USABILITY REMOTE EVALUATION FOR WWW

Marco A. A. Winckler          Carla M. D. S. Freitas          José Valdeni de Lima

Universidade Federal do Rio Grande do Sul, Instituto de Informática,
Caixa Postal 15064 - 90.450-071 Porto Alegre, RS, Brazil
{winckler,carla,valdeni}@inf.ufrgs.br

ABSTRACT

In the last few years many remote methods have been proposed for usability evaluation. For evaluation of WWW interfaces, remote methods present some advantages over traditional ones. In general, they are cheaper and faster than the traditional ones due to the fact that neither users nor evaluators need to move to take part in the tests. However, they fail to provide more accurate data for usability analysis. This paper presents a faster and low cost method for WWW evaluation that provides better data for analysis. The novelty of this method is due to the merging of tracking of users' actions and subjective information provided through questionnaires in the same evaluation tool.

Keywords

Remote usability evaluation, WWW interfaces.

INTRODUCTION

The first experiences in evaluating WWW interfaces have applied traditional usability methods, inspection methods and user testing in usability laboratories [3]. These methods are useful but they are too expensive to be applied in most WWW projects [1].

Some studies have pointed out remote usability evaluation methods [2] as a solution for the WWW. The success using remote methods is due to two important factors: they reach real users remotely located and consider the WWW idiosyncrasies. The WWW interfaces are quite different from other types of interfaces because of their contents, design and technology are changing everyday. So, fast and low cost methods are needed to cope with that evolution. In general, remote methods are cheaper and faster than the traditional ones because few equipments are required and the users can stay at home or at their workplaces while taking part in the experiments.

The main problem with the remote methods is that the users and evaluator are often located at different sites. Although automated tools could get users interactions, it is very difficult to know what the users are trying to do. So, such remote methods may be successful or not, according to the tools used to gather user's action and comments.

Perhaps the most suitable remote instrumentation for the WWW would be log files from WWW servers and questionnaires. Sullivan [5] stressed out an inferential method using WWW log files that reveals useful but lacking to accomplish some real users' tasks. In contrast, data from questionnaire [6] are useful, but it is a hard task to identify usability problems with them because they fail to provide more accurate data. Looking for fast, low cost, and easy-of-use evaluation method for WWW interfaces, we have developed a method that combines questionnaire and automated gathering of user interactions. The goal of this method consists in getting data from real users while they are working in some specific tasks remotely.

REMOTE USABILITY TESTING METHOD

In this method, the questionnaires are both used to lead the users to complete some task, which the evaluators want to study, and to get subjective opinions about the interface.

Since the method is based on information collected during user tasks, it needs special tools. So, we have developed some tools that work together with the WWW servers, reading the log files.

The evaluation process is organized as follow. When a user loads a WWW interface he/she is invited to do some tasks using that interface. If the user agrees, it will be presented some questions while his/her interactions over the interface are captured. The questions must address real user's task or some part of the interface that must be analyzed. For example, if the evaluator wants to know how easily the users can find a phone number, the question must ask them to do that using the interface. While he/she is looking for the phone number, the navigation is recorded in a WWW log file. After the user has sent his/her answer, the evaluator can analyze the success or failure of his/her actions to complete the tasks. The evaluator knows the user's goal (complete the task), so he/she can identify some browsing patterns. He can also check if the user is doing the correct actions to complete the task or going in a wrong way, or better yet, if nothing has been done about the task.

In order to persuade the users to take part in the evaluation, they can receive an electronic bonus, as a software registration, privileged information, or other similar electronic gift.

We have built three tools that work together with the WWW log server using Java technology (see Figure 1). They are a monitor, a test manager and a visualization tool.
The monitor supports dialog with the users presenting the tasks to them and getting their inputs. The Test Manager coordinates all monitors that are running in parallel. The visualization tool organizes the gathered data for the analysis process. The users' interactions over the subject interface and over the monitor presenting the tasks are stored separated, but the visualization tool unifies them.

![Diagram](image)

Figure 1. Tools and communication structure.

We can summarize the method in steps:

- the evaluator select some tasks to investigate;
- he/she launches the tools giving the task list as a parameter for the test manager tool;
- the users are invited to take part in the test. If they agree they can get an electronic bonus;
- while the users are taking the test, data is collected and stored in the two logs;
- after getting enough user sessions, the data is analyzed using the visualization tools.

The analysis can be done looking for which links the user passes to complete each task as well as their answers for the questions posed to them. When the task requires a user's answer, the evaluator must compare if the user reached the right page for the question and if the answer was correct.

To identify usability problems we use metrics such as navigation patterns (cyclic, random or sequential navigation), time between interactions and number of jumps.

We assume that in well-designed interfaces the users can quickly reach his/her goals passing through few links. If there are cyclic patterns the evaluator must infer that something is going wrong; the users did not accomplish the proposed task or a usability problem happened. Random navigation could have the same cause as a cyclic pattern, but also the user might be looking for other information than the answer for the proposed question. Comparing the users' answers in two or more tests the evaluator can infer what really happened.

**FINAL COMMENTS**

Our current tools suffer from cache and proxy effects. These problems were partially solved asking the user to turn off cache and proxy access while taking the test. In the future we must use more accurate gathering tools free from these undesirable effects.

In fact, this method is not too efficient when compared to the traditional methods such as user testing on usability labs. Otherwise, regarding the current log analysis methods this approach reveals better results, because the evaluators can infer what the users are trying to do to complete a task.

This method was applied in a case study (the evaluation of our Institute website) where we could identify some bad navigation patterns associated to usability problems [7]. By using it we could find 25 different problems such as too many links to complete a simple task and pages that could not be reached by the presented links.

This method was originally proposed for WWW evaluation as part of a methodology for remote evaluation [8], but it can be adapted to other kinds of remote testing.

**ACKNOWLEDGMENTS**

We thank Marcelo Pimenta for the helpful comments, and CNPq, the Brazilian Council for Research and Development, for supporting this work.

**REFERENCES**


