Multi-objective optimization of multimodal transportation networks
Interpretation of the Pareto set from a case study in Amsterdam

TRAIL Research School, October 2012

Authors
Ties Brands MSc. and Prof. dr. ir. Eric van Berkum
Centre for Transport Studies, University of Twente, the Netherlands

© 2012 by T. Brands, E.C. van Berkum and TRAIL Research School
Abstract

We define the optimization of infrastructure planning in a multimodal network context as a multi-objective network design problem, rather than evaluating a pre-defined set of network scenarios. This provides insight into the extent to which facilitating better transfers between modes can contribute to various aspects of sustainability, namely accessibility, operation subsidies, use of urban space and climate impact. For a real life case study the Pareto set is estimated by a genetic algorithm, showing that minimizing the use of urban space clearly competes with minimizing operations subsidies. Furthermore, travel time and climate impact are rather in line with each other. Finally it is shown that the Pareto set is strongly influenced by the frequency of one specific train line, indicating that increasing line frequency more effective than opening new park and ride facilities or new train stations.

Keywords

Multimodal networks, multi-objective optimization, genetic algorithm, multimodal modelling, sustainability
1 Introduction

Highly urbanized regions in the world nowadays face well known problems in the traffic system, like congestion, use of scarce space in cities by vehicles and infrastructure and the emission of greenhouse gases. A shift from the car to public modes is likely to alleviate these problems, but investments in public transport (PT) infrastructure require large financial recourses. To take more advantage out of the existing PT infrastructure, facilitating an easy transfer from private modes (bicycle and car) to PT modes (bus, tram, metro, train) may stimulate the use of PT, while limiting the budget that is needed. Therefore, we focus on network developments that enable multimodal trips, like opening new park and ride facilities, opening of new train stations or opening new or changing existing transit lines. The car and bicycle network are assumed to be given.

When such transportation network developments are planned, the decision is often based on an evaluation of a few pre-defined scenario’s based on expert judgment, usually using multiple criteria. However, the best from these scenarios is still likely to have room for improvement. Another method is to optimize a network with respect to accessibility, taking bounds into account for externalities, for example an emission reduction target or a budget constraint. This results in one optimal network solution, but does not provide insight in the dependencies between objectives, i.e. the extent to which the objectives are opposed or aligned, neither is information provided on the possibilities to improve the network further if the budget is slightly increased. Another common method is to combine the objectives beforehand using certain weighting factors (e.g. using a weighted sum), where the weighting factors represent the compensation principle between the objectives which policy makers are willing to accept. However, setting these weights is not trivial: if these are determined in advance, uncertainty concerning these weighting factors is not incorporated and the sensitivity of the outcome to these weighting is not known in advance. For these reasons in this paper a multi-objective optimization approach is adopted, which enables us to identify trade-offs between objectives by studying the Pareto optimal set (Coello et al. 2006).

The resulting mathematical problem is known as the multi-objective network design problem, and received a lot of attention in the literature, in many different versions. One subclass of problems is the transit network design problem, which has been studied in various ways, as reviewed by (Guihaire and Hao 2008). This includes greedy algorithms, evolutionary algorithms and design meetings involving expert judgments. In addition to that, the unimodal road network design problem has, for example, been studied in general by (Mathew and Sharma 2006; Chen et al. 2010). Applications in a multimodal context are less common, but they do exists, for example road link capacity and bus routes (Miandoabchi et al. 2011) or pricing of private and public links (Hamdouch et al. 2007).

Section 2 defined the problem in more detail. Section 3 addresses the solution method: it describes the mathematical optimization technique and the way multimodal trips are modelled. In section 4 the case study is introduced and the results for the case are presented. Finally, section 5 contains the conclusions.
2 Problem definition

2.1 Bi-level problem

The transportation network design problem is often solved as a bi-level optimization problem, for example (Viti et al. 2003; dell'Olio et al. 2006; Tahmasseby 2009). In our research, the network design problem is regarded as a bi-level system as well. In our case the problem is discrete. The upper level represents a network authority that wants to optimize system objectives. In the lower level the travelers minimize their own generalized costs in the multimodal network, which results in a stochastic user equilibrium. The upper level reflects the behaviour of the planner, acting as a government, optimizing system objectives. The network design in the upper level interacts with the behaviour of the travelers in the network: the lower level. Each traveler minimizes his or her own generalized costs (e.g. travel time, expense), by making individually optimal choices. This is put into operation by a traffic model, which assumes user equilibrium. This results in a network state (for example travel times and loads) for each solution, from which the objective function values can be derived. This equilibrium is a constraint for the upper level problem.

2.2 Network and demand definition

The network is defined as a directed graph \( G \), consisting of nodes \( N \) and links \( A \). Transportation zones \( Z \) are a subset of \( N \) and act as origins and destinations. Total fixed transportation demand \( D \) is stored in a \([Z \times Z]\) matrix. Furthermore, transit service lines \( L \) are defined as ordered subsets \( A_l \) of \( A \) and transit stations or stops \( S \) are defined as a subset of \( N \).

2.3 Decision variables

Decision variables in this multimodal network design problem are related to transfer facilities or to PT facilities and are defined in table 1. These decision variables can be related to design dilemmas. Park and ride facilities cost money to operate, but increase the service area of PT stations for travelers with a car available. Both new stations as new express train statuses are a typical example of the trade-off between speed for through travelers and the area coverage of the PT system. Finally, increasing the frequency of PT is mainly a trade-off between costs and reduction of waiting time (and thus reduction of travel time).

<table>
<thead>
<tr>
<th>Decision variable</th>
<th>Formulation</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Park and Ride facility at station ( s )</td>
<td>( p_s \in {0,1} )</td>
<td>This binary variable indicates whether it is possible to park the car at a station ( s ). At existing stations with park and ride facility, this variable is fixed to 1. At candidate locations, this variable can take values 0 and 1.</td>
</tr>
<tr>
<td>Existence of station ( s )</td>
<td>( t_s \in {0,1} )</td>
<td>This binary variable indicates whether transit vehicles call at station ( s ) or not. At existing stations this variable is fixed to 1, at candidate locations this variable can take values 0 and 1.</td>
</tr>
</tbody>
</table>
This binary variable indicates whether transit vehicles of express lines call at station \( s \) or not. At existing stations this variable is fixed to 1, at candidate locations this variable can take values 0 and 1.

\[ f_i \in F_i \]

\( F_i \) contains possible values for the frequency of transit line \( l \). Existing transit lines can either be fixed (\( F_i \) contains only 1 element) or free (\( F_i \) contains 2 or more elements). In the latter case 0 may also be included. For candidate transit lines \( F_i \) always contains at least 2 elements, including 0.

For every potential network development, a decision variables is defined in advance. Network developments are only included as a candidate location if spatial and physical constraints are met. The car and bicycle networks are assumed to be fixed. To some extent, predefinition of candidate locations weakens the power of the optimization approach, because it restricts the solutions space when compared to a complete free design of park and ride, stations and transit lines. If a potential good candidate solution is not included in this set, it will never be chosen in the final network design. However, it still explores a much bigger region than the current practice of evaluation of a few scenarios (the feasible set consists of 4.9 billion solutions). This includes different combinations of measures, that may have a bigger benefit than the sum of the benefits of the individual measures. Furthermore, we think that the current transportation network is not likely to change strongly, as infrastructure in developed countries in general developed only gradually in recent years. Finally, high calculation times make a limited size of the decision space desirable.

### 2.4 Objective functions

The values of the objective functions are calculated based on loads and costs in the network, which are stored in link characteristics and in \( |Z| \times |Z| \) matrices. The objectives are operationalized by total travel time, number of car trips to urban zones (to represent use of urban space for parking), \( \text{CO}_2 \) emissions and exploitation costs (see table 2). Investment costs are not considered, because the chosen decision variables typically involve higher exploitation costs instead of high investment costs. All four objectives are to be minimized.

#### Table 2: Definition of objective functions and list of symbols

<table>
<thead>
<tr>
<th>Policy objective</th>
<th>Measured by</th>
<th>Formulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessibility</td>
<td>Total travel time</td>
<td>( \sum_{ijm} T_{ijn} D_{ijm} )</td>
</tr>
<tr>
<td>Climate impact</td>
<td>( \text{CO}_2 ) emissions</td>
<td>( \sum_{abf} q_{abf} a \delta_{abf} E_{abf}^{\text{CO}<em>2} (\nu</em>{a}) k_{a} )</td>
</tr>
<tr>
<td>Use of urban space</td>
<td>Number car trips to and from urban zones</td>
<td>( \sum_{i \in Z, \ j \in M} D_{ijm} + \sum_{i, j \in Z, \ m \in M} D_{ijm} )</td>
</tr>
<tr>
<td>Cost efficiency</td>
<td>Operation subsidies (operation costs – operation revenues)</td>
<td>( \sum_{b \in B, \ f \in F} [C_b \sum_{a} q_{abf} f_{ab}] - \sum_{b \in B, \ f \in F} [\Delta_{bf} f_b \sum_{a \in A} k_a P_{bf}] )</td>
</tr>
</tbody>
</table>
With:

- $T_{ijm}$: Travel time from origin $i$ to destination $j$ with mode or mode chain $m$ (min)
- $D_{ijm}$: Transportation demand from origin $i$ to destination $j$ with mode or mode chain $m$
- $q_{ab}$: Flow on link $a$ for vehicle type $b$ (veh per hour)
- $\delta_{ad}$: Road type indicator, equals 1 if link $a$ is of road type $d$, 0 otherwise
- $E_{eb}^{\text{CO}_2}(v_{ab})$: CO$_2$ emission factor of vehicle type $b$ on road type $d$, depending on average speed of link $a$ for vehicle type $b$ $v_{ab}$ (grams/(veh*km))
- $k_a$: Length of link $a$ (km)
- $Z_U$: Set of highly urban zones
- $M_O$: Set of modes (including mode chains) that start the trip with a car leg
- $M_D$: Set of modes (including mode chains) that end the trip with a car leg
- $B_{PT}$: Set of vehicle types that are part of the public transport system
- $A_l$: Set of links that are traversed by line $l$
- $C_b$: Exploitation costs for vehicle type $b$ (euros per vehicle*hour)
- $t_{ab}$: Travel time in link $a$ for vehicle type $b$
- $\Delta_{bl}$: PT vehicle type indicator, equals 1 if line $l$ is of vehicle type $b$, 0 otherwise
- $f_b$: Fare for using PT of vehicle type $b$ (euros per km)
- $p_{la}$: Passenger flow in transit line $l$ on link $a$ (passenger per hour)

3 Solution method

3.1 Upper level

The problem is hard to solve and is computationally too expensive to be solved exactly, so we rely on a genetic algorithm. This class of algorithms is often used to solve multi-objective problems, because they do not end up in a local minimum, they do not require the calculation of a gradient and are able to produce a diverse Pareto set (Deb 2001).

More specifically, we use the NSGA-II algorithm, developed in (Deb et al. 2002). NSGA-II has been successfully applied by researchers to solve multi-objective optimization problems in traffic engineering and proved to be efficient for this type of problems (Sharma et al. 2009; Wismans et al. 2011). It is a multi-objective optimization algorithm that optimizes multiple objectives (4 in this paper) simultaneously, searching for a set of non-dominated solutions, i.e. the Pareto optimal set. It is a genetic algorithm, based on the principles of natural selection within evolution, combining solutions to new solutions (crossover), where the solutions with
Multi-objective optimization of multimodal transport networks

A higher fitness value have a larger chance to survive over worse solutions. In the next generation, these enhanced solutions are recombined again, until no progress is made any more or until the maximum number of iterations is reached. Within NSGA-II, the mating selection is done by binary tournament selection with replacement. In addition to this mating process, a random mutation operator is applied to a limited number of solutions from each generation, to promote the exploration of different regions in the solution space. Furthermore, NSGA-II contains elitism, to preserve good solutions in an archive. If the number of non-dominated solutions grows bigger than the archive, the archive only contains the best non-dominated solutions based on the defined fitness value.

The fitness value is calculated in two steps. In the first step (non-dominated sorting), the solutions are ranked based on Pareto dominance. All solutions in the Pareto set receive rank 1. In the next step, these solutions are extracted from the set and all Pareto solutions in the remaining set receive rank 2, etc. In the second step, the solutions are sorted within these ranks based on their crowding distance. Crowding distance calculation requires sorting of the population according to each objective value. The extreme values for each objective are assigned an infinite value, assuring that these values survive. All intermediate solutions are assigned a value equal to the absolute difference in the function values of two adjacent solutions. Concluding, the crowding distance value (and thus the fitness value) is higher if a solutions is more isolated, promoting a more diverse Pareto optimal set. For details on the algorithm, the reader is referred to (Deb et al. 2002).

3.2 Lower level

To be able to assess a multimodal network in a suitable way, a multimodal traffic assignment model is applied in the lower level (see fig. 1). This includes a nested logit mode choice model, a car assignment model and a PT assignment model. In the following section, each component is described.

![Figure 1: Multimodal traffic assignment model used in the lower level with N iterations](image-url)
3.2.1 Modal split
In this approach various combinations of access mode, PT and egress mode are seen as separate modes. These combinations of modes are also called trip chains. Apart from these PT related modes, the car mode is considered as a separate mode. The bicycle is not considered as a separate mode, because the focus here is on interregional trips, where the bicycle can only play a minor role on its own. Therefore, the bicycle trips are not included in the demand data. However, the bicycle is considered to be an important mode as an access and egress mode.

Depending on the costs per mode, a distribution over the modes is calculated by using a nested logit model (Ben-Akiva and Bierlaire 1999). This step splits the total OD matrix $D_{ij}$ into several OD matrices $D_{ijm}$, one for every mode. Within the nested logit model, we use two nests: one for the mode car and one for all trip chains with PT as a main mode. The latter nest contains the trip chains that include walking, bicycle and car as access mode as well as trip chains that contain walking, bicycle and car as egress mode. By doing so, the most attractive PT option mainly competes with the private car. In the case a second, evenly attractive PT mode is added to the choice set, the use of this new option mainly contains former PT travelers, rather than former car travelers.

The costs of a mode are calculated by adding the different cost components to one generalized cost value. In the case of car these costs consist of travel time and of distance to represent fuel costs and other variable costs, for example maintenance costs. In the case of PT these costs consist of travel time, waiting time, transfer penalty and of distance to represent the ticket price.

3.2.2 Car assignment
The car-only trips are assigned to the network using the standard capacity dependent user equilibrium assignment of Frank-Wolfe. The costs of car depend on the flow following a BPR curve.

3.2.3 Public transport assignment
The PT assignment method (including various access and egress modes) includes multiple routing based on the principles of optimal strategies, as developed by (Spiess and Florian 1989), without capacity restrictions.

As indicated before, costs are calculated for several trip chains, consisting of an access mode, PT as main mode and an egress mode. The number of travelers using a certain chain is determined in the modal split step, so for the PT assignment algorithm the access and egress modes are fixed. The route choice algorithm consists of two steps: stop choice and line choice. For stop choice, a set of candidate stops is specified by defining an access mode specific search radius. All stops within this distance are added to the set of candidate stops. Moreover, a minimum number of stops to be found can be set, i.e. if the number of stops within the search radius is smaller than this number, the search radius is extended until the minimum number of stops has been found. Within this set of candidate stops, the distribution of travelers among stops is calculated using formula 1: a multinomial logit model using the generalized costs of the stops in the candidate set.

$$P_s = \frac{e^{-\theta G_s}}{\sum_{s \in S_0} e^{-\theta G_s}}$$

(1)

- $P_s$: fraction of travelers that choose stop $s$
- $G_s$: generalized costs when using stop $s$
- $S_0$: set of candidate stops at the origin
- $\theta$: logit parameters for stop choice
The generalized costs when using a certain stop depends on the distribution of travelers among the transit lines serving that stop. This distribution is calculated by formula 2. It is a combination of a logit model depending on generalized costs of a line alternative and on the frequency of that line. The latter is to model the situation in which a fast vehicle has just departed and the traveler prefers to take the next slower vehicle over waiting for the next fast vehicle.

\[ P_{ls} = \frac{f_s e^{-\lambda G_{ls}}}{\sum_{x \in L_s} f_s e^{-\lambda G_{ls}}} \]  

\( P_{ls} \): fraction of travelers that choose line \( l \) at stop \( s \)
\( f_s \): frequency of line \( b \)
\( G_{ls} \): generalized costs when using line \( l \) at stop \( s \)
\( L_s \): set of candidate lines at stop \( s \)
\( \lambda \): service choice parameter

So concluding, the PT demand is assigned to the PT network, using multiple routes and multiple access and egress modes, without capacity restrictions. For every combination of access and egress modes a separate OD matrix is taken as input.

4 Case study

The optimization framework is applied to a case study in the Amsterdam metropolitan area, which covers a large part of the Randstad (fig. 2). It contains a detailed multimodal network, including bicycle links, car links, transit lines (including distinction between local services and express services). This enables a detailed modeling of the trip chain. On the other hand, the number of zones is limited, to ensure fast calculation times.

Figure 2: the area of the case study
4.1 Decision variables

In the case study, 26 decision variables are defined. The decision variables correspond to the following possible measures:

- 6 candidate locations for new park and ride facilities
- 3 candidate locations for new train stations
- 4 candidate locations for express train status of train station
- Frequency of transit lines, from which
  - 5 train lines
  - 6 bus lines
  - 2 candidate locations for tram line extension

The remaining of the network is unchanged. This includes the networks of access and egress modes as well as the car network.

4.2 Results

The resulting Pareto set gives insight in the interdependencies and tradeoffs between objectives. Furthermore, decision variables can be identified that strongly influence the result. In total 2960 solutions are calculated during the execution of the algorithm. From these solutions, 339 appear to be non-dominated or Pareto optimal with respect to the set of calculated solutions. Note that these Pareto set is an approximation of the real Pareto set, since it is impossible to calculate all solutions and thus the true Pareto set is not known.

4.2.1 Trade-offs

The plots in figures 3 to 5 show 2 dimensional plots of the 4 dimensional Pareto set. Each plot shows 2 objective functions and every solution in the Pareto set is plotted by one dot based on the corresponding objective values. Note that each plot contains dominated solutions if we look at only those 2 objectives, but in the 4 dimensional space the plots only contains Pareto solutions.

![Figure 3: relation between urban space used and operation costs](image)

Figure 3: relation between urban space used and operation costs
Figure 3 shows the relation between the amount of urban space used by (parking of) cars and the operation costs of the PT and park and ride system. A clear relation can be identified: if the number of cars in the city centre is to be decreased, the PT system needs a higher budget. In the lower end of the graph it can be observed that the first reduction from 79 to 77.5 thousand urban cars trips can be achieved more easily than the second reduction from 77.5 to 76.8 thousand trips.

Figure 4 shows the relation between travel time and climate impact, i.e. CO$_2$ emission. These two objectives are roughly in line with each other: it is very well possible to minimize both climate impact and travel time simultaneously. In the multimodal context this makes sense, because using the multimodal decision variables in this paper, reducing travel time is possible by improving the PT system, which causes a modal shift from car to PT, achieving smaller CO$_2$ emissions. However, in the left lower corner if the figure it can be seen that still a tradeoff exists between the objectives, which might be caused by solutions with such high PT quality that the additional CO$_2$ emissions of the transit vehicles are not compensated any more by modal shift from car.
Finally, figure 5 shows the relation between travel time and PT operation subsidies. Here the relation is less clear, although a weak negative relation can be identified. These 2 objectives interact in a complex way for two reasons. Firstly, a better PT system increases costs, but if it is very beneficial for the travelers the PT revenues also increase and ultimately PT subsidies may decrease. Secondly, increasing the quality of PT cause a shift to PT modes, but due to the used logit model as a mode choice model, this shift already slightly takes place when the PT option is still slower than the car option, eventually increasing total travel time.

4.2.2 Important decision variables

Figure 6 shows the Pareto set for the objectives CO$_2$ emission and operation subsidies, distinguishing between solutions that have a frequency of 0 or 2 for one specific train line in the network and solutions that have a frequency of 4 or 6 for that train line. When looking at the Pareto set, 2 clusters can be identified based on the value of the corresponding decision variable, namely the frequency of this main train line. Apparently, the frequency of this train line strongly determines the shape of the set. This clearly shows that a high frequency on this line is beneficial for CO$_2$ emissions and a low frequency on this line is beneficial for lower operation subsidies. However, some overlap between the two categories can still be seen: within this overlap area other decision variables can still play a distinctive role with respect to the values for operation subsidies and CO$_2$ emission.
Conclusions

In this paper an optimization framework is set up to design a multimodal transportation network, based on a predefined set of candidate locations for network developments. The framework is applied to a real world case study with 26 decision variables. The result of the optimization process is an estimation of the Pareto optimal set, consisting of 339 possibly optimal solutions.

Analysis of this Pareto set leads to the following observations. Firstly, a strong trade-off exists between the use of urban space by cars and operation subsidies: a high quality, but expensive public transport system is needed to attract former car users. Secondly, it is possible to reduce travel time while also reducing climate impact. This is caused by the multimodal decision variables, that promote the sustainable mode of public transport, by improving the PT network. Travel time and operation subsidies do not show a clear relationship, probably because solutions with low travel times do not necessarily include high shares of PT: the car is able to provide reasonable travel times as well. Finally it is shown that the Pareto set is strongly influenced by the frequency of one specific train line, indicating a key role for that train line in the network of this case study. This may indicate that the quality of the public transport network itself is more important to achieve sustainability objectives by promoting multimodal trips than the quality of the transfer facilities. The low score for new train stations indicates that, based on the objectives in this case study, higher speeds for through travellers are to be preferred. However, further analysis of the influence of other variables on the objectives is desired in the future.

Further effort will be put into analyses and methods to make the Pareto set more useful as decision support tool. This includes further identification of effective and ineffective decision variables, interrelation between decision variables and reduction of the Pareto set to a smaller amount of solutions, that is easier to interpret.
Acknowledgements

We thank the SAR program of NWO for funding this research.

References


