

Position Paper: Explainable Search of Multimodal Itineraries

Jan Fabian Ehmke¹, Thomas Horstmannshoff¹

Abstract. Travel websites like Google Maps promise the seamless creation of multimodal itineraries according to individual traveler preferences. The engines of these websites are based on efficient techniques of mathematical optimization and can search multimodal itineraries in negligible runtime as long as the number of traveler preferences is small. However, due to the large number and variety of mobility services, the solution space of multimodal itineraries is often complex. Hence, the search process is not transparent to individual traveler, and it is hard for them to assess the impact of their preferences on the set of created multimodal itineraries. In this position paper, we propose to make the search of individualized multimodal itineraries more explainable. Demonstrating the idea of solution sampling and using a large data set of multimodal mobility services, we derive information on the solution space and demonstrate how these can support travelers in the adaptation of their preferences and in the creation of a set of individualized itineraries.

Keywords: Explainable OR, Multimodal Mobility, Sampling, Customer Empowerment.

1 Introduction

In recent years, the number and variety of mobility services has risen considerably. Especially innovative mobility services such as bike, car and ride sharing services have emerged and complement traditional mobility services such as scheduled trains and buses. Following recent mobility surveys, travelers expect a reasonable choice of itineraries created from their individual preferences such as travel time, price, number of transfers, mode, waiting time and walking distance [1-5]. Since travelers require door-to-door travel information, individual mobility services need to be combined to form a complete itinerary, considering the preferences of individual travelers.

Travel websites such as Google Maps promise the seamless creation of a set of individualized itineraries, but usually remain very simple in discovering and considering individual preferences. The reason is that the search of multimodal itineraries requires quantifications of traveler preferences. Then, these can be considered through constraints of established mathematical models and be solved with efficient solution techniques of multimodal network search. However, state-of-the-art solution techniques are limited in the number of preferences they can consider efficiently. Delling et al. [6]

¹ Management Science Group, Otto von Guericke University Magdeburg, Universitätsplatz 2, 39016 Magdeburg, Germany

highlight that taking more preferences into account at the same time has an enormous impact on the runtime of multimodal path-finding approaches.

With regard to the consideration of multiple traveler preferences, from a mathematical-optimization point of view, hierarchical objective functions are common. These functions optimize for one objective first (e.g. minimize travel times) and focus on secondary objectives subsequently (e.g. minimize the number of transfers). However, as it is well-known for hierarchical objective functions, depending on how tight the constraints are, there may be either no feasible solution at all or a very large number of solutions [7]. Hence, travelers could face either an empty or an excessively large set of itineraries. Since the search is based on an “invisible” mathematical model, travelers cannot interact with the search engine in a reasonable way and often perceive the creation of itineraries as a “black-box” process. However, if travelers knew more – in a simplified and aggregated way – about the characteristics of the solution space, they could interact with the search engine to define their individual preferences and adapt them if necessary. In information systems literature, this extended interaction of customer and search engine is called “customer empowerment” [8].

In this position paper, we propose the idea of the individualized creation of multimodal travel itineraries. Based on efficient mathematical modeling and solution techniques and using a large data set of German mobility services, we demonstrate the potential and limitations of “black-box” mathematical optimization and discuss how to include more complex traveler preferences. With solution sampling, we present a simple idea to identify characteristics of the solution space and empower travelers to restrict or enhance their preferences interactively.

2 Multimodal Itinerary Creation

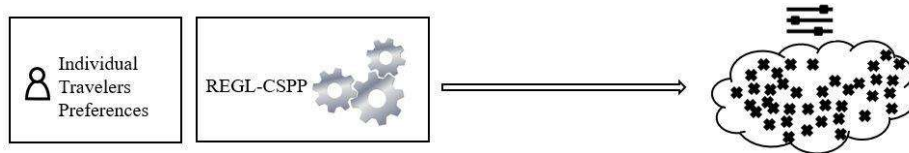


Fig. 1. Common way of creating multimodal travel itineraries

Figure 1 illustrates the common way of searching multimodal itineraries. Given formalized traveler preferences, a multimodal routing algorithm such as REGL-CSPP, introduced by [9], is run to determine a probably extensive number of multimodal itineraries (each feasible itinerary is indicated by a cross). It is then the traveler’s burden to dynamically reduce or extend the retrieved set of itineraries by using filtering and sorting techniques without having a clear idea about the characteristics of the solution space. As these are hardly transparent for the traveler, it is usually quite difficult for them to restrict or extend the set of itineraries systematically.

We propose a data-driven decision support system that uses *solution sampling* in conjunction with searching multimodal travel itineraries. Figure 2 visualizes our idea: By systematically re-running a multimodal routing algorithm, we retrieve valuable

meta-information about the solution space characteristics, which can be used to explain the solution space to travelers as well as empower them to reduce/extend the retrieved set of itineraries. This implies identifying promising solutions on the one hand, but also partially infeasible solutions on the other hand, allowing the traveler to reflect on the importance and impact of their preferences. For instance, in case the traveler sets the upper bound for travel time to 4 hours, a favored solution taking 4 hours and 5 minutes might still be relevant. In addition, solution space characteristics can be used to derive typical relationships between preferences. For instance, for some types of itineraries, there might be a linear relationship between travel time and price, which can help the traveler to adjust their preferences systematically.

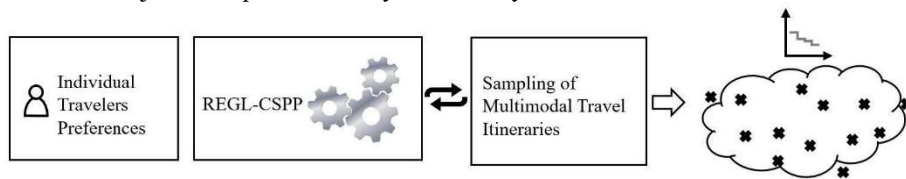


Fig. 2. Sampling of multimodal travel itineraries

In the end, the proposed decision support system could be integrated in a multimodal routing app to enable an explainable search of multimodal itineraries for the traveler. In the following, based on real-world multimodal mobility service data from Germany, we will detail our idea of a more explainable search through solution sampling.

3 Sampling Framework

To guide travelers in their search of multimodal travel itineraries, we focus on retrieving solution space characteristics through solution sampling. Solution sampling builds on the repeated execution of efficient solution techniques for the multimodal search of itineraries (in our case: REGL-CSPP). Our framework consists of four steps:

1. Two-dimensional analysis of each traveler preference
2. Merge of the created sets of itineraries
3. Removal of dominated itineraries
4. Dynamic reduction/extension of the set of itineraries

For step 1, given a request for a specific source s and a target t of an itinerary, for all parameters $i \in I$ with I being the set of quantifiable traveler preferences, we identify “reasonable” parameter intervals $[l_i, u_i]$. The lower bound l_i of the parameter i can be easily determined by running REGL-CSPP without any constraints. Algorithmic details are provided in [10]. The upper bound u_i is derived from the maximum value for the respective parameter after applying REGL-CSPP for a search based on all other parameters under investigation. Then, we search all two-dimensional relationships between all parameter combinations $i, j \in I \mid i \neq j$ within the identified intervals by systematically re-running REGL-CSPP according to a predefined sampling density. For instance,

to identify the two-dimensional solution set $S_{traveltime,price}$, we set the parameter “travel time” as the objective and rerun REGL-CSPP with different upper bounds for the parameter “price” starting with the identified upper bound u_{price} . By iteratively softening the restriction (e.g., by dividing the interval of price into k equally sized buckets) and re-running REGL-CSPP with that specific restriction, we can identify itineraries with different prices that are all optimal with regard to travel time.

For step 2, we merge all itineraries searched based on two-dimensional sets $S_{i,j} \forall i, j \in I \mid i \neq j$ into one comprehensive set S_{all} . Then, in step 3, dominated itineraries are removed. An itinerary s_1 dominates an itinerary s_2 if s_1 is no worse in any parameter $i \in I$ than s_2 [6]. All remaining itineraries with optimal parameter combinations create the Pareto Front.

For step 4, the set of non-dominated itineraries S_{all} is reduced if needed. In case the number of created itineraries in S_{all} is sufficient, i.e. it is in a predefined interval $[c_{min}, c_{max}]$, e.g. $[1, 10]$, the set of itineraries is presented to the traveler in conjunction with detailed information on solution space characteristics. The set of created itineraries can be reduced by applying fuzzy logic approaches to multimodal route planning as demonstrated by [6]. The main idea of fuzzy logic in this case is to identify itineraries from the set of Pareto-optimal solutions that show negligible differences, which are likely not relevant for the traveler. Therefore, undesirable and likely irrelevant itineraries can be dismissed, e.g. an itinerary where a traveler would have to pay significantly more to save only a few seconds of travel time.

4 Exemplary Results

To demonstrate our approach of systematic sampling, we create information about the search space for multimodal itineraries between several major cities in Germany. The experimental setting considers a large amount of data on current mobility services and different modes of transportation as well as innovative ridesharing services such as “BlaBlaCar”. Modelled as a time-expanded network, the entire network for one specific day of operation consists of approximately 6 million nodes and 200 million arcs. As exemplary traveler preferences, we consider travel time, price and number of transfers. We estimate price information for public transit systems, as these are not given in their respective published GTFS data. We analyze long-distance trips between large cities in Germany assuming a start time for the itinerary at 9 am.

As result of solution sampling, Figure 3 shows information on the relationship between the preferences “price” and “travel time” as well as between “price” and “number of transfers” for long-distance trips between Berlin and Munich, Nuremberg and Hamburg as well as Dortmund and Berlin. For price versus travel time, to ensure the comparability, we normalized the parameters between 0 and 1. A travel time of “0” indicates the itinerary with the smallest occurrence for travel time, whereas a travel time of “1” represents the itinerary with the highest occurrence for travel time.

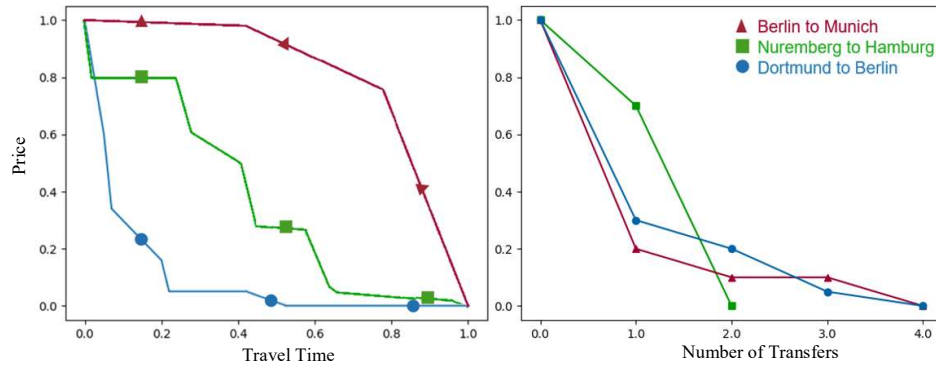


Fig. 3. Exemplary results for three different itineraries

Significant differences in the relationship between the parameters travel time and price can be discovered. For itineraries between Berlin and Munich (indicated by triangles), a traveler has to invest quite a significant amount of time until large cost savings can be realized. In comparison, prices are much lower when investing only slightly more travel time for traveling between Dortmund and Berlin (indicated by circles). Between Nuremberg and Hamburg (indicated by rectangles), a roughly linear relationship between travel time and costs can be discovered.

Minor differences can be seen for the relationship between the parameters “number of transfers” and “price”. For itineraries between Berlin and Munich as well as between Dortmund and Berlin, a traveler has to be willing to transfer at least once in order to achieve substantial price savings. Further transfers do not result in significant savings. For itineraries between Nuremberg to Hamburg, travelers realize the highest savings in case of transferring twice.

The identification of these aggregated relationships between relevant traveler preferences can be communicated to travelers and help guiding them in the solution space. Thereby, we empower travelers to interactively individualize the set of itineraries. For instance, it can be of high relevance for a price-sensitive traveler to have an idea which of their individual preferences has a high impact on the price for the respective itinerary and thereby help assessing which of the preferences should be relaxed/restricted to reduce/extend the set of itineraries under investigation.

Table 1: Results of Solution Sampling

Sampling density	Mode Choice	\emptyset Dominated solutions removed [%]
$k = 5$	All	23.5
	Without car	19.8
$k = 20$	All	89.1
	Without car	87.8

Table 1 compares the impact of different sampling densities (5 and 20) as well as different mode choice configurations. For the latter, either the travelers are willing to use all available modes, or the traveler does not want to drive a car. It can be seen that with a small sampling density of $k = 5$ about 20% of the retrieved solutions can be removed as they are being dominated by another solution. With a larger sampling density, up to 90% of the retrieved solutions are of no relevance for the traveler and hence are removed before showing the results to them.

5 Conclusion and Outlook

In recent years, there has been tremendous progress in the field of searching multimodal travel itineraries. While the search of individualized itineraries for a given and limited set of preferences is possible, this often comes as a black-box search, and it is hard for travelers to interact with the search engine in a way that can help finding more individualized itineraries. We propose the idea of solution sampling to benefit from the state-of-the-art algorithms in network search while deriving characteristics of the solution space that can be helpful to make the search process more explainable to travelers.

Although search engines and mathematical optimization form an important part in finding individualized multimodal itineraries, a more comprehensive framework is needed to provide travelers and customers in general with more “white-box” solutions. Alt et al. [8] present a framework from the information-systems perspective not only suited for the search of itineraries, but also for services from finance, health and education. Embedding mathematical optimization and methods from data mining in this framework can help creating individualized and especially transparent options, combining the power of model-based optimization and data-driven decision making.

References

1. Grotenhuis, Jan-Willem; Wiegmans, Bart W.; Rietveld, Piet (2007): The Desired Quality of Integrated Multimodal Travel Information in Public Transport. Customer needs for time and effort savings. In: *Transport Policy* 14 (1), 27–38.
2. Ehmke, Jan Fabian; Mattfeld, Dirk C.; Albrecht, Linda (2016): Position Paper: Combining Mobility Services by Customer-Induced Orchestration. In: *Proceedings of Rec-Tour@RecSys 2016*.

3. Esztergár-Kiss, Domokos; Csiszár, Csaba (2015): Evaluation of Multimodal Journey Planners and Definition of Service Levels. In: *International Journal of Intelligent Transportation Systems Research* 13 (3), 154–165.
4. Stopka, Ulrike (2014): Identification of User Requirements for Mobile Applications to Support Door-to-Door Mobility in Public Transport. In: David Hutchison, Takeo Kanade und Josef Kittler (Eds.): *Human-Computer Interaction. Applications and Services: 16th International Conference, HCI International 2014, Part 3*, 8512, 513–524.
5. Willing, Christoph; Brandt, Tobias; Neumann, Dirk (2017): Intermodal Mobility. In: *Business & Information Systems Engineering* 59 (3), 173–179.
6. Delling, Daniel; Dibbelt, Julian; Pajor, Thomas; Wagner, Dorothea; Werneck, Renato F. (2013): Computing Multimodal Journeys in Practice. In: David Hutchison, Takeo Kanade und Josef Kittler (Hg.): *Experimental Algorithms. 12th International Symposium, SEA 2013, Proceedings*, 7933, 260–271
7. Ehrgott, Matthias (2005): *Multicriteria Optimization*. 2nd edition. Springer.
8. Alt, Rainer; Ehmke, Jan Fabian; Haux, Reinhold; Henke, Tino; Mattfeld, Dirk Christian; Oberweis, Andreas et al. (2019): Towards customer-induced service orchestration - requirements for the next step of customer orientation. In: *Electron Markets* 29 (1), 79–91.
9. Barrett C., Bisset K., Holzer M., Konjevod G., Marathe M., Wagner D. (2008) Engineering Label-Constrained Shortest-Path Algorithms. In: Fleischer R., Xu J. (eds) *Algorithmic Aspects in Information and Management. AAIM 2008. Lecture Notes in Computer Science*, vol 5034. Springer, Berlin, Heidelberg.
10. Horstmannshoff, Thomas; Ehmke, Jan Fabian (2020): Creation of Individualized Sets of Multimodal Travel Itineraries. In: *Proceedings of 22nd EURO Working Group on Transportation Meeting (EWGT)*, to appear.