# Pricing network edges to cross a river

**Document status and date:**
Published: 01/01/2004

**DOI:**
10.26481/umamet.2004009

**Document Version:**
Publisher's PDF, also known as Version of record

# Pricing Network Edges to Cross a River

Alexander Grigoriev *      Stan van Hoesel *      Anton F. van der Kraaij *

Marc Uetz *      Mustapha Bouhtou †

April 5, 2004

## Abstract

We consider a Stackelberg pricing problem in directed networks: Tariffs (prices) have to be defined by an operator, the leader, for a subset of the arcs. Clients, the followers, choose paths to route their demand through the network selfishly and independently of each other, on the basis of minimal total cost. The problem is to find tariffs such as to maximize the operator's revenue. We consider the case where each client takes at most one tariff arc to route the demand. The problem, which we refer to as the river tarification problem, is a special case of problems studied previously in the literature.

We prove that the problem is strongly $\mathcal{NP}$-hard. Moreover, we show that the polynomially solvable case of uniform tarification yields an $m$–approximation algorithm, and this is tight. We suggest a new type of analysis that allows to improve the result to $\mathcal{O}(\log m)$, whenever the input data is polynomially bounded. We furthermore derive an $\mathcal{O}(m^{1-\varepsilon})$–inapproximability result for problems where the operator must serve all clients, and we discuss some polynomial special cases. Finally, a computational study with instances from France Télécom suggests that uniform pricing performs better in practice than theory would suggest.

## 1   Introduction

The general setup for the tarification problem involves two non-cooperative groups, operators that set tariffs, the *leaders* of the Stackelberg game, and clients that have to pay these tariffs, the *followers* of the Stackelberg game. More precisely, we assume that a network is given, and a subset of the arcs is owned by an operator. Being the owner of an arc, the operator can set the tariffs on these arcs for renting capacity to one or several clients. Each client wishes to route a certain amount of flow (a commodity) on a path connecting two vertices. Such a path can involve one or several arcs belonging to the operator, and we assume that each client selfishly selects a path with minimum total cost to route his demand. Before the clients select their paths, however, the operator has to set the tariffs, which he does in order to maximize total revenue.

In general, tarification strategies can be used to distribute traffic over a network in order to minimize some measure of system performance that does not coincide with the objectives of the individual clients. In this context, the structural analysis of (Nash) equilibria versus system optima has been heavily researched recently. The papers by Roughgarden and Tardos [13], and Cole et al. [2, 3] are some of the recent references. In the mentioned references, it is the phenomenon of network congestion what makes the clients interact with each other, thus introducing a Game Theoretic setting. In the same direction, Fortz and Thorup [5, 6] analyze pricing strategies empirically in order to prevent overloaded IP links in internet traffic routing. Notice that, in these

---

*Dept of Quantitative Economics, Maastricht University, P.O.Box 616, 6200 MD Maastricht, The Netherlands. E-mail: {a.grigoriev, s.vanhoesel, a.vanderkraaij, m.uetz}@ke.unimaas.nl

†France Télécom R&D, 39-40 rue du Général Leclerc, F-92131 Issy-Les-Moulineaux, France.   E-mail: mustapha.bouhtou@francetelecom.com

contexts, pricing need not take place before the routing of the traffic (for example, marginal cost pricing).

The problem we consider here is different in two aspects. First, we assume that there is no congestion, hence the clients do not influence each other. They choose minimum cost paths to route their commodities, independent of each other. The Game Theoretic setting is only introduced by the fact that there exist an operator trying to maximize revenue using high tariffs, and the clients try to avoid high tariffs by choosing minimal cost paths. Second, the pricing is assumed to take place before the users choose their paths, so we are faced with a Stackelberg game, where the operator (leader) first sets the tariffs, and then, subject to these tariffs, the clients (followers) react selfishly.

A natural formulation of the problem, referred to as the (general) tarification problem, is the bilevel formulation of Labbé et al. [11]: At one level the operator strives to maximize his revenue, while at the other level the clients seek to minimize the cost of routing their demands. In this formulation, the objective functions of the operator and the clients are bilinear. Roch et al. [12] show that the problem is strongly $\mathcal{NP}$-hard, even when restricted to a single client. In the same paper, a polynomial time $\mathcal{O}(\log m)$–approximation algorithm for the problem with a single client is proposed, where $m$ is the number of tariff arcs.

In this paper, we consider the problem with multiple clients. However, we assume that the path taken by any client utilizes at most one tariff arc. Several applications of this particular tarification problem, to which we refer as the river tarification problem (RTP) are briefly discussed in Section 2. Section 3 describes the model in detail. In Section 4, we show that the river tarification problem is strongly $\mathcal{NP}$-hard. The quality of uniform tarification policies, where all arcs are priced with the same tariff, is analyzed in section 5. Uniform tarification is well-known to be solvable in polynomial time, even for the general tarification problem [14]. We show that uniform tarification is an $m$–approximation, and we also show that this analysis is tight. We then propose a new type of analysis, not based on the usual geometric interval arguments, which allows to improve this to a $\mathcal{O}(\log m)$–approximation whenever the input is polynomially bounded (in terms of $m$). If the operator is forced to serve all clients, we show in Section 6 that the river tarification problem is not approximable to within a factor $\mathcal{O}(m^{1-\varepsilon})$, unless $\mathcal{ZPP} = \mathcal{NP}$. We discuss some polynomially solvable special cases in Section 7, and finally, we empirically analyze the quality of uniform tarification policies in Section 8.

## 2 Applications

Given the restriction that each client uses at most one of the tariff arcs, we can essentially restrict ourselves to specific traffic networks that resemble the situation of a town that is divided by a river. To cross the river, several possibilities exist, such as bridges, tunnels, or boats. Local government may decide to put tolls for each crossing on several of these transportation possibilities, thus creating an instance of the river tarification problem. Besides its application in transportation networks, we illustrate some actual applications of the river tarification problem in telecommunication networks, as they occur for a large global operator.

**Network structure.** First, the network structure may be such that we know a priori that the path of each client takes at most one tariff arc. Consider for example the international interconnections market, where several operators offer connections inside a particular country. If we focus on the market for this particular country, we can assume that it is not profitable for any client to enter the country twice. This gives rise to an instance of the river tarification problem.

**Autonomous systems.** Consider a TCP/IP network. Whenever an autonomous system (represented by some subnetwork) has to transit data, the data may enter and exit the autonomous system at different points. Clients have to pay a price for sending data through the autonomous system, yielding revenue for its owner. The data flow can be modelled such that once it is routed through the autonomous system, it does not pass a second time. See Figure 1(a), where the clients have the choice between two different autonomous networks ($AS1$ and $AS2$) with two entry and exit points to route their demand from source to target. This situation can be modelled as a river

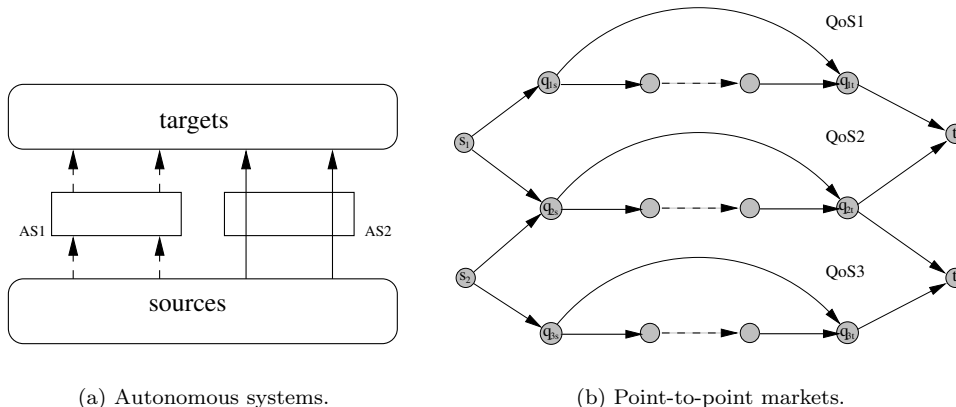(a) Autonomous systems.  (b) Point-to-point markets.

Figure 1: Applications of the river tarification problem.

tarification problem, introducing a tariff arc for each entry-exit combination.

**Point-to-point markets.** Consider a telecommunications operator that is offering bandwidth capacity between two points $A$ and $B$ at a certain price. Other operators are active in this market as well. Their prices for bandwidth capacity are known. Clients can choose between different levels of Quality of Service (QoS) from each operator, and clients have an ordered preference for the QoS-levels. We can model this problem as an instance of the river tarification problem, too. Figure 1(b) shows a small example with two customers, represented by two commodities $(s_1, t_1)$ and $(s_2, t_2)$. The operator has three QoS levels, represented by the subnetwork between the nodes $q_{is}$ and $q_{it}$, where $i \in \{1, 2, 3\}$. In this example, customer $(s_1, t_1)$ is interested in two QoS levels, namely QoS1 and QoS2, whereas customer $(s_2, t_2)$ is interested in QoS2 and QoS3. The preference of each customer $k$ with regard to each QoS level is determined by the cost of the edge from the source $s_k$ to the node $q_{is}$, $i \in \{1, 2, 3\}$, smaller cost indicating a higher preference for the QoS level. The prices of other operators for the same QoS level is given by the cost on the (fixed cost) arcs $(q_{is}, q_{it})$, $i \in \{1, 2, 3\}$. The revenue for the operator for each QoS level $i$, $i \in \{1, 2, 3\}$ is then determined by setting appropriate tariffs on the tariff arcs (dashed arcs in the figure).

## 3   Model

An instance of the general tarification problem is a directed graph $G = (N, A)$, where the arc set $A$ is partitioned into a set of $m$ tariff arcs $T \subseteq A$ and a set of fixed cost arcs $F = A \setminus T$. There are $n$ clients (or commodities) $k \in \{1, \ldots, n\}$, where each client $k$ has a demand $d_k$ that has to be routed from source node $s_k$ to target node $t_k$. Because there is no congestion involved, we may assume without loss of generality that all demand values $d_k$ are scaled to be integral. We define for a commodity $k$ the set of all possible paths from $s_k$ to $t_k$ by $P_k$. The tariff on a tariff arc $a \in T$ is denoted by $\tau_a$, and the vector of all tariffs is given by $\tau = (\tau_a)_{a \in T}$. The cost of a fixed cost arc $a \in F$ is denoted by $c_a$.

The clients route their demands from source to destination according to a path with minimal total cost, where the total cost of a path is defined as the sum of the tariffs and fixed costs on the arcs of the path. Whenever the client has a choice among multiple paths with the same total cost but with different revenues for the operator, we assume that the client takes the path that is most profitable to the operator. This can always be achieved with arbitrary precision by reducing all tariffs by some small value $\varepsilon$. We assume that an $\{s_k, t_k\}$-path exists that consists only of fixed cost arcs for every client $k \in \{1, \ldots, n\}$, since the problem is otherwise unbounded.
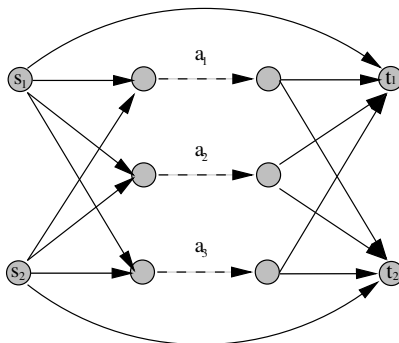
3

Figure 2: River tarification problem with $n = 2$ and $m = 3$.

The revenue associated with a path $p \in P_k$ induced by a client $k$ with demand $d_k$ is defined by

$$\pi_p(\tau, d_k) = d_k \sum_{a \in T \cap p} \tau_a \,, \tag{1}$$

whereas the fixed cost of the path $p$ is given by

$$c_p(d_k) = d_k \sum_{a \in F \cap p} c_a \,. \tag{2}$$

Then $l_p(\tau, d_k) := c_p(d_k) + \pi_p(\tau, d_k)$ is the total cost of the path $p \in P_k$ for a client $k$. Obviously, $l_p(\tau, d_k)$ is a piecewise linear, non-decreasing function in $\tau$, and linear, non-decreasing in $d_k$. The (general) tarification problem can now be described by the following bilevel program; see [11, 1].

$$\begin{aligned} \max_{\tau} \quad & \sum_{k \in K} \pi_{p_k^*}(\tau, d_k) \\ s.t. \quad & p_k^* = \arg\min_{p \in P_k} l_p(\tau, d_k) \qquad \text{for all } k \in \{1, \dots, n\}. \end{aligned} \tag{3}$$

For the bilevel program in (3), at the upper level the revenue of the operator is maximized, while at the lower level the clients take the shortest path, given the tariffs determined at the upper level. Since the demands are known a priori, objective functions of both the operator and the clients in (3) are linear and hence we are dealing with a so-called linear-linear bilevel program. Jeroslow [9] showed that linear-linear bilevel problems are $\mathcal{NP}$-hard in general. For references on bilevel programming, see Dempe [4] and Vicente and Calamai [15] who have compiled an annotated bibliography on this subject.

The additional assumption in the river tarification problem is that, independent of the tariffs, any client routes his demand only on paths which include at most one tariff arc. This will allow us to assume without loss of generality that the network has a very special structure. Replacing shortest paths using only fixed cost arcs by direct arcs, and possibly introducing some dummy arcs with zero or infinite cost, we obtain a shortest path graph model (SPGM) as defined by Bouhtou et al. [1]: All tariff arcs are disjoint, and there exists an arc from any source node $s_k$ to the tail node of any tariff arc, and from the head node of any tariff arc to any target node $t_k$. Moreover, there exists a a fixed cost arc $(s_k, t_k)$ for all $k = 1, \dots, n$, and the cost $c_k$ is the highest acceptable price for client $k$. Figure 2 shows the shortest path graph model of an instance of the river tarification problem with three tariff arcs and two clients. The tariff arcs $a_i, i \in \{1, 2, 3\}$ are given by the dashed arcs in the network. We may also assume without loss of generality that all fixed cost arcs incident with the target nodes $t_k$ have zero cost (by adding their costs to the fixed cost arcs incident with $s_k$).

The essential parameters that define an instance of river tarification problem are therefore the number of tariff arcs $m$, the number of clients $n$, the demand values $d_k, k \in \{1, \dots, n\}$, and the costs $c_a$ of the fixed cost arcs $a \in F$. Due to the fact that any path taken by a client involves
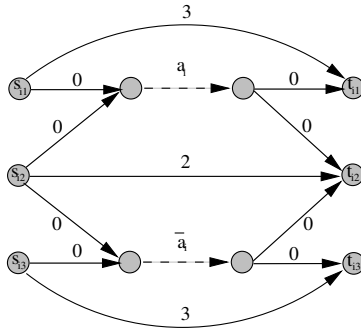
Figure 3: Subnetwork for variable $x_i$, $i \in \{1, \ldots, n\}$.

exactly one fixed cost arc with non-zero cost, we may assume without loss of generality that the costs $c_a$ of the fixed cost arcs $a \in F$ are integral. Moreover, due to the integrality of the costs of the fixed cost arcs, it is immediate that any reasonable solution will adopt only tariffs which are integral, too. Notice, however, that this is not true for the general tarification problem, where a path chosen by a client can consist of several tariff arcs. There, even if the costs of the fixed cost arcs are integral, it may be that an optimal solution utilizes fractional tariffs.

## 4   Complexity

Roch et al. [12] show that the general tarification problem is $\mathcal{NP}$-hard in the strong sense, even when restricted to a single client, using a reduction from the $\mathcal{NP}$-complete problem 3-SAT [7]. Their reduction works for tarification problems where paths are allowed to use (and indeed, must use) several tariff arcs. We show that the tarification problem with multiple clients, but restricted to at most one tariff arc per path, is $\mathcal{NP}$-hard in the strong sense, too.

We also use a reduction from 3-SAT. Therefore, consider a boolean function $f : \{0, 1\}^n \to \{0, 1\}$ on $n$ variables $x_1, \ldots, x_n$, in conjunctive normal form. Such a function $f$ is the conjunction of $m$ clauses $C_k$,

$$f = \bigwedge_{k=1}^{m} C_k \,, \tag{4}$$

each clause $C_k$ being the disjunction of three literals, $C_k = (\ell_{k1} \vee \ell_{k2} \vee \ell_{k3})$. Any literal $\ell_{kj}$ represents either a variable $x_i$, or its negation $\bar{x}_i$, $i \in \{1, \ldots, n\}$. Then $f$ is satisfiable if there exists a truth assignment $x_1, \ldots, x_n$ such that at least one literal per clause is true.

Any function of the form (4) can be polynomially transformed to an instance of the river tarification problem as follows. For each variable $x_i$, $i \in \{1, \ldots, n\}$, we construct a constant-size subnetwork as shown in Figure 3. Each of these subnetworks has three clients with unit demand, with origin-destination pairs $\{s_{ij}, t_{ij}\}$, $j \in \{1, 2, 3\}$. Moreover, each subnetwork has two tariff arcs, $a_i$ representing the truth assignment $x_i = 1$, and $\bar{a}_i$ representing $x_i = 0$, as depicted in Figure 3.

An upper bound on the cost of routing commodities 1 and 3 is given by fixed cost arcs $(s_{i1}, t_{i1})$ and $(s_{i3}, t_{i3})$, both with cost 3. For commodity 2, the upper bound on the cost is given by a fixed cost arc $(s_{i2}, t_{i2})$, with cost 2. The maximal revenue for each subnetwork is thus given by setting one of the tariffs to 2, and the other to 3, yielding a revenue of $2 \cdot 2 + 3 = 7$. In all other cases, the revenue is not more than 6.

Next, for each clause $C_k$, $k \in \{1, \ldots, m\}$, we create a *clause-commodity* $k$ with origin destination pairs $\{s_k, t_k\}$, with unit demand. Whenever a variable $x_i$ ($\bar{x}_i$, respectively) appears as one of clause $C_k$'s literals, we connect $s_k$ to $s_{i1}$ ($s_{i3}$, respectively), and $t_{i1}$ ($t_{i3}$, respectively) to $t_k$, using arcs of zero cost. In addition, we introduce a fixed cost arc $(s_k, t_k)$ with cost 2, defining an upper bound of 2 for the cost of routing clause-commodity $k$. The so-defined instance of the river

tarification problem has $2n$ tariff arcs, $3n + m$ commodities (or clients), and $7m + 11n$ fixed cost arcs, hence the transformation is indeed polynomial. Example 1 illustrates the transformation.

**Example 1.** The boolean function $f = (x_1 \vee x_2 \vee x_3) \wedge (\bar{x}_1 \vee \bar{x}_2 \vee x_4)$ on 4 variables $x_1, \ldots, x_4$ can be transformed to the network displayed in Figure 4. A tariff arc $a_i$ (resp. $\bar{a}_i$) represents the truth assignment $x_i = 1$ (resp. $x_i = 0$), $i \in \{1, \ldots, 4\}$.

**Theorem 1.** *The river tarification problem is strongly $\mathbb{NP}$-hard.*

*Proof.* Consider the polynomial transformation defined previously. We show that a satisfying truth assignment for $f$ exists if and only if the revenue for the river tarification problem is equal to $2m + 7n$. Suppose there exists a satisfying truth assignment. Then at least one literal in each clause $C_k$, $k \in \{1, \ldots, m\}$ is fulfilled. Take for each $C_k$ one fulfilled literal and set the tariff on its corresponding tariff arc to 2. The complementary tariff arc receives tariff 3. Due to the fact that we start with a valid truth assignment, this is indeed well-defined, since either $x_i = 1$ or $x_i = 0$, but never both. This way, the total revenue from all clause-commodities is $2m$. For all remaining subnetworks, if any, the two tariffs can be set arbitrarily to 2 and 3, respectively. The revenue from all subnetworks is thus $7n$. Hence, the total revenue is $2m + 7n$.

Conversely, suppose there exists a set of tariffs such that the total revenue is $2m + 7n$. The maximal possible revenue created by all except the clause-commodities is $7n$, only achievable by setting one tariff arc per subnetwork to 2 and the other to 3. On the other hand, the maximal possible revenue created by the clause-commodities is $2m$. Hence, in order to achieve a revenue of exactly $2m + 7n$, we must have both, $2m$ from the clause-commodities, and $7n$ from the remaining commodities. We now define a truth assignment, setting 'true' all literals that correspond to tariff arcs of cost 2. This is a well-defined truth assignment, since we know that in each subnetwork one tariff is 2 and the other 3. Moreover, each of the $m$ clause-commodities contributes a revenue of 2, hence it must use a path with a tariff-arc of cost 2. In other words, for each clause at least one literal is 'true', hence the truth assignment satisfies all clauses. □

The reduction used for the proof of Theorem 1 shows that the river tarification problem remains $\mathbb{NP}$-hard even for unit demands, a fixed number of tariff values and when the operator is forced to use tariffs such that he serves (a given subset of) all clients.

## 5 The quality of uniform tarification policies

The uniform tarification problem (UTP) is the same problem as the general tarification problem, with the additional restriction that all tariffs are required to be identical. As shown by van Hoesel et al. [14], the uniform tarification problem can be solved in polynomial time, even in the general setting where clients may use paths with several tariff arcs. The algorithm described in van Hoesel et al. [14] uses the parametric shortest path algorithm of Young et al. [16] and Karp and Orlin [10] to determine the tariff values (i.e. breakpoints) for which the shortest path tree changes for any client. Calculating the revenue for the operator at each breakpoint and maintaining the best solution yields the optimal uniform tarification policy in polynomial time.

We are interested in the loss that can be experienced by adopting a uniform tarification policy. Apart from pure theoretical interest, the question is motivated out of the desire to – in the long run – extend to models with more than one operator. In that situation, we have a Game Theoretic setting between operators for the determination of the tariffs. In that context, the quality of an operator's strategy moves into the foreground - uniform tarification being a strategy that can be implemented efficiently (i.e., in polynomial time).

Therefore, denote by $\Pi^{UTP}$ the revenue for an optimal uniform tarification, and by $\Pi^{\mathrm{OPT}}$ the revenue for an optimal non-uniform tarification. By definition, $\Pi^{\mathrm{UTP}} \leq \Pi^{\mathrm{OPT}}$.

**Lemma 1.** *If an optimal tarification for the river tarification problem with revenue $\Pi^{\mathrm{OPT}}$ utilizes at most $r$ different tariffs, then for the optimal uniform tarification, $\Pi^{\mathrm{UTP}} \geq \Pi^{\mathrm{OPT}}/r$.*
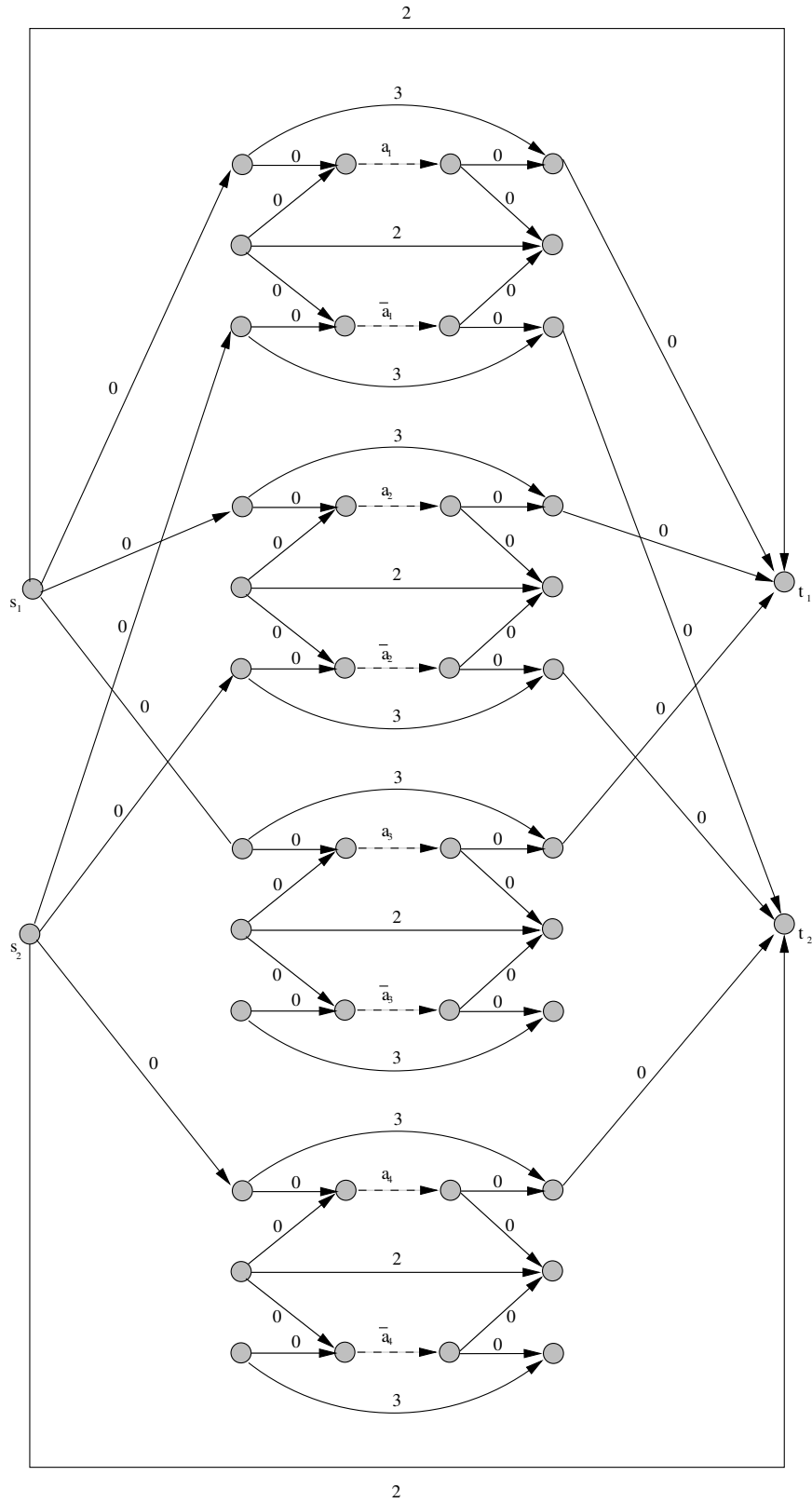
Figure 4: Network for $(x_1 \vee x_2 \vee x_3) \wedge (\bar{x}_1 \vee \bar{x}_2 \vee x_4)$.
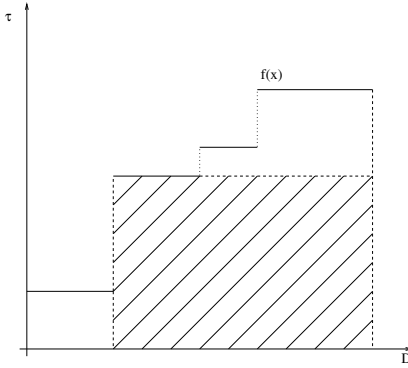
Figure 5: Staircase function $f(x)$ with inscribed rectangle.

The proof of this lemma is quite trivial once the structure of an optimal solution to the river tarification problem is clear. To this end, consider an optimal non-uniform tarification with tariffs $\tau_1 \leq \cdots \leq \tau_m$, and let $D_i$ be the total demand on an arc $a_i$ with tariff $\tau_i$, $i \in \{1, \ldots, m\}$. By $D = \sum_{i=1}^{m} D_i$ we denote the total demand served by the operator. Then the revenue created by this solution is the area under the following 'staircase' function $f : [0, D] \to [0, \infty[$, illustrated in Figure 5.

$$f(x) = \tau_i \quad \text{for all } x \text{ with } \sum_{j < i} D_j \leq x < \sum_{j \leq i} D_j, \quad i \in \{1, \ldots, m\}. \tag{5}$$

*Proof of Lemma 1.* Consider any of the rectangles inscribed under the graph of function $f(x)$, with area $T_i := \tau_i \cdot \sum_{j \geq i} D_j$. Then it holds that $\Pi^{\mathrm{UTP}} \geq T_i$ for all $i \in \{1, \ldots, m\}$, since the area of any such rectangle is a lower bound for the revenue yielded by the optimal uniform tariff $\Pi^{\mathrm{UTP}}$. This claim is true because, starting with the optimal set of tariffs $\tau_1, \ldots, \tau_m$, and levelling all tariffs to some value $\tau_j$, creates a revenue loss of at most $\sum_{i:\tau_i < \tau_j} D_i \tau_i + \sum_{i:\tau_i > \tau_j} D_i(\tau_i - \tau_j)$. (Notice that this does not hold for the general tarification problem.) Hence, if only $r$ different tariffs are utilized, we consider the $r$ (inclusion-)maximal rectangles under function $f$, say $T_{i_1}, \ldots, T_{i_r}$, and get $r \cdot \Pi^{\mathrm{UTP}} \geq \sum_{j=1}^{r} T_{i_j} \geq \Pi^{\mathrm{OPT}}$.

$\square$

Since $r \leq m$, Lemma 1 yields the following theorem. Tightness of the result will be shown below, using Example 2.

**Theorem 2.** *Uniform tarification is an m-approximation for the river tarification problem, and this bound is tight.*

We next derive an improved bound on the quality of uniform tarification policies for the case of bounded input data. To this end, denote by $\tau_{\max}$ an upper bound on the maximal possible tariff. For example, a trivial value for $\tau_{\max}$ is given by the maximal cost of any fixed cost arc, $\tau_{\max} = \max_{a \in F} c_a$. Then, if $D$ is the total demand captured by the operator in an optimal non-uniform tarification,

$$U := D \cdot \tau_{\max} \tag{6}$$

is an upper bound for the optimal revenue $\Pi^{\mathrm{OPT}}$.

**Lemma 2.** *If $\Pi^{\mathrm{OPT}}$ is the revenue for an optimal tarification for the river tarification problem, and $U$ is the upper bound on $\Pi^{\mathrm{OPT}}$ as defined in (6), then for the optimal uniform tarification, $\Pi^{\mathrm{UTP}} \geq \Pi^{\mathrm{OPT}}/(1 + \log U)$.[1]*

---

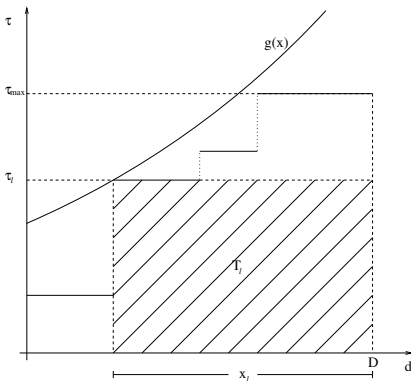[1]With $\log x$ we denote the natural logarithm, i.e., $e^{\log x} = x$.

Figure 6: Illustration for the proof of Lemma 2.

*Proof.* Indeed, we will even prove a slightly stronger result than claimed in Lemma 2. First recall the definition of the staircase function $f$ in (5), as well as the inscribed rectangles, with areas $T_i = \tau_i \cdot \sum_{j \geq i} D_j$. Let $\ell$ be the index of the maximal of these $m$ rectangles, with area $T_\ell$. Let $x_\ell := \sum_{j \geq \ell} D_j = T_\ell / \tau_\ell$. We show

$$\Pi^{\mathrm{UTP}} \geq \frac{\Pi^{\mathrm{OPT}}}{1 + \log(U/T_\ell)} \,. \tag{7}$$

Lemma 2 then follows, because we know that the optimal tariff only adopts integral values, and $T_\ell \geq 1$. To prove (7), we define the function

$$g(x) := \frac{T_\ell}{D - x} \text{ for } x \in [0, D) \,. \tag{8}$$

We claim that $g(x) \geq f(x)$ for $x \in [0, D)$. To see this, take any $x$ with $\sum_{j < i} D_j \leq x < \sum_{j \leq i} D_j$, then $f(x) = \tau_i$ by definition. However,

$$g(x) = \frac{T_\ell}{D - x} \geq \frac{T_\ell}{D - \sum_{j < i} D_j} = \frac{T_\ell}{\sum_{j \geq i} D_j} = \frac{T_\ell}{T_i / \tau_i} \geq \tau_i = f(x) \,,$$

where the first inequality follows by choice of $x$, and the last follows by choice of $\ell$ as the index of the largest rectangle.

Hence, the area under the staircase function, which equals $\Pi^{\mathrm{OPT}}$, can be upper bounded in terms of the area defined by the function $g(x)$, as depicted in Figure 6. To compute this area, we partition it into three parts, namely the rectangle $T_\ell$ itself, the area under $g(x)$ on the domain $x \in [0, D - x_\ell]$, as well as the area to the right of $g(x)$ on the domain $\tau \in [\tau_\ell, \tau_{\max}]$. The latter is the integral of the function $D - g^{-1}(\tau) = T_\ell / \tau$ on the domain $[\tau_\ell, \tau_{\max}]$. We thus obtain the following.

$$
\begin{aligned}
\Pi^{\mathrm{OPT}} \leq \ T_\ell \ + \int_0^{D - x_\ell} \frac{T_\ell}{D - x} \, dx + \int_{\tau_\ell}^{\tau_{\max}} \frac{T_\ell}{\tau} \, d\tau \ &= \ T_\ell [1 + \log D + \log \tau_{\max} \\
&\qquad - \log \tau_\ell - \log x_\ell] \\
&= \ T_\ell \left[ 1 + \log U - \log T_\ell \right] \\
&= \ T_\ell \left[ 1 + \log(U/T_\ell) \right] \,,
\end{aligned}
$$

and since $T_\ell \leq \Pi^{\mathrm{UTP}}$, claim (7) follows.

$\square$

Notice that claim (7) confirms the following geometric intuition: The closer the staircase function $f(x)$ is to the straight line $x \mapsto (\tau_{\max}/D) \cdot x$, the closer is $T_\ell$ to $U/4$, which yields
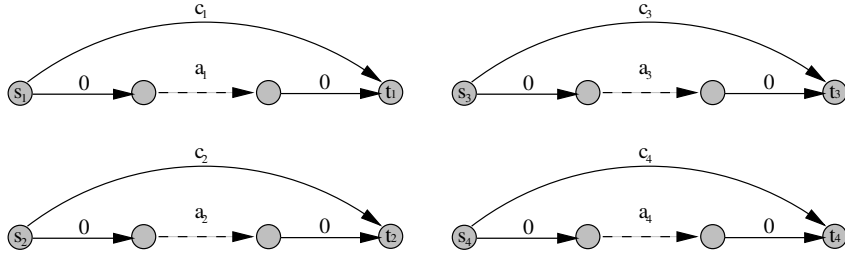
Figure 7: The analysis of uniform tarification policies is tight.

an approximation ratio of $(1 + \log 4) \approx 2.4$ for uniform tarification. Geometric intuition indeed suggests a ratio of roughly 2, the additional 0.4 being caused be the difference between the functions $g(x)$ and $f(x)$. In Section 8, we compare the quality of uniform versus non-uniform tarification, based on instances obtained from France Télécom.

**Theorem 3.** *If both demand per client $d_k$, $k \in \{1, \dots, n\}$, and maximal tariff $\tau_{\max}$ are polynomially bounded in terms of $m$ (the number of tariff arcs), then uniform tarification is an $\mathcal{O}(\log m)$-approximation for the river tarification problem.*

*Proof.* Since $d_k \in \text{poly}(m)$ for all $k$ and $\tau_{\max} \in \text{poly}(m)$, we have that $1 + \log U \in \mathcal{O}(\log m)$. Thus, the theorem follows from Lemma 2.

$\square$

Finally, let us show tightness of the bounds in Lemmas 1 and 2, and thus also tightness of Theorem 2.

**Example 2.** Given $m$ commodities and $m$ tariff arcs. Every commodity is operating its own subnetwork with one tariff arc, that is, the entire network consists of $m$ disjoint subnetworks and each of them contains one commodity and one tariff arc. Fix an integer $b \geq 3$ and let the demand in subnetwork $k$ be given by $d_k = b^k - b^{k-1}$, $k \in \{1, \dots, m\}$. This way, the total demand $D$ equals $b^m - 1$. Moreover, the maximal revenue for subnetwork $k$ is limited by a fixed cost arc $(s_k, t_k)$, with cost $c_k = b^{2m-k}$. Hence, the maximal tariff $\tau_{\max}$ equals $b^{2m-1}$.

In the optimal solution, the tariff for each subnetwork $k$ is set to its maximal value, $b^{2m-k}$. Each subnetwork therefore contributes a revenue of $b^{2m} - b^{2m-1}$, and $\Pi^{\text{OPT}} = m(b^{2m} - b^{2m-1})$. The optimal uniform tarification consists in setting the tariff on all tariff arcs to $b^m$. This way, every unit of demand creates a profit of $b^m$, yielding a total revenue of $b^{2m} - b^m$. Other (reasonable) uniform tariffs would be values $b^{2m-k}$, $k \in \{1, \dots, m-1\}$. This yields a total revenue of $b^{2m} - b^{2m-k}$, which is less. Therefore, we obtain

$$\Pi^{\text{UTP}}/\Pi^{\text{OPT}} = \frac{b^{2m} - b^m}{m(b^{2m} - b^{2m-1})} \leq \frac{b^{2m}}{m(b^{2m} - b^{2m-1})} = \frac{1}{m} \cdot \frac{b}{b-1}.$$

Now, observe that in the optimal solution $m$ different tariffs are utilized. Lemma 1 (Theorem 2, respectively) suggests that uniform tarification provides an $m$–approximation. Example 2 proves that this is best possible, since $b$ can be chosen arbitrarily large.

Moreover, Lemma 2, respectively claim (7), suggests that uniform tarification yields a $(1 + \log(U/T_\ell))$–approximation. Here, $T_\ell$ is the largest rectangle under the function $f$; see the proof of Lemma 2. In Example 2, $U = \tau_{\max} \cdot D = (b^m - 1) \cdot b^{2m-1}$ and $T_\ell = b^{2m}$. We thus have $U/T_\ell = (b^m - 1)/b$. Hence, $(1 + \log(U/T_\ell)) = 1 + \log(b^m - 1) - \log b \leq 1 + (m-1)\log b \leq m \log b$. Hence, Lemma 2 yields that uniform tarification is an $\mathcal{O}(m)$–approximation on this example (the hidden constant being $\log b$), while the same Example 2 shows that $\mathcal{O}(m)$ is indeed best possible. Notice that this does not contradict the claim of Theorem 3, since in Example 2 neither tariffs nor demands are polynomially bounded in terms of $m$, the number of tariff arcs.
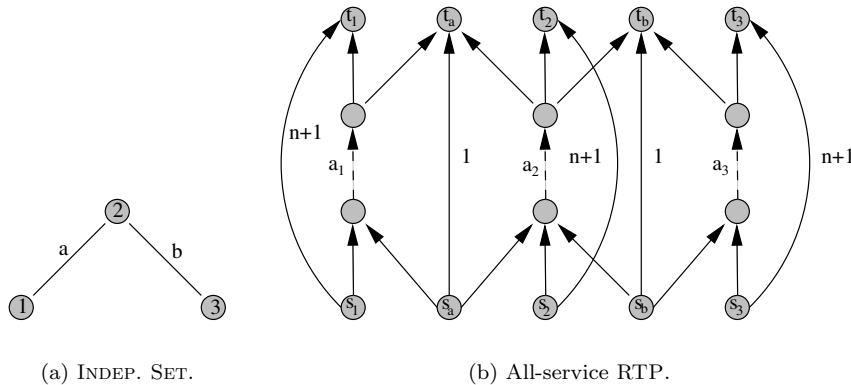
(a) Indep. Set.  (b) All-service RTP.

Figure 8: Reduction of Independent Set to all-service RTP.

# 6 Inapproximability of the river tarification problem

In this section, we consider the following variation of the river tarification problem: The operator must set tariffs in order to capture the demand of all clients, that is, tariffs must be such that no client $k$ is forced to use the arc $(s_k, t_k)$. We refer to this problem as the *all-service* river tarification problem. NP-hardness of this problem follows by our previous reduction presented in Section 4. We show:

**Theorem 4.** *For any $\varepsilon > 0$, the existence of a polynomial time approximation algorithm for the all-service river tarification problem with with $n$ clients and $m$ tariff arcs*

- *with worst case ratio $O(m^{1/2-\varepsilon})$ implies $\mathcal{P} = \mathcal{NP}$;*

- *with worst case ratio $O(m^{1-\varepsilon})$ implies $\mathcal{ZPP} = \mathcal{NP}$;*

- *with worst case ratio $O(n^{1/4-\varepsilon})$ implies $\mathcal{P} = \mathcal{NP}$;*

- *with worst case ratio $O(n^{1/2-\varepsilon})$ implies $\mathcal{ZPP} = \mathcal{NP}$.*

*Proof.* The proof uses an approximation preserving reduction from Independent Set [7] to the all-service RTP.

So assume we are given a graph $G = (V, E)$, and the problem is to find a maximum cardinality subset $V' \subseteq V$ of vertices such that no two vertices in $V'$ are connected by an edge. The transformation works as follows. For every vertex $v \in V$ we introduce a client with origin-destination pair $\{s_v, t_v\}$ and demand $d_v = |E|$, and a corresponding tariff arc $a_v$. We connect the source $s_v$ to the tail of the tariff arc $a_v$, and the head of $a_v$ to the destination $t_v$, using zero cost fixed cost arcs. Moreover, there is a fixed cost arc $(s_v, t_v)$ with cost $(|V| + 1)$ for all vertices $v \in V$. For every edge $e \in E$ we introduce a client with origin-destination pair $\{s_e, t_e\}$ and unit demand. The upper bound on the cost of routing this demand is given by the fixed cost arc $(s_e, t_e)$ with cost 1. For all edges $e \in E$ and all vertices $v \in V$ with $v \in e$, we furthermore introduce fixed cost arcs $(s_e, \text{tail}(a_v))$ and $(\text{head}(a_v), t_e)$, with zero cost. This transformation results in an instance of the all-service RTP with $|V|$ tariff arcs, and $|V| + |E|$ clients. Figure 8 gives an example of such a transformation for a graph $G = (V, E)$ with 3 nodes and 2 edges.

We claim that $G$ has an independent set of cardinality at least $k$ if and only if there exists a tariff policy for the all-service RTP with a total revenue of $|V||E|(k + 1) + |E|$.

First, assume that $G$ has an independent set $V'$ of cardinality $k$. For all $v \in V'$, set the tariff on the corresponding tariff arc $a_v$ to $|V| + 1$, and all other tariffs to 1. By the definition of an independent set, for any edge $e = (v, u) \in E$ at least one of the vertices, $v$ or $u$, is not in $V'$. Therefore, the tariff of at least one of the tariff arcs, $a_v$ or $a_u$ is 1. All clients corresponding to an

11

edge $e$ can thus be served, using one of the tariff arcs $a_v$ or $a_u$. The clients $(s_v, t_v)$ corresponding to the vertices $v \in V$ are also served, since the upper bound of $|V| + 1$ is not exceeded with the so-defined tariffs. Hence, all demands are served. The revenue consists of $|E|$ from all clients corresponding to the edges $E$ of $G$, $|E|(|V|+1)k$ from the clients corresponding to the independent set $V'$, and $|E|(|V| - k)$ from the clients corresponding to $V \setminus V'$. That yields a total revenue of $|E||V|(k + 1) + |E|$.

Conversely, assume that there exists a set of tariffs that captures all demands, such that the revenue is $|E||V|(k+1)+|E|$. We will show that this implies that the graph $G$ has an independent set of cardinality at least $k$. Since all demands are captured at this tarification strategy, for any edge $e = (v, u) \in E$, the tariff on at least one of the arcs, $a_v$ or $a_u$, is 1. Consider the set of vertices $V' := \{v \in V : t_{a_v} > 1\}$. By definition, no pair of nodes $v, u \in V'$ is connected by an edge. Hence, $V'$ is an independent set in $G$. Let $k' := |V'|$. The revenue is equal to $|E| + |E|(|V| - k') + |E|(|V| + 1)k' = |E||V|(k' + 1) + |E|$, which by assumption is at least as large as $|E||V|(k + 1) + |E|$. This implies that $k' \geq k$ and thus that $V'$ is an independent set in $G$ of cardinality $k' \geq k$.

Now, let us assume that we have an $\alpha$-approximation algorithm $\mathcal{A}$ for the all-service RTP, with $\alpha \geq 1$. Consider any instance $G = (V, E)$ of INDEPENDENT SET, and the all-service RTP resulting from the above reduction. We can assume that both the optimal solution and the solution produced by $\mathcal{A}$ only utilize tariff values 1 or $|V| + 1$, because any tariff greater than 1 and not equal to $|V| + 1$ can be turned into $|V| + 1$ with a revenue gain. So $\Pi^{\text{OPT}} = |E||V|(k + 1) + |E|$ for some $k$, and $\Pi^{\mathcal{A}} = |E||V|(k' + 1) + |E|$ for some $k'$. The first part of the proof yields that the maximal independent set of $G$ has size $k$, and algorithm $\mathcal{A}$ can be used to find an independent set of size at least $k'$. Moreover,

$$\frac{1}{\alpha} \leq \frac{|E||V|(k' + 1) + |E|}{|E||V|(k + 1) + |E|} = \frac{1 + \frac{1}{|V|} + k'}{1 + \frac{1}{|V|} + k} \leq \frac{2 + k'}{1 + k},$$

hence $k' \geq (k + 1)/\alpha - 2$. In other words, we have an $\mathcal{O}(\alpha)$–approximation algorithm for the INDEPENDENT SET problem.

It is now well known from work of Håstad [8] that the INDEPENDENT SET problem cannot have a polynomial time approximation algorithm with worst case guarantee $\mathcal{O}(|V|^{1/2-\varepsilon})$ unless $\mathcal{P} = \mathcal{NP}$, and that it cannot have a polynomial time approximation algorithm with worst case guarantee $\mathcal{O}(|V|^{1-\varepsilon})$ unless $\mathcal{ZPP} = \mathcal{NP}$. Since the number of tariff arcs $m$ in our transformation equals $|V|$, the first two claims of the theorem follow. Since the number of clients $n$ in our transformation equals $|V| + |E| \in \mathcal{O}(|V|^2)$, the second two claims follow. $\qquad\square$

It is probably worthy to mention that this inapproximability result still holds when all input data is polynomially bounded (in terms of $m$); this follows directly from our reduction.

## 7 Polynomially solvable special cases

Since the river tarification problem is a special case of the general tarification problem, the respective results known for the general tarification problem hold. Among these are those discussed by Labbé et al. [11]: the case with multiple commodities, but one tariff arc (single arc linear tarification problem); and the case with one client where the path in the network taken by the client in the optimal solution is known a priori (fixed path linear tarification problem). Furthermore some special cases are discussed by Van Hoesel et al. [14]: the case with multiple commodities and a linear pricing strategy, where the number of tariff arcs is constant (bounded arcs linear tarification problem) and the case with multiple commodities and parametric tarification (parametric tarification problem).

In addition, the river tarification problem is also polynomially solvable if the number of clients $n$ is bounded from above by a constant. In that case, the number of assignments of clients to tariff arcs is bounded by $m^n$ which is a polynomial for fixed $n$. Consider therefore the following
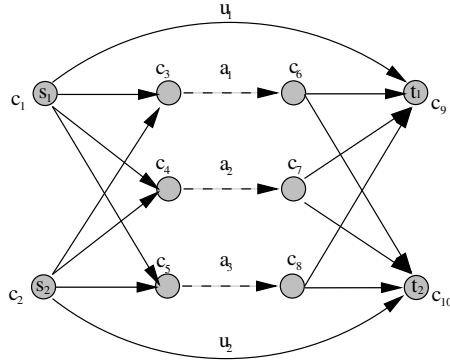
Figure 9: Example of a special cost structure ($n = 2, m = 3$).

formulation, where the path taken by each client in the network is denoted by $p_k^* \in P_k$. Recall that $P_k$ represents the set of all possible paths taken by a client $k \in K$.

$$
\begin{aligned}
\max_{\tau} \quad & \sum_{k \in K} \pi_{p_k^*}(\tau, d_k) \\
s.t. \quad & l_p(\tau, d_k) \geq l_{p_k^*}(\tau, d_k) \quad \forall k \in K, \forall p \in P_k \\
& \tau_a \geq 0 \quad \quad \quad \quad \quad \forall a \in T
\end{aligned} \tag{9}
$$

Since for each client, there are at most $m + 1$ paths in the network, $|P_k|$ is bounded by $m + 1$. Hence, the number of constraints is polynomial in the input data. Therefore, if we solve $m^n$ instances of (9), we can retrieve the optimal solution in polynomial time.

Finally, for specific cost structures of the network, the river tarification problem is polynomially solvable, too. Denote by $u_k$ the upper bound on the cost for each client $k \in \{1, \ldots, n\}$. For each node $v \in N$ in the shortest paths network define a value $c_v \geq 0$, and assume that the cost of each fixed cost arc $(i, j)$ is given by $(c_i + c_j)$, except for the arcs $(s_k, t_k)$, which represent the upper bounds $u_k$. See Figure 9 for an example. We can now transform the costs of fixed cost arcs such that the cost of using a tariff arc is the same for all commodities, by incorporating the costs $c_{s_k}$ and $c_{t_k}$ in the upper bound: $u_k \leftarrow u_k - (c_{s_k} + c_{t_k})$. For this new instance, in an optimal solution all clients use the same tariff arc. We can thus find the optimal solution of this special case of RTP by solving $m$ instances of the single arc linear tarification problem.

## 8   Numerical Results

As stated previously, whenever the function that describes the total revenue in an optimal non-uniform solution, i.e. the staircase function defined in (5), is close to a straight line, geometric intuition suggests a worst-case ratio for uniform tarification of approximately 2. The worst case Example 2 crucially hinges on a (staircase) function that approximates a hyperbola. Thus, it can be conjectured that the empirical performance of uniform tarification policies outperforms the theoretical bounds we have found. This is indeed confirmed in the following numerical experiments, displayed in Table 1. The study is based on instances obtained from France Télécom.

These instances represent telecommunication networks for the international interconnections market, as described in Section 2. We compare the optimal solutions for both uniform tariffs ($\Pi^{UTP}$) and non-uniform tariffs ($\Pi^{OPT}$). The optimal non-uniform solution is calculated using the model and mixed integer programming formulation described in Bouhtou et al. [1]. The value of $\Pi^{UTP}$ is calculated using the same formulation, requiring that all tariffs be equal. As such, we do not compare the actual computation times, but are just interested in the effectiveness of the optimal uniform tarification policies. Table 1 gives a brief description of each network, stating the number of nodes, arcs, tariff arcs and clients. The optimal non-uniform and uniform solution

Table 1: Quality of Uniform Tarification on France Télécom instances.

| Instance | $\|N\|$ | $\|A\|$ | $m$ | $n$ | $\Pi^{OPT}$ | $\Pi^{UTP}$ | % |
|---|---|---|---|---|---|---|---|
| RTN1 | 29 | 94 | 7 | 15 | 841 | 624 | 74% |
| RTN2 | 29 | 98 | 6 | 21 | 4099 | 3496 | 85% |
| RTN3 | 59 | 206 | 10 | 13 | 1118 | 880 | 79% |
| RTN4 | 59 | 204 | 10 | 20 | 2217 | 1512 | 68% |
| RTN5 | 49 | 120 | 9 | 21 | 74948 | 55968 | 74% |
| RTN6 | 33 | 116 | 15 | 12 | 28166 | 20328 | 72% |

values are displayed in the columns $\Pi^{OPT}$ and $\Pi^{UTP}$. The final column is the approximation ratio.

# 9 Conclusion

Several questions remain for further research. It remains open at this point if there exists a better approximation algorithm than just uniform tarification, or if the problem is inapproximable even without the all-service assumption. Moreover, it would be interesting to analyze which type of results can be extended to the general tarification problem.

# Acknowledgment

# References

[1] M. Bouhtou, S. van Hoesel, A. F. van der Kraaij, and J.L. Lutton. Tariff optimization in networks. *Meteor Research Memorandum*, RM03011, 2003.

[2] R. Cole, Y. Dodis, and T. Roughgarden. How much can taxes help selfish routing? In *Proceedings of the 4th Annual ACM Conference on Electronic commerce*, pages 98–107, 2003.

[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] S. Dempe. Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. *Optimization*, 52:333–359, 2003.

[5] B. Fortz and M. Thorup. Increasing internet capacity using local search. *To appear in Computational Optimization and Applications*.

[6] B. Fortz and M. Thorup. Internet traffic engineering by optimizing OSPF weights. In *Proc. 19th IEEE Conf. on Computer Communications (INFOCOM)*, pages 519–528, 2000.

[7] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman, San Francisco, 1979.

[8] J. Håstad. Clique is hard to approximate within $n^{1-\varepsilon}$. *Acta Mathematica*, 182:105–142, 1999.

[9] R.G. Jeroslow. The polynomial hierarchy and a simple model for competitive analysis. *Mathematical Programming*, 32:146–164, 1985.

[10] R.M. Karp and J. B. Orlin. Parametric shortest path algorithms with an application to cyclic staffing. *Discrete Applied Mathematics*, 3:37–45, 1981.

[11] M. Labbé, P. Marcotte, and G. Savard. A bilevel model of taxation and its application to optimal highway pricing. *Management Science*, 44:1608–1622, 1998.

[12] S. Roch, G. Savard, and P. Marcotte. Design and Analysis of an approximation algorithm for Stackelberg network pricing. *Optimization Online*, 2003.

[13] T. Roughgarden and É. Tardos. How bad is selfish routing? *Journal of the Association for Computing Machines*, 49(2):236–259, 2002.

[14] S. van Hoesel, A. F. van der Kraaij, C. Mannino, G. Oriolo, and M. Bouhtou. Polynomial cases of the tarification problem. *Meteor Resarch Memorandum*, RM03053, 2003.

[15] L.N. Vicente and P. H. Calamai. Bilevel and multilevel programming: A bibliography review. *Journal of Global Optimization*, 5:291–306, 1994.

[16] N. E. Young, R. E. Tarjan, and J. B. Orlin. Faster parametric shortest path and minimum-balance algorithms. *Networks*, 21:205–221, 1991.