Research Article

A Tree-Based Model for Setting Optimal Train Fare Zones

J. Alberto Conejero,[1] Cristina Jordán,[2] and Esther Sanabria-Codesal[3]

[1] *Instituto Universitario de Matemática Pura y Aplicada, Universitat Politècnica de València, 46022 València, Spain*
[2] *Instituto Universitario de Matemática Multidisciplinar, Universitat Politècnica de València, 46022 València, Spain*
[3] *Departamento de Matemática Aplicada, Universitat Politècnica de València, 46022 València, Spain*

Correspondence should be addressed to J. Alberto Conejero; aconejero@upv.es

Given a railway line with n stops and the number of travelers between each pair of stops, we show how to split these stops into k different fare zones in order to maximize the benefit obtained from the sale of tickets to the travelers. We present a method to obtain this solution that is based on finding the longest path in a weighted root tree. This method improves in terms of efficiency the combinatorial method, where all the possible distributions have to be considered for deciding which is the optimal one.

1. Introduction

In regional or local transportation networks a fare system where the stops are grouped into different zones is usually applied. This is the case for subway, bus, and regional railway transportation networks. The price of a ticket between two stops depends on the number of zones that a passenger must traverse during the travel. The stops and the number of zones are determined by the company. The assignment of the stops to certain zones usually depends on their distance to the center of the network. This is usually an arbitrary decision and it could attend to different reasons. The maximization of benefits can be addressed either to social welfare or to profit. In this last case, an analysis of the flow of travelers between every single pair of stops is required. One may wonder whether a change of fares in order to increase the income would result in a decrease of the demand; however this depends on several factors such as the type of area, the trip purposes, the distance of the journey, and the ticket types, among others. Moreover, one also has to distinguish between short, medium, and long-term elasticities [1]. In our work, we will consider that the number of travelers between each pair of stops is constant and independent of the change of the fare prices.

1.1. Related Work. In 1980 Webster and Bly published a collaborative report on [2] identifying many factors which influence demand of travelers. This has been of great value to public transport operators and transport planners. This has been recently updated in [1]. This same year Cervello analyzed the effect of flat rates versus differentiated ones [3]. In particular, he studied the impact of the variability of fares depending on the distance and time-of-day. Moreover, the differentiated fares can be further subdivided into zonal fares, distance-based fares, sectional fares, and time-based fares [4, 5]. Zonal-based fares are the most common fare systems in regional railway and underground transportation networks in Spain.

Several studies have been conducted in the last thirty years in order to study fare optimization; see, for instance, [6–9]. The objective functions in the above studies were to maximize either total profit and/or social welfare. In some of the cases elasticities in the demand of the transportation service were included. The solution of these problems was given by the maximization of certain profit function subject to certain restrictions. Sometimes, the maximization can be split into two levels using a bilevel model approach, with the operator model at the upper level and the user at the lower one; see [10, 11]. For further information we refer the reader to [12, Section 2.2].

The problem of fare zone design for local public transportation networks using a graph theoretical approach was considered by Hamacher and Schöbel in [13]; see also [14, 15]. In these works, the design is obtained in order to minimize either the maximal or average deviation between the fare zone system and the distance tariff for all the pairs of stops. Some heuristics algorithms were provided in this line in [15].

In this work we present an algorithm to find the optimal fare zone system for a line of transport, where the stops are located one after the other. Our model is based on finding the longest path between the root and one of the leaves on a certain weighted directed tree with root. Weighted graphs often appear in the modeling of problems of logistics [16], scheduling and transport routing [17], and management in the air industry [18, 19] and of car-rental companies [20]; see also [21].

Apart from this problem, the use of discrete mathematics in railway transportation problems is not new and has been considered for different problems [22–24]. This is done, for instance, in the *rolling stock planning* phase [25, 26] and in its integration with the *timetabling phase* [27–29], or in the assignment tracks to the trains [30, 31]. In particular, graph models have been used for choosing a fare planning model that maximizes the revenue [32, 33].

1.2. Preliminaries on Weighted Graphs and Trees. We recall that a directed weighted graph is a 3-tuple $G = (V, E, w)$ where V is the set of *nodes*, E is the set of *arcs* (ordered pairs of nodes), and $w : E \rightarrow \mathbb{N}_0$ assigns to every arc $(u, v) \in E$, with $u, v \in V$, a positive integer amount (or zero) denoted by $w(u, v)$, or just as $w_{u,v}$. A *path* on a graph is defined as an alternate sequence of nonrepeated nodes and arcs of the form $u_1, (u_1, u_2), u_2, (u_2, u_3), u_3, \ldots, u_{n-1}, (u_{n-1}, u_n), u_n$ for some nodes $u_1, u_2, \ldots, u_n \in V$ and $(u_1, u_2), (u_2, u_3), \ldots, (u_{n-1}, u_n) \in E$. A directed tree is a directed graph which would be a tree if the directions on the arcs were ignored and all the arcs are directed away from a particular node, which is called the root of the tree. A nondirected graph is a tree if it is connected (every pair of nodes are joined by a path) and it verifies that $|E| = |V| - 1$.

The solution of the longest path problem on a graph $G = (V, E, w)$ consists of finding the path between a certain pair of nodes $u, v \in V$ whose sum of the weights of the arcs is maximized. Despite the fact that this is in general an NP-hard problem, its solution can be computed in a linear time on directed acyclic graphs. For further information on graphs and trees, we refer the reader to [21, 34–36].

1.3. Organization of the Paper. The paper is organized as follows: in Section 2 we introduce the notation and we calculate the total number of admissible solutions to the problem using a combinatorial approach. Section 3 is devoted to the complete analysis of the case of 2 zones. A general algorithm for the construction of a directed rooted tree for an arbitrary number of stops and zones is provided in Section 4. Then the solution to our problem is reduced to find the longest path from the root to one of the leaves in this tree. In Section 5 we give details about the size of the tree constructed

in Section 4 and how to label their nodes. This will help us to study the computational cost of this problem that is shown in Section 6. We conclude this section with an example of application.

2. Notation

Our algorithm for finding the optimal solution considers partial admissible solutions of the problem and compares among themselves in order to choose the optimal one. For dealing with this partial distributions and with all the admissible solutions to the problem, we introduce a special notation for referring to them. Besides, as we want to get the optimal revenue for the company that supplies the service, then the number of passengers between every pair of stops must be taken into account.

Firstly, let us consider a railway line where we know the location of n stops. For determining a fare system of k zones, we assume that there is at least one stop at every zone; therefore $n \geq k$. The case $n = k$ has a trivial solution since we have a unique stop at every zone so that the cases of interest will be the ones with $n > k$.

Secondly, let us consider n stops, namely, $\{s_1, s_2, \ldots, s_n\}$. These stops are assumed to be ordered in the railway line; namely, $s_1 \prec s_2 \prec s_3, \ldots, \prec s_n$, with s_1 being the first one and s_n the last one. We will also assume that every single stop s_i is located at only one of the k zones, and we denote by z_i its zone, with $1 \leq i \leq n$ and $1 \leq z_i \leq k$.

We assume the following statements about the distribution of stops into zones. The first one indicates that two consecutive stops in the line must be in the same zone or in contiguous zones.

(A1) For every $1 \leq i \leq n-1$, either $z_{i+1} = z_i$ or $z_{i+1} = z_i + 1$.

Among all possible distributions generated under assumption (A1), only some of them will be admissible distributions of stops into zones. They will be the ones verifying assumptions (A2) and (A3).

(A2) Consider $z_1 = 1$ and $z_n = k$; that is, the first and the last city of the line are located at the first and at the last zone, respectively.

(A3) For every $1 < i < k$ there exists some $1 < j < n$ such that $z_j = i$; that is, we have at least one stop at every zone.

We introduce a notation for referring to particular distributions of n stops into k zones. We also consider partial distributions that deal with $n' < n$ stops distributed into k zones.

Definition 1. We define a partial distribution of n' stops into k zones as a k-tuple of the form $[y_1, y_2, \ldots, y_k]$ where $y_1 + y_2 + \cdots + y_k \leq n$ with $y_i \geq 0$ for all $1 \leq i \leq k$, and such that if there exists $y_{i_0} = 0$, then $y_i = 0$ for $i \geq i_0$.

The k-tuple $[y_1, y_2, \ldots, y_k]$ indicates that stops $s_1, \ldots, s_{y_1}$ are in zone 1, stops $s_{y_1+1}, \ldots, s_{y_1+y_2}$ are in zone 2, and so on, until we have that stops $s_{y_1+\cdots+y_{k-1}+1}, \ldots, s_{y_1+\cdots+y_k}$ are located in zone k.

If we distribute all the n stops we refer to this distribution as an admissible one.

Definition 2. We define an admissible distribution of n stops into k zones as a k-tuple of the form $[y_1, y_2, \ldots, y_k]$ where $y_1 + y_2 + \cdots + y_k = n$ and $y_i \geq 1$ for all $1 \leq i \leq k$.

The total number of admissible distributions can be computed using a combinatorial approach.

Proposition 3. *Let n be the number of stops to be distributed into k zones. The total number of admissible distributions of this n stops into k zones is $(n-1)!/(n-k)!(k-1)!$.*

Proof. By assumption (A2), $z_1 = 1$ and $z_n = k$. The other $n-2$ stops must be distributed into the k zones. By assumption (A3) we have to assign at least one stop of these $n-2$ to each zone i for $2 \leq i \leq k-1$. Let us count how many admissible distributions can we find rearranging the others $n-2$ stops.

Let us consider two types of symbols, $\mathcal{S}$ and $\mathcal{C}$, where $\mathcal{S}$ stands for a stop and $\mathcal{C}$ for a change of zone. A string of n symbols $\mathcal{S}$ and $k-1$ symbols $\mathcal{C}$ represents an admissible distribution of n stops into k zones whenever there is at least

  (i) one $\mathcal{S}$ before the first $\mathcal{C}$,

  (ii) one $\mathcal{S}$ after the last $\mathcal{C}$,

  (iii) one $\mathcal{S}$ between every pair of symbols $\mathcal{C}$.

If we assume the necessity of having at least one symbol $\mathcal{S}$ at every zone, then we can rephrase the problem to calculate how many strings can be constructed with $n-k$ symbols $\mathcal{S}$ and $k-1$ symbols $\mathcal{C}$. For instance, a string of the form

$$\underbrace{\mathcal{S}\mathcal{S}\cdots\mathcal{S}}_{i_1}\mathcal{C}\underbrace{\mathcal{S}\mathcal{S}\cdots\mathcal{S}}_{i_2}\mathcal{C}\cdots\mathcal{C}\underbrace{\mathcal{S}\mathcal{S}\cdots\mathcal{S}}_{i_k}, \tag{1}$$

with $i_1, \ldots, i_k \geq 0$ and $i_1 + i_2 + \cdots + i_k = n-k$, corresponds to an admissible distribution $[i_1 + 1, i_2 + 1, \ldots, i_k + 1]$. The total number of strings constructed with two different elements $\mathcal{S}$ and $\mathcal{C}$, with $\mathcal{S}$ appearing $n-k$ times and $\mathcal{C}$ appearing $k-1$ times, is given by

$$P_{n-1}^{n-k,k-1} = \frac{(n-1)!}{(n-k)!\,(k-1)!}, \tag{2}$$

which is the number of permutations with repetition of $n-1$ elements with $n-k$ of one type and $k-1$ of the other type. For n much more greater than k, this is of order $\mathcal{O}(n^{k-1})$. □

Given an admissible distribution $[y_1, y_2, \ldots, y_k]$ we can compute the benefit obtained from the passengers tickets. For every pair of stops s_i, s_j with $1 \leq i, j \leq n, i \neq j$, we denote by $p_{i,j}$ the number of passengers that depart from s_i and arrive at s_j on a certain period of time. A passenger with a return ticket between two stops is counted twice, one on each sense. We assume that these passengers pay two times the price of a single ticket and no discount is granted to them.

Once we have an estimation of the number of passengers between every single pair of stops, we can compute which will be the benefit obtained by the sale of tickets after applying a certain assignment of stops to the zones. As we have indicated in the Introduction, we are assuming that a change of the location of the stops in the zones does not modify the number of travelers between every pair of stops.

We consider a fare system for the sale of tickets based on the number of zones. The price of a ticket depends on how many zones a passenger crosses in his/her trip, that is, a counting zone tariff system. If a passenger remains at the same zone, then the fare is h_0. If it is required to visit 2 zones, that is, the initial and final stops are at contiguous zones, then the price would be $h_0 + h_1$. When a passenger visits 3 zones, then the cost of the ticket would be $h_0 + h_2$, and so on. For simplicity, when we compute the benefit coming from a certain distribution of stops into zones we omit the fixed value h_0 since it is applied to all the tickets. For maximizing the benefit, the idea is to find a distribution that globally forces the passengers to traverse the highest number of zones, taking into account the aforementioned restrictions set on the distribution of stops into zones. This is the underlying idea on which the following algorithms are based.

3. The Case of 2 Zones

The simplest case of just a single zone is not a problem for any number n of stops. In this section we solve the problem of findig a distribution of n stops into 2 zones in order that the revenue obtained from the passengers tickets is maximized. First, we introduce the trees that model this problem in terms of $n \in \mathbb{N}$. Later, we will show how to compute the optimal solution using the structure of this tree.

Some figures will help us to show the construction of the trees used to model different statements of the problem. A dashed line will be used for separating admissible distributions (nodes) with the last stop in different zones. Besides, a node with a thick line around it will be used to indicate the admissible distributions at every step, that is, the ones verifying assumptions (A1), (A2), and (A3).

3.1. Definition of the Tree for n Stops and 2 Zones, $n \geq 3$. We start at the first stop s_1 that must be located at zone 1. This is represented by the partial admissible distribution $[1, 0]$. Then we can add one more stop, s_2. By assumption (A1), s_2 can be assigned either to zone 1 or to zone 2. The first case corresponds to the partial distribution $[2, 0]$ and the second one to the unique admissible distribution $[1, 1]$. A pair of arcs departing from $[1, 0]$ and arriving at $[2, 0]$ and $[1, 1]$ shows that these distributions are generated from $[1, 0]$; see Figure 1.

A third stop, s_3 can be added. Taking into account (A1) we have two options. On the one hand, the distribution $[2, 0]$ lets us generate two new distributions: $[3, 0]$, if we add the third stop to zone 1, and $[2, 1]$, if we add it to zone 2. On the other hand, from distribution $[1, 1]$ we can only add s_3 to the second zone since there is no third zone in this case. This enables us to consider the distribution $[1, 2]$, too. However, the unique admissible distributions with 3 stops will be $[2, 1]$ and $[1, 2]$, and not $[3, 0]$. In Figure 2 we have represented this case.

We can proceed by induction using (A1) to construct a tree for each number n of stops. Suppose that we have done
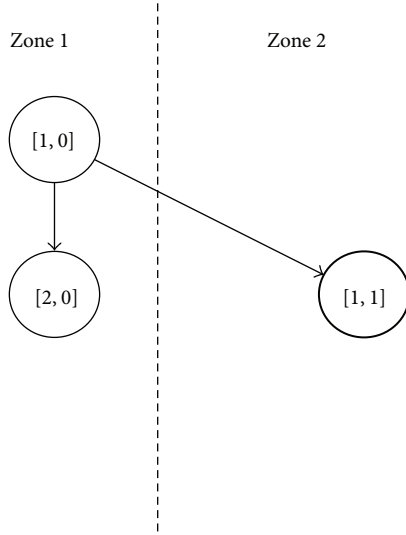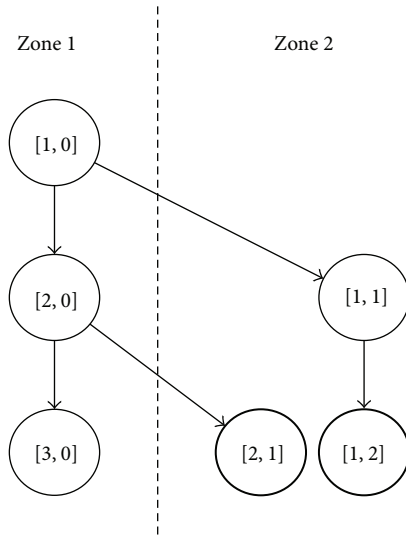
FIGURE 1: Tree for 2 zones and 2 stops.

FIGURE 2: Tree for 2 zones and 3 stops.

the construction up to $n - 1$ stops. We will see how to get the partial distributions for n stops. There are two kinds of nodes representing a distribution of $n - 1$ stops attending to its form as follows:

(i) $[n - 1, 0]$,

(ii) the nodes of the form $[i, j]$ with $1 \leq i, j \leq n$ and $i + j = n - 1$.

On the one hand, from $[n-1, 0]$ we can construct two new distributions $[n, 0]$ and $[n - 1, 1]$. On the other hand, from a node $[i, j]$ of the second form we can only generate one more distribution, $[i, j + 1]$; see Figure 3.

With this procedure we construct a tree with a root at $[1, 0]$; see Figure 3. Fixing a value $n \in \mathbb{N}$, the nodes $[i, j]$ with $i + j = n$ will be the leaves of the tree. In particular,

every leaf, except $[n, 0]$, corresponds to one of the admissible distributions of n stops into 2 zones, bijectively.

We point out that in our representations of the trees generated under this procedure, all nodes with the same value for $i + j$ are displayed horizontally as we can also see in Figure 3. Moreover, the nodes of the form $[i, j]$ with $j = 0$ are set to the left of the dashed line, since they correspond to partial distributions that uniquely have stops in zone 1. The rest of the nodes are set to the right of the dashed line since they have at least one stop in zone 2.

3.2. Optimal Solution for n Stops and 2 Zones. For finding the optimal arrangement of n stops into 2 zones, we assign weights to all arcs in order that if we take an admissible distribution $[i, j]$ with $i + j = n$ and $j \neq 0$, then the sum of the weights of the arcs in the path that connects $[1, 0]$ with $[i, j]$ will represent the total benefit obtained from the passengers that have to change of zone during their travel.

Theorem 4. *Let T be a tree constructed under assumptions (A1), (A2), and (A3), for representing the admissible distributions of n stops into 2 zones. Let us assume that the number of passengers between every pair of stops $(p_{i,j})_{1 \leq i, j \leq n}$ is given.*

Then there is an assignment of weights to the edges in order that the sum of the weights of the longest path from the root to $[1, 0]$ to a leaf of the form $[i, j]$ with $1 \leq i, j \leq n$ and $i + j = n$ returns the maximal benefit for distributing n stops into 2 zones.

Proof. Take a tree T constructed under assumptions (A1), (A2), and (A3), for representing the admissible distributions of n stops into 2 zones. Fix i, j with $i + j < n$; then an arc connecting a node representing the partial distribution $[i, j]$ with a node $[i + 1, j]$ or $[i, j + 1]$ represents that we added one stop to the partial admissible distribution $[i, j]$, namely, s_{i+j+1}.

The weight that we assign to this arc will represent the increment of benefit due to the changes of zone of the passengers that travel between each one of the stops $\{s_1, \ldots, s_i\}$ and s_{i+j+1}. The following formula gives the value of the weights for all the arcs in the tree:

$$
w(u, v) = \begin{cases}
0 & \text{for } u = [i, 0], \\
& v = [i + 1, 0], \\
& \text{with } 1 \leq i \leq n - 1, \\
h_1 \sum_{\ell=1}^{i} p_{\ell, i+1} + p_{i+1, \ell} & \text{for } u = [i, 0], \\
& v = [i, 1], \\
& \text{with } 1 \leq i \leq n - 1, \\
h_1 \sum_{\ell=1}^{i} p_{\ell, i+j+1} + p_{i+j+1, \ell} & \text{for } u = [i, j], \\
& v = [i, j + 1] \text{ with} \\
& 1 \leq i, j \leq n - 1, \\
& 2 \leq i + j \leq n - 1.
\end{cases}
\tag{3}
$$

In this way, the optimal solution is given by the weight of the longest path among all the paths from the root $[1, 0]$ to
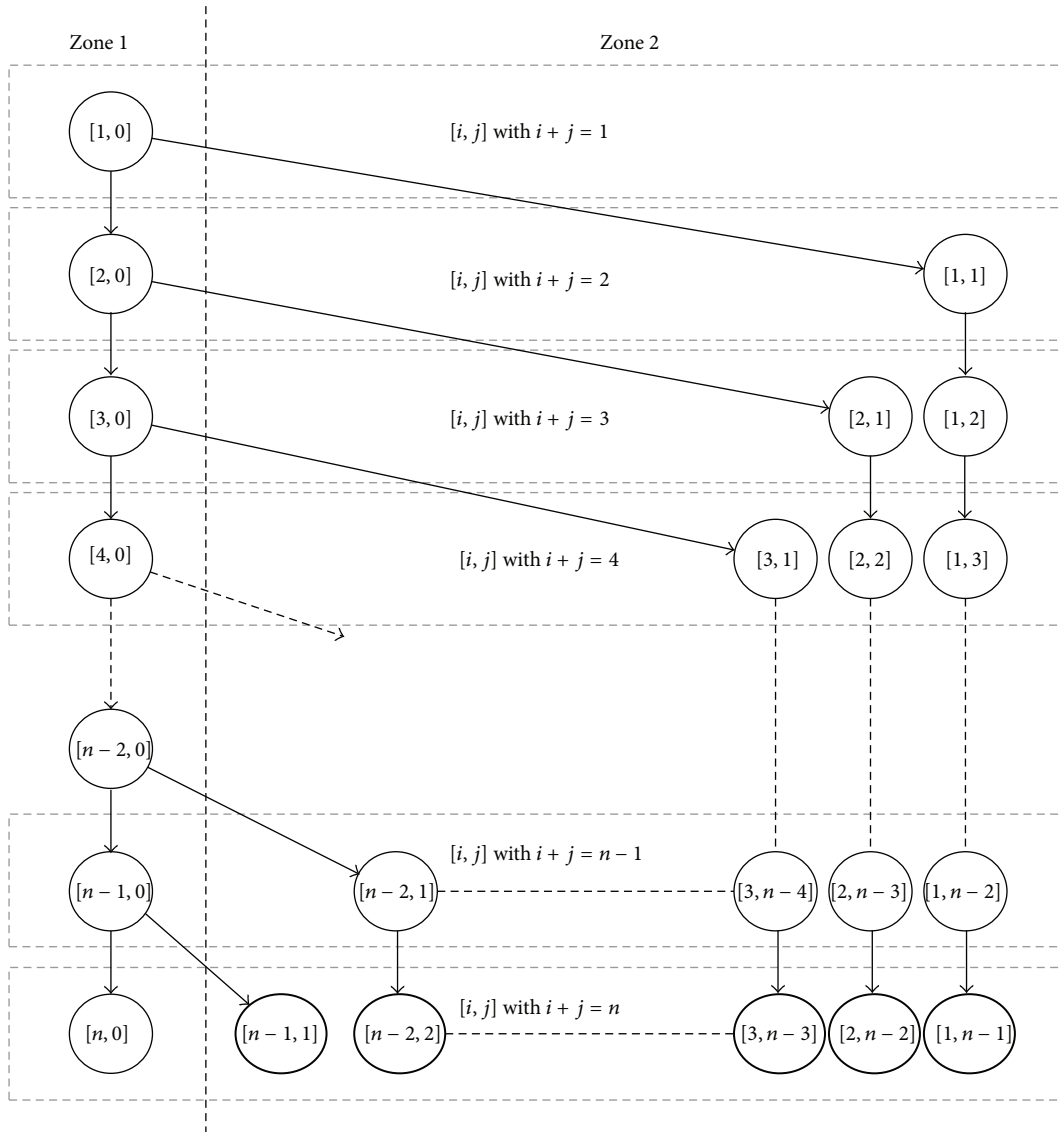
FIGURE 3: Tree for 2 zones and n stops.

any vertex $[i, j]$ with $i + j = n$. By properties of trees, once determining $[i, j]$ such a path is unique. □

4. The General Case of k Zones

In this section we will see how to construct a tree for the case of n stops and k zones, with $k \geq 2$ and $n \geq 3$. The algorithm is mainly based on assumption (A1), which permits inductively defining the nodes and the arcs as we have seen in Section 3.

From Step 1 up to Step $k - 1$, at the substeps $q.1$ and $q.2$ for $1 \leq q \leq k - 1$, we define two new nodes from a node $[i_1, i_2, \ldots, i_q, 0, \ldots]$, which are $[i_1, i_2, \ldots, i_q + 1, \ldots]$ and $[i_1, i_2, \ldots, i_q, 1, \ldots]$. We also define two arcs connecting node $[i_1, i_2, \ldots, i_q, \ldots]$ with each of them and assign its weight using formulas (4), (5), (6), (7), and (8). When a node is of the form $[i_1, i_2, \ldots, i_k]$ with $i_k \neq 0$, then we connect this node with a single node $[i_1, i_2, \ldots, i_k + 1]$ and assign a weight to this arc using formula (9).

An example of a tree for 5 stops and 3 zones is represented in Figure 4.

4.1. Definition of the Tree for n Stops and k Zones with $n > k \geq 2$

Step 0. Consider the following.

(0.1) Define $i_0 = 0$. (*This index will be used from Step 3 onwards*).

(0.2) Start a tree with a unique node $[\underbrace{1, 0, 0, \ldots, 0}_{k}]$. This will be its root.

(0.3) Define a list L. We initialize the list with the node $[1, 0, 0, \ldots, 0]$.

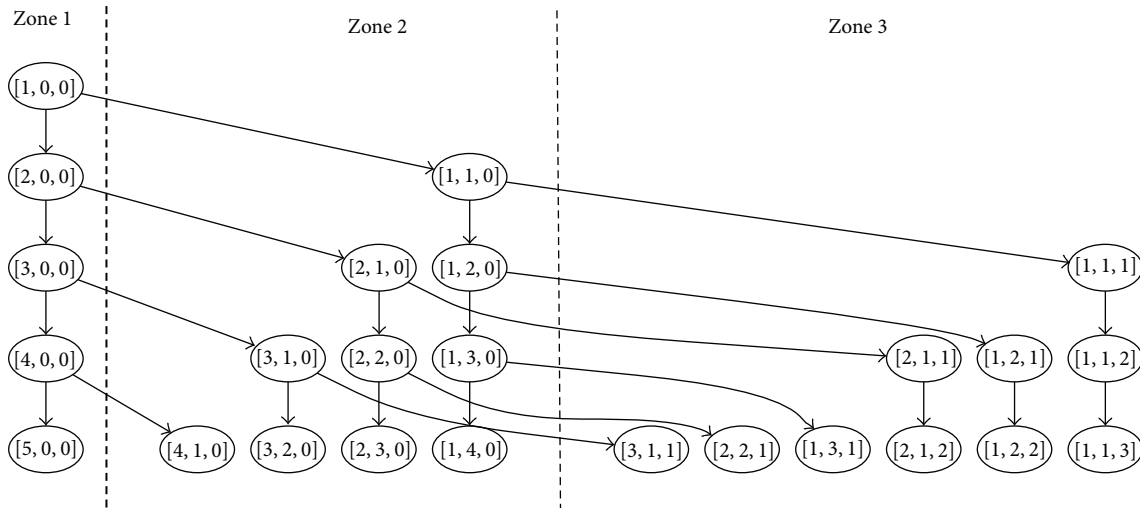*Step 1.* We take the first node in the list L, namely, $[i_1, i_2, i_3, \ldots, i_k]$.

FIGURE 4: Tree for 3 zones and 5 stops.

If $i_j = 0$ for $2 \leq j \leq k$ then consider the following.

(1.1) If $i_1 \leq n - k + 1$ then

    (i) define two new nodes $[i_1 + 1, 0, 0, \ldots, 0]$ and $[i_1, 1, 0, 0, \ldots, 0]$ that are added to L;

    (ii) define the arcs $([i_1, 0, 0, \ldots, 0], [i_1 + 1, 0, 0, \ldots, 0])$ with weight 0 and $([i_1, 0, 0, \ldots, 0], [i_1, 1, 0, \ldots, 0])$ with weight

$$h_1\left(\sum_{\ell=1}^{i_1} p_{\ell, i_1+1} + p_{i_1+1, \ell}\right); \tag{4}$$

    (iii) remove $[i_1, 0, 0, \ldots, 0]$ from the list L and return to the beginning of this step.

(1.2) Else if $n - k + 1 < i_1 \leq n - 1$ then

    (i) define two new nodes $[i_1 + 1, 0, 0, \ldots, 0]$ and $[i_1, 1, 0, 0, \ldots, 0]$ that are added to L;

    (ii) define the arcs $([i_1, 0, 0, \ldots, 0], [i_1 + 1, 0, 0, \ldots, 0])$ and $([i_1, 0, 0, \ldots, 0], [i_1, 1, 0, \ldots, 0])$ with weight 0;

    (iii) remove $[i_1, 0, 0, \ldots, 0]$ from the list L and return to the beginning of this step.

(1.3) Else if $i_1 = n$ then

    (i) remove $[i_1, 0, 0, \ldots, 0]$ from the list L and return to the beginning of this step.

Step 2. If $i_2 \neq 0$ but $i_j = 0$ for $j = 3, \ldots, k$ then consider the following.

(2.1) If $i_1 + i_2 \leq n - k + 2$ then

    (i) define two new nodes $[i_1, i_2 + 1, 0, \ldots, 0]$ and $[i_1, i_2, 1, 0, \ldots, 0]$ that are added to L;

    (ii) define the arc $([i_1, i_2, 0, \ldots, 0], [i_1, i_2 + 1, 0, \ldots, 0])$, with weight

$$h_1\left(\sum_{\ell=1}^{i_1} p_{\ell, i_1+i_2+1} + p_{i_1+i_2+1, \ell}\right), \tag{5}$$

    and the arc $([i_1, i_2, 0, \ldots, 0], [i_1, i_2, 1, 0, \ldots, 0])$ with weight

$$h_2\left(\sum_{\ell=1}^{i_1} p_{\ell, i_1+i_2+1} + p_{i_1+i_2+1, \ell}\right) + h_1\left(\sum_{\ell=i_1+1}^{i_1+i_2} p_{\ell, i_1+i_2+1} + p_{i_1+i_2+1, \ell}\right); \tag{6}$$

    (iii) remove $[i_1, i_2, 0, \ldots, 0]$ from the list L and return to Step 1.

(2.2) If $n - k + 2 < i_1 + i_2 \leq n - 1$ then

    (i) define two new nodes $[i_1, i_2 + 1, 0, 0, \ldots, 0]$ and $[i_1, i_2, 1, \ldots, 0]$ that are added to L;

    (ii) define the arcs $([i_1, i_2, 0, \ldots, 0], [i_1, i_2 + 1, 0, \ldots, 0])$ and $([i_1, i_2, 0, \ldots, 0], [i_1, i_2, 1, \ldots, 0])$, both of them with weight 0;

    (iii) remove $[i_1, i_2, 0, \ldots, 0]$ from the list L and return to Step 2.

(2.3) If $i_1 + i_2 = n$ then

    (i) remove $[i_1, i_2, 0, \ldots, 0]$ from the list L and return to Step 2.

(*For $q = 3, \ldots, k - 1$ the steps are defined following the next pattern:*)

Step q. If $i_q \neq 0$ and $i_j = 0$ for $j = q + 1, \ldots, k$ then consider the following.

(q.1) If $i_1 + i_2 + \cdots + i_q \leq n - k + q$ then

  (i) define two new nodes $[i_1, i_2, \ldots, i_q + 1, 0, \ldots, 0]$, $[i_1, i_2, \ldots, i_q, 1, \ldots, 0]$ that are added to L;
  (ii) define the arc $([i_1, i_2, \ldots, i_q, 0, \ldots, 0], [i_1, i_2, \ldots, i_q + 1, 0, \ldots, 0])$, with weight

$$\sum_{r=1}^{q-1} h_{q-r} \left(\sum_{\ell = i_0 + i_1 + \cdots + i_{r-1} + 1}^{i_0 + i_1 + \cdots + i_r} p_{\ell, i_1 + i_2 + \cdots + i_q + 1} + p_{i_1 + i_2 + \cdots + i_q + 1, \ell} \right), \quad (7)$$

  and the arc $([i_1, i_2, \ldots, i_q, 0, \ldots, 0], [i_1, i_2, \ldots, i_q, 1, 0, \ldots, 0])$ with weight

$$\sum_{r=1}^{q} h_{q-r+1} \left(\sum_{\ell = i_0 + \cdots + i_{r-1} + 1}^{i_0 + \cdots + i_r} p_{\ell, i_1 + i_2 + \cdots + i_q + 1} + p_{i_1 + i_2 + \cdots + i_q + 1, \ell} \right); \quad (8)$$

  (iii) remove $[i_1, i_2, \ldots, i_q, 0, \ldots, 0]$ from the list L and return to Step 1.

(q.2) If $n - k + q < i_1 + i_2 + \cdots + i_j \leq n - 1$ then

  (i) define two new nodes $[i_1, i_2, \ldots, i_q + 1, 0, \ldots, 0]$ and $[i_1, i_2, \ldots, i_q, 1, \ldots, 0]$ that are added to L;
  (ii) define the arcs $([i_1, i_2, \ldots, i_q, 0, \ldots, 0], [i_1, i_2, \ldots, i_q + 1, 0, \ldots, 0])$ and $([i_1, i_2, \ldots, i_q, \ldots, 0], [i_1, i_2, \ldots, i_q + 1, 0, \ldots, 0])$, both of them with weight 0;
  (iii) remove $[i_1, i_2, \ldots, i_q, 0, \ldots, 0]$ from the list L and return to Step 1.

(q.3) If $i_1 + i_2 + \cdots + i_q = n$ then

  (i) remove $[i_1, i_2, \ldots, i_q, 0, \ldots, 0]$ from the list L and return to Step 1.

Step k. Since $i_k \neq 0$ then consider the following.

(k.1) If $i_1 + i_2 + \cdots + i_k \leq n - 1$ then

  (i) define a new node $[i_1, i_2, \ldots, i_{k-1}, i_k + 1]$ and add it to L;
  (ii) define the arc $([i_1, i_2, \ldots, i_{k-1}, i_k], [i_1, i_2, \ldots, i_{k-1}, i_k + 1])$ with weight

$$\sum_{r=1}^{k-1} h_{k-r} \left(\sum_{\ell = i_0 + \cdots + i_{r-1} + 1}^{i_0 + \cdots + i_r} p_{\ell, i_1 + i_2 + \cdots + i_k + 1} + p_{i_1 + i_2 + \cdots + i_k + 1, \ell} \right); \quad (9)$$

  (iii) remove $[i_1, i_2, \ldots, i_{k-1}, i_k]$ from the list L and return to Step 1.

(k.2) If $i_1 + i_2 + \cdots + i_k = n$ then

  (i) remove $[i_1, i_2, \ldots, i_{k-1}, i_k]$ from the list L and we are done.

4.2. Optimal Solution for n Stops and k Zones. With a similar argument as in the proof of Theorem 4, we can get the optimal solution for k zones.

Theorem 5. *Let us consider n stops to be distributed into k zones and assume that the number of passengers between every pair of stops $(p_{i,j})_{1 \leq i, j \leq n}$ is given. Let T be a weighted tree constructed under assumptions (A1), (A2), and (A3) as it is indicated in Section 4.1.*

Then the longest path from the root to $[1, 0, \ldots, 0]$ to a leaf of the form $[i_1, \ldots, i_k]$ with $1 \leq i_1, \ldots, i_k \leq n$ and $i_1 + \cdots + i_k = n$ returns the maximal benefit for distributing n stops into k zones.

Remark 6. We point out that in substeps $1.2, \ldots, k - 1.2$ some nodes defined are useless for finding the solution, since they do not belong to any path from the root $[1, 0, 0, \ldots, 0]$ to any admissible distribution. However, their definition will simplify the enumeration of the nodes as we will see in the next section.

5. Some Considerations regarding the Nodes

The algorithm in Section 4 lets us construct a tree that models our problem. As we have seen in the previous section, the construction of the nodes and the arcs joining them can be done inductively. However, it could be of interest to find an explicit formula for calculating the total number of nodes in each tree; see, for instance, the case of 3 zones and 5 stops in Figure 5 and the case of 4 zones and 6 stops in the figure in the Supplementary Material available online at http://dx.doi.org/10.1155/2014/384321.

Suppose that we have n stops and k zones, with $n > k \geq 2$. At the qth step of the construction of the tree, $q = 1, \ldots, n$, we have all admissible distributions of q stops distributed in a row. More precisely we have all distributions of the form $[i_1, \ldots, i_k]$ with $\sum_{\ell=0}^{k} i_\ell = q$ and if $i_\ell = 0$, then the next indexes are null; that is, $i_{\ell+1} = \cdots = i_k = 0$.

We recall that for every $q < n$, every node labelled in the step q as $[i_1, \ldots, i_k]$ is split into one or two nodes at step $q + 1$. More precisely, a node $[i_1, \ldots, i_j, 0, \ldots, 0]$, with $i_1, \ldots, i_j \neq 0$ and $j < k$, gives birth to $[i_1, \ldots, i_j + 1, 0, \ldots, 0]$ and $[i_1, \ldots, i_j, 1, \ldots, 0]$. A node at zone k in the step q of the form $[i_1, \ldots, i_k]$, with $i_k \neq 0$, is only connected with the node $[i_1, \ldots, i_k + 1]$ at the zone k in the step $q + 1$. All these nodes can be enumerated using properties of Pascal's triangle; see, for instance, [37].

Theorem 7. *Let us denote by $z_{q,j}$ the number of nodes at zone j at Step q of the construction of the tree that models the problem for n stops into k zones. Then*

$$z_{q,j} = \binom{q-1}{j-1} = \frac{\prod_{l=1}^{j-1} (q-l)}{(j-1)!} \quad \text{for } 1 \leq q \leq n, \ 1 \leq j \leq k.$$

$$(10)$$

Proof. Let us denote by $z_{q,j}$ the number of nodes at zone j at Step q of the construction of the tree, which is defined for
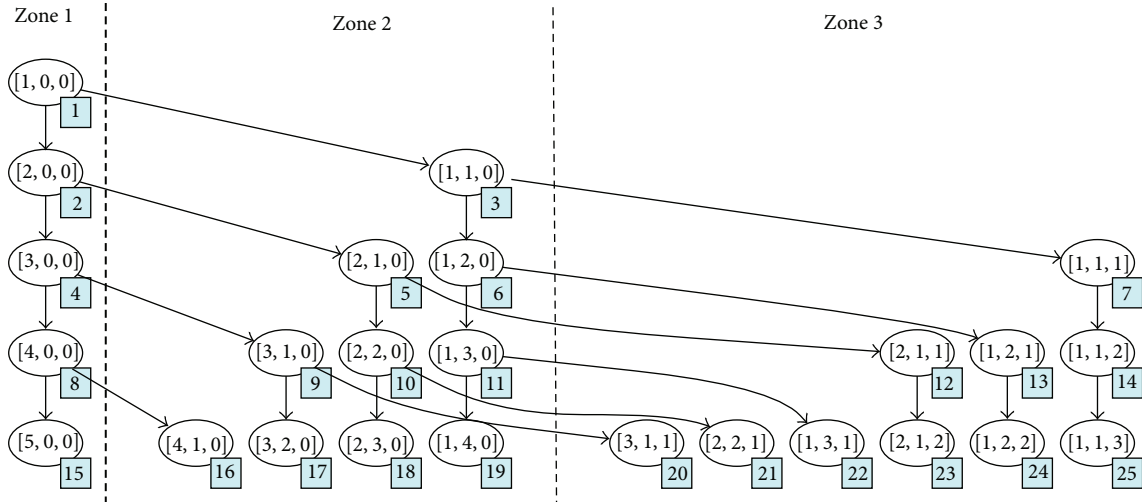
FIGURE 5: Tree for 3 zones and 5 stops with labels at the nodes.

all pairs (q, j) with $1 \leq q \leq n$ and $1 \leq j \leq \min\{k, q\}$. The sequence $\{z_{q,j}\}_{j,q}$ can be inductively defined as

$$
\begin{cases}
z_{q,1} = 1 & \forall 1 \leq q \leq n, \\
z_{q,q} = 1 & \forall 1 \leq q \leq k, \\
z_{q,j} = z_{q-1,j-1} + z_{q-1,j} & \text{if } 1 < q \leq n, \ q < j \leq k.
\end{cases}
\tag{11}
$$

This recurrence is similar to the one used for generating the combinatorial numbers in Pascal's triangle [37]. Therefore, we can identify the numbers $z_{q,j}$ with the combinatorial numbers expressed in (10). □

As the number of nodes in a row is represented by combinatorial numbers, in order to assign labels to the nodes, we can use an existing connection with the partial sums by rows of Pascal's triangle. Furthermore, explicit formulas for these numbers can be provided using generating functions [38].

Proposition 8 (see [38]). *Let us consider the functions*

$$
G_j(x) := \frac{1}{x^{j-1}} \left(\sum_{i=0}^{j-1} \frac{x^i}{(1-x)^{i+1}} - \frac{(2x)^{j-1} - 1}{2x - 1} \right) \quad \text{for } 1 \leq j \leq k.
\tag{12}
$$

If we denote by $G_{j,i}$ the ith coefficient of the Taylor expansion of G_j at $x = 0$, then the following relation holds:

$$
\sum_{\ell=1}^{j} z_{q,\ell} = \sum_{\ell=1}^{j} \binom{q-1}{\ell-1} = G_{j,q-j} \quad \text{for } j \leq q.
\tag{13}
$$

Example 9. Let us analyze with more details the basic cases of 2 and 3 zones with n stops. For 2 zones, the total number of nodes of the tree generated for distributing n stops is

$$
\sum_{q=1}^{n} \binom{q-1}{0} + \sum_{q=2}^{n} \binom{q-1}{1} = \frac{n^2 + n}{2}.
\tag{14}
$$

Similarly, for 3 zones and n stops the number of nodes needed to define the tree is

$$
\sum_{q=1}^{n} \binom{q-1}{0} + \sum_{q=2}^{n} \binom{q-1}{1} + \sum_{q=3}^{n} \binom{q-1}{2} = \frac{n^3 + 5n}{6},
\tag{15}
$$

and for 4 zones and n stops we have a tree with the following number of nodes:

$$
\sum_{q=1}^{n} \binom{q-1}{0} + \sum_{q=2}^{n} \binom{q-1}{1} + \sum_{q=3}^{n} \binom{q-1}{2} + \sum_{q=4}^{n} \binom{q-1}{3}
$$
$$
= \frac{n^4 - 2n^3 + 11n^2 + 14n}{24}.
\tag{16}
$$

Remark 10. The formulas in Proposition 8 and in Example 9 can be obtained taking into account that the sum of combinatorial numbers along a diagonal of Pascal's triangle can be expressed as an arithmetic sequence of higher order [39]. The general formula for the number of nodes in the problem of k zones and n stops is given by

$$
\sum_{j=1}^{k} \sum_{q=1}^{n} \binom{q-1}{j-1}
\tag{17}
$$

that can be also rewritten as

$$
n + \frac{1}{2!} \prod_{l=0}^{1} (n - l) + \frac{1}{3!} \prod_{l=0}^{2} (n - l)
$$
$$
+ \cdots + \frac{1}{(k-1)!} \prod_{l=0}^{k-2} (n - l) + \frac{1}{k!} \prod_{l=0}^{k-1} (n - l).
\tag{18}
$$

Therefore, using properties of arithmetic sequence of higher order [39] and if n is much more greater than k, we affirm that the number of nodes of the tree for the problem of k zones and n stops is of order $\mathcal{O}(n^k)$.

We can also enumerate all admissible distributions of a tree using a sequential order. As we have already seen, this can be done recursively. Here we provide explicit formulas for the cases of 2 and 3 zones.

Example 11. The number corresponding to a node representing the admissible distribution corresponding to the case of 2 zones and n stops $[i_1, i_2]$, with $i_1 + i_2 \leq n$, can be easily defined from the sequence $\{a_j\}_{j \geq 1}$ with $a_j = (j^2 + j)/2$:

$$
\begin{cases}
1 + a_{i_1 - 1} & \text{for } i_1 \neq 0, \ i_2 = 0, \\
1 + a_{i_1 + i_2 - 1} + i_2 & \text{for } i_1, i_2 \neq 0.
\end{cases}
\tag{19}
$$

For the case of 3 zones and n stops, the number assigned to a (partial) admissible distribution $[i_1, i_2, i_3]$, with $i_1 + i_2 + i_3 \leq n$, can be easily defined from the sequence $\{a_j\}_{j \geq 1}$ with $a_j = (j^3 + 5j)/6$. The label of $[i_1, i_2, i_3]$ is

$$
\begin{cases}
1 + a_{i_1 - 1} & \text{for } i_1 \neq 0, \ i_2, i_3 = 0, \\
1 + a_{i_1 + i_2 - 1} + i_2 & \text{for } i_1, i_2 \neq 0, \ i_3 = 0, \\
1 + a_{i_1 + i_2 + i_3 - 1} + i_2 \\
\quad + \sum_{l=1}^{i_3} (i_1 + i_2 + i_3 - l) & \text{for } i_1, i_2, i_3 \neq 0.
\end{cases}
\tag{20}
$$

Remark 12. We point out that in the algorithm of Section 4, at some steps of the form $(*.2)$, we were introducing additional nodes for simplifying the notation as it is indicated at the beginning of this section; see also Remark 6, despite the fact that they are not considered for finding the optimal solution. Since no admissible distribution is accessible from one of these nodes, then we can remove them and reduce the number of nodes that we are considering in our model. Nevertheless, we have preserved them for the clarity of the enumeration.

6. Discussion and Results

As we have already mentioned in Section 2, for n stops to be distributed into k zones, we have to evaluate a number of $\binom{n-1}{k-1}$ possible solutions. The cost of generating all these solutions is of order $\mathcal{O}(n^k)$, since k nested loops from 1 to n are required. For each one of these possible solutions one has to consider the $n(n-1)$ arrangements of two different stops and to compute the income received from the passengers between every pair of stops (in both directions). Then the cost is of order $\mathcal{O}(n^{k+2})$.

If we think in terms of the tree constructed in Section 4.1 for modeling the problem, each possible solution to the problem of the optimal assignment of stops into zones is associated with a leaf of that tree, or in other words to a node in the zone k at step n. This is consistent with the definition of $z_{n,k}$, which is in fact $\binom{n-1}{k-1}$. So as to get the benefit from a certain distribution $[i_1, \ldots, i_k]$, of n stops into k zones, we have to compute the length (in terms of weight) of the unique path that connects the root $[1, 0, \ldots, 0]$ with the node associated with $[i_1, \ldots, i_k]$. Each path from the root to a leaf consists of $n - 1$ edges. The tree structure used

for representing the distributions of stops into zones lets us obtain the weight of each of them by computing once the weight of each arc in order to get the optimal solution to our problem. This improves the computational cost of the solution with respect to the combinatorial approach. Let us see in detail the benefits of using the tree-based approach.

Firstly, as we have indicated in the previous section, the number of nodes is of order $\mathcal{O}(n^k)$. This also holds for the number of arcs, since the number of arcs in a tree is equal to the number of nodes minus 1.

Secondly, the computational cost of the calculations for assigning a weight to a single arc is of order $\mathcal{O}(n)$: for every arc with an end node of the form $[i_1, i_2, \ldots, i_j, 0, \ldots, 0]$ we need at most $2(i_1 + i_2 + \cdots + i_{j-1})$ sums plus at most $j - 1$ products in order to determine the benefit from the passengers between the first $i_1 + \cdots + i_{j-1}$ and the last one; see (7), (8), and (9). These computations are of order $\mathcal{O}(n)$. The same estimation of the computational cost is obtained if the end node is of the form $[i_1, i_2, \ldots, i_k]$ with $i_k \neq 0$. Therefore, the cost of assigning weights to all arcs of the tree is of order $\mathcal{O}(n^{k+1})$.

Finally, for finding the longest path from the root to a leaf we need $\mathcal{O}(n^k)$ operations because we only need to consider each arc only once as we can see right now. Consider the $z_{n,k}$ leaves of the tree, which are located at level n. Each one is connected to a different node in the level $n - 1$. But some of the nodes at level $n - 1$ share their adjacent node at level $n - 2$. For instance, a pair of nodes of the form $[i_1, \ldots, i_{k-1}, 1]$, $[i_1, \ldots, i_{k-1} + 1, 0]$ at level $n - 1$, share its adjacent node at level $n - 2$, which is $[i_1, \ldots, i_{k-1}, 0]$. This also holds for nodes of the form $[i_1, \ldots, i_j, 1, 0, \ldots, 0]$, $[i_1, \ldots, i_j + 1, 0, 0, \ldots, 0]$ for $j = 2, \ldots, k - 2$.

Our strategy will be to start with the leaves and find their adjacent nodes at level $n - 1$. Then we find the adjacent nodes to them at level $n - 2$. For every node at level $n - 2$, we choose between the two only possibilities we have and record which is the path to a leaf that provides a longer path. Inductively, we repeat the process with nodes at an upper level until we arrive at the root. Therefore, once the tree is defined, for finding the distribution with highest revenue, we have to perform a number of computations of order $\mathcal{O}(n^k)$, similar to the order of the number of arcs in the tree.

All this together gives that we need $\mathcal{O}(n^{k+1})$ computations to find the solution using the tree-based approach for finding the solution to our problem, which is smaller than the computational cost of the combinatorial approach, $\mathcal{O}(n^{k+2})$.

For illustrating these results we have considered the problem of splitting 10 stops into 2, 3, 4, 5, and 6 zones. In particular we have considered an estimation of the number of the passengers of the railway line Valencia-Castellón over the month of August 2013. The number of passengers between each pair of stops is given in Table 1.

We present in the table the most efficient distributions of the 10 stops into different number of zones. For 3 zones the optimal distribution is $[7, 1, 2]$ and for 4 zones $[6, 1, 1, 2]$. Besides, for 5 and 6 zones the results are the ones shown in Figures 6 and 7.

TABLE 1

|	A	B	C	D	E	F	G	H	I	J
A	—	9260	5616	13692	3254	13804	1468	4564	19586	29611
B	5689	—	48	2402	397	1526	678	83	3545	3378
C	1334	28	—	788	53	388	410	230	558	573
D	3485	1354	524	—	225	479	580	294	2193	2922
E	507	17	3	402	—	94	16	393	361	1125
F	3463	934	306	629	806	—	1052	2753	3706	6300
G	576	236	240	390	340	510	—	1048	1802	1799
H	1301	344	54	495	509	1353	257	—	4014	6741
I	13065	4124	1334	3566	931	4669	1778	5249	—	867
J	33181	9914	4395	9494	3752	26694	5931	17326	1824	—
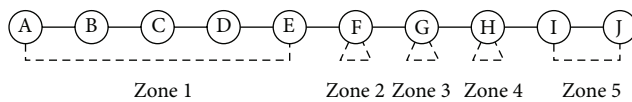
FIGURE 6: Optimal distribution for the 10 stops of Table 1 into 5 zones. This corresponds with [5, 1, 1, 1, 2].
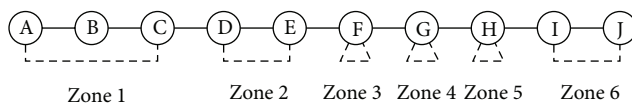

FIGURE 7: Optimal distribution for the 10 stops of Table 1 into 6 zones. This corresponds with [3, 2, 1, 1, 1, 2].

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] N. Paulley, R. Balcombe, R. Mackett et al., "The demand for public transport: the effects of fares, quality of service, income and car ownership," *Transport Policy*, vol. 13, no. 4, pp. 295–306, 2006.

[2] F. V. Webster and P. H. Bly, "The demand for public transport. Report of an international collaborative study," Transport and Road Research Laboratory, Crowthorne, Berkshire, UK, 1980.

[3] R. Cervero, "Flat versus differentiated transit pricing: what's a fair fare?" *Transportation*, vol. 10, no. 3, pp. 211–232, 1981.

[4] D. Fleishman, N. Shaw, A. Joshi, R. Freeze, and R. Oram, "Fare policies, structures, and technologies," Tech. Rep. 10, Transportation Research Board, 1996.

[5] V. R. Vuchic, *Urban Transit: Operations, Planning and Operations*, John Wiley & Sons, New York, NY, USA, 2004.

[6] G. Kocur and C. Hendrickson, "Design of local bus service with demand equilibrium," *Transportation Science*, vol. 16, no. 2, pp. 149–170, 1982.

[7] P. M. Schonfeld, "Multiple period optimization of bus transit systems," *Transportation Research Part B*, vol. 25, no. 6, pp. 453–478, 1991.

[8] S. I.-J. Chien and L. N. Spasovic, "Optimization of grid bus transit systems with elastic demand," *Journal of Advanced Transportation*, vol. 36, no. 1, pp. 63–91, 2002.

[9] L. N. Spasovic, M. P. Boile, and A. K. Bladikas, "Bus transit service coverage for maximum profit and social 440 welfare," *Transportation Research Record*, vol. 1451, pp. 12–22, 1994.

[10] W. H. K. Lam and J. Zhou, "Optimal fare structure for transit networks with elastic demand," *Transportation Research Record*, no. 1733, pp. 8–14, 2000.

[11] G.-S. Yoo, D.-K. Kim, and K. S. Chon, "Frequency design in urban transit networks with variable demand: model and algorithm," *KSCE Journal of Civil Engineering*, vol. 14, no. 3, pp. 403–411, 2010.

[12] T. Feng-Ming, *Optimizing fare structure and service frequency for an intercity transit system [Ph.D. thesis]*, New Jersey Institute of Technology, Newark, NJ, USA, 2009.

[13] H. W. Hamacher and A. Schöbel, "Design of zone tariff systems in public transportation," *Operations Research*, vol. 52, no. 6, pp. 897–908, 2004.

[14] H. W. Hamacher and A. Schobel, "On fair zone design in public transportation," in *Computer Aided Transit Scheduling*, J. R. Daduna, J. Branco, and J. M. Pinto, Eds., vol. 430 of *Lecture Notes in Economics and Mathematical Systems*, pp. 8–22, Springer, Berlin, Germany, 1995.

[15] L. Babel and H. Kellerer, "Design of tariff zones in public transportation networks: theoretical results and heuristics," *Mathematical Methods of Operations Research*, vol. 58, no. 3, pp. 359–374, 2003.

[16] C. Haksever, B. Render, R. Russell, and R. Murdick, *Service Management and Operations*, Prentice Hall, Upper Saddle River, NJ, USA, 2000.

[17] S. Raff, "Routing and scheduling of vehicles and crews: the state of the art," *Computers & Operations Research*, vol. 10, no. 2, pp. 63–211, 1983.

[18] D. Teodorovic, *Airline Operation Research*, Gordon and Breach Science, New York, NY, USA, 1988.

[19] S. Yan and C. H. Tseng, "A passenger demand model for airline flight scheduling and fleet routing," *Computers and Operations Research*, vol. 29, no. 11, pp. 1559–1581, 2002.

[20] J. A. Conejero, C. Jordan, and E. Sanabria-Codesal, "An iterative algorithm for the management of an electric-car-rental service," *Journal of Applied Mathematics*, vol. 2014, Article ID 483734, 11 pages, 2014.

[21] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall, 1993.

[22] N. Ben-Khedher, J. Kintanar, C. Queille, and W. Stripling, "Schedule optimization at SNCF: from conception to day of departure," *Interfaces*, vol. 28, no. 1, pp. 6–23, 1998.

[23] A. Caprara, L. Kroon, M. Monaci, M. Peeters, and P. Toth, "Passenger railway optimization," in *Handbooks in OR & MS*, C. Barnhart and G. Laporte, Eds., vol. 12, Elsevier Science, 2006.

[24] H. Bisheng, S. Rui, H. Shiwei, and X. Yue, "High-speed rail train timetabling problem: a time-space network based method with an improved branch-and-price algorithm," *Mathematical Problems in Engineering*, vol. 2014, Article ID 641562, 15 pages, 2014.

[25] V. Cacchiani, A. Caprara, and P. Toth, "Solving a real-world train unit assignment problem," *Mathematical Programming B*, vol. 124, pp. 207–231, 2010.

[26] V. Cacchiani, A. Caprara, and P. Toth, "A Lagrangian heuristic for a train-unit assignment problem," *Discrete Applied Mathematics*, vol. 161, no. 12, pp. 1707–1718, 2013.

[27] F. Barber, L. Ingolotti, A. Lova et al., "Integrating timetabling, network and line design," Tech. Rep. ARRIVAL-TR-0151, 2008.

[28] A. Marin, J. A. Mesa, and F. Perea, "Integrating robust network design and line planning under failures," in *Robust and Online Large-Scale Optimization*, R. K. Ahuja, R. H. Mohring, and C. D. Zaroliagis, Eds., pp. 273–292, Springer, Berlin, Germany, 2009.

[29] M. Michaelis and A. Schöbel, "Integrating line planning, timetabling, and vehicle scheduling: a customer-oriented heuristic," *Public Transport*, vol. 1, no. 3, pp. 211–232, 2009.

[30] M. Demange, G. Di Stefano, and B. Leroy-Beaulieu, "On the online track assignment problem," *Discrete Applied Mathematics*, vol. 160, no. 7-8, pp. 1072–1093, 2012.

[31] X. Liu, S. He, R. Song, H. Li, L. Wang, and Y. Sun, "Synthetic optimization model and algorithm for railway freight center station location and wagon flow organization problem," *Mathematical Problems in Engineering*, vol. 2014, Article ID 246791, 12 pages, 2014.

[32] R. Borndörfer, M. Karbstein, and M. E. Pfetsch, "Models for fare planning in public transport," *Discrete Applied Mathematics*, vol. 160, no. 18, pp. 2591–2605, 2012.

[33] M. Neumann, "Fare planning for public transport," *Public Transport*, vol. 1, pp. 211–232, 2009.

[34] A. Dolan and J. Aldous, *Networks and Algorithms: An Introductory Approach*, John Wiley & Sons, Chichester, UK, 1993.

[35] J. R. Evans and E. Minieka, *Optimization Algorithms for Networks and Graphs*, Marcel Dekker, New York, NY, USA, 1992.

[36] J. L. Gross and J. Yellen, *Graph Theory and Its Applications*, Chapman & Hall/CRC Press, New York, NY, USA, 2006.

[37] "The On-Line Encyclopedia of Integer Sequences," Pascal's triangle, https://oeis.org/wiki/Pascal_triangle.

[38] The On-Line Encyclopedia of Integer Sequences. Bernoulli's triangle, 2014, https://oeis.org/wiki/Bernoulli%27s_triangle.

[39] J. Alonso, "Arithmetic sequences of higher order," *The Fibonacci Quarterly*, vol. 14, no. 2, pp. 147–152, 1976.