

Timelines for Mobile Web Usability Evaluation

F.Paternò, A.G.Schiavone, P.Pitardi

CNR – ISTI, HIIS Laboratory

Via Moruzzi 1

56124, Pisa, Italy

{fabio.paterno, antonio.giovanni.schiavone, pierpaolo.pitardi}@isti.cnr.it

ABSTRACT

In the field of Web usability evaluation, one potential effective approach is the usage of logging tools for remote usability analysis (i.e. tools capable of tracking and recording the users' activities while they interact with a Web site), and then presenting the recorded data to usability experts in such a way to support detection of possible usability problems. In the design of such automated tools, in addition to the problems related to user behavior recording, another important issue is the choice of meaningful visual representations in order to support the usability expert analysis.

In this paper we present a timeline-based system for interactive events visualization, which has been exploited in a proxy-based mobile Web usability evaluation tool. We discuss how such visualizations can be exploited in finding usability issues and the type of problems that can be detected through them. We show various ways to compare timelines related to users sessions with ideal timelines representing optimal behavior.

CCS Concepts

• Human-centered computing → Visualization → Visualization systems and tools • Human-centered computing → Human computer interaction (HCI) → HCI design and evaluation methods → Usability testing.

Keywords

Remote usability evaluation; Tool for mobile Web logging analysis; Timelines.

1. INTRODUCTION

In recent years, the World Wide Web and its related technologies have undergone significant changes that have affected both the technology used and the features offered to users: as result of these changes, current Web applications are much more complex than in the past. At the same time, several new types of devices able to run Web applications have become increasingly popular. Indeed, the usage of mobile devices (smartphones and tablets) has surpassed that of desktop computers, and new types of devices able to connect and browse the Web (such as videogame consoles, smart TVs, smart watches, etc.) are becoming more and more widespread. In such a dynamic and complex environment, it is increasingly crucial to develop usable Web applications in order to obtain good user

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org. AVI '16, June 07 - 10, 2016, Bari, Italy
Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM 978-1-4503-4131-8/16/06...\$15.00

DOI: <http://dx.doi.org/10.1145/2909132.2909272>

experience in all the possible applications' contexts of use. For several years researchers have conducted studies about usability analysis of Web applications, offering various methodologies and tools for the measurement and the evaluation of usability [1, 2]. In this respect, one of the most used and effective methods is remote usability evaluation, i.e. tools able to support usability evaluation in which users and evaluators are separated in time and/or space. One approach to remote evaluation is based on tools able to record users' interactions with Web applications while they accomplish one or more tasks, and allow evaluators to analyze the logs remotely. In this way, the recording can be performed while users are in their daily environments, and so their interaction with the Web application is more natural and not affected by their being in an unfamiliar environment such as a research laboratory. There are many advantages in using these types of tools: they can reduce both evaluators' effort when setting up evaluations and users' effort when participating in an evaluation. Furthermore, performing users' tests in their daily environment allows user interaction to be more natural and the collected data more realistic. Lastly, the collected data can be analyzed by usability experts at any time, and be further processed.

Nevertheless, usability experts who rely on this type of tools can run into several problems. For example, depending on the recorded data types and the recording level of detail, it is possible that user interactions may generate a huge amount of data that is difficult to manage and understand. Moreover, still depending on the recording level of detail, the data set collected may not be very significant, and therefore it is possible that the resulting representation of the recorded interactions turns to be unclear. Therefore, in order to improve the overall effectiveness of remote usability evaluation tools, it is necessary not only to improve the aspects concerning recording user interactions, but also to design and implement new and more efficient ways to represent the logged data. Such representations should aim to simplify the usability experts' evaluation, facilitating data understanding, comparison and selection. At the same time, they should allow them to abstract from many details in order to be able to obtain significant pieces of information. In this paper we present a proposal for usage data representation and visualization: this proposal is mainly based on timelines, which allow usability experts to conduct more comprehensive user interaction analysis, thus facilitating understanding. We present an interactive representation of such timelines, which can be compared with those corresponding to optimal behavior in order to detect deviations that can reveal potential usability problems.

2. RELATED WORK

The most known commercial software for remote Web analysis is Google Analytics: it provides statistical information about the Web site (e.g. the number of visitors and page views) and visitors (e.g. geographical origin of requests, type of device used). This information is graphically displayed through several types of chart

(line charts, histograms, pie charts, etc.). Other two popular commercial software packages are Crazy Egg and Click Tale. Both of them gather data relating to the user interactions, are able to provide various types of reports in textual format, and to graphically highlight the points of the page in which the user has clicked or touched the screen through heat maps. Such tools are focused on a quantitative analysis of the interaction (number of clicks, areas of greatest interaction), rather than providing information useful to analyze the usability of the Web site. In addition, the navigation and data analysis is performed primarily through textual information (e.g. tables) or static graphical representations (in-page visualization and heat maps).

Several research groups have proposed solutions for remote usability evaluation: Landay et al. [3] made a first attempt to graphically highlight potential usability issues. In fact, their tool was able to gather clickstream data, which were successively presented as an interactive graph, whose nodes represented the Web pages visited, and arrows indicated the transitions between them. The proposed usage of a graph is surely effective to indicate relationships between pages, but not suitable for expressing more complex types of relationships such as the intra-page interactions. WUP [4] was a tool that supported remote usability evaluation of Web sites and considered formal use situations. The tool enabled the evaluator to perform a comparison between actual user behavior and an optimal sequence of actions. This tool included the use of timelines in assessing usability issues in order to represent usage data. Unfortunately, such tool had a low level of interactivity, limited to the ability to scroll and zoom the different timelines. In addition, the graphical representations were unintuitive. WELFIT [5] is a tool for the automatic analysis of usability able to carry out evaluation tests remotely and non. Data acquisition is performed on the client side, though a script must be included in every page by the developers. The tool processes the collected data in order to identify usage patterns and indicate potential user interface design problems: detected issues are shown to usability experts through some statistical reports that do not help to clarify the causes of the usability issues. Lettner et al. [6] proposed an approach for automatically extracting and grouping interaction sequences from users in mobile environments based on a framework able to annotate the source code and thus to define some app's states already during the development phase. To visualize results, the authors used Sankey diagrams, enriched with histograms of the number of interactions that happened at a certain state. Although the data display is very clear and compact, the information provided is only related to transitions between predefined states: therefore, it is not possible to analyze the user behavior in more detail, or highlight usability critical issues. Most of the aforementioned research projects are mainly focused on the evaluation of usability of desktop applications, while one of the most interesting fields of research seems to be the study of usability of mobile applications, which are becoming predominant.

3. VISUALIZING INTERACTION DATA

3.1 Introduction

In this paper we present a solution aiming to bridge the gaps highlighted so far. This solution is part of a new proxy-based Web usability evaluation tool that is able to record the behavior of a user while interacting with any Web application (through either desktop or mobile devices) in order to complete the execution of one or more selected tasks. The data on user behavior are collected through a logger, previously injected into the Web page through a proxy, and stored in a database. Each user interaction is represented as a sequence of events that are generated directly by the user (e.g.

tap, pinch, mousemove, click, etc.) or the browser in response to user actions (e.g. page resize, mobile device orientation change). Indeed, our solution is able to analyze all typical events of touch-based mobile devices (tap and double tap, pinch, pan, swipe, press, rotate, orientation changes) and, for those events for which such information is relevant, also their direction (e.g. pinch out, pan down, swipe left). Furthermore, it is also able to gather information from the smartphone sensors. For example, if the user is about to perform a task through a device with GPS (such as a smartphone), the injected script request permission to share its localization, if the user agrees, geo-localization changes will also be recorded. In any case, the accelerometer information is also gathered in order to detect whether the user is moving even in indoor environments. These data are collected anonymously and only information about the device and browser are recorded. In the tool, usability experts can create and delete usability evaluations and the tasks associated with them, as well as analyze the corresponding data collected during user tests, and eventually share access with other experts using the same tool. In addition, usability experts have the ability to define some custom events (e.g. a click on a particular button or the transmission of data collected through a specified form), which have particular relevance for the Web application usability analysis, and which will subsequently be recorded by the tool in addition to the predefined events. The analysis of the recorded data is performed individually for each task, and is carried out through two different views: Overview and Timelines.

The Overview view allows usability experts to have a high-level overview of the data collected during the execution of a specific task. The aim of this view is to “zoom out” from the details of the recorded data and look at the “big picture” of the user behavior while performing a task. The view is structured as follows: at the top there is a task summary, containing some task-related information such as its name, description, the Web page URL where the task performance starts and indication of whether its execution is skippable or not. Three different pie charts indicate the percentage of the use of different operative systems, browsers and device types used during the evaluation. Each user session is considered as a sequence of events: all events recorded during the execution of all tasks are grouped into categories. For instance, any event corresponding to interactions with task description or instructions for starting, finishing, or skipping a task are grouped into the “Task Management” category. The view provides statistics about the most populated categories, both in tabular form and through histograms. The aim is to provide a general indication about the types of events most common for accomplishing any task. If the usability expert has defined custom events for that specific evaluation, the Overview provides a summary of custom event occurrences, indicating for each of them in tabular form the session in which the event was triggered.

The Overview also presents a histogram showing the completion time of the task under analysis for all the users who have performed it: this is useful to quickly compare all sessions of task execution and get some clues on the actual usability of the Web application. For instance, a session of excessively long duration may indicate that the user had trouble finding the requested information or performing the actions required to complete the task, while an excessively short completion time may indicate that the user has found a faster, unanticipated way to complete the task. In both cases, the usability expert has to deepen the analysis, and make a more thorough assessment of the data, as permitted by the Timeline view.

Task One: Plan Travel timelines

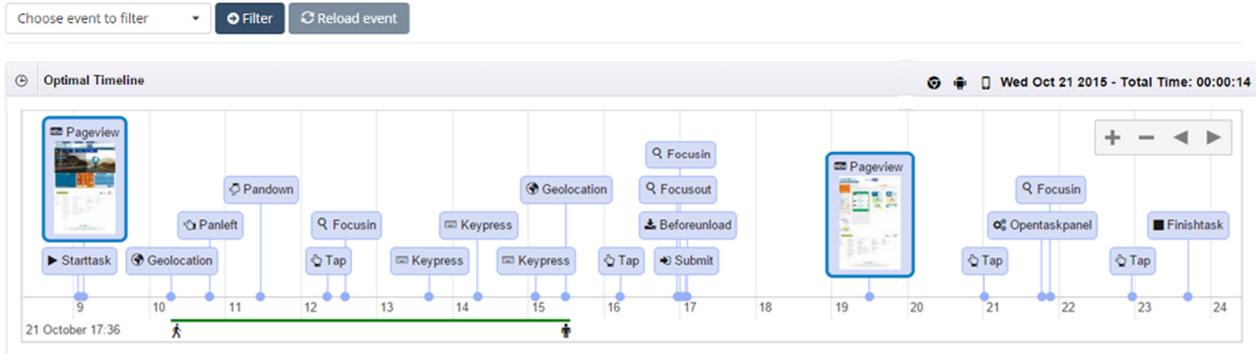


Figure 1. Optimal timeline representing interaction with a form.

3.2 Timelines

Timelines are the second view and enable usability experts to have a closer, detailed look at the recorded data. Every event is characterized by its name, timestamp as well as other relevant information that varies depending on the type of event (e.g. the coordinates of the corresponding point on the page where it was triggered, the associated DOM element, etc.). The problem with displaying event sequences is similar to the micro-interactions visualization problem [7], with the difference that our data have an important temporal aspect: because of this, an effective way to display this kind of data is through timelines. In the proposed solution, the logged interactions are displayed as a chronologically-ordered sequence of timelines, one for each recorded session of task execution. One of these timeline has a special role, as it is a timeline recorded by the usability expert who created the evaluation session: its purpose is to show an optimal execution of the task considered, representing a de facto touchstone for all other recorded timelines. Due to this special role, it is always displayed as the first element in the timeline sequence and explicitly labeled as "Optimal Timeline". Previous work [8] has considered the use of task models for representing optimal user behavior. Such a solution is more expressive since task models can provide an abstract description of multiple sequences of events that are all optimal in terms of use of the considered application. However, it requires knowledge in task modelling that some designers do not have, and thus the use of a event sequences obtained by simply using the application according to the designer view seems more practical and efficient.

Each timeline is inserted inside a container whose header provides some information about the user session represented and is shown in a compact manner. In the upper left corner there is the timeline label, while in the upper right corner there are three icons (see Figure 1) representing the browser, the operating system and the type of device (desktop PC, smartphone or tablet) used to perform the task. Hovering over one of the icons will pop up a tooltip containing more detailed information (for instance, which particular model of smartphone has been used during the task execution). Furthermore, in the top right corner the date and the starting time of the recording are also indicated, along with the total interaction time. Each event that populates the timeline is graphically represented by its own box containing the event type and an icon that illustrates in a simple way both the event type and, when meaningful, the event direction (e.g. a downwards arrow to indicate a downwards scroll event). The boxes are arranged chronologically (according to the event timestamps) left-to-right and connected to the time axis located horizontally at the base of the timeline. Furthermore, the box for each event of type "Pageview" (i.e. an event triggered after loading a new Web page) includes a miniature of the loaded page: clicking on it brings up a screenshot of the page. This is an important feature in the design of our timelines since it enables the usability expert to have a better understanding of the actual user interface accessed at that specific time, and thereby better analyze the corresponding sequence of events performed.

The timeline views can be modified through the buttons on the timeline, or via mouse clicks or mobile device gestures. In this way they can be shifted back and forth along the time axis, as well as

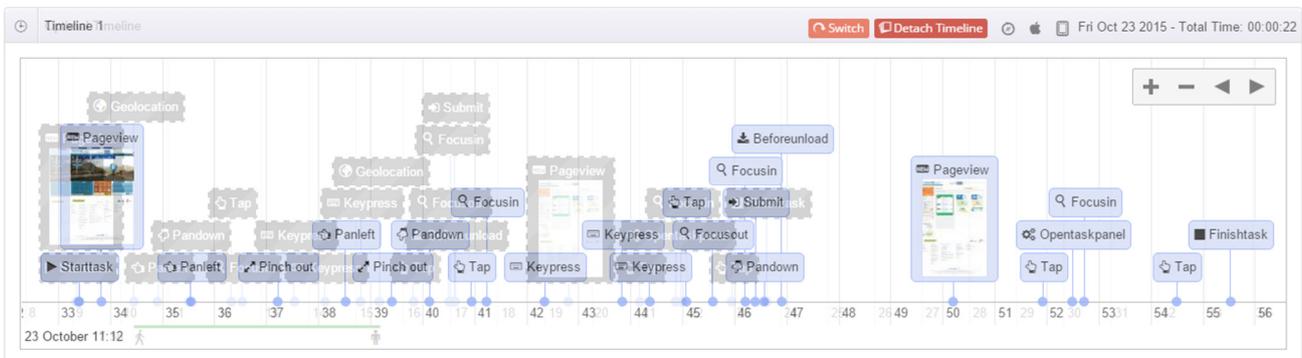


Figure 2. User Timeline overlapping with Optimal Timeline, which is visible in transparency and with reversed colors.

zoomed in and out. In this case, event boxes are dynamically rearranged along the timeline in order to obtain the best visualization. Figure 1 shows an example visualization of an optimal timeline representing a set of events generated on a mobile device while the user was moving, the recorded data regards the interaction with a form. The timeline shows a miniaturized view of the page screenshot at the start of the task, where the user centered the form with some pan events; then she tapped inside a form field to set the focus on it, and finally started to enter some values, and submitted the form values. In response, the browser unloaded the current Web page and loaded a new one, whose miniaturized screenshot is shown. Lastly, the user opened the task panel and tapped on the “Finish task” button in order to indicate that the task was completed.

In a usability test various users can perform the same task, thus generating many timelines. Consequently, some timelines can fall at a considerable distance from the optimal one in the list of timelines, thus making their comparison difficult, and forcing the usability expert to frequently scroll up and down along the Web page that contains the Timeline view. In order to avoid this inconvenient situation and facilitate the work of the usability expert, we have introduced a feature that provides the possibility of automatically shifting a timeline and overlapping it over the optimal one: in this case, one of the two overlapping timelines will always be visible in transparency, facilitating the comparison (see Figure 2). The foreground and background timelines can be distinguished because the foreground has event representations with black labels on a light blue background while the background timeline has white labels over a gray background with wavy borders. It is also possible to reverse the overlapping order, placing the optimal timeline in the foreground and the other one in the background, as well as to detach them, automatically returning the analyzed timeline to its original position. The two overlapped timelines can also be “locked” and simultaneously moved back and forth along the time axis, as well as zoomed in and out concurrently. The name of the timeline in the foreground is indicated in the upper left corner.

The overlapping timelines feature is one of the main elements of our proposal: in fact, this feature allows the expert to analyze, simultaneously and at a single glance, the evolution of two distinct sessions (the optimal one and the user one). In this way, the usability evaluator can easily distinguish relevant information, which is not immediately understandable otherwise. For example, experts can check if the two logs follow the same sequences of events or diverge, see how they diverge, and gather some elements to help understand the reasons for such differences. Furthermore, if the Overview reveals that a user session has an excessively long task completion time with respect to the optimal, the overlapping functionality allows the two sessions to be analyzed in detail. In fact, the expert can check whether this difference in timing depends only on the fact that the user interacts at a slow pace, or that she is performing incorrect actions within the page, or moving back and forth amongst other pages on the same site. Figure 2 shows an example of overlapping timelines: the same optimal timeline shown in Figure 1 is in the background, while a new timeline, corresponding to the performance of the same task by another user, is presented in the foreground. It shows that the duration of Timeline 1 is much longer than the optimal case (note how the second Pageview event occurs later than in the background). Through a closer look, we can also notice that the user corresponding to Timeline 1 had to zoom the page in order to easily interact with the form (as revealed by a sequence of pinch out and

pan events). In the case that the timelines are too crowded, making them difficult to read, we have also provided the functionality to reduce the number of displayed events, by filtering them according to their type. In fact, through a drop down menu located above the timeline bar beside the task name (see Figure 1), it is possible to select one or more event types, and through the “Filter” button remove the selected types from all the timelines, thus reducing the number of displayed events. The “Reload Events” button restores the original condition, redrawing all the timelines with all the events. The tool also provides the possibility to show the route taken by the user during the test: this information is shown as a map indicating the route. In mobile environments, this feature can provide additional information to the usability expert, by providing information about the environment in which the task was performed, and thereby allowing a better reconstruction of the user behavior.

4. CONCLUSIONS AND FUTURE WORK

We have presented a proposal for an interactive visualization system for displaying events data collected through a remote usability evaluation tool. Our solution allows usability experts to easily perform an accurate analysis of user interactions in order to detect usability issues, providing them with both a general overview of user behavior and several functionality to analyze the collected temporal data in detail.

Future work will be dedicated to improving the visualization part of the tool. We also aim to provide further support to analyze user behavior and identify associated usability issues.

5. REFERENCES

- [1] Ivory M.Y., Hearst M.A., “The state of the art in automating usability evaluation of user interfaces.” *ACM Computing Surveys* Volume 33 Issue 4, December 2001, pp. 470–516.
- [2] Hilbert D. M., Redmiles David F., “Extracting Usability Information from User Interface Events”, *ACM Computing Surveys*, volume 32, Issue 4, December 2000, pp. 384–421.
- [3] Landay J.A, Waterson S. J., Hong J. I., Sohn T., Heer J., Matthews T., “What did they do? understanding clickstreams with the WebQuilt visualization system”, *Proceedings AVI 2002*, pp.22-24, ACM Press, Trento, Italy
- [4] Carta T., Paternò F., Santana V.F., “Web usability probe: a tool for supporting remote usability evaluation of web sites”, *Proceedings INTERACT*, September 05-09, 2011, Lisbon.
- [5] Santana V.F., Baranauskas M.C.C., “WELFIT: a remote evaluation tool for identifying web usage patterns through client-side logging”, *Int. J. Hum. Comput. Stud.*, 76 (2015), pp. 40–49.
- [6] Lettner F., Grossauer C., Holzmann. C., “Mobile Interaction Analysis: Towards a Novel Concept for Interaction Sequence Mining”, *Proceedings MobileHCI 2014*, ACM Press, pp. 359-368.
- [7] Breslav, S., Khan, A., Hornbæk, K. “Mimic: visual analytics of online micro-interactions”, *Proceedings AVI '14*, ACM Press, pp. 245–252.
- [8] Paganelli L., Paternò F., *Tools for Remote Usability Evaluation of Web Applications through Browser Logs and Task Models*, Behavior Research Methods, Instruments, and Computers, The Psychonomic Society Publications, 2003, 35 (3), pp.369-378, August 2003.