

Analysis of OSM data through OSM-Notes user posting

Toshikazu Seto^{1,*}, Hiroshi Kanasugi¹ and Yuichiro Nishimura²

¹ Center for Spatial Information Science, The University of Tokyo, Tokyo, Japan;
tosseto@csis.u-tokyo.ac.jp, yok@csis.u-tokyo.ac.jp

² Division of Humanities and Social Sciences, Nara Women's University, Nara, Japan;
nissy_yu@cc.nara-wu.ac.jp

* Author to whom correspondence should be addressed.

This abstract was accepted to the Academic Track of the State of the Map 2019 Conference in Heidelberg after peer-review.

Data from Volunteered Geographic Information (VGI) and public surveying are essentially different; while a great quality and quantity of data is available for VGI, discussions about data quality are scarce [1]. As the use of VGI in business expands, attempts have been made to develop quality verification tools (e.g. KeepRight and Validation OSM) with the understanding that, in addition to position accuracy, data quality equally depends upon the diversity and interaction of the number of users involved in data generation [2].

OSM-Notes is a capability for describing errors and discussing OSM data and fixme, although there are few research cases [3]. This feature is new and was only added to the [openstreetmap.org](https://www.openstreetmap.org) website in April 2013. It allows users to specify and comment on any point on the map, and the history of comments can be accumulated and closed when the problem is solved. Unlike the `fixme` tag (allowing contributors to mark objects and places that need further attention in the form of a "note to self" or request for additional mapping resources), this function does not directly associate with OSM data, and there is no need to have an OSM account to create one. In this research, the OSM-Notes feature is mainly viewed as data that can be examined speedily from OSM data globally in terms of the content of the notes posted and the location of users. The purpose of the research is to analyze the context of OSM-Notes for quantitative and qualitative GIS approach [4].

This analysis was conducted on data dumped from Planet OSM as April 2019. Since `planet-notes.osn` (about 782 MB) is a special binary format, we used an enhanced parser able to separately output open (unresolved discussions) and closed (resolved discussions) based on "osn2osm". After converting to `.osm` format data, the set was combined with Natural Earth's border data and population data and treated as spatial data. As a result, there were OSM-Notes postings in 237 countries: 415,433 open and 129,887 closed records.

By counting the number of OSM-Notes postings by country based on whether they are open or closed, we determined that the majority were in the United States, Germany, and Russia. Moreover, it became clear that Japan, Canada, Korea, and Taiwan are the regions where there are a large number of submissions in urban areas (based on Natural Earth's

Seto et al. (2019). Analysis of OSM data through OSM-Notes user posting

In: Minghini, M., Grinberger, A.Y., Juhász, L., Yeboah, G., Mooney, P. (Eds.). Proceedings of the Academic Track at the State of the Map 2019, 15-16. Heidelberg, Germany, September 21-23, 2019. Available at <https://zenodo.org/communities/sotm-2019>
DOI: [10.5281/zenodo.3387685](https://doi.org/10.5281/zenodo.3387685)



definition). In addition to these, many OSM users have posted from many countries, including Iraq, Ukraine, and Ecuador, and it is clear that active discussions are being held by many contributors. In these countries, mapping for humanitarian assistance is commonplace, as other reference resources are unavailable, so it is necessary to improve data quality through the use of OSM-Notes. OSM-Notes even has a function that allows non-OSM users to post, as is the case with 50% or more of the posts in Spain, Korea, etc.

This data can be used to analyze urban trends and spatial features within a single country. For example, according to the analysis for Japan, OSM-Notes has many posts about the location of shops and POIs (Points of Interest), and suggestions based on Maps.ME, Facebook, and Pokémon Go. This is considered to be the main reason that anonymous posting is permitted. It is also worth noting that very few users post only a single note. Thus, by analyzing OSM-Notes, it is possible to grasp hot issues between users of OSM data.

Overall, we found that postings to OSM-Notes are more frequent in cities where OSM mapping is active. On the other hand, how to accept bug reports from applications other than OSM (and not users' discussions) which edit areas where OSM data is limited remains an open problem. We need to consider the OSM data update process for multiple applications. Moreover, how to accept it is an important issue in considering the process of constructing OSM data. However, because the functional relationship between OSM-Notes and the data on OSM cannot be specified directly, the nature of the feature is also difficult to grasp directly compared to the `fixme` tag.

References

- [1] Senaratne, H., Mobasheri, A., Ali, A. L., Capineri, C., & Haklay, M. (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1), 139-167.
- [2] Haklay, M. (2010). How good is volunteered geographical information? A comparative analysis of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design*, 37(4), 682-703.
- [3] Seto, T., Iwasaki, N., & Nishimura, Y. (2017). Evaluation of the data update frequency and user interaction of OSM in Japan: case study of OSM-Note. In *State of the Map 2017*.
- [4] Cope, M., & Elwood, S. (2009). *Qualitative GIS: A Mixed Methods Approach*. SAGE publications, London, UK.