Lecture Notes in Geoinformation and Cartography

LNG&C

Georg Gartner
Haosheng Huang *Editors*

Progress in Location-Based Services 2014

Springer

Lecture Notes in Geoinformation and Cartography

Series editors

About the Series

The Lecture Notes in Geoinformation and Cartography series provides a contemporary view of current research and development in Geoinformation and Cartography, including GIS and Geographic Information Science. Publications with associated electronic media examine areas of development and current technology. Editors from multiple continents, in association with national and international organizations and societies bring together the most comprehensive forum for Geoinformation and Cartography.

The scope of Lecture Notes in Geoinformation and Cartography spans the range of interdisciplinary topics in a variety of research and application fields. The type of material published traditionally includes:

- proceedings that are peer-reviewed and published in association with a conference;
- post-proceedings consisting of thoroughly revised final papers; and
- research monographs that may be based on individual research projects.

The Lecture Notes in Geoinformation and Cartography series also includes various other publications, including:

- tutorials or collections of lectures for advanced courses;
- contemporary surveys that offer an objective summary of a current topic of interest; and
- emerging areas of research directed at a broad community of practitioners.

More information about this series at http://www.springer.com/series/7418

Georg Gartner · Haosheng Huang
Editors

Progress in Location-Based Services 2014

Springer

Editors
Georg Gartner
Haosheng Huang
Department of Geodesy and Geoinformation
Vienna University of Technology
Vienna
Austria

ISSN 1863-2246 ISSN 1863-2351 (electronic)
ISBN 978-3-319-11878-9 ISBN 978-3-319-11879-6 (eBook)
DOI 10.1007/978-3-319-11879-6

Library of Congress Control Number: 2014953956

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Recent years have witnessed rapid advances in location-based services (LBS) with the continuous evolvement of mobile devices and communication technologies. LBS have become more and more popular not only in citywide outdoor environments, but also in shopping malls, museums, and many other indoor environments. They have been applied for emergency services, tourism services, intelligent transport services, gaming, assistive services, etc.

This book provides a general picture of recent research activities related to this field. Such activities emerged in the last years, especially concerning issues of outdoor/indoor positioning, smart environment, spatial modeling, personalization, context awareness, cartographic communication, novel user interfaces, crowdsourcing, social media, big data analysis, usability, and privacy. The innovative and contemporary character of these topics has led to a great variety of interdisciplinary research and studies, from academia to business, from computer science to geodesy.

The contributions in this book are a selection of peer-reviewed full papers submitted to the 11th International Symposium on Location-Based Services in Vienna (Austria) in November 2014, organized by the Research Group Cartography, Vienna University of Technology. We are grateful to all colleagues who helped with their critical reviews. Please find a list of their names in the "Reviewers" section.

The conference series on LBS has been held at

- 2002—Vienna, Austria
- 2004—Vienna, Austria
- 2005—Vienna, Austria
- 2007—Hong Kong, China
- 2008—Salzburg, Austria
- 2009—Nottingham, UK
- 2010—Guangzhou, China
- 2011—Vienna, Austria

- 2012—Munich, Germany
- 2013—Shanghai, China
- 2014—Vienna, Austria

The conferences themselves were a response to an increased interest in providing anyone, anything, anytime, and anywhere services. These conferences together offer a general overview of how LBS-related research has been evolving in the last years. The contributions of this book reflect the recent main areas of interest, including wayfinding and navigation, outdoor and indoor positioning, spatial-temporal data processing and analysis, usability, and application development.

We would like to thank our colleagues Manuela Schmidt, Felix Ortag, Florian Ledermann, and Günther Retscher for their help during the production of this book.

Vienna, Austria, September 2014

Georg Gartner
Haosheng Huang

Contents

Reviewers

The production of this book would not have been possible without the professional help of our scientific committee. We would like to thank all the following experts who have helped to review the papers published in this book.

Suchith Anand, University of Nottingham, UK
Gennady Andrienko, Fraunhofer IAIS/City University London, Germany/UK
Thierry Badard, Laval University, Canada
Rex Cammack, University of Nebraska at Omaha, USA
Cristina Capineri, Siena University, Italy
William Cartwright, RMIT University, Australia
Matt Duckham, University of Melbourne, Australia
Claire Ellul, University College London, UK
Peter Fröhlich, Telecommunications Research Center Vienna, Austria
Georg Gartner, Vienna University of Technology, Austria
Haosheng Huang, Vienna University of Technology, Austria
Mike Jackson, University of Nottingham, UK
Bin Jiang, University of Gävle, Sweden
Jie Jiang, National Geomatics Center of China, China
Markus Jobst, Austrian Federal Office for Metrology and Surveying, Austria
Hassan Karimi, University of Pittsburgh, USA
Farid Karimipour, University of Tehran, Iran
Jukka Krisp, University of Augsburg, Germany
John Krumm, Microsoft Research, USA
Chun Liu, Tongji University, China
Liqiu Meng, Technische Universität München, Germany
Xiaolin Meng, University of Nottingham, UK
Peter Mooney, National University of Ireland Maynooth, Ireland
Takashi Morita, Hosei University, Japan
Hans-Berndt Neuner, Vienna University of Technology, Austria
Ed Parsons, Google, UK
Michael Peterson, University of Nebraska at Omaha, USA

Martin Raubal, Swiss Federal Institute of Technology Zurich, Switzerland
Karl Rehrl, Salzburg Research, Austria
Günther Retscher, Vienna University of Technology, Austria
Tapani Sarjakoski, Finnish Geodetic Institute, Finland
Manuela Schmidt, Vienna University of Technology, Austria
Johannes Schöning, Hasselt University, Belgium
Volker Schwieger, University of Stuttgart, Germany
Stefan van der Spek, Delft University of Technology, The Netherlands
Josef Strobl, University of Salzburg, Austria
Nico van de Weghe, Ghent University, Belgium
Stephan Winter, University of Melbourne, Australia
Kefei Zhang, RMIT University, Australia
Sisi Zlatanova, Delft University of Technology, The Netherlands

Part I
Wayfinding and Navigation

Is OSM Good Enough for Vehicle Routing? A Study Comparing Street Networks in Vienna

Anita Graser, Markus Straub and Melitta Dragaschnig

Abstract As a result of OpenStreetMap's (OSM) openness and wide availability, there is increasing interest in using OSM street network data in routing applications. But due to the heterogeneous nature of Volunteered Geographic Information (VGI) in general and OSM in particular, there is no universally valid answer to questions about the quality of these data sources. In this paper we address the lack of systematic analyses of the quality of the OSM street network for vehicle routing and the effects of using OSM rather than proprietary street networks in vehicle routing applications. We propose a method to evaluate the quality of street networks for vehicle routing purposes which compares relevant street network features as well as computed route lengths and geometries using the Hausdorff distance. The results of our case study comparing OSM and the official Austrian reference graph in the city of Vienna show close agreement of one-way street and turn restriction information. Comparisons of 99,000 route pairs with an average length of 6,812 m show promising results for vehicle routing applications with OSM, especially for route length computation where we found median absolute length differences of 1.0 %.

Keywords OpenStreetMap (OSM) · Volunteered geographic information (VGI) · Quality assessment · Routing · Street networks

1 Introduction

Vehicle routing applications used in route planning, navigation, and fleet management software depend heavily on the quality of the underlying street network data. Errors such as missing streets, wrong or missing turn restrictions or one-way street information lead to wrong route choices and wrong distance estimations.

A. Graser (✉) · M. Straub · M. Dragaschnig
Dynamic Transportation Systems, Mobility Department, AIT Austrian Institute of Technology, Vienna, Austria
e-mail: anita.graser@ait.ac.at

G. Gartner and H. Huang (eds.), *Progress in Location-Based Services 2014*,
Lecture Notes in Geoinformation and Cartography,
DOI 10.1007/978-3-319-11879-6_1

Historically, street network data for vehicle routing applications was only available through a limited number of vendors or official government sources. With the increasing popularity of Volunteered Geographic Information (VGI) projects such as OSM, there is a growing interest in using such free and open data sources in routing applications for different modes of transport.

The adoption of OSM in professional settings is hindered by concerns about the unknown quality of OSM data. Besides simple omission of objects, potential users are also concerned about active vandalism. One important factor is that OSM quality is not consistent between regions. Some countries such as Germany and Austria have large communities of contributors (Neis 2012), which has been found to correlate positively with higher data quality (Neis et al. 2012), while other countries have only smaller groups of contributors. Additionally, OSM quality shows an urban-rural divide (Thaller 2009; Zielstra and Zipf 2010) with better quality in urban regions. It is therefore necessary to evaluate the map in the area of interest with respect to quality for a specific application before OSM can be used. Existing studies describe evaluation methods for quality aspects such as positional accuracy and attribute completeness (for an extensive, non-exhaustive list see OSM Wiki 2014). However, to date there has been no systematic analysis of the quality of the OSM street network for vehicle routing and the effects of switching vehicle routing applications from established proprietary or governmental street networks to OSM.

Aiming to fill this gap, this paper presents a comparative method which evaluates OSM street network quality in comparison to a reference street network. The method is based on comparisons of street network features such as turn restrictions and one-way streets as well as comparisons of routes calculated on both street networks. Route comparisons analyze both route length and route geometry to evaluate the effects of different street networks on the results of vehicle routing applications.

Section 2 presents an overview of related work, followed in Sect. 3 by a description of the methodology used in this study. Section 4 presents results of the comparison of OSM and the official Austrian reference graph in the analysis area. The paper closes in Sect. 5 with a discussion of results and an outlook for future work.

2 Related Work

A first systematic approach for analyzing the quality of OSM is presented by Haklay (2010). He compares OSM to Ordnance Survey UK calculating positional accuracy and comparing network length in regular grid cells covering the study area. Ather (2009) extends Haklay's work, comparing completeness of street names. Subsequent papers compare OSM to other street network datasets such as Ordnance Survey Ireland (Ciepluch et al. 2011), TeleAtlas/TomTom Multi-Net (Thaller 2009; Zielstra and Zipf 2010), or Navteq (Ludwig et al. 2011).

Ludwig et al. (2011) and Koukoletsos et al. (2012) follow a different approach based on matching street objects of OSM and street objects in a reference dataset. This enables them to compare attributes and geometries of specific street objects in two network datasets. Other studies, such as Neis et al. (2012), present OSM-internal checks on validity and topology of the street network.

Work focusing on routing-specific aspects of street networks includes Thaller (2009), who compared three routing examples calculated by OpenRouteService (which uses OSM) and TomTom personal navigation devices, based on route length, travel time estimation, and route choice.

Zielstra and Hochmair H (2012) compare shortest path lengths for 1,000 pedestrian routes of "typical pedestrian walking distance" per city. Routes were calculated using OSM and other free and commercial street network datasets (TomTom, Navteq, TIGER, ATKIS, as well as a combination of networks) in two German and two US cities. They found that OSM provided the most complete data source and the shortest routes—only outperformed by a combination of all available datasets.

On the topic of vehicle routing, Ludwig et al. (2011) found that the "oneway" attribute was missing from 28.1 % of features in inhabited areas and 48.8 % of features in uninhabited areas of Germany when compared to the Navteq street work. Similarly, "speed limit" was found missing for 80.7 % of objects in inhabited areas and 92.6 % of objects in uninhabited areas. Neis et al. (2012) compared the number of turn restrictions found in OSM and TomTom and showed that the number of turn restrictions in OSM is significantly lower (21,000 instead of 176,000 for Germany in June 2011) than in TomTom MultiNet dataset.

As shown in Graser et al. (2014), comparing the number of turn restrictions can cause misleading results due to differences in how street networks are modeled with respect to network generalization and representation of driving restrictions. While a comparison of turn restriction counts between the official Austrian reference graph "Graph Integration Platform" (GIP) and OSM for the greater Vienna region in December 2012 found 2,500 turn restrictions in GIP but only 691 (27.6 %) in OSM, a systematic routing-based comparison showed that 1,515 (60.6 %) of the 2,500 GIP turn restrictions had a matching representation in OSM. Similarly, 10,499 (87.8 %) of the 11,964 one-way streets in GIP could be matched to one-way streets in OSM.

None of the studies so far offer a systematic evaluation of how exchanging an established street network dataset with OSM affects the output of vehicle routing applications with respect to resulting route length and route geometry.

3 Methodology

Our approach comprises the following steps: After the initial preparation of routable graphs, we compare the street networks based on network completeness, similarity of turn restriction and one-way street information, and vehicle routing results.

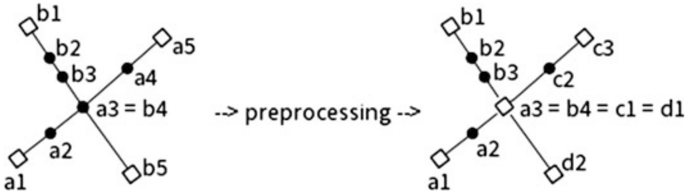


Fig. 1 OSM preprocessing step splitting edges at intersections

3.1 Preparing Routable Graphs

The OSM street network consists of nodes and edges but OSM edges are not necessarily split at each intersection. Instead, in the OSM representation, edges are considered to be connected if they share a common node at the point of intersection. Therefore, OSM is not routable without preprocessing, which splits edges at the appropriate intersections as depicted in Fig. 1. The GIP street network, on the other hand, is modeled using nodes and edges which are split at intersections and connected through explicit turn relations. Without a turn relation, even GIP edges sharing a common intersection node are not considered to be connected. For more details on the different approaches to street network modeling used in OSM and GIP, including a matching of OSM highway tag values and GIP functional road classes, see Graser et al. (2014).

Another aspect where the modeling approaches of OSM and GIP differ is the handling of features such as driving permissions and turn relations. While GIP tends to explicitly define driving permissions for various modes of transport, OSM tends to use conventions and explicit restrictions; i.e. the OSM tag combination vehicle = no and bicycle = yes evaluates to a ban on all vehicles except bicycles. Similarly, GIP turn restrictions are modeled implicitly through missing turn relations, while in OSM, all turn maneuvers are allowed at an intersection as long as there is no explicit restriction relation specified.

In order to create a correct OSM routing graph, edges that share a node at their intersection have to be split up at the intersection node. It is worth noting that the data preparation has to ensure that edges which do not share a common node at the point of intersection are not split in order to avoid creating junctions where there should not be any. This is especially relevant at overpasses and underpasses created by bridges, tunnels or similar network features. Turn restrictions are created from OSM tag combinations that define turn maneuvers and finally, each driving direction and turn restriction is labeled with the modes of transport they concern.

3.2 Street Network Comparison

The comparison of street networks is divided into three parts: (1) assessment of street network completeness; (2) comparison of turn restriction and one-way street information relevant for vehicle routing; (3) comparison of routing results.

The first step is a general comparison of the length of the street networks of OSM and GIP determined by calculating the total sum of the length values of all street network graph edges. This is the most common test for **completeness of street networks** used in Haklay (2010) and numerous subsequent publications. This test can only provide a rough estimate of data completeness since it assumes that both datasets contain similar types of information. Before applying the test, it is therefore necessary to remove road classes which are not represented in both datasets.

In a second step, we compare **turn restriction and one-way street** information relevant for vehicle routing applications. The routing-based comparison method presented in Graser et al. (2014) compares forbidden maneuvers of driving against the one-way street direction and turning at a turn restriction of one street network with routing results calculated on the second street network (see Fig. 2) to test whether both street networks contain matching driving restrictions. Similarity between forbidden turn maneuver and route generated on the second street network is determined using the Hausdorff distance (Hausdorff 1914). A Hausdorff distance above 10 m is interpreted as a correctly modeled turn restriction. Additionally, similarity between a maneuver describing driving against the one-way direction and the route generated on the second street network is determined using length comparison. If the one-way information is present in both street networks, the generated route has to find a way around the driving restriction and will therefore be considerably longer than the forbidden maneuver, which is generated by extracting a 10 m long section from the center of a one-way street (see forbidden maneuver in Fig. 2c, d). A route length above 20 m is interpreted as a correctly modeled one-way street.

The **routing comparison** step of the street network comparison procedure examines routes calculated between identical start and end points. A regular grid is created and used to distribute start and end points in the study area. For each cell

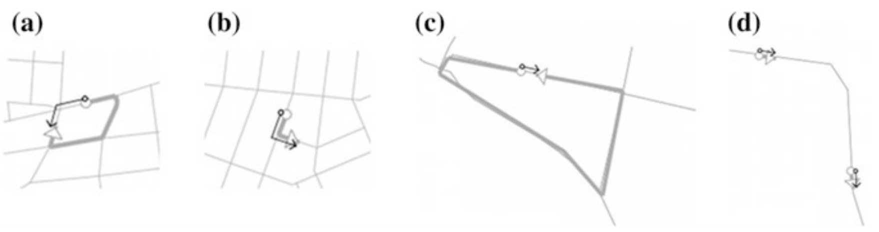


Fig. 2 **a** Correctly modeled turn restriction; **b** Missing turn restriction; **c** Correctly modeled one-way street; **d** Missing one-way restrictions; (*narrow black arrows*: forbidden maneuver; *wide grey arrows*: routes generated on the comparison graph)

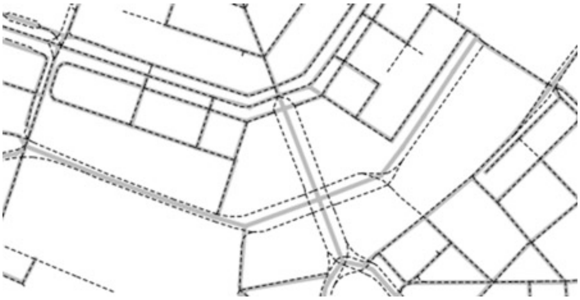


Fig. 3 Details and network generalization differences between OSM (*dashed black lines*) and GIP (*wide grey lines*) at Schwarzenbergplatz

pair consisting of a source and target cell, *n* routes are calculated. Before the routes are calculated, it is necessary to select route start and end points. Distributing start or end points randomly within the cells would lead to ambiguous situations, e.g. if the points end up in the middle between two edges or at an intersection where it is unclear which street should be selected for the route start or end. To minimize these ambiguous situations, we first choose a random network edge within the cell and then select the center point of the edge as start or end point for the route. To select start and end points in the second dataset, a simple map matching routine is applied as follows: the start and end points generated on the first dataset are each matched onto the 13 nearest junctions in the graph and of all incoming and outgoing edges of these junctions, the one with the minimum normal distance is chosen. Finally, the routes are calculated using shortest distance routing with Dijkstra's algorithm (Dijkstra 1959).

The evaluation starts by computing length differences between the routing results on OSM and the reference graph. The distribution of length differences provides a first assessment (see also Fig. 5b). Systematic differences can be observed if routes in one network are systematically shorter than in the other network. Systematic differences might be due to (1) higher road density in one network; (2) lack of driving restrictions; or (3) lack of necessary connections.

Even if the resulting length differences are small or non-existent, this evaluation step only confirms or disproves that route calculations on both graphs result in routes of the same length. While this might be sufficient for certain kinds of vehicle routing applications which only focus on the resulting distance estimates, further evaluations are necessary for applications which depend on calculating correct route geometries, because the first evaluation step cannot confirm whether both route calculations result in the same routes in respect to route geometry. Therefore, in the second part of the routing comparison procedure, the Hausdorff distance is calculated to assess route similarity based on route geometry since it describes the difference between two route geometries independent of the route length.

4 Results

The datasets used in this study are the raw OSM XML data provided by Geofabrik (2013) for March, 19th 2013 and the effective GIP export for routing motorized traffic (called "MIV export") within a 10 × 10 km study area (Fig. 4). In the following comparisons, the GIP export is used as the reference street network graph which OSM is compared with. We want to point out that the methodology could just as well be applied to commercial street network data by providers such as TomTom or Nokia HERE. Since the GIP export does not contain unpaved roads which would be equivalent to the OSM type "track", we removed streets of type "track" from the OSM network.

4.1 Street Network Completeness

A preliminary assessment of OSM and GIP street networks shows that the OSM street network is 1,402 km long and thus 210 km (+17.6 %) longer than the GIP export which is 1,192 km long. Since by removing unpaved roads we ensured that both datasets represent the same road classes, this difference is largely due to the more generalized nature of the GIP export street network which is optimized for routing applications and only contains road center lines as shown in Fig. 3. No generalization was applied to the OSM dataset.

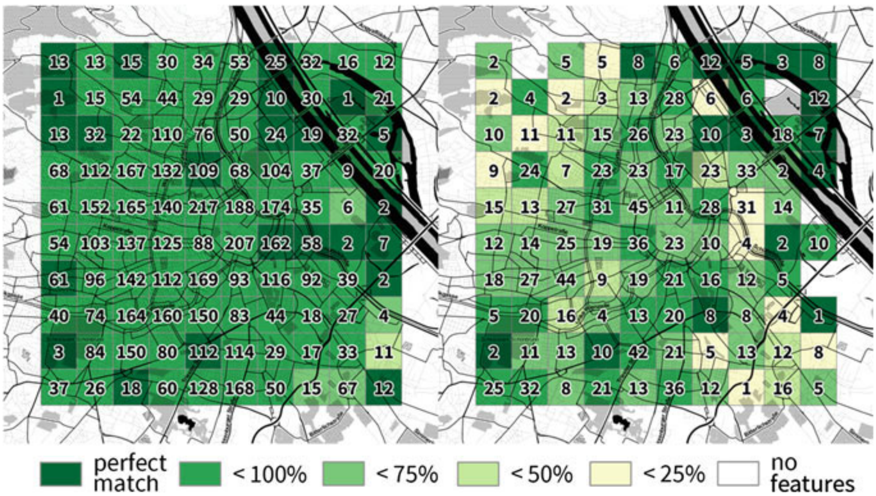


Fig. 4 Spatial distribution of matching one-way streets (*left*), and turn restrictions (*right*); absolute counts (value in the cell) and ratio of matching features (*color*)

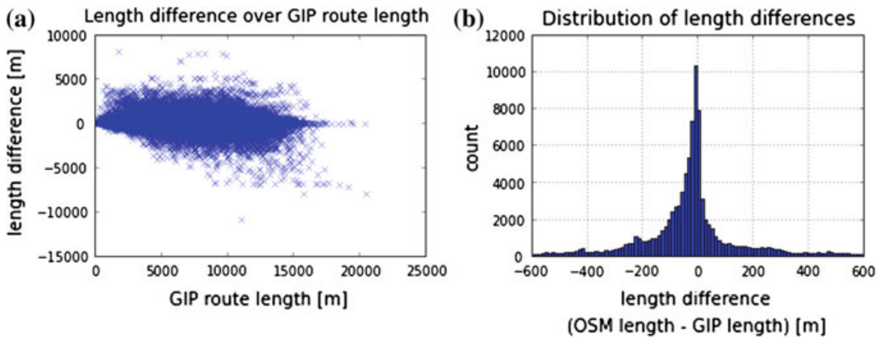


Fig. 5 **a** Length differences depending on GIP route length for individual routes; **b** Distribution of length differences for individual routes

4.2 One-Way Streets and Turn Restrictions

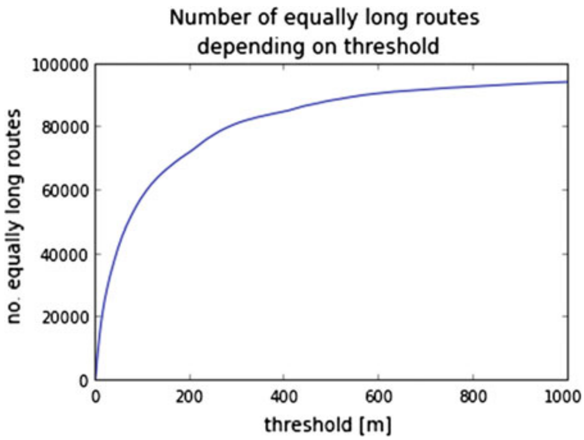
The comparison of one-way streets shows that 6,289 (95.4 %) of the 6,595 GIP one-way streets in the study area can be matched to a one-way street in OSM. Similarly, 842 (68.3 %) of the 1,232 GIP turn restrictions have a matching representation in OSM.

Figure 4 depicts the spatial distribution of one-way streets and turn restrictions in the analysis area. The rate of matching features is color-coded using darker shades for cells with more matches and lighter shades for cells with fewer. The number written inside the cell states the number of occurring one-way streets or turn restrictions in the respective cell. The numbers in the turn restriction map clearly show that turn restrictions are much less common than one-way streets. Additionally, some cells do not contain a single one and are therefore omitted from the turn restriction map. While agreement about one-way streets is high with 91 out of 100 cells with a match ratio better than 80 %, only 30 of the 96 cells which contain turn restrictions reach the same ratio of 80 % matching features in OSM and GIP.

4.3 Routing Comparison

In this case study, we used a grid with 100 1 × 1 km cells and calculated ten route pairs consisting of an OSM route and a GIP route per cell pair. This leads to a total of 99,000 route pairs with an average GIP route length of 6,812 m (min: 54 m; max: 20,465 m). We calculate each route pair's length difference as OSM route length minus GIP route length. Negative difference values therefore stand for shorter OSM routes. Based on all 99,000 routes, the mean length difference is −15.5 m and the median length difference –17.3 m. These results show that OSM routes tend to be shorter than the corresponding GIP routes. Figure 5a depicts the relation of length

Fig. 6 Number of route pairs with equally long OSM and GIP routes depending on threshold



difference and GIP route length per route pair. High negative length differences are found for long GIP routes over 5 km length while positive length differences are also found for shorter routes. Figure 5b depicts the overall distribution of length differences per route pair and clearly shows the trend of shorter OSM routes in the shift towards negative length differences.

We have seen the trend towards shorter OSM routes, but how often can OSM and GIP routes be considered equally long for the purpose of vehicle routing applications? Figure 6 presents the number of route pairs with equally long OSM and GIP routes depending on the threshold chosen to define "equally long": at a threshold tolerance of ±10 m, 15,874 (16.0 %) of the total 99,000 routes are considered equally long. For a threshold of ±25 m this value rises to 29,325 (29.6 %), growing to 58,373 (59.0 %) for a threshold of ±100 m. Additional evaluations of absolute length differences in relation to GIP route length show that the median OSM route length deviates by 1.0 % from the corresponding GIP route length.

To gain a better understanding of the spatial distribution of route pairs with similar route length and those with bigger deviations, we further evaluate the length difference values grouped by route starting cell. Mean route length by cell varies between 5 and 9 km depending on whether the cell is located in the center of the analysis area or around its border. Figure 7b shows that, in most cells, OSM routes are shorter than GIP routes, confirming our previous interpretation of individual route results, while Fig. 7a depicts the same median length difference values plotted over the median GIP route length. This confirms the intuition that higher length difference values are found for cells with longer GIP routes.

Figure 8 depicts the length differences for all route pairs starting in the respective cell. Cells in the center of the grid generally show lower median length difference values than cells around the border of the analysis area. While cells with high negative median difference values—indicating that OSM routes starting there are considerably shorter than the respective GIP routes-cluster in the northwest, cells

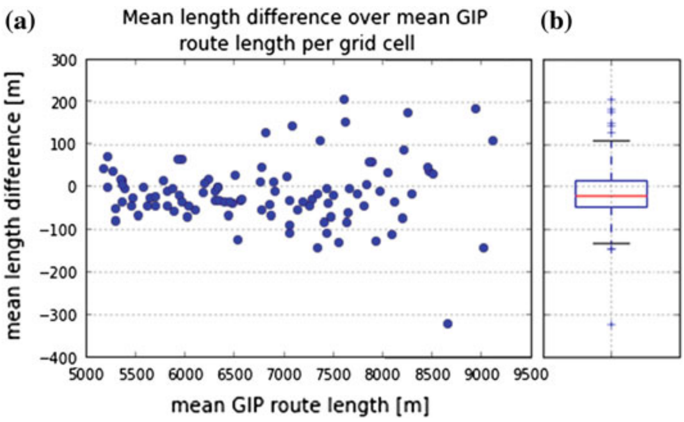


Fig. 7 **a** Median length difference over median GIP route length per route starting grid cell; **b** Distribution of median length differences per route starting grid cell

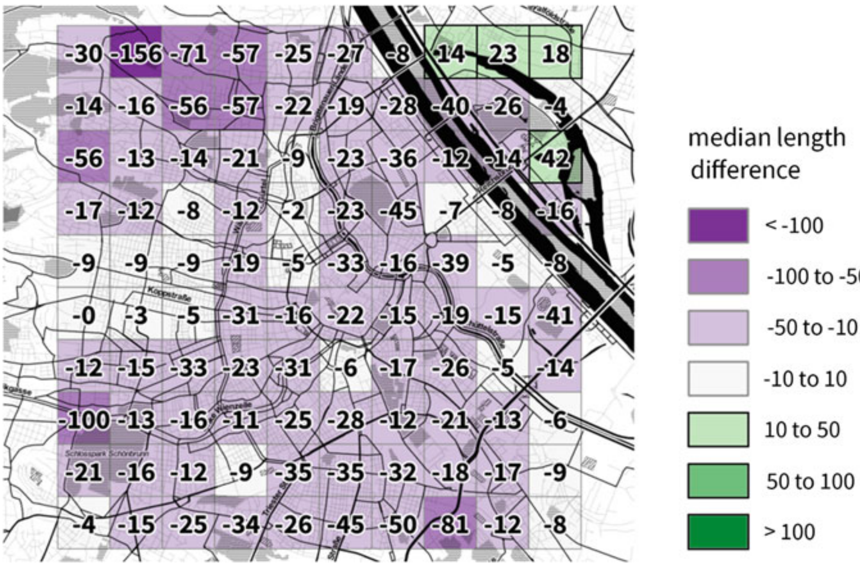


Fig. 8 Median length difference in meters for all route pairs starting in the given cell

with positive values are found in the northeast of the analysis area, north of the river Danube where most of the routes have to cross the Danube bridges to the southern part of Vienna. In any case, it has to be noted that route length differences are accumulated along the whole route, and the underlying street network deviations causing the differences are therefore not necessarily located in the route starting cell.

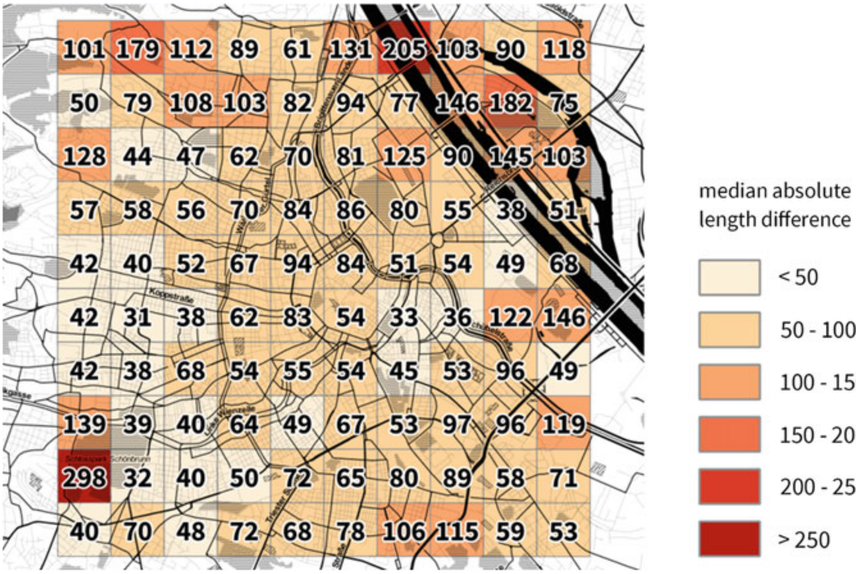


Fig. 9 Median absolute length difference in meters for all route pairs starting in the given cell

Figure 9 shows the median absolute length difference thus highlighting areas with bigger length differences independent of whether OSM or GIP routes tend to be shorter. As before, higher values are found at the borders of the analysis area but the focus of highest difference values has shifted to areas southwest of Schlosspark Schönbrunn (298 m difference) and along the Danube (205 m difference). Closer inspection of the routes starting in the grid cell near this park shows that the cell is mostly covered by the park and contains only a very limited number of street edges (< 20 in either datasets). As a result, the algorithm picking route start and end points ends up picking from this small set of edges over and over for each cell pair and thus a single network difference can affect multiple routes.

To gain a better understanding of the sources of the length differences, 25 routes were inspected manually. For all randomly selected and inspected routes with very large differences (eight kilometers) the reasons were that the automatic matching process selected topologically different start or end edges, e.g. motorway links instead of motorway exit. For further studies, it is recommended to take special care that the chosen start and end edges of the routes match to the same logical edges in the road graphs.

Other causes for length differences were map defects or inaccuracies, such as missing or wrong information about where motor vehicles are allowed to drive, different one-way information, missing or wrong turn restrictions, and different lengths of dead end streets.

After these length-based comparisons, the following sections present the results of comparing route geometries. In this study, similarity of route pair geometries was

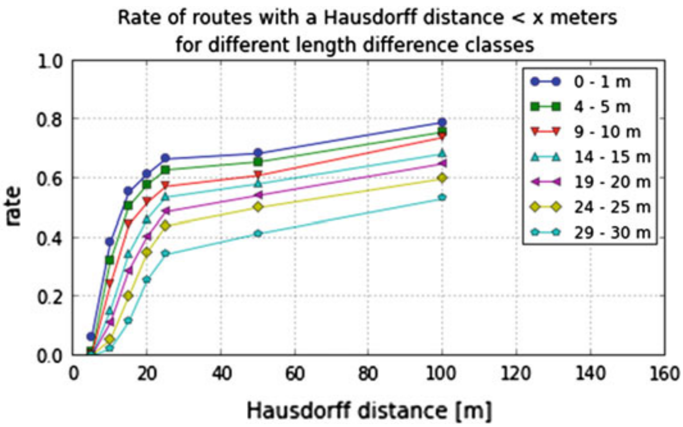


Fig. 10 Rate of routes with a Hausdorff distance under a given threshold

calculated as the Hausdorff distance between OSM and GIP route since it describes the difference between two route geometries independent of the route length. Figure 10 shows how the rate of route pairs with OSM and GIP route geometries which are considered similar grows as the Hausdorff distance threshold, which is used to define “similar”, is increased. The figure shows separate curves for several length difference classes from the top-most curve which represents only route pairs with a length difference of 0–1 m to the curve at the bottom which represents route pairs with a length difference of 29–30 m. These results show a steady increase of similar routes up to a threshold of 25 m which slows down considerably afterwards.

Based on these results, Fig. 11 depicts the relation between length difference of OSM and GIP routes and the rate of route pairs with similar route geometries defined by a Hausdorff distance less than 25 m. A threshold of 25 m was chosen since it is well below the size of typical city blocks, thus eliminating routes which deviate by one city block while allowing for smaller deviations caused by different levels of street network generalization. The graph in Fig. 11 shows a linear

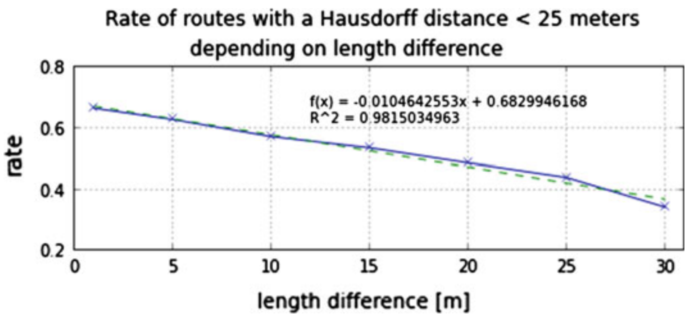


Fig. 11 Rate of routes with a Hausdorff distance under 25 m over length difference

relationship with a high coefficient of correlation $R^2 = 0.98$. Since calculating Hausdorff distance is computationally expensive, we use the resulting linear relationship

$$-0.0104642553 * \text{absolute\_length\_difference} + 0.6829946168 \quad (1)$$

to estimate the total number of route pairs with both a length difference less than 25 m and a Hausdorff distance less than 25 m. Of the total 99,000 routes in the sample, 16,903 (17.1 %) route pairs fall into this class and are therefore considered to be a perfect match for the purpose of this study.

5 Conclusion and Future Work

The comparison of OSM and GIP for vehicle routing application presented in this work investigates the influence of switching from GIP to OSM on resulting shortest path route length and route geometry. The study covers comparisons of street network completeness, turn restriction and one-way street information as well as 99,000 routes with a mean GIP route length of 6,800 m in the city of Vienna.

Route length comparison results show that for 59.0 % of routes, the computed OSM route length is within a tolerance of 100 m of the corresponding GIP route length, and for 29.6 % of routes, OSM route length is within a tolerance of 25 m of the corresponding GIP route length. Observed route length differences vary by location but it has to be noted that route length differences are accumulated along the whole route and, as a result, locating the street network deviations causing the differences therefore is no trivial task. OSM routes tend to be shorter than GIP routes which could be explained by two factors: first, the OSM network could be denser than the GIP network and thus contain more “shortcuts”, be they right or wrong; and second, the OSM street network could contain fewer driving restrictions and thus be more connected. While comparisons of street network length show that the OSM street network within the analysis area is 17.6 % longer than the GIP street network, the differences are mostly due to the more generalized nature of the GIP export and not due to additional connecting streets in the OSM street network. Regarding driving restrictions, a comparison of one-way streets shows that 95.4 % of the 6,595 GIP one-way streets in the study area can be matched to a one-way street in OSM and similarly, 68.3 % of the 1,232 GIP turn restrictions have a matching representation in OSM. Differences in the remaining turn restrictions and one-way streets will influence route length and geometry deviations. Based on the results of an evaluation of absolute length differences relative to GIP route length, vehicle routing applications that compute route length based on OSM instead of GIP would result in routes with a median absolute length difference of 1.0 % relative to the original GIP route length.

To further evaluate the similarity of routing results, Hausdorff distance between OSM route geometry and the corresponding GIP route geometry was calculated.

Results of this evaluation show that 17.1 % of route pairs have both a length difference less than 25 m as well as a Hausdorff distance less than 25 m and are therefore considered to be a perfect match for the purpose of this study. It has to be noted that due to the varying quality of OSM, applying the analysis methods presented in this study in other geographic regions might result in significantly different results.

Expected correlations between shorter route length and a better agreement of routing results both in respect to route length and geometry should be investigated in subsequent studies. Further work is planned to evaluate the effect of migrating to OSM on specific vehicle routing applications, such as floating car data systems that require routing between successive vehicle positions, which are sampled at intervals up to two minutes, leading to considerably shorter routes than the ones evaluated in this study.

Acknowledgements This work is funded by the Austrian Climate and Energy Fund within the "Electric Mobility flagship Projects" program (project "emporA 2") and the KLIEN initiative of the Austrian bmvit (project no. 839,478 "Crossing Borders").

References

Ather A (2009) A Quality Analysis of OpenStreetMap Data. Master's thesis, University College London

Ciepluch B, Mooney P, Jacob R, Zheng J, Winstanely AC (2011) Assessing the quality of open spatial data for mobile location-based services research and applications. Arch Photogrammetry, Cartography Remote Sens 22:105–116

Dijkstra EW (1959) A note on two problems in connexion with graphs. Numer Math 1:269–271

Geofabrik (2013) http://www.geofabrik.de/data/download.html. Accessed 19 March 2013

Graser A, Straub M, Dragaschnig M (2014) Towards an open source analysis toolbox for street network comparison: indicators, tools and results of a comparison of OSM and the official Austrian reference graph. Transactions in GIS 18:510–526. doi:10.1111/tgis.12061

Haklay M (2010) How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environ Plan 37(4):682–703. doi:10.1068/b35097

Hausdorff F (1914) Grundzüge der Mengenlehre

Koukoletsos T, Haklay M, Ellul C (2012) Assessing Data Completeness of VGI through an Automated Matching Procedure for Linear Data. Trans GIS 16(4):477–498

Ludwig I, Voss A, Krause-Traudes M (2011) A Comparison of the Street Networks of Navteq and OSM in Germany. Springer, Advancing Geoinformation Science for a Changing World, pp 65–84

Neis P (2012) Distribution of Active Users in OpenStreetMap – Oct-Nov 2012. http://neis-one.org/2012/11/active-users-osm-nov12/. Accessed 11 May 2014

Neis P, Zielstra D, Zipf A (2012) The Street Network Evolution of Crowsourced Maps: OpenStreetMap in Germany 2007-2011. Future Internet 4:1–21. doi:10.3390/fi4010001

Thaller D (2009) Die Open-Source-Plattform "OpenStreetMap", eine Konkurrenz für Geodatenhersteller?. Master's thesis, Universität Wien. Fakultät für Geowissenschaften, Geographie und Astronomie

OSM Wiki (2014) Research. http://wiki.openstreetmap.org/w/index.php?title=Research&oldid=995511. Accessed 10 May 2014

Zielstra D, Hochmair H H (2012) Comparison of Shortest Path Lengths for Pedestrian Routing in Street Networks Using Free and Proprietary Data. Proceedings of Transportation Research Board–Annual Meeting, Washington, DC, USA, 22–26 January 2012

Zielstra D, Zipf A (2010) A Comparative Study of Proprietary Geodata and Volunteered Geographic Information for Germany. AGILE 2010. Guimaraes, Portugal