Recent attempts have shown that many studies are impossible to replicate in certain scientific disciplines. This has led to what today is referred to as the replication crisis. The underlying causes of the replication crisis are many, one of them being that often experimental conditions aren't clearly reported. [7]

The replication crisis has been especially apparent in scientific fields with human subjects, such as medicine or psychology [7]. However, in his article *Should Computer Scientists Experiment More?*, Tichy fear that computer science could follow a similar trajectory if experiments aren't incorporated more in computer science research [18]. In the bachelor thesis by Omstedt et al. [5], different traffic demands are generated through simulation, and the algorithms are compared by analyzing their effectiveness during these varying conditions. This is clearly an example of experimentation in science.

However, the authors are enabling reproducibility in some cases and less in others. The experiments and simulations were conducted using a road traffic simulation program Simulation of Urban Mobility (SUMO). The conditions during experiments are clearly described, and code to generate simulations is provided in a referenced GitHub repository [6]. Nonetheless, it is not explicitly stated by the authors that an already trained model is provided. This might not be particularly relevant to this research due to the relatively small data sets used for training. Although, it can be of relevance when data sets are larger since the resources needed for reproduction then can get significantly larger.

For instance, Google's model BERT-large [8] used for natural language processing (NLP), uses around 350 million parameters and was trained on 64 TPU chips for four days at a cost of $7,000. Recently, Google released their model openGPT3 with 175 billion parameters. The best results from BERT-large could not be reproduced on a single CPU, and models such as openGPT3 are too large to be used in production. Furthermore, DeepMind's AlphaGo has even more extreme costs, where the best result required 1,920 CPUs and GPUs to play a single game of Go and had an estimated experimental reproduction cost of $35,000,000. [15]

These extreme costs constitute a barrier to entry when it comes to AI research, and especially for reproduction since it is difficult to justify spending vast costs on reproduction studies. Nonetheless, larger models do contribute to better performance, which is a necessary contribution to the scientific field. However, if the reproduction of experiments is going to be viable in these cases, providing a pre-trained model would be a prerequisite.

I would consider this bachelor thesis being highly ethical from a utilitarian point of view. The utilitarian point of view states that one should strive to increase the amount of happiness in the world. Furthermore, it could be argued that both environmental and economic flourishing contributes to this cause. By examining the algorithms that govern our traffic conditions, these concerns can be incrementally improved, and a utilitarian would thus view this contribution as highly ethical.

When examining the statistical results of the paper, not much discussion is needed. A clear argument is provided as to why the authors use average squared waiting time (ASWT) as a measure of the effectiveness of the algorithms. Furthermore, the results are presented clearly and concisely. However, some focus could have been dedicated toward investigating the statistical significance of the averages obtained. One could have done this by examining confidence intervals, or calculating a p-value through a student t-test and comparing averages to these results.

3 Suggestions

One criterion for a master thesis is that the work should present the author's foundational knowledge in the specific topic, as well as on current research and more advanced in-depth knowledge. The authors have been able to present advocate foundational knowledge in their literature study. However, one could argue that the demonstration of in-depth knowledge and insight into current research is not sufficient for a master thesis. Especially the area of related work seems insufficient. The paper does refer to three previous papers made in the field of study [3] [19] [16]. While the authors do deviate somewhat from this research, it is not enough to be viewed as a significant scientific contribution, which would be expected of a thesis on a master's level.

New research in the subject area has been published since the authors contributed with their bachelor thesis. Relevant research by Liang et al. [11] was published in Mars 2018, only two months before the release of the bachelor thesis. The paper is especially relevant to examine further since the authors draw upon more current research in their field of study. By doing so, the paper attempts to improve upon prior attempts to use reinforcement learning in traffic control systems, and thus make a more significant scientific contribution.

The literature study in the paper [11] examines previous research and notes that reinforcement learning has been used while attempting to improve current traffic control systems. Furthermore, the authors found that results in previous research suggested that the use of reinforcement algorithms has suffered from two different problems. Either, the state space has been defined by the number of vehicles or the waiting queue length [17] [2]. However, more recent research concludes that both these types of state spaces are too limited to capture the complexity of a real traffic situation [9]. Or, the state space includes more complex information, such as the vehicle's speed and waiting time, which is the case in the bachelor thesis of interest [5].

In the thesis, Omstedt et al. conclude that the relatively poor performance of the reinforcement algorithm of the study compared to the deterministic algorithm during higher traffic demands could be caused by the curse of dimensionality [5]. Although, the fact that complexity grows exponentially when the number of states increases in a traditional reinforcement learning model, has long been considered general knowledge in the field. Furthermore, the fact that deep reinforcement learning can be used to counter the curse of dimensionality that is associated with Q-learning, is generally known in the field of AI and has been for a relatively long time [14].

If the thesis of interest would have been performed in higher education on a master's level, by students with specialized knowledge in the field of reinforcement learning, one could expect the authors to conclude these findings earlier by doing a more thorough literature study. This is in fact done by Liang et al. [11], which presents previous research that suggests beneficial aspects of using deep reinforcement learning in traffic control systems [10] [12].

If the thesis work aspires to further the development of current research, using deep reinforcement learning would be one way to achieve this, but Van Der Pol [12] advances even further in the author's master thesis. The author presents research exploring the use of a multi-agent algorithm in combination with deep reinforcement learning, to achieve cooperation between traffic lights. In this setting, traffic lights do not act as individual agents to maximize their own reward function, but rather optimize the reward function of the entire system. While both approaches have been explored in previous work, combining the two had not previously been done. [12]

Both papers by Liang et al. [10] and Van Der Pol [12] were accessible in 2018 when the bachelor thesis was written. If this was a thesis on a master's level, one would advise the authors to include such previous research and build upon it, since it would be expected to work with the cutting edge research on a master's level.

References

[1] *5 questions for Richard Geddes on how to improve US infrastructure policy*. June 2021. URL: https://www.aei.org/economics/5-questions-for-richard-geddes-on-how-to-improve-us-infrastructure-policy/.

[2] Monireh Abdoos, Nasser Mozayani, and Ana LC Bazzan. "Holonic multi-agent system for traffic signals control". In: *Engineering Applications of Artificial Intelligence* 26.5-6 (2013), pp. 1575–1587.

[3] Baher Abdulhai, Rob Pringle, and Grigoris J Karakoulas. "Reinforcement learning for true adaptive traffic signal control". In: *Journal of Transportation Engineering* 129.3 (2003), pp. 278–285.

[4] *Biden signs $1 trillion bipartisan infrastructure bill into law, unlocking funds for transportation, broadband, utilities*. Nov. 2021. URL: https://www.cnbc.com/2021/11/15/biden-signing-1-trillion-bipartisan-infrastructure-bill-into-law.html.

[5] Erik Björck and Fredrik Omstedt. "A comparison of algorithms used in traffic control systems". Bachelor's Thesis. 2018. URL: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-229709.

[6] Erik Björck and Fredrik Omstedt. *Traffic Control System Comparison*. May 2020. URL: https://github.com/Xaril/traffic-control-system-comparison.

[7] J. Boye. "Expermimental Methodology". Lecture.

[8] Jacob Devlin et al. "Bert: Pre-training of deep bidirectional transformers for language understanding". In: *arXiv preprint arXiv:1810.04805* (2018).

[9] Wade Genders and Saiedeh Razavi. "Using a deep reinforcement learning agent for traffic signal control". In: *arXiv preprint arXiv:1611.01142* (2016).

[10] Li Li, Yisheng Lv, and Fei-Yue Wang. "Traffic signal timing via deep reinforcement learning". In: *IEEE/CAA Journal of Automatica Sinica* 3.3 (2016), pp. 247–254.

[11] Xiaoyuan Liang et al. "Deep reinforcement learning for traffic light control in vehicular networks". In: *arXiv preprint arXiv:1803.11115* (2018).

[12] Elise van der Pol. "Deep reinforcement learning for coordination in traffic light control". In: *Master's thesis, University of Amsterdam* (2016).

[13] *Public Goods*. July 2021. URL: https://plato.stanford.edu/entries/public-goods/.

[14] John Rust. "Using randomization to break the curse of dimensionality". In: *Econometrica: Journal of the Econometric Society* (1997), pp. 487–516.

[15] Roy Schwartz et al. "Green ai". In: *Communications of the ACM* 63.12 (2020), pp. 54–63.

[16] Matt Stevens and Christopher Yeh. "Reinforcement Learning for Traffic Optimization". In: 2016.

[17] Samah El-Tantawy, Baher Abdulhai, and Hossam Abdelgawad. "Design of reinforcement learning parameters for seamless application of adaptive traffic signal control". In: *Journal of Intelligent Transportation Systems* 18.3 (2014), pp. 227–245.

[18] Walter F Tichy. "Should computer scientists experiment more?" In: *Computer* 31.5 (1998), pp. 32–40.

[19] Kaidi Yang, Isabelle Tan, and Monica Menendez. "A reinforcement learning based traffic signal control algorithm in a connected vehicle environment". In: *17th Swiss Transport Research Conference (STRC 2017)*. STRC. 2017.