To Find or To be Found, That is the Question in Mobile Information Retrieval

Dik Lun Lee
Department of Computer Science and Engineering
Hong Kong University of Science and Technology
dlee@cse.ust.hk

ABSTRACT

Web search engines today require users to specify their interests using keywords. By and large, they are successful because the web is so large and so diversify that any reasonable keywords can return some useful documents, whether or not the most relevant or most complete results are returned is a different question. Information filtering aims at matching dynamic information sources against a relatively static user profile to pick up documents that are of interest to the users. It has been studied in the information retrieval community, albeit with less intensity compared to search.

In the mobile scenario, letting interesting information find you is often more important than finding the interesting information when it is needed. The proactive and context-aware nature of information push will manifest itself into all levels and all aspects of mobile information retrieval, turning the mobile device from a place where user commands are issued to a place where information is automatically collected for the user. This paper discusses user profiling methods for monitoring user actions to extract content-based and location-based user profiles, wireless data dissemination architectures that cater for ad hoc location-based information publishing, and applications on mobile advertisements.

Categories and Subject Descriptors

H.3.3 [**Information Systems**]: Information Storage and Retrieval—*information search and retrieval*

General Terms

Algorithms, Design

Keywords

mobile information retrieval, user profiling, location-based, data broadcast, mobile advertisements

1. INTRODUCTION

SIGIR 2008 Workshop on MobIR '08
July 24, Singapore

Mobile information retrieval, in its simplest realization, is the provision of a search interface suitable for the small form factor of mobile devices. Even a simple interpretation like this poses a lot of problems due to the difficulty of input and information display on mobile devices. Since traditional web search requires users to express their interests in keywords, it is very difficult to reduce the number of keystrokes required. Automatic input completion and query suggestion methods reduce but not eliminate the burden on input [3]. Despite the challenges, this simple scenario does not seem to lead to a fundamental paradigm shift from traditional IR.

Web search today is query and user driven. Therefore, it requires the user to express his/her interests in keywords. Keyword matching has been known to have low precision and recall due to the problem of terminology mismatch. Search engines, especially enterprise search engines, utilize long synonym lists to help resolve the problem, but obviously it is a labor-intensive, non-adaptive solution. Although information filtering has been studied in the IR community, the work has been focused on relatively static, keyword-based subscription profiles and relatively homogeneous textual information sources.

In the mobile scenario, it is more desirable to let interesting information find the users than to require the users to query for the information they need. The proactive and context-aware nature of information push manifests itself into all levels and all aspects of mobile information retrieval, including simplified user interface, comprehensive user profiling and just-in-time, just-in-place delivery methods.

This paper discusses two important issues in mobile IR that we are working on: user profiling methods for extracting content and location interests, and wireless data dissemination that caters for ad hoc location-based information publishing (or data dissemination for the small, or *DDFTS*). We will use mobile advertisement as a business case for *DDFTS*.

2. USER PROFILING

In order to proactively deliver highly relevant information to the user's mobile device, the system must be able to understand the user, and a good way to do this is to monitor the user's actions.

An important aspect of MIR is location because it reflects indirectly what the user might be interested in according to his/her current location.[1] One way to infer the user's interest is to assume that the user is interested in information

[1]Time is also an important aspect that reflects the user interest, but it won't be discussed in this paper since the same issue is also applicable to web search.

around his/her current location. While it is true most of the time, it is not difficult to find counter-examples either. Consider the following types of queries:

Local: John finished his meeting early. He wants to find a cinema nearly and the movies that are being shown.

Non-location-based: He wants to find the reviews for the movies (movie reviews don't have to be submitted or published in a place near him).

Non-local, location-based: Not being able to find any good movie, he wants to pick up some information on a family trip to Singapore that he has been planning for in the last few days. He wants to see if Four Seasons Hotel, his favorite hotel for vacation, is available in Singapore.

Location-based and community-based: Since he has not visited Singapore before, he wants to find out what his family can do in Singapore. Since he has absolutely no ideas about Singapore, he needs to rely on information from others who visited Singapore before.

It would be very useful if the system knows what John has been doing in the past few days. Granted that he might have done a lot of things, planning for a vacation to Singapore must be one of the activities and as such some clues must have been left in his profile. If the system is able to show "vacation to Singapore" as one of his activities, John only needs to click the button to continue his planning activity (e.g., the travel agency's homepage would be opened automatically).

Tracking a user's current location or the locations that he has visited is useful for answering local queries and for collecting a user's physical activities (e.g., that John stops at Starbucks at 8am every weekday). This approach requires localization using GPS or other means [6] and the capturing of location semantics [2]. Location tracking is not suitable for non-local queries, which pose different levels of difficulty for the reasoning engine.

A user's web searching or web surfing activities, on the other hand, reflect more the user's interests that last for days. For example, tasks like researching materials for writing this paper are performed over weeks, resulting in many search queries issued and many pages downloaded. By analysing the clickthroughs, the system can generate a group of concepts that the user is interested in.

Figure 1(a) depicts the components involving in capturing user clickthroughs and the derivation of user interests [7]. Figs. 1(b) and (c) are two example queries together with the location and content concepts extracted. In essence, concepts are extracted from web contents (in the case of search engine, web snippets) using typical data mining techniques. They indicate the conceptual space around the user's immediate interests (beach and Southeast Asia). The concepts are further classified into different types. A geography ontology can be used to extract location concepts while a name ontology can be used to extract people names, etc.[2] We can observe that the extracted concepts are quite informative about the user's queries. When the user further clicks on the specific pages, his/her interests would be further revealed.

[2]Due to the focus of this paper, all non-location concepts are classified as content concepts in the examples.

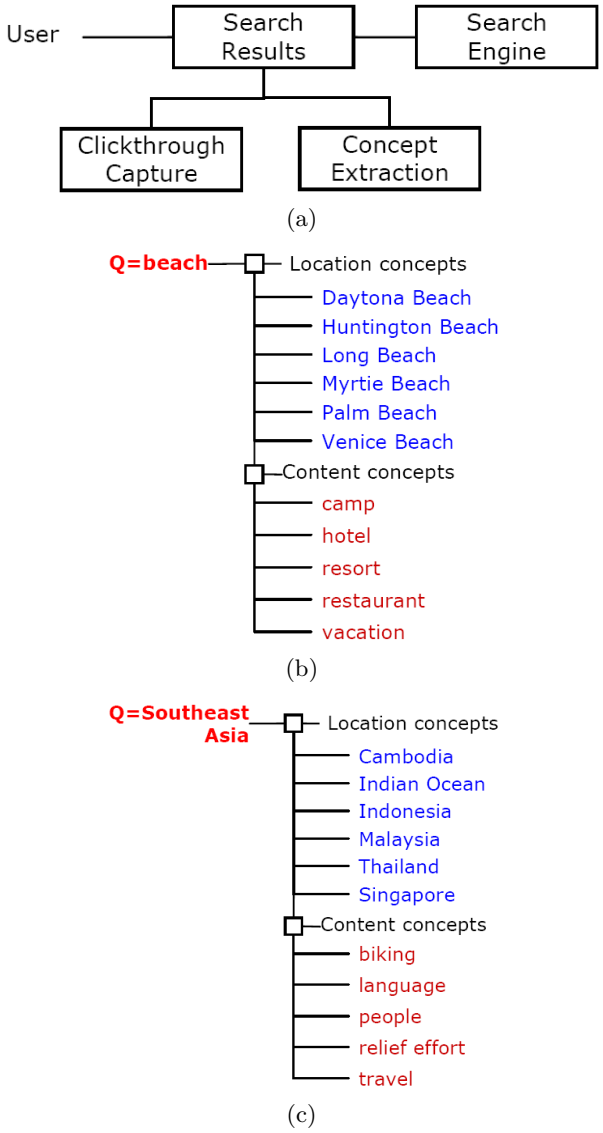

Figure 1: (a) Capturing a user's conceptual interests through search monitoring, and examples of location and content concepts extracted for (b) query "beach," and (c) query "Southeast Asia."

Clickthrough is a kind of implicit feedback, meaning that it does not require conscientious user actions and hence is user friendly. The drawback is that the system must infer from extremely noisy data what the user is really interested in at the current moment. For example, are there any relationship between beach and Southeast Asia? The relationship is not obvious in the examples but then a sequence of clicks may related them together, and hence revealing more specific interests about the user (e.g., finding a beach in Singapore).

3. PUSH-BASED INFORMATION DISSEMINATION

3.1 The Google Audio Experience

Most people know Google as a search engine company. However, increasingly Google has become a global online advertising company. Google AdWords and Google AdSense are well-known, novel advertising models. What is less well known is Google Audio that targets at radio stations (and voice channels such as phone directory services). Radio stations are interesting in that they are local and appeal to audience of different interests and demographics. Contrast a country music program on a Nashville (USA) radio station and a politics program in Washington D.C. (USA). The different appeals to different audiences are very clear. Therefore, Google Audio is an excellent example for location-based, broadcast-based advertising system. Figure 2 shows the major players of the system.

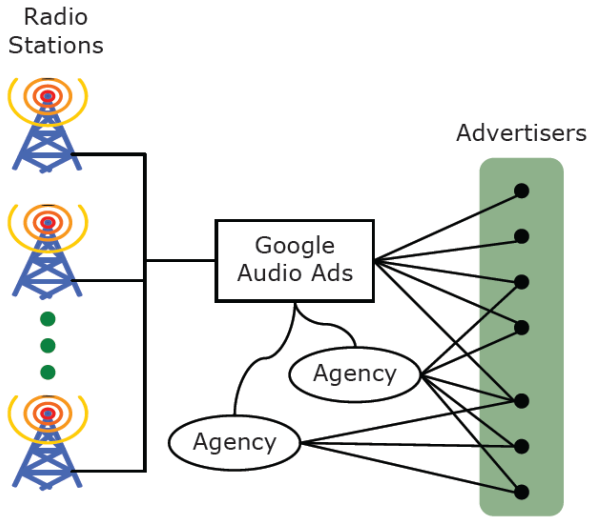

Figure 2: Google Audio.

Traditionally, advertisers who want to advertise on radios have to deal with individual radio stations directly or via agencies. It incurs a larger delay between the time a campaign is planned and the actual launch of the ads. It is even harder and causing more pain if an advertiser wants to coordinate his/her campaign across multiple stations. Google Audio serves as a bridge between advertisers and radio stations. Each radio station in the ad network has a local server managing the slots that the station can sell (the inventory). Advertisers create the voice ads and specify criteria on demographic and program content. Google Audio will perform an automatic match between the requests and inventory and reserve the inventory for the advertisers.

From a research point of view, Google Audio is a simplified case of location-based data broadcast because the radio stations don't move and their broadcast program and audience demographics are rather stable. The location granularity of a radio station is large (e.g., a city). The challenge to deliver fine-grain location-based, personalized ads to the customers remains to be met.[3]

3.2 Data Dissemination for the Small

Traditionally, advertisers who can afford the price and overhead to advertise on radios or TV are fairly large (it won't be your grandma's candy store around the corner). Google Audio makes it easy and potentially less expensive for small advertisers to advertise on these traditional media. However, as mentioned earlier, radio is not considered truly local-based because of its large geographic coverage. All, the number of radio stations, and hence the inventory size, are relatively small, confining audio ads to relatively large advertisers. In the rest of this subsection, we will discuss data dissemination where the publication channels are abundant, mobile, adaptive and inexpensive.

3.2.1 *Bluetooth-Based Data Broadcast*

Bluetooth is the the most common communication method on mobile phones; more phones are equipped with Bluetooth than WiFi. Compared to WiFi, Bluetooth is more suitable for automatic service discovery and the configuration and operational efforts are small. Because of the relatively short range (around 30 meters in practice) of Bluetooth, it is particularly suitable for location-based advertising and messaging.

The idea of Bluetooth-based data broadcast is very similar to that of Google Audio, except that radio stations are replaced with Bluetooth front-ends and advertisers can schedule their ads to be sent out from any Bluetooth front-ends. Figure 3 shows the basic architecture.

Advertisers connect to the Ad Manager to upload their ads, which could be text, audio and video messages. The advertisers schedule their ads according to the available inventory (ad slots in terms of location and time) and define the scheduling rules (location and time constraints for displaying the ads). Ad Server receives ads and scheduling rules and tries to schedule the ads to the proper Ad Front-end. Ad Server could profile the users by recording their Bluetooth names and addresses, accepted and rejected ads, and the time and place of their actions.

There are a couple of interesting implementation issues. Timeouts in client discovery has been found to be a problem in Bluetooth devices [5]. Specifically, to maximize utilization and reach, the system must be able to discover all Bluetooth clients in the range and broadcast all relevant advertisements to them before they walk out of range. The use of multiple Bluetooth front-ends to speed up client discovery and the selection and scheduling of the most relevant ads to be disseminated are interesting questions.

3.2.2 *Heterogeneous Multi-Channel Environment*

[3]Eric Schmidt, CEO of Google, has been using GPS-enable car radios and GPS-enable mobile phones as examples for the viability of delivering highly targeted audio ads to the customers. See, for example, the link http://blogs.zdnet.com/micro-markets/?p=131

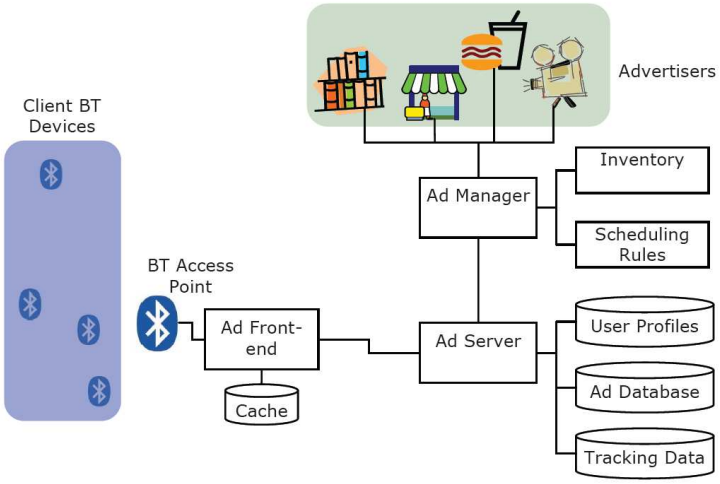

Figure 3: Bluetooth-based data broadcast system.

Heterogeneous Multi-Channel Environment (HMCE) refers to a system with multiple data broadcast channels and an index channel for clients to discover and select the desired data items from the channels [1]. The backend of HMCE is the same as that of a Bluetooth-based system. The main differences are that the client may be within the range of many front-ends and that HMCE is client-driven in that the client decides what to receive from the broadcast channels by specifying their own filtering criteria on the data channels.

Since a client could receive data from different front-ends, it would be very inefficient for the client to monitor the data sent out from each front-end. Instead, HMCE provides an index channel to index all channels within a wide geographic area and the mobile client will start with searching the index channel. If a match is found, the index will tell the client the exact data channel(s) for downloading the complete data. There are many different configurations under HMCE and the system must consider scheduling of both the index and data on the respective channels.

4. CONCLUSIONS

Mobile phones nowadays are powerful enough for performing a lot of data processing and storing a lot of data. They can store all the data that we can imagine in a mini SD card and track every place that we have visited in our lifetime. The question would then be what to do with this power. The points made in this paper are that data input on mobile devices will remain to be difficult for years to come, that data pushing will be the dominant way to bring interesting information to the user, and that effective user profiling is a must to make all these happen.

In this paper, we discuss methods for capturing users' content and location interests, the need for data dissemination for the small or *DDFTS*, and architectures for building such a system. Mobile advertisement is used as an example to show the viability of *DDFTS*. There are many challenging issues: the separation of noise from signal, the classification of content and location interests into a semantic representation, and the use of the proper representation for different contexts and user tasks. Privacy and security are not discussed in this paper at all. It is assumed that all user profiles are stored on the mobile devices, which users typically are more comfortable with compared to capture user profiles on the server [4].

5. ACKNOWLEDGMENTS

This work was supported by grants from the Research Grant Council, Hong Kong SAR, China (Grant Nos. 615806 and 615707).

6. REFERENCES

[1] A. Y. Ho and D. L. Lee. Data indexing for heterogeneous multiple broadcast channel. In *Mobile Data Management*, pages 274–283, 2004.

[2] H. Hu and D. L. Lee. Semantic location modeling for location navigation in mobile environment. In *Mobile Data Management*, pages 52–61, 2004.

[3] M. Kamvar and S. Baluja. Query suggestions for mobile search: understanding usage patterns. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1013–1016, New York, NY, USA, 2008. ACM.

[4] D. Kern, M. Harding, O. Storz, N. Davis, and A. Schmidt. Shaping how advertisers see me: user views on implicit and explicit profile capture. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems*, pages 3363–3368, New York, NY, USA, 2008. ACM.

[5] T. Kindberg and T. Jones. "merolyn the phone": A study of bluetooth naming practices (nominated for the best paper award). In *Ubicomp*, pages 318–335, 2007.

[6] D. L. Lee and Q. Chen. A model-based wifi localization method. In *InfoScale '07: Proceedings of the 2nd international conference on Scalable information systems*, pages 1–7, 2007.

[7] K. W.-T. Leung, W. Ng, and D. L. Lee. Personalized concept-based clustering of search engine queries. *IEEE Transactions on Knowledge and Data Engineering*, Apr. 2008.