Available online at www.sciencedirect.com

ScienceDirect

Applied Mathematics and Computation 189 (2007) 1–8

APPLIED MATHEMATICS AND COMPUTATION

www.elsevier.com/locate/amc

Optimal design of signal-controlled road network

Suh-Wen Chiou

Department of Information Management, National Dong Hwa University 1, Sec. 2, Da Hsueh Road, Shou-Feng, Hualien 97401, Taiwan

Abstract

An optimal design of signalized road network is considered where total travelers' delay is minimized subject to user equilibrium flow. This problem can be formulated as an optimization problem by taking the user equilibrium traffic assignment as a constraint. In this paper, a projected conjugate gradient method is presented to solve the signalized road network problem with global convergence. Numerical examples are investigated on simple grid road network. As it shows, the proposed method achieved substantially better performance than did traditional approaches when solving the signalized road network design problem with equilibrium flows.

Keywords: Conjugate gradient; Signalized road network; Equilibrium constraints; Optimization

1. Introduction

An optimal design of signalized road network is considered while the route choice of users is taken into account. This problem can be formulated as an optimization problem by taking the user equilibrium traffic assignment as a constraint. In the past decades, many researchers via the techniques of optimization have investigated this problem [1,6,9,2]. In this paper, a bilevel programming program is considered to formulate an optimal design of signal-controlled road network. In the upper level program, the performance index is defined as the sum of a weighted linear combination of rate of delay and number of stops per unit time for all traffic streams, which is evaluated by the traffic model from TRANSYT [7]. The corresponding mathematical approximations of the performance index and the average delay to a vehicle at the downstream junction in the TRANSYT model have been obtained. At the lower level the user equilibrium traffic assignment obeying Wardrop's first principle can be formulated as a minimization problem. Because the user equilibrium assignment constraint is nonlinear, which leads the signalized road network design problem to be a non-convex problem, only locally optimal solutions can therefore be found.

In this paper, a projected conjugate gradient (PCG) approach is proposed to determine the optimal signal settings and network flows with global convergence. Numerical computations are conducted on a grid network with signalized junctions where the proposed PCG method outperforms traditional methods on various

E-mail address: chiou@mail.ndhu.edu.tw

0096-3003/$ - see front matter
doi:10.1016/j.amc.2006.11.034

sets of initials and sets of demand scalars. The rest of the paper is organized as follows. In the next section, a signalized road network is formulated where users' route choice is taken into account. In Section 3, a projected conjugate gradient method is developed with global convergence. In Section 4, a grid network with signal-controlled junctions under various sets of initials is considered where numerical computations for the proposed PCG and traditional methods are conducted. Conclusions and discussions for this paper are remarked in Section 5.

2. Problem formulation

2.1. Notation

Notation used throughout this paper is summarized below.

$G(N,L)$ denotes a directed road network, where N is the set of signal controlled junctions and L is the set of links

$\Psi = (\zeta, \theta, \phi)$ denotes the set of signal setting variables, respectively for the reciprocal of cycle time, start and duration of greens, where $\theta = [\theta_{jm}]$ and $\phi = [\phi_{jm}]$ represents the vector of starts θ_{jm} and durations of green ϕ_{jm} for signal group j at junction m as proportions of common cycle time

Λ_a represents the duration of effective green for link a

g_{jm} represents the minimum green for signal group j at junction m

$\bar{c}_{jlm}$ represents the clearance time between the end of green for group j and the start of green for incompatible group l at junction m

s_a represents saturation flow rate on link a

$\Omega_m(j,l)$ represents a collection of numbers 0 and 1 for each pair of incompatible signal groups at junction m; where $\Omega_m(j,l) = 0$ if the start of green for signal group j proceeds that of l and $\Omega_m(j,l) = 1$, otherwise

D_{a} represents the rate of delay on link a

S_{a} represents the number of stops per unit time on link a

W denotes the set of OD pairs

T denotes the travel demands for OD pairs

R_{w} denotes the set of paths between OD pair w

h denotes vector of path flow

q denotes vector of link flow

δ denotes the link-path incidence matrix, where $\delta^w_{ap} = 1$ if path p between OD pair w uses link a and $\delta^w_{ap} = 0$ otherwise

c denotes the link travel times

2.2. Signalized road network problem

The signalized road network design problem can be formulated as to

$$\underset{\Psi}{\mathrm{Min}} \quad Z = Z_0(\Psi, q^*(\Psi)) = \sum_{a \in L} D_a W^1_{aD} + S_a W^1_{aS} \tag{1}$$

$$\text{subject to} \quad \zeta_{\min} \leqslant \zeta \leqslant \zeta_{\max}, \tag{2}$$

$$g_{jm}\zeta \leqslant \phi_{jm} \leqslant 1 \quad \forall j, m, \tag{3}$$

$$q_a \leqslant s_a \Lambda_a \quad \forall a \in L, \tag{4}$$

$$\theta_{jm} + \phi_{jm} + \bar{c}_{jlm}\zeta \leqslant \theta_{lm} + \Omega_m(j,l), \quad j \neq l, \ \forall j, l, m, \tag{5}$$

where W^1_{aD} and W^1_{aS} are respectively link-specific weighting factors for the rate of delay and the number of stops per unit time used in TRANSYT. The first constraint is for the common cycle time and the constraints (3)–(5) are for the green phase, link capacity and clearance time at each junction. Also the equilibrium flows $q^*(\Psi)$ are found by solving the following traffic assignment problem.

$$\underset{q}{\text{Min}} \quad \sum_{a\in L}\int_0^{q_a} c_a(x,\Psi)\mathrm{d}x \tag{6}$$

$$\text{subject to} \quad \sum_{p\in R_w} h_p^w = T_w \quad \forall w\in W, \tag{7}$$

$$q_a = \sum_{w\in W}\sum_{p\in R_w} h_p^w \delta_{ap}^w \quad \forall a\in L, \tag{8}$$

$$h_p^w \geqslant 0 \quad \forall p\in R_w,\ \ w\in W. \tag{9}$$

3. A solution method for signalized road network design problem

In this section, an effective search for solving problems (1)–(9) is developed, for which a search direction of descent is generated and a new iterate is created. The search process will be terminated at a KKT point or a new search direction can be generated. In the following, a projected conjugate gradient method is proposed to obtain a descent search direction.

3.1. A projected conjugate gradient method

Lemma 1 (Fletcher & Reeves' conjugate gradient method). *Consider a continuously differentiable function $Z:\Re^n\to\Re$ to generate a sequence of iterates $\{x_k\}$, according to*

$$x_{k+1} = x_k + \alpha_{k+1}d_{k+1}, \tag{10}$$

where d_{k+1} is a search direction and α_{k+1} is the step length which minimizes Z along d_{k+1} from x_k. Let $d_1 = -\nabla Z(x_0)$ and the superscript t *denote matrix transpose operation, d_{k+1} is decided by*

$$d_{k+1} = -\nabla Z(x_k) + \frac{\nabla Z^t(x_k)\nabla Z(x_k)}{\nabla Z^t(x_{k-1})\nabla Z(x_{k-1})}d_k, \quad k=1,2,3,\ldots \tag{11}$$

Then for the sequence of points $\{x_k\}$ generated by the conjugate gradient method

$$Z(x_k) > Z(x_{k+1}), \quad k=0,1,2,3,\ldots, \tag{12}$$

whenever $\nabla Z(x_k)\neq 0$.

The direction generated by (11) for an unconstrained nonlinear problem is a descent direction which strictly decreases the objective function value provided that the corresponding gradient value is not zero. Suppose that the first-order partial derivatives for the objective function in problems (1)–(9) evaluated at (Ψ_0, q_0) exist, which can be expressed as

$$\nabla Z(\Psi_0) = \nabla_\Psi Z_0(\Psi_0, q_0) + \nabla_q Z_0(\Psi_0, q_0)\nabla q(\Psi_0), \tag{13}$$

where the first-order derivatives with respect to signal settings and flows are derived from Chiou [3] and the second item are from the sensitivity analysis for network flows in Patriksson [5]. Let A denote the coefficient matrix and B the constant vector in constrains (2)–(5) the problems (1)–(9) can be rewritten as

$$\underset{\Psi}{\text{Min}} \quad Z = Z_1(\Psi) \tag{14}$$

$$\text{subject to} \quad A\Psi^{\mathrm{t}} \leqslant B. \tag{15}$$

In the followings, we apply Fletcher & Reeves' conjugate gradient method to a linear constraint set as given in (14) and (15) by introducing a matrix in projecting the gradient of the objective function onto a null space of active constraints as in (2)–(5) with equality in order to effectively search for implementable points.

Theorem 1 (*Projected conjugate gradient* (*PCG*) *method*). *Consider the problem in* (14) *and* (15) *a sequence of feasible iterates $\{\Psi_k\}$ can be generated according to*

$$\Psi_{k+1} = \Psi_k + \alpha_{k+1}H_{k+1}d_{k+1}, \tag{16}$$

where d_{k+1} is the conjugate gradient direction determined by (11) and α_{k+1} is the step length minimizing Z_1 along d_{k+1} for which Ψ_{k+1} is within the feasible region defined by (2)–(5). Suppose that M_{k+1} has full rank at Ψ_k, which is the gradient of active constraints with equalities in (2)–(5) and the projection matrix H_{k+1} is of the following form:

$$H_{k+1} = I - M^t_{k+1}(M_{k+1}M^t_{k+1})^{-1}M_{k+1}. \tag{17}$$

A modified search direction s_{k+1} can be determined in the following form:

$$s_{k+1} = H_{k+1}d_{k+1}. \tag{18}$$

Then the sequence of feasible points $\{\Psi_k\}$ generated by the projected conjugate gradient method monotonically decreases the performance value,

$$Z_1(\Psi_k) > Z_1(\Psi_{k+1}), \quad k = 1, 2, 3, \ldots, \tag{19}$$

whenever $H_{k+1}\nabla Z_1(\Psi_k) \neq 0$ and $\nabla Z_1(\Psi_k)$ is from (13).

Proof. Following the results of Lemma 1, we have

$$\nabla Z_1^{\mathrm{t}}(\Psi_k)d_{k+1} = -\nabla Z_1^{\mathrm{t}}(\Psi_k)\nabla Z_1(\Psi_k) < 0, \quad k = 1, 2, 3, \ldots. \tag{20}$$

Multiply Eq. (20) by projection matrix H_{k+1}, it becomes

$$\begin{aligned}\nabla Z_1^{\mathrm{t}}(\Psi_k)H_{k+1}d_{k+1} &= -\nabla Z_1^{\mathrm{t}}(\Psi_k)H_{k+1}\nabla Z_1(\Psi_k)\\ &= -\nabla Z_1^{\mathrm{t}}(\Psi_k)H^{\mathrm{t}}_{k+1}H_{k+1}\nabla Z_1(\Psi_k)\\ &= -\|H_{k+1}\nabla Z_1(\Psi_k)\|^2 < 0.\end{aligned} \tag{21}$$

Thus for sufficiently small β, $\beta > 0$, we have

$$Z_1(\Psi_k) > Z_1(\Psi_k + \beta s_{k+1}). \tag{22}$$

Because by definition α_{k+1} is the step length which minimize Z_1 along s_{k+1} from Ψ_k, it implies

$$Z_1(\Psi_k) > Z_1(\Psi_k + \beta s_{k+1}) \geqslant Z_1(\Psi_k + \alpha_{k+1}s_{k+1}) = Z_1(\Psi_{k+1}), \tag{23}$$

which completes this proof. □

Theorem 2 (*PCG method as $H_{k+1}\nabla Z_1(\Psi_k) = 0$*). *In Theorem 1, when $H_{k+1}\nabla Z_1(\Psi_k) = 0$, if all the Lagrange multipliers corresponding to the active constraint gradients with equalities in (2)–(5) are positive or zeros, it implies the current Ψ_k is a KKT point. Otherwise choose one negative Lagrange multiplier, say μ_j, and construct a new $\widehat{M}_{k+1}$ of the active constraint gradients by deleting the jth row of M_{k+1}, which corresponds to the negative component μ_j, and make the projection matrix of the following form*:

$$H_{k+1} = I - \widehat{M}^t_{k+1}(\widehat{M}_{k+1}\widehat{M}^t_{k+1})^{-1}\widehat{M}_{k+1}. \tag{24}$$

The search direction then is determined by (18) and the results of Theorem 1 hold.

Proof. Let μ_j be a negative component of the Lagrange multiplier and H_{k+1} defined in (17), we show $H_{k+1}\nabla Z_1(\Psi_k) \neq 0$. By contradiction, suppose

$$H_{k+1}\nabla Z_1(\Psi_k) = 0 \tag{25}$$

and let $w_{k+1} = -(\widehat{M}_{k+1}\widehat{M}^{\mathrm{t}}_{k+1})^{-1}\widehat{M}_{k+1}\nabla Z_1(\Psi_k)$, then (25) can be rewritten as

$$0 = \nabla Z_1(\Psi_k) + \widehat{M}^{\mathrm{t}}_{k+1}w_{k+1}. \tag{26}$$

For any $\mu_j < 0$, there exists a corresponding jth row, r_j, of the active constraint in (2)–(5) and $\hat{w}_{k+1}$ such that

$$0 = H_{k+1}\nabla Z_1(\Psi_k) = \nabla Z_1(\Psi_k) + \widehat{M}^{\mathrm{t}}_{k+1}\hat{w}_{k+1} + \mu_j r^{\mathrm{t}}_j. \tag{27}$$

We subtract (27) from (26) and it follows:

$$0 = \widehat{M}_{k+1}^{\mathrm{t}}(w_{k+1} - \hat{w}_{k+1}) - \mu_j r_j^{\mathrm{t}} \tag{28}$$

since $\mu_j \neq 0$ which contradicts the assumption that $\widehat{M}_{k+1}$ has full rank. Thus $H_{k+1}\nabla Z_1(\Psi_k) \neq 0$. □

Corollary 1 (Stopping condition). *If Ψ_k is a KKT point for problem in (14) and (15) then the search process may stop; otherwise a new descent direction at Ψ_k can be generated according to Theorems 1 and 2.*

3.2. PCG solution scheme

Consider the signalized road network optimal design problem in (1)–(9), a PCG solution scheme is given below:

Step 1: Start with Ψ_k, set index $k = 0$.
Step 2: Solve a traffic assignment problem with signal settings Ψ_k, find the first-order derivatives by (13).
Step 3: Use the projected conjugate gradient method to determine a search direction by (18). Go to Step 4.
Step 4: If $H_{k+1}\nabla Z_1(\Psi_k) \neq 0$, find a new Ψ_{k+1} in (16) and let $k \leftarrow k+1$. Go to Step 2. If $H_{k+1}\nabla Z_1(\Psi_k) = 0$ and all the Lagrange multipliers corresponding to the active constraint gradients are non-negative, Ψ_k is the KKT point and stop. Otherwise, find the most negative Lagrange multiplier and cancel the corresponding constraint and find a new projection matrix and go to Step 3.

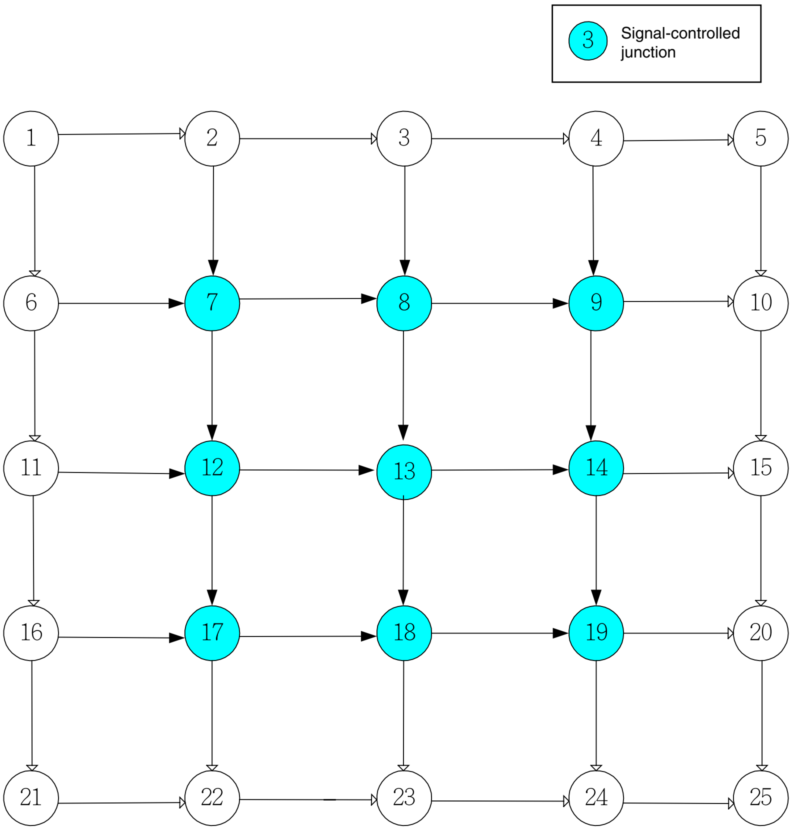

Fig. 1. 25-node network.

4. Numerical example and computational comparisons

In this section, numerical experiments are conducted twofold. Firstly, numerical computations are carried out for showing the effectiveness and robustness of the proposed PCG as compared to those of traditional methods, e.g. LCA [4] and SAB [8] in a signal-controlled grid network with distinct sets of initials. Secondly, numerical comparisons are made continuously with various congestions by monotonically increasing travel demand scalars.

4.1. Grid network with signal-controlled junctions

A 5 * 5 grid-size network shown in Fig. 1 is considered for illustrating the effectiveness of the proposed PCG method in optimal design of signalized network where 4-pair OD trips and 9 signalized junctions are taken into account. Trip rates are set at 100 veh/h and link capacity is set 1800 veh/h. Numerical results with two sets of distinct initials are summarized in Table 1. As it seen in Table 1, the proposed PCG method achieved near global optimum of 40 veh within 2.6% of SO which is short for system optimization. The proposed PCG method significantly outperforms traditional method LCA in performance index values (PI) approximately by 39% and 31% and by 11% and 9% over the SAB method. Regarding the computational times taken by the proposed PCG method as it obviously seen from Table 1, it requires less computational efforts either in CPU times or in solving the corresponding traffic assignments as it does for LCA or SAB.

4.2. Grid signalized network with increasing congestions

The second investigation is to test robustness of the proposed PCG method on increasingly congested grid signalized networks when compared to conventional methods. The increasing traffic congestion is caused by continuously increasing scalars to base travel demands. Computational results for the proposed PCG method are summarized in Table 2, where results are expressed in the relative difference percentage to SO at two sets of

Table 1
Computational results for 5 * 5 grid network

Variable/algorithm	First set of signal settings				Second set of signal settings			
	LCA	SAB	PCG	SO	LCA	SAB	PCG	SO
ϕ_{11}/ζ	45	41	57	55	55	39	48	58
ϕ_{21}/ζ	52	55	35	50	60	60	56	42
ϕ_{12}/ζ	37	66	52	55	65	40	44	50
ϕ_{22}/ζ	60	30	40	50	50	59	60	50
ϕ_{13}/ζ	37	35	55	49	50	48	49	45
ϕ_{23}/ζ	60	61	37	56	65	51	55	55
ϕ_{14}/ζ	44	54	56	60	58	56	60	56
ϕ_{24}/ζ	53	42	36	45	57	43	44	44
ϕ_{15}/ζ	40	41	49	54	55	46	52	54
ϕ_{25}/ζ	57	55	43	51	60	53	52	46
ϕ_{16}/ζ	55	53	48	52	60	56	52	60
ϕ_{26}/ζ	42	43	44	53	55	43	52	40
ϕ_{17}/ζ	53	48	44	49	60	49	60	55
ϕ_{27}/ζ	44	48	48	56	55	50	44	45
ϕ_{18}/ζ	56	56	47	53	62	50	54	50
ϕ_{28}/ζ	41	40	45	52	53	49	50	50
ϕ_{19}/ζ	60	48	49	49	55	51	49	49
ϕ_{29}/ζ	37	48	43	56	60	48	55	51
$1/\zeta$	107	106	102	115	125	109	114	110
#	108	32	16	21	128	35	14	24
PI (veh)	65	45	40	39	58	44	40	39
CPU time (s)	15.2	7.6	5.5	6.1	17.6	7.2	5.1	5.8

ϕ_{jm}/ζ denotes the duration of greens for signal group j at junction m measured in s, $1/\zeta$ denotes the common cycle time measured in s, and # denotes the number of equilibrium assignment problems solved.

Table 2
Computational results for 5 ∗ 5 grid network with congestion

Scalar/algorithm	First set of signal settings			Second set of signal settings		
	LCA	SAB	PCG	LCA	SAB	PCG
1.0	66.7	15.4	2.6	48.7	12.8	2.6
1.1	68.1	14.8	2.4	61.7	12.1	2.3
1.2	56.9	16.1	2.1	49.8	15.9	2.2
1.3	61.2	15.8	2.0	50.1	14.7	2.3
1.4	60.5	16.3	1.9	52.9	17.6	2.1
1.5	73.1	18.8	1.9	53.9	15.6	1.9
1.6	71.8	14.5	2.0	56.6	16.4	2.0
1.7	74.1	13.5	2.2	54.4	19.5	2.1
1.8	70.9	15.4	1.9	55.9	20.3	1.9
1.9	75.2	17.8	2.0	61.2	19.8	2.0
2.0	72.5	21.3	2.1	70.3	21.6	2.1
2.1	75.4	22.7	2.2	72.3	22.8	2.2
2.2	76.1	21.7	1.8	66.9	20.7	1.8
2.3	77.2	20.4	1.9	68.4	19.8	1.9
2.4	75.6	19.8	1.9	73.1	18.0	1.9
2.5	77.1	21.9	1.9	74.6	17.3	1.9
2.6	74.5	24.6	2.0	77.5	21.8	2.0
2.7	76.3	23.7	1.9	71.3	22.5	1.9
2.8	77.4	22.6	1.9	69.1	21.4	1.9
2.9	75.2	20.5	2.0	71.5	24.3	2.0

distinct initials. As it shown in Table 2, the proposed PCG method achieved the system optimum within 3% by consistently yielding the least relative difference percentages at 20 sets of travel demands. On the other hand, the conventional methods, like LCA and SAB, achieved the system optimum with much higher values of the relative difference percentages than those did the proposed PCG method. Furthermore, taking the LCA method for example, the values of the relative difference percentages were larger as traffic congestion became severe and the values on average were as high as 72 and 63 for two sets of initials. Consider the performance of SAB, it achieved the system optimum with 19%, which seems relatively better than those did LCA but still far less robust than those did the proposed PCG method.

The implementations for carrying out the computational efforts on LCA, SAB and PCG methods have been conducted on SUN SPARC Ultra II workstation under operating system Unix SunOS 5.5.1 using C++ compiler. The stopping criterion for these solutions is set when the relative difference in the performance index value between the consecutive iterations less than 0.05%.

5. Conclusions and discussions

In this paper, we presented a newly projected conjugate projection method (PCG) with global convergence to effectively solve the signalized road network problem. A 5 ∗ 5 grid road network with signal-controlled junctions is used to numerically demonstrate the robustness and superiority of the proposed method to conventional approaches in solving optimal design of signalized road network. As it reported from numerical comparisons under traffic congestion at various sets of initials, the proposed PCG method consistently outperformed other conventional methods with obvious significance and less computational efforts have been taken.

It is envisaged to test the proposed PCG method on a wide range of general road signalized networks. Further investigations will be made for the signalized road network design problem with link capacity expansions and elastic travel demands.

Acknowledgement

The author is grateful for support from Taiwan National Science Council via grant NSC 95-2416-H-259-014.

References

[1] M. Abdulaal, L.J. LeBlanc, Continuous equilibrium network design models, Transportation Research B 13 (1979) 19–32.
[2] H. Ceylan, M.G.H. Bell, Traffic signal timing optimisation based on genetic algorithm approach, including drivers' routing, Transportation Research B 38 (2004) 329–342.
[3] S-W. Chiou, TRANSYT derivatives for area traffic control optimization with network equilibrium flows, Transportation Research B 37 (2003) 263–290.
[4] B.G. Heydecker, T.K. Khoo, The equilibrium network design problem, in: Proceedings of AIRO'90 Conference on Models and Methods for Decision Support, Sorrento, 1990, pp. 587–602.
[5] M. Patriksson, Sensitivity analysis of traffic equilibria, Transportation Science 38 (2004) 258–281.
[6] S. Suh, T.J. Kim, Solving nonlinear bilevel programming models of the equilibrium network design problem: a comparative review, Annals of Operations Research 34 (1992) 203–218.
[7] R.A. Vincent, A.I. Mitchell, D.I. Robertson, User guide to TRANSYT, version 8. TRRL Report, LR888, Transport and Road Research Laboratory, Crowthorne, 1980.
[8] H. Yang, S. Yagar, Traffic assignment and signal control in saturated road networks, Transportation Research A 29 (1995) 125–139.
[9] H. Yang, M.G.H. Bell, Models and algorithms for road network design: a review and some new developments, Transport Reviews 18 (1998) 257–278.