A MATHEURISTIC FOR OPTIMIZING SKIP–STOP OPERATION STRATEGIES IN RAIL TRANSIT LINES

FRANCISCO A. ORTEGA¹, JUAN A. MESA², RAMÓN PIEDRA-DE-LA-CUADRA³, MIGUEL A. POZO⁴

¹Higher Technical School of Architecture, Universidad de Sevilla, Spain
²Higher Technical School of Engineering, Universidad de Sevilla, Spain
³Department of Mathematics, Universidad de Cordoba, Spain
⁴Faculty of Mathematics, Universidad de Sevilla, Spain

ABSTRACT

In many transit systems, operators use skip–stop strategies to reduce travel time of particular train services by not stopping (skipping) at less densely populated stations. This decision of omitting some stops reduces the travel time for the users within the vehicle and increases the speed of operation, favouring the provision of new transit services where they are more necessary. In this work, the best A/B stop–skip patterns for a set of transit services along a railway corridor are determined by means of a three-phase methodology that includes the formulation of a nonlinear integer programming inspired in the multiple knapsack problem and the application of a heuristic algorithm based on mathematical properties (matheuristic).

Keywords: matheuristics, optimization, railways, stop–skip strategy.

1 INTRODUCTION

The skip–stop (or limited-stop) mechanism is a strategy that some transport companies follow to reduce travel times without the need to increase the fleet of vehicles. It consists of privileging a majority of passengers by offering shorter travel times, as a result of having previously selected a group of low-activity stations, where the trains will not stop to leave/pick up passengers.

Distinguishing between express and local stations appears first in the Northwestern Elevated of Chicago by July 1900. The skip–stop service was also first developed for the Chicago Metro system in 1947, and later implemented in Philadelphia and New York (see [1]). In 1947, system of express and local schedules provided by the Chicago Transit Authority (CTA) had become a nuisance for users, because the really available services for riders were hard to comprehend. In order to stop the fall in demand that was being happening, the CTA planned a clever way of running express service on its two-track lines. This proposal was known as the A/B skip–stop plan (see [2]). As illustrative examples, the skip–stop operation mode has been used (and/or is being used) in practice in Santiago de Chile since 2001 [3], in Bogota’s Transmilenio system [4], in Los Angeles’s Metro Rapid system [5], in Singapur’s Transit Link [6] and in the bus line that connects the East zone of Seville area with its historic centre [7]. Figure 1 shows four train services that run along a railway corridor with seven stations. The services appear classified in types A and B and the stations in types A, b and A/b. The horizontal sections in the polygonal lines represent the stopping time of the trains at the stations.
The travel time between stations along a railway line consists of five components, usually identified as phases of acceleration, constant speed, inertia, braking and downtime. Several studies have shown that skip–stop operations can:

- improve passenger waiting and in-vehicle times
- save operating costs (note that skipping stops allows vehicles to return to their depots in a shorter period of time; as a result, vehicles can be reused sooner) and
- reduce fuel (or power), as a result of not accelerating or decelerating at skipped stations.

Nevertheless, the prevalent purpose of introducing stop–skip patterns is not actually to save time, as discussed in [8, 9]. The main aim is to better distribute passenger loads during peak periods, when trains are at their highest capacity level. The skip–stop services are especially suitable for those transit routes with unbalanced demand, so the stops with higher demand would be served by more vehicles, in order to improve the overall serviceability of the route. The idea consists of trains stopping at stations so that there’s more of an opportunity for folks to get on the trains at those stops. In order to avoid confusions and misunderstandings to passengers when skip–stop services are implemented, different means are usually used, like providing information boards and verbal indications at the stations.

As the Directive (EU) 2016/797 of the European Parliament and of the Council of 11 May 2016 on the interoperability of the rail system within the European Union [10] points out, passengers must be provided with easily understandable travel information about rules applicable to them both in railway stations and in trains. For example, the two types of buses that operate in the east of Seville and share the same route (express line and normal line) are differentiated by signals on the front panel display of the vehicle. In the same way, the trains of their two respective lines that operate with skip–stop patterns in Santiago de Chile are also visually identified with green and red signals; hence, passengers know in advance what colour they should choose. Additionally, this information is also provided on screens at
the station platforms, as well as through the public address system when trains are going to arrive at the stations.

Regarding the stop–skipping patterns for a one-way single track, the fundamental approaches are divided into deterministic (see, for instance, [11], [3]) and stochastic approaches ([12]). The deterministic form is derived from the description and analysis given by [13] in which stations along a line are classified into three groups, A, B and AB. The consideration of only two types of stations simplifies the degree of diversity, and travellers can more easily memorize the options to be able to configure their own routes between nodes of the transit network. The trains in line A stop at the A and AB stations, while the trains belonging to line B stop at the B and AB stations. When they intend to alight at a B station, passengers boarding at an A station will need to transfer at an Ab station onto line B. Thus, this disadvantage might affect the attractiveness of stop–skipping schedules.

The skip–stop operation scheme has been widely applied in bus transit services. Reference [14] proposed a real-time deadheading strategy to determine the dispatching time, deadhead vehicles and skip stations to minimize the total passenger cost. A heuristic algorithm was used to solve the model for operating the MBTA Green line. Reference [12] focused on the real-time stop–skipping control problem and presented an enumeration method with fast solving speed. Reference [15] studied the service reliability of a route in the city of Dalian and an optimized deadheading strategy for a part of the route by means of a heuristic algorithm. The advantages and disadvantages of four kinds of operating strategies were analysed in [16], and a nonlinear integer programming model was developed to solve the real-time dynamic transit operation problem, in a setting where the benefits of the operators and passengers were balanced. The skip–stop operation mode on rail transit lines has been described in the literature; for instance, see [17,18].

Different mathematical tools have been used to solve the skip–stop service problems:

- Dynamic programming was adapted for this purpose in [19] and [20].
- Greedy algorithms have been investigated for solving multiple train problems in [21].
- Fuzzy mathematical programming was the method used in [22] for the Taiwan’s high-speed rail.
- Formulations in terms of nonlinear integer programming were proposed in [23], [24], [4] and [25] for solving dynamic versions of the skip–stop service problem.
- Metaheuristic genetic algorithms (GAs) have extensively been used for solving skip–stop scheduling problems. See, for instance, [26], [27], [28], [29] and [30].
- Other metaheuristics such as Tabu search method ([31]) or bee colony algorithm ([6]) have also been used for this context.

Matheuristics are heuristic algorithms made by the interoperation of metaheuristics and mathematic programming (MP) techniques [32]. Matheuristics are optimization algorithms inspired in (or derived from) a mathematical model. An essential feature of the matheuristics is the implementation in some part of the solution procedure of characteristics or properties derived from a mathematical model. Metaheuristics topic has attracted the interest of researchers, as shown in the publication of monographs in journal special issues ([33,34]).

We propose, in this work, to determine a skip–stop scheme through a three-phase methodology. In the first, we find the optimal strategy of skipping stops for a given train fleet and, in the second phase, we determine, by means of a matheuristic procedure, the optimal allocation for train itineraries. For this last purpose, we will develop the concept of proximity between
configurations of train itineraries and, in accordance with Hall’s method ([35]), design a matheuristic that optimizes the skip–stop strategy.

In Section 2, a methodology of three phases for determining an optimal skip–stop scheme for train schedules is introduced. The first phase consists of formulating a nonlinear integer programming inspired in the multiple knapsack problem (MKP). The second phase is a matheuristic procedure adapted from the Hall’s method. The third phase is a greedy algorithm of comparing and replacing. A computational experience, which illustrates the proposed procedure, is carried out in Section 3. Finally, conclusions are summarized in Section 4.

2 METHODOLOGY
Let \( I \) be the set of stations of a railway corridor and let \( S \) be a train service set. We define the binary variable \( y^s_i \); \( i \in I, s \in S \). If \( y^s_i = 1 \), then station \( i \in I \) is visited by transit service \( s \in S \). This variable \( y^s_i \) will allow us to construct the solution vector: \( Y = (y^s_1, y^s_2, \ldots) \).

In order to deal with the demand the binary variables \( x^s_{ij} \); \( i, j \in I, s \in S \), which take value 1 when the train \( s \) stops at both stations \( i \) and \( j \), are defined.

Assume that travel demand from station \( i \) to station \( j \) depends on the time window considered and, subsequently, on the train \( s \) that transits during the time period considered. Hence, we can assume a preliminary study in which the potential demand from station \( i \) to \( j \) when train \( s \) passes through \( i \) has been estimated. This demand will be an initial data that we will denote by \( p^s_{ij} \); \( i, j \in I, s \in S \). This value will be associated to the population available to board train \( s \) at station \( i \) with destination \( j \) if the number of intermediate stops was 0.

Define a new variable \( n^s_{ij} \); \( i, j \in I, s \in S \) that will be the number of effective intermediate stops between stations \( i \) and \( j \) for train service \( s \). If stations \( i \) and \( j \) were consecutive stops along the railway corridor, then \( n^s_{ij} = 0, \forall s \in S \). Otherwise, \( n^s_{ij} \) will be an integer value, superiorly bounded by the real number of intermediate stations between \( i \) and \( j \). Now we can estimate the real number of travellers that will depend on the potential demand \( p^s_{ij} \); \( i, j \in I, s \in S \) and the number of intermediate stops \( n^s_{ij} \); \( i, j \in I, s \in S \). The greater number of intermediate stops introduced, the fewer number of travellers will be interested in the train service. We propose to use the following mathematical expression:

\[
w^s_{ij} = \frac{p^s_{ij}}{\lambda \cdot n^s_{ij} + 1}; \forall i, j \in I, j > i \text{ [where positive parameter } \lambda > 0 \text{ must be calibrated].}
\]

Therefore, if \( n^s_{ij} = 0 \) then the real demand \((w^s_{ij})\) coincides with the initial value \( p^s_{ij} \). Any other positive value of \( n^s_{ij} \) will suppose a decrease in the effective number of travellers with respect to the initial value.

By means of the following linear integer programming model, we formulate the problem of maximizing the number of passengers for a transit line in which an indeterminate number of intermediate stops along the line can be omitted.

**Objective and constraints:** Maximize the number of passengers boarding trains at stations.

\[
\max \sum_{s \in S} \sum_{i,j \in I, j > i} w^s_{ij} \cdot x^s_{ij}
\]  

subject to:

\[
w^s_{ij} = \frac{p^s_{ij}}{\lambda \cdot n^s_{ij} + 1}; \forall i, j \in I, j > i
\]
The objective function (1) maximizes the number of passengers for a transit line using a generalized multiple knapsack model; note that each train can be assumed like a backpack that may or may not pick up the demand for OD pair trips in their corresponding temporary windows. Constraints (2) identify the actual demand according to the number of intermediate stations. Constraints (3) prevent the capacity $c^s$ of train $s$ from being exceeded when it stops at each station $i \in I$. Constraints (4) imply that if it is decided to pick up travellers from an origin–destination pair, the respective service $s$ will have to stop at both stations. Constraints (5) indicate the nature of the variables used in the model.

The KP is a classic problem of combinatorial optimization that has been widely studied for more than a century (see, for example, [36]). It consists in selecting objects with the objective of filling a knapsack so that they provide the greatest profit without exceeding the storage capacity of the own knapsack. The MKP is a generalization of the standard KP where, instead of considering only a knapsack, it is about filling several knapsacks of different capacities. The problem of MKP is strongly NP-complete and, due to its computational complexity, the need for using heuristic algorithms for generating good solutions is justified (see, for instance, [37]). Previous model can be considered a variant of the MKP model, where each service may be assimilated with a different knapsack that stores passengers boarding the train from stations, as long as the capacity of the vehicle allows it.

**FIRST PHASE:** Taking these precedents into account, we propose the heuristic shown in Table 1 for solving the optimization problem in order to determine the most effective deployment of skip–stop services along the rail corridor.

Once the model is solved, we will obtain a set of optimal solutions (train services) that will indicate the stops that each train must make in its corresponding service, in order to globally maximize the number of passengers in the transit system. The solutions obtained can be very different from each other. But remember that we do not want each service to be different from one another in general but to divide the services into type A and type B; so in most cases, the optimal solutions obtained from the previous optimization model would not be the final solution to our problem, because our skip–stop scheme must only contain two different types of services. That is why we must develop a second phase. For that purpose, we propose a heuristic which transforms the optimal solutions of the KP into the best possible configuration for our skip–stop problem.
SECOND PHASE: The solutions obtained in the first phase are binary sequences where 1 in the $i$th position means train stops at stop $i$ and 0, otherwise. From Hamming/rectangular/Euclidean metrics, we can calculate the matrix $W$ of inter-distances between pair of binary sequences. In this way, we can classify the sequences according to a concept of proximity. This proximity is one-dimensional in nature. Therefore, we can construct a $W$ matrix of inter-distances (from any metric, like Hamming or rectangular or euclidean) between pairs of service sequences and, based on the method published by Hall in 1970 (where a spatial interpretation of maximum eigenvectors of the matrix $B = D – W$ is made) [35], we will obtain the relative position on the OX axis of the representative points. This relative position will allow us to establish a classification of trains and stations in types A and B. Matheuristic shown in Table 2 is inspired in the above-mentioned work.
THIRD PHASE: Note that coordinates of eigenvector $v_{max}$ are values included in interval $[-1,1]$. The $i$th point corresponding to $i$th coordinate of $v_{max}$ indicates the relative position of the $i$th train service within interval $[-1,1]$. The observed proximity between points will allow us to classify both trains and stations in types A and B. For this third phase, we propose the application of the greedy algorithm of comparing and replacing shown in Table 3.

**Table 3: Heuristic 3.**

**STEP 1:** Denote by $i$ and $j$ the trains corresponding to the two points farthest from each other in the previous distribution. Let train $i$ be included in type A and train $j$ in type B.

**STEP 2:** For each station $k$ do

- If train $i$ (type A) stops (i.e. is equal to 1 in $k$th coordinate) and train $j$ doesn’t stop (=0 in $k$) then INCLUDE station $k$ in type A set.
- If train $i = 0$ and train $j=1$ in $k$th coordinate then INCLUDE station $k$ in type B set.
- If train $i = 1$ and train $j=1$ in $k$th coordinate then INCLUDE station $k$ in type AB set.

**STEP 3:** For each intermediate train $m$ do

- Compute the number of coincident coordinates with respect to trains $i$ and $j$.
- Choose, between $i$ and $j$, the train where a higher number of coincidences was reached with $m$ (assume, for example, $i$)
- Force coincidences in the binary sequences (changing values 0 to values 1) until trains $i$ and $m$ belong to the same type.
3 COMPUTATIONAL EXPERIENCE

In order to illustrate the developed methodology, let us suppose a railway corridor with five stations where four trains circulate.

**PHASE I.** Assume that, as a result of the optimization procedure, the optimal sequence of skip–stop operations for the four trains is represented by means the sequences:

\[ s=1: (1,1,0,0,1) \]
\[ s=2: (0,1,0,1,0) \]
\[ s=3: (0,1,1,1,0) \]
\[ s=4: (0,0,1,1,1) \]

From these data, the distance matrix \( W \) between each pair of sequences can be built. For Hamming’s distance, the rows of matrix \( W \) are the following:

\[
\begin{array}{c}
(1,1,0,0,1) \\
(0,1,0,1,0) \\
(0,1,1,1,0) \\
(0,0,1,1,1)
\end{array}
\Rightarrow
\begin{bmatrix}
0 & 3 & 4 & 4 & 0 \\
3 & 0 & 1 & 3 & 0 \\
4 & 1 & 0 & 2 & 0 \\
4 & 3 & 2 & 0 & 0
\end{bmatrix}
\]

**PHASE II.** According to the above-mentioned methodology, let us build from \( W \), matrices \( D \) and \( B \).

\[
D = \begin{pmatrix}
11 & 0 & 0 & 0 \\
0 & 7 & 0 & 0 \\
0 & 0 & 7 & 0 \\
0 & 0 & 0 & 9
\end{pmatrix} \quad B = D - W = \begin{pmatrix}
11 & -3 & -4 & -4 \\
-3 & 7 & -1 & -3 \\
-4 & -1 & 7 & -2 \\
-4 & -3 & -2 & 9
\end{pmatrix}
\]

The eigenvalues of matrix \( B \), ordered from highest to lowest, are 14.8482, 11.3273, 7.82446 and 0. Note that the eigenvalue 0 is always sold by the construction of the matrix \( B \). The meaning of each eigenvalue is the value of the objective function, as is referred in [35]. Since our interest is the maximization of inter-distances, for better appreciating the existing differences between binary sequences, we select the maximum eigenvalue (14.8482) and calculate its associated eigenvector. The coordinates of this eigenvector provide us the position on the OX axis of the four points representative of the binary sequences.

The eigenvector corresponding to the highest eigenvalue 14.8482 is:

\[-0.848468, 0.128783, 0.312184, 0.407501\]

which indicates the relative positions of train sequences 1, 2, 3 and 4 on the OX axis. In Fig. 2, these four points are graphically represented in interval \([-1,1]\).

Figure 2: Relative positions in segment \([-1,1]\) of four binary sequences associated to trains.
PHASE III. Note that the transit services, corresponding to the two points farthest from each other in Fig. 2, are sequences 1 and 4. Let train 1 be included in type A and train 4 in type B.

According to this initial assignment of trains type A and type B, the stations are classified as follows:

<table>
<thead>
<tr>
<th>Train</th>
<th>Type</th>
<th>Station</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train 1</td>
<td>(1,1,0,0,1) Type A</td>
<td>Station 1: Type A</td>
</tr>
<tr>
<td>Train 2</td>
<td>(0,1,0,1,0) Type ?</td>
<td>Station 2: Type A</td>
</tr>
<tr>
<td>Train 3</td>
<td>(0,1,1,1,0) Type ?</td>
<td>Station 3: Type B</td>
</tr>
<tr>
<td>Train 4</td>
<td>(0,0,1,1,1) Type B</td>
<td>Station 4: Type B</td>
</tr>
</tbody>
</table>

When the intermediate trains are compared with their prototypes, some modifications are required. Since the point corresponding to service 2 is closer to the point corresponding to service 4 than the point corresponding to service 1 (see Fig. 2), we must assign type B to service 2. To achieve a total match between services 2 and 4, it is necessary to modify both the configurations of the sequences of 2 and 4.

<table>
<thead>
<tr>
<th>Train</th>
<th>Type</th>
<th>Station</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train 1</td>
<td>(1,1,0,0,1) Type A</td>
<td>Station 1: Type A</td>
</tr>
<tr>
<td>Train 2’</td>
<td>(0,1,1,1,1) Type B</td>
<td>Station 2: Type A/B</td>
</tr>
<tr>
<td>Train 3’</td>
<td>(0,1,1,1,0) Type B</td>
<td>Station 3: Type B</td>
</tr>
<tr>
<td>Train 4’</td>
<td>(0,1,1,1,1) Type B</td>
<td>Station 4: Type B</td>
</tr>
</tbody>
</table>

Similar reasoning can be applied to the point corresponding to service 3. There are initially four coordinates that coincide when we compare the binary sequences associated with services 3 and 4. To achieve total match, we modify one more coordinate of service 3.

<table>
<thead>
<tr>
<th>Train</th>
<th>Type</th>
<th>Station</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train 1</td>
<td>(1,1,0,0,1) Type A</td>
<td>Station 1: Type A</td>
</tr>
<tr>
<td>Train 2’</td>
<td>(0,1,1,1,1) Type B</td>
<td>Station 2: Type A/B</td>
</tr>
<tr>
<td>Train 3’</td>
<td>(0,1,1,1,1) Type B</td>
<td>Station 3: Type B</td>
</tr>
<tr>
<td>Train 4’</td>
<td>(0,1,1,1,1) Type B</td>
<td>Station 4: Type B</td>
</tr>
</tbody>
</table>

Finally, it is possible to determine an optimal classification of trains and stations for the implementation of a skip–stop strategy.

4 CONCLUSIONS

The skip–stop operation represents a low-cost approach to improve the operation speed into transit networks without additional investments in infrastructure are required. A three-phase methodology for determining an optimal skip–stop scheme for train schedules has been introduced. Matheuristic procedure includes the formulation of a nonlinear integer programming inspired in the MKP, the application of an algorithm adapted from Hall’s method and, finally, the implementation of a greedy algorithm of comparing and replacing. To illustrate the performance of the proposed procedure, the methodology has been applied to a laboratory case.

As a future research, we propose to generalize the mathematical model by considering the possibility of trains not stopping at stations and, additionally, by taking transshipments into account, as feature which affects the passenger’s behaviour.
ACKNOWLEDGEMENTS
This research has been partially supported by the Spanish projects MTM2015-67706-P and MTM2016-74983-C2-1-R (MINECO/FEDER, UE).

REFERENCES


