Statistical Analysis of the User Experience

Avi Harel - Ergolight

isENBIS Meeting

Usability is a key dimension in interactive systems, affecting the user's satisfaction, productivity and safety. Traditionally, the term usability denotes the ease with which people can employ a particular tool or other human-made object, in order to achieve a particular goal. Another important aspect of usability is the capability to cope with exceptional situations, especially, of risky situations. Such situations contribute to negative user experience, and should be eliminated by design.

Usability assurance is achieved by proper design, feedback and testing. Common practices for usability testing are based on inspecting user representatives and on feedback questionnaires. These practices cannot reveal critical failure modes typical to remote users, in real life, especially in fast interactions. This article presents a method for usability failure diagnostics, by statistics of logs of the users' activity.

Log files include time stamps, enabling measuring the time elapsed between consecutive user actions. For the analysis we use a memoryless navigation model, which is basically a Markov process, in which the nodes represent user interface states, and the arcs represent state transitions. In our navigation model, we also attribute arcs with average time intervals, obtained from the log files. In web applications, the states are pages on screen.

Statistical analysis of log files is key to cost effective usability analysis and to remote usability testing. Such analysis should go beyond standard Google analytics, Webtrend or similar traffic tracking monitors. In our work we show how data can be collected, analyzed and interpreted with a focus on usability. The statistical models we use range from simple t-tests to Markov chains and Bayesian Networks.

Our navigation model assumes two types of negative user experience: Confusion and irritation. Instances of confusion are characterized by hesitation; therefore, the expected time interval in case of confusion should be longer than the average. Instances of user irritation are characterized by fast response, resulting in the expected time interval being shorter than the average.

Also, the navigation model assumes that following a negative user experience, the user will try to revert to the previous state or to get help. The controls actuating such actions are called User Problem Indicators (UPI). Examples of UPIs in Windows applications are the Esc key, and the Help, Cancel and Undo features. In web applications, UPIs consists of backward navigation, rescuing to the home page or to a main menu item, or site exit.

The rate of UPIs may be used for scoring UI states and web pages. A usability index may be defined as the proportion of the normal transactions through the state or page. For dynamic systems, such as those implementing the SOA paradigm, the usability index enables prediction of usability barriers, for example, due to extreme system overload. However, it should be noticed that users may actuate UPIs intentionally, as a means for exploring the system capability and behavior, by recalling for information or by trial-and-error sessions. This implies that for the diagnosis, the UPIs should be weighed and selected by measures of the user's hesitation or irritation. This is achieved by statistical analysis of the transition time intervals associated with UPIs.

The time interval between two consecutive actions may be attributed to the combination of the time it takes to perceive and decide how to respond to the first action, and the time it takes to search, select and execute the second action. A measure of the first mental activity associated with a given control may be obtained by averaging the elapsed time of all transitions from that control. Similarly, a measure of the second mental activity associated with a given control may be obtained by averaging the elapsed time of all transitions to that control.

Besides operating the system, the users may be doing many different things. The user side activities are not recorded in the log files, resulting in high noise level in the data. To minimize the effect of such noise, high values of interval length should be attributed with low weigh. Averaging by medians enables filtering out such noise. Alternatively, we can transform the time intervals to frequencies. This transformation reduces the effect of such noise.

The navigation model includes assumptions about the effect of the user motivation on the rate of UPIs. For example, if a page is difficult to read or comprehend, then highly motivated users will spend more time than the average trying to read the content, and their disappointment should be reflected by a high UPI rate. Also, if the page is mostly nuisance, for example, a page that includes animated ads, then low-motivated users will spend less time than the average on this page, and subsequently will activate a UPI. Bayesian Networks are useful to verify such assumptions.

To decide that the reading time of a particular state or page is too short or too long, we can compare the average reading time of two samples: one consisting of transactions terminated with UPIs and the other including the normal transactions. If they differ significantly, we can conclude that reading is a problem for that page. Simple t-tests can be employed for the statistical decision.

The navigation model described above was implemented in WebTester, a tool that extracts usability diagnostic data of web pages from server log files. The tool implements a seven-layer model, providing usability reports on usability barriers. The main benefit of this tool is the ability to pinpoint particular problems of particular web pages.

The face validity of this tool is high, but not perfect. Therefore, usability professionals should be required to examine the reports, to confirm that the diagnosis is valid and to add the value of their expertise in recommendations for how to improve the page design.

References

A. Harel, R.S. Kenett and F. Ruggeri, **Modeling Web Usability Diagnostics on the basis of Usage Statistics** (A. Harel, R. Kenett and F. Ruggeri) in: Statistical Methods in eCommerce Research, W. Jank and G. Shmueli editors, Wiley.

A. Harel, R.S. Kenett and F. Ruggeri, Decision support for user interface design: Usability diagnosis by time analysis of the user activity , *COMPSAC*, 2008.