Traffic Signal Control for Isolated Intersection Based on Coordination Game and Pareto Efficiency

Yang Zhao[1], Yiwen Liang[2], Jianming Hu[3], *Member, IEEE*, and Zuo Zhang[3], *Member, IEEE*

Abstract— Nowadays, with the rapidly increasing of the vehicles number, traffic congestion becomes a severe problem in almost all large cities. The applications of Intelligent Transportation Systems makes it possible for a signal control system to collect real-time data of an intersection, helping to improve the traffic signal control strategy and traffic efficiency. This paper proposes an algorithm based on coordination game and Pareto efficiency, by using vehicles position, speed and acceleration information, to control traffic signal for an isolated intersection in real time. The proposed algorithm is simulated on the SUMO microscopic traffic simulation platform and compared to a Webster's fixed-time plan and an actuated control algorithm to evaluate the performance in different traffic flow. The simulation results demonstrate that the algorithm based on coordination game and Pareto efficiency is more efficient than the Webster's fixed-time plan and actuated control algorithm in average queue length, average total delay and average travel time.

I. INTRODUCTION

With the development of the economy, the number of vehicles increase rapidly, causing a severe traffic congestion problem. The intersection plays a critical role in urban traffic system, so an efficient traffic control strategy can really ease the traffic pressure. However, the existing urban traffic signal control (TSC) systems cannot fully achieve the requirements of the optimal traffic control. Hence, it is necessary to design a time-variant TSC system to adjust the real-time traffic demand. The theoretical research of the TSC algorithm could date back to 1950s. From then on, TSC algorithms proposed by many researchers could be categorized as one of the following: fixed-time plan (FP), actuated method (ACT) and adaptive algorithm [1].

The FP, using the predetermined time cycle and green duration for each phase, does not consider the real-time traffic data. Therefore, it is suitable for the relatively stable and regular traffic flows. The most famous FP method is proposed by Webster in 1958. This formula based on the flow rate of each lane in an intersection is useful to find an optimal cycle and appropriate duration for green time in each phase [2]. However, because the traffic system is a dynamic system, the FP plan cannot fit all real-time conditions.

The ACT, also called real-time traffic-responsive control, came into practice with the help of the development of the sensor technologies. This method need install the vehicle detectors before the intersection stop lines. By collecting the data from the detectors, the signal control system will make different traffic control action under a certain control strategy based the collected real-time information. The ACT method, choosing to change to another phase or extend the current phase duration based on real-time traffic demand and predefined logic rules, is widely used in current traffic system. Though the ACT method was proved to perform better than FP method in most situations, it does not work when the traffic saturation is high. The extension of the current phase only depends on vehicles of the current phase, without considering the queue length other phases. Therefore, it cannot achieve the optimal strategy of the intersection.

Parameters like time, day, season, weather, and many unpredictable situations such as accidents are highly influential on traffic system. Hence, the adaptive TSC system was creadted to take into account more information to achieve more efficient signal controller. Sydney coordinated adaptive traffic system (SCATS) and Split cycle offset optimization technique (SCOOT) both belongs to this category. SCATS detects the traffic flows and headways at the stop bars. Based on information gathered in one-minute intervals, green times are reallocated to the phases of greatest need. The system benefits to Sydney over fixed time control in terms of reduced delay, reduced accidents, fuel saved and air pollution reduction are quantified [3]. And the SCOOT system adjusts the signal timings in frequent, small increments to match the latest traffic situation and is designed for general application within urban traffic control systems [4].

Although the adaptive method can fit different traffic situations better, it need more real-time information in the whole network. The limitation of traditional traffic sensors is that they are the pointed detectors that only provide instantaneous vehicle information only when the vehicle passes the detector and cannot continuously provide the position, speed and acceleration data of the vehicles [5]. Therefore, the Intelligent Transportation System (ITS) introduces wireless communications among smart vehicles and roadside infrastructure to collect continuous real-time data. Under this environment, the traffic control system can use high speed information transactions between vehicles, vehicles and infrastructure, rather than the sensor at fixed location, to get access to detailed vehicle information, such as position,

[1]Yang Zhao is with Department of Automation, Tsinghua University, 100084 Beijing, China zhaoy-18@mails.tsinghua.edu.cn

[2]Yiwen Liang is with School of Information and Computer Engineering, Northeast Forestry University, 150040 Harbin, China

[3]Jianming Hu and Zuo Zhang are with Department of Automation, Tsinghua University, 100084 Beijing, China, and also with Beijing National Research Center For Information Science And Technology, 100084, Beijing, China hujm@mail.tsinghua.edu.cn, zhangzuo@tsinghua.edu.cn

speed and acceleration.

Game theory is the study of mathematical models of strategic interactions between rational decision makers [6]. It has applications in all fields of social science, military, as well as in logic and computer science. In the field of the transportation, game theory has been applied to traffic flow optimization [7], cooperative adaptive cruise control systems [8], intersection traffic control [9] and route guidance [10]. Besides, Game theory is considered a suitable method to solve the intersection traffic signal control problem [11].

In this paper, we propose an isolated intersection traffic signal control algorithm based on coordination game. We will view each phase of the isolated interaction as the game player and use the estimated queue length as reward function. This algorithm is a cycle-free control strategy. To consider all phase reward at the same time, we will view all the rewards as a vector and choose the optimal action using Pareto efficiency. The game theory based algorithm is then simulated on the SUMO [12] microscopic traffic simulation platform and compared to the Webster's FP and ACT. To evaluate the performance of the game theory based algorithm, average queue length, average total delay and average travel time will be calculated.

The rest of this paper is organized as follows. In section II, we will review briefly for some game theory based signal controllers. In section III, we will introduce the signal control model based on coordination game and describe how to control a signalized intersection under the game theory based algorithm. In section IV, the simulation will be shown and we will discuss and summarize the results of the simulation. In section V, we will conclude the whole paper.

II. RELATED WORK

In order to minimize the queue length and to maximize the intersection efficiency, many approaches are proposed to solve the TSC problem. Game theory, the basic theory of coordinating the actions of intersection controllers, has been introduced into traffic coordinated control problem in recent years. In 1995, Bazan et al. [13] viewed the intersection as two-player games in which agents have different and incomplete knowledge and update their beliefs with time. In 2003, the theory of differential game is used to model the TSC system [14]. In this paper, dynamic traffic assignment and traffic signal control are formulated as a leader-follower differential game, in which a leader and multi-follower participates, under the open-loop information structure. After that, Villalobos et al. [15] proposed a non-cooperative game model for signal control problem and solved the model using ϵ-Nash equilibrium and Extraproximal method.

Besides the non-cooperative game model, which may suffer from the local optimal solution, Tan et al. [16] were the first to use Nash bargaining to optimize a two-phase traffic signal, which was modeled by a two-player coordination game model. In 2016, Hossam M. Abdelghaffar et al. [11] develop the Nash bargaining algorithm, which uses a cycle-free control strategy to optimize four-phase isolated signalized intersection traffic signal timings.

Moreover, Xia Xinhai et al. [17] combined the game theory and reinforcement learning to propose a traffic signal control agent interaction model. Dong et al. [18] [19] applied the game theoretic framework for the multi-intersection.

However, most of these work only consider two-phase information at one time, which may lead to the local optimal solution. Another limitation of current game theory based algorithm is that they used either non-cooperative game model or cooperative model. Our work will model the four-phase isolated intersection with non-cooperative game framework and solve it by using Pareto efficiency, which is a concept from cooperative game model, to guarantee the solution is at least Pareto Optimality.

III. SIGNAL CONTROL MODEL BASED ON COORDINATION GAME

In this section, we will describe the coordination game model and Pareto efficiency solution for two players, which is shown in section III-A, and the four-phase traffic signal model based on coordination game, as is shown in section III-B.

A. Two-Player Coordination Game and Pareto Efficiency Solution

In game theory, coordination games are a class of games with multiple pure strategy Nash equilibrium in which players choose the same or corresponding strategies [20]. When a game is a coordination game, the following inequalities hold in the payoff matrix shown in Table I. We use P_i to denote the players and use A_i, R_i to denote the actions and rewards of player 1, a_i, r_i to denote the actions and rewards of player 2. For player P_1, $R_1 > R_3$ and $R_4 > R_2$, and for player P_2, $r_1 > r_2$ and $r_4 > r_3$. This means strategy profiles (A_1, a_1) and (A_2, a_2) are pure Nash equilibrium and this can be extend to more than two players.

TABLE I
TWO PLAYERS PAYOFF MATRIX

	$P_2(a_1)$	$P_2(a_2)$
$P_1(A_1)$	(R_1, r_1)	(R_2, r_2)
$P_1(A_2)$	(R_3, r_3)	(R_4, r_4)

If $R_1 > R_4$ and $r_1 > r_4$, we say (R_1, r_1) Pareto dominates the (R_4, r_4) (of coures it also Pareto dominates the (R_2, r_2) and the (R_3, r_3)). Thus, the best action pairs is (A_1, a_1) though the game still have two pure Nash equilibrium. In this situation, we say the (A_1, a_1) is Pareto efficiency solution. Similarly, if $R_4 > R_1$ and $r_4 > r_1$, the (A_2, a_2) is the Pareto efficiency solution.

However, in most situations, we will have $R_1 > R_4$, $r_1 < r_4$ or $R_1 < R_4$, $r_1 > r_4$, which means it is hard to say which action pair is more efficient. To solve this problem, mixed strategy Nash equilibrium was proposed. For the coordination game above, a mixed Nash equilibrium is given by a probability p to play A_1 and $1-p$ to play A_2 for player P_1, and another probability q to play a_1 and $1-q$ to

play a_2 for player P_2, where p and q can be calculated as (1) and (2).

$$p = \frac{r_4 - r_3}{r_1 - r_2 - r_3 + r_4} \tag{1}$$

$$q = \frac{R_4 - R_2}{R_1 - R_2 - R_3 + R_4} \tag{2}$$

Unfortunately, this solution is not good enough. On one hand, the mixed Nash equilibrium solution is Pareto dominated by the two pure Nash equilibrium since the players will fail to coordinate. On the other hand, the solution is useful when the coordination game is also a repeat game, but in our model, it is not.

In our situation, the controller can be viewed as a traffic management who aims to reduce the vehicle queue length. That means we can choose the Pareto solution in the payoff matrix to get the largest payoff in the whole system. To get the Pareto solution, for every action pair (A, a), we define the objective function as the weighted sum of payoff of all players under this action pair, i.e:

$$W(A, a) = w_1 u_1(A, a) + w_2 u_2(A, a) \tag{3}$$

where $u_i(a_1, a_2)$ is the payoff of player i when the player 1 chooses the action a_1 and player 2 chooses the action a_2 and $w_1, w_2 \geq 0$ are predefined, determining which player is more important. The Pareto solution can be obtained by solve the following optimization problem, i.e:

$$(A^*, a^*) = \underset{(A,a)}{\arg\max} \, W(A, a) \tag{4}$$

Besides, we also do not want to let one player's benefit is much larger than the others'. We introduce the reject region bounded by two line passing through the $(0, 0)$ and we would ignore the Pareto solution in the rejcet region. The geometry interpretation of (4) is shown in Fig. 1.
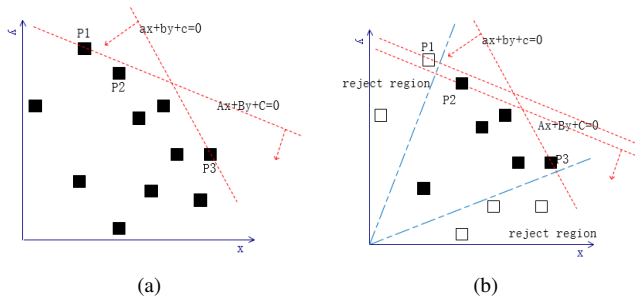

Fig. 1. The geometry interpretation for the Pareto solution. 5a shows that P1 and P3 is the different solution under different non-nagetive weights. But in 5b, the P1 is in the rejcet region bounded by the two blue line, so the real solution is the P2 if we choose (A, B) as the weights.

In our case, the weights represent our preference to choose keep action or switch action, and the reject region will balance the queue length for each phase.

B. Four-Phase Traffic Signal Model Based on Coordination Game

For four-phase isolated intersection, we use the standard NEMA phasing as shown in Fig. 2. And the yellow time for
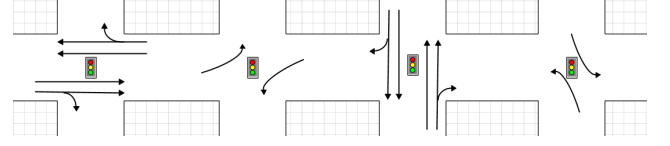

Fig. 2. Standard NEMA phasing

each phase is predefined as 3 seconds. According to game theory, we will define the game players, strategies and reward functions as follow:

- **Players:** All four phases, denoted by P_1, P_2, P_3, P_4, where P_1 represents the current green phase. In the period of signal cycle, every second in green time will have a non-cooperative game process and controller will choose the corresponding action.
- **Strategies:** Keep current phase, A_1, or change to next phase, A_2. Firstly, we must note that the phase cannot be kept all the time. To avoid this situation, we predefine the minimum green time for each phase as 5 seconds and the maximum green time for each phase as 30 seconds. Every phase in a signal cycle will at least obatins minimum green time and at most obatins maximum green time. Secondly, note that the strategy profiles $(keep, keep, keep, keep)$ and $(change, change, change, change)$ are two pure Nash equilibrium. And that is why this model is based on coordination game.
- **Payoff functions:** Estimated queue length of each phase in a few time interval after applying a specific action, denoted by $R_P(t, \Delta t)$. This measurement can be estimated if we know position, speed and acceleration of all vehicles in SUMO.

The reward function for each player (phase) is calculated as the following equation:

$$R_P(t, \Delta t) = \max_{l \in P} \{q_l(t) + A_l(t, \Delta t) - D_l(t, \Delta t)\} \tag{5}$$

where l is the corresponding lane for each player, $q_l(t)$ is the current queue length (number of vehicle) associated with corresponding lane l at time t, $A_l(t, \Delta t)$ is the estimated number of arrival vehicle in lane l between the time interval $[t, t + \Delta t]$, $D_l(t, \Delta t)$ is the estimated number of departure vehicles in lane l between the time interval $[t, t + \Delta t]$, $D_l(t, \Delta t)$.

In SUMO, if the vehicle i speed v_i^t is less than a predefined threshold speed v_{th} at time t and the distance d_i^t to the intersection is less than a predefined threshold distance d_{th}, we will view this vehicle is in queue. Thus,

$$q_l(t) = \sum_{i \in l} n_i^t \tag{6}$$

$$n_i^t = \begin{cases} 1 & \text{if } v_i^t < v_{th} \text{ and } d_i^t < d_{th} \\ 0 & \text{else} \end{cases} \tag{7}$$

For $A_l(t, \Delta t)$, we will use the kinematics formulas to

estimate as (8), where $\mathrm{sgn}(x)$ is the symbolic function.

$$A_l(t, \Delta t) = \sum_{i \in l} \mathrm{sgn}(\min\{0, \Delta t - \frac{v_i^t}{a_i^t}\}) \qquad (8)$$

For $D_l(t, \Delta t)$, the caculation could be split into three conditions. For the phase which is in green phase and hope to keep the current phase (condition c_1), we could simply use the saturation flow rate to estimate; for the phase which is in red phase and hope to switch to the green phase (condition c_2), we could not use the saturation flow rate (the green phase will just begin), so we will use kinematics formulas to estimate; for the else condition, the number of departure vehicle is zero. Hence, the formula of the $D_l(t, \Delta t)$ is shown as the (9):

$$D_l(t, \Delta t) = \begin{cases} Q_{sl}\Delta t & \text{for cond } c_1 \\ \sum_{i \in l} \mathrm{sgn}(\max\{0, \Delta L\}) & \text{for cond } c_2 \\ 0 & \text{else} \end{cases} \qquad (9)$$

where Q_{st} is the saturation flow rate for lane l, ΔL could be estimated as follow:

$$\Delta L = v_i^t(\Delta t - t^y) + \frac{1}{2}a_i^t(\Delta t - t^y)^2 - d_i^t \qquad (10)$$

where t^y is the yellow light duration.

To solve the four players coordination game, we can view the P_2, P_3, P_4 as a cooperative player P_c who wants to choose switch action, and the payoff function could be calculated by

$$R_{p_c}(t, \Delta t) = aR_{p_2}(t, \Delta t) + bR_{p_3}(t, \Delta t) + cR_{p_4}(t, \Delta t) \quad (11)$$

where $a, b, c \geq 0$ satisfy

$$a + b + c = 1 \qquad (12)$$

Hence, this problem is converted to the two-player coordination game and can be solved by the Pareto efficiency based method introduced in section III-A.

IV. SIMULATION AND RESULTS

In this section, we will describe the simulation and the results. Section IV-A introduces the test intersection. Section IV-B describes the measures of effectiveness and simulation parameters and section IV-C is the results of the experiments and our discussion.

A. Test Intersection

The simulation will be tested on an intersection with four roads comprised of three lanes each, as shown in Fig. 3.

The traffic demand origin-destination (O-D) matrix is provided in Table II, which are set in the SUMO simulation platform.
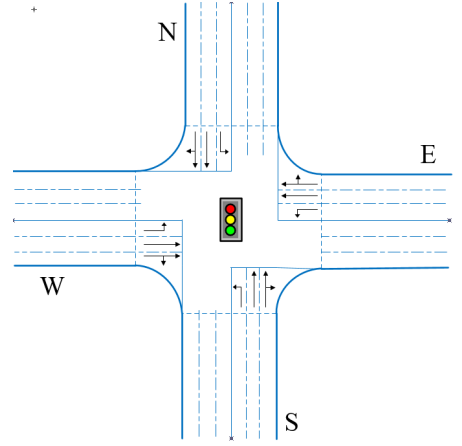

Fig. 3. Test Intersection.

TABLE II

O-D TRAFFIC DEMAND MATRIX

Zone	W	E	N	S	Total
W	-	900	225	225	1350
E	900	-	225	225	1350
N	113	113	-	450	676
S	113	113	450	-	676
Total	1123	1123	900	900	4046

B. Measures of Effectiveness and Simulation Parameters

The following measures of effectiveness (MOE) are used to evaluate the performance of our algorithm, Webster's FP and ACT method.

- **Average Queue Length (veh):** the sum of the vehicles in queue for each lane each second divied by the simulation duration and number of the lane.
- **Average Travel Time (s/veh):** the sum of all vehicles' travel time divied by the number of vehicles.
- **Average Total Delay (s/veh):** the sum of all vehicles' delay divied by the number of vehicles.

The important simulation parameters' value are set as follow:

Threshold speed $\quad s_{th} = 3.6\,(\mathrm{km/h})$

Threshold distance $\quad d_{th} = 100\,(\mathrm{m})$

Saturation flow rate $\quad Q_{sl} = 1650\,(\mathrm{veh/h/lane})$

Yellow time $\quad t^y = 3\,(\mathrm{s})$

Max vehicle speed $\quad v_{max} = 50\,(\mathrm{km/h})$

Max green time $\quad t^g_{max} = 30\,(\mathrm{s})$

Min green time $\quad t^g_{min} = 5\,(\mathrm{s})$

C. Results and Discussion

The Webster's FP and the ACT method will be used as the base line to evaluate the performance of our approach. Three scenarios are simulated: one for the orignal O-D matrix shown in Tabel II; second is for a lower demand which is the demand of 80% of the orignal O-D matrix; the last is
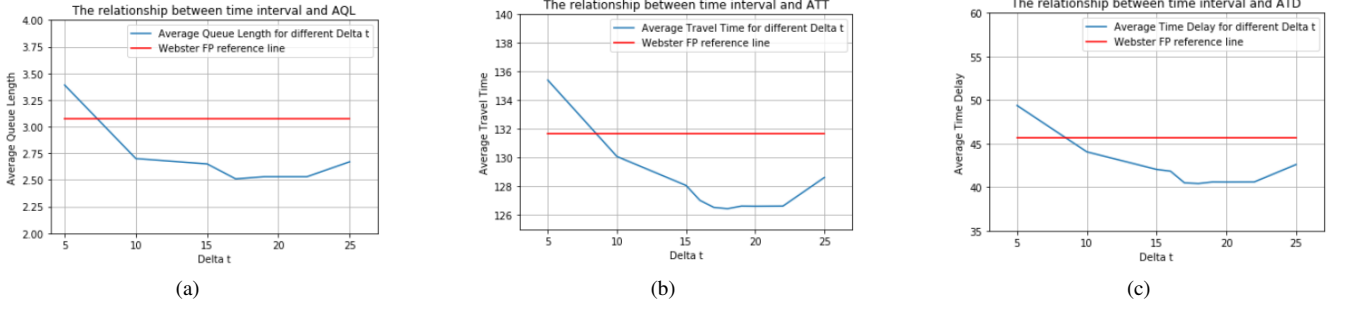
(a)   (b)   (c)

Fig. 4.   The results of simulation for different Δt and different measurements under the orignal traffic demand. The red line is a reference results obtained by the Webster's FP. 5a shows the relationship between Δt and the average queue length. 5b shows the relationship between Δt and the average travel time. 5c shows the relationship between Δt and the average total delay.

TABLE III

MOE FOR DIFFERENT TRAFFIC DEMAND AND DIFFERENT METHODS

Demand	Lower Demand			Orignal Demand			Higher Demand		
Method	Webster	ACT	GT	Webster	ACT	GT	Webster	ACT	GT
Average Queue Length (veh)	1.65	1.48	1.53	3.07	2.67	2.51	10.11	11.57	7.57
AQL Improvement (%)	+7.30	-3.37		+18.24	+5.99		+25.12	+34.57	
Average Trival Time (s/veh)	119.05	116.77	117.71	131.68	127.30	126.53	211.54	233.62	193.52
ATT Improvemen (%)t	+1.10	-0.08		+3.90	+0.60		+8.52	+17.20	
Average Total Delay (s/veh)	33.16	30.87	31.82	45.64	41.26	40.49	125.47	147.54	107.45
AQL Improvement (%)	+4.04	-0.03		+11.22	+1.86		+14.36	+27.17	

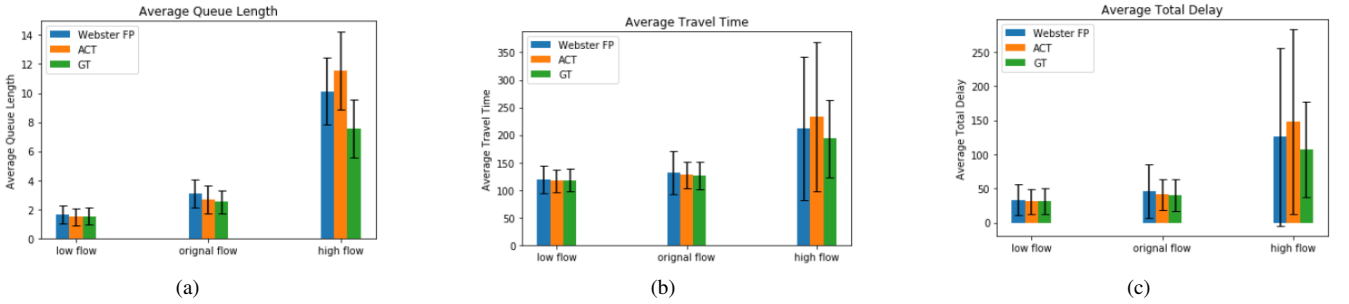(a)   (b)   (c)

Fig. 5.   The results of simulation for different methods and different measurements. 5a shows the average queue length and standard deviation for FP, ACT and GT under different traffic demand. 5b shows the average travel time and standard deviation for FP, ACT and GT under different traffic demand. 5c shows the average total delay and standard deviation for FP, ACT and GT under different traffic demand.

for a higher demand which is the demand of 120% of the orignal O-D matrix.

First, we will study the effect for the different time interval Δt under the orignal O-D matrix demand. Beacuse our yellow time for each phase is set as 3 seconds and the max green time t_{max}^{g} is set as 30 seconds, the Δt is set to vary from 5 seconds to 25 seconds. To evaluate the performance, we use the results of the Webster's FP as the reference and the experiments' results are shown as Fig. 4.

Fig. 4 shows that our approach performs better than the Webster's FP approach for almost all the Δt in average queue length, average travel time and average total delay, especially when Δt varies from 16 to 22. This results prove that our approach is not sensitive to the parameter Δt.

Next, we will give the explaination of the results shown in Fig. 4. When Δt is small, which means it is close to the yellow time, the payoff function of next phase corresponding to the switch function increase slowly for too short effective green time. When Δt is large, the estimation used kinematics formulas will have large bias beacuse all the variables we use are detected at one moment.

According to the results obtained above, we set the parameter Δt as 17 to continue our experiments. The simulation results for different methods and different traffic demand are shown in Tabel III. We also calculate the improvement of our game theory approach compared with the Webster approach and ACT method in Tabel III. And we can find that the average values for all three MOE of our approach is better except the ACT approach under lower traffic demand. However, the Fig. 5 shows the average and the standard deviation for different measurement and different traffic demand and we can find that the standard deviation of our approach is less than the other two methods, which indicates that our approach is more stable.

To show the reason why our approach works, Fig. 6 shows the average queue length for every lane by using Webster's FP, ACT method and our approach under the orignal traffic demand. It is obvious that although our approach is not the best for almost every lane, the last results show that our approach is the best. On the other hand, Fig. 6 shows that the Webster's FP works better for phase with higher traffic flow and ACT method performs better for phase with lower traffic flow, which makes the large difference between different phase under the same method. However, for our game theory based approach, the average queue length of almost all the lane is on the same level, no matter how the level of traffic flow is. Hence, we can view our approach as a balanced method to guarantee the traffic signal system is more efficient.
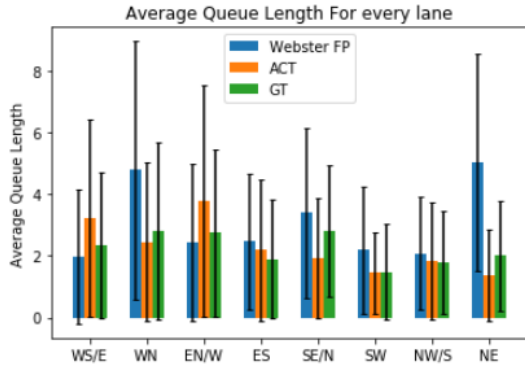

Fig. 6.   The average queue length for every lane of this test intersection for different methods under the orignal traffic demand.

The simulation results show that, for our coordination game method, not only the average value, but also the standard deviation are better than the Webster's FP and ACT method in average queue length, average travel time and average total delay. What's more, our approach is a real-time, stable, balanced and non-sensitive to parameters traffic signal control algorithm based on the discussion above.

V. CONCLUSIONS

This paper proposed a real-time coordination game based isolated intersection traffic signal control algorithm and used the SUMO simulation platform to evaluate this approach in a test intersection compared with the Webster's FP and ACT method in average queue length, average travel time and average total delay. The results show that it is possible for our approach to make the isolated intersection more efficient. Besides, we also discussed the effect of the parameter in our approach and gave the explaination of the experiments' results. Moreover, we used the simulation results to show the reason why our approach works.

In the future, we will test our approach on an real intersection and real traffic demand to demonstrate the effectiveness of our coordination game based approach. What's more, we hope to extend our algorithm to multi-intersection, which is more benefit to the urban traffic signal control.

REFERENCES

[1] D. Zhao, Y. Dai, and Z. Zhang, "Computational intelligence in urban traffic signal control: A survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 4, pp. 485–494, 2012.

[2] F. V. Webster, "Traffic signal settings," Tech. Rep., 1958.

[3] A. Sims *et al.*, "Scat the sydney co-ordinated adaptive traffic system," in *Symposium on Computer Control of Transport 1981: Preprints of Papers*. Institution of Engineers, Australia, 1981, p. 22.

[4] P. Hunt, D. Robertson, R. Bretherton, and R. Winton, "Scoot-a traffic responsive method of coordinating signals," Tech. Rep., 1981.

[5] Y. Feng, K. L. Head, S. Khoshmagham, and M. Zamanipour, "A real-time adaptive signal control in a connected vehicle environment," *Transportation Research Part C: Emerging Technologies*, vol. 55, pp. 460–473, 2015.

[6] P. Hammerstein and R. Selten, "Game theory and evolutionary biology," *Handbook of game theory with economic applications*, vol. 2, pp. 929–993, 1994.

[7] K.-H. N. Bui and J. J. Jung, "Cooperative game-theoretic approach to traffic flow optimization for multiple intersections," *Computers & Electrical Engineering*, vol. 71, pp. 1012–1024, 2018.

[8] I. H. Zohdy and H. Rakha, "Game theory algorithm for intersection-based cooperative adaptive cruise control (cacc) systems," in *2012 15th International IEEE Conference on Intelligent Transportation Systems*. IEEE, 2012, pp. 1097–1102.

[9] M. Elhenawy, A. A. Elbery, A. A. Hassan, and H. A. Rakha, "An intersection game-theory-based traffic control algorithm in a connected vehicle environment," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*. IEEE, 2015, pp. 343–347.

[10] L. Jun, "Study on game-theory-based integration model for traffic control and route guidance," *Tian Jin University*, 2003.

[11] H. M. Abdelghaffar, H. Yang, and H. A. Rakha, "Isolated traffic signal control using a game theoretic framework," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2016, pp. 1496–1501.

[12] 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. [Online]. Available: http://elib.dlr.de/80483/

[13] A. L. Bazan, "A game-theoretic approach to distributed control of traffic signals." in *ICMAS*, 1995, p. 439.

[14] L. Zhen-long, "A differential game modeling approach to dynamic traffic assignment and traffic signal control," in *SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme-System Security and Assurance (Cat. No. 03CH37483)*, vol. 1. IEEE, 2003, pp. 849–855.

[15] I. A. Villalobos, A. S. Poznyak, and A. M. Tamayo, "Urban traffic control problem: a game theory approach," *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 7154–7159, 2008.

[16] T. Linglong, Z. Xiaohua, H. Dunli, S. Yanzhang, and W. Ren, "A study of single intersection traffic signal control based on two-player cooperation game model," in *2010 WASE International Conference on Information Engineering*, vol. 2. IEEE, 2010, pp. 322–327.

[17] X. Xinhai and X. Lunhui, "Traffic signal control agent interaction model based on game theory and reinforcement learning," in *2009 International Forum on Computer Science-Technology and Applications*, vol. 1. IEEE, 2009, pp. 164–168.

[18] H. Dong and Z. Dai, "A multi intersections signal coordinate control method based on game theory," in *2011 International Conference on Electronics, Communications and Control (ICECC)*. IEEE, 2011, pp. 1232–1235.

[19] Z. Dai, H. Dong, and Q. Wang, "A multi-intersection coordinated control algorithm based on game theory and maximal flow," in *IECON 2013-39th Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2013, pp. 3258–3263.

[20] M. J. Osborne and A. Rubinstein, *A course in game theory*. MIT press, 1994.