2007 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology - Workshops

Investigating User Browsing Behavior

Ganesan Velayathan
The Graduate University for Advanced Studies,
National Institute of Informatics
2-1-2, Hitotsubashi,Chiyoda,Tokyo,JAPAN
ganesan@nii.ac.jp

Seiji Yamada
National Institute of Informatics
2-1-2, Hitotsubashi,Chiyoda,Tokyo,JAPAN
seiji@nii.ac.jp

Abstract

This paper describes our efforts to investigate factors in user browsing behavior to automatically evaluate Web pages that the user shows interest in. We developed a client site logging tool to monitor and log the user's browsing behavior. We performed user experiment using ten participants to collect the browsing behavior, and evaluated the behaviors by performing classification learning using C4.5. We generated common user browsing behavior rules and evaluated these common rules against the individual participant data. This paper reports those findings.

1. Introduction

In recent years, systems enabling Web personalization for individuals have gained particular attention from commercial companies such as My Yahoo [1], and Google Personalized Search [2] and researchers. Web personalization can be described as actions "*that make the Web experience of a user personalized to the user's taste*" [4]. Most previous research and tools for Web personalization have relied upon overt methods of asking users for their answers in order to construct user profiles. Building a user profile that adapts to a user's daily interests is a challenging task. This is because it is hard to predict which Web sites most interest the user without asking the user to interact explicitly with the system. A user must interact with the system explicitly and tell the system each time a webpage is relevant to the task or not. This is laborious; it requires time and inclination, and users often forget to notify the system. A less intrusive approach to the construction of user profiles is required.

With this in mind, this paper proposes a new method of automatically discovering and judging user interest based on user browsing behavior. **User browsing behavior** is defined here as **the habitual actions performed by users when browsing or searching**, such as clicking links, bookmarking, printing or selecting text, and so on. If we can detect and learn these patterns of user behavior, we can use these tendencies to perform automatic page evaluation without placing any burden on the user. We are also hypothesizing that user browsing strategies and habits do not change greatly over time. Furthermore, we believe that browsing behavior is unlikely to change even if a user is browsing in a different language environment, which means a language-independent method of evaluating Web pages can be constructed. By using this method, we believe that a highly accurate automatic self-constructed user profile can be created. This study is similar to the work done by Fox et al [5]; however, the scope of their work was restricted to search engine result pages and subsequent clickthroughs. The approach of building a customized browser is similar to Claypool's Curious Browser [3]; however, we focus on keeping the browser as similar as possible to Internet Explorer.

Starting from these considerations, we undertook construction of the GINIS Framework[13]. We used the GINIS Framework to conduct experiments with ten participants. This is a follow up work to our previous work which can be found in [13, 14, 15]. This paper does not present in detail all the ideas behind this research. For better understanding of this work, ideas on building the client side logging mechanism can be found in [13], and the statistical evaluations and the rules generated can be found in [15].

This paper focuses on evaluation of the common rules generated using C4.5. We evaluated in detail the rules generated from the collected data and tested the rules with test data from each individual user. We generated common rules using all ten participant's data, and used each participant's individual data to test the fitness of the common rules. Furthermore, we generated common rules for five participants (participants A-E) and tested the generated rules against the other five participants (participants F-J) , and vice versa.

[1] http://my.yahoo.com/
[2] http://labs.google.com/

0-7695-3028-1/07 $25.00 © 2007 IEEE
DOI 10.1109/WI-IATW.2007.122

2 Current Practice and Research

2.1 Web Navigation

In the field of Web navigation, many previous studies were based on data found in server logs for analyzing various aspects of user navigation. However, data available in server logs only shows user behavior for a single site, and it is difficult to collect and aggregate all the logs to discover common rules of user behavior. Shahabi and Chen [11] have pointed out that Web server logs might be inaccurate because Web usage data from the server side are not reliable.

Little research has been done relating to client side web navigation. One of the latest studies is the work presented by Weinreich et al [16]. Weinreich focused on three aspects of Web navigation: changes in the distribution of navigation actions, speed of navigation, and within-page navigation.

2.2 Personalization

Personalization is the process of presenting the right information to the right user at the right moment [12]. In most cases, personalization begins with building a user profile based on user interest. Interest data can be collected either explicitly by asking for feedback from the user regarding preferences or implicitly by observing user behavior, such as time spent reading a Web page. Research such as Chaffee's[2] proposes building user preferences automatically by studying user browsing history.

2.3 Implicit Feedback

Implicit feedback techniques unobtrusively obtain information about users by observing their natural interactions with the system [7]. The user behaviors most extensively investigated as sources of implicit feedback include reading time, as well as saving, printing and selection behaviors. The primary advantage to using implicit techniques is that such techniques remove the user burden from providing explicit feedback [6].

Morita and Shinoda [8] investigated reading time as a feedback factor, Seo and Zhang [10] considered user browsing patterns, and Kelly and Belkin [7] investigated the use of reading time and scrolling as an indication of relevance. Fox et al. [5] found that dwell time, position, scroll count, and exit type are predictive actions of relevance judgments for individual Web pages and that dwell time, number of results listings, and exit type are more predictive of overall session satisfaction.

However, there are many other user browsing behaviors to consider. This point remains in question, continuing to motivate research towards its resolution.

3 Methodology

When browsing or searching the Web, we interact with the browser's user interface. For example, we print, save, bookmark, move the mouse, click a link, and so on. We use the term **navigation action** here to describe the individual "components" of user behavior in performing actions using the browser directly. We use the term **user behavior** to describe the result achieved by performing these navigation actions. For example, the navigation actions "Hit Backspace Key," "Click Back Arrow on Menu," or "Use Back Button on Mouse" all constitute the same "Move Back" user behavior.

A preliminary study was first made, monitoring one user's browsing behavior using video and screen shot captures. After carefully studying these screen shots and taking into consideration the technical programming restrictions, we chose around 40 specific user behaviors and 70 navigation actions and built a monitor for the browser. The logging attributes were designed based on the work presented by Catledge and Pitkow [1], with some additions.

During the experimental stage, we gathered this information at the navigation action level, after which we preprocessed the log and construed user behavior. So, for example, if a user performed "Click Somewhere on Browser" and "Drag Mouse while Holding Left Mouse Button" navigation actions, this would be collectively recorded as " Text Highlight."

The experiment for collecting user behavior was divided into two separate parts. The first part was the learning part, during which the users were prompted to answer every time a "Next Page" navigation behavior was undertaken (by using any navigation action to move from a page to a new page). On this questionnaire, users could choose "like," "dislike," or "unknown" in regard to the page indicated. The default answer was set to "unknown," as users tend to click without really considering the question whenever they are busy with deadlines etc.

The learning engine refined the logs from the raw log database and stored it in the Behavioral Database (BDB). Classification learning (building a decision tree using C4.5) was performed based on the information from the BDB, and user interest rules were created and stored in the Rule Database (Rule DB). During the testing stage, the prediction engine compared the user's new behavior with the behavioral patterns stored in the Rule DB, and the user's implicit interest in a particular page was predicted based on this comparison. Figure 1 show an overview of the steps of the system built for this research.

The GINIS Framework consists of four main modules: a client interface to detect and log user behavior (the browser), a database to store the user log information (the raw logger), a learning engine, and a prediction engine.

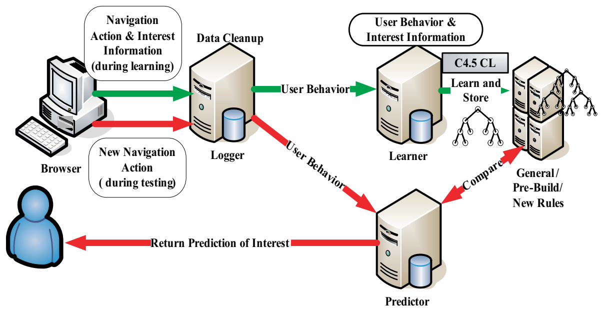

Figure 1. Overview of GINIS Frameworks

4 Experiments and Results

4.1 Overview

We conducted systematic experiments using the GINIS Framework with ten participants (6 male and 4 female) to gather data in order to discover new rules linking user behavior and interest. C4.5[9] was used as the classification algorithm. Verification was performed using the 10-fold validation method. The total number of behaviors logged from the users was 64,312. There was an average 5.2 user behaviors per page, with the total visited pages during this experiments being 12,368 pages.

After gathering all ten participants' behavior logs, and removing the instances of inconsistency, C4.5 classification learning was performed. The C4.5 error rate was set at 25%. Here, a "case" means a set of user behaviors or navigation action data and the evaluation of the Web page when such behaviors take place. The term "inconsistency" means that the cases of training data have the same type of behavior occurrence but different kind of user evaluation.

4.2 Classification Learning and Rules

As a result, 2249 cases were used as training data, and within these 1885 were judged as "of interest" and 364 were judged as "not of interest" by the user. Of these evaluations, 2005 pages (89.15%) were correctly classified, and 244 pages (10.85%) were incorrectly classified.

Twenty-six rules were output in total by C4.5. Five of these were found to be rules governing "of interest", and 21 of these were rules governing "not of interest." The default class was "of interest". Details of generated rules are not shown due to lack of space. Statistical evaluation of the generated rules can be found in our previous work [15]. This workshop paper focuses on evaluating the fitness of the generated rules.

To further the evaluation of the generated rules, we performed fitness evaluation using two different methods. The first method is by generating common rules using data from all ten participants and testing the fitness of generated rules with each individual participant's data.

However, since evaluation of the test data was performed with the basic data used to create the rules, other methods of evaluation are required. We furthered the evaluation of our experiment by building new rules using five participants' data and evaluating the rules' fitness with the other five participants' data. This evaluation process was divided into two parts. The first part was where the rules using the C4.5 classifier system was generated using participants A-E and tested against participants F-J. The second part was where the rules were generated using the data of participants F-J and tested against participants A-E.

Figure 2 indicates the results of this fitness test. The bar chart indicates the total data used for each participant (the collected data from each participant which was used for testing the fitness of the rules). The non-dotted line (color: red) indicates the error value of the fitness test (in %) of each data. The dotted line (color: green) indicates the fitness of A-E data tested against the rules generated using F-J data, and also the rules generated using A-E data and tested against F-J data. These were combined together into one graph due to lack of space.

4.3 Discussion

As with the first evaluation part, testing of the learning part was performed using the same data. T resulted in high rule fitness. An average of 11.10% rule fitness error was obtained. As for the second part, further evaluation of new rules generated, this resulted in average overall fitness error of 18.02%.

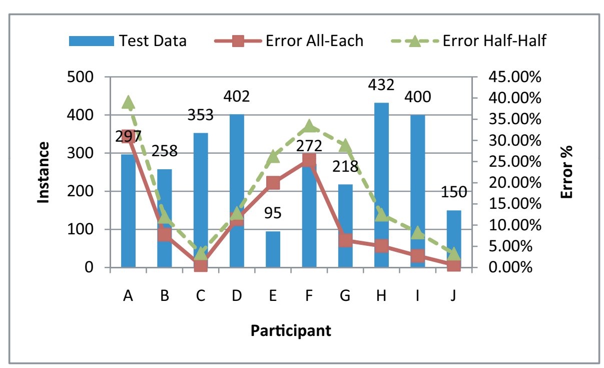

Figure 2. Testing Common Rule's Fitness against Participant's Data

5 Present Limitations and Future Work

scroll The situation and the nature of occurrence of scrolling behavior were not considered in detail.

mouse Mouse locus data were not used to constitute behavior.

order The order of navigation/behavior was not considered.

session Actions and behaviors were recorded over a single page, not over a sequence of navigation.

comparison No comparisons were made with other classification algorithms.

Considering these present restrictions, we are working towards building more user behavior from the available collected experimental data. In particular, mouse locus data and scroll action cover almost 50% of all collected data, and further analyzing these behaviors will result in better user behavior creation, thus further improving classification learning results and enabling better rule creation.

6 Conclusions

We have presented here a method to automatically detect and log user behavior at the client-side by creating a client-side browser. We used this framework to conduct an open environment field study. Based on the generated common rules from the users we conducted rule fitness evaluation. We successfully confirmed that all the rules generated by C4.5 classification learning algorithm have 60% or higher fitness, where the error for all the tests was lower than 40%. As for the average rule fitness error, in total both of our evaluation methods indicates that the fitness error is less than 20%. From this, we can conclude that the rules generated using C4.5 have high fitness.

References

[1] L. D. Catledge and J. E. Pitkow. Characterizing browsing strategies in the world-wide web. In *Proceedings of the Third International World-Wide Web Conference on Technology, Tools and Applications*, pages 1065–1073, New York, NY, USA, 1995. Elsevier North-Holland, Inc.

[2] J. Chaffee and S. Gauch. Personal ontologies for web navigation. In *CIKM '00: Proceedings of the 9th International Conference on Information and Knowledge Management*, pages 227–234, New York, NY, USA, 2000. ACM Press.

[3] M. Claypool, P. Le, M. Wased, and D. Brown. Implicit interest indicators. In *IUI '01: Proceedings of the 6th International Conference on Intelligent User Interfaces*, pages 33–40, New York, NY, USA, 2001. ACM Press.

[4] M. Eirinaki and M. Vazirgiannis. Web mining for web personalization. *ACM Trans. Inter. Tech.*, 3(1):1–27, 2003.

[5] S. Fox, K. Karnawat, M. Mydland, S. Dumais, and T. White. Evaluating implicit measures to improve web search. *ACM Transactions on Information Systems (TOIS)*, 23(2):147–168, 2005.

[6] B. J. Jansen and M. D. McNeese. Evaluating the effectiveness of and patterns of interactions with automated searching assistance: Research articles. *Journal of the American Society for Information Science and Technology*, 56(14):1480–1503, 2005.

[7] D. Kelly and N. J. Belkin. Reading time, scrolling and interaction: exploring implicit sources of user preferences for relevance feedback. In *SIGIR '01: Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 408–409, New York, NY, USA, 2001. ACM Press.

[8] M. Morita and Y. Shinoda. Information filtering based on user behavior analysis and best match text retrieval. In *SIGIR '94: Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 272–281, New York, NY, USA, 1994. Springer-Verlag New York, Inc.

[9] J. R. Quinlan. Induction of decision trees. *Machine Learning*, 1:81–106, 1986.

[10] Y.-W. Seo and B.-T. Zhang. Learning user's preferences by analyzing web-browsing behaviors. In *AGENTS '00: Proceedings of the 4th International Conference on Autonomous Agents*, pages 381–387, New York, NY, USA, 2000. ACM Press.

[11] C. Shahabi and Y.-S. Chen. An adaptive recommendation system without explicit acquisition of user relevance feedback. *Distributed and Parallel Databases*, 14(2):173–192, 2003.

[12] M. Speretta and S. Gauch. Personalized search based on user search histories. In *WEBIST'05: Proceedings of the International Converence on Web Information Systems and Technologies*, pages 270–277, 2005.

[13] G. Velayathan and S. Yamada. The ginis framework: Interaction-based evaluation of web pages. In *iiWAS'05:Proceedings of the 7th International Conference on Information Integration and Web-based Applications and Services*, pages 303–312, 2005.

[14] G. Velayathan and S. Yamada. Behavior-based web page evaluation. In *WI-IATW '06: Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pages 409–412, Washington, DC, USA, 2006. IEEE Computer Society.

[15] G. Velayathan and S. Yamada. Can we find common rules of browsing behavior? In *QUERY LOG ANALYSIS, WWW'07: Query Log Analysis: Social and Technological Challenges, A workshop at the 16th International Conference on World Wide Web*, New York, NY, USA, 2007. ACM Press.

[16] H. Weinreich, H. Obendorf, E. Herder, and M. Mayer. Off the beaten tracks: exploring three aspects of web navigation. In *WWW '06: Proceedings of the 15th International Conference on World Wide Web*, pages 133–142, New York, NY, USA, 2006. ACM Press.