TRANSPORTATION NETWORK REDUCTION

Summary. Network reduction problem is formulated as follows: We are given a transportation network \( T \), a set of important origin – destination relations \( R \) and a number \( q \) greater than 1. The goal is to find a subnetwork \( S \) of the given network \( T \) such that all shortest paths between all origin – destination pairs from \( R \) using only subnetwork \( S \) are not longer than \( q \)-multiple of the corresponding distance in original network \( T \). A mathematical model and an exact algorithm of just mentioned task is presented.

1. INTRODUCTION

Designers of transportation systems are often face with the following problem. Transportation processes takes place on a given transportation network. Transportation demand is defined by an origin destination matrix – OD matrix – representing the quantity of transportation flow which has to be transported from every origin to every destination. Some origin destination relations are more important than others – obviously most important relations are those with largest transportation flows. Transportation has to be provided by several transportation lines – e.g. tram lines, trolley bus lines, bus lines etc. Transportation lines cannot be build on all segments of transportation network. Therefore only a part of the given network – a subnetwork can be chosen for providing transportation service. The chosen subnetwork has to have the property, that all important origin – destination paths using only chosen subnetwork are not longer than the \( q \)-multiples of corresponding distances in the original network. Problem was first formulated by J. Černý, several heuristic results were published by Černá, Černý, Zimmermann and Peško in [1 - 3] and [5], Zimmermann proved, that network reduction problem is NP-hard.
2. MATHEMATICAL MODEL OF NETWORK REDUCTION PROBLEM

Essential elements of a real transportation network are arc vertices (important points and crossings) and communications between adjacent vertices. Every communication can be one way or two way street. If all communications in the studied transportation network are two way streets then an undirected graph can be used as an appropriate mathematical model. However, we can use also a directed graph in which a two way communication is represented by two arcs with the same end points and with opposite orientation.

Let $G = (V, A, c)$ be a digraph where the vertex set $V$ represents the set of all important points and all road crossings of the studied transportation network, the arc set $A$ is the set of all road communications between adjacent points resp. crossings and where $c: A \rightarrow \mathbb{R}$ is a real function representing the length of the corresponding road communication. Suppose that we are given a set $R$ of important transportation relations, i.e. $R$ is the set of ordered pairs $r = (o_r, d_r)$ of the type origin, destination. In practice, $R$ can be determined using an O – D matrix or by an expert who knows from where to where most intensive transportation streams flow.

The goal is to determine the cheapest subset of transportation network – the cheapest transportation subnetwork – which will be used for providing transport from $o_r$ to $d_r$.

Precise formulation for “the cheapest subnetwork” is “the subgraph $G' = (V', A', c)$ of $G = (V, A, c)$ with the smallest total arc cost $\sum_{a \in A'} c(a)$ in which the length of the shortest $(o_r, d_r)$-path in the subnetwork $G'$ should not be greater than the $q$-multiple of the length of the shortest $(o_r, d_r)$-path in the subnetwork $G$ for every $r \in R$, where $q$ is a given fixed number, $q > 1$.”

Denote:

- $n$ – the number of all vertices in the digraph $G = (V, A, c)$, i.e. $n = |V|$
- $c_{ij}$ – if $(i, j) \in A$, $c_{ij} = c(i, j)$ – the cost of the arc $(i, j)$, otherwise $c_{ij} = \infty$
- $R$ – the set of important relations – the set of important $(o_r, d_r)$ pairs
- $o_r, d_r$ – origin and destination of the relation $r$
- $\ell_r$ – the length of the $(o_r, d_r)$-shortest path in original network $G$
- $q$ – a given fixed number, $q > 1$
- $x_{ij}^{(r)}$ – a decision variable equal to $x_{ij}^{(r)} = 1$ if the arc $(i, j)$ lies on a $(o_r, d_r)$-path in the subnetwork $G'$, otherwise $x_{ij}^{(r)} = 0$
- $y_{ij}$ – a decision variable, $y_{ij} = 1$ if the arc $(i, j)$ is chosen into subnetwork $G'$, otherwise $y_{ij} = 0$

The total cost of all chosen arcs can be expressed as $\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} y_{ij}$. Therefore the problem of transportation network reduction can be formulated as follows:

Minimize
\[ \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} y_{ij} \]  \hspace{1cm} (1)
Subject to
\[ \sum_{j=1}^{n} x_{ij}^{(r)} = 1, \quad \sum_{i=1}^{n} x_{ij}^{(r)} = 1 \]  \hspace{1cm} (2)
\[ \sum_{i=1}^{n} x_{ik}^{(r)} = \sum_{j=1}^{n} x_{kj}^{(r)} \]  \hspace{1cm} (3)
\[ \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij}^{(r)} \leq q \ell_r \]  \hspace{1cm} (4)
\[ x_{ij}^{(r)} \leq y_{ij} \]  \hspace{1cm} (5)
\[ x_{ij}^{(r)} \in \{0,1\} \]  \hspace{1cm} (6)
Transportation network reduction

Conditions (2) and (3) characterize an \((a_r, d_r)\)-path for every relation \(r\). Conditions (2) say that only one arc comes from origin \(a_r\) and only one arc enters into destination \(d_r\) for every relation \(r\). Conditions (3) say that the number of arcs entering into the vertex \(k\) is equal to the number of arcs coming out from the vertex \(k\). Conditions (4) ensure that every shortest \((a_r, d_r)\)-path in reduced network \(G'\) is not longer than the shortest \((a_r, d_r)\)-path in original network \(G\).

Let us notice that none of constraints (2) – (6) requires that \(y_{ij} \in \{0,1\}\). However, constraints (5) together with the fact that all coefficients \(c_{ij}\) in objective function (1) are positive and (1) – (6) is a minimization problem ensure that \(y_{ij} \in \{0,1\}\) (namely \(y_{ij} = 1\) if there is a \(r\) such that \(x_{ij}^{(r)} = 1\) and \(y_{ij} = 0\) if it holds \(x_{ij}^{(r)} = 0\) for all \(r \in R\)).

Mathematical model (1) – (6) is important from a theoretical point of view, however large number of variables \(|V|^2 |R| + |V|^2\) and large number of constraints reduce its application in practice to very small instances. Moreover, theoretical results and practical experiences show that the number of arcs in digraphs created by surface transportation systems as roads with relations is much higher than the number of nodes. Nevertheless, in the air transport you would not use a network reduction algorithm, as you can always create a new connection between two airports.

Therefore we have developed the following procedure.

Let \(m = |A|\) be the number of arcs of digraph \(G = (V, A, c)\). Let
\[
A = \{a_1, a_2, ..., a_m\}, \quad c_j \text{ be the cost of the arc } a_j \text{ for } j = 1, 2, ..., m.
\]

We can compute a set of \(k_r\) shortest paths not longer than \(q\ell_r\) for every relation \(r \in R = \{r_1, r_2, ..., r_{|R|}\}\). Palúch’s multilabel algorithm [4] can be used of computing sets of \(k_r\) shortest paths. Let us enumerate just found shortest paths as follows:

\[
I_1 = \{1, 2, ..., k_1\}, \quad I_2 = \{k_1 + 1, k_1 + 2, ..., k_1 + k_2\}, \quad I_3 = \{k_1 + k_2 + 1, k_1 + k_2 + 2, ..., k_1 + k_2 + k_3\},
\]

\[
I_{|R|} = \{\sum_{i=1}^{|R| - k|R|} k_i + 1, \sum_{i=1}^{|R| - k|R|} k_i + 2, ..., \sum_{i=1}^{|R| - k|R|} k_i + k_{|R|}\} - k_{|R|} \text{ shortest paths for relation } r_{|R|}.
\]

It is convenient to denote by \(P\) the total number of just created shortest paths – \(P = \sum_{i=1}^{|R|} k_i\). The main idea of our procedure is to choose exactly one path from every set \(I_1, I_2, ..., I_{|R|}\) such that the total cost of all arcs contained at least in one \(I_r\) (where \(r = 1, 2, ..., |R|\)) is minimal. Then the set of all chosen arcs will create the arc set of the reduced network.

Define constants \(b_{ij}\) for \(i = 1, 2, ..., |P|\) and for \(j = 1, 2, ..., m\) as follows
\[
b_{ij} = \begin{cases} 1 & \text{if the path } i \text{ contains the arc } j \\ 0 & \text{otherwise} \end{cases}
\]

Let \(z_i\) for \(i = 1, 2, ..., P\) be a binary decision variable with following meaning
\[
z_i = \begin{cases} 1 & \text{if all arcs of the path } i \text{ are selected into the reduced network} \\ 0 & \text{otherwise} \end{cases}
\]

Denote by \(y_j\) for \(j = 1, 2, ..., m\) a binary decision variable
\[
y_j = \begin{cases} 1 & \text{if arc } j \text{ is selected into the reduced network} \\ 0 & \text{otherwise} \end{cases}
\]
Then the total cost of all chosen arcs will be equal to \( \sum_{j=1}^{m} c_j y_j \).

Now we can write the following mathematical model:

\[
\begin{align*}
\text{Minimize} & \quad \sum_{j=1}^{m} c_j y_j \\
\text{Subject to} & \quad \sum_{j \in I_r} z_j = 1 \quad \text{for all } r \in R \\
& \quad b_{ij} z_i \leq y_j \quad \text{for all } i = 1,2,\ldots,P \\
& \quad \text{and for all } j = 1,2,\ldots,m \\
& \quad z_i \in \{0,1\} \quad \text{for all } i = 1,2,\ldots,P
\end{align*}
\]

(7) (8) (9) (10)

Conditions (8) ensure that for every relation \( r \in R \) exactly one path is chosen from the set \( I_r \) of precomputed shortest paths. Constraints (9) ensure that \( y_j=1 \) if there exists at least one chosen path containing the arc \( a_j \). Obligatory constraint makes sure that \( z_i \) is a binary variable. Remember that no similar condition for variables \( y_j \) are necessary since from the fact, that \( z_i \) is binary together with (9) and (7) it follows, that \( y_j \) can acquire only values from the set \( \{0,1\} \).

Remark. The number of constraints (9) is \( P \cdot m \). We can construct an equivalent mathematical model by replacing conditions (9) by

\[
\begin{align*}
\sum_{i=1}^{P} b_{ij} z_i & \leq L y_j \quad \text{for all } j = 1,2,\ldots,m \\
y_j & \in \{0,1\} \quad \text{for all } j = 1,2,\ldots,m
\end{align*}
\]

(11) (12)

where \( L \) is a large number (it suffices \( L = P \)). Mathematical model (7), (8), (10), (11), (12) has number of constraints reduced. However it depends on concrete solver which model is more advantageous. Remember that if we create the model (7)-(10) for all \( (o_r,d_r) \)-shortest paths not longer than \( q \ell_r \) for every relation \( r \in R = \{r_1,r_2,\ldots,r_{|R|}\} \) we obtain exact solution of network reduction problem.

3. COMPUTER RESULTS

The computations via model (7)-(9) have been carried out on a PC with OS Linux, 8-core XEON, 3 Ghz, RAM 16 GB. We used the Python API for Gurobi MILP solver [8].

We use a subset of Slovakia maps data of region Žilina as input data in our experiments. We use data in OSM format downloaded from [6] and published under Open Data Commons Open Database License (ODbL) [7]. Our model of transportation network contains 2664 vertices and 3411 arcs. We placed 73 bus-stops into some vertices of directed graph and simulate transportation flow between pairs of those bus-stops. Model of transportation network is shown on Fig. 1.

We chose at random 50 and 200 relations (different pairs of bus-stops) to simulate two different transportation demands and search 10, 20, 30, 40 and 50 shortest paths between those pairs of bus-stops. Greater number of shortest paths means greater parameter \( q \) (ratio between shortest paths in reduced and full model of transportation network). Computation of our largest problem in Gurobi solver [8] did not exceed 30 seconds. All results are shown in Tab. 1. Reduced transportation network for 200 relations is shown on Fig. 2.
Fig. 1. Model of transportation network in region Žilina, Slovakia
Bild 1. Verkehrsnetzmodell im Region Žilina, Slowakei

Table 1

<table>
<thead>
<tr>
<th>Number of shortest paths</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 relations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time [s]</td>
<td>0.131</td>
<td>0.267</td>
<td>0.435</td>
<td>1.065</td>
<td>1.845</td>
</tr>
<tr>
<td>Cost [m]</td>
<td>55856</td>
<td>54593</td>
<td>54358</td>
<td>54187</td>
<td>53826</td>
</tr>
<tr>
<td>200 relations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time [s]</td>
<td>0.454</td>
<td>2.493</td>
<td>3.096</td>
<td>16.147</td>
<td>20.808</td>
</tr>
<tr>
<td>Cost [m]</td>
<td>78291</td>
<td>77220</td>
<td>76729</td>
<td>75371</td>
<td>74944</td>
</tr>
</tbody>
</table>

Fig. 2. Reduced transportation network from problem with 200 relations
Bild 2. Reduziertes Verkehrsnetz für das Problem mit 200 Relationen
4. CONCLUSION

Computer experiments showed that our method based on mathematical model (7) – (10) is applicable to large real world contrary to first model (1) – (6). Another advantage of our approach is that a person can control the process of creating every set $I_r$ of shortest paths in a man – machine process and he can exclude some paths inconvenient from a reason which cannot be simply modelled by mathematical or computer tools.

Acknowledgement

The research was supported by the Slovak research grants VEGA 1/0296/12 „Public Service Systems with Fair Access to Service“, APVV-0760-11 „Designing of Fair Service Systems on Transportation Networks“ and by the Czech Science Foundation Grant No. P 402/12/2147 „Economically Optimal Processes on Networks“.

Authors are thankful to the unknown reviewer for carefully reading this paper, recovery several unpleasant errors and several useful suggestions.

References

7. Open Data Commons Open Database License (ODbL). Available at: http://opendatacommons.org/licenses/odbl/.

Received 23.11.2013; accepted in revised form 02.06.2015