Research Issues in Mobile Querying*

Martin Breunig[1] Christian S. Jensen[2]

Michael Klein, Andre Zeitz, Georgia Koloniari, Johannes Grünbauer, Pedro José Marrón, Christoforos Panieyiotoa, Susanne Boll, Simonas Saltenis, Kai-Uwe Sattler, Manfred Hauswirth, Wolfgang Lehner, Ouri Wolfson

[1] University of Osnabrück, Research Centre for Geoinformatics and Remote Sensing, Eichendorffweg 30, D-49377 Vechta, Germany
mbreunig@uni-osnabrueck.de

[2]Department of Computer Science, Aalborg University, Fredrik Bajers Vej 7E, DK-9220 Aalborg Øst, Denmark
csj@cs.aau.dk

Abstract. This document reports on key aspects of the discussions conducted within the working group. In particular, the document aims to offer a structured and somewhat digested summary of the group's discussions. The document first offers concepts that enable characterization of "mobile queries" as well as the types of systems that enable such queries. It explores the notion of context in mobile queries. The document ends with a few observations, mainly regarding challenges.

1 Characterizing the Subject Area

As the participants have quite different backgrounds, a first step was to reach some degree of common understanding of what qualifies as a "mobile query." In any setting where mobile querying occurs, there is *data* (which are being queried), a *query*, and an *originator* (who issues the query). The term "mobile" may apply to each of these:

1. The data being queried may be mobile, meaning that the data reside on a mobile device. This setting brings about issues to do with the limited capabilities of

* This document summarizes the results of the working group discussions on "Mobile Queries" held during the Seminar on Mobile Information Management in October 2004 at Dagstuhl, Germany.

Dagstuhl Seminar Proceedings 04441
Mobile Information Management
http://drops.dagstuhl.de/opus/volltexte/2005/165

mobile devices. In particular, small footprint database management systems have been developed for use on mobile devices such as PDAs. Another issue that occurs when a query concerns data on multiple mobile devices is that of efficiently identifying the relevant data. This is an issue addressed in (spatial) Peer-to-Peer database research. More specific issues relate here to, e.g., indexing and distribution of indexes. Next, the data being queried may be mobile in the sense that it concerns mobile objects. In particular, the data may be spatial or geographic and may represent the changing locations of mobile objects. In this case, indexing with the purpose of supporting efficient querying without jeopardizing update performance is an example research topic.

2. It is also possible for the mobility to be part of the query. This is the case when location is part of the query. A continuous query that returns the *k* nearest points of interest to a moving object is an example of this.

3. Finally, it is possible for the originator of the query to be mobile. This occurs when it is a mobile user who issues the query.

Next, it is instructive to introduce concepts that enable a characterization of the systems within which the mobile querying takes place and that enable the mobile querying. Inspired by concepts from distributed information systems, the group identified three characterizations of mobile systems.

1. Distribution. A system may be characterized by how and the degree to which its *data*, *processing*, and *control* are distributed. For example, database management systems with no centralized control, and thus a high degree of local autonomy, are often termed federated database management systems.

2. Heterogeneity. The degree heterogeneity of the technologies employed by a mobile system is an important determinant of system complexity. One important aspect of heterogeneity relates to the database schema. Specifically, there may be one, global schema, there may be multiple schemas for the same data, or the data may be unstructured.

3. Scale. It was argued that the sheer scale of a system may introduce entirely new issues and research challenges, for which reason a characterization of the mobile system by its scale is important. It should be noted that scale and heterogeneity are typically correlated.

2 Context in Mobile Querying

The working group took as its outset a usage scenario where a mobile user receives some kind of service. To deliver this service, the underlying system issues mobile queries. This scenario differs substantially from a desktop computing/office scenario. Specifically, the mobile scenario is often characterized by less convenient humanmachine interaction because the system lacks a convenient keyboard, has only a small screen, if any, and is not the primary focus of the user (moving about safely in traffic may be the primary focus). Next, the general setting of the user changes much more frequently, and is much more varied, than in the desktop computing/office scenario.

This new setting offers new challenges and opportunities. It is important to be able to deliver the "right" service, and thus query result, to the right user at the right time. What is right may be inferred from past user interaction and knowledge of the user's current setting and objectives, which may be inferred from sensors, e.g., location and speed sensors.

It was also observed that not only the conventional one-time queries are of relevance, but that continuous or active queries might well be important for the assumed scenario. A continuous query is one that is registered and that remains active until it is de-registered. For example, a continuous query may retrieve and display all points of interest of a certain type, e.g., gas station, restaurants, hotels, that are within a certain distance from the user.

Based on this line of thinking, the group decided to explore the notion of *context* in mobile querying. At a high, functional level of abstraction, context is *data that adds value to a service being delivered to a user*. Figure 1, explained next, illustrates how the group decided to relate context to mobile querying.

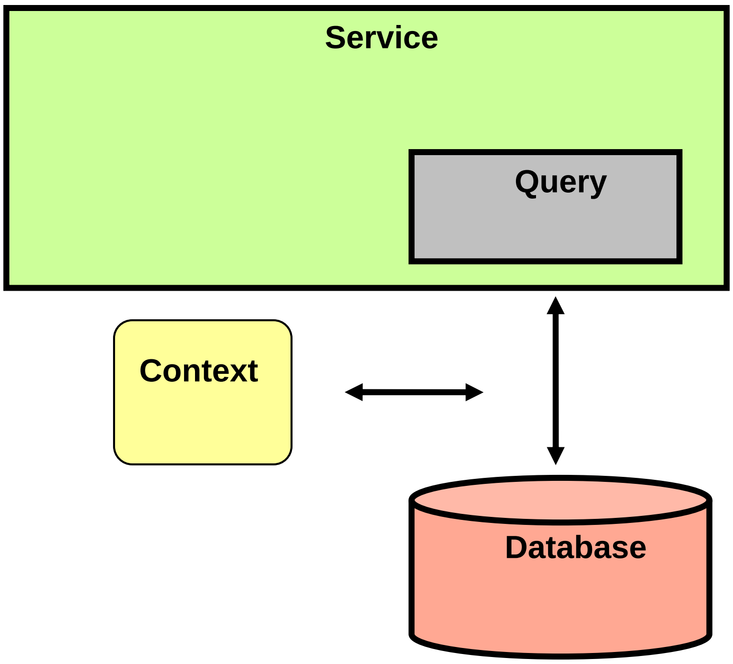

Figure 1: The Role of Context in Mobile Querying

The use of a mobile service involves mobile queries. In particular, a service issues queries against a database. Queries from the service are "intercepted" on their way from the service to the database. Upon its interception, a query is modified to take into account context. Subsequently, the modified query is issued to the database, and the result is returned to the service.

Consider as an example a mobile service user who is interested in the "nearest" gas stations with competitive gas prices. (Mobile services that identify such gas stations actually do exist!). We assume that the user's current position is supplied to the service and that the underlying database contains the relevant information about gas stations. The service may retrieve, say, the up to ten nearest currently open gas stations with attractive prices that are within a 10 mile radius. In this example, the context may consist of:

- Information about the user's past purchases of gas.
- Information about the user's past driving. For example, a GPS receiver in the car may have logged the user's position every second when the car has been moving since the user purchased the car, some two years ago. Based on this, the on-board system maintains a list of frequently used routes along with usage metadata.
- Information about the variability of the gas prices of the different gas stations.
- Information about the fuel level in the car.

With this context available, the service can be optimized. First, if the fuel level is not very low, there is some flexibility in when to get gas. The service then uses the past-driving information to determine, or predict, where the user is currently going and along which route. It is now possible for the service to search for really cheap gas stations along the route that are within range according to the fuel level, that involve only a small or no detour for the user, that are familiar to the user, and that will neither have changed their prices nor will be closed by the time the user will be able to arrive. The resulting gas stations are likely to be the ones that are the most interesting to the user.

One immediate consequence of the group's conceptualization of context is that there may be a need for functionality that explains the results of queries: The same query issued by a service at different times may yield different results. This functionality may bear resemblance to explanatory functionality known from expert systems.

The group explored the nature of context at a more concrete level of abstraction, which led to the following observations.

It was argued that the concept of "user profile" is closely related to, or perhaps part of, context. A user profile expresses what an individual user is interested in, while context more generally captures the "current situation." It was also argued that a user profile is relatively static, while (other) aspects of a current situation are typically much more dynamic. For example, a user preference can be that the user prefers a certain chain of gas stations, while the much more volatile current location of the user is part of the current situation. It was also pointed out that the concept of "preferences," which is employed by a variety of standard software and by web sites, is related to context.

It was agreed that, as a general consideration, context is complex, and there was agreement that the specific content of context varies from service to service, and possibly quite a bit so. However, it was also found that the notion of location in time (e.g., current location) appears particularly relevant across mobile services. As "sensors" such as GPS enable automatic acquisition of location, the use of location as context is then highly relevant. It was also found that the history of interaction between the use and the service, or information reflecting this interaction, is frequently an important part of context. For example, when feeding content to a tourist, it can be helpful if information about previous content fed to the user is taken into account.

Specifics on how to organize context and on general techniques for how to modify a query based on available context were solicited, and the group also discussed the operations one would like to apply to context. Techniques for realizing "standard" operations, such as inserting into and modifying context, were discussed. This included techniques that would automatically discover and acquire relevant context. Here, it was noted that context is open-ended by nature. Next, it was pointed out that mechanisms for recommendation of parts of context, perhaps those best characterized as preferences, would be of interest. Parallels were drawn to recommender systems.

The group also explored two approaches to the management of context that were considered as extremes. First, the "*SAP approach*" (this naming refers to the framework nature and very large schema of the SAP system), involves the development of a *general framework* for context that can then be customized to fit the needs of any service. This schema-based approach entails the design of a very large database schema and very careful data modeling. With this approach, there may be possibilities for plugging in additional schema elements as needed by a new service. Second, the "*Google approach*" (this naming was chosen because Google is a prominent search engine) can be characterized as follows: No schema is designed in advance, and the approach is ad hoc (by means of keyword searches and ranking functions are used, the search engine identifies context), open, and is simple in important senses.

3 Challenges and Suggestions

On the basis of its discussions, the working group concluded that many challenges exist in relation to mobile querying in general and in relation to the use of context in mobile querying, in particular.

When adopting a more concrete setting, issues arise with respect to determining what constitutes the data, what constitutes the context, and what constitutes the queries. Challenges are brought about the by inherent, continuous change, including mobility, in data and context. And challenges arise due to potentially very large numbers of mobile users and objects.

Different members of the working group contributed the more specific *challenges*, listed next, to mobile querying.

- Contending with very high volumes of queries and updates (including the handling of peaks).

- Identifying basic functionality, then composing a database management system of basic units of functionality, covering retrieval, update, streaming, etc.

- Identification of basic operators for aggregation, data modeling, etc.

- Building a kind of middleware that hides mobility from the user (distributed platform on which you can develop your system, like ebay).

Next, the working group encouraged fellow researchers to explore the following directions in their work on mobile querying. These may be good starting points:

- *Ranking.* Giving the user a list of possible answers (e.g., does the user really want the nearest points of interest, or are other criteria, not mentioned explicitly by the user, also important?).
- *Learning from data retrieval systems*. Like Google, e.g., searching for similar "things" (points of interest), applying data retrieval techniques to geospatial data.
- *Learning from how humans do it*. For example, users ask their environment ("what is the best restaurant in the area?") Humans obtain recommendations from social networks.

Finally, in its discussions, the working group found it helpful to distinguish among different types of application areas for mobile querying. The following types of application areas illustrate some of the diversity among applications.

- *Everyday applications*. (Involves ordinary citizens conducting their daily business. Examples include the gas station scenario, inter modal transportation services, and inter-vehicle communication.)
- *Public service applications*. (Examples include positioning of vehicles, individuals, or mobile phones, e.g., in connection with criminal investigations).
- *Disaster protection*. (Examples include landslide, fire, flood, environmental, and bridge monitoring.)