Smart Traffic Management System for smart cities using Reinforcement Learning Algorithm

D. Venkata Siva Reddy, R.Vasanth Kumar Mehta

Abstract— ***Traffic congestion at junctions or on roads may be seen due to many reasons like slow driving, increased vehicle queue etc. In some emergency cases when vehicles are stopped for longer period traffic jam may occur. Adaptive traffic control system is a traffic management strategy used to control the traffic by facilitating the signals to instantaneously adjust to the present traffic demand. Adaptive traffic signal functions by utilizing both hardware and software coordination. Q-learning needs already designed precise form of the environment for selecting action. As a substitute, we can adopt dynamic communication system to find the interaction between state, action and rewards of that particular environment. The present traffic signal works based on the pre specified traffic flow data to extract short time anticipation which helps to evaluate the consequence on the signal controlling system.***

If single model is used then at each and every time adaptive traffic light control agents need to collect photographs of the existing condition of the traffic and generate control signals. We have implemented occurrence, replay and ideal mechanisms to improve the consistency of the algorithm. The main aim in designing the algorithm is to control overcrowded traffic and for this we have incorporated the dynamic network with the linear signal arrangement.

Keywords: Traffic control, Reinforcement learning, Deep learning, Value-function method and Artificial neural networks.

1. INTRODUCTION

Drastic Urban development has simultaneously resulted in rapid increase in the road transportation. As a result the concept of transportation has become an important criterion in maintaining and also developing the smart cities. The first priority is to reduce the possibility of over jamming of vehicles and also see smooth passage of automobile without causing disturbances to other passerby. The other important goal is to control the air pollutants that are emitted by the motor vehicles into the atmosphere causing drastic effect on the atmosphere and also result in serious health hazards. Based on the various types of work done, different suggestions are deduced. In order to fabricate perfect traffic control system, the whole process is categorized into certain sets. Vehicle actuated signal control is one such group where depending on the traffic situation the traffic signals are controlled. This is done by using inductive loop detectors. Next is adaptive signal control. Here based on the situation of the traffic the signal time is managed automatically. Last method is pre timed signal controlling system where based on the traffic demand for some remarkable reasons the traffic signal (green) is fixed for stipulated time period.

Revised Version Manuscript Received on March 10, 2019.

D. Venkata Siva Reddy, Research Scholar, Dept. of Computer Science & Engineering, SCSVMV University, Kanchipuram, Tamil Nadu, India.(E-Mail: lionshivareddy@gmail.com)

Dr. R.Vasanth Kumar Mehta, Associate Professor, Dept. of Computer Science & Engineering, SCSVMV University, Kanchipuram, Tamil Nadu, India. (E-Mail: Vasanth.mehta@kanchiuniv.ac.in)

Here we are trying to work on adaptive signal controlling system using machine learning and AI. We tried to apply Reinforcement learning technique which is a type of machine learning and also comes under the AI branch. We found the results satisfactory. Here the main advantage is that prior knowledge regarding the situation is not required but rather it will set the requirements as it get interacted with the environment and acts accordingly. Reinforcement signal is essential to find the behavior. The advantage of the reward is that the algorithm of reinforcement learning can be used to learn universal requirements once or can keep on adapting. The reason in applying reinforcement learning in traffic controlling system because it requires only slight knowledge about domain application. The agent studies the performance or modifications that happen in the environment and act by applying trial and error method. It obtains scalar reward after receiving each action. The reward that is received depends on the performance of the action. The objective of the agent is to find out an finest control policy so that the reward obtained will be maximum passing through frequent dealings with the respective environment.

Figure1.1: Reinforcement Learning in Traffic Control System.

Compound optimization error can be successfully handled by using reinforcement learning technique and thus Deep learning technique has drawn enormous fascination. Seeing the various merits in combining reinforcement with deep learning method we have decided to work on setting proper and effective traffic controlling system using these effective techniques.

2. RELATED WORK

Our work is mainly related with adaptive traffic control technique and so we focused on the previous works that are done in this subject. There are array of significant works that

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

are done in developing effective methods in traffic controlling. Although these results appeared to be assuring but carry certain drawbacks when its application comes to practical circumstances. Refined traffic model has helped in extending reinforcement learning with new representative and reward functions. This possibly may be applicable to tricky and also to actual traffic problems. On observation it was clear that the light control in traffic maintenance comes under Markov decision.

Keyarsalan et al employed fuzzy ontology. He applied this system to manage the traffic light signals at remote junctions with the help of vision system. To accomplish this they used the software technology and also the knowledge of neural networks to get the required traffic records. The proposed system is comparatively more acceptable compared to other proposals. Q-learning method was applied on the available fixed environments by Abdoos et al. Here they have mainly worked on the length or the queue of the vehicles in traffic junction. Thus the proposal is mainly done based on the calculation of the queue. From the result it is concluded that the Q learning works efficiently when applied on fixed time in particular traffic. Multi-agent reinforcement was applied by Arel etal to attain successful traffic control strategy. This method mainly intends to reduce the usual interruption, jamming and the possibility of overcrowding in the traffic circles. Zhu et al. initiated reinforcement learning method which is appropriate in controlling proper traffic signal. The signal lights are designed in such a way that these lights act as smart instrument in controlling the complex traffic situation. We conclude that the results account that the proposal shows good performance and helps in estimating and controlling the traffic based on the vehicles average delay, total stops expected etc.

Richter devised a method which he decided to be partially observable Markov's Decision Problem (POMDP) for traffic related problems. He utilized policy gradient methods to assure confined junction with in a partial observable environment. Since there has been certain improvement in Deep learning, it has now gained its means in traffic controlling process. Earlier researchers used deep stacked auto encoders (SAE) neural networks to guess Q-values. Here each Q-value is related to each accessible signal phase. In the Richter proposed method the speed and the queue of the vehicles will be calculated each and every time so that the proposed algorithm gets adjusted to the environment. Lillicarp, J. Hunt and etal, have given means to use deep Q-network to plot Q-values from particular condition. The important points that were considered during their study were position, speed of the motor vehicle and also the traffic signal point but in our work we are planning to use the traffic visuals collected from the important traffic junctions as our input data. We are also trying to suggest a method which can directly record the probability distribution from the provided input snaps through deep policy gradient to the action prevailing at that period of time.

3. REINFORCEMENT LEARNING

Traffic system with signal control was initially studied by Thorpe in 1997 using reinforcement learning technique. He actually used RL method to minimize the time required to discharge a fixed volume of traffic through a road network. Thorpe used neural network for light based value function to calculate the cars stopping time. Here to get the output the controllers must consider massive situations. State action, reward state action are different types of reinforcement learning used by Thorpe and could test only single light controller using 4x4 grid. Here the controller is able to give decision by consider only one direction of the traffic junction. Thus it seems that the Thorpe's concept cannot be applied to real time project in dealing traffic congestion near traffic signal junctions. Thus we are slightly changing our concept by applying reinforcement method which is model based and instead of light based we used car based value functions.

4. DEEP LEARNING AND DEEP Q-LEARNING

Reinforcement learning helps the system to decide whether the choice taken is precise or not. Reinforcement learning generally predicts its outcome only after number of repetitions (iterations) hence this method has chances of producing right decisions. Thus it has achieved recognition in most of the machine learning projects. The agent in RL method tries to produce superb action for the particular given state. Trial and error method helps in extracting more knowledge thus is capable to decide whether to give positive or negative reward to the given set of state action pair.

Deep learning also comes under machine language. Deep learning prepares the system to perform calculations iteratively. The state of the environment and the data obtained is generally categorize, forms clusters and finally forecast predictions based on the acquired information. The working procedure of the deep learning resembles to that of human brain. Neural networks forms the center of the deep neural network and it consists of nodes (artificial neurons) which are mainly responsible in carrying out computations. There are numerous layers and every layer possesses certain nodes.

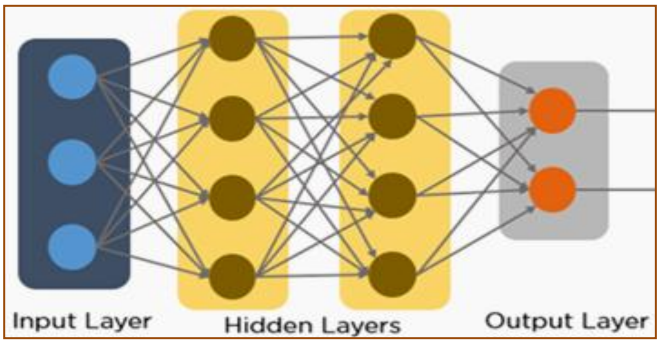

Figure 4.1: Neural networks

The neural networks generally comprise of three layers namely first is the input layer, second is the hidden layer and last is the output layer. The input layer collects the input of different configuration. Hidden layer execute feature extraction, mathematical and other computations. The final layer produces output. Now the expected output is compared with the generated output and then if there are any errors then by employing cost function it is estimated. Now the estimated

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

error is back propagated to reduce the obtained errors so that the output obtained will be almost similar to expected.

5. RESULTS

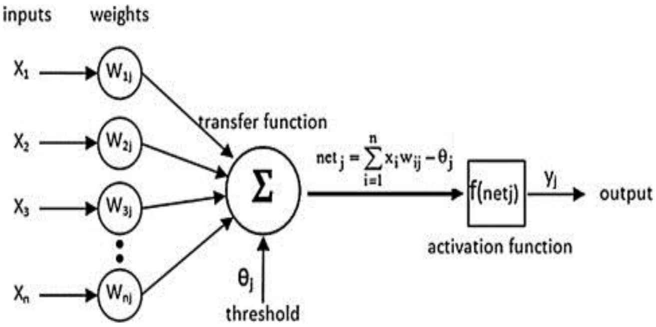

Figure 4.2: Artificial neural network.

The output of one layer is considered as the input of the subsequent layer in deep neural networks. At each and every layer a nonlinear transformation is pertained to study and attain descriptive factors. Deep Q-learning Network (DQN) utilizes this assistance of deep learning to signify the agent's examination as a sign in knowing an optimal control policy. The DQN method combines deep neural network function roughly with Q-learning to study action value function. π represents policy which will inform the agent type of action to be chosen for every input state. When non-linear function is applied to the neural networks then it is found that it will roughly coincide with the reinforcement learning. When high aspect of the continuous and action space is considered some sort of congregation is seen.

The main cause for this dilemma may be due to repeated states in reinforcement learning duties have connection and the fundamental policy of the agent is altering regularly, because of small variations in Q-values. Deep Q Learning Network gives some answer to overcome these difficulties and as a result the algorithm can execute considerably. In case of interrelated states DQN applies experience replay concept. As a result every time the agent's knowledge is stored into data set. Here data set is represented by D and the state is represented as st, action that is selected as at and as a result of action the reward that is received is represented as rt. The state that is considered in the next step is represented as st+1.The DQN employ small sets consisting random sampling which are uniform and are capable of repeating the prior contact at the time of training. This reverses the irregular connections seen among the successive samples. There is another alternative method called policy gradient which has shown good result in handling convergence problem.

Artificial intelligence is used to device algorithm that can be applicable in real time traffic data collection. Since the algorithm does the work automatically there will be no need of engineer to manually input or monitor the data. Five different tools are used for collecting the traffic information automatically using deep learning method. It is capable in assembling and mining the data related to the traffic condition.

6. DESCRIPTION OF TEST-BEDS

A computer generated two stage signal controlling system is connected in the remote traffic signal test-bed near the junction of two two-lane roads. Individual Poisson process helps in extracting the information regarding the onset time of the vehicle on the four probabilities based on the pre-distinct average arrival rates.

Generally if considered practically the agent works constantly but when we consider linear signal system the independent Q-learning agents that control two phase junctions separately consists of test beds. Poisson processes entities are used to fabricate vehicle onset on each approach to the system. Miniature level arrangement is used to copy the exact traffic movements. In order to accomplish this road connection is divided into blocks so that the vehicles will move on one block for each time but in some cases if there is no place available on downstream connection then in such cases the vehicles which enter will be struck there until there is a clearance to move ahead near the traffic circle. This may result in heavy traffic jam. If we consider isolated junctions then the state information obtained consists of the length of the queue and intervene time phase on the four approaches. If multi agent is adopted then there will be an advantage to get added state details because there will be widen chances of interaction among the agents. As a result the scrutiny of each individual agent on the traffic vision will become even more effective.

In addition to local queue lengths, different array of upstream and downstream queue lengths and the conclusion of signal modify controlling upstream and downstream movements are also estimated as state elements. The increase of elements to the state definition noticeably elevates the dimension of the state space. Thus a balance is required between the advantage and the effect seen due to the collected information on dealing the problem. Evaluating the length of the queue may be competently accomplished using video imaging technology along with artificial neural network or other pattern recognition techniques.

7. CONCLUSION

Traffic overcrowding can be resolved with our proposed algorithm for adaptive signal control using reinforcement learning. This method helps to obtain required information automatically from the real time traffic records. Initially we tested our proposal on small function and then expanded for more complicated situations. We employed multi agents to tackle synchronization problem faced between the different agents. The algorithm which we proposed can gather the important traits automatically by using convulsion neural networks. Further we are interested to work on the importance of green waves in reinforcement learning. The various patterns i.e. stationary and non-stationary traffic patterns may be studied further to make the traffic control system even more practical in use.

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*

REFERENCES

1. R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction. MIT Press, 1998.
2. M. Duggan, J. Duggan, E. Howley, and E. Barrett. An autonomous network aware vm migration strategy in cloud data centres. In Cloud and Autonomic Computing (ICCAC), 2016 International Conference on , pages 24–32. IEEE, 2016.
3. W. Genders and S. Razavi, "Using a deep reinforcement learning agent for traffic signal control," November 2016, [Online]. Available: https://arxiv.org/abs/1611.01142.
4. T. Tieleman and G. Hinton, "Lecture 6.5-rmsprop: Divid e the gradient by a running average of its recent magnitude," COURSERA: Neural networks for machine learning , vol. 4, no. 2, 2012.
5. T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez , Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," February 2016, [Online]. Available: https://arxiv.org/abs/1509.02971.
6. D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of sumo simulation of urban mobility,"International Journal On Advances in Systems and Measurements,vol. 5, no. 3 & 4, pp. 128–138, December 2012.
7. R. Wunderlich, C. Liu, and I. Elhanany, "A novel signal-scheduling algorithm with quality-of-service provisioning for an isolated intersection," IEEE Transactions on Intelligent Transportation Systems , vol. 9, no. 3, pp. 536–547, September 2008.
8. P. Balaji, X. German, and D. Srinivasan. Urban traffic signal control using reinforcement learning agents. IET Intelligent Transport Systems, 4(3):177–188, 2010.
9. G. Barto and S. Mahadevan. Recent advances in hierarchical reinforcement learning. Discrete Event Dynamic Systems, 13(4):341–379, 2003.
10. J. Baxter, P. L. Bartlett, and L. Weaver. Experiments with infinite-horizon, policy-gradient estimation.Journal of Artificial Intelligence Research, 15:351–381, 2001.
11. E. Brockfeld, R. Barlovic, A. Schadschneider, and M. Schreckenberg. Optimizing traffic lights in a cellular automaton model for city traffic. Physical Review E, 64(5):056132, 2001.
12. Y. K. Chin, N. Bolong, A. Kiring, S. S. Yang, and K. T. K. Teo. Q-learning based traffic optimization in management of signal timing plan. International Journal of Simulation, Systems, Science & Technology, 12(3):29–35, 2011.
13. M. Abdoos, N. Mozayani, and A. L. Bazzan. Holonic multi-agent system for traffic signals control.
14. Engineering Applications of Artificial Intelligence, 26(5):1575–1587, 2013.
15. B. Abdulhai, R. Pringle, and G. J. Karakoulas. Reinforcement learning for true adaptive traffic signal control. Journal of Transportation Engineering, 129(3):278–285, 2003.
16. Arel, C. Liu, T. Urbanik, and A. Kohls. Reinforcement learning-based multi-agent system for network traffic signal control. IET Intelligent Transport Systems. 4(2):128–135, 2010.