

Location Based Social Networks – Definition, Current State of the Art and Research Agenda

Oliver Roick^{*}

The growing pervasiveness of smartphones allowing users to permanently locate themselves and publish their position to the web has fostered the augmentation of social networks such as Facebook and Twitter with locational components or the creation of new social networks explicitly around geographic information, such as Foursquare. These services run on top of large data bases of geographic information that is primarily created and maintained through voluntary effort by the users of the services. Therefore, these data sets can be seen as special types *User Generated Content* or *Volunteered Geographic Information (VGI)* (GOODCHILD 2007).

VGI has become a serious alternative data source to geographic data provided by governmental and commercial institutions. Several studies of *VGI* – especially on the case of OpenStreetMap – have shown that the data quality in terms of completeness and positional accuracy is comparable to the quality of governmental and commercial data, yet still depending on the area being investigated (NEIS ET AL. 2012; HAKLAY 2010; GIRRES & TOUYA 2010). Further work has demonstrated the applicability of OpenStreetMap as foundation for several application scenarios, such as Routing (NEIS & ZIPF 2008) disaster relief (NEIS ET AL. 2010) and the creation 3D city models (OVER ET AL. 2010; GOETZ & ZIPF 2012). Only recently, several major businesses – including Apple and Foursquare – switched from Google Maps to OpenStreetMap as their primary data source for base maps and therefore further proved the maturity of *VGI*.

^{*}Heidelberg University, Chair of GIScience, Berliner Straße 48, 69120 Heidelberg, E-Mail: roick@uni-heidelberg.de

Based on these studies as well as the growing popularity of location based services and social networks we argue that user generated data from social networks are a valuable data source for different application scenarios. Social networks provide up-to-date information on highly dynamic data such as *Points of Interest* with businesses opening, closing and moving on a daily basis.

In order to fully exploit and employ data from social networks in-depth research is required on different aspects. First, there is a pressing need to investigate what data is available in social networks and how it can be accessed and retrieved. Second, data quality – primarily in terms of data set completeness as well as positional and attribute accuracy – and its evolution over time have to be understood. Methods already applied in researching data quality from crowdsourced geographic data (NEIS ET AL. 2012; HAKLAY 2010; GIRRES & TOUYA 2010) have to be evaluated for suitability and improved if necessary.

However, in order to allow for a full exploitation of geographic information from social networks data fusion is of utmost importance. Social networks are focused thematically or on certain user activities, so their data sets are complementary. Therefore, data fusion is the first step in the processing chain when mining information from the data sets of social networks. Data fusion, in general, is a two-fold process: First, features that represent the same real world object are matched. Second, the matched features are merged. Here their geometries and attribute fields and values have to be integrated, depending on certain criteria (COBB ET AL. 1998).

Several methods for matching geographical data have been presented before, which employ geometric proximity or semantic similarity measures (RUIZ ET AL. 2011). However, especially in urban areas a high density of geometries can be expected. Further, user generated data cannot be expected to have complete attribute sets and the information is mostly provided in an unstructured way through user generated *tags*. Hence, in order to fuse geographic data from social networks matching methods are required that employ geometric proximity and semantic distances in a combined manner. In terms of data fusion methods have to be improved, which allow for merging geometries and attributes by weighting the influence of certain features depending on their data quality.

The related methodologies, current state of art and future tasks will be discussed within the workshop and first results of our work will be presented.

References

COBB, M.A.; CHUNG, M.J.; FOLEY, III, H.; PETRY, F.E.; SHAW, K.B. & MILLER, H.V. (1998): A Rule-based Approach for the Conflation of Attributed Vector

- Data. *Geoinformatica*, 2(1): 7–35.
- GIRRES, J.F. & TOUYA, G. (2010): Quality Assessment of the French OpenStreetMap Dataset. *Transactions in GIS*, 14(4): 435–459.
- GOETZ, M. & ZIPF, A. (2012): OpenStreetMap in 3D – Detailed Insights on the Current Situation in Germany. In: *The 15th AGILE International Conference on Geographic Information Science*. Avignon, France.
- GOODCHILD, M. (2007): Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4): 211–221.
- HAKLAY, M. (2010): How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design*, 37: 682–703.
- NEIS, P.; SINGLER, P. & ZIPF, A. (2010): Collaborative mapping and Emergency Routing for Disaster Logistics - Case studies from the Haiti earthquake and the UN portal for Afrika. In: A. CAR; G. GRIESEBNER & J. STROBL (Hrsg.), *Geospatial Crossroads GI Forum '10. Proceedings of the Geoinformatics Forum Salzburg*.
- NEIS, P.; ZIELSTRA, D. & ZIPF, A. (2012): The Street Network Evolution of Crowdsourced Maps: OpenStreetMap in Germany 2007–2011. *Future Internet*, 4(1): 1–21. URL <http://www.mdpi.com/1999-5903/4/1/1/>.
- NEIS, P. & ZIPF, A. (2008): OpenRouteService.org is three times Open: Combining OpenSource, OpenLS and OpenStreetMaps. In: *GIS Research UK (GISRUK 08)*. Manchester, England.
- OVER, M.; SCHILLING, A.; NEUBAUER, S. & ZIPF, A. (2010): Generating web-based 3D City Models from OpenStreetMap: The current situation in Germany. *CEUS*, 34(6): 496–507.
- RUIZ, J.J.; ARIZA, F.J.; UREÑA, M.A. & BLÁZQUEZ, E.B. (2011): Digital map conflation: a review of the process and a proposal for classification. *Int. Journal of Geographical Information Science*, 25(9): 1439–1466.