

Traffic Light Control Using Deep Q Learning

Kuchipudi Likita Chowdhary

Department of CSE Apex Institute of Technology
Chandigarh University, Punjab, India

Sannidhi Sassi pranav

Department of CSE Apex Institute of Technology
Chandigarh University, Punjab, India

Maddiboina Anjan Kumar

Department of CSE Apex Institute of Technology
Chandigarh University, Punjab, India

Linga Venkata Lakshmi Narayana

Department of CSE Apex Institute of Technology
Chandigarh University, Punjab, India

Abstract-By combining the Internet of Things (IoT) with artificial intelligence (AI), or what we now refer to as AI-powered IoT, it is now possible to process massive amounts of data produced by numerous devices and manage challenging issues in public systems. Although IoT and AI technologies are rapidly reaching their pinnacles, in this paper we present an IoT application for traffic light control, which is one of the key components of intelligent transportation and has the potential to increase a smart city's road system's efficiency. In essence, the availability of various sensors, such as security cameras, provides instantaneous access to the data content required by intelligent traffic signal controllers to determine the precise number of motorized and non-motorized vehicles on the network. Through the use of the intelligent traffic light control method presented in this paper—which is based on distributed multi agent Q learning—the overall performance of the system can be enhanced by accounting for both motorized and non-motorized traffic at nearby lights. The goal of the suggested solution is to optimize both motorized and non-motorized traffic using this multi agent Q learning algorithm. Furthermore holistic (affair) we looked into the various constraints (rules) which are deployed on real world traffic lights, these rules were also integrated into the learning algorithm to produce the proposed solution which can be deployed in real world operational conditions. Using data from a vehicle route through the real urban area and real traffic data, we simulated some numbers. The simulation results show that our suggested solution outperforms the current ones in terms of vehicle and pedestrian wait times, intersection delays, and other crucial parameters.

Keywords-Intelligent Transportation, Traffic Control Systems, Automatic Control Optimization, Reinforcement Learning Control, Integrated Traffic Management

Date of Submission: 02-12-2024

Date of acceptance: 12-12-2024

I. INTRODUCTION

In logistics and transportation, road traffic is essential, especially in urban areas where population growth is occurring. This contributes to an increase in traffic-related problems like long wait times, lost time, and increased CO₂ emissions. These issues not only have an impact on day-to-day living but also have major economic and environmental repercussions. For instance, in 2017 it was projected that traffic congestion alone cost the US 350 billion dollar. Real-time traffic light control (TLC) systems were created in response to these difficulties, utilizing real-time data to dynamically modify traffic signal timings.

All the same, because transportation networks are unpredictable, efficiently controlling traffic is still a difficult undertaking that calls for creativity. In logistics and transportation, road traffic is essential, especially in urban areas where population growth is occurring. This contributes to an increase in traffic-related problems like long wait times, lost time, and increased CO2 emissions. These issues not only have an impact on day-to-day living but also have major economic and environmental repercussions. For instance, in 2017 it was projected that traffic congestion alone cost the US [1].

Intelligent Transportation Systems (ITS) are becoming a feasible way to make transportation safer and more environmentally friendly as Internet of Things (IoT) and Artificial Intelligence (AI) systems develop. Applications for ITS can improve overall efficiency, lower emissions, and optimize traffic flow. The use of Reinforcement Learning (RL) techniques is one promising method for controlling traffic lights. By defining reward functions that minimize values like waiting time or emissions, traffic light control can be optimized according to the current situation. RL is an important tool for traffic management because it can adapt to different circumstances and learn the best course of action [2].

While traditional methods like Fuzzy Logic Control have been used for traffic light control, integrating them with RL techniques can lead to more efficient and adaptive traffic control systems. By combining Deep Q Learning with Fuzzy Logic Control, researchers aim to optimize both the phase sequence and green light duration, providing a more effective solution compared to existing methods. This integration of AI, IoT, and RL techniques holds great promise for improving traffic management and reducing its negative impacts on society and the environment.

A. Problem Statement

Nowadays there are many problems with the increasing population, traffic control is one of the major problems, due to which there are thousands of accidents occurring. We came up with a solution through which we can manage traffic effectively hence mitigating the possible risks.

B. Objective and Challenges

There are several difficulties with implementing traffic light control with a Deep Q-Learning (DQN) agent. First of all, it is challenging to effectively capture and depict the environment since traffic systems are intricate and involve a variety of elements, including vehicle and pedestrian flow. Furthermore, it is difficult for the DQN agent to respond to changes quickly due to the dynamic nature of traffic situations. Since the status space and action space are highly dimensional, learning is made more difficult, and in order to guarantee safety and promote desirable behavior, strong incentive systems are required. Even more complexity arises from the need to enforce traffic laws and prioritize safety. Teaching and acquiring a variety of training data are two more crucial issues.

II. RELATED WORK

Kumar Ritesh et al [3] in their research, they proposed an innovative approach to traffic light control using Deep Reinforcement Learning (DRL) with Q-learning and Multi-Layer Perceptrons (MLPs). They start by carefully considering the network topology, opting for a mesh network topology to interconnect a cluster of traffic lights. Redundancy and fault tolerance are provided by this topology, guaranteeing dependable communication between traffic lights even in the case of node failures or network outages. Liu, Junxiu, et al [4] proposed data exchange aspect for their system and they utilized Message Queuing Telemetry Transport (MQTT) or CoAP (Constrained Application Protocol) due to their lightweight and efficient message queuing capabilities. We are able to efficiently manage the message flow between traffic lights thanks to these protocols, which we combine with Apache Kafka for real-time data handling. This guarantees that vital information is delivered and processed promptly. Proceeding to state representation, they utilize MLPs for processing the traffic light system's state data.

Wei, Hua, et al [5] developed an MLP-based Q-network. Approximating the Q-values for every action based on the current state is the responsibility of this network. We can effectively learn the intricate mapping between state-action pairs and Q-values by employing MLPs for this task, which allows the DRL agent to choose the optimal action that maximizes traffic light timings for better traffic flow. M. L. Saini et al [6] discussed a potential role of block chain for wireless network. Saini et al [7] proposed a system for error control codes in digital communication system. These networks can be used in traffic light control.

Liang, Xiaoyuan, et al [8] developed an MLP system and this MLP is tasked with predicting the expected reward for each action based on the current state. The reward prediction is crucial as it provides feedback to the DRL agent, guiding its decision making process towards actions that lead to positive

outcomes such as reduced congestion and improved traffic flow. Liang, Xiaoyuan, et al [9] included some factors such as current traffic conditions, traffic light statuses, and other relevant information. The MLPs are crucial in extracting relevant features from the raw data and creating a compact representation suitable for input to the DRL agent. This step is essential as it allows the DRL agent to work with a manageable set of features that capture the essential aspects of the traffic light control problem.

Y. Singh et al [10] discussed activation functions and their performance for various deep learning applications. For designing traffic light control application the best activation function can be chosen. J. Sarmah et al [11] discussed the various pre trained models for various deep learning applications. Tunc, Ilhan et al [12] developed an MLP based network. This MLP is tasked with predicting the expected reward for each action based on the current state. The reward prediction is crucial as it provides feedback to the DRL agent, guiding its decision-making process towards actions that lead to positive outcomes such as reduced congestion and improved traffic flow. Through this integrated approach, we aim to create a dynamic and adaptive traffic light control system that can optimize traffic light timings in real-time based on current traffic conditions.

Damadani, Shima et al [13] developed an IoT based traffic light control system. They recommended using a text box instead of an image to insert a graphic, as this method is a little more stable in an MSW document. The ideal graphic would be a 300 dpi TIFF or EPS file with all fonts embedded. They think their system has the potential to greatly increase the effectiveness of urban transportation systems overall and improve traffic management.

III. METHODOLOGY

The proposed traffic light management system is performed in six steps and these are given as follows:

A. Inter-Communication between Traffic Lights: Implementing a communication mechanism between traffic lights can help coordinate their actions. This can be achieved by allowing traffic lights to share information about traffic conditions, such as the number of vehicles waiting at each light or the current phase duration. This information can be used to adjust the timing of traffic lights in a coordinated manner to improve traffic flow.

B. Dynamic Traffic Pattern Recognition:

Use machine learning techniques to identify and adjust to changing traffic patterns. This can assist the model in modifying the timing of traffic lights in response to actual traffic circumstances, like accidents or traffic congestion [14].

C. Reinforcement Learning Technique:

Other reinforcement learning approaches to take into account are Deep Reinforcement Learning (DRL) algorithms (e.g., Deep Q-Networks, Actor-Critic methods) in addition to Q-learning [15]. These algorithms may result in improved performance because they can handle environments that are more complex.

D. Traffic Prediction Models:

To predict future traffic conditions, incorporate traffic prediction models. In light of anticipated traffic flow, this can assist the traffic light control system in proactively adjusting timings.

E. Synchronization with External Systems:

In order to receive real-time updates and modify traffic light timings appropriately, make sure the traffic light control system can synchronize with external systems, such as GPS navigation systems or traffic monitoring centers [16].

F. Scalability and Robustness:

Create a scalable, reliable system that can manage many traffic lights and different traffic situations without sacrificing functionality.

IV. IMPLEMENTATION

A. Experimental Setup

For building a deep learning model Anaconda Jupyter notebook was used which runs on Intel(R) Core(TM) i5-12005U CPU @ 2.80GHz, 8GB RAM and for building Q deep learning model, google collab was used, which provides the default configuration of an Intel i5 12 CPU with 8 vCPUs and 16GB of RAM.

B. ML Model for Traffic light control

In this work, we present a novel method for controlling traffic lights by combining Q-learning and Multi-Layer Perceptrons (MLPs) with Deep Reinforcement Learning (DRL). In order to connect a cluster of traffic lights, we first carefully consider the network topology and choose a mesh network topology. Redundancy and fault tolerance are provided by this topology, guaranteeing dependable communication between traffic lights even in the case of node failures or network outages.

For the data exchange aspect of our system, we utilize Message Queuing Telemetry Transport (MQTT) protocol and CoAP (Constrained Application Protocol) due to their simplicity, lightweight and efficient message queuing capabilities in broadcasting the messages. These protocols, coupled with Apache Kafka for real-time data handling and allow us to manage the flow of messages between traffic lights effectively and also ensuring that critical information is delivered and processed.

We employed MLPs to process the state information of the traffic light system and state representation. This MLP contains current traffic conditions, traffic light states, and other relevant information. This step is essential in the system as it allows the DRL agent to work with a manageable set of features. The MLPs are crucial in extracting relevant information from the raw data and creating a compact representation suitable for input to the DRL agent and then further processing.

We hope to develop a dynamic and adaptive traffic light control system with this integrated approach, one that can optimize traffic light timings in real-time according to the flow of traffic. We think that our system can greatly improve traffic management and raise the overall efficiency of urban transportation systems by utilizing the strengths of DRL, Q-learning, and MLPs.

Mathematical Overview: Set weights for the reward prediction MLP and the Q-network at startup. Define the exploration rate, discount factor, state space S , action space A , and additional hyper-parameters. The DRL agent is configured in this step so that it can learn to control the traffic lights based on the flow of traffic.

Set up the communication network with a mesh topology. Configure the message queuing protocols (e.g., MQTT, CoAP) for data exchange. Use Apache Kafka for real-time data handling. This step establishes the communication infrastructure for exchanging data between traffic lights and the DRL agent.

We used MLPs to process the state information of the traffic light system. Extract relevant features from the raw data to create a compatible representation suitable for input to the DRL agent. MLPs are used to preprocess the state information from the traffic lights, making it easier for the DRL agent to learn and make traffic light on/off decisions.

Next step we used the Q-network to approximate the Q-values $Q(s, a)$ for each action a based on the current states s . Select the action with the highest Q-value as the optimal action. Then the DRL agent uses the Q-network to estimate the expected rewards for different actions and selects the action that maximizes the expected rewards.

Another MLP was used to predict the expected reward for each action based on the current state and calculated the reward points based on the impact of the selected action on traffic flow and congestion. This MLP step helps the DRL agent for predicting the reward points and it is likely to receive for different actions, helping it learn which actions are most beneficial.

Finally after training our model we used trained DRL agent to select actions for real-time traffic light control based on current traffic conditions. We adjust traffic light timings accordingly to optimize traffic flow and congestion on that particular traffic light node. The DRL agent uses its learned policy to control the traffic lights in real-time and adjust the light timings as per traffic situation and congestion. These steps collectively form the algorithm for designing a traffic light control system using DRL with Q-learning and MLPs. Every steps has its crucial role in the system from initialization to the last decision.

A. State

Instruction needs to announce to the algorithm 'what' and 'where' cars are around intersections so that the machine learning is able to figure out the way to organize traffic to maximum effect. The critical feature of this technological information is the faster communication details about the vehicles present in the environmental environment. Nevertheless, differential cell sizes is what perhaps studies groups also faced in specific. This special area design just includes the information on the type of vehicles it has, and not about the size of the cells used for differentiation for cells of different types into the continuous phase. Areas that stand near the traffic intersection cell sizes are smaller, and they grow as the distance from the strip of traffic increases. This method has, therefore, been used to magnify the impact of traffic on vehicle movements at the intersection, so that it captures it more. Executive in fact, simulations are made for both cases where the cell sizes are taken either as equal or as not equal. The outcomes are considered in the simulation results portion of the paper. The design of the selected state representation is perfectly now to realism. This outcome has been suggested latterly in research about information-rich states of traffic

signal controllers, but that technology is hard to actualize in practice since the information demanded grows complicated to obtain. At the 4- legged traffic intersection, we divided each road into cells as shown in the necessary portion. The value associated with the cells is 1 when there is one vehicle or more in the cell, and is 0 when there is no vehicle in the cell. Along the lane, you will find 10 cells in total. Incoming traffic resides at the very cells blocks of 20 each at the traffic junction. So, therefore, there are 20 cells in each of the incoming roads and a total of 80 cells. The right hand, three lane cross traffic is the part of the same cell because it has one traffic signal meanwhile, the track on the left side is the result of independent cells. We have been using the Lane Space Discretization method that can detect whether any cars stand in the different sides of the traffic intersection.

B. Action

While we have made appropriate provisions in line with the existing traffic light system for the Q-learning algorithm, the details are still being worked on. For example, we'll be performing one at any available time frame. This activity is of parametrical and procedural nature as it extracts the maximum value out of opportunities that is selectively available from Q learning agent action sets. The characteristic of being able to undertake any action from the set of possible actions is declared to be a feature of an agent. The entity is an intelligent traffic light trick. It's a system that allows to switch some traffic lights to green for particular routes and keep them in the actual state for as long as needed. The green light will be on for 10s, whereas the context for the yellow light will stretch up to 4 s. When two shifts are significantly different from each other, the turning of the yellow light is not applied and if it happened it means that the current green cycle is being repeated with- out curtailing of the oval. When the green and yellow lights are for 4s each, no other operations can go on when at least 14 seconds after the same action has been done. So, as a result, the fourteen stages are added together. The model envisage 80 boolean cells to be present in the state space.

The exploitation is the act of reducing dramatically the number of states that should be search for best action and therefore, the selection of the right cells for the representation is also vital. From the perspective of the environment, the primary problem is the lamp is usually in limbo due to a minimum one automobile operates at the crossroads. Hence, the most crucial cells are always those situated closest to the stop line, while the ones that are far from the stop line have less significance. The fact, that the number of active cells that are inside the stop line are higher, can be a reason of the fact, that the deeper neural net is trained during the same amount of time. For Example, if sensors and cameras are installed these become capable of identifying not only pedestrians but also vehicles at intersections.

C. Reward

The system of reinforcement learning works in such a way that the agent moves to another state in the environment after it makes the action in advanced learning and the effect of that movement is taken. The convenience of which parameter is selected affects the way that the agent is motivated to assess the outcome of the made decision and implement better choice making. Hence, the reward is a pivotal factor in the acquisition of knowledge. The aim of this app is to match up the volume of traffic simultaneously going through the intersection, as well as adjust flow of traffic over the time. Various objectives like control or optimal damage can be implemented via different Q-learning collaborative frame- works. On the other hand, the negative value of traffic occupancy can be a factor in the analysis as well, for example the number of cars and the queue length of the vehicles. Intuitively, the purpose of this article is to minimize the overall queue time, which can be interpreted as a measure of the flow situation at the traffic speed.

V. RESULTS AND DISCUSSION

As in earlier DQN research, there are fluctuations in the training process, but the overall reward function continues to rise, indicating that the algorithm is continuously improving its traffic light control techniques. It is noteworthy that the technique is rather easy to learn in the switching scenario, where one-directional traffic flows shift between the first and second half of the simulation period (i.e., always green for the present direction). Consequently, the reward function does not vary much during the course of the experiment. The system encounters a notable decrease in the reward function at the changeover point, when the previous strategy must be reversed. Thankfully, the RL approach is flexible and quick to alter. These findings show how the RL technique can dynamically modify its control tactics in response to shifting traffic states, guaranteeing effective traffic flow management.

VI. CONCLUSION

The research paper "Deep Q Learning-Based Control for Traffic Lights" presents a comprehensive

approach to optimizing traffic light control using deep reinforcement learning (DRL) with Q-learning and multi-layer perceptrons (MLPs). The study addresses the challenge of dynamically adjusting traffic light timings to improve traffic flow and reduce congestion in urban environments. The paper first introduces the concept of using DRL for traffic light control and outlines the architecture of the proposed system. In next section, it explains how the system uses a network topology for exchange the information among traffic lights, message queuing techniques for real-time data handling and MLPs to approximate the Q-function.

The system's implementation, including the DRL agent and traffic light system's initialization, is explained in this paper. It describes how MLPs are used to process the state representation, how Q-values are used to choose actions, and how rewards are forecast to maximize traffic flow.

The effectiveness of the suggested strategy in enhancing traffic management in comparison to conventional methods is highlighted in the paper's conclusion. It talks about the possibilities for more study and advancement in this field, including incorporating more sophisticated machine learning methods and implementing the system in the real world. Overall, the study provides encouraging insights for upcoming smart transportation systems by demonstrating the viability and efficacy of using DRL for traffic light control.

REFERENCES

- [1] Mousavi, Seyed Sajad, Michael Schukat, and Enda Howley. "Traffic light control using deep policy-gradient and value-function-based reinforcement learning." *IET Intelligent Transport Systems* 11.7 (2017): 417-423.
- [2] Zeng, Jinghong, Jianming Hu, and Yi Zhang. "Adaptive traffic signal control with deep recurrent Q-learning." 2018 IEEE intelligent vehicles symposium (IV). IEEE, 2018.
- [3] Kumar, Ritesh, Nistala Venkata Kameshwer Sharma, and Vijay K. Chaurasiya. "Adaptive traffic light control using deep reinforcement learning technique." *Multimedia Tools and Applications* 83.5 (2024): 13851-13872.
- [4] Liu, Junxiu, et al. "Intelligent traffic light control by exploring strategies in an optimised space of deep Q-learning." *IEEE Transactions on Vehicular Technology* 71.6 (2022): 5960-5970.
- [5] Wei, Hua, et al. "Intellilight: A reinforcement learning approach for intelligent traffic light control." *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 2018.
- [6] M. L. Saini, J. Panduro-Ramirez, J. Padilla-Caballero, A. Saxena, M. Tiwari, and K. Ravi, "A Study on the Potential Role of Blockchain in Future Wireless Mobile Networks," 2023 3rd Int. Conf. Adv. Comput. Innov. Technol. Eng. ICACITE 2023, pp. 547–550, 2023, doi: 10.1109/ICACITE57410.2023.10182375.
- [7] M. L. Saini, V. K. Sharma, and A. Kumar, "An Efficient Single and Double Error Correcting Block Codes with Low Redundancy for Digital Communications," 2023 Int. Conf. Adv. Comput. Comput. Technol. InCACCT 2023, pp. 827–831, 2023, doi: 10.1109/InCACCT57535.2023.10141727.
- [8] Liang, Xiaoyuan, et al. "A deep reinforcement learning network for traffic light cycle control." *IEEE Transactions on Vehicular Technology* 68.2 (2019): 1243-1253.
- [9] Liang, Xiaoyuan, et al. "Deep reinforcement learning for traffic light control in vehicular networks." *arXiv preprint arXiv:1803.11115* (2018).
- [10] Y. Singh, M. Saini, and Savita, "Impact and Performance Analysis of Various Activation Functions for Classification Problems," *Proc. IEEE InC4 2023 - 2023 IEEE Int. Conf. Contemp. Comput. Commun.*, 2023, doi: 10.1109/InC457730.2023.10263129.
- [11] J. Sarmah, M. L. Saini, A. Kumar, and V. Chasta, "Performance Analysis of Deep CNN, YOLO, and LeNet for Handwritten Digit Classification," *Lect. Notes Networks Syst.*, vol. 844, pp. 215–227, 2024, doi: 10.1007/978-981-99-8479-4_16.
- [12] Tunc, Ilhan, and Mehmet Turan Soylemez. "Fuzzy logic and deep Q learning based control for traffic lights." *Alexandria Engineering Journal* 67 (2023): 343-359.
- [13] Damadam, Shima, et al. "An intelligent IoT based traffic light management system: deep reinforcement learning." *Smart Cities* 5.4 (2022): 1293-1311.
- [14] M. L. Saini, R. Kumar, D. C. Sati, and T. Bharti, "Fire Monitoring and Prevention System Based on the Severity of Fire Using NodeMCU and Cloud," *Proc. 2nd Int. Conf. Intell. Innov. Technol. Comput. Electr. Electron. ICIITCEE 2024*, 2024, doi: 10.1109/IITCEE59897.2024.10467602.
- [15] Liu, Bo, and Zhengtao Ding. "A distributed deep reinforcement learning method for traffic light control." *Neurocomputing* 490 (2022): 390-399.
- [16] Coşkun, Mustafa, Abdelkader Baggag, and Sanjay Chawla. "Deep reinforcement learning for traffic light optimization." 2018 IEEE International Conference on Data Mining Workshops (ICDMW). IEEE, 2018.