Intelligent Traffic Control Based on Multi-armed Bandit and Wireless Scheduling Techniques

Chanwoo Park and Jungwoo Lee
School of Electrical Engineering and Computer Sciences
Seoul National University, Seoul 151-744, Korea
Email: cpark@wspl.snu.ac.kr and junglee@snu.ac.kr

***Abstract*—Intelligent Transportation System (ITS) researches, including vehicular communications, have been making great advancements to improve road safety and traffic flow efficiency. In this paper, we propose two new traffic control systems. In the first method, we assume a system with fully autonomous cars and infrastructure to avoid collision completely. Vehicles communicate with the access point in both random access mode and polling mode, and the movement of the automobiles will be coordinated by the infrastructure using IEEE 802.11 DCF/PCF mechanisms. In the second method, there is a given set of lanes with unknown reward statistics and we consider the lanes as a multi-armed bandit. We use multi-armed bandit algorithm to choose the best lane to drive in and to maximize the total expected reward while minimizing the regret. Traffic congestion is very difficult to predict and deal with because it is a function of many unknown factors such as number of cars, weather, road constructions, and accidents. The proposed algorithms are designed for urban road networks to ease the congestion, and make it more predictable at the same time. We find that the first algorithm makes the traffic system able to balance efficiency and fairness and the second algorithm helps vehicles choose the best lane with minimizied regret.**

***Keywords*-ITS; Vehicular Network; IEEE 802.11 DCF/PCF; MAB.**

I. Introduction

Driverless cars are vehicles with fully automated driving capabilities [1]. In many urban environments, a rapid increase in the number of cars has caused severe problems such as traffic congestion, air pollution and road safety. Researchers have been putting a lot of effort into developing new types of transportation systems (e.g., driverless cars) as a solution to this problem. Researchers first pondered the idea of driverless cars in the 1970s [2]. Since then, there have been many prototypes of driverless cars tested and lots of research and development on driverless cars going on. VisLab (Artificial Vision and Intelligent Systems Laboratory) has successfully completed the rally of 13,000 km from Milan to Shanghai on driverless vehicles in 2010. There has been active research on vehicle network going on to develop interactive system enabling a number of new services for road safety, mobility and efficiency such as Vehicle Infrastructure Integration (VII) [3] and the California Partners for Advanced Transit and Highways (PATH) [4].

In many cities, especially in large metropolises, traffic congestion during rush-hours is one of major problems. Traffic congestion is a very tricky problem to deal with not only because it makes trip times longer and increase vehicular queuing but also because there are too many variables on road networks, such as number of cars at a certain time, weather of that day, unexpected road construction and car accidents etc. Because of this uncertainty of road networks, it is very difficult to predict and deal with the traffic congestion properly.

There are two basic assumptions in this paper. First, vehicles are fully self-driven, which means each vehicle knows its destination and drives from one place to another without input from a human operator. Secondly, we assume that the system is established to completely avoid collisions between cars. This central system manages the car network to make the road environment collision-free.

Based on these assumptions, we propose an algorithm which is used for traffic control when there is no traffic light at intersections. The basic principle of this algorithm is that the system gives priority to the lane which has the longer queue of cars so that more congested lanes can be relieved more quickly. This will make travel times during rush-hour more predictable. There are two values we have to consider when it comes to traffic control without traffic lights, which are flow efficiency and fairness between users. The proposed algorithm in this paper uses IEEE 802.11 DCF/PCF Mechanisms to balance these two values.

The rest of this paper is organized as follows: Section II introduces the proposed system using IEEE 802.11 legacy DCF/PCF and its performance. After describing the second proposed system which is based on multi-armed bandit algorithm in Section III, the paper concludes with Section IV.

II. Traffic Control Based on IEEE 802.11 DCF/PCF Mechanisms

The system algorithm and the system model are introduced.

A. IEEE 802.11 legacy DCF/PCF

The IEEE 802.11 standard makes it mandatory for all stations to implement the Distributed Coordination Function (DCF), a form of Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) [5]. CSMA is a contention-based protocol which makes sure that all stations first sense the medium before transmitting. The main goal is not to have stations transmitting at the same time, which results in collisions and corresponding retransmissions. Probabilistically speaking, they have the same opportunities when stations contending for

 ISBN: 978-1-61208-241-7

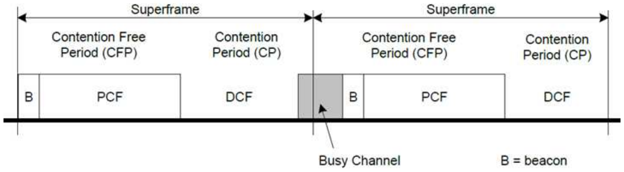

Fig. 1. IEEE 802.11 legacy DCF/PCF operation between beacon intervals

medium access in DCF mode. Each station has its own random back off timer when contending so they can achieve fairness in the long-term.

There is another optional access method, namely, the point coordination function (PCF) based on poll-and-response mechanism. PCF is intended to transmit real-time information such as VoIP, streaming video. In PCF mode, a point coordinator within the Access Point (AP) controls which stations can transmit during a certain given period of time, which is called the Contention Free Period (CFP). The point coordinator will take a look through all stations which are operating in PCF mode and poll them one at a time. Therefore, PCF is a contention-free protocol enabling stations to transmit data frames continuously.

AP sends beacon frames at regular intervals so that the IEEE 802.11 protocol makes stations alternate between the use of DCF and PCF in a single interval. With DCF, stations will compete for the channel access by using CSMA. For the following CFP, the stations will wait for a poll from the point coordinator before sending data frames as shown in Fig. 1. Therefore, DCF is basically a protocol based on random contention so it aims for fairness while PCF is a protocol controlled by the point coordinator trying to give opportunities to stations which need to be served first. In the following section, we discuss how we can apply this IEEE 802.11 DCF/PCF mechanism to traffic control system.

B. System Model

There is a four-way intersection and each road has eight lanes. In the ith lane, cars are generated by Poisson distribution with expected number (arrival rate) λ_i every time slot. This system assumes that all cars are driverless and safely controlled by the car network system so that collision avoidance system is perfectly implemented. In each direction, the first lane is dedicated for left-turns, the second and third lanes are for cars going straight and the fourth lane is only for right-turns. This is described in Fig. 2.

1) Contention-Free Period (CFP): The system divides each repeat interval into two parts, Contention Period and Contention-Free Period just like IEEE 802.11 DCF/PCF. In Contention-Free Period, we have total sixteen lanes as incoming channel to the intersection and each lane has its own fixed route to pass through the intersection. Some of the routes can overlap each other; so, traffic control is needed on the overlapping spot. Let Q_i denote the number of cars queued up before entering the intersection in the ith lane ($i = 1, \cdots, 16$) and Q_i is updated at the very beginning of each time slot.

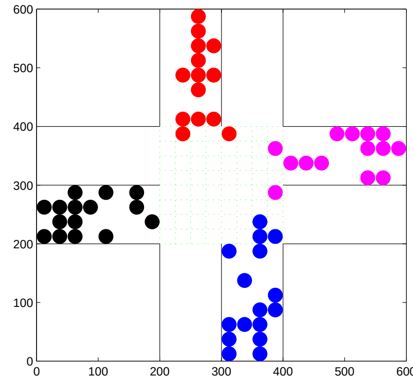

Fig. 2. Basic simulation model

When the ith route and the jth route intersect and there are two cars, one from ith lane and another one from the jth lane, going to the intersecting point, the system has to decide which car it will let go first. The decision will be made based on which one is larger between Q_i and Q_j. If there are n lanes ($x_1, \cdots, x_n$) whose routes crossing each other, the system will give priority to the xth lane, which satisfies,

$$x = \underset{i \in x_1,\dots,x_n}{\arg\max}\ Q_i \tag{1}$$

2) Contention Period (CP): If the system always give priority to the most congested lane, the car on the road which is relatively free of traffic will have to wait until its lane become the most congested. It is not fair to force the cars on the less congested lane to wait for too long just because the other lane is very busy. This motivates us to introduce contention period. In CP, if there are n cars coming to the overlapping spot at the same time, the n cars will take turns to pass through the overlapping spot no matter how many cars are queuing up in each lane. When you design a system you can make CFP longer if your main goal is to ease the congestion, or you can make CP longer if you aim for fairness. The proposed traffic control system with/without traffic light are shown in Fig. 3 and Fig. 4, respectively.

C. Simulation Results

1) Proposed traffic system vs. Traditional traffic system : Now we compare the proposed traffic system with traditional traffic system which are characterized by the existence of traffic lights. We put different weight on each lane with different Poisson expectation, i.e. λ_{North} = 1/2, λ_{South} = 1/4, λ_{East} = 1/8, λ_{West} = 1/16, so that each lane has different level of congestion. We measure the travel time for a car to pass through the intersection. One cycle (repeat interval) of the system consists 60 time slots with 30 time slots of CFP and 30 time slots of CP. The result is shown in Table I.

When the proposed system is used, the elapsed time of a single car to pass through the intersection is reduced by 34.4 % on average, which means total traffic flow become smoother.

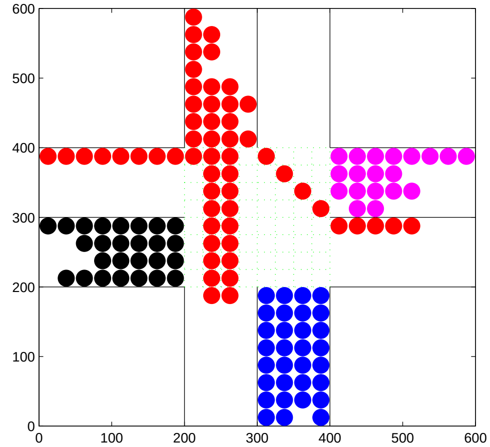

Fig. 3. Traditional traffic system with traffic lights

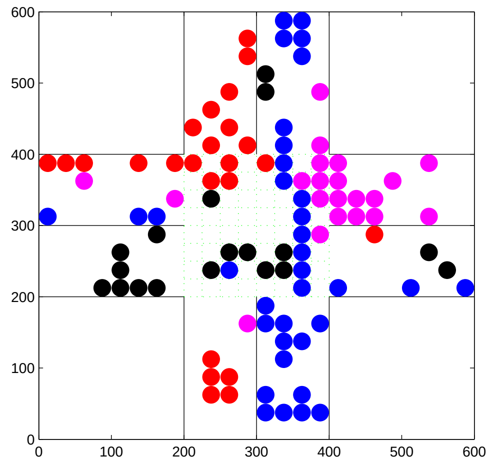

Fig. 4. Proposed traffic system without traffic lights

Another notable result is that variance between travel times of different users is significantly decreased by 90.5 %. This reduced variance means that the travel time become much more predictable even though each road has different traffic density. Fig. 5 shows this result in Gaussian distribution. We can confirm that the travel time for each user become shorter and a lot more predictable by using the proposed traffic system.

2) Traffic density and the system performance: Now we analyze the relationship between the traffic density and the system performance. We can expect by intuition that the proposed system will work more efficiently if there is less traffic on the road. For example, there are probably few cars on the road during late nights or early dawns, which means they need not wait before entering the intersection. In this simulation, we measure the travel time as a function of traffic density. As you can see in Fig. 6, the travel time in the proposed system barely increase until traffic density reaches

TABLE I
TRAVEL TIME WITH/WITHOUT TRAFFIC LIGHT

	Traditional system (with traffic light)		Proposed system (without traffic light)	
(unit: time slot)	Variance ($x10^4$)	Average ($x10^2$)	Variance ($x10^4$)	Average ($x10^2$)
Left turn	36.85	11.23	0.78	4.05
Straight	145.2	22.43	12.74	16.09
Right turn	36.44	11.20	4.53	9.30

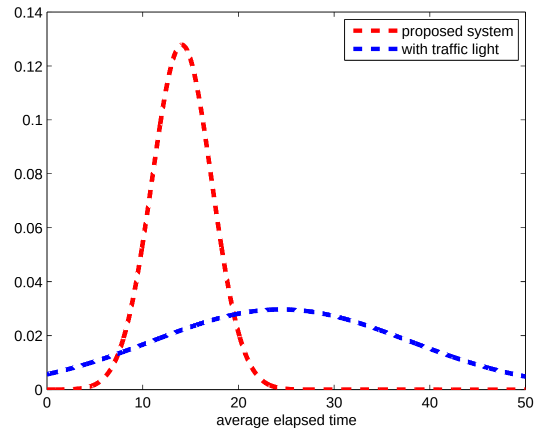

Fig. 5. Gaussian distribution model of travel time

a certain number, 0.2 cars per time slot in this graph. That means road capacity in the proposed system is able to let cars pass through the intersection without stoppage until the traffic density reaches 0.2 cars per time slot. After the traffic density of 0.2, the travel time in the proposed system starts to increase almost linearly. After applying linear estimation we find the slope of estimated line, and it is shown in Table II. Even traffic density become higher than 0.2, the slope of the travel time with respect to traffic density in the proposed system is still lower than traditional system.

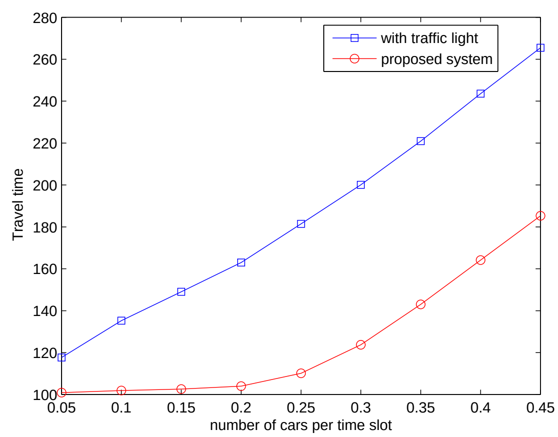

Fig. 6. Relationship between traffic density and travel time

TABLE II
SLOPE OF TRAVEL TIME ACCORDING TO TRAFFIC DENSITY

	With traffic light	Proposed system	note
Left turn	21.7	16.8	22.6% reduced
Straight	43.3	27.8	35.8% reduced

3) Efficiency vs. Fairness: The proposed algorithm is based on CFP/CP of the IEEE 802.11 DCF/PCF mechanisms. The proportion of CFP and CP can be adapted according to the degree of traffic congestion. The longer CFP the system has, the more efficient the system becomes so that it can ease the traffic congestion faster. On the other hand, the longer CP ensures fairness between users, which means the users on the less congested road do not have to suffer for the sake of overall system efficiency. In our simulation, we set a repeat interval to be 60 time slots and divide the repeat interval into CFP/CP. In Fig. 7, we plot the variance of the travel time as we increase the proportion of CFP to CP. The graph shows that the higher proportion of CFP in the repeat interval makes variance of the travel time smaller, so the travel time becomes more predictable as we expected.

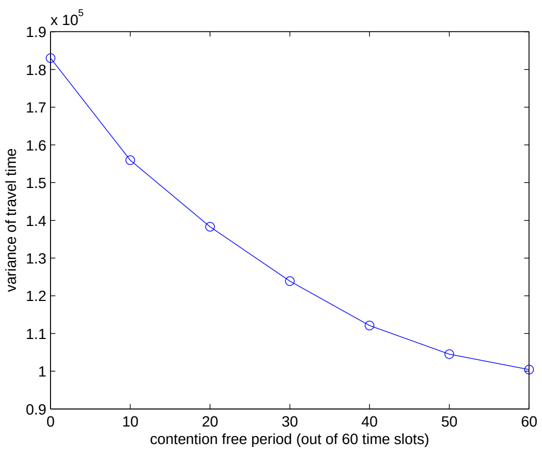

Fig. 7. Contention free period vs. variance of travel time

III. TRAFFIC CONTROL BASED ON MULTI-ARMED BANDIT ALGORITHM

A. Multi-armed bandit Policy

Multi-armed bandit (MAB) problems are a class of sequential resource allocation problems concerned with allocating one or more resources among several alternative projects. Such problems are paradigms of a fundamental conflict between making decisions that yield high current rewards, versus making decisions that sacrifice current gains with the prospect of better future rewards [6]. A policy is an algorithm that chooses the next machine to play based on the sequence of past plays and obtained rewards. Let T_i(n) be the number of times machine i has been played during the first n plays. Then the regret of a certain policy after n plays is defined by

$$\mu^* n - \mu_j \sum_{j=1}^{K} E[T_j(n)] \quad \text{where } \mu^* = \max_{1 \le i \le K} \mu_i \tag{2}$$

and E[·] denotes expectation. Thus, the regret is the expected loss due to the fact that the policy does not always play the best machine [7].

B. System Model

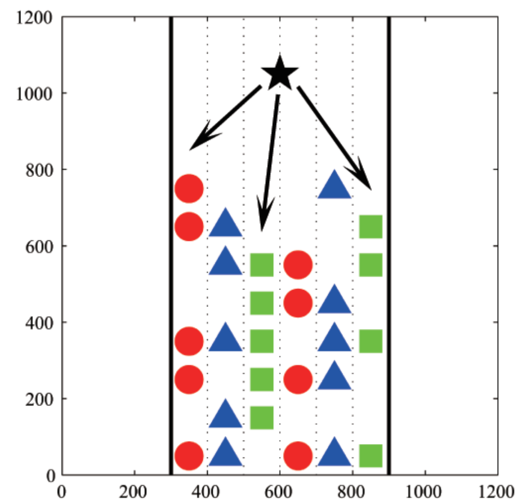

Fig. 8. Lane Selection Algorithm using MAB

We have a road with 6 lanes and each lane is randomly congested according to a given distribution as shown in Fig. 8. If a vehicle can move forward within a certain time period because there is no traffic congestion, we consider that the vehicle receives a *reward*. The vehicle chooses one lane to drive in and obtains a reward drawn i.i.d. over time from a distribution with unknown mean. Different lanes may have different reward distributions. The objective of the proposed algorithm is to find a policy that maximizes the total expected reward and to converge to the best lane while minimizing the regret.

Upper Confidence Bound (UCB) algorithm is used in this simulation as a sequential lane selection policy. The vehicle will drive in lane i that maximizes the priority index below,

$$\text{Priority}_i = \bar{x}_i + \sqrt{\frac{2 \ln n}{n_i}} \tag{3}$$

where $\bar{x}_i$ is the average reward obtained from lane i, n_i is the number of times lane i has been played so far, and n is the overall number of plays done so far.

C. Simulation Results

We have four lanes to choose from in the simulation and each lane has its own reward distribution, $R_i = \{0.1, 0.3, 0.5, 0.7\}$, where R_i means the probability that a vehicle can move forward without being stuck in traffic. The goal of this simulation is to find the best lane while minimizing the regret. In the simulation, policy UCB performs better than policy ε-GREEDY, as it is shown in Fig. 9. We can see that the lane a vehicle chooses and drives in converges to the best lane more quickly when it is played by policy UCB. The regret of policy UCB is also less than policy ε-GREEDY.

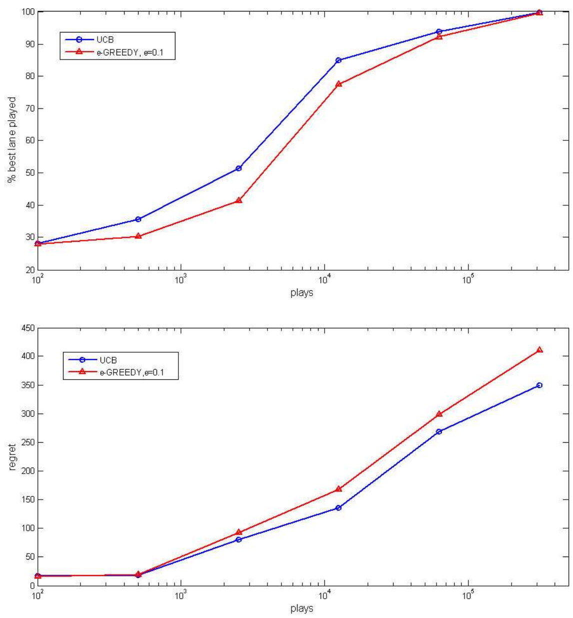

Fig. 9. Performance of different lane selection policies

IV. Conclusion

With the advent of driverless car technologies, a new intelligent transport system can be developed to make traffic system more efficient and safe with the help of vehicle-to-vehicle and vehicle-to-infrastructure communications. As part of this intelligent system, we have introduced a new traffic system model specifically designed for urban road networks. The proposed system based on wireless scheduling technique has no traffic lights at intersections, and uses IEEE 802.11 DCF/PCF mechanisms to control the traffic especially during rush hours. Each cycle of traffic control is divided into contention free period (CFP) and contention period (CP). In CFP, the system will try to clear up the most congested lane while the system will address the fairness issue in CP. We can achieve a proper balance between efficiency of the system and user fairness by using the proposed algorithm. The proposed algorithm can easily accommodate emergency traffic by giving the highest priority to emergency vehicle such as police cars, fire trucks, and ambulances. In the second proposed system, we utilize multi-armed bandit algorithm to tackle the lane selection problem when a vehicle faces exploration vs. exploitation dilemma. We apply UCB algorithm to the traffic system for maximization of the total expected reward. Other advanced policies for MAB problems can also be employed. It appears that the proposed traffic control system is promising to alleviate traffic congestion in urban road systems.

Acknowledgment

This research was supported in part by Basic Science Research Program (2010-0013397) and Mid-career Researcher Program (2010-0027155) through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology, Seoul R&BD Program (JP091007, 0423-20090051), Power Generation & Electricity Delivery of the KETEP grant funded by the Korea government Ministry of Knowledge Economy (No. 2011T100100151), the Institute of New Media & Communications (INMAC), and the BK21 program.

References

[1] R. Benenson, S. Petti, T. Fraichard, and M. Parent, ”Towards urban driverless vehicles,” *International Journal of Vehicle Autonomous Systems*, Volume 6, Number 1-2, pp.4-23, 2008

[2] S. Tsugawa, T. Yatabe, T. Hirose, and S. Matsumoto, An automobile with artificial intelligence, in Proc. IJCAI, pp. 893-895, 1979

[3] S.E. Shladover, *Preparing the Way for Vehicle-Infrastructure Integration*, California. Dept. of Transportation, 2005

[4] J.A. Misener and S.E. Shladover, ”PATH investigations in vehicle-roadside cooperation and safety: a foundation for safety and vehicle-infrastructure integration Research,” *IEEE ITSC*, 2006

[5] IEEE Std. 802.11-1999, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, Reference number ISO/IEC 8802- 11:1999(E), IEEE Std. 802.11, 1999 edition, 1999

[6] A. Mahajan and D. Teneketzis, ”MULTI-ARMED BANDIT PROBLEMS,” in Foundations and Applications of Sensor Management, pp. 121-151, Springer-Verlag, 2007

[7] P. Auer, N. Cesa-Bianchi, and P. Fischer, ”Finite time analysis of the multiarmed bandit problem,” *Machine Learning*, Volume 47, pp.235-256, 2002

 ISBN: 978-1-61208-241-7