Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Multi-Routes Algorithm using Temperature Control of Boltzmann Distribution in Q value-based Dynamic Programming

Shanqing Yu, Shingo Mabu, Manoj Kanta Mainali, Shinji Eto, Kaoru Shimada, Kotaro Hirasawa
Graduate School of Information, Production and Systems, Waseda University,
Hibikino 2-7, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0135, Japan,
Tel/Fax: +81 93 692-5261,
E-mail: iamasmilingfish@fuji.waseda.jp, mabu@aoni.waseda.jp, mkmainali@fuji.waseda.jp,
eto@aoni.waseda.jp, k.shimada@aoni.waseda.jp, hirasawa@waseda.jp

***Abstract*—In this paper, we propose a heuristic method trying to improve the efficiency of traffic systems in the global perspective, where the optimal traveling time for each Origin-Destination (OD)pair is calculated by extended Q value-based Dynamic Programming and the global optimum routes are produced by adjusting the temperature parameter in Boltzmann distribution. The key point is that the temperature parameter for each section is not identical, but constantly changing with the traffic of the section, which enables the diversified routing strategy depending on the latest traffics. In addition, the simulation results show that comparing with the Greedy strategy and constant temperature parameter strategy, the proposed method, i.e., temperature parameter control strategy of the Q value-based Dynamic Programming with Boltzmann distribution, could reduce the traffic congestion effectively and minimize the negative impact of the information update interval by adopting suitable temperature parameter control strategy.**

***Index Terms*—Q value-based Dynamic Programming, Boltzmann Distribution, Temperature Parameter, Greedy Strategy**

I. Introduction

Modern metropolises with overcrowded traffic have been long suffering from traffic problems such as traffic congestion and traffic accidents. Nowadays, Vehicle Navigation Systems are widely applied to improve the efficiency of traffic systems and solve the problem of the traffic congestion, while ensuring traffic safety, since the developed Navigation Systems not only provide static map data and optimal routes, but also inform the drivers of the real time traffic conditions on the map [1][2][3][4].

However, the traditional guiding strategy in Navigation Systems which informs every driver of the shortest path according to the current updated information may turn out to be a suboptimal strategy in some cases, since it may cause some negative behavioral phenomena like concentration and overreaction [5][6], where concentration means that the traffic volume centers on the same optimal route which consequently causes the traffic jam, and overreaction means that the vehicles on the optimal route would turn to another route after the update of the information, which results in the unexpected low traffic volume on the previous optimal route. In addition, the time delay between route generation and traffic information update also depress the efficiency of Navigation System.

In order to avoid the disadvantages of the traditional guiding strategy and improve the efficiency of traffic systems in global perspective, we have already proposed a global optimum traffic routing strategy - Q value-based Dynamic Programming with Boltzmann Distribution [7][8], where Q value-based Dynamic Programming [9] and Boltzmann distribution [10] are combined to minimize the total traveling time of all vehicles considering the traffic volumes. In this paper, we extend the original algorithm and propose a novel method in which the optimal traveling time for each Origin-Destination pair is calculated by an extended Q value-based Dynamic Programming and the global optimum routes are produced by adjusting the temperature parameters in Boltzmann distribution. The key point is that the temperature parameter in the proposed method is no longer constant nor identical for each section, which enables the diversified routing strategy depending on the latest traffics. The simulation results show that the proposed method reduces the traffic congestion effectively and minimizes the negative impact of the information update interval of the traditional methods by adopting suitable temperature parameters.

This paper is organized as follows: In the next section, the outline of Q value-based Dynamic Programming with Boltzmann Distribution is reviewed, while the details of the proposed routing strategy are described in section 3. Section 4 shows the simulations, in which the comparison among Greedy strategy, constant temperature parameter strategy and the proposed method is carried out under various traffic conditions with different information update intervals. Section 5 is devoted to conclusions.

II. Overview

A. *Q value-based Dynamic Programming*

Q value-based Dynamic Programming is used to calculate the optimal traveling time to each destination from every intersection of the road networks because it has the following two distinguished advantages over other shortest-path search algorithms [11][12][13].

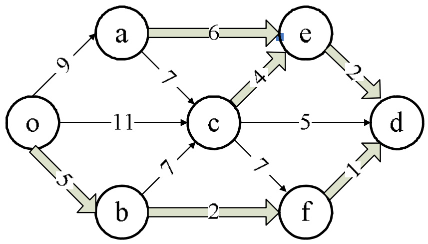


Fig. 1. A simple road network

• The conventional methods search for only the optimal route, but can't search for the second best or the third best one, while Q value-based Dynamic Programming can tell us which intersection is the best one or second best one to move in the next.

• Q value-based Dynamic Programming is less computationally intensive and easy to search for alternative optimal routes when the traveling time of the road sections changes. Most of the optimal route search algorithms would delete the original solution and recompute everything from scratch, when the traveling time of the road networks changes, while Q value-based Dynamic Programming exploits the available recent information and updates the solution with a minimum number of computations.

The Q value–$Q_d(i,j)$, which is defined as the minimum traveling time to destination d, when a vehicle bound for destination d moves to intersection j at intersection i, is calculated iteratively based on the following equations.

$$Q_d(i,j) \longleftarrow t_{ij} + \min_{k \in A(j)} Q_d(j,k), \qquad j \in A(i) \qquad (1)$$

$$Q_d(d,j) = 0, \qquad j \in A(d) \qquad (2)$$

where,
$i, j \in I$: set of suffixes of intersections
$d \in D$: set of suffixes of destinations
t_{ij} : traveling time from intersection i to intersection j
$A(i)$: set of suffixes of intersections moving directly from intersection i

B. Q value-based Dynamic Programming with Boltzmann Distribution

In this section, how to generate the optimal route from origin to destination for the drivers is explained using Q value-based Dynamic Programming with Boltzmann Distribution. Firstly, $P_d(i,j)$, i.e., the probability that the vehicle bound for destination d moves to intersection j at intersection i is calculated using Boltzmann distribution as follows.

$$P_d(i,j) = \frac{e^{-\frac{Q_d(i,j)}{\tau}}}{\sum_{j \in A(i)} e^{-\frac{Q_d(i,j)}{\tau}}}, \qquad (3)$$

where,
τ : parameter called temperature

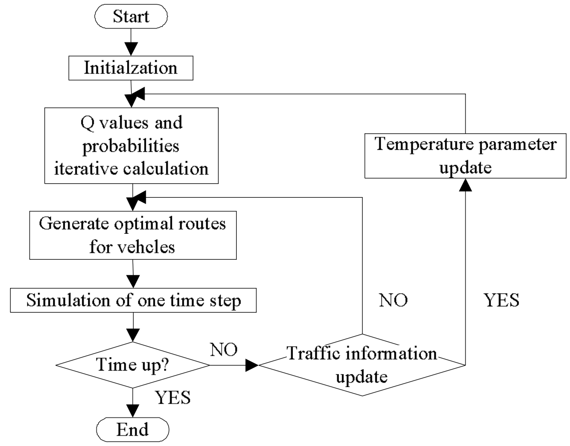


Fig. 2. Flow chart of the proposed method

Secondly, the optimal routes are generated based on the above probability. For example, take the road network in **Fig.1** describing that $Q_d(o,a) = 9$; $Q_d(o,b) = 5$; $Q_d(o,c) = 11$. Suppose $\tau = 2$, then, $P_d(o,a) = 0.1142$; $P_d(o,b) = 0.8438$; $P_d(o,c) = 0.0420$. The vehicle going to destination d from origin o has 11.42 percent probability to select (o,a), 84.38 percent to choose (o,b) and 4.20 percent to pick up (o,c).

III. Proposed Routing Strategy

Fig.2 shows the flow chart of the proposed method and the details of each module will be explained in the following subsections.

A. Extended Q value-based Dynamic Programming with Boltzmann Distribution

In the conventional Q value-based Dynamic Programming, Q values of the intersection pairs are updated using **Eq.(1)**. Therefore, the optimal route is selected by calculating $arg\ minQ_d(i,j)$, that is, greedy strategy. However, when **Eq.(3)** of the Q value-based Dynamic Programming with Boltzmann Distribution is considered, which intersection should be selected as the next intersection is determined based on a certain probability. Further more, if we consider **Eq.(3)** when updating $Q_d(i,j)$, we can obtain the following extended Q value-based Dynamic Programming with Boltzmann Distribution.

$$Q_d^{(n)}(i,j) = t_{ij} + \sum_{k \in A(j)} P_d^{(n-1)}(j,k) Q_d^{(n-1)}(j,k), \qquad (4)$$
$$d \in D,\ i \in I - d,\ j \in A(i)$$

$$P_d^{(n)}(i,j) = \frac{e^{-\frac{Q_d^{(n)}(i,j)}{\tau_{ij}}}}{\sum_{j \in A(i)} e^{-\frac{Q_d^{(n)}(i,j)}{\tau_{ij}}}}, \qquad (5)$$
$$d \in D,\ i \in I - d,\ j \in A(i)$$

$$Q_d^{(n)}(d,j) = 0, \qquad d \in D,\ j \in A(d) \tag{6}$$

$$P_d^{(n)}(d,j) = 0, \qquad j \neq d,\ d \in D,\ j \in A(d) \tag{7}$$

$$P_d^{(n)}(d,d) = 1.0, \qquad d \in D \tag{8}$$

where,
τ_{ij} : temperature parameter from intersection i to intersection j

Q values and Probabilities for all the pairs of adjacent intersections are initialized as follows.

$$Q_d^{(0)}(i,j) = 0, \qquad d \in D,\ i \in I-d-B(d),\ j \in A(i) \tag{9}$$

$$Q_d^{(0)}(i,d) = t_{id}, \qquad d \in D,\ i \in B(d) \tag{10}$$

$$Q_d^{(0)}(d,j) = 0, \qquad d \in D,\ j \in A(d) \tag{11}$$

$$P_d^{(0)}(i,j) = 0, \qquad d \in D,\ i \in I-d-B(d),\ j \in A(i) \tag{12}$$

$$P_d^{(0)}(d,j) = 0, \qquad j \neq d,\ d \in D,\ j \in A(d) \tag{13}$$

$$P_d^{(0)}(d,d) = 1.0, \qquad d \in D \tag{14}$$

where,
$B(i)$: set of suffixes of intersections moving directly to intersection i

IV. Temperature Parameter Control Strategy

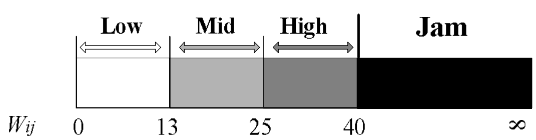


Fig. 3. Four levels of the traffic volume of the section in terms of the average waiting time

A. Traffic Volume

In this paper, we simply divide the traffic volumes of the sections into four levels, i.e., Low, Mid, High and Jam based on the average waiting times of the sections as shown in **Fig.3**. The average of the waiting time W_{ij} of section S_{ij}, i.e., the section from intersection i to intersection j is calculated by the following equation every update interval.

$$W_{ij} = \frac{\sum_{k \in V_{ij}} W_{ij}^k}{N_{ij}}, \tag{15}$$

where,
V_{ij} : set of suffixes of vehicles which go through section S_{ij} during the last interval
W_{ij}^k : the length of time that the kth vehicle waited in section S_{ij}
N_{ij} : number of vehicles that go through section S_{ij} during the last interval

B. Strategy

In Boltzmann distribution, there is a very important parameter named temperature. Basically, the probability in Boltzmann distribution is likely to be inversely related to Q values. In addition, the parameter "temperature" has an influence on the updating of Q values.

The temperature parameter of **Eq.(3)** is always constant and identical for every section in the road networks. In that sense, all the probabilities of all intersections have the same strategy. When the "temperature" is very high, Boltzmann distribution is identical to the random distribution in which each intersection has equal opportunities to be selected, on the other hand, when the "temperature" approaches 0, only the optimal route is available, just like greedy strategy.

However, in the new proposed method using **Eq.(4)** and **Eq.(5)**, each section has its own temperature parameter which enables the diversified routing strategy depending on the latest traffics. In addition, if the temperature parameter of the adjacent sections sharing the same intersection are not identical to each other, the probability for each section to be selected depends on its own temperature parameter. When the temperature parameter of a section is much lower than other sections, no matter how much Q value is, the section is less probable to be chosen. For example, in the road network shown in **Fig.1**, suppose $\tau(o,a) = \tau(o,c) = 2$ and $\tau(o,b) = 0.5$, then, $P_d(o,a) = 0.7288$; $P_d(o,b) = 0.0029$; $P_d(o,c) = 0.2681$. Even if $Q_d(o,b)$ equals 5 which is much smaller than $Q_d(o,a)$ and $Q_d(o,c)$, the vehicle going to destination d from origin o has only 0.29 percent probability to select section (o,b) and most of the vehicles may choose the second best solution (o,a).

Based on the above principle, we developed the following temperature parameter control strategy, in which the temperature parameter of the section controls its traffics, in order to avoid the vehicles moving into the sections with overcrowded traffics and consequently inducing the traffic congestion.

$$\tau_{ij} = \begin{cases} \tau_a, & If \quad W_{ij} \in Low, \\ \tau_b, & If \quad W_{ij} \in Mid, \\ \tau_c, & If \quad W_{ij} \in High, \\ \tau_d, & If \quad W_{ij} \in Jam, \end{cases} \tag{16}$$

where,
$\tau_a > \tau_b > \tau_c > \tau_d$: constant

V. Simulation

A. Traffic Simulator

Our simulation has been done in a 7×11 road network, where the length of the road sections is initialized in the range of 9 to 25 as shown in **Fig.4**. Each road section is bidirectional and has two driveways in which vehicles going to turn left preferentially choose the left one, while vehicles going to turn right and turn around prefer the right one. The signal control follows the regulation shown in Table 1, in which the time delay between neighboring intersections is 3 times steps.

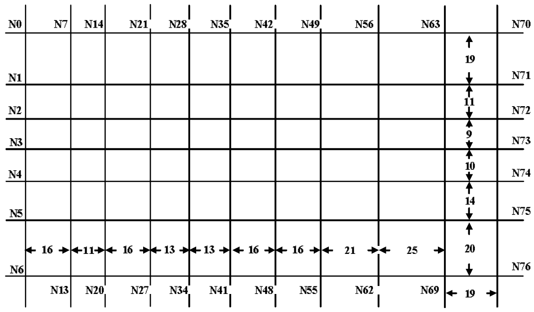


Fig. 4. Simulation road network

TABLE I
SIGNAL CONTROL OF SIMULATOR

Direction \ Signal	Red (time steps)	Yellow (time steps)	Green (time steps)
Horizontal	6	2	7
Vertical	9	2	4

Every section in the road network could be selected as origin and destination. The vehicle arrival rate to each section is assumed to follow the following Poisson distribution.

$$P_{ij}(n) = \frac{(\lambda_{ij}T)^n e^{-\lambda_{ij}T}}{n!}, \qquad n = 0, 1, ... \tag{17}$$

where,

$P_{ij}(n)$: the probability that n vehicles arrive at section S_{ij} during time T

λ_{ij}: (number of vehicles/unit time step) the rate that vehicles arrival at section S_{ij}

B. Parameter Sensitivity Analysis

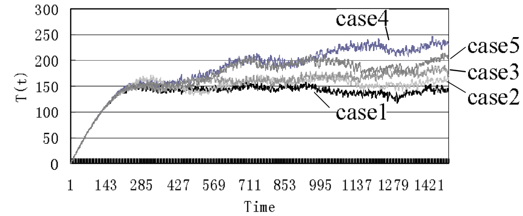
TABLE II
DIFFERENT PARAMETERS IN CONTROL STRATEGY

parameter	case1	case2	case3	case 4	case 5
τ_a	15.0	1.2	20.0	30.0	15.0
τ_b	14.5	1.0	19.0	20.0	10.0
τ_c	14.0	0.9	18.0	10.0	5.0
τ_d	13.0	0.8	17.0	5.0	1.0

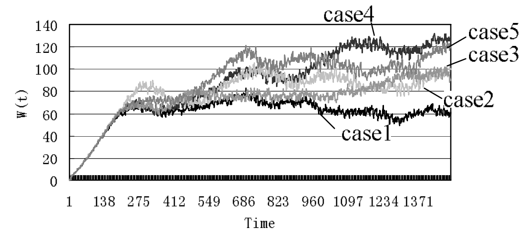
TABLE III
PARAMETER SETTING OF PARAMETER SENSITIVITY ANALYSIS

Item	Value
Maximum time step N	1500
Information update interval	60
Initial number of vehicles	1200
λ vehicle arrival rate at each section	$2\ vehicles/100\ time\ steps$

In the proposed method, temperature parameter τ_{ij} is adjusted according to the traffics from intersection i to intersection j based on **Eq.(13)**. Therefore, simulations with various τ_a, τ_b, τ_c and τ_d in **Eq.(13)** have been done to study their impact on the system performances (see Table 2). The dynamic average traveling time ($T(t)$) and dynamic average waiting time ($W(t)$) of all the current vehicles in the traffic simulator are calculated by the following equations for evaluating the system performances.



a: $T(t)$ with different parameters in control strategy



b: $W(t)$ with different parameters in control strategy

Fig. 5. Comparison of different parameters in control strategy

$$T(t) = \frac{\sum_{k \in V(t)} T_k}{N(t)}, \tag{18}$$

$$W(t) = \frac{\sum_{k \in V(t)} W_k}{N(t)}, \tag{19}$$

where,

t : time

$V(t)$: set of suffixes of vehicles traveling in the traffic simulator at time t

T_k : the length of time that the kth vehicle traveled from its start to time t

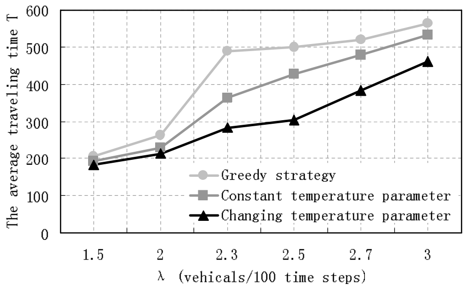
W_k : the length of time that the kth vehicle waited from its start to time t

$N(t)$: number of vehicles in the traffic simulator at time t

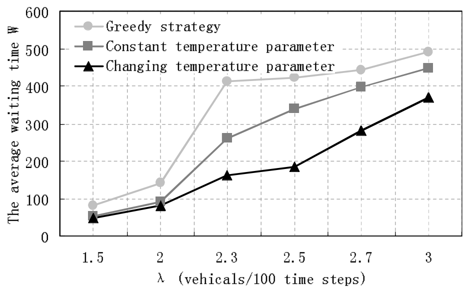
The simulation results in **Fig.5** using the date in Table 3 demonstrate the comparison among five cases with different combinations of parameters in Table 2. It reveals that using unsuitable parameters, i.e., too small in case 2, too large in case 3 and too different in case 4 and 5 would reduce the efficiency of the system. So, case 1 is used for the following simulations.

C. Experiment 1

In this experiment, Greedy strategy, constant temperature parameter strategy and the proposed method, i.e., temperature

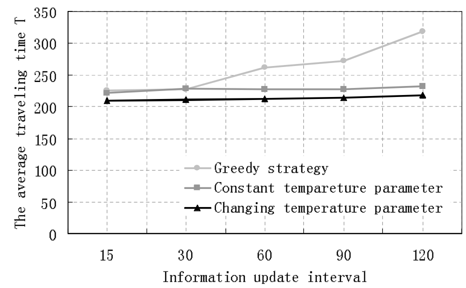


a: The average of traveling time $T(t)$ under different traffic conditions

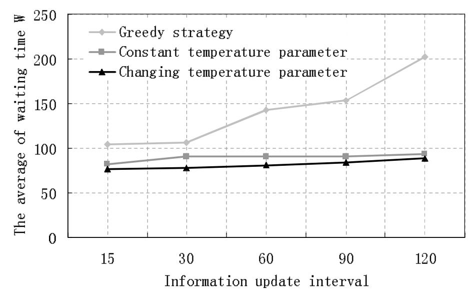


b: The average of waiting time $W(t)$ under different traffic conditions

Fig. 6. Comparison of three methods under different traffic condition



a: The average of traveling time $T(t)$ with different information update interval



b: The average of waiting time $W(t)$ with different information update interval

Fig. 7. Comparison of three methods with different information update interval

parameter control strategy of the Q value-based Dynamic Programming with Boltzmann distribution are compared under various traffic conditions. The parameters setting except λ is shown in Table3.

Fig.6 shows the average traveling time (T) and waiting time (W) of all the vehicles in the traffic simulator during the whole simulation time period under different λ. Here, each section shares the same arrival rate λ, which describes how many vehicles arrive at the section for every time unit. It is obvious from simulations that the new proposed method performs much better than Greedy strategy and constant temperature parameter strategy especially when the traffic system is suffering from overcrowd traffics.

D. Experiment 2

The traffic information including the average of the waiting times of the sections, Q values and optimal routes are updated every update interval. In this subsection, the comparison of the three methods is done under different update intervals which is shown by using T and W in **Fig.7**. The parameters setting is the same as the one shown in Table3 except the intervals of the information update. From the curves in **Fig.7**, it is easy to figure out that different intervals have a small influence on the efficiency of the proposed method, while Greedy strategy become worse when the update interval increases. Although the constant temperature parameter strategy has the same advantage, it does not perform as well as the new proposed method.

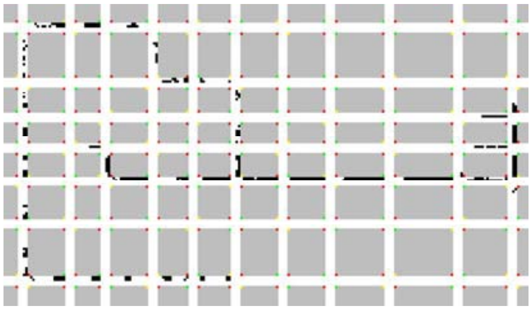
E. Experiment 3

In this subsection we developed a very special traffic condition, in which only section $(N0, N1)$ and $(N0, N7)$ have traffic inputs with the arrival rate of 100 $vehicles/100\ time\ steps$ and the destination is only intersection $N74$, to observe the negative behavioral phenomena caused by Greedy strategy and how the proposed method reduces it. The experiment is done on 2000 time steps with the number of initial vehicles being zero and the interval of the information update is 60.

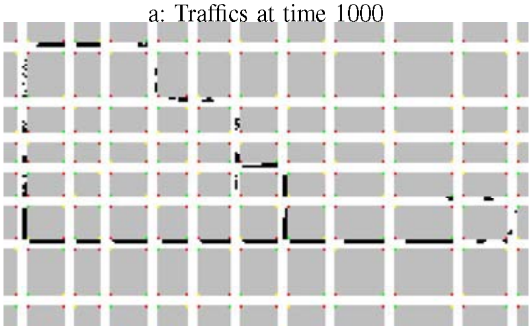
Fig.8 shows the traffics at time 1000 and 2000 when using Greedy strategy, in which each small black square denotes a vehicle. The figures show that the vehicles concentrate on the shortest pathes with unexpected low traffic volumes on other sections in the road network with Greedy strategy and consequently the traffic congestion occurs on the shortest pathes. In addition, **Fig.9** shows that our proposed method distributes the traffics to many sections effectively and save the traveling time for all the vehicles in the traffic system as shown in **Fig.10**.

VI. Conclusion

In this paper, a multi-routes algorithm is proposed by adjusting the temperature parameters of the sections depending on the latest traffics trying to improve the efficiency of the traffic systems in the global perspective. The simulation results



a: Traffics at time 1000



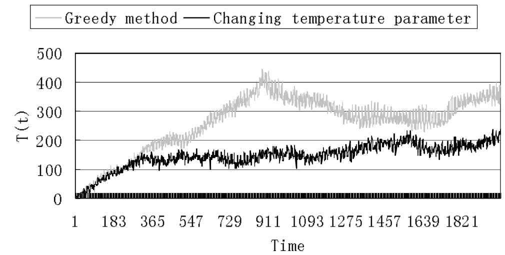
b: Traffics at time 2000

Fig. 8. Traffics by Greedy strategy

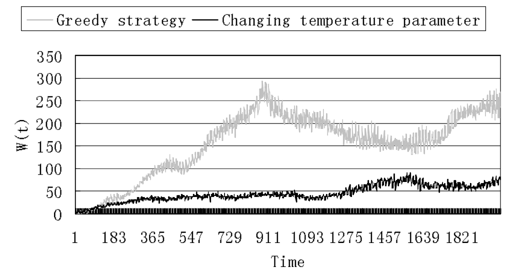
Fig. 9. Traffics by proposed method at time 1000

demonstrate its effectiveness on reducing the traffic congestion and minimizing the negative impact of the information update interval comparing with other conventional methods.

However, the temperature parameter controlling strategies still remain immature and we will study further development in the future.



a: $T(t)$ in a special traffic



b: $W(t)$ in a special traffic

Fig. 10. Comparison with Greedy strategy and Temperature parameter controlled method

REFERENCES

[1] M. A. P. Taylor, J. E. Woolley and R. Zito, "Integration of the global positioning system and geographical information systems for traffic congestion studies", Transportation Research, Vol. 8, pp. 1-6, 2000.

[2] H. Al-Deek, M. Martello, A. D. May and W. Sanders, "Potential benefits of invehicle information systems in a real life freeway corridor under recurring and incident induced congestion", In Proc. of the IEEE First VNIS Conference, Toronto, 1989.

[3] D.E. Boyce, "Route guidance systems for improving urban travel and location choices", Transportation Reseach, Vol. 22A (4), pp. 275-282,1988.

[4] B. J. Kanninen, "Intelligent transportation systems : an economic and environmental policy assessment", Transportation Reseach, Vol. 30A (1), pp. 1-10, 1996.

[5] M. Ben - Akiva, "Dynamic network models and driver information systems", Transportation Reseach, Vol. 25A (5), pp. 251-266, 1991.

[6] R. Arnott, "Does providing information to drivers reduces traffic congestion ?", Transportation Reseach, Vol. 25A (5), pp. 309-318, 1991.

[7] S. Yu, H. Wang, F. Ye, S. Mabu, K. Shimada and K. Hirasawa, "A Q value-based Dynamic Programming algorithm with Boltzmann distribution for optimizing the global traffic routing strategy", In proc. of the SICE 2008 Annual Conference, pp. 619-622, Tokyo, 2008/8.

[8] S. Yu, F. Ye, H. Wang, S. Mabu, K. Shimada, S. N. Yu, and K. Hirasawa, "A global routing strategy in dynamic traffic environments with combination of Q value-based Dynamic Programming algorithm and Boltzmann distribution", In proc. of the SICE 2008 Annual Conference, pp. 623-627, Tokyo, 2008/8.

[9] M. K. Mainali, K. Shimada, S. Mabu and K. Hirasawa, "Optimal Route Based on Dynamic Programming for Road Networks", Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.12 No.4, pp. 546-553, 2008.

[10] E. A. Guggenheim, "Boltzmann's Distribution Law", North-Holland Publishing Comp., Amsterdam. 1955.

[11] E. Dijkstra, "A Note on Two Problems in Connection with Graphs", Numerische Mathematik, Vol. 1, pp. 269-271, 1959.

[12] G. R. Jagadeesh, T. Srikanthan and K. H. Quek, "Heuristic Techniques for Accelerating Hierarchical Routing on Road Networks", IEEE Transactions on Intelligent Transportation Systems, Vol. 3, No.4, pp. 301-309, Dec. 2002.

[13] J. L. Bander and C. C. White, III, "A Heuristic Search Algorithm for Path Determination with Learning," IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans" , Vol. 28, No. 1, pp. 131-134, Jan. 1998.