VGI Edit History Reveals
Data Trustworthiness and User Reputation

Fausto D'Antonio
University of L'Aquila
Via G. Gronchi 18, 67100
L'Aquila, Italy
fausto.dantonio@gmail.com

Paolo Fogliaroni
Vienna University of Technology
Gusshausstr. 27-29, 1040
Vienna, Austria
paolo@geoinfo.tuwien.ac.at

Tomi Kauppinen
Aalto University School of Science
Department of Media Technology
FI-00076 Aalto, Finland
tomi.kauppinen@aalto.fi

Abstract

Volunteered Geographic Information (VGI) is an approach to crowdsource information about geospatial features around us. People around the world are engaged with typing in their observations about the world (like locations of shops, cafeterias), or to semi-automatically gather them with mobile devices (like hiking paths or roads). In this process people might make mistakes, for instance assign misleading tags to features or provide over simplistic boundaries for features. In this paper we study what kinds of things might contribute to assess trustworthiness of data, and reputation of contributors for VGI. We present a model for analysing the different factors, and a method for automatically creating the trust and reputation scores.

1 Introduction

Volunteered Geographic Information (VGI) [5] is an approach to crowdsource information about geospatial features around us. Recently, VGI is gaining increasing attention and (web) services relying on it are becoming ubiquitous.

Thanks to the greater engagement of contributors, coverage and precision of VGI data is quickly approaching the level granted within professional Geographic Information Systems (GIS), as shown by several comparative studies with official national datasets [6, 14]. However, while in professional GIS data quality is granted by certified authorities, the assessment of VGI data quality remains an open challenge [11, 7, 13].

A basic method to assess the quality of a VGI dataset consists in comparing it against a professionally–generated ground–truth dataset. However, this approach suffers major drawbacks. First, it requires access to professional datasets that, in the best case, is expensive and, in the worst case, is not possible at all. Moreover, it does not provide a quality assessment procedure that is universally valid (e.g., think of cases where a ground–truth dataset is not available at all).

As suggested in [1], a different approach consists in assessing the quality of VGI data through a proxy measure: trustworthiness. Trustworthiness is defined [12] as a "bet about the future contingent action of others". In this sense, trustworthiness is strictly related to the concept of (others') reputation.

This paper presents ongoing work on an evaluation model of volunteers' reputation and data trustworthiness that derives the coveted information from VGI data, without requiring a comparison with external sources. We draw inspiration from the work in [7] and extend it by (i) relating data trustworthiness and user reputation and (ii) accounting for the relevance of data editing. (iii) Finally, our model accounts for atomic editing operations, rather than for composite editing patterns.

2 Related work

Quality assessment of VGI is still rather new research topic. As suggested by Flanagin and Metzger [3], there is a critical need for identifying methods and techniques to evaluate the VGI quality.

One rather standard approach is to compare VGI datasets to authoritative, ground-truth datasets, as done, for example, by Mooney, Corcoran and Winstanley [11] who analysed characteristics of polygons contributed by OpenStreetMap users. According to the results, volunteers seem to be able to more easily trace outlines of water features compared to forest features.

A different approach is undertaken by Bishr and Janowicz [1] that promote the notion of informational trust to be used as a proxy measure for quality. Their proposal was one of the first examples for using trustworthiness for quality assessment in VGI.

Keßler, Trame and Kauppinen apply the Bishr and Janowicz proposal to use trust as a proxy measure, in [8] they used trust and provenance for studying contribution patterns in the case of OSM. An extension of this work is [7] in which, Keßler and De Groot, provided a few indicators that influence trust and that were basically derived from data provenance.

The work presented in [7] has the purpose to build a model that depends mostly on provenance data; so that there is no need of a reference comparison dataset; trustworthiness is associated to each feature and represents the proxy value of data quality. In this work Keßler introduce the user reputation issue and leave it for future refinement.

Keßler used five parameters for trustworthiness evaluation. (1) Versions, they are an important source of provenance information. (2) Users, the higher is the number of users that works on a feature the higher is the trustworthiness value. (3)

Confirmations, all the revisions that were made in the neighbourhood of a feature are taken into account. (4) Tag corrections, a semantic change over a feature decreases the feature trustworthiness. (5) Rollbacks, restoring a feature's previous state also decreases the feature trustworthiness.

3 Model Overview

In this section we give an overview of the main constituents and mechanisms underlying our trustworthiness evaluation model. More details are given in next sections.

The model we introduce can work with any VGI system that provides the following, basic requirements:

- The system supports (directly or indirectly) feature versioning
- which is expressible as a sequence of basic editing operations of the type: creation, modification, and deletion.

3.1 Feature Versioning

Geographic features in a VGI system are subject to repeated changes over time. A change is operated by a contributor and brings a feature representation into a new state or version. History of a feature's changes is referred to as feature provenance [8] and feature versioning is a particular interpretation of it.

Similarly to the work presented in [7] we base trustworthiness evaluation on provenance information. This approach is a realization of the "many eyes principle" [6] that assumes that incorrect, wrong, or malicious information about a feature get corrected over versions contributed by different volunteers. The main underlying idea is that a high number of lay contributors reporting on the same feature and an iterative adjustment of the feature information provides a valid alternative to field expertise.

Our approach consists in assigning a trustworthiness value to each version of a feature. Thus, in order to be compliant with our model, a VGI system must provide provenance information directly in the form of feature versioning. Alternatively, versioning must be derivable from feature provenance.

3.2 Editing types

While the approach presented in [7] derives trustworthiness values from editing patterns[1], our approach relies upon atomic editing types:

Creation When a new feature is inserted into the dataset there are no previous versions that the feature can be compared with. Thus, it is assigned a trustworthiness value equal to its author reputation (cf. Section 5.1).

Modification The modification of an existing feature yields a new version whose trustworthiness depends on the author reputation and on the compatibility with previous versions. That is, if previous versions are similar to current version the latter is associated a high trustworthiness. Contrarily, dissimilarities are rated with low scores. Note that a modification also affects trustworthiness of previous versions.

[1] An editing pattern is a sequence of atomic editings that can be interpreted as a unique high–level change. For example, it has been shown [8] that emerging editing patterns in OpenStreetMap (http://www.openstreetmap.org/) are confirmations, corrections, and rollbacks.

Deletion A deletion occurs when a real–world feature somehow disappears (e.g., a building is demolished). Notably, this editing type yields a new "void" version. Trustworthiness of this version is set equal to its author reputation and previous versions are not affected.

Studying how our approach behaves with respect to editing patterns is left for future work. However we would like to anticipate one notable situation. A deletion and a creation happening consecutively must be deemed a modification: this is the case that a feature is deleted and recreated rather than being modified. This also covers the case of non–genuine deletions.

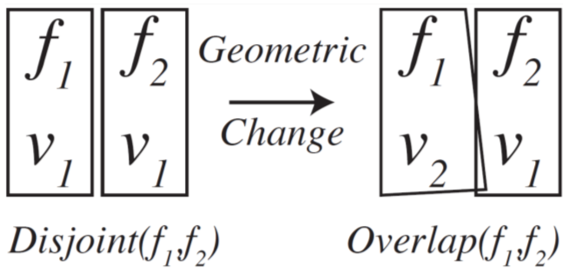
4 Measuring Editing Relevance

We argue that diverse edits must be weighted differently. More specifically, the more closely the version resulting from a change fits the feature in the real world, the higher the trustworthiness of this version.

Since a direct comparison with real world cannot be performed, we suggest evaluating the level of fitness by clustering the versions of a feature according to three main characteristics and comparing versions assigned to the same cluster.

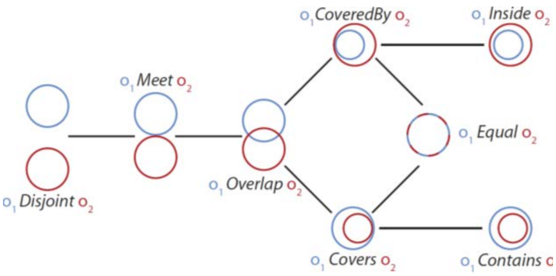
A spatial feature consists of two components: spatial and semantic. Moreover, the spatial characteristic can be further refined into qualitative and geometrical. The reason for such a finer distinction relies on the fact that a small change in the geometry of a feature may correspond to a notable change in the qualitative spatial relations holding among this feature and its neighbours.

Figure 1: Geometric and qualitative spatial change.



For example, let us consider the scenario depicted in Figure 1 where two features f_1 and f_2 at version v_1 are *disjoint* but very near to each other. A geometric change occurs that slightly modifies the geometry of f_1 as depicted on the right side of the figure. This small geometric change modified the topological relation holding among the two feature into overlap. Qualitatively, this is a notable change since according to the theory of *conceptual neighbourhoods* [4] the two relations are not close. Indeed, as shown in Figure 2, one cannot switch from the *disjoint* relation to the *overlap* relation without going through the relation *meet*. Conversely, big geometric changes may not alter the qualitative arrangement of features. Thus, the consideration of both the geometric and qualitative aspect allows for mitigating the evaluation of spatial changes.

Figure 2: Conceptual neighbour graph of 9-Intersection model [2].



Accordingly, we distinguish the following characteristics:

Semantic The semantic or thematic aspect describes, by means of textual tags, the function of a feature in the world. When a semantic change occurs we evaluate its relevance by considering the semantic distance between the tag associated to the new version and the tags associated to versions in the same cluster. This can be done, for example, by considering the shortest path on a wordnet [10] graph[2]. Alternatively, the shorter conceptual distance in an ontology including both previous and altered tag can be used.

Geometric The geometric aspect describes the shape and position of the feature in the world. The relevance of geometric changes is evaluated with respect to a series of quantitative variances like, area, perimeter, and vertices number and position.

Qualitative (Spatial) This aspect addresses qualitative spatial relations of different types (e.g., topological, directional, distance) holding among a feature and its neighbours. The relevance of the change depends on the distance on the conceptual neighbourhood graph between the relations holding for the new version of the feature with respect to those occurring on previous versions.

5 Reputation and Trustworthiness

We denote trustworthiness and reputation by T and R, respectively. As done in [9], we bound the values of the two parameters between 0 and 1: $0 \leq T; R \leq 1$. We associate trustworthiness to each version v of a feature f: by $T(f_v)$, we denote the trustworthiness value of version v of feature f.
The reputation of a user u changes over time t: by $R(u,t)$, we denote the reputation of user u at time t.

[2] http://graphwords.com/

5.1 Reputation

User reputation depends on the trustworthiness of all the feature version his editing produced and is defined as the average of such values:

$$R(u,t) = \frac{\sum_{f_i \in F(u,t)} T(f_i)}{|F(u,t)|} \tag{1}$$

where $F(u,t)$ is the set of all the feature versions edited by user u until time t.

5.2 Trustworthiness

The overall trustworthiness value T of a feature version f_v accounts for three different effects: direct T_{dir}, indirect T_{ind}, and temporal T_{time}.

Direct Effect The parameter T_{dir} is the expression of the level of similarity with respect to the characteristics discussed in Section 4 between the current feature version f_v and previous ones. Accordingly, its value depends on three factors: semantic $T_{dir,s}$, geometric $T_{dir,g}$, and qualitative $T_{dir,q}$.
Direct effect is modelled as:

$$T_{dir} = w_s \cdot T_{dir,s} + w_g \cdot T_{dir,g} + w_q \cdot T_{dir,q} \tag{2}$$

where w_s, w_g, and w_q are weights used to balance the influence of the three characteristics.
To assure $0 \leq Tdir \leq 1$ we enforce:

$$w_s + w_g + w_q = 1 \tag{3}$$

and

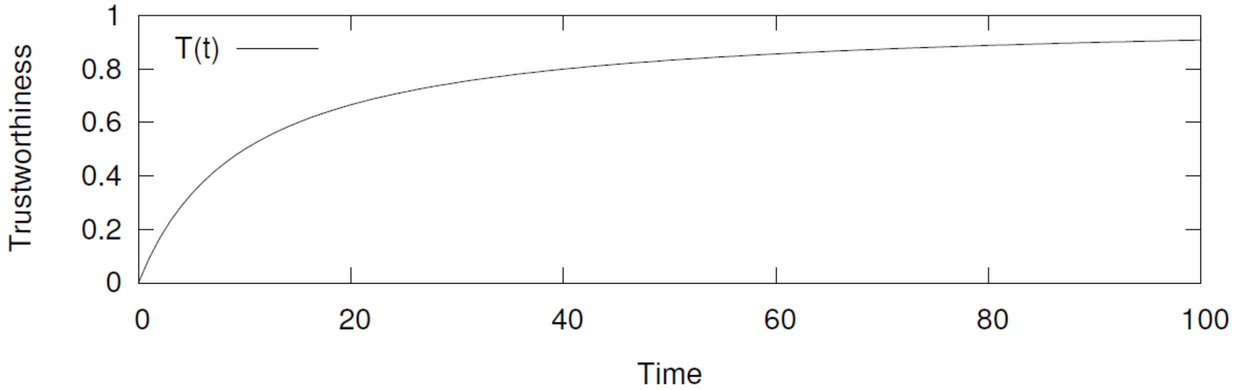
$$0 \leq T_{dir,s}, T_{dir,g}, T_{dir,q} \leq 1 \tag{4}$$

Indirect Effect The parameter T_{ind} models contributions on the overall trustworthiness value T that do not directly depend on the feature version f_v itself. For example, this can be used to account for confirmations [8]: the fact that a user contributes information about f_v's neighbours can be interpreted as a confirmation that f_v has a high fitness level with respect to the feature in the real world. Hence, f_v's trustworthiness must be increased.
Also in this case we account for three different factors, expression of the characteristics defined in Section 4:

$$T_{ind} = w_s \cdot T_{ind,s} + w_g \cdot T_{ind,g} + w_q \cdot T_{ind,q} \tag{5}$$

Similarly to direct effect, we assure T_{ind} falls in the interval [0; 1] by enforcing conditions similar to those reported in Equations 3 and 4.

Figure 3: Time effect on trustworthiness.



Temporal Effect The parameter T_{time} accounts for the effect of time on features trustworthiness. Namely, the longer a feature version f_v persists over time, the higher the probability that f_v has a high fitness level with respect to the real feature. Thus, if a feature version remains unaltered over time, its trustworthiness must be increased. We allow this by modelling time effect as:

$$T_{time} = \frac{t_v}{t_f + c} \quad (6)$$

where t_f is the life time of feature f (all versions), t_v is the life time of version v of feature f, and c is a parameter taking positive values that can be used to adjust the slope of the resulting curve (cf. Figure 3). So, when t_v approaches infinity also t_f does, c becomes negligible, and T_{time} approaches 1.

Accordingly, the overall trustworthiness T is defined as:

$$T = w_{dir} \cdot T_{dir} + w_{ind} \cdot T_{ind} + w_{time} \cdot T_{time} \quad (7)$$

where w_{dir} and w_{ind} are weights used to balance the importance of the direct and indirect effect, respectively, and such that:

$$\mathrm{w}_{dir} + \mathrm{w}_{ind} = 1 \quad (8)$$

Temporal effect is weighted by

$$w_{time} = 1 - (w_{dir} \cdot T_{dir} + w_{ind} \cdot T_{ind}) \quad (9)$$

in order to assure that the overall trustworthiness value T increases as time passes with a pace following the curve depicted in Figure 3.

6 Conclusions and Outlook

Inspired by the work in [7], we introduced a model to evaluate VGI user reputation and VGI feature trustworthiness as a proxy measure for data quality.

We anchored our model to basic editing types and discussed how changes among feature versions can be evaluated grounding upon three characteristics: semantic, geometric, and qualitative.

We provided a high–level formulation for reputation and trustworthiness and discussed how the latter is a product of direct and indirect effects as well as a function of time.

The model is still under development, yet finer detailed than what was possible to discuss in this short paper. In the next phases we plan to implement the model and to study its behaviour using OpenStreetMap historical data. Also, a comparison with other trustworthiness evaluation models will be carried out.

References

[1] M. Bishr and K. 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 640*, 2010.

[2] M. J. Egenhofer, A formal definition of binary topological relationships, in *Foundations of data organization and algorithms*, Springer, 1989, pp. 457-472.

[3] A. J. Flanagin and M. J. Metzger, The credibility of volunteered geographic information. *GeoJournal*, 72.3-4, pp.137-148, 2008.

[4] C. Freksa, Conceptual neighborhood and its role in temporal and spatial reasoning, in M. Singh e L. Travé-Massuyès, editors, *Decision Support Systems and Qualitative Reasoning*, pp. 181-187. North Holland, Amsterdam, 1991.

[5] M. F. Goodchild, Citizens as sensors: the world of volunteered geography, *GeoJournal*, vol. 69, n. 4, pp. 211-221, 2007.

[6] M. Haklay, How good is volunteered geographical information? A comparative study of OpenStreetMap

and Ordnance Survey datasets, *Environment and planning. B, Planning & design*, vol. 37, n. 4, p. 682, 2010.

[7] C. Keßler, J. Trame, and T. Kauppinen, Tracking editing processes in volunteered geographic information: The case of OpenStreetMap, in *Identifying objects, processes and events in spatio-temporally distributed data (IOPE), workshop at conference on spatial information theory*, 2011.

[8] C. Keßler and R. T. A. de Groot, Trust as a proxy measure for the quality of Volunteered Geographic Information in the case of OpenStreetMap, in *Geographic Information Science at the Heart of Europe*, Springer, 2013, pp. 21-37.

[9] N. Mezzetti, A socially inspired reputation model, in *Public Key Infrastructure*, Springer, 2004, pp. 191-204.

[10] G. A. Miller, WordNet: a lexical database for English, *Communications of the ACM*, vol. 38, n. 11, pp. 39-41, 1995.

[11] P. Mooney, P. Corcoran, and A. C. Winstanley, Towards quality metrics for openstreetmap, in *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2010.

[12] P. Sztompka, Trust: A sociological theory, Cambridge University Press, 1999.

[13] D. Zielstra, H. H. Hochmair, and P. Neis, Assessing the Effect of Data Imports on the Completeness of OpenStreetMap - A United States Case Study, *Transactions in GIS*, vol. 17, n. 3, pp. 315-334, 2013.

[14] D. Zielstra and A. Zipf, A comparative study of proprietary geodata and volunteered geographic information for Germany, in *13th AGILE international conference on geographic information science*, Leuven, 2010.