Towards a model for an Internet content pre-caching agent for small computing devices

Andreas Komninos

University of Strathclyde
Livingstone Tower, 26 Richmond st.,
Glasgow, G1 1XH, UK
andreas@cis.strath.ac.uk

Mark D. Dunlop

University of Strathclyde
Livingstone Tower, 26 Richmond st.,
Glasgow, G1 1XH, UK
mark.dunlop@cis.strath.ac.uk

Abstract

In the near future Internet access will be commonly available in two forms to most palmtops: wide bandwidth of a low-cost land-based connection while docked, and limited capacity expensive wireless connections while mobile. The coming 3G networks will help the bandwidth but the disparity of connection cost and bandwidth between docked and undocked devices will continue.
This paper proposes a model for automatically pre-loading (or pre-caching) palmtop computers with websites likely to be of use in order to support the user's daily activities. To achieve this, the model extracts information from a user's diary and feeds this into a predictive system.

1 Introduction

Small personal computing devices are available to almost everyone today. These are presently distinguished between those used primarily for communication, such as mobile phones, and those known as Personal Digital Assistants (PDAs). However, the distinctive line between these two categories slowly starts to disappear.
A characteristic, which signifies this blending of devices, is the incorporation of PIM functions into many of today's mobile phones. Most importantly, the typical calendar application, which is central to PDAs, is now available in most modern mobile phones. Another application which is currently available on PDAs and is slowly appearing on mobile phones is the Web Browser. Until recently, surfing the web on a mobile phone has been restricted to WAP purpose-built sites, mainly due to the slow speeds available through the GSM network (~9.2Kbps) and the limited display capabilities of mobile phones. With the advent of wider bandwidth communication standards (GPRS), at least part of the problem is being addressed.
The distinction line between PDA and mobile phone is expected to disappear with the coming of 3G networks and 3G mobile devices, which will provide several communication services currently limited to desktop computers. Perhaps the most important of these is broadband connection to the Internet and its vast amount of resources. There is however some concern regarding the costs of this new service. The 2.5-G GPRS transmission standard has been around for a while, but it is still very expensive to use. The same is expected for 3G networks, once they become widely available, with analysts estimating that users will be reluctant to accept the high costs of 3G-network usage.[i]

A Komninos, M D Dunlop, "Towards a model for an Internet content pre-caching agent for small computing devices", proceedings of HCI International 2003, Crete, June 2003

Draft paper - see proceedings for final version

2 Predicting content needs

It would be possible for a user to pre-fetch all of the data that would be required through their low-cost, land-based (and often broadband) internet connection and feed it into their mobile device. The problem with this scenario would be that users would have to spend a considerable amount of time browsing for and saving websites and then transferring to the mobile device. Furthermore, documents might need extra processing in order to appear properly in the small device.
An intelligent software agent could examine the user's electronic calendar and try to estimate the kind of activities they will perform and, possibly, what kind of internet content they may need, in order to support those. The web content downloaded through this predictive system, would be stored in the user's desktop, processed and then transferred over to the mobile device.
The predictive system can be enhanced over time to learn and remember a users' preference for different categories of information. Such a predictive system should be able to obtain information from the user directly and indirectly. The majority of information should be obtained through indirect means, in order to minimise interference with the user's other activities. However, the system should maintain its ability to directly interact with the user, in order to resolve any possible uncertainties, which will be essential to the initial stages of the system's training process.

3 Agent structure

Words such as "learn", "guess", "remember" have been previously used in our theory regarding the way the predictive system should work. These prompt an insight of the human memory as a starting point to model the pre-fetching system. Cognitive psychology can offer paradigms for a flexible, adaptable and interconnected "knowledge-base" structure.

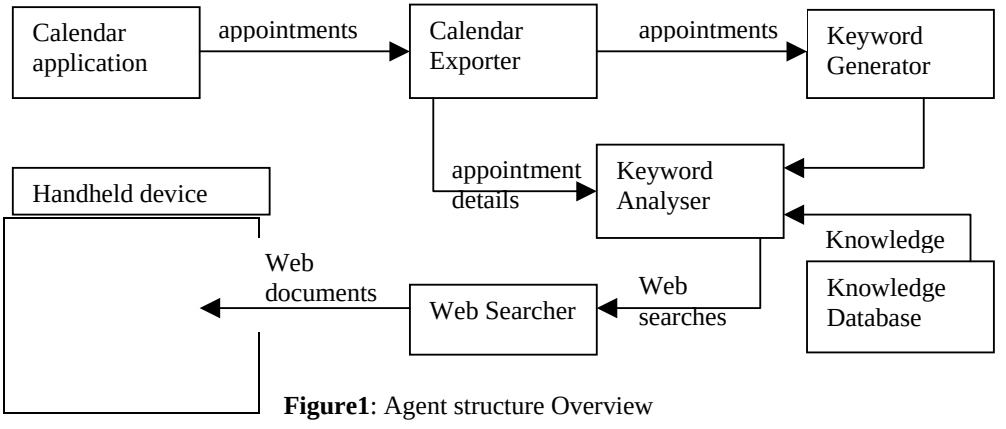

Figure1: Agent structure Overview

In the human brain, various tasks, such as storing information into working memory, are performed with the simultaneous work of various centres. A paradigm of this is the central executive, which uses other agent-type centres, like the rehearsal loop, in order to process and store information in the working memory area. This means that the central executive, which is the "main processor" for the system can offload results from previous tasks to other, smaller capacity agents, so it is free to engage in other difficult processing tasks[ii]. From this paradigm, we adopt a modular form for the agent's structure, such as depicted in figure 1.
The advantages of such a structure are obvious. Instead of relying on one main executive to perform tasks in a serial manner, it is possible to have several "sub-agents" working

collaboratively in order to perform tasks with enhanced efficiency, in order to minimise the total running time of the retrieval process.

4 Prediction management

The agent should have at least some knowledge of the world it operates in, in order to perform effectively and produce meaningful results. Ideally the agent should be able to augment its knowledge level. This can be achieved either by allowing the agent to add further knowledge to what it already possesses, or by allowing the agent to obtain knowledge about the knowledge it already holds. We shall call the latter meta-knowledge and refer to it later on. Also the agent should hold some knowledge about how to extract keywords from the appointment entries. This is an important aspect of the agent, since, without proper keywords, there can be no satisfactory web searches.

We conducted a study in the way users write in their calendars, by obtaining actual calendar entries and analysing their content. From this, it has been determined that users typically fall into three categories (hasty, meticulous and average), according to their style of input. From this study, we can formulate some rules for inferring additional information from hastily written entries (e.g. it has been found that when an entry's subject is the name of a person, it is 90% certain that the user will be meeting that person) and for determining which words and in which context can become keyword candidates. Keywords will be compared to a database of known items to identify their significance. In his work "A calendar with common sense"[iii], Erik T. Mueller has describes a potentially useful database for common-sense knowledge items, which could be useful starting point. Information particular to the user, such as names of friends or colleagues from the user's phone book, where she lives, etc, should also be used as part of the rules.

Once a keyword has been identified as such, the agent should be able to associate it with several other keywords, in order to form search phrases for submission to web search engines. The users in our conducted survey have identified several desirable searches that they would like for particular entry categories (an example is shown later in table 1). Other items/searches can be obtained through statistics held by web search sites (Google, for example, have a useful search suggestion tool to recommend related keywords[iv]).

Another type of rule that we will use, is the actual location, where a particular keyword has been found. For example, if an unknown word is found in the "location" field of an appointment entry, it can be assumed that this keyword is a location and we should conduct location-related searches for that word. Also, words in the title field of the appointment entry are likely to be of more importance than words present in the notes field, and keywords from the title could be combined with keywords in the notes field, in order to derive more specific searches.

In the case of uncertainty (whose percentage threshold will be determined through experimentation) during the attempt to apply the rules in a particular keyword, the agent should ask for clarification from the user. This knowledge will be kept for future reference, as the agent learns about items relative to the user.

The success of the conducted searches will be measured through relevance feedback (section 5.2), and this will in itself form another rule, which will be remembered for later use.

5 Memory Organisation

Through the conducted user survey, it was been noted that appointment items tend to fall within a few category types, the most frequent of which are meetings, tasks and reminders. This evidence is further supported by the fact that most modern electronic calendars provide facilities for categorising appointments. Early work into calendar usage by Kelley & Chapanis[v] and Kincaid

et.al[vi], shows similar results to our findings. It should be therefore possible, instead of providing the agent a list of related search terms for each possible keyword, to cluster keywords into categories and associate search terms with each category. The use of clustering with keywords should mean that the knowledge database is somehow smaller in size, while remaining just as efficient.

These category clusters will consist of a category descriptor (against which the keywords will be compared) and a list of related search terms. In [2], it is explained that when reading words in a text, the human brain does not actually read the words letter by letter. Instead, some of the letters of the word are read and then the brain "guesses" what letters should come between the read letter and the next one. From relevant experiments, it is apparent, the brain appears to be using a ranking mechanism for the knowledge it holds, and uses the most frequently used (higher scoring) items when guessing.

Table 1: Sample Keyword category

Category Name	Category descriptor	Related Terms	Weight
"Location"	Travel, fly, population, city, town, train, bus	Map	0.7
		Hotel	0.8
		Airport	0.6
		Car Hire	0.4
		Museum	0.3

Following this ranking paradigm from the brain system, the agent should be able to make distinctions between search terms that are important to the user and those which are rarely used. This can be determined using relevance feedback techniques, such as Rocchio's algorithm[vii]. Initially, search terms will have a score (weight) associated with them, which will be pre-defined, something that can be achieved with the usage of the TF/IDF (term frequency/inverse document frequency) algorithm through the analysis of user statements of desired searches and keyword appearance in related web documents. As the users continue to make use of the agent, these weights will be adjusted according to the relevance feedback that will be obtained from the user.

Explicit feedback should be obtained rarely and with discretion, so as not to impede the user or interfere in their activities. It is expected that the majority of information will be collected through implicit feedback, which is unobtrusive and also can provide accurate results.

There are several heuristics[viii], which can be used to implicitly measure relevance. Some examples may be whether a user has clicked a document title to see the document, the amount of time they have spent reading the document or whether they have followed any hyperlinks from that document. Each of these heuristics should have its own weight, as it is for example obvious that when one spends some time reading a document obviously indicates much more interest than simply having opened the document.[ix]

By combining this relevance feedback information with observation of the user behaviour, it is possible to gain the meta-knowledge that we talked about previously. Meta-knowledge is important because it can allow the agent to choose which part of its knowledge it will use, in order to provide results that are useful to the user.

6 Summary

In this paper we have described a model for the management and organisation of knowledge for a pre-caching agent, based on paradigms from the human working memory system. The agent will initially be equipped with some general knowledge that is necessary, in order to support its

function. However, some of this knowledge will be “forgotten” and new knowledge will be “learned”, as the agent adapts to the user’s preferences and needs.
The heuristics for the evaluation and estimation of the users needs are based not only on findings of previous research, but also on the study of actual calendar entries and their users. These heuristics will provide a realistic rulebase, which will support the agent’s evolutionary process.
The agent’s structure is such that will allow a flexible approach to the user’s needs, by evaluating and adapting to their style of input and personal preferences. This will be achieved through the constant evaluation and weighting of the appropriate factors, using relevance feedback techniques.
.

7 Future work

Currently, the model has not yet been fully implemented into a piece of executable software, so the overall functionality of the model remains to be examined. In the course of this examination, it will be interesting to observe not only the functionality of the proposed system, but also the behaviour of its users and the evolution of their interactions with this system.
In our study of calendar users, it was indicated by the majority of the users that they would be willing to change their input styles, in order to help the program achieve better results. Will they adhere to this statement? Will there be changes in the way detailed entries are made? Will people begin to insert keywords, or even search phrases in their calendars?
The flexible capacities of such a model could be utilised in other scenarios, where the pre-caching of documents is desirable. For example, this system could be enhanced to allow the inclusion of items such as recent emails or text messages and documents that may be relative to a situation, such as a meeting. We believe that the system could prove a useful aid for the increasing number of people who use their mobile computing devices to support their daily activities.

8 References

[i] CLAIRE WOFFENDEN (2000), Users 'will not pay for high 3G costs', http://www.vnunet.com/News/1112409

[ii] DANIEL REISBERG, Cognition – exploring the science of the mind, 2nd edition, W.W. Norton & co, New York. p. 15-17

[iii] ERIK T. MUELLER (2000), A calendar with common sense, ACM Intelligent User Interfaces, p.p. 198-201

[iv] GOOGLE AdWord suggestion tool, https://adwords.google.com/select/main?cmd=KeywordSandbox

[v] KELLEY, J. F., & CHAPANIS, A. (1982). How Professional Persons Keep their Calendars: Implications for Computerization. *Journal of Occupational Psychology*, 55, pp. 241-256.

[vi] KINCAID, C. M., DUPONT, P. D., & KAYE, A. R. (1985). Electronic Calendars in the Office: An Assessment of User Needs and Current Technology. *ACM Transactions on Office Information Systems*, 3(1), pp. 89-102.

[vii] J.J. Rocchio (1971), Relevance feedback in information retrieval, in the SMART retrieval system, Prentice Hall,

[viii] YOUNG-WOO SEO & BYOUNG-TAK ZHANG (2000), Learning users’ preferences by analyzing web-browsing behaviours, ACM Agents

[ix] YOUNG-WOO SEO & BYOUNG-TAK ZHANG (2000), A reinforcement learning agent for personal information filtering, ACM Agents