Collecting a Ground Truth Dataset for OpenStreetMap

Carsten Keßler
Institue for Geoinformatics, University of Münster, Germany
carsten.kessler@uni-muenster.de

1 Introduction

The quality of OpenStreetMap (OSM) and volunteered geographic information (VGI) in general has already been discussed extensively in the literature. Researchers have looked at this issue from different angles such as credibility [2], trust [1], provenance [12, 9], precision [4], and communities [5]. Comparative studies often use commercial datasets or datasets from national mapping agencies for reference [4, 3, 6]. However, in order to fully evaluate how well OpenStreetMap reflects the streets, buildings, and different kind of amenities out there, such reference datasets are not sufficient, as none of them has a scope as broad as OSM.

In a recent study, we therefore decided to collect a ground truth dataset by hand [7]. In this case, the goal was to evaluate whether it is possible to assess feature quality based on provenance information; however, any other kind of study on the quality of OSM data (and VGI in general) is facing the same problem: How to obtain reliable reference data that have the same thematic and spatial scope as the VGI dataset under consideration? In the remainder of this abstract, I will therefore discuss different options to collect such a ground truth dataset.

2 Using Existing Data Sources

Most previous comparative studies on VGI data quality use commercial or administrative reference data because of their documented quality requirements, update cycles, and well-defined thematic scope. Unfortunately, this inevitably leads to a quality measurement by the standards of the reference dataset. While this may be acceptable for studies focusing on precision, any analysis that looks deeper into thematic aspects of OpenStreetMap will quickly reach the limits of the reference data due to the broad scope of OSM.

This situation might change with the growing number of social location-based services such as Foursquare, Facebook, or Yelp.[1] These services collect place descriptions from users "checking in" to share their current location with their social network [11]. The tags and descriptions provided by users of those services have reached a significant level of detail and are therefore a valuable source to complement administrative and commercial datasets.

Leveraging this potential, however, is still hampered by a number issues. As all of the above-mentioned services and their competitors are in a commercial market, the data they collect is their main asset. Therefore, even though APIs exist to access the data, this access is very limited, and copying larger parts of the data is generally against the respective terms of service. Such services could hence be seen as part of a distributed infrastructure for information about places [8], which is limited in that it can only be employed for feature-specific look-ups and smaller studies. Implementing such an approach, however, is also challenging in terms of place reconciliation, i.e., matching the correct places from different services to each other.

[1] See http://foursquare.com, https://www.facebook.com/about/location, and http://yelp.com.

3 Starting from Scratch

As mentioned above, we have collected our own ground truth dataset for a recent study [7]. While our dataset consisted of only 74 features that were pre-selected in OSM based on certain characteristics, the conducted field survey was still a considerable effort, especially because it was completed by a single person. Considering the effort spent, the product consists of a small and highly local selection of features. Scaling this approach to the point where a comparative study can actually produce significant results without a local focus clearly calls for a community-driven effort.

Crowdsourcing the problem in the spirit of VGI and OSM is an obvious solution. The question, however, is how to motivate people who have already contributed to the OSM dataset itself, as collecting the same data again, which clearly seems like a double effort. Gaining new contributors – which could eventually be converted to VGI – most likely requires approaches that go beyond the mere data collection and "seeing your feature on the map" aspect. One potential way to accomplish this goal is through gamification, which has already been shown in different studies to work well for the collection of VGI [13, 10]. The potential of such an approach has not least been demonstrated by the success of Foursquare, that has turned the collection and description of POIs into a game based on check-in counts and "mayorships".

Leveraging this principle would then require a separate interface (both in the sense of user interface and API) that allows researchers to create tasks within the game. Once completed by the players of the game, these tasks would allow for an ad-hoc quality comparison of a feature in a VGI dataset, and its counterpart in the ground truth reference dataset.

The two approaches for collecting a ground truth dataset introduced in this abstract are still in its early stages. They may be accomplished by or combined with additional methods of data

collection and integration. Moreover, the development of initial prototypes, along with usability tests and systematic evaluation, is required.

References

[1] Mohamed Bishr and Krzysztof Janowicz. Can we Trust Information? – The Case of Volunteered Geographic Information. In *Towards Digital Earth Search Discover and Share Geospatial Data Workshop at Future Internet Symposium, volume*, volume 640, 2010.

[2] Andrew J Flanagin and Miriam J Metzger. The credibility of volunteered geographic information. *GeoJournal*, 72(3-4):137–148, 2008.

[3] Jean-François Girres and Guillaume Touya. Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS*, 14(4):435–459, 2010.

[4] Mordechai Haklay. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and planning. B, Planning & design*, 37(4):682, 2010.

[5] Mordechai Haklay, Sofia Basiouka, Vyron Antoniou, and Aamer Ather. How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus' Law to Volunteered Geographic Information. *Cartographic Journal, The*, 47(4):315–322, 2010.

[6] Marco Helbich, Christoph Amelunxen, and Pascal Neis. Comparative Spatial Analysis of Positional Accuracy of OpenStreetMap and Proprietary Geodata. In *Int. GI_Forum 2012. Salzburg. Austria*, 2012.

[7] Carsten Keßler and René Theodore Anton de Groot. Trust as a Proxy Measure for the Quality of Volunteered Geographic Information in the Case of OpenStreetMap. In *Proceedings of AGILE 2013*.

[8] Carsten Keßler, Krzysztof Janowicz, and Mohamed Bishr. An Agenda for the Next Generation Gazetteer: Geographic Information Contribution and Retrieval. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, GIS '09, pages 91–100, New York, NY, USA, 2009. ACM.

[9] Carsten Keßler, Johannes Trame, and Tomi Kauppinen. Tracking editing processes in volunteered geographic information: The case of OpenStreetMap. In *Proceedings of Workshop on Identifying Objects, Processes and Events in Spatio-Temporally Distributed Data (IOPE)*, 2011.

[10] Sebastian Matyas, Peter Kiefer, and Christoph Schlieder. Changing the rules: Acquiring Quality Assured Geospatial Data With Location-based Games. In Ruck Thawonmas Akinori Nakamura Mitsuyuki Inaba, Koichi Hosoi and Masayuki Uemura Grey, editors, *Proceedings of the 4th annual Asian GAME-ON Conference on Simulation and AI in Computer Games (GAME-ON ASIA'2012)*, pages 44–50. EUROSIS, 2012.

[11] Grant McKenzie and Martin Raubal. Ground-truthing spatial activities through online social networking data. In *Extended Abstracts of the Seventh International Conference on Geographic Information Science (GIScience 2012), Columbus, OH*, 2012.

[12] Peter Mooney and Padraig Corcoran. Characteristics of Heavily Edited Objects in OpenStreetMap. *Future Internet*, 4(1):285–305, 2012.

[13] Carman Neustaedter, Anthony Tang, and Tejinder K. Judge. Creating scalable location-based games: lessons from Geocaching. *Personal Ubiquitous Comput.*, 17(2):335–349, 2013.