Mobile Applications of Focused Link Discovery

Darren Wei-Che Huang
Shlomo Geva, Kelly Y. Itakura
Queensland University of Technology
Brisbane, Australia
{s.geva,kelly.itakura,w2.huang}@qut.edu.au

Andrew Trotman
Department of Computer Science
University of Otago
Dunedin, New Zealand
andrew@cs.otago.ac.nz

Abstract *Interaction with a mobile device remains difficult due to inherent physical limitations. This difficulty is particularly evident for search, which requires typing. We extend the One-Search-Only search paradigm by adding a novel link-browsing scheme built on top of automatic link discovery. A prototype was built for iPhone and tested with 12 subjects. A post-use interview survey suggests that the extended paradigm improves the mobile information seeking experience.*

Keywords: Focused Link Discovery, Wikipedia, Mobile Information Seeking, User Studies Involving Documents.

1. Introduction

1.1. Background and Motivation

The Internet can now be accessed through mobile devices. Services such as email, web search, blogs, and instant messaging are all available on smart phones. Marketing campaigns by communications companies have persuaded us that we should access information from wherever we are, and we now can, and do. Because of this the nature of mobile information seeking has changed. However, search and navigation on a small mobile device remains awkward [11], especially as many service providers have not yet adapted to the small screen of a mobile device.

Recently there has been a substantial improvement in mobile provision by popular providers such as *Facebook*, *Skype*, *Twitter*, *Wikipedia*, *Windows Live Messenger*, and the search engines. They each have mobile versions with enhanced interaction. It is not, for example, uncommon to see users using mobile maps to look for shops, petrol stations, and restaurants, or social software to check for contacts. But, problematically, the screen on a smart phone is small.

Focused link discovery research has examined the automatic identification of new anchor texts in an existing document for linking to a *best entry point* (BEP) from where a user should start reading in a target document. This is an example of how existing technology can be adapted to facilitate better mobile search on a small screen[1].

This paper discusses a novel mobile knowledge seeking application that uses link discovery to alleviate the small screen search problem. To navigate via links, a four-page link browsing scheme is proposed. The first of these four pages is the document view displayed exactly as usual, but with additional hypertext links optionally added. The second displays the article's anchor texts. The third displays details of target documents. The final page is a history view. New links are identified using previously published link discovery algorithms already shown to be good on the Wikipedia (our test collection) [8].

We tested our prototype application with 12 participants and present our findings: that the four page scheme with link discovery enhances the information seeking experience.

1.2. Related Work

Arvola et al. [2] proposed a framework for evaluating passage and XML retrieval search engines. They used a mobile browsing scenario to simulate how the relevant content should be extracted and localized (to reduce the effort of reading on a small screen).

They suggested interaction followed two phases: the fetch phase, in which the user enters a query into a relevance ranking search engine; and the browse phase which presents a set of non-overlapping passages in a retrieved document.

In their interface, the effort of finding relevant content within a document is decreased because they insert links to the next relevant passage at each relevant passage. In this way the mobile user is directed to the relevant parts of a document rather than the beginning of the document. This framework removed the first vertical scroll to the first relevant passage, but also preserves the original document order and structure. The user can also determine if there is any further relevant information in the document without scrolling to the bottom.

Their motivation was the design of an evaluation framework (with metrics) for focused retrieval. They do not investigate how focused retrieval helps to alleviate mobile search difficulties.

The concept of fetching relevant documents and browsing relevant passages proposed herein is similar to that of Arvola et al. but we improve mobile information seeking using focused link discovery. This,

Proceedings of the 16th Australasian Document Computing Symposium, Canberra, Australia, 2 December 2011.

coupled with our four-page browsing scheme, makes it possible for the user to browse related content without performing additional searches – without typing new queries on a small mobile keyboard or screen.

Mobile search is still an unsolved problem. In 2010 Steve Jobs stated that "On mobile, search hasn't happened. People aren't searching on their phones." [3].

While Jobs raises questions about the utility of mobile search, Microsoft, Google, and Yahoo! all see the value of mobile search services across multiple devices. They optimized their search interfaces to include localised search, addresses, and maps. Users now expect information to be a single click away from an interactive mobile interface. Making it so reduces the effort involved in search and increases access to information. The shift to mobile has altered the way users think about, and use, search.

We use the Wikipedia as a sample document collection in our experiments. Several existing mobile Wikipedia applications already exist. For example, there is a mobile version of the English Wikipedia supporting different mobile devices (iPhone, Android, Palm WebOS, Opera Mini and NetFront). There are also third-party applications including: *Wapedia*, *Sevenal*, *iWiki*, *Boopsie*, *WikiPock* and *HAWHAW*.

Our application is novel because it emphasizes two aspects. The first is the automatic identification of new links that are inserted from displayed documents to Wikipedia documents. The second is a new way to interact with those links.

Mobile applications that augment the information on the device have already been proposed. Phone-derived GPS coordinates are an essential part of many augmented reality applications [6,7]. Our application augments text documents by inserting new links.

2. Focused Link Discovery

Instead of requiring the user to enter queries on a tiny keyboard, we re-cast search as browse. To do so it is necessary to augment documents by adding new links that point to related content.

The branch of IR known as Automatic Link Discovery studies the automatic identification and maintenance of hypertext links in a collection of documents. Successful algorithms eliminate the human effort required to build a highly accurate link-graph, and to keep the graph up-to-date[1].

Going beyond ordinary document level link discovery, Focused Automatic Link Discovery identifies where, within the target document, the user should start reading. This location is known as the best entry point, or BEP. BEP links are more versatile than whole-document links (especially for a mobile device) because the user does not need to scroll (or skim read) through irrelevant content in order to find relevant content. It is reasonable to expect focused automatic link discovery to be beneficial in reducing user effort when finding information while on the move.

Central to this research is the evaluation of these automatically identified links. The INEX link-the-wiki track offers a platform for researchers to collaboratively assess link discovery systems [9]. In addition to having examined document-to-document links (which began in 2007), the evaluation examined anchor-to-BEP links in 2008 and 2009.

3. Conceptual Framework

In order to demonstrate the applicability of focused automatic link discovery to a mobile environment we designed and built a simulation of an iPhone application. The conceptual flow of the application is illustrated in Figure 1. When an article is loaded into the system, the text is parsed by the text scanner. The anchor indicator then identifies potential anchor terms and phrases, which are then filtered by the anchor filter. The filtering process, for example, resolves potentially overlapping anchors such as "apple", "computer", and "apple computer". This is a form of contextual disambiguation.

The link detector identifies candidate target documents for the anchors and produces link candidates. The candidates are then selected and filtered by the filtering component which ranks candidate links, categorizes relevant XML elements and documents, and prepares a summary of each target document.

4. Mobile Information Seeking

4.1. Mobile User Interface

This section describes our user interface for small mobile screens. Our application starts with the *One Search Only* screen [10]. This screen, also known as the *Home* screen is shown in Figure 2 (left).

There are three ways to start the information seeking process. The first is to enter a query into the search box (depicted, left). The second is to enter a URL. The third is to select a bookmark. The *Transfer* button (top right) is used to apply focused automatic link discovery to the currently displayed document.

Figure 2 (middle) shows a Wikipedia infobox while the table of contents with links is displayed in Figure 2 (right) for quick access to document subsections. New links in the infobox have been identified by our focused automatic link discovery system. Since the anchors on a small screen are themselves small, they are difficult to click. One of our aims is to reduce this difficulty; we discuss our solution in the next section.

4.2. Mobile Link Browsing Scheme

Our goal is to simplify the operation of information seeking on mobile devices. Link discovery identifies related content while a simple navigation process allows the user to browse through that content. Our interface is simple and intuitive and alleviates the diffi-

[1] Link discovery algorithms are beyond the score of this paper, the interested reader is referred to the INEX track overviews [8].

culties associated with the more common query-typing, result list reading, and document skim-reading iterative process. An initial search is necessary, but it is expected to be the only search consequently it is a *One Search Only* system [10].

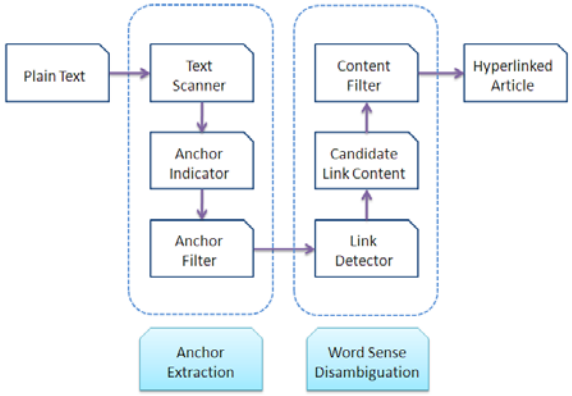

Figure 1. Conceptual flow of mobile link discovery

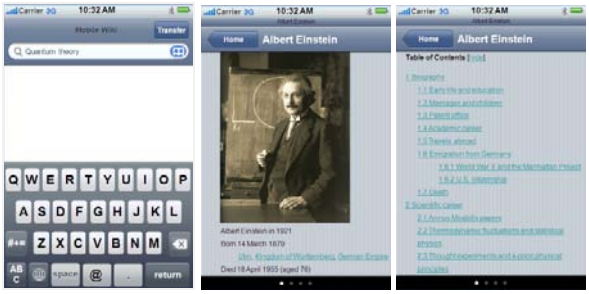

Figure 2. Home and Article content screens

There are two possible responses to the one search as shown in Figure 3: either a relevance ranked list of documents (case 1); or if there is only one result, then that document is displayed (case 2). In case 1, the user can scroll up and down the results list and choose a document to view. That document is displayed on screen with its original links shown. By clicking the *transfer* button new links are added. For case 2, the article and links are immediately displayed on screen.

Figure 4 shows the conceptual flow of the link-browsing scheme. There are 4 screens. Articles can be displayed in 3 ways: (1) the article content; (2) the article's anchor terms; (3) and the article's links. And there is also a history screen (4). The user switches between them using the iPhone *swipe* gesture.

In the first screen, the article content screen, the article is not shown from the beginning, but rather the screen it automatically scrolled to the most relevant content (the BEP[2]). This happens regardless of whether the article is displayed as a consequence of a search or a link-click. Because of this the user can immediately read the relevant portion of the articles without ever scrolling.

The second screen displays the set of anchor-texts. These can be listed in either document order or alphabetical order.

[2] Algorithms for computing the BEP are beyond the scope of this contribution. The interested reader is referred to the proceedings of INEX between 2007 and 2010

The third screen displays the titles and a summary of the target documents that are linked-to by the current article. Each summary, prepared by the Content Filter, can be hidden or displayed by clicking an icon on the right hand side of each link.

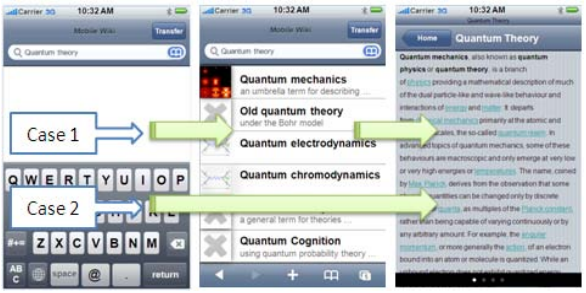

Figure 3. Search screen for mobile information seeking

The fourth screen shows the browsing history, and is a list of all the articles the user has viewed. This history list can be used to return to any document the user has viewed in the current session.

On the article screen and the anchor terms screen, the anchor to target relationship is many-to-many. That is, many anchors may target the same document, but also each anchor may target multiple documents. To help navigate this we introduce the *anchor-targets* page.

The bottom left of Figure 4 shows an example *anchor-targets* page. A list of document titles and summaries is given allowing the user to choose which target they prefer.

Several navigation features are common to many of the screens. There is *Back* button to return to the previous screen. The round icon on the *anchor terms*, *anchor links* and *anchor-targets* page indicates whether the document has previously been viewed or not. At the top of each navigation-page there is a *Home* button that returns to the search page.

5. Evaluation

We implemented the simulation of the iPhone application using Xcode and evaluated it using 12 volunteers (Nielsen & Landauer [12] suggest 5 users sufficient to find most usability problems). Our goal was to test the user experience using the 4-page interface, The links themselves were pre-discovered offline using the algorithm of Huang et al. [8]. The user testing experimental software did not actually discover links in real-time, but rather restricted users to a set of pre-linked pages and one level deep link following. The interface however was not restricted in design and recursive browsing can be implemented in a straight forward manner. None of the volunteers had used this mobile link discovery application previously. Half had an IT degree while the others had a non-IT background. Each was asked to experiment and to become familiar with the tool.

We were specifically seeking qualitative information on how the small screen affects the operation of

the mobile link browsing system (and therefore how to improve it) before undertaking a full implementation.

A questionnaire was prepared and an interview survey was conducted. The questions included prior experience of existing mobile search applications, the experience of using mobile version of Wikipedia, the use of the link browsing scheme, the experience of using the four-layer navigation process, how the link browsing scheme improved mobile information seeking, which layers were used most and most useful, and the pros and cons of these four layers.

Figure 4. Conceptual flow of the link browsing scheme

Table 1. Usability survey results

Questions	Strong	Average	Weak
Current Mobile Search			
How often do you use it?	5	2	5
How are they?		7	5
Mobile Wikipedia			
How often do you use it?	1	4	7
Is it useful for browsing information?	2	8	2
4-Layers Link Browsing Scheme			
Easy to understand and use?	9	2	1
Is it useful for mobile information seeking?	9	2	1
Article Content Layer – positive	10	2	
Anchor Terms Layer – positive	10	2	
Anchor Links Layer – positive	11		1
History Layer – positive	11		1

Table 1 summarises the results. Most of our subjects were aware of specialised mobile search applications but many only used these for getting general information (e.g. bus/movie timetable, weather, maps and news). Most did not directly use mobile Wikipedia except when linked-to by sources such as Google.

There was strong support for the four-layer navigation and so it is reasonable to conclude that it is an effective method of interaction. The *article content* screen and the *anchor terms* screen were viewed positively. Through informal discussion participants indicated that it was easier and quicker to find relevant anchors in the list of anchor terms than in the *article content* screen.

Participants indicated that multiple links per anchor was better than one-to-one linking. This is of particular note because it is not the normal state of affairs for the web. Multiple targets per anchor could be implemented in JavaScript and might be of particular value to information sources such as the Wikipedia – however it is not seen there.

The *anchor links* and *history* screens were the layers most frequently used.

These results suggest that the link browsing scheme is practicable for mobile users, and an improvement on conventional mobile search.

6. Conclusions and Future Work

Using a mobile phone to search the Internet is awkward because the screen is small and typing is difficult. In order to make it easier to find information using these devices we suggested turning search into browse.

If related documents are extensively interlinked then the quantity of queries needed in order to satisfy an information need could be reduced. The user would be able to click rather than forced to search. But documents are not extensively linked and it is highly unlikely that authors will change them to be so.

To address this we suggested the use of automatic link discovery technology. This technology adds links between documents by identifying anchor text in one and a related best entry point in another.

But clicking links on a small screen is also difficult. To address this we introduced a novel multi-screen link browsing scheme.

To test our theories we designed and implemented a simulation for an iPhone application with the two technologies: focused automatic link discovery; and a four-screen link browsing scheme. The four-screens were the article view, the anchor terms; the article's links, and the history screen. Navigation between screens was with the iPhone *swipe* gesture. We note that link discovery algorithms are beyond the scope of this paper and so we use a previously published algorithm previously shown to be effective.

The application was tested with 12 subjects of varying backgrounds. A post use survey was conducted.

The survey results show that the link browsing scheme is easy to understand and easy to use. Most of the subjects believed the link browsing scheme was useful for information seeking in a mobile environment. It is conceptually simple and practical, suggesting that further research into the *One Search Only* paradigm is warranted.

But, the experiment we conducted was with only a small user base (of 12) and consequently can only be considered as a proof of concept. A substantially larger number of participants are necessary before any firm conclusions can be drawn.

Our application functions in one of two possible modes. The first is as a mobile (link-enhanced) version of the Wikipedia. The second is as an application that can add links from any page to the Wikipedia (similar to the Wikify at the University of Waikato).

We believe this kind of service will be invaluable in the future as other Wikipedia like resources become available (such as dictionaries).

In future work we will examine whether or not our interface is generally useful or just for the Wikipedia. We will do this by turning our prototype iPhone application into a fully functional application and rolling-out to a larger user base.

References

[1] Huang, W. C., Xu, Y., Trotman, A, and Geva, S. Mobile Link Browsing: Efficient Interaction with Mobile Contents, In Joint HCSNet-EII Workshop 2008, Australia.

[2] Arvola, P., Kekalainen, J. and Junkkari, M. Expected reading effort in focused retrieval evaluation, Information Retrieval: Focused Retrieval and Result Aggregation, 13(5), 2010, 460-484.

[3] Subramanian, K. Clash of the Titans: The Battle to Become the Mobile Search Leader, June 2010, http://techcrunch .com/2010/06/27/mobile-search-clash/.

[4] eMarketer The Search Wars Are Going Mobile, 2007, http://www.emarketer.com/Article.aspx?R=1005083.

[5] Wikipedia Help: Mobile Access, 2010, http://en.wikipedia. org/wiki/Help:Mobile_access.

[6] Jones, M., Buchanan, G. and Thimbleby, H. Sorting Out Searching on Small Screen Devices, *In Proceedings of the 4th International Symposium on Mobile HCI*, Springer, 2002, 81-94.

[7] Sweeney, S., Crestani, F. Supporting Searching on Small Screen Devices Using Summarization, *Lecture Notes in Computer Science*, Udine, Italy, 2954. Springer, Berlin, 187-201.

[8] Huang, W.C., Trotman, A, and Geva, S. Overview of INEX 2009 Link the Wiki Track, Lecture Notes In Computer Science.

[9] Huang, W. C., Trotman, A, and Geva, S. The Methodology of Manual Assessment in the Evaluation of Link Discovery, In Proceedings of the 14th Australasian Document Computing Symposium, Sydney, Australia, 4 December 2009.

[10] Huang, W. C., Trotman, A., and Geva, S. (2008) Efficient Information Seeking on a Mobile Device, Proceedings of SIGIR 2008 Workshop on Mobile Information Retrieval (MobIR'08)

[11] Dunlop, M., and Brewster, S., (2002). The Challenge of Mobile Devices for Human Computer Interaction. Personal Ubiquitous Comput. 6:4(235-236)

[12] Nielsen, J., and Landauer, T. K. (1993) A mathematical model of the finding of usability problems. In Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems (CHI '93), pp 206-213