International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.8, Issue.10, October 2020 E-ISSN: 2347-2693

An Agent-Based Traffic Signal Control Using Reinforcement Learning Algorithm

O.E. Taylor[1*], P. S. Ezekiel[2], V.T. Emmah [3]

[1,2,3]Dept. of Computer Science, Rivers State University, Port Harcourt, Nigeria

[*]Corresponding Author: tayonate@yahoo.com, Tel.: +2348034448978

DOI: https://doi.org/10.26438/ijcse/v8i10.1722 | Available online at: www.ijcseonline.org

Received: 05/Oct/2020, Accepted: 12/Oct/2020, Published: 31/Oct/2020

***Abstract*—** Traffic light control has been a significant test in most major roads in Nigeria. The control of traffic has been so poor in certain spots in Nigeria to such an extent that more timing is being distributed to zones with lesser vehicles while little timing is being allotted to zones of more vehicles. This paper presents an Agent-based system to determine the control of traffic light signals using Reinforcement Learning algorithm by applying Deep Q Learning Techniques. The Reinforcement learning algorithm was trained using a Deep Q-learning technique with a total of 4 input layers, a batch size of 100, learning rate of 0.001 and a training epoch of 800 and a gamma of 0.97. The learning environment was made up with a maximum number of steps of 5400, total numbers of car generated to be 1000, green light duration in 10, yellow light duration to be 4. The number of actions taken by the agent equals 4 on 80 different states. The system helps in reducing traffic congestion by adapting to the learning environment, therefore knowing lanes with more vehicles during and without rush hours. By this, system optimizes the green time effectively by allocating more time to lane with more vehicles during and with rush hours, therefore, reducing the average cumulative delays and average cumulative queued length of vehicles. The result showed that system is efficient in traffic signal control with an average queued vehicle length of 5 to 20 vehicle

***Keywords*—**Reinforcement learning, Deep Learning, Traffic, Agent, Environment, Stimulation

I. INTRODUCTION

Traffic light control has been a significant test in most major roads in Nigeria. The control of traffic has been so poor in certain spots in Nigeria to such an extent that more timing is being distributed to zones with lesser vehicles while little timing is being allotted to zones of many vehicles. This by itself causes clog and postponements during times of heavy traffic. Ineffectively timing traffic light control have made life entirely deplorable for most Nigerians, subsequently causing elevated level of apprehension, making commuter not to keep to time and furthermore making them showing up late at night to their different destinations. Gridlock has been expanding in a significant part of the world, created or not, and everything demonstrates that it will keep on deteriorating, speaking to an undoubted threat to the nature of urban life. Its primary articulation is a dynamic decrease in rush hour gridlock speeds, bringing about increments in venture times, fuel utilization, other working expenses and natural contamination, as contrasted and a continuous traffic stream.

Traffic light in the urban territories is getting expanding complex with the exponential development in vehicle check. Development of the street system to oblige the expanded vehicle tally is certainly not a socially plausible alternative and is fundamental to build the use of the current foundation through appropriate guideline of traffic stream. Traffic lights were acquainted with control of the traffic stream, in this way improving the security of street clients. Notwithstanding, traffic lights make bottleneck for traffic stream in paths that don't have the option to proceed during a particular stage and improvement of signal timing is required to diminish the general defer experienced by all vehicles at the crossing point. Advancement can be acted in disconnected (pre-planned) or on the web (versatile) way [1]. Traffic light is one of the most significant mechanical methods for directing traffic stream, improving impediment, and improving its wellbeing and even vitality preservation and emanation decrease. At present, traffic light control issue not just makes some long-memories clog wonder at peak time, yet in addition has evident capacity of preparing in peak time. So as to facilitate the traffic pressure, discerning investigation and control are considered as a significant device. It's encouraging and improvements are continually keeping up with the occasions, joined by data innovation, PC innovation, and framework science [2]. This paper presents a reinforcement learning algorithm in controlling traffic signal.

II. RELATED WORK

Urban traffic signal control using reinforcement learning agents [1] conveyed multi-specialist based traffic light control for advancing green planning in an urban blood vessel street system to diminish the absolute travel time and defer experienced by vehicles. The proposed multi-operator engineering utilizes traffic information gathered by sensors at every convergence, putting away chronicled traffic examples and information conveyed from specialists in adjoining crossing points to figure green time for a stage. The boundaries like loads, edge esteems utilized in figuring the green time was calibrated by online fortification learning with a goal to lessen by and large deferral. Paramics programming was utilized as a stage to recreate 29 signalized convergences at Central Business District of Singapore and test the exhibition of their proposed multi-specialist traffic light control for various traffic situations. Their proposed multi-operator support learning (RLA) signal control indicated noteworthy improvement in mean time deferral and speed in contrast with other traffic light framework like progressive multi-specialist framework (HMS), agreeable group (CE) and impelled control.

Evaluation and Application of Urban Traffic Signal Optimizing Control Strategy Based on Reinforcement Learning [2] proposed a methodology for signal control plot improvement as showed by the unmistakable traffic stream traits, they parceled the sub territories subject to the three key limits of cycle length, vein coordination signal balance, and green split, a ton of different leveled control computations reliant on help learning was created to upgrade and improve the current sign arranging plan. In the sign control road sort out, each intermingling has its effect go, and the intersection point and region inside this range are essentially affected by it.

Evaluating reinforcement learning state representations for adaptive traffic signal control [3] tried to comprehend the information necessities and the exhibition contrasts in various state portrayals for support learning traffic light control. They displayed three state portrayals, from low to high-goal, and look at their presentation utilizing the offbeat preferred position entertainer pundit calculation with neural system work estimation in reproduction. Their Results show that low-goal state portrayals (e.g., inhabitance and normal speed) perform indistinguishably from high-goal state portrayals (e.g., singular vehicle position and speed). Their outcomes demonstrate actualizing support learning traffic light regulators might be conceivable with regular sensors, for example, circle indicators, and don't require modern sensors, for example, cameras or radar.

Asynchronous *n*-step Q-learning adaptive traffic signal control [4] applied support learning procedures with work guess to prepare a versatile traffic light regulator. They utilized the non-concurrent n-step Q-learning calculation with a two shrouded layer counterfeit neural system as our fortification learning specialist. They built up a unique stochastic busy time reenactment to test the specialist's exhibition. They thought about against customary circle locator activated and straight Q-learning traffic light control strategies, their fortification learning model built up a prevalent control strategy, decreasing mean all out postponement by up 40% without trading off throughput. In any case, they discovered that their proposed model somewhat builds delay for left turning vehicles contrasted with the impelled regulator, as an outcome of the prize capacity, featuring the requirement for a suitable prize capacity which really builds up the ideal approach.

The Real-time Traffic Signal Control System for the Minimum Emission using Reinforcement Learning in V2X Environment [5] concentrated on the reason of V2X condition, which makes changes in rush hour gridlock stream and the discharge were dissected dependent on infinitesimal traffic data. They said fortification learning model is built dependent on Deep Learning which learns the continuous traffic data and presentations the ideal traffic light. The exhibition of their framework was examined through minute traffic test system - Vissim. Their proposed framework is required to contribute on investigating the traffic stream and the natural impacts. Likewise, it is required to contribute on developing the green keen urban communities with an approach of independent vehicle activity in future V2X condition.

Policy Analysis of Adaptive Traffic Signal Control Using Reinforcement Learning [6] created two support learning versatile traffic light regulators, investigates their educated approaches, and contrasted them with a Webster's regulator. The offbeat Q-learning and favorable position entertainer pundit versatile calculations are utilized to create support learning traffic light regulators utilizing neural system work estimate with two activity spaces. They utilized a total measurement state portrayal (i.e., vehicle line and thickness), their proposed fortification learning traffic light regulators build up the ideal approach in a dynamic, stochastic traffic micro- simulation. Their outcomes show that the support learning regulators expands red and yellow occasions in any case accomplish better execution looked at than the Webster's regulator, lessening mean lines, halted time, and travel time. The support learning regulators show objective arranged conduct, building up a strategy that prohibits numerous stages found in a custom stage cycle (i.e., ensured turning developments) rather than picking stages that augment reward, instead of the Webster's regulator, which is compelled by repeating rationale that reduces execution.

Simulation and Optimization of Traffic in a City [7] developed a Reinforcement Learning algorithm which learns the waiting time of traffic lights (green and red) of vehicles at each point of intersection. They developed a green light district simulator using java programming language for testing the efficiency of their system which allows them to be able to edit infrastructures using the mouse and also to set various frequency signals and creating various traffic patterns. Their simulator itself was developed based on a cellular automaton model, and therefore it can be used creating different amounts of detail.

Simulation of modern Traffic Lights Control Systems using the open source Traffic Simulation Sumo [8] developed an agent based traffic logic algorithm which takes in the length of traffic jam as input data and they tested it on an Open Source Simulation (Sumo). Their algorithms solves the problem of traffic jam by looking into the coming lane, checking the numbers of queened vehicles. If the number of queued vehicles in the incoming lane is longer than that of the current lane, a green light signal will be given to the incoming lane. The also implemented and Optical Information System (OIS) sensors. The OIS sensor uses detector for simulation of traffic light. The OIS sensors looks at all the lanes. The OIS sensor do not only check the number of queued vehicles but it also checks mean speed and the halting duration.

A dynamic and automatic traffic light control expert system for solving the road congestion problem [9]. Developed a dynamic and automatic traffic light control expert system coupled with a model for stimulation. The stimulating model is made of six sub models written in Arena in other to have a better analysis of traffic problem. The model uses arrival and departure time in stimulating the incoming and going vehicles on roads. Each of the sub models represents a road that is made up of three intersection points. After testing, they got their average waiting time of cars at every intersection point to be 65s and the green light signal to be 125s.

The impact of traffic-light-to-vehicle communication on fuel consumption and emissions [10]. Presents a sensitivity analysis and identification of gear choice and distance from the traffic light at which cars are being informed as a key factor which can be influencing. They carried out some experiments which shows that a suboptimal gear choice can void the benefits of the adaptation of speed. They also present a scale-up simulation using a real world inner city road network.

III. METHODOLOGY

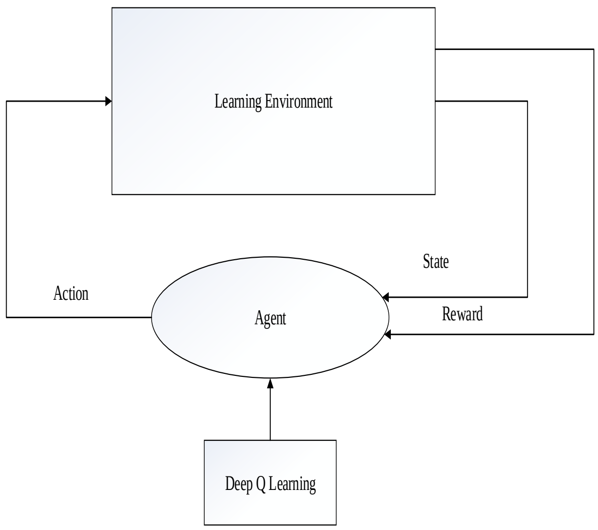


Figure 1: Architecture of the proposed system

This system uses a Reinforcement learning algorithm in training a building/training an agent that will stimulate the control of traffic. The following are some basic explanation of some terms used in training/ building the agent:

Environment: The *environment* is typically a set of states the agent is attempting to influence via its choice of actions or a task or simulation and the Agent is an AI algorithm that interacts with the environment and tries to solve it.

Agent: This is an Artificial Intelligent algorithm that is been trained to interact/learn in the learning/training environment in other to accomplish a specific task. The ability of the agent in looking forward to future reward, gamma = 0.97

Action: These are the steps or methods taken by the agent to interact with the learning environment, and thus change positions in other to achieve a better result. The number of actions taken by the agent is 4. Which is North, South, East and West, making it a total of four moving lane.

State: The states are different positions taken by the agent in the environment. The number of states at which the agent moves is 80.

Reward: Reward determines the outcome or result of the agent after performing various actions at different states in the learning environment.

Deep Q Learning: This is a reinforcement learning algorithm which is being used in learning a policy, telling the agent what actions to take at certain defined conditions. The algorithm uses an episode of 50, a total layer of four (4), batch size of 100, learning rate = 0.001 and a training epoch of 800.

$$Q(s, a) = r(s, a) + \gamma \max_{a} Q(s', a) \quad .. (1)$$

$$Q(s, a) \rightarrow \gamma Q(s', a) + \gamma^2 Q(s'', a) \ldots\ldots \gamma^n Q(s''^{\ldots n}, a) \ldots (2)$$

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R_{t+1} + \gamma \max_{a} Q(S_{t+1}, a) - Q(S_t, A_t)] \ldots (3)$$

Where:

Q is the Q learning factor

s and a are actions carried out by the agent on a particular state

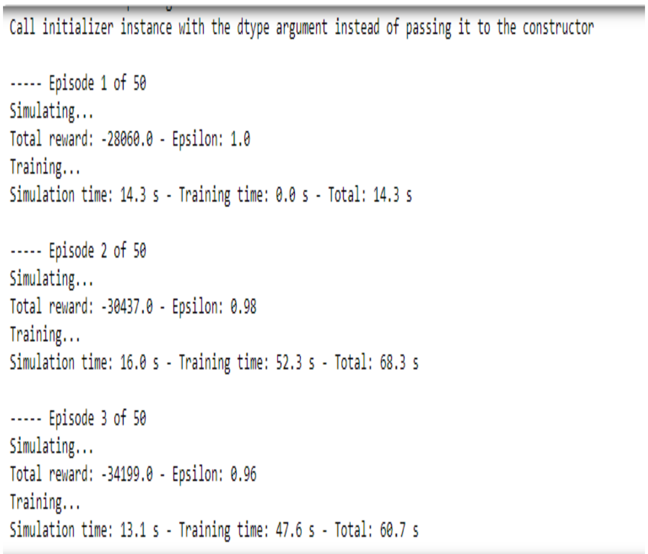
γ is the gamma

α is the rate at which the agent learns in the environment

t is the time taken by the agent in completing one action in a state

IV. RESULTS AND DISCUSSION

This paper uses an Agent-based approach in controlling the flow of traffic using Reinforcement Learning algorithm on a Sumo GUI environment. The agent tries to choose a better traffic phase at an intersection point in other to optimize a better traffic efficiency. The agent was designed with a state representation in other to identify the positions of various vehicles in the learning environment. The actions of the agent was defined by a set of configurations with a fixed duration of time. The agent was tested on a Simulation of Urban Mobility (Sumo) environment to replicate a four way intersection point. Sumo is an open source programming, exceptionally convenient and constantly street traffic recreation bundle intended to deal with enormous street systems. Sumo has a traffic signal control Interface was being utilized, offering access to a running street traffic recreation, it permits to recover estimations of reproduced objects and to control their conduct on-line. The Reinforcement learning algorithm was trained using a Deep Q-learning technique with a total of 4 input layers, a batch size of 100, learning rate of 0.001 and a training epochs of 800 and a gamma of 0.97. The learning environment was made up with a maximum number of steps of 5400, total numbers of car generated to be 1000, green light duration in 10, yellow light duration to be 4. The number of actions taken by the agent equals 4 on 80 different states. The system helps in reducing traffic congestion by adapting to the learning environment (The traffic environment which has 4 moving lane from which vehicle pass from), therefore knowing lanes with more vehicles during and without rush hours. By this, system optimizes the green time effectively by allocating more time to lane with more vehicles during and with rush hours, therefore, reducing the average cumulative delays and average cumulative queued length of vehicles. The performance of the agent shows the ability of the agent to determine the average waiting time, and the average cumulative delay of vehicles at the four-intersection point, therefore making the agent to take a better decision in allocating the green light signals to the lane with highest waiting vehicles. The System had efficiency in traffic signal control by having an average queued vehicle length of 5 to 20 vehicles.


```
Call initializer instance with the dtype argument instead of passing it to the constructor

----- Episode 1 of 50
Simulating...
Total reward: -28060.0 - Epsilon: 1.0
Training...
Simulation time: 14.3 s - Training time: 0.0 s - Total: 14.3 s

----- Episode 2 of 50
Simulating...
Total reward: -30437.0 - Epsilon: 0.98
Training...
Simulation time: 16.0 s - Training time: 52.3 s - Total: 68.3 s

----- Episode 3 of 50
Simulating...
Total reward: -34199.0 - Epsilon: 0.96
Training...
Simulation time: 13.1 s - Training time: 47.6 s - Total: 60.7 s
```

Figure 2: The number of first 3 training episode.

Episode are states that come in between an initial-state and a terminal-state. Here, the goal of the agent is to maximize the total number of rewards received during an episode.

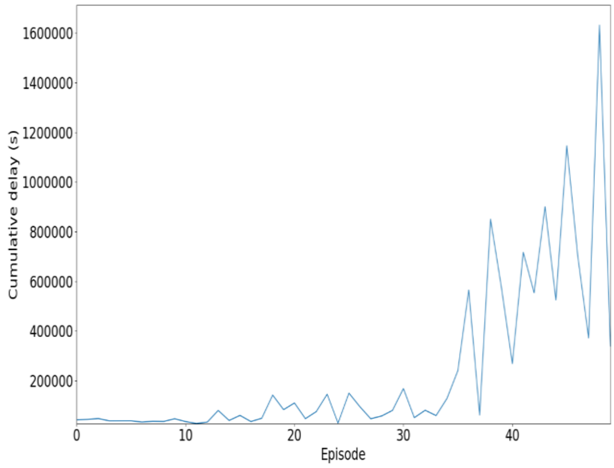


Figure 3: Cumulative delays of the vehicles during and episode

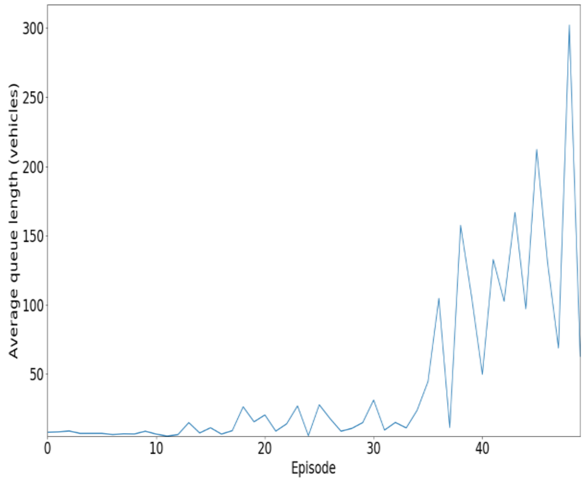


Figure 4: Graphical representation of episode vs cumulative delay. This is a graphical representation summing the total delays of vehicles at each episode.

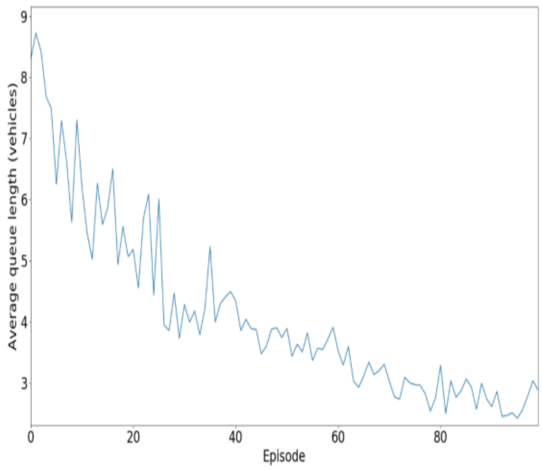


Figure 5: The average queued length of vehicles on each episode

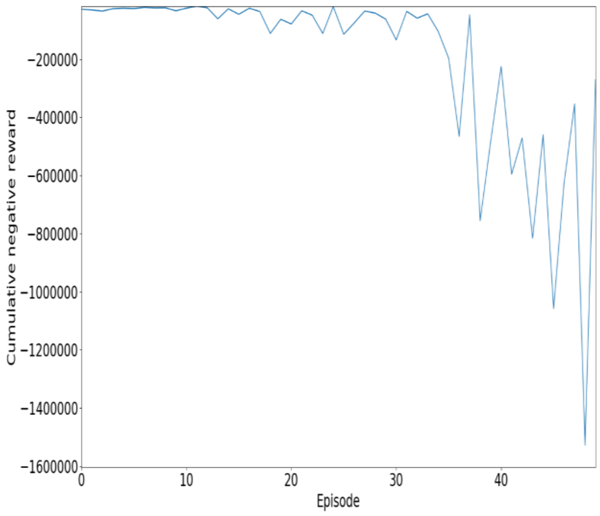


Figure 6: The number of rewards gotten by the agent within 0-50 episode

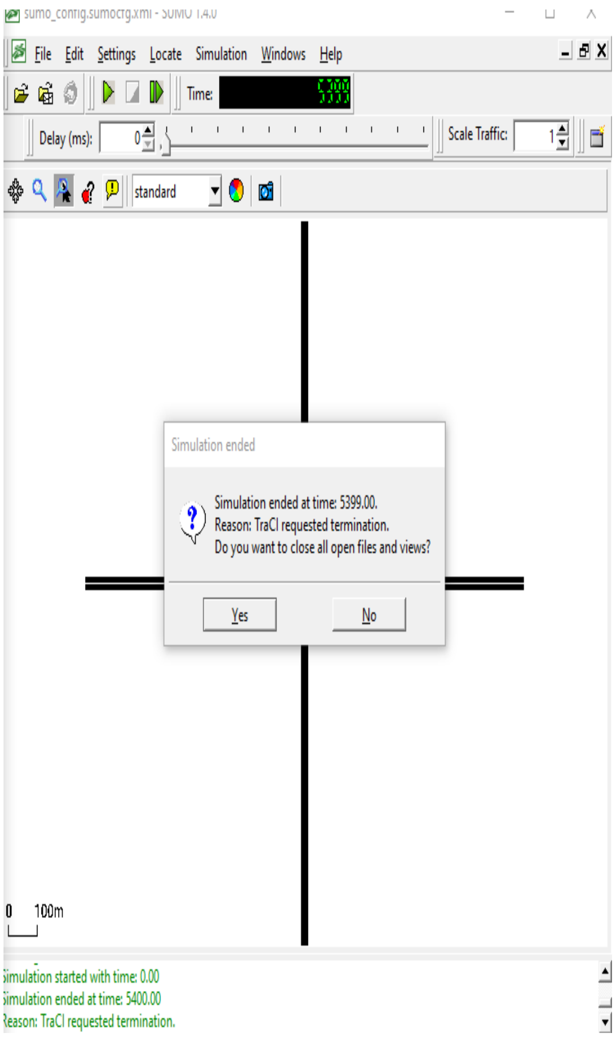


Figure 7: Testing environment on Sumo environment with a stimulation time of 5399 secs.

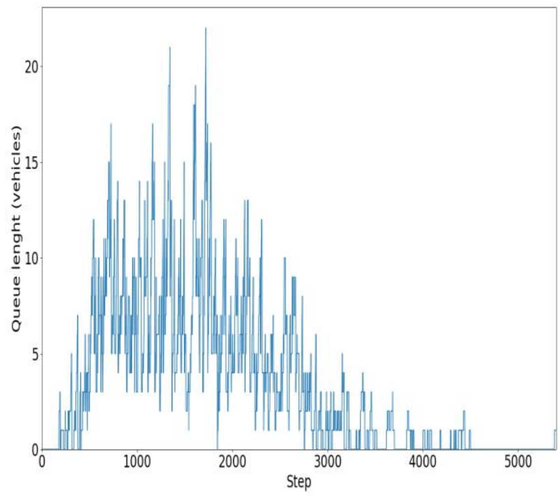


Figure 8: The average queued length of vehicles tested in Sumo GUI at different steps

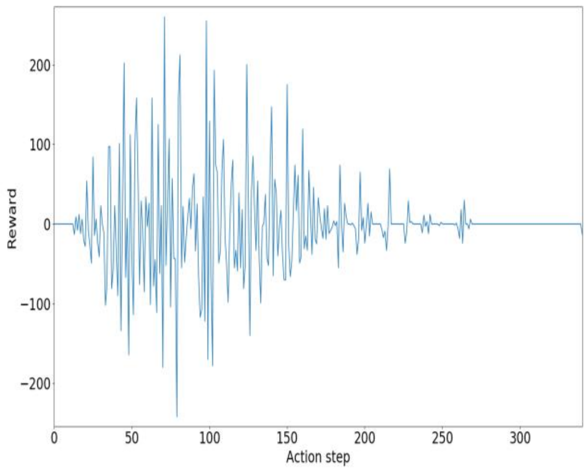


Figure 9: The reward gotten by the agent on the Sumo GUI testing environment at different action steps.

V. CONCLUSION AND FUTURE SCOPE

In this paper presents an Agent based system to determine the control of traffic light signals using Reinforcement Learning algorithm by applying Deep Q Learning Techniques. The agent tries to choose a better traffic phase at an intersection point in other to optimize a better traffic efficiency. The Reinforcement learning algorithm was trained using a Deep Q-learning technique with a total of 4 input layers, a batch size of 100, learning rate of 0.001 and training epochs of 800 and a gamma of 0.97. The learning environment was made up with a maximum number of steps of 5400, total numbers of car generated to be 1000, green light duration in 10, yellow light duration to be 4. The number of actions taken by the agent equals 4 on 80 different states. The system helps in reducing traffic congestion by adapting to the learning environment (The traffic environment which has 4 moving lane from which vehicle pass from), therefore knowing lanes with more vehicles during and without rush hours. By this, system optimizes the green time effectively by allocating more time to lane with more vehicles during and with rush hours, therefore, reducing the average cumulative delays and average cumulative queued length of vehicles. The System had efficiency in traffic signal control by having an average queued vehicle length of 5 to 20 vehicles. This paper can further be extended by taking it into a real system where agents will control the flow of traffic, and it can also by using Double Q Learning Technique in training the agent.

REFERENCES

[1]. P.G. Balaji, X. German, D. Srinivasan "Urban traffic signal control using reinforcement learning agents", IET Intelligent Transport Systems, **Vol.**4 **issue**.3 **pp.177-188, 2010**.

[2].W. Yizhe, X. Yang, Y. Liu, H. Liang "Evaluation and Application of Urban Traffic Signal Optimizing Control Strategy Based on Reinforcement Learning", Journal of Advanced Transportation, **Vol.**2018 **issue**.1494 **pp. 1-9**, **2018**.

[3]. G. Wade and R. Saiedeh "Evaluating reinforcement learning state representations for adaptive traffic signal control", Procedia Computer Science **Vol.** 130, **pp. 26-33**, **2018**.

[4].G. Wade and R. Saiedeh "Asynchronous *n*-step Q-learning adaptive traffic signal control", Journal of Intelligent Transportation Systems**, Vol**.23 **issue**.4, **pp.319-331**, **2019**.

[5]. J. Kim, J. Sangchul, K. Kwangsik, L. Seungjae "The Real-time Traffic Signal Control System for the Minimum Emission using Reinforcement Learning in V2X Environment" Intellian Association of Chemical Engineering **Vol**. 72, **pp.91-96**, **2019**.

[6]. G. Wade and R. Saiedeh "Policy Analysis of Adaptive Traffic Signal Control Using Reinforcement Learning" Journal of Computing in Civil Engineering **Vol.**34 i**ssue** 1 **pp.1-5**, **2020**.

[7]. M. Wiering, J. Vreeken, J. Veenen, A. Koopman "Simulation and Optimization of Traffic in a City" Conference paper, IEEE, Intelligent Vehicle Symposium **2004**.

[8]. K. Daniel, B. Elmar, M. Jurgen, R. Julia, C. Rossel, T. Wolfram, W. Peter, W. Richard "Simulation of modern Traffic Lights Control Systems using the open source Traffic Simulation" Proceedings of the 3rd Industrial Simulation Conference **2005**.

[9]. W. Wen "A dynamic and automatic traffic light control expert system for solving the road congestion problem" Expert System Applications, **Vol**.34 **issue.**4, **pp. 2370-2381 2008**.

[10]. M. Killat, H. Hartenstein, R. Luz, S. Hausberger, T. Benz "The impact of traffic-light-to-vehicle communication on fuel consumption and emissions", Conference of Internet of Things (IOT). **2010**.