Transportation Research Part C 64 (2016) 58–71

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Automatic calibration of agent-based public transit assignment path choice to count data ☆

Manuel Moyo Oliveros *, Kai Nagel

Technische Universität Berlin, Transport Systems Planning and Transport Telematics, Salzufer 17-19, 10587 Berlin, Germany[1]

ARTICLE INFO

Article history:
Received 15 May 2014
Received in revised form 5 January 2016
Accepted 5 January 2016
Available online 2 February 2016

Keywords:
Transit assignment
Micro-simulation
Transit calibration
Multi-agent simulation

ABSTRACT

This work describes the calibration of a schedule-based transit assignment inside an iterative microscopic agent-based simulation. The calibration challenge implies that the behavioral rules should be modified in order to move the simulation closer to observed passenger counts. First, route choice set of agents is enriched with travel parameter utilities randomization. Secondly, the calibration interacts directly into the performance evaluation of individual daily plan of activities, so that the plan is also evaluated for its contribution to the count reproduction. In this way, appropriate plans from the calibration perspective can persist along simulation iterations. The Berlin public transport system with day-based counts is used as test scenario. The results show that the calibration approach can work with large scale scenarios, and that it is able to deal with the inter-temporal aspects implied by counts.

1. Introduction

The implementation of a microscopic agent-based model for public transport simulation represents a helpful instrument for mobility research. If the simulation includes also a travel behavioral approach, it constitutes an advantageous analysis tool for the knowledge of passengers' route decisions and to recognize the effects of those decisions on the whole transport system.

In microscopic models, a special challenge has been calibration. More technically: Given a set of macroscopic observations, how should the physical or behavioral microscopic rules of the agent-based simulation be modified in order to move the simulation closer to the observations? The example considered here is a situation where one has daily passenger counts for certain or all transit lines, demand for 1% of the population from trip diaries, but no route information from the survey. The task is to find passenger connections and possibly modify the passenger demand such that the simulation matches the counts.

This topic belongs to a class of problems that are quite common in agent-based simulations. Agent-based simulations were always built around the notion of "emergence", that is, they are expected to be particularly useful where certain macroscopic properties, in our case congestion, vehicle overloading, and resulting delay patterns, cannot be derived in analytical ways from the microscopic input data (including the behavioral rules), and in consequence one needs to run the simulation

☆ Conference on Agent-Based Modeling in Transportation Planning and Operations.

* Corresponding author. Tel.: +49 30 314 23308; fax: +49 30 314 26269.
E-mail address: manuel.moyo-oliveros@alumni.tu-berlin.de (M. Moyo Oliveros).

[1] http://www.vsp.tu-berlin.de.

http://dx.doi.org/10.1016/j.trc.2016.01.003

in order to obtain them. However, because the connection from input data to emergent properties is by simulation, the mathematical connection is not as well established as in normal numerical modeling.

2. Related works

The use of smart cards, sensors or surveys as data sources for passenger measurements allows the availability of transportation usage information for research purposes (Pelletier et al., 2011; Min et al., 2016). Several works for demand estimation of large scenarios based on passenger counts can be found. Rongviriyapanich et al. (2000) use on–off count data for two bus routes in Tokyo to evaluate OD estimation techniques used originally for road networks. The Entropy Maximization Method is reported as the most practical of all considered techniques, if a prior OD matrix is available. In the same way, Fung (2005) uses boarding and alighting counts of the Hong Kong metro network for station-to-station OD matrix assignment calibration and validation. Random choice function coefficients generation methods in Monte Carlo simulation are also presented. Farrol and Livshits (1998) calibrate a scenario of the Greater Toronto Area based on a survey in 1996. Their results present the adjusted weights to access time, wait time, and the transfer penalty in an EMME/2 implementation. Parveen et al. (2007) present the G-EMME/2 tool to estimate also the Toronto transit network with volumes obtained from on-board surveys. A genetic algorithm represents chromosomes as sets of parameter values generated randomly. The minimization of a "misfit" function optimizes the travel time of the five parameters: boarding time, wait-time factor, wait-time weight, auxiliary time weight, and boarding-time weight. Lu (2008) uses automatically-obtained passenger boarding and alighting counts for a bus line in Columbus, Ohio to review five OD flow estimation methods. The results show that the output quality depends to a great extent on the quality of the base OD matrix. Li (2009) presents the statistical inference of large transit route matrices using on–off counts. Considering a given occupancy of a passenger on a stop, the probability of alighting on a posterior stop of the transit line is calculated with a Markov chain model. A Bayesian analysis draws inference about unknown parameters. Wahba and Shalaby (2011) calibrate the transit scenario of Toronto with the integration of a genetic algorithm engine. Passengers travel experiences are assessed. Thus, the learning mechanism adjusts waiting and in-vehicle time. Moreover, the calibration defines nine parameters related to trip purpose and vehicle type. Ji et al. (2015) present a heuristic expectation maximization model to estimate passenger OD flows using information from onboard surveys. The model is evaluated with empirical comparison to ground-truth observations and with the estimation realized with the iterative proportional fitting estimation procedure. The initial estimation is improved with the inclusion of larger boarding and alighting data vectors obtained with automatic passenger count technologies.

Most demand estimation models focus on aggregated OD flows. That is a disadvantage, considering the research opportunities that can be given with passenger mobility data on an individual level. The present work makes a contribution with a microscopic agent-based approach that addresses the estimation at a fully disaggregated level, catapulting the possibilities of detailed or collective travel behavior analysis.

3. Simulation methodology

An important challenge for the agent-based approach for the recent past has been computational: Finding an implementation that is both close to the agent-based concepts and fast enough for real world scenarios. For transit micro-simulation, the open source framework MATSim (Multi-Agent Transport Simulation) is employed in this work. Its capacity to simulate large scale scenarios and its modular architecture make it appropriate for calibration research. Its simulation approach follows these steps:

- Initialization: Synthetic travelers with activity-based day plans are assumed to be given; see below how they were generated for the present study. This includes generating suitable connections through the public transit network for each trip that is part of the passenger's plan. The calculation of routes and more transit simulation details are described by Rieser (2010). The routing process looks for least compound cost paths with a trade-off between walk time, waiting time, in-vehicle travel time, travel distance, and vehicle change counts. These travel preferences are represented respectively in an economic approach by the parameters presented in Table 1.
- The "effective value" includes the marginal Opportunity Cost of Time, which needs to be added to all "direct" marginal (dis)utilities of time (de Serpa, 1971; Jara-Diaz, 2003; Börjesson and Eliasson, 2014). The opportunity cost of time represents the implicit punishment for not performing an activity. In MATSim, it is approximately given by the parameter called "*performing*", and its default is −6/h. The utilities are taken as dimensionless quantities; −6/h, say, means "minus six utils per hour" or "minus 6/3600 utils per second". Note that the original approach weights all temporal contributions equally.
- Synthetic reality (also known as network loading): All plans are executed by moving agents in the multimodal network simultaneously in a simulation of the physical system ("synthetic reality") which includes a detailed simulation of the public transit system. The events triggered by the simulation are useful for modeling travel incidents such as boarding delays, failures because of overloading, and delays because of late incoming vehicles.