Geographic relevance – different notions of geographies and relevancies

Tumasch Reichenbacher and Stefano De Sabbata

Department of Geography, University of Zurich, Switzerland

tumasch.reichenbacher@geo.uzh.ch; stefano.desabbata@geo.uzh.ch

Introduction

Geographic information is increasingly used in mobile contexts. Mobility constrains several aspects of this mobile usage, such as limited screen estate, number of desired interactions, or availability due to battery time. As a consequence designers and developers of mobile geographic applications strive to reduce the amount of information delivered to users and aim at serving just useful content rather than a larger amount of unspecific data. This is to address the problem of information overload as well as to provide information that is relevant. Ultimately this is the goal of any information retrieval system.

Thus, we want to set off by studying how relevance of information is defined and handled in information science/information retrieval (IR), Geographic Information Retrieval (GIR), Mobile Information Retrieval (MIR), Location Based Services (LBS), and Geographic Relevance (GR). The objective is to reveal commonalities as well as differences among these fields, mainly between GIR and GR, and eventually show potential mutual benefits.

Relevance in information retrieval

The term relevance has been widely used in information science and particularly in Information Retrieval (IR) since the first introduction of IR systems in the 1950s and 1960s and refers to seeking for information stored in collections of documents driven by documents' relevance (or pertinence) to a user's information need. Thus, the classic document IR systems (e.g. web search engines) mainly deal with the semantic similarity between a set of query terms and the terms found in documents. The concept of relevance behind the first generation of IR systems was not only based on a single dimension, but it was also binary (i.e., each document could be either relevant or not relevant) [1]. However, more recently many researchers [e.g., 2, 3] claimed that relevance is a multidimensional concept, a system of relevancies on different levels with degrees of relevance rather than a single, binary relevance reflected in ranks for the retrieved documents, based on the similarity of the document and the query.

Relevance in geographic information retrieval

The traditional approach of IR does not take much account of explicit or implicit spatial references of information. This circumstance may lead to poor or unsatisfactory results of queries that involve spatial information, such as toponyms or spatial relations. The research field of Geographic Information Retrieval (GIR) at the intersection of traditional IR and Geographic Information Science (GIScience) aims at not only employing the theme, but also the geographic scope of documents when retrieving documents [4]. Therefore, relevance of documents is also defined by their spatial footprints. Examples of this approach can be found in [5]. GIR is typically employed in web browser searches for documents or images in unstructured collections implying that the geographic information is implicitly represented in web documents (e.g., as toponyms) that can be first geo-parsed and then geo-referenced through GIR processing.

The extension of GIR to mobile usage has led to a new field termed Mobile Information Retrieval (MIR). In MIR, information retrieval algorithms developed in GIR are applied in a mobile environment. Therefore the geographic context (mainly location and time) becomes part of the query processed by a spatio-temporal IR algorithm. Examples of these methodologies are presented in [6, 7]. The MIR approach proposed by [6] can be seen as the use of binary functions on different relevance dimensions (e.g., space, time, visibility).

Relevance in location based services

Related to MIR are Location-Based Services (LBS) that employ some notion of spatial relevance. LBS are mobile applications that aim to offer spatial information processing capabilities on the basis of the user's location [8]. LBS usually employ only simple spatial concepts to assess binary values of relevance (e.g., buffers around the user's position as a binary information filter). Yet, apart from spatial context, there are several other context dimensions only scarcely addressed in research that give rise to contextual information needs and influence the relevance of geographic objects in such usage situations.

Geographic Relevance

To overcome this research gap several researchers in GIScience [e.g., 9, 10, 11] proposed more recently the concept of Geographic Relevance (GR) offering a different perspective to the relevance of geographic information. We define GR as a quality of an entity in geographic space or its representation in an information system, i.e. an object, document, or image. This quality is expressed as the relation between an entity or its representation and the actual context of using the representation. The relevance may refer to different levels of complexity of the represented entities. First, relevance can refer to discrete objects (e.g., point, linear or areal features), or properties of these objects. Second, relevance can also denote spatial, temporal, or spatio-temporal relations between objects. Third, relevance may relate to groups of objects (e.g., clusters, structure, configuration). And finally, relevance can relate to functions of objects, or places.

More pragmatically, GR aims to assess the relevance of an object that is a representation of a geographic entity within a computer system or information database. This object can be a collection of documents or an entry in a database describing a point of interest. Still, even if the object is a single document, the objective is to approximate the relevance of the entity, not to judge the relevance of a geo-referenced document or a document reporting geographic information.

Next we will analyse fundamental differences as well as shared properties between GR and the previous concepts of relevance.

Differences and commonalities between relevance conceptualisations

Even though our definition of GR subsumes many of the general characteristics of information relevance as well as concepts from GIR, GR is characterised by the following qualities that differentiate it from them.

As opposed to IR and GIR that are predominantly concerned with retrieving documents, GR works on geographic objects representing reality and represented in geographic databases or as maps. The assessment of GR based on spatial representations means that the information objects are spatially structured (in the physical world) compared to unstructured digital document collections on the Internet (the informational world) that are the focus of GIR. Therefore, the

notion of GR implies an enhanced engagement with the physical world, and we can observe a relationship between the concepts in the mind, the real objects in physical space, and their representation on a visuo-spatial display. The engagement with reality further entails a super ordinate user activity for which the information seeking or retrieval action is only a part and which may, through the affordances of objects or places, trigger further information needs.

The situatedness of mobile users also requires GR to integrate more than just topical and spatial dimensions of the usage context. Thus GR extends the context dimensions to time (e.g., temporal validity of geographic objects), personal mobility, and environmental factors. Typically GIR does not include temporal relevance and spatio-temporal relevance in the sense of time geography. On the contrary, the personal mobility opportunities and limitations, combined with the spatio-temporal availability of the geographic entities, are very important factors in the estimation of the usefulness of an entity in a given situation, just as it is to a certain extent in MIR.

Moreover, as mentioned above, the entities taken into account in GR assessment are spatially structured. That is, they are not simply separate objects placed in an abstract Euclidean space at a certain distance from the location representing the position of the user. They are spatially organised and related in certain ways. This spatial layout can have historical, geographical, economical, or social reasons. The neighbourhood of entities in space can often be referred to these reasons, which shape the cities we live in, and are also the cause of more complex geographic phenomena. For example, entities can be part of spatial hierarchies or clusters that developed within the city as it has evolved over time. Entities can follow co-location rules considered common for a given city, which can be completely meaningless in another city, or in a neighbouring country. It can be misleading to assume a priori that these spatial structures are not part of the situation the user is in, or to assume that a user would not consider them as important facets of relevance. In fact, these spatial structures have so far not been taken into account as part of relevance assessment.

Outlook to mutual benefits

GR can be seen as an attempt to join the concepts of geographic related relevance developed within GIR with the recent advances of context-aware mobile geovisualisation and LBS. This has brought us a new concept of relevance, which is clearly different from its precursors by its tighter engagement with the physical world — in contrast with the informational world of the documents — and the dynamics of the user's activity and mobility.

This closer relationship to the physical world has given GR the opportunity to fully unveil facets of relevance that previously had been only partially studied (e.g. the personal mobility, introduced in MIR) or not covered at all (e.g., the spatial structure of the entities in the user's surrounding environment). Nevertheless, the usage of these new facets should not remain restricted within GR, but can be transferred back to MIR and to GIR in general.

In fact, if these criteria are useful to judge the relevance of an entity (in the physical world), then a document (in the informational world) referring to that entity and reporting about these facets (e.g., a spatial cluster the entity is part of) is more informative to the user than a document that does not report this information. This implies that the first document is more relevant than the second. Thus, these new criteria can be useful in the assessment of the relevance of a document reporting geographic information. It is then clear how the advances in the study of GR can also change the understanding of relevance in MIR and GIR.

References

[1] Mizzaro, S., 1998. Relevance: The whole history. Historical studies in information science, 221–244.

[2] Saracevic, T. *Relevance reconsidered*. in *Second Conference on Conceptions of Library and Information Science (CoLIS 2)*. 1996. Copenhagen.

[3] Cosijn, E. and P. Ingwersen, Dimensions of relevance. *Information Processing and Management*, 2000. **36**: p. 533-550.

[4] Purves, R.S. and C.B. Jones, Geographic Information Retrieval (GIR). *Computers, Environment and Urban Systems*, 2006. **30**(4): p. 375-377.

[5] Jones, C.B**.** , et al., Spatial information retrieval and geographical ontologies an overview of the spirit project. SIGIR 2002: Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Tampere, Finland, 2002, p. 387–388.

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

[7] Carpineto, C., et al., Mobile information retrieval with search results clustering: Prototypes and evaluations. *Journal of the American Society for Information Science and Technology*, 2009. **60**(5).

[8] Shiode, N., et al., *The impact and penetration of location-based services. Telegeoinformatics: Location Based Computing and Services*. 2003, London: Taylor and Francis.

[9] Raper, J., Geographic Relevance. *Journal of Documentation*, 2007. **63**(6): p. 836 - 852.

[10] Reichenbacher, T. *The Importance of Being Relevant*. in *XXII International Cartographic Conference*. 2005. A Coruna, Spain.

[11] Zipf, A. and K.-F. Richter, Using Focus Maps to Ease Map Reading: Developing Smart Applications for Mobile Devices. *Künstliche Intelligenz*, 2002. **4**(2): p. 35-37.

[12] De Sabbata, S., *Criteria of Geographic Relevance*, in *GIScience 2010 - 6th International Conference on Geographic Information Science*. 2010: Zurich, Switzerland.