Generation of a tailored routing network for disabled people based on collaboratively collected geodata

Pascal Neis a,*, Dennis Zielstra b

a Department of Geography, Geoinformatics Research Group, University of Heidelberg, Berliner Str. 48, D-69120 Heidelberg, Germany
b Geomatics Program, University of Florida, 3205 College Avenue, Fort Lauderdale, 33314 FL, USA

A B S T R A C T

The generation of a routing network for disabled people inherits a number of prerequisites that need special consideration. Widespread routing applications that rely on commercial or governmental geodata sources are not feasible for this specific task, due to the lack of detailed information about features such as sidewalks, surface conditions or road incline. In recent years the research community has experienced a strong increase in studies related to routing applications tailored to disabled people in which the lack of a sophisticated dataset played a major role. This study proposes an algorithm for the generation of a disabled people friendly routing network, based on collaboratively collected geodata provided by the OpenStreetMap (OSM) project. This new representation of a routing graph can be used in numerous applications and maps dedicated to people with disabilities. The algorithm is tested and evaluated for selected areas in Europe, resulting in newly generated extended networks that include sidewalk information. The results have shown that the success of the final implementation of the introduced algorithm depends highly on the attribute quality of the OSM dataset.

© 2013 Elsevier Ltd. All rights reserved.

Introduction

Routing and navigation applications on the Internet, in cars or on personal smartphones are omnipresent. Most common devices and applications rely on geodata provided by one of the well-known proprietary data providers such as Naveq® or TomTom®. These providers offer routing network data which is suitable for motorized and (for selected cities) non-motorized path finding applications. People with special needs, however, who rely on a more specialized dataset, cannot utilize the provided commercial geo-information and require highly detailed ground-truth data. Commercial geo-data providers do not offer this detailed information due to the high costs that arise during the collection and the maintenance of the data.

In the past few years the number of freely available and open source geo-information platforms on the Internet has increased tremendously. These new data sources are oftentimes referred to as Volunteered Geographic Information (VGI; Goodchild, 2007). As the name implies, most of these platforms rely on the contributions of non-professional volunteers that collaboratively collect geodata. A number of possible motivational factors that trigger VGI project contributions has been identified in a recent study, including the desire to make geospatial information freely available to everyone, learning new technologies, relaxation and recreation, self-expression or just pure fun (Budhathoki & Haythornthwaite, 2012). The contribution patterns found in VGI projects tend to be more casual in comparison to the contributions made to Public Participation Geographic Information Systems (PPGIS) in which volunteers collect geodata for a particular purpose, such as to improve landuse planning or discuss policy issues and decision making (Brown, 2012). One of the biggest and most established projects in the realm of VGI is OpenStreetMap1 (OSM). In contrast to the aforementioned proprietary data providers, the OSM project data is distributed under an Open Data Commons Open Database License (ODbL2). This particular license allows interested Internet users to download, copy, distribute, transmit and adapt the collected geodata, free of charge, as long as OSM and its contributors are credited in the final project.

Despite early concerns about the credibility and reliability of VGI (Flanagin & Metzger, 2008) several studies demonstrated the potential of OSM in a variety of applications in recent years. OSM data has been utilized to develop a number of Location Based Services

---

2 http://opendatacommons.org/licenses/odbl/ (accessed on 27 November 2013).
Routing network requirements for disabled people

Several studies in the past have highlighted the prerequisites that the geodata source of choice has to fulfill to be considered for a potential navigation system for pedestrians (Gaisbauer & Frank, 2008), wheelchair users (Charles, Kincho, Jean-Claude, & John, 2002; Kasemsuppakorn & Karimia, 2009) or blind people (Kammoun et al., 2010). Oftentimes the customized system and its corresponding data are created through extensive surveys. A specification by the German Institute for Standardization (Deutsches Institut für Normung (DIN)) provides a foundation for this particular type of information. DIN 18024-1 describes the accessibility requirements for disabled people in public transit infrastructure and buildings. The standards include a number of recommendations for different handicap types, which also help to define the target user group for which our study was conducted: (Source: DIN 18024-1):

- Wheelchair users
- Blind and visually impaired people
- Deaf and hearing impaired people
- Walking impaired people
- People with other handicaps
- Elderly people
- Children and people of short or tall stature

Based on the specification, some of recommended parameters that need to be implemented in the final dataset can be surface information, incline and width of a street segment. However, based on a number of different studies, other parameters for a disabled friendly routing network have been determined (Beale, Field, Briggs, Picton, & Matthews, 2006; Ding et al., 2007; Kasemsuppakorn & Karimia, 2009; Matthews, Beale, Picton, & Briggs, 2003; Menkens et al., 2011; Sobek & Miller, 2006). Table 1 summarizes all parameters based on the findings of the studies, the DIN 18024-1 and some newly defined parameters based on our research.

In some of the studies the desired geodata was traced from satellite imagery (Kasemsuppakorn & Karimi, 2008; Kasemsuppakorn & Karimia, 2009), while others developed tools that generated a network by utilizing a buffer method (Karimi & Kasemsuppakorn, 2012), implementing pedestrian GPS traces (Kasemsuppakorn & Karimi, 2013), developing a binary image processing method to retrieve a pedestrian network (Gaisbauer & Frank, 2008; Kim, Park, Bang, & Yu, 2009) or presented an automated method to generate a sidewalk network from building blocks (Ballester, Pérez, & Stuiver, 2011).

Collaboratively collected geodata: the OpenStreetMap project

User-Generated Content (UGC) (Anderson, 2007) and particularly Volunteered Geographic Information (VGI) (Goodchild, 2007) have become a widely known Internet phenomenon in recent years. The OSM project, initiated in 2004, is the most successful VGI project based on collaboratively collected and freely available geodata (Goetz, 2012a; Mooney, Corcoran, & Winstanley, 2010; Neis, Goetz, & Zipf, 2012). Most contributors collect the geodata by utilizing GPS

---

Based on VGI is the increased data collection efforts by the OSM applications such as Location Based Services or 3D applications compared to commercial or governmental dataset distributors. Zielstra and Hochmair (2011b, 2012) have shown that the OSM project can be treated and evaluated individually (an application has different requirements to the dataset and needs on which side? comparison can differ in terms of data quality and number of active contributors to map features of the real world (e.g. a bus or tram line of the public transportation network). Attribute information about objects is added by applying Tags consisting of a key–value pair. A comprehensive list of OSM key–value pairs for a large number of map features is available on one of the OSM related wiki pages. However, it needs to be noted that this list does not represent a strict specification or standardization, which means that each contributor can assign keys or values based on her/his own understanding and preference. Brando and Bucher (2010) and Girres and Touya (2010) criticized this tagging procedure in OSM and suggested that the data quality can be improved by using predefined specifications for objects and their corresponding tags. Nevertheless, the current tagging implementation is an essential part of the open approach to data contributions in OSM (Neis, Goetz et al., 2012).

The OSM project has three different object types that allow the active contributors to map features of the real world (Ramam, Topf, & Chilton, 2010). A Node object represents a point feature with its latitude and longitude coordinates, whereas a Way object is utilized to represent streets or closed line areas (i.e. polygons) such as landuse information or buildings. The Relation object contains information on how two or more objects are related to each other (e.g. a bus or tram line of the public transportation network). Attribute information about objects is added by applying Tags consisting of a key–value pair. A comprehensive list of OSM key–value pairs for a large number of map features is available on one of the OSM related wiki pages. However, it needs to be noted that this list does not represent a strict specification or standardization, which means that each contributor can assign keys or values based on her/his own understanding and preference. Brando and Bucher (2010) and Girres and Touya (2010) criticized this tagging procedure in OSM and suggested that the data quality can be improved by using predefined specifications for objects and their corresponding tags. Nevertheless, the current tagging implementation is an essential part of the open approach to data contributions in OSM (Neis, Goetz et al., 2012).

The default OSM dataset is not applicable for routing or navigation purposes. Renz and Wölfli (2010) and Schmitz, Zipf, and Neis (2008) introduced different methods on how to generate a routing network based on OSM data. These initial concepts were implemented in the first processing step of the disabled friendly routing network generation.

### Summary of required parameters for the generation of a routing network for disabled people.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of street</td>
<td>Ways which can be used for a routing network for disabled people</td>
<td>8</td>
</tr>
<tr>
<td>Sidewalk</td>
<td>Has the street a sidewalk, and if yes on which side?</td>
<td>1–8</td>
</tr>
<tr>
<td>(Sidewalk) Width</td>
<td>Width of the street/sidewalk</td>
<td>1–8</td>
</tr>
<tr>
<td>(Sidewalk) Surface</td>
<td>Surface of the street/sidewalk</td>
<td>1,2,4–8</td>
</tr>
<tr>
<td>(Sidewalk) Smoothness</td>
<td>Smoothness of the street/sidewalk</td>
<td>1,4–8</td>
</tr>
<tr>
<td>(Sidewalk) Slope/Incline</td>
<td>Incline of the street/sidewalk</td>
<td>1,2,4–8</td>
</tr>
<tr>
<td>(Sidewalk) Camber</td>
<td>Camber of the street/sidewalk</td>
<td>1,2,4,7</td>
</tr>
<tr>
<td>(Sidewalk) Curb/Kerb</td>
<td>Sliced curb (height)</td>
<td>1–4–6–8</td>
</tr>
<tr>
<td>(Sidewalk) Curvature</td>
<td>Curvature of the street/sidewalk</td>
<td>2,7</td>
</tr>
<tr>
<td>Lighting</td>
<td>Is the street lighted?</td>
<td>4,7,8</td>
</tr>
<tr>
<td>Tactile Paving</td>
<td>Is tactile paving available?</td>
<td>7,9</td>
</tr>
<tr>
<td>Steps</td>
<td>Number of steps</td>
<td>1–3–5–8</td>
</tr>
<tr>
<td>Step height</td>
<td>Height of the individual steps</td>
<td>3,7</td>
</tr>
<tr>
<td>Ramp</td>
<td>Is a ramp (at the steps) available?</td>
<td>1–3–6–8</td>
</tr>
<tr>
<td>Handrail</td>
<td>Is a handrail railing (at the steps/ramp) available?</td>
<td>7</td>
</tr>
<tr>
<td>Crossing</td>
<td>Crossing (with/without Traffic signals)</td>
<td>1,2,7,8,9</td>
</tr>
<tr>
<td>General Access</td>
<td>General access information of the street/sidewalk</td>
<td>3,8</td>
</tr>
</tbody>
</table>

**Notes:** 1Matthews et al. (2003); 2Beale et al. (2006); 3Soheki and Miller (2006); 4Bing et al. (2007); 5Kasemuppakorn and Karimia (2009); 6Menkens et al. (2011); 7DIN 18024-1; 8Our research.
The creation of the routing graph is followed by the identification of the relevant OSM tags. Nearly all of the aforementioned special requirements for disabled people (Section 2.1) are mapped in OSM in some way or another. The representation of sidewalks in the OSM database plays a major role in this particular case. A sidewalk is only mapped as a separate feature if the sidewalk is not in close proximity to the street (Ramm et al. 2010). In all other cases the information of the sidewalk is part of the street object, e.g. sidewalk:left:surface = good. There are multiple OSM values with different key combinations that can be utilized for our purpose. Table 2 matches the prerequisites of a disabled friendly routing network (Table 1) with the corresponding OSM Tags. Overall only two parameters shown in Table 1 cannot be found in the OSM mapping schema: the camber and curvature of a sidewalk.

**Generation**

The generation of the sidewalk routing network consists of several geometric processes. Fig. 1 illustrates the individual steps of the algorithm. In Step 1 junctions are created, which consist of three ways and one node. Each way has a sidewalk declaration in the OSM database. In Step 2 a temporary line running parallel to each way segment is generated for each side at which a sidewalk exists. The newly generated lines represent the temporary paths for pedestrians and wheelchair users. During the generation of these temporary paths the way type, documented in the OSM database, is taken into account too. For instance, the temporary line for a tertiary road will be created with a distance of 5 m, while in the case of a residential road a distance of 3.5 m will be applied. The distances are based on guidelines provided by the German “Forschungsgesellschaft für Straßen- und Verkehrswesen (FGSV),” which include detailed information about the construction of roads and other infrastructure. Furthermore each sidewalk parameter (e.g. surface or width) is transferred from the initial line to the temporarily generated sidewalk line. In Step 3 the final sidewalk geometries are connected to their corresponding junction node. If a connection between two sidewalks crosses a way of the initial OSM network, a crossing between the two sidewalks will be created (see Fig. 1, Step 3). The last step (Step 4) removes all ways of the initial network that have a newly generated sidewalk representation. The final image in Fig. 1 shows the routing network generated by the algorithm as an overlay on an OSM basemap.

**Evaluation**

The prototype of the algorithm was tested for all capital cities of the 50 sovereign European states. For each city a test region was extracted from the OSM dataset using a circular polygon with a radius of 2 km around each city center. The position of the city center, especially in the Northeast and Southwest areas (Fig. 2a), lies in one city district. Next to Berlin proved to have good sidewalk information coverage with a decline in information concentration when moving away from the center, especially in the Northeast and Southwest areas (Fig. 2a). Most sidewalk information in Riga (Fig. 2b) lies in one city district east of the Daugava River, whereas in London (Fig. 2c) the majority of the required information is only distributed along the main roads.

To evaluate the efficiency of the presented algorithm, 100 shortest paths between random start and end points in each test area were calculated. For comparison purposes two paths were generated for each city. The first path was computed on the regular street network graph, whereas the computation of the second path was based on the newly generated sidewalk graph (Table 3). Next to the total length comparison between both paths, indicating potential detours due to errors of omission or commission, a buffer comparison method introduced by Goodchild and Hunter (1997) was applied to test if the computed route geometries of the sidewalk graph differ from the routes of the regular street network. A buffer of 10 m on each side of the generated routes was applied and evaluated in Table 4. The parsing, processing and generation of the sidewalk network, was implemented in JAVA programming language and took less than 8 s for each city. Table 4 contains more information and general statistics for each of the three test areas. The values provided in the "Generated Sidewalk Network Length" column contain the total length of all features with at least one sidewalk Tag in the OSM dataset. If a street has a sidewalk on both sides the length of the feature is only counted once.

![Fig. 1](http://wiki.openstreetmap.org/wiki/Nominatim) (accessed on 5 October 2013).

![Fig. 2](http://planet.osm.org) (accessed on 5 October 2013).

![Table 2](http://wiki.osm.org/wiki/Osmosis) (accessed on 5 October 2013).

![Table 3](http://wiki.openstreetmap.org/wiki/Osmosis) (accessed on 5 October 2013).

![Table 4](http://wiki.openstreetmap.org/wiki/Osmosis) (accessed on 5 October 2013.)

Notes.

a Additional high-way-values: primary, primary_link, secondary, secondary_link, tertiary, tertiary_link, unclassified, living_street, pedestrian, residential, service, track, footway, cycleway, bridleway, steps ("only if accessible for pedestrians/wheelchairs").

b Additional surface-values: paved, asphalt, concrete, paving_stones, cobblestone, concrete_plates.

c Additional smoothness-values: excellent, good, intermediate, bad, very_bad.

d Currently a proposed OSM tag.
the percentage of overlap between the buffers was determined. The results showed that the largest total length difference can be found in London, combined with the lowest polygon overlap value, indicating slightly different routes between the two generated networks.

Next to the aforementioned factors, it is important to evaluate whether the computed path, based on the newly presented approach, exists only along major street types, such as primary or secondary roads, or if it also contains footways or sidewalks, i.e. ways that are not accessible to motorized traffic. Fig. 3 illustrates the number of road features that were utilized during the generation of the routes based on the regular road network in each city. Additionally, the corresponding percentage of footway information that was implemented in the total route length was computed. The results show that the generated routes for Riga and London have a higher percentage while Berlin reveals the lowest value in this comparison. These results can be compared to the percentage of footway information utilized during generation of the tested routes based on the newly generated sidewalk network (Fig. 4). All three diagrams show an improvement in the number of footway features. Although London includes less sidewalk information in the OSM dataset in comparison to Berlin, the tested area in London still shows a similar or slightly better result. Similarly good results can
be reported for Riga, where the test dataset only contains about 36% of the sidewalk information that is needed. However, the majority (89%) of the calculated routes implement more than 60% of footway or sidewalk information.

Further, the quantity of the previously introduced crucial tags for a disabled friendly routing network was evaluated (cf. Tables 1 and 2). Table 6 shows the percentages of features that were tagged with the additional information that is needed to create the desired network. Some of the introduced tags shown in Table 2 were missing entirely in the three tested areas. However, an additional visual inspection of the selected 50 datasets showed that some contributors use a point as decimal mark while others prefer to use a comma. Others switch between meter and centimeter units when collecting information about the width or the sloped curb of a sidewalk. Other contributors again attach the units of their measurements directly to the value of the object. Besides these errors in naming conventions when tagging an object in OSM, other information in the database is sometimes not interpretable. For instance the key incline, which describes the slope of a street, was used for about 78,000 ways (according to an OSM tag information webpage). 42% of the values of this particular key include information such as ‘up’ and 26% are tagged with ‘down’. This additional information, whether the slope value was taken when going ‘down’ or ‘up’ the road, renders useless when generating a routing network for wheelchair users. This means that almost 68% of the information retrieved from the incline tag uses a temporary value such as “up” or “down” which indicates that further information is needed.

A similar issue can be detected when utilizing the key ‘sloped_curb’. The OSM wiki contains detailed information about how the kerb of a sidewalk should be tagged. For our analysis the key ‘sloped_curb’ was implemented due to its importance on the wheelchair routing webpage. Several other documentations also recommend using the key ‘kerb’, sometimes also referred to as ‘curb’. Next to the different naming conventions, a second ambiguity with this particular tag arises when determining the exact

### Table 6
Completeness of disabled routing related sidewalk information.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Percentage in mapped sidewalks</th>
<th>Berlin</th>
<th>Riga</th>
<th>London</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lighting</td>
<td>20.6%</td>
<td>74.2%</td>
<td>91%</td>
<td></td>
</tr>
<tr>
<td>Smoothness</td>
<td>1.6%</td>
<td>0%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Surface</td>
<td>28.2%</td>
<td>44.4%</td>
<td>9.8%</td>
<td></td>
</tr>
<tr>
<td>Width</td>
<td>0%</td>
<td>0%</td>
<td>2.1%</td>
<td></td>
</tr>
<tr>
<td>Number of mapped crossings</td>
<td>79</td>
<td>253</td>
<td>458</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3. Percentage of footway feature lengths.

Fig. 4. Percentages of footway and sidewalk information in routes with sidewalks.

Limitations

During the development of the network and the testing process of the algorithm several problems occurred when utilizing the OSM dataset. As the evaluation of the algorithm has shown, the geodata quality has to be tested for the individual use case. This means that the algorithm can only generate an adequate network if the corresponding sidewalk information is available in the area of interest. A second major issue is the completeness and variety of keys and values that the OSM contributors can apply to the individual objects. The collected information in the tested areas for instance showed, that some contributors use a point as decimal mark while others prefer to use a comma. Others switch between meter and centimeter units when collecting information about the width or the sloped curb of a sidewalk. Other contributors again attach the units of their measurements directly to the value of the object. Moreover, there are different naming conventions when tagging an object in OSM, other information in the database is sometimes not interpretable. For instance the key incline, which describes the slope of a street, was used for about 78,000 ways (according to an OSM tag information webpage). 42% of the values of this particular key include information such as ‘up’ and 26% are tagged with ‘down’. This additional information, whether the slope value was taken when going ‘down’ or ‘up’ the road, renders useless when generating a routing network for wheelchair users. This means that almost 68% of the information retrieved from the incline tag uses a temporary value such as “up” or “down” which indicates that further information is needed.

A similar issue can be detected when utilizing the key ‘sloped_curb’. The OSM wiki contains detailed information about how the kerb of a sidewalk should be tagged. For our analysis the key ‘sloped_curb’ was implemented due to its importance on the wheelchair routing webpage. Several other documentations also recommend using the key ‘kerb’, sometimes also referred to as ‘curb’. Next to the different naming conventions, a second ambiguity with this particular tag arises when determining the exact

---

9 http://taginfo.openstreetmap.org/keys/incline#values (accessed on 5 October 2013).
location of the kerb information. Where should the contributor add this information? Should a node be added to the start and the end of a way or should it be added as a tag to the way (e.g. ‘sidewalk:start:kerb’ and ‘sidewalk:end:kerb’)? A standardized tagging convention in this particular case would improve the OSM quality significantly.

However, one of the main questions that arise is: Do contributors map this detailed information worldwide although it is not being rendered in the OSM standard maps? At least in recent years the volunteers started collecting detailed information beyond the scope of regular streets or buildings. A few years ago, the OSM dataset did not provide any turn restriction or detailed address information for navigation applications. After the community was introduced to applications that utilize this information, there was an increase in mapping and tagging efforts for these particular attributes.

Conclusions and future work

In this article we introduced a newly developed algorithm that generates a routing network for disabled people from a freely available and collaboratively collected geodataset, provided by the OSM project. The newly created network proved to have several advantages over traditional routing networks and is highly adaptable. The variety of supported attributes during the network generation allows the algorithm to be used for different use cases such as routeplanners or personal navigation assistants for people with disabilities. Furthermore, the new representation of a sidewalk network can be implemented in several types of online, offline and printed maps.

During the development of the prototype of the algorithm several issues occurred with the applied VGI dataset. In some cases the provided information proved to be unfeasible due to contributor collection errors or the lack of information in the selected test area. Therefore it needs to be noted that the preferred type of information and its corresponding quality have to be tested for each individual case where OSM data will be utilized (cf. Mondzeh & Sester, 2011; Mooney et al., 2012). However, the proposed algorithm and its generated network for pedestrians and disabled people provide room for new research projects based on the current findings, such as the combination with OSM 3D city models (Goetz, 2012b) or indoor (Goetz, 2012a), blind (Kammoun et al. 2010) and tactile (Pielot & Boll, 2010) routing applications.

Furthermore, several improvements to the algorithm are feasible. During the generation of the sidewalk network it could be useful to consider building information, which is also available in the OSM project database, to position the sidewalks correctly between the road and a row of houses, similar to the work introduced by Ballester et al. (2011). Some required tags, such as the incline of a road, are currently not widely mapped by the volunteers of the OSM project. In this particular case, the combination of the 2D way geometry from OSM together with a Digital Elevation Model (DEM) could result in a strong improvement (cf. Beale et al. 2006).

Lastly, combining the suggested generated network with the original OSM data topology would allow the development of a multi modal routing graph that implements sidewalk and public transportation network information, e.g. to plan a route for wheelchair users. Also, barriers such as street lamps or road signs in the middle of a sidewalk should be taken into account during the creation of the new sidewalk network.

References


Neis Pascal graduated with a Master of Science (MSc) degree from the University of Applied Sciences Mainz, Germany in 2008. He is currently working toward a Ph.D. degree at the Chair of GIScience at the University of Heidelberg. His current research interests include the analysis of user-generated geodata.

Zielstra Dennis is a PhD candidate with GIS concentration at the University of Florida, USA. His doctoral research focuses on Volunteered Geographic Information and other Web 2.0 content, with particular interest in pedestrian travel behavior.