

Juraj Cenek *

EVALUATION OF A HEURISTICAL ALGORITHM FOR TRANSIT LINE PLANNING

Transit line routing is a part of strategic planning of transportation services in a region. As such it is vital for providing services of a standard quality for an acceptable price. From a mathematical point of view line planning is an NP hard problem which is difficult to solve and so mostly heuristic methods or an interactive approach supported by simulations are used. The paper describes shortly a new heuristic algorithm, evaluates its quality on a design of set of regional lines and compares the designed plan with a current transportation services in a real region chosen as an example.

Keywords: *Transportation planning, transit lines, optimization, heuristics.*

1. Introduction

The line routing problem is a challenge, which every government or delegated agency providing basic transportation service must face. Solving such a problem is not straightforward while quite an amount of criteria from passenger's point of view as well as from transport company's view must be respected. Some approaches addressing the problem known so far such as discussed in [1], [2], [3] classify the line routing problem as an NP-hard problem, often described by a non-linear model. Solution techniques are usually heuristics. An interactive process and computer simulation may be used as well.

This article evaluates a new heuristic algorithm for designing a set of regional line routes and presents results of tests of algorithm on real input data.

2. Problem description

State and regional administration arrange for a basic public transportation system. Such a service can be described as a periodical process of providing transportation service to population on set of routes in certain times. First step of designing a mass transportation system is a transit lines planning or, in other words, setting routes on which transportation services will be offered. The transportation service should be offered in public interest, which means that every inhabitant has a fair access to schools, administration offices, healthcare services, etc...

3. Data model and input parameters

The line planning is based on transportation infrastructure data and data about transportation flows. Transportation infrastructure

consists of nodes (for example towns, crossings, villages) and edges (for example roads, rails, generally connections between two network nodes). Nodes and edges can be specified by topographic and descriptive data. Topographic data are, for example, longitude and latitude of a node, identification of nodes from and to which certain edge goes, or shape description of an edge. Other data describe properties of relevant objects such as edge length, town name, population etc. Another input for algorithm is a subset of the node set representing attraction centers. Attraction centers are towns where major transportation flows origin or end. In reality these are district or county centers, or other regional centers, where services such as health care, education, administration offices, etc. are located.

Transportation demand data are defined by a so-called OD (origin - destination) square matrix, where every row and every column represents one town/village. Its items define how many passengers want to travel from the town represented by the current row to the town represented by the current column during a certain time period. Travel demand is time dependent, i.e. it usually changes during a day, a week, or even a year. While planning regional line routes, it is sufficient to distinguish only two time periods with different travel demands, namely a working day and a weekend day. The OD matrixes can be derived from the statistics about travelers obtained from the travel tickets and smart cards [4] or from the data that potential travelers input to the Internet explorer when planning their trips. Every approach to obtaining OD matrix brings up some problems. Data from internet planners provide data about occasional trips, while regular trips connections can be memorized. Data obtained from tickets database do not contain data about monthly-ticket owners, fixed-trip tickets, etc.

Every edge is assigned by a weight defining how many passengers are interested in passing this edge while traveling, i.e. the weight

* Juraj Cenek
INPROP, Zilina, Slovakia, E-mail: cenek@inprop.sk

corresponds to transportation flow on the edge. Weights can be computed as follows. At the beginning they are set to zero. Then we find the shortest path between towns corresponding to each item of the OD matrix and add the value of the current matrix item to the weight of each edge belonging to the shortest path.

4. Planning algorithm

The very core of the algorithm is based on tracking routes with maximal transportation flow. A line route is created by selecting the edge from an attraction centre with the maximal weight and successive extending the route by adding edges incident with the last added edge and having the maximal weights. The route creation process finishes if the last added edge weight is zero, or the last added edge ends in an attraction centre, or the difference between the line length and the shortest path length from the line initial node to the line terminal node is more than (described later). This way created route is a new line route. To prevent creating a very same route in the next step, all route edge weights are reduced by the minimal weight of an edge contained in the route. Such a line tracks the main flow of transportation demands, so the condition of minimizing transfer counts is respected. As initial edge weights were set on the edges of the shortest paths, the requirement of minimizing traveling times of passengers is respected as well. The proposed algorithm can be applied on a multimodal network as well, but output routes created for such a network must be manually or algorithmically corrected to prevent occurrence of unwanted transfers.

Inputs

Necessary inputs for the algorithm are:

- Vector infrastructure data, representing a graph G defined by a set of nodes V and a set of edges $H: G = (V, H)$, where nodes represent crossings and towns (network nodes), and edges represent links of the infrastructure.
- Q - set of ordered pairs (r, s) , where $r \in V, s \in V$ are towns, that have nonzero transportation demand
- OD (origin - destination) matrix - matrix of transportation flows between every pair of towns in the given region $O = [p^{rs}]$, where $(r, s) \in Q$
- Set of attraction centres C
- Maximal allowed increment of the line length d_e according to the shortest path length
- Maximal byway length d_d
- $p(v)$ - the weight of node v that represents population count of the node. If node v does not represent a town, then $p(v) = 0$.

Variables that are used in the algorithm description:

- c_i - the weight of edge h_i , specifying transportation flow on the edge
- $R_{r,s} = \{h_1, h_2, \dots, h_n\}$ - the shortest path from node r to node s consisting of edges h_1, h_2, \dots, h_n
- L - set of all lines
- $l = \{h_1, h_2, \dots, h_m\}$ - line

Formulation

The algorithm begins with the initialisation, where every item value of matrix O is added to the weight of every edge included in the shortest path from the node representing current row to the node representing current column. The graphical representation of the resulting edge weights (where the width of the line representing a given edge is proportional to the edge weight) shows passengers movement on the network. Such a map can help to select the set of attraction centres C as another algorithm input. Attraction centres are towns, which are incident with the biggest passenger flows - the widest lines. The described process can be formulated as follows:

[Initialisation - step 1] Assign transportation demands described by matrix O to graph G edges:

- For every edge $h_i \in H$ set $c_i = 0$.
- For every pair $(r, s) \in Q$: calculate shortest path $R_{r,s}$, for every edge $h_i \in R_{r,s}$ set $c_i = c_i + p^{rs}$.

The core of the algorithm can be described by following steps:

[Paths creation - step 2]

Let u denotes edge h_j with maximal weight out of all edges, that begin in an attraction centre: $u = h_j: c_j = \max\{c_i: h_i = (v, w), v \in C, w \in V\}$. Initialise constructed line $l = \{u\}$.

If the weight of edge u is not positive (greater than zero), go to *Finish*.

[Paths continuation - step 3]

Let u_1 denotes edge h_j , with maximal weight out of all edges, that begin in terminal node w of edge $u: u_1 = h_j: c_j = \max\{c_i: h_i = (w, k), k \in V, h_i \notin l\}$.

If such an edge does not exist, go to *Finish*.

Insert edge into constructed line $l: l = l + \{u_1\}$.

Check, if in range of d_d around terminal node k of edge u_1 exists node t , that satisfies condition $p(t) > 0$. If such a node exists, and edge the node was reached by does not have the greatest weight of edges outgoing from node k , and edges of the shortest path from node k to node t do not belong to the constructed link, add edges of the shortest path from node k to node t and edges of the backward path from node t back to node k to the edges of line l .

If the terminal node of edge u_1 belongs to the set of attraction centres, go to *Finish*.

If line l is longer than the shortest path from the initial line node to terminal node k of edge u_1 by more than d_e , go to *Finish*. $u = u_1$. Go to step 3.

Add line l to set of lines $L: L = L + \{l\}$.

Subtract minimal weight out of line l edges from the weights of all edges of line l : for every edge $u_1 \in l$ set $c_i = c_i - \min\{c_j: u_j \in l\}$.

Go to *Initialisation* step 1.

[Finish] - end of algorithm

5. Results

The algorithm was tested on real data of Nitra and Zilina counties for a multimodal infrastructure consisting of the railroad and

road networks. Infrastructure data of Zilina county were defined by 774 railroad and 4644 road edges, 315 towns and villages, and 12 attraction centers. Data of Nitra county were specified by 852 railroad and 8887 road edges, 354 towns and villages, and 9 attraction centers. Input data about transportation flows were derived from data of railroad and bus ticket sales of a respective county.

The algorithm was used for both counties to generate a set of line routes. Manual corrections were done for lines that were routed using unwanted edges. There are various reasons for such corrections. Either input infrastructure data are not precise enough or did not correspond to real conditions (for example a road was not in an appropriate condition for bus traffic) or edge included correctly in infrastructure as feasible was excluded after transportation provider's revision for some operational or other limits. Other reasons for additional corrections may include routing a line beyond town borders, or a case that settled area is separated from a transportation system and no standard infrastructure link exists to connect this area to the transportation system (for instance settlement is separated by a river with no road bridge and only a bridge for pedestrians connects a bus stop with residential areas which is not included in infrastructure data).

Differences against actual line routes were consulted with public transportation planners of the mentioned counties and were admitted as one of possible solutions according to transportation service provider's point of view as well as to passenger preferences.

A set of lines with the following characteristics was generated for Zilina county:

- 113 bus lines,
- 7 train lines,
- vehicle-kilometers calculated for one year (Table 1).

Table 1

| Vehicle-kilometers [1000. km/year] | Working days | | Weekends and holidays | |
|---------------------------------------|--------------|-----------|-----------------------|-----------|
| | Train | Bus | Train | Bus |
| Local transportation | 3942,642 | 44378,650 | 1390,652 | 19597,552 |
| Inter center transportation | 21146,290 | 21877,930 | 6958,764 | 10069,278 |

A set of lines with following characteristics, was generated for Nitra county:

- 185 bus lines,
- 24 train lines,

- vehicle - kilometers calculated for one year (Table 2).

Table 2

| Vehicle-kilometers [1000. km/year] | Working days | | Weekends and holidays | |
|---------------------------------------|--------------|-----------|-----------------------|-----------|
| | Train | Bus | Train | Bus |
| Local transportation | 1578,882 | 58910,164 | 1010,858 | 22086,892 |
| Inter center transportation | 5457,458 | 9444,058 | 1407,68 | 3768,97 |

6. Comparison

There are two different points of view to evaluate proposed line routing by:

- government or delegated agency providing basic transportation service
- passengers.

As the discussed algorithm delivers only line routes, it is not possible to evaluate costs for transportation companies providing services on proposed lines. Even a comparison of estimated vehicle-kilometers from Table 1 and Table 2 with current transportation system vehicle-kilometers is not correct, while the sum of lines in a current set of lines serving respective area is about 3 times higher than number of lines proposed by the algorithm (as discussed below).

Passenger's top priority is usually to spend least time in a mass transportation system, which corresponds with the length of transportation connection. This can be quantified for a whole plan as a sum of distances needed for every transportation demand to reach the destination (passenger-kilometers). Another priority is to minimize number of transfers between different lines. To calculate mentioned sums, passenger behavior must be simulated. To simplify simulating passenger's decisions which connection to choose to reach destination, pattern behavior was created as follows:

- for the purpose of calculating minimal number of transfers needed to reach destination, passenger chooses a connection with least number of transfers needed.
- for the purpose of calculating minimal person-kilometers, passenger chooses a connection with shortest path to reach the destination.

Characteristics in Table 3 compare person-kilometers, transfer counts and line counts in a current transportation system with proposed line plan. Results were evaluated by assuming above mentioned pattern behaviour serving each travel demand.

Table 3

| | Number of transfers | | Person- kilometers | | line count | |
|----|---------------------|---------|---------------------|-------------|---------------------|------|
| | Current real system | Plan | Current real system | Plan | Current real system | Plan |
| NR | 1731000,998 | 1692424 | 1493862,153 | 1612937,267 | 574 | 190 |
| ZA | 2451155,470 | 2532336 | 2079109,163 | 2459013,813 | 576 | 158 |

Calculation of above mentioned characteristics brings up a variety of problems. First of all, it is selection of a set of lines considered to be lines serving transportation demand in a relevant region from a complete set of offered transportation lines. It is impossible to precisely decide which lines do and which don't serve the region. Passengers choose their transportation route by its speed and so a long distance line can't be excluded from the set of lines serving the relevant region if it leads its way through the region and stops at places of origin and destination of passenger's journey. Such lines are primarily designed to serve travel demands in a higher transportation system such as long distance transportation system, interstate transportation system, etc. Lines of higher level transportation system are designed to connect its stops by shortest or fastest route, so it might not visit local centers along its route, and connect towns, that are not connected by the local transportation system. The line plan designed by above mentioned algorithm has a star topology (lines are designed from center following travel demand flows ending in terminal nodes, where no further travel demands are requested) and long distance or interstate (tangential) lines connect towns on different spikes of the star topology, so it improves local transportation services and tangential travel demands are served more efficiently. Such lines give preference to the current transportation system over the proposed line plan while tangential lines serve some travel demands quicker (shorter path) with fewer line transfers. The discussed situation is illustrated in Fig. 1.

However above discussed long distance (tangential) line is a connection designed for quick transportation between bigger centers and it's a part of line plan of a higher level transportation

system. The line plan designed by algorithm is designed to serve only a limited area of a county and doesn't include any connection to the surrounding areas. That is the reason, why the proposed plan doesn't include lines that would connect towns on different lines. The proposed line plan was designed as a basic transportation service for transportation to schools, nursery schools, work, administration offices, health-care or shopping centers. All mentioned destinations can be summed into the term attraction center as mentioned in the introduction.

The proposed line route plan for Nitra county (NSK) needs approx. 3% less line transfers to service all transport demands comparing to the current real transportation system, while 7.4% more person-kilometers are needed. This means that passenger in the proposed line plan for NSK will need to transfer lines less often while riding longer than in the current transportation system in average.

The proposed plan for Zilina county (ZSK) needs approx. 3.3% more transfers to service all transport demands comparing to the current real transportation system, while 18.4% more person-kilometers are needed. This means that passenger in the proposed line plan for ZSK will need to transfer lines more often while riding longer than in the current transportation system in average.

Difference in number of line transfers between the proposed line route plan and current transportation system is within 3.3% for tested areas, which confirms that the algorithm gives results comparable with the real system even without the support of lines

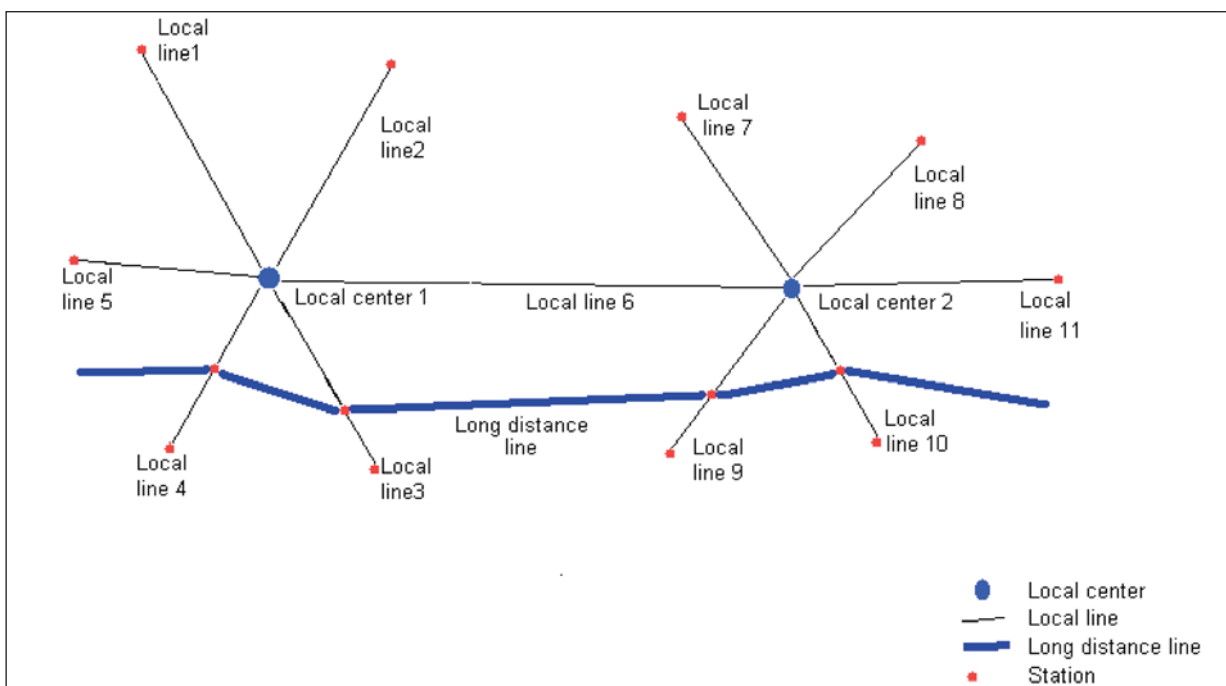


Fig. 1 demonstration of tangential line

from the higher level transportation system. Higher amount of person-kilometers for both tested areas as well as higher amount of line transfers in ZSK in the proposed plan compared to the current transportation system can be explained by a higher number of lines in a current set of lines serving the relevant area than the number of lines in the proposed line plan. In both tested regions the set of selected lines of the current transportation system contained approx. 3-times more lines than the proposed plan.

7. Conclusions

The results of the heuristical line planning algorithm in terms of a number of line transfers for service of transportation demands in a region are comparable with results of the current transporta-

tion system. Differences in person-kilometers in the current and proposed line plans are caused by absence of lines of a higher level transportation system (long distance lines, interstate lines) in the proposed plan, as they are not designed to serve primarily local transportation demands. The very same algorithm can be used to design a line routing plan for a higher level transportation system, which would give a set of transit lines. Completing the proposed line route plan with a set of transit lines would make the results comparable with a selection of lines from the current line plan, however we can allege that the heuristical line planning algorithm gives solutions of a comparable quality to the current transportation planning with a perspective of obtaining better results with a more detailed description of a set of transportation services offered in a region.

References

- [1] CERNA, A., CERNY, J.: *Teorie rizeni a rozhodovani v dopravnich systemech [Management Theory and Decision-Making in Traffic Systems]*. Praha : Institut Jana Pernera, 2004, ISBN: 80-86530-15-9
- [2] QUAK, C. B.: *Bus Line Planning - A Passenger-Oriented Approach of the Construction of a Global Line*, 2003 [online]. Delf University of Technology : Delf : Netherlands [Cit. 2008-10-21]. <<http://www.st.ewi.tudelft.nl/~roos/buslineplanning.pdf> >
- [3] FAN, W., MACHEMEHL, R. B.: Optimal Transit Route Network Design Problem, *J. of Transportation Engineering*, ISSN 0733-947X CODEN JTPEDI. 2006, vol. 132, No.1, pp. 40–51, American Society of Civil Engineers, Reston, VA
- [4] CENEK, J.: Line Routing Algorithm, *J. of Information, Control and Management Systems*, vol. 8, No. 1.