Optimus Web: Selective Delivery of Desktop or Mobile Web Pages

Nádia Fernandes, Tiago Guerreiro, Diogo Marques, Luís Carriço
Faculdade de Ciências da Universidade de Lisboa
Campo Grande, Edifício C6, 1749-016 Lisboa, Portugal
nadiaf@fc.ul.pt, tjvg@di.fc.ul.pt, dmarques@lasige.di.fc.ul.pt, lmc@di.fc.ul.pt

ABSTRACT
We propose a system that selectively delivers the most accessible representations of a website to the end-user. To do so, we resort to a Web proxy that, using the results of an automatic evaluator, changes HTTP requests and delivers the most suitable available representation. The accessibility evaluator uses WCAG 2.0 [1] guidelines and considers pages after browser interpretation. We performed a user study with 19 blind participants where different representations of website were used. Results show that user performance correlates with the computed accessibility levels, and that for different websites either mobile or desktop versions may turn out to be more accessible. Expert users seem to perform independently from the level of accessibility of the pages. These results suggest that a proxy that selectively delivers Web page representations is feasible and that future work should also consider the profile of the user for selection.

Author Keywords
Selective Delivery, Mobile pages, Web accessibility, Automatic evaluation, Blind people

ACM Classification Keywords
H.1.2.Models and Principles: User/Machine Systems - Human factors; H.52.Information Interfaces and Presentation: User Interfaces - Theory and Methods

INTRODUCTION
Users with visual disabilities face significant challenges on the Web. Although research efforts and public policies have tried to ameliorate the situation [2], change in the Web tends to outpace accessibility remedies [5]. A Web page is no longer a static document delivered by a server. Instead, it is a browser's interpretation of both content and logic, at a moment in time. It changes and adapts.

Such advances have made the job of accessibility technologies harder. But new opportunities have also been created. Since Web sites now commonly adapt to mobile browsers (i.e., mobile representation of page), and mobile versions tend to be tidier and more structured [6], they could potentially be more easily interpreted by the assistive technology.

We investigate if this gap could be bridged by forcibly delivering mobile representation to desktop set-ups. Indeed, previous work suggests that mobile representations are not always more accessible [3]. We therefore propose selectively delivering either mobile or desktop versions, depending on the results of an automated accessibility evaluator, designed to interpret Web sites as a browser would (QualWeb evaluator) [4]. We built a Web proxy that automates this process, at the client-side. An experimental study of 19 blind users was conducted to validate this approach. We found that user performance was, in general, better in the Web site representation where less accessibility barriers were found. However, for expert users, this relationship was weak or non-existent.

BACKGROUND
The approach suggested in this work is automatic selection and delivery of the most accessible web content to users. Other approaches aim at the same goal. Transcoding is a mechanism of content manipulation to improves page presentation [10]. For instance, delivering different pages adapted to the size of different devices or simplifying the page's text for dyslexic users [7]. Transcoding can be used to render more accessible pages. However, effective approaches require deep knowledge of page structure, of the way it is used, and preferably of its content [10], which certainly is not possible for a large number of cases.

Hoehl and Lewis [6] suggested a lighter mechanism, based on a proxy, that simply delivers mobile representations of the pages to blind users. The approach is founded on the presupposed simpler stand of those representations. However, a study of Fernandes et al [3] comparing the accessibility levels of mobile and desktop representations, showed otherwise. Different accessibility levels were in fact identified for each representation, but with no clear advantage for either of them.

The idea we present in this paper is to merge the aforementioned projects to selectively deliver the most accessible page to the user. Therefore, the proxy should be able to automatically evaluate the accessibility of a page.

W4A '15, May 18 - 20, 2015, Florence, Italy
Copyright 2015 ACM 978-1-4503-3342-9/15/05…$15.00
http://dx.doi.org/10.1145/2745555.2746668

We acknowledge this as a controversial subject. Recently, Vigo et al. [8], for example, questioned the accuracy of automatic evaluation. They showed that a sample of existing automatic evaluators have a relatively low evaluation accuracy, especially in what pertains coverage and completeness. Still, they find it a valid complementary approach for accessibility evaluation. And an evaluation method per se when others are unavailable or not feasible.

This is the case for the proposal at hand. Moreover, none of the tools used in their study fully considered the Web pages as they would be interpreted by a browser: after computing changes to the HTML source introduced by client-side code (e.g. JavaScript). QualWeb evaluator [4], which we used, takes that into account and verifies their compliance with Content Accessibility Guidelines (WCAG) 2.0 [1]. Expectedly it should render better levels of completeness.

METHODOLOGY

We conducted a user study with 19 blind participants aimed at understanding and comparing their performance, behaviours and opinions regarding desktop and mobile versions of Web pages. Data collection lasted three weeks in June-July 2014. Each participant took part in two sessions, performing the same tasks in three different Web sites. The tasks included selecting an item, using a search mechanism, and identifying an element or a piece of information. In each session, the proxy forced the delivery of either mobile or desktop representations. The order of the sessions was counterbalanced.

Website selection. The Web sites were selected based on a preliminary questionnaire of Web habits across the participant pool. We identified the type of sites participants used and its popularity, including sites that were never used. Based on these we identified candidates that had a desktop and mobile representation.

We then evaluated the accessibility of these webpages. Accessibility levels were calculated with QualWeb and verified via expert evaluation to account for extra barriers. We considered accessibility to be high when the total number of accessibility problems detected were less than 80, medium between 80 and 120, and low more than 120.

Finally we selected three Web sites, with varying degrees of popularity and different level of accessibility in its desktop version: an online newspaper (very popular, high accessibility), an e-commerce site (somewhat popular, low accessibility), and a cooking recipe catalogue (unknown to participants, medium accessibility).

Participants. A total of 19 participants (2 females) were recruited, ages ranging from 26 to 56 years old (M=47.15, SD=10.40), with light visual perception at most. The participants were recruited from a training centre for blind people. Twelve (12) participants were congenitally blind. All had recently finished or were enrolled in a short-term computer skills courses at the centre. The courses are organized in three levels, according to ability. The levels in which participants were placed (5 in advanced level, 11 at intermediate level, 3 introductory users) correlated with self-reported expertise level (Spearman, r=0.906, p=0.000). All participants reported using JAWS as their screen reader.

Apparatus. The study took place in a quiet room, in the training centre where the participants were recruited from. The tasks were performed using a Windows 8.1 personal computer with JAWS 14.0, speakers, and a keyboard with special rubber markers on F and J. The proxy service ran locally in the computer participants used. It was pre-set at the beginning of the session to deliver either mobile or desktop representations. For evaluation purposes, pages were pre-rendered, thus ensuring loading times were similar. Audio and video of the sessions was recorded.

Procedure. In the first session, participants were briefed about the study and its length, knowing that they would be performing tasks on Web sites. They were also administered a preliminary background questionnaire.

In each session, participants were asked to do a specific task in each of the three Web sites. They had to perform navigation, identification, selection and search tasks. A researcher assigned to monitor the test indicated the URL of the site to the participant and the task to perform (e.g., in the newspaper site the user had to find the politics section, find specific news and identify a specific piece of information on the news). The URLs were consistent across different sessions; the proxy determined internally which version to present and redirected accordingly, based on the pre-defined counterbalancing. Participants were not told which version would be delivered.

Participants were asked to think-aloud. Particularly, they were asked to report, in each step of a task, i.e. in each intermediate page, if they experienced any barriers. After each task, the participants were asked for general comments and suggestions. To accomplish all the tasks the users navigated thru 19 webpages (within those 3 websites).

Design and Analysis. The study used a within-subjects design (3x2) where all users performed a task in all Web sites (newspaper, recipes, e-commerce) in both Versions (i.e., desktop and mobile). Shapiro-Wilk normality test was applied to assess the normality of continuous data (i.e., Time). Only the data of the recipes site was normally distributed. Parametric tests were used for normal continuous variables; non-parametric alternatives were used otherwise. The materials and dataset for this study are available[1].

RESULTS

We focus on performance differences while using mobile and desktop versions of a Web page, and how they relate to

[1] http://goo.gl/fQxVye

accessibility level. A strong relationship indicates that delivering the best-score of the representations is likely to improve blind people's efficiency in completing the types of tasks we evaluated.

Mobile pages are not always more accessible. Participants visited a total of 19 webpages, from which 12 (63%) were more accessible in their mobile version, according with QualWeb. These results go in line with those reported in [3] and suggest that the mobile representation is not, in many cases, more accessible. Figure 1 shows the mean task completion time for each site per version. Participants took less time to complete the task in the mobile representations of the newspaper and e-commerce sites. However, they were faster in the desktop version of the recipes site. In this particular case, we observed that the search function in the mobile representation caused some difficulties.

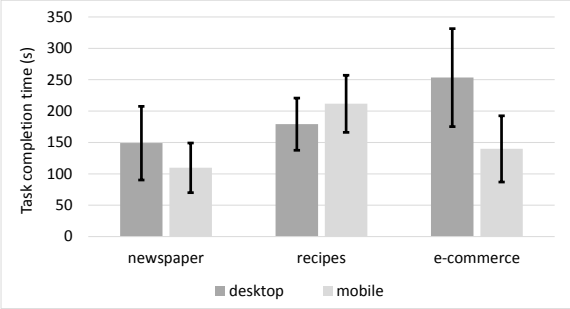

Figure 1 – Average task completion times in each Web site, for the desktop and mobile representations. Error bars represent 95% confidence intervals.

No significant effects were verified in the newspaper and recipes sites. The e-commerce site presents significant differences between the results of mobile and desktop versions (Wilcoxon-test, $Z = -2.54$; $p = 0.011 < 0.05$). This indicates a significant improvement in time performance using the mobile versions for e-commerce site.

Different representations, different accessibility problems. The type of problems that the participants faced the most were: 1) irrelevant content before task content; 2) insufficient feedback to the performed tasks; 3) link destination unclear; 4) no alternative descriptions in lists and tables; 5) heading structure violated; 6) HTML inconsistencies; 7) skip links missing; 8) title of the pages not defined; 9) lists defined without the correct HTML elements; 10) broken links; and 11) users cannot understand and use the content. We observed a pronounced reduction of problems 1, 3, 5, 9, 10 in mobile representations and increasing of 11. Figure 2 presents the total number of barriers found by the users in each version. Only in the recipes site the users experienced less accessibility problems in the desktop version (i.e., existence of problem 11 in mobile version).

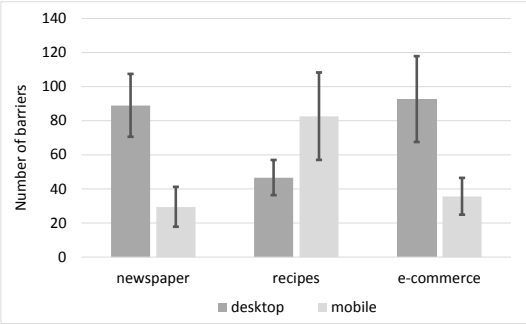

Figure 2 – Average number of barriers in each web site, for desktop and mobile representations. Error bars 95% confidence intervals.

Comparing the barriers detected by the users and the ones detected by QualWeb evaluator, we were able to verify that QualWeb is not yet able to automatically detect problems 9 and 10. The former was particularly serious for newspaper and e-commerce Web sites; at least one of the pages per task has several lists. The latter, affecting the recipes site, is not completely addressed by WCAG 2.0 guidelines.

Expert users avoid accessibility barriers. The most used navigation strategies by expert users were: trying to locate a specific link or header in the list of elements; trying the virtual search or other shortcut available from the screen reader; and if those strategies did not work, trying to navigate through the elements. Expert users were able to avoid accessibility problems and were 20-65% more efficient than non-experts (i.e., medium and basic users). Basic users pass through almost all possible elements in the page, until they find the one they are searching. They encountered, approximately, 70% of the accessibility barriers. Still, they perform 14-30% more efficiently than medium users in desktop representations. In mobile representations, basic users performed 30-40% less efficiently than medium users. Medium users applied a mix of basic and experts users' strategies. Since they are not completely proficient in the aforementioned navigation strategies, they had to perform extra tries to be able to perform them, and sometimes without success. They were disturbed by approximately 47% of accessibility barriers.

Accessibility barriers influence time performance. Differences in performance and number of barriers found by different levels of users were verified. No correlation between the total number of barriers and performance was established for basic users, taking into consideration the reduced number (3). For medium level users the total number of accessibility problems and performance showed to be correlated (Spearman, $r = 0.502$; $p=0.048 < 0.05$). Experts present no relation between performance and the number of barriers.

IMPLICATIONS OF THIS RESEARCH

Client-side selective delivery allows for more accessible content delivery. Initial results show that accessibility is relatively balanced between mobile and desktop representations in all the 19 pages used by the participants. Looking at performance, mobile representations are in general more efficient and contain less barriers. This is probably related with the reduction of content in those pages. An exception rises when in the desktop representations the extra content can be a guide to the task. This strengthens the argument for the selective delivery of representations to the user. We can claim that the distribution of the more suitable representation can be decided "on-the-fly" (by a proxy) before page presentation.

Provide personalized delivery. Considering the different navigation strategies of the users the page selection should address the user level. For instance: basic users may need of more content in the page to understand the tasks and decide the navigation approach, so they can prefer desktop pages in most situations. On the contrary, expert user may prefer mobile representations with less content to navigate.

Better understanding of coping mechanisms. The better performance of the expert participants and the lack of correlation with the total number of barriers show that experts are skilled to overcome and avoid accessibility problems. They can navigate through the page without noticing the majority of the problems, as already found by *Vigo and Harper* [9]. In opposition, non-expert users (i.e., medium and basic users) experience those accessibility barriers. This way, it is an important asset to enable non-experts users to acquire those navigation strategies faster. Medium level users seem to be acquainted with shortcuts and strategies but they fall short in applying them correctly which leads to even worse performance. This way, instead of non-experts users having to learn by themselves, a cooperative method with the knowledge of experts could be developed. Additionally, the knowledge of the navigation strategies could help developers/designers to understand how and where to introduce the more important parts of the pages so that they are not missed or jumped over by the expert users.

Improvement of accessible design and evaluation guidelines. The sessions performed allowed to understand the key barriers that disturbed users in each of the representations. Developers/designers should address carefully those problems. In addition to WCAG techniques new heuristics can be added to improve accessibility conformance. This is particularly relevant in the case where the guidelines do not completely address the problems verified, as also mentioned in [9].

CONCLUSIONS/FUTURE WORK

This paper presents a study to understand and compare differences of performance and behaviour of 19 blind participants using different versions (i.e., mobile and desktop) of three Web sites. We provided facts of: 1) selectively delivering mobile and desktop version of the Web sites can leveraged to improve accessibility and thus performance; 2) users' level should be considered; and 3) new guidelines and heuristics should be introduced to help users and developers/designers. Moreover, we are able to recommend sharing navigation knowledge of expert users with non-expert users and that developers/designer should consider those strategies in the page design/development. As future work, longitudinal studies are required to assess long term benefits and user acceptance of the approach. Further, the user ability level showed to be relevant and user models should be included in the selective delivery process.

ACKNOWLEDGMENTS

We thank Fundação Raquel and Martin Sain and all the users that participated in these studies.

REFERENCES

1. Caldwell, B., Cooper, M., Reid, L.G., and Vanderheiden, G. Web Content Accessibility Guidelines (WCAG) 2.0. 2008. http://www.w3.org/TR/WCAG20/.
2. Comission, E. eAccessibility - Opening up the Information Societ. 2008. http://www.eaccessibility-impacts.eu/.
3. Fernandes, N. and Rodrigues, A. and Hijon, R. and Duarte,C and Carriço, L. Web Accessibility of Mobile and Desktop Representations. *Proc. of BCS HCI* (2014).
4. Fernandes, N., Costa, D., Neves, S., Duarte, C., and Carriço, L. Evaluating the accessibility of rich internet applications. Proc. of International Cross-Disciplinary Conference on Web Accessibility (W4A '12). ACM (2012).
5. Harper, S. and Yesilada, Y. *Web Accessibility*. Springer, London, 2008.
6. Jeffery Hoehl, C.L. SimpleWebAnywhere: Mobile Web on the Desktop. *Proc. of CSUN*, (2012).
7. Rello, L., Baeza-Yates, R., Bott, S., and Saggion, H. Simplify or help? *Proc. of W4A*, ACM (2013).
8. Vigo, M. and Brown, J. Benchmarking Web Accessibility Evaluation Tools : Measuring the Harm of Sole Reliance on Automated Tests Categories and Subject Descriptors. *Proc. of W4A*, ACM (2013).
9. Vigo, M. and Harper, S. Coping tactics employed by visually disabled users on the web. *International Journal of Human Computer Studies 71*, 11 (2013).
10. Yesilada, Y. and Harper, S. Experiential Transcoding : An Eye Tracking Approach. *Proc. of W4A*, ACM (2013).