Updating high-res demand model with low-res GSM OD matrices.

Rafał Kucharski[1,*], Justyna Mielczarek[1], Arkadiusz Drabicki[1], and Andrzej Szarata[1]

[1]Politechnika Krakowska, ul. Warszawska 24, 31-155 Kraków, Poland
[*]corresponding author: rkucharski(at)pk.edu.pl

Abstract

In this paper we propose a method to update a high-density yet outdated transport demand model, with a low-density yet actual mobility data. We address the common problem where on one hand the detailed transportation demand model, based on sound, yet limited sample becomes outdated and on the other the low-density big datasets of revealed trips are available (in particular the OD matrices from GSM providers). We propose the method which keeps the rich mobility description from the demand model and fuse it with the demand revealed from the OD matrices. Since building transport demand model is costly and time consuming, data-driven updating methods are needed.

In particular, we use land-use variables, trip generation formulas, structure of trip purposes and mode choice from the classic four-step demand model. We fuse it with two-level destination choice model inferred from revealed OD matrices where we first decompose the trip generation into identified destination strata and then apply the destination choice model within classes. Finally, we tune the trip volumes to match the values in the revealed OD matrices and validate the results against the traffic and passenger counts with a transparent and non-invasive calibration procedure.

We illustrate the method with the case of updating the regional demand model of over Małopolska region (Poland) with over 250 communes. We update it using demand revealed in GSM trip matrices between 20 districts of Małopolska. Thanks to the proposed data-fusion we not only updated the demand model to the actual values, but also enhanced realism of regional mobility patterns reproduced in the model. Specifically, we better captured: long distance trips, attractiveness of main cities, trans-regional mobility, transparency of the modelling procedure. As a result, the costly and outdated regional demand model was updated with dataset available at hand. Proposed method can be applied at the urban scale, where zoning is typically more detailed than BTS stations.

Problem background

Traditionally, travel demand is analysed through the mobility surveys. They rely on face-to-face travel diaries, often improved by using telephones, internet or personal GPS devices. Typically, number of respondents is relatively low, yet for each of them the detailed mobility pattern is traced and used to formulate the demand model. The central disadvantages are costs, frequency (typically 5 to 10 years) and limited sample size.

In parallel, data in big sets became available so that mobility can be observed from numerous sources. Notably, cell phones are traced as they move between the BTS receiving stations. Although cell phones are widespread, estimating mobility from their displacements remains a challenge and raises numerous issues. One of them is the representativeness, which can be fairly alleviated when the market penetration is known. Another one is accuracy, which is technically limited by the GSM infrastructure. In practice this yields at least an order of magnitude lower zoning for revealed OD matrices than for the demand model. Finally, the mobility description is missing in the revealed trips, thus trip purposes, mode and route choices etc cannot be directly read.

In this paper we address this problem by fusing the two available data sources with objective to keep the rich mobility description from travel surveys and exploit the completeness of revealed trips.

Method

We use the classic, four-step demand model to estimate trip volumes between zones. We update it with the OD matrix where trips between main zones are revealed. Main zones (20 districts of the case-study) are much bigger than the zones inside them (250 communes of the case-study).

Trip generation volumes from the demand model are stratified into destination classes. In the case study we stratified destinations into: central agglomeration (Kraków), main cities (Tarnow, Nowy Sacz), other districts, within district, and outside of region. Destination strata shares for each zone are read directly from revealed OD matrices (inferred from the main zone level). Trip destinations are partially predefined at the above stratification and complemented with the gravity model reproducing actual behaviour within the strata. Mode choices are modelled in-line with the demand model, yet each of destination strata can be independently parametrized. Similarly, the route and path choice models can be different for various destination classes. Each strata is a separate four-step demand model summed up into resulting OD matrix.

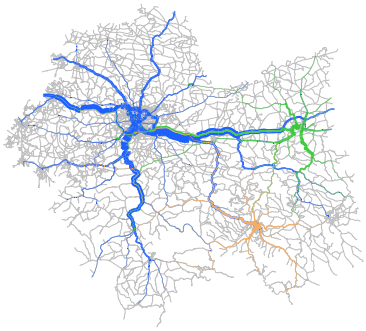

Figure 1: Traffic flows to central metropolis (Kraków - blue) and remaining two cities of the region. Trip volumes estimated from the demand model and catchment areas preserved from the GSM matrices better then in the original demand model.

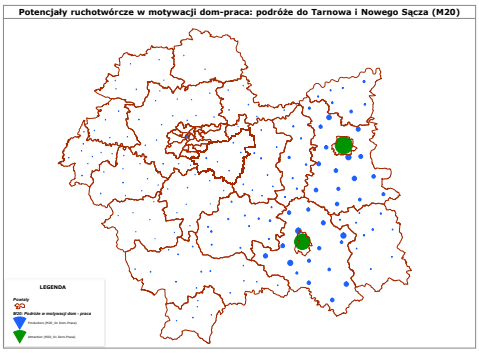

Figure 2: Trip generation from communes (blue) to two main cities in the region (green). Catchment areas preserved from the GSM matrices.

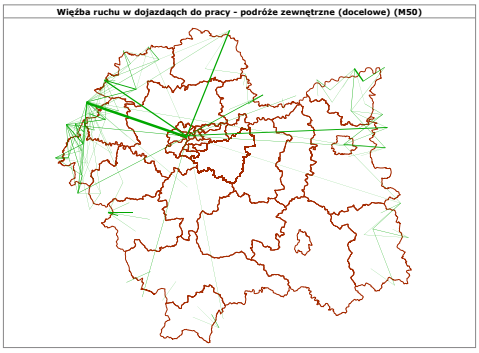

Figure 3: Trip volumes from the communes outside the region. Stronger connections in the western part and high share of trips to Kraków preserved from the GSM matrices and not captured in the original demand model.

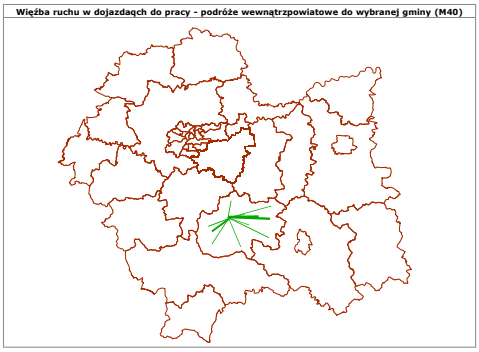

Figure 4: Intra-district trip volumes between communes. Diagonal values of the GSM matrix allowed to realistically estimate the mobility inside the main-zones.

We validate the model first by matching the revealed trip volumes at the trip generation level. Then we parametrize the mode and route choice models to obtain a good fit against traffic and passenger counts. Such transparent calibration allows to obtain a good fit without invasive OD matrix estimation.

Results

We illustrate the method with the Małopolska, 3M inhabitants region of Poland. The demand model was built few years back using data from over 10 000 surveyed trips. Małopolska was divided into over 250 communes (gmina) and the GSM provider revealed the mobility only between over 20 districts (powiat). Due to high market penetration matrices can be treated as representative (uniformly distributed across income, age, etc.). Matrices reveal number of trips between districts (powiaty) in Małopolska and outside, the value on the diagonal is inter-district mobility (between BTS and within district).

We applied the proposed method and obtained the updated model where: long-distance trips are identified and thoroughly analysed, catchment areas of metropolis and big-cities are preserved, most of spatial distribution is solved implicitly already at the generation level, external traffic at outskirts of Małopolska is preserved, good fit was obtained with a traceable calibration procedure.