Available online at www.sciencedirect.com

SCIENCE DIRECT®

Transportation Research Part B 39 (2005) 361–383

TRANSPORTATION RESEARCH PART B

www.elsevier.com/locate/trb

Bilevel programming for the continuous transport network design problem

Suh-Wen Chiou *

Department of Information Management, National Dong Hwa University, 1, Sec 2, Dahsueh Road, Shoufeng, Hualien 97401, Taiwan, ROC

Received 10 March 2002; received in revised form 22 September 2003; accepted 13 May 2004

Abstract

A Continuous Network Design Problem (CNDP) is to determine the set of link capacity expansions and the corresponding equilibrium flows for which the measures of performance index for the network is optimal. A bilevel programming technique can be used to formulate this equilibrium network design problem. At the upper level problem, the system performance index is defined as the sum of total travel times and investment costs of link capacity expansions. At the lower level problem, the user equilibrium flow is determined by Wardrop's first principle and can be formulated as an equivalent minimization problem. In this paper we exploit a descent approach via the implementation of gradient-based methods to solve CNDP generally where the Karush–Kuhn–Tucker points can be obtained. Four variants of gradient-based methods are presented and numerical comparisons are widely made with the previous on three kinds of test networks. The proposed methods have achieved substantially better results in terms of the robustness to the initials and the computational efficiency in solving equilibrium assignment problems than did others especially when the congested road networks are considered.

Keywords: Network design problem; Bilevel programming; Gradient-based method

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

0191-2615/$ - see front matter
doi:10.1016/j.trb.2004.05.001

1. Introduction

A Continuous Network Design Problem (CNDP) is to determine the set of link capacity expansions where Wardrop's first principle is followed. The measure of system performance can be described as the sum of total travel times and investment cost of link capacity expansions. CNDP is one of the most computationally intensive problems in transportation field (Yang and Bell, 1998). Abdulaal and LeBlanc (1979) were the first ones who proposed the Hooke-Jeeves algorithm to solve CNDP and tested that algorithm on a medium-sized realistic network. Gershwin and Tan (1979) formulated the CNDP as a constrained optimization problem in which the constrained set was expressed in terms of the path flows and performed their method on small networks. Marcotte (1983) and Marcotte and Marquis (1992) presented heuristics for CNDP on the basis of system optimal approach and obtained good numerical results. However, these heuristics have not been extensively tested on large-scale networks generally.

As for the development of solution methods to CNDP for practical use, Suwansirikul et al. (1987) proposed a simple heuristic called the Equilibrium Decomposed Optimization (EDO) and performed this heuristic on several example networks. The computing efficiency of EDO results from the decomposition of the CNDP into a set of interacting optimization problems and at each iteration only one user equilibrium needs to be calculated when updating the improvements of all links of the network. By approximating the first-order derivative of the objective function with respect to the decision variable, EDO employed one-dimensional search to locate good solutions for CNDP. Using the results from sensitivity analysis for equilibrium network flows (Tobin and Friesz, 1988), Friesz et al. (1990) and Suh and Kim (1992) have respectively proposed various solution methods to CNDP. Furthermore, using the marginal function (Chen and Florian, 1995) defined by the optimal value function of the lower level equilibrium problem, Meng et al. (2001) transferred the bilevel program of CNDP into a single level continuously differentiable optimization problem. Also an augmented Lagrangian method for solving CNDP has been proposed and good results have been obtained from various numerical tests.

Since the bilevel program for CNDP is non-convex and non-differentiable, regarding the development of solution methods for bilevel programming problem (BLPP), Luo et al. (1996), Marcotte and Zhu (1996), Shimizu et al. (1997) and Outrata et al. (1998), respectively characterized the optimality conditions and derived the corresponding solution methods for BLPP where the non-smooth approach (Clarke, 1983) has been taken into account. Regarding the sensitivity-based approach applied to BLPP, Falk and Liu (1995) investigated theoretic analysis for general non-linear BLPP and proposed a descent approach in terms of the bundle method to solve the non-linear bilevel programming problem where the gradient of the objective function can be obtained when the subgradient information of the lower level are available. Regarding the technique of sensitivity analysis applied to BLPP in transportation field, Yang and Yagar (1995) presented a sensitivity analysis based algorithm (SAB) to solve the signal control and traffic assignment problem where a linearized sub-problem was formulated at current signal settings. Chiou (1999) explored a mixed search procedure to solve an area traffic control optimization problem confined to equilibrium network flows, where good local optima can be effectively found via the gradient projection method. Yang (2001) has also conducted a good survey of recent advances in transportation bilevel programming problems. In this paper, following the theoretic analysis of BLPP presented by Falk and Liu and adopting the results for sensitivity analysis for equilibrium flows, we propose a BLPP model and four variants of gradient-based methods to generally solve CNDP. The Karush–Kuhn–Tucker points for CNDP can be effectively identified via the application of gradient-based methods for locally optimal search. Numerical tests are conducted extensively on various types of networks where the comparison results are made for the algorithms presented and the previous.

The rest of this paper is organized as follows. In Section 2, we summarize fundamental mathematical preliminaries related to BLPP and the sensitivity analysis for equilibrium network flows. In Section 3, we state BLPP formulation of CNDP and in Section 4, new solution algorithms based on the descent approach via the implementation of gradient-based methods are proposed. In Section 5, numerical experiments are conducted on various types of networks. Conclusions and discussions for this paper are drawn in Section 6.

2. Preliminaries

Throughout this paper, vectors are treated as column vectors with transpose denoted as a ‘t’ superscript. For a scalar-valued function f of vector $x \in \Re^n$, the row vector of derivatives is denoted as $\nabla f(x) = \left[\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right]$. For a vector-valued function $g : \Re^n \to \Re^m$, the Jacobian is denoted by the $m \times n$ matrix $\nabla g(x) = [\nabla g_1(x)^{\mathrm{t}}, \ldots, \nabla g_m(x)^{\mathrm{t}}]^{\mathrm{t}}$. The notation $\nabla f_a(a, b)$ denotes the gradient vector of f with respect to the first argument.

Let $F : \Re^n \times \Re^q \to \Re, f : \Re^n \times \Re^q \to \Re^n$ be continuously differentiable in (x, y), $g : \Re^n \times \Re^q \to \Re^m$ twice continuously differentiable in (x, y) and concave in x, and $h : \Re^n \times \Re^q \to \Re^p$ linear affine in x and continuously differentiable in y. A general bilevel programming problem can be formulated as to

$$\mathbf{Min} \quad F(x(y), y) \tag{1}$$

$$\text{s.t.} \quad f(x,y)(z-x) \geqslant 0, \quad x \in K(y), \ \forall z \in K(y) \tag{2}$$

$$y \in S$$

where S is the feasible set of y, and $K(y)$ is the feasible constraint set for x given by

$$K(y) = \{x : g(x,y) \geqslant 0, h(x,y) = 0\} \tag{3}$$

For a fixed y°, if x° solves problem (2), it follows that the Karush–Kuhn–Tucker conditions hold at x°, i.e. there exists Lagrange multipliers $\pi^\circ \in \Re^m$, $\mu^\circ \in \Re^p$ such that $(x^\circ, \pi^\circ, \mu^\circ)$ satisfies the following equations:

$$\begin{aligned} & f(x^\circ, y^\circ) - \nabla g(x^\circ, y^\circ)^{\mathrm{t}} \pi^\circ - \nabla h(x^\circ, y^\circ)^{\mathrm{t}} \mu^\circ = 0 \\ & \pi^{\circ\mathrm{t}} g(x^\circ, y^\circ) = 0 \\ & \pi^\circ \geqslant 0 \end{aligned} \tag{4}$$

Moreover, the following conditions are assumed to hold at x° for problem (2) with fixed y° (Fiacco, 1976; Tobin, 1986).

A1. The Second Order Sufficient Conditions for a locally unique solution (SOSC)
If $d^t\nabla f(x^\circ, y^\circ)d > 0$, for $d \neq 0$, such that

$$\nabla g_i(x^\circ, y^\circ)d \geqslant 0 \text{ for all } i \text{ such that } g_i(x^\circ, y^\circ) = 0$$

$$\nabla g_i(x^\circ, y^\circ)d = 0 \text{ for all } i \text{ such that } \pi_i^\circ > 0$$

$$\nabla h_i(x^\circ, y^\circ)d = 0 \text{ for all } i$$

A2. The Linear Independence condition (LI)
If $\nabla g_i(x^\circ, y^\circ)$, i such that $g_i(x^\circ, y^\circ) = 0$, and $\nabla h_i(x^\circ, y^\circ)$, $i = 1, \ldots, p$ are linear independent.
A3. The Strict Complementary Slackness condition (SCS)
If $\pi_i{}^\circ > 0$ when $g_i(x^\circ, y^\circ) = 0$; then it follows the results given by Fiacco (1976, Theorem 2.1) and Tobin (1986, Theorem 3.1):

R1. For y in a neighborhood of y°, there exists a unique once continuously differentiable vector function $w(y) = [x(y)^t, \pi(y)^t, \mu(y)^t]^t$ such that $[x(y^\circ)^t, \pi\,(y^\circ)^t, \mu(y^\circ)^t]^t = [x^{\circ^t}, \pi^{\circ^t}, \mu^{\circ^t}]^t$ where $x(y)$ is a locally unique minimum of problem (2) with associated unique Lagrange multiplier $\pi(y)$ and $\mu(y)$; also for y in the neighborhood of y° the SCS and LI of binding constraints hold at $x(y)$.

R2. The first-order approximation of solution to problem (2) at (x°, y°) can be given by (Tobin, 1986, Corollary 3.1)

$$w(y) = w(y^\circ) + \nabla w(y^\circ)(y - y^\circ) \tag{5}$$

where $\nabla w(y^\circ) = [J_w]^{-1}[-J_y]$, J_w and J_y respectively represent the Jacobian of the following system with respect to w and y evaluated at $(w(y^\circ), y^\circ)$.

$$\begin{aligned}
&f(x,y) - \nabla g(x,y)^t\pi - \nabla h(x,y)^t\mu = 0\\
&\pi^t g(x,y) = 0\\
&h(x,y) = 0
\end{aligned}$$

Regarding the sensitivity analysis for equilibrium network flows, Tobin and Friesz proposed a restricted equilibrium problem for which the path flows are positive. Analogous to Eq. (5), the first-order approximation of equilibrium link flow x at (x°, y°) is expressed as

$$\tilde{x}(y) = x^\circ + \nabla x(y^\circ)(y - y^\circ) \tag{6}$$

where the details of $\nabla x(y)$ can be found in Tobin and Friesz (1988, p. 248). For y in the neighborhood of y°, if x° is the solution of problem (2), BLPP in (1)–(2) can be approximated by

$$\begin{aligned}
\textbf{Min}\quad & F(x(y), y)\\
\text{s.t.}\quad & y \in S
\end{aligned} \tag{7}$$

and $x(y)$ is replaced by $\tilde{x}(y)$ as given in Eq. (6).

Let $\widetilde{F}(y) = F(x(y), y)$, A be the coefficient matrix for y and B the constant vector, the feasible set S for y in BLPP (7) can be given by

$$S = Ay \leqslant B$$

Therefore the BLPP (7) can be regarded as a single level differentiable constrained optimization in the following way:

$$\mathop{\mathbf{Min}}_{y} \quad \widetilde{F}(y) \tag{8}$$
$$\text{s.t.} \quad Ay \leqslant B$$

3. Problem formulation

The following notation is used for CNDP formulation:

- E the set of links in the network
- W the set of OD pairs
- T the vector of fixed OD pair demands, $T = [T_w]\ \forall w \in W$
- R_w the set of paths between OD pair w
- h the vector of path flows, $h = [h_p^w]\ \forall p \in R_w,\ \forall w \in W$
- x the vector of equilibrium link flows, $x = [x_a]\ \forall a \in E$
- y the vector of link capacity expansions, $y = [y_a]\ \forall a \in E$
- u the vector of upper bound for link capacity expansions, $u = [u_a]\ \forall a \in E$
- θ the conversion factor from investment cost to travel times
- $c(x,y)$ the vector of link travel times, $c(x,y) = [c_a(x_a, y_a)]\ \forall a \in E$
- $G(y)$ the vector of investment costs, $G(y) = [G_a(y_a)]\ \forall a \in E$
- Δ the link-path incidence matrix, $\Delta = [\delta_{ap}^w]\ \forall a \in E, p \in R_w, w \in W$, where $\delta_{ap}^w = 1$ if path p between OD pair w uses link a, and $\delta_{ap}^w = 0$ otherwise
- Λ the OD-path incidence matrix

Thus CNDP can be formulated in terms of the BLPP as follows:

$$\mathop{\mathbf{Min}}_{y} \quad F(x,y) = \sum_{a\in E}(c_a(x_a(y), y_a)x_a(y) + \theta G_a(y_a)) \tag{9}$$
$$\text{s.t.} \quad 0 \leqslant y_a \leqslant u_a \quad \forall a \in E$$

where $x(y)$ is the equilibrium flow defined by the following equivalent minimization problem:

$$\mathop{\mathbf{Min}}_{x} \quad z = \sum_{a\in E} \int_0^{x_a} c_a(w, y_a)\,\mathrm{d}w \tag{10}$$
$$\text{s.t.} \quad \sum_{p\in R_w} h_p^w = T_w \quad \forall w \in W$$
$$x_a = \sum_{w\in W}\sum_{p\in R_w} h_p^w \delta_{ap}^w \quad \forall a \in E$$
$$h_p^w \geqslant 0 \quad \forall p \in R_w,\ w \in W$$

Analogous to problem (7), the CNDP can be re-expressed as a single level differentiable optimization problem with simple range constraint for y when the first-order differentiable derivatives of $x(y)$ are available and $x(y)$ can be replaced by the linear form $\tilde{x}(y)$.

$$\begin{aligned} &\mathop{\mathbf{Min}}_{y} \quad \widetilde{F}(y) \\ &\text{s.t.} \quad 0 \leqslant y_a \leqslant u_a \quad \forall a \in E \end{aligned} \tag{11}$$

According to the assumptions A1–A3, and the corresponding results R1 and R2, a local optimum for CNDP (11) can be found by the descent approach. A descent approach via the implementation of gradient-based methods is adopted in finding a search direction d at a local solution y° along which the value of objective function $\widetilde{F}(y)$ decreases while the equilibrium flow pattern is approximated by (6). Additionally, with the linear independence constraint qualification satisfied in problem (11), the Karush–Kuhn–Tucker points are also identified (see Bazaraa et al., 1993, pp. 192–193).

4. Solution algorithms

Since the CNDP is non-convex, only local optimal solutions can be found. In this section new algorithms based on the gradients for CNDP (11) are presented. The first one is based on Rosen's gradient projection method (Bazaraa et al., 1993) for which the Karush–Kuhn–Tucker solution of (11) can be identified. Further gradient-based methods like Fletcher and Reeves' conjugate gradient method, quasi-Newton method of Broyden, Fletcher, Goldfarb and Shanno (BFGS) and PARalell TANgents (PARTAN; Luenberger, 1989) version of Rosen's gradient projection method are also taken into account. Detailed descriptions of these gradient-based solution algorithms can be found in Bazaraa et al. (1993), Avriel (1976), Polak (1997) and Norcedal and Wright (1999).

4.1. Rosen's gradient projection method

Following the application steps of gradient projection method presented by Chiou (1999, p. 282), the implementation steps for conducting Rosen's gradient projection method to solve CNDP (11) are as follows:

Step 1. Given initial values of y^0, start with step $k = 0$. Let I_0 be the set of indices of binding constraints.

Step 2. At step k, with fixed y^k the equilibrium flow x^k in problem (10) can be solved when the absolute difference value of z between successive iterations is within the stopping criterion $\tilde{\varepsilon}$.

Step 3. Calculate $\nabla x(y^k)$ and obtain the gradient $\nabla\widetilde{F}(y^k)$.

Step 4. Calculate the projection matrix and decide the descent direction d^k. If $d^k = 0$, go to Step 5. Otherwise, find α_{opt} which minimizes $\widetilde{F}(y^k + \alpha d^k)$ subject to $0 \leqslant \alpha \leqslant \alpha_{\max}$ where $\alpha_{\max}$ can be found in Chiou (1999, p. 282). Let $y^{k+1} = y^k + \alpha_{\mathrm{opt}} d^k$, set $k \leftarrow k + 1$ and return to Step 2.

Step 5. Check Lagrange multiplier μ vector. If $\mu \geqslant 0$ then y^k is Karush–Kuhn–Tucker point and stop. Otherwise find μ_i the most negative component of vector μ and set $I_0 \leftarrow I_0 - \{i\}$. Set $k \leftarrow k + 1$ and return to Step 2.

4.2. Conjugate gradient projection method

According to Fletcher and Reeves' conjugate gradient method, the implementation steps for solving CNDP (11) with simple range constraints are given as follows:

Step 1. Given initial values of y^0, start with step $k = 0$. Find x^0 with fixed y^0 for problem (10) when satisfying the stopping criterion $\tilde{\varepsilon}$. Calculate $\nabla x(y^0)$ and $\nabla \widetilde{F}(y^0)$. Let $d^0 = -\nabla \widetilde{F}(y^0)^{\mathrm{t}}$.

Step 2. If $k = 0$ go to Step 3. Otherwise, with fixed y^k find x^k satisfying the stopping criterion $\tilde{\varepsilon}$. Calculate $\nabla x(y^k)$ and $\nabla \widetilde{F}(y^k)$.

Step 3. At each step k: find α_{opt} minimizing $\widetilde{F}(y^k + \alpha d^k)$ subject to $0 \leqslant \alpha \leqslant \alpha_{\max}$. Set projection on the feasible set $0 \leqslant y^k \leqslant u$ as $y^{k+1} = [y^k + \alpha_{\mathrm{opt}} d^k]^U$ and $d^{k+1} = -\nabla \widetilde{F}(y^{k+1})^{\mathrm{t}} + w^k d^k$ with $w^k = \frac{\|\nabla \widetilde{F}(y^{k+1})\|^2}{\|\nabla \widetilde{F}(y^k)t\|^2}$ where

$$[\bullet]^U = \begin{bmatrix} \min\{\max\{0, y_1^k + \alpha_{\mathrm{opt}} d_1^k\}, u_1\} \\ \vdots \\ \min\{\max\{0, y_n^k + \alpha_{\mathrm{opt}} d_n^k\}, u_n\} \end{bmatrix}$$

d_j^k and y_j^k respectively denote the jth component of d^k and y^k, and $0 \leqslant y_j^k \leqslant u_j \; \forall j \in E$.

Step 4. For a predetermined stopping tolerance $\varepsilon > 0$, if $\|\widetilde{F}(y^{k+1}) - \widetilde{F}(y^k)\| < \varepsilon$ then stop and y^{k+1} is the solution for problem (11). Otherwise, set $k \leftarrow k + 1$ and return to Step 2.

4.3. Quasi-Newton projection method: algorithm of BFGS

According to quasi-Newton method of BFGS, the implementation steps for solving CNDP (11) with simple range constraints for the link capacity expansions are given as follows:

Step 1. Given initial values of y^0, start with step $k = 0$ and set the positive definite as $H_0 = I$.

Step 2. With fixed y^k find x^k satisfying stopping criterion $\tilde{\varepsilon}$ for problem (10). Calculate $\nabla x(y^k)$ and $\nabla \widetilde{F}(y^k)$.

Step 3. At each step k, find the search direction $d^k = -H_k \nabla \widetilde{F}(y^k)^{\mathrm{t}}$. Also,

Step 3-1. find the projection $y^{k+1} = [y^k + \alpha_{\mathrm{opt}} d^k]^U$ on the feasible set $0 \leqslant y^k \leqslant u$, where $y^k + \alpha_{\mathrm{opt}} d^k$ is the minimum of $\widetilde{F}(y^k + \alpha d^k)$ subject to $0 \leqslant \alpha \leqslant \alpha_{\max}$, and

Step 3-2. set $p_k = y^{k+1} - y^k$ and calculate $q_k^t = \nabla\widetilde{F}(y^{k+1}) - \nabla\widetilde{F}(y^k)$ then compute $H_{k+1} = H_k - \frac{H_k p_k p_k^t H_k}{p_k^t H_k p_k} + \frac{q_k q_k^t}{q_k^t p_k}$.

Step 4. For a predetermined stopping tolerance $\varepsilon > 0$, if $\|\widetilde{F}(y^{k+1}) - \widetilde{F}(y^k)\| < \varepsilon$ then stop and y^{k+1} is the solution. Otherwise, set $k \leftarrow k + 1$ and return to Step 2.

4.4. Rosen's gradient projection method with PARTAN

Applying the techniques of parallel tangents (PARTAN) (Luenberger, 1989) to Rosen's gradient projection method given in Section 4.1, we conduct the search procedures for CNDP (11) as given as follows:

Steps 1–3 and 5 are the same as given in Section 4.1.

Step 4. If $k = 0$ go to Step 4 in Section 4.1; otherwise go to Step 4′.

Step 4′. Calculate the projection matrix and decide the descent direction d^k. If $d^k = 0$, go to Step 5. Otherwise find α_{opt} and set $v^{k+1} = y^k + \alpha_{opt} d^k$.

Step 4′-1. Conduct PARTAN line search and find λ^* such that

$$\widetilde{F}(y^{k+1} = y^{k-1} + \lambda^*(v^{k+1} - y^{k-1})) = \mathop{\mathbf{Min}}_{0 \leqslant \lambda \leqslant \lambda_{\max}} \{\widetilde{F}(y^{k-1} + \lambda(v^{k+1} - y^{k-1}))\}$$

where $\lambda_{\max} = \min\{\underline{\lambda}, \bar{\lambda}\}$,

$$\underline{\lambda} = \min\left\{\frac{y_j^k}{|d_j^k|} \text{ if } d_j^k < 0\ \forall j \in E\right\} \quad \text{and} \quad \bar{\lambda} = \min\left\{\frac{u_j - y_j^k}{d_j^k} \text{ if } d_j^k > 0\ \forall j \in E\right\}.$$

Step 4′-2. Set $k \leftarrow k + 1$ and return to Step 2.

Corollary 1 (via the implementation of gradient-based method). *Let y° locally solve CNDP (11) and the range constraints for link capacity expansions satisfy the linear independence constraint qualification, then y° is a Karush–Kuhn–Tucker point for (11). Furthermore, following the implementation of gradient-based methods in determining the feasible direction at current y^k, which are given in Sections* 4.1–4.4, *this search procedure will terminate only at a KKT point, which is regarded as an approximate solution algorithm for the single level differentiable equilibrium network design problem (11).*

5. Computational results

Numerical experiments are widely conducted in this section for comparing the results yielded by the proposed algorithms with those obtained by other alternatives on three kinds of example networks. Following the numerical illustrations conducted by Meng et al. (2001, pp. 95–97), the 16-link network and Sioux Falls city aggregated network are chosen for the first two demonstrations. The detailed descriptions for alternatives in solving CNDP can be referred to Meng et al. (2001), Friesz et al. (1992) and Suwansirikul et al. (1987). Additionally, the sensitivity analysis based (SAB) algorithm proposed by Yang and Yagar (1995) is considered as an alternative to solve CNDP. A good lower bound of CNDP can be found by solving the network design problem confined to the system optimal network flows (SO), where in the problem (10) the marginal travel time cost function is used instead. Also, following the implementations conducted by Harker and Friesz (1984) and Suwansirikul et al., an inexact solution produced by the iterative optimization-assignment (IOA) method is calculated as an upper bound for CNDP. In this section four variants of gradient-based algorithms proposed for solving CNDP (11): Gradient Projection (GP) method, Conjugate Gradient projection method (CG), Quasi-NEWton projection method (QNEW) and PARTAN version of gradient projection method (PT), are considered for numerical illustrations. In the following numerical experiments the stopping tolerance in the objective function value for problem (11) is set $\varepsilon = 0.01$, and the stopping criterion in performing equilibrium traffic assignment (10) is set $\tilde{\varepsilon} = 0.005$.

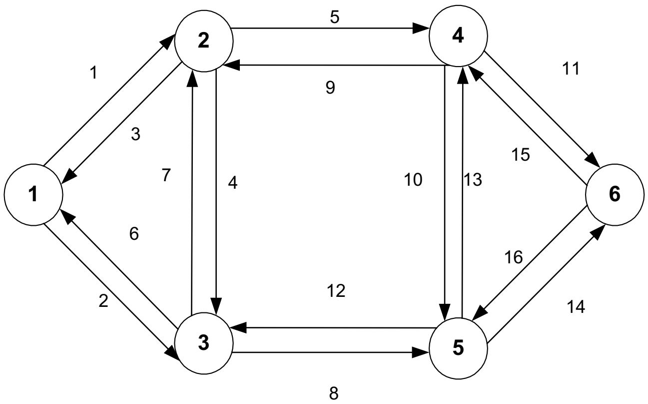

Fig. 1. 16-link network.

Table 1
Abbreviations of heuristics to CNDP

Abbreviation	Names of heuristics	Sources
IOA	Iterative Optimization-Assignment algorithm	Allsop (1974)
HJ	Hooke-Jeeves algorithm	Abdulaal and LeBlanc (1979)
EDO	Equilibrium Decomposed Optimization	Suwansirikul et al. (1987)
MINOS	Modular In-core NOnlinear System	Suwansirikul et al. (1987)
SA	Simulated Annealing algorithm	Friesz et al. (1992)
SAB	Sensitivity Analysis Based algorithm	Yang and Yagar (1995)
AL	Augmented Lagrangian algorithm	Meng et al. (2001)
GP	Gradient Projection method	This paper
CG	Conjugate Gradient projection method	This paper
QNEW	Qusai-NEWton projection method	This paper
PT	PARATAN version of gradient projection method	This paper

Table 2
Comparison of results for algorithms for case I of 16-link network

Variable/algorithm	MINOS	HJ	EDO	IOA	SA	AL
y_1	0	0	0	0	0	0
y_2	0	0	0	0	0	0
y_3	0	1.2	0.13	0	0	0.0062
y_4	0	0	0	0	0	0
y_5	0	0	0	0	0	0
y_6	6.58	3.00	6.26	6.95	3.1639	5.2631
y_7	0	0	0	0	0	0.0032
y_8	0	0	0	0	0	0
y_9	0	0	0	0	0	0
y_{10}	0	0	0	0	0	0
y_{11}	0	0	0	0	0	0.0064
y_{12}	0	0	0	0	0	0
y_{13}	0	0	0	0	0	0
y_{14}	0	0	0	0	0	0
y_{15}	7.01	3.00	0.13	5.66	0	0.71701
y_{16}	0.22	2.80	6.26	1.79	6.7240	6.7561
F	211.25	215.08	201.84	210.86	198.10378	202.9913
#	–	54	10	9	18300	2700

Note: where F denotes the value of objective function for CNDP, # denotes the number of equilibrium assignment problems solved. The initials of y were set 0.0 and the upper bound for y was set 10.0 for EDO, IOA, SA and AL.

Table 3
Comparison of results for algorithms for case I of 16-link network (continued)

Variable/algorithm	SAB	GP	CG	QNEW	PT	SO
y_1	0	0	0	0	0	0
y_2	0	0	0	0	0	0
y_3	0	0	0	0	0	0
y_4	0	0	0	0	0	0
y_5	0	0	0	0	0	0
y_6	5.8352	5.8302	6.1989	6.0021	5.9502	5.9979
y_7	0	0	0	0	0	0
y_8	0	0	0	0	0	0
y_9	0	0	0	0	0	0
y_{10}	0	0	0	0	0	0
y_{11}	0	0	0	0	0	0
y_{12}	0	0	0	0	0	0
y_{13}	0	0	0	0	0	0
y_{14}	0	0	0	0	0	0
y_{15}	0.9739	0.87	0.0849	0.1846	0.5798	0.1449
y_{16}	6.1762	6.1090	7.5888	7.5438	7.1064	7.5443
F	204.7	202.24	199.27	198.68	200.60	193.39
#	6	14	7	12	16	25

Note: The initials value of y were set 0.0 and the upper bound 10.0.

Table 4
Comparison of results for algorithms for case II of 16-link network

Variable/algorithm	MINOS	HJ	EDO	IOA	SA	AL
y_1	0	0	0	0	0	0
y_2	4.61	5.40	4.88	4.55	0	4.6153
y_3	9.86	8.18	8.59	10.65	10.1740	9.8804
y_4	0	0	0	0	0	0
y_5	0	0	0	0	0	0
y_6	7.71	8.10	7.48	6.43	5.7769	7.5995
y_7	0	0	0.26	0	0	0.0016
y_8	0.59	0.90	0.85	0.59	0	0.6001
y_9	0	0	0	0	0	0.001
y_{10}	0	0	0	0	0	0
y_{11}	0	0	0	0	0	0
y_{12}	0	0	0	0	0	0.1130
y_{13}	0	0	0	0	0	0
y_{14}	1.32	3.90	1.54	1.32	0	1.3184
y_{15}	19.14	8.10	0.26	19.36	0	2.7265
y_{16}	0.85	8.40	12.52	0.78	17.2786	17.5774
F	557.14	557.22	540.74	556.61	528.497	532.71
#		134	12	13	24300	4000

Note: The initials of y were set 0.0 and the upper bound for y was set 20.0 for EDO, IOA, SA and AL.

Table 5
Comparison of results for algorithms for case II of 16-link network (continued)

Variable/algorithm	SAB	GP	CG	QNEW	PT	SO
y_1	0.0189	0.1013	0.1022	0.0916	0.101	0.1165
y_2	2.2246	2.1818	2.1796	2.1521	2.1801	2.1467
y_3	9.3394	9.3423	9.3425	9.1408	9.3339	9.3447
y_4	0	0	0	0	0	0
y_5	0	0	0	0	0	0
y_6	9.0466	9.0443	9.0441	8.8503	9.0361	9.0424
y_7	0	0	0	0	0	0
y_8	0.0175	0.008	0.0074	0.0114	0.0079	0
y_9	0	0	0	0	0	0
y_{10}	0	0	0	0	0	0
y_{11}	0	0	0	0	0	0
y_{12}	0.0816	0.0375	0.0358	0.0377	0	
y_{13}	0	0	0	0	0	0
y_{14}	0.0198	0.0089	0.0083	0.0129	0.0089	0
y_{15}	2.1429	1.9433	1.9483	1.9706	1.9429	1.7995
y_{16}	18.9835	18.9859	18.986	18.575	18.9687	18.9875
F	536.084	534.017	534.109	534.08	534.02	512.013
#	45	31	16	11	33	29

Note: The initials value of y were 0.0 and the upper bound 20.0.

The first example network shown in Fig. 1 consists of 6 nodes and 16 links, where two OD pairs from nodes 1 to 6 with demand T and from nodes 6 to 1 with demand $2T$ are considered. Two cases of travel demand levels are used for illustration where $T = 5.0$ and 10.0. The travel time and investment cost functions used in problems 9 and 10 are adopted from Suwansirikul et al. (1987, p. 259) together with the details of data input for each link. The abbreviations for test heuristics are summarized in Table 1. The comparison results for cases I and II are given in Tables 2–5 respectively where the results of Tables 2 and 4 are summarized from the previous work (Meng et al., 2001, p. 98; Suwansirikul et al., 1987, p. 259). As it seen from Table 3, the gradient-based algorithms yielded similar results as those did EDO, SA and AL, but took much less computational efforts in solving equilibrium traffic assignment problems. Among the gradient-based algo-

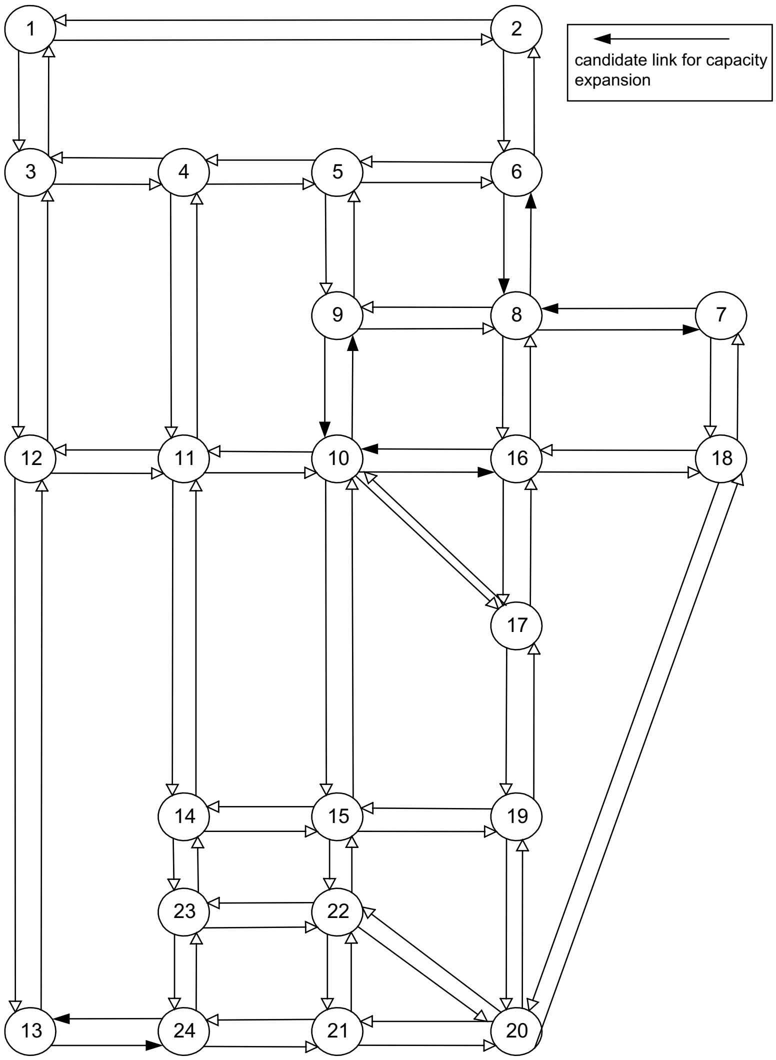

Fig. 2. Sioux Falls network.

rithms, CG and QNEW achieved good solutions as those did SA, which was recognized as a global optimum for CNDP. The network design problem with system optimal network flows was also solved. As it seen from Table 3, the relative difference percentages with respect to SO in the objective function value for the gradient-based solution methods were within 6. For case II example, as it shown in Tables 4 and 5, again the results produced by gradient-based methods were close to those did by the globally optimal method SA with much less computational efforts in solving equilibrium assignment problems. The solutions achieved by the gradient-based algorithms were substantially better than those did by MINOS, HJ, EDO and IOA and slightly better

Table 6
Comparison of results for algorithms on Sioux Falls city network

Variable/algorithm	HJ	HJ	EDO	SA	AL	IOA	SO
Initial value of y_a	2.0	1.0	12.5	6.25	12.5	12.5	12.5
y_{16}	4.8	3.8	4.59	5.38	5.5728	4.6875	4.9274
y_{17}	1.2	3.6	1.52	2.26	1.6343	3.9063	5.0358
y_{19}	4.8	3.8	5.45	5.50	5.6228	1.2695	2.6211
y_{20}	0.8	2.4	2.33	2.01	1.6443	1.6599	1.9953
y_{25}	2.0	2.8	1.27	2.64	3.1437	2.3331	3.4475
y_{26}	2.6	1.4	2.33	2.47	3.2837	2.3438	3.1620
y_{29}	4.8	3.2	0.41	4.54	7.6519	5.5651	8.1272
y_{39}	4.4	4.0	4.59	4.45	3.8035	4.6862	6.7257
y_{48}	4.8	4.0	2.71	4.21	7.3820	5.4688	4.2947
y_{74}	4.4	4.0	2.71	4.67	3.6935	6.2500	5.0114
F	81.25	81.77	83.47	80.87	81.752	87.34	77.763
#	58	108	12	3900	2700	31	24

Note: The upper bound for y was 25.0 for EDO, SA, AL, IOA and SO.

Table 7
Comparison of results for algorithms on Sioux Falls city network (continued)

	SAB	GP	CG	QNEW	PT	SAB	GP	CG	QNEW	PT
Initial value of y_a	12.5	12.5	12.5	12.5	12.5	6.25	6.25	6.25	6.25	6.25
y_{16}	5.7392	4.8693	4.7691	5.3052	5.0237	3.9131	5.4277	5.0769	4.9776	4.7921
y_{17}	5.7182	4.8941	4.8605	5.0541	5.2158	3.8059	5.3235	5.1244	5.0287	5.0827
y_{19}	4.9591	1.8694	3.0706	2.4415	1.8298	2.6608	1.6825	1.5291	1.9412	2.0046
y_{20}	4.9612	1.5279	2.6836	2.5442	1.5747	2.7253	1.6761	1.5027	2.1617	1.3947
y_{25}	5.5066	2.7168	2.8397	3.9328	2.7947	2.9609	2.8361	2.5706	2.6333	2.6430
y_{26}	5.5199	2.7102	2.9754	4.0927	2.6639	2.9906	2.7288	2.7372	2.7923	2.8031
y_{29}	5.8024	6.2455	5.6823	4.3454	6.1879	4.2311	5.7501	4.8474	5.7462	5.3823
y_{39}	5.5902	5.0335	4.2726	5.2427	4.9624	4.4014	4.9992	4.6927	5.6519	5.4699
y_{48}	5.8439	3.7597	4.4026	4.7686	4.0674	4.7756	4.4308	4.7897	4.5738	5.0102
y_{74}	5.8662	3.5665	5.5183	4.0239	3.9199	4.8253	4.3081	4.1864	4.1747	4.4771
F	84.21	82.71	82.53	83.07	82.53	84.34	82.57	82.80	83.08	82.72
#	11	9	6	4	7	10	10	7	5	9

Note: The upper bound for y was 25.0.

than that did by SAB. The relative difference percentages with respect to SO in the value of objective function for the gradient-based method solutions were within 4.4.

The second example network is chosen from a 'benchmark' problem of aggregated Sioux Falls city network as shown in Fig. 2. The travel time and link investment cost functions used in problems 9 and 10 are adopted from Suwansirikul et al. (1987, pp. 261–262) where the convex investment function form is adopted from Abdulaal and LeBlanc (1979, 28), together with the data input details. Comparison results for the gradient-based algorithms and the other alternatives are shown in Tables 6 and 7. As it has been mentioned in Suwansirikul et al., the multiple local

Table 8
Comparison of results (∗) for algorithms on Sioux Falls network with scaling factors

Scalar/algorithm	SAB	GP	CG	QNEW	PT	EDO	IOA	SO
0.80	51.76	48.38	48.78	48.84	48.81	49.51	53.58	45.85
	(14)	(10)	(3)	(4)	(9)	(7)	(28)	(17)
0.85	57.57	55.85	55.47	55.68	55.51	56.17	58.18	52.04
	(11)	(11)	(6)	(5)	(9)	(10)	(28)	(4)
0.90	66.31	63.65	63.61	63.85	63.66	64.95	67.31	58.85
	(18)	(8)	(8)	(5)	(6)	(7)	(31)	(45)
0.95	76.27	72.31	72.52	71.78	71.93	72.96	76.48	67.43
	(11)	(9)	(7)	(4)	(8)	(8)	(19)	(11)
1.00	84.21	82.71	82.53	83.07	82.53	83.47	87.34	77.76
	(11)	(9)	(6)	(4)	(7)	(12)	(31)	(24)
1.05	95.53	95.39	95.89	93.73	95.42	95.97	105.04	88.57
	(9)	(11)	(8)	(5)	(8)	(7)	(31)	(29)
1.10	107.58	107.65	107.86	108.64	107.58	108.93	113.71	102.92
	(11)	(6)	(10)	(4)	(6)	(8)	(31)	(29)
1.15	124.49	123.97	124.67	124.74	122.64	127.76	131.81	117.43
	(16)	(5)	(12)	(11)	(11)	(8)	(25)	(6)
1.20	144.86	141.53	141.04	141.62	142.27	149.39	150.99	134.89
	(9)	(11)	(10)	(7)	(9)	(12)	(31)	(24)
1.25	170.74	164.12	162.16	163.62	162.05	165.12	179.89	151.88
	(14)	(9)	(7)	(5)	(7)	(13)	(25)	(17)
1.30	189.43	188.16	182.93	184.97	183.35	190.48	208.44	178.19
	(10)	(5)	(7)	(5)	(9)	(12)	(31)	(11)
1.35	221.92	211.93	209.63	214.94	210.54	215.64	232.09	204.55
	(11)	(7)	(9)	(5)	(7)	(15)	(37)	(20)
1.40	247.8	246.04	242.95	242.74	241.08	253.19	279.38	232.31
	(15)	(9)	(6)	(5)	(7)	(17)	(16)	(15)
1.45	275.67	278.43	276.87	278.17	276.51	284.39	328.61	267.93
	(12)	(11)	(10)	(7)	(9)	(18)	(22)	(36)
1.50	336.05	326.31	318.25	309.55	320.71	335.52	377.83	306.94
	(9)	(9)	(8)	(4)	(6)	(19)	(40)	(23)
1.55	387.16	361.79	361.33	363.41	364.47	393.62	435.91	359.47
	(11)	(12)	(9)	(7)	(10)	(19)	(37)	(37)
1.60	452.01	433.64	408.45	409.04	431.11	427.56	475.08	406.06
	(14)	(9)	(9)	(9)	(11)	(19)	(40)	(46)

Note: ∗ measured in hours and the number in parentheses denotes the equilibrium assignment problems solved. The initials of y were set 12.5 and EDO was performed by bisection search.

optima exist due to the non-convexity of CNDP and evidently each method leads to a different solution to the CNDP and all the possible local optima are within the solution bound produced by IOA and SO methods. Two sets of initial values of link capacity expansion are used for testing the robustness of proposed methods to the initials. As it seen from Table 7, the difference in the objective function value between the two sets of initials of link capacity expansion was not significant for the proposed gradient-based methods. In comparison with the previous results generated by HJ, EDO and AL, it has been found that HJ and AL slightly outperformed than did the gradient-based algorithms while the computational efforts incurred by HJ and AL were much more intensive. Regarding the performance made by the gradient-based algorithms, the results produced by GP, CG and PT were slightly outperformed than those did by QNEW; however, QNEW took the least iterations in solving the equilibrium assignment problems. Furthermore, regarding the relative difference percentages with respect to SO, at the initials of $y = 12.5$, those did by GP, CG, QNEW and PT were about 6.36, 6.13, 6.82 and 6.13; and at the initials of $y = 6.25$, they were about 6.18, 6.48, 6.84 and 6.37 respectively.

To investigate the capability of the proposed methods in solving CNDP for congested networks, extensive numerical tests have been considered by scaling the travel demands and the corresponding results are shown in Tables 8 and 9. The proposed gradient-based algorithms achieved good solutions with consistently less computational efforts in solving equilibrium assignment problems as compared to those did SAB and EDO when the travel demands became increased. Furthermore, as shown in Table 9, the proposed gradient-based algorithms outperformed other alternatives by giving much less average for 17 scaled travel demand tests.

Table 9
Comparison of results in percentage (relative to SO) for algorithms on Sioux Falls network

Scalar/algorithm	SAB	GP	CG	QNEW	PT	EDO	IOA
0.80	12.89	5.52	6.39	6.52	6.46	7.98	16.86
0.85	10.63	7.32	6.59	6.99	6.67	7.94	11.80
0.90	12.68	8.16	8.09	8.50	8.17	10.37	14.38
0.95	13.11	7.24	7.55	6.45	6.67	8.20	13.42
1.00	8.29	6.37	6.13	6.83	6.13	7.34	12.32
1.05	7.86	7.70	8.26	5.83	7.73	8.35	18.60
1.10	4.53	4.60	4.80	5.56	4.53	5.84	10.48
1.15	6.01	5.57	6.17	6.22	4.44	8.80	12.25
1.20	7.39	4.92	4.56	4.99	5.47	10.75	11.94
1.25	12.42	8.06	6.77	7.73	6.70	8.72	18.44
1.30	6.31	5.60	2.66	3.80	2.90	6.90	16.98
1.35	8.49	3.61	2.48	5.08	2.93	5.42	13.46
1.40	6.67	5.91	4.58	4.49	3.78	8.99	20.26
1.45	2.89	3.92	3.34	3.82	3.20	6.14	22.65
1.50	9.48	6.31	3.68	0.85	4.49	9.31	23.10
1.55	7.70	0.65	0.52	1.10	1.39	9.50	21.26
1.60	11.32	6.79	0.59	0.73	6.17	5.29	17.00
Average	8.75	5.78	4.89	5.03	5.17	7.99	16.19

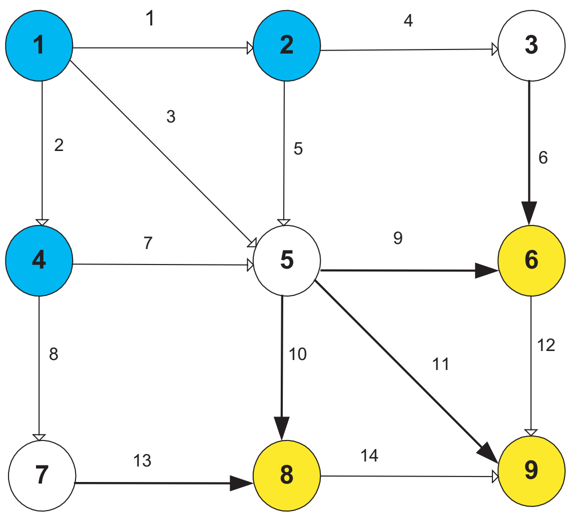

Fig. 3. 9-node grid network.

Table 10
OD trip matrix for 9-node grid network (in vehicles/per hour)

	6	8	9
1	120	150	100
2	130	200	90
4	80	180	110

Table 11
Data for 9-node grid network

$c_a = A_a + B_a(\frac{x_a}{K_a+y_a})^4, F(x,y) = \sum_{a\in E}(c_a(x_a,y_a)x_a + \theta_a y_a)$

Link a	A_a	B_a	K_a	θ_a
1	2.00	0.300	280.0	–
2	1.50	0.225	290.0	–
3	3.00	0.450	280.0	–
4	1.00	0.150	280.0	–
5	1.00	0.150	600.0	–
6	2.00	0.300	300.0	1.5
7	2.00	0.300	500.0	–
8	1.00	0.150	400.0	–
9	1.50	0.225	500.0	1.0
10	1.00	0.150	700.0	1.0
11	2.00	0.300	250.0	1.5
12	1.00	0.150	300.0	–
13	1.00	0.150	350.0	1.0
14	1.00	0.150	220.0	–

Table 12
Comparison of results for algorithms on 9-node grid network

Variable/algorithm	SAB	GP	CG	QNEW	PT	EDO	SO
Initial value of y_a	0	0	0	0	0	0	0
y_6	0	0	0	0	0	0.0002	0.0002
y_9	0	0	0	0	0	0.0002	0.0002
y_{10}	0	0	0	0	0	0.0002	0.0002
y_{11}	0	0	0	0	0	0.0002	0.0002
y_{13}	128.087	128.103	137.445	137.445	127.793	137.445	0.0002
F (*)	4110.02	4110.02	4109.55	4109.55	4110.05	4109.56	3819.75
#	15	12	8	4	9	17	17
Initial value of y_a	250	250	250	250	250	250	250
y_6	0	0	0	0	0	0.0002	0.0002
y_9	0	0	0	0	0	0.0002	0.0002
y_{10}	0	0	0	0	0	0.0002	0.0002
y_{11}	0	0	0	0	0	0.0002	0.0002
y_{13}	128.827	137.445	130.564	130.564	127.79	137.445	0.0002
F (*)	4109.95	4109.55	4109.55	4109.55	4110.05	4109.56	3819.75
#	15	10	7	5	9	17	17

Note: * the objective function value is measured in hours and the upper bound on y was 500. The EDO was performed by the bisection search.

Table 13
Comparison of results (*) on 9-node grid network with scaling factors

Scalar/algorithm	SAB	GP	CG	QNEW	PT	EDO	SO
0.8	3043.18(14)	3042.91(7)	3042.91(5)	3042.91(4)	3043.27(6)	3042.91(12)	2973.22(15)
0.9	3549.36(14)	3549.02(8)	3549.02(5)	3549.02(4)	3549.45(9)	3549.02(16)	3394.02(18)
1.0	4109.95(15)	4109.55(10)	4109.55(7)	4109.55(5)	4110.05(9)	4109.56(17)	3819.75(17)
1.1	4634.97(13)	4632.38(11)	4632.38(9)	4632.38(6)	4632.38(10)	4666.00(17)	4256.44(17)
1.2	5137.64(15)	5135.28(11)	5135.29(9)	5135.29(4)	5135.29(9)	5195.64(19)	4694.92(16)
1.3	5700.01(14)	5703.21(9)	5697.22(8)	5697.22(4)	5697.22(8)	5731.73(16)	5163.53(16)
1.4	6240.73(13)	6244.02(10)	6237.98(9)	6237.98(5)	6237.98(7)	6269.12(14)	5664.13(17)
1.5	6858.20(14)	6781.51(11)	6781.51(9)	6781.51(5)	6781.51(8)	6844.09(15)	6228.01(15)

Note: * the objective function value is measured in hours and the number in parentheses denotes the equilibrium assignment problems solved. The initials of link capacity expansions were set 250.

Among the proposed gradient-based algorithms, CG and QNEW generated the lowest value in the average of the relative difference percentages, which obviously indicates a promising local optimum can be found via the gradient-based algorithms with reasonably computational efforts.

The third test example is a grid network as shown in Fig. 3, which contains 9 nodes and 14 links and is adapted from Yang et al. (2001) for simultaneous estimation for OD matrices and travel-cost coefficient in congested traffic networks. In this numerical test, it includes 9 OD pair travel

Table 14
Comparison of results in percentage (relative to SO) on 9-node grid network

Scalar/algorithm	SAB	GP	CG	QNEW	PT	EDO
0.8	2.35	2.34	2.34	2.34	2.36	2.34
0.9	4.58	4.57	4.57	4.57	4.58	4.57
1.0	7.6	7.59	7.59	7.59	7.6	7.59
1.1	8.89	8.83	8.83	8.83	8.83	9.62
1.2	9.43	9.38	9.38	9.38	9.38	10.67
1.3	10.39	10.45	10.34	10.34	10.34	11.00
1.4	10.18	10.24	10.13	10.13	10.13	10.68
1.5	10.12	8.89	8.89	8.89	8.89	9.89
Average	7.94	7.79	7.76	7.76	7.76	8.30

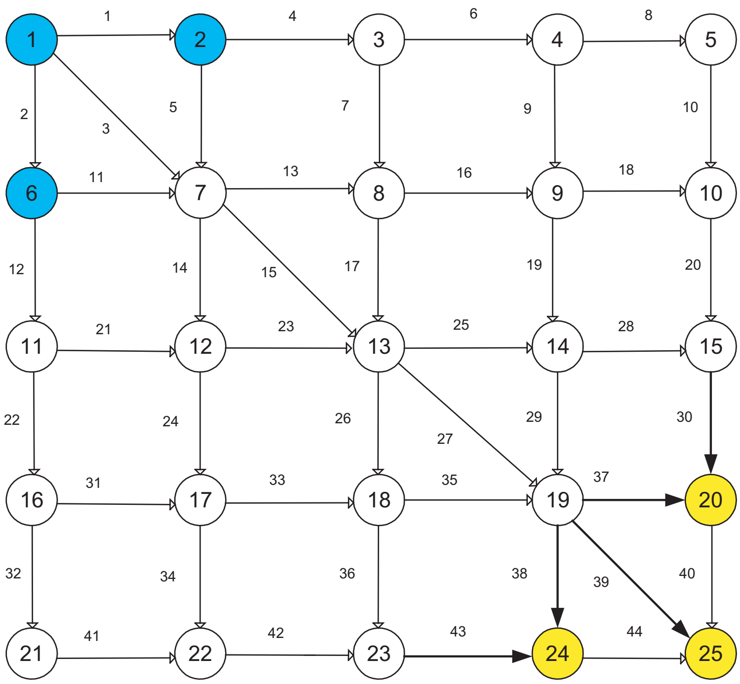

Fig. 4. 25-node grid network.

demands and 5 selected links for capacity expansions. Descriptions for OD travel demands and the objective function used in problems 9 and 10 together with the data input parameters for each link are given in Tables 10 and 11 respectively. As it seen from Table 12, all the solution heuristics yielded nearly identical results and the proposed gradient-based methods again took consistently less computational efforts in solving equilibrium assignment problems as compared to other alternatives. Furthermore, the computational results for scaling travel demand tests are summarized in

Table 15
OD trip matrix for 25-node grid network (in vehicles/per hour)

	20	24	25
1	120	150	100
2	130	200	90
6	80	180	110

Table 16
Data for 25-node grid network

$c_a = A_a + B_a(\frac{x_a}{K_a+y_a})^4, F(x,y) = \sum_{a\in E}(c_a(x_a,y_a)x_a + \theta_a y_a)$

Link a	A_a	B_a	K_a	d_a	Link a	A_a	B_a	K_a	θ_a
1	2.0	0.300	280	–	23	2.0	0.300	500	–
2	1.5	0.225	290	–	24	1.0	0.150	700	–
3	3.0	0.450	280	–	25	1.5	0.225	500	–
4	2.0	0.300	280	–	26	1.0	0.150	700	–
5	1.0	0.150	600	–	27	2.0	0.300	250	–
6	1.0	0.150	280	–	28	1.5	0.225	500	–
7	1.0	0.150	600	–	29	1.0	0.150	700	–
8	1.0	0.150	280	–	30	1.0	0.150	300	1.0
9	1.0	0.150	600	–	31	2.0	0.300	500	–
10	2.0	0.300	300	–	32	1.0	0.150	400	–
11	2.0	0.300	500	–	33	2.0	0.300	500	–
12	1.5	0.225	290	–	34	1.0	0.150	700	–
13	2.0	0.300	500	–	35	1.5	0.225	500	–
14	1.0	0.150	600	–	36	1.0	0.150	700	–
15	3.0	0.450	280	–	37	1.5	0.225	500	1.0
16	1.5	0.225	500	–	38	1.0	0.150	700	1.0
17	1.0	0.150	600	–	39	2.0	0.300	250	1.5
18	1.5	0.225	500	–	40	1.0	0.150	300	–
19	1.0	0.150	600	–	41	1.0	0.150	350	–
20	2.0	0.300	300	–	42	1.0	0.150	350	–
21	2.0	0.300	500	–	43	1.0	0.150	220	1.0
22	1.0	0.150	400	–	44	1.0	0.150	220	–

Table 13 where the relative difference percentages with respect to SO are given in Table 14. As it reported in Table 14, again the proposed gradient-based heuristics outperformed SAB and EDO in solving congested networks by yielding lower values of the relative difference percentages.

The final numerical test for demonstration is conducted on a larger grid network with 25 nodes and 44 links as shown in Fig. 4, which is expanded from the 9-node grid graph and served as a general testing for computational efficiency and superiority of the proposed methods. Descriptions for data input are given in Tables 15 and 16. Computational results for the 25-node graph are summarized in Tables 17–19. Again, the gradient-based algorithms gave nearly identical solutions to CNDP while the EDO heuristic converged to slightly higher values. Further tests are conducted by scaling the travel demands. As it seen from Tables 18 and 19, the proposed gradient-based algorithm again converged to the similar values as those did SAB and EDO with less com-

Table 17
Comparison of results for algorithms on 25-node grid network

Variable/algorithm	SAB	GP	CG	QNEW	PT	EDO	SO
Initial value of y_a	0	0	0	0	0	0	0
y_{30}	0	0	0	0	0	0.0007	0.0007
y_{37}	0	0	0	0	0	0.0007	0.0007
y_{38}	0	0	0	0	0	0.0007	0.0007
y_{39}	0	0	0	0	0	0.0007	0.0007
y_{43}	477.836	479.778	493.953	493.953	478.253	499.324	242.266
F (*)	10155.4	10155.2	10154.5	10154.5	10155.4	10168.2	9515.92
#	15	11	8	6	9	16	19
Initial value of y_a	250	250	250	250	250	250	250
y_{30}	0	0	0	0	0	0.0007	0.0007
y_{37}	0	0	0	0	0	0.0007	0.0007
y_{38}	0	0	0	0	0	0.0007	0.0007
y_{39}	0	0	0	0	0	0.0007	0.0007
y_{43}	482.478	484.505	481.145	481.145	479.83	499.324	242.266
F (*)	10154.9	10154.8	10155.0	10155.0	10155.2	10168.2	9515.92
#	15	10	9	5	7	17	18

Note: * the objective function value is measured in hours and the upper bound on y was 1500.

Table 18
Comparison of results (*) on 25-node grid network with scaling factors

Scalar/algorithm	SAB	GP	CG	QNEW	PT	EDO	SO
0.8	7590.19(11)	7590.06(8)	7590.19(5)	7590.19(4)	7590.36(7)	7589.71(17)	7434.38(16)
0.9	8888.89(11)	8888.77(8)	8888.91(5)	8888.91(5)	8889.07(8)	8888.25(18)	8473.65(16)
1.0	10154.9(15)	10154.8(10)	10155.0(9)	10155.0(5)	10155.2(7)	10168.2(17)	9515.92(18)
1.1	11452.0(14)	11440.9(9)	11441.1(8)	11441.1(6)	11441.3(8)	11452.6(18)	10588.8(17)
1.2	12643.1(15)	12632.5(9)	12626.3(7)	12626.3(5)	12626.3(8)	12642.9(16)	11665.3(15)
1.3	13833.7(13)	13831.6(7)	13831.7(6)	13831.7(5)	13832.1(8)	13834.1(15)	12771.4(16)
1.4	15195.7(14)	15185.6(8)	15185.7(7)	15185.7(4)	15186.4(7)	15198.3(17)	13920.2(16)
1.5	16371.6(16)	16354.6(7)	16354.6(5)	16354.6(4)	16354.8(6)	16371.2(16)	15078.5(15)

Note: * the objective function value is measured in hours and the number in parentheses denotes the equilibrium assignment problems solved. The initials of link capacity expansions were set 250.

putational efforts in solving equilibrium assignment problems. Furthermore, the results presented in Table 19 again showed the superiority of the proposed algorithms by yielding consistently lower values in the relative difference percentages than did SAB and EDO.

The implementations reported above have been conducted on Sun SPARC Ultra 30 with the operating system UNIX OS 5.7. Computer programs were coded in C++language and linked with the LEDA library (Mehlhorn et al., 1998).

Table 19
Comparison of results in percentage (relative to SO) on 25-node grid network

Scalar/algorithm	SAB	GP	CG	QNEW	PT	EDO
0.8	2.10	2.09	2.10	2.10	2.10	2.09
0.9	4.90	4.90	4.90	4.90	4.90	4.89
1.0	6.71	6.71	6.72	6.72	6.72	6.85
1.1	8.15	8.05	8.05	8.05	8.05	8.16
1.2	8.38	8.29	8.24	8.24	8.24	8.38
1.3	8.32	8.30	8.30	8.30	8.31	8.32
1.4	9.16	9.09	9.09	9.09	9.10	9.18
1.5	8.58	8.46	8.46	8.46	8.46	8.57
Average	7.04	6.99	6.98	6.98	6.99	7.06

6. Conclusions and discussions

In this paper a bilevel programming technique was used to formulate the continuous equilibrium network design problem. A single level optimization problem was considered when the first-order differentiable derivatives of the equilibrium link flow with respect to the link capacity expansions were available under restrict condition of the sensitivity analysis for perturbed problem. A descent approach via the implementation of gradient-based methods was adopted in finding a search direction at a local solution along which the value of objective function decreased while the equilibrium flow pattern was linearly approximated. Moreover, with the linear independence constraint qualification satisfied in problem (11), the Karush–Kuhn–Tucker points were also identified. In this paper we presented four variants of gradient-based algorithms in generally solving CNDP where the good local optima can be found. Numerical comparisons have been made with the previous on three kinds of general example networks.

Firstly, for the 16-link example network, the proposed methods produced similar results as those did the SA method, which can be regarded as a globally optimal solution to CNDP, with much less computational efforts. Furthermore, regarding the relative difference percentages with respect to SO in the value of objective function, the solutions yielded by the gradient-based methods were less significant than those did by the SAB method. Secondly, for the Sioux Falls realistic test network, a benchmark problem for testing heuristics for CNDP, the proposed heuristics again performed as good as did HJ and SA and demonstrated the robustness to the initials when solving CNDP. Furthermore, consider the capacity of the proposed methods in solving CNDP for congested networks, extensive numerical tests have been conducted and the proposed gradient-based methods achieved substantially better results with consistently less computational efforts in solving equilibrium assignment problems than those did the SAB and EDO when the travel demands became increased. Among the proposed gradient-based algorithms, CG and QNEW generated the lowest value in the average of the relative difference percentages, which obviously indicates that a promising local optimum can be found via the gradient-based algorithms with reasonably computational efforts. Thirdly, further tests have been also conducted on grid networks. Again the superiority and computational efficiency in solving CNDP by the proposed gradient-based algorithms were demonstrated. Good solutions were also obtained by implementing the proposed gradient-based algorithms for congested networks when lower values in the relative difference percentages with respect to SO have been found with less computational efforts than those did the SAB and EDO.

By giving the restriction of strong mathematical conditions on the lower level problem, e.g. the strict complementary slackness condition (SCS), the bilevel network problem can be re-expressed as a single level differentiable optimization problem. In this paper, we provided an efficient way to solve the bilevel network problem generally by adopting a descent approach via the implementation of gradient-based methods. In comparison with the results obtained by other alternatives in solving CNDP, which is a special case of the BLPP, the proposed methods have yielded substantially good performance in terms of the robustness to the initials and computational efficiency in solving equilibrium assignment problems for congested networks. Consider a general case of the BLPP 1, 2, the techniques of non-differentiable optimization need to be investigated furthermore for solving the BLPP without the SCS restriction. Solution methods such as the subgradient optimization approach and bundle method are being taken into account.

Acknowledgments

This author would like to express her gratitude to two anonymous referees for their thorough review on the earlier versions of the paper and thankful for their helpful comments. The work reported has been partially sponsored by UK Engineering and Physical Science Research Council (EPSRC) via grant GR/L81468 and partially funded by Taiwan National Science Council Grant (NSC 92-2416-H-259-005).

References

Allsop, R.E., 1974. Some possibilities for using traffic control to influence trip distribution and route choice. In: Transportation and Traffic Theory (Proceedings of the Sixth International Symposium on Transportation and Traffic Theory). Elsevier, New York, pp. 345–375.

Avriel, M., 1976. Nonlinear Programming: Analysis and Methods. Prentice-Hall, New Jersey.

Abdulaal, M., LeBlanc, L., 1979. Continuous equilibrium network design models. Transportation Research B 13, 19–32.

Bazaraa, M.S., Sherali, H.D., Shetty, C.M., 1993. Nonlinear Programming: Theory and Algorithms. John Wiley, New York.

Chiou, S.-W., 1999. Optimization of area traffic control for equilibrium network flows. Transportation Science 33 (3), 279–289.

Clarke, F., 1983. Optimization and Nonsmooth Analysis. John Wiley, New York.

Chen, Y., Florian, M., 1995. A coordinate descent method for the bi-level O–D matrix adjustment problem. International Transactions in Operational Research 2, 165–179.

Fiacco, A.V., 1976. Sensitivity analysis for nonlinear programming using penalty methods. Mathematical Programming 10, 287–311.

Falk, J., Liu, J., 1995. On bilevel programming, part I: general nonlinear cases. Mathematical Programming 70, 47–72.

Friesz, T.L., Cho, H.-J., Mehta, N.J., Tobin, R.L., Anandalingam, G., 1992. A simulated annealing approach to the network design problem with variational inequality constraints. Transportation Science 26 (1), 18–26.

Friesz, T.L., Tobin, R.L., Cho, H.-L., Metha, N.J., 1990. Sensitivity analysis based algorithms for mathematical programs with variational inequality constraints. Mathematical Programming 48, 265–284.

Gershwin, S.B., Tan, H.N., 1979. Hybrid optimization: optimal static traffic control constrained by drivers' route choice behavior. Laboratory for Information and Decision System Report LIDS-p-870, Massachusetts Institute of Technology.

Harker, P.T., Friesz, T.L., 1984. Bounding the solution of the continuous equilibrium network design problem. In: Proceedings of the Ninth International Symposium on Transportation and Traffic Theory. VNU Science Press, Delft, Netherlands.

Luenberger, D.G., 1989. Linear and Nonlinear Programming. Addison-Wesley, Reading, MA.

Luo, Z.-Q., Pang, J.-S., Ralph, D., 1996. Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge, UK.

Marcotte, P., 1983. Network optimization with continuous control parameters. Transportation Science 17, 181–197.

Marcotte, P., Marquis, G., 1992. Efficient implementation of heuristics for the continuous network design problem. Annals of Operations Research 34, 163–176.

Marcotte, P., Zhu, D.-L., 1996. Exact and inexact penalty methods for generalized bilevel programming problems. Mathematical Programming 74, 141–157.

Mehlhorn, K., Naher, S., Uhrig, C., 1998. The LEDA User Manual, Version 3.7. Max-Planck-Institut fur Informatik, 66213 Saarbrucken, Germany.

Meng, Q., Yang, H., Bell, M.G.H., 2001. An equivalent continuously differentiable model and a locally convergent algorithm for the continuous network design problem. Transportation Research B 35, 83–105.

Norcedal, J., Wright, S.J., 1999. Numerical Optimization. Springer-Verlag, New York.

Outrata, J., Kocvara, M., Zowe, J., 1998. Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results. Kluwer Academic Publishers, Boston, MA.

Polak, E., 1997. Optimization Algorithms and Consistent Approximations. Springer-Verlag, New York.

Suh, S., Kim, T.J., 1992. Solving nonlinear bilevel programming models of the equilibrium network design problem: a comparative review. Annals of Operations Research 34, 203–218.

Shimizu, K., Ishizuka, Y., Bard, J., 1997. Nondifferentiable and Two-level Mathematical Programming. Kluwer Academic Publishers, Boston, MA.

Suwansirikul, C., Friesz, T.L., Tobin, R.L., 1987. Equilibrium decomposed optimization: a heuristic for continuous equilibrium network design problem. Transportation Science 21, 254–263.

Tobin, T.L., 1986. Sensitivity analysis for variational inequality. Journal of Optimization Theory and Application 48, 191–204.

Tobin, T.L., Friesz, T.L., 1988. Sensitivity analysis for equilibrium network flows. Transportation Science 22, 242–250.

Yang, H., 2001. Special issue: Transportation bilevel programming problems: recent methodological advances. Transportation Research B 35.

Yang, H., Bell, M.G.H., 1998. Models and algorithms for road network design: a review and some new developments. Transport Reviews 18, 257–278.

Yang, H., Yagar, S., 1995. Traffic assignment and signal control in saturated road networks. Transportation Research A 29 (2), 125–139.

Yang, H., Meng, Q., Bell, M.G.H., 2001. Simultaneous estimation of the origin-destination matrices and travel-cost coefficient for congested networks in a stochastic user equilibrium. Transportation Science 35 (2), 107–123.