IEICE TRANS. INF. & SYST., VOL.E93–D, NO.11 NOVEMBER 2010

LETTER

Estimating Mobile-Friendliness Scores of Web Pages

Jihwan SONG[†*a)], ***Student Member***, **Xing XIE**[††], **Yoon-Joon LEE**[†], ***and*** **Ji-Rong WEN**[††], ***Nonmembers***

SUMMARY Mobile devices such as cell phones and personal digital assistants (PDAs) are becoming increasingly popular tools to access the Internet. Unfortunately, the experience of users attempting to access web pages with these mobile devices has been less than satisfactory because of their small display areas, slow communications links and low computing power. In this paper, we propose a trained scorer to estimate the mobile-friendliness scores of web pages, providing an indication of their suitability for mobile devices. These scores help mobile-friendly pages receive higher ranks in search results when mobile users seek information on the web. Our experiments show that the search results re-ranked by our mobile-friendliness scores increase mobile user satisfaction.
key words: *mobile search, usability study, mobile-friendliness*

1. Introduction

Despite the rapidly increasing number of mobile users [1], general-purpose search engines may not put mobile-friendly pages high in their search results. This is because search engines do not explicitly consider the mobile-friendliness of web pages in their page rankings. Here, *mobile-friendliness* means mobile users can obtain information from the page without any major inconvenience, even though mobile devices have certain inherent limitations like small display areas, less powerful CPUs and slow communications links [2]. In other words, mobile-friendliness indicates the suitability of a given page for mobile devices. Although there are "mobile" search engines, they usually focus on finding only mobile-specific pages (written in WML, cHTML, etc.). Such mobile-specific pages are undoubtedly mobile-friendly, but they are very few in number. Consequently, mobile search engines frequently fail to meet mobile users' information needs.

In this paper, we propose a method for estimating mobile-friendliness scores of web pages. The *mobile-friendliness score* indicates in a quantitative manner how suitable a web page is for mobile devices, and thus can help search engines find mobile-friendly pages on the web. The contributions of our work include:

- A usability study was conducted to find out which pages are mobile-friendly and why. Here, 16 people participated and 400 web pages were examined.
- We derived 21 factors affecting mobile-friendliness from the results of the usability study. We call these factors *mobile-friendliness factors*. Also, we propose a trained scorer, the *Mobile-Friendliness Scorer* (MF-Scorer), for estimating mobile-friendliness scores.
- In our experiments, 40 people participated in estimating mobile-friendliness scores for 1,200 web pages. These scores were then compared with the scores generated by the MFScorer and by mobiReady**, a popular tool for measuring the mobile-readiness of web content. Our experimental results show that our mobile-friendliness scores are more similar to human-generated scores than mobiReady's.

2. Related Work

Web Content Accessibility Guidelines (WCAG) 2.0***, HTML 4.0 Guidelines for Mobile Access**** and Mobile Web Best Practices 1.0***** have been published to promote mobile-friendly web page design. Although we can derive a metric for calculating mobile-friendliness scores of web pages from these guidelines, many guidelines are impractical, some too abstract, and some unnecessary.

W3C MobileOK Checker****** and mobiReady are now being developed to evaluate the mobile-friendliness of web pages. MobileOK Checker only tests whether web pages follow the guidelines of Mobile Web Best Practices 1.0 and does not provide a quantitative measurement of their mobile-friendliness. mobiReady estimates mobile-friendliness scores like ours does, but our experimental results reveal a gap between scores generated by people and by mobiReady.

3. Usability Study

We performed a usability study to determine what factors could hinder obtaining information from desktop-oriented web pages with mobile devices. First, we selected 100 infor-

Manuscript received February 9, 2010.
Manuscript revised July 10, 2010.
†The authors are with Division of Computer Science, Korea Advanced Institute of Science and Technology, Daejeon, Republic of Korea.
††The authors are with Web Search and Mining Group, Microsoft Research Asia, Beijing, China.
*This work was done when the author was visiting Microsoft Research Asia.
a) E-mail: jhsong@dbserver.kaist.ac.kr
DOI: 10.1587/transinf.E93.D.3124

**http://ready.mobi/
***http://www.w3.org/TR/WCAG20/
****http://www.w3.org/TR/NOTE-html40-mobile/
*****http://www.w3.org/TR/mobile-bp/
******http://validator.w3.org/mobile/

Table 1 Obstacles to obtaining information from desktop-oriented web pages: causes and effects.

Summary	Cause	Effect
Page layout broken (35.1%)	– Page layout designed by using absolute position, absolute units or tables. – Page or table(s) has background images. – Page has overly large images.	– Users may not easily grasp main content. – Users may have a negative impression of the web page.
Functions not working (23.0%)	– Page has unsupported tags (IFRAME, ...), unsupported scripts (Java script, VBScript, some HTML Script (onmouse, onkey), ...), or unsupported objects (movie formats, sound formats, Flash movies, Java applets, image formats other than JPEG/GIF, ...)	– Users may not be able to use page functions. – Users may not be able to see some content related to non-working functions.
Frequent horizontal scrolling (22.1%)	– Page has high-resolution images, many columns or frames. – Page has unnecessary horizontal spaces.	– Users may need to horizontally scroll on every line when reading text. – User may miss some content rendered beyond the right boundary of a page.
Low primary content understandability (8.5%)	– Page has complex navigational links at the top. – Page has too many decorative images at the top that are not related to primary content.	– Users may miss or have difficulty grasping primary content.
Low image text readability (4.5%)	– Page has text in images that are scaled down.	– Users may have difficulty reading text in the scaled-down image text. Figure 1 shows examples of image text.
Long page loading time (3.9%)	– Page has many large images. – Page size itself is too large.	– Users may need great patience. – Communication costs may increase when transmitting a large page by wireless networks. – Page may not be fully loaded because it is too large.
Too much vertical scrolling (2.9%)	– Page is too long.	– Users may have difficulty accessing the primary content. – User concentration is weakened.

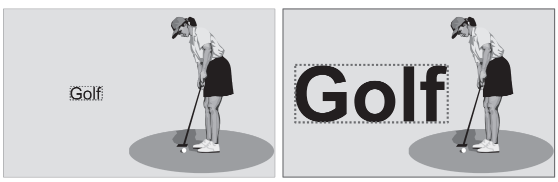

Fig. 1 Examples of image text (dotted boxes indicate text regions); the left image text "Golf" may not be recognized (bad readability) when the image size is reduced; in contrast, the right image text "Golf" may have good readability even when the size of the image is reduced.

mational web pages [3]; we randomly selected 10 queries from *Google 2002 to 2006 Year-End Zeitgeist*[†], *Yahoo! 2006 Top Searches*[††] and *the Top of Live Search 2006*[†††]; and we entered these queries into Google and collected 100 links from the results pages. (The links were saved to an HTML file for the subjects who participated in our study to easily open them.) Then, we asked nine people to give us reasons why they sometimes felt inconvenienced or frustrated when they opened the selected web pages in the browsers of mobile devices[††††]. Here, the nine subjects were all the MSRA interns who majored in computer science; they all had many experiences to access web pages via mobile devices. We analyzed their valuable 900 comments and summarize them as listed in Table 1.

After the first step of the usability study, we performed a second study to determine whether the degree of mobile-friendliness of a web page can be quantitatively calculated. We selected 300 web pages in a way similar to that of the first usability study. Additionally, we generated four questions based on the results of the first study. For each of the 300 web pages, seven people gave scores from 1 to 4 (1=poor, to 4=excellent) for four questions, and we regarded the average of the scores as a mobile-friendliness score of the page. These seven subjects were completely different members from the previous study (threes majored in computer science and the others majored in biology, English literature, electronics, and industrial design). If most of the subjects gave similar scores for each web page, the mobile-friendliness score of that page might be quantitatively calculated from our study. To check whether they gave similar scores or not, we used a two-way ANOVA as listed in Table 2. *Rows* refers to the seven mobile-friendliness scores generated for each of the 300 pages, and *Cols* refers to the 300 mobile-friendliness scores generated by each of the seven subjects. The results of the two-way ANOVA show that the subjects gave similar scores to each page.

Table 2 Results of a two-way ANOVA.

	SS	df	MS	F	P-value	F crit
Rows	1250.46	299	4.18	18.32	0	1.15
Cols	2.37	6	0.40	1.73	0.11	2.10
Error	409.63	1794	0.23			
Total	1662.42	2099				

4. Estimating Mobile-Friendliness Scores

From the results of the usability study in Sect. 3, we derived 21 factors that affect mobile-friendliness:

- **F1** (*page width*): minimum page width (px) without horizontal scrolling
- **F2** (*page height*): minimum page height (px) without vertical scrolling
- **F3** (*text size*): html document size (Bytes)

[†]http://www.google.com/intl/en/press/zeitgeist/
[††]http://buzzlog.buzz.yahoo.com/topsearches2006/lists/
[†††]http://livesearch.spaces.live.com/blog/cns!8560B877FE8E9138!1245.entry?_c=BlogPart
[††††]Microsoft Windows Mobile 6.1 compatible devices and their Internet Explorer Mobile browsers were used for all usability study and experiments in this paper.

- **F4** (*foreground image count*): number of foreground images
- **F5** (*background image count*): number of background images
- **F6** (*foreground image total size*): summation of sizes (Bytes) of the foreground images
- **F7** (*background image total size*): summation of sizes (Bytes) of the background images
- **F8** (*foreground image average size*): $\frac{F6}{F4}$ (Bytes)
- **F9** (*background image average size*): $\frac{F7}{F5}$ (Bytes)
- **F10** (*frame count*): number of frames
- **F11** (*column count*): number of columns
- **F12** (*maximum image width*): maximum width (px) of all images
- **F13** (*maximum image height*): maximum height (px) of all images
- **F14** (*wide image count*): number of images wider than the width of mobile device displays (e.g., default of 220 px)
- **F15** (*absolute unit count*): number of absolute units (e.g., 120 px) used in the attribute *width* in tags *col*, *colgroup*, *hr*, *table*, *td* and *th (thread)*
- **F16** (*layout table tag count*): number of empty table tags such as $\langle tr\rangle\langle/tr\rangle$ and $\langle td\rangle\langle/td\rangle$
- **F17** (*small and transparent image count*): number of transparent images and images of less than 10 pixels
- **F18** (*unsupported tag count*): number of tags such as *object*, *applet*, *script* and *iframe* that are not supported by mobile devices
- **F19** (*image text count*): number of images that include text and satisfy the following two conditions: (1) The width of the image is greater than the width of the display area of a mobile device and (2) $\frac{H_t W_d}{W_i} \geq 8$ (px), where H_t is the average height (px) of all text regions, W_d is the width (px) of the display area of a mobile device and W_i is the width (px) of the image
- **F20** (*top link ratio*): $\frac{A_l}{A_m} \times 100$, where A_l is the area (px) of all the links on the top of a page and A_m is the area (px) of the display of a mobile device
- **F21** (*top image ratio*): $\frac{A_i}{A_m} \times 100$, where A_i is the area (px) of all the images on the top of a page and A_m is the area (px) of the display of a mobile device

We propose a trained scorer, the Mobile-Friendliness Scorer (MFScorer), based on a support vector machine [4], [5] as shown in Fig. 2. We made a learning-based scorer rather than other methods (e.g., rule-based) because the number of mobile-friendliness factors is not small and many kinds of mobile devices exist in the world. MFScorer regards the estimation of mobile-friendliness scores as a multi-class classification problem; i.e., it classifies web pages into four classes: poor (score 1), bad (score 2), good (score 3) and excellent (score 4). For 21 mobile-friendliness factors (training features), MFScorer was trained with 300 mobile-friendliness scores (a training set) from our usability study. After training, if a user enters the URL of a web page in MFScorer, MFScorer calculates the values of 21 mobile-friendliness factors from the page and estimates the mobile-friendliness score (1=poor, 2=bad, 3=good and 4=excellent) of that page.

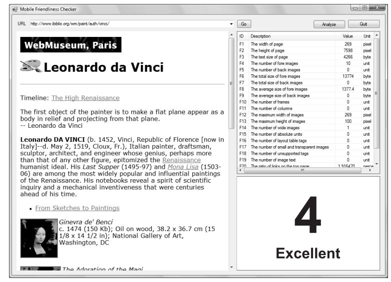

Fig. 2 Our trained scorer, Mobile-Friendliness Scorer.

5. Experiments

In the experiments, we show that MFScorer provides greater satisfaction to mobile users than does the previous work, mobiReady, when each of the mobile-friendliness scores of MFScorer and mobiReady are applied to the search results, respectively.

First, we selected 40 queries in the same way as in Sect. 3 and then entered the queries into Google. From the search results, we took the 30 top-ranked web pages (including ranking information) for each of the queries. We then asked 40 people to estimate the mobile-friendliness scores of those pages, with MFScorer and mobiReady estimating the mobile-friendliness of the same pages. Here, 53% of the 40 subjects majored in computer science and the others had different backgrounds such as mathematics, chemistry, physics, etc. After scoring the pages, we re-ranked the 30 web pages for each of the queries by the mobile-friendliness scores generated by the 40 people, by MFScorer and by mobiReady. To do this, the following value F was used:

$$F = (1-\alpha) \times \frac{R_{max} - (R-1)}{R_{max}} + \alpha \times \frac{S - S_{min}}{S_{max} - S_{min}} \quad \text{where } 0 < \alpha < 1 \qquad (1)$$

In Formula (1), R and R_{max} are an original rank and maximum rank, respectively; S is a mobile-friendliness score; S_{min} and S_{max} are the lower and upper bounds of a mobile-friendliness score, respectively; and α is a weight-value (i.e., the original ranks are given greater consideration when α goes to 0 and mobile-friendliness scores are given greater consideration when α goes to 1). In the experiment, R_{max} was set to 30, since only the top 30 pages were used. Because people gave mobile-friendliness scores from 1 to 4, the S_{min} and S_{max} of people were set to 1 and 4. In the case of MFScorer, its S_{min} and S_{max} were set to 1 and 4, respectively. Similarly, the S_{min} and S_{max} of mobiReady were set to 1 and 5, respectively. We re-ranked the 30 web pages for each of the queries in decreasing order of their F value.

For each of the queries, we selected the top 10 web pages generated by users as relevant pages. Then, by comparing the relevant pages with the top 10 pages of MFScorer

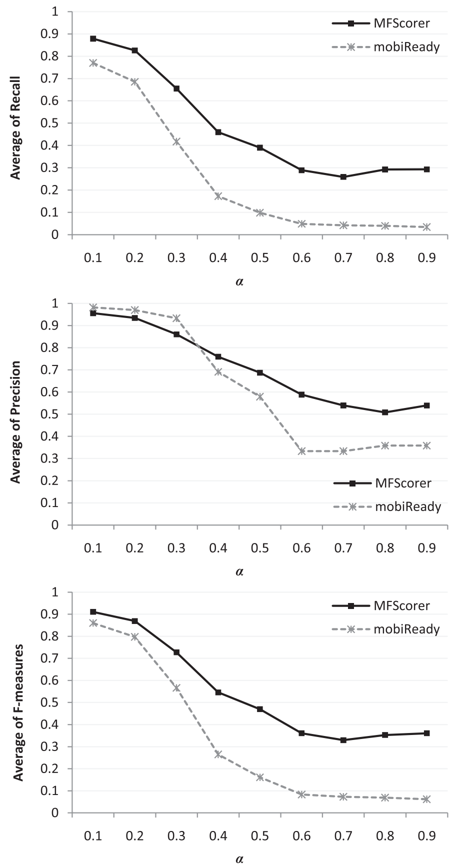

Fig. 3 Averages of recalls, precisions and F-measures of MFScorer and mobiReady with varying α between 0.1 and 0.9.

and mobiReady, we calculated recalls, precisions and F-measures for each of the queries. Figure 3 shows the averages of the recalls (top), precisions (middle) and F-measures (bottom) of MFScorer and mobiReady for all the queries with varying α from 0.1 to 0.9. In the figure, the averages of recalls of MFScorer were better than those of mobiReady. In the case of the averages of precisions, mobiReady outperformed MFScorer slightly between 0.1 and 0.35 of α, but MFScorer had much better performance than mobiReady in the other range of α. Lastly, the averages of the F-measures of MFScorer performed better than those of mobiReady at any value of α.

From the results so far, the mobile-friendliness scores estimated by MFScorer satisfy mobile users more than those generated by mobiReady when the scores are considered as a measure of mobile-user satisfaction.

6. Conclusion

In this paper, we proposed a trained scorer called MFScorer for estimating the mobile-friendliness scores of web pages. These scores can help search engines assign high ranks to mobile-friendly web pages when users search for information with search engines on mobile devices. Through our in-depth usability study, we derived 21 factors affecting mobile-friendliness. These factors are more practical than those obtained in previous works because their values can be calculated by machine. Also, we obtained 300 mobile-friendliness scores from the usability study, and they were used for a training set of MFScorer. Our experimental results show that search results re-ranked by MFScorer's mobile-friendliness scores are more relevant than those re-ranked by mobiReady's scores.

Acknowledgments

We are grateful to the MSRA interns and KAIST graduate students who participated in our usability study.

References

[1] M. Chae and J. Kim, "What's so different about the mobile Internet?," Commun. ACM, vol.46, no.12, pp.240–247, 2003.
[2] M.J. Albers and L. Kim, "User web browsing characteristics using palm handheld for information retrieval," Proc. IEEE Technology & Teamwork Conference, pp.125–135, Cambridge, MA, Sept. 2000.
[3] D.E. Rose and D. Levinson, "Understanding user goals in web search," Proc. 13th Int. Conference on World Wide Web, pp.13–19, NY, USA, May 2004.
[4] B. Boser, I. Guyon, and V. Vapnik, "A training algorithm for optimal margin classifiers," Proc. 15th Annual Workshop on Computational Learning Theory, pp.144–152, PA, USA, July 1992.
[5] C. Cortes and V. Vapnik, "Support-vector network," Mach. Learn., vol.20, no.3, pp.273–297, 1995.