Your Browser is Watching You: Dynamically Deducing User Gaze from Interaction Data

David Hauger[1], Alexandros Paramythis[1], and Stephan Weibelzahl[2]

[1] Institute for Information Processing and Microprocessor Technology
Johannes Kepler University, Altenbergerstrasse 69, A-4040, Linz, Austria
{hauger, alpar}@fim.uni-linz.ac.at

[2] National College of Ireland, Dublin, Ireland
School of Computing, National College of Ireland, Mayor Street, IFSC, Dublin 1, Ireland
sweibelzahl@ncirl.ie

Abstract. Most user models in adaptive hypermedia systems rely on server-side data collection, such as pages visited or links selected. The work reported in here seeks to expand the inferences that can be made on the client side, by observing and analyzing patterns in the users' interactive behavior such as mouse moves, clicks, text selection and scrolling. These patterns can be used to estimate with high levels of accuracy what the user is reading on a page. In an empirical study involving gaze tracking of 13 participants we validated these patterns and analyzed different prediction models. We established that a high frequency of mouse moves on a page, clicks on text and the mouse being in motion strongly contribute to the accuracy of predictions of what users have read.

Key words. interaction monitoring, gaze modeling, eye-tracking, empirical study, mouse activity

1 Interaction Data in Adaptive Hypermedia

Traditionally adaptive hypermedia systems (AHS) use server-side data collection for inferences about what the users have read, are interested in, etc. The users' client-side behavior has long been identified as a potential additional source of information towards the goal of increasing the accuracy of assumptions regarding what they really read while at a page. Previous work exploited client-side data collection to identify a general "level of activity" [1], the reading time [2], or learning types [3] based on patterns in the mouse behavior. Rather than making inferences on pages as a whole, it is possible to treat fragments of a page separately [4]. We decided to go one step further, and examine whether it is possible to increase the accuracy of predicting what a user was looking at while at a page, by identifying and interpreting specific patterns in the user's interactive behavior.

We introduced a number of hypotheses on usage patterns users might show and, based on their observation, what additional inferences could be made on eye gaze.

In order to test our hypotheses, we designed a study that would allow us to compare users' reading behavior when encountering different types of text, with their interactive behavior while reading these texts through a browser. Reading behavior was determined through eye-tracking, whereas interactive behavior was recorded through the purposely developed JavaScript library. Participants were given four tasks to perform, each based on a different type of text typically encountered online (instructions and additional information for a board game, a set of search results, a health-related article, and a set of news items). All material was presented through a TFT screen, running a 1280x1024 resolution. The preliminary results presented below are based on 13 participants (6 male, 7 female) completing the first task.

The study itself was organized as follows: After answering a questionnaire on demographic data and prior knowledge, participants completed a reading speed test. Next, participants were instructed to read a series of four different types of text on the screen. The first text, and the one addressed in this paper, introduced participants to the Game of Go. The text was split into seven pages. It comprised text (ca. 7010 words), graphics (11) and pictures (5).

While participants were reading, two types of data were recorded. Firstly, the JavaScript library, embedded in each web page, recorded all mouse and keyboard behavior, including the mouse position, mouse clicks, keys pressed and scrolling events. Secondly, using the eye-tracker, we recorded the actual gaze position.

3 Results

In total 112 page requests have been observed with a page being visited for 2 to 1096 seconds with a mean of 122 (σ=116s). On average each user spent 17.5 minutes on the information on the game of Go. For this study we have defined a "mouse move" to be any set of changes in the mouse pointer's position, preceded and followed by at least one second of idle time (i.e., time during which the mouse pointer's position does not change).

Even without looking at particular interaction patterns, the current mouse position can give a very rough indication of where the user is looking: mouse and eye position are correlated, both horizontally (r=.101, N=89739) and vertically (r=.250, N=89739). The vertical correlation's being much higher than the horizontal correlation might be due to users not typically following their gaze with the mouse from left to right when reading, but "pointing" at the paragraph.

The first step towards improving upon this baseline correlation was the insight that some people use the mouse a lot, while others "park" it as long as they do not need it. On pages where the mouse is moved frequently it should be easier to estimate where people are looking. For pages where users moved their mouse frequently, *the mouse pointer's position is strongly correlated with the position of the users' gaze*. The frequency ratio (FR) denoting the proportion of time moving the mouse on a page was used as a filter for selecting pages. We applied filter settings between 25% and 75%.

Is it possible to improve prediction of gaze by considering the frequency of mouse usage on a page? The higher the percentage of mouse movements, the lower the distance between mouse and gaze positions (r=-.299; N=89739). When events are filtered by the level of frequency of mouse movements, the correlation increases even further (see Table 1). The more restrictive the filter, the higher the correlation. In accordance with the baseline, the correlations in the vertical direction are higher than in the horizontal. The correlations get even higher when limiting the analysis to events when the mouse is actually in motion, both in horizontal (e.g., $r_{>25\%}$=.746, N= 36202 and $r_{>75\%}$=.777, N= 21270) as well as in vertical (e.g., $r_{>25\%}$=.521, N= 36202 and $r_{>75\%}$=.580, N= 21270) direction. Further to the above, results also showed that some events like clicks on non-link text and text selections are very well suited to identify whether the related paragraph has been read, although the improvement of prediction they afford comes at the expense of limited coverage.

In summary, we found that certain patterns in users' interactive behavior can be used to approximate their gaze position. Based on these results we are currently developing a prediction algorithm that is used by the client-side monitoring mechanism to deduce what parts of pages users have focused on. The algorithm uses weighted interaction pattern-based indicators that combine several of the reported results. We plan to implement it as a JavaScript library that can then be integrated into any Adaptive Hypermedia System to improve the granularity and accuracy of user models with respect to what users have really looked at while on a page.

Table 1. Correlations of x and y positions of mouse cursor and eye gaze, depending on the frequency of mouse usage. Regression model: eye pos. = (weight * mouse pos.) + constant

	Frequency of mouse moves	Correlation		Regression model: mouse on eye		
		$r_{\text{eye vs mouse}}$	N	constant	weight	sig.
vertical	baseline	.250	89739	345.133	0.228	.000*
	frequency > 25%	.608	39134	211.082	0.567	.000*
	frequency > 50%	.658	21906	173.328	0.613	.000*
	frequency > 75%	.746	16360	165.461	0.666	.000*
horizontal	baseline	.101	89739	577.499	0.075	.000*
	frequency > 25%	.393	39134	385.193	0.386	.000*
	frequency > 50%	.493	21906	241.912	0.604	.000*
	frequency > 75%	.560	16360	188.254	0.727	.000*

Acknowledgments. The work reported in this paper is partially funded by the "Adaptive Support for Collaborative E-Learning" (ASCOLLA) project, supported by the Austrian Science Fund (FWF; project number P20260-N15).

References

1. Goecks, J., Shavlik, J.W.: Learning Users' Interests by Unobtrusively Observing Their Normal Behavior. In: Int. Conf. on Intelligent User Interfaces – Proceedings of the 5th Int. Conf. on Intelligent User Interfaces. (2000) 129-132
2. Spada, D., Sánchez-Montañés, M.A., Paredes, P., Carro, R.M.: Towards Inferring Sequential-Global Dimension of Learning Styles from Mouse Movement Patterns. In: AH. (2008) 337-340
3. Cha, H.J., Kim, Y.S., Park, S.H., Yoon, T.B., Jung, Y.M., Lee, J.H.: Learning Styles Diagnosis based on User Interface Behaviors for the Customization of Learning Interfaces in an Intelligent Tutoring System. In: Intelligent Tutoring Systems. Vol. 4053. (2006) 513-524
4. Hauger, D.: Using Asynchronous Client-Side User Monitoring to Enhance User Modeling in Adaptive E-Learning Systems. In: UMAP 09: Adaptation and Personalization for Web 2.0, Trento, Italy (2009)