Edge Pricing of Multicommodity Networks for Heterogeneous Selfish Users

George Karakostas *
Department of Computing and Software
McMaster University
karakos@mcmaster.ca

Stavros G. Kolliopoulos†
Department of Informatics
and Telecommunications
University of Athens
and
Department of Computing and Software
McMaster University
sgk@di.uoa.gr

Abstract

We examine how the selfish behavior of heterogeneous users in a network can be regulated through economic disincentives, i.e., through the introduction of appropriate taxation. One wants to impose taxes on the edges so that any traffic equilibrium reached by the selfish users who are conscious of both the travel latencies and the taxes will minimize the social cost, i.e., will minimize the total latency. We generalize previous results of Cole, Dodis and Roughgarden that held for a single origin-destination pair to the multicommodity setting.

Our approach, which could be of independent interest, is based on the formulation of traffic equilibria as a nonlinear complementarity problem by Aashtiani and Magnanti [1]. We extend this formulation so that each of its solutions will give us a set of taxes that forces the network users to conform, at equilibrium, to a certain prescribed routing. We use the special nature of the prescribed minimum-latency flow in order to reduce the difficult nonlinear complementarity formulation to a pair of primal-dual linear programs. LP duality is then enough to derive our results.

1. Introduction

In the *selfish routing* setting, we are given a directed network $G = (V, E)$ and a set of k *classes of users* (commodities), each with its own origin and destination, and with a fixed total demand (traffic) rate per class $d_i > 0$, $i = 1, \ldots, k$. Individual users are thought as carrying each an infinitesimal amount of a commodity. We are also given a nonnegative latency function l_P describing the delay experienced by users wishing to travel on the path P as a function of the total flow through the edges of the path. In this work we assume that the *additive model* holds, i.e., for every edge e there is a latency function $l_e(f_e)$ that describes the latency on this edge due to the flow f_e that crosses it. Then the latency for a path is defined as $l_P(f) := \sum_{e \in P} l_e(f_e)$. Each user tries to selfishly route his flow so that his path cost is minimized. A *traffic equilibrium* is an assignment of traffic to paths so that no user can unilaterally switch her flow to a path of smaller cost. Wardrop's principle [12] for selfish routing postulates that

> at equilibrium, for each origin-destination pair the travel costs on all the routes actually used are equal, or less than the travel costs on all nonused routes.

A natural question then is, can one use Wardrop's principle in order to design a network that, at equilibrium, will force the user routes to follow a prescribed pattern? Although it may be possible to design such a network by designing the topology, or the latency functions, we will assume that these parameters are already fixed. Another possibility is to use economic disincentives in the form of *taxation* on the edges of the network, so that the users' path cost has both a travel time and a budgetary component. Without taxation, users experience only their own traffic delay as their cost. With taxation users are also charged for the right to use a path. This technique has been studied by the

* Research supported by an NSERC Discovery grant.
† Research partially supported by NSERC Research Grant 227809-00.

traffic community for a long time (cf. [3] and the references therein), and will be the focus of this work as well. Our goal will be to design edge taxes that impose to the users a routing that optimizes the network performance. The measure of this performance in this paper is the *social cost* (or *total latency*), defined as $\sum_{\text{path } P} f_P l_P(f) = \sum_{e \in E} f_e l_e(f_e)$ for a flow f that routes f_P units of traffic through path P.

Although it must obey Wardrop's principle at equilibrium, the unregulated choice of paths by individual users may incur a social cost which in general can be higher than the social optimum. In fact the latency of an equilibrium with no taxation can be arbitrarily larger than the social optimum [11]. Through appropriate edge taxation, we would like to force *all* equilibria on the network to induce flow that minimizes the social cost $\sum_{e \in E} f_e l_e(f_e)$. We refer to a set of edge taxes that achieves this as *optimal taxes*. This definition of optimal taxes is stricter than the definition in [3], where the optimal taxes are just required to force *some* equilibrium flow to minimize the social cost. The difference is significant: we will need strict monotonicity of the edge latency functions to fulfill the stricter definition, and simple monotonicity to fulfill the weaker one. In [3] the authors can also meet the stricter definition of optimal taxes under the same assumption of strict monotonicity. Each selfish user of class i using path P will experience the following path cost:

$$\text{path cost}(P) := \text{latency}(P) + a(i) \cdot \text{taxation}(P).$$

The taxation(P) is the sum of taxes along the edges of the path. The factor $a(i) > 0$, denotes the sensitivity of user class i to the taxes. In the *homogeneous* case all user classes have the same sensitivity to the taxation (i.e., $a(i) = 1$, for all i), while in the *heterogeneous* case $a(i)$ can take different positive values for different classes.

The homogeneous case has been extensively studied. Optimal taxes can be defined using the so-called marginal cost pricing [2]. Various ramifications, extensions and related computational results have been obtained over the years. See [3] for a wealth of related references. Cole, Dodis and Roughgarden [3] were the first to consider the natural generalization into the heterogeneous case. The heterogeneity expresses the fact that different users trade off money and time in a different manner. A small value of $a(i)$ corresponds to a large sensitivity to time and a large value indicates sensitivity to money. The main result of [3] is that for an instance where all users share the same origin-destination pair, there is an optimal set of taxes. Such a set can be efficiently computed if a flow that minimizes the social cost is efficiently computable or more generally if it is given [3].

The general setting of selfish routing assumes multiple origin-destination pairs for the selfish users. Extending the results of [3] to this multicommodity setting is therefore a natural question which was also stated in [3] as an open problem. In this work we resolve this open problem. Specifically we show that optimal taxes exist while generalizing the work in [3] in two ways:

1. The commodities (user classes) can have different origin-destination pairs.
2. The latency functions l_e do not have to be continuous for the optimal taxes to exist. We only assume that a flow $\hat{f}$ of minimum total latency exists.

Similarly to [3] the set of optimal taxes is computable in polynomial time if $\hat{f}$ is given. We focus on the case where the number of user classes is finite, in other words the function a takes only finitely many distinct values. In fact the main proof of [3] is given for the finite case and then a limiting argument is used to approximate the infinite case. We leave this issue for future work. We believe that the finite case captures what is important about the problem.

The proof of existence in [3] relied on a fixed-point theorem tailored to the intricate combinatorial structure of the problem at hand. In contrast, our approach for showing the existence of optimal taxes uses only general tools from mathematical programming, like linear programming duality and the formulation of the traffic (Wardrop) equilibrium as a nonlinear complementarity problem in [1]. We extend the latter formulation in order to express the constraint that optimal taxes are part of the solution. In fact, the new formulation is a special case of a more general formulation for the setting of the problem with side constraints, as described in [7]. In general, the existence of solutions to a nonlinear complementarity formulation is a very difficult problem. We use the fact that the prescribed flow $\hat{f}$ optimizes the social cost, to reduce the nonlinear complementarity problem at hand to a *linear* one, that is easy to solve. In fact this linear complementarity problem corresponds to the complementary slackness conditions for a pair of primal-dual *linear programs*. LP duality is then enough, not only to prove that optimal taxes exist, but also to compute them (in [3] LP-duality was also used for the constructive part of computing the optimal taxes in the single-commodity case). This 'linearization' technique is our main technical contribution. Although in this work it derives its power from the special nature of $\hat{f}$ as a global optimizer of the so-

cial cost, it may be useful in pricing networks that induce other prescribed flows to their selfish users, see Section 5 for a discussion.

The complementarity formulation we employ is powerful enough to allow the demands to be *elastic*. In this case the flow rate for each class is not constant, but is given by a function of the cost of the paths used by the users. This topic has received considerable attention in the transportation literature (see, e.g., [5]). This setting captures the more realistic scenario, where the destination-client of a commodity adjusts its demand according to the transfer time from the origin-server. For example, the download volume from a Web server by a client depends on whether there is a quick connection to the server or not. We outline possible extensions of our work in the elastic demand setting and propose some relevant open problems in Section 4.

The main result of this work together with several extensions were also discovered independently by Fleischer, Jain, and Mahdian [4], and appear also in these proceedings.

2. Preliminaries

The model: Let $G = (V, E)$ be a directed network (possibly with parallel edges but with no self-loops), and a set of *users*, each with an infinitesimal amount of traffic (flow) to be routed from an origin node to a destination node of G. Moreover, each user α has a positive *tax-sensitivity* factor $a(\alpha) > 0$. We will assume that the tax-sensitivity factors for all users come from a finite set of possible positive values. We can bunch together into a single *user class* all the users with the same origin-destination pair and with the same tax-sensitivity factor; let k be the number of different such classes. We denote by $d_i, \mathcal{P}_i, a(i)$ the total flow of class i, the flow paths that can be used by class i, and the tax-sensitivity of class i, for all $i = 1, \ldots, k$ respectively. We will also use the term 'commodity i' for class i. Set $\mathcal{P} \doteq \cup_{i=1,\ldots,k} \mathcal{P}_i$. Each edge $e \in E$ is assigned a *latency function* $l_e(f_e)$ which gives the latency experienced by any user that uses e due to congestion caused by the total flow f_e that passes through e. In other words, as in [3], we assume the additive model in which for any path $P \in \mathcal{P}$ the latency is $l_P(f) = \sum_{e \in P} l_e(f_e)$, where $f_e = \sum_{e \ni P} f_P$ and f_P is the flow through path P. If every edge is assigned a per-unit-of-flow tax $\beta_e \geq 0$, a selfish user in class i that uses a path $P \in \mathcal{P}_i$ experiences total cost of

$$\sum_{e \in P} l_e(f_e) + a(i) \sum_{e \in P} \beta_e$$

hence the name 'tax-sensitivity' for the $a(i)$'s: they quantify the importance each user assigns to the taxation of a path.

Let $\hat{f}$ be a flow that minimizes the total latency $\sum_e f_e l_e(f_e)$. Note that, although in certain cases (e.g., when the functions $f_e l_e(f_e)$ are convex) the flow $\hat{f}$ can be computed efficiently, for more general latency functions it may be extremely difficult to compute $\hat{f}$ (see Section 4 in [3]). We will assume that $\hat{f}$ is given to us off-line and that it induces finite latency on every edge.

Given edge latency functions l_e, we define a set of new path latency functions

$$l'_P(f) := \frac{l_P(f)}{a(i)} = \frac{\sum_{e \in P} l_e(f_e)}{a(i)}, \ \forall P \in \mathcal{P}_i, \forall i. \quad (1)$$

A function $g(x)$ is *positive* if $g(x) > 0$ when $x > 0$. We assume that the functions l_e are strictly increasing, i.e., $x > y \geq 0$ implies $l_e(x) > l_e(y)$, and that $l_e(0) \geq 0$. This implies that $l_e(f_e) > 0$ when $f_e > 0$, i.e., the function l_e is positive. We only need the strict monotonicity for Corollary 1 to hold, see Remark 1. For the rest of the paper it is enough for the l_e functions to be nondecreasing and positive. Similar assumptions on monotonicity are made in [3].

In what follows we will use heavily the notion of a nonlinear complementarity problem. Let $F(x) = (F_1(x), F_2(x), \ldots, F_n(x))$ be a vector-valued function from the n-dimensional space $\mathbb{R}^n$ into itself. Then the nonlinear complementarity problem of mathematical programming is to find a vector x that satisfies the following system:

$$x^T F(x) = 0, \quad x \geq 0, \quad F(x) \geq 0.$$

3. Existence and computation of optimal taxes

The traffic (or Wardrop) equilibria[1] for a network can be described as the solutions of the following mathematical program (see [1] p. 216):

$$\begin{aligned} (T_P(f) - u_i) f_P &= 0 \quad \forall P \in \mathcal{P}_i, i = 1 \ldots k \\ T_P(f) - u_i &\geq 0 \quad \forall P \in \mathcal{P}_i, i = 1 \ldots k \\ \sum_{P \in \mathcal{P}_i} f_P - d_i &= 0 \quad \forall i = 1 \ldots k \\ f, u &\geq 0 \end{aligned}$$

where T_P is the cost of a user that uses path P, f_P is the flow through path P, and $u = (u_1, \ldots, u_k)$ is the vector of shortest travel times (or generalized costs) for the commodities. The first two equations model

1 [3] uses the term Nash equilibria.

Wardrop's principle by requiring that for any origin-destination pair i the travel cost for all paths in $\mathcal{P}_i$ with nonzero flow is the same and equal to u_i. The remaining equations ensure that the demands are met and that the variables are nonnegative. Note that the formulation above is very general: every path $P \in \mathcal{P}_i$ for every commodity i has its own T_P (even if two commodities share the same path P, each may have its own T_P).

If the path cost functions T_P are positive, [1] shows that the system above is equivalent to the following nonlinear complementarity problem (Proposition 4.1 in [1]):

$$\begin{aligned} (T_P(f) - u_i) f_P &= 0 \quad \forall i,\ \forall P \in \mathcal{P}_i \\ T_P(f) - u_i &\geq 0 \quad \forall i,\ \forall P \in \mathcal{P}_i \\ u_i(\sum_{P \in \mathcal{P}_i} f_P - d_i) &= 0 \quad \forall i \\ \sum_{P \in \mathcal{P}_i} f_P &\geq d_i \quad \forall i \\ f, u &\geq 0 \end{aligned} \tag{CP}$$

In our case the costs T_P are defined as

$$\sum_{e \in P} l_e(f_e) + a(i) \sum_{e \in P} b_e,\ \forall i,\ \forall P \in \mathcal{P}_i,$$

where b_e is the per-unit-of-flow tax for edge e, and $a(i)$ is the tax sensitivity of commodity i. In fact, it will be more convenient for us to define T_P slightly differently:

$$T_P(f) := l'_P(f) + \sum_{e \in P} b_e, \quad \forall i,\ \forall P \in \mathcal{P}_i,$$

where $l'_P(f)$ is defined in (1). It is easy to see that there is a one-to-one correspondence between the solution sets of (CP) with the two path cost definitions (solution $(f_{P_1}, \ldots, f_{P_{|\mathcal{P}|}}, u_1, \ldots, u_k)$ of (CP) with the first definition of T_P corresponds to solution $(f_{P_1}, \ldots, f_{P_{|\mathcal{P}|}}, \frac{u_1}{a(1)}, \ldots, \frac{u_k}{a(k)})$ with the second definition of T_P, and vice-versa).

Let $\hat{f}$ be a flow that satisfies the users' demands and minimizes the social cost $\sum_{e \in E} f_e l_e(f_e) = \sum_i \sum_{P \in \mathcal{P}_i} f_P l_P(f)$, i.e., $\hat{f}$ is a solution of the following mathematical program:

$$\begin{aligned} \min \sum_{e \in E} f_e l_e(f_e) \quad & \text{s.t.} \\ \sum_{P \in \mathcal{P}_i} f_P &= d_i \quad \forall i \\ f_e &= \sum_{P \in \mathcal{P}: e \in P} f_P \quad \forall e \in E \\ f_P &\geq 0 \quad \forall P \end{aligned} \tag{MP}$$

Note that (MP) in general may be very difficult (or impossible) to solve efficiently, but we will assume that a solution $\hat{f}$ has already been computed and given to us.

Our main result is to show that there is a set of per-unit taxes b which forces the users of the non-cooperative setting of (CP) to induce on the edges the same edge flow as $\hat{f}$, i.e. we show the following

Theorem 1 *Consider the selfish routing game with the latency function seen by the users in class i being*

$$T_P(f) := \sum_{e \in E} l_e(f_e) + a(i) \sum_{e \in P} b_e,\ \forall i,\ \forall P \in \mathcal{P}_i.$$

If for every edge $e \in E$ l_e is a strictly increasing function with $l_e(0) \geq 0$, then there is a vector of per-unit taxes $b \in \mathbb{R}_+^{|E|}$ such that, if $\bar{f}$ is a traffic equilibrium for this game, $\bar{f}_e = \hat{f}_e$, $\forall e \in E$,. Therefore $\bar{f}$ minimizes the social cost $\sum_{e \in E} f_e l_e(f_e)$.

In order to prove Theorem 1, it is enough to prove the following

Theorem 2 *If $\forall e \in E$ the functions $l_e(\cdot)$ are strictly monotone and $l_e(0) \geq 0$, then there is $b \in \mathbb{R}_+^{|E|}$, such that if $\bar{f}$ solves (CP) with*[2]

$$T_P(f) := l'_P(f) + \sum_{e \in P} b_e,\ \forall i,\ \forall P \in \mathcal{P}_i,$$

then $\bar{f}_e = \hat{f}_e$, $\forall e \in E$.

Our plan for proving Theorem 2 will be to find b such that

- all solutions of (CP) induce the *same edge flow*
- there is at least one solution of (CP) with edge flow $\hat{f}_e$, $\forall e \in E$

The first goal is easy to achieve, due to the following uniqueness result of [1]:

Theorem 3 (Theorem 6.2 in [1]) *Suppose that the functions T_P in (CP) are defined as follows:*

$$T_P(f) := \sum_{e \in P} t_e^i(f),\ \forall P \in P_i,\ \forall i,$$

where $t_e^i(f)$ is the cost function for edge e and commodity i,

$$t_e^i : \mathbb{R}_+^{|\mathcal{P}|} \to \mathbb{R}_+.$$

If t is strictly increasing, then for all solutions of (CP) the vector u and the vector of induced edge flows are unique.[3]

---

2 Note that we changed the definition of T_P to the more convenient one. This will not affect our computations, as discussed above.

3 Note that we may still have solutions with different path flows. This uniqueness result concerns only the edge flows.

In our case, $t_e^i(f) = \frac{l_e(f_e)}{a(i)} + b_e$, which is strictly increasing ($l_e(\cdot)$ is strictly increasing, and b_e is a positive constant). Therefore we get the following corollary:

Corollary 1 *All solutions of (CP) induce the same edge flow.*

Remark 1 The uniqueness result of Corollary 1 is the only place where the strict monotonicity of the latency functions is needed. We note that if we follow the definition of optimal taxes by Cole et al. [3], then we do not need this uniqueness property, and our results hold for nondecreasing l_e; in what follows we show that for the set of taxes whose existence we prove, there is *some* solution (f^*, u^*) of (CP) with $f_e^* = \hat{f}_e,\ \forall e \in E$. On the other hand, according to our definition of optimal taxes, *all* solutions of (CP) must induce the optimal edge flow, therefore Corollary 1 is needed.

Note that the uniqueness result of Corollary 1 holds independently of our choice of b. Hence we can concentrate our efforts on finding a b that will steer the (unique) edge flow of a solution of (CP) towards $\hat{f}$. In order to achieve this, *we include it in (CP) as a requirement*. Hence we get the following nonlinear complementarity problem, in which we require that the edge flow of any solution is at most $\hat{f}_e$:

$$
\begin{aligned}
f_P(T_P(f) - u_i) &= 0 \quad \forall i,\ \forall P \in \mathcal{P}_i \qquad \text{(BIG CP)}\\
T_P(f) &\geq u_i \quad \forall i,\ \forall P \in \mathcal{P}_i\\
u_i(\sum_{P\in\mathcal{P}_i} f_P - d_i) &= 0 \quad \forall i\\
\sum_{P\in\mathcal{P}_i} f_P &\geq d_i \quad \forall i\\
b_e(f_e - \hat{f}_e) &= 0 \quad \forall e \in E\\
f_e &\leq \hat{f}_e \quad \forall e \in E\\
f_P, b_e, u_i &\geq 0 \quad \forall P, e, i
\end{aligned}
$$

If (f^*, b^*, u^*) is a solution of (BIG CP) then (f^*, u^*) also solves (CP) with

$$T_P(f) := l'_P(f) + \sum_{e\in P} b_e^*,\ \forall P \in \mathcal{P}_i,\ \forall i.$$

The crucial observation is that this solution not only satisfies all the demands, but also induces the same edge flow as $\hat{f}$:

Lemma 1 *Let (f^*, b^*, u^*) be any solution of (BIG CP). Then $\sum_{P\in\mathcal{P}_i} f_P^* = d_i,\ \forall i$ and $f_e^* = \hat{f}_e,\ \forall e \in E$.*

Proof: The proof of the first part is essentially the same as the proof by contradiction of Proposition 4.1 in [1]. Suppose that $\sum_{P\in\mathcal{P}_i} f_P^* > d_i \geq 0$ for some i. Then $u_i^*(\sum_{P\in\mathcal{P}_i} f_P^* - d_i) = 0 \Rightarrow u_i^* = 0$ and there is a path $P \in \mathcal{P}_i$ such that $f_P^* > 0$. Since $f_P^* \neq 0$, the $T_P()$ function is positive, $T_P(f^*) > 0 = u_i^*$. Because (f^*, b^*, u^*) is a solution of (BIG CP), we have that $f_P^*(T_P(f^*) - u_i^*) = 0 \Rightarrow f_P^* = 0$, a contradiction. Hence $\sum_{P\in\mathcal{P}_i} f_P^* = d_i,\ \forall i$.

Since f^* is part of a solution for (BIG CP), $f_e^* \leq \hat{f}_e,\ \forall e \in E$. The following claim is a result of the special nature of $\hat{f}$ as a minimizer of the social cost:

Claim 1 *Let $\bar{f}$ be a flow that satisfies the following set of constraints:*

$$
\begin{aligned}
\sum_{P\in\mathcal{P}_i} f_P &= d_i \qquad \forall i \in \{1, \ldots, k\}\\
f_e &= \sum_{P\in\mathcal{P}: e\in P} f_P \quad \forall e \in E\\
f_e &\leq \hat{f}_e \qquad \forall e \in E\\
f_P &\geq 0 \qquad \forall P \in \mathcal{P}
\end{aligned}
$$

Then $\bar{f}_e = \hat{f}_e,\ \forall e \in E$.

Proof of Claim: By the definition of $\hat{f}$ through (MP), we have that

$$\sum_{e\in E} \hat{f}_e l_e(\hat{f}_e) \leq \sum_{e\in E} \bar{f}_e l_e(\bar{f}_e) \qquad (2)$$

Since $\bar{f}_e \leq \hat{f}_e$ and $l_e(\cdot)$ is increasing we obtain that $l_e(\bar{f}_e) \leq l_e(\hat{f}_e)$. Since $l_e(\cdot)$ is nonnegative for all $e \in E$, we obtain that

$$\bar{f}_e l_e(\bar{f}_e) \leq \hat{f}_e l_e(\hat{f}_e),\ \forall e \in E.$$

If for some e, $\bar{f}_e < \hat{f}_e$, then because $l_e(\cdot)$ is increasing $0 \leq l_e(\bar{f}_e) \leq l_e(\hat{f}_e)$; because $l_e(\cdot)$ is positive $l_e(\hat{f}_e) \neq 0$. From these two facts

$$\bar{f}_e l_e(\bar{f}_e) < \hat{f}_e l_e(\hat{f}_e)$$

and then $\sum_{e\in E} \bar{f}_e l_e(\bar{f}_e) < \sum_{e\in E} \hat{f}_e l_e(\hat{f}_e)$ which contradicts (2). □

Since f^* satisfies the constraints of Claim 1, we have that $f_e^* = \hat{f}_e,\ \forall e \in E$. □

Lemma 1 implies that if there is a solution (f^*, b^*, u^*) for (BIG CP), then we can use b^* as the per-unit taxes to ensure that (f^*, u^*) is also a solution to the (CP). The fact that $f_e^* = \hat{f}_e,\ \forall e$ together with the edge flow uniqueness property of Corollary 1 will imply Theorem 2, provided that such a (f^*, b^*, u^*) *exists*.

In general, it is quite difficult to prove existence of solutions to nonlinear complementarity problems (see [8], [9], [10]), and usually these results do not apply to the complementarity problem at hand, or they hold for special cases of the latency functions l_e. We

note that these latency functions l_e are actually the source of nonlinearity for (BIG CP). Surprisingly, the special nature of $\hat{f}$ will help us 'linearize' (BIG CP), and show it equivalent to a primal-dual pair of *linear* programs (LP). Then the well-understood duality theory of LP will give us the existence result we are looking for.

The key observation is again Lemma 1. Since for every solution (f^*, b^*, u^*) of (BIG CP) we have $f_e^* = \hat{f}_e$, $\forall e$, and $l_e(f_e)$ depends only on f_e, the set of solutions for (BIG CP) will not change if we replace $l_e(f_e)$ with $l_e(\hat{f}_e)$ in (BIG CP) to get the following linear complementarity problem (BIG CP′) :

$$
\begin{aligned}
f_P\Big(\sum_{e\in P}\frac{l_e(\hat{f}_e)}{a(i)} + \sum_{e\in P} b_e - u_i\Big) &= 0 \quad \forall i,\ \forall P\in\mathcal{P}_i \\
\sum_{e\in P}\frac{l_e(\hat{f}_e)}{a(i)} + \sum_{e\in P} b_e &\geq u_i \ \forall i,\ \forall P\in\mathcal{P}_i \\
u_i\Big(\sum_{P\in\mathcal{P}_i} f_P - d_i\Big) &= 0 \quad \forall i \\
\sum_{P\in\mathcal{P}_i} f_P &\geq d_i \ \forall i \qquad \text{(BIG CP}')\\
b_e(f_e - \hat{f}_e) &= 0 \quad \forall e\in E \\
f_e &\leq \hat{f}_e \ \forall e\in E \\
f_P, b_e, u_i &\geq 0 \quad \forall P, e, i
\end{aligned}
$$

Lemma 2 *(BIG CP) is equivalent to (BIG CP′).*

Proof: If (f^*, b^*, u^*) is a solution of (BIG CP), then it is also a solution of (BIG CP′) due to Lemma 1. Conversely, if $(\bar{f}, \bar{b}, \bar{u})$ is a solution of (BIG CP′), then we can prove the equivalent to Lemma 1 for (BIG CP′) (we only have to notice that $l_e(\hat{f}_e) \geq l_e(\bar{f}_e) > 0$ whenever $\bar{f}_e > 0$, hence if there is $P\in\mathcal{P}_i$ such that $\bar{f}_P > 0$, then $\sum_{e\in P}\frac{l_e(\hat{f}_e)}{a(i)} + \sum_{e\in P}\bar{b}_e > 0$). □

Proof of Theorem 2: Due to the complementary slackness conditions, (BIG CP′) is equivalent to the following primal-dual pair of LPs:

$$
\begin{aligned}
&\min \sum_i \sum_{P\in\mathcal{P}_i} f_P l'_P(\hat{f}) \quad \text{s.t.} && \text{(LP)}\\
&\sum_{P\in\mathcal{P}_i} f_P \geq d_i && \forall i\in\{1,\ldots,k\}\\
&f_e = \sum_{P\in\mathcal{P}: e\in P} f_P && \forall e\in E\\
&f_e \leq \hat{f}_e && \forall e\in E\\
&f_P \geq 0 && \forall P
\end{aligned}
$$

$$
\begin{aligned}
&\max \sum_i d_i u_i - \sum_{e\in E}\hat{f}_e b_e \quad \text{s.t.} && \text{(DP)}\\
&u_i - \sum_{e\in P} b_e \leq l'_P(\hat{f}) && \forall i,\ \forall P\in\mathcal{P}_i\\
&b_e, u_i \geq 0 && \forall e\in E, \forall i
\end{aligned}
$$

The objective function of (LP) is bounded from below by 0 and there is a feasible solution $(\hat{f})$, therefore duality theory implies that there is a solution (f^*, b^*, u^*) to this pair of LPs, which is also a solution of (BIG CP) due to Lemma 2. Then for per-unit taxes $b := b^*$, Theorem 2 holds.

□

Fleischer et al. [4] use directly the (LP)-(DP) pair to prove the equivalent of Theorem 1. Note that if there are more than one solutions to (LP)-(DP), each one of them can be used to define a selfish routing game for which Theorem 1 holds. Also, note that since every solution $f^*, (b^*, u^*)$ of (LP)-(DP) is also a solution of (BIG CP), (f^*, u^*) is a solution of (CP) when b^* is used as the tax set. Therefore, for the set of taxes whose existence and calculation come from (LP)-(DP), the corresponding selfish routing game has at least one equilibrium.

Computation of optimal taxes: If we are given $\hat{f}$, the computation of a set of optimal per-unit taxes involves just the solution of the linear program (DP). Similar results can be also found in [3] for the single origin-destination pair case.

4. Elastic demands

If the demands in (CP) are not the constants d_i but are given by functions $D_i(u)$ of the generalized costs u, then these demands are called *elastic*. The complementarity formulation of the traffic equilibria of [1] in its full generality allows for the demands to be elastic. Using this insight we formulate the analogous complementarity problems to the ones in Section 3. The main complication is that the minimum-latency flow $\hat{f}$ cannot be considered a priori given before some selfish routing game starts. At an equilibrium the u_i achieve some concrete value which in turn fixes the demands. These demands will then determine the corresponding minimum-latency flow $\hat{f}$.

In this general case (CP) becomes:

$$\begin{aligned}(T_P(f)-u_i)f_P &= 0 \quad \forall i,\ \forall P\in\mathcal{P}_i \qquad \text{(CPE)}\\ T_P(f)-u_i &\geq 0 \quad \forall i,\ \forall P\in\mathcal{P}_i\\ u_i(\sum_{P\in\mathcal{P}_i} f_P - D_i(u)) &= 0 \quad \forall i\\ \sum_{P\in\mathcal{P}_i} f_P - D_i(u) &\geq 0 \quad \forall i\\ f,u &\geq 0\end{aligned}$$

In this section we study linear latency functions $l_e(f_e)=\alpha_e f_e+\beta_e,\ \forall e\in E$, where $\alpha_e>0, \beta_e\geq 0,\ \forall e\in E$. Suppose that we are given a vector u^* of generalized costs. Then the social optimum $\hat{f}^*$ for the particular demands $D_i(u^*)$ is the solution of the following quadratic program:

$$\begin{aligned}\min \sum_{e\in E}\alpha_e\hat{f}_e^2+\beta_e\hat{f}_e \quad \text{s.t.} \qquad \text{(QP)}\\ \sum_{P\in\mathcal{P}_i}\hat{f}_P \geq D_i(u^*) \qquad \forall i\\ \hat{f}_e=\sum_{P\in\mathcal{P}:e\in P}\hat{f}_P \quad \forall e\in E\\ \hat{f}_P\geq 0 \qquad \forall P\end{aligned}$$

By Theorems 3.1, 3.3 of [6] $\hat{f}^*$ solves QP iff $\hat{f}^*$ solves the following linear program:

$$\begin{aligned}\min \sum_{e\in E}(2\alpha_e\hat{f}_e^*+\beta_e)\hat{f}_e \quad \text{s.t.} \qquad \text{(LP2)}\\ \sum_{P\in\mathcal{P}_i}\hat{f}_P \geq D_i(u^*) \qquad \forall i\\ \hat{f}_e=\sum_{P\in\mathcal{P}:e\in P}\hat{f}_P \qquad \forall e\in E\\ \hat{f}_P\geq 0 \qquad \forall P\end{aligned}$$

The dual of (LP2) is:

$$\begin{aligned}\max \sum_i D_i(u^*)\mu_i \qquad \text{s.t.} \qquad \text{(DP2)}\\ \mu_i\leq\sum_{e\in P}(2\alpha_e\hat{f}_e^*+\beta_e) \quad \forall i,\forall P\in\mathcal{P}_i\\ \mu_i\geq 0 \qquad \forall i\end{aligned}$$

Note that if the functions $D_i(u)$ are bounded, and

$$K_1:=k\cdot\max_i\max_{u\geq 0}\{D_i(u)\}$$

then the solutions $\hat{f}^*$, μ^* of (LP2), (DP2) are upper bounded as follows:

$$\begin{aligned}\hat{f}_P^*\leq D_i(u^*)\leq K_1,\ \forall P\in\mathcal{P}_i\\ \mu_i\leq\sum_{e\in P}(2\alpha_e\hat{f}_e^*+\beta_e)\leq\sum_{e\in P}(2\alpha_e K_1+\beta_e),\ \forall i\end{aligned}$$

The Karush-Kuhn-Tucker conditions for (LP2)-(DP2) imply that $(\hat{f}^*,\mu^*)$ is a pair of primal-dual solutions iff they are solutions of the following complementarity problem:

$$\begin{aligned}(\sum_{e\in P}(2\alpha_e\sum_{P\ni e}\hat{f}_P+\beta_e)-\mu_i)\hat{f}_P &= 0 \quad \forall i,\forall P\in\mathcal{P}_i\\ \sum_{e\in P}(2\alpha_e\sum_{P\ni e}\hat{f}_P+\beta_e)-\mu_i &\geq 0 \quad \forall i,\forall P\in\mathcal{P}_i\\ \mu_i(\sum_{P\in\mathcal{P}_i}\hat{f}_P-D_i(u^*)) &= 0 \quad \forall i\\ \sum_{P\in\mathcal{P}_i}\hat{f}_P-D_i(u^*) &\geq 0 \quad \forall i\\ \hat{f},\mu &\geq 0\end{aligned}$$

By incorporating this complementarity problem into the analogue of (BIG CP) we get the following complementarity problem (recall that $l_e()$ are now linear):

$$\begin{aligned}f_P(T_P(f)-u_i) &= 0 \quad \forall i,\forall P\in\mathcal{P}_i\\ &\qquad \text{(GENERAL CP)}\\ T_P(f)\geq u_i &\quad \forall i,\forall P\in\mathcal{P}_i\\ u_i(\sum_{P\in\mathcal{P}_i}f_P-D_i(u)) &= 0 \quad \forall i\\ \sum_{P\in\mathcal{P}_i}f_P-D_i(u) &\geq 0 \quad \forall i\\ b_e(f_e-\hat{f}_e) &= 0 \quad \forall e\in E\\ f_e\leq\hat{f}_e &\quad \forall e\in E\\ (\sum_{e\in P}(2\alpha_e\hat{f}_e+\beta_e)-\mu_i)\hat{f}_P &= 0 \quad \forall i,\forall P\in\mathcal{P}_i\\ \sum_{e\in P}(2\alpha_e\hat{f}_e+\beta_e)-\mu_i &\geq 0 \quad \forall i,\forall P\in\mathcal{P}_i\\ \mu_i(\sum_{P\in\mathcal{P}_i}\hat{f}_P-D_i(u)) &= 0 \quad \forall i\\ \sum_{P\in\mathcal{P}_i}\hat{f}_P-D_i(u) &\geq 0 \quad \forall i\\ f_P,b_e,u_i,\hat{f}_P,\mu_i &\geq 0\end{aligned}$$

where $\hat{f}_e=\sum_{P\ni e}\hat{f}_P$.

The framework that is in place now is similar to the setting in Section 3. Let (f,u) be a solution to (CPE). Under the assumptions that the $l_e(\cdot)$, $D(\cdot)$ functions are both strictly monotone and positive the results in [1] establish that the solution u is unique and that f induces a unique edge flow. The missing part of the puzzle is an existence result for (GENERAL CP). Let $(f^*,b^*,u^*,\hat{f}^*,\mu^*)$ be such a hypothetical solution. Then $\hat{f}^*$ is a minimum latency flow solution for the demand vector $D(u^*)$. Moreover $f_e^*\leq\hat{f}_e^*,\ \forall e\in E$. The

analogue of Theorem 2 would establish that after setting $b = b^*$ in (CPE), any solution $(\bar{f}, \bar{u})$ would satisfy $\bar{f}_e = f_e^*$ and $u = u^*$. Therefore $\bar{f}_e \leq \hat{f}_e^*$, $\forall e \in E$. Under the existing assumptions on $l_e(\cdot)$, we would obtain that any *equilibrium flow* $\bar{f}$ for the selfish routing game where the users are conscious of the modified latency

$$T_P(f) := l_P(f) + \sum_{e \in P} b_e^*, \quad \forall i, \ \forall P \in \mathcal{P}_i,$$

is a minimum-latency solution for *the demand vector reached in the same equilibrium.* Therefore the b^* vector would be the vector of the optimal taxes. The heterogeneity of the user classes can be handled in a manner similar to Section 3. We pose the existence of a solution to (GENERAL CP) and the handling of more general latency functions as open problems.

5. Further applications and discussion

The formulation of Wardrop equilibria as solutions to a (nonlinear) complementarity problem introduced by Aashtiani and Magnanti [1] provides us with a very general setting for tackling network design problems for non-cooperative users. It can handle diversity on the disutility functions (T_P's) of the users, elastic demands, and a wide variety of edge latency functions (l_e), even *non-continuous* ones. In this work we have made only a limited usage of this generality: our T_P's are the specific disutility functions dictated by the heterogeneous taxation-sensitivity problem tackled by [3], and in the elastic case we considered only linear edge latency functions. We believe that this approach can be used in many other settings as well.

The 'linearization' of the nonlinear complementarity problem (BIG CP) also seems to have further implications for network design. Note that the crucial observation was Lemma 1 and especially the simple fact of Claim 1. We can use the same arguments for other flows $\hat{f}$ with nice properties, provided $\hat{f}$ is a global optimizer for the social cost objective function $\sum_{e \in E} f_e l_e(f_e)$. For example, suppose that we would like to impose on our network a set of taxes, so that the users will be compelled to induce a flow that will minimize the maximum path latency over all paths in $\mathcal{P}$. Let $\hat{B}$ be the solution of the following program:

$$\begin{array}{rl} \min B & \text{subject to:} \\ \sum_{P \in \mathcal{P}_i} f_P = d_i & \forall i \in \{1, \ldots, k\} \\ f_e = \sum_{P \in \mathcal{P}: e \in P} f_P & \forall e \in E \\ \sum_{e \in P} l_e(f_e) \leq B & \forall P \in \mathcal{P} \\ f_P \geq 0 & \forall P \end{array}$$

We assume again that the l_e's are strictly monotone and positive. Let F be the set of flows that achieve the optimum $\hat{B}$ in the program above. If we pick $\hat{f} \in F$ to be a flow in F that also minimizes $\sum_{e \in E} f_e l_e(f_e)$ over all flows in F, then Lemma 1 holds (*proof sketch:* if there is flow f with $f_e \leq \hat{f}_e, \forall e \in E$ and $f_e < \hat{f}_e$ for some e, then f would be in F and still achieve a smaller social cost, a contradiction). Hence, if we are given $\hat{f}$, then we can compute a set of per-unit edge taxes that will minimize the maximum latency of all paths.

The independent work of [4] makes explicit several such applications in various settings.

Acknowledgement: We thank Jiming Peng for reading an earlier draft and for valuable comments.

References

[1] H. Z. Aashtiani and T. L. Magnanti. Equilibria on a congested transportation network. *SIAM Journal of Algebraic and Discrete Methods*, 2:213–226, 1981.

[2] M. Beckmann, C. B. McGuire, and C. B. Winsten. *Studies in the Economics of Transportation.* Yale University Press, 1956.

[3] R. Cole, Y. Dodis, and T. Roughgarden. Pricing network edges for heterogeneous selfish users. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, pages 521–530, 2003.

[4] L. Fleischer, K. Jain, and M. Mahdian. Taxes for heterogeneous selfish users in a multicommodity network. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, 2004.

[5] D. W. Hearn and M. B. Yildirim. A toll pricing framework for traffic assignment problems with elastic demand. In M. Gendreau and P. Marcotte, editors, *Transportation and Network Analysis: Current Trends. Miscellanea in honor of Michael Florian.* Kluwer Academic Publishers, 2002.

[6] J. J. Judice and G. Mitra. Reformulation of mathematical programming problems as linear complementarity problems and investigation of their solution methods. *Journal of Optimization Theory and Applications*, 57:123–149, 1988.

[7] G. Karakostas and S. G. Kolliopoulos. Selfish routing in the presence of side constraints. Technical Report CAS-03-13-GK, Department of Computing and Software, McMaster University, 2003.

[8] S. Karamardian. The nonlinear complementarity problem with applications, part 1. *Journal of Optimization Theory and Applications*, 4:87–98, 1969.

[9] M. Kojima. A unification of the existence theorems of the nonlinear complementarity problem. *Mathematical Programming*, 9:257–277, 1975.

[10] N. Megiddo and M. Kojima. On the existence and uniqueness of solutions in nonlinear complementarity theory. *Mathematical Programming*, 12:110–130, 1977.

[11] T. Roughgarden and É. Tardos. How bad is selfish routing? *Journal of the ACM*, 49:236–259, 2002.

[12] J. G. Wardrop. Some theoretical aspects of road traffic research. *Proc. Inst. Civil Engineers, Part II*, 1:325–378, 1952.