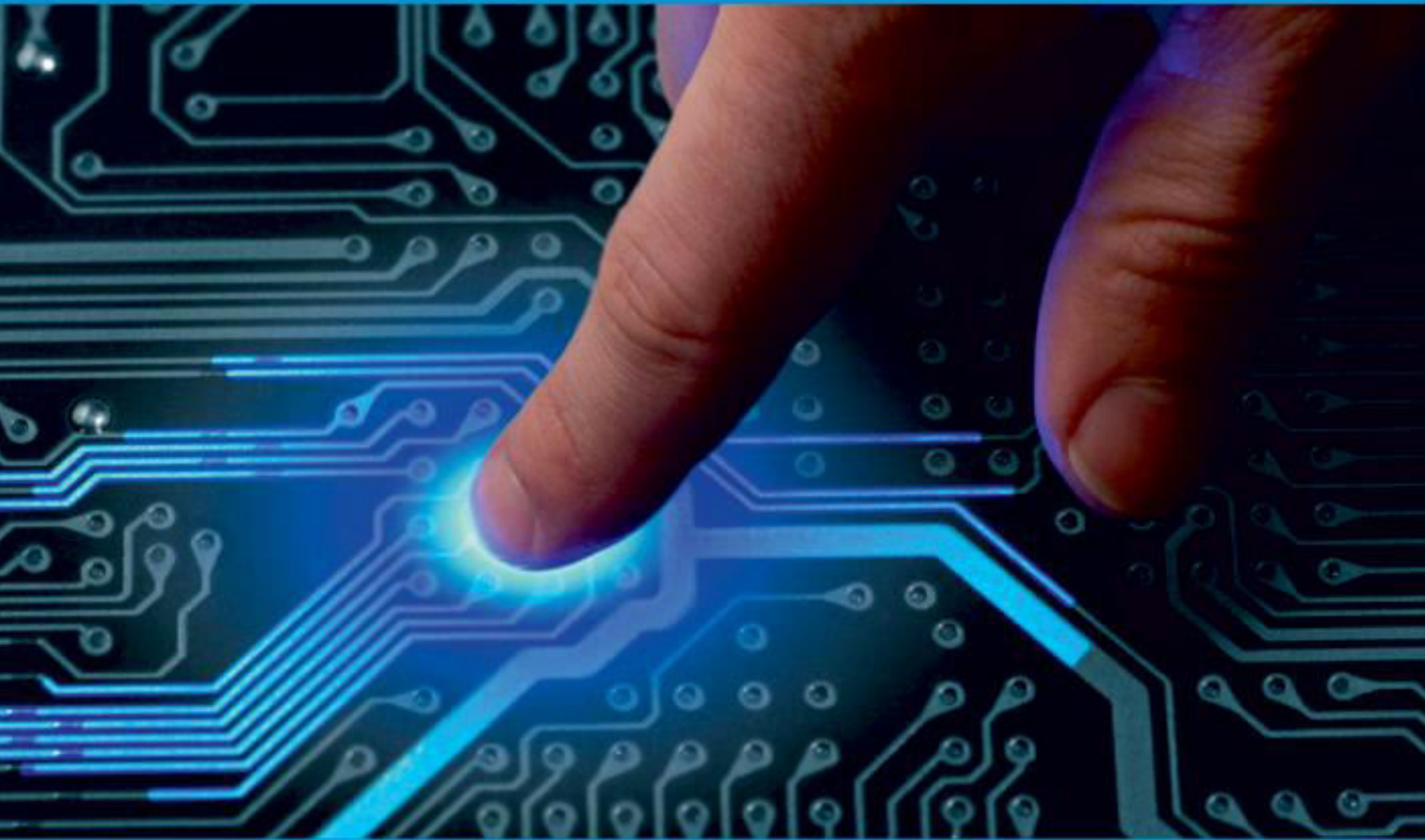
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 10, October 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.625

© 2024 IJIRCCE | Volume 12, Issue 10, October 2024| DOI: 10.15680/IJIRCCE.2024.1210066

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Reinforcement Learning based Approach for Traffic Signal Control

Subha Indu S[1], Dheeraj.S.Prabhu[2]

Assistant Professor, Department of Software System, Sri Krishna Arts and Science College Coimbatore, India[1]

Student, Department of Software System, Sri Krishna Arts and Science College Coimbatore, India[2]

ABSTRACT: The paper presents a innovative solution to the classic problem of adaptive traffic signal control (TSC), leveraging Artificial Intelligence instead of conventional optimization or basic rule-based methods. Reinforcement Learning (RL), a rapidly advancing field within Machine Learning, offers crucial advantages such as scalability, generalization, and real-time applicability for addressing traffic signal management challenges. However, these benefits also place greater responsibility on researchers to carefully design the state representation and reward system. The effectiveness of the RL algorithm largely depends on these abstractions, which determine whether it can successfully solve the problem. This study introduces a new approach to state representation and reward structure that enhances the generalizability and scalability of the TSC solution. The feasibility of the proposed method is validated through a simulation study using a high-fidelity microscopic traffic simulator. The results indicate that this Deep Reinforcement Learning-based approach is a strong contender for real-time traffic signal control.

I. INTRODUCTION

The rapid growth in population and urbanization has led to increasing traffic demands in cities worldwide. Traffic signal controls (TSCs) are implemented to manage traffic flow and reduce congestion at intersections. During periods of heavy traffic, vehicles often slow down or come to a stop in lanes, leading to longer queues of waiting vehicles. Congestion in a single lane can create a ripple effect, impacting traffic flow in other lanes and at adjacent intersections. Three primary factors contribute to worsening congestion. First, when the number of vehicles entering an intersection exceeds those leaving it, congestion builds up. Second, cross-blocking occurs when vehicles are unable to proceed through an intersection, even with a green light, because the downstream lane is already full. Lastly, green idling happens when a green signal is activated at an intersection with no vehicles present to take advantage of it. The application of Deep Reinforcement Learning (DRL) for traffic light control has been explored to a limited extent. Many researchers have employed Value-based algorithms in Single-Agent approaches, such as Deep Q-Learning and its on-policy variant SARSA However, compared to Policy-based methods, these approaches do not guarantee convergence, even to local optima, which can make them more prone to convergence issues during training and ultimately affect performance. Additionally, the majority of reward strategies used in these studies rely on low-level features like queue length, or do not incorporate traffic-related features at all (Thorpe and Anderson, 1996). This reliance on low-level features complicates the process of credit assignment and difficult the ability of the model to generalize effectively.

II. SIGNIFICANCE OF DEEP REINFORCMENT LEARNING FOR TRAFFIC SIGNAL CONTROL

The implementation of Deep Reinforcement Learning (DRL) has generated significant interest in the field of artificial intelligence, marked by several notable successes. In 2013, DeepMind showcased DRL's capabilities by applying it to a variety of Atari 2600 games, achieving superhuman performance. In 2016, the same organization used DRL to train AlphaGo, which went on to defeat multiple world champions in the game of Go. Since then, DRL has found applications across diverse domains, including robotics, natural language processing, healthcare, business management, Industry 4.0, smart grids, computer vision, and transportation, particularly in traffic signal control (TSC) and autonomous vehicles.

Given these advancements, this article reviews the limited research on DRL in TSC, driven by the objective of developing solutions that exceed human intelligence. The advantages of DRL make it particularly appealing for traffic signal management, as outlined below:

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- Adaptability to Real-Time Conditions: DRL enables agents to respond to real-time traffic situations that can change unpredictably due to unforeseen events like adverse weather or accidents.
- Model-Based and Model-Free Approaches: DRL can operate using either a model-based or model-free framework, allowing agents to learn autonomously without needing prior knowledge of the environment. The model-based approach involves creating a model of the operational environment to select actions based on feedback, while the model-free approach bypasses the creation of a model. The latter is often preferred for TSC due to its lower complexity and reduced computational requirements.
- Comprehensive Reward Representation: DRL allows for reward structures that align with system goals and performance metrics, incorporating various factors influencing system performance. For instance, rewards can be adjusted based on average waiting times, queue lengths, or vehicle throughput at intersections.
- Addressing the Curse of Dimensionality: DRL effectively mitigates the curse of dimensionality that traditional reinforcement learning approaches face, especially in TSC, where the state space is extensive due to the multitude of factors influencing a traffic network.

III. REINFORCEMENT LEARNING

Reinforcement Learning (RL) is a dynamic and extensively studied area within Machine Learning (ML), recognized for its significant achievements in various domains such as board games , robotics (), and video games Unlike Supervised Learning (SL), which focuses on mapping input x to output y using training samples, RL operates on a different learning paradigm. One of the main challenges of SL is the continuous demand for training data, which is often unavailable in many control scenarios. In contrast, RL creates its own training data by interacting with the environment in real-time.

Reinforcement Learning (RL) is a prominent area of research within Machine Learning, characterized by its ability to enable agents to learn from interactions with their environment. Unlike Supervised Learning, which requires labeled training data to map inputs to outputs, RL allows agents to generate their own training data through trial and error. This learning process is typically modeled as a Markov Decision Process (MDP), where the agent strives to maximize cumulative rewards by following a policy. The total reward is calculated over time, factoring in a discount rate that prioritizes immediate rewards over future ones. During each interaction, the agent observes the current state of the environment, selects a action based on its policy, and receives feedback in the form of a reward, which informs its future decisions. This reward system acts as a crucial guiding principle for the agent, helping it refine its policy to achieve optimal performance.

IV. POLICY GRADIENT ALGORITHM

The Policy Gradient algorithm is a fundamental approach in Reinforcement Learning that focuses on optimizing the policy directly than relying on value functions. This method works by parameterizing the policy, typically as a neural network, which outputs a probability distribution over possible actions given a specific state. The objective of the Policy Gradient algorithm is to maximize the expected cumulative reward by adjusting the parameters of the policy based on the feedback received from the environment. During training, the algorithm collects trajectories of states, actions, and rewards, using these experiences to compute gradients that indicate how to adjust the policy parameters. This adjustment is often achieved through techniques such as the REINFORCE algorithm, which applies the gradient ascent method to update the policy in the direction that increases the likelihood of actions leading to higher rewards. By iteratively refining the policy through these updates, the Policy Gradient algorithm can effectively learn complex behaviors in dynamic environments, making it particularly useful for problems with high-dimensional action spaces.

V. ENVIRONMENT

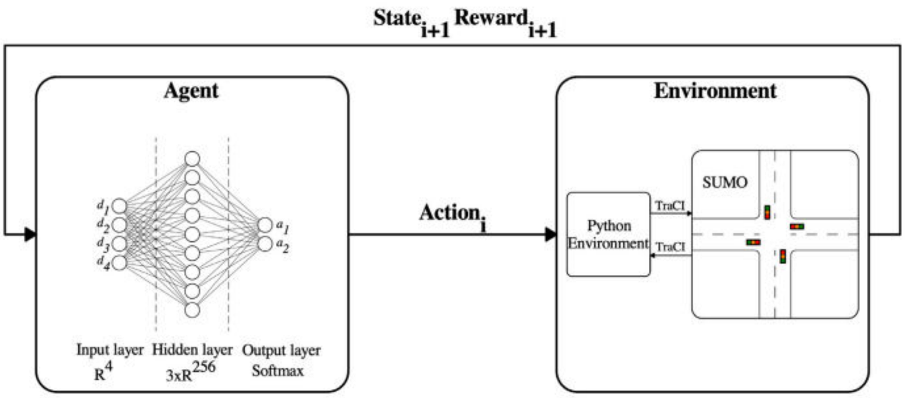
The problem addressed in this study is Traffic Signal Control (TSC) at a simple intersection, where vehicles can approach from all four directions, with one lane designated for each direction. For the sake of simplicity, left turns are excluded, allowing vehicles to either turn right or proceed straight out of the network. The training environment is simulated using SUMO (Simulation of Urban Mobility) (Lopez et al., 2018), which is an ideal platform for this type of modeling. SUMO supports both real-world transportation networks and synthetic environments, offering extensive customization and scalability options. Additionally, it features an excellent Python interface known as TraCI, which is essential for Deep

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Reinforcement Learning applications since many development tools and libraries, such as TensorFlow, PyTorch, and Keras, are built in Python. Figure 1 illustrates the integration of SUMO into the proposed RL training loop, highlighting how the environment is structured within SUMO and controlled through the Python training framework via TraCI.



The Python training loop is structured to comply with OpenAI Gym standards, allowing the agent's decisions to be translated into TraCI commands through the Python environment. These commands then initiate modifications in the SUMO infrastructure. The changes are monitored by TraCI, which collects relevant data and presents it back to the Python environment to accurately represent the state of the infrastructure to the agent.

Deep Reinforcement Learning (DRL) offers significant advantages, including real-time applicability, generalization, and scalability, which enable the trained neural network to effectively handle previously unseen scenarios. While a great deal of focus is placed on the algorithms to achieve high levels of generalization and scalability, the importance of the state vector that conveys the environment's conditions to the agent should not be overlooked. This aspect of DRL is critical for the algorithm's ability to assign credit, or determine how different pieces of information impact the decision-making process. Therefore, the design of the control task's abstraction is essential for developing a scalable and generalizable solution.

Consequently, state representation is one of the most vital elements, as the clarity and precision of the state descriptor directly affect the success of the training process. Moreover, the responsibility for crafting the state descriptor lies entirely with the researcher, meaning that the outcome heavily relies on their understanding of the control problem's abstraction. Similarly, careful consideration must also be given to the design of the reward signal in DRL, as it serves as the sole guide for the agent in identifying optimal behavior. In this paper, we aim to develop a state vector and reward scheme for the TSC problem that are easily generalizable and possess strong scalability potential.

VI. STATE

Image state representations are prevalent in traffic signal control problems. Many researchers prefer Discrete Traffic Signal Encoding (DTSE), in which each lane at an intersection is segmented into cells that extend from the end of the lane. The size of these cells is determined by the average length of vehicles. Each cell is populated with various types of data, such as speed, acceleration, vehicle position, and signal phase. This structured data can then be input into a Convolutional Neural Network (CNN) with (n) channels, where each cell functions like a pixel and (n) represents the different types of information contained in each cells.

Another common representation is the use of raw RGB images, which can incorporate roadside information that may be beneficial in certain scenarios. Alternatively, feature-based value vectors are frequently employed for state representation. In this approach, each lane is characterized by a set of values that might include phase duration, average speed, cumulative waiting time, phase cycle, and queue length. The primary advantage of this method is that the necessary data can be easily obtained from the intersection using loop detectors or road sensors.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

In this study, we adopt a feature-based approach that includes a single value for each lane: the count of vehicles currently present in that lane. This measure differs from the queue length, which is often the standard feature used in TSC state representations. Estimating the queue length accurately can be challenging, and it tends to provide delayed information regarding the lane's traffic load. By contrast, our chosen value offers a direct reflection of the current traffic load on each lane, allowing for more immediate decision-making related to traffic signals. To improve the training process and mitigate issues related to vanishing and exploding gradients, these vehicle counts are normalized to a range of [0, 1]. This straightforward relationship between vehicle count and traffic load makes it easier to inform traffic light management decisions.

VII. ACTION

In the established SUMO infrastructure, traffic can enter from four lanes at the intersection, where vehicles are permitted only to turn right or proceed straight out of the network. For this setup, we employ a simple discrete action framework consisting of two distinct actions. The first action activates the traffic signal for the horizontal lanes, turning them green while the vertical lanes are set to red; this configuration is referred to as North-South Green (NSG). The second action reverses this arrangement, turning the horizontal lanes red and the vertical lanes green, known as East-West Green (EWG). It is also essential to note that the selected traffic signal configuration remains unchanged for a duration of 5 seconds, as measured by SUMO's internal clock. Only after this period can the agent make a new selection for the traffic signals.

VIII. TRAINING

Training a traffic signal control system (TSC) involves a systematic approach to develop and refine algorithms that can optimize traffic flow at intersections. Initially, the training process begins by simulating a traffic environment, often using platforms like SUMO, where various traffic scenarios are generated to reflect real-world conditions. The TSC algorithm interacts with this simulated environment by making decisions based on the current state of traffic, which includes factors like vehicle counts, speeds, and waiting times. As the algorithm executes actions, it receives feedback in the form of rewards or penalties, which inform it about the effectiveness of its decisions. This feedback is crucial for adjusting the algorithm's parameters and improving its performance over time. Techniques such as reinforcement learning are commonly employed, allowing the TSC system to learn from trial and error, thereby enhancing its ability to adapt to dynamic traffic conditions. The ultimate goal of this training process is to create a robust TSC that can efficiently manage traffic flow, reduce congestion, and improve overall safety at intersections.

IX. LIMITATION

Traffic signal control systems (TSC) face several limitations that can affect their effectiveness and adaptability. One significant challenge is their reliance on fixed-time or pre-timed signal plans, which may not adequately respond to dynamic traffic conditions, leading to inefficiencies and increased congestion. Additionally, many traditional TSC systems lack the ability to incorporate real-time data, such as unexpected traffic patterns or incidents, which can hinder their responsiveness to changing circumstances. Furthermore, the complexity of urban environments, with varying traffic volumes and behaviors, can complicate the implementation of TSC algorithms, making it difficult to achieve optimal performance across different scenarios. Moreover, the integration of advanced technologies, such as artificial intelligence or machine learning, often requires substantial infrastructure investments and can pose challenges in terms of data privacy and security. These limitations highlight the need for more adaptive and intelligent traffic management solutions to improve overall traffic flow and safety.

X. CONCLUSION

In conclusion, the integration of reinforcement learning (RL) and the Policy Gradient algorithm in traffic signal control systems (TSC) offers a promising approach to address the complexities of managing urban traffic flow. By utilizing RL, TSC can dynamically adapt to real-time traffic conditions, learning optimal control strategies through interactions with the environment. The Policy Gradient algorithm, in particular, enhances this adaptability by enabling the direct optimization of the traffic control policy, allowing for the effective handling of high-dimensional action spaces and

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

complex state representations. This combination not only facilitates improved traffic management but also enhances the system's ability to generalize across different scenarios, making it a scalable solution for diverse urban environments. As the research progresses, the implementation of these advanced techniques is expected to lead to significant improvements in traffic efficiency, reduction of congestion, and overall enhancement of road safety, paving the way for smarter and more responsive traffic management systems.

REFERENCES

[1] Liang, X., Du, X., Wang, G., & Han, Z. (2018). "Deep reinforcement learning for traffic light control in vehicular networks." *IEEE Transactions on Vehicular Technology,*

[2] Wei, H., Zheng, G., Yao, H., & Li, Z. (2018). "Intellilight: A reinforcement learning approach for intelligent traffic light control." *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*

[3] Van der Pol, E., & Oliehoek, F. A. (2016**).** "Coordinated deep reinforcement learners for traffic light control." *Workshop on Machine Learning for Traffic and Transportation*Deep Reinforcement Learning based approach for Traffic.

[4] Genders, W., & Razavi, S. (2016). "Using a deep reinforcement learning agent for traffic signal control.

[5] Gao, J., Shen, Y., Liu, J., & Ito, M. (2017). "Adaptive traffic signal control: Deep reinforcement learning algorithm with experience replay and target network

[6] Chu, T., Wang, J., & Chen, L. (2019). "Multi-agent deep reinforcement learning for large-scale traffic signal control." *IEEE* Transactions on Intelligent Transportation System

[7] Zheng, Y., Yao, H., Xu, X., & Li, Z. (2019). **"**Deep reinforcement learning for intelligent traffic light control in IoT-based urban traffic systems." *IEEE Internet of Things Journal*

[8] Nishi, T., Takahashi, Y., & Phien, H. N. (2018). "Traffic signal control based on deep Q-learning with continuous action space for large-scale urban traffic networks." *IEEE Transactions on Intelligent Transportation Systems,*

[9] Choe, T. E., Bae, I. H., & Choi, H. K. (2018). "Deep learning-based traffic signal control considering traffic congestion." *Journal of Advanced Transportation, 2018*.

[10] Mousavi, S. S., Schukat, M., & Howley, E. (2017). "Traffic light control using deep policy-gradient and value-function-based reinforcement learning." *IET Intelligent Transport Systems,*