Estimating Aggregated Origin-Destination Matrices from Automatic Fare Collection

Joana Hora\textsuperscript{a,b,*}, Teresa Galvão\textsuperscript{a,b}, Ana Camanho\textsuperscript{a,b}

\textsuperscript{a}Faculdade de Engenharia da Universidade do Porto, 4200-465 Oporto, Portugal
\textsuperscript{b}INESC TEC, 4200-465 Oporto, Portugal

Abstract

This study addresses the estimation of Origin-Destination (OD) matrices from Automated Fare Collection (AFC) data, following an aggregated approach. The methodology includes the implementation of the Trip Chaining Method (TCM) to estimate the alighting-stop of each AFC record, followed by the identification of transfers amid consecutive AFC records. The identification of transfers allows for the subsequent aggregation of trip-legs into full-trips. The set of full-trips is then used to build OD matrices. This methodology was applied to the case study of Porto, considering 20,000 smart cards over the year of 2013. The results were analyzed through a quantitative approach, complemented by the visualization of OD matrices built at two geographic levels: a detailed perspective using Public Transportation (PT) stops as matrix entries, and a top-level perspective using the PT geographic zones. Data were segmented by type of day, seasonality, and user frequency.

Keywords: Automatic Fare Collection, Origin-Destination estimation, Trip Chaining Method, Transfers

1. Introduction

Until the end of the 20th century, data for the management of PT was mainly gathered from surveys. Aligned with the technological advent of the last decades, PT operators started to invest in Intelligent Transportation Systems technologies and applications, such as Automated Vehicle Location, AFC, Mobile Data Terminals, Automatic Passenger Counter, wireless communications, among others. These new sources of data bring forward a huge potential for the development of innovative optimization techniques, the continuous improvement of PT systems, in addition to upholding the enhanced understanding of impacts arriving from decision-making processes. However, for most cases, the full potential from such large data sets remains unexploited.

This work aims to extract reliable OD matrices from entry-only AFC data. To that end, a methodology combining the TCM and the identification of transfers is proposed. The adequacy of the methodology is demonstrated with its application to the case study of Oporto. The TCM was firstly proposed by Barry et al. (2002) and has been applied to case studies worldwide, including the studies by Alsger et al. (2015), Gordon et al. (2013), and Trépanier et al. (2007). The implementation of the TCM in Oporto’s city includes the by studies Nunes et al. (2015) and Hora et al. (2017).

<table>
<thead>
<tr>
<th>Nomenclature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT</td>
<td>Public Transportation</td>
</tr>
<tr>
<td>OD</td>
<td>Origin-Destination</td>
</tr>
<tr>
<td>AFC</td>
<td>Automatic Fare Collection</td>
</tr>
<tr>
<td>TCM</td>
<td>Trip Chaining Method</td>
</tr>
</tbody>
</table>

\* Corresponding author. Tel.: +351 936 793 388; 
E-mail address: joana.hora@gmail.com

Peer-review under responsibility of the scientific committee of the 23rd EURO Working Group on Transportation Meeting.
2. Methodology

When a smart card has two or more AFC records in the same day, the algorithm estimates the alighting locations of each trip-leg by applying the TCM, considering the following main assumptions (complementary assumptions to build the algorithm are not detailed in this short version):

1. passengers start the next journey stage at or near the alighting location of their previous trip;
2. passengers end the last trip of the day at or near the boarding location of the first trip of the day;
3. passengers can only alight in the sequence of stops not yet traveled by the directed route they boarded;
4. passengers travel out of the PT system when the walking distance between consecutive AFC transaction records is estimated to be higher than a specified threshold.

The next step of the methodology comprises the identification of transfers. This identification concerns distinguishing if the alighting stop of an AFC record corresponds to a transfer within a sequence of trip-legs, or if it corresponds to the destination of a full-trip. The main assumptions adopted to identify transfers between two consecutive AFC records are detailed next (complementary assumptions to build the algorithm are not detailed in this short version):

1. Passengers will not transfer to another vehicle of the same route in which they are traveling, regardless of its direction;
2. Passengers are not willing to walk more than a specified threshold to transfer to another route;
3. Passengers are not willing to wait for more than a specified threshold to transfer to another route;
4. When passengers travel out of the PT system, the next AFC record is the beginning of a new trip.

3. Results

The results were analyzed (1) by type of day considering the categories of workdays, Sundays & holidays, and Saturdays; (2) by seasonality considering the categories of regular service days and reduced service days; (3) passenger frequency considering the categories of frequent and occasional passengers. Each category is analyzed quantitatively as well as with the visualization of OD matrixes at the stop level and at the geographic area level (considering the geographic zones currently used by the operators).

Acknowledgments

This work is financed by the ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 Programme and by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia within project POCI-01-0145-FEDER-032053 (PTDC/ECI-TRA/32053/2017).

References


