
Enhancing Mobile Browsing and Reading

Chen-Hsiang Yu

MIT CSAIL
32 Vassar St
Cambridge, MA 02139
chyu@mit.edu

Robert C. Miller

MIT CSAIL
32 Vassar St
Cambridge, MA 02139
rcm@mit.edu

Abstract

Although the web browser has become a standard interface for information access on the Web, the mobile web browser on the smartphone does not hold the same interest to mobile users. A survey with 11 mobile users shows that only 18% of the participants like mobile web browsers, whereas 82% of them like other mobile applications. This research focuses on understanding mobile users' difficulties and proposes innovative ideas to enhance mobile web browsing. This research enhances mobile browsing and reading in three directions: (1) dynamically generating mobile web sites for browsing (2) using orientation sensor information to detect natural interactions and text-to-speech (TTS) to continue reading between different activities, and (3) providing a speech interface to ease web navigation and supporting dialog programming for repetitive tasks. The *Read4Me* Browser is a prototype system built to demonstrate the proposed ideas.

Keywords

Mobile browsing, mobile reading, Read4Me

ACM Classification Keywords

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

General Terms

Design, Experimentation

Introduction

According to Pew Internet Mobile Access 2010 [9], eight in ten American adults (82%) currently own a cellphone of some kind and 32% of American adults use their cellphone to access the Internet, an increase from the 25% in 2009. The market research firm Nielsen forecasts that there will be more smartphones than feature phones in the U.S. market by the end of 2011 [8].

Although the web browser has become a standard interface for people to access information on the Web, mobile web browsers do not hold the same interest to mobile users for several reasons, including fat finger issues for input and mobile navigation, limited display for web page rendering, and difficulties in continuing to read while switching activities, such as standing up to walk, or getting into a car to drive. In order to provide better user experiences, some web sites develop a mobile version for mobile web browsers, but the development process is pricy and time-consuming. Even though the web sites can be customized for mobile web browsers, the mobility capability of the mobile phone raises more issues for mobile users. For example, the phones are used with a short attention span and the users are prone to switch activities while using the device. If you have repetitive tasks to run, such as checking weather, emails or schedule, the fat finger problem is still an issue on small display devices.

In this paper, we analyze the above issues and propose solutions for some of them in three directions. First, instead of investing effort to recreate a new mobile web site for mobile users, this paper proposes that mobile web browsers can perform a transformation on any traditional web site by using its web page information

to provide organized content for better mobile browsing. The transformation process could be done automatically on the fly. Second, whereas focusing on the display of a smartphone may be difficult while walking or driving, Vadas's study showed that audio is an acceptable modality for mobile comprehension of text. [10] Although audio has been used in screen readers for blind or visually impaired people [1,4,5,6], there was no clear suggestion of how to combine the audio modality with visual presentation for sighted people on mobile devices. To support the requirement of frequent context switches on the mobile web browser, this paper proposes to combine orientation sensor information with text-to-speech (TTS) to enable continuity.

Third, although web automation [3] and programming by demonstration systems (PBD) [2] have successfully demonstrated the potential for repeating tasks on PCs, there is no clear model which can be used directly on mobile web browsers. Without keyboard and mouse, the condition becomes worse. This paper proposes that mobile web browsers should provide a speech interface to ease web navigation and support dialog programming for the repetitive tasks.

Read4Me Browser is a mobile web browser built to demonstrate the proposed ideas. In the following sections, we will talk about these ideas and analyze the design of the system. Discussion and future work are also added at the end of this paper.

A survey of mobile usage

In addition to mobile web browsers, mobile applications have become another popular option to access information on Web. In order to provide better design



Figure 1: Mobilize NYTimes web site automatically by using "Mobilize Site" feature of the Read4Me

for mobile users, a survey was conducted to understand mobile users' experiences and opinions. This paper followed a user-centered design process similar to [11] and conducted a formative, pre-design evaluation on a sample of lead users. These users were randomly recruited on a college campus by asking people if they ever used a smartphone. The ones who had used smartphones and were willing to join the survey were the participants. 11 participants joined the survey. The participants' age ranges are: 20-30 (64%), 30-40 (9%), 40-50 (9%), >50 (18%). 20 questions were asked in the survey and each survey took about 30-45 minutes. There were 7 males and 4 females and all of them had experience using a smart phone, a mobile web browser and mobile applications.

The results show that many participants spent less than 1 hour per day using mobile web browsers (55%) and mobile applications (55%). The top three activities for using the smartphone are while waiting, walking and while lying in bed. To our surprise, driving is also a situation in which participants use mobile web browsers (9%) and mobile applications (18%). In terms of the purpose, participants reported that they use mobile web browsers for e-mail (73%), reading news (64%) and search (27%) and use mobile applications for social networking (73%), weather (64%) and playing games (36%). Although 82% of the participants like the design of mobile applications, only 18% of the participants like mobile web browsers. The reasons for disliking mobile web browsers include: bad content layout on the mobile device, interface makes it slow to skim and browser, etc. The survey results show that mobile web browsers have a lot of space to improve.

Read4Me Browser

This research attempts to enhance mobile browsing in three directions: generating mobile web sites dynamically, using orientation sensor information with text-to-speech (TTS) techniques to support continuous reading while switching activities, and providing speech interfaces to ease web navigation and support dialog programming for repetitive tasks. Read4Me Browser was developed to demonstrate these ideas and is currently implemented on the Android OS 2.2 platform.

Mobilize web sites automatically

Although some web sites already provide mobile versions for mobile users, there are still many web sites that only have one type of web page, which makes mobile browsing difficult on these sites. As mentioned previously, the development process of having a mobile web site is time-consuming and pricy. Having a way to generate mobile web sites from traditional web sites would provide many benefits to both of content providers and mobile users. Different from the previous work [7], Read4Me Browser uses existing information of the web site to create a mobile application-like mobile web site dynamically.

When Read4Me Browser navigates to a web site that is only designed for PCs, users can click on the "Mobilize Site" option from the menu to transform it into a mobile version. The algorithm behind the scenes involves the Read4Me Browser sending the current URL to a back-end server, which is dedicated to parse this web page information to generate a new mobile web site automatically with a mobile application-like user interface (Figure 1). Compared with the traditional development process, which takes a few weeks or

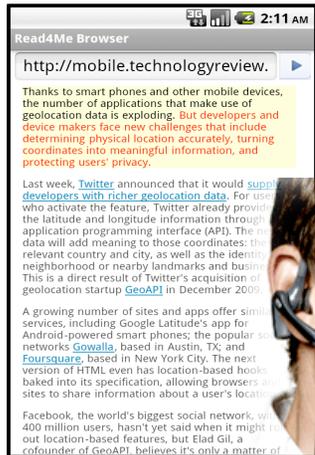


Figure 2: Reading sentences for hands-free condition

months to create a mobile web site, the “Mobilize Site” feature of Read4Me Browser creates mobile web sites in real time.

Natural interaction with TTS for continuous reading

Reading a long article on mobile web browser is not convenient, but frequent context switching, like standing up or getting in a car, make it even worse. This paper proposes using orientation sensor information to detect natural interactions and text-to-speech (TTS) to solve this issue, i.e., the browser can automatically detect certain conditions and react appropriately so the user can continue reading, even while switching from a hands-on to a hands-free condition.

Read4Me Browser defines two operational states, Look & Point and Hear & Say, and uses orientation information to provide seamless transitions between different conditions. While the first state represents the condition of looking and pointing at a mobile screen, the second state supports hands-free conditions, such as walking or driving. In the Look & Point state, a user can read web pages as usual and use his finger to touch the screen for scrolling, clicking, zooming, etc. On the other hand, when switching to the Hear & Say state, users can listen to web page content, which is parsed by Read4Me Browser and read using TTS. In the Hear & Say state, Read4Me Browser reads one sentence at a time and highlights the sentence being read in red so that the user can easily identify it when the user switches back to the Look & Point state (Figure 2).

Two conditions can trigger switching from the Look & Point state to the Hear & Say state: (1) the user brings the phone up to his ear (2) the user docks the phone inside his car or puts the phone in his pocket. On the other hand, when the user places the phone in a viewable position, such as a reading position, the phone will switch to the Look & Point state (Figure 3).

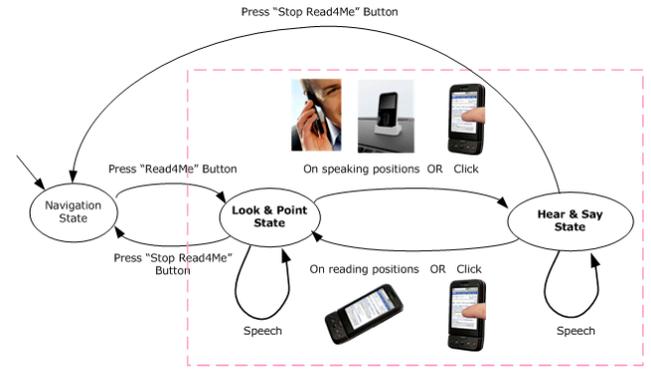


Figure3. The scenario of seamless handover

Speech Interface and Dialog Programming

The fat finger problem is a well-known issue for mobile users, which affects the user’s ability to select a textbox in a webpage, or follow a link. Instead of typing and clicking, Read4Me Browser allows users to speak to it. There are a few benefits to using speech in mobile browsing. For example, it is easier to ask the browser to search for keywords rather than reading through the whole page, and it is easier to tell the browser to click on a tiny link rather than using a fat finger.

Read4Me Browser takes these observations into consideration. When reading a web page in the Look & Point state, users can say “search for XXX” to locate

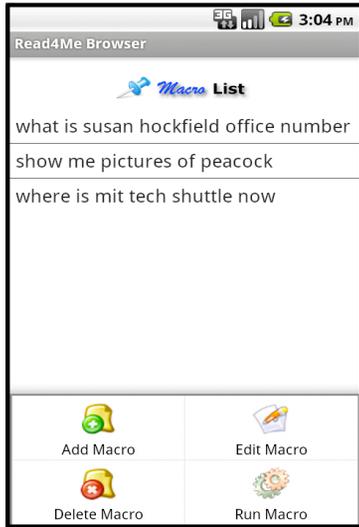


Figure 4: Macro Viewer

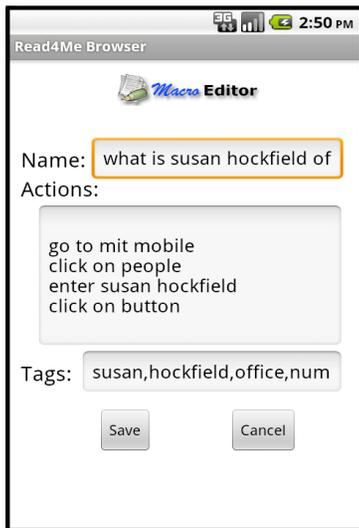


Figure 5: Macro Editor

the XXX term in the current web page. In addition, if the user plans to go to the New York Times or MIT mobile web site, instead of typing `http://www.nytimes.com` or `http://m.mit.edu` into the address bar or choosing them from a bookmark, users can say "go to new york times" or "go to mit mobile".

Read4Me Browser currently implements three commonly used commands: "go", "click" and "enter". Because Read4Me uses the speech recognition feature in the Android SDK, the successful recognition rate is low if you only use one word as a command. To improve the success rate, a usability test was conducted on each of these commands. The experimental results show that adding a preposition behind specific verbs, such as "go" and "click", can enhance the speech recognition to acceptable levels. The success rate of "go" vs. "go to" is 10% vs. 97% and "click" vs. "click on" is 55% vs. 80%. However, the "enter" command can be recognized well without adding a preposition. (Success rate is 90%) These experiments were based on comparing the spoken commands with the first recognized result from the Android speech recognizer. Based on the testing results, Read4Me Browser uses the commands listed in Table 1.

Table 1: Read4Me browser speech commands

Commands	Description
go to	Direct browser to go to indicated URL
click on	Click on objects on web page, including links, buttons, etc.
enter	Enter data into the input field

Although Read4Me Browser only supports three commands at this moment, it can already be used to perform some interesting tasks, such as showing pictures of a given term or finding someone's contact information on the web. Users say one instruction at a time to make Read4Me Browser execute it.

One drawback of the above design is that if a user needs to repeat the same or similar task daily, he needs to say the same instructions again. For some tasks, this is fine, but some tasks are tedious. To save users time, Read4Me Browser extends the above speech interface to support dialog programming, i.e., saving speech commands as a macro. Users can say a high-level intention and Read4Me Browser finds the most similar macro script to run. When there is no macro to satisfy the intention, Read4Me Browser will learn the process by asking for a user demonstration, which will be saved as a macro script for future reference. Users can also use a macro viewer to view all recorded scripts (Figure 4) and use a macro editor to edit these scripts (Figure 5). Figure 6 shows macro script examples and Figure 7 shows the execution of a macro script.

Restrictions and Challenges

This paper attempts to enhance mobile browsing and reading in three directions, but there are still a few challenges that need more investigation. First, although Read4Me Browser can use orientation information to provide seamless transitions between activities for continuous reading, the orientation sensor is too sensitive to interpret the user's intention successfully all the time. An investigation should be conducted to provide better guidelines for this kind of usage. Second, using the speech interface to automate web navigation

Ex. 1: show me pictures of peacock
go to google image
enter peacock
click on button

Ex. 2: where is mit tech shuttle now
go to mit mobile
click on shuttle schedule
click on tech shuttle

Figure 6: Macro script examples

is helpful, but for dialog programming, we still need to study a good way for users to parameterize scripts. For example, if the browser has already learned “show me pictures of peacock”, could the browser learn to satisfy the request “give me hedgehog pictures” by replacing “peacock” with “hedgehog” in the existing script? We would like to design an algorithm for achieving this parameterization.

Conclusion and Future Work

Mobile browsing has many limitations originating from its physical design. Our research attempts to enhance mobile browsing and reading in three directions: generating smartphone friendly mobile web sites, using orientation sensor information to detect natural interactions and TTS for mobile reading, and providing speech interfaces and dialog programming. Read4Me Browser is a prototype system demonstrating the above ideas. Future work will focus on conducting an experiment on orientation data to provide better guidelines for natural interaction, and investigate a new algorithm for reasoning and parameterizing recorded speech macros.

Acknowledgements

We thank MIT UID group and reviewers’ valuable suggestions. This work is supported in part by Quanta Computer as part of the TParty project.

References

- [1] Bigham, J., Prince, C. and Ladner, R. Webanywhere: A Screenreader On-the-go. *W4A '08*, 2008.
- [2] Cypher, A. Watch What I do – Programming by Demonstration. Chapter 9. The MIT Press.
- [3] Darris Hupp and Robert C. Miller. Smart Bookmarks: Automatic Retroactive Macro Recording on the Web. *UIST*, 2007.
- [4] Freire, A. et. al. Using screen readers to reinforce web accessibility education, *Proceedings of the 12th annual SIGCSE conference on Innovation and technology in computer science education*, 2007.
- [5] Gaal, A. et. al. Towards an Open Source Screen Reader: Screenreader Usability Extensions (SUE). *In Proceedings of ICCHP*, 2008.
- [6] Miyashita, H. et. al. Aibrowser for multimedia: introducing multimedia content accessibility for visually impaired users. *Proceedings of SIGACCESS on Computers and accessibility*, 2007.
- [7] Nichols, J., Hua, Z., and Barton, J., Highlight: A System for Creating and Deploying Mobile Web Applications. *Proc.UIST2008*, ACM, 2008.
- [8] Nielsen: US Smartphone Penetration to Be over 50% in 2011, GPS Business News, 2010.
- [9] Aaron Smith. Mobile Access 2010, Pew Internet & American Life Project, 2010.
- [10] Vadas, K., Patel, N., Lyons, K., Starner, T. and Jacko, J. Reading on-the-go: a comparison of audio and hand-held displays. *In MobileHCI*, 2006.
- [11] Yu, C.H. and Miller, R.C. Enhancing web page readability for non-native readers. *Proc.CHI2010*, ACM, 2010.

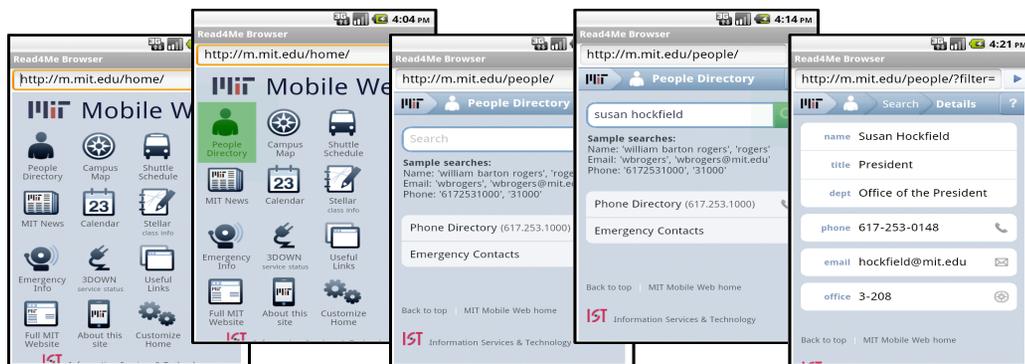


Figure 7: Running a macro to find MIT President’s contact information.