ADVANCECRT

Centre for Research Training

Science Foundation Ireland For what's next

COORDINATING AUTONOMOUS AND CONVENTIONAL VEHICLES TO ACHIEVE RELIABLE TRAVEL TIMES USING DEEP REINFORCEMENT LEARNING TECHNIQUES

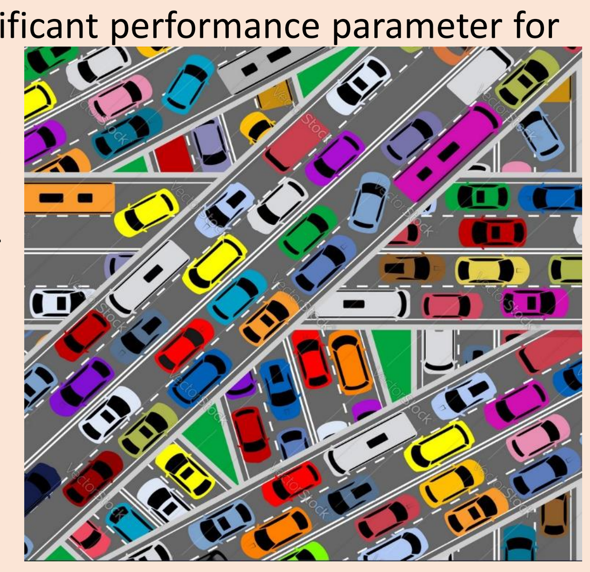
Shaguna Gupta *(guptas2@tcd.ie)*

Supervisor: Prof. Vinny Cahill

Trinity College Dublin, School of Computer Science & Statistics

1. Introduction: Travel Time Reliability

Travel time in transportation is the time taken to move from one place to another with respect to goods, services, people or vehicles. This time can be very different for similar/identical routes depending upon various factors. Travel time has become a significant performance parameter for traffic networks and evaluation of traffic flow control.



Source: https://www.vectorstock.com/royalty-free-vector/traffic-congestion-on-roads-vector-1324369

Largely, these days unexpected driver behaviour, improper use of road infrastructure and road blockages among other factors result in long vehicular queues on roads causing congestion and affecting travel schedules. Travel delay is one of the ways to characterize congestion. Unexpected delay in traffic causes unpredictability in travel times on a daily basis [1]. This underlines the need to consider travel time in terms of its variability. Thus, the quality of traffic service on highways and urban roadways can be assessed using travel time reliability.

According to FHWA, *travel time reliability* is defined as "consistency or dependability in travel time, as a measure from day to day and or across different times of the day'' [2].

According to Transportation Economics Committee of the Transportation Research Board (TRB), *travel time reliability* can be defined as "a measure of the dispersion (or spread) of the travel time distribution" [3]. Further, travel time variability is directly influenced by the variability in the congestion which thus imposes the need to not define congestion only in terms of averages.

2. Quantification: Ways to Measure Travel Time Reliability (TTR)

As travel time reliability is a key performance characteristic for transportation systems, metrics to gauge the reliability levels have been developed. Further, a number of reliability measures have been introduced to quantify the extent of variability in travel times. There are two basic approaches to assessing reliability: Mathematical methods for measuring reliability are based on the conventional User Equilibrium (UE) route choice principle [4]. Empirical methods [2] are based on travel time distributions. Empirical reliability measures involve performance indicators such as the 95th percentile travel time, Buffer Index (BI) and Planning Time Index (PTI).

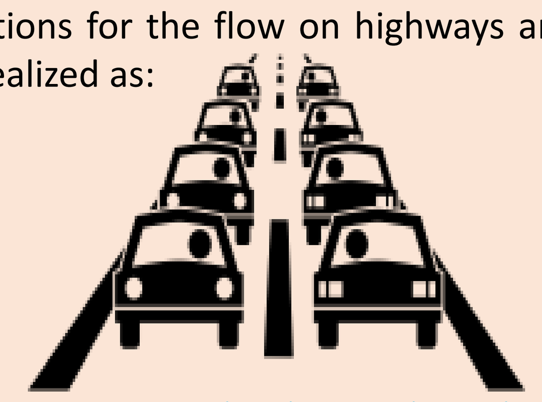
- The ***95th percentile*** represents the travel time on the heaviest traffic days depending upon the length of the route. It is the 95th percentile of the measured travel time.
- The ***Buffer Index*** is the ratio of buffer time to the average travel time where buffer time is the extra time allocated to allow for any unexpected delay. In other words, buffer time is the difference between 95th percentile travel time and the average travel time. As it increases, reliability gets worse.
- The ***Planning Time Index*** is the ratio of 95th percentile travel time to free flow travel time. It accounts for typical delay as well as unexpected delay. The idea is that low variability would means high TTR.

3. Proposed Method: Coordination among Vehicles

For complex traffic situations involving autonomous and human drivers, it becomes even more difficult to predict travel times and the behavior of vehicles. Thus, it is imperative to use data and technology in an integrated manner to achieve consistency in travel times so that inconvenience due to variable travel times is eliminated and reduced at large in case of public transits [5].

In addition, for coordination, pattern of traffic can be examined using machine learning algorithms based on a reward system. A smooth traffic flow can be achieved by training algorithms to learn strategies for coordinating traffic. There is a requirement of dynamic traffic flow which can be obtained by advanced solutions such as automated routing solutions for the flow on highways and urban intersections [6]. Overall synchronization of traffic can be realized as:

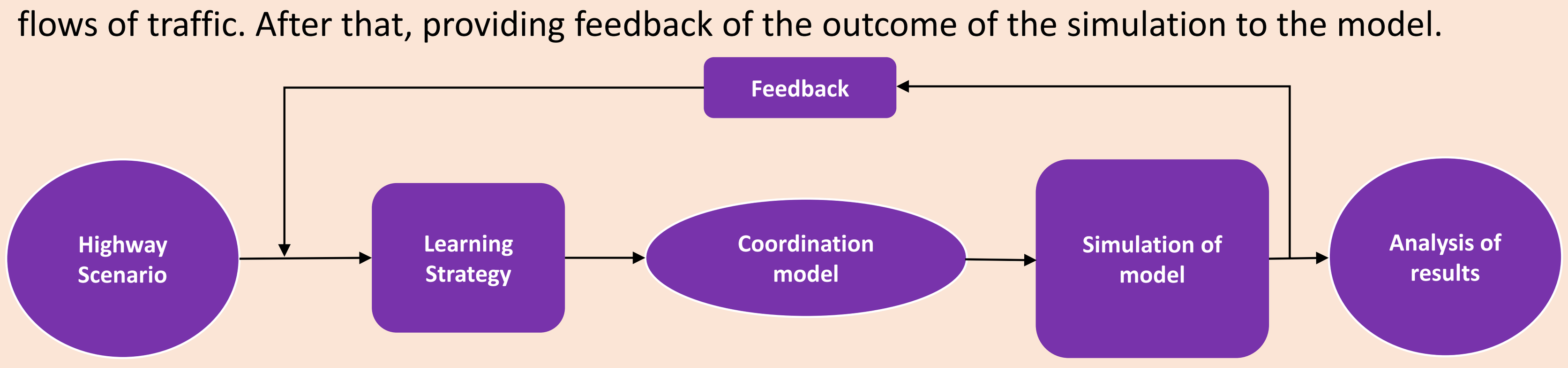
- Scheduling among vehicles:
  - Self-Organization
  - Centralized Control System
- Real-time traffic feedback usage:
  - Vehicle-to-Vehicle Communication
  - Vehicle-to-Infrastructure Communication



Source: https://thenounproject.com/term/traffic-jam/859866/

4. Approach: Progressive Steps

The proposed approach would involve building a ML-based coordination model using state-of-the-art techniques as the control layer to suggest driver's actions and alternate routes in congestion scenarios based on different circumstances on highways [7] The next step would be simulating this model over platforms inclusive of autonomous vehicle trajectories on highways and, dynamic flows of traffic. After that, providing feedback of the outcome of the simulation to the model.



As the above diagram suggests, our overall goal is to learn a good strategy to keep the flow of traffic moving even in bottleneck scenarios such as lane closures, unexpected blockages. Mainly, the focus is on the optimization of the traffic model in order to attain vehicle coordination.

Further, our future objective is to take Floating Car Data (FCD) from the individual drivers firstly to optimize traffic flow on highways and then in urban settings in which traffic routes are diverse.

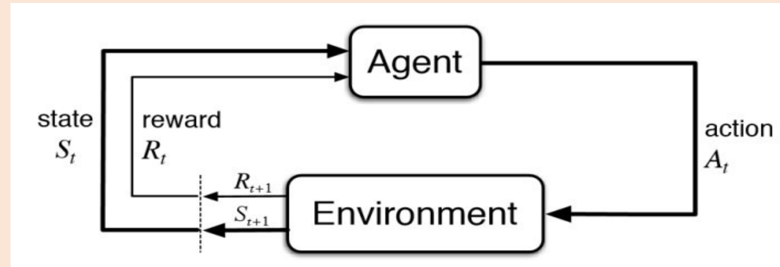
5. Technique: Reinforcement Learning Methods

Reinforcement Learning (RL) is a system for learning a good policy in order to make appropriate decisions using feedback. It is a machine learning method that is concerned with how agents should take actions in an environment and aims to maximize their cumulative reward [8].

Two main approaches to represent model-free reinforcement learning [9] are:

- Policy-based methods and,
- Value-based Methods.

Some well-known model-free RL techniques are compared in the table below:



Reinforcement Learning Environment Model

Source: http://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf

Techniques	On-policy	Off-policy	Action Space	State Space	Source
Monte-Carlo	✓	✓	Discrete	Discrete	A. Barto & M. Duff., "Monte carlo matrix inversion and reinforcement learning". In NIPS, pages 687-694, 1994.
SARSA	✓		Discrete	Discrete	G. A. Rummery & M. Niranjan, "Online Q-Learning using Connectionist Systems" (1994).
Q-Learning		✓	Discrete	Discrete	C. J. Watkins & P. Dayan, "Q-learning". Machine learning, 8(3-4):279-292, (1992).
DQN		✓	Discrete	Continuous	V. Mnih et al., "Human-level control through deep reinforcement learning". Nature, 518(7540):529, 2015.
DDPG		✓	Continuous	Continuous	T. P. Lillicrap et al., "Continuous control with deep reinforcement learning". arXiv:1509.02971, 2015.
TRPO	✓		Continuous	Continuous	J. Schulman et al., "Trust region policy optimization". In Icml, volume 37, pages 1889-1897, 2015.
TD3		✓	Continuous	Continuous	S. Fujimoto, H. Hoof, D. Meger, "Addressing Function Approximation Error in Actor-Critic Methods". arXiv:1802.09477, 2018.
PPO	✓		Continuous	Continuous	J. Schulman et al., "Proximal policy optimization algorithms". arXiv:1707.06347, 2017.
A2C/A3C	✓		Continuous	Continuous	V. Mnih et al., "Asynchronous methods for deep reinforcement learning". In International conference on machine learning, pages 1928-1937, 2016.
SAC		✓	Continuous	Continuous	T. Haarnoja et al., "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor". arXiv:1801.01290, 2018.

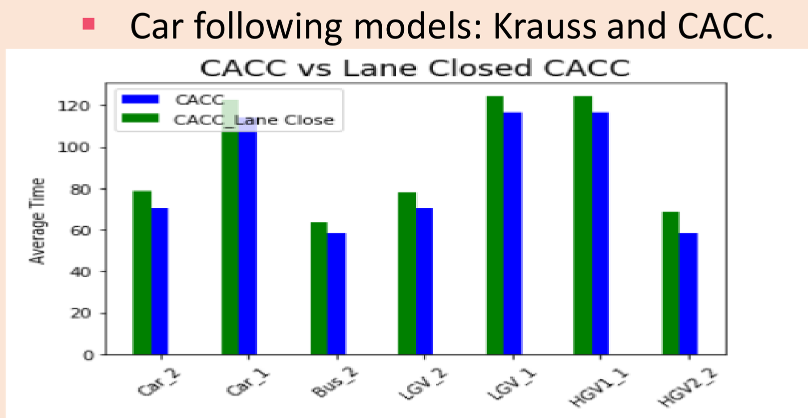
Table: Reinforcement Learning techniques comparison based on RL taxonomy defined by OpenAI

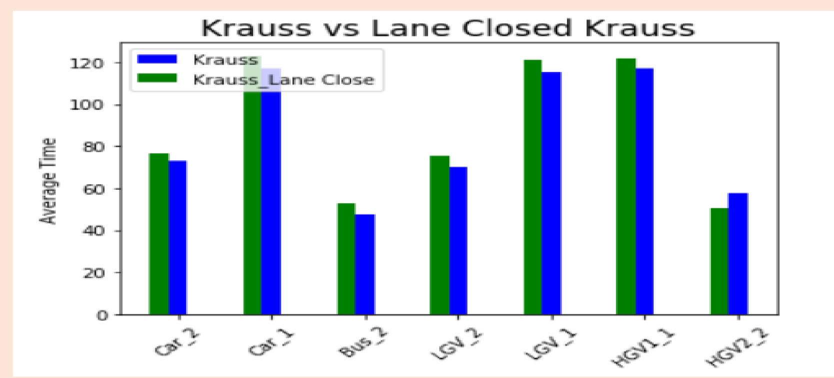
Source: https://spinningup.openai.com/en/latest/spinningup/rl_intro2.html

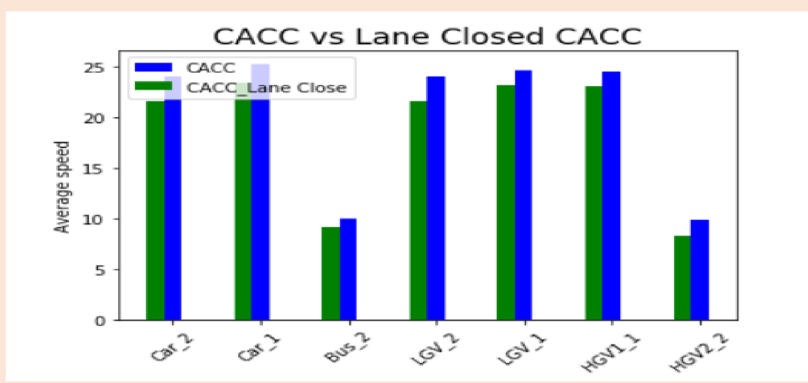
6. Example Scenario: Simulation of Urban Mobility (SUMO) Scenario

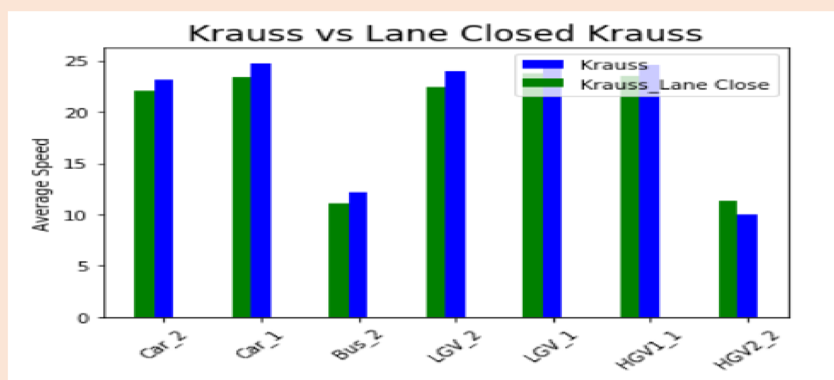
SUMO is an open source, highly portable, microscopic and continuous multi-modal traffic **simulation package** designed to handle large networks **[10]**.

- Segment: Flow of traffic is chosen from M50 motorway segment 11-12.
- Traffic Data: Real traffic data is taken between 15:00hrs and 16:00hrs as on 25th ,September 2020.
- Volume: 8993 vehicles.
- Car following models: Krauss and CACC.









The results show that the average speed of the vehicles decreases significantly due to lane closure for the type of vehicles passing through the route for both the car following models i.e. CACC and Krauss. The average time taken by the vehicles increase in case of lane closure scenario for both the models.

This set of experiment is our baseline for computing travel time and TTR on the segment considered. Next, we will compare them for different scenarios using existing car following models.

7. Conclusion

- Traffic congestion is a serious problem in society. Congestion causing poor traffic performance has negative impacts.
- As traffic volumes and congestion grows on highways and urban roadways, it becomes extremely challenging to maintain dependable schedules and reliable travel times.
- In order to achieve our goal, coordination between autonomous and conventional vehicles can reduce traffic congestion by using real time traffic feedback.
- Overall, traffic congestion can be addressed by implementing optimization techniques such as central control or self-organizing technologies which adapts naturally within the varying traffic conditions.
- The goal is to propose a good learning strategy in order to make the coordination easier and dynamic among vehicles based on different traffic scenarios.
- The implementation of such a technique would enable adapting varying traffic situations and could be applied to large scalable extents.

8. Acknowledgements & References

[1] Afrin, Yodo, "Survey of Road Traffic Congestion Measures towards a Sustainable and Resilient Transportation System", *Sustainability* **2020**, *12*, 4660.

[2] FHWA Report 2006. "Travel Time Reliability: Making it there on time, All the time" US Department of Transportation, Federal Highway Administration. https://ops.fhwa.dot.gov/publications/tt_reliability/ttr_report.htm (Jan 2006).

[3] Transportation Economic Committee of Transportation Research Board (TRB). "Benefit-Cost Analysis: Travel Time Reliability". http://bca.transportationeconomics.org/benefits/travel-time-reliability

[4] Lee, S., Moon, B., Asakura, Y.(2000). "Reliability Analysis Calculation on Large Scale Transport Networks''. Reliability of Transport Networks, eds., Research Studies Press, 203-216.

[5] H. Liu, L. Xiao, X. Kan, "Using Cooperative Adaptive Cruise Control (CACC) to Form High-Performance Vehicle Streams. FINAL REPORT" UC Berkeley Research Reports 2018. https://escholarship.org/uc/item/8pw857gb

[6] S. Devi, T. Neetha, "Machine Learning based traffic congestion prediction in a IoT based Smart City", International Research Journal of Engineering and Technology (IRJET), vol:04, Issue:05, May-2017.

[7] Ata A. et al., 'Modelling smart road traffic congestion control system using machine learning techniques', Neural Network World 2/2019, 99–110.

[8] Nguyen et al., "Deep Reinforcement Learning for Multiagent Systems: A Review of Challenges, Solutions, and Applications", IEEE Transactions on Cybernetics, vol. 50, no. 9, September 2020.

[9] OpenAI Spinning up. "Reinforcement Learning Introduction" https://spinningup.openai.com/en/latest/spinningup/rl_intro.html

[10] "SUMO User Documentation", Online Webpage. https://sumo.dlr.de/docs/index.html