You've Got Mail! Calendar, Weather and More: Customizable Phone Access to Personal Information

Manuel A. Pérez-Quiñones
Department of Computer Science
Virginia Tech
perez@vt.edu

Jochen Rode
Department of Computer Science
Virginia Tech
jrode@vt.edu

ABSTRACT
We present a design and a prototype of a system that provides access to calendar, email, weather, and news information over a phone using a VoiceXML interface. The system provides quick access to personal information, while enabling but not requiring interactivity. As one major application, we envision the system being used while commuting to work. At home, users define their preferences regarding content and order of presentation using a website. On the road, the personalized audio "feed" plays like a radio news show. Instead of commercials, the user is reminded of today's meetings, deadlines and listens to the email inbox.

Author Keywords
Personal Information Management, phone, voice access, customization

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
Ubiquitous computing has presented a future where the computers, as traditionally known, would disappear into the environment [7]. There have been other variations of this theme, among them: information appliances, where computers will have narrow but well defined functionality; pervasive computing where computation is available in the environment at all times. Many of these ideas were expressed before the boom of the internet and/or the quick adoption of the cell phone. Both of these (internet and cell phone) provide access to information at any time; together they provide access to information *anytime, anywhere*.

Within the realm of Information Technology, Personal Information Management (PIM) is a term used to refer to a collection of information tasks, all related to the use of computer devices to manage a personal calendar, to do list, appointment book, email, and other general day-to-day tasks. The most common computational devices used for PIM are the desktop, laptop, and PDA. Cell phones and other small devices (e.g. the iPod which includes a calendar, address book, and notes) are being added to the mix of devices that provide support for PIM.

With the wide variety of tasks considered part of PIM and the wide selection of devices used today, PIM is anything but ubiquitous. Synchronization of information plagues the use of this information. Compatibility across devices is another issue; not all devices support all data types (e.g. having pictures in the address book). Finally, connectivity of devices also interferes with the use of information. Some devices have connectivity to the network (e.g. via 802.11, Bluetooth or using the phone network) while others are surrogates of the desktop/laptop computer (e.g. the iPod and PDAs without network connection).

In this paper we present a voice-enabled system that provides access to personal information while away from the office or home. The system has two major design goals. First, it allows ubiquitous access without requiring data synchronization. The data is always up to date and available over a phone and over the web. Second, the system can be pre-programmed and customized in order to allow it to be used without requiring interactivity. This may be desirable in situations that require a strong focus on the primary task, such as driving to work.

In the balance of the paper we first illustrate the history and motivation for this work by presenting and discussing a sample usage scenario, then describe the system architecture, and conclude with thoughts about potential future research directions for voice-enabled PIM.

The work described here has its roots in a system that was primarily built to support remote access to a personal calendar and has been in use for almost one year. The current work extends the original design by providing access to email, weather, and news and adding customization functionality.

RELATED WORK

Ubiquitous Computing

Mark Weiser argued that the most profound technologies are those that disappear into the environment [7]. This provocative and stimulating piece of work sparked a lot of ideas and discussions about what ubiquitous computing is and should be. The most recent trend is the work done by Paul Dourish [3] where tangible computing and social computing are both explained in terms of their embodiment principles.

VoiceXML

VoiceXML is a markup language created by a consortium of companies that includes Lucent, Motorola, IBM, and AT&T. The language is XML-based and requires only a text editor to create VoiceXML pages. The language includes tags that are specifically geared towards building voice interfaces and to support telephony applications. The ease of editing a VoiceXML file and the unique features of the language makes it relatively easy to create a voice interface for phone-based applications. Furthermore, the designers of the language have opted to use the World Wide Web infrastructure for delivery of VoiceXML content. VoiceXML pages are served by an HTTP server (e.g., Apache). The voice server communicates back to the web server via calls to server-side scripts (PHP, ASP, Java servlets, CGIs), much like a web browser does. The promise of a voice-enabled web is discussed in some detail in two recent articles [2, 4, 8].

The next logical extension of the Internet is to voice-enable the web so that we can access information from phones and cellular phones that are not equipped with web browsers. This will extend the reach of web-based information to match the reach of the global phone network. This is a particularly attractive solution since cell phones are very popular and in use in countries and rural areas where connectivity to the Internet is difficult to achieve. Furthermore, the mobile phone is already used as a "device proxy" [5]. It has become the preferred way to stay in contact with the main office while mobile workers are on travel. Furthermore, a phone based information access system makes use of today's "most pervasive device" [1]. Many systems have been reported in the literature that provide access to information via the phone [1, 9, 6]. However, most attempt to provide full functionality support to manage the information through the phone. In our system, we limit the functionality to just the minimum that supports the most frequent use of the phone while away from the office. The goal is to provide access to but not management of information.

CONTEXT OF USE

To better understand the context of use of the system, consider the following scenario: It is an early weekday morning and Alfred is driving to work. While he is slowly moving through the dense city traffic he uses his hands-free cell phone to call into his company's voice system. Alfred listens to the weather forecast, his new emails, and personal calendar. He has gotten used the system which allows him to be up-to-date even before he arrives at the office. Alfred loves that he can spend the 30 minute of commute in a meaningful way. Today his calendar informs him about a meeting at 8:30am which he had forgotten about. Since the traffic will not allow him to arrive on time, he simply says: "cancel meeting" and the system guides him in canceling the meeting and automatically sends all participants an email with a small "Sorry" note attached as a sound file. The remainder of his commute Alfred listens to news headlines from the BBC and his favorite geek news outlet Slashdot.

The scenario above highlights the main motivations for our work. In normal circumstances, the user would just listen to the personalized audio feed as if it were a radio broadcast. He or she can mentally prepare for the day without having to interact and be distracted by the system. The user has pre-customized the voice system to best fit his or her needs. However, in special circumstances, such as the situation described in the scenario, the user can perform basic important interactions to react to changes in the environment. Extraordinary circumstances may develop at any time and location such as receiving a phone call from the kids' school requiring a parent to change the day's schedule on the spot. This often means that the set of devices available is not the full complement of computational equipment. In the scenario, Alfred is driving and thus does not have access to his laptop or PDA (even if he has them with him).

Rather than replicating the full functionality of PIM applications (email client, calendar) with a voice interface, we have focused on planning for those interactions that are crucial and we reserve nonessential and complex tasks for other modes of interaction such as working in front of a desktop computer. For example, in response to a change of schedule, there is often a simple follow up: postpone or cancel all or some of the remaining events for the day in question. The result of the postponement is often rescheduling, but the initial communication required for all participants of the postpone meetings is to cancel the event. The reschedule can take place at a later time.

Another key component, not highlighted in the scenario, is the requirement that the information presented over the phone is always up to date to the minute. If this system requires synchronization, its effectiveness decreases as for example, an important email may not be read on time.

EARLY PROTOTYPE

For nine months, one of the authors has used a prototype of a voice-enabled calendar system. In the period of April 2004-December 2004, there have been 64 phone calls to the system; an average of about 7 calls per month. In this period of time, a small set of types of activities have emerged where the phone access to the calendar has been essential. These include:

- Upon return from a trip to a conference, refresh memory of day's activities
- Handling of unexpected events (car trouble, traffic, weather, urgent personal situations) that requires cancellation of day's events
- Remote query (e.g. while at lunch), handling the question: "Are you available later today?"

Note that in the handling of these situations, the user was not always accessing his calendar from his own cell phone. At times, he had borrowed a phone to call and check his calendar.

In all of these cases, the most important feature of the system was the pervasiveness (or ubiquity) of the information. Unlike the PDA, which requires synchronization, the information accessed over the phone was always up to date. Also unlike the PDA, the information is available from any phone, so even when the user did not have his phone at hand, as long as he remembered his voice calendar phone number, his calendar information was always available.

For a related project, in order to investigate the dimensions of time management and gauge the potential interest in voice-enabled applications we conducted an online survey of calendar management. We found out that from 98 respondents, 67 said they would use a voice-enabled calendar system if one was available. This evidence encouraged us to design the system here described.

The phone access to calendar information was so effective, that we decided to extend the design to cover other personal and personalized information (email, community events, weather, news) while retaining the main design goals. The following section describes the architectural design of the system and how it is integrates with the users' desktop-based PIM activities.

SYSTEM ARCHITECTURE

Figure 1 illustrates the information flow and the technologies enabling the system. Before a user can listen to a personalized audio show, he or she must define preferences of information content, depth and order using a customization application (Figure 2). This software is implemented as a web application based on the server-side programming language PHP. Most of the customization choices are stored on the web server as a file in XML format. Only the authentication information (pre-configured caller-IDs and PIN numbers) are contained within a MySQL database, for increased query performance upon incoming phone calls. Users maintain their personal calendars via a desktop application like Apple iCal which is capable of exporting calendaring information in the standardized iCal/vCal data exchange format. Every time the calendar information is updated on the desktop, the calendar software automatically transfers an iCal file to the user's account on the web server via WebDAV. Thus, the server copy of the calendar file is always up to date.

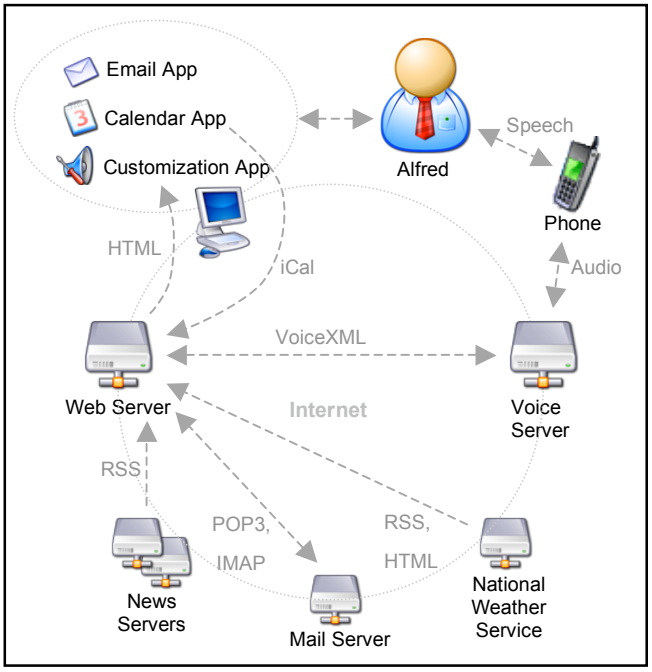

Figure 1: Architecture of the Voice system

When a user calls into the system, the voice server first authenticates the person via the caller ID (or optionally a user-entered "home" phone number) and a 4-digit PIN number. It then answers with a custom-generated VoiceXML program which the built-in Text-to-Speech engine transforms into audible output. A PHP script generates the VoiceXML for each individual user based on the user's preferences as well as the data the web server receives in real-time (such as e-mail from the mail server) or periodically in a store-and-forward fashion (such as a news feed from the BBC) from other sources.

A web-based customization application allows the user to select and configure the channels he or she would like to listen to and also pre-program the order of presentation. The following channels are currently implemented (although adding new content is quite straightforward):

- *Personal calendar:* Reads the user's personal calendar published from iCal
- *Event calendar:* Reads community and campus related events. Users can "subscribe" to the event categories matching their interests (e.g. movies, sports)
- *Email:* Reads sender and subject of new emails and (soon) on demand the email's contents
- *Weather:* Reads current temperature and forecast based on a user-defined ZIP-code (we are interested in transforming this into a automated, location-based service)
- *News:* A collection of news channels that read story headlines and short descriptions received via RSS.

The configuration on this website needs to be one only once but can be adjusted if the user's preferences change.

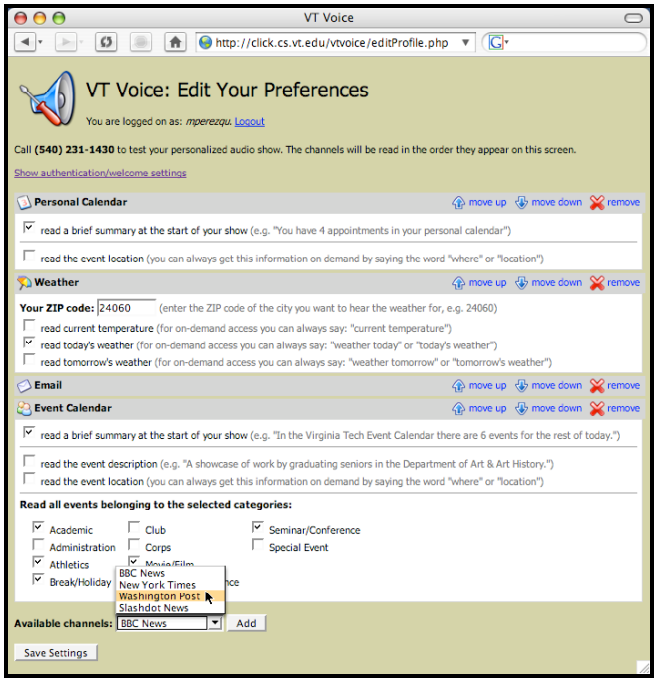

Figure 2: The web application used to define user's preferences. The E-Mail channel appears collapsed.

DISCUSSION, CONCLUSIONS AND FUTURE WORK

While our voice system is currently still in the prototype stage we are planning to expand it into a pilot project, soon to be used by members of the Virginia Tech community to evaluate the efficacy of the idea of pre-customized phone access to personal information in an ecologically valid setting.

Until now our work has concentrated on the design and implementation of the infrastructure and user interfaces required to enable voice output of personal and personalized information. Our next step will be to enhance the system by enabling interactivity. As discussed before, the focus will be on providing important features such as event cancellation, and notification (such as sending an email with a voice attachment) but *not* on implementing the voice interface equivalents of desktop applications.

Two major weaknesses to our approach are cost and the quality of the transmission. In the long run, listening to news headlines on the cell phone may not be economically feasible. Furthermore, the quality of current-day text-to-speech engines is still not comparable to human-produced contents. Although the quality problem could be addressed at least for news headlines by providing pre-recorded sound files instead of computer generated audio, the cost issue remains. Mobile audio players, such as Apple's iPod may provide a solution to this dilemma. In fact, the current trend to provide audio streams for download, known as *podcasting*, may provide the technical base for an integrated solution that addresses the cost as well as the quality problem. However, online access through phone or other means would still be required to provide up-to-date content such as access to email. We see a considerable potential for synergy effects if the cell phone inherits bigger storage capacity from hard-disk audio players.

ACKNOWLEDGEMENTS

We thank B. Collier Jones and Rob Capra for their ideas and valuable feedback throughout this work and David Vignoni for creating the "lively" icon library we used for user interface design and illustrations.

REFERENCES

1. Anerousis, N. and Panagos, E. (2002). "Making Voice Knowledge Pervasive." IEEE Pervasive Computing, Volume 1 Issue 2, pp. 42-48.
2. Danielsen, P. J. (2001) "The Promise of a Voice-Enabled Web", IEEE Computer, August 2000, Volume 33, Number 8, pp. 104-106.
3. Dourish, P. (2001). Where the Action Is: Foundations of Embodied Interaction. MIT Press.
4. Lucas, B. (2000) "VoiceXML for Web-based distributed conversational applications", Communications of the ACM September 2000, Volume 43, Issue 9.
5. Perry, M, O'hara, K., Sellen, A., Brown, B., and Harper, R. "Dealing with mobility" ACM Transactions on Computer- Human Interaction (TOCHI), December 2001, Volume 8 Issue 4.
6. Schmandt, C. (1993). "Phoneshell: the telephone as computer terminal.", MULTIMEDIA '93: Proceedings of the first ACM international conference on Multimedia, pp. 373--382. Anaheim, California, United States. ACM Press.
7. Weiser, M. (1991). "The computer for the 21st century." Scientific American, 265(3), 94-104.
8. World Wide Web Consortium, "Voice Extensible Markup Language (VoiceXML), Version 2.0," www.w3.org/TR/2001/WD-voicexml-2020011023 (current May 2002).
1. Yankelovich, N. (1994). "Talking vs. taking: speech access to remote computers." CHI '94: Conference companion on Human factors in computing systems, pp. 275-276. Boston, Massachusetts, United States. ACM Press.