A Comparative Study of Pedestrian Accessibility to Transit Stations Using Free and Proprietary Network Data

Revised version submitted: November 15, 2010
Word Count: 4857 words plus 7 figures/tables at 250 words each = 6607 words

By
Dennis Zielstra
University of Florida, Geomatics Program
3205 College Avenue
Ft. Lauderdale, FL-33314
Phone: (954) 577-6392
E-mail: dzielstra@ufl.edu

Hartwig H. Hochmair
University of Florida, Geomatics Program
3205 College Avenue
Ft. Lauderdale, FL-33314
Phone: (954) 577-6317
Fax: (954) 475-4125
E-mail: hhhochmair@ufl.edu

ABSTRACT

Availability of a transit service is a key factor in a traveler's choice on transportation mode. Transit service is only a realistic option if the service is available at or near the locations at the time a person plans to travel. Whereas there are a variety of measures for transit availability, such as service frequency, this paper focuses on the spatial aspect of pedestrian accessibility to transit stations, that is, on service coverage. Service areas are commonly used to visualize accessibility for pedestrians to transit systems and to analyze the potential ridership. Since the service area for a station is defined over the maximum network walking distance from a transit station, a complete street network that includes pedestrian segments, i.e., shortcuts, is of high importance for a realistic assessment of service areas. Whereas most proprietary geodata providers solely concentrate on car related geodata, public domain street data and Volunteered Geographic Information (VGI), such as OpenStreetMap (OSM), provide a potential valuable source for pedestrian data. In this paper we compare the amount of pedestrian related data between freely available sources, i.e., OSM and/or TIGER, and proprietary providers, i.e., Tele Atlas and/or NAVTEQ, and analyze its effect on modeling transit accessibility for pedestrians. The analysis is conducted for five US and four German cities, which allows to identify differences between these two countries.

INTRODUCTION

The importance of accessibility to transit services and its role to transit ridership has been frequently reported in the transportation literature (*4, 3, 1*). Therefore identifying the percentage of population served by a transit system is a key performance measure for transit systems (*7, 14*). Pedestrian access to a transit service is strongly associated with transit ridership (*12*), where accessibility and availability can be measured both in spatial and temporal dimensions (*20*). If transit cannot be accessed, other aspects of transit quality, such as cost, travel time, crowding, and safety become irrelevant for the mode choice. Accessibility to transit provides an incentive to walk to transit stations, which increases the proportion of Americans who obtain more than 30 minutes of daily physical activity, which is recommended by the Surgeon General (*2*). Accessibility is often visualized through network service areas which are areas that can be reached from a given network facility within a specified impedance, such as distance or travel time. Service areas can be used to define optimal stop spacing, identify redundancy and gaps in the system, and understand and predict demand for transit (*5*). Since Euclidean buffers overestimate the service area of a stop, buffers based on network distances are preferable (*18, 12, 23, 13*), although they do not account for off-street shortcuts (*6*). The latter limitation is caused by street network data sets with missing pedestrian paths. This is the actual focus of the paper which has two related objectives. The first objective is to compare the portion of pedestrian segments between a variety of free and commercial street data sets for selected cities in the US and Germany and thus provide specific pedestrian related information about the datasets available for this study. The second objective is to analyze to which extent the integration of the pedestrian-only segments, e.g., shortcuts and alleys, affects the size of service areas generated around bus and metro stations.

The high demand for freely available spatial data within recent years has boosted the availability of Volunteered Geographic Information (VGI) (*8*) on the Internet. The development of Web 2.0, the Global Positioning System (GPS) and its integration into cell phones, photo cameras, and other mobile devices, allows the Web community to interact with each other, provide information to central sites, and thus become a significant source of geographic information. The OpenStreetMap (OSM) project (available at openstreetmap.org), has the goal to create a detailed map of the world based on VGI in vector data format. The information is collected by many participants, collated on a central database and distributed in multiple digital formats through the World Wide Web. It provides free downloads of spatial data from a large selection of themes, including roads, transit lines, tourist sites, or land use layers. Whereas most commercial providers of street data focus primarily on car navigation, OSM data also include pedestrian and bicycle paths, which makes them a potential rich data source used for off-street routing. The availability of pedestrian segments and their effect on service areas generated around transit stations will be analyzed for OSM, TIGER/Line, Tele Atlas and NAVTEQ street datasets. The geographic area of analysis includes five US cities (Atlanta, Chicago, Miami, New York, and San Francisco) and four German cities (Berlin, Cologne, Hamburg, and Munich). Service areas will be separately analyzed for bus facilities and metro/light rail facilities.

The remainder of the paper is structured as follows: The next section reviews previous findings on the size of service areas in transit systems. This is followed by an introduction to the commercial and free street data sets and the analysis method used in this study. The next section provides results on the amount of pedestrian segments in data sets and the size of service areas in different networks. This is followed by a discussion of the results and aspects of future work.

SERVICE AREAS FOR TRANSIT SYSTEMS

Whereas nearness of proximity of the trip origin to the next transit station is one of the most important deciding factors for public transportation users to walk to transit stations or not, other factors related to density (e.g., employment density, residential density), land-use (e.g., land use diversity), and design (e.g. number of park and ride spaces at a transit stations, transit service levels, sidewalk density), were found to be significant effects as well (*3*). The most common standard measures of walking distance to transit stations used in the public transit industry to generate service areas are 400 meters (0.25 miles) buffers around bus stops (*18, 23*) and 800 meters (0.5 miles) around rail stations (*19, 16*) with most passengers (75-80%) walking the given distances or less. A more recent study suggests a larger catchment area around transit stations (*6*), with a 85th percentile of walking distances to bus service of around 550 meters from the origin and 615 meters to the destination. For metro service, distances were found to be 750 and 695 meters, and for light rail 1200 and 1100 meters, respectively. Travelers aim to minimize the distance and time of the walking portion of their trip (*19, 11*), where the maximum walking distance is affected by demographic variables (income, education), characteristics of the pedestrian environment, such as topography and safety of intersections, or population density among others (*17, 23*).

The Transit Capacity and Quality of Service Manual (TCQSM) addresses transit availability at a planning level within three measures, which are service frequency, hours of service, and service coverage (*15*). These fixed-route transit service measures can be determined at transit stops, along route segments and corridors, or throughout a system. System service coverage, which is of interest for this paper, can be measured, for example, as route miles per square mile, or percentage of the system area served. Since land uses and population and job densities may vary greatly from one system to another, the transportation literature suggests to use transport supportive areas instead, which look at how much of the area that would typically produce the majority of a system's ridership, i.e., the densest areas, are served (*21*). Whereas catchment areas are often associated with potential ridership, we restrict this exploratory analysis to the geometric aspect of the catchment area, i.e., how the size of the catchment area is affected through integration of pedestrian segments in the street model. This is considered as a first indicator of relevance for consideration of pedestrian segments for ridership analysis and multi-modal routing applications.

ANALYSIS METHODOLOGY

Free and Proprietary Street Data Sets

The computation of a service area is based on the underlying street network, which can be created from a fast growing selection of free street data, or commercial data sets. The quality of geodata and especially VGI has become a major research area over the past years. For example, Haklay (*10*) and Zielstra and Zipf (*24*) analyzed the quality of OpenStreetMap data for England and Germany, in particular the aspects of completeness, positional accuracy, and attribute accuracy. For Germany, a comparison of coverage between Tele Atlas with OSM data conducted for several periods during 2009 revealed the largest discrepancies and slowest development for car navigation related data with total length differences of up to 43%. This indicates that Tele Atlas at that time specialized in car related data. The comparison of pedestrian navigation related data where the total length differences over time could be reduced from 27% to almost 10%

provided further evidence for the assumption that OSM specializes on smaller streets and alleys. The completeness of OSM data reveals some spatial pattern, which is a clear decrease in the completeness of the OpenStreetMap dataset away from the metropolitan areas towards the surrounding rural areas. However, within more densely populated areas, OSM data offer more overall data than the commercial provider (*24*).

The major motivation for voluntary contribution to VGI data collection efforts in European countries lies in the fact that geospatial data layers are generally not provided for free by the federal agencies. This motivation may not be as high in the US where numerous base layers are publicly available, including TIGER/Line data from the US Census Bureau. For the area of the United States the OpenStreetMap project imported the Census TIGER/Line datasets. Therefore, comparing coverage of OpenStreetMap with TIGER/Line data, as will be done in this paper, provides some insight about the activity of the OpenStreetMap community in the US.

As far as commercial data sets goes, we were able to obtain NAVTEQ Discover Cities data for the states of Florida and Georgia. This data collection integrates into the street network also pedestrian specific routing segments that include shortcuts through parks, plazas, or publicly accessible buildings for selected cities, including Miami and Atlanta. For Tele Atlas we were able to obtain the Multinet Dataset for all cities that were analyzed in the US and Germany. The point data for bus and metro/rapid trail stops were obtained from local transit authorities for US cities, and extracted from OSM data for German cities. The exact area used for US cities was delimited by the counties a city consists of, such as Miami-Dade County for Miami. These counties were further clipped to the urban area layer, as defined by the US Census Bureau, to focus on the more densely populated regions in metropolitan areas. Boundaries of German cities were adopted from the Tele Atlas boundary layer. TABLE 1 gives an overview of street datasets available for the different cities, areas of cities, and the number of transit stations used.

TABLE 1 Characteristics of data sets used for different cities.

		Commercial		**Free**		**Area [km^2]**	**# Bus**	**# Metro/ Rail**
		NAVTEQ	Tele atlas	OSM	Tiger			
US	Atlanta	√	√	√	√	879	9419	38
	Chicago		√	√	√	3047	11800	381
	Miami	√	√	√	√	988	11679	27
	New York		√	√	√	774	12462	493
	San Francisco		√	√	√	122	8066	166
Germany	Berlin		√	√		891	3358	1203
	Cologne		√	√		405	645	326
	Hamburg		√	√		755	2931	333
	Munich		√	√		310	1375	673

To be able to assess the effect of pedestrian-only segments on the size of service areas, all network segments that were only accessible to pedestrians (and not cars) were extracted first. Due to different attributes and road classification schemes between different data sets, the extraction methods had to be adapted for the four datasets. For example, for NAVTEQ data a query over street attributes formulated as "[Auto Access] = 'N' AND [Pedestrian Access] = 'Y'" was used to extract pedestrian-only data. For TIGER/Line, pedestrian data can be extracted through filtering the corresponding MAF/TIGER Feature class code (MTFCC), in this case code S1710, which stands for walkway/pedestrian trail. This extracted collection was then either merged with the car navigation data or not, which gave two different street networks of different complexity. For the data analysis, the relative increase in service area, based on integration of pedestrian-only data compared to the street data set, are computed for each city. Whereas it is tempting to ask which of the analyzed datasets is the best one to generate service areas for pedestrians, it must be noted that one dataset may be more complete on the street data, and another one with pedestrian data, so that a combination of different datasets gives possibly the most comprehensive solution for this task. The goal of this paper is to demonstrate the effect of pedestrian-only data within a given data source and assess whether their integration leads to a noticeable increase in service area, thus accessibility.

Selected Method for Measuring Service Areas

The size of service areas can be measured in two ways. The first method is to measure the area of polygons created in the neighborhood around the facility. Whereas the polygon method is effective in visualizing the general extent of services areas, the exact area results for this analysis method depend on a variety of analysis settings that specify how exactly polygons are generated. If the polygons are not trimmed after their generation, then the polygon method connects all neighboring points on the street network that are the exact threshold distance away from the nearest facility. These polygon boundaries are curved inside towards the facility if they would intersect with streets that lie beyond the threshold distance. The generation of these polygons follows mathematically sound rules, but may lead to shapes and indentations that appear to be somewhat counterintuitive. An example for such indentation is given in FIGURE 1a, with bus stations visualized as dots, and a used distance threshold of 400m. The effect of such indentations can be reduced by defining a trim distance to be applied after the polygons have been created. Setting a trim distance trims the polygons containing the edges at the periphery of the service area further to be within the specified distance of these outer edges, thus provides polygons that are closer tied to the actual street shape. If using this setting, the proportional contribution of peripheral edges depends on the selected trim distance. FIGURE 1b and c show the same situation as before, now using a trim distance of 10m (b) and 30m (c), respectively. Considering this biasing effect for polygon measures we decided to use another measure instead, which is the total length of street segments that are reached within the impedance thresholds. In other words, these are the segments that overlap with polygons. This measure is independent of the trim distance chosen and therefore better suited for quantitative comparison of service areas. As can be seen, the portion of streets overlapping with service area polygons does not change throughout figures a-c, which is a desirable characteristic. This suggested measure has also the advantage that there is no need to remove inaccessible areas, such as water bodies, which should be done when using the polygon measure method. ESRI's ArcGIS 9.3.1 Geographic Information System was used for all spatial analysis procedures in this paper.

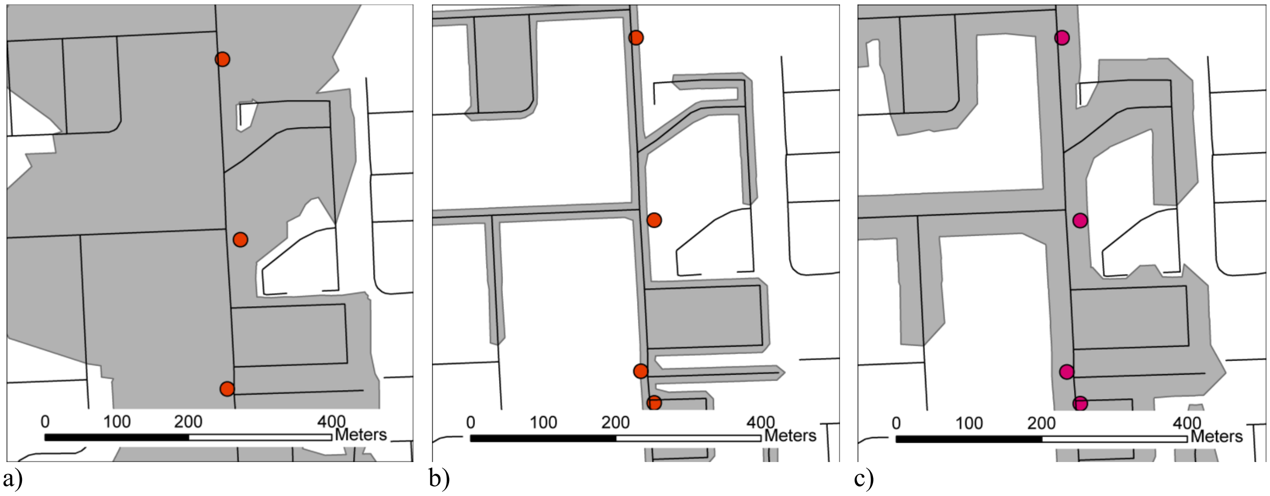


FIGURE 1 Generating polygon based service areas: No trimming (a), 10m (b) and 30m (c) trim distance.

RESULTS

This section gives an overview of the total length of streets and pedestrian-only segments found in the 9 urban street networks, and its effects on service areas.

Analysis of Pedestrian Segments in Data Sets

To give some reference lengths for comparison, the first row in FIGURE 2 shows the overall length of all used street types for entire urban street networks, which comprises segments that are accessible to cars/pedestrians and pedestrians only. Differences of street lengths between the four data providers are small for US cities (FIGURE 2a), and maximum lengths found vary between the different data providers. The somewhat higher overall length value for TIGER in Chicago seems to be the result of a particular classification scheme in TIGER, where private streets in industrial areas, i.e., quarries, are also classified as local neighborhood and rural streets, whereas such streets are excluded in the other data sources. In Germany, a more distinct pattern can be found, where total lengths are clearly higher for OSM than for Tele Atlas for each analyzed city. Differences range between 13% for Cologne and 44% for Munich (FIGURE 2b).

The second row in FIGURE 2 shows the length of pedestrian-only segments. Independent of the data source, the portion of pedestrian-only segments relative to total segment lengths is small for US cities. The range lies between 0.03 % (TIGER data in Atlanta) and 8.97 % (OSM data for San Francisco), as shown in TABLE 2. Among these relatively small numbers, there is no clear pattern on which data source provides the best coverage for pedestrian-only data (FIGURE 2c). OSM is ahead for Chicago and San Francisco, the latter one being the city where the US OSM project was launched (which may be a determining factor for the prominence of OSM in that area). The other three cities are best covered by commercial data providers. NAVTEQ provides most pedestrian-only segments for Atlanta and Miami, which is probably because these two cities are part of the pedestrian oriented NAVTEQ Discover Cities project. Tele Atlas is ahead in New York, mostly, because it covers in more detail the pedestrian paths in the Central Park compared to the other data sources.

In Germany, where OSM and Tele Atlas data were analyzed, a generally higher percentage of pedestrian-only segments can be found for both data sets compared to US cities (TABLE 2). This result suggests that the urban structure of the selected German cities is more pedestrian friendly and therefore provides more shortcuts and alleys. Expressed in relative numbers, the portion of pedestrian-only segments for Tele Atlas varies between 7% (Munich) and 17% (Cologne) and for OSM between 26% (Cologne) and 31% (Berlin). The difference between OSM and Tele Atlas data is more pronounced for pedestrian-only data than for the entire street network (compare the relative size of bars between FIGURE 2d and d). OSM has between 3.2 (Berlin) and 5.6 (Munich) times as many pedestrian-only data as Tele Atlas (see FIGURE 2d), which reveals the OSM focus on off-street data capture.

For each of the five US cities the overall length of pedestrian-only segments in OSM is a manifold of the corresponding TIGER data lengths (compare green with black bars in FIGURE 2c). The ratio OSM/TIGER lies between 2.8 (New York) and 113 (San Francisco). This is an indication that, at least for some US regions, the OSM community has been very active and did not just import TIGER/Line data.

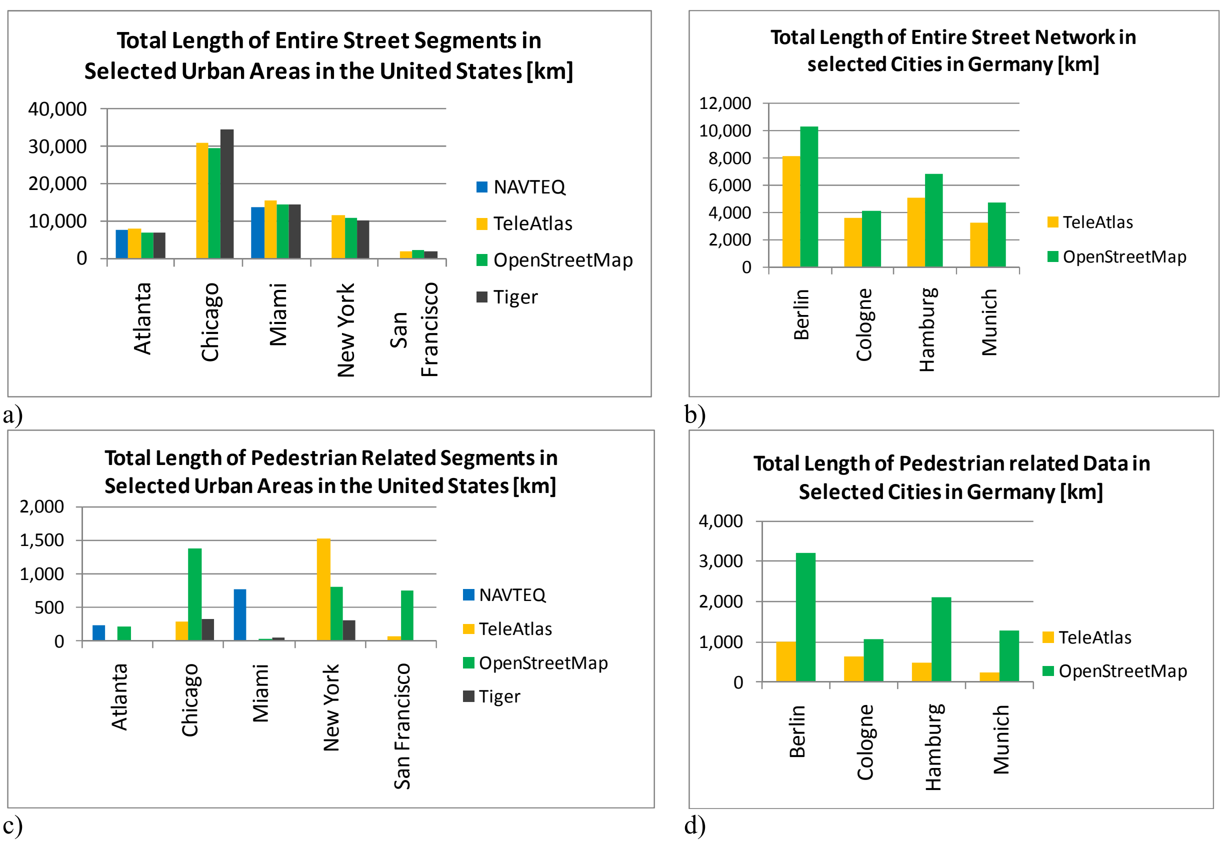
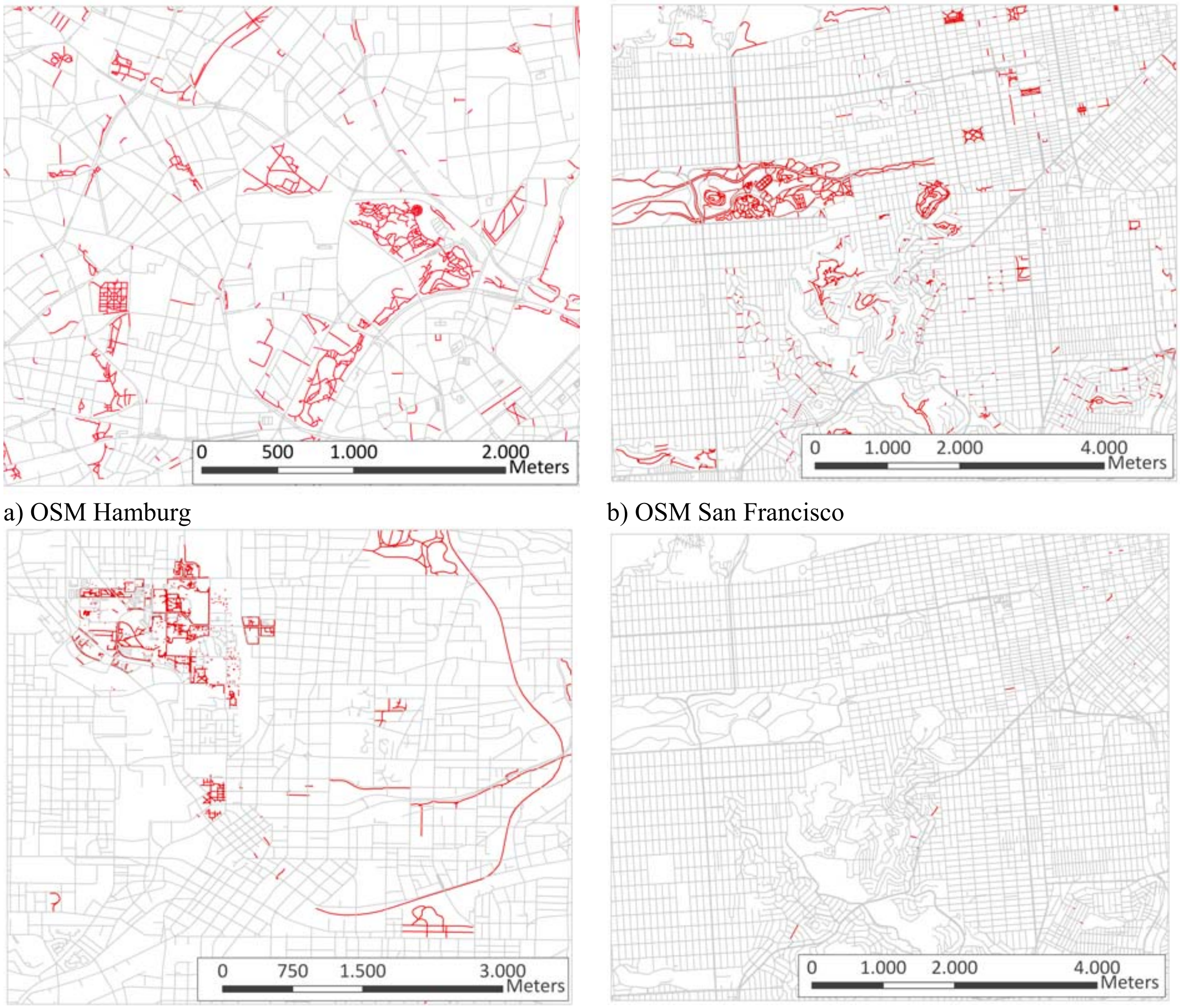


FIGURE 2 Total length of streets (a b) and pedestrian-only segments (c, d) for selected US urban areas and German cities.

TABLE 2 Percentage of pedestrian-only segments in datasets rounded to the nearest 0.1%.

		NAVTEQ	Tele Atlas	OSM	TIGER
US	Atlanta	0.9	0.1	0.9	0.0
	Chicago	-	0.3	1.4	0.3
	Miami	1.6	0.0	0.1	0.1
	New York	-	3.6	2.1	0.8
	San Francisco	-	1.0	9.0	0.1
Germany	Berlin	-	12.4	30.9	-
	Cologne	-	17.1	25.5	-
	Hamburg	-	9.3	30.5	-
	Munich	-	7.0	27.0	-

To visualize some of these observed differences and regional effects, FIGURE 3 maps street data (in gray) overlaid with pedestrian-only data (in red) for selected cities. FIGURE 3a shows the OSM dataset for Hamburg with a high density of pedestrian-only segments. The distribution of pedestrian-only segments is relatively homogenous throughout the network and not concentrated around parks only. FIGURE 3b shows part of San Francisco, which is the analyzed US city that has the highest percentage of OSM pedestrian-only segments (9.0%). Pedestrian-only paths are more concentrated around parks and around the bay area (the latter one not shown here). FIGURE 3d below shows the situation for TIGER data in the San Francisco area. As can be seen only a few isolated pedestrian-only segments are integrated, and the difference to the OSM coverage can be easily noticed. FIGURE 3c shows Atlanta downtown as an example for a city with a smaller OSM pedestrian-only coverage of 0.9%, where the segments are mostly found in the downtown area.

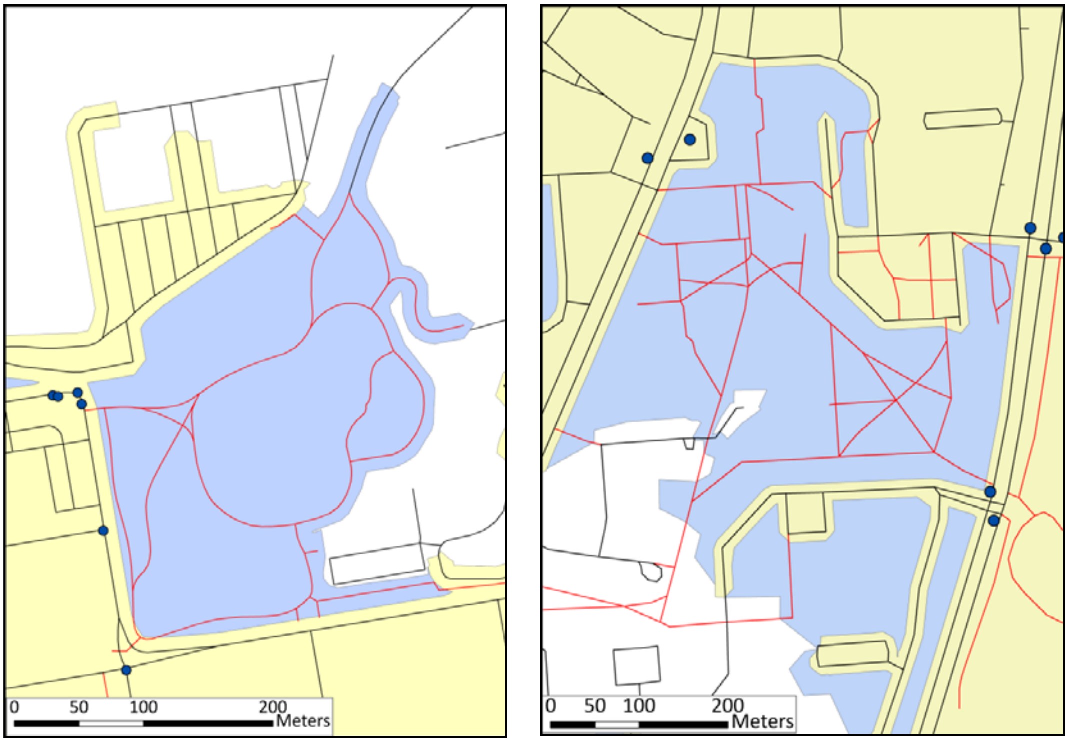


a) OSM Hamburg
b) OSM San Francisco
c) OSM Atlanta
d) TIGER San Francisco

FIGURE 3 Visualization of pedestrian-only segments on top of streets for selected cities.

Analysis of Service Areas

As an example, FIGURE 4 demonstrates the effect of pedestrian-only segments on the generation of service areas for San Francisco (FIGURE 4a) and Berlin (FIGURE 4b). Dots show the locations of bus stops and lines show OSM street data. Gray lines indicate segments that can be traversed both by car and pedestrian, whereas lines in red are pedestrian-only segments (shortcuts, alleys, or paths in parks). For the generation of service areas in a city, two networks are used. The first network is built from segments passable by cars ("car network"), whereas the second network comprises in addition to this also pedestrian-only segments ("combined network"). Areas in yellow visualize service areas generated on the car network using a threshold distance of 400 meters. Areas in blue show additional service areas covered when using the combined network instead. To numerically capture the relative increase in service area due to consideration of pedestrian-only segments we first measure the total length of segments in the combined network that are located in blue areas (following the principle explained in connection with FIGURE 1). This number is then divided by the total length of segments in the combined network that are located in yellow or blue areas.



a) b)

FIGURE 4 Effect of pedestrian-only segments for generation of service areas in San Francisco (a) and Berlin (b).

TABLE 3 shows for the 9 cities the relative growth of service areas through the integration of pedestrian-only segments. Values in the "Bus" column are based on a 400 m threshold distance, and values in the "Metro" column on an 800 m threshold distance. Mean growth rates of service areas for US and German cities are given below.

TABLE 3 Relative growth of service area when integrating pedestrian-only segments, given in %.

		NAVTEQ		Tele Atlas		OSM		TIGER	
		Bus	Metro	Bus	Metro	Bus	Metro	Bus	Metro
US	Atlanta	4.6	18.8	0.7	0.8	2.9	11.3	3.7	5.5
	Chicago	-	-	0.3	0.5	8.1	20.6	0.8	0.9
	Miami	4.5	2.9	0.2	0.2	0.3	0.0	0.4	0.0
	New York	-	-	1.7	2.6	9.3	8.1	1.6	2.1
	San Francisco	-	-	3.0	3.2	18.5	17.1	0.5	0.5
Mean		***4.5***	***10.7***	***1.2***	***1.4***	***7.8***	***11.5***	***1.4***	***1.8***
Germany	Berlin	-	-	11.5	21.4	59.4	76.8	-	-
	Cologne	-	-	8.2	4.7	70.2	84.8	-	-
	Hamburg	-	-	13.5	16.3	94.9	111.3	-	-
	Munich	-	-	7.4	6.9	98.7	117.2	-	-
Mean				***10.2***	***12.3***	***80.8***	***97.5***		

Some general patterns can be observed, as follows. The effect of pedestrian-only segments on the size of service area is generally larger for the 800 m than for the 400 m distance. This indicates that the accessibility to metro and light-rail stations benefits more from the integration of pedestrian-only segments than accessibility to bus stops. The average growth of service areas, both for bus and metro, is higher for the OSM data set when compared to the commercial Tele Atlas data set (for all 9 cities) and TIGER/Line data (for the five US cities). This stronger growth seems to be caused by the higher average percentage of pedestrian-only segments in the OSM data set compared to the two other data sets (compare corresponding columns in TABLE 2). Considering only Atlanta and Miami, where data sets from all four sources were available, NAVTEQ data result in the largest growth of service area, both for bus and transit facilities. Also here, the higher percentage of pedestrian-only segments in NAVTEQ data compared to the other data sources seems to be the cause for this strong increase in service area (see rows for Atlanta and Miami in TABLE 2).

Growth rates are between 8.3 and 10.3 times larger for Germany than for US, where this effect is similar for Tele Atlas and OSM data. This shows that in analyzed German cities integration of pedestrian-only paths has a much larger impact on service areas for pedestrians than for US cities, which, again can most likely be attributed to the higher percentage of pedestrian-only paths in German cities. Whereas these initial findings are indicative of a positive correlation between percentage of pedestrian-only segments in a network and service area growth, one must be careful with this conclusion. Comparing pedestrian-only percentages of all 9 cities with their average service area growth rates computed from the different data sources gives a strong significant positive correlation (Pearson's r) for bus stops ($r=0.93$, $n=9$, $p<0.01$) and metro stops ($r=0.94$, $n=9$, $p<0.01$). However, when considering correlations within the groups of US and German cities separately, only one corresponding correlation can be found to be significantly positive, i.e., percentage of pedestrian-only segments and service area growth for bus facilities in the US ($r=0.94$, $n=5$, $p<0.01$). This indicates that besides percentage of pedestrian-only paths also the spatial distribution of pedestrian-only paths within a network affects the growth of the service area. A concentration of such segments around the transit network will more strongly affect growth of service areas than with equally distributed pedestrian-only segments in an urban area.

DISCUSSION AND FUTURE WORK

Based on the data sets available for the different cities, the results show that pedestrian-only segments increase service areas around transit facilities up to 20% in US cities, and up to 117% in German cities (TABLE 3). These computations are based on the lengths of streets and paths within service areas, and not the area measurement itself. These generally high rates suggest that integration of pedestrian-only segments can lead to a more realistic assessment of service areas when compared to using networks that contain only streets that are passable by cars. The maximum growth rates for US and Europe were found for OSM data analysis in those cities where NAVTEQ data were not available to us for the analysis. Since NAVTEQ data revealed the highest growth rates for US cities where available (Atlanta, Chicago), there is the possibility that growth rates can even exceed the maximum identified growth rate of 20% for US cities, such as Chicago or San Francisco. For Europe the same conclusions cannot be made at this point since no NAVTEQ data were available for comparison at the time of the study.

Further inspection of growth numbers is needed to better understand the benefits of pedestrian-only segments for service area modeling. First, results suggest that some urban street networks do not provide a significant number of shortcuts that facilitate transit access for pedestrians. For example, NAVTEQ data, which provide the most comprehensive pedestrian data set for Miami (TABLE 2), yields a mere 4.5% service area growth for bus stops and 2.9% for metro and light-rail stations. Second, especially for pedestrian oriented cities with plazas, pedestrian paths and alleys, the choice of the pedestrian-only data set greatly affects how computed service areas will change. OSM data provide a free and relatively comprehensive option for cities where commercial pedestrian data sets are not yet available, such as for the analyzed German cities. For Atlanta and Miami, NAVTEQ Discover Cities data, which claim to include detailed pedestrian paths, show a better coverage than OSM. TIGER/Line data were found to have fewer pedestrian-only segments in the five US cities than OSM, which gave a smaller growth rate for service areas in most cases. This finding indicates that the OSM community in the US is active and helps to develop a comprehensive network of pedestrian paths, as can be particularly seen in the San Francisco and Chicago areas. Future work will elaborate on whether any systematic pattern can be found that captures the different percentage of reported pedestrian paths and growth rates, respectively, between the analyzed cities, where the urban structure may be a determining factor.

Some further conclusions of the findings are that the considerations of pedestrian-only segments in network models will help to refine the assessments of overlapping service areas in transit networks (*6*). Since walk access to transit has a high effect on mode choice, refined service areas using methods demonstrated in this paper, will also help to improve the accuracy of predictive models on transit ridership. However, in order to estimate the impact of pedestrian-only paths on potential transit ridership, the current analysis needs to be combined with household and employment densities (*21, 15*). Whereas this study looked at one data source at a time, more complete road street networks could be obtained by combining different data sets, where the integration of heterogeneous (free) data sets and their generalization for analysis tasks is still a challenge (*22*), and the assessment of VGI data quality, especially OSM data, is an ongoing issue of high importance for successful geo-applications (*9, 10*).

ACKNOWLEDGEMENTS

The authors thank NAVTEQ and Tele Atlas for the generosity to provide us with sample data sets of the US street network. Alexander Zipf facilitated the use of Tele Atlas data for Germany for this study.

References

(*1*) Beimborn, E. A., Greenwald, M. J., and Jin, X. Accessibility, Connectivity, and Captivity: Impacts on Transit Choice. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. *1981*, 2003, pp. 1-9.

(*2*) Besser, L. M. and Dannenberg, A. L. Walking to Public Transit: Steps to Help Meet Physical Activity Recommendations. *American Journal of Preventive Medicine*, Vol. *29*, 4, 2005, pp. 273-280.

(*3*) Cervero, R. Walk-and-Ride: Factors Influencing Pedestrian Access to Transit. *Journal of Public Transportation*, Vol. *7*, 2001, pp. 1-23.

(*4*) Cervero, R. and Radisch, C. Travel Choices in Pedestrian Versus Automobile Oriented Neighborhoods. *Transport Policy*, Vol. *3*, 3, 1996, pp. 127-141.

(*5*) El-Geneidy, A. M., Strathman, J. G., Kimpel, T. J., and Crout, D. The effects of bus stop consolidation on passenger activity and transit operations. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. *1971*, 2006, pp. 32-41.

(*6*) El-Geneidy, A. M., Tétreault, P. R., and Surprenant-Legault, J. *Pedestrian access to transit: Identifying redundancies and gaps using a variable service area analysis*. Transportation Research Board - 89th Annual Meeting, Washington, D.C. Transportation Research Board of the National Academies, 2010.

(*7*) Fielding, G., Glauthier, R., and Lave, C. Performance indicators for transit management. *Transportation*, Vol. *7*, 4, 1978, pp. 365-379.

(*8*) Goodchild, M. F. Citizens as Voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0 (Editorial). *International Journal of Spatial Data Infrastructures Research (IJSDIR)*, Vol. *2*, 2007, pp. 24-32.

(*9*) Goodchild, M. F. Spatial accuracy 2.0. In J.-X. Zhang and M. F. Goodchild (Eds.), *Spatial Uncertainty, Proceedings of the Eighth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, *No. 1*, World Academic Union, Liverpool, 2008, pp. 1-7.

(*10*) Haklay, M. How good is Volunteered Geographical Information? A comparative study of OpenStreetMap and Ordnance Survey datasets *Environment and Planning B, Planning and Design*, Vol. *37*, 4, 2010, pp. 682 – 703.

(*11*) Hochmair, H. H. Grouping of Optimized Pedestrian Routes for Multi-Modal Route Planning: A Comparison of Two Cities. In L. Bernard, H. Pundt and A. Friis-Christensen (Eds.), *The European Information Society - Taking Geoinformation Science One Step Further*, *Lecture Notes in Geoinformation and Cartography*, Springer, Berlin, 2008, pp. 339-358.

(*12*) Hsiao, S., Lu, J., Sterling, J., and Weatherford, M. Use of geographic information system for analysis of transit pedestrian access. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. *1604*, 1997, pp. 50-59.

(*13*) Kimpel, T., Dueker, K., and El-Geneidy, A. Using GIS to measure the effect of overlapping service areas on passenger boardings at bus stops. *Urban and Regional Information Systems Association Journal*, Vol. *19*, 1, 2007, pp. 5-11.

(*14*) Kittelson & Associates. *Development of a Transit Level of Service (TLOS) Indicator - Final Report* Tallahassee, FL, Florida Department of Transportation. Available from http://www.dot.state.fl.us/transit/Pages/tlosfinalrpt.pdf, 1999.

(*15*) Kittelson & Associates. *Transit Capacity and Quality of Service Manual, 2nd Edition - TCRP Report 100* Wahsington, D.C., Transportation Research Board. Available from http://www.trb.org/Main/Blurbs/Transit_Capacity_and_Quality_of_Service_Manual_2nd_153590.aspx, 2003.

(*16*) Kuby, M., Barranda, A., and Upchurch, C. Factors influencing light rail station boardings in the United States. *Transportation Research Part A*, Vol. *38*, 3, 2004, pp. 223-247.

(*17*) Loutzenheiser, D. R. Pedestrian Access to Transit: Model of Walk Trips and Their Design and Urban Form Determinants Around Bay Area Rapid Transit Stations. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. *1604*, 1997, pp. 40-49.

(*18*) O'Neill, W., Ramsey, D., and Chou, J. Analysis of transit service areas using geographic information systems. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. *1364*, 131-139, 1992.

(*19*) O'Sullivan, J. and Morrall, J. Walking Distances to and from Light-Rail Transit Stations. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. *1538*, 1996, pp. 19-26.

(*20*) Polzin, S., Pendyala, R. M., and Navari, S. Development of time-of-day-based transit accessibility analysis tool. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. *1799*, 2002, pp. 35-41.

(*21*) Pushkarev, B. and Zupan, J. M. *Public Transportation and Land Use Policy*, Indiana University Press, Bloomington, IN, 1977.

(*22*) Scheider, S. and Schulz, D. Specifying Essential Features of Street Networks. In S. Winter (Ed.), *Conference on Spatial Information Theory (COSIT'07)*, *LNCS 4736*, Springer, Berlin, 2007, pp. 169-185.

(*23*) Zhao, F., Chow, L.-F., Li, M.-T., Ubaka, I., and Gan, A. Forecasting transit walk accessibility: Regression model alternative to buffer. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. *1835*, 2003, pp. 34-41.

(*24*) Zielstra, D. and Zipf, A. *OpenStreetMap Data Quality Research in Germany*. Sixth international conference on Geographic Information Science (GIScience), Zurich, Switzerland, 2010.