Chapter 4.13
Applying Commonsense Reasoning to Place Identification

Marco Mamei
Università di Modena e Reggio Emilia, Italy

ABSTRACT

Recent mobile computing applications try to automatically identify the places visited by the user from a log of GPS readings. Such applications reverse geocode the GPS data to discover the actual places (shops, restaurants, etc.) where the user has been. Unfortunately, because of GPS errors, the actual addresses and businesses being visited cannot be extracted unambiguously and often only a list of candidate places can be obtained. Commonsense reasoning can notably help the disambiguation process by invalidating some unlikely findings (e.g., a user visiting a cinema in the morning). This paper illustrates the use of Cyc—an artificial intelligence system comprising a database of commonsense knowledge—to improve automatic place identification. Cyc allows to probabilistically rank the list of candidate places in consideration of the commonsense likelihood of that place being actually visited on the basis of the user profile, the time of the day, what happened before, and so forth. The system has been evaluated using real data collected from a mobile computing application.

DOI: 10.4018/978-1-60960-818-7.ch4.13

1. INTRODUCTION

The recent diffusion of GPS-equipped handheld devices and smart phones is opening new scenarios in the development of context-aware services.

Other than consolidated location-based services (Hu et al., 2006), one interesting research direction is to build applications continuously logging the GPS signal to identify the places that

matter to the user and to create a diary of user whereabouts (Ashbrook & Starner, 2003; Bicocchi et al., 2008, Hightower et al., 2005, Miluzzo et al., 2008).

These applications run on handheld devices and typically perform two separate steps:

1. The application runs segmentation and clustering algorithms on the log of GPS traces. This allows indentifying where the user spends most of his/her time in terms of geographic coordinates (longitude and latitude).
2. The application translates geographic coordinates into more descriptive address and business being there (this step typically make use of geocoding and yellow-pages services). In fact, a simple list of coordinates is only partially informative and the need of translating from positions to places (i.e., adding semantic meaningful tags to the discovered coordinates) has been widely recognized (Hightower, 2003). A diary containing information like "*the user was at home*" rather than "*the user was at coordinates (10.873, 44.630)*" would be naturally much more informative and easy to use in applications.

The content of such a whereabouts diary can provide relevant context information to a number of other services. For example a match making application could use the information in the diary to match two persons as compatible given the fact that they frequent almost the same places.

The key challenge to make the diary concrete and usable is precision. Precision is the most important factor to evaluate the diary. If entries in the diary are wrong applications built on top of it risk of becoming useless.

Unfortunately, there are several factors affecting and undermining the precision of diary applications:

1. *GPS signal errors:* urban canyons, multipath fading, atmospheric conditions and sheer number of visible satellites create errors in GPS reading making it difficult to precisely retrieve the user location.
2. *Map and reverse-geocoding errors:* the maps used to reverse-geocode (i.e., from coordinates to address) the user location can contain inaccuracies. A number of maps, for example, only store the first and the last street numbers and evenly spread the numbers in between along the street segment. This of course can create errors when the address corresponding to some physical coordinates is retrieved.
3. *Averaging errors:* a number of algorithms to extract visited place from the log of GPS readings involve detecting drop outs and clusters in the GPS readings. The coordinates belonging to a "hot-spot" are then averaged together to identify the place where the user spent some time. This averaging can create errors in the user actual location.

Even though several techniques like tracking, filtering and map-constrain (Fox et al., 2003) notably reduce the errors, these problems are unavoidable and a some-meters error in localization is always present. In cities with packed buildings and businesses one next to the other, even small errors in coordinates can create large "semantic" errors in the places being visited (e.g., a pub next to a toy-shop).

One possibility to solve such ambiguities (at least from a probabilistic point of view) comes from commonsense reasoning (Lenat & Guha, 1990; Liu & Singh, 2004). Commonsense reasoning refers to a class of artificial intelligence projects that attempt to assemble a comprehensive ontology and database of everyday common sense knowledge, with the goal of enabling applications to perform human-like reasoning.

Using commonsense reasoning is it possible to infer that a small boy does not go to the university,