2010

Bounding the inefficiency of logit-based stochastic user equilibrium

Xiaolei Guo
University of Windsor

Hai Yang

Tian-Liang Liu

Bounding the Inefficiency of Logit-based Stochastic User Equilibrium

Xiaolei Guo[a], Hai Yang[b] and Tian-Liang Liu[c]

[a] Corresponding Author. Department of Civil Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China. Email: guoxl@umn.edu Tel: 612-625-0249

[b] Department of Civil Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China. Email: cehyang@ust.hk

[c] Research Center for Contemporary Management, School of Economics and Management, Tsinghua University, Beijing, 100084, China. Email: liutl@sem.tsinghua.edu.cn

Abstract. Bounding the inefficiency of selfish routing has become an emerging research subject. A central result obtained in the literature is that the inefficiency of deterministic User Equilibrium (UE) is bounded and the bound is independent of network topology. This paper makes a contribution to the literature by bounding the inefficiency of the logit-based Stochastic User Equilibrium (SUE). In a stochastic environment there are two different definitions of system optimization: one is the traditional System Optimum (SO) which minimizes the total actual system travel time, and the other is the Stochastic System Optimum (SSO) which minimizes the total perceived travel time of all users. Thus there are two ways to define the inefficiency of SUE, i.e. to compare SUE with SO in terms of total actual system travel time, or to compare SUE with SSO in terms of total perceived travel time. We establish upper bounds on the inefficiency of SUE in both situations.

Keywords: Transportation; Selfish routing; Inefficiency; Stochastic user equilibrium

1. Introduction

A common behavioral assumption in traffic network modeling is that every user chooses a route that she perceives as being the shortest under the prevailing traffic conditions. In other words, every traveler tries to minimize her own (perceived) travel time. This selfish routing assumption leads to the deterministic user equilibrium (UE) traffic assignment when users are assumed to have perfect information, or their perceived travel times are exactly their actual ones. More realistically, the perceived travel time may be considered as a random variable distributed across the population of users, i.e. each user may perceive a different travel time over the same link. Then the selfish routing assumption results in the stochastic user equilibrium (SUE) traffic assignment (Sheffi, 1985). In contrast to uncoordinated selfish travel behaviors, system optimization is to minimize the total system travel time which measures the overall network performance under fixed demand. A system optimum (SO) flow pattern has the maximum efficiency by definition. Not surprisingly, selfish routing generally does not yield an SO flow pattern, which implies that UE and SUE are typically inefficient.

There has been an increasing interest recently in trying to quantify and bound the inefficiency of Nash equilibrium or UE in transportation context. Koutsoupias and Papadimitriou (1999) proposed to analyze the inefficiency of equilibria from a worst-case perspective. The term "price of anarchy" was coined to characterize the degree of inefficiency (Papadimitriou, 2001), which is the ratio of the worst social cost of a Nash Equilibrium to the cost of an optimal solution. Roughgarden (2003) proved that the worst-case inefficiency due to selfish routing is independent of the network topology. Several authors analyzed the bound on the inefficiency of equilibria for more general classes of cost functions and model features such as toll pricing (e.g. Chau and Sim, 2003; Correa et al, 2004; Roughgarden and Tardos, 2004; Han et al., 2008; Han and Yang, 2008; Yang et al., 2008). Roughgarden (2005) summarized the latest developments of this research subject. Nevertheless, in the context of traffic networks, the various studies up to date focused on the case of deterministic UE, the inefficiency of SUE was, however, ignored so far.

This study is intended to make a contribution to the above emerging literature by determining the worst-case inefficiency of the logit-based SUE. The logit SUE model is an important one in transportation science that addresses suboptimal user route choices or difference in the costs perceived by different users. Before discussing the inefficiency of SUE, we should mention that there are two different system optimum definitions in a stochastic environment:

one is the aforementioned conventional SO which minimizes the total actual system travel time, and the other is the *stochastic system optimum* (SSO) which minimizes the total *perceived* travel time of all users (Maher et al., 2005; Stewart, 2007), or equivalently maximizes the net economic benefit (Yang, 1999). As a result, there are two ways to define the inefficiency of SUE, i.e. to compare SUE with SO in terms of total actual system travel time, or to compare SUE with SSO in terms of total perceived travel time (or equivalently in term of network economic benefit). We study the inefficiency of SUE in both situations.

The remainder of the paper is organized as follows. For completeness, Section 2 gives a brief review of bounding the inefficiency of deterministic UE. In Sections 3 and 4, we make use of the equivalent variational inequality (VI) formulation of the logit-based SUE, and compare SUE with SO and SSO, respectively, to bound its inefficiency with the two alternative definitions of total system travel time. In Section 5, we discuss the tightness of the inefficiency bounds established. Some concluding remarks are given in Section 6.

2. Review of bounding the inefficiency of deterministic UE

We consider a transportation network described as a strongly connected, directed network (N, A) where N and A denote the sets of nodes and links, respectively. Let W denote the set of all Origin-Destination (OD) pairs, R_w be the set of all paths between OD pair $w \in W$, d_w be the travel demand between OD pair $w \in W$, f_{rw} be the flow on path $r \in R_w$, $w \in W$, and v_a be the flow on link $a \in A$. The following relationships and constraints hold

$$v_a = \sum_{w \in W} \sum_{r \in R_w} f_{rw} \delta_{ar}, \quad a \in A \tag{1}$$

$$d_w = \sum_{r \in R_w} f_{rw}, \quad w \in W \tag{2}$$

$$f_{rw} \geq 0, \quad r \in R_w, \quad w \in W \tag{3}$$

where δ_{ar} is equal to 1 if path r uses link a and 0 otherwise. Let link flow vector be $\mathbf{v} = (v_a, a \in A)^{\mathrm{T}}$ and path flow vector be $\mathbf{f} = (f_{rw}, r \in R_w, w \in W)^{\mathrm{T}}$, then the feasible set of link flows is given by $\Omega_v = \{\mathbf{v} \mid \text{there exists an } \mathbf{f} \text{ such that (1)-(3) hold}\}$, and the feasible set of path flows is given by $\Omega_f = \left\{ \mathbf{f} \mid \text{constraints (2)-(3) hold} \right\}$. In this paper, we consider separable link cost (travel time) function $t_a(v_a)$, $a \in A$, which means that the travel time of one link depends on the flow on the link only. It is assumed that $t_a(v_a)$ is a nondecreasing function of v_a for all $a \in A$. Let c_{rw} be the travel time along path $r \in R_w$, $w \in W$, which is the sum of travel times on all links that constitute the path. We thus have

$$c_{rw} = \sum_{a \in A} t_a(v_a) \delta_{ar}, \quad r \in R_w, \quad w \in W \tag{4}$$

The total system travel time $T(\mathbf{v})$ is given by

$$T(\mathbf{v}) = \sum_{a \in A} t_a(v_a) v_a = \sum_{w \in W} \sum_{r \in R_w} c_{rw} f_{rw} \tag{5}$$

It is well known (e.g., Smith, 1979; Dafermos, 1980) that the UE problem can be formulated as an equivalent VI problem, namely to find $\mathbf{v}^{\text{ue}} \in \Omega_v$ such that

$$\sum_{a \in A} t_a\left(v_a^{\text{ue}}\right)\left(v_a^{\text{ue}} - v_a\right) \le 0, \text{ for any } \mathbf{v} \in \Omega_v \tag{6}$$

With the UE link flow solution $\mathbf{v}^{\text{ue}}$, the total system travel time under UE is given by $T_{\text{UE}} = T\left(\mathbf{v}^{\text{ue}}\right) = \sum_{a \in A} t_a\left(v_a^{\text{ue}}\right) v_a^{\text{ue}}$. On the other hand, the SO problem that minimizes the total system travel time is given by

$$\min_{\mathbf{v} \in \Omega_v} \sum_{a \in A} t_a(v_a) v_a \tag{7}$$

Let $\mathbf{v}^{\text{so}}$ denote the link flow solution to the SO problem, then the minimum system travel time is given by $T_{\text{SO}} = T\left(\mathbf{v}^{\text{so}}\right) = \sum_{a \in A} t_a\left(v_a^{\text{so}}\right) v_a^{\text{so}}$. Define the following ratio

$$\rho^{\text{ue}} = \frac{T_{\text{UE}}}{T_{\text{SO}}} = \frac{T\left(\mathbf{v}^{\text{ue}}\right)}{T\left(\mathbf{v}^{\text{so}}\right)} \tag{8}$$

Clearly, it holds that $\rho^{\text{ue}} \ge 1$. This ratio is called the *inefficiency*, or *price of anarchy*, of the selfish user equilibria (Papadimitriou, 2001).

The way to bound the inefficiency of UE is also used later when we analyze the SUE case. Therefore, for completeness, a brief outline of bounding ρ^{ue} is given here, based on the geometric proof due to Correa et al (2005).

The bounding method is based on the VI formulation (6). Let $\mathbf{v} \in \Omega_v$ be an arbitrary feasible link flow, then from VI (6) it follows

$$\sum_{a\in A} t_a\left(v_a^{\text{ue}}\right)v_a^{\text{ue}} \le \sum_{a\in A} t_a\left(v_a^{\text{ue}}\right)v_a = \sum_{a\in A} t_a\left(v_a\right)v_a + \sum_{a\in A}\left(t_a\left(v_a^{\text{ue}}\right) - t_a\left(v_a\right)\right)v_a \tag{9}$$

We now consider how to upper bound the last term of the right-hand side of (9) in terms of the left-hand side of (9). Note that each link cost function is nondecreasing and thus $\left(t_a\left(v_a^{\text{ue}}\right) - t_a\left(v_a\right)\right)v_a \le 0$ for $v_a \ge v_a^{\text{ue}}$, we only need to focus on the term $\left(t_a\left(v_a^{\text{ue}}\right) - t_a\left(v_a\right)\right)v_a$ for which $v_a < v_a^{\text{ue}}$. In this case, $\left(t_a\left(v_a^{\text{ue}}\right) - t_a\left(v_a\right)\right)v_a$ is equal to the area of the shaded rectangle in Figure 1, and $t_a\left(v_a^{\text{ue}}\right)v_a^{\text{ue}}$ is the area of the large rectangle in Figure 1.

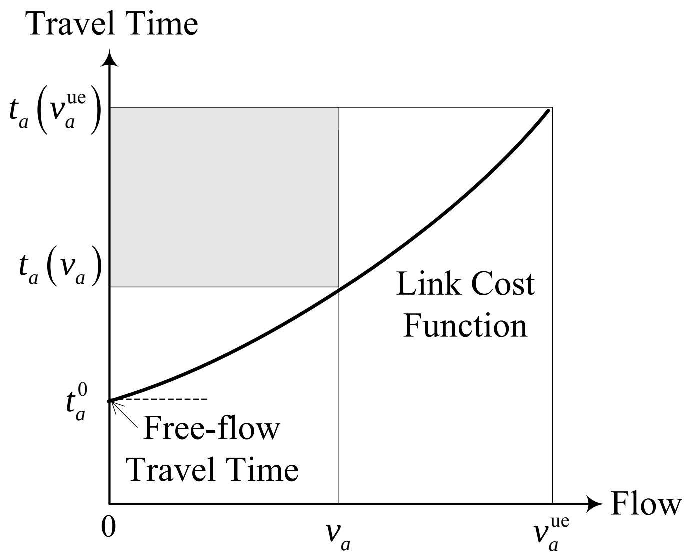

Figure 1. Geometric illustration of the definition of $\gamma(\mathcal{C})$

We need to upper bound the area of the shaded rectangle in terms of the area of the large rectangle. To do this, for each link cost function $t_a(\cdot)$ and nonnegative link flow $z_a \ge 0$, we define the following parameter

$$\gamma_a\left(t_a, z_a\right) = \max_{v_a \ge 0} \frac{\left(t_a\left(z_a\right) - t_a\left(v_a\right)\right)v_a}{t_a\left(z_a\right)z_a}, \quad a \in A \tag{10}$$

Here, $0/0 = 0$ by convention. Since $\left(t_a\left(z_a\right) - t_a\left(v_a\right)\right)v_a \le t_a\left(z_a\right)v_a$ if $0 \le v_a \le z_a$ and $\left(t_a\left(z_a\right) - t_a\left(v_a\right)\right)v_a \le 0$ if $v_a > z_a$, we have $0 \le \gamma_a\left(t_a, z_a\right) \le 1$. For a given class $\mathcal{C}$ of link cost functions (e.g., polynomials of a certain degree), we let

$$\gamma(\mathcal{C}) = \max_{t_a \in \mathcal{C},\, z_a \ge 0} \gamma_a(t_a, z_a) \tag{11}$$

With this definition, we have the following lemma.

Lemma 1. *Let* $\mathbf{v}^{\text{ue}}$ *be the UE link flow with separable link cost functions drawn from a given class* $\mathcal{C}$, *and let* $\mathbf{v}$ *be an arbitrary nonnegative link flow. Then*

$$\sum_{a \in A} \left(t_a\left(v_a^{\text{ue}}\right) - t_a(v_a) \right) v_a \le \gamma(\mathcal{C}) T\left(\mathbf{v}^{\text{ue}}\right) \tag{12}$$

Proof. From definitions (10) and (11), with z_a replaced by v_a^{ue}, we have

$$\begin{aligned} \sum_{a \in A} \left(t_a\left(v_a^{\text{ue}}\right) - t_a(v_a) \right) v_a &\le \sum_{a \in A} \gamma_a\left(t_a, v_a^{\text{ue}}\right) t_a\left(v_a^{\text{ue}}\right) v_a^{\text{ue}} \\ &\le \sum_{a \in A} \gamma(\mathcal{C}) t_a\left(v_a^{\text{ue}}\right) v_a^{\text{ue}} = \gamma(\mathcal{C}) T\left(\mathbf{v}^{\text{ue}}\right) \end{aligned}$$

which completes the proof. ♦

With Lemma 1, substituting (12) into (9) gives rise to

$$T\left(\mathbf{v}^{\text{ue}}\right) \le T(\mathbf{v}) + \gamma(\mathcal{C}) T\left(\mathbf{v}^{\text{ue}}\right), \text{ for any } \mathbf{v} \in \Omega_v \tag{13}$$

Let $\mathbf{v} = \mathbf{v}^{\text{so}}$ in (13), we have the following theorem.

Theorem 1. *Let* $\mathbf{v}^{\text{ue}}$ *be the UE link flow with separable link cost functions drawn from a given class* $\mathcal{C}$, *and* $\mathbf{v}^{\text{so}}$ *be an SO link flow, then*

$$\rho^{\text{ue}} = \frac{T\left(\mathbf{v}^{\text{ue}}\right)}{T\left(\mathbf{v}^{\text{so}}\right)} \le \frac{1}{1 - \gamma(\mathcal{C})} \tag{14}$$

Theorem 1 simply states that the upper bound, $\left(1 - \gamma(\mathcal{C})\right)^{-1}$ on the inefficiency, ρ^{ue}, of UE, or the worst-case inefficiency of UE is independent of the network topology but dependent on the class of cost functions only.

3. Bounding the inefficiency of SUE compared with SO

We now consider the case of stochastic user equilibrium. In a SUE model, each user is a utility-maximizer, and each path, r, $r \in R_w$, $w \in W$, is an alternative associated with some random utility function U_{rw}. A given path's utility is primarily related to its travel time, then U_{rw} is given by

$$U_{rw} = -\theta C_{rw} = -\theta c_{rw} + \xi_{rw}, \quad r \in R_w, \quad w \in W \tag{15}$$

where C_{rw} is the random perceived travel time along the path, θ is a positive unit scaling parameter, c_{rw} is the actual travel time along the path as defined before, $-\theta c_{rw}$ is the measured utility, and ξ_{rw} is a random term associated with the path under consideration and can be considered to represent the unobservable or unmeasurable factors of utility. Let P_{rw} denote the probability of users choosing path r, $r \in R_w$, $w \in W$, which is also the share of users choosing the path, then the utility maximization (perceived travel time minimization) principle implies that

$$P_{rw} = \Pr\left(U_{rw} \geq U_{kw}, \forall\, k \in R_w\right), \quad r \in R_w, \quad w \in W \tag{16}$$

This choice probability has the following properties

$$0 \leq P_{rw} \leq 1, \quad r \in R_w \quad \text{and} \quad \sum_{r \in R_w} P_{rw} = 1, \quad w \in W$$

If the random term ξ_{rw} in (15) is assumed to be normally distributed, one would obtain the probit-based route choice model. However, the probit-based model does not entail a closed-form expression of the path choice probability and thus makes our subsequent analysis of inefficiency analytically intractable. Hence we consider the logit-based route choice model only. The logit-based model assumes that the random terms of the utility functions associated with all paths are independently and identically distributed Gumbel random variables. The choice probability is then given by

$$P_{rw} = \frac{\exp\left(-\theta c_{rw}\right)}{\sum_{k \in R_w} \exp\left(-\theta c_{kw}\right)}, \quad r \in R_w, \quad w \in W \tag{17}$$

and the path flow assignment is given by

$$f_{rw} = d_w P_{rw}, \quad r \in R_w, \quad w \in W \tag{18}$$

It is well-known (Fisk, 1980) that the above logit-based SUE model can be formulated as the following equivalent minimization problem

$$\min_{\mathbf{f}\in\Omega_f} Z(\mathbf{f}) = \sum_{a\in A}\int_0^{v_a} t_a(\omega)\mathrm{d}\omega + \frac{1}{\theta}\sum_{w\in W}\sum_{r\in R_w} f_{rw}\ln f_{rw} \tag{19}$$

Denote SUE path flow as $\mathbf{f}^{\text{sue}} \in \Omega_f$, and the corresponding SUE link flow $\mathbf{v}^{\text{sue}} \in \Omega_v$, the total system travel time under SUE is given by $T_{\text{SUE}} = \sum_{w\in W}\sum_{r\in R_w} c_{rw}\left(\mathbf{f}^{\text{sue}}\right) f_{rw}^{\text{sue}} = \sum_{a\in A} t_a\left(v_a^{\text{sue}}\right) v_a^{\text{sue}}$.

Similar to that of UE, the *inefficiency* of SUE compared with SO is defined as

$$\rho^{\text{sue}} = \frac{T_{\text{SUE}}}{T_{\text{SO}}} \tag{20}$$

To find an upper bound on the inefficiency of SUE (compared with SO here or compared with SSO later), we need the equivalent VI formulation for the logit-based SUE model, which is given in the following lemma.

Lemma 2. *If the separable link cost function, $t_a(v_a)$, $a \in A$, is monotonically increasing with link flow, a logit-based SUE problem with fixed OD demand is equivalent to the following variational inequality, i.e., find $\mathbf{f}^{\text{sue}} \in \Omega_f$, such that*

$$\sum_{w\in W}\sum_{r\in R_w}\left(c_{rw}\left(\mathbf{f}^{\text{sue}}\right) + \frac{1}{\theta}\ln f_{rw}^{\text{sue}}\right)\left(f_{rw} - f_{rw}^{\text{sue}}\right) \geq 0, \text{ for any } \mathbf{f} \in \Omega_f \tag{21}$$

Proof. It suffices to prove that minimization problem (19) is equivalent to VI (21). With the assumption of monotonically increasing link cost function, problem (19) of minimizing a strictly convex function over a compact (closed and bounded) set guarantees the existence and uniqueness of a path flow solution $\mathbf{f}^{\text{sue}} \in \Omega_f$. In addition, the entropy-type objective function ensures that the optimum is achieved at an interior point. A necessary and sufficient condition for $\mathbf{f}^{\text{sue}} \in \Omega_f$ to be the unique optimal solution to problem (19) is that

$$\left[\nabla_{\mathbf{f}} Z\left(\mathbf{f}^{\text{sue}}\right)\right]^{\mathrm{T}}\left(\mathbf{f} - \mathbf{f}^{\text{sue}}\right) \geq 0, \text{ for any } \mathbf{f} \in \Omega_f \tag{22}$$

Using $v_a = \sum_{w\in W}\sum_{r\in R_w} f_{rw}\delta_{ar}$, substituting

$$\left[\nabla_{\mathbf{f}} Z\left(\mathbf{f}^{\text{sue}}\right)\right]^{\mathrm{T}} = \left[\cdots, \frac{1}{\theta} + \frac{1}{\theta}\ln f_{rw}^{\text{sue}} + \sum_{a\in A} t_a\left(v_a^{\text{sue}}\right)\delta_{ar}, \cdots\right] = \left[\cdots, \frac{1}{\theta} + \frac{1}{\theta}\ln f_{rw}^{\text{sue}} + c_{rw}\left(\mathbf{f}^{\text{sue}}\right), \cdots\right]$$

into (22), and in view of $\sum_{w\in W}\sum_{r\in R_w}\frac{1}{\theta}\left(f_{rw} - f_{rw}^{\text{sue}}\right) = \sum_{w\in W}\frac{1}{\theta}\left(d_w - d_w\right) = 0$, we have VI (21). This completes the proof. ♦

With the equivalent VI formulation of the logit-based SUE given in Lemma 2, we can move on to bound the inefficiency of SUE. Let $\mathbf{f}^{\text{so}} \in \Omega_f$ be a system-optimal path flow, and let $\mathbf{f} = \mathbf{f}^{\text{so}}$ in VI (21), we have

$$\sum_{w \in W} \sum_{r \in R_w} \left(c_{rw}\left(\mathbf{f}^{\text{sue}}\right) + \frac{1}{\theta} \ln f_{rw}^{\text{sue}} \right) \left(f_{rw}^{\text{so}} - f_{rw}^{\text{sue}} \right) \geq 0 \tag{23}$$

Denote $c_{rw}^{\text{sue}} = c_{rw}\left(\mathbf{f}^{\text{sue}}\right)$ and $c_{rw}^{\text{so}} = c_{rw}\left(\mathbf{f}^{\text{so}}\right)$, (23) gives

$$\sum_{w \in W} \sum_{r \in R_w} c_{rw}^{\text{sue}} f_{rw}^{\text{sue}} \leq \sum_{w \in W} \sum_{r \in R_w} c_{rw}^{\text{sue}} f_{rw}^{\text{so}} + \frac{1}{\theta} \sum_{w \in W} \sum_{r \in R_w} \left(f_{rw}^{\text{so}} - f_{rw}^{\text{sue}} \right) \ln f_{rw}^{\text{sue}}$$

which is equivalent to

$$T_{\text{SUE}} \leq T_{\text{SO}} + \sum_{w \in W} \sum_{r \in R_w} \left(c_{rw}^{\text{sue}} - c_{rw}^{\text{so}} \right) f_{rw}^{\text{so}} + \frac{1}{\theta} \sum_{w \in W} \sum_{r \in R_w} \left(f_{rw}^{\text{so}} - f_{rw}^{\text{sue}} \right) \ln f_{rw}^{\text{sue}} \tag{24}$$

In view of

$$\sum_{w \in W} \sum_{r \in R_w} \left(c_{rw}^{\text{sue}} - c_{rw}^{\text{so}} \right) f_{rw}^{\text{so}} = \sum_{a \in A} \left(t_a\left(v_a^{\text{sue}}\right) - t_a\left(v_a^{\text{so}}\right) \right) v_a^{\text{so}}$$

we can rewrite (24) as

$$T_{\text{SUE}} \leq T_{\text{SO}} + \sum_{a \in A} \left(t_a\left(v_a^{\text{sue}}\right) - t_a\left(v_a^{\text{so}}\right) \right) v_a^{\text{so}} + \frac{1}{\theta} \sum_{w \in W} \sum_{r \in R_w} \left(f_{rw}^{\text{so}} - f_{rw}^{\text{sue}} \right) \ln f_{rw}^{\text{sue}} \tag{25}$$

With parameters $\gamma_a\left(t_a, z_a\right)$, $a \in A$ and $\gamma(\mathcal{C})$ defined by (10) and (11), respectively, we have similar result as in Lemma 1

$$\sum_{a \in A} \left(t_a\left(v_a^{\text{sue}}\right) - t_a\left(v_a\right) \right) v_a \leq \gamma(\mathcal{C}) T_{\text{SUE}}, \text{ for any } \mathbf{v} \geq 0 \tag{26}$$

Let $\mathbf{v} = \mathbf{v}^{\text{so}}$ in (26), then we obtain

$$\sum_{a \in A} \left(t_a\left(v_a^{\text{sue}}\right) - t_a\left(v_a^{\text{so}}\right) \right) v_a^{\text{so}} \leq \gamma(\mathcal{C}) T_{\text{SUE}} \tag{27}$$

With (27), we have an upper bound on the second term of the right-hand side of (25) in terms of T_{SUE}. Now we seek an upper bound on the third term.

Lemma 3. *Consider the following maximization problem*

$$\max \ Z(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \left(y_i - x_i \right) \ln x_i \tag{28}$$

subject to

$$\sum_{i=1}^{n} x_i = d \tag{29}$$

$$\sum_{i=1}^{n} y_i = d \tag{30}$$

$$x_i, y_i \geq 0,\ i = 1, 2, \cdots, n \tag{31}$$

where $d > 0$ *is a constant. The optimal value of this problem is* $Z_{\max} = kd$, *where* k *solves equation* $ke^{k+1} = n-1$, *with* e *being the base of natural logarithm.*

Proof: Note that the objective function (28) is linear in $\mathbf{y}$, and in view of the constraints on $\mathbf{y}$, it is not difficult to see that the optimal vector $\mathbf{y}$ has one component equal to d and the others equal to 0 (if the optimal vector $\mathbf{y}$ is not unique, there exists such an optimal corner solution). Without loss of generality, let $\mathbf{y} = (d, 0, ..., 0)^{\mathrm{T}}$ be an optimal $\mathbf{y}$ vector, then the objective function (28) is simplified to be

$$\max\ Z(\mathbf{x}) = (d - x_1)\ln x_1 - \sum_{i=2}^{n} x_i \ln x_i \tag{32}$$

and the KKT necessary conditions for the optimality of $\mathbf{x}$ are

$$\frac{\partial Z(\mathbf{x})}{\partial x_i} + \lambda \leq 0,\quad x_i \geq 0,\quad \left(\frac{\partial Z(\mathbf{x})}{\partial x_i} + \lambda\right) x_i = 0,\quad i = 1, 2, ...n$$

where λ is the Lagrange multiplier associated with the equality constraint (29). In view of $\partial Z(\mathbf{x})/\partial x_1 = -\ln x_1 + d/x_1 - 1$ and $\partial Z(\mathbf{x})/\partial x_i = -\ln x_i - 1$, $i = 2, ...n$, the optimal solution must have $x_i > 0,\ i = 1, 2, \cdots, n,$ because $x_i = 0$ would give $\partial Z(\mathbf{x})/\partial x_i = +\infty$, which obviously violates the optimality condition. Then the KKT conditions reduce to

$$-\ln x_1 + \frac{d}{x_1} - 1 + \lambda = 0,\quad x_1 > 0 \tag{33}$$

$$-\ln x_i - 1 + \lambda = 0,\quad x_i > 0,\quad i = 2, \cdots, n \tag{34}$$

Combining (33) and (34), we have

$$\frac{d}{x_1} = \ln \frac{x_1}{x_2} \tag{35}$$

$$x_i = x_2,\quad i = 3, \cdots, n \tag{36}$$

Substituting (35) into objective function (32), we have the optimal objective value

$$Z_{\max} = (d - x_1)\ln x_1 - (n-1) x_2 \ln x_2 \tag{37}$$

In view of $(n-1)x_2 = d - x_1$, and making use of (35), we obtain

$$Z_{\max} = d\left(\frac{d}{x_1} - 1\right) \tag{38}$$

Let $k = d/x_1 - 1$, then we have $Z_{\max} = kd$, and

$$x_1 = \frac{1}{1+k}d\ ,\ \ x_2 = \frac{k}{(n-1)(1+k)}d \tag{39}$$

Substituting (39) into (35) gives that k solves $ke^{k+1} = n-1$. This completes the proof. ♦

From Lemma 3, it follows immediately that

$$\sum_{r\in R_w}\left(f_{rw}^{\text{so}} - f_{rw}^{\text{sue}}\right)\ln f_{rw}^{\text{sue}} \le k_w d_w,\ w\in W \tag{40}$$

where k_w solves $k_w e^{k_w+1} = |R_w| - 1$ and $|R_w|$ is the number of feasible paths between OD pair $w \in W$. Substituting (27) and (40) into (25) yields

$$T_{\text{SUE}} \le T_{\text{SO}} + \gamma(\mathcal{C})T_{\text{SUE}} + \frac{1}{\theta}\bar{k}D \tag{41}$$

where $D = \sum_{w\in W} d_w$ is the total traffic demand and $\bar{k}$ is the average of $k_w,\ w\in W$ weighted by OD demand:

$$\bar{k} = \sum_{w\in W}\left(\frac{d_w}{D}\right)k_w,\ \text{ where } k_w \text{ solves } k_w e^{k_w+1} = |R_w| - 1 \tag{42}$$

If we define $\bar{c} = T_{\text{SO}}/D$ as the average travel time of all network users at system optimum, then (41) can be rewritten as

$$T_{\text{SUE}} \le T_{\text{SO}} + \gamma(\mathcal{C})T_{\text{SUE}} + \frac{1}{\theta\bar{c}}\bar{k}T_{\text{SO}}$$

which gives rise to

$$T_{\text{SUE}} \le \left(\frac{1}{1-\gamma(\mathcal{C})}\right)\left(1 + \frac{1}{\theta\bar{c}}\bar{k}\right)T_{\text{SO}} \tag{43}$$

The term $1/\theta\bar{c}$ in (43) needs to be further addressed. The logit model parameter θ, in its original meaning, is inversely proportional to the standard error of the distribution of the perceived path travel times (Sheffi, 1985), and the logit model assumes that all paths in the network has the same standard error. Specifically, $\theta = \pi / (\sqrt{6}\sigma)$, where σ is the common standard deviation of the perceived path travel times. Then we have

$$\frac{1}{\theta \bar{c}} = \frac{\sqrt{6}\sigma}{\pi \bar{c}} \tag{44}$$

To provide more sensible results, we define $\bar{c}_0$ as the average free-flow travel time for all OD pairs, then it is clear that $\bar{c}_0 \le \bar{c}$. Furthermore, we define a ratio $\varsigma = \sigma / \bar{c}_0$. Then replacing $\bar{c}$ with $\bar{c}_0$ in (44) yields

$$\frac{1}{\theta \bar{c}} \le \frac{\sqrt{6}\sigma}{\pi \bar{c}_0} = \frac{\sqrt{6}}{\pi}\varsigma \tag{45}$$

where $\varsigma = \sigma / \bar{c}_0$ measures the standard deviation of perceived travel time as percentage of the average free-flow travel time. Intuitively, the ratio $\varsigma = \sigma / \bar{c}_0$ is like a coefficient of variance, which represents the relative travel time perception error of users. Considering that in reality, users' absolute perception error σ may increase as the path travel time increases, the relative error ς may better reflect users' perception randomness.

Substituting (45) into (43), we have the following theorem.

Theorem 2 *Let T_{SUE} be the total system travel time under logit-based stochastic user equilibrium, and T_{SO} be the minimum total system travel time, then*

$$\rho^{\text{sue}} = \frac{T_{\text{SUE}}}{T_{\text{SO}}} \le \left(\frac{1}{1-\gamma(\mathcal{C})}\right)\left(1 + \frac{\sqrt{6}}{\pi}\bar{k}\varsigma\right) \tag{46}$$

Comparing (46) with (14), we find that the bound of ρ^{sue} is generally larger than the bound of ρ^{ue}, which means that the worst-case inefficiency of SUE is generally worse than that of UE. Note that this comparison is made in the respect of the worst-case inefficiency, a specific SUE can be more efficient than UE on a network.

The bounding result given by (46) depends on three dimensionless parameters, namely $\gamma(\mathcal{C})$, $\bar{k}$ and ς. As mentioned for the deterministic UE case, $\gamma(\mathcal{C})$ is defined exclusively by the class of link cost functions, $\bar{k}$ (or more essentially, k_w) is determined by the number of available paths and thus reflects the degree of network complexity, and ς represents the relative travel time perception error of users. If users' perception error is zero or the travel time is deterministic, then $\varsigma = 0$ and (46) becomes $\rho^{\text{sue}} \le \left(1-\gamma\left(\mathcal{C}\right)\right)^{-1}$, thus we have the inefficiency bound of the standard deterministic UE. In addition, if $|R_w| = 1$ for all $w \in W$, i.e. each OD pair has only one feasible path and hence travel time perception error has no effect on route choice, we have $k_w = 0,\ w \in W$ from (42) and thus $\bar{k} = 0$. In this case, we have $\rho^{\text{sue}} \le \left(1-\gamma\left(\mathcal{C}\right)\right)^{-1}$ as well.

Whenever $|R_w| > 1$, the value of k_w (and hence $\bar{k}$) is very limited. As seen in Table 1, k_w is only marginally larger than $\log_{10}|R_w|$. For a sufficiently complex network with the number of paths between each OD pair being $100 \le |R_w| \le 1000$, k_w takes only a limited value between 2.63 and 4.42. This observation shows that the network size has very limited effect on the cost inefficiency bound of stochastic user equilibrium.

Table 1. Numerical values of k_w with increasing $|R_w|$

$\|R_w\|$	1	10	10^2	10^3	10^4
k_w	0	1.10	2.63	4.42	6.36

4. Bounding the inefficiency of SUE compared with SSO

In a stochastic traffic assignment environment, besides the total actual system travel time, the total *perceived* travel time of all users is also a useful system performance index, as it reflects the net economic benefit. For the logit-based stochastic traffic assignment model, the total perceived travel time of all users can be given in a closed-form expression as (Maher et al. 2005)

$$F(\mathbf{f}) = \sum_{a\in A} t_a(v_a)v_a + \frac{1}{\theta}\sum_{w\in W}\sum_{r\in R_w} f_{rw}\ln f_{rw} - \frac{1}{\theta}\sum_{w\in W} d_w \ln d_w \tag{47}$$

and its opposite, $-F(\mathbf{f})$, can be regarded as the net economic benefit (the direct utility corresponding to the aggregate demand minus the total travel time incurred by all users in the network) (Yang, 1999). The stochastic system optimization (SSO) problem is to minimize the total perceived travel time (or equivalently, to maximize the net economic benefit), namely

$$\min_{\mathbf{f}\in\Omega_f} F(\mathbf{f}) \tag{48}$$

Let $\mathbf{f}^{\text{sso}} \in \Omega_f$ solve the SSO problem (48), then $F_{\text{SSO}} = F(\mathbf{f}^{\text{sso}})$ is the minimum total perceived travel time of all users, and correspondingly, $-F_{\text{SSO}}$ is the maximum net economic benefit (or consumer surplus) of the network. Specifically,

$$F_{\text{SSO}} = \sum_{a\in A} t_a\left(v_a^{\text{sso}}\right)v_a^{\text{sso}} + \frac{1}{\theta}\sum_{w\in W}\sum_{r\in R_w} f_{rw}^{\text{sso}} \ln f_{rw}^{\text{sso}} - \frac{1}{\theta}\sum_{w\in W} d_w \ln d_w \tag{49}$$

On the other hand, uncoordinated selfish travel behaviors of users will result in an SUE flow pattern $\mathbf{f}^{\text{sue}} \in \Omega_f$, with $F_{\text{SUE}} = F(\mathbf{f}^{\text{sue}})$ being the total perceived travel time of all users at equilibrium. Specifically

$$F_{\text{SUE}} = \sum_{a\in A} t_a\left(v_a^{\text{sue}}\right)v_a^{\text{sue}} + \frac{1}{\theta}\sum_{w\in W}\sum_{r\in R_w} f_{rw}^{\text{sue}} \ln f_{rw}^{\text{sue}} - \frac{1}{\theta}\sum_{w\in W} d_w \ln d_w \tag{50}$$

By definition, we have $F_{\text{SUE}} \ge F_{\text{SSO}}$. However, unlike the previous cases, in which we use the ratios $T_{\text{UE}}/T_{\text{SO}} \ge 1$ and $T_{\text{SUE}}/T_{\text{SO}} \ge 1$ to measure the inefficiency of UE and SUE compared with SO, here we can not use the ratio $F_{\text{SUE}}/F_{\text{SSO}}$ to measure the inefficiency of SUE compared with SSO. The reason is that F_{SUE} and F_{SSO} may be negative as can be seen from (49)-(50), which means that the ratio $F_{\text{SUE}}/F_{\text{SSO}}$ may be meaningless (consider the case $F_{\text{SUE}} > 0 > F_{\text{SSO}}$). Consequently, instead of using the ratio of F_{SUE} to F_{SSO}, we shall use the difference between F_{SUE} and F_{SSO}, namely the term $F_{\text{SUE}} - F_{\text{SSO}} \ge 0$, which is the absolute efficiency loss of SUE compared with SSO. We can have a clearer understanding of the term $\left(F_{\text{SUE}} - F_{\text{SSO}}\right)$ from an economic viewpoint: since $\left(-F_{\text{SSO}}\right)$ and $\left(-F_{\text{SUE}}\right)$ are, respectively, the maximum possible net economic benefit and the net economic benefit realized at equilibrium, the difference between the two, equal to $\left(F_{\mathrm{SUE}}-F_{\mathrm{SSO}}\right)$, naturally represents the welfare loss caused by uncoordinated selfish routing behaviors of users.

To bound the inefficiency of SUE compared with SSO, we shall give an upper bound on the welfare loss $\left(F_{\mathrm{SUE}}-F_{\mathrm{SSO}}\right)$ in terms of some meaningful measure. To this end, we make use of the equivalent VI formulation of the logit-based SUE. The manipulation is quite similar to that in last section. Specifically, let $\mathbf{f}=\mathbf{f}^{\mathrm{sso}}$ in VI (21), we have

$$\sum_{w\in W}\sum_{r\in R_w}\left(c_{rw}\left(\mathbf{f}^{\mathrm{sue}}\right)+\frac{1}{\theta}\ln f_{rw}^{\mathrm{sue}}\right)\left(f_{rw}^{\mathrm{sso}}-f_{rw}^{\mathrm{sue}}\right)\geq 0 \tag{51}$$

Denote $c_{rw}^{\mathrm{sue}}=c_{rw}\left(\mathbf{f}^{\mathrm{sue}}\right)$ and $c_{rw}^{\mathrm{sso}}=c_{rw}\left(\mathbf{f}^{\mathrm{sso}}\right)$, (51) gives

$$\sum_{w\in W}\sum_{r\in R_w}c_{rw}^{\mathrm{sue}}f_{rw}^{\mathrm{sue}}+\frac{1}{\theta}\sum_{w\in W}\sum_{r\in R_w}f_{rw}^{\mathrm{sue}}\ln f_{rw}^{\mathrm{sue}}\leq\sum_{w\in W}\sum_{r\in R_w}c_{rw}^{\mathrm{sue}}f_{rw}^{\mathrm{sso}}+\frac{1}{\theta}\sum_{w\in W}\sum_{r\in R_w}f_{rw}^{\mathrm{sso}}\ln f_{rw}^{\mathrm{sue}}$$

which is equivalent to

$$F_{\mathrm{SUE}}\leq F_{\mathrm{SSO}}+\sum_{w\in W}\sum_{r\in R_w}\left(c_{rw}^{\mathrm{sue}}-c_{rw}^{\mathrm{sso}}\right)f_{rw}^{\mathrm{sso}}+\frac{1}{\theta}\sum_{w\in W}\sum_{r\in R_w}f_{rw}^{\mathrm{sso}}\left(\ln f_{rw}^{\mathrm{sue}}-\ln f_{rw}^{\mathrm{sso}}\right) \tag{52}$$

In view of

$$\sum_{w\in W}\sum_{r\in R_w}\left(c_{rw}^{\mathrm{sue}}-c_{rw}^{\mathrm{sso}}\right)f_{rw}^{\mathrm{sso}}=\sum_{a\in A}\left(t_a\left(v_a^{\mathrm{sue}}\right)-t_a\left(v_a^{\mathrm{sso}}\right)\right)v_a^{\mathrm{sso}}$$

we can rewrite (52) as

$$F_{\mathrm{SUE}}-F_{\mathrm{SSO}}\leq\sum_{a\in A}\left(t_a\left(v_a^{\mathrm{sue}}\right)-t_a\left(v_a^{\mathrm{sso}}\right)\right)v_a^{\mathrm{sso}}+\frac{1}{\theta}\sum_{w\in W}\sum_{r\in R_w}f_{rw}^{\mathrm{sso}}\left(\ln f_{rw}^{\mathrm{sue}}-\ln f_{rw}^{\mathrm{sso}}\right) \tag{53}$$

Now we only need to provide upper bounds on the two terms of the right-hand side of (53). For the first term, its upper bound can be obtained in the same way as we obtain (27) in last section. Specifically, let $\mathbf{v}=\mathbf{v}^{\mathrm{sso}}$ in (26), we simply have

$$\sum_{a\in A}\left(t_a\left(v_a^{\mathrm{sue}}\right)-t_a\left(v_a^{\mathrm{sso}}\right)\right)v_a^{\mathrm{sso}}\leq\gamma\left(\mathcal{C}\right)T_{\mathrm{SUE}} \tag{54}$$

From Gibbs' inequality (or the property of the Kullback-Leibler divergence between two discrete probability distributions), we readily have a zero upper bound of the second term of the right-hand side of (53). Namely, the following inequality holds

$$\sum_{r\in R_w} f_{rw}^{\text{sso}}\left(\ln f_{rw}^{\text{sue}} - \ln f_{rw}^{\text{sso}}\right) \le 0,\ w\in W \tag{55}$$

with equality if and only if $f_{rw}^{\text{sue}} = f_{rw}^{\text{sso}},\ r\in R_w,\ w\in W$. Substituting (54) and (55) into (53) gives the following theorem.

Theorem 3 *For a logit-based stochastic traffic assignment model, let* T_{SUE} *and* F_{SUE} *be, respectively, the total actual system travel time and the total perceived travel time of all users under stochastic user equilibrium, and let* F_{SSO} *be the minimum possible total perceived travel time of all users, then*

$$\frac{F_{\text{SUE}} - F_{\text{SSO}}}{T_{\text{SUE}}} \le \gamma\left(\mathcal{C}\right) \tag{56}$$

Theorem 3 states that the welfare loss, $F_{\text{SUE}} - F_{\text{SSO}}$, of SUE compared with SSO is not larger than a fraction of the total actual system travel time, T_{SUE}, under SUE. Like the deterministic UE case, the fraction or $\gamma\left(\mathcal{C}\right)$ is independent of network topology, but depends solely on the class of link cost functions.

5. On the tightness of the inefficiency bounds

We begin with our discussion of the tightness of the inefficiency bounds by presenting all the results in Theorems 1-3 into the following similar expressions for a comparison.

Theorem 1: bounding the efficiency loss of deterministic UE

$$\frac{T_{\text{UE}} - T_{\text{SO}}}{T_{\text{UE}}} \le \gamma\left(\mathcal{C}\right) \tag{57}$$

Theorem 2: bounding the efficiency loss of SUE compared with SO

$$\frac{T_{\mathrm{SUE}}-T_{\mathrm{SO}}}{T_{\mathrm{SUE}}} \le \frac{\gamma(\mathcal{C})+\sqrt{6}\bar{k}\varsigma/\pi}{1+\sqrt{6}\bar{k}\varsigma/\pi} \tag{58}$$

Theorem 3: bounding the welfare loss of SUE compared with SSO

$$\frac{F_{\mathrm{SUE}}-F_{\mathrm{SSO}}}{T_{\mathrm{SUE}}} \le \gamma(\mathcal{C}) \tag{59}$$

All the three major results involve the common parameter, $\gamma(\mathcal{C})$, that depends on the given class, $\mathcal{C}$, of link cost functions under consideration. Specific expression of $\gamma(\mathcal{C})$ can be obtained for the following practical class of link cost functions:

$$t_a(v_a) = t_a^0 + \alpha_a (v_a)^p,\ a \in A \tag{60}$$

where t_a^0 is a constant free-flow travel time, $\alpha_a \ge 0$ is a link-specific non-negative parameter, and $p \ge 0$ reflects the degree of congestion sensitivity of the link costs. In this case one can easily obtain (Roughgarden, 2005)

$$\gamma(\mathcal{C}) = \left(\frac{p}{p+1}\right)\left(\frac{1}{p+1}\right)^{\frac{1}{p}} \tag{61}$$

with $\gamma(\mathcal{C}) \to 0$ as $p \to 0$ (without traffic congestion), and $\gamma(\mathcal{C}) \to 1$ as $p \to +\infty$ (with severe congestion). For the widely used BPR type link cost function with $p = 4$, we have $\gamma(\mathcal{C}) = 0.5350$.

The bound in (57) for the deterministic UE is tight and can be furnished by a simple example with one OD pair connected by two parallel links (Roughgarden, 2003). Let the link travel time functions be $t_1(v_1) = 1$ and $t_2(v_2) = (v_2)^p$, $p > 0$, and let the OD demand be $d = v_1 + v_2 = 1$. The UE solution is $v_1^{\mathrm{ue}} = 0$ and $v_2^{\mathrm{ue}} = 1$ with $T_{\mathrm{UE}} = 1$. The SO solution is $v_1^{\mathrm{so}} = 1 - (p+1)^{-1/p}$ and $v_2^{\mathrm{so}} = (p+1)^{-1/p}$ with $T_{\mathrm{SO}} = 1 - p(p+1)^{-(p+1)/p} < 1$. Therefore, $(T_{\mathrm{UE}} - T_{\mathrm{SO}})/T_{\mathrm{UE}} = p(p+1)^{-(p+1)/p}$, which is consistent with $\gamma(\mathcal{C})$ given in (61).

The bound in (58) for SUE is somewhat complicated, because the system inefficiency under SUE is due to the combined effects of congestion externality and users' perception randomness, which, in general, cannot be decoupled. This also renders that it is generally difficult to find a specific instance to furnish the established upper bound in a general congested network, which requires inequality (27) and (40) to take equality simultaneously. Nevertheless, we can shed some light on the tightness of the inefficiency bound by considering certain special cases. First, when users' perception error is zero ($\theta \to +\infty$ or $\varsigma = \sigma/\bar{c}_0 = 0$, as mentioned earlier), SUE reduces to the deterministic UE and the bound (58) reduces to (57), and hence it is tight. Second, in a network without congestion or when link costs are all constants, the system inefficiency due to congestion externality becomes immaterial and the system efficiency loss is solely due to users' perception randomness. In this case we have $\gamma(\mathcal{C}) = 0$ and the bound in (58) reduces to

$$T_{\text{SUE}} \le \left(1 + \frac{\sqrt{6}\bar{k}\varsigma}{\pi}\right) T_{\text{SO}} \tag{62}$$

One can easily construct a simple example for which equality holds for (62). Consider a network having one OD pair connected by 10 parallel links as shown in Figure 2. In this case $\bar{k}$ solves for $ke^{k+1} = n - 1$ with $n = 10$ (or $\bar{k} = 1.101$). Suppose $\varsigma = \sigma/\bar{c}_0 = 0.5$ in the logit-based route choice model, where $\bar{c}_0$, by definition, is the (minimum) free-flow OD travel time between the single OD pair. Without loss of generality, we can simply let link 1 has the shortest constant travel time and let $t_1 = 1.0$. Thus, $\sigma = \varsigma\bar{c}_0 = 0.5$ and the logit model parameter $\theta = \pi/\left(\sqrt{6}\sigma\right) = 2.565$. Recall the proof of Lemma 3, to make (62) an equality it suffices to construct an appropriate constant travel time $t_2 = t_3 = \cdots = t_{10} > 1.0$ such that the resulting SUE link flow equal to the optimal solution given by (39), i.e. $v_1 = d/(1+k)$ and $v_2 = v_3 = \cdots = v_{10} = d\,k/(n-1)(1+k)$. Let the OD demand be $d = 1$, we have $v_1 = 0.4760$ and $v_2 = v_3 = \ldots = v_{10} = 0.0582$. Indeed this SUE link flow pattern can be generated by choosing $t_2 = t_3 = \cdots = t_{10} = 1.819$ with $\theta = 2.565$ in the logit model. The corresponding $T_{\text{SUE}} = v_1 t_1 + 9 v_2 t_2 = 1.429$. In view of $T_{\text{SO}} = t_1 \times d = 1.0$, we have $T_{\text{SUE}} = 1.429 T_{\text{SO}}$. This exactly attains the equality in (62), which, in this specific example, is $T_{\text{SUE}} \le \left(1 + \sqrt{6} \times 1.101 \times 0.5/\pi\right) T_{\text{SO}} = 1.429 T_{\text{SO}}$.

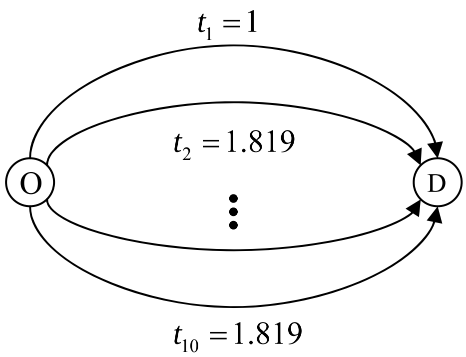

Figure 2. A network such that the bound in Theorem 2 is tight

We now move on to examine the tightness of the bound given in (59), which compares SUE with SSO in terms of total perceived system travel time. We give the following corollary based on Theorem 3.

Corollary 1. $F_{\text{SUE}} - F_{\text{SSO}} = \gamma(\mathcal{C}) T_{\text{SUE}}$ *if and only if* $\gamma(\mathcal{C}) = 0$.

Proof: First if $\gamma(\mathcal{C}) = 0$, then Theorem 3 gives $F_{\text{SUE}} - F_{\text{SSO}} \le 0$. Note that we always have $F_{\text{SUE}} \ge F_{\text{SSO}}$ by the definition of SSO. Thus we have $F_{\text{SUE}} = F_{\text{SSO}}$, which simply gives $F_{\text{SUE}} - F_{\text{SSO}} = \gamma(\mathcal{C}) T_{\text{SUE}}$ in view of $\gamma(\mathcal{C}) = 0$. On the other hand, suppose that $F_{\text{SUE}} - F_{\text{SSO}} = \gamma(\mathcal{C}) T_{\text{SUE}}$ holds. From the derivation of Theorem 3, $F_{\text{SUE}} - F_{\text{SSO}} = \gamma(\mathcal{C}) T_{\text{SUE}}$ means that inequality (55) takes equality, and thus $\mathbf{f}^{\text{sue}} = \mathbf{f}^{\text{sso}}$ and $F_{\text{SUE}} = F_{\text{SSO}}$, which simply gives $\gamma(\mathcal{C}) = 0$ from $F_{\text{SUE}} - F_{\text{SSO}} = \gamma(\mathcal{C}) T_{\text{SUE}}$. This completes the proof. ♦

Corollary 1 states that the bound in (59) is tight if and only if link costs are constants or $\gamma(\mathcal{C}) = 0$. In this case, $\mathbf{f}^{\text{sue}} = \mathbf{f}^{\text{sso}}$ or the SUE is fully efficient in terms of minimizing total perceived system travel time. This can be also seen from the SUE objective function (19) and the SSO objective function (47): when link costs are constants, the two objective functions are equivalent (the third term of (47) is constant and thus can be omitted). For a general class of link cost functions (networks with congestion effects), $\gamma(\mathcal{C}) \neq 0$, the inefficiency bound (59) is not tight or we always have $F_{\text{SUE}} - F_{\text{SSO}} < \gamma(\mathcal{C}) T_{\text{SUE}}$. Even so, this bounding result is still attractive in view of the property that $\gamma(\mathcal{C})$ is independence of network topology.

We conclude this section by offering a remark on the choice of the two social optimum concepts (SO and SSO) and the corresponding SUE inefficiency bounds given by (58) and (59). Clearly, choice of either one depends on the source (or the analyst's interpretation) of users' perception randomness. If users' perception randomness is due to imprecise information about the actual travel times, then the inefficiency or deviation of SUE should be measured against the deterministic SO. If, however, users' perception randomness is due to their different tastes or preferences for diversity in route choice, then SSO should be taken as the optimum criterion, because users' variety-seeking behaviors are considered as a fraction of the net economic benefit.

6. Conclusion

We have defined the inefficiency of SUE in two different ways, i.e. comparing SUE with SO in terms of total actual system travel time, or comparing SUE with SSO in terms of total perceived system travel time. For both notions, we provided upper bounds on the inefficiency of the logit-based SUE, based on its equivalent VI formulation and the properties of the divergence between two discrete path flow distributions.

When comparing SUE with SO in terms of total actual system travel time, the inefficiency bound of SUE depends on both the class of cost functions and the degree of perception error and the network complexity. Nevertheless, it is found that the effect of network complexity in terms of number of available paths is rather limited. Unlike the price of anarchy of the deterministic UE established in the literature, the inefficiency bounds established for the SUE is generally not tight unless either the users' perception error is zero or the network has constant link travel times.

When comparing SUE with SSO in terms of total perceived travel time (or equivalently net economic benefit), we established an upper bound on the welfare loss of SUE and found that the welfare loss is not larger than a fraction of the total actual system travel time under SUE, and the "fraction" is independent of network topology, but depends solely on the class of link cost functions. We also found that the established inefficiency bound is tight only when link costs are constants (in this case SUE coincides with SSO).

Acknowledgements

The research described here was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. HKUST6215/06E). The third author was partially supported by a grant (20080440357) from China Postdoctoral Science Foundation. We are sincerely grateful for two anonymous reviewers for their insightful comments on an earlier version of the paper.

References

Chau, C. K. and Sim K. M. (2003) The price of anarchy for non-atomic congestion games with symmetric cost maps and elastic demands. *Operations Research Letters* 31, 327-334.

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

Correa, J. R., Schulz A. S., and Stier-Moses N. E. (2005) On the inefficiency of equilibria in congestion games. Extended abstract in *Proceedings of the 11th Conference on Integer Programming and Combinatorial Optimization,* Vol. 3509 of *Lecture Notes in Computer Science*, 167-181. Springer-Verlag, Berlin.

Dafermos, S. (1980) Traffic equilibrium and variational inequalities. *Transportation Science* 14, 42-54.

Fisk, C. (1980) Some developments in equilibrium traffic assignment. *Transportation Research* 14B, 243-255.

Han, D., Lo, H.K., Sun, J. and Yang, H. (2008) The toll effect on price of anarchy when costs are nonlinear and asymmetric. *European Journal of Operational Research* 186 (1), 300-316.

Han, D. and Yang, H. (2008) The multi-class, multi-criterion traffic equilibrium and the efficiency of congestion pricing. *Transportation Research Part E* 44 (5), 753-773.

Koutsoupias, E. and Papadimitriou C. H. (1999) Worst-case equilibria. In *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*, Trier, Germany, Vol. 1563 of *Lecture Notes in Computer Science*, 404-413. Springer, Berlin.

Maher M., Stewart K., and Rosa A. (2005) Stochastic social optimum traffic assignment. *Transportation Research* 39 B, 753–767.

Papadimitriou, C. H (2001) Algorithms, games, and the Internet. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing (STOC)*, Heraklion, Greece, 749-753. ACM Press, New York, NY.

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

Roughgarden, T. (2005) *Selfish Routing and the Price of Anarchy*. The MIT Press, Cambridge.

Roughgarden, T. and Tardos E. (2004) Bounding the inefficiency of equilibria in nonatomic games. *Games and economic behavior* 47, 389-403.

Sheffi, Y. (1985) *Urban Transportation Networks*. Prentice-Hall, Englewood Cliffs, NJ.

Smith, M. J. (1979) The existence, uniqueness and stability of traffic equilibria. *Transportation Research* 13B, 295-304.

Stewart, K. (2007) Tolling traffic links under stochastic assignment: Modelling the relationship between the number and price level of tolled links and optimal traffic flows. *Transportation Research* 41A, 644-654.

Yang, H. (1999) System optimum, stochastic user equilibrium and optimal link tolls. *Transportation Science* 33 (4), 354–360.

Yang, H., Han, D. and Lo, H.K. (2008) Atomic splitable selfish routing with polynomial cost functions. *Networks and Spatial Economics* 8 (4), 443-451.