Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Information Sciences 177 (2007) 4109–4121
INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL
www.elsevier.com/locate/ins

Comparative tests of solution methods for signal-controlled road networks

Suh-Wen Chiou *

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

Received 23 February 2006; received in revised form 16 March 2007; accepted 17 March 2007

Abstract

In this paper, we consider a signal-controlled road network with link capacity expansions. This network design problem can be formulated as a constrained optimization subject to equilibrium flows. A projected Quasi-Newton method is proposed to find good local optimal solutions. Numerical calculations are conducted using a real data road network and large-scale grid networks.

Keywords: Optimization; Signal-controlled road network; Link capacity expansions; Projected Quasi-Newton method

1. Introduction

In this paper, we consider a signal-controlled road network design problem (SIGNDP) where the signal settings and link capacity expansions are simultaneously determined. A simple heuristic was proposed by [18] for a network design problem with link capacity expansions. Friesz, Tobin, Cho and Mehta [11] were the first ones who presented gradient-based methods for the network design problem. In particular, Meng, Yang and Bell [15] developed a bilevel program for the network design problem and proposed an augmented Lagrangian method. Recently, Gao, Sun and Shan [12] constructed a bilevel model for a transit system and employed the sensitivity analysis to solve a continuous network design problem. Ban, Liu, Ferris and Ran [5] furthermore formulated a mathematical program with complementarity constraints for a general network design problem and conducted numerical computations. Utilizing a fuzzy control approach, Chou and Teng [9] analyzed an urban road network with signal settings. Ceylan and Bell [7] also proposed a genetic algorithm (GA) to deal with an area traffic control problem. Despite such contributions, however, a comparative analysis of solution algorithms for the SIGNDP does not seem to have been fully tested on large-scale networks.

A network design problem can be formulated as an optimization program with equilibrium constraints (OPEC), which has been discussed by Shimizu, Ishizuka and Bard [17]. Applications of the OPEC to real cases

* Tel.: +886 3 8633108; fax: +886 3 8633100.
E-mail address: chiou@mail.ndhu.edu.tw

0020-0255/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2007.03.022

have been widely investigated (see Anandalingam and Friesz [4], Bard [6], Amouzegar and Moshirvaziri [3] and Lu, Shi and Zhang [13]). In this paper, the SIGNDP can be formulated as an OPEC, where the signal timing plan can be defined by the common cycle time, the start and duration of greens. A novel algorithm called the projected Quasi-Newton method is proposed and numerical calculations are conducted using a real data road network and large-scale grid networks. The rest of the paper is organized as follows. In the following section, we introduce required notation. In Section 3, a novel algorithm called the projected Quasi-Newton method is presented. A near global optimum for the SIGNDP can be easily identified when the PARTAN search technique is furthermore utilized. In Section 4, numerical computations are conducted using a real data network and large-scale grid networks. Conclusions and discussions for this paper are presented in Section 5.

2. Problem formulation

The SIGNDP can be formulated as an OPEC in the following way.

2.1. Notation

$G(N, L)$ represents a directed road network, where N is the set of nodes and L is the set of links
y represents a vector of link capacity expansions
u represents a vector of upper bounds for link capacity expansions
γ represents a conversion factor from investment cost to travel delays
$G(y)$ represents a vector of investment costs
$\Psi = (\zeta, \theta, \phi)$ represents a set of signal setting variables, respectively, for the reciprocal of cycle time, start and duration of greens, where $\theta = [\theta_{jm}]$ and $\phi = [\phi_{jm}]$, respectively, represent 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 represents a vector of durations of effective greens
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 signal group j and the start of green for incompatible signal group l at junction m
s_a denotes the saturation flow on link a, $\forall a \in L$
$\Omega_m(j, l)$ 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, $\forall a \in L$
S_a represents the number of stops per unit time on link a, $\forall a \in L$
W denotes a set of origin-destination pairs
T denotes the travel demands for origin-destination pairs
$R = [R_w]$ denotes set of paths between origin-destination pair w, $\forall w \in W$
f denotes path flows
q denotes link flows
δ denotes a link-path incidence matrix
Δ denotes an origin-destination-path incidence matrix
c denotes link flow travel times
C denotes path flow travel times
π denotes minimum travel times

2.2. A user equilibrium traffic assignment

A user equilibrium traffic assignment problem can be solved by the following variational inequality. Determine values q such that

$$c^{\mathrm{t}}(q)(z - q) \geqslant 0 \tag{1}$$

for all $z \in K = \{q : q = \delta f,\ T = \Delta f,\ f \geqslant 0\}$ where the superscript t denotes matrix transpose.

2.3. A parametric variational inequality formulation

For a user equilibrium traffic assignment with link capacity expansions and signal settings, a parametric variational inequality for (1) can be formulated as follows. Determine values $q(y,\Psi)$ such that

$$c^{\mathrm{t}}(q,y,\Psi)(z-q(y,\Psi)) \geqslant 0 \tag{2}$$

for all $z \in K(y,\Psi) = \{q(y,\Psi) : q(y,\Psi) = \delta f(y,\Psi), \quad T = \Delta f(y,\Psi), \quad f(y,\Psi) \geqslant 0\}$

2.4. Sensitivity analysis

Following Patriksson [16], the sensitivity analysis of (2) can be carried out in the following way. Let the changes in link or path flow be denoted by dq or df and the changes in path flow travel time be denoted by $\mathrm{d}C_{\mathrm{p}}$. Introduce

$$\widetilde{K} = \{\mathrm{d}q : \exists \mathrm{d}f \text{ such that } \mathrm{d}q = \delta\mathrm{d}f, \quad \Delta\mathrm{d}f = 0, \text{ and } \mathrm{d}f \in K_0\}$$

where

$$K_0 = \left\{ \mathrm{d}f : \begin{array}{lllll} \text{(i) } \mathrm{d}f_{\mathrm{p}} \quad \text{free}, & \text{if } f_{\mathrm{p}} > 0 & & & \\ \text{(ii) } \mathrm{d}f_{\mathrm{p}} = 0, & \text{if } C_{\mathrm{p}} > \pi_w & & & \\ \text{(iii) } \mathrm{d}f_{\mathrm{p}} = 0, & \text{if } C_{\mathrm{p}} = \pi_w, & \text{and} & f_{\mathrm{p}} = 0 & \text{with} \quad \mathrm{d}C_{\mathrm{p}} > 0^{\forall p \in R_w,\ w \in W} \\ \text{(iv) } \mathrm{d}f_{\mathrm{p}} > 0, & \text{if } C_{\mathrm{p}} = \pi_w, & \text{and} & f_{\mathrm{p}} = 0 & \text{with} \quad \mathrm{d}C_{\mathrm{p}} \leqslant 0 \end{array} \right\}$$

Therefore the directional derivatives can be obtained by the following affine variational inequality. Determine $\mathrm{d}q \in \widetilde{K}$ such that

$$(\nabla_y c(q,y,\Psi)\,\mathrm{d}y + \nabla_\Psi c(q,y,\Psi)\,\mathrm{d}\Psi + \nabla_q c(q,y,\Psi)\,\mathrm{d}q)^{\mathrm{t}}(z-\mathrm{d}q) \geqslant 0 \tag{3}$$

for all $z \in \widetilde{K}$ where $\nabla_y c$, $\nabla_\Psi c$ and $\nabla_q c$ are gradients evaluated at (q,y,Ψ) when the changes in decision variables $(\mathrm{d}y, \mathrm{d}\Psi)$ are specified.

2.5. An OPEC formulation

Let P be the performance index for the SIGNDP and Π be the constraints of signal settings, which can be expressed as a linear system of inequalities for minimum greens, maximum cycle time and capacity constraints. Let $S(y,\Psi)$ define the solution set of equilibrium flows, which can be obtained from (2). A SIGNDP can be formulated as follows:

$$\begin{array}{ll} \underset{y,\Psi,q}{\mathrm{Min}} & P = P_0(y,\Psi,q) \\ \text{subject to} & 0 \leqslant y \leqslant u, \Psi \in \Pi \\ & q \in S(y,\Psi) \end{array} \tag{4}$$

Let W^1_{aD} and W^1_{aS} denote link-specific weighting factors, respectively, for the rate of delay and the number of stops per unit time. The model (4) can be specified in the following manner:

$$\begin{array}{ll} \underset{y,\Psi,q}{\mathrm{Min}} & P = P_0(y,\Psi,q) \\ & = \sum_{a \in L} D_a W^1_{aD} + S_a W^1_{aS} + \gamma G_a(y_a) \\ \text{subject to} & \zeta_{\min} \leqslant \zeta \leqslant \zeta_{\max} \\ & g_{jm}\zeta \leqslant \phi_{jm} \leqslant 1, \quad \forall j,m \\ & q_a \leqslant s_a \Lambda_a, \quad \forall a \in L \\ & \theta_{jm} + \phi_{jm} + \bar{c}_{jlm}\zeta \leqslant \theta_{lm} + \Omega_m(j,l), \ j \neq l, \quad \forall j,l,m \\ & q \in S(y,\Psi) \end{array} \tag{5}$$

3. A solution method

As it widely recognized that the problem (5) is non-convex, in this section, a novel algorithm called the projected Quasi-Newton method for the SIGNDP is developed. The generalized gradients for (5) are firstly established.

3.1. The generalized gradients

Due to non-convexity of (5), the generalized gradients of (5) can be expressed in the following way. Introduce

$$\partial P(y^*, \Psi^*) = co\{\lim_{k\to\infty} \nabla P(y_k, \Psi_k) : (y_k, \Psi_k) \to (y^*, \Psi^*), \nabla P(y_k, \Psi_k) \text{ exists}\} \tag{6}$$

According to Clarke [10], the generalized gradient is a convex hull (abbreviated to *co*) of all points of the form $\lim \nabla P(y_k, \Psi_k)$ where the subsequence $\{y_k, \Psi_k\}$ converges to the limit value (y^*, Ψ^*). The gradients for the objective function of (5) can be given as follows. For $\varepsilon \in (y, \Psi)$

$$\nabla_\varepsilon P(y_k, \Psi_k) = \nabla_\varepsilon P_0(y_k, \Psi_k, q_k) + \nabla_q P_0(y_k, \Psi_k, q_k) \nabla_\varepsilon q(y_k, \Psi_k) \tag{7}$$

where the gradient of equilibrium flow with respect to the decision variables can be obtained from (3).

3.2. A projected Quasi-Newton method

A projected Quasi-Newton method for problem (5) can be established in the following way:

Theorem 1 (Quasi-Newton method). *Let $X_1 = (y_1, \Psi_1)$ locally solve (5), and Q_1 denote a positive definite symmetric matrix. For $k = 1, \ldots, n$, let $X_{k+1} = X_k + \alpha_k d_k$, where*

$$d_k = -Q_k \nabla P^{\mathrm{t}}(X_k) \tag{8}$$

is a descent direction at X_k with α_k for problem (5). For $k = 1, \ldots, n-1$, Q_k is given as follows:

$$Q_{k+1} = Q_k - \frac{Q_k u_k u_k^{\mathrm{t}} Q_k}{u_k^{\mathrm{t}} Q_k u_k} + \frac{v_k v_k^{\mathrm{t}}}{v_k^{\mathrm{t}} u_k} \tag{9}$$

where $u_k = X_{k+1} - X_k$ and $v_k^{\mathrm{t}} = \nabla P(X_{k+1}) - \nabla P(X_k)$. If $\nabla P(X_k) \neq 0$ for $k = 1, \ldots, n$, then $Q_1, \ldots, Q_n$ are symmetric and positive definite so that $d_1, \ldots, d_n$ are descent directions.

Proof. For $k = 1$, Q_1 is symmetric and positive definite, then

$$\nabla P^{\mathrm{t}}(X_1) d_1 = -\nabla P^{\mathrm{t}}(X_1) Q_1 \nabla P(X_1) < 0$$

it implies d_1 is a descent direction. We assume that the result holds true for $k \leqslant n-1$ and then show that it holds true for $k+1$. Let x be a non-zero vector such that

$$x^{\mathrm{t}} Q_{k+1} x = x^{\mathrm{t}} Q_k x - \frac{(x^{\mathrm{t}} Q_k u_k)^2}{u_k^{\mathrm{t}} Q_k u_k} + \frac{(x^{\mathrm{t}} v_k)^2}{v_k^{\mathrm{t}} u_k} \tag{10}$$

Since Q_k is a symmetric positive definite matrix, there exists a positive definite matrix $Q_k^{\frac{1}{2}}$ such that $Q_k = Q_k^{\frac{1}{2}} Q_k^{\frac{1}{2}}$. Let $p = Q_k^{\frac{1}{2}} x$, and $q = Q_k^{\frac{1}{2}} u_k$, then $x^{\mathrm{t}} Q_k x = p^{\mathrm{t}} p$, $u_k^{\mathrm{t}} Q_k u_k = q^{\mathrm{t}} q$, and $x^{\mathrm{t}} Q_k u_k = p^{\mathrm{t}} q$. Substituting in (10), we have

$$x^{\mathrm{t}} Q_{k+1} x = p^{\mathrm{t}} p - \frac{(p^{\mathrm{t}} q)^2}{q^{\mathrm{t}} q} + \frac{(x^{\mathrm{t}} v_k)^2}{v_k^{\mathrm{t}} u_k} = \frac{(p^{\mathrm{t}} p)(q^{\mathrm{t}} q) - (p^{\mathrm{t}} q)^2}{q^{\mathrm{t}} q} + \frac{(x^{\mathrm{t}} v_k)^2}{v_k^{\mathrm{t}} u_k} \tag{11}$$

According to Schwartz inequality,

$$(p^{\mathrm{t}} p)(q^{\mathrm{t}} q) - (p^{\mathrm{t}} q)^2 \geqslant 0$$

Because Q_k is positive definite, it implies

$$q^{\mathrm{t}} q = u_k^{\mathrm{t}} Q_k u_k > 0, \quad \forall u_k \neq 0$$

To show that

$$x^{\mathrm{t}}Q_{k+1}x > 0$$

it suffices to show that $v_k^{\mathrm{t}}u_k > 0$ and $v_k \neq 0$. Because

$$\begin{aligned} v_k^{\mathrm{t}}u_k &= (\nabla P(X_{k+1}) - \nabla P(X_k))^{\mathrm{t}}(X_{k+1} - X_k) = (\nabla P(X_{k+1}) - \nabla P(X_k))^{\mathrm{t}}\alpha_k d_k \quad (\text{since } \nabla P^{\mathrm{t}}(X_{k+1})d_k = 0) \\ &= -\alpha_k \nabla P^{\mathrm{t}}(X_k)d_k = \alpha_k \nabla P^{\mathrm{t}}(X_k)Q_k \nabla P(X_k) > 0 \end{aligned}$$

(since Q_k is positive definite and is a descent direction with $\alpha_k > 0$).

Suppose $v_k = 0$, it will contradicts $v_k^{\mathrm{t}}u_k > 0$, thus $v_k \neq 0$. Therefore

$$x^{\mathrm{t}}Q_{k+1}x > 0, \quad \forall x \neq 0$$

It implies Q_{k+1} is positive definite and follows that

$$\nabla P^{\mathrm{t}}(X_{k+1})d_{k+1} = -\nabla P^{\mathrm{t}}(X_{k+1})Q_{k+1}\nabla P(X_{k+1}) < 0, \quad \forall \nabla P(X_{k+1}) \neq 0$$

Therefore d_{k+1} is a descent direction at X_{k+1}. □

The search direction generated by the Quasi-Newton method for an unconstrained non-linear problem is a descent direction which strictly decreases the objective function value provided that the corresponding gradient value is not zero. In problem (5), let A and B, respectively, denote the coefficient matrix and constant vector associated with the signal settings. Using the first order partial derivatives (7) for the objective function, the problem (5) can be re-expressed in the following manner.

$$\begin{aligned} \min_{y,\Psi} \quad & P = P_1(y, \Psi) \\ \text{subject to} \quad & 0 \leqslant y \leqslant u \\ & A\Psi^{\mathrm{t}} \leqslant B \end{aligned} \tag{12}$$

Theorem 2 (A projected Quasi-Newton method). *In problem (12), a sequence of iterates $\{y_k, \Psi_k\}$ can be generated according to*

$$y_{k+1} = y_k + \alpha_k G_k d_k^y \tag{13}$$

$$\Psi_{k+1} = \Psi_k + \alpha_k G_k d_k^\Psi \tag{14}$$

where $d_k = \begin{pmatrix} d_k^y \\ d_k^\Psi \end{pmatrix}$ is the descent direction determined by (8) and α_k is the step length which minimizes P_1 along d_k. The projection matrix G_k is of the following form.

$$G_k = I - M_k^{\mathrm{t}}(M_k M_k^{\mathrm{t}})^{-1}M_k \tag{15}$$

In problem (12), M_k is the gradient of active constraints, where the active constraint gradients are linearly independent and thus M_k has full rank. The search direction s_k can be determined in the following form:

$$s_k = G_k d_k \tag{16}$$

Then the sequences of points $\{y_k, \Psi_k\}$ are generated by the projected Quasi-Newton method such that

$$P_1(y_k, \Psi_k) > P_1(y_{k+1}, \Psi_{k+1}), \quad k = 1, 2, \ldots \tag{17}$$

whenever $G_k \nabla P_1(y_k, \Psi_k) \neq 0$.

Proof. Following the results of Theorem 1, we have

$$\nabla P_1^{\mathrm{t}}(y_k, \Psi_k)d_k = -\nabla P_1^{\mathrm{t}}(\Psi_k)Q_k \nabla P_1(\Psi_k) < 0, \quad k = 1, 2, 3, \ldots. \tag{18}$$

Multiply Eq. (18) on both sides by the projection matrix G_k, it becomes

$$\nabla P_1^t(y_k,\Psi_k)G_kd_k = -\nabla P_1^t(y_k,\Psi_k)G_kQ_k\nabla P_1(y_k,\Psi_k) = -\nabla P_1^t(y_k,\Psi_k)G_k^tG_kQ_k^{\frac{1}{2}}Q_k^{\frac{1}{2}}\nabla P_1(y_k,\Psi_k)$$
$$= -\left\|G_kQ_k^{\frac{1}{2}}\nabla P_1(y_k,\Psi_k)\right\|^2 < 0 \tag{19}$$

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

$$P_1(y_k,\Psi_k) > P_1((y_k,\Psi_k)+\beta s_k) \tag{20}$$

Because by definition α_k is the step length which minimizes P_1 along s_k from (y_k,Ψ_k), it implies

$$P_1(y_k,\Psi_k) > P_1((y_k,\Psi_k)+\beta s_k) \geqslant P_1((y_k,\Psi_k)+\alpha_k s_k) = P_1(y_{k+1},\Psi_{k+1}) \tag{21}$$

which completes this proof. □

Theorem 3 (A projected Quasi-Newton method, when $G_k\nabla P_1(y_k,\Psi_k)=0$). *Following the results of Theorem 2, when $G_k\nabla P_1(y_k,\Psi_k)=0$ and if all the Lagrange multipliers corresponding to the active constraint gradients are positive or zeros, it implies the current (y_k,Ψ_k) is a local optimal solution. Otherwise choose one negative Lagrange multiplier, say μ_j, and construct a new $\widehat{M}_k$ of the active constraint gradients by deleting the jth row of M_k, which corresponds to the negative component μ_j, and make the projection matrix of the following form*:

$$G_k = I - \widehat{M}_k^t(\widehat{M}_k\widehat{M}_k^t)^{-1}\widehat{M}_k \tag{22}$$

The search direction then can be determined by (16) and the results of Theorem 2 hold.

Proof. This proof is separated as in the following two parts. First, suppose all the Lagrange multipliers are positive or zeros, and by definition we have

$$0 = G_k\nabla P_1(y_k,\Psi_k) = (I - M_k^t(M_kM_k^t)^{-1}M_k)\nabla P_1(y_k,\Psi_k) = \nabla P_1(y_k,\Psi_k) + M_k^t w \tag{23}$$

where $w = -(M_kM_k^t)^{-1}M_k\nabla P_1(y_k,\Psi_k)$ corresponds to the Lagrange multipliers μ matching the active constraint gradients M_k. If $\mu \geqslant 0$, and the active constraint gradients are linearly independent, then (y_k,Ψ_k) is a local optimal solution to (12) and the current iterate (y_k,Ψ_k) is a local optimal point. Or choose one negative Lagrange multiplier, say μ_j, and construct another $\widehat{M}_k$ of the active constraint gradients by deleting the jth row of M_k, and make a new projection matrix of the form given in (22). The new search direction thus can be determined by (16). Provided that $G_k\nabla P_1(y_k,\Psi_k)\neq 0$, then the results of Theorem 3 hold. We only need to show that $G_k\nabla P_1(y_k,\Psi_k)\neq 0$. By contradiction, suppose that $G_k\nabla P_1(y_k,\Psi_k)=0$. By the definition of G_k and letting $\widehat{w} = -(\widehat{M}_k\widehat{M}_k^t)^{-1}\widehat{M}_k\nabla P_1$, we have

$$0 = G_k\nabla P_1(y_k,\Psi_k) = (I - \widehat{M}_k^t(\widehat{M}_k\widehat{M}_k^t)^{-1}\widehat{M}_k)\nabla P_1(y_k,\Psi_k) = \nabla P_1(y_k,\Psi_k) + \widehat{M}_k^t\widehat{w} \tag{24}$$

Since $\widehat{M}_k^t\widehat{w} = \widehat{M}_k^t\overline{w} + \mu_jA_j^t$ where A_j is the jth row of active constraint gradients, from (23) we get

$$0 = \nabla P_1(y_k,\Psi_k) + \widehat{M}_k^t\overline{w} + \mu_jA_j^t \tag{25}$$

Subtracting (25) from (24), it follows

$$0 = \widehat{M}_k^t(\widehat{w}-\overline{w}) - \mu_jA_j^t \tag{26}$$

Since $\mu_j \neq 0$, it violates that the assumption $\widehat{M}_k$ of full rank. Therefore $G_k\nabla P_1\neq 0$. □

According to the results of Theorems 2 and 3, the iterative process in search for the solution of problem (12) can either be terminated at a local optimal point or a new search direction can be generated.

Corollary 4 (Stopping condition). *If (y_k,Ψ_k) is a local solution for (12) then the search process may stop; otherwise a new descent direction at (y_k,Ψ_k) can be generated according to Theorems 2 and 3.* □

3.3. Variants of QNEW

In order to effectively solve the problem (12), a number of variants of QNEW can be developed as follows.

3.3.1. QNEW-1

QNEW-1 is a locally optimal search for a full optimization with respect to the link capacity expansions, the common cycle time, the start and duration of greens at each junction. It can be specified in the following steps.

Step 1. Start with (y_k, Ψ_k) and set index $k = 1$.
Step 2. Solve a user equilibrium problem with fixed link capacity expansions and signal settings. Calculate the first order derivatives for the user equilibrium with parameters by solving variational inequality (3).
Step 3. Conduct the projected Quasi-Newton method to determine a search direction.

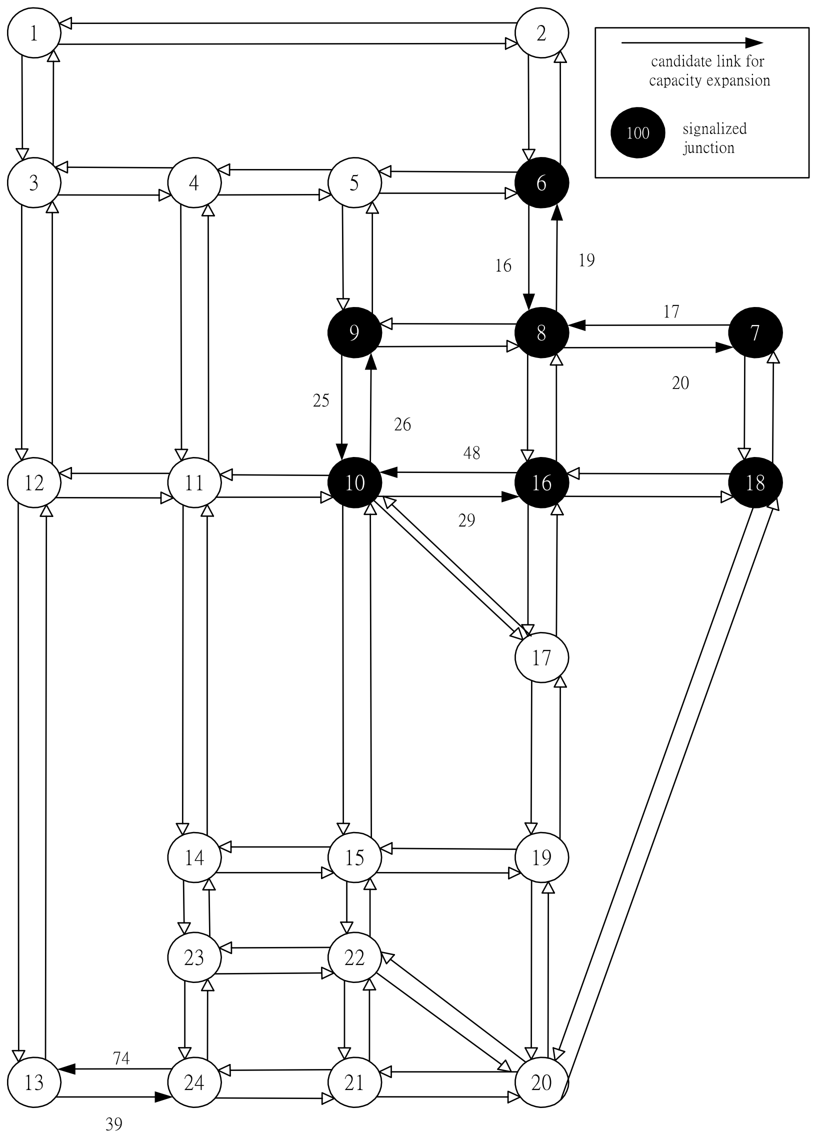

Fig. 1. Sioux Falls network.

Step 4. If $G_k \nabla P_1(y_k, \Psi_k) \neq 0$, find a new (y_{k+1}, Ψ_{k+1}) by (13) and (14). Replace k by $k+1$ and go to Step 2. If $G_k \nabla P_1(y_k, \Psi_k) = 0$ and all the Lagrange multipliers corresponding to the active constraint gradients are positive or zeros, (y_k, Ψ_k) is the local optimal point and stop; otherwise, according to Theorem 3, find a new projection matrix and go to Step 3.

3.3.2. QNEW-2

Conduct QNEW-1 for signal settings with respect to the start and duration of greens only and link capacity expansions for given common cycle time, a locally optimal search is performed again at each junction.

3.3.3. QNEW-3

Conduct a global search for signal settings with respect to the start of greens only and link capacity expansions for given common cycle time and the duration of green for all signal groups at each junction. By doing so, a new solution point can be easily identified in another part of the feasible region. After such point is found, conduct QNEW-1 again and locate new signal settings and link capacity expansions.

3.3.4. QNEW-4

A mixture of steps can be performed alternately by carrying out QNEW-1, QNEW-2 and QNEW-3 until the difference of the values of the performance index between successive iterations is negligible.

3.3.5. QNEW-P

In order to effectively conduct the local search in signal settings and link capacity expansions, the PARTAN search technique [14] is particularly taken into account in QNEW-1.

Table 1
Computational results for Sioux Falls network at 1st set initials

	SAB	GA	MIX	QNEW-1	QNEW-2	QNEW-3	QNEW-4	QNEW-P
Initial value of $1/\zeta$	75	75	75	75	75	75	75	75
Initial value of ϕ_{jm}/ζ	32	32	32	32	32	32	32	32
Initial value of y_a	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5
Initial value of PI	135	135	135	135	135	135	135	135
$1/\zeta$	108	110/entry>	118	106	126	110	120	114
ϕ_{16}/ζ	46	50	39	57	42	49	52	59
ϕ_{17}/ζ	40	42	52	48	48	54	55	53
ϕ_{18}/ζ	59	52	44	46	49	56	53	50
ϕ_{19}/ζ	55	45	54	50	52	54	55	56
$\phi_{1,10}/\zeta$	49	48	52	54	49	48	51	50
$\phi_{1,16}/\zeta$	49	51	55	55	53	54	55	49
$\phi_{1,18}/\zeta$	47	52	48	52	48	48	52	48
y_{16}	4.8	3.1	5.2	4.9	5.5	4.1	4.2	3.9
y_{17}	3.9	3.3	5.3	5.5	3.6	4.5	5.3	5.7
y_{19}	1.2	2.6	3.6	3.4	2.2	3.6	2.4	1.9
y_{20}	1.6	2.7	1.7	2.6	3.4	2.7	2.7	2.6
y_{25}	2.3	2.6	2.3	2.6	3.1	2.9	2.5	3.1
y_{26}	2.3	2.9	2.7	3.7	3.2	2.1	2.2	2.9
y_{29}	5.5	4.3	5.7	5.4	4.6	4.3	5.7	5.2
y_{39}	4.6	4.1	4.9	5.5	4.8	4.4	5.1	5.5
y_{48}	5.4	4.7	5.3	4.5	3.8	1.9	4.7	4.5
y_{74}	6.5	4.2	3.9	4.1	3.5	2.8	4.2	4.1
PI (in veh)	90.2	94.3	86.7	82.1	82.7	81.4	81.1	80.8
Cpu time (in s)	44	65	39	12	13	14	15	13
#	64	76	62	35	36	37	39	36

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

4. Numerical calculations

In this section, numerical computations are conducted by QNEW in signal-controlled networks where the effectiveness and robustness of QNEW are widely investigated. In the following numerical computations, comparisons are made for the sensitivity analysis based method (SAB) [19], the GA method [7] and the mix search heuristic (MIX) [8]. Using typical values found in practice, the minimum green time is 7 s, the clearance times are 5 s and the maximum cycle time is 180 s.

4.1. Sioux Falls city network

The first example network is chosen from a Sioux Falls city network as shown in Fig. 1. The non-signalized link travel time and link investment cost functions used are adopted from [1, pp. 261–262]. The convex

Table 2
Computational results for Sioux Falls network at 2nd set initial

	SAB	GA	MIX	QNEW-1	QNEW-2	QNEW-3	QNEW-4	QNEW-P
Initial value of $1/\zeta$	135	135	135	135	135	135	135	135
Initial value of ϕ_{jm}/ζ	60	60	60	60	60	60	60	60
Initial value of y_a	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5
Initial value of PI	144	144	144	144	144	144	144	144
$1/\zeta$	130	128	110	112	112	116	118	124
ϕ_{16}/ζ	53	51	47	54	42	47	52	55
ϕ_{17}/ζ	58	54	50	55	41	51	56	61
ϕ_{18}/ζ	57	60	44	48	59	56	53	58
ϕ_{19}/ζ	62	59	51	50	55	54	57	55
$\phi_{1,10}/\zeta$	59	48	52	51	49	48	51	52
$\phi_{1,16}/\zeta$	55	51	53	55	53	54	55	55
$\phi_{1,18}/\zeta$	57	60	49	52	48	49	51	52
y_{16}	4.9	3.9	5.4	4.7	5.3	3.1	5.2	4.6
y_{17}	3.6	3.8	5.2	5.1	3.6	3.8	5.3	5.2
y_{19}	2.6	2.7	1.6	3.4	3.2	2.6	1.6	2.4
y_{20}	1.7	2.7	2.6	2.6	2.2	2.7	2.6	2.1
y_{25}	2.3	3.6	2.3	2.3	3.4	2.9	2.1	1.6
y_{26}	2.3	3.9	2.2	2.9	3.8	2.6	2.2	2.3
y_{29}	5.1	4.1	5.5	5.4	4.6	4.3	5.7	5.4
y_{39}	4.8	4.2	4.9	4.6	4.1	4.4	4.8	5.6
y_{48}	5.6	4.7	4.4	4.5	5.8	4.7	4.8	4.8
y_{74}	5.2	4.2	4.3	4.7	3.6	4.8	4.1	4.7
PI (in veh)	95.7	90.8	83.5	84.1	83.2	83.1	82.1	81.9
Cpu time (in s)	45	69	42	14	14	14	15	13
#	66	79	65	36	36	38	39	36

Table 3
Computational results for Sioux Falls network with scaling demands

	SAB	GA	MIX	QNEW-1	QNEW-2	QNEW-3	QNEW-4	QNEW-P
1.0	88.72	84.34	82.57	84.18	83.72	82.34	82.11	82.08
1.1	92.35	85.12	83.09	84.77	84.08	82.29	82.08	82.07
1.2	95.41	90.45	84.17	85.41	84.12	83.11	83.09	83.01
1.3	97.33	89.76	85.43	87.12	86.55	84.67	84.12	83.06
1.4	100.21	93.44	89.08	90.19	89.08	85.16	85.11	83.25
1.5	110.54	99.18	91.33	92.34	90.16	87.26	85.66	84.56
1.6	112.77	102.56	92.67	96.15	95.22	90.54	88.17	85.26
1.7	129.98	105.41	95.44	99.05	98.34	92.33	91.08	88.16
1.8	127.09	116.78	103.65	106.45	102.31	99.08	97.49	90.71
1.9	135.21	121.09	110.87	112.33	110.99	106.35	103.71	93.44

investment function form is adopted from Abdulaal and LeBlanc [1, p. 28]. Computational results are summarized in Tables 1–4. As it observed in Tables 1 and 2, QNEW-P significantly improved the PI values of 80.8 veh and 81.9 veh with rates of 40% and 43% at two sets of initial signal settings. In particular, when the traffic road network becomes more congested as it seen in Tables 3 and 4, QNEW-P again outperformed other algorithms while requiring much less computing overhead.

4.2. Large-scale grid networks

Three kinds of large-scale grid networks are illustrated for investigations on the robustness and effectiveness of QNEW-P and other algorithms, which can be denoted by $n/m/k/x/y$, respectively, representing for the

Table 4
Computational results in cpu time for Sioux Falls network with scaling demands

	SAB	GA	MIX	QNEW-1	QNEW-2	QNEW-3	QNEW-4	QNEW-P
1.0	31.9	32	23.6	10.9	11.3	11.3	11.6	11.7
1.1	42.3	35.2	23.9	11.7	14.8	12.9	12.8	12.1
1.2	55.1	40.5	34.7	12.4	14.2	13.1	15.9	13.1
1.3	47.3	39.7	35.3	17.1	16.5	19.6	24.2	13.6
1.4	60.1	43.4	39.8	20.9	19.8	25.6	29.1	13.5
1.5	71.4	39.8	39.3	22.3	19.6	27.6	35.6	14.9
1.6	82.7	42.6	42.6	26.5	22.8	30.5	38.7	15.6
1.7	89.9	45.4	45.4	29.5	18.4	32.3	41.8	18.6
1.8	97.9	46.8	43.6	33.4	22.3	39.8	44.7	20.7
1.9	105.2	52.9	40.8	32.3	25.9	41.5	47.1	23.4

Table 5
Computational results in PI values on $n/m/k/x/y = 225/420/42/9/10$ grid network

	GA	MIX	QNEW-1	QNEW-2	QNEW-3	QNEW-4	QNEW-P
1st set	60.6	56.7	56.2	56.1	54.7	51.2	50.1
2nd set	65.7	53.9	54.8	55.1	53.6	50.8	48.2
3rd set	61.1	54.3	55.1	55.2	53.4	48.1	47.8
4th set	70.3	56.2	56.1	56.4	54.1	49.1	49.2
5th set	71.2	55.9	55.8	54.9	53.2	50.2	47.5
6th set	67.8	58.4	55.3	54.8	52.9	51.1	49.5
7th set	66.5	55.5	56.1	56.3	53.1	52.3	49.9
8th set	68.1	56.1	56.2	55.7	52.9	51.4	50.1
9th set	69.2	53.2	54.1	55.2	53.3	52.5	49.6
10th set	70.3	54.8	53.9	55.9	53.8	52.5	51.7
Average/std.	67.1/3.7	55.5/1.5	55.4/0.9	55.6/0.6	53.5/0.6	50.9/1.5	49.4/1.3

Where std. stands for standard deviation.

Table 6
Computational results in PI values on $n/m/k/x/y = 625/1200/76/16/20$ grid network

	GA	MIX	QNEW-1	QNEW-2	QNEW-3	QNEW-4	QNEW-P
1st set	139.1	132.7	135.2	133.8	129.7	118.4	110.1
2nd set	145.7	143.9	144.8	145.1	137.6	120.8	118.2
3rd set	141.1	144.3	142.1	141.2	138.4	121.9	117.9
4th set	149.3	146.2	136.1	138.4	124.1	119.1	112.2
5th set	151.2	149.9	145.8	142.9	133.2	120.2	117.5
6th set	157.8	148.4	152.3	141.8	129.9	122.1	119.5
7th set	156.5	155.5	146.1	147.3	137.1	125.3	122.9
8th set	148.1	146.1	151.2	153.7	142.9	124.4	120.6
9th set	159.2	153.2	154.1	152.2	146.3	127.5	119.6
10th set	155.3	154.8	153.9	145.9	133.8	126.5	118.7
Average/std.	150.3/7	147.5/6.7	146.2/6.9	144.2/6	135.3/6.6	122.6/3.2	117.7/3.8

number of nodes, links, origin-destination trips, signal-controlled junctions and selected links with capacity expansions. Computational results are summarized in Tables 5–10. Ten sets of distinct initial data are taken into account where trip rates range from 1.0 to 2.5 veh/h and saturation flows rate from 1000 to 1800 veh/h. Initial data for link capacity expansions are set from zeros to 25 and selected signal settings are set with equally divided green splits with variable cycle times from 60 to 180 s.

As it seen from Tables 5–10, QNEW-P again shows the superiority in solving the SIGNDP. For example, as it seen in Table 5, QNEW-P substantially outperformed the well-known GA and MIX by 26% and 11% on average. Also, as it observed in Table 6, again, QNEW-P outperformed GA and MIX nearly by 22% and

Table 7
Computational results in PI values on $n/m/k/x/y = 2025/3960/120/25/30$ grid network

	GA	MIX	QNEW-1	QNEW-2	QNEW-3	QNEW-4	QNEW-P
1st set	207.1	204.4	185.2	193.8	166.7	173.4	150.1
2nd set	215.7	214.9	194.8	195.1	177.6	170.8	148.2
3rd set	204.1	207.3	192.1	191.2	178.4	171.9	151.9
4th set	209.3	216.2	186.1	188.4	174.1	169.1	152.2
5th set	215.2	219.9	195.8	192.9	183.2	170.2	147.5
6th set	217.8	214.4	182.3	191.8	179.9	172.1	149.5
7th set	206.5	205.5	186.1	177.3	177.1	165.3	152.1
8th set	218.1	214.1	191.2	193.7	182.9	174.4	150.5
9th set	229.2	225.2	194.1	192.2	176.3	167.5	149.1
10th set	225.3	224.8	193.9	195.9	173.8	166.5	148.1
Average/std.	214.8/8.2	214.7/7.4	190.2/4.8	191.2/5.3	177/4.8	170.1/3	149.9/1.7

Table 8
Computational results in cpu time (s) on $n/m/k/x/y = 225/420/42/9/10$ grid network

	GA	MIX	QNEW-1	QNEW-2	QNEW-3	QNEW-4	QNEW-P
1st set	10.6	9.7	10.6	9.7	10.2	9.6	9.9
2nd set	9.7	10.9	9.7	10.9	9.5	10.8	10.2
3rd set	11.1	9.3	10.1	9.3	10.1	9.9	9.5
4th set	10.3	10.2	9.3	10.2	9.3	10.5	10.1
5th set	11.2	9.9	10.7	9.7	10.2	9.2	9.7
6th set	9.8	9.4	9.4	9.4	9.7	9.4	9.5
7th set	10.5	10.5	9.1	10.1	10.1	10.1	10.5
8th set	9.8	10.1	10.2	9.4	9.5	9.1	10.1
9th set	9.2	10.2	9.2	9.2	9.2	9.7	10.2
10th set	10.3	10.8	9.3	9.8	10.3	10.1	10.1
Average	10.3	10.1	9.8	9.8	9.8	9.8	10

Table 9
Computational results in cpu time (s) on $n/m/k/x/y = 625/1200/76/16/20$ grid network

	GA	MIX	QNEW-1	QNEW-2	QNEW-3	QNEW-4	QNEW-P
1st set	31.6	29.7	31.2	32.7	30.6	35.7	30.1
2nd set	29.7	30.9	29.7	31.9	29.5	31.9	28.7
3rd set	31.6	29.3	30.5	33.3	31.6	29.3	30.6
4th set	30.3	30.2	30.3	30.2	30.3	29.2	29.3
5th set	41.2	29.9	33.2	29.9	31.2	29.9	31.2
6th set	29.8	31.4	29.8	33.4	29.8	36.4	29.5
7th set	32.5	30.5	32.1	32.5	42.5	28.5	31.5
8th set	29.9	30.1	29.9	30.1	29.9	40.1	29.9
9th set	27.2	30.2	28.2	33.2	33.2	38.2	28.2
10th set	33.3	31.8	31.3	32.3	35.3	39.8	30.3
Average	31.7	30.4	30.6	31.9	32.4	33.9	29.9

Table 10
Computational results in cpu time (s) on $n/m/k/x/y = 2025/3960/120/25/30$ grid network

	GA	MIX	QNEW-1	QNEW-2	QNEW-3	QNEW-4	QNEW-P
1st set	99.6	109.7	110.6	119.7	120.2	129.6	109.9
2nd set	98.7	104.9	109.7	120.9	129.5	130.8	101.2
3rd set	101.1	99.3	110.1	119.3	120.1	129.9	99.5
4th set	105.3	100.2	109.3	120.2	119.3	135.5	100.1
5th set	111.2	99.9	110.7	119.7	130.2	129.2	99.7
6th set	109.8	99.4	119.4	119.4	129.7	129.4	99.5
7th set	100.5	100.5	119.1	120.1	120.1	131.1	100.5
8th set	99.8	103.1	120.2	119.4	129.5	129.1	99.1
9th set	99.2	102.2	119.2	117.2	129.2	129.7	96.2
10th set	107.3	105.8	119.3	115.8	130.3	140.1	98.1
Average	103.3	102.5	114.8	119.2	125.8	131.4	100.4

20%, respectively. The minimum cost flow algorithm of Edmonds-Karp's capacity scaling [2] with a polynomial time bound of $O(m \log n(m + n \log n))$ was utilized for implementing the traffic assignments for the SIGNDP. Computational efforts on algorithms mentioned in this paper were carried out on SUN SPARC SUNW, 900 MHZ processor under operating system Unix SunOS 5.8.

5. Conclusions and discussions

In this paper, we presented an optimal scheme for SIGNDP and carried out a series of comparative analyses for various algorithms. The proposed QNEW-P consistently outperformed other algorithms while exhibiting similar computing overhead. Regarding the issues about road works occurring on one side of link, a user equilibrium assignment with non-separable cost mappings is being considered and we will discuss this issue in a subsequent paper.

Acknowledgements

Many thanks go to referees' helpful comments on earlier versions of this paper and 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 13B (1979) 19–32.
[2] R.K. Ahuja, T.L. Magnanti, J.B. Orlin, Network Flows: Theory, Algorithms, and Applications, Prentice-Hall, New Jersey, 1993.
[3] M.A. Amouzegar, K. Moshirvaziri, Strategic management decision support system: an analysis of the environmental policy issues, Environmental Modeling and Assessment 6 (2001) 297–306.
[4] G. Anandalingam, T. Friesz, Hierarchical optimization: an introduction, Annals of Operations Research 34 (1992) 1–11.
[5] J.X. Ban, H.X. Liu, M.C. Ferris, B. Ran, A general MPCC model and its solution algorithm for continuous network design problem, Mathematical and Computer Modelling 43 (2006) 493–505.
[6] J. Bard, Practical Bilevel Optimization, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.
[7] H. Ceylan, M.G.H. Bell, Traffic signal timing optimization based on genetic algorithm approach, including drivers routing, Transportation Research Part B 38 (2004) 329–342.
[8] S.-W. Chiou, Optimization of area traffic control for equilibrium network flows, Transportation Science 33 (1999) 279–289.
[9] C.-H. Chou, J.-C. Teng, A fuzzy logic controller for traffic junction signals, Information Sciences 143 (2002) 73–97.
[10] F.F. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.
[11] T.L. Friesz, R.L. Tobin, H-L. Cho, N.J. Mehta, Sensitivity analysis based algorithms for mathematical programs with variational inequality constraints, Mathematical Programming 48 (1990) 265–284.
[12] Z. Gao, H. Sun, L.L. Shan, A continuous equilibrium network design model and algorithms for transit systems, Transportation Research Part B 38 (2004) 235–250.
[13] J. Liu, C. Shi, G. Zhang, On bilevel multi-follower decision making: general framework and solutions, Information Sciences 176 (2006) 1607–1627.
[14] D.G. Luenberger, Linear and Nonlinear Programming, Addison-Wesley, Reading, MA, 1989.

[15] Q. Meng, H. Yang, M.G.H. Bell, An equivalent continuously differentiable model and a locally convergent algorithm for the continuous network design problem, Transportation Research Part B 35 (2001) 83–105.
[16] M. Patriksson, Sensitivity analysis of traffic equilibria, Transportation Science 38 (3) (2004) 258–281.
[17] K. Shimizu, Y. Ishizuka, J. Bard, Nondifferentiable and Two-level Mathematical Programming, Kluwer Academic Publishers, Boston, 1997.
[18] C. Suwansirikul, T.L. Friesz, R.L. Tobin, Equilibrium decomposed optimization: a heuristic for continuous equilibrium network design problem, Transportation Science 21 (1987) 254–263.
[19] H. Yang, S. Yagar, Traffic assignment and signal control in saturated road networks, Transportation Research 29A (1995) 125–139.