WAP-Speech: Deriving Synergy between WAP and the Spoken Dialog Interface

Tin-Hang Lo and Helen M. Meng

Human-Computer Communications Laboratory
Department of Systems Engineering & Engineering Management
The Chinese University of Hong Kong,
Hong Kong SAR, China.
`{thlo, hmmeng}@se.cuhk.edu.hk`

ABSTRACT

We present a prototype that integrates WAP-based information access with spoken dialog interaction. WAP phone usability is encumbered by keypad entry for typing text, and also by the small screen-size that requires much scrolling during browsing or menu selection. The spoken dialog interface offers a desirable complement whereby commands may be spoken instead of typed, and the dialog design may present a menu hierarchy for user selection. We refer to this integrated interface as *WAP-speech*. Our prototype system offers two information services on an Enterprise Java Beans platform. (i) The weather information service allows the user to traverse a three-tier menu on the WAP phone to retrieve the weather information of a city. Alternatively, the user may interact with the spoken dialog interface and the requested weather information is pushed to his WAP phone. (ii) The stock alert service allows the user to specify the alert condition for a Java-based software agent to monitor a real-time financial information feed. When the pre-specified condition is met, the alert message is automatically generated and pushed to the user's WAP phone. Preliminary usability studies indicate that the *WAP-speech* interface can significantly shorten task completion time when compared to the WAP-only interface. The software agent architecture design is suitable for asynchronous WAP applications, where it performs off-line tasks and generates WAP push messages.

Keywords

WAP, spoken dialog interface

1. INTRODUCTION

In recent years we see the emergence of wireless Web-based data services. It is estimated that there are currently 4.4 million Wireless Application Protocol (WAP) pages worldwide [1], and subscribers of NTT i-mode in Japan has reached 60 million [2]. However, the popularity of Web access on the mobile phone is encumbered by usability problems in both input and output. To input text with the WAP phone keypad is a cumbersome process, to input the word "Hong Kong" requires keying in the number sequence "44-666-66-4-0-55-666-66-4".[1] Keying in non-alphabetic text for languages, such as Chinese, special steps need to be learned. As regards the output, the small WAP phone screen supports only four to six lines of text, each with 20 characters. Selecting an item from a menu list, the user needs to scroll through many screens of snippets of information. The situation is aggravated by information latency from the 14.4kpbs WAP transmission rate. It is easy for the user to lose track during WAP browsing.

To circumvent these usability problems, we propose to integrate WAP with a spoken dialog interface. Users can input commands simply by speaking. The hierarchy of menus typical of WAP applications may be transformed into goal-oriented spoken dialogs. Hence users may either traverse the WAP menu tree to retrieve personal information, or issue spoken requests via the spoken dialog interface and upon finishing the dialog, the requested information is pushed onto their WAP screens. We believe that the spoken dialog interface may expedite the interaction process for information access. From the perspective of a pure spoken displayless dialog interaction, the WAP screen also serves as a desirable complement. In a displayless setting, the information must be presented through speech synthesis. Since speech is transient, it takes the user sometime to listen to the entire message. In comparison, the small visual display can present information in a more succinct manner. Hence we believe that such *WAP-speech* integration presents many advantages. The only constraint hindering tight integration is that the current 2G environment requires explicit context switching between the WAP and the spoken dialog interface [3]. However, this constraint will be relaxed with the onset of 3G [4]. Furthermore, spoken dialog development will proliferate with the W3C's effort in developing VoiceXML 2.0 [5].

We have implemented a prototype system integrating WAP with a spoken dialog interface, to which we will refer as the *WAP-speech* interface. Our system offers weather and stocks information services. These services, together with our system architecture and usability statistics, will be presented in the following sections.

---

[1] The hyphen denotes a pause in keypad entry.

2. The WAP-Speech Weather Information Service

2.1 The WAP-based Interface

We have implemented a WAP-based information service using a three-tier hierarchical menu. For example, to reach the Hong Kong weather report, the user has to perform (i) one scroll and three clicks to select "Asia" in the *REGION* menu; (ii) one click and two scrolls to select "China" in the *COUNTRY* menu; and (iii) keying in the string "Hong Kong" in the *CITIES* text-box, which requires 17 key presses (see Figure 1).

Figure 1. The three-tier hierarchical menu : (i) Selecting Asia from the region menu , (ii) Selecting China from the country menu, (ii) Keypad entry for "Hong Kong" in the text box for city names.

Another four clicks (OK, HONG KONG, FOLLOW LINK and CURRENT FORECAST) are required to obtain the real-time weather forecast (see Figure 2).

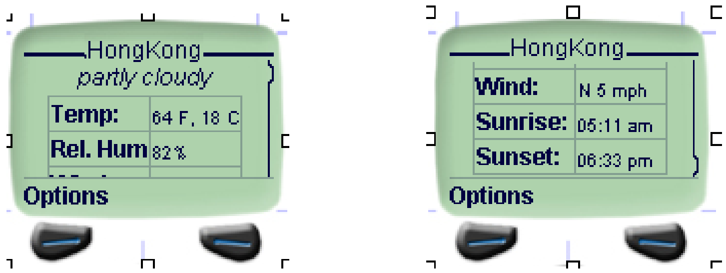

Figure 2. Weather information for Hong Kong is presented in two screen dumps on the WAP phone.

Since keypad entry is rather cumbersome, we improve the system usability by means of a substring matching technique. This is illustrated in Figures 3 and 4. The user can replace 17 key presses with two key presses and one click to retrieve the Hong Kong weather information.

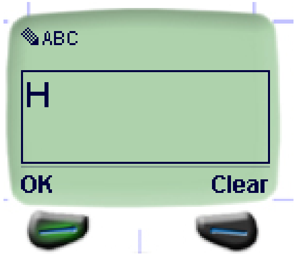

Figure 3. User types in the letters 'H' and 'O'.

Figure 4. Substring matching brings the user to the first city name in our application whose spelling begins with 'HO'. The user can simply click the link to get the weather information for Hong Kong.

2.2 The Spoken Dialog Interface

We have developed a trilingual spoken dialog interface supporting the languages used in Hong Kong (Cantonese, Putonghua and English). Speech complements the WAP phone input/output interface, and the dialog can be designed to embed the structure of the three-tier menu hierarchy (see Table 1).

System	Welcome to the WAP-speech global weather service. Please say the name of a geographical region. Possible options include: Asia, Africa, Europe…
User	**Asia** ***(user can barge in while listening to the menu items list)***
System	Please say the name of a country. Possible options include: China, Indonesia, Japan…
User	**China**
System	Please say the name of a city. Possible options include: Abagqi, Aihui, Akqi…..
User	**Hong Kong**
System	Would you like to select an additional city?
User	**No**
System	Thank you for using the WAP-speech weather information service. The following information has been pushed to your WAP personal profile. Hong Kong. Current temperature is sixty four degrees Fahrenheit or eighteen degrees Celsius. Relative humidity is eight two percent. Wind blows to the north at five miles per hour. Sunrise is at five eleven a m. Sunset is at six thirty two pm.

Table 1. Spoken dialog interface that corresponds to the WAP-based weather information service presented in the previous section.

Screen dumps in Figure 5 shows the new weather information pushed to the user. When the user clicks on "My Profile" (central screen-dump in Figure 5), he is connected to the city selected previously during the spoken dialog interaction, i.e. "Hong Kong".

Figure 5. The weather information requested by the user in the spoken dialog interaction is ultimately pushed to his WAP phone as shown in the three screen dumps above.

3. The WAP-Speech Stock Alert Service

We have also implemented a *WAP-speech* stock alert service. Our stock information is retrieved from a dedicated satellite feed provided by Reuters Hong Kong. Upon the user's request, a software agent is launched to monitor the satellite feed for a user-specified condition. A typical spoken dialog is presented in Table 2. This spoken dialog interface is also trilingual.

System	Welcome to the WAP-speech stock alert service. Please say the stock code.
User	**zero zero zero eight.**
System	Zero zero zero eight. Pacific Century Cyberworks. Current price is at two point five dollars per share. At what price would you like to set the information alert service?
User	**Two point nine dollars.**
System	Two point nine per share. Would you like to be alerted above or below this threshold?
User	**Above.**
System	Thank you for using our stock alert service. A software agent has been launched for you.

Table 2. An example dialog illustrating our spoken dialog interface for the WAP-based stock alert service.

When the user-specified condition is met, the alert message is automatically pushed to the WAP phone (see Figure 6).

Figure 6. The WAP-speech stock alert service. The alert message is shown in two screen dumps.

4. System Implementation

Our system architecture is illustrated in Figure 7. We have adopted the Enterprise Java Bean (EJB) platform. If only WAP is used for information access, user requests via the WAP phone are sent to the remote EJB clients in the front-end host machine. Each service has its remote EJB client (with IP address of corresponding application server) which invokes service remotely by RMI[1]. At server end, the EJB container initiates a new EJB or calls a Java-based software agent to perform services, such as retrieving weather information, or monitoring the Reuters satellite feed. Upon task completion, the software agent can push a message to the user's WAP phone by means of the WAP1.2.1 push technology. The push message is accompanied by an URI[2] linking to the new WML page. The alerted user can access the WML page by simply clicking the embedded URI in the push message. IBM has a similar idea called "Voice In, WAP Out" in their DEMOmobile in 2000 [6] which implemented text messaging for user alerts. They demonstrated a caller making inventory inquiries over the cell phone by interacting with a speech-enabled web server. The report requested by the caller was sent back as a text message to the cell phone in less then ten seconds. In our system WAP-speech, we used the WAP push function in WAP v1.2.1 for sending text messages to alert the user. WAP version 1.2.1 was the latest release at the time of implementation in 2001. Each of these text messages specify a URI, on which the user can click to be transferred to an automatically generated WAP page with the requested output information.

The front-end machine also hosts the spoken dialog system. Our trilingual spoken dialog interface is implemented on the SpeechWorks 6.0 platform that offers barge-in capabilities. It configures the overall user query and sends the command to the appropriate remote EJB client. Thereafter all interactions are identical to WAP-only information access as described above.

[1] Remote Method Invocation
[2] Universal Resource Indicator

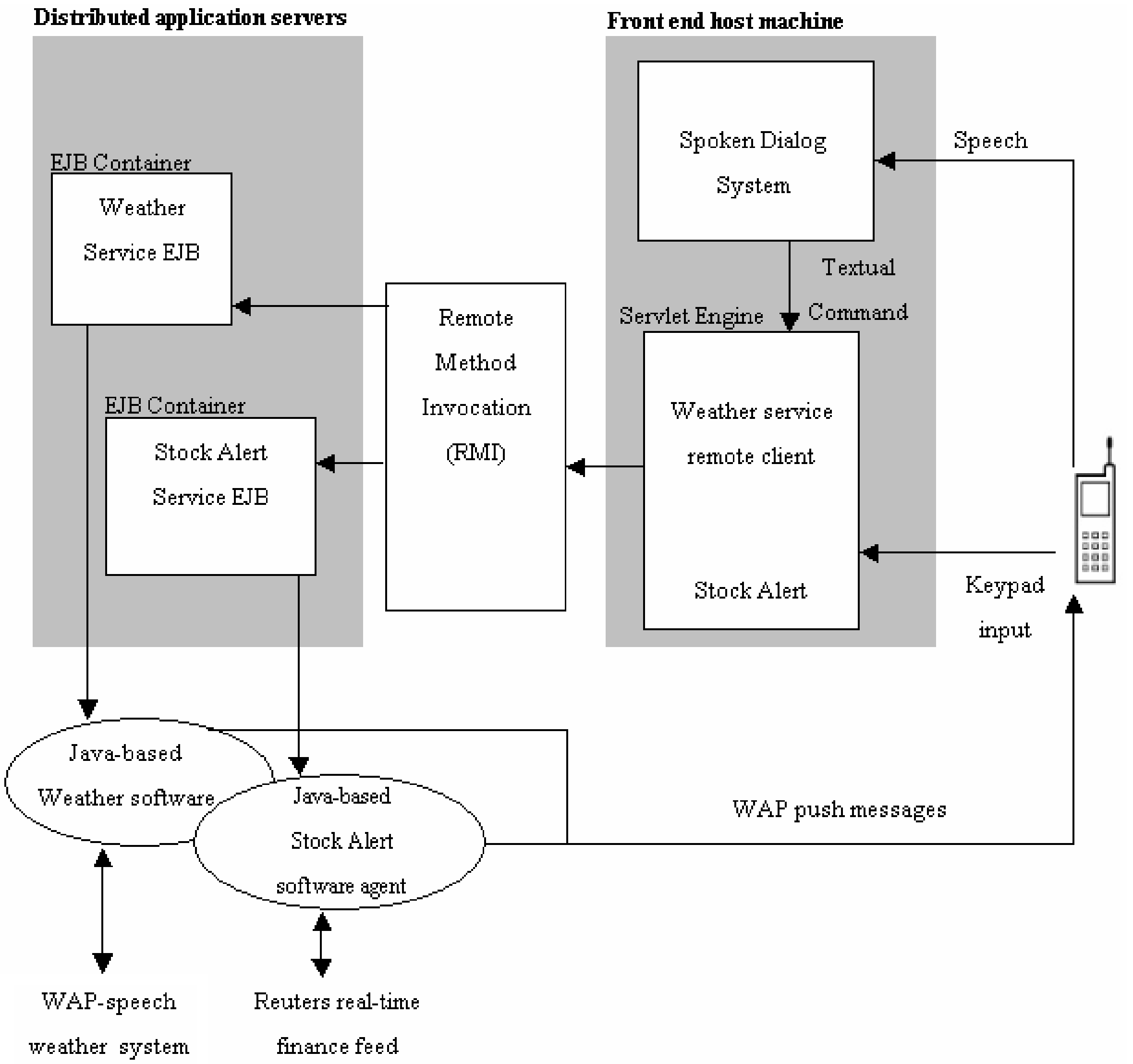

Figure 7. System architecture supporting WAP and spoken dialog interaction for Web information access.

We have conducted a preliminary evaluation based on the task completion duration using (i) keypad entry; (ii) keypad entry with substring matching; and (iii) *WAP-speech*. Users are asked to access the weather information for a city (see Table 3). It shows substring matching in keypad entry significantly expedites WAP-based information access. The duration is shortened further when *WAP-speech* is used.

	WAP (keypad entry)	WAP (keypad entry with substring matching)	WAP-speech
Duration for Task Completion (seconds)	188	86	60

Table 3. Task completion durations for weather information access using (i) keypad entry on the WAP phone; (ii) keypad entry with substring matching on the WAP phone; and (iii) WAP-speech, i.e. WAP integrated with spoken dialog interaction.

5. Conclusions and Future Work

We have integrated WAP-based information access with spoken dialog interaction. Usability of the WAP phone interface is encumbered by the keypad entry and small screen size for input and output respectively. The spoken dialog interface offers a desirable complement whereby commands may be spoken instead of typed, and the spoken dialog may present a hierarchy of menus sequentially for user selection. We refer to this integrated interface as *WAP-speech*. We have implemented two prototype services based on an EJB platform. The weather information service allows the user to traverse a three-tier menu for the weather information of the city of choice. Alternatively, the user may interact with the spoken dialog system and the requested weather information is pushed to his WAP phone. The stock alert service allows the user to specify the alert condition for a software agent to monitor a real-time financial information feed. When the pre-specified condition is met, the alert message is automatically generated and pushed to the user's WAP phone. Preliminary usability studies indicate the *WAP-speech* interface can significantly shorten task completion time when compared to the WAP-only interface. Future work will be devoted to the development of spoken dialog interface in VoiceXML 2.0 and multi-access to Web information via Web, WAP and voice browsing.

6. ACKNOWLEDGMENTS

This project is supported by SpeechWorks International Ltd. and Reuters Hong Kong. We thank Mr. Arthur Goh and Mr. Steven Lee from SpeechWorks for various helpful suggestions and comments.

7. REFERENCES

1. WAP forum. *WAP specification evolves with advancements in the wireless Internet experience* http://www.wapforum.org/new/20001212276WAP.htm**.**
2. M. Ramsay and J. Nielsen, *WAP Usability Report*, Nielsen Norman Group, December 2000.
3. AT&T. *Wireless PocketNet® Service.* http://www.attws.com/personal/pocketnet/index.jhtml;dsessionid=DP45RJRDQDDGTR1M5W0SFEY
4. Office of Telecommunications. *Bluffers Guides : Mobile Telephony* http://www.oftel.gov.uk/publications/news/on52/blgu0601.htm.
5. S. McGlashan, D. Burnett, P. Danielsen, J. Ferrans, A. Hunt, G. Karam, D. Ladd, B. Lucas, B. Porter , K. Rehor, S. Tryphonas, Voice Extensible Markup Language (VoiceXML) Version 2.0, World Wide Web Consortium, 2001 http://www.w3.org/TR/2001/WD-voicexml20-20011023/
6. IBM. *Voice In, WAP Out" in DEMOmobile* http://www-3.ibm.com/software/speech/news/20000908-es.html.
7. D. W. Augustsson, *WAP Push Architectural Overview Version 08-Nov-1999*, Wireless Application Protocol Forum, Ltd., 1999.
8. H. H. Liao, *WAP Push Message Version 16–August-1999*, Wireless Application Protocol Forum, Ltd., 1999.