INCLUDING USERS' SEMANTICS IN EVALUATING THE CREDIBILITY OF CROWDSOURCED LANDSCAPE DESCRIPTIONS

A. Forati [a,*], S. Soleimani [b], F. Karimipour [a] and M.R. Malek [b]

[a] Faculty of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran
(forati, fkarimipr)@ut.ac.ir
[b] Faculty of Geodesy and Geomatics Engineering, K.N. Toosi University of Technology, Tehran, Iran
ssoleimani@mail.kntu.ac.ir, m.malek@kntu.ac.ir

Commission II, WG II/4

KEYWORDS: Landscape Description, Crowdsourcing, Credibility, Users' Semantics

ABSTRACT:

Landscape refers to the visible features of an area of land often considered in terms of their aesthetic appeal. The spatial configurations, composition conditions and perception situations of a landscape may be described by people, which reveal how they see and percept the landscape; thus they are different from person to person depending on the way they think or experience their surroundings. Especially, when the landscape descriptions are acquired through crowdsourcing processes, it is affected by users' semantics, which results in descriptions with varying credibilities. This paper proposes an approach to consider the user's semantics in evaluating the creditability of crowdsourced landscape descriptions.

1. PROBLEM DEFINITION

Landscape refers to the visible features of an area of land often considered in terms of their aesthetic appeal. The notion of landscape is also illuminated by the concepts of land, nature, space, and temporal elements as well as people's attitude towards them (Litton 1968, Derungs and Purves, 2013). There are many different perspectives in order to define landscape, such as recreational, silvicultural, hydrological, ecological, acoustical, and wildlife, to name a few (McGarical, 1995). However landscape definition mostly refers to the visible features of an area of land to be considered as recreational areas.

Landscape description provides insight into the worlds of perception. Knowing how people see and perceive a landscape imparts additional information about different characteristics of landscape that could be used in aesthetic analysis, environmental protection, visual pollution, way finding, and landscaping. The landscape descriptions may be categorized into (1) spatial configurations; (2) composition conditions; and (3) perception situations. *Spatial configurations* of a landscape refer to how the available elements are arranged. Distance between elements, direction of elements respect to an origin, visible sections of each element from specific points of view, and topological relations between elements are among metrics that define landscape configuration (Litton 1968, Yaouanc et al., 2009). Figure 1 illustrates some of the spatial configurations of a landscape. *Composition conditions* mostly quantify the distribution, variety, and amount of features such as plants and animals in an area without considering spatial qualities (McGarigal et al., 1995; Ritchie et al., 2009). As composition metrics includes the environmental structure of landscape, the environmental pollution and even animal migration could be discussed in this category. Finally, *perception situations* and knowledge of landscape determine the way people perceive various landscapes (Bell, 2012). Observers, depending on their outlooks, have specific feelings when they are in a certain type of landscape, which among others, highly corresponds to personal factors.

The main issue in acquiring the landscape descriptions is that perception of land varies from person to person depending on the way they think or experience their surroundings. There are evidences that parameters like gender, age, education, profession, and even health conditions influence landscape descriptions provided by people (Schirpke et al., 2013).

This is more challenging when the landscape descriptions are acquired through crowdsourcing processes (Howe, 2006), by which a broad spectrum of heterogeneous data resources (which is affected by users' semantics) could be generated at a rapid rate. In addition, as crowdsourced data is mostly based on human experience of geography, deploying perception-based parameters to express their spatial quality is more efficient than measurement-based parameters used in case of official spatial data.

Generally, crowdsourcing and participating is a relatively new method in the field of landscape science. However, Brown and Brabyn (2012) conducted a research on different compositional values of the landscape using public participation GIS. This study attempts to focus on the compositional values such as wilderness, aesthetic, spiritual, recreation, and economic that are important to all people, without considering the semantic of the participants. However, according to different categories of landscape descriptions, the semantic of the observer who describes a view plays a role in her/his descriptions.

Among several perception-based parameters suggested in the literature, Flanagin and Metzger (2008) introduce the concept of *credibility*, as a perceptual variable, for evaluating collaborative productions. "Although there is no clear definition of credibility, it is generally thought to be the believability of a source or message, which is composed of two primary

---

* Corresponding author

dimensions: trustworthiness and expertise" (Flanagin and Metzger, 2008). This paper proposes an approach to consider the user's semantics in evaluating the credibility of crowdsourced landscape descriptions. We draw inspiration from the Kessler et al. (2013)'s definition of trustworthiness and extend it by relating data trustworthiness, user reputation and accounting social group reputation to.

Distance

Direction

Topology

Figure 1. Some of the configuration conditions in landscape description

2. SEMANTIC ISSUES IN LANDSCAPE DESCRIPTION

The landscape is a combination of natural and conceptual phenomena. Spatial configurations and composition conditions are essential to describe the landscape. However, cross-cultural issues are still questionable in understanding how people perceive the landscapes (Purves, 2008). The landscape description almost acts as a mental tool to help people share their experiments and thoughts (Cohen, 2013).

People may possess different perspectives when facing a specific landscape. Furthermore, their mental possession affects the way they define and describe a landscape. Some assessment tools associated with physical and mental features of landscape, called *metrics* (Yaouanc et al., 2009), could be combined to provide a formal environment for differentiating the structures of landscapes. These metrics could be easily used by all expert and non-expert people. However, the credibility of acquired data would vary due to users' semantics.

As discussed, landscape has different aspects when it comes to be described as a mental and physical phenomenon. For visual aspect, there are some metrics namely, topology relations, direction relations, distance, solid angle, and salience (Soleimani, 2015) that almost depend on user's mindset, experiences, skills, and even culture. To illustrate, thinking about a feature in a landscape as a salient, one is dependent to the observer preference. Moreover, according upon various human's skills, some observer could estimate solid angle and distance more precisely than others. Obviously, such differences affects the credibility of the collected data them.

The crowdsourced data is more based on human cognition and experience than on measurements (Karimipour and Azari, 2015), thus describing a landscape through a crowdsourcing process may undermine the credibility of the results due to varieties in humans' landscape cognition as well as their semantics.

Therefore, in order to evaluate the credibility of users' descriptions of landscapes, one should consider the users' subjective cast of minds. For instance, in case of spatial configuration of landscape, the descriptions provided by the engineer participants may be considered more credible. More precisely, it seems that participant with measuring skills provide more credible "distance" and "area", participant with a geography background may produce more credible "direction", and those who are expert at mapping activities produce more credible topological data.

3. PROPOSED APPROACH

As discussed, the semantics of users, as sensors in crowdsourcing process, have an effective role in data credibility. This paper defines the credibility of landscape description based on the users' semantics. We rank the users for different metrics of landscape description, and correspondingly, the credibility of the data they provide.

A web-based platform (Figure 2) is designed to collect the required information about the landscape description. In order to produce the visual data of a landscape, participants have to fill out the predefined metrics embedded in the designed web-page, such as available salient elements and their estimated visual area. Some other information including coordinates, time, and even weather could be gathered automatically considering the available information in the geo-tag image uploaded by the participant.

Besides, a set of user information is acquired for the desired credibility assessment. As a matter of fact, knowing user's age, gender, education, profession, interests, and disease, to name but a few, could give us an insight on how credible is the data he/she could produce.
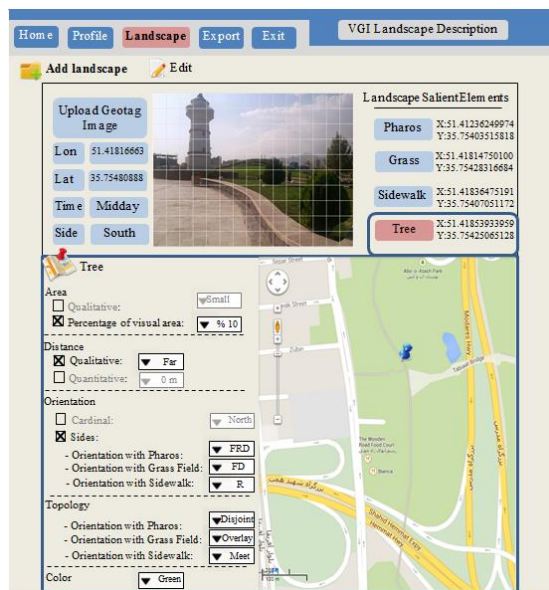
Figure 2. The web-based platform designed to collect the crowdsourced landscape descriptions

The steps of the proposed approach are as follows (Figure 3):

Classification: In order to evaluate the credibility of the data provided, the users are classified into different groups such as civil engineers, surveyors, doctors, artists, to name a few, and initial credibility values are assigned to each group for each aspect of landscape description (i.e., configuration, composition, and perception issues). This specifies the users' ability to describe various aspects of landscape. The credibility of these aspects directly relate to the users semantics and conceptions. For example, we expect that the "distance" provided by an engineer participant who is always interacting with dimensions, incontrovertibly, has more credibility compared to the data acquired by an artist; while the introduced "color" by an artist participant could have a more credibility rather than the one suggested by an engineer.

Credibility evaluation: The crowdsourced landscape description acquired by a user will be assigned the relevant credibility values: For an already registered user, his/her own credibility values will be applied. For a new user, the initial credibility would be the one assigned to the corresponding group for the desired aspect of landscape description; and if the user is not a member of any classes, a fuzzy inference process will be applied to assign relevant initial credibility values, which as we see in the third step, will be updated during the data acquisition process. By integrating the data provided by different users, a final description of that aspect will be specified with a certain credibility.

Update: The above credibility resulted from integrating the crowdsourced data will be used to update the initial credibility values assigned to the individual registered users, as well as different user groups through established methods suggested in the literature (for example, Chatterjee et al., 2008; Adler et al., 2011; Adler and Alfaro, 2007).
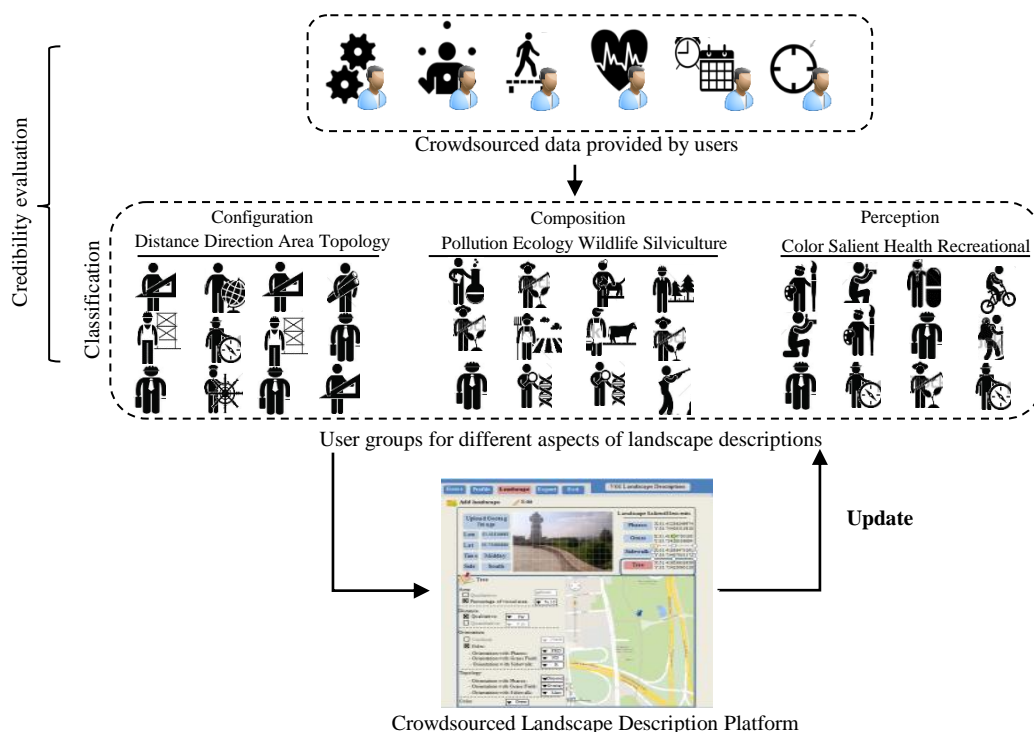

Figure 3. The proposed approach to include the users' semantics in landscape description

4. CONCLUSION

This paper proposes an approach to include the users' semantics in evaluating the credibility of landscape descriptions acquired through crowdsourced processes. The initial outcomes of implementing the proposed approach show an improvement in the final results. However, grouping and assigning proper initial credibility values are the main concerns to be considered in future. We are also working on generalizing the approach to be deployed for other types of crowdsourced data.

REFERENCES

Adler, B.T. and Alfaro., L.D., 2007. A Content-Driven Reputation System for the Wikipedia, Proceedings of the 16th International Conference on World Wide Web, Banff, Alberta, Canada, pp. 261-270.

Adler, B.T., Alfaro, L.D., Mola-Velasco, S.M., Rosso, P. and West, A.G., 2011. Wikipedia Vandalism Detection: Combining Natural Language, Metadata and Reputation Features, Computational Linguistics and Intelligent Text Processing, LNCS Vol. 6609, pp. 277-288.

Bell, S., 2012. Landscape: Pattern, Perception and Process. Routledge.

Brown, G. and Brabyn, L., 2012. An Analysis of the Relationships between Multiple Values and Physical Landscapes at a Regional Scale using Public Participation GIS and Landscape Character Classification. Landscape and Urban Planning, 107(3), 317-331.

Chatterjee, K., Alfaro, L.D. and Pye, I., 2008. Robust Content-Driven Reputation, Proceedings of the 1st ACM Workshop on AISec, Alexandria, Virginia, USA, pp. 33-42.

Cohen, R. (Ed.), 2013. The Development of Spatial Cognition, Psychology Press.

Derungs, C., and Purves, R.S., 2013. From Text to Landscape: Locating, Identifying and Mapping the Use of Landscape Features in a Swiss Alpine Corpus, International Journal of Geographical Information Science, 28(6), pp. 1-22.

Flanagin A and Metzger M., 2008. The Credibility of Volunteered Geographic Information, GeoJournal, 72(3), pp. 137-148.

Howe, J., 2006. The rise of crowdsourcing. Wired magazine, 14(6), 1-4.

Karimipour, F. and Azari, O., 2014. Citizens as Expert Sensors: One Step Up on the VGI Ladder, In: Progress in Location-Based Services, Lecture Notes in Cartography and Geoinformation, Springer, pp. 213-222.

Keßler, C., Theodore, R. and de Groot, A., 2013. Trust as a Proxy Measure for the Quality of Volunteered Geographic Information in the Case of OpenStreetMap, In: Vandenbroucke, D., et al. (eds.), Geographic Information Science at the Heart of Europe, Lecture Notes in Geoinformation and Cartography, Springer, pp. 21-37.

Litton, R.B., 1968. Forest Landscape Description and Inventories - a Basis for Landplanning and Design.

McGarigal, K. and Marks, B.J., 1995. Spatial Pattern Analysis Program for Quantifying Landscape Structure. Technical Report PNW-GTR-351. US Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Purves, R.S. and Edwardes, A.J., 2008. Exploiting Volunteered Geographic Information to Describe Place. Proceedings of the GIS Research UK 16th Annual Conference, pp. 252-255.

Ritchie, L.E., Betts M.G., Forbes, G., and Vernes, K., 2009. Effects of Landscape Composition and Configuration on Northern Flying Squirrels in a Forest Mosaic, Forest Ecology and Management, 257(9), pp. 1920-1929.

Schirpke, U., Tasser, E., and Tappeiner, U., 2013. Predicting Scenic Beauty of Mountain Regions. Landscape and Urban Planning, Vol. 111, pp. 1-12.

Soleimani, S., 2015. Landscape Description in Volunteered Geographic Information (VGI) Using Spatial and Temporal Relationships, MSc Thesis, KNT University, Tehran, Iran (In Persian).

Yaouanc, L., Saux, J.M. and Claramunt, C., 2010. A Semantic and Language-based Representation of an Environmental Scene. Geoinformatica, 14(3), pp. 333-352.