The Ubiquitous Office: A Nomadic Search and Access Solution

ユビキタスオフィス：外出先からの文書の検索とアクセスを実現するソリューション

Victor Poznanski* Steffan Corley*

Philip Edmonds* Anthony Hull* Michio Wise* Morgan Willis* Ryoichi Sato* Claire Green*

Abstract

There is an unprecedented amount of information available at our fingertips when we are at our desks. However, typical employees spend less time at their desks than ever before. In this article, we present our vision of the Ubiquitous Office, which offers mobile workers nomadic access to the key information and services that they need. By nomadic, we mean that they can travel light, taking nothing more than their mobile phone with them, and are still able to use network-connected peripherals around them. We have developed a prototype mobile solution including printers and mobile phones that tackles two important technical issues for the Ubiquitous Office: ergonomic search and browsing of documents on small appliances and secure transmission of documents to nearby peripherals.

今日我々は自分のデスクに居ながらにして未曾有の量の情報を手に入れることが可能である。しかしながら典型的なビジネスマンが自分のデスクで過ごす時間は以前にも増して少なくなってきている。この論文では外出中のモバイルワーカに重要な情報やサービスへのアクセスを提供するユビキタスオフィスのビジョンについて紹介する。ここで言う "nomadic" とは，携帯電話以上のものを持ち歩くことなく，軽い装備で出張しながら，身の回りにあるネットワーク接続機器を利用できる，ということを意味する。我々はプリンタと携帯電話を使ったモバイルソリューションのプロトタイプを開発した。このプロトタイプではユビキタスオフィスの 2 つの技術的課題に取り組んでいる。つまり小型情報機器によるドキュメントの人間工学的なサーチとブラウジング，および手近な周辺機器への安全なドキュメント転送である。

Introduction

Having the right information at your fingertips is crucial to the operation of most corporations. In 2001, the average employee spent between 38.7 and 150 hours per year searching [1]. Deploying a good search solution can double productivity and save hundreds of thousands of dollars per annum in medium sized companies [1]. However, the average employee is also spending less time working at a PC behind a desk. IDC predicts that 66% of US workers (105 million) will be mobile by the end of 2006. In Europe, 100 million workers will be mobile by the end of 2007 [2]. These people have similar information needs on the move.

There is a spectrum of solutions to help people with their mobile information needs. At one end of the spectrum are functionally-complex portable computing devices that provide for most of their needs (e.g., laptops and portable printers). However, in this article we will focus on the other end of the spectrum, in which the mobile worker carries physically and computationally lightweight devices that rely on opportunistic local connectivity to other devices to achieve his purpose. The approach where a person travels light and connects wirelessly to the Internet and to appliances available around them is known as ***nomadic computing***.

We have developed a technology vision for the nomadic worker of the future that we call the Ubiquitous Office. We imagine such workers would carry nothing more than a

* Sharp Laboratories of Europe, Ltd.

mobile appliance such as a phone (or a wireless PDA). Through their phones, workers would have access to a range of web services, such as local maps and news. Most importantly, they could securely and effortlessly find their critical email and office files at any time of the day, read them, and relay them to nearby printers, monitors, and projectors that are enabled for Ubiquitous Office functionality. They could go on to use their mobile appliance as a remote control, say, to control a presentation or decide on print options.

This vision is illustrated in **Fig. 1**.

Most mobile operators already provide parts of the solution. For example, Vodafone Spain and Jazztel recently announced a VPN offering for mobile phones (see section 3 for an explanation of VPNs). However, in Europe, these solutions usually assume the use of a portable PC rather than a phone. Mobile phones are also becoming increasingly capable of displaying documents [3]. Even so, the combination of a limited keyboard and QVGA screen still makes searching for files a challenge. Beyond the user interface issues is the major bottleneck of the connection itself. Even though GPRS (a European packet data service) promises data rates of up to 114kbps, actual rates are far lower. On an uncongested network, downloads typically run at 30kbps, around 5 minutes per MB at a cost of around $5. 3G networks improve on this, but will prove more expensive. When the user wants to print or display a document on a fixed appliance, rather than view it on his phone, the ideal option would be for the file to arrive at the appliance via fast and cheap land-lines. The question becomes: how can we achieve this seamlessly and securely?

In the remainder of this article, we will delve into these issues and describe some of the solutions we developed as part of a prototype Ubiquitous Office.

In section 2, we concentrate on the problem of search on low-bandwidth small appliances. In section 3, we concentrate on the problem of secure, seamless access. In section 4, we summarize our work and discuss possibilities for the future.

1. Search

A core problem for both nomadic and desk-based workers is how to find the documents that they need. Enterprise search companies such as Verity, Autonomy and Fast have made it their business to supply large-scale search solutions. For smaller enterprises, various companies provide desktop document management solutions that include search, such as Scansoft's PaperPort and Sharp's Sharpdesk and SharpFind solutions.

The fundamental technology used in search engines [4] has changed little since its inception. However, modern corporate search engines form part of tailored solutions aimed at improving a company's productivity. For example, search technology is incorporated into classification systems that assist with workflow automation, and routers that alert users to specific information as soon as it arrives. Products targeted at smaller companies are more generic: PaperPort and Sharpdesk both provide a simple integrated solution for the management of scanned images and PDF.

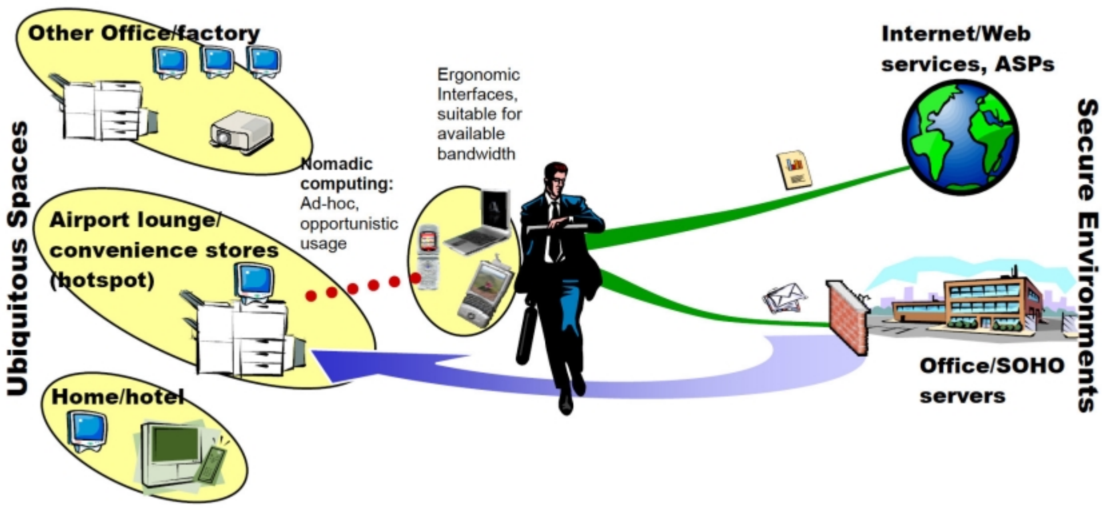

Fig. 1 The ubiquitous office vision.

While these solutions exist for search on the desktop, the problems of finding a document in a nomadic environment are somewhat different. The first difference is the scenarios under which a nomadic user will typically be searching for a document. In particular, users on the move are more likely to use their appliances for well-defined goals, such as finding a particular document, rather than doing background research. They may well be working under time pressure (for example, they are in a hurry to print a document before or during a meeting). If it is too complex or time-consuming to obtain the information they need using a mobile appliance, users will either give up or try to find a desktop computer.

The other major difference is that the limitations of mobile appliances mean that actually finding a document is far more difficult. These limitations include:

- Small screen size: it is difficult to view the contents of large folders or read through long lists of search results.
- Limited input facilities: it is difficult to input a detailed query.
- Inability to accurately and quickly display many common document types (e.g., Microsoft Office documents), making it harder to know when the right documents have been found.

One of the aims of our Ubiquitous Office prototype was to improve the mobile search and browsing process to overcome these limitations. Before we detail our solutions, we briefly consider some related work.

1•1 Related Work

The problems of search and browsing on mobile devices have been discussed in relatively recent workshops [5,6]. Some authors have considered the implications of porting existing browsing and search applications to mobile devices [7,8]. Major portals and search engines have also simplified their user interfaces for browsing and searching the (wireless) web on a mobile device [9,10].

Jones et al. [11] have suggested a mobile search user interface where results lists are improved by using key phrases automatically extracted from documents to represent the documents. They asked users to classify the documents into the correct location in a subject hierarchy, basing their decisions either on the titles or the key phrases. They found that automatically extracted key phrases were as good as titles (but not better) for the categorization task. This result suggests that showing key phrases in addition to titles may be useful in an information retrieval interface.

1•2 Ubiquitous Office Prototype

Our Ubiquitous Office prototype uses three concepts to make it faster and easier for users to find the documents they need. (1) "QuickLinks" makes it far quicker to browse for documents, (2) "Search Assistant" makes search easier by recommending key phrases from the user's documents that may improve the user's search, and (3) query-based summarization allows users to quickly preview search results. We give more detail on each of these technologies below.

1•2•1 QuickLinks

Companies often use elaborate directory structures to store documents. Even if a user knows where a document is, it can take a significant number of clicks to reach that document. Where the document is being accessed across a slow mobile connection, each click can take a few seconds, resulting in up to a minute to obtain a single document. Worse, users often get lost in the directory structure and fail to find the document they require on the first attempt.

Our QuickLinks system provides a solution by observing user's browsing behaviour and automatically providing dynamic bookmarks - shortcuts to directories or files they may wish to revisit in the near future.

The QuickLinks user interface is shown in **Fig. 2** (left).

1•2•2 Search Assistant

Searching is often an interactive process: the user enters a short query and then gradually refines it. On a desktop computer, it is simple to refine (lengthen) the query or browse through quite a large number of results. However, this is not practical on a small appliance with limited input and screen space. Even the process of adding more words can be cumbersome.

The Search Assistant is designed to allow users to enter a short query, and then to make recommendations as to how to improve the query by adding key phrases from documents in the results list. Natural language information extraction techniques are used to extract the phrases from the user's documents (similar extraction techniques are described in [12] and [13]). The user can select one of these phrases to extend the query (and thereby get a shorter results list) without having to input any more text. An example of the Search Assistant in action is shown in **Fig. 2** (right).

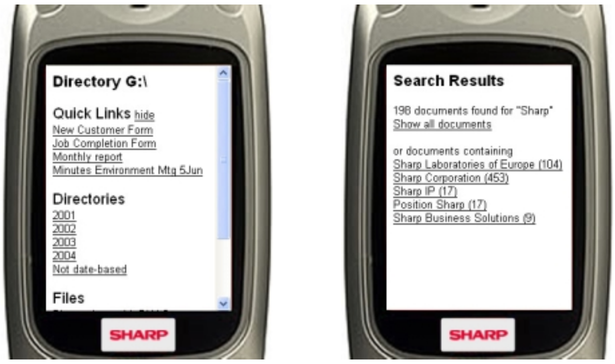

Fig. 2 QuickLinks during a browsing session (left) and displaying phrases for the query “Sharp” in Search Assistant (right).

Fig. 3 When the user selects a document, they are shown a short textual summary to help them identify whether this is the document they intended.

1•2•3 Summarization

The last problem facing users is that, once they have found a document, it is very difficult for them to quickly determine whether it is the right document, and often impractical to display the whole document. Our solution is to offer the user a textual summary of the document, showing just the most relevant sentences from the document that relate to the query [14,15]. This summary is tailored to fit the display size of a mobile device.

The summarization user interface is shown in **Fig. 3**.

The search and browsing functionality described above is hosted on a web server within the company and accessed using a VPN (see **Fig. 4**). The user interface is built on standard web technologies and can be displayed in any web browser. We built versions optimized for display on phones and PDAs, but could not deploy the phone version because mobile VPNs were not yet available in the UK.

2. Secure Access

Most companies are protected by firewalls. In brief, a firewall protects a company's internal network by preventing unauthorized connections to the network from the Internet. Typically, the authorized way for mobile workers to access their company network from their device (mobile phone or laptop) when outside the company is by using a virtual private network (VPN). A VPN serves two purposes - it enforces a security policy (usually that both the connecting device and the user are known, and that only certain network services are accessed) and it encrypts the transmission across the Internet (forming a tunnel between the company network and the device), so that no one listening in on the communication can understand it or interfere with it.

Firewalls pose a problem for nomadic users who would like a peripheral, such as a printer, to connect back to their office network and print one of their files. Of course, one way around this is for the user to load the entire file onto their phone and then transmit it locally to the printer, but the file can be very large, and mobile data rates are typically slow and transmission costs expensive. A better solution would be to allow a public peripheral to access your network temporarily, but how many companies would give arbitrary access rights to a public appliance that could easily be tampered with? Even if companies did allow this, a user would still have to configure the appliance for his company's VPN and identify himself; both of which would be cumbersome processes using the typical keypads and buttons available on printers and monitors.

Early work by Xerox and more recent work by NTT have tackled some of these problems, and we will discuss this before going on to describe our solution.

2•1 Related Work

Xerox's Satchel prototype [16,17] demonstrated a context-aware system for nomadic computing. Users could use a Nokia Communicator to browse (but not search) their documents, stored on their company network. Once they had selected a document, they were offered a variety of "document services". If the communicator could make infrared contact with a Satchel-enabled printer, these services included printing.

In order to print, the Communicator passed a "token" to the printer. This token allowed the printer to contact the user's company network and retrieve the document to be printed.

Satchel used a proprietary gateway to access documents within a corporate network through a firewall. While one of their aims was compliance with security policies, the Satchel inventors themselves say that they "do not know [...] whether a system with the current security features would be widely acceptable by corporate information technology managers" [16]. A proprietary solution introduces the risk of new security holes. One of our primary design aims was to work with existing infrastructure (e.g., VPNs), rather than introduce our own.

Similarly, Satchel required a program to run on the Nokia Communicator, meaning that the system would only work with devices that had the Satchel client pre-installed or could be reprogrammed outside the factory. Our solution is intended to work with any modern mobile phone.

NTT have proposed a system that also allows nomadic access to peripherals [18]. The system is built around a completely new security infrastructure and all the user needs to carry is a smart card. The problem for such approaches is that it is difficult to persuade companies to replace their existing infrastructure in this way.

HP have also done work in this area [19], but have not concentrated on the security issues.

2•2 Access in the Ubiquitous Office Prototype

A key design goal for our Ubiquitous Office prototype was to be able to integrate it with existing security infrastructure, but still allow relatively safe nomadic access. Our access solution is as follows (for the sake of clarity, we assume that the user is printing rather than displaying the document):

- The user connects from their phone to the corporate network using his company's existing VPN. A private web server running within his corporate network serves our search-and-browse user interface using standard low-bandwidth web protocols (i.e., XHTML over HTTPS); any modern phone or PDA with a web browser can display the user interface.
- The user finds the documents he wants to print (using the search and browsing interface described in section 2) and selects them for printing.
- A one-time code is generated - this code serves as the permission to print these documents.
- The documents are strongly encrypted (a good method is AES [20]) and copied, together with the one-time code, to a secondary server situated within the company's demilitarized zone (a semi-public area of the corporate network which normally houses the company's web server and other public servers). This secondary server is used as a holding area.
- An "access identifier" is generated and returned to the mobile phone. This access identifier contains the location (URL) of the holding area, the decryption key, and the one-time code.
- In order to print the document, the user transfers the access identifier from the mobile phone to the printer (see below). The printer then unpacks the access identifier to obtain the URL of the holding area,

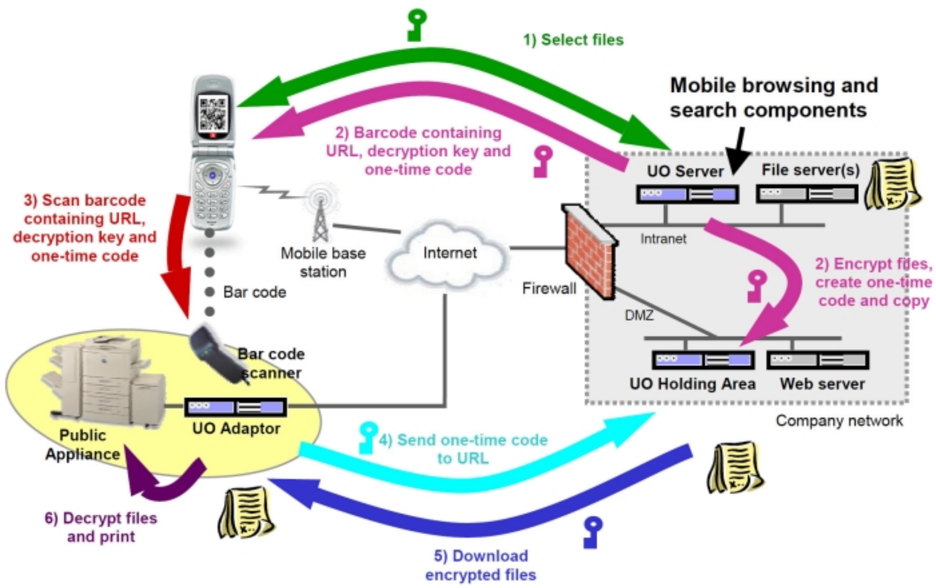

Fig. 4 Illustration of how documents are made available to public appliances in the Ubiquitous Office prototype.

connects to the holding area (using standard secure web protocols) and presents the one-time code as authorization. In return, the holding area returns the encrypted documents, deletes them and removes the one-time code from its list of current codes.

- Finally, the printer uses the decryption key to decrypt the documents and prints them.

This solution is illustrated in **Fig. 4**. It can be easily integrated with an existing corporate network and is also relatively secure. The untrusted public printer has access to only a holding area in the public area of the company's network, and no direct access to the internal network. This holding area only stores documents that are soon to be printed, in an encrypted form; it deletes them immediately after they have been printed. Even if a hacker did gain access to the holding area, the damage (in terms of exposure of company secrets) would be minimal.

We built two versions of our system: the first used infrared transmission to send the access identifier from the mobile phone to the printer. The second version encoded the access identifier as a 2-dimensional bar code (we used QR code [21]) and displayed it on the screen of the mobile phone. The user simply holds the mobile phone in front of a bar code reader attached to the printer. The bar code is scanned by the printer and decoded to recreate the access identifier. This solution is simple for the user and will work with most mobile phones.

We have performed informal evaluations of the system with a variety of users within Sharp and have received very positive feedback. One future direction for us is a larger scale deployment to test the system more thoroughly and gain further insight into user needs.

Conclusion

Many researchers have provided visions of a ubiquitous future, but they are often impractical because they require an instantaneous and radical change to current infrastructure. We have suggested a compromise to help us along the way: build secure systems on top of the existing infrastructure.

We have outlined the major problems for search in a nomadic environment: bandwidth considerations, small displays with limited interaction, and a focus on known documents. Our solution to the bandwidth problem is to develop systems that can work at lower bandwidths by transmitting browse and search results in formatted plain text with summary information and allowing the high-bandwidth nomad-enabled appliances to access the documents over a fast connection and then to process or render them in detail. For small displays, such as those of mobile phones, we have developed three techniques: QuickLinks to speed up access to recently used documents, the Search Assistant to help the user find particular files with minimal text input, and query-based summarization to reduce complex documents to textual representations of their most relevant parts. We have prototyped this system using PDAs and mobile phones. Informal user studies have suggested that the system is useful. In the future, we would like to adapt the technology to work on other appliances and develop special-purpose search techniques.

Acknowledgements

The authors would like to thank their colleagues in Sharp Corporation at the Digital Home Appliance Group for helpful discussions. We would particularly like to thank K. Doi-san for his input to the ideas presented in section 3 and Dr. T. Chiba for his ongoing support for this work.

References

[1] ROI of the Google Search Appliance: Intranet Deployments, Google, 2003. Available: http://www.google.com/appliance/pdf/google_roi_intranet.pdf.

[2] IDC expects a rise in mobile workers in Europe, Telecomworldwire, Oct 2003. Available: http://www.findarticles.com/cf_0/m0ECZ/2003_Oct_14/108842609/p1/article.jhtml.

[3] Sharp Develops Electronic Document Display System for Zoomable Viewing of Business Documents on Mobile Phones, Sharp, December 2003. Available: http://sharp-world.com/corporate/news/031209.html.

[4] C.J. van Rijsbergen, Information Retrieval, London: Butterworths, 1979.

[5] M.K. Leong and G. Loudon (eds.), Workshop on Mobile Personal Information Retrieval, held at SIGIR 2002, Tampere, Finland.

[6] Y. Maarek, A. Soffer and B.W. Chang (eds.), Workshop on Mobile Search, held at WWW 2002, Honolulu, Hawaii.

[7] N. Milic-Frayling and R. Sommerer, "SearchMobil: Search Support Interface for PDAs", Workshop on Mobile Personal Information Retrieval, held at SIGIR 2002, Tampere, Finland.

[8] H. Lee and A.F. Smeaton, "Searching the Fischlar-NEWS Archive on a Mobile Device" , Workshop on Mobile Personal Information Retrieval, held at SIGIR 2002, Tampere, Finland.

[9] http://www.google.com/options/wireless.html.

[10] http://mobile.yahoo.com/.

[11] S. Jones, M. Jones, S. Deo, "Using Keyphrases as Search Result Surrogates on Small Screen Devices", Personal and Ubiquitous Computing, vol. 8, no. 1, pp. 55-68, 2004.

[12] C. Jacquemin and D. Bourigault (2003), "Term Extraction and Automatic Indexing", in R. Mitkov (ed.), Handbook of Computational Linguistics, pp. 599-615, Oxford University Press, Oxford, 2003.

[13] A. Mikheev, M. Moens and C. Grover, "Named Entity recognition without gazetteers", In Proc. of EACL, Bergen, Norway, 1999.

[14] J. Kupiec, J.O. Pedersen and F. Chen, "A trainable document summarizer", Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '95), Seattle, WA, pp. 68-73, 1995.

[15] A. Tombros and M. Sanderson, "Advantages of Query Biased Summaries in Information Retrieval", Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '98), Melbourne, Australia, pp. 2-10, 1998.

[16] M. Lamming, M. Eldridge, M. Flynn, C. Jones and D. Pendlebury, "Satchel: Providing Access to Any Document, Any Time, Anywhere", ACM Transactions on Computer-Human Interaction (TOCHI), vol. 7, no. 3, pp. 322-352, 2000.

[17] M. Flynn, D. Pendlebury, C. Jones, M. Eldridge and M. Lamming, "The Satchel System Architecture: Mobile Access to Documents and Services", Mobile Networks and Applications, vol. 5, no. 4, pp. 243-258, 2000.

[18] E. Niwano, M. Kanbe and S. Yamamoto, "Model-based Ubiquitous Card Network for Nomadic Information Sharing", presented at JavaOne 2002, San Francisco, CA. Slides available: http://servlet.java.sun.com/javaone/sf2002/conf/speakers/14290-bio.en.jsp.

[19] T. Kindberg and J. Barton, "A Web-Based Nomadic Computing System", Computer Networks, vol. 35, no. 4, pp. 443-456, 2001.

[20] Advanced Encryption Standard (AES) (FIPS 197), 2001. Available: http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf.

[21] QR Code Standard (ISO/IEC 18004), 2000. Available: http://www.iso.org/.

(received June 9, 2004)