What and to what extent you have read: understanding your daily online learning process

Lei Zhang
Georgia State University
Atlanta,Georgia,U.S.A
lzhang14@student.gsu.edu

Ying Zhu
Georgia State University
Atlanta,Georgia,U.S.A
yzhu@cs.gsu.edu

ABSTRACT

In this paper, we present our research on detecting habitual reading region on the screen when a user is doing his daily online reading. Instead of traditional eye-tracking based solution, we developed an unintrusive user attention tracking(UUAT) solution. In UUAT, we developed a Reading-Clue technology, which helps to pinpoint user reading attention when necessary. At the same time, user reading behavioral data(mouse click and mouse scroll) is collected in background. With these behavioral data, UUAT calculates a user's preferred reading region. Our experiment indicates that the accuracy of UUAT attention tracking is as high as 93.2%.

Author Keywords

Attention Tracking; Reading; Mouse movements; Scrolling;

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Personal Data Mining[12](PDM) has been proposed as a complement to commercial/business data mining. While commercial data mining[4] often aims to discover new knowledge to improve business opportunity or maketing strategies, PDM is proposed for a users individual good. In PDM, user behavioral data is collected by a neutral utility, and it is under the user's full control. The user can choose third-party tools(local) or service(cloud) to analyze his personal data, aiming to improve individual productivity or life quality. Besides being mined alone, individual's behaviroal data can be analyzed with data from other users to find correlations among the data or similar users(e.g. User with similar interests).

As the first step of PDM, detailed and objective raw personal data should be collected. Self-Tracking[13, 14] is getting increasingly popular to log personal data, so that future analysis can be conducted for a users good in health. Self-Tracking has been mainly involved in personal physical data, such as, weight, heartbeat, diet and sports. Moreover, the way of logging data has been mainly referred to manual logging, e.g. a user input his weight before going to bed everyday. However we argue that the reign of personal data is broder than physical data and the ways to collect personal data can be extended to an automatic and unaware way.

In this research, we consider a users daily reading as a source of personal data. An immediate reason is that understanding a user's daily reading helps to build a users personal knowledge database, let alone "For many self-trackers, the goal is unknown"[14]. To be specific, we do not consider reading of an artilce as a simple-and-plain data item (e.g one record to log the tittle, url, timestamp,etc). Instead, we consider it to be richer informational process , where potentially we can discover much more knowledge. For example, during the reading process, if a user spent more unit reading time on a specific paragraph than average(no distraction), this indicate that the user might have pondered on the information in that paragraph or have reading struggles on paragraph one. In either case, it might indicate that the user has encountered new knowledge and he takes time to digest this new information. Another example is that, if a user skims very fast over a paragraph, it is very likely that he is not interested in the information provided by that paragraph. From this point of view, it might be a mileage for personal knowledge acquisition, e.g., today is the day I first come to know the concept of shale gas(I read and pondered on the paragraph), although I have read the same phrase years ago(I skimmed over that paragraph very fast). To enable this type of knowledge discovery and to record this type of personal data, it is unlikely to adopt the existing user-aware solutions(e.g. eye-trackiing or self-tracking), a smart and intelligent collection method is required to obtain personal reading data.

In this research, we focus on how to collect detailed and objective reading behavioral data, so that PDM can be conducted on an accurate and correct basis. To be specific, we focus on a users daily online reading activity. Our goal is as follows: in a finer granularity, we accurately identify what texts a users read and to what extent he has spent on specific parts of an article. To realize this goal, we developed a browser-based attention tracking system, namely UUAT, to collect daily reading behavioral data. UUAT collects a user's interaction data during his reading, then it computes user's behavioral reading region and finally provides reading details of that reading event. In this way, we enable a more extensive PDM applications which we present in future work section. Our contributions are as follows: 1. To the best of our knowledge, we first compute the user preferred reading region in an unaware but accurate way. 2. We monitor a user's reading and produce detail, meaningful, and objective data, which can be used as input for deep PDM.

RELATED WORK

The idea of quantified self-tracking is well known proposed in [14]. In [14], the author advocates collecting individual quantified data. He presents cases where individual data collection helped to solve person-specific problems. For example, Barbier used her daily logged personal data to find a way to cure her insomnia[14], Seth Roberts, a Caltech professor, analyzed his personal data and find an optimum diet(flaxseed oil) to improve his math performance[14]. In this work, the author advocates a diversified personal data collection, as long as the data is quantified, albeit you might not know the goal of your data collection immediately. The current state-of-art personal data mainly includes personal physical data, such as weight, glucose, heartbeat, sports, food/medication consumption. So the existing pervasive methods of personal data collection are mainly by a user himself, such as the methods adopted by PatientsLikeMe[1], a user/patient input his biological data into his computer(usually a spreadsheet) and save it for future use. Our research adopts an user-unaware, automatic method to collect a user's personal data.

Our first research goal is the computation of a user's preferred reading region, the literature in this research field can be mainly divided into two catagories. Eye-Tracking based research and mouse/keyboard tracking based research.

[11] is a typical eye-tracking based research. The author used eye-tracking to find out that a user's attention region and cursor are closely related, both of them can be modeled as a normal distribution. We argue that eye-tracking based methods are also tagged as intrusive and expensive. On the one hand, the behavior of subjects wearing a headset might be biased from his natural behavior. On the other hand, it is impractical to popularize an eye-tracking based application to a large scale, especially when a pervasive self-tracking is referred.

Compared to eye-tracking method, our research is closer to the work of [10, 8, 9, 3, 7, 2] The work in [2] developed a vertical heatmap bar for long webpage viewers, indicating the dwell time of each specific part of web page. This paper collects similar data as us (user scroll), but aims at a different goal: to navigate user when he want to read backwards. The work in [7] developed compared the different reading assistance technologies, and find out the improvement on reading from each technology. Our work is different from [2, 7] that we do not intend to assist a user's reading process(navigation or improve reading efficiency), we aims to develop an "observer", which can get objective and accurate user reading behavioral data.

The authors of [10] collected user mouse data, at the same time they collected users eye gaze data, they proved that the user mouse data can indicate user behavior, e.g. which part of the screen is more attractive, the user distraction behavior can be detected and the user experience can be predicted by mouse analysis in an accuracy of 80

By literature review we distinguish our research from existing works as follows: the data collection in our research is designed to be more scalable, so our data collection method can be tagged as "unintrusive", unlike the eye-gaze technology adopted in [10, 9]. Meanwhile our goal is personal data collection instead of usability test, we proposed a new CHI technology to accurately infer users preferred reading region on the screen, so that we can collect high quality data that can describe a users daily reading behavior objectively.

The review for knowledge discovery

WHERE DO YOU READ?

To understand online reading in an user-unaware way, it is necessary to determine a user's Reading Region(RR). A Reading Region is defined as the screen area where a user's reading attention effectively covers. Only those texts in the RR are possibly read and learnt by a user. In this section we will first model RR and introduce our method to compute behavioral RR and realtime RR.

Model of Reading Region

Psycologically, felt involvement[6] influences attention and comprehension. Felt involvement is low when a person starts reading any text. but it increases rapidly as reading continues until felt involvement reaches a peak value. After that, felt involvement decreases due to a decrease in the attentional resources. The hypothesis in [6] can also be proved by [5].

The previous research on user attention tracking [5] have disclosed some facts when a user is doing daily reading on long articles and documents. Vertically, in most of the time, the reading progresses towards the end of the screen in a line-by-line manner. However, the reading does not start from the first line on the screen to the last line of the screen. Generally, a user has a preferred reading region on screen, which in this paper, we name it as behavioral RR(BRR), as indicated in figure 1. A BRR can be identified by two parameters Δ and w_0. Δ indicates the offset from the head of the screen and w_0 indicates the vertical span of the whole region. If a user wants to read the contents outside the RR, a scroll action is expected. In this research, we take aforementioned assumption and model the user attention region as displayed in figure 1.

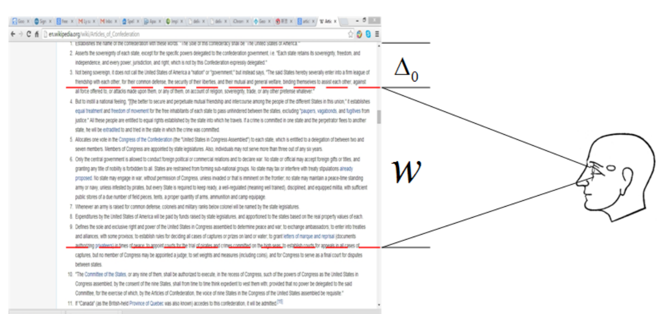

Figure 1. The preferred reading region on the screen

Taking this model into consideration, our next research question is: What behavioral data should be collected and analyzed to compute RR?

Eye tracking data can be easily mentioned to answer the question, especially when a "visual" reading region is the goal. However, we argue that the eye tracking data has twofold major drawbacks: 1. The eye tracking collection requires extra hardware, which means the eye tracking based solution is not scalable and very costly. 2. What is even worse, eye tracking is "intrusive". A subject person is aware that he is wearing a headset device and he is under a camera coverage, his behavior is very likely to be different from how he behaves when he is doing reading alone. For these two reasons, we can not use eye tracking data. To meet the requirement of inattentive but objective, the only option left is the user interaction data collected when the reading is ongoing. In the next subsection we will discuss how to collect and analyze user interaction data to get the user preferred region.

Calculating BRR by Interaction Data

When a user is doing daily reading, since the keyboard input rarely happens during the reading process, we are more interested in the mouse-generated data. There are two categories of mouse-generated data, click and scroll. We consider the mouse-generated data to be informational behavioral data. To be specific, an intentional mouse click indicates the position of user's instantaneous reading attention, while a page scroll indicates how much contents has been moved out of reading region.

In this subsection, we answer the question of how to infer BRR by mouse-generated data. To be specific, according to figure 1, RR can be identified by two parameters Δ and w_0. So the computation of BRR can be transformed to how to compute Δ and w_0.

Now we examine a typical online reading event, during which a user finishes reading of a whole article. We suppose the article is relatively long, which requires many scroll actions. The whole reading process can be illustrated by figure 2

In figure 2, a user conducts a scroll action when he has finished reading contents of current reading region. At ith scroll action, a w_i of page height will be moved out of region. So when the article reading is finished, we have equation (1). In equation (1) , Δ and w_0 are defined in figure 1. Ideally, equation (2) should hold and the parameter w_0 can be calculated in equation (3). In practice, it is unlikely for equation (2) to strictly hold, however, equation (1) can still be applied.

$$H = \sum_{i=1}^{n} w_i + \Delta + w_0 \tag{1}$$

$$w_1 = w_2 = ... = w_n = w_0 \tag{2}$$

$$w_0 = \frac{\sum_{i=1}^{n} w_i}{n} \tag{3}$$

With equation 3, we can compute w_0 parameter of a user's behavioral reading region(BRR). The computation of Δ can not be accomplished if only the scroll data is considered. Here we take the mouse click data into consideration.Since we take the assumptions that the reading process progresses line by line and the mouse click is an indication of instantaneous user attention, we take the first click after each scroll as the clue to compute Δ.

We note the click data K as a sequence $K = (K_0, K_1, ..., K_n)$, where there are n scrolls and the component K_i is the click sequence collected after i_{th} scroll. Furthermore, $K_i = (K_i^0, K_i^1, ...)$ and K_i^j consists of timestamp and coordinates, $< t_{k_i^j}, x_{k_i^j}, y_{k_i^j} >$. So we can compute Δ as following equation:

$$\Delta = \frac{\sum_{i=0}^{n} y_{k_i^0}}{n+1} \tag{4}$$

Once we compute Δ and w_0, we can have a good estimation of BRR. BRR can be considered as relatively stable. During the reading process in real time, BRR can be used to estimate the reading region if not enough information is provided. Due to the variation, equation (2) does not always hold, a real-time reading region(RRR) should be computed at each scroll action, especially when the variation among scrolls are relatively large.

Calculating RRR by Interaction Data

Since equation (2) does not always hold, if the variance of w_i is relatively large(actually it is), we need to consider the question how to infer user reading region on each single scroll. Figure 3 displays the events happened during reading between two scroll actions. Here solid dots are collected mouse click data with the position on the corresponding place on the screen. Since we take the assumptions that the reading process progresses line by line and the mouse click is an indication of instantaneous user attention, we can calculate a minimum reading region after i_{th} scroll action, $w_i^{\min}$, which can be identified by highest click and lowest click(as indicated in figure 3).

However, $w_i^{\min}$ is a minimum approximate of w_i, we can have a more accurate approximate if we take the neighboring scroll actions into consideration. In figure 4, we put the click data both before ith scroll(purple dots) and after $i+1$th(yellow dots). We can see there is a gap between two neighbor minimum approximates($g1$ and $g2$ in figure4). So we can have a moderate estimate of w_i:

$$w_i = w_i^{\min} + (g_1 + g_2)/2 \tag{5}$$

It should be noted that, ideally, $w_i^{\min} = w_i = w_0 (1 \leq i \leq n)$ and $g_1 = g_2 = 0$. But in fact, it is unlikely that a user starts reading and finishes reading from the exact same locations, click data has to be considered to compute RRR. If there is no click data available, we use BRR to replace RRR.

WHAT AND TO WHAT EXTENT YOU HAVE READ

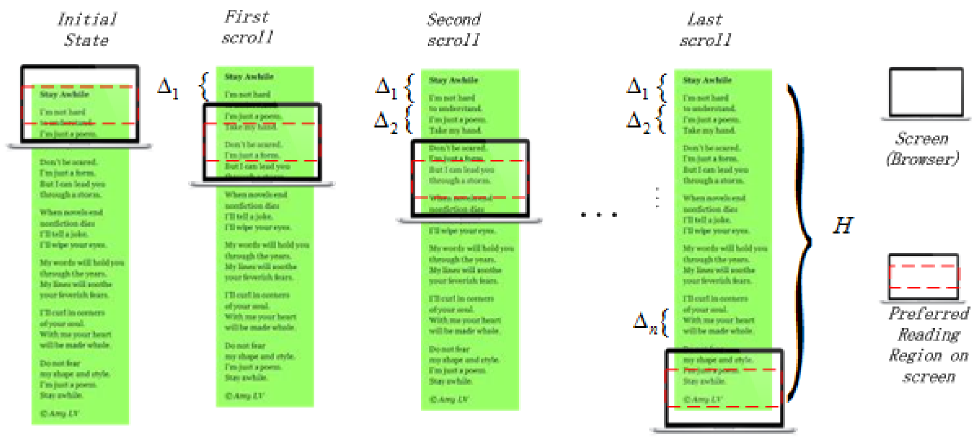

Figure 2. Reading of a long article in a browser

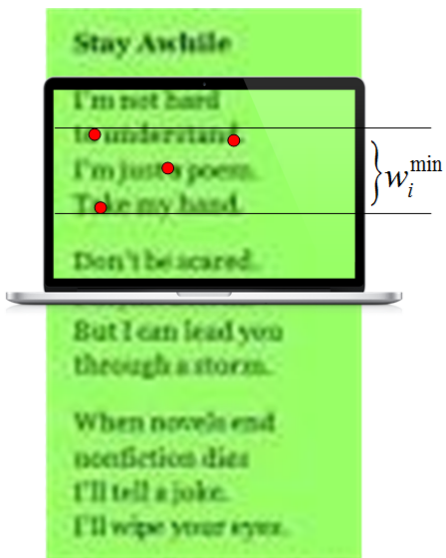

Figure 3. Mouse click and reading region

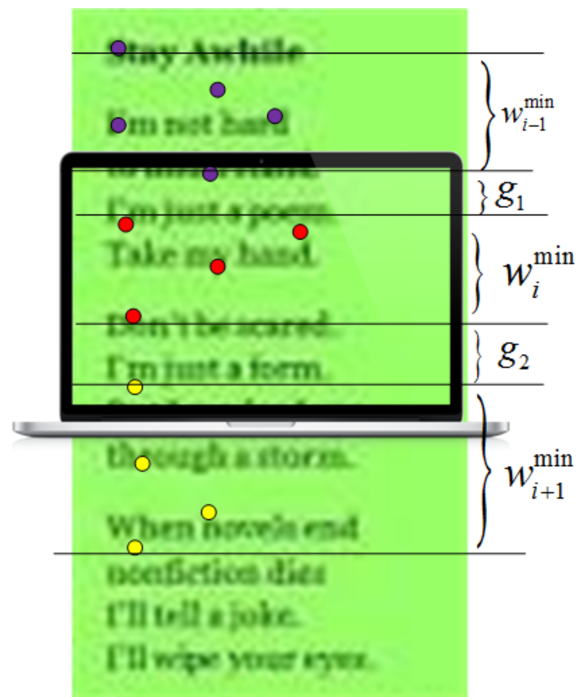

Figure 4. Mouse click and reading region

Once we compute the BRR and RRR, we can answer our second research question: what and to what extents a user has read. Here we want to answer this question in a finer granularity. We consider a whole article to be a bad granularity, instead, we answer this question to the granularity of paragraph.

In order to formulate our solutions here, we introduce the following notations. Given two points $p_1(x_1, y_1)$ $p_2(x_2, y_2)$ on a screen at time instant t, the context identified by the two points can be notated as : $C(p_1, p_2)$, which is indicated as dark background texts in figure 5. The corresponding time spent reading this context is $T(p_1, p_2)$.

For each scroll span w_i, p_s is defined as the left top point of that screen area, p_e is defined as the right bottom point of that screen area. Intuitively, we can have the reading time of this screen $T(w_i)$ as follows:

$$T(w_i) = T(p_s, p_e) \tag{6}$$

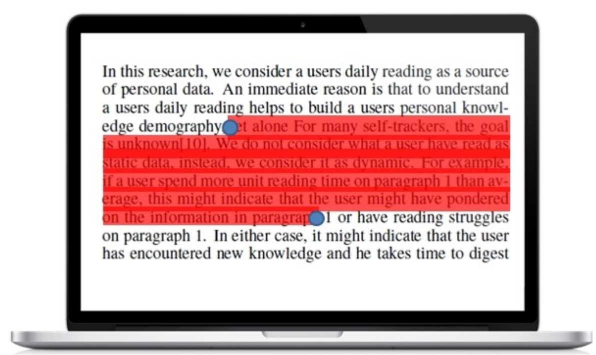

Figure 5. Contexts between points

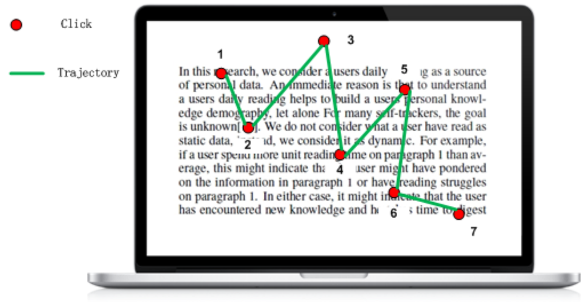

Figure 6. Mouse Click and Text Extraction

However, it is only a coarse-grid estimation of the dwell time, because the context $C(p_1, p_2)$ is relatively large. By taking the mouse click data into consideration, we can have a fine-grid estimation of dwell time. As we stated before, we consider mouse click as indication of real time user attention. However, there might be some "noise" in mouse click data, which can not be considered as a clue of user reading. Figure 6 illustrates a typical reading scenario where mouse click and mouse trajectory are displayed. Apparently, point/click p_3 is considered to be a "noise", since the position of p_3 does not indicate any text contents. This might indicate a occasional subconscious click. Except p_3, all the other six points locate in a certain text area. However, not all of these six points can be used as reading attention indication.

In real time, we note each click as $k(t, x, y)$, where t is timestamp, x, y are the cooridnate of the point where click happened. The intentional click happens in a sequence of $k_1k_2k_3...k_n$. So, based on our assumption of line-by-line reading, we have the following equations hold, in equation 7, $1 \leq i < j < k \leq n$.

$$\begin{cases} C(k_i, k_j) \cap C(k_l, k_s) = \emptyset \\ t_i < t_j \leq t_l < t_k \end{cases} \tag{7}$$

According to equation 7 we can filter out P_5(in figure 5) as noise since $C(k_1, k_2) \cap C(k_4, k_5) \neq \emptyset$.

In practice, the information if a click is intentional or not is not given, so we developed algorithm 1to find the maximal intentional click set(ICS) where equation 7 holds.

DESIGN AND IMPLEMENTATION OF UUAT

To verify our ideas, we designed and implemented an Unobtrusive User Attention Tracking(UUAT) system, which can accomplish two tasks after collecting user behavioral data: 1. Calculate the user reading region, BRR or RRR. 2. Infer what content(texts) and how long a user has read in an article. In this section we introduce our UUAT system and present our implementation details.

Since our goal is to analyze daily online reading, we prefer Browser-Client(BS) architecture to Client-Server(CS) architecture, for the reason that BS architecture is more scalable and platform-independent. We choose browser plugin/extension as our implementation architecture, for the following considerations: 1. Compatability. The plugin technology has been supported by all the mainstream web browsers. 2. Agility. The development cost and deployment cost of a plugin-based solution is very low. 3. Privacy. All the collected data is under full control of a user himself. The user can keep all data at his own computer and analyze it locally, or he can choose a cloud service to analyze his data.

Algorithm 1: Find the maximal ICS(intentional click set)

Input: The set of clicks $K = \{k_1, k_2, \ldots, k_n\}$
Output: The largest subset that fullfils equation 7
for $i \leftarrow 1$ ***to*** n **do**
    $S_i \leftarrow \{k_i\}$
for $i \leftarrow 1$ ***to*** n **do**
    if $S_i = \emptyset$ **then**
        continue;
    for $j \leftarrow 1$ ***to*** $i - 1$ **do**
        if $S_j \cup \{k_i\}$ *fulfills equation 7* **then**
            $S_j \leftarrow S_j \cup \{k_i\}$
            $S_i \leftarrow \emptyset$
$max \leftarrow ||S_1||$
for $i \leftarrow 1$ ***to*** n **do**
    if $||S_i|| > ||S_max||$ **then**
        $max \leftarrow i$
return $S_{\max}$

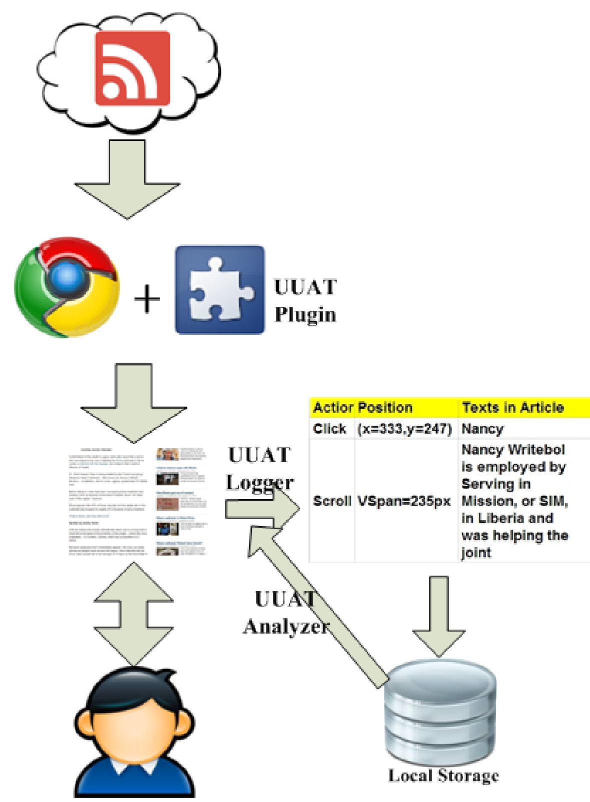

Figure 7. UUAT architecture

Figure 7 illustrates the architecture of UUAT as a chrome plugin. Once user requested an article from a website, the browser will first get the HTML contents for that article, then the html contents will be passed to UUAT and JavaScript code will be embedded into the original HTML page. When a user is reading this article, his behavioral data will be captured and saved at a local position. In real time, user behavior can be analyzed with the help of historical data(if there is historical data) and reading context data can be collected.

UUAT collects three types of raw interaction data: MouseClick, Cursor and Scroll. The click and cursor are relatively easy to capture, by assigning handlers to *document.onclick* and *document.onmousemove* event listeners, UUAT can easily log the position where the click happens and the cursor trajectory. The logging of user scroll is subtle. There are many cases (also in our collected data) a user scroll the screen more than one time in a very short time, during which no actual reading happened. Actually this happens very often when the user scrolls to a new position and aligns his reading region with the article contents before his scroll. To detect a real user scroll action, we adopted a look-backl strategy to rule out those "adjust-use" scrolls, which is shown in the following code piece.

```
var lastScrlStmp;

window.onscroll = logScroll(event);

function logScroll(event) {
        var timeStamp = new Date().getTime();
        lastScrlStmp = timeStamp;
        setTimeout(ifScroll,MIN_INTERVAL);
}

function ifScroll(){
        var curTmStmp = new Date().getTime();
        var diff = curTmStmp - lastScrlStmp;

        //real scroll detected
        if( diff > MIN_INTERVAL) {
                logScrollData();
        }
}
```

Behavioral Data Analysis

The goals of behavioral data analysis are twofold: 1. Combined with historical data, to calculate a preferred reading region so that in the future when less behavioral data is provided, an estimation of user reading can still be calculated by historical reading model. 2. For this reading process, to calculate what a user have read and to what extent he has read.

EXPERIMENT AND VERIFICATION

Experimental Design and Procedure

REFERENCES

1. Patients Like Me, howpublished = `http://www.patientslikeme.com/`, note = Accessed: 2010-09-30.
2. Atterer, R., and Lorenzi, P. A heatmap-based visualization for navigation within large web pages. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*, ACM (2008), 407–410.
3. Atterer, R., and Schmidt, A. Tracking the interaction of users with ajax applications for usability testing. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2007), 1347–1350.
4. Berry, M. J., and Linoff, G. *Data mining techniques: for marketing, sales, and customer support*. John Wiley & Sons, Inc., 1997.
5. Buscher, G., Biedert, R., Heinesch, D., and Dengel, A. Eye tracking analysis of preferred reading regions on the screen. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, ACM (2010), 3307–3312.
6. Celsi, R. L., and Olson, J. C. The role of involvement in attention and comprehension processes. *Journal of consumer research* (1988), 210–224.
7. Hornbæk, K., and Frøkjær, E. Reading patterns and usability in visualizations of electronic documents. *ACM Transactions on Computer-Human Interaction (TOCHI) 10*, 2 (2003), 119–149.
8. Huang, J., White, R. W., and Dumais, S. No clicks, no problem: using cursor movements to understand and improve search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2011), 1225–1234.
9. Lagun, D., and Agichtein, E. Viewser: Enabling large-scale remote user studies of web search examination and interaction. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, ACM (2011), 365–374.
10. Navalpakkam, V., and Churchill, E. Mouse tracking: measuring and predicting users' experience of web-based content. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2012), 2963–2972.
11. Nielsen, J., and Pernice, K. *Eyetracking web usability*. New Riders, 2010.
12. Ozzie, R. E., Gates III, W. H., Flake, G. W., Bergstraesser, T. F., Blinn, A. N., Brumme, C. W., Cheng, L., Connolly, M., Dani, N. V., Glasgow, D. A., et al. Personal data mining, Apr. 19 2011. US Patent 7,930,197.
13. Swan, M. Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified self-tracking. *International journal of environmental research and public health 6*, 2 (2009), 492–525.
14. Wolf, G. The data-driven life. *The New York Times 28* (2010), 2010.