Can we Trust Information? - The Case of Volunteered Geographic Information

Mohamed Bishr[1] and Krzysztof Janowicz[2]

[1] m.bishr@uni-muenster.de
Institute for Geoinformatics, University of Muenster , Germany
[2] jano@psu.edu
Department of Geography, The Pennsylvania State University, USA

Abstract. In this paper we take a fresh look at the problem of information quality for user contributed content. We assert that the traditional quality criteria for assessing the quality of geographic information are difficult to apply to Volunteered Geographic Information. The notion of *informational trust* is introduced and linked to the established notion of interpersonal trust. We then propose to use informational trust and reputation as proxy measures for information quality and outline the spatial and temporal dimensions of trust that have to be considered.

1 Introduction

For the vision of a Digital Earth, user generated content is an essential repository of up-to-date information ranging from maps and pictures, to local observations. Such user generated geospatial information is collectively refered to as VGI [1]. However, its integration into Spatial Data Infrastructures, the Social Web, or the Digital Earth comes at a price. As discussed in previous work [2], the voluntary contributors of such applications are not necessarily experts, have different backgrounds, and varying motivations to contribute data. Therefore, not all provided information is valuable or may even be misleading. Given the tremendous flow of VGI and the lack of knowledge about the traditional GI quality criteria (completeness, lineage, accuracy, and consistency), the problem of information quality is of increasing importance. To address these challenges, we proposed trust as a potential proxy measure for the quality of geospatial information[2]. So far, most existing work on the notion of trust in social networks has neglected the spatial and temporal dimension of trust and their implications on user generated content. This paper outlines new insights on this proposal.

2 Trust and Reputation

Trust is a widely studied phenomenon, it has been a core resaerch topic in sociology, political science, economics, philosophy, as well as computer science [2–7]. In our discussion we refer to trust as a social phenomenon such as discussed by Luhmann, Golembiewski, and others [8, 9]. In Sztompka [7], trust is defined

as *a bet about the future contingent actions of others*, which is also the definition we adopt here. This definition has two components, belief and commitment. The belief that a particular person will act in a favourable way and my commitment to a certain action based on that belief. In this work, we particularly refer to *interpersonal trust* [10] as a social tie between a trustor and a trustee.

Several researchers argue that trust holds only between people. Trusting a company like Lufthansa to take you to your destination is, in fact, trusting the people behind the company. Therefore, one could argue that trust in entities is based on trust in the persons responsible for these entities. Following this argumentation, we propose the notion of *people-object transitivity of trust* which differs from the trust transitivity commonly applied to Web-based social networks [6]. In our view, *interpersonal trust* implies the transition of trust from the trustee to information entities conveyed by the trustee. The trustor can then assert trust directly in the information conveyed by the trustee. We call this *informational-trust*; where a trusting tie between a trustor and an information entity such as VGI is mediated by interpersonal trust between the VGI originator and the VGI consumer.

We further propose to extend this notion by spatial and temporal characteristics of the trust phenomenon. In Buskens' work [3] distance is used as a proxy measure for social network density of buyers and suppliers' networks. Partners in proximity are always preferred partners [11], implying a higher degree of trust. Buskens also asserts that the cost to the buyer as a result of abusing trust by the trustee increases with distance. This means that the larger the distance the more difficult it is to resolve problems which also leads to lower trust. A similar conclusion is provided by Lyons [12] – distance has a positive effect on the probability of a subcontractor and a customer governing their relation by a formal contract, such that the probability increases with distance, implying less trust. A related explanation is provided by Lorenz [13], where the subjects indicated that personal contact is important for establishing trust, and as such geographic distance was necessary in easing this personal contact increasing trust with distance proximity. Also, it is intuitive to assert that trust decays and develops over time, or that trust develops slowly, but can easily be tarnished if abused, showing a temporal affinity to the concept of trust. Some researchers tried to formalize this temporal effects on trust; see for example [14]. VGI is spatial and temporal in nature - by extending the notion of trust with spatial and temporal dimensions we can develop trust models that can be used for filtering and triage of VGI and provide higher quality information to consumers.

Finally we briefly address the notion of reputation and relate it to trust. We view reputation as the perception of trustworthiness of a person by a community [15]. Reputation of a person does not belong to that person, but to the community. It depends on many factors including previous behavior, community perception of the person, the capacity of the community to sanction bad behavior, and propagation through word of mouth. Reputation can then be used to assert the degree of trust to be placed in a person. This is achieved because

reputation helps a trustor assess how likely is the trustee to honour trust. As such, reputation is an additional factor in determining the trust of VGI.

3 Informational trust as a proxy measure of information quality

In fact we have no objective measures of the accuracy of, say, google's search results and hardly any benchmarks to compare to, yet people continue to rely on google. Our argument here is that the quality of google's results is assessed based on whether or not it fits the purpose of the users at a given time. This is an intuitive assessment that most of us make in such a situation. Hence, information quality always depends on the current context, e.g., a problem that needs to be solved. If a larger group of people find some VGI fit for their purpose we can assert that this VGI is of higher quality compared to others (at least for the given purpose). The volunteers providing VGI can then act as markers for the quality of their contributions. They mediate between the VGI consumers and the consumed VGI so that if person A can identify the VGI contributor B as having a reputation of being trustworthy, then A can assert that the VGI by B is itself trustworthy. This does not contradict our theoretical understanding of trust given the earlier discussion in section 2 on our proposed notion of *informational trust*.

This process requires spatial and temporal aspects of trust and reputation models for VGI triage. In previous work [16, 17], we have discussed models for VGI quality assessment based on the spatial and temporal notion of trust and reputation. The spatial and temporal dimensions of informational trust can be used to make assertions such as that people in proximity of VGI observations provide more accurate information (e.g. the locals know better, the proximate spectator sees more, and so on) or that the trust in some VGI develop and decay over time. One can postulate someone contributing VGI about a significant change in her local urban area, and intuitively assume that her location (with respect to the changed entity) as well as the observation time affect the trust we might vest in her observation.

4 Example and Discussion

Our previous work [17] demonstrates the integration of space into the informational trust models. This model, however, does not include a temporal dimension. The use case for the model is that of a newly urbanized area in Muenster, Germany where we postulate that the urban modellers will depend on VGI to collect information about urban growth volunteered by the local residents. The local residents carry hand-held devices that can log their position while making observations about their urban environment. The system also stores social relations between the residents where they assign trust ratings to each others in a social network of contributors.

During reporting by volunteers the observed entities will naturally overlap. So assume that volunteers n_2 and n_6 in figure 1 are making two separate observations, e.g. n_2 observes that in this location a new building is being constructed, let us call this observation m_{1a} and n_6 later makes another observation about the same location with the same conclusion, let us call this observation m_{1b}. The system uses proximity of the location put by the volunteer on the map and then suggests to the volunteer n_6 that m_1a and his own observation m_{1b} might be identical. If the user confirms that both are identical, both m_{1a} and m_{1b} are amalgamated so that the observed event m_1 is recorded in the system and is counted as reported twice by volunteers n_2 and n_6. The resulting network model is a hybrid model of a bipartite graph between volunteers and their contributed observations about urban development and a social network among the volunteers, see figure 1.

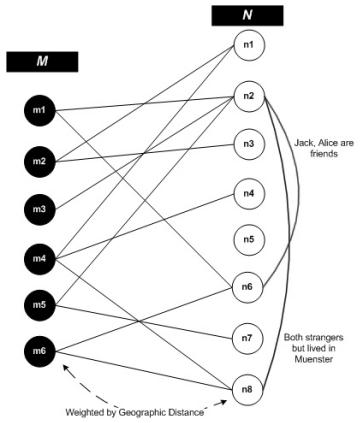

Fig. 1. *An affiliation network of volunteers N and the VGI they contribute M with links between a volunteer and VGI entities weighted by distance between the volunteers and the observed entity's actual location. Links connecting N white nodes are a separate agent-agent social network, making this model a hybrid affiliation-one mode social network.*

Equation 1 is one part of the trust model presented in previous work [17] and used here to demonstrate the basic assumptions underlying our model. The trust ratings of a member of M is denoted t_{m_h} and is computed by the given equation where t_{n_g} is the specific trust associated with a given volunteer. This t_{n_g} is computed through the model in earlier equations using trust ratings on the social network and acts as an indicator of user reputations. While $r_{(n,m)_i}$ is the ith rating from one n to another m. Then the equation is weighted by the distance c_i using a log function to smooth the sharp variations resulting from using raw distance. The resulting rating t_{m_h} is a rating of how much some amalgamated observation m is to be trusted and is then used to sift through

the VGI entities and curate them by which ones are more trustworthy than others. It is to be noted that trust ratings on a social network can be biased, so that friends would rate each others higher regardless of the actual quality of the contributions. This is a risk that needs to be mitigated in the system by opening ratings of people to all other people in the social network to create a reputation effect. Hence, the group dynamics resulting in overall user reputations would make the system more resilient to individual biases.

$$t_{m_h} = \sum_{i=1,g=1}^{k} \frac{t_{n_g} r_{(n,m)_i}}{\log c_i} \text{ where } c_i > 1 \tag{1}$$

We also postulate that the integration of the temporal dimension in this model will enable us to evaluate the VGI based on temporal currency of the information. For example, if two observations by two different volunteers about the same phenomenon contradict each others then two scenarios could be possible: first, the time difference between both observations is sufficiently large that the phenomenon could have potentially changed, or secondly the time difference is sufficiently small as to warrant doubt on the trustworthiness of one or both observations. The temporal dimension is added in analogy to the spatial distance presented in equation 1, and will require that we know how both the spatial and the temporal distance interact. This needs to be determined by human participants tests and simulations and is ongoing work.

We conclude by stating that the temporal and spatial dimensions then can lead to a multitude of innovative trust and reputation solutions for VGI. To elucidate, let us assume that in the presented examples some disaster or emergency arises in the study area. In this case through the integration of the spatial and temporal dimensions we can infer a new form of reputation which we call *event-reputation* of volunteers using distance and time. In this case distance and time become very important so that a volunteer close to an event (spatial dimension) at the time of the disaster's occurrence (temporal dimension) is given an *event-reputation* for a given period of time concerning reporting about this event. This event-reputation is not only ephemeral, but can also be extended further given crowed feedback on the quality of the reports from this particular volunteer through the course of the disaster to make it a permanent reputation for this volunteer about a certain local area.

5 Conclusions

In this paper we revisited our earlier proposal of using trust as a proxy measure of the quality of user contributed content. Beyond the earlier proposal this paper introduces the notion of informational trust as a derivative of interpersonal trust. This approach resolves the ontological conflicts about the nature of trust as a social tie and trust in information like VGI. We then extended this notion of trust by a spatial and temporal dimension. One implication of this is that the location of volunteers with respect to the observed phenomena in VGI impacts

the informational trust of their observations. The temporal dimension of informational trust plays a similar role where the time between different observations of volunteers impacts informational trust.

6 Acknowledgements

The authors of this paper would like to thank Werner Kuhn for coining the term proxy measure of quality as a way to describe the basic premise of this work. We also thank Simon Scheider for his insightful suggestion on the notion of *informational trust*. Some of the ideas in this paper were inspired through a brief email exchange with some members of SwiftRiver and Ushahidi project team.

References

1. Goodchild, M.: Citizens as sensors: the world of volunteered geography. GeoJournal **69**(4) (2007) 211–221
2. Bishr, M., Kuhn, W.: Geospatial information bottom-up:a matter of trust and semantics. In: AGILE, Aalborg, Denmark, Springer-Verlag (2007)
3. Buskens, V.W.: Social networks and trust. Springer (2002)
4. Fukuyama, F.: Trust: The social virtues and the creation of prosperity (1996)
5. Gambetta, D.: Can We Trust Trust? In: Trust: Making and breaking cooperative relations. Basil Blackwell New York (1990)
6. Golbeck, J.: Computing and Applying Trust in Web-based Social Networks. PhD thesis (2005)
7. Sztompka, P.: Trust: A Sociological Theory. Cambridge University Press (1999)
8. Luhmann, N.: Trust and Power. John Wiley and Sons (1979)
9. Golembiewski, R., McConkie, M.: The centrality of interpersonal trust in group processes. Theories of group processes (1975) 131–185
10. Mayer, R., Davis, J., Schoorman, F.: An integrative model of organizational trust. Academy of management review **20**(3) (1995) 709–734
11. Nohria, N., Eccles, R.G.: Networks and organizations: structure, form, and action. Harvard Business School Press (1992)
12. Lyons, B.: Contracts and specific investment: An empirical test of transaction cost theory. Journal of Economics and Management Strategy **3**(2) (1994) 257–278
13. Lorenz, E.: Neither Friends nor Strangers: Informal Networks of Subcontracting in French Industry. In: Trust: Making and Breaking Cooperative Relations. Basil Blackwell, New York (1988)
14. Marsh, S.: Formalising Trust as a Computational Concept. PhD thesis (1994)
15. Mezzetti, N.: A socially inspired reputation model. Lecture Notes in Computer Science (2004) 191–204
16. Keßler, C., Janowicz, K., Bishr, M.: An agenda for the next generation gazetteer: Geographic information contribution and retrieval, In 17th ACM SIGSPATIAL – International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL GIS 2009), ACM press (November 2009)
17. Bishr, M., Mantelas, L.: A trust and reputation model for filtering and classifying knowledge about urban growth. GeoJournal **72**(3) (2008) 229–237