WEBUSE: AN APPROACH FOR WEB USABILITY EVALUATION

Alexander Nikov  
Fatih University, Turkey  
nikov@fatih.edu.tr  
Permanent address: Technical  
University of Sofia, Bulgaria  
nikov@tu-sofia.acad.bg

Svetla Vassileva  
Technical University of Sofia, Bulgaria  
vassileva@tu-sofia.acad.bg

Sofia Anguelova  
Technical University of Sofia, Bulgaria  
anguelova@gmx.net

Sylvia Tzvetanova  
Technical University of Sofia, Bulgaria  
tzvetanova@tu-sofia.acad.bg

Stefka Stoeva  
Technical University of Sofia, Bulgaria  
stoeva@gmx.net

SUMMARY
An approach for web site usability evaluation (WebUse) is developed. It uses data collected from objective web site usability evaluations like physiological parameters, time, etc. and from subjective parameters evaluated by usability questionnaires. The objective data are measured in interval scale, as the objective data in ordinal or nominal scale. For aggregating these mixed data a neuro-fuzzy model is employed. WebUse approach is implemented in a MS EXCEL software tool. A case study for evaluation of a web site by this approach is carried out. The analysis of experimental results showed that the approach supports the analysis of web sites and allocation of usability problems. A comparison of WebUse approach with a conventional evaluation approach showed significant better performance for WebUse. The main advantage of WebUse is the combination of objective and subjective usability parameters which improves the reliability of the usability evaluations and thus more usability-related bottlenecks of web sites can be discovered. Further advantage of WebUse is its ability of handling missing data for some usability criteria.

1 INTRODUCTION
The development of good quality web pages requires sophisticated methods for design and evaluation. Using practical methods for testing web site design may reduce the development costs while integrating usability. Most of the approaches evaluate the usability by questionnaires. Based on analysis of results they give recommendations for web design improvement. A number of tools and techniques
support the evaluation of the usability of web sites: Questionnaire for User Interface Satisfaction [1]; Perceived Usefulness and Ease of Use [2]; Nielsen's Attributes of Usability [7]; Nielsen's Heuristic Evaluation [8]; Computer System Usability Questionnaire [5]; After Scenario Questionnaire [4]; Practical Heuristics for Usability Evaluation [10]; Purdue Usability Testing Questionnaire [6]; etc. [12]. These tools are based only on subjective evaluation of usability by checklists. In order to increase the reliability of evaluation results an approach for web usability evaluation (WebUse) combining the subjective with objective usability evaluation is proposed.

2 DESCRIPTION OF WebUse APPROACH

WebUse approach combines results from objective experimental usability studies (mostly on-line data) and from subjective usability evaluation by usability checklists (off-line data). The usability evaluation of the web site is done on the basis of results obtained during and after execution of specific pre-defined representative tasks with interactive web systems. A web usability index is determined on the basis of objective and subjective evaluations (cf. Fig. 1 and Table 1) as follows:

1. Objective evaluation: measurements and information gathered for:
   - User's Background,
   - Physiological Parameters,
   - Interactions,
   - Environment Parameters,
   and

2. Subjective evaluation: Users opinions filled in the relevant checklists for each of the following subjective criteria:
   - Navigation Quality,
   - Information Quality,
   - Interface Design Quality,
   - Overall System Quality and
   - User's Attitude.

Based on hierarchic structure proposed the objective and subjective data gathered are aggregated to quantitative web usability index by a hybrid model - the QuickFBP algorithm [9], which combines fuzzy logic with neural networks. Important advantage of this model is that it calculates the usability index when: 1) the data are measured in scales of different type like interval scale (objective data) and ordinal/nominal scale (subjective data); 2) some of the data is missing, which could happen especially for objective measurements. For supporting the practical application of WebUse approach a MS EXCEL software tool with user-friendly interface was developed. After entering the data gathered the program computes the web usability index.

3 EXPERIMENTAL STUDY

The WebUse approach was experimentally studied by testing and evaluating the web site (MFsite) of the Motley Fool company, delivered within CIFter Project funded by NIST (National Institute of Standards and Technology). The test website is a frozen-in-time original version of the Motley Fool website. This is a self-contained archive: Links to external sites have been replaced with local links. The dynamic elements found in the tested Motley Fool site have been replaced with static pages. This static snapshot contains about 10,000 pages.

3.1 Test participants

The test participants were 30 undergraduate students at the Technical University of Sofia. The average age of the students was 22.6 years. The number of men was equal to the number of women. Their experience in using web sites, according to their answers in the checklists was from "missing" (2 persons have visited a web site for the first time) to "good" (about half of the students have used Internet more than two years). The average for whole the test group was 4.4 years computer experience and 2.1 years...
Internet experience. 35% of the students affirmed that they prefer reading literature before using an unknown software product, including interactive systems and web sites. More then one-third of the test participants were not personally interested in the MF site domain of “investment strategies”, but found the test site interesting. The English knowledge of the majority of the test participants was good enough to perform the tasks.

Figure 1: WebUse evaluation hierarchic model

3.2 Test environment
The experiment took place in a computer laboratory of Technical University Sofia. The difference to the real working environment, where the user is at his/her own working place or at home, was that the working space was greater and the number of present persons was higher.

3.3 Testing procedure
The participants received documents containing the explanations of the tasks to be carried out with the tested web site and the checklists to be filled in after completing the tasks. The tasks were the same for all participants.

Table 1: WebUse hierarchical structure

<table>
<thead>
<tr>
<th>Level 0</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1. Gender</td>
<td>1.2. Age</td>
<td>1.3. Computer experience / Internet experience, skills</td>
<td>1.4. Learning style</td>
</tr>
<tr>
<td>1.1.1. User’s background</td>
<td>1.2.1 ECG</td>
<td>1.3.1. Access time</td>
<td>1.4.1. PC configuration</td>
</tr>
<tr>
<td>1.2. Physiological Parameters</td>
<td>1.2.2 EMG</td>
<td>1.3.2. Successfully performed tasks / Task performance</td>
<td>1.4.2. Screen / Monitor</td>
</tr>
<tr>
<td>1.2.1. ...</td>
<td>1.2.3 Respiratory signals</td>
<td>1.3.3. Errors</td>
<td></td>
</tr>
<tr>
<td>1.2.1. ...</td>
<td>1.2.4 Other</td>
<td>1.3.4. Number of interactions</td>
<td></td>
</tr>
<tr>
<td>1.2.1. ...</td>
<td></td>
<td>1.3.5. Ratio total access time / errors</td>
<td></td>
</tr>
<tr>
<td>1.3. Interactions</td>
<td>1.3.1. ...</td>
<td>1.4.1. ...</td>
<td></td>
</tr>
<tr>
<td>1.3.1. ...</td>
<td>1.4. Environment</td>
<td>1.4.1. ...</td>
<td></td>
</tr>
<tr>
<td>1.3.5. ...</td>
<td>1.4. ...</td>
<td>1.4.5. ...</td>
<td></td>
</tr>
<tr>
<td>1.4. Environment</td>
<td>1.4. ...</td>
<td>1.4.5. ...</td>
<td></td>
</tr>
<tr>
<td>1.4. ...</td>
<td>1.4.5. ...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 3.3.1 Test tasks

The usability of the web site was measured and evaluated by three tasks for finding specific information within the site. The first task "MAIN PAGE" demands overall orientation in the site structure. The second task "DETAIL" requires a kind of "in depth" study of the site: Does the main page function well? Does it function well enough to come „deeper“ into the site? The third task "BROKER" consists in finding a specific function on the web site pages. The complexity of the three tasks was approximately equal. On Fig. 2 and Fig. 3 are shown two screenshots of the evaluated web site.
3.3.2 Questionnaires

The questionnaires used in our study are based on standard questionnaires like Post Test Questionnaire [11], QUIS (Questionnaire for User Interface Satisfaction) [1] and NAU (Nielsen’s Attributes of Usability) [7].
4 EXPERIMENTAL RESULTS

WebUse was applied to most of the evaluation criteria for the web site evaluation except the physiological criteria. Currently we had not technical possibility to study them, but it is planned for our further studies. Each set of ratings for every usability criterion was estimated by a concordance coefficient. The outliers were eliminated.

All data collected was processed by MS EXCEL software tool implementing the WebUse approach. On Table 2 are shown the summary values determined by WebUse: the values calculated by the conventional weighted mean approach and the values of experts’ evaluations.

The most important objective evaluations are as follows: The total maximum time for performing the test was 45 minutes: 15 min for completing the tasks and 30 min for filling in the checklists and giving additional background information. The effectiveness was determined by the percentage of the participants who have solved successfully the tasks: 94%. The efficiency was determined by the time for solving each task successfully: 5-14 minutes. The average access time for all three tasks was 9,4 min. The average number of errors was 4,4. The average number of interactions for the performed tasks was 13,7. The ratio between the total access time for the three tasks and the errors registered for this time interval was 2,8 min/error. The lowest evaluation values of objective criteria (cf. Fig. 4) are for user’s background (2.1) followed by environment evaluation (3.0). The reasons are the usability problems: insufficient Internet experience, low speed of Internet access and the comfort on the workplaces.
Table 2: Evaluation results within the interval scale [1, 5], where 1 is the worst and 5 is the best rating.

<table>
<thead>
<tr>
<th>Usability Criteria</th>
<th>WebUse Evaluation</th>
<th>Weighted Mean Evaluation</th>
<th>Expert Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level 0</td>
<td>Level 1</td>
<td>Level 2</td>
</tr>
<tr>
<td>1. OBJECTIVE EVALUATION</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1. User’s Background</td>
<td>2.10</td>
<td>3.28</td>
<td>4.58</td>
</tr>
<tr>
<td>1.2. Interactions</td>
<td>4.91</td>
<td>4.21</td>
<td>4.50</td>
</tr>
<tr>
<td>1.3. Environment Parameters</td>
<td>3.00</td>
<td>3.41</td>
<td>3.00</td>
</tr>
<tr>
<td>2. SUBJECTIVE EVALUATION</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1. Information Quality</td>
<td>3.06</td>
<td>3.02</td>
<td>3.58</td>
</tr>
<tr>
<td>2.3. Interface Design Quality</td>
<td>3.15</td>
<td>3.58</td>
<td>4.58</td>
</tr>
<tr>
<td>2.4. Overall System Quality</td>
<td>3.65</td>
<td>3.23</td>
<td>3.25</td>
</tr>
<tr>
<td>2.5. User’s Attitude</td>
<td>2.96</td>
<td>2.35</td>
<td>3.92</td>
</tr>
</tbody>
</table>

Figure 4: Objective evaluations by WebUse

The lowest subjective evaluations (cf. Fig. 5) are for the criteria user's attitude (2.96), Information quality (3.06) and Interface design quality (3.15). Regarding user’s attitude usability drawbacks of the web site are insufficient joy of use, satisfaction and comfort of use. Regarding the information quality the bottlenecks are presence of errors and incorrectness, and insufficient information highlighting. In relation to interface design quality the drawback of the web site concerns improper use of blank spaces. Generally the subjective evaluation is lower than the objective evaluation (cf. Fig. 6). Here are the most usability problems requiring relevant design improvements.
A comparison of the performance of WebUse approach with the conventional weighted mean approach was carried out (cf. Fig. 7). The accuracy of both approaches was calculated based on expert evaluations. WebUse approach showed significant better accuracy (96%) than weighted mean approach (86%). Also the subjective WebUse evaluation (96%) is significant better than weighted mean evaluation (87%) (cf. Fig. 8). The objective weighted mean evaluation (91%) is non-significant better than WebUse objective evaluation (88%) (cf. Fig. 8).
5 CONCLUSIONS

WebUse approach integrates objective data with subjective data for improving usability evaluation reliability. It is based on a hybrid neuro-fuzzy model for aggregating usability data gathered. The experimental study confirmed the following advantages of WebUse:
1) Ability of aggregation of usability data measured in different scales;
2) Ability to handle missing data for some usability data;
3) Significant higher accuracy in comparison with conventional weighted mean approach and thus higher reliability of evaluation results for allocating usability problems.

As we were not able to collect physiological data, the inclusion of this kind of data may result in better performance of WebUse. The application of tools like ErgoLight [3] could improve the collection of logging data.

WebUse approach presents a step toward creation of a common industry format for usability studies.

ACKNOWLEDGMENT

The present study is related to the international project CIFter: Common Industry Format for reporting the results of usability studies with users and Testing of usability Evaluation Reports, funded by U.S. National Institute of Standards and Technology (NIST), http://zing.ncsl.nist.gov/cifter/

REFERENCES