Improving collective intelligence and exploration in a VGI like context through communication of uncertainty information

Jochen Schiewe

HafenCity University Hamburg, Lab for Geoinformatics and Geovisualization (g2lab)
E-Mail: jochen.schiewe@hcu-hamburg.de

Motivation

In times of Volunteered Geographic Information (VGI) data availability is not the bottleneck for geographic applications anymore. Instead, aspects of data quality (in the following, the more general term "uncertainty" will be used) become more and more important. This is mainly due to the fact that data are now coming from a huge number of different and unknown sources, leading to a heterogeneous distribution of uncertainty.

On one hand, uncertainty modeling and visualization for expert users is already an established topic in research and development; in particular, if exploration of data is the goal of usage. On the other hand, in a VGI like context producers and users (in the following summarized as "prosumers") share hardly any uncertainty information. It is well accepted that the number of people that actively work with data at certain place improves data quality (also known as "collective intelligence" or "Linus' Law"); however, even this number is not communicated in known map environments. An attempt in this direction at a very low level is implemented in *OpenStreetMap* with the option for users to press a "like" or "dislike" button for the current region of interest.

Consequently, within our work we try to go one step further: As many VGI is also used in (explorative) applications where uncertainty matters, we see the *need for describing, sharing and using uncertainty information for prosumers in an usable manner*.

Concept

In order to motivate and ease the **generation** of uncertainty information by both, volunteer producers and users, we propose a strongly reduced and simplified set of parameters. Comber et al. (2007) or Boin & Hunter (2007) argue that the application of typical metadata catalogues is not an operational method. Alternatively, we propose a generally applicable and standardized interface that allows the (not in all cases mandatory) input of the following information:

- Differentiation between producer and consumer (user);

- evaluation of geometric uncertainty on an ordinal ("five stars") scale which is well-known from other evaluation scenarios like *Amazon*;

- evaluation of attribute uncertainty on an ordinal ("five stars") scale;

- free comments.

In order to allow for an input of producers and users, this set of uncertainty information will not be attached to single objects, but linked to the respective position using a marker symbol. Besides this active input done by prosumers, also the date and a counter will be automatically recorded in order to produce information related to up-to-dateness and amount of usage.

A further advantage of keeping the description of uncertainty information in such a rather easy and standardized model is the possibility to generate a follow-up **visualization** in a fast and automatic manner. According to the results of MacEachren et al. (2012), who investigated the intuitiveness of visual variables for representing uncertainty information, we prefer iconic representations, i.e. bulls-eye location for spatial uncertainty and smileys for attribute uncertainty (figure 1).
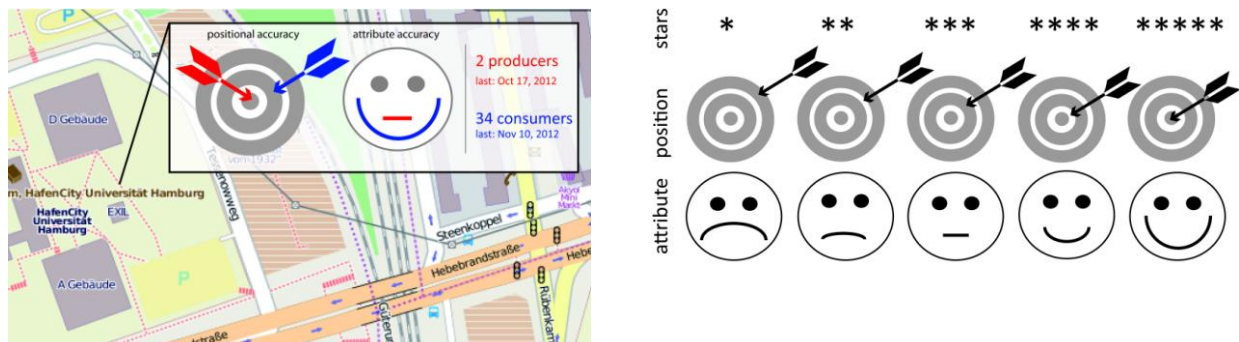

Figure 1: Visualization of uncertainty information using an ordinal scale

Evaluation

So far, no reliable empirical evidence exists about the actual motivation of prosumers to generate, share and use uncertainty information. To overcome this gap, we are currently conducting a qualitative study in form of a formalized questionnaire which is based on the above outlined concept of uncertainty communication. This survey aims for answering the following key questions:

- What are typical applications that demand for uncertainty information?
- Is the proposed depth of uncertainty information appropriate?
- Is the proposed visualization format (i.e., iconic vs. simple five stars display) clear enough?

Results of this study that firstly addresses producers (e.g., coming from the *OpenStreetMap* community) will be presented. Based on this, derived enhancements of the outlined concept will also be addressed.

References

Boin, A.T. & Hunter, G.J. (2007): What communicates quality to the spatial data consumer. 5th International Symposium in Spatial Data Quality, Enschede (The Netherlands), 8 p.

Comber, A.J., Fisher, P.F. & Wadsworth, R.A. (2007): User-focused metadata for spatial data, geographical information and data quality assessments. 10th AGILE International Conference on Geographic Information Science, Aalborg (Denmark), 13 p.

Grira, J., Bédard, Y. & Roche, S. (2009): Spatial data uncertainty in the VGI world: going from consumer to producer. Geomatica, 64(19): 61-71.

MacEachren, A.M., Roth, R.E., O'Brien, J., Li, B., Swingley, D. & Gahegan, M. (2012): Visual Semiotics & Uncertainty Visualization: An Empirical Study. IEEE Transactions on Visualization and Computer Graphics, 18(12): 2496-2505.