# Revenue Maximizing Tariff Zone Planning for Public Transport Companies

S. Müller[a], K. Haase[b], L. Reyes-Rubiano[a*]

[a] Chair of Operations Management, Otto von Guericke University Magdeburg, Magdeburg, Germany
{sven.mueller, lorena.reyes}@ovgu.de

[b] Institute for Transport Economics, University of Hamburg, Hamburg, Germany
knut.haase@uni-hamburg.de

*Extended abstract submitted for presentation at the 11th Triennial Symposium on Transportation Analysis conference (TRISTAN XI)*
*June 19-25, 2022, Mauritius Island*

March 27, 2022

---

Keywords: Public transportation, revenue management, contiguity, districting, mixed integer programming

# 1 INTRODUCTION

Public transport companies are under pressure to reduce operational costs and increase the number of passengers and revenue. The design of the tariff system has been proven valuable to impact revenues in public transportation. There exist different tariff systems in public transport (Schöbel & Urban, 2020). This paper studies the zoning and pricing problem to maximize expected revenue ($R$) through a counting zones tariff system. The literature on combining zoning and pricing problems is scarce. Thus, we contribute to this literature by a new mixed-integer programming (MIP) model and a MIP-based heuristic method to solve the revenue maximizing tariff zone problem (RMTZP), yielding a counting zones tariff system. It is assumed that (i) the price per zone takes denumerable values, (ii) the number of public transport trips depends on the price (system), and (iii) public transport passengers always choose the time-shortest path. Our new model formulation is: (1) flexible to adjust to any objective function; (2) not limited to a predefined number of tariff zones; we impose contiguity of the tariff zones using the properties of primal and dual graphs, (3) coming with a new set of constraints that ensures contiguity and forces tariff zones to a desired spatial pattern (rings or stripes) without altering the model structure; (4) able to optimally solve instances of up to 120 districts (stops) within reasonable time using off-the-shelf solvers.

# 2 Modeling

We consider a public transport graph $\mathcal{G}^{\text{PT}} : (\mathcal{I}, \mathcal{A}, \tau_{ij})$ with nodes $i \in \mathcal{I}$, arcs $(i, j) \in \mathcal{A}$, and travel time $\tau_{ij}$. The set of nodes $\mathcal{I}$ represents the public transport stops and the set of arcs $\mathcal{A}$ indicates the public transport connections between adjacent nodes (gray arcs and nodes, Figure 1). We impose contiguity of each tariff zone using the properties of primal and dual graphs (Validi *et al.*, 2020). We assume that each stop is located in a unique (artificial) district. Let $\mathcal{G}^{\text{BO}} : (\mathcal{N}, \mathcal{B})$ be the district border graph, with nodes $\mathcal{N}$ and arcs $\mathcal{B}$. Here, we consider $\mathcal{G}^{\text{BO}}$ as the dual to $\mathcal{G}^{\text{PT}}$. The problem includes a set $\mathcal{S}_{ij}$ that contains the arcs that belong to the time-shortest path through $\mathcal{G}^{\text{PT}}$ of each origin-destination (O-D) tuple from $i \in \mathcal{I}$ to $j \in \mathcal{I}$. Due to the properties of planar dual graphs, for each arc $(i, j) \in \mathcal{A}$ there exist two intersecting arcs $((n, m), (m, n)) \in \mathcal{B}$. Let us define set $\mathcal{D}_{ij}$ denoting the border arcs $(n, m)$ corresponding to

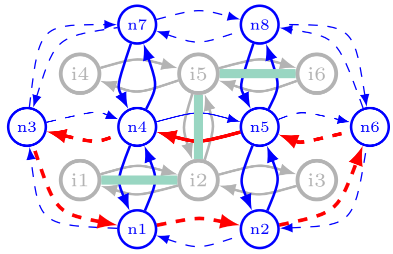


Figure 1 – *Graphs of the RMTZP. Note,* $\mathcal{G}^{BO}$ *contains more nodes than necessary for a dual graph.* $\mathcal{G}^{BO}$ *is shown in blue while* $\mathcal{G}^{PT}$ *is displayed in light gray. Shortest path from stop (district) i1 to stop (district) i6 is highlighted in green. Assume t=2 tariff zones are optimal for O-D tuple i1-i6 (*$X_{i1,i6,2}=1$*) and the tariff zone border is between node i2 and i5, then* $Y_{n4,n5}+Y_{n5,n4}=1$*. This in turn induces and artificial flow along the border arcs to ensure contiguous zones (red).*

$(i,j) \in \mathcal{A}$. Now, let us define the set $\mathcal{OD}_{ij}$ containing the border arcs along the time-shortest path from $i \in \mathcal{I}$ to $j \in \mathcal{I}$.
Following, we summarize the variables and constraints of our MIP model **P1** to solve RMTZP:

**Parameters**

$T_{ij}$ : Maximum number of tariff zones along the shortest path from $i \in \mathcal{I}$ to $j \in \mathcal{I}$ ($T_{ij} \leq | \mathcal{S}_{ij} | +1$)

$r_{ijt}(\pi_{pt})$ : Expected revenue if $t = 1, \ldots, T_{ij}$ tariff zones are visited on the shortest path from $i \in \mathcal{I}$ to $j \in \mathcal{I}$ and given price per zone $\pi_{pt}$.

$a_n$ : Feasible node degree of the border node $n \in \mathcal{N}$

$b_n$ : Amount of artificial outflow or inflow at border node $n \in \mathcal{N}$

$u$ : Total sum of outflows over all border nodes $n \in \mathcal{N}$ ($u = \sum_{n|b_n \geq 0} b_n$)

**Decision variables**

$X_{ijt}$ = 1, if $t = 1, \ldots, T_{ij}$ tariff zones are visited on the shortest path from $i \in \mathcal{I}$ to $j \in \mathcal{I}$ (0, otherwise), with $T_{ij}$ as the maximum number of tariff zones along the shortest path from $i$ to $j$

$Y_{nm}$ = 1, if a tariff zone border is established along border arc $(n,m) \in \mathcal{B}$ (0, otherwise)

$W_{nm}$ = 1, if there is a flow along the border arc $(n,m) \in \mathcal{B}$ (0, otherwise)

$$\textbf{Maximize } R(p) = \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{I}} \sum_{t=1}^{T_{ij}} r_{ijt}(\pi_{pt}) \cdot X_{ijt} \tag{1}$$

**subject to**

$$\sum_{t=1}^{T_{ij}} X_{ijt} = 1 \qquad \forall\ i \in \mathcal{I}, j \in \mathcal{I} \tag{2}$$

$$\sum_{t=1}^{T_{ij}} t \cdot X_{ijt} - \sum_{(n,m)\in\mathcal{OD}_{ij}} Y_{n,m} = 1 \qquad \forall\ i \in \mathcal{I}, j \in \mathcal{I} \tag{3}$$

$$W_{nm} - W_{mn} = Y_{nm} \qquad \forall\ (n,m) \in \mathcal{B} \mid n < m \tag{4}$$

$$\sum_{m\in\mathcal{B}} W_{nm} - \sum_{m\in\mathcal{B}} W_{mn} = b_n \qquad \forall\ n \in \mathcal{N} \tag{5}$$

$$\sum_{m\in\mathcal{B}} W_{nm} + \sum_{m\in\mathcal{B}} W_{mn} \leq a_n \qquad \forall\ n \in \mathcal{N} \tag{6}$$

$$X_{ijt} \in \{0,1\} \qquad \forall\ i \in \mathcal{I}, j \in \mathcal{I}, t = 1, ..., T_{ij} \tag{7}$$

$$Y_{nm} \in \{0,1\} \qquad \forall\ (n,m) \in \mathcal{B} \tag{8}$$

$$W_{nm} \in \{0,1\} \qquad \forall\ (n,m) \in \mathcal{B} \tag{9}$$

The objective (1) maximizes the total expected revenue in the service area for a given price system $p \in \mathcal{P}$. Constraints (2) select the number of tariff zones visited for each O-D tuple. Constraints (3) couples the number of tariff zone $t \cdot X_{ijt}$ and the number of tariff zone borders $Y_{nm}$ along the shortest path from $i \in \mathcal{I}$ to $j \in \mathcal{I}$. The number of tariff zones visited from stop $i$ to stop $j$ is equal to 1 plus the number of tariff borders crossed along the shortest path from $i$ to $j$. Equations (4) couple of $Y_{nm}$ and $W_{nm}$. If a tariff zone border is established between adjacent stops $i$ and $j$, i.e., $Y_{nm}$ =1 with $(n, m) \in \mathcal{D}_{ij}$, then there is a flow $W_{nm} \in \{0, 1\}$ along $(n, m) \in \mathcal{B}$. Equations (5) are flow constraints that indicate outflow minus inflow must be equal to $b_n$ at border node $n \in \mathcal{N}$. Values of $b_n$ depend on the desired spatial pattern of the tariff zone. The ring pattern occurs when $b_n$=0 and $a_n$=2, and a striped pattern $b_n \geq 1$ and $a_n$=2. Additionally, the stripe pattern is enforced by establishing a set of nodes $n \in \mathcal{N}$ as the source nodes of the artificial flow and another set of nodes $n \in \mathcal{N}$ as the sink nodes for the artificial flow. The flow conservation constraints (5) ensure contiguous tariff zones. Constraints (6) control the number of adjacent tariff zones at node $n$. Finally, the domains of the variables are given by (7), (7), and (9).

Relaxation of RMTZP: The size of the RMTZP is mainly influenced by the number of O-D tuples. The expected $R$ for each O-D tuple can be computed given a pricing system and a maximum number of tariff zones. We declare a subset $\mathcal{C}$ which contains $\gamma\cdot \mid \mathcal{I} \mid$ of O-D tuples with highest $R$; where $\gamma \geq 1$ and determines the size of the subset $\mathcal{C}$. We propose a MIP-based heuristic to find $\mathcal{C}$. Then, the model **P1** is solved by considering O-D tuples $\in \mathcal{C}$ instead of all $i \in \mathcal{I}$ and $j \in \mathcal{I}$.

**Price problem:** The zone problem **P1** depends on a given price system $p \in \mathcal{P}$. The trip price under a counting zone tariff system depends on the visited zones along the trip and the price per zone. Let $r_{ijt}(\pi_{p_t})$ represent the $R$ on the shortest path from $i \in \mathcal{I}$ to $j \in \mathcal{I}$ given price system $p \in \mathcal{P}$, with $\pi_{p_t}$ as the price per zone when visiting $t$ zones under price system $p_t \in \mathcal{P}$. $R$ in **P1** is given by Equation (1). Therefore, we solve $\mid \mathcal{P} \mid$ independent zoning problems, select the solution with maximum $R$, and benchmark these solutions against the optimal solutions of the RMTZP.

# 3 Computational experiments

In this section, we present the design of computational experiments and results. We evaluate the performance of the MIP model and our MIP-based heuristic under different problem sets. The performance is evaluated in terms of $R$ and CPU time. We generate a set of artificial instances representing the service area of a city. Total demand depends on the number of inhabitants at each zone $i$, defined as a uniform [10,000; 20,000]. The total number of trips from $i$ to $j$ follows a gravity model, while public transport shares follow a multinomial logit model.

## 3.1 Comparison of the MIP-based heuristic against the MIP model

We solve all problem sets to optimality and compare the MIP model against the MIP-based heuristic, considering the constraints to enforce tariff zones to have a ring pattern. Problem sets consider different realistic sizes of $\mid \mathcal{I} \mid= \{49, 81, 121\}$, $\mid T_{ij} \mid$=7, and $\mid \mathcal{P} \mid$=10, and a network connectivity level given by {0.25, 0.5, 0.75, 1}. The larger the network connectivity level the larger is $\mid \mathcal{A} \mid$ for given $\mid \mathcal{I} \mid$.

Table 1 shows $R$ of the solution obtained with MIP model and MIP-based heuristic over 10 seeds, and $\gamma$ over the values { 2, 4, 8, 10, 15, 20, 25, 30}. Our first results show that the MIP model provides a solution with higher $R$ than the MIP-based heuristic solution. On average, our MIP-based heuristic under estimates the optimal $R$ by 2.54% but it is faster by 42.97%.

## 3.2 Impact of enforcing a desired spatial pattern

This section is devoted to determine the impact of enforced spatial patterns on $R$. This numerical study is focused on the MIP model **P1** and the instance with $\mid \mathcal{I} \mid= 49$. We determine the solution

| | MIP | | MIP-based heuristic | | | |
|---|---|---|---|---|---|---|
| $\mathcal{I}$ | **Expected revenue 10000 monetary units (a)⋆** | **CPU time seconds (b)⋆** | **Expected revenue 10000 monetary units (c)⋆** | **CPU time seconds (d)** | **Gap (a) vs (c)** | **Gap (b) vs (d)** |
| 49 | 1855.93 | 104.41 | 1846.85 | 117.15 | 0.48 | -16.05 |
| 81 | 1939.49 | 4087.39 | 1899.39 | 570.80 | 2.06 | 87.47 |
| 121 | 2012.58 | 386384.54 | 1911.02 | 175053.97 | 5.08 | 57.48 |
| | | | **Average** | | **2.54** | **42.97** |

Table 1 – *Comparison of R and CPU time of the MIP-based heuristic with respect to the optimal solution. All problems are solved to optimality. (⋆) Average values over all seeds and network connectivity levels*

| | Ring pattern | | No pattern | |
|---|---|---|---|---|
| $a_n$ | $R$ **(10000 monetary units)** | **CPU time** (seconds) | $R$ **(10000 monetary units)** | **CPU time (seconds)** |
| 2 | 1965.36 | 125.76 | 1977.92 | 10559.98 |
| 3 | 1965.36 | 1639.02 | 1977.92 | 9051.35 |
| 4 | 1989.42 | 3308.62 | 1992.07 | 12980.14 |

Table 2 – *Average R of the MIP model varying: constraints to enforce a ring pattern, and values of* $a_n$.

for different values of border node degree $a_n$ given by $\{2, 3, 4\}$. The value of $a_n$ indicates the maximum number of adjacent tariff zones that gather at a district border node $n \in \mathcal{N}$. Table 2 shows the average $R$ and CPU times of solutions with and without a ring pattern. Results demonstrate that solutions with a lower $R$ are obtained when constraints are included to enforce a desired spatial pattern compared to cases with no desired pattern. Solutions without any spatial pattern have an $R$ that is on average 0.43% higher than solutions with a ring pattern. The CPU time for the solutions with no spatial pattern is 85.12% higher than for solutions with the ring pattern.

# 4 Conclusion

In this paper, we investigate how to design a counting zones tariff system to maximize the revenues of public transportation service companies. The price for a trip depends on the price per tariff zone and the number of visited zones. This approach is well-known and accepted by passengers and practitioners. We design an MIP model and an MIP-based heuristic to design an optimal counting zones tariff system and solve the price problem. Our approach is based on the properties of dual and primal graphs enabling contiguity of tariff zones and enforce spatial patterns of the zones. The proposed methods can solve instances of reasonable sizes. Our approach is flexible enough to enforce the counting zones tariff system to any spatial pattern. The results show that enforcing tariff zones to a specif spatial pattern reduces $R$ and CPU time.

Currently, we are working on testing our MIP-based heuristic with real data from the San Francisco Bay Area with 1,415 districts (i.e., stops). Due to the irregular shape of the San Francisco Bay Area, we are interested in studying how to enforce tariff zones to follow a desired spatial pattern. In addition, we plan to analyze price systems with discounts.

# References

Schöbel, Anita, & Urban, Reena. 2020. Cheapest Paths in Public Transport: Properties and Algorithms. *In: 20th Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2020)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

Validi, Hamidreza, Buchanan, Austin, & Lykhovyd, Eugene. 2020. Imposing contiguity constraints in political districting models. *Preprint, http://www. optimization-online. org/DB_HTML/2020/01/7582. html.*