EJTIR

Issue 14(1), 2014
pp. 46-65
ISSN: 1567-7141
www.ejtir.tbm.tudelft.nl

Pareto-improving toll and subsidy scheme on transportation networks

F. Xiao[1],
School of Transportation & Logistics, Southwest Jiaotong University, P.R.China, National United Engineering Laboratory of Integrated and Intelligent Transportation

H.M. Zhang[2]
Department of Civil & Environmental Engineering, University of California, Davis

This paper presents an original study on the economics of a link-based Toll and Subsidy Scheme (TSS) on a general transportation network. Different from a traditional congestion pricing scheme, the combination of toll and subsidy is found to be able to serve more planning purposes simultaneously, such as efficiency, fairness, and public acceptance. We first demonstrate that on a one-origin or one-destination network, a pareto-improving, system-optimal and revenue-neutral TSS always exists and can be obtained by solving a set of linear equations. Recognizing that such a scheme may not always exist for a multi-origin network, we then define the maximum-revenue problem with pareto-improving constrains to find the maximum possible revenue collected by the toll and subsidy scheme with optimal arc flows and non-increasing origin-destination travel costs. We discover that the problem is actually the dual problem of a balanced transportation problem, which can thus be solved efficiently by existing algorithms. At the end of the paper, a numerical example with a small synthetic network is provided for the comparison of toll and subsidy scheme with other existing toll schemes in terms of OD travel disutilities.

Keywords: *toll and subsidy; self-sustainable; pareto-improving; revenue-neutral; no-toll equilibrium; system-optimal.*

1. Introduction

The concept and study of road pricing has a long history that dates back to the 1920s (Smith, 1979; Dafermos, 1973; Pigou, 1920; Beckmann et al., 1955). The economic theory behind it is the use of a price mechanism to reallocate a scarce resource, road capacity, so as to reduce the efficiency loss due to travelers' selfish-routing behavior (Roughgarden, 2005; Correa et al., 2004; Roughgarden, 2003; Anderson and Renault, 2003; Roughgarden and Tardos, 2002). Among existing pricing schemes, marginal-cost pricing is known to achieve a socially optimal (SO) flow pattern even though travelers still minimize their own travel costs (Yang and Huang, 1998; Yan and Lam, 1996; Akamatsu and Kuwahara, 1989; Smith, 1979). The marginal toll is easy to calculate, although its realization may not be straightforward. This pricing scheme, however, has

[1] A: No.111 2nd Ring Rd, North 1st Section, Jinniu district, Chengdu, 610031, China T: +862887600752 F: +862887600165 E: xiaofeng@swjtu.edu.cn
[2] A: One Shields Avenue, Davis, California, 95616, US T:+1(530)754-9203 F:+1(530)752-7872 E: hmzhang@ucdavis.edu

two obvious flaws: first, in most cases it increases the travel disutility of road users unless the tolls collected were redistributed back to the road users in some way; second, it results in two kinds of unfairness: 1) the anonymous pricing scheme usually incurs undesirable benefit distribution over the population because of road users' differences in their value-of-time, and 2) people traveling from different OD pairs may experience substantially different changes of travel disutility after the toll is imposed, depending on the levels and locations of tolls.

Motivated by reducing the total cost incurred by road users, Dial (2000, 1999) proposed the minimal-revenue congestion toll and provided a fast solution algorithm for obtaining the toll in a multi-origin network with fixed demand. Compared with the first-best, marginal cost toll, it lowers the total travel cost of the road users but some of the travelers may still experience an increase in travel cost. And for travelers from different OD pairs, their travel disutility may still vary significantly. Another direction to lower the travel cost is to directly return the revenue collected by the marginal-cost toll to the travelers by a monthly "credit allowance" (Kockelman and Kalmanje, 2005; Gulipalli and Kockelman, 2008). The allowance is uniformly redistributed to the drivers regardless of their travel distances. Such a redistribution scheme would pit long-distance commuters against short-distance commuters because the former subsidizes the latter, which causes another form of inequity other than price discrimination.

To address the fairness issue, "Pareto-improving" pricing schemes were introduced, which not only improve the performance of the network (not necessarily to the level of the first-best solution), but also reduce the travel disutility of travelers from every OD pair (Law-phongpanich and Yin, 2009; Yang and Zhang, 2002). Unfortunately, under general network settings, the link-based pareto-improving congestion toll scheme does not always exist if the toll rate is restricted to be nonnegative. Moreover, since the whole problem is usually formulated as an optimization problem with equilibrium and equity constraints, it is often difficult to obtain the global optimum, if there is one, in a large network. To overcome the intrinsic shortcoming of the single pricing scheme, most recently Guo and Yang (2010) studied a combination of congestion pricing and revenue refunding schemes in general transportation networks. By implementing an additional OD and class-based refunding scheme, the two targets of pareto-improving and system optimal can be simultaneously achieved. The pareto-improving OD-based refunding scheme is attractive in the way that it always exists despite of the network topology. But detecting every traveler's trip OD on everyday might be a problem in reality. The same person may have multiple trips within one day, each has different origin and destination, making the refunding procedure even more complicated. Besides, there could be privacy issues when the OD information is extracted for the refunding purpose for individual travelers.

Road pricing schemes aim at obtaining a desirable flow distribution over the network by charging a toll on all or a selective set of links, hence increasing the travel costs of all or some of the travelers. We can call such schemes as managing traffic with a stick. Instead of forcing travelers to shift from an overly utilized road to an underutilized road with a stick (i.e. toll), can we offer them a carrot (i.e. subsidy) when they use the underutilized road? Both schemes can change travelers' route choices but the effects on the traveler's surplus are opposite. Compared with traditional congestion tolls, the combination of toll and subsidy has several advantages: 1) the scheme faces less public resistance because everyone might benefit from it; 2) it adds additional degree of freedom to achieve simultaneously different planning goals, such as equity, efficiency and self-sustainability; and 3) it functions also as a revenue redistribution scheme. It is well known that allocating the revenue collected by the toll could be a controversial issue. To road users, the congestion toll is merely a variant of a lump-sum tax. In contrast, the toll and subsidy scheme is able to directly return the revenue to road users. The scheme itself can break even, which means the total revenue given out to the road users could all be covered by the revenue collected from the users.

In the literature the toll and subsidy strategies has been investigated within various contexts. For example, Liu et al. (2009) studied the possibility and conditions of pareto-improving congestion pricing schemes on a bi-modal network. They considered a continuously distributed VOT under static network equilibrium and concentrated on the revenue-neutral pricing. Nie and Liu (2010) extended their work by focusing on the existence of a self-financing and pareto-improving toll scheme under different distributions of VOT. They considered a two-mode simple network, and the system is not necessarily optimal under the toll schemes. Hearn and Ramana (1998) studied the toll pricing framework on a general transportation network with identical travelers. Under what they call the "Robinhood" scheme, network users collect a payment on some of the links and pay a toll on others, indicating that it is equivalent to a toll and subsidy scheme. They claimed that the system-optimal toll set can be obtained by solving a linear program problem. And under the "Robinhood" scheme, the total revenue collected by the system optimal toll can be zero. Recently Chen and Yang (2012) investigated a bi-objective optimization problem where a toll cum rebate scheme is utilized to obtain both system optimum and emission minimization. Yet in their study, the toll cum rebate scheme is not required to be pareto-improving in terms of travelers' cost.

This paper further explores the properties of the toll and subsidy scheme (TSS) on a general network with fixed demand and identical travelers. We try to answer several intriguing questions that have never been answered before: does there exist a toll and subsidy scheme that can simultaneously achieve the following objectives: socially optimal, pareto-improving and revenue-neutral? If the answer is yes, does it depend on the network topology? Moreover, if it exists, can we find a fast solution algorithm to obtain it? And if not, can we relax some assumptions and obtain a second-best solution? To answer those questions, we start with discussing an alternative first-best pareto-improving TSS. Then the existence of a system optimal, pareto-improving and revenue-neutral TSS will be proven on a single origin or single destination network. For a network with multiple origins and destinations, we develop the model and the algorithm to solve the TSS. Numerical examples with sample networks are also provided to demonstrate the resulting equilibrium and effectiveness of the algorithm, and to compare the TSS with other existing toll schemes in terms of OD travel disutility.

2. Definitions

We consider a directed network $G=(N,A)$, denoted by a set of consecutively numbered nodes, N, a set of consecutively numbered links, A and a set of OD pairs, W. A directed link (i,j) has two endpoints i and j. On this network, each OD pair is connected by a set of paths through the network. The OD demands are denoted by a column vector d with entries $d_w, w \in W$.

Let $x=(x_{ij})_{ij\in A}$ and $t=(t_{ij})_{ij\in A}$ represent the link flow and link travel time vectors, we have $t_{ij}=t_{ij}(x_{ij})$. For the no-toll equilibrium (NTE) to be unique, we assume $t_{ij}(\bullet)$ is strictly increasing and convex. $\rho=(\rho_{ij})_{ij\in A}$ is denoted as the column vector of link-based toll rates. When ρ_{ij} is positive, users traveling on link (i,j) will pay ρ_{ij} amount of toll, while if ρ_{ij} is negative, users will receive $|\rho_{ij}|$ amount of subsidy for using link (i,j). $c_{ij}=t_{ij}(x_{ij})+\rho_{ij}$, represents the generalized link travel cost (Here without loss of generality we assume that the factor to transfer time to money, which we call the "value of time", is equal to 1). We also define μ as the column vector of OD travel disutilities.

Throughout the paper, the superscript "m" represents the marginal-cost (MC) based solution, "$\sim$" over the symbol represents system-optimal (SO) solution and "$-$" over the symbol represents the solution under NTE.

3. The alternative System-Optimal (SO) and Pareto-Improving (PI) TSS

It's well known that if the toll rate of every link is charged to its marginal travel cost, the link flow pattern is system-optimal (SO), i.e. the total system cost is minimized. The toll internalizes the external cost of road users and is always nonnegative, which takes the form

$$\rho_{ij}^{m} = x_{ij}\frac{dt_{ij}(x_{ij})}{dx_{ij}} \geq 0 \tag{1}$$

However, if we do not require the toll rate to reflect the marginal cost, there are an infinite number of schemes that can produce the optimal link flows when the OD demands are fixed. Given any SO TSS $\tilde{\rho}$, one straightforward way to obtain an alternative TSS is to proportionally change all the links' travel costs, following the equation below

$$\rho(\alpha) = (\alpha - 1)(\tilde{t} + \tilde{\rho}) + \tilde{\rho},\ 0 < \alpha \leq 1 \tag{2}$$

Under such a TSS, the link travel cost becomes

$$c = \tilde{t} + \rho(\alpha) = \alpha(\tilde{t} + \tilde{\rho}) \tag{3}$$

Here we assume $0 < \alpha \leq 1$. When $\alpha = 1$, the toll pattern reduces to original system-optimal TSS, $\tilde{\rho}$; and when $\alpha \to 0$, all the toll rates become negative and all the generalized link travel costs approach to 0. From Eq. (3) we observe that all the link travel costs under the alternative TSS are changed by the same percentage, α. Correspondingly, all the path travel costs are also changed by α, so that the equilibrium remains unchanged. Therefore, the link flows are still optimal under the TSS defined by Eq. (2). With this observation, we give Lemma 1.

Lemma 1. *On a general network with fixed OD demands, we can always find a TSS that is pareto-improving and system-optimal. Furthermore, under such a TSS, every link retains a positive generalized travel cost.*

Proof. To prove Lemma 1 we only need to construct such a TSS based on a general network. Suppose under some scheme, $\tilde{\rho}$, the system is optimal and every link's generalized travel cost is positive (for example, the marginal-cost TSS). If all the OD travel disutilities under $\tilde{\rho}$ are less than the OD travel disutilities under NTE, the SO TSS $\tilde{\rho}$ is already pareto-improving. Otherwise, there must exist one OD pair which experiences a travel disutility increase under $\tilde{\rho}$, i.e.

$$min\{\frac{\bar{\mu}_w}{\tilde{\mu}_w}, w \in W\} < 1 \tag{4}$$

We also know that as long as a TSS $\rho(\alpha)$ follows Eq. (2), it is system-optimal because it will not change the link flow pattern under $\tilde{\rho}$; To be pareto-improving, let

$$\epsilon = min\{\frac{\bar{\mu}_w}{\tilde{\mu}_w}, w \in W\} \tag{5}$$

Since we reduce the travel cost on every link proportionally by adjusting the toll rate until the travel disutility of the OD pair which experiences the highest percentage increase of travel disutility is equal to the NTE travel disutility, all the other OD pairs will thus experience nonincreasing travel disutility under $\rho(\epsilon)$, which indicates that $\rho(\epsilon)$ is pareto-improving. And because ϵ is always greater than 0, from Eq. (3) we know that the generalized travel cost of every link under $\rho(\epsilon)$ is still positive.□

Though a SO and PI TSS can be easily derived in this way, the flaws are also obvious: first, the spatial inequity still exists, because such a TSS changes all the OD travel disutility by the same percentage. Due to the OD distribution and network topology, the travel disutility changes from NTE to SO among different OD pairs may still vary significantly; second, decreasing all the link costs by a same percentage could come with a price: considering the extreme case that the lowest OD demand experiences the highest percentage of travel disutility increase under the marginal-cost TSS, to be pareto-improving, we have to subsidize all the other OD pairs by giving out money to a large number of road users who just experience very small percentage of travel disutility increase or even travel disutility decrease. The revenue collected by the TSS could easily go negative so that for supporting such a TSS, a large amount of external subsidy may be needed. Thus one may ask that can we find a TSS which is both *pareto-improving* and *self-sustainable* (the revenue is nonnegative)? We shall answer this question in the following sections, by first looking into the case of single-origin networks, then extending the result to multi-origin, multi-destination networks.

4. The case of a single-origin network

4.1 A network with one OD pair

We first consider the simplest case in which there is only one OD pair in a general network. We have the following conclusion for this special case

Lemma 2.*There always exists one SO, PI and revenue-neutral TSS, $\hat{\rho}$, when the network has only one OD pair*[1] .

Proof. We know that the system cost under marginal-cost TSS is always less than or equal to the system cost under NTE. That is

$$t(\tilde{x})^T \tilde{x} \le t(\overline{x})^T \overline{x} \tag{6}$$

Because there is only one OD pair, the OD travel disutility under NTE equals

$$\overline{\mu} = \frac{1}{d} t(\overline{x})^T \overline{x} \tag{7}$$

And the OD travel disutility under the marginal-cost TSS becomes

$$\tilde{\mu} = \frac{1}{d}(t(\tilde{x})^T \tilde{x} + (\rho^m)^T \tilde{x}) \tag{8}$$

If $\tilde{\mu} \le \overline{\mu}$, the marginal-cost TSS is just the scheme we want to find, with $\alpha = 1$. Otherwise, from Eqs. (7) and (8) we have

[1]We note that Lemma 1 can also be directly derived from Guo and Yang's work Guo and Yang(2010), because if there is only one OD, the TSS proposed here is equivalent to a toll plus OD-based refunding scheme defined in their paper.

$$(\rho^m)^T \tilde{x} > 0 \tag{9}$$

Let

$$\alpha = \frac{t(\tilde{x})^T \tilde{x}}{t(\tilde{x})^T \tilde{x} + (\rho^m)^T \tilde{x}} \tag{10}$$

Under a TSS defined by (2) we know that now the OD travel disutility

$$\hat{\mu} = \alpha \frac{1}{d}(t(\tilde{x})^T \tilde{x} + (\rho^m)^T \tilde{x}) = \frac{1}{d} t(\tilde{x})^T \tilde{x} < \overline{\mu} \tag{11}$$

which implies that the TSS is pareto-improving. And from the previous discussion, since $\hat{\rho}$ satisfies definition (2), it will not change the flow pattern under the marginal-cost toll, which means the link flow pattern is system-optimal. Moreover, utilizing (2) and (10), the total revenue collected by $\hat{\rho}$, R, can be calculated by

$$R = \hat{\rho}^T \tilde{x} = (\alpha - 1) t(\tilde{x})^T \tilde{x} + \alpha (\rho^m)^T \tilde{x} = 0 \tag{12}$$

which implies that the TSS $\hat{\rho}$ is revenue-neutral. The proof is complete.□

Actually in the one OD case, by proportionally lowering down the travel cost on every link, the system cost saved by the TSS is totally returned to the road users, so that every road user is able to experience a reduction of travel disutility. However, if there are multiple OD pairs, since some of the link flows are shared by different OD pairs, it is not easy to find an anonymous link-based TSS that is SO, PI and revenue-neutral. Furthermore, if the numbers of OD pairs are sufficiently large, such a TSS may not even exist. In the following, we will first discuss a special network, the single-origin or single-destination network, and then study the general situation with multiple origins and destinations.

4.2 The SO-NTE TTS on a single-origin network

We first define a TSS as a SO-NTE TSS under which *the link flow pattern is system-optimal but the OD travel disutilities are the same with those under NTE*. As long as the NTE is not fully efficient, we know that the revenue of SO-NTE TSS is always positive. We also notice that according to Eq. (2) we can always reduce the toll rate on each link in the network so as to return the revenue to travelers. Therefore, once we find a SO-NTE TSS, we can always find an alternative TSS that is pareto-improving and revenue-neutral.

We associate every node in the network with a node potential, π_i. π_i can be calculated by the definition that π_i is equal to the length of the shortest path distance from the origin node o to node i. We can see that under this definition, all the other nodes' potentials are relative values, depending on the node potential of the origin node, π_o, which can be arbitrarily chosen. If π_w is the node potential of the destination node w, from the definition, $\pi_w - \pi_o = \mu_w$ is the OD travel disutility of the OD pair w under equilibrium. For a given set of node potentials π, we define the reduced cost of an link (i, j) as $c_{ij}^{\pi} = (t_{ij} + \rho_{ij}) + \pi_i - \pi_j$. We know that if a TSS reproduces the SO link flow pattern, $\tilde{x}$, the reduced costs have to be zero on the flow-bearing links and nonnegative on unused links, that is

$$\begin{cases} c_{ij}^{\pi} = 0, & if\ \tilde{x}_{ij} > 0 \\ c_{ij}^{\pi} \geq 0, & Otherwise \end{cases} \tag{13}$$

To find the SO-NTE TSS (if it does exist), we construct the following Maximum-Revenue problem with Pareto-Improving constraints (MRPI)

$$\max_{(\rho,\pi)} \sum_{ij} \tilde{x}_{ij}\rho_{ij} \tag{14}$$

subject to

$$\tilde{t}_{ij} + \rho_{ij} + \pi_i - \pi_j = 0,\ \forall \tilde{x}_{ij} \geq 0,\ (i,j) \in A \tag{15}$$

$$\tilde{t}_{ij} + \rho_{ij} + \pi_i - \pi_j \geq 0,\ \forall \tilde{x}_{ij} = 0,\ (i,j) \in A \tag{16}$$

$$\pi_w - \pi_o \leq \bar{\mu}_w,\ \forall w \in W \tag{17}$$

The objective function is maximizing the total revenue collected by the TSS. (15)-(16) restrict the network flow pattern to be optimal. Constraint (17) states that the TSS ρ has to be pareto-improving. The revenue reaches its upper-bound when constraint (17) is binding, indicating that if the solution to the problem is its upper-bound, the solution is just the SO- NTE TSS we are looking for. Otherwise, the SO-NTE TSS does not exist. To transform the greater-than-or-equal constraints to equality constraints, we define $s_{ij} \geq 0$ as a nonnegative slack variable associated with each link $(i,j) \in A$ and $\lambda_w \geq 0$ as a nonnegative slack variable associated with each destination node $w \in W$. In addition, we define a zero-flow indicator z_{ij} to be

$$z_{ij} = \begin{cases} 1, & if\ x_{ij} = 0, \\ 0, & Otherwise. \end{cases} \tag{18}$$

By utilizing the zero-flow indicator and slack variables, the problem can be rewritten as

LP1:

$$\max_{(\rho,\pi,s,\lambda)} \sum_{ij} \tilde{x}_{ij}\rho_{ij} \tag{19}$$

subject to

$$\tilde{t}_{ij} + \rho_{ij} + \pi_i - \pi_j - z_{ij}s_{ij} = 0,\ \forall (i,j) \in A \tag{20}$$

$$\bar{\mu}_i - \pi_i - \lambda_i = 0,\ \forall i \in W \tag{21}$$

$$s, \lambda \geq 0 \tag{22}$$

$$\rho, \pi \in \mathcal{R} \tag{23}$$

We notice that here the toll rates and node potentials are unconstrained, which means that by solving LP1, the link generalized travel cost could be negative if the toll rate goes negative and sufficiently low. Then naturally, one may ask is it possible that the TSS induces some negative cycle? To answer the question we give the lemma below

Lemma 3. *Under the TSS scheme provided by LP1, the network contains no negative cycle.*

Proof. From the definition of reduced cost c_{ij}^{π} , for any directed cycle P , we have

$$\begin{aligned}\sum_{(i,j)\in P} c_{ij}^{\pi} &= \sum_{(i,j)\in P} ((t_{ij}+\rho_{ij})+\pi_i-\pi_j) \\ &= \sum_{(i,j)\in P} (t_{ij}+\rho_{ij}) + \sum_{(i,j)\in P} (\pi_i-\pi_j) \\ &= \sum_{(i,j)\in P} (t_{ij}+\rho_{ij})\end{aligned} \tag{24}$$

From Eq. (21) we know that at the equilibrium, $c_{ij}^{\pi} = z_{ij}s_{ij} \geq 0, \forall(i,j)\in A$, which implies That $\sum_{(i,j)\in P} c_{ij}^{\pi} \geq 0$. From Eq.(24) we have that $\sum_{(i,j)\in P} (t_{ij}+\rho_{ij}) = \sum_{(i,j)\in P} c_{ij}^{\pi} \geq 0$, thus we conclude that as long as the TSS scheme ρ is the solution of LP1, the network does not contain a negative cycle.□

Obviously, the upper-bound of LP1 is equal to the difference of system costs between NTE and SO.

4.3 Dual linear program

To find the dual problem of LP1, we first restate LP1 in matrix notation. Let

$I = |A|\times|A|$, the identity matrix;

$J = |W|\times|W|$, the identity matrix;

$\Delta = (\delta_{l,ij})_{ij\in A}^{l\in N}$, the network node-link incidence matrix, where

$$\delta_{l,ij} = \begin{cases} 1, & if\ i = l, \\ -1, & if\ j = l, \\ 0, & otherwise \end{cases}$$

$b = \Delta\tilde{x} = (b_n)_{n\in N}$ trip demand to each node n, where

$$\begin{cases} b_i > 0, & if\ i\ is\ an\ origin\ node, \\ b_i < 0, & if\ i\ is\ a\ destination\ node, \\ b_i = 0, & otherwise. \end{cases}$$

$\Lambda = (\sigma_{i,j})_{j\in W}^{i\in N}$, diagonalization of the network node-OD incidence matrix, where

$$\sigma_{i,j} = \begin{cases} 1, & if\ i\ is\ an\ origin\ node, \\ -1, & if\ i\ is\ a\ destination\ node, \\ 0, & otherwise. \end{cases}$$

From the above definitions, we immediately have

$$\Lambda d = \Delta\tilde{x} = b \tag{25}$$

$Z = (z_{a,a'})_{a'\in A}^{a\in A}$ diagonalization of z indicators, where

$$z_{a,a'} = \begin{cases} z_a, & if\ a = a', \\ 0, & otherwise. \end{cases}$$

Then the primal problem LP1 can be rewritten as

$$\max_{(\rho,\pi,s,\lambda)} \tilde{x}^T \rho \tag{26}$$

subject to

$$\begin{pmatrix} I & \Delta^T & -Z & 0 \\ 0 & -\Lambda^T & 0 & J \end{pmatrix} \begin{pmatrix} \rho \\ \pi \\ s \\ \lambda \end{pmatrix} = \begin{pmatrix} -\tilde{t} \\ \bar{\mu} \end{pmatrix} \tag{27}$$

$$s, \lambda \geq 0 \tag{28}$$

$$\rho, \pi \in \mathcal{R} \tag{29}$$

The dual problem to Eqs.(26)-(29) is

DP1:

$$\min P(y_1, y_2) = -\tilde{t}^T y_1 + \bar{\mu}^T y_2 \tag{30}$$

subject to

$$\begin{pmatrix} I & 0 \\ \Delta & -\Lambda \\ -Z & 0 \\ 0 & J \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \begin{matrix} = \tilde{x} \\ = 0 \\ \geq 0 \\ \geq 0 \end{matrix} \tag{31}$$

From the first constraint we readily obtain that $y_1 = \tilde{x}$. The fourth constraint restricts y_2 to be nonnegative. And since $y_1 = \tilde{x}$, $-Zy_1 = 0$ always holds. Thus the third constraint is redundant. The second constraint gives that $\Lambda y_2 = \Delta y_1 = \Delta \tilde{x} = b$. We know that $|\Lambda| = |y| =$ the number of Destination nodes, thus y_2 is unique. From Eq. (25), y_2 is just equal to the OD demand vector d. Substituting all the constraints, the objective function becomes

$$P(y) = \bar{\mu}^T d - \tilde{t}^T \tilde{x} \tag{32}$$

The first term in Eq. (32) is the total travel cost under user equilibrium and the second term is the total travel cost under SO. Thus the result of primal linear program LP1 is always the upper-bound of the problem, which implies that the SO-NTE TSS always exists, as well as the SO, PI and revenue-neutral TSS.

The proof is similar for the single-destination multi-origin network. The algorithm is exactly the same except that we need to set the potential of the destination node to be 0 and the potentials of other nodes as the shortest path distance from this node to the destination node. To sum up, we give the following theorem

Theorem 1.*On a single-origin or single-destination network, we can always find a TSS,* $\hat{\rho}$ *, which has the following properties:*

i) Every road user experiences the same percentage improvement of the travel disutility;

ii) The link flow pattern is socially optimal;

iii) The revenue collected by the toll is 0, i.e. the TSS is break-even (i.e., revenue neutral).

Under the SO-NTE TSS given by LP1 we have

$$\lambda_i = 0, \ \forall i \in W \tag{33}$$

Thus the SO-NTE scheme can be solved by the following group of equations

$$\begin{cases} \tilde{t}_{ij} + \rho_{ij} + \pi_i - \pi_j = 0, & \forall x_{ij} > 0 \\ \pi_i = \overline{\mu}_i, & \forall i \in W \\ \pi_o = 0 \end{cases} \tag{34}$$

The extreme situation happens when every node in the subgraph except the origin node is a destination node. Then the SO-NTE TSS is unique.

Theorem 1 gives us some good news: compared with the minimal-revenue pricing scheme proposed by Dial (1999), the advantages of the TSS here are obvious: it not only further lowers the revenue and as a consequence lowers the travel cost of each traveler, but also mitigates OD-specific price discrimination. Furthermore, the TSS does not require every link to be associated with a toll when the number of destination nodes is relatively small. Assume the number of nodes in the network is n and w of them are destination nodes, then the freedom of the problem (34) is $(n-1-w)$, which means at most $(n-1-w)$ links can be free links.

4.4 A numerical example

Here we provide a numerical example (Figure 1) of the famous network of Braess' Paradox (Braess, 1968).

Suppose node 1 is the origin node, nodes 2, 3, 4 are all destination nodes. The OD demands $d_{12} = 2$, $d_{13} = 1.5$, $d_{14} = 4$. Solving SO and NTE we have the optimal link travel time and OD travel disutility under NTE

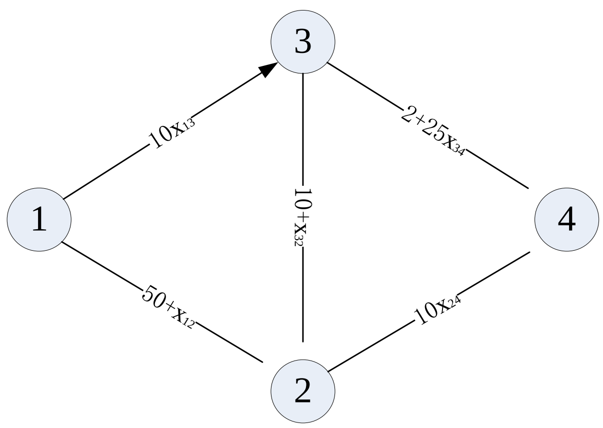

Figure 1 Network of Braess' paradox

$$\begin{pmatrix} \tilde{t}_{13} \\ \tilde{t}_{12} \\ \tilde{t}_{34} \\ \tilde{t}_{24} \\ \tilde{t}_{32} \end{pmatrix} = \begin{pmatrix} 26.96 \\ 54.80 \\ 31.89 \\ 28.04 \\ 10.00 \end{pmatrix}, \quad \begin{pmatrix} \bar{u}_{14} \\ \bar{u}_{12} \\ \bar{u}_{13} \end{pmatrix} = \begin{pmatrix} 79.22 \\ 53.30 \\ 42.01 \end{pmatrix} \tag{35}$$

We also have the total system costs under NTE and SO, which are respectively 486.47 and 452.74. Setting $\alpha = 452.74/486.47 = 0.93$. We calculate the objective OD travel disutility under the pareto-improving and revenue-neutral TSS

$$\begin{pmatrix} \hat{\mu}_{14} \\ \hat{\mu}_{12} \\ \hat{\mu}_{13} \end{pmatrix} = \begin{pmatrix} 73.72 \\ 49.60 \\ 39.10 \end{pmatrix} \tag{36}$$

which are also the node potentials for destination nodes 2, 3, 4 (we set the node potential of the origin node 1 to be 0). By solving a group of linear equations, we obtain the pareto- improving, system-optimal and revenue-neutral TSS, which are shown in the parentheses in Figure 2. The underlined number associated with each node is the node potential under the TSS.

5. The case of multi-origin networks

5.1 The MRPI problem

Similar with the one-origin network, for the multi-origin network we can still formulate the MRPI problem for finding a SO and PI TSS with the lowest possible external subsidy required. The SO and PI TSS is not necessarily self-sustainable now, since we have only one single set of toll rates to deal with the trips from all the origins. However, in case the maximum revenue is negative, it's still meaningful to look for a SO and PI TSS, with the minimum possible subsidy required, which is able to eliminate the deadweight loss of the selfish-routing behavior and make everyone better-off simultaneously. By using superscript k as origin index, we first enrich the notations as below.

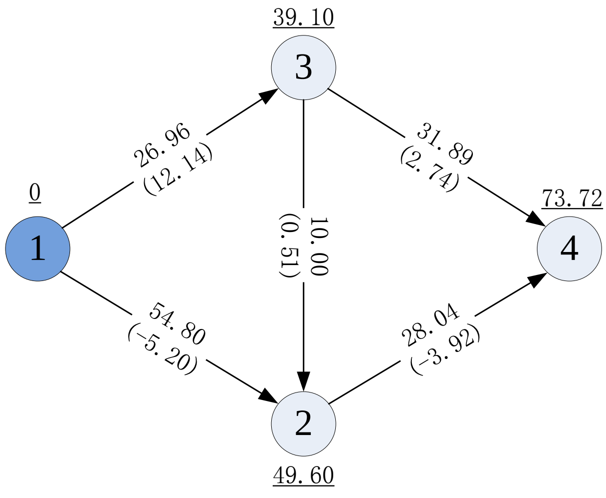

Figure 2 Resulting toll rates

K : origin-node set,

π^k : $(\pi_n^k)_{n\in N}^{k\in K}$, node potentials for origin k,

s^k : $(s_{ij}^k)_{ij\in A}^{k\in K}$, link slack variables for origin k,

λ^k : $(\lambda_w^k)_{w\in W^k}^{k\in K}$, OD slack variables for each destination nodes with origin k,

b^k : $(b_n^k)_{n\in N}^{k\in K}$, the OD matrix,

Λ^k : $(\sigma_{i,j}^k)_{j\in W^k}^{i\in N}$, the diagonalization of the network node-OD incidence matrix for origin k,

Z^k : $(z_{a,a'}^k)_{a'\in A}^{a\in A}$, the diagonalization of z^k indicators,

Without loss of generality, we investigate the two-origin case, since the general pattern for n origins can then be easily revealed from the two-origin case. The primal linear program for the two-origin case can be written in matrix notation in the following form. Here we add a reasonable assumption that the generalized link travel cost after the TSS cannot be negative, i.e. $\rho+\tilde{t}\geq 0$, so that people will not experience lower travel cost when they travel longer.

$$\max_{(\rho,\pi,s,\lambda)} \tilde{x}^T\rho \tag{37}$$

subject to

$$\begin{pmatrix} I & \Delta^T & 0 & -Z^1 & 0 & 0 & 0 \\ I & 0 & \Delta^T & 0 & -Z^2 & 0 & 0 \\ 0 & (-\Lambda^1)^T & 0 & 0 & 0 & J^1 & 0 \\ 0 & 0 & (-\Lambda^2)^T & 0 & 0 & 0 & J^2 \end{pmatrix} \begin{pmatrix} \rho \\ \pi^1 \\ \pi^2 \\ s^1 \\ s^2 \\ \lambda^1 \\ \lambda^2 \end{pmatrix} = \begin{pmatrix} -\tilde{t} \\ -\tilde{t} \\ \overline{\mu}^1 \\ \overline{\mu}^2 \end{pmatrix} \tag{38}$$

$$s,\lambda \geq 0 \tag{39}$$

$$\rho+\tilde{t}\geq 0 \tag{40}$$

Let $\theta=\rho+\tilde{t}\geq 0$, which is the generalized link travel cost. Substituting θ, we have the equivalent formulation

LPn:

$$\max_{(\theta,\pi,s,\lambda)} \tilde{x}^T\theta \tag{41}$$

subject to

$$\begin{pmatrix} I & \Delta^T & 0 & -Z^1 & 0 & 0 & 0 \\ I & 0 & \Delta^T & 0 & -Z^2 & 0 & 0 \\ 0 & (-\Lambda^1)^T & 0 & 0 & 0 & J^1 & 0 \\ 0 & 0 & (-\Lambda^2)^T & 0 & 0 & 0 & J^2 \end{pmatrix} \begin{pmatrix} \theta \\ \pi^1 \\ \pi^2 \\ s^1 \\ s^2 \\ \lambda^1 \\ \lambda^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \overline{\mu}^1 \\ \overline{\mu}^2 \end{pmatrix} \quad (42)$$

$$\theta, s, \lambda \geq 0 \quad (43)$$

From Lemma 1 we know that the SO and PI TSS always exists and the problem has an upper-bound which is equal to the difference of system travel time costs between NTE and SO. Thus the linear program LPn and its dual are both feasible and bounded and have the same value. From Eqs. (51)-(43), the dual problem becomes

$$\min P(y) = (\overline{\mu}^1)^T y_2^1 + (\overline{\mu}^2)^T y_2^2 \quad (44)$$

subject to

$$\begin{pmatrix} I & I & 0 & 0 \\ \Delta & 0 & -\Lambda^1 & 0 \\ 0 & \Delta & 0 & -\Lambda^2 \\ -Z^1 & 0 & 0 & 0 \\ 0 & -Z^2 & 0 & 0 \\ 0 & 0 & J^1 & 0 \\ 0 & 0 & 0 & J^2 \end{pmatrix} \begin{pmatrix} y_1^1 \\ y_1^2 \\ y_2^1 \\ y_2^2 \end{pmatrix} \begin{matrix} \geq \tilde{x} \\ = 0 \\ = 0 \\ \geq 0 \\ \geq 0 \\ \geq 0 \\ \geq 0 \end{matrix} \quad (45)$$

From the first three constraints, we have $\Lambda^1 y_2^1 + \Lambda^2 y_2^2 \geq \Delta\tilde{x}$. The last two are nonnegative constraints. The other constraints are just redundant. The problem can be transformed into a simplified version

DPn:

$$\min P(y) = (\overline{\mu}^1)^T y_2^1 + (\overline{\mu}^2)^T y_2^2 \quad (46)$$

subject to

$$\Lambda^1 y_2^1 + \Lambda^2 y_2^2 = \Delta\tilde{x} \quad (47)$$

$$y_2^1, y_2^2 \geq 0 \quad (48)$$

where constraint (47) becomes an equation because for the minimization problem DPn with all the coefficients positive, the constraint is always binding. By observing DPn, we notice that y_2^1 and y_2^2 can be interpreted by the OD flows respectively belonging to origin 1 and origin 2. Thus the dual problem of LPn is actually equivalent to a **balanced transportation problem**: finding the optimal pattern of the distribution of goods from several points of origin to several different destinations, with the fixed OD travel costs. In addition, if the solution of DPn is greater than the total system cost under SO, the SO and PI TSS is self-sustainable, and vice versa.

Theorem 2.*On a multi-origin network with fixed demands, the MRPI problem is equivalent to a single-commodity balanced transportation problem with the same demand to each node and the OD travel costs equal to the OD travel disutilities at NTE.*

Thus it becomes very easy to find out if the PI, SO and revenue-neutral TSS exists. We just need to solve DPn and compare the value to the total system cost under SO. If it's greater, the PI, SO and revenue-neutral TSS exists, and vice versa. Since the transportation problem is a special case of a linear program, it can be readily solved by any linear programming algorithms, like the simplex method. However, the special structure of it allows us to solve it by faster, more specialized algorithms than the simplex. The only difficulty of solving the original problem LPn is to find Z^k , the origin-specific zero-flow indicators, since it is not explicitly given by the traditional traffic assignment software packages. Fortunately, there is a convenient way to calculate it. The procedure has been proposed by Dial (2000), which follows a simple logic: finding the origin-stratified node potentials $\tilde{\pi}_i^k$ under the system-optimal marginal-cost link toll rates ρ^m and corresponding link flow pattern, $\tilde{x}$. Then

$$z_{ij}^k = \begin{cases} 0, if \ \tilde{x}_{ij} + \rho_{ij}^m + \tilde{\pi}_i^k - \tilde{\pi}_j^k = 0, \\ 1, \qquad otherwise. \end{cases} \tag{49}$$

5.2 Lowest toll rates

Actually the MRPI problem may have multiple solutions. All of them can maximize the total revenue with decreasing OD travel disutilities but the toll rate levels associated with the links may vary quite differently. For example, if one path contains two links which are complementary to each other, one can raise the toll rate on one link as high as she/he wants and just reduce the credit rate on the other one correspondingly, without changing the travel cost of the whole path. Practically, people may want to find the SO and PI toll rate set with the lowest-possible absolute values. To find the lowest toll rates we can follow two steps and in each step we need to solve a linear program: In the first step, we solve the dual problem of MRPI. We assume the value of MRPI is P_{min} and the total SO travel cost is $\widetilde{TC}$. We then set

$$TC = \begin{cases} \widetilde{TC}, \ if \ P_{\min} \geq \widetilde{TC}, \\ P_{\min}, \quad otherwise. \end{cases} \tag{50}$$

We also define another variable $m \geq 0$ as the upper-bound of the toll rates, i.e. $\theta \leq m + \tilde{t}$; In the next step, to find the lowest toll rates we solve the following linear program

LCPI:

$$\min_{(\theta,\pi,s,\lambda,m)} m \tag{51}$$

subject to

$$\begin{pmatrix} I & \Delta^T & 0 & -Z^1 & 0 & 0 & 0 & 0 \\ I & 0 & \Delta^T & 0 & -Z^2 & 0 & 0 & 0 \\ 0 & (-\Lambda^1)^T & 0 & 0 & 0 & J^1 & 0 & 0 \\ 0 & 0 & (-\Lambda^2)^T & 0 & 0 & 0 & J^2 & 0 \\ \tilde{x}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{-1} \end{pmatrix} \begin{pmatrix} \theta \\ \pi^1 \\ \pi^2 \\ s^1 \\ s^2 \\ \lambda^1 \\ \lambda^2 \\ m \end{pmatrix} \begin{matrix} =0 \\ =0 \\ =\bar{\mu}^1 \\ =\bar{\mu}^2 \\ =TC \\ \le \tilde{t} \end{matrix} \qquad (52)$$

$$\theta, s, \lambda, m \ge 0 \qquad (53)$$

where $\mathbf{-1}$ is a column vector with all the entries equal to $\mathbf{-1}$. The value of the problem LCPI actually gives the lowest-possible upper-bound of the toll rates that are either with the minimum subsidy or revenue-neutral. And because the upper-bound of the toll rate is minimized based on the maximized revenue, implicitly, the lower-bound of the toll rate is also limited. The TSS itself remains socially-optimal and pareto-improving.

5.3 A numerical example

For comparison, we use the multi-origin network from Dial (2000) in our numerical example. The network node-OD incidence matrix, the system-optimal link time and marginal-cost tolls are respectively shown in Table 1 and Figure 3.

Table 1: OD matrix

O/D	3	4	Sum
1	10	20	30
2	30	40	70
Sum	40	60	

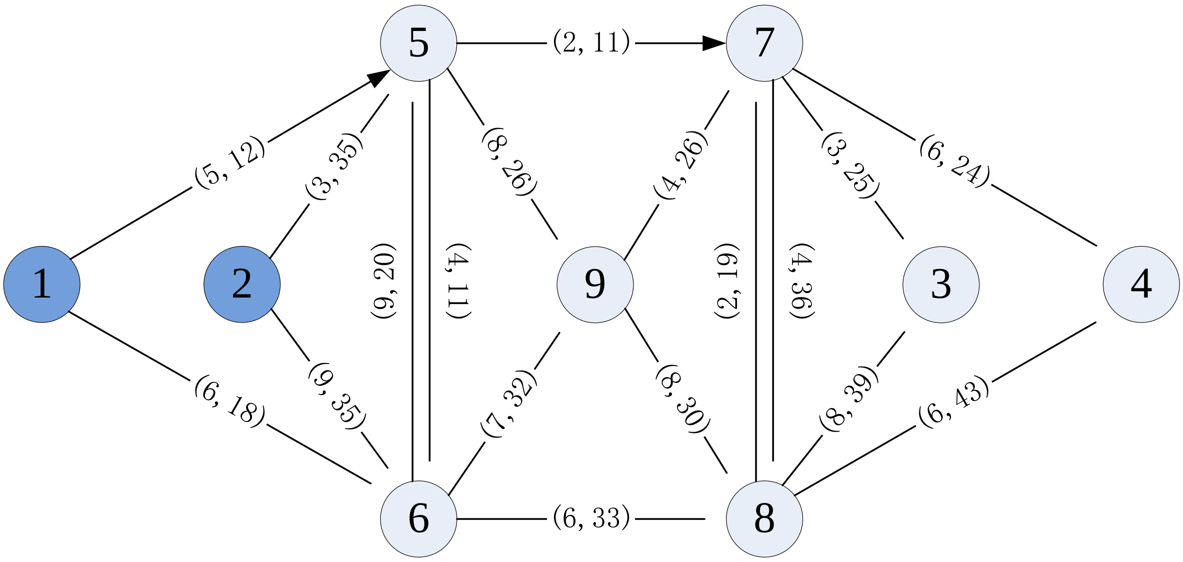

Figure 3: The 9-node network

Figure 4 shows the NTE link time.

By solving NTE we have

$$\begin{pmatrix} \bar{\mu}_{13} \\ \bar{\mu}_{14} \\ \bar{\mu}_{23} \\ \bar{\mu}_{24} \end{pmatrix} = \begin{pmatrix} 24.85 \\ 23.70 \\ 24.16 \\ 25.03 \end{pmatrix} \quad (54)$$

Substituting the NTE OD travel costs and OD demands into DPn, we have

$$\min P(y) = 24.85 y_{13} + 23.70 y_{14} + 24.16 y_{23} + 25.03 y_{24} \quad (55)$$

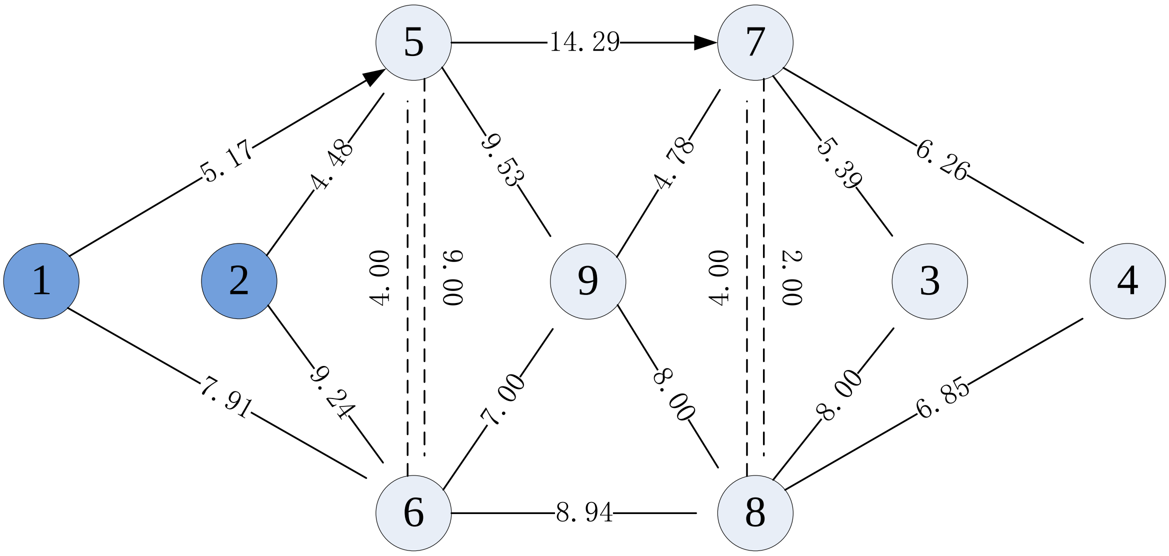

Figure 4: NTE link time

subject to

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_{13} \\ y_{14} \\ y_{23} \\ y_{24} \end{pmatrix} = \begin{pmatrix} 30 \\ 70 \\ 40 \\ 60 \end{pmatrix} \quad (56)$$

$$y \geq 0 \quad (57)$$

The value to the problem is

$$P_{\min}(y) = 2428.3 \quad (58)$$

which is also the value of the primal problem LPn. By solving the SO problem we know that the total system-optimal travel cost is $2253.9 < P_{\min}$, which implies that for this 9-node network, the pareto-improving and revenue-neutral TSS exists. And the maximum revenue the TSS is able to collect is 2428.3 − 2253.9 = 174.4.

To find the toll rates, we first obtain the origin-specific flow-bearing subgraphs and node potentials, which are shown in Figure 5. Without loss of generality, we set the potentials of the origin nodes to be 0.

The solution to LPn is not unique, which is good because this leaves us more room to achieve other goals, like pursuing the lowest possible toll rates, the TSS with least number of links to be charged or the TSS with equity or emission constraints.

Figure 6 shows the solution with the lowest-possible toll rates, which, as we have discussed, is the solution to LCPI. All the toll rates in the network are between −4.33 and 2.39. Compared with the MC toll and MR toll in Table 2, we observe that the TSS produces a nonincreasing travel cost for every OD pair. The revenue of the TSS is zero.

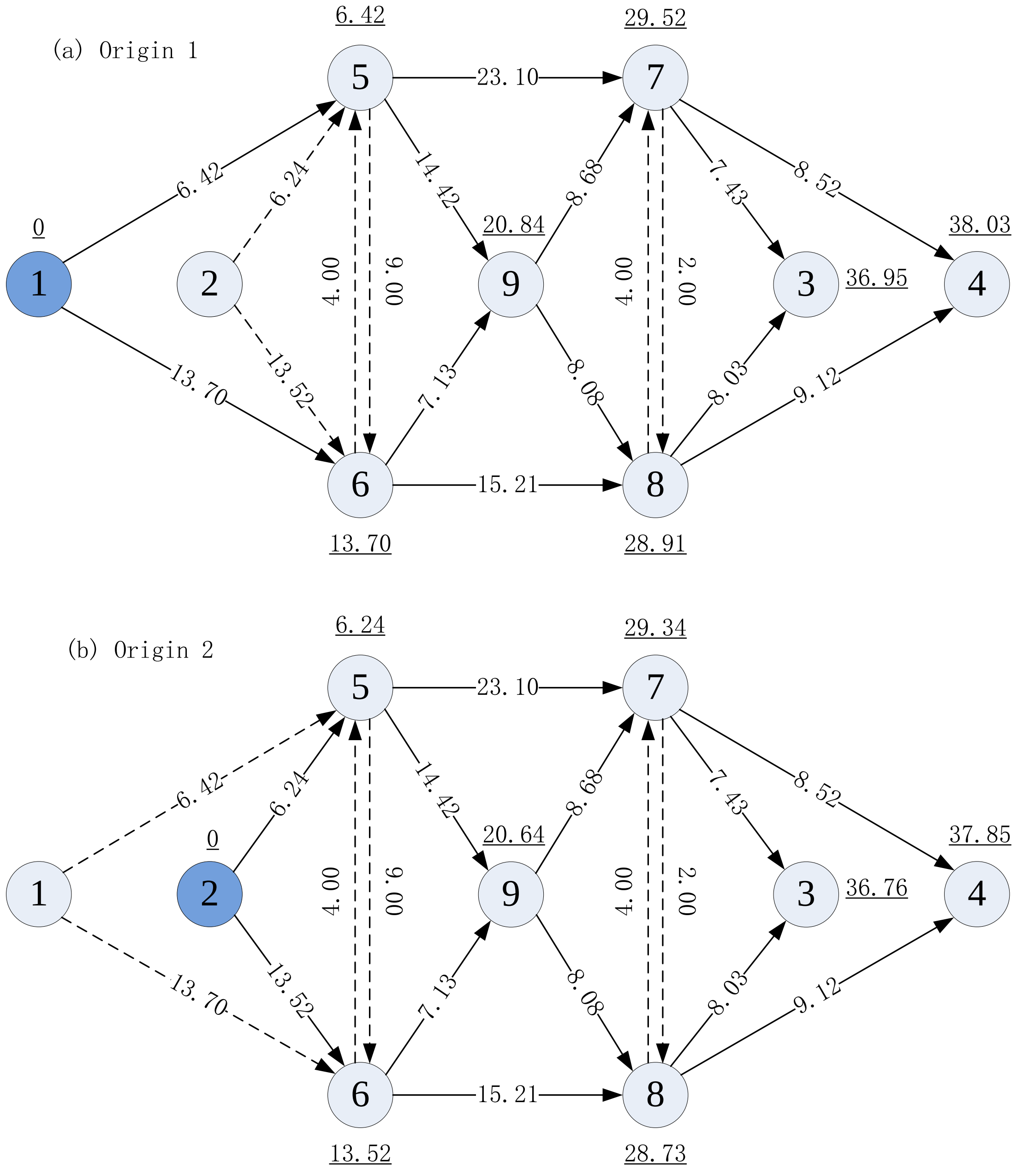

Figure 5: Subgraphs, node potentials and generalized cost under MC toll

It's worth noting that in this case if we restrict the toll rate to be nonnegative, making the problem to be a SO and PI toll problem, the primal problem becomes infeasible. Therefore, this 9-node network does provide a counter example to show that the nonnegative SO and PI toll does not always exist in a multi-origin, multi-destination network.

Table 2 Comparison of OD travel disutilities: NTE, MC Toll, MR Toll and TSS

OD pair	NTE	MC Toll	MR Toll	TSS	
1-3	24.85	36.95	30.6	21.08	0.85
1-4	23.70	38.03	29.21	23.70	1.00
2-3	24.16	36.76	32.96	20.92	0.87
2-4	25.03	37.85	31.57	23.54	0.94

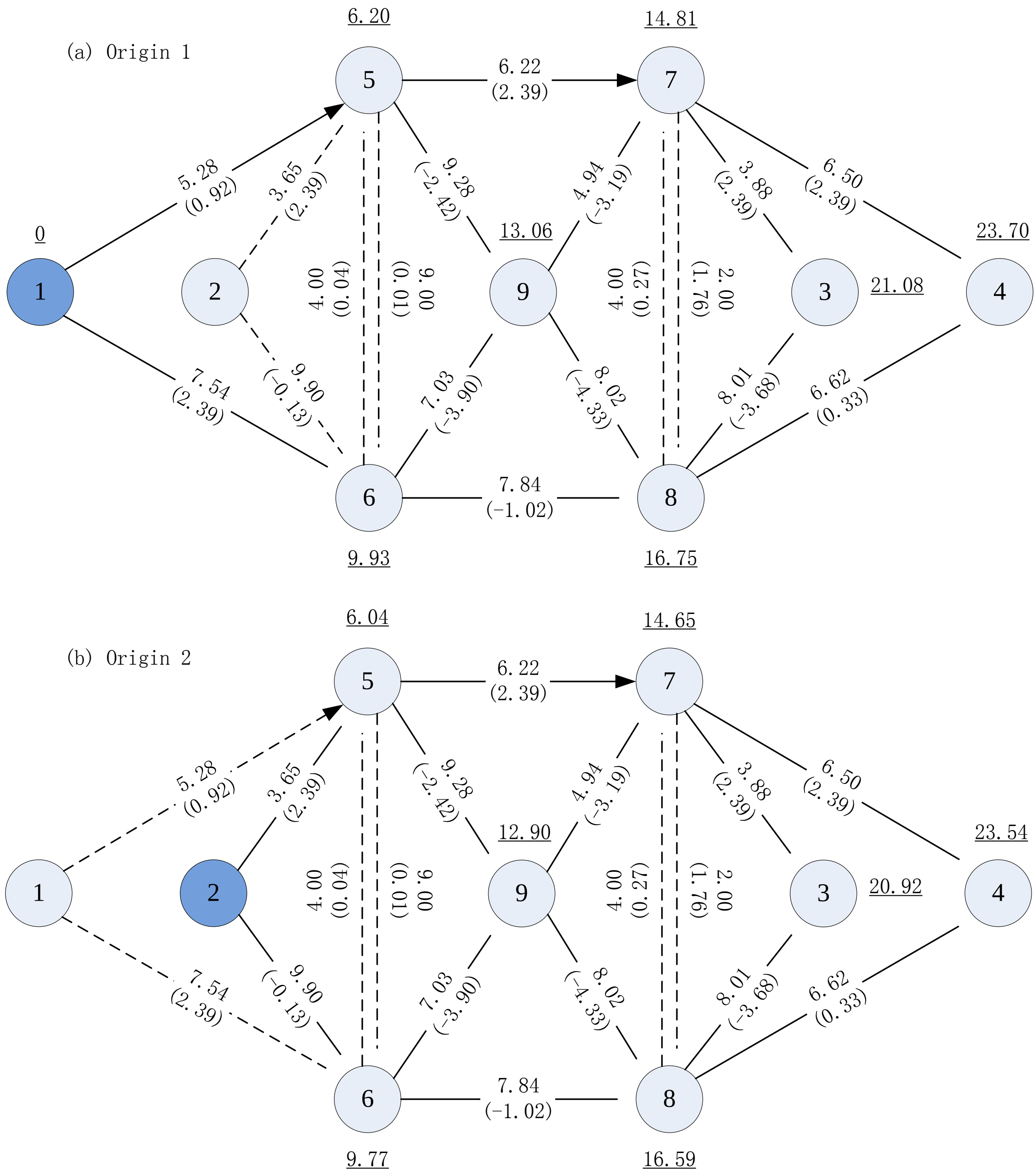

Figure 6 The pareto-improving and revenue-neutral TSS and node potentials

6. Conclusions

This paper investigates the properties of an innovative toll and subsidy scheme on a transportation network with fixed OD demands. We have two major findings from this study:

first, for one-origin or one-destination network, we can always find a TSS to be system-optimal, pareto-improving and revenue-neutral; second, for multi-origin and multi- destination network, the SO and PI toll scheme can be calculated by solving the MRPI problem. The toll scheme may not always be self-sustainable, depending on the network structure. The dual of MRPI is equivalent to the balanced transportation problem, so that we can solve the MRPI problem much faster than the traditional simplex method.

Several advantages can be found of the TSS over traditional tolling schemes where the tolls are restricted to be non-negative. First, the additional freedom allowed by offering a toll on some links let us to achieve multiple planning objectives all at once. Second, the resulting travel disutility for each OD pair is usually lower in the TSS than in a tolling scheme, and third, the TSS offers a natural (anonymous) way of distributing the revenue that can be less controversial, thus improving its chances of public acceptance.

The study of the link-based TSS is still in its infancy. Future extensions of this study may discover more potentials of this system and lead to broader applications. We are currently pursuing the following extensions:

1. *Bi-objective TSS design.* We know that for a multi-origin network, the pareto-improving TSS cannot always be self-sustainable. Under extreme situations the extra subsidy could be very high. Thus the planner may consider to reduce the subsidy by allowing a tolerable increase in OD travel disutility. Then the model has to be extended to tackle the following dual objectives: minimizing the subsidy and the OD travel disutility.

2. *Multi-class travelers and elastic demand.* Needless to say, considering value-of-time differentiation and elastic demand make the problem more complicated. For elastic demand, the link flows can no longer be optimal under the TSS because the OD demands are changed. And for multi-class travelers the system planner has to make a balance between travel time cost and monetary cost. And it will be harder to achieve the goal of pareto-improving.

3. *Other planning purposes.* In most cases the SO, PI and revenue-neutral TSS is not unique, providing extra flexibility for the system planner to realize further planning purposes, like finding the maximum number and optimal location of the free links. Capacity and environmental factors can also be involved as additional constraints.

Acknowledgement

The research was supported by the National Natural Science Foundation of China (71201135) and the grant from the Major Program of National Social Science Foundation of China (13&ZD175). The authors, however, are solely responsible for the contents of this paper.

References

Akamatsu, T. and Kuwahara, M. (1989). Optimal toll pattern on a road network under stochastic user equilibrium with elastic demand. In: *Selected Proceedings of the 5th World Conference on Transport Research*, 1, 259–273.

Anderson, S. and Renault, R. (2003). Efficiency and surplus bounds in Cournot competition. *Journal of Economic Theory*, 113 (2), 253–264.

Beckmann, M., McGuire, C. and Winsten, C., (1956). Studies in the Economics of Transportation. Yale University Press, New Haven, CT, No. 226 pp.

Braess, D. (1968). Uber ein paradoxon aus der verkehrsplanung. *Mathematical Methods of Operations Research*, 12 (1), 258–268.

Chen, L. and Yang, H. (2012). Managing congestion and emissions in road networks with tolls and rebates. *Transportation Research Part B*, 46 (8), 933 – 948.

Correa, J., Schulz, A. and Stier-Moses, N. (2004). Selfish routing in capacitated networks. *Mathematics of Operations Research*, 29 (4), 961–976.

Dafermos, S. (1973). Toll patterns for multiclass-user transportation networks. *Transportation Science*, 7 (3), 211.

Dial, R. (1999). Minimal-revenue congestion pricing part I: A fast algorithm for the single-origin case. *Transportation Research Part B*, 33 (3), 189–202.

Dial, R. (2000). Minimal-revenue congestion pricing Part II: An efficient algorithm for the general case. *Transportation Research Part B*, 34 (8), 645–665.

Gulipalli, P. and Kockelman, K. (2008). Credit-based congestion pricing: A Dallas-Fort Worth application. *Transport Policy*, 15 (1), 23–32.

Guo, X. and Yang, H. (2010). Pareto-improving congestion pricing and revenue refunding with multiple user classes. *Transportation Research Part B*, 44 (8-9), 972–982.

Hearn, D. and Ramana, M. (1998). Solving congestion toll pricing models. *Equilibrium and Advanced Transportation Modeling*, 109–124.

Kockelman, K. and Kalmanje, S. (2005). Credit-based congestion pricing: a policy proposal and the public's response. *Transportation Research Part A*, 39 (7-9), 671–690.

Lawphongpanich, S. and Yin, Y. (2010). Solving the Pareto-improving toll problem via manifold suboptimization. *Transportation Research Part C*, 18 (2), 234-246.

Liu, Y., Guo, X. and Yang, H. (2009). Pareto-improving and revenue-neutral congestion pricing schemes in two-mode traffic networks. *Netnomics*, 10 (1), 123–140.

Nie, Y. M. and Liu, Y. (2010). Existence of self-finacing and pareto-improving congestion pricing: Impact of value of time distribution. *Transportation Research Part A*, 44 (1), 39 – 51.

Pigou, A. (1920). Wealth and welfare. Macmillan and co., limited.

Roughgarden, T. (2003). The price of anarchy is independent of the network topology. *Journal of Computer and System Sciences*, 67 (2), 341–364.

Roughgarden, T. (2005). Selfish routing and the price of anarchy. The MIT Press.

Roughgarden, T. and Tardos, É. (2002). How bad is selfish routing? *Journal of the ACM (JACM)*, 49 (2), 259.

Smith, M. (1979). The marginal cost taxation of a transportation network. *Transportation Research Part B*, 13 (3), 237–242.

Yan, H. and Lam, W. (1996). Optimal road tolls under conditions of queueing and congestion. *Transportation Research Part A*, 30 (5), 319–332.

Yang, H. and Huang, H. (1998). Principle of marginal-cost pricing: how does it work in a general road network? *Transportation Research Part A*, 32 (1), 45–54.

Yang, H. and Zhang, X. (2002). Multiclass network toll design problem with social and spatial equity constraints. *Journal of Transportation Engineering*, 128 (5), 420-428.