Can we use OpenStreetMap POIs for the Evaluation of Urban Accessibility?

S. Steiniger[1], M.E. Poorazizi[2], D.R. Scott[1], C. Fuentes[1], R. Crespo[3]

[1]CEDEUS & Pontificia Universidad Católica de Chile A, Avda. Vicuña Mackenna 4860, Macul - Santiago - Chile
Email: ssteiniger@uc.cl

[2]University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada
Email: mepooraz@ucalgary.ca

[3]Universidad Bernardo O'Higgins, Avenida Viel 1497 Ruta 5 Sur - Santiago, Chile
Email: ricardo.crespo@ubo.cl

Abstract

High urban accessibility, measured in terms of easy and fast to reach public and private services, such as schools, parks and shops, is considered an important indicator for quality of life. We developed a web-based platform to measure urban accessibility for transportation planning. Several cities that we aim to study are in Chile where data on urban services is scarce. We evaluated if OpenStreetMap (OSM) can be used as a source of Points of Interest (POI) to evaluate urban accessibility. Based on a field survey in Santiago de Chile, we found that completeness of the OSM POI database is geographical very diverse (7% - 73%) and therefore accessibility scores change significantly. However, scores in some areas did not change much when POIs were added to the OSM database. Given the lack of alternative data sources we recommend using OSM, but suggest evaluating POI completeness in areas with average and low accessibility scores.

1. Urban Accessibility

Urban accessibility measures how easy or difficult it is to visit places of one's day-to-day activities by measuring travel time. As daily activities are considered things such as going to work or school, visiting the library or a park, doing (grocery) shopping, going out to a restaurant, or visiting a movie theatre. Traditionally, transportation planning has focused on measuring accessibility as accessibility to work places, as work-related trips generate high travel demand. Following newer but also classic accessibility measurement theory and practice (e.g., Talen and Anselin 1998; Geurs and van Wee 2004), we developed a web-based platform that measures urban accessibility by counting activity opportunities within a certain travel time – or "travel-shed". The platform should later be used to evaluate how accessibility changes when investments into urban infrastructure or changes to public transit schedules are made.

To evaluate urban accessibility based on activity opportunities and with respect to different modes of travel (e.g. walking, biking, public transit) different data sources are needed: (a) activity opportunities in form of Point of Interest (POI) data, (b) road and pedestrian path network information, and (c) bus and subway/LRT schedules if accessibility with public transit is of interest. In this paper we aim to evaluate how completeness of the (OpenStreetMap) POI database affects calculated walkability scores.

2. Why OpenStreetMap as a Data Source

Our first platform prototype was implemented to measure accessibility in Calgary, Canada (Steiniger *et al.* 2013). For this prototype we used road network data from OpenStreetMap and Points of Interest (POI) were received from MapQuest's proprietary online database.

When we changed our case study to the metropolitan area of Santiago de Chile, an urban area formed by about 37 municipalities with more than 5 Million inhabitants, it became clear that the MapQuest POI database has a focus on North America and Europe and contains much less POI data for regions such as South America. Observing that Google's database also missed a significant number of POI for some areas of Santiago de Chile we chose OpenStreetMap (OSM) as our POI data source. OSMs open database licence and global database coverage helped in making this choice too. Subsequently we used OSM's Overpass API (Olbricht 2015) in our Santiago implementation to obtain POI data.

3. Evaluating Accessibility and OSM Completeness

Given the relatively few POI objects found in the MapQuest POI database, and noting that the Google's database seems also to miss quite some information for some areas of Santiago de Chile, it was seen necessary to evaluate the completeness of OSM's POI database and evaluate the effect of completeness on accessibility score calculation. We note here that a full coverage of all POI is not likely to be achieved for any database, as some areas of Santiago are considered as too dangerous to enter for non-residents. Hence, as a further question, it is of interest what level of POI completeness needs to be reached to obtain a reliable estimate of urban accessibility. An initial hypothesis was that data completeness may be similarly low across Santiago, and therefore accessibility scores will raise by the same amount if the database is completed. For our experiments we concentrate on walkability, as a subcomponent of general accessibility. In the next two subsections we explain briefly how walkability scores are calculated and then present our approach to completeness evaluation.

3.1 Evaluating Urban Walkability

Adopting the basic method from Walkscore.com (2011) our accessibility score is calculated in three basic steps for a particular location (x,y) (see Steiniger *et al.* 2013): (i) calculation of an accessibility area, i.e. a walkshed, (ii) retrieval of activity opportunities (i.e. POI) within that area, and (iii) calculation of a score. Thereby the contribution of each POI to the total score, i.e. its weight, is based on the POI type, number of POI of the same type, and perhaps distance to the location (x,y). The method returns a "walkscore" between 0 (not walkable) and 100 (very walkable).

The POI types considered are: (1) grocery stores, (2) restaurants, (3) shopping (shopping & business), (4) cafés and bars/pubs, (5) banks (ATMs), (6) parks, (7) schools, (8) books (libraries & book stores), and (9) entertainment (cinemas, sport venues, museums). Important is that a POI may not contribute to a total score if the area/shed contains (too) many other POI of the same type. For instance, to reach the maximum score of 100 there needs to be only one ATM and one Park, but five shops/businesses. We note that the original WalkScore.com implementation ("Street Smart") used a 1-mile circle for POI evaluation while our approach utilizes the road network to obtain a walkshed. Accessibility evaluation is implemented employing OGC Web Processing Services (WPS) with GeoServer, and by using OpenTripPlanner, PostGIS and Python scripts for network, area and score analysis (see Poorazizi *et al.* 2015 for details).

3.2 Evaluation of the OSM POI Database

The evaluation of accuracy and completeness of the OpenStreetMap database has been lately a frequent topic in the GIS literature, whereby different feature types and geographical regions were analysed (e.g. Girres and Touya 2010; Arsanjani *et al.* 2015; Camboim et al. 2015; among others). However, for our particular purpose and geographical focus we performed our own completeness evaluation. Therefore we selected 11 ground survey sites in different parts of the Santiago metropolitan area. Stratified sampling was applied with the objective to cover areas with (i) low and high accessibility scores calculated in an initial step, and (ii) areas that are equally geographically distributed.

A team of two persons surveyed each area for objects that would fall into the 9 POI categories mentioned above, while also looking for missing streets and pedestrian paths. Some of the survey sites were considered too dangerous to go there in person. Here, we used Google's StreetView tool – which is available for almost all areas of Santiago - to map POI on screen. In total about 570 unmapped POI were found in the survey, and added to the OSM database. Afterwards we calculated walkability scores (assuming a 15 min walk at 5km/h) for a point grid/lattice covering Santiago, with a spacing of dx=320m and dy=370m. Thereby we calculated scores with the original and with the updated OSM POI database, and evaluated scores only for those grid points that fell within one of the 11 survey areas.

4. Results and Recommendations

The resulting walkability scores for the survey areas, and for Santiago, aren't distributed normal/Gaussian (we observed a mixed distribution), so it is actually not recommendable to present average values. However, to obtain an idea about score variability we present average values below. Figure 1 (left) shows a walkability score map for Santiago, calculated with the original OSM POI database.

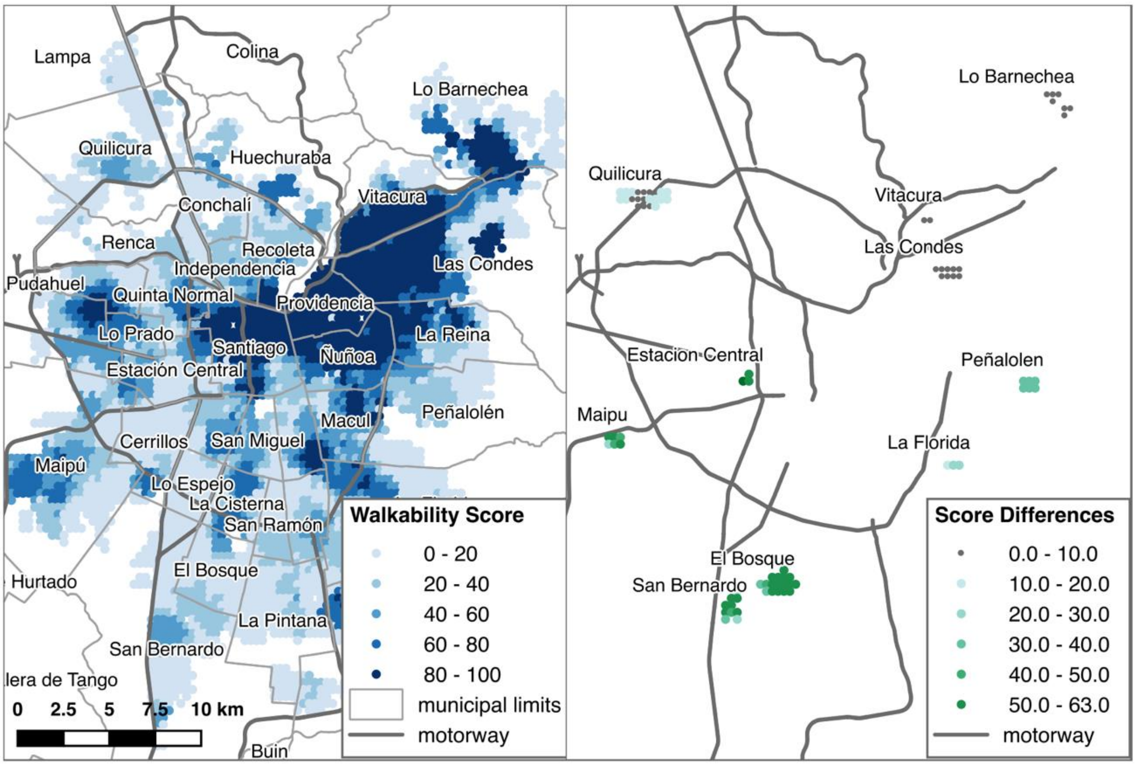

Figure 1. Walkability scores for Santiago de Chile (left) and score differences for 10 of 11 surveyed areas after OSM database updates (right).

POI completeness was in-between 7% to 73%. About half of the survey areas had a POI completeness of 20% or less. Higher percentages of completeness have been reported for other OSM feature types, such as roads (Girres and Touya 2010: 37-45% in France) and land-use (Arsanjani *et al.* 2015: 40-60 % for some urban areas in Germany). This difference can be explained by the fact that roads, paths and land cover are usually visible in satellite images, which are often used by volunteers for mapping on the desktop PC. In contrast, when POI should be mapped on the desktop PC, then the area needs to be covered by a street-image database like Google StreetView. However, a basic set of POI may exist already due to country-wide imports of governmental datasets, containing for instance schools or hospitals. We note that in our survey we found only 2 unmapped pedestrian pathways.

Looking at the walkability scores, the score average did not change at all for two of our 11 survey areas, despite additions to the POI database. The maximum (average) score change found was 58 score points – which is significant considering the score limits [0-100]. Summarizing over all 82 sample grid points the score average rose by about 50 walkscore points (see also Figure 1 right). The two areas that did not experience any change were (a) a Central Business District (Las Condes) – in which 23 POI where added to an existing 62 POI, and (b) a suburban urban sprawl neighbourhood (Lo Barnechea) – in which 6 POI where added to 13 existing POI.

Interestingly there is no clear linear relation between the number of POI added to the database versus score increase. However, it can still be said that: If a neighbourhood was previously scoring low, then scores will raise if more POI are added. But, if a score is already high, then adding new POI to the database may not raise the score at all, due to maximal count limits for POI types during score calculation. Our conclusions are therefore as follows: (i) OpenStreetMap can be used for POI-based urban accessibility analysis, whereby database completeness has a significant influence. (ii) There is most likely no general accessibility score offset/bias for a larger city, since city districts have a high variability in POI completeness. (iii) If OpenStreetMap is used, then POI completeness needs to be evaluated for areas with low or average scores (< 70) before further conclusions can be drawn. Given the (currently few) data points, a POI database completeness of 85% or higher is needed for reliable walkability score estimation. In a next step we will investigate if there are correlations between completeness and socio-demographic factors. If correlations exist, then this would allow developing models for the estimation of completeness.

Acknowledgements

The authors acknowledge financial support by the project “AccesoBarrio” (Conicyt/Fondecyt 1150239) and the Centro de Desarrollo Urbano Sustentable (Conicyt/Fondap/15110020).

References

Arsanjani JJ, Mooney P, Zipf A. and Schauss A, 2015, Quality Assessment of the Contributed Land Use Information from OpenStreetMap Versus Authoritative Datasets. In: Arsanjani JJ et al. (Eds.), *OpenStreetMap in GIScience*, LNGC. Springer, 37–58.

Camboim SP, Bravo JVM and Sluter CR, 2015, An Investigation into the Completeness of, and the Updates to, OpenStreetMap Data in a Heterogeneous Area in Brazil. *ISPRS International Journal of Geo-Information 4*, 1366–1388.

Geurs KT and van Wee B, 2004, Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography*, 12:127–140.

Girres JF and Touya G, 2010, Quality Assessment of the French OpenStreetMap Dataset. *Transactions in GIS* 14, 435–459.

Olbricht RM, 2015, Data Retrieval for Small Spatial Regions in OpenStreetMap. In: Arsanjani JJ et al. (Eds.), *OpenStreetMap in GIScience*, LNGC. Springer, 101–122.

Poorazizi ME, Steiniger S and Hunter AJS, 2015, A service-oriented architecture to enable participatory planning: an e-planning platform. *International Journal of Geographical Information Science*, 29:1081–1110.

Steiniger S, Poorazizi ME and Hunter AJS, 2013, WalkYourPlace - evaluating neighbourhood accessibility at street level. In: Ellul C et al. (Eds.), *Proceedings of the 29th Urban Data Management Symposium*. ISPRS International Archives of Photogrammetry, Remote Sensing, and Spatial Information Science, London, UK.

Talen E and Anselin L, 1998, Assessing spatial equity: An evaluation of measures of accessibility to public playgrounds. *Environment and Planning A*, 30:595–614.

WalkScore, 2011, *Walk Score Methodology*. http://pubs.cedeus.cl/omeka/document/82