Web Mining for Usability Evaluation

Eduardo Zamora[1], Cristian Rusu[1], and Silvana Roncagliolo[1]

[1] Pontificia Universidad Católica de Valparaíso, Escuela de Ingeniería Informática
Av. Brasíl No. 2241, Valparaíso, Chile
eduardo.zamora@gmail.com, {cristian.rusu, silvana}@ucv.cl

Abstract. Web usability evaluations are critical in a website development process, and many techniques have been developed to assist these evaluations. This paper shows a rather new approach to usability evaluations, based on Data Mining and Knowledge Database Discovery. It takes classical and well known usability evaluation techniques and complements them with Web Usage Mining, in order to create a well balanced and complete evaluation process ideal for websites. A case study proves the efficiency of the new approach and reveals practical information and conclusions.

Keywords: Web usage mining, Web usability, Usability evaluation.

1 Introduction

The key for building a successful website is what everyone is searching now. Usability seems to be one of the most promising solutions. Usability is a concept that has been around for many years but, just as global warming, it has been proved to be an important issue only now, when we already are in chaos.

Developing usable web sites may be difficult, since there is no well defined way to do it. There are well recognized parameters, guidelines, rules and recommendations that can help us in our process, but all these cannot guarantee a usable site. This is what makes usability testing and evaluation so important, the need to have real feedback from who matters the most, the users.

The paper presents an approach that integrates classic evaluation techniques such as user testing, card sorting, heuristic evaluations and questionnaires and complements them with a relatively new technique provided by the field of data mining, *Web Usage Mining* (WUM). With this approach a more complete web usability evaluation is possible combining data collected directly from users and usage data collected from the server. With some theory and a case study we will intend to show how WUM improves the reach and efficacy of web usability evaluations in finding usability problems.

In the present paper we show an integrated approach of classic techniques and WUM, in contrast with other papers and studies where WUM is usually used as a stand alone activity to find usability problems or with the objective of finding users preferences for commercial purposes [1], [6], [9], [14]. The current state of Web Usage Mining and the Web itself does not allow it to be 100% trustworthy since there

XXXIV Conferencia Latinoamericana de Informática (CLEI 2008), pp. 619-628, 2008.

are still shortcomings that will be presented in this paper [15]. Even so, mining proved to be accurate enough and very useful in usability evaluations.

2 Web Usage Mining

Web Mining can be defined as the discovery of useful information on World Wide Web through data mining techniques [7]. There are three types of information that can be analyzed, giving place to three different areas of Web Mining: *Web Content Mining* (WCM), *Web Structure Mining* (WSM), and *Web Usage Mining* (WUM). In WCM we analyze the content of the site, meaning the textual information. In WSM we analyze the structure of the site, meaning the hyperlink configuration and the connection between pages. Finally, in WUM we analyze the information about the interaction between the user and the website, available in server log files.

WUM is the process of discovering and interpreting user access patterns in web information systems, mining the data collected from their interaction with the system [8]. This is a very common definition, but it may fall a little short since patterns is not all that can be found, and in the current state of the art significant patterns are one of the most difficult things to find. Both tools used in the present case study revealed very few usage patterns and yet they were very useful. There is currently little free software capable of delivering trustable usage patterns in an easy way.

Nowadays the viability of WUM is still being argued and criticized, mostly because how inaccurate server log file data can be. So in this paper there will be no over or underestimating of WUM capabilities, instead there will be real results from real evaluations.

Typical applications of WUM have actually little to do with usability. Most commercial applications today pursue commercial goals, to attract more users and to make the most of the existing ones. To achieve these goals WUM provides a way to study the website users, finding out their preferences, habits and (maybe with good luck) goals. In other words, WUM is inserted in a marketing environment and CRM (*Customer Relation Management*). Examples of such applications are *WebTrends*, *AlterWind* and *Clicktracks*.

Usability may help to achieve the same goals, by enhancing the user experience, but the main focus is different. By designing a usable system developers expect to satisfy users needs better and more efficiently, and thus users will use the system more and faithfully. The focus is on the user and not on the system, and WUM can help by delivering real usage from real users of the site.

3 Website Usability Evaluation

Usability is defined by ISO as "*The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"* [12]. In this case the product is a website, the users are the visitors, and the context of use is of course… online (Intranet may be

included). A less technical definition characterizes a usable website as the one where users can find what he needs fast and effectively, always with satisfaction.

With the growing importance of online activities such as transactions, information search, teaching and learning, tax declaration and others, a number of new concepts like e-learning and e-government have emerged. The "e-life" is becoming more present and its consequences affect not only businessmen and the CEOs (*Chief Executive Officer*), but also common people: older people that will have to submit an electronic vote (e-voting), farmers that will have to sell their products online to find better markets, housewives shopping online, etc. When we realize that the web is not only for the literates anymore, it becomes clear that we must focus the design on users, to reach them all, and good usability is a key factor to have in mind.

Usability evaluations can be done within the entire lifecycle of the site. There are plenty of usability evaluation methods to go around; an extensive list can be found in [4]. Evaluations can be divided into two types: inspections and tests.

Usability inspections usually have several experts reviewing the site in order to find potential usability problems. Inspections are useful to find high level flaws in design rules and consistency; they can also be cheaper and easier to do [3]. Most common inspections are heuristic evaluations and cognitive walkthrough. Tests usually involve real life users trying out the site under specific conditions, doing certain tasks. The interaction between the users and the system (the website in this case) is recorded and analyzed, in order to find usability problems. Testing works well to find specific domain problems and workflow issues.

The need for better and more complete evaluations has motivated the creation of automating software for usability evaluations. A good review of the state of the art in this area can also be found in [4]. In the automation of collecting data from the user-website interaction Web Usage Mining appears as an important candidate.

4 Web Usage Mining and the Usability Evaluation

We have tried to integrate classical methods for web usability evaluations (usability tests and inspections) with a rather non-conventional technique (WUM), with the purpose to find a balance and complement, which could make the evaluation process more complete, bringing developers better and more accurate results. That is why we used usability tests (with tasks and users), card sorting, heuristic evaluation, questionnaires and WUM. These five activities cover together most aspects of websites such as content, structure, navigation, flow, layout, and others [5].

As mentioned above, WUM uses web server logs to find information, but server log file credibility is still under heavy judgment, and various opinions can be found in this respect [2], [5], [11], [13]. The main two views in this subject are the following:

- Server access logs have some shortcomings, but they bring enough accurate data to implement solutions based on the information mined from them.
- Server access logs shortcomings render them useless for mining information, and any information extracted will be inaccurate and not truthful enough.

Discovering which point is correct or where the truth lies is one of the main motivations behind the investigation and work presented here. By this point, and with the data collected from the evaluations, the conclusion is that the truth lies somewhere in between, since log files were not useless, but also not too accurate. Because of this, it is recommended to always insert WUM in a bigger and more complete process of usability evaluation, complementing other activities and no as the only activity to perform.

Having this in mind, an important question rises: how exactly can WUM be used in a web usability evaluation? We propose the following approach:

- WUM must be performed before the rest of the activities, since its results may help to guide the evaluators:
  1. By developing questions that can be used in other activities,
  2. By developing hypothesis that can be proved in other activities.
- The result of all evaluation activities will find a way to complement in the analysis stage, being WUM of great importance to complement the results:
  3. By helping prove hypothesis developed in other activities,
  4. By delivering unique data that can help improving the site in technical aspects (such as uptime, accessibility and speed).

To sum up, evaluators should begin the usability evaluation by analyzing the logs with WUM and then study the results, since they may reveal problems directly and also help the evaluators design or focus the following evaluation activities and questionnaires. This approach may seem fairly obvious, but most published cases study show a CRM-study approach, or a stand alone WUM usage to find problems based on assumptions.

5 Case Study: The Website of Pontificia Universidad Catolica de Valparaiso

The website of Pontificia Universidad Católica de Valparaiso (PUCV, www.pucv.cl) was analyzed. PUCV is a Chilean university with 80 years of existence and over 14.000 students currently enrolled, being one of the main universities of the country. The website is structured to be a portal, linking the different areas of the university and main related and partner sites in one main interface.

The activities done in this usability evaluation process are the ones described previously and the results will be presented summed up and summarized, showing only the more relevant data. The evaluation process involved 34 participants with ages between 18 and 26. Usability tests and card sorting counted each with entry and exit questionnaires to collect demographic data from users as well as user's opinions, feelings and ideas. Heuristic evaluation was performed by 4 evaluators.

The logs analyzed involved the accesses recorded between September 2 and October 4 of 2007, and the software used to mine the logs were *AterWind log*

analyzer professional (evaluation version) and *Analog*. The main data about hits shows:

- Total hits: 14.568.556
- Total visitors: 289.644
- Total unique ips: 108.997
- Total transfered data : 123,386 GB

This shows a high number of visits and a good amount of activity on the server. The results also showed a clear tendency in relation to when the visits occur. Most visits occur Monday around 4 pm and the hit count decreases as the week advances.

The top visited areas are shown in Fig. 1. The top searched items were (being Google the top search engine used): (1) *pucv*, (2) *universidad catolica de valparaíso*, (3) *percentil*, (4) *ucv*, (5) *estadistica descriptiva*.

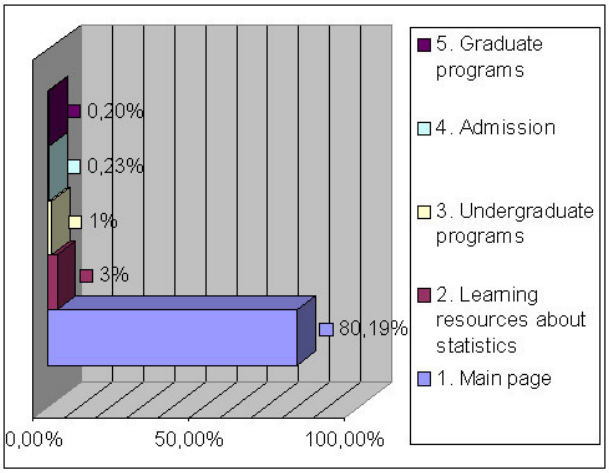

Fig. 1. Top visited areas of PUCV website.

The above mentioned data may help to better understand the use that the site receives. Other relevant data will be presented along with the usability problems found, to show how this data helped the evaluation.

5.1 Usability problems found performing only WUM

1. Accessibility problems when accessing through search engines

The problem was found looking at two relevant data from WUM, the second item on the top areas visited ("*Learning resources about statistics*") and the third item on the top searches ("*percentile*" which means "*percentage*"). These two items correspond to an area containing statistics learning resources for teachers. The high position of this item makes you wonder why statistics learning resources contents are so demanded in comparison with other items like "*admission*", and the answer presented itself clearly after searching the item in a search engine like Google. The

mentioned area of the site appears within the first results whenever you make a search related to that content (in Spanish of course).

Now, being first on web search results is far from being a problem, it is actually a good opportunity except when that search leads the user to a dead end page isolated from the rest of the site and with no identification of where the user is. So, even if the user likes the content presented in the page there is no easy way to access more content or navigate through the site. We all know dead end pages should be avoided, but that was not exactly the problem. The explanation behind the dead end page is frame usage. The statistics learning resources area is built with frames, and when the engine finds the content, it returns only the frame that contains it, stripping the page from the navigational frame and the upper "identification" frame.

2. *Actualization problems when accessing through search engines*

Similar to the previous one, this problem has to do with search engines and it was found analyzing WUM data that may seem out of place or strange. A common search topic was also "*physics*". Searches led to a specific page that was unexpectedly high in hit count. This page, just as the previous one from problem 1 was stripped from all navigation devices but a home link in the bottom that led to problem 2, a fully functional version of the site from 5 years ago. Users could have been navigating this old site for a while before they realized that the information presented in here was old and obsolete. This problem also appeared in tests with users, when they used the internal search of the page.

3. *Problem with server load and data transferred*

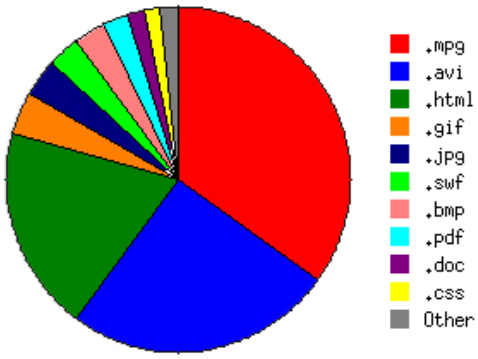

Fig. 2. Server traffic by file extension.

This may not be strictly a usability issue, but it concerns the server and thus it concerns the site. WUM showed that 60% of the data transferred by the server was video files (.mpg, .avi) and most users interviewed admitted they didn't know about the existence of video files in the site (see Fig. 2). The site has some videos for faculties and programs presentation, and one for PUCV introduction. The videos are heavy and they are not visible online, so if a user wants to see them, he/she has to

download them. With the means provided by the new technology this old kind of video presentation has no excuse and just wastes bandwidth, space, and time.

5.2 Usability problems found complementing WUM and specific usability evaluations

In contrast with the previous usability problems, these were found by comparing and complementing results of all activities including WUM. The reader may notice that the first problems are more technical related and the problems that follow are more interaction related.

1. Accessibility problems for Firefox users

This issue was found contrasting data from the questionnaires with data from WUM. Questionnaires showed that 50% of the interviewed users used *Firefox* and half of them used it as their main browser, but data from WUM showed that percentage of visits with *Firefox* is only 12%. Why didn't the rest of the users use *Firefox* to access the site? Well, one option is clear if you look at the images in Fig. 3 and 4. Using *Firefox* critical information cannot be seen!

2. Problems with the search tool

A large number of usability problems were found regarding the internal search tool of the site, this problem appeared in heuristic evaluation, testing and questionnaires and we can summarize them, saying that it is confusing, malfunctional, frustrating and misleading. Usability testing with users showed that from all the search options presented, users tended to select "*Advanced Search*" which didn't actually allow to search, but instead taught users uninteresting information about web searches. WUM here allowed confirming the problem by showing that "*Advanced Search*" was one of the top 15 areas accessed and it was far up from the other search options.

3. Problems with search in the library section

This problem also arose in various activities and showed that internal search in the library section was very deficient. WUM shows that this section is within the top ten hit areas of the site, yet it had one of the more problem-intense tasks on the tests. Users took an average of 5 minutes to find a book and only 67% were able to find it. Having a top ten area being one of the most troublesome ones is definitely something to work on.

4. Problems with site structure and navigation

Various issues found on the evaluation pointed to this problem. The main impression that users stated about entering the site for the first time is confusion. There is just too much information displayed in one single interface and it has no hierarchy or clear visual structure. The navigation is not guided and a visitor can easily get lost in the mess.

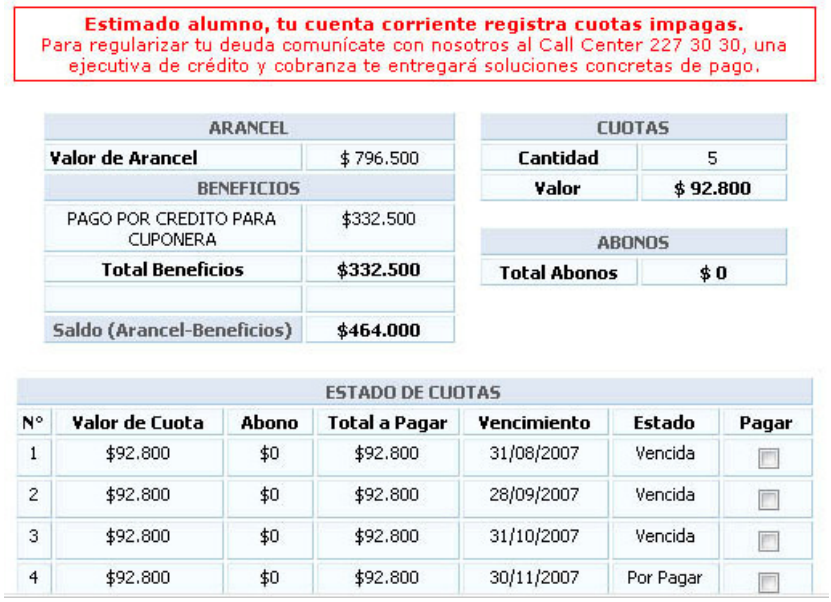

Fig. 3. Layout in *Explorer*.

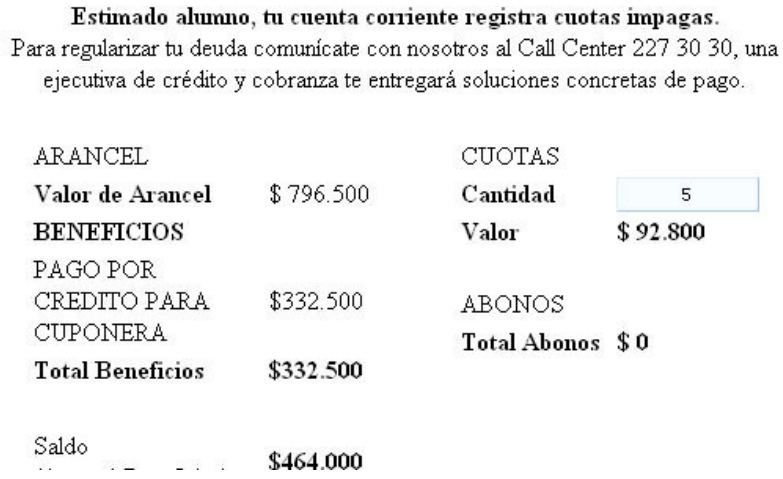

Fig. 4. Layout in *Firefox*.

While analyzing the WUM results it seemed strange there were such big difference between main page hits and the hits for the rest of the areas, 80% against a 3% of the following item (which shouldn't have even been there) and a 1% of the most visited area which does contain relevant information about PUCV and their business goals

(for comparison you can visit [10]). This data together with short page view average time and short visits make you wonder what do the users visit within the site and how. The following hypothesis was developed: visitors get to the main page and get discouraged by the layout and design, they leave quickly thinking that will not find what they are looking for easily, or users find what they look for very quick, but not very often.

Tests proved the hypothesis to be true, it showed that for complex tasks such as find a form to retire from a subject the users behave like the first option, and for common and repetitive tasks they behave like the second one. A shortcoming for WUM was found here, since these repetitive tasks were accessing through the portal their web mail and account on an e-learning system, but these accesses are not recorded in this server, thus they are invisible to the testers.

6 Conclusions

Most nowadays cases study focus on marketing studies and CRM integration of Web Usage Mining with the objective of discover user's preferences and profiles in order to gain more consumers or make the most out of the existing ones. But in this commercial revolution we must not forget the original and true intention of the web: deliver and spread information. Usability plays an important role in this respect; helping to better understand and deliver information on the web, to make the most out of all kind of systems, to the inclusion of people with disabilities to the online age, and many other aspects that may have little to do with profit and commerce.

Keeping this in mind, we have tried a new approach to the use of data mining, an approach that may help companies as well as universities and please consumers as well as readers or researchers. The case study provided useful information that helped unveil and confront some myths and questions surrounding the use of WUM. At this respect we can now say:

- WUM provides enough useful and trustful data and it can be an important part of a web usability evaluation process.
- It's shortcoming does not render it useless, but they do affect it's reach.

So WUM is not perfect, as every technique has pros and cons, which may be:

Pros:

- There is technical information easily accessible: logs are not hard to set up and use, so you don't really have to waste them.
- Problems can be found just by using WUM: as presented in this paper, running the mining first and paying good attention to the results may help you find problem in the site directly from this data.
- There are free and easy to use tools that provide a good set of useful data.

Cons:

- Analysis of WUM results can be confusing: data is not always presented in a well summarized and executive format, so you may need to dig deep into yours site's structure to realize what the data means.

- A site is not always designed to be analyzed: this may make the analysis more difficult since you may have to investigate what the links presented in the report mean and what they contain. Giving a page meaningful and well representative links, titles and names can save a lot of time in the analysis and in the general evaluation process.
- Access data is server dependant and may be incomplete: in the case of PUCV web site, here we found an 80% to 3% gap on accesses, but this may have not been exactly like that, because the most accessed links were not registered in the analyzed server, but in their respective servers.

A great deal of work is left to do in this area, building better applications and software free of these shortcomings, providing more precise and complete data. WUM as proved to be useful and improving its results in the future may position this technique as a must have for every complete web usability evaluation.

7 References

1. Bracke P.: Web usage mining at an academic health sciences library: an exploratory study. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=521513, Arizona Health Sciences Library
2. Haigh S., Megarity J.: Measuring Web Site Usage: Log File Analysis. Information Technology Services, National Library of Canada (1998)
3. Halabi L.: Expert usability review vs. usability testing, www.Webcredible.co.uk, January 2007.
4. Ivory, M., Hearst, M.: The State of the Art in Automating Usability Evaluation of User Interfaces, ACM Computing Surveys, 33 (4), 470--516 (2001)
5. Levi M., Conrad F.: Usability Testing of World Wide Web Sites. U.S. Department of Labor, Bureau of Labor Statistics (2002)
6. Ortega J.: Análisis de sesiones de la web del Cindoc: una aproximación a la minería de uso web. El profesional de la información, 14 (3), 190--198 (2005)
7. Silva M.: Web Mining: definiciones y aplicaciones, Webmining.cl (2001)
8. Shahabi C., Banaei-Kashani F.: Efficient and Anonymous Web-Usage Mining for Web Personalization. Department of Computer Science, Integrated Media Systems Center, University of Southern California, USA.
9. Zorrilla M., E. Menasalvas E., Marín D., Mora E., Segovia J.: Web Usage Mining Project for Improving Web-based Learning Sites. LNCS 3643, Springer, Berlin (2005)
10. AlterWind sample report. http://www.alterwind.com/loganalyzer/samplereport/
11. Assessing Web Site Usability from Server Log Files, Tec-Ed inc. (1999)
12. Guidance on Usability, ISO 9241-11. (1998)
13. Web Site Access Logs: Guess (Literally) Who's Coming to Visit. www.Morebusiness.com, October 1999
14. White paper: Web analytics and usability testing: understanding the total costumer experience. Techsmith Corporation (2005)
15. Web Site Access Logs: Guess (Literally) Who's Coming to Visit, http://www.morebusiness.com/getting_started/website/d935705350.brc