User-Centric Similarity and Proximity Measures for Spatial Personalization

Yanwu Yang, Chinese Academy of Sciences, China
Christophe Claramunt, Naval Academy Research Institute, France
Marie-Aude Aufaure, Ecole Centrale Paris, France
Wensheng Zhang, Chinese Academy of Sciences, China

ABSTRACT

Spatial personalization can be defined as a novel way to fulfill user information needs when accessing spatial information services either on the web or in mobile environments. The research presented in this article introduces a conceptual approach that models the spatial information offered to a given user into a user-centered conceptual map, and spatial proximity and similarity measures that considers her/his location, interests and preferences. This approach is based on the concepts of similarity in the semantic domain, and proximity in the spatial domain, but taking into account user’s personal information. Accordingly, these spatial proximity and similarity measures could directly support derivation of personalization services and refinement of the way spatial information is accessible to the user in spatially related applications. These modeling approaches are illustrated by some experimental case studies.

Keywords: Conceptual Map, Similarity Measure, Spatial Personalization, Spatial Proximity, User-Centric

INTRODUCTION

Personalization offers many opportunities to improve the way information is delivered to a given user. Web personalization should help to remove irrelevant information, to adapt web content and structure, and thus to improve the quality of web services and user’s satisfaction. E-commerce and distance learning are some of domains where web personalization has been successfully applied, while personalization has been also applied to other domains, particularly where the information is complex and diverse per nature.

Spatial web personalization can be considered as an emerging research area of interest as the amount of geo-referenced data distributed either on the web or mobile environments grows exponentially. Let us consider that spatial data represents any item of information that is in one
way or the other related to a given geographical location. Namely, geographical entities that appear on a map, or items that represent an information or an object located on Earth (e.g., a photograph, a given landmark).

The development of spatial Web services has been widely active over the past few years, together with significant progresses on spatial information retrieval and location-based services that open novel perspectives for real-time diffusion of geographical data (Harle & Hopper, 2008; Kwon & Shin, 2008). With respect to whether a given user is located in the virtual (such as the web) or in the physical environment, spatially related information services are closely related to the fields of spatial web information retrieval and location-based services.

In the past decade, the progressive integration of spatial information within web pages has been widely investigated to facilitate spatially related information retrieval on the web (Ding et al., 2000; Gravano et al., 2003). Two complementary strategies have been applied: a technique based on the spatial distribution of web links to a given page, and another based on the distribution of spatial references in a given web page. With consideration of user’s location, (Gravano et al., 2003) identified whether a given query submitted to a search engine should be oriented towards either “local” or “global” pages. For example, a query oriented to some general biological information on “wildflowers” should consider global information, while a search for some “houses for sale” after local information, as this later case is locally oriented per definition.

More specifically, location-based spatial queries usually consider spatial queries that retrieve information based on the users’ current locations which is likely to be mobile. When considering location-based spatial queries, a given user (and thus the entities of interest) is likely to be in displacement in the physical environment. Its movement can be represented as a sequence of time-stamped locations. Over the past few years, there has been several works oriented towards mobile spatial queries and sequential patterns (Jayaputera & Taniar, 2005), including nearest neighbor (NN), k nearest neighbor, continuous k nearest neighbor (Zhao et al., 2008; 2009), reverse nearest neighbor (Lee et al., 2008) and range search (Xuan et al., 2008). When considering road networks, (Zhang et al., 2004) studied spatial queries in the presence of obstacles, where the distance between two points is defined as the length of the shortest path that connects them without crossing any obstacles. (Zhang et al., 2003) considered a moving query as a query which is reevaluated only when the query exits the validity scope. Taking location dependency (Goh & Taniar, 2004a) into account, Goh and Taniar (2004b; 2005; 2007) built user profiles based on past mobile visiting data, filters and to mine mobile pattern and association rules from mobile users. (Wang et al., 2006; 2008) presented an approach that derives group patterns of mobile users based on their movement data, where a group pattern is defined as a group of users that are within a distance threshold from one another for at least a minimum duration.

Although spatial information is diverse and multiple, similarities in the semantic domain are likely to emerge as information is often delivered per domains of interest. Similarly, spatial entities of similar interest are often located at proximity as stated by the First Law of Geography (Tobler, 1970). The notion of similarity has been considered in web engineering as an important component for personalization (Baeza-Yates & Ribeiro-Neto, 1999; Mobasher et al., 2000; Kumar et al., 2007) and clustering (Zhang et al., 2008). Similarity measures have been used to identify the degree of similarity between two information entities. Similarity measures between trajectories on road networks are crucial to discover sequential patterns in moving trajectories and search for similar trajectories on road networks. Similarity between trajectories can be defined as non-metric distance functions based on trajectory properties such as longest common subsequence, or spatio-temporal functions.

To the best of our knowledge, the role of space when deriving similarities in an information space has been hardly (if any, not fully)
Efficient Algorithm for Mining High Utility Pattern Considering Length Constraints
www.igi-global.com/article/efficient-algorithm-for-mining-high-utility-pattern-considering-length-constraints/228935?camid=4v1a

The Role of Schema and Document Matchings in XML Source Clustering
Pasquale De Meo, Giacomo Fiumara, Antonino Nocera and Domenico Ursino (2012). *XML Data Mining: Models, Methods, and Applications* (pp. 125-153).
www.igi-global.com/chapter/role-schema-document-matchings-xml/60907?camid=4v1a