Journées SMAI MODE 2018

Stochastic Optimization of Public Transport Schedules

Sofia ZAOURAR-MICHEL
NAVER LABS Europe

Résumé.
We propose a data-driven method to optimize public transport schedules. Using transit data to construct scenarios that reflect the uncertainty of the system, we compute schedules that minimize the expected waiting time during transfers. We model the problem as a two-stage stochastic program and propose two equivalent formulations: a mixed integer linear program (MILP) and a mixed integer quadratic program (MIQP). The MILP version is solved exactly using a generic solver but does not scale well; whereas the MIQP has a partially separable structure that we exploit to design an efficient local search heuristic. We provide results of the two approaches and compare them to the state of the art using transit data collected from Nancy, France.

Mots-clefs : Data-driven optimization, stochastic optimization, mixed integer programming

A major weakness of public transport, as perceived by passengers is the time lost during transfers [1]. This latter can be significant: for example it represents an average of 23% of travel time for multi-modal trips in the UK [2]. In the Operations Research community, numerous works have been done in the domain of schedule synchronization. The goal is to coordinate the timetables of the different lines in order to minimize the waiting time at the connections (see e.g. the surveys [3, 4]). Most of the Operations Research approaches focus on the theoretical timetables and give the same importance to all the connections. If we observe the real usage of the system however, it is clear that some connections have more impact than others because of their frequency and volume of transferring passengers. That is why transportation authorities often incorporate their expert knowledge and experience of the system to design the timetables. However, with the constant growth and sophistication of public transport, these approaches reach their limit.
A natural alternative is to leverage transit data, that is accessible to transportation authorities, to precisely quantify and model transfer waiting times while accounting for the stochasticity of the system. Among the first data-driven approaches to schedule optimization, [5] used the queries to an on-line trip planner to approximate the real usage of the system. The idea was to compute shifts of the schedules that minimize the expected waiting times across a number of scenarios. A scenario is a given realisation of the uncertain parameters of the problem: the number of transferring passengers at each connection and the actual arrival and departure time of the vehicles at each stop. With some simplifying assumptions, [5] modeled the problem as a two-stage stochastic linear program and solved it exactly.
We extended this work in [6] by using the real transit data and taking into account the fact that passengers adapt their behaviour to the change of schedules. This led to a two-stage stochastic linear program with mixed-integer variables (MILP) that we solved exactly using Coin-OR Cbc solver. While these approaches give finer and more realistic models, the size of the problem explodes with the complexity of the transportation network and the number of

considered scenarios. To allow a better scaling, we proposed in [7] an equivalent formulation of the waiting time minimization problem as a mixed-integer quadratic program (MIQP). Although theoretically less appealing than the MILP model, this formulation allows us to exhibit a partial separability property that we exploit to design an efficient heuristic. The idea is that for a given modification of the schedules, the evaluation of its impact on the waiting times can be done in parallel for the different scenarios. Even within a scenario, we show that the evaluation of certain aggregated transfers is separable and can be done in parallel. Moreover, it turns out that the optimization subproblem that we need to solve for the evaluation has a closed-form solution, making the evaluation very fast. We take advantage of these properties to design an efficient parallel local search algorithm that explores random modifications of the current schedules, looking for ones that reduce the waiting time. We compare the three approaches [5], [6] and [7] using transit data collected from the city of Nancy (France). We observe that using the MILP model leads to a significant reduction of the expected waiting time compared to the LP model, but requires much more computing resources. The local search heuristic gives expected waiting times that are only slightly larger than the optima computed by the MILP, while being able to scale to much larger problems.

Références

[1] R. H. Pratt, N. J. Pedersen, J. J. Mather. *Traveler Response to Transportation System Changes: Transit Scheduling and Frequency*. Transportation Research Board, 2004.

[2] R. Gallotti, M. Barthelemy. *Anatomy and efficiency of urban multimodal mobility*. Scientific reports, Nature Publishing Group, 2014.

[3] V. Guihaire, J.-K. Hao. *Transit network design and scheduling: A global review*. Transportation Research Part A: Policy and Practice, 2008.

[4] V. Poorjafari and W.L Yue. *Synchronisation of Public Transport Services: Comparison of Methods*. Proceedings of the Eastern Asia Society for Transportation Studies, 2013.

[5] R. Nair, C. Coffey, F. Pinelli, F. Calabrese. *Large-Scale Transit Schedule Coordination Based on Journey Planner Requests*. Transportation Research Record, 2013.

[6] S. Zaourar-Michel, B. Chidlovskii. *Stochastic optimization of public transport schedules to reduce transfer waiting times*. IEEE International Smart Cities Conference, 2016.

[7] S. Zaourar-Michel. *Efficient Decomposition Method for the Stochastic Optimization of Public Transport Schedules*. NIPS workshop on Machine Learning for Intelligent Transportation Systems, 2016.