Criteria of Geographic Relevance

S. De Sabbata

Department of Geography, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
Email: stefano.desabbata@geo.uzh.ch

1. Introduction

In the last few years, remarkable advances in mobile computing, telecommunication services, and positioning systems have radically changed the way we seek for geographic information. Smart-phones, mobile internet, and vast amounts of geographic data are becoming available to an ever-growing number of users. This new scenario calls for new methods and systems able to provide relevant geographic information to mobile users.

To handle this new challenge, the concept of Geographic Relevance (GR) has been proposed by Raper (2007). GR can be defined as a quality of an entity in the geographic space (such as shops and museums) or its representation (i.e., an object, document, or image), expressed as the strength of the relationship between entity and actual context of use (Reichenbacher et al. 2009).

This definition encompasses the concept of wireless/mobile relevance proposed by Coppola et al. (2002), since it entails a situational relevance (Wilson 1973, Saracevic 2007) that deals with the user context and the objects in the physical world.

The relationship between the context of a mobile user and a geographic object in her environment is very complex. It involves not only the user's interest, but also her position, time schedule, current activity and knowledge of the environment. It also involves the category the geographic object belongs to, together with its location, time validity, and affordance. Thus, to assess the geographic relevance of an entity with respect to a given usage context, it is necessary to analyse this relationship using an appropriate set of criteria of relevance.

The aim of this paper is to raise the issue of defining this set of criteria of geographic relevance. The next two sections present the criteria of relevance proposed in Information Retrieval (IR), and how they can be applied to GR. In addition, four new criteria strictly related to the nature of GR are proposed.

2. Criteria of Relevance

According to Mizzaro (1997), in the period 1959–1976, many researchers started looking into criteria other than topicality adopted by experts in judging the relevance of documents. Since the 1980s, studies in this area have focused on the criteria taken into account by the users of information systems.

Barry and Schamber (1998) compared the criteria identified in two independent studies (Schamber 1991, Barry 1994) and developed a list of 10 common criteria. All these criteria are fully applicable to GR, either as criteria to judge the relevance of a geographic entity (depth/scope/specificity, availability of information/sources of information, and affectiveness) or its representation (accuracy/validity, clarity, currency, tangibility, reliability/quality of sources, accessibility, and verification).

More recently, other studies have proposed different criteria of relevance. Xu and Chen (2006) suggested a five-factor (topicality, novelty, reliability, understandability, and scope) model of relevance and argued that novelty is a major criterion of relevance. This criterion, also identified by Barry (1994), can be crucial in GR. In fact, it is reasonable to think that in most scenarios a user would tend to search for geographic entities unknown to him rather then well-known ones.

Savolainen and Kari (2006) studied the criteria of relevance used in web searching, taking into account also the literature on image and video information seeking. They analysed the user-defined criteria when accepting hyperlinks during web browsing, and identified three new interesting criteria that can be applied to GR, namely: familiarity, variety, and curiosity.

In (da Costa Pereira et al. 2009), the authors proposed coverage and appropriateness as new criteria of relevance for personalized IR. The coverage criterion measures "how strongly the user interest is included in a retrieved document", whereas the criterion of appropriateness measures how focused the document is on the topic the user is interested in. Both criteria can be useful in judging the relevance of a geographic entity, for example in judging how the affordance of a geographic entity matches the current activity of the user.

3. Criteria of Mobile Relevance

Schamber (1991), interviewing 30 users of a weather information system, identified the criterion of geographic proximity — the closer an entity is, the more relevant it is. This is one of the main criteria used in Geographic Information Retrieval (GIR) and plays a central role in Location Based Systems (LBSs). Moreover, Schamber (1991) also identified two categories of criteria related to the representation of information (i.e., dynamism and presentation quality), that would have a straightforward application in a map-based system.

Bierig and Göker (2006) analysed the criteria employed by users to evaluate the relevance of local events, and found that temporal proximity causes "an overall large effect on users' perception of usefulness". Given the strong interaction between location, time, and topicality, the authors also concluded that it is necessary to include the "interactive behaviour" between these three elements within the model.

Mountain and MacFarlane (2007) described four filters for Mobile Information Retrieval (MIR), that is four criteria to filter information based on the location of a mobile user. The criterion of spatial proximity is strictly related to the geographic proximity mentioned by Schamber (1991). The criterion of (spatio-)temporal proximity states that entities that can be reached in a short period of time are more relevant than those that are temporally distant. The visibility criterion is based on the assumption that a user in a mobile environment is interested in what she can see. Nevertheless, the authors also suggest the opposite criterion of concealed places, assuming an interest on entities that are close, but not visible. Finally, based on the search-ahead filter proposed by Mountain and MacFarlane (2007), the criterion of directionality is here defined as: a user is more interested in what she is going to find in her path than in those entities that she has already passed.

4. Criteria of Geographic Relevance

All the criteria discussed in the two previous sections have a clear application to GR, and will be fundamental in the implementation of a GR-based system. Here four new criteria, not yet discussed in literature, are introduced that bring an important contribution to GR.

The geographic entities considered in a relevance judgment do not exist as independent entities, but rather they exist within a specific geographic context. Entities are commonly part of more complex phenomena which have to be taken into account. For example, if a shop is part of a shopping centre, it is unlikely that a user would not consider this fact when judging the relevance of that shop.

Based on this assumption and based on well established geographical concepts and methods, I propose the following additional criteria of GR:

- **hierarchies**: degree of separation within a spatial hierarchy — based on the evidence that "geographic units are cognitively and empirically organized into a nested hierarchical form (e.g., school districts)" in Golledge (2002: 8);
- **clusters**: whether a geographic entity is part of a spatial cluster of related or unrelated entities — based on the methods discussed in (Han et al. 2001);
- **co-location patterns**: whether an entity satisfies a co-location pattern typical for its category — based on the methods discussed in (Shekhar and Huang 2001);
- **spatio-temporal association rules**: whether an entity satisfies a rule, involving spatial, temporal and/or other attributes, identified within a related collection of entities — based on the methods discussed in (Koperski and Han 1995).

The analysis of co-location patterns can be used to identify correlated services or businesses commonly located close to each other. For example, if it is common to have restaurants close to hotels, a user would probably consider this co-location in judging the usefulness of a hotel under examination.

Likewise, it is reasonable to suppose that common spatio-temporal rules (identified within a set of given entities) are involved in a relevance judgment, since they can be assumed to be known to the user. A system can also take into account these rules to infer missing information, i.e. to simulate an educated guess made by the user. For example about the opening hours of a shop given its location, or about how expensive a hotel is given its location and services.

Table 1. Preliminary list of criteria of geographic relevance.

Properties	Geography	Information	Presentation
topicality	spatial proximity	specificity	accessibility
appropriateness	temporal proximity	availability	clarity
coverage	spatio-temporal proximity	accuracy	tangibility
novelty	directionality	currency	dynamism
	visibility	reliability	presentation quality
	hierarchies	verification	
	clusters	affectiveness	
	co-location	curiosity	
	association rules	familiarity	
		variety	

5. Conclusions

In this paper, I have presented a review of the criteria of relevance proposed in literature and suggested the introduction of four new criteria: hierarchies, clusters, co-location, and association rules. These criteria take into account the geographic context of the entities involved in the assessment of GR. The complete list of criteria identified in this preliminary discussion is presented in Table 1. The criteria are classified into four sets: the criteria used in judging the entity by means of its properties (Properties) and by means of its geographical essence (Geography), and the criteria used in judging the representation of the entity within the system by means of the available information (Information) and by means of the way the information is presented to the user (Presentation).

As a next step, a web-based questionnaire will be developed in order to evaluate the criteria presented. In it, the users will be asked to explain which criteria they would use to judge the relevance of geographic entities in a set of given scenarios. Further steps will include the implementation of the identified criteria of GR in a geographic mobile information system and a user study of the system prototype.

References

Barry C L, 1994, User-defined relevance criteria: An exploratory study. *Journal of the American Society for Infromation Science*, 45(3), 149–159.

Barry C L and Schamber L, 1998, Users' criteria for relevance evaluation: a cross-situational comparison. *Journal Information Processing and Management*, 34: 219–236.

Bierig R and Göker A, 2006, Time, location and interest: an empirical and user-centred study. *Proceedings of the 1st international conference on Information interaction in context*.

Coppola P, Della Mea V, Di Gaspero L and Mizzaro S, 2002, From e-relevance to w-relevance. *Mobile Personal Information Retrieval, Proceedings of the SIGIR2002 Workshop,* 64–72.

da Costa Pereira C, Dragoni M and Pasi G, 2009, Multidimensional relevance: A new aggregation criterion. *Journal of Advances in Information Retrieval, Proceedings of the 31st European Conference on IR Research*.

Golledge R G, 2002, The nature of geographic knowledge. *Annals of the Association of American Geographers*, 92(1): 1–14.

Han J, Kamber M and Tung A K H, 2001, Spatial clustering methods in data mining: A survey. *Geographic Data Mining and Knowledge Discovery*.

Koperski K and Han J, 1995, Discovery of spatial association rules in geographic information databases. *Advances in spatial databases*:47–66.

Mizzaro S, 1997, Relevance: The whole history. *Journal of the American Society for Information Science*, 48(9): 810–832.

Mountain D and MacFarlane A, 2007, Geographic information retrieval in a mobile environment: evaluating the needs of mobile individuals. *Journal of Information Science*, 33(5): 515–530

Raper J, 2007, Geographic relevance. *Journal of Documentation*, 63(6):836–852.

Reichenbacher T, Crease P and De Sabbata S, 2009, The concept of geographic relevance. *Proceedings of the 6th International Symposium on LBS & TeleCartography*.

Saracevic T, 2007, Relevance: A review of the literature and a framework for thinking on the notion in information science. part ii: nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology*, 58(13):1915–1933.

Savolainen R and Kari J, 2006, User-defined relevance criteria in web searching. In *Journal of Documentation*, 62(6): 685–707.

Schamber L, 1991, Users' criteria for evaluation in multimedia information seeking and use situations. *Information Processing and Management*, 30(2), 205–221.

Shekhar S and Huang Y, 2001, Discovering spatial co-location patterns: A summary of results. *Advances in Spatial and Temporal Databases*: 236–256.

Wilson P, 1973, Situational relevance. *Information storage and retrieval*, 9(8):457– 471.

Xu Y and Chen Z, 2006, Relevance judgment: What do information users consider beyond topicality?. *Journal of the American Society for Information Science and Technology*, 57(7): 961–973.