Pricing a Segmented Market Subject to Congestion

Maxime Fortin* Patrice Marcotte† Gilles Savard*

*GERAD, École Polytechnique de Montréal
C.P. 6079, Succursale Centre-ville
Montréal, QC, H3C 3A7, Canada
{maxime.fortin,gilles.savard}@polymtl.ca

†CRT, Université de Montréal
C.P. 6128, Succursale Centre-Ville
Montreal, QC, H3C 3J7, Canada
marcotte@iro.umontreal.ca

1 Introduction

The optimal setting of prices, taxes or subsidies on goods and services can be naturally modeled as a bilevel program. Indeed, bilevel programming is an adequate framework for modeling optimization situations where a subset of decision variables is not controlled by the main optimizer (the *leader*), but rather by a second agent (the *follower*) who optimizes its own objective function with respect to this subset of variables.

In this presentation we address the problem of setting profit-maximizing tolls on a congested transportation network involving several user classes. At the upper level, the firm (leader) sets tolls on a subset of arcs and strives to maximize its revenue. At the lower level, each user minimizes its generalized travel cost, expressed as a linear combination of travel time and out-of-pocket travel cost. We assume the existence of a probability density function that describes the repartition of the value of time (VOT) parameter throughout the population. This yields a bilevel optimization problem involving a bilinear objective at the upper level and a convex objective at the lower level. Since, in this formulation, lower level variables are flow densities, it follows that the lower level problem is infinite-dimensional.

We devise a two-phase algorithm to solve this nonconvex problem. The first phase aims at finding a good initial solution by solving for its global optimum a discretized version of the model. The second phase implements a gradient method, starting from the initial point obtained in the initial phase.

2 The Bilevel Model

We consider a multicommodity network where each commodity $k \in \mathcal{K}$ is associated with an origin-destination pair of a transportation network G having node set $\mathcal{N}$ and arc set $\mathcal{A}$. The set $\mathcal{A}$ is partitioned into the subset A_1 of toll arcs and the subset A_2 of toll-free arcs. For each O-D pair k, let d_k be the demand associated with the O-D pair k and P_k the set of paths connecting its origin to its destination. We denote by P the set of paths connecting any given O-D pair in the network G.

We associate with each arc $a \in A$ a generalized travel time composed of a travel delay and a travel cost. Let α be a parameter representing the inverse of *value of time* that converts one money unit into one time unit. In our multi-class setting, we assume that each traveler has its own value α characterized by a continuous density function h. We assume that $h(\alpha)$ is positive over the open interval $(0, \alpha_{\max})$, null at its endpoints and integrable over the closed interval $[0, \alpha_{\max}]$ [1].

We denote by $v_p(\alpha)$ the flow density on path $p \in P$. The set of feasible path flow vectors $v(\alpha) = \{v_p(\alpha)\}_{p \in P}$ is given by

$$Y(\alpha) = \Big\{ v(\alpha) \geq 0 \,\big|\, \sum_{p \in P_k} v_p(\alpha) = d_k h(\alpha), \ \forall k \in K \Big\}. \tag{1}$$

The total path flow vector $\overline{v}$ is then

$$\overline{v} = \{\overline{v}_p\}_{p \in P}$$

with

$$\overline{v}_p = \int_0^{\alpha_{\max}} v_p(\alpha) d\alpha.$$

The set of feasible total path flow vectors is the compact, finite-dimensional polyhedron (see [3])

$$\overline{Y} \quad = \quad \Big\{ \overline{v} \geq 0 \,\big|\, \sum_{p \in P_k} \overline{v}_p = d_k, \ \forall k \in K \Big\}. \tag{2}$$

The total flow on arc a, denoted x_a, is the sum of path flows going through arc a.

For each arc a, $d_a(x_a)$ represents the delay on arc a, and c_a the fixed part of the travel cost. Let us define $A(p)$ as the set of arcs that compose path p. The travel delay and the fixed cost on path p are then expressed as

$$D_p(\overline{v}) = \sum_{a \in A(p)} d_a(x_a)\,, \quad C_p = \sum_{a \in A(p)} c_a\,. \tag{3}$$

If t_a denotes the toll vector on an arc a (with $t_a \equiv 0, \forall a \in A_2$), the generalized cost π_p on path p is a linear function of travel delay and travel time:

$$\pi_p(\alpha, \overline{v}, t) = D_p(\overline{v}) + \alpha(C_p + \sum_{a \in A_1(p)} t_a)\,, \tag{4}$$

where $A_1(p)$ is defined as $A(p) \cap A_1$.

[1] We could have also considered distinct density functions h^k for each commodity k.

For fixed t, a flow density vector is an equilibrium (almost everywhere) if and only if it satisfies the infinite-dimensional variational inequality

$$\text{VI}: \qquad v \in \overline{Y} = \left\{ v \in \left\{ \ell^2(0, \alpha_{\max}) \right\}^{|P|} : v(\alpha) \in Y(\alpha), \quad \forall \alpha \in [0, \alpha_{\max}] \right\}$$
$$\langle D(\overline{v}) + \alpha C, v - y \rangle \leq 0, \quad \forall y \in \overline{Y} \tag{5}$$

where $C = \{C_p + \sum_{a \in A_1(p)} t_a\}_{p \in P}$ and the operator $\langle \cdot, \cdot \rangle$ denotes the scalar product of vectors function in $\left\{ \ell^2(0, \alpha_{\max}) \right\}^{|P|}$, i.e.,

$$\langle \Phi, \Psi \rangle = \int_0^{\alpha_{\max}} \langle \Phi(\alpha), \Psi(\alpha) \rangle \, d\alpha.$$

In the case where $D = (D_p)_{p \in P}$ is the gradient of some function d, one can show that v is an equilibrium solution if and only if it satisfies the first-order optimality conditions associated with the infinite-dimensional mathematical program

$$\min_{\overline{v} \in \overline{Y}} d(\overline{v}) + \int_0^{\alpha_{\max}} \alpha \langle C, v(\alpha) \rangle \, d\alpha.$$

The price setting model is then formulated as the following bilevel program:

$$\text{BP:} \qquad \max_{t, x, v(\alpha)} \quad \sum_{a \in A_1} t_a x_a \tag{6}$$
$$x_a = \sum_{p | a \in A(p)} \int_0^{\alpha_{\max}} v_p(\alpha) d\alpha, \quad \forall a \in A \tag{7}$$
$$\min_{x, v(\alpha)} \quad d(\overline{v}) + \int_0^{\alpha_{\max}} \alpha \langle C, v(\alpha) \rangle \, d\alpha \tag{8}$$
$$v(\alpha) \in Y(\alpha), \quad \forall \alpha \in [0, \alpha_{\max}]. \tag{9}$$

3 Theoretical properties of the lower level program

In this section, we focus our attention on the lower level program (8-9). This involves some notation and a key assumption. In the sequel we set, for the sake of clarity,

$$T_p = C_p + \sum_{a \in A_1(p)} t_a \,, \tag{10}$$

and $\pi_p(\alpha) = \pi_p(\alpha, \overline{v}, t)$. Since $\overline{v}$ and t are fixed, this should cause no confusion. Let $\{Y_i, \ i = 1, ..., N\}$ denote the set of extreme points of the polyhedron $\overline{Y}$ and define

$$D_i(\overline{v}) = \langle D(\overline{v}), Y_i \rangle \quad \text{and} \quad C_i = \langle C, Y_i \rangle \,, \quad i = 1, ..., N.$$

We introduce an important assumption that can always be enforced through a suitable perturbation of the arc cost functions.

Nondegeneracy assumption: $C_i \neq C_j$ for all distinct extreme points Y_i and Y_j of $\overline{Y}$.

Theorem 1 *[3] If the delay function D is strictly monotone on the compact polyhedron $\overline{Y}$ and the nondegeneracy assumption holds, then the equilibrium solution is unique (almost everywhere).*

Theorem 2 *[1] If the delay function D is strictly monotone on the compact polyhedron $\overline{Y}$ and the nondegeneracy assumption holds, then the function $\overline{v}(t)$ is continuous in t.*

For fixed t, the infinite-dimensional variational inequality (VI) can be solved by a linearization algorithm (see Marcotte and Zhu [3]). At each major iteration of this algorithm, one solves, for fixed $\overline{v}$, the parametric linear program

$$\mathrm{LP}(\alpha): \qquad \min_{y(\alpha)\in Y(\alpha)} \langle D(\overline{v}) + \alpha C, y(\alpha)\rangle \,, \tag{11}$$

whose solution is unique (almost everywhere) and yields a descent direction for the gap function (see [3]), under the nondegeneracy assumption. Now, in order to construct an algorithm for the bilevel program, we need to perform a sensitivity analysis of the flow patterns with respect to the toll vector t. This is achieved by partitioning the set of paths according to a domination criterion.

Definition 1 *We say that a path $p' \in P$ is **dominated** if*

$$\pi_{p'}(\alpha) > \min_{p\in P} \{\pi_p(\alpha)\}, \quad \forall \alpha \in [0, \alpha_{\max}]\,. \tag{12}$$

Otherwise, we say that p' is undominated.

Definition 2 *We say that a path $p' \in P$ is* weakly undominated *if there exists only one value $\alpha \in [0, \alpha_{\max}]$ such that*

$$\pi_{p'}(\alpha) = \min_{p\in P} \{\pi_p(\alpha)\}\,. \tag{13}$$

Definition 3 *We say that a path $p' \in P$ is* strongly undominated *if there exists a nonempty open interval $(\alpha_1, \alpha_2) \subseteq [0, \alpha_{\max}]$ such that the equality (13) holds for all $\alpha \in (\alpha_1, \alpha_2)$.*

The three domination situations are illustrated in Figure 1, where each straight line represents the perceived cost associated with a given path. In this example, path p_3 is weakly undominated since its perceived cost is minimal for a single value α_1 of the VOT parameter, while paths p_2, p_4 and p_5 are strongly undominated. Path p_1 is dominated.

The interval $[0, \alpha_{\max}]$ can be partitioned into $\bigcup_{j=0,\dots,M} \left[\alpha_{i_j}, \alpha_{i_{j+1}}\right]$ with $0 = \alpha_{i_0} < \alpha_{i_1} < \dots < \alpha_{i_M} < \alpha_{i_{M+1}}$ such that Y_{i_j} is strongly undominated on $\left[\alpha_{i_{j-1}}, \alpha_{i_j}\right]$ with

$$\alpha_{i_j} = \frac{D_{i_{j+1}}(\overline{v}) - D_{i_j}(\overline{v})}{C_{i_j} - C_{i_{j+1}}}. \tag{14}$$

The values α_{i_j} are called the **critical points** associated with the lower level program (8-9). Geometrically, each critical point corresponds to the intersection point of two lines which represent the perceived costs associated with two strongly undominated paths, as illustrated in Figure 1. It can be shown that the solution of $\mathrm{LP}(\alpha)$ is given by:

$$y(\overline{x}, \alpha) = Y_{i_k} h(\alpha) \quad \text{ for } \alpha \in \left[\alpha_{i_{k-1}}, \alpha_{i_k}\right].$$

Under the nondegeneracy assumption, this solution is unique, except at the critical points.

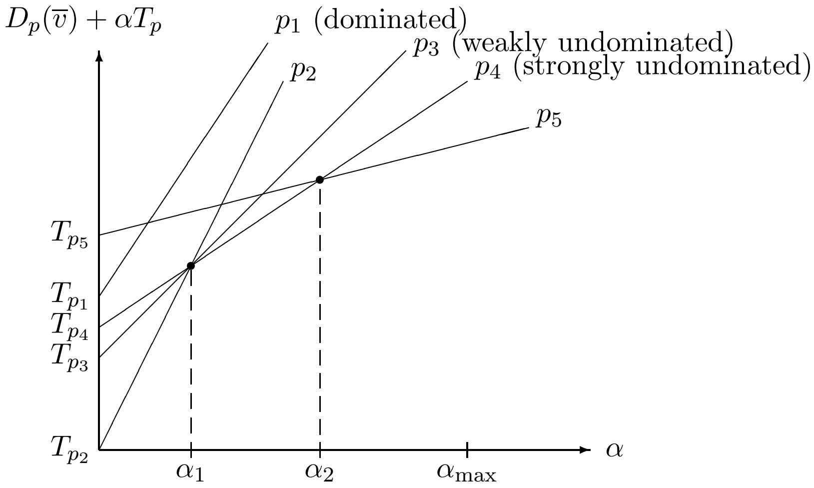

Figure 1: Dominated and undominated paths

4 The upper level program

Consider the upper level program (6-7). The following lemma states the invariance of the set of strongly undominated paths set in a neighborhood of a given toll vector t.

Lemma 1 *Let t and t' be two toll vectors. Let us denote by $(\mathcal{SU}, \mathcal{WU}, \mathcal{D})$ and $(\mathcal{SU}', \mathcal{WU}', \mathcal{D}')$ the sets of strongly undominated paths, weakly undominated paths and dominated paths induced by t and t' respectively. Then, there exists $\delta > 0$ such that*

$$\|t - t'\| < \delta \quad \Rightarrow \quad \mathcal{SU} \subseteq \mathcal{SU}'\,, \mathcal{D} \subseteq \mathcal{D}'\,. \tag{15}$$

Let us introduce explicitly origin-destination indices. For each O-D pair k, we associate with the optimal paths $p_1^k, ..., p_{M_k}^k$ the critical points $0 = \alpha_0^k, \alpha_1^k, ...\alpha_{M_k}^k, \alpha_{M_k+1}^k = \alpha_{\max}$. Under the assumptions that the set $\mathcal{WU}$ is empty and the VOT density h is continuous, the profit function R is differentiable at t and its partial derivative at $\overline{t}$ with respect to $t_a, a \in A_1$ is given by

$$\frac{\partial R(\overline{t})}{\partial t_a} = x_a(\overline{t}) + \sum_{a \in A_1} \overline{t}_a \frac{\partial x_a(\overline{t})}{\partial t_a}, \tag{16}$$

where

$$\frac{\partial \alpha_i^k}{\partial t_a} = \frac{1}{C_i^k - C_{i+1}^k} \cdot \frac{\partial \left(D_{i+1}^k - D_i^k\right)}{\partial t_a} - \frac{D_{i+1}^k - D_i^k}{\left(C_i^k - C_{i+1}^k\right)^2} \cdot \frac{\partial \left(C_i^k - C_{i+1}^k\right)}{\partial t_a} \tag{17}$$

$$\frac{\partial C_i^k}{\partial t_a} = \mathbb{1}\left\{a \in A(p_i^k)\right\} \tag{18}$$

$$\frac{\partial D_i^k(\overline{x})}{\partial t_a} = \sum_{a \in A(p_i^k)} d_a'(\overline{x}_a) \frac{\partial \overline{x}_a}{\partial t_a} \tag{19}$$

$$\frac{\partial \overline{x}_a}{\partial t_a} = \sum_{k \in K} \sum_{i=1}^{M_k} \mathbb{1}\left\{a \in A(p_i^k)\right\} \left(h(\alpha_i^k) \frac{\partial \alpha_i^k}{\partial t_a} - h(\alpha_{i-1}^k) \frac{\partial \alpha_{i-1}^k}{\partial t_a} \right). \tag{20}$$

5 A two-phase algorithm

In this section, we briefly outline a two-phase algorithm for solving the price-setting problem. In the first phase, we solve a discrete approximation of the model, using an exact branch-and-cut algorithm. This is achieved by extending the mixed-integer formulation of Labbé, Marcotte and Savard ([2]) to a formulation where the VOT density function is discretized (coarsely) and congestion functions are approximated by step functions. The resulting model (see [1]) is a large 0-1 MIP with $(|\tilde{K}| \times m) + (m \times w)$ binary variables, where $|\tilde{K}|$ is the new number of commodities (each commodity being duplicated according to the number of discrete classes), m is the number of toll arcs and w the number of steps in the approximated delay function. In the second phase of the algorithm, a local ascent algorithm, based on the gradient information derived in Section 4, is initiated at the solution obtained in the first phase. Whenever a linesearch fails to move away from the current solution[2], a small step ('serious step') is taken in the gradient direction. A pseudo-code description of the resulting algorithm is given below.

ALGORITHM `Toller`

Phase 1 Step 0: find an initial point t_0 and set $k = 0$;

Phase 2 Step 1: compute $g_k = \nabla R(t_k)$; if $\|g_k\| \le \epsilon$ then **stop**;
Step 2: is the linesearch successful at t_k ?
Yes: $\lambda_k \in \arg\min_{\lambda \ge 0} R(t_k + \lambda g_k)$;
No: $\lambda_k = 1/(k+1)$;
Step 3: set $t_{k+1} = t_k + \lambda_k g_k$, replace k by $k+1$ and return to step 1.

[2]This happens when the function is not differentiable at the current point, or is only differentiable in a very small neighborhood of the current point.

6 Numerical experiments

The performance of the algorithm has been thoroughly tested. In this abstract, we present results corresponding to a set of 10 randomly generated problems. Exhaustive results and sensitivity analysis will be presented at the conference and can be found in [1]. The networks considered contain 200 arcs (20 controlled by the leader), 30 nodes and 10 O-D pairs. By bilevel programming standards, these are very large instances.

It is clear that both the value and the computing time of the Phase 1 solution increase with the quality of the approximation, while the computing time is adversely affected. In order to assess the trade-off between 'value' and 'CPU', we ran four scenarios. In scenarios A and B, Phase 1 is not implemented and Phase 2 is initiated with tolls on all toll arcs set to either 0 or -10. Scenarios C and D implement Phase 2 in increasing coarseness of the approximations. Table 1 contains statistics pertaining to the four scenarios. The columns refer to the number of arcs used in the final solution, the number of toll arcs used, the improvement over the base scenario A, and the CPU time. For each scenario, the first line provides the mean (over 10 test cases) and the second line the standard deviation. As can be expected, the quality of the discretization in Phase 1 improve the quality of the solution. Computations were performed on a PC powered by a Pentium 1.2 Ghz processor.

scenario		#arcs	#tolls	improv.	CPU
A	ave	43.70	4.10	0	10.77
	std	5.95	1.91	0	12.73
B	ave	39.50	4.00	-3.60	2.12
	std	4.22	1.89	13.25	2.40
C	ave	44.60	4.00	3.30	9.68
2 classes, 4 steps	std	6.19	1.94	12.24	13.17
D	ave	44.90	3.80	4.60	8.99
4 classes, 4 steps	std	5.90	1.69	12.49	8.60

Table 1: Numerical results

References

[1] M. Fortin,"Tarification avec segmentation de la demande et congestion", M.Sc.A. thesis, École Polytechnique, Montréal (2004).

[2] M. Labbé, P. Marcotte and G. Savard,"A bilevel model of taxation and its application to optimal highway pricing", *Management Science* 44, 608–1622 (1998).

[3] P. Marcotte and D.L. Zhu, "Equilibria with infinitely many differentiated classes of customers", in *Complementarity and variational problems. State of the art*, J.-S. Pang and M. Ferris (eds), 234-258, SIAM (1997).