Spatial Filters for Mobile Information Retrieval

David M Mountain
giCentre, Dept Information Science,
Northampton Square, City University,
London, EC1V 0HB, UK
+44 (0) 20 7040 4044
dmm@soi.city.ac.uk

ABSTRACT
This paper introduces the concept of *spatial filters* as an approach to increasing the relevance of the information retrieved by users of mobile information systems. This approach applies post-query filters to remove results that are deemed not relevant given some aspect of a mobile individual's spatial behaviour. The aim is to reduce the volume of results returned to users of mobile information systems, to avoid the need for manual filtering. These filters have been implemented in a location-based service (LBS) using georeferenced document collections from small numbers of content providers. Feedback from end-users in operational settings suggests that they found information that had been filtered in this way to be more relevant than unfiltered information. An opportunity for future research is to consider how these filters can be applied to the task of combining rankings based upon thematic and spatial criteria, for documents retrieved from unstructured or semi-structured collections, such as those handled by geographic information retrieval systems.

Categories and Subject Descriptors
H.3.3 [**Information Storage And Retrieval**]: Information Search and Retrieval – *Retrieval models*

General Terms
Algorithms, Design, Experimentation, Human Factors.

Keywords
Geographic information retrieval; mobile information retrieval; mobile computing; location-aware; location-based services; spatial filters.

1. INTRODUCTION

1.1 Background
Recent advances in mobile telecommunications technology have seen the functionality of mobile devices evolve from predominantly voice exchange to the capacity to exchange and process digital information. Mobile devices have increased in sophistication in terms of processing power, memory, storage capacity, screen resolution and battery performance. This environment has fostered the development of mobile web browsers. Many content providers format their content for mobile web browsers, however, use of the mobile web remains low. Given that our information needs do not cease when we leave our desks, it is curious that despite the widespread use of Internet search engines on desktop machines, use of the same tools from mobile devices has failed to take off in the same way.

Whilst the constraints of the interface - and the real or perceived cost of mobile Internet access - may contribute in part to this slow take-up, an alternative explanation for this limited usage has been a lack of consideration for the distinct information needs of mobile users when designing mobile search systems. Research by Oestrem [1] suggests that usage of mobile devices is characterized by larger numbers of short sessions per day, compared with the fewer, longer sessions typical of desktop usage. In addition, mobile users often use their devices whilst moving through, and interacting with, dynamic environments, suggesting they are less able to devote as their attention to formulating queries and manually filtering results. This suggests the need for automated filtering algorithms, to improve the precision of information retrieved.

1.2 Research Context
There are two fields of research particularly relevant to this paper. The first is the emerging field of location-based services (LBS) which have been defined as search tools that tailor retrieved information based upon the location at which a query was made [2, 3], or more informally as getting "answers to where the questions are" [4]. Thus far, the majority of LBS have developed tailored applications for specific regions using customised content, often commercial points of interest databases for "where's my nearest" applications [5], or multimedia information that has been geocoded [6, 7]. Despite our familiarity with desktop searching of Internet resources, few applications have attempted to use this un- and semi-structured content for LBS, since the majority of these documents contain no explicit spatial referencing. Recently, the field of geographic information retrieval (GIR) has demonstrated algorithms that can retrieve and rank documents from an unstructured collections on the basis of theme and geographic criteria [8, 9, 10,11]. This promise of a vast collection of spatially referenced information offers a great opportunity for LBS, and could result in much richer and more dynamic content than until now has been the case.

2. APPROACH: SPATIAL FILTERS

2.1 Defining spatial filters
The idea of applying *spatial filters* for mobile information retrieval followed a period of user requirements analysis as part of the WebPark project [7] where visitors to a National Park were accompanied on their visit by a member of the team and the questions that they raised recorded as they moved within the park. 60% of these questions were found to have some spatial component which was often stated implicitly in their question, for

GIR'07, November 9, 2007, Lisbon, Portugal.
ACM 978-1-59593-828-2/07/0011

example: “Is ... found around here?”; “Is this…?” and “Is that …?” (whilst pointing); “Will we pass …?”.

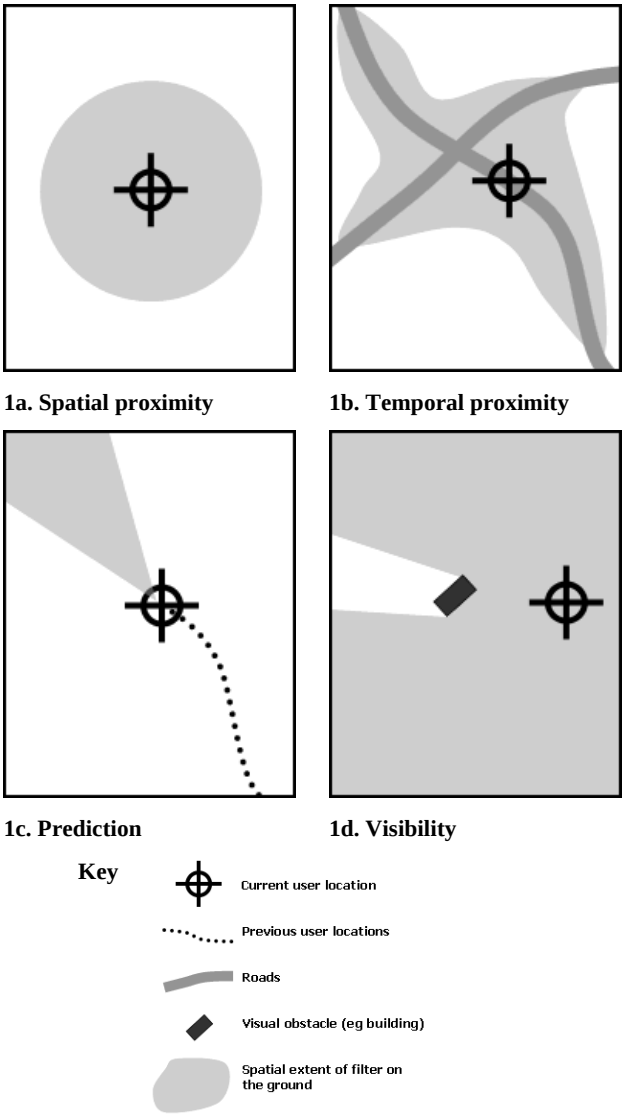

Figure 1: Spatial filters for mobile search

A driving aim of the WebPark mobile information system was to allow users to define alternative spatial components to their query quickly and easily. Based upon the user requirements study, four alternative spatial filters were proposed, as shown and described in fig 1 and table 1. *Spatial proximity* is perhaps the most familiar filter, where following a query given some thematic criteria (e.g. free text searching, or a search by category) a post-query filter removes all results where the *spatial footprint* [8] associated with a document is beyond a specified Euclidean distance of an individual’s current location. From the perspective of geographic information retrieval, this filter expresses the spatial relation “near”, with the spatial component of the query being the user’s current location. An alternative to distance in space is distance in time (*temporal proximity*), where results are filtered on the basis of the time taken to travel from the user’s current location to the spatial footprint associated with each document. Whereas a spatial proximity filter can retrieve relevant results within (e.g.) 1km, a temporal proximity filter can retrieve relevant results within (e.g) 10 minutes travel time. Travel time is dependent upon dynamic factors such as an individual’s mode of transport and time of day. Analysis of large volumes of GPS track logs has proved a promising approach to defining the region of space accessible from a given location [12]. This filter defines an alternative understanding of the spatial relation “near”, expressed as distance in time rather than distance in metres.

Table 1: The spatial relations, criteria and example query scenarios associated with different spatial filters.

Filter	Spatial relation	Criteria	Example query scenario
Spatial proximity	Near	Euclidean distance	Breaking news stories within 1000m of present location
Temporal proximity	Near	Travel time	Toilets within 10 mins of present location
Prediction	Near	Likely future locations	Restaurants I may pass in the next hour
Visibility	Visible	Viewshed	Mountain peaks I can see from here.

A third filter, *prediction*, aims to retrieve information that is relevant to a user’s likely future location. Prediction surfaces can be calculated using the recent spatial behaviour exhibited by a mobile individual which take account of speed, directional trends, and sinuosity [13]. The prediction filter can also be seen as an alternative form of the spatial relation “near”, but defined as closeness to likely future locations, rather than known present location. The fourth proposed filter, *visibility*, takes account of the tendency for the queries from mobile individuals to be inspired by their surrounding environment, commonly expressed informally in the question “What’s that?”.

2.2 Implementing post-query spatial filters

Of the four spatial filters described, two were implemented as part of the WebPark [7] mobile information system: spatial proximity and prediction. The thematic component to the queries in the WebPark system was defined by a category search, where users selected particular categories of information that interested them (fig 2). To define the spatial component of the query, users could use buttons to select between three options: “search around my position” (spatially proximity), “ahead of me” (prediction) and “search in all park” (unfiltered results).

The initial thematic query identified information as either relevant or irrelevant, depending upon the category to which it belonged. The spatial filter was then applied to this set of results that were deemed relevant given this thematic criteria, and results were ranked according to user-defined spatial criteria, either distance from current location (using the spatial proximity filter) or likely future locations (using the prediction filter).

3. EVALUATION

The final, summative round of user testing took place at the end of the WebPark project within the Swiss National Park in June

2004 [14]. Over 100 end-users were asked to evaluate the system in an operational setting. Participants were visitors to the park recruited at the visitor information centre. Each was given a short briefing about how to use the system in the morning, asked to use the system during the course of that day, and complete an evaluation questionnaire when returning the device in the evening. Evaluating the system with real users in an operational setting in this way was designed to assess the impact of the system on the end user, and avoid the criticism associated with systems that focus evaluation on algorithmic performance in laboratory or simulated settings [15].

2a. Query interface, showing options for spatial filters

2b. Ranked result set.

Figure 2: WebPark search interface, demonstrating queries formulated using thematic and spatial criteria.

Feedback related to the performance of the geographic filters suggests that they hold great potential as an approach to handling the spatial component of queries for mobile individuals. Filtered information was considered much more relevant than unfiltered information. Whilst only 54% of visitors considered unfiltered information to be either relevant or very relevant, roughly 90% considered information retrieved using one of the two implemented filters to be relevant. The "around me" filter performed best, with two-thirds of participants regarding the information retrieved to be "extremely relevant". Using the "search ahead" filter, roughly half of participants considered the information to be "extremely relevant". User's were also asked how beneficial they considered two (unimplemented) proposed filters to be. Responses suggest that the visibility filter would provide the greatest benefit to mobile users, with 44% considering it "extremely beneficial" or "beneficial". A travel time (temporal proximity) filter is also considered to be of use, with a similar number (45%) considering this "extremely beneficial" or "beneficial", although fewer (16%) chose the top category. Interestingly one third of respondents did not answer this question, suggesting that the use of these geographic filters may still be difficult to imagine, and hence hard to express an opinion.

4. CONCLUSIONS

Spatial filters implemented in an LBS have been shown to increase the relevance of information retrieved by mobile users. At present these have been used to rank documents from bespoke georeferenced collections, imported from a single content provider. Future research will use these to combine rankings based upon thematic and spatial criteria [16], for documents retrieved from unstructured or semi-structured collections.

5. REFERENCES

[1] Ostrem, J., *Palm OS user interface guidelines.* 2002.

[2] Brimicombe, A. and Y. Li, *Mobile Space-Time Envelopes for Location-Based Services.* Transactions in Gis, 2006. **10**(1): p. 5-23.

[3] Jiang, B. and X. Yao, *Location-based services and GIS in perspective.* Computers Environment and Urban Systems, 2006. **30**(6): p. 712-725.

[4] Haller, R., *(pers. comm.) Discussion on the purpose of location-based services*, D. Mountain, Zernez.

[5] upmystreet, *www.upmystreet.com (visited 20 Aug* 2007).

[6] AmbieSense, www.ambiesense.com. *(visited 20 Aug* 2007).

[7] WebPark, *http://www.webparkservices.info. (visited 20 Aug* 2007).

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

[9] Silva, M.J., Martins, B., Chaves, M., Afonso, A.P. and Cardoso, N., *Adding geographic scopes to web resources.* Computers Environment and Urban Systems, 2006. **30**(4): p. 378-399.,

[10] Li, Z., Wang, C., Xie, X., Wang X. and Ma, W.-Y., *Indexing implicit locations for geographical information retrieval.* in *Workshop on Geographic Information Retrieval, SIGIR.* 2006. Seattle, USA.

[11] Arampatzis, A., van Kreveld, M., Reinbacher, I., Jones, C.B., Vaid, S., Clough, P., Joho, H. and Sanderson, M., *Web-based delineation of imprecise regions.* Computers Environment and Urban Systems, 2006. **30**(4): p. 436-459.

[12] Mountain, D.M., J.L.Y. Tsui, and J.F. Raper. *Modelling accessibility via transportation networks based upon previous experience: a Geographically Weighted Regression approach*. in *Geocomputation 2007*. in press. National University of Ireland, Maynooth, Ireland: National University of Ireland.

[13] Mountain, D.M., *Exploring mobile trajectories: An investigation of individual spatial behaviour and geographic filters for information retrieval*, in *Information Science*. 2005, City University: London.

[14] Krug, K. and W. Abderhalden, *Report on testing and validation*. 2004, WebPark: Zurich. p. 22.

[15] Scholtz, J., *Metrics for evaluating human information interaction systems.* Interacting with Computers, 2006. **18**(4): p. 507-527.

[16] Beard, K. and V. Sharma, *Multidimensional ranking for data in digital spatial libraries.* International Journal on Digital Libraries, 1997. **1**(2): p. 153-160.