Information Access for Context-Aware Appliances

Gareth J. F. Jones
Department of Computer Science
University of Exeter
Exeter EX4 4PT, U.K.
G.J.F.Jones@exeter.ac.uk

Peter J. Brown
Department of Computer Science
University of Exeter
Exeter EX4 4PT, U.K.
P.J.Brown@exeter.ac.uk

ABSTRACT

The emergence of networked context-aware mobile computing appliances potentially offers opportunities for remote access to huge online information resources. Information access in context-aware information appliances can utilize existing techniques developed for effective information retrieval and information filtering; however, practical physical and operational features of these devices and the availability of context information itself suggest that the document selection process should make use of this contextual data.

1 INTRODUCTION

There is currently increasing interest in the development of *mobile computing* applications. Until recently the role of mobile devices was generally confined to the provision of a range of applications for users while they were away from their normal office environment. For example, *PDAs (Personal Digital Assistants)* typically provide a diary, address book and basic word processing. However, the establishment of the *Wireless Application Protocol (WAP)* [4] means that PDAs and other devices can easily be connected to the Internet. WAP is currently appearing on mobile phones, but looks set to be added to any number of information appliances in the near future. The mobile delivery of online material has many potential benefits in enabling users to have immediate timely access to information to assist them in their current situation.

An important feature of many networked mobile information appliances[1] in general, is that they have some sense of their *context*, e.g. their current location, the current time, their current operator, their direction of travel, or the ambient temperature. Context-awareness can potentially combine a large number of context parameters, but in this discussion we generally restrict ourselves to the simpler case of a single context variable. Utilizing context information in the provision of online information in these environments gives rise to the concept of *context-aware information access (CAIA)*.

In this paper we explore the relationship between established *information retrieval (IR)* and *information filtering (IF)* techniques and CAIA, highlighting the novel features of this retrieval environment, and describe our current work which is aimed at beginning to address these issues. We suggest that the characteristics of *context-aware appliances (CAAs)* mean that effective CAIA should look beyond standard IR or IF techniques, and explore methods by which the features associated with being *context-aware* can be incorporated into retrieval models.

2 BACKGROUND TO THE DEVELOPMENT OF CONTEXT-AWARE INFORMATION ACCESS

In this section we review the background to CAIA by outlining pertinent existing work, give definitions for IR and IF as used in this discussion, and outline relevant details of the physical characteristics and operational environments of CAAs.

2.1 Existing Research in CAIA

Existing research in CAIA has been restricted to small scale pilot studies, e.g. [2]. In these applications matching between user interests and available documents has been restricted to simple Boolean matching of fields. These methods have sufficed for small homogeneous data sets; however, we believe that the realisation of CAIA in large scale heterogeneous environments, such as those potentially available in WAP applications, will require more sophisticated retrieval strategies.

2.2 Information Retrieval and Information Filtering

Before exploring CAIA we first outline the key features of IR and IF as used in this discussion. IR is generally viewed as the interactive retrieval, in response to a user search request, of potentially relevant documents from largely static document collections. The output of an IR system is usually a ranked list of documents with brief summaries indicating the contents of each document. By contrast, IF is based on the use of profiles (long-term requests) indicating long standing topics of user interest to select for delivery to the user from among a stream of new documents. Users are thus presented with individual documents which the IF system "believes" may be of interest. A detailed comparison of IR and IF is contained in [1].

[1] In this discussion an information appliance is taken to be any independent computing device. This may be a general purpose PDA or a specialised device such as a digital camera.

To appear in the Proceedings of ACM SIGIR 2000: 23rd Annual International ACM SIGIR Conference, Athens, July 2000.

2.3 Context-Aware Appliances (CAAs)

A key feature of CAAs is that their context will change over time. This change in context can exhibit various features, for example it may be frequent or infrequent, rapid or slow, smooth or erratic, and predictable or unpredictable. In addition, the portability of CAAs has significant implications for interaction in the information access process. For example, the display of WAP enabled mobile phones is a *microbrowser* enabling users to view around 6 short lines of text and simple graphics. This small browser means that, by comparison with standard desktop computer based IR systems, it is difficult for users to browse among a long list of retrieved documents and potentially frustrating to examine more than a few documents for relevant information. The quality of this display is likely to increase rapidly over the next few years, including the introduction of colour, much improved definition, and full-motion video; however it will inevitably always be very small.

2.4 Issues for CAIA

The physical limitations of many CAAs suggest that CAIA applications should aim for high precision in order to reduce the need for browsing. Fortunately, for interactive IR tasks in mobile CAAs high precision will probably often be the main requirement. The user is likely to be seeking information relevant to their current context, e.g. suitable directions to a destination, and satisfying this immediate information need is probably more important than providing the user with all available relevant information. While the user may often, in principle, be happy to accept any resulting reduction in overall recall that this strategy may lead to, experience with existing search engines suggests that often the user's search request will not be sufficiently detailed to yield such high precision. Further, the small physical size of CAAs is unlikely to encourage lengthy text entry, although they typically include various features such as dictionary-based word completion or speech recognition to assist data entry.

IF may appear better suited to CAIA than IR since detailed profiles are often available enabling high precision to be achieved. However, in addition to the loss of interactively, we may encounter another practical problem. IF applications may often require high recall e,g. share price information; in order to achieve this precision is likely to suffer. While operating in an IF mode a CAA is likely to be of secondary interest to the user while they are engaged in some other activity. Experience suggests that they are likely to become frustrated very rapidly if their attention is repeatedly drawn to a CAA often providing non-relevant information [5] in order to achieve high recall.

Our current research is exploring the augmentation of standard IR and IF in CAAs to take their *current context* into account in the process of information delivery. The following section introduces some methods we are exploring and describes a simple example illustrating one approach that we are currently investigating.

3 CONTEXT-AWARE INFORMATION ACCESS

The current context can be used in various ways in the information access process. A simple method of using context is a Boolean filter on potentially relevant documents. For example, documents may be filtered so that only those whose time context matches a Boolean constraint are passed to a subsequent query-document matching stage. Alternatively, a more sophisticated strategy may be adopted to take context into account as part of the calculation of the query-document matching score.

Context-Aware Matching Functions In standard IR and IF the query-document matching score is usually calculated based on the terms contained in the query and the document, often with output document rank determined in part by term weighting functions. In this case all documents are treated as potentially equally likely to be relevant with the final ranking determined only by the contents. Thus regardless of the query's (and usually the user's) context or the context of individual documents, the same ranked list will be returned in response to a query.

Assuming that the contexts associated with the query and documents are correlated with relevance, this situation can be extended to take context into account in the matching process. Consider the example shown in Figure 1. The user is currently at location l, documents are associated with locations, perhaps describing geographical features or relevant historical events. If the user issues a query to this document archive, it would be reasonable to take into account the value of l and the context of each document in the calculation of the matching score.

This concept may be combined into a retrieval framework as follows. Assume that each term i in each document j has a *combined weight* $cw(i,j)$ which combines standard concepts of collection frequency, term frequency and document length compensation [6]. The standard matching score $ms(j)$ between a search query and a document j would be given by,

$$ms(j) = \sum_i cw(i,j)$$

In order to take the context of the query and document into account we can introduce a function $f(c,j)$ representing the difference in context c between them. This function can be incorporated into the calculation of $ms(j)$ as follows,

$$ms(j) = f(c,j) \sum_i cw(i,j)$$

In a simple case difference in context could be the distance between them as shown in Figure 1 for examples $d1$, $d2$ and $d3$.

The most appropriate form of $f(c,j)$ would depend on the application. Some possibilities are as follows:

$$\begin{array}{l} \text{if } (d <= maxD) \\ \quad f(c,j) = 1.0 \\ \text{else} \\ \quad f(c,j) = 0.0 \end{array}$$

where $maxD$ is the maximum distance from the context of the query that relevant documents are believed to be located. Alternatively a decaying function might be used, e.g.

$$f(c,j) = -k_1 d + k_2$$

or

$$f(c,j) = k_1 \exp^{-d} + k_2$$

where k_1 and k_2 are empirically determined constants.

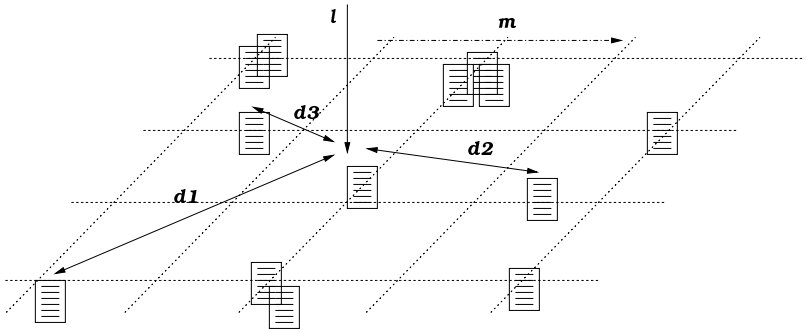

Figure 1: Example of location and motion based context-aware information retrieval.

By incorporating context into the matching score using a function of this type precision may be improved without precluding the retrieval of documents with high query-document matching scores for which there is a mismatch in context. In order to explore the effects of context users could be offered the opportunity to enter a given context manually to simulate a retrieval result.

An important practical aspect of context is that it usually changes gradually, and frequently predictably; hence it is often possible to forecast the next context [3], and potentially to incorporate this information into the query-document matching process. Thus the weighting of a document associated with a likely next context could be increased, whereas that of a context which is becoming more remote might be reduced. This is illustrated in Figure 1 by the arrow m illustrating the direction of motion which can be used to predict the most likely next locations.

A Multiple Context Example Consider a slightly more complex example. A user may wish to enquire about convenient restaurants based on their current location. In order to provide a best answer to this enquiry the system could identify potentially suitable restaurants by comparing menus, pricing, reviews, etc with the preferences of the user. Taking into account context the system would further investigate the distance to each potential restaurant and the possible means of transport with associated timetables, giving an indicating of time to travel which will depend on the current time. It could further take into account details such as the weather, if it is raining it may be better to travel further and for longer via public transport than to walk to a nearby restaurant. Some of the contextual data used in deciding the information to deliver to the user will be continuous while others will be discrete. Clearly realizing systems of this type would enquire online access to a variety of information sources and effective integration of the information obtained.

Context-Aware Documents In order to make CAIA systems work, documents must have a notion of context associated with them, e.g. a location or time with which they are most closely associated. This context might be assigned manually, e.g. in the case of a map, or automatically by a process of shallow information extraction, e.g. extracting proper nouns such as place names. In the latter case physical proximity of location contexts might be found using information extracted from a map. This need for context information requires more "intelligence" in the indexing process than is usually adopted for IR and IF.

Relevance Feedback While mobile CAAs may have very limited scope for browsing, the application of basic relevance feedback methods is fairly straightforward. Users can be offered the opportunity to mark documents as relevant. This information can be used in a standard way to modify term weights and expand the search query, and potentially modify values associated with context functions such as k_1 and k_2.

4 CONCLUSIONS AND CURRENT WORK

The work described in this paper is part of our ongoing work in context-aware information management. Our current work is concentrated on further exploration of the CAIA paradigm, investigating the retrieval requirements of mobile users and the development of suitable query-document matching models. An important aspect of this work which we are keen to develop is the establishment of an evaluation framework and test set to explore CAIA applications.

REFERENCES

[1] N. J. Belkin and W. B. Croft. Information filtering and information retrieval: two sides of the same coin? *Communications of the ACM*, 35(12):29-38, 1992.

[2] P. J. Brown, J. .D. Bovey, and X. Chen. Context-aware Applications: from the laboratory to the marketplace. *IEEE Personal Communications*, 4(5):58-64, 1997.

[3] S. K. Das and S. K. Sen. Adaptive location prediction strategies based on a hierarchical network model in a cellular mobile environment. *The Computer Journal*, 42(6):473-486, 1999.

[4] S. Mann. *Programming Applications with the Wireless Application Protocol.* Wiley, 2000.

[5] B. J. Rhodes. The Wearable Remembrance Agent: A system for augmented memory. *Personal Technologies*, 1:218-224, 1997.

[6] S. E. Robertson, S. Walker, S. Jones, M. M. Hancock-Beaulieu, and M. Gatford. Okapi at TREC-3. In D. K. Harman, editor. *Proceedings of the Third Text REtrieval Conference (TREC-3)*. pages 109-126, 1995.