Using Effective Reading Speed to Integrate Adaptivity into Web-based Learning

Muan Hong Ng[1], Wendy Hall[1], Pat Maier[2], and Ray Armstrong[3]

[1]Intelligence, Agents and Multimedia Research, University of Southampton, SO17 1BJ Southampton, United Kingdom.
{mhn99r, wh}@ecs.soton.ac.uk
[2] Center of Learning and Teaching, University of Southampton, SO17 1BJ Southampton, United Kingdom.
pjm@soton.ac.uk
[3] Rheumatology Unit, Mail Point 78, Southampton General Hospital, SO16 6YD, Southampton, United Kingdom.
rayarmstrong@btinternet.com

Abstract. It has always been difficult to determine to what extent a user has read a page especially in the area of educational adaptive hypermedia systems. We propose the use of an individual's effective reading speed to estimate how much of a page a user has read during their browsing activity. This method is currently used to apply history-based link annotation in a medical web-based learning application, JointZone. A validation test of this work has shown a positive result in approximating user's reading value when compared to conventional methods.

1 Introduction

A difficult problem faced by many adaptive hypermedia developers is to identify whether a page has been read. A generally unreliable solution is to use the viewing of a page as an indicator that the page has been read and understood (e.g. InterBook [2]). This method is obviously inefficient as the student might only be surfing through the pages without paying much attention to the content. In MANIC [11], the system assumes a high studied rating for a page if a user spends an optimal amount of time on it. This optimal time that is generated, based on the length of the content, is static for every user irrespective of their individual reading speed. However, Psychological research has shown that there are "astonishing differences in the rate of reading speed" among individuals [10]. It has also been found that two individuals reading at the same speed can have different rates of comprehension [5]. Motivation for this research comes from the users' unpredictable browsing behavior on the web, particularly the speed used in reading web pages. Research has shown that users frequently skim read the web pages instead of reading them in detail [4]. Studies have also shown that there is a general decline in the level of comprehension (especially in the recall of specific details) with faster reading as compared to reading at a normal speed [3, 7, 8]. Hence the reading speed, the time spent and the level of understanding on each page varies among individuals. The work in this paper captures this aspect of users' differences in reading speed and uses this to calculate how well they have read a web page based on the display time of the page.

2 Calculating the Effort Index

JointZone [13] is an interactive website for Rheumatology for both undergraduate medical students and practicing doctors. It contains approximately 80 documents, 750 images and 30 interactive case studies. The user model in JointZone contains an *effort index* (0-100%) for each page in the domain. This index is the system's estimation of the 'degree of effort spent' by a user on each page. The effort index is actually calculated by comparing the display time of a web page with an individual's optimal reading time. The optimal reading time for a page is identified for each user depending on the individual's reading rate, comprehension rate and prior knowledge of the domain. These user's characteristics are obtained when they use the system for the first time. Prior their first entry, a user will be asked to complete a knowledge test (10 multiple choice questions based on the general concept of the domain) to give a measurement of their prior knowledge (0-100%). Then each user has to take a 'reading speed test' where their reading rate is captured in number of words read per minute. The reading time is recorded through a start and stop button on the web page. This is followed by a comprehension test (six true or false questions on general ideas of the text) to give a score of their comprehension rate (0-100%). The combined measure of reading rate and comprehension gives the index of *effective reading speed* [6]. This index is commonly used in the commercial product for speed-reading courses [9, 12]. We have however, modified the equation by adding the factor of prior knowledge (see Equation 1).

$$\text{Effective reading speed} = \text{reading rate} \times \frac{1}{2}\left(\text{comprehension rate} + \text{prior knowledge}\right) \quad (1)$$

$$\text{Optimal reading time for a page} = \frac{\text{total number of words in the page}}{\text{Effective reading speed}} \quad (2)$$

$$\text{Effort Index Estimation, } G(x) = e^{\frac{-(x-optimalTime)^2}{2\sigma^2}}, \text{ where } \sigma = 1.23 \quad (3)$$

During a browsing session the optimal reading time for any page in the domain is calculated in real time using Equation 2. The effort index on each page is then estimated using a Gaussian function (see Equation 3) by comparing the actual time spent on a page (x) with the optimal reading time. If 'x' approaches the optimal reading time, a high effort index is assumed. However, one problem still exists, as it is difficult to determine if a user has indeed read a page when it is displayed on the screen. To tackle this, a heuristic cutoff point is used to give a zero effort index to cases where display time falls below eight second or three times the optimal reading time. On the other hand, the use of individual optimal reading time enable us to skew the effort index based on a user's prior knowledge since a good student who skims a page will gain a higher effort index (with lower optimal reading time) as compared to a poor student who spent the same amount of time on the page.

3 Validation Test

A validation test was carried out to study the correlation between the effort index and the users' understanding of the content itself. Thirty subjects took part in the evaluation. Prior to the experiment, all users are instructed to complete a series of reading tests to capture their individual effective reading speed. The analysis for thirty users shows a mean effective reading speed of 116 word/min (sd = 58.74), mean prior knowledge of 30.67% (sd = 15.52) and mean comprehension rate of 63% (sd = 15.93). In the experiment, they were asked to read a page (length = 443 words) in the domain followed by a performance test to examine how much they understand the page. This test contains five multiple-choice questions, which asked users to recall some important aspects in the text (main factual type) [3]. All users spent an average of 3.19 minutes reading the page (sd = 1.12 minutes). The effort index for each user was estimated by the system using Equation 3. The score of the performance test was then compared with the effort index. The Pearson correlation test showed a significant correlation between the effort index and the performance test ($r = 0.521$, $p = 0.003$). The performance score was also compared with a separate calculation of effort index using a standard optimal reading time for all users (2.25 minute using the average on-screen reading rate of 200wpm). The Pearson correlation test shows a non-significant correlation of 0.056 ($p = 0.767$). Hence, we can conclude that each user has a different optimal reading time, which gives a relatively better approximation of the user's understanding of the domain.

Fig. 1. An example of the use of effort index in the history-based link annotation in a personalized site map

4 History-based Link Annotation

The effort index is used as a basis to form the history-based link annotation [1] in the JointZone web-based medical learning. As shown in Figure 1, this adaptive feature is applied on a personalised site map to provide feedback to the user on the system's assumption of his/her reading value for each page. This site map acts as a navigational support mechanism for the users and helps them to quickly gain an overview of the domain on pages they have read or not read.

5 Conclusions and Future Work

The individual effective reading speed used in calculating the effort index on web pages gives us a relatively good approximation in judging how well a page has been read. However, the cognitive issues in measuring the user's understanding have yet to be addressed in this work. Hence we cannot claim that the effort index completely represents a user's understanding of the text on the page. Nevertheless, the validation test showed that our method gives a comparatively more accurate approximation than existing methods. With this encouraging result, we will continue to expand the use of history-based link annotation with the effort index. We will also undertake a set of evaluations to specifically analyze this adaptive procedure within JointZone, to discover if this feature aids the user.

Acknowledgements

Arthritis Research Campaign, UK, educational project grant A0549.

References

1. Brusilovsky, P.: Adaptive Hypermedia. Journal on User Modeling and User Adapted Interaction, 11 (2001) 87-110
2. Brusilovsky, P., Eklund, J., and Schwarz, E.: Web-based Education For All: A Tool for Developing Adaptive Courseware. In: Proceedings of Seventh International World Wide Web Conference, 14-18 April, 30 (1-7) (1998) 291-300
3. Dyson, M. and Haselgrove, M.: The Effects of Reading Speed and Reading Patterns on Our Understanding of Text Read From Screen. Journal of Research in Reading 23 (2000) 210-223
4. Horton, W., Taylor, L., Ignacio, A and Hoft, N. L.: The Web page Design Cookbook. John Wiley, New York (1996)
5. Huey, E. B.: The Psychology and Pedagogy of Reading. The Macmillan Company (1908)
6. Jackson, M. D. and McClelland, J.L.: Processing Determinants of Reading Speed. Journal of Experimental Psychology: General 108 (1979) 151-181
7. McConkie, G., Rayner, K. and Wilson, S.: Experimental Manipulation of Reading Strategies. Journal of Educational Psychology 65 (1973) 1-8
8. Poulton, E. C.: Time for Reading and Memory. The British Journal of Psychology 49 (1958) 230-245
9. ReadingSoft.com found at http://www.readingsoft.com
10. Romanes, G. J.: Mental Evolution in Animals. Appleton & Co, New York (1885).
11. Stern, M. and Woolf, B. P.: Adaptive Content in an Online Lecture System. In: Proceedings of the International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, Trento Italy (2000).
12. TurboRead Speed Reading found at http://www.turboread.com
13. JointZone – A Study for Rheumatology found at http://www.iam.ecs.soton.ac.uk/users/mhn99r/learn/userlogin/index.jsp