With Whom Transport Operators Should Partner? An Urban Mobility and Services Geolocation Data Analysis

Marta Campos Ferreira[1,2(✉)], Teresa Galvão Dias[1,2], and João Falcão e Cunha[1,2]

[1] Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
{mferreira, tgalvao, jfcunha}@fe.up.pt
[2] INESC-TEC, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

Abstract. Automated Fare Collection (AFC) systems produce a large amount of very detailed data, which analysis may be very useful to authorities and transport planners to define future service delivery strategies. Such analysis can be further improved by relating to other data sources, such as points-of-interest (POI) data. As a result public transport operators are able to identify the city service providers with whom it would be more interesting to establish partnerships and propose joint value propositions benefiting both service providers. The objective of such partnerships is to attract new customers and retain those that already exist by providing combined offers, discounts or loyalty schemes. The potential of such analysis is demonstrated by using data related to the city of Porto, Portugal. This study relies on two different data sources: AFC system data and points-of interest data. Urban mobility data is used to identify mobility patterns of different segments of passengers and points-of-interest data is used to analyse the type of services that are likely to concentrate around public transport stations. The results allowed to identify the potential city services to establish partnerships according to the mobility profiles of passengers and the concentration levels of services around public transport stations.

Keywords: Public transport · AFC systems · Points-of-Interest

1 Introduction

The pervasive adoption of Automated Fare Collection (AFC) systems by transport operators worldwide broadens the range of new possibilities beyond fare collection. Such systems produce a large amount of very detailed data regarding on-board transactions, which have been the subject of several research studies. Examples of such studies include: estimation of the origin-destination matrix of the passenger trips [1]; analysis of individual characteristics of travel behaviour to develop user-tailored travel time estimates [2]; validation of estimated travel behaviour [3]; understanding the travel patterns of regular public transport users [4]; and refining public transport

J. C. Ferreira et al. (Eds.): INTSYS 2018, LNICST 267, pp. 133–143, 2019.
https://doi.org/10.1007/978-3-030-14757-0_10

operations, planning, and strategic decisions [5]. Analysis of AFC systems data to investigate other services than mobility related has not been much explored.

As the number and size of cities grow, cities are increasingly facing challenges to develop sustainable modes of transport such as public transport. Several initiates have been carried on to attract more people to public transport and move towards more sustainable mobility. Private car license purchase to access metropolitan areas during peak hours in Singapore [6], vehicle restrictions through congestion pricing in Latin American cities [7] and cities organizing days without cars on the streets are examples of such initiatives. However, these initiatives are not enough and some even proved to cause negative externalities such as excessive mileage accumulation, resource depletion and losses in time and productivity [8].

Therefore, public transport has to be more appealing. The use of mass media to convey the benefits of public transport has proved to have a positive impact on people's attitude to the use of public transport [9]. Another study [10] revealed that newsletter marketing campaigns and sending free bus tickets to households had a positive impact in bus use after the campaign period. Price changes have a significant impact on the number of people that choose the public transport for leisure activities, but do not have much influence on work trips [11]. Therefore, one of the main challenges of public transport is how to address an audience so different and with distinctive characteristics. The composition of public transport customers is extremely varied, including different ages, behaviours, routines, and needs and also different motives for travelling (work, school, leisure, shopping). Thus, we propose a holistic view of the public transport ecosystem, which is composed by several stakeholders such as transport operators, travellers and city services (e.g. hospitals, schools, stores, restaurants, theatres, and gyms).

A multiservice approach is proposed to encourage the use of public transport services, consisting of partnerships between transport operators and city service providers [12]. Such partnerships may include marketing campaigns, combined packages and discounts, offered exclusively to public transport customers. Some cities, like Berlin[1] and Hamburg[2], have already implemented these type of initiatives, but very oriented towards tourists. They offer combined packages of free public transport and discounts at partners. Most of these partners include museums, art-galleries and city tours and very few partnerships are established with restaurants, shopping and stores. The question is: are these the most interesting partners that transport operators should cooperate with? Are these partners also interesting to locals living in the city?

Research work related to this topic are scarce. [13] introduce a case of a service (chain of electronic stores) exposure to different demographic segments during week days and weekends, using data from Porto, Portugal. Another study [14] based on data from the city of Montreal, Canada, explores possible commercial partnerships that might benefit from the characteristics of smartcards. However, this study only consider few services of the city and does not complement the analysis with smartcard data.

---

[1] www.visitberlin.de/en/berlin-welcome-card.

[2] www.hvv.de/en/tickets/single-day-tickets/hamburg-card/.

To explore partnerships between city services and transport operators it is vital to identify the city service providers with whom it would be more interesting to establish partnerships and propose joint value propositions benefiting both service providers. This study explores two data sources from the city of Porto, Portugal: AFC system and points-of-interest (POI) data. Urban mobility data is used to identify mobility patterns of different segments of passengers and POI data is used to analyse the type of services that are likely to concentrate nearby public transport stations. Results allow us to identify potential city services to establish partnerships with, according to mobility profiles of passengers and concentration levels of services around public transport stations. The service exposure of a selected mobility profile is also presented.

The paper is organized as follows: next section describes the datasets that were analysed. Section 3 describes the methodology followed. Section 4 details the main results and Sect. 5 presents the main conclusions and future work.

2 Data

In order to understand the services that are interesting to establish partnerships with, it is important to analyse the demand of each station and the services that are likely to concentrate near stations. To perform the analysis two data sources, from the city of Porto, are used as an illustrative example: AFC system and POI data.

2.1 Metropolitan Area of Porto Public Transport

The public transport network of the Metropolitan Area of Porto (MAP) is composed by an area of 1,575 km2, serving 1.75 million of dwellers. The network is consists of 3.959 stops and 11 operators: 128 bus lines (72 public and 56 private), 81 metro stations and 19 train stations. The electronic ticketing system is an open (ungated) system, composed by ticket readers along the platforms at each metro/train station and at each bus vehicle, and handheld devices for inspectors. The fare media consists of contactless cards, called Andante, which are accepted by all 11 operators. During the year of 2013, a total of 136.32 million journeys were performed [15].

Andante is an entry-only AFC system and the fares are defined according to a zonal structure. The MAP network is divided into 48 geographic travel zones as represented in Fig. 1(a)) and the price to pay for the journey depends on the number of zones crossed from its origin and final destination. The Andante travel cards can be one of the two types: occasional tickets (OT) and monthly subscription (MS). Passengers charge the OT cards with the number of zones they want to cross during a certain journey. In the case of MS it can be used to travel in a set of adjacent zones that were previously chosen by the passenger. Usually, MS are used by locals living in the city who use public transport on a regular basis. OT are mostly used by occasional passengers and tourists. Andante is also a system based on time, allowing passengers to travel for a certain period of time, which increases with the number of zones included in the ticket. To start a journey, passengers tap the Andante card on a ticket reader. Passengers must validate the travel card in the beginning of the journey and whenever changing vehicles during the journey. Therefore, a journey can be composed of more than one journey stage performed in different routes and/or vehicles.

2.2 AFC Data

The Andante system records a validation every time a passenger taps the travel card on the ticket reader. This happens at the beginning of each journey stage, and whenever passengers change vehicles during the journey. Concerning each validation several attributes are recorded, of which are used in this analysis: travel card id, metro station or bus stop where the validation occurred, zone and corresponding coordinates, operator (public bus operator (STCP) or metro operator (MP)), and type of travel card (OT or MS). The data used as an illustrative example of the multiservice approach are the validations recorded in 2013. In this year, 133.979.203 validations were recorded and were performed by 3.017.357 different travel cards, with 45 average validations per card, and in 3.042 different stops. Table 1 shows the number of validations that were recorded, per month, during the year of 2013. Despite the richness of data the researchers had access to, the customer cannot be identified, which is desirable from a privacy perspective. It is not possible to link the travel id card to the customers' name, address, phone number or any other private information.

Table 1. Number of validations during 2013 per month

Month	Nr of validations (millions)	Month	Nr of validations (millions)
January	11.733	July	11.193
February	10.524	August	8.389
March	10.922	September	11.518
April	11.532	October	13.092
May	13.107	November	11.231
June	10.796	December	9.938

2.3 POI Data

The second dataset used in this analysis was extracted from Google Places API. This database has very detailed information regarding the location and types of POI. Google Places API allows to scan the centre point and range of an area, but it is limited to 20 results at a time. To cover the areas of the city with the highest concentration of POI, an area of about 400 km^2 was manually estimated, represented by the green line in Fig. 1 (a). This area was then divided into a matrix, resulting in 69.696 scan areas, with each cell having 75 m side. The POI scan was performed after calculating the centre point of each cell. A total of 33.330 POI were retrieved, after filtering duplicate results at each cell scan.

The Google Places API returned the POI categorised according to 97 categories. These were then manually grouped in 11 high-level categories. Of the attributes provided by the Google Places API, the following were selected to perform the analysis: name (the place's name), type (the place's category) and geographic coordinates.

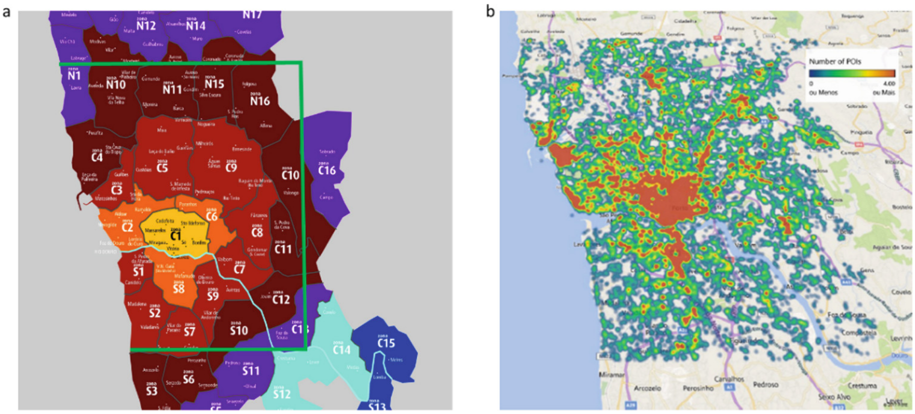

Fig. 1. (a) All public bus stops and metro stations of MAP divided by zones; (b) Heat map of POI distribution

The POI were classified taking into consideration the following high-level categories (number of establishments): Shopping (11.855), Eat and drink (8.513), Service (5.650), Healthcare (3.028), Lodging (1.011), Lifestyle and beauty (997), Education (773), Religion (595), Leisure and culture (440), Government (274), Transportation (194). Comparing Fig. 1 (a) and (b) it is possible to observe that zones C1, C3, C4, C2 and S8 are those with the highest number of POI per unit area.

Shopping high-level category includes places categorized as stores, clothing stores, grocery or supermarket, electronics store, jewelry and shoe store, shopping mall, book and liquor stores. Eat and drink includes restaurants, cafes, bakeries and bars. Service category comprise services such as finance, accounting, car repair, travel agencies, banks, laundries, florists, and insurance agencies. Healthcare includes doctors, dentists, pharmacies and hospitals. Lodging refers to places categorized as lodging. Lifestyle and beauty includes hair care, beauty salons, gyms and spas. Education refers to schools and universities. Religion includes churches, cemeteries and places of worship. Leisure and culture category includes parks, museums, art galleries, movie theatres, libraries and stadiums. Government category comprises city hall, courthouses, police and fire stations. Finally, transport includes some transit stations, subway and train stations.

3 Methodology

The flowchart of the proposed research method in this study is presented on Fig. 2. First, boarding stops/stations of each journey stage can be extracted from AFC system data. The boarding stop/station is sufficient for the analysis, since at a certain time of the day it will be the alight stop/station [1]. Second, the type of POI located around each stop/station can be retrieved from the POI dataset. The distance between the POIs and the stops/stations is calculated a posteriori. Third, the passenger mobility profile is analyzed taking into account the nearest POIs from the travelled stops/stations.

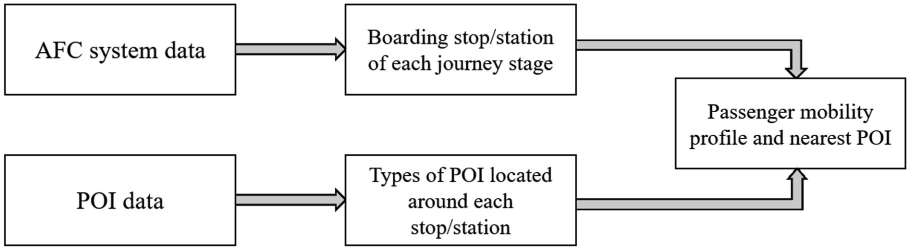

Fig. 2. Flowchart of the research method

The analysis of both datasets followed a granular model, from superior granularity analysis to inferior granularity. The choice criterion to narrow the sample was to choose the sub-samples with the highest number of validations comparing to the others. First we analyzed the full AFC dataset of the year of 2013. Then the month of October was selected and all travel zones were analyzed. During this month 13.570.975 validations were recorded, representing the highest number of validations when compared to other months (see Table 1). The validations were performed with 714.836 different travel cards, in 2.465 different stops and 76 lines. Then we analyzed the zone C1 (shown in yellow in Fig. 1(a)), which accounts for the highest number of validations (48%) during the month of October when compared to the other zones. Zone C1 encompasses the center of the city, being very frequented by tourists and locals. From this we further narrowed the sample and selected the station with the highest number of validations of zone C1 during October – Trindade. Finally we analyzed an individual mobility profile as a case of service exposure. When narrowing the sample, we could observe that despite the number of OT cards is much higher than the number of MS cards, they account for very few number of validations. Therefore, we calculated a validations per card (VPC) ratio, given by:

$$VPC\ ratio = \frac{\sum validations}{\sum travelcards} \tag{1}$$

To measure the agglomeration of POIs around public transport stations, we calculated the distance between POIs and stations using a max radial geodistance of 600 m for each station. This parameter value was based on the Public Transport Accessibility Levels methodology, which assumes an area of 8 min. at 4.8 km/h as the longest distance a person would normally walk to access a bus service [16]. In this work, to perform the analysis, we used the MongoDB, which comes with Node.js, and used Microsoft Excel Power Map to produce the visualization elements.

4 Results

This section presents the main results regarding the analysis of the AFC and POI datasets and present an example of service exposure of a selected mobility profile.

4.1 Analysis of AFC Data

During the year of 2013, 133.979.203 of validations were performed with 2.516.476 different OT travel cards and 500.881 MS travel cards (see Table 2). Despite the high number of OT cards, they only account for 26,1% of the total validations, which corresponds to an average of 14 annual validations per OT card against 198 annual validations per MS card. The same rationale applies to the month of October, zone C1 and Trindade station.

Table 2. Number of validations, travel cards and VPC ratio

	2013 (year)	October (month)	C1 zone October	Trindade station (October)
Nr of validations (% of validations)	133.979.203 (100%)	13.570.975 (100%)	6.513.385 (100%)	1.163.976 (100%)
MS card	99.010.631 (73,9%)	10.418.602 (76,8%)	5.008.001 (76,9%)	787.462 (67,7%)
OT card	34.968.571 (26,1%)	3.152.373 (23,2%)	1.505.384 (23,1%)	376.514 (32,3%)
Nr of cards (% of cards)	3.017.357 (100%)	714.836 (100%)	542.764 (100%)	231.104 (100%)
MS card	500.881 (16,6%)	164.149 (23,0%)	134.102 (24,7%)	68.466 (29,6%)
OT card	2.516.476 (83,4%)	571.183 (79,9%)	408.662 (75,3%)	162.638 (70,4%)
VPC ratio	44	19	12	5
MS card	198	63	37	12
OT card	14	6	4	2

During October 13.570.975 of validations were recorded, from which 76,8% were performed with MS cards and 23,2% with OT cards. Usually, MS cards are used by locals living in the city who use public transport on a regular basis, while OT cards are mostly used by occasional passengers and tourists. Therefore, the main users of the public transport service are people who live in the city and use the public transport regularly.

4.2 Analysis of POI Data

Each stop of zone C1 was measured with each POI considering a max radial geodistance of 600 m, and those above that geodistance were discarded, which resulted on a 7.138 POI sample (see Table 3). The same was done for Trindade station, resulting on 988 POI located around it. This station is located in the city centre, in zone C1, and is the most important one where all metro lines pass by, and passengers change line or vehicle. Additional calculations were made for POIs located between 200 m and 300 m, and 0 m and 100 m around this station.

Analysis of Table 3 shows that eat and drink, shopping, and services are the kind of services that tend to concentrate around public transport stations located in C1 at a maximum distance of 600 m. These are followed by healthcare services and lodging. Education, lifestyle and beauty, religion and leisure and culture show lower levels of concentration around stations. At a distance of 300 m to Trindade station shopping and service are the categories with more services near the station, followed by healthcare, lodging and eat and drink. For a distance of 100 m eat and drink and healthcare are the categories with more services near the station. The overall analysis shows that regardless of analyzing a zone (composed of several stations), a particular station or different distances, the type of services that tend to concentrate around the stations is practically the same.

Table 3. Number and percentage of POI

Category	Nr of POI total sample (%)	Nr of POI C1 zone (0 m–600 m) (%)	Nr of POI Trindade st (0 m–600 m) (%)	Nr of POI Trindade st (200 m–300 m) (%)	Nr of POI Trindade st (0 m–100 m) (%)
Shopping	11.855 (35,6%)	1.821 (25,5%)	284 (28,7%)	32 (22,5%)	5 (17,9%)
Eat and drink	8.513 (25,5%)	1.516 (21,2%)	187 (18,9%)	18 (12,7%)	8 (28,6%)
Service	5.650 (17,0%)	1.446 (20,3%)	172 (17,4%)	36 (25,4%)	2 (7,1%)
Healthcare	3.028 (9,1%)	692 (9,7%)	93 (9,4%)	21 (14,8%)	7 (25,0%)
Lodging	1.011 (3,0%)	733 (10,3%)	159 (16,1%)	21 (14,8%)	1 (3,6%)
Lifestyle and beauty	997 (3,0%)	252 (3,5%)	27 (2,7%)	2 (1,4%)	2 (7,1%)
Education	773 (2,3%)	245 (3,4%)	23 (2,3%)	2 (1,4%)	0 (0%)
Religion	595 (1,8%)	125 (1,8%)	8 (0,8%)	2 (1,4%)	1 (3,6%)
Leisure and culture	440 (1,3%)	189 (2,6%)	19 (1,9%)	4 (2,8%)	0 (0%)
Government	274 (0,8%)	78 (1,1%)	10 (1,0%)	4 (2,8%)	0 (0%)
Transport	194 (0,6%)	41 (0,6%)	6 (0,6%)	0 (0%)	2 (7,1%)
Total	33.330 (100%)	7.138 (100%)	988 (100%)	142 00%)	28 100%)

4.3 Case of Service Exposure

Performing an individual analysis allows to understand the value of the proposed study. To perform the analysis, an individual mobility profile was selected from the AFC dataset. The selection of this profile followed the following criteria: travel card with validations at Trindade station and with an average of 4 validations per day (2 journeys with 2 stages each). Considering 22 working days per month and 4 validations per day, we selected travel cards with 88 validations. This resulted on a 432 MS travel cards sample, from which an individual mobility profile was randomly selected. The most frequent locations of the travel card is calculated by the ratio between the number of validations in each station/stop and the total number of validations of the travel card, paired with stations/stops' geolocations.

The selected mobility profile is represented in Fig. 3(a). The circles indicate the stations where the travel card was validated. The larger the circles, the more validations were performed at the station. Analysis of the AFC data shows that the user travels daily south-north-south, being Trindade the main cross point with a quarter of all its validations. The services that are located around the stations travelled by user, up to a distance of 100 m, were calculated. This resulted in 281 services, represented in Fig. 3 (b). Once again, the services that tend to concentrate around the stations are shopping (27,8%), eat and drink (22,1%), services (18,1%), healthcare (13,5%), lodging (6,4%) and lifestyle and beauty (4,6%). Therefore, customized services offerings could be targeted to this person in order to meet his/her mobility profile and the type of services located near the places frequented by the user. From this analyses it is possible to know exactly the restaurants, stores and services that are near the stations travelled by this person and discounts or deals of his/her interest could be sent.

Fig. 3. (a) Mobility profile of a selected user; (b) Number of POI per category, at 100 m, matched with the mobility profile

5 Conclusions and Future Work

Cities all over the world are increasingly facing challenges to develop sustainable modes of transport such as public transport. Taking into consideration that every journey has a motive, such as school, work, leisure, entertainment, we advocate for a multiservice approach to attract new customers to the public transport and retain the ones that already exist. It involves creating partnerships between public transport operators and city service providers, such as service packages, discounts, loyalty schemes, and marketing campaigns. It is expected that these initiatives will make public transport more attractive by offering benefits and service packages exclusively to its customers.

To explore these partnerships it is vital to understand the public transport usage and identify the city service providers with whom it would be more interesting to establish partnerships and propose joint value propositions benefiting both service providers. Two data sources from the city of Porto, Portugal, are used as an demonstrative example: AFC system and points-of-interest data.

The overall analysis performed showed that the majority of customers using the public transport network are locals living in the city, who use the public transport on a regular basis. Regarding the agglomeration of services nearby public transport stations, shopping, eat and drink and services like accounting, travel agencies, banks, laundries and florists are the services that tend to concentrate around public transport stations. These are followed by healthcare service providers such as doctors, dentists and pharmacies and by lodging services. Education, religion, leisure and culture, government and transport are the services that display lower concentration levels around public transport stations. The methodology followed in this paper can be replicated with other datasets (different cities or countries) and results may be compared.

From a managerial perspective, most of the partnerships that already exist, in some cities, are very oriented towards tourists and consist of discounts and combined packages of public transport and leisure and culture services such as museums, art galleries and touristic attractions. However, the analysis performed revealed that these are not the services that are likely to concentrate nearby public transport stations. Partnerships with service providers that tend to concentrate nearby stations, such as stores, shopping and restaurants are not being exploited to its full potential. Moreover, combined service offerings geared to locals living in city is missing. It is expected that these initiatives make public transport more attractive by providing benefits and services exclusively accessible for passengers.

This study allowed to comprehend the complexity and dynamic of a city and to identify interesting topics to be addressed in future research. Further and richer datasets could be added to the analysis, such as datasets with information about service providers like ratings, satisfaction index and influx hours. The users' interests and preferences could also be added to the mobility profiles in order to define and target effective service offerings. Such analysis can be the basis to the design and development of recommender systems that promote the use of public transport services. Moreover, its offer may be adjusted by taking into account the location of POIs that are not being served by public transport.

Acknowledgements. The authors thank Transportes Intermodais do Porto for providing the AFC system data necessary for this work.

References

1. Nunes, A.A., Dias, T.G., e Cunha, J.F.: Passenger journey destination estimation from automated fare collection system data using spatial validation. IEEE Trans. Intell. Transp. Syst. **17**(1), 133–142 (2016)
2. Lathia, N., Smith, C., Froehlich, J., Capra, L.: Individuals among commuters: building personalised transport information services from fare collection systems. Pervasive Mob. Comput. **9**(5), 643–664 (2013)
3. Munizaga, M., Devillaine, F., Navarrete, C., Silva, D.: Validating travel behavior estimated from smartcard data. Transp. Res. Part C Emerg. Technol. **44**, 70–79 (2014)
4. Lee, S.G., Hickman, M.: Travel pattern analysis using smart card data of regular users. In: 90th Annual Meeting of the Transportation Research Board, no. 520 (2011)
5. Van Oort, N., Cats, O.: Improving public transport decision making, planning and operations by using big data: cases from Sweden and the Netherlands. In: Proceedings of IEEE Conference on Intelligent Transportation Systems, ITSC, vol. 2015–Octob, pp. 19–24 (2015)
6. Seik, F.T.: A unique demand management instrument in urban transport: the vehicle quota system in Singapore. Cities **15**(1), 27–39 (1998)
7. Mahendra, A.: Vehicle restrictions in four latin American cities: is congestion pricing possible? Transp. Rev. **28**(1), 105–133 (2008)
8. Chu, S.: Car restraint policies and mileage in Singapore. Transp. Res. Part A Policy Pract. **77**, 404–412 (2015)
9. Beale, J.R., Bonsall, P.W.: Marketing in the bus industry: a psychological interpretation of some attitudinal and behavioural outcomes. Transp. Res. Part F Traffic Psychol. Behav. **10** (4), 271–287 (2007)
10. Taniguchi, A., Fujii, S.: Promoting public transport using marketing techniques in mobility management and verifying their quantitative effects. Transp. (AMST) **34**(1), 37–49 (2007)
11. Ibraeva, A., de Sousa, J.F.: Marketing of public transport and public transport information provision. Procedia - Soc. Behav. Sci. **162**(Panam), 121–128 (2014)
12. Campos Ferreira, M., Galvão Dias, T.: How to encourage the use of public transport? A multiservice approach based on mobile technologies. Lect. Notes Bus. Inf. Process. **201**, 314–325 (2015)
13. Campos Ferreira, M., Costa, V., Dias, T.G., Falcão E Cunha, J.: Understanding commercial synergies between public transport and services located around public transport stations. Transp. Res. Procedia **27**, 125–132 (2017)
14. Páez, A., Trépanier, M., Morency, C.: Geodemographic analysis and the identification of potential business partnerships enabled by transit smart cards. Transp. Res. Part A Policy Pract. **45**(7), 640–652 (2011)
15. Transportes Intermodais do Porto. Relatório e Contas (2013)
16. TfL: Transport assessment best practice, guidance document (2010)