Implicit Interest Indicators

Mark Claypool, Phong Le, Makoto Waseda and David Brown
Computer Science Department
Worcester Polytechnic Institute
100 Institute Road
Worcester, MA 01609

{claypool—dcb}@cs.wpi.edu

ABSTRACT

Recommender systems provide personalized suggestions about items that users will find interesting. Typically, recommender systems require a user interface that can "intelligently" determine the interest of a user and use this information to make suggestions. The common solution, "explicit ratings", where users tell the system what they think about a piece of information, is well-understood and fairly precise. However, having to stop to enter explicit ratings can alter normal patterns of browsing and reading. A more "intelligent" method is to use *implicit ratings*, where a rating is obtained by a method other than obtaining it directly from the user. These implicit interest indicators have obvious advantages, including removing the cost of the user rating, and that every user interaction with the system can contribute to an implicit rating.

Current recommender systems mostly do not use implicit ratings, nor is the ability of implicit ratings to predict actual user interest well-understood. This research studies the correlation between various implicit ratings and the explicit rating for a single Web page. A Web browser was developed to record the user's actions (implicit ratings) and the explicit rating of a page. Actions included mouse clicks, mouse movement, scrolling and elapsed time. This browser was used by over 70 people that browsed more than 2500 Web pages.

Using the data collected by the browser, the individual implicit ratings and some combinations of implicit ratings were analyzed and compared with the explicit rating. We found that the time spent on a page, the amount of scrolling on a page and the combination of time and scrolling had a strong correlation with explicit interest, while individual scrolling methods and mouse-clicks were ineffective in predicting explicit interest.

1. INTRODUCTION

One way that intelligent user interfaces can be "intelligent" is to understand the intentions of the user. The high-level is goal is to understand via interpreting sequences of actions. The low-level goal is to understand simple actions, such as scrolling down the text on a Web page or bookmarking a Web page. Intelligently understanding the interest of a user in an item is critical for many systems that use intelligent user interfaces, particularly recommender systems [9, 17, 7, 5] that provide personalized suggestions about items of interest.

In order to adaptively recommend information a system must have "ratings" on each item from each user. The most common and obvious solution is for the interface to use *explicit ratings*, where users tell the system what they think about some object (e.g., a music CD) or piece of information (e.g., a Newspaper article). Explicit ratings are well-understood, fairly precise [18], and are common in everyday life, due to movie reviews, restaurant 'stars', *etcetera.*

However:

- Having to stop to enter explicit ratings can alter normal patterns of browsing and reading;
- Unless users perceive that there is a benefit from providing ratings, they may stop providing them [4]. Hence, users may continue to read, resulting in system use, but no ratings at all [1];
- Research on the GroupLens system [16] found that with explicit ratings, users were reading a lot more articles than they were rating; and
- Collaborative filtering requires many ratings to be entered for every item in the system in order to provide accurate predictions (i.e., the "sparsity" problem) [16].

Hence, explicit ratings, while common and trusted, may not be as reliable as is often presumed. The solution? Use *implicit ratings.* An implicit rating is a rating that is obtained by a method other than obtaining it 'directly' from the user. Obvious advantages of implicit ratings are:

- they remove the cost of the user examining and rating items;
- potentially, every user interaction with the system (and, sometimes, the absence thereof) can contribute to an implicit rating.

Although each implicit rating is likely to be less accurate than an explicit rating, they:

- can be gathered for "free";

IUI'01, January 14-17, 2001, Santa Fe, New Mexico, USA.

- can be combined with other implicit ratings for a more accurate rating; and
- can be combined with explicit ratings for an enhanced rating (countering, for example, the "what I say is not what I want" problem).

We believe that the capture and use of implicit ratings has significant benefits yet poses significant challenges that have yet to be investigated.

The main objective of this research is to collect, measure and evaluate the predictive power of some promising implicit interest indicators. We concentrate on interest/approval indicators for a single, current Web page, based on a single behavioral sign or a pattern of behavior. To accurately gather implicit interest indicators, we developed a Web browser, called *The Curious Browser*, that allows us to capture users actions as they browse the Web. We deployed the browser in a user study with over 80 people browsing over 2500 Web pages.

We analyzed the individual implicit ratings and some combinations of implicit ratings and compared them with the explicit ratings. We found that the time spent on a page, the amount of scrolling on a page and the combination of time and scrolling had a strong correlation with explicit interest, while individual scrolling methods and mouse-clicks were ineffective in predicting explicit interest. Moreover, implicit interest indicators may be as effective as explicit interest indicators in terms of accurate coverage while having none of the user-costs from explicitly requesting user interest.

The contributions of this work are:

- Experimentally-based statistical analysis of the correlation between the implicit interest indicators of mouse activity, keyboard activity and time with explicit interest.
- A categorization of implicit interest indicators.
- A Web browser that records a variety of implicit interest indicators.[1]
- The dataset from the user experiments.[2]

The rest of this paper is as follows: Section 2 describes related work in gathering implicit interest indicators; Section 3 describes a general categorization of interest indicators; Section 4 details our approach towards gathering implicit interest indicators; Section 5 describes our user study experiments and results; Section 6 analyzes the results from the experiments; Section 7 presents our conclusions; and Section 8 mentions some possible future work.

2. RELATED RESEARCH

We have divided related work in implicit ratings into three categories: work that discusses the concept and application of implicit ratings, work that uses the time spent accessing an item as an implicit rating, and work that uses marking an item as an implicit rating.

[1]Download from: `http://perform.wpi.edu/`

[2]Download from: `http://perform.wpi.edu/`

2.1 Concepts

Nichols [13] discusses the costs and benefits of using implicit ratings for information filtering applications. He categorizes implicit ratings by the actions a user may perform, such as "Examine" for reading a whole item, or "Save" for saving, bookmarking or printing an item. He observes that the limited evidence suggests that implicit ratings may have great potential, but that there has been little experimental work evaluating their effectiveness. He identifies that properly understood implicit ratings may be used in several ways: the first is to provide more ratings upon which to base predictions, and the second is as a check on explicit ratings to decide when to ignore them or not. We propose to provide experimental evaluation of the effectiveness of implicit ratings.

Oard and Kim [14] build upon work by Nichols [13] by categorizing implicit ratings, dividing them into "Examination", where a user studies an item, "Retention" where a user saves an item for later use, and "Reference" where a user links all or part of an item into another item. They suggest two strategies for using implicit ratings. Our work proposes to experimentally evaluate one of their two strategies using implicit ratings from one of the three categories proposed.

2.2 Experiments on Examining

Morita and Shinoda [12] study the amount of time spent reading a Usenet News article. They examined users in a carefully controlled experimental environment in which users were not allowed to interrupt their reading and only read a carefully chosen news domain. They find that the 'time' people spend reading Net News articles is the primary indication of them having interest in it. However, they find no correlation between reading time and message length or reading difficulty level. We propose to extend the study of implicit ratings into a less well-controlled environment, with more types of implicit ratings, to see if their statistically significant results still hold. In addition, the "controlled" nature of their experiments may have reduced the accuracy of their studies, since in our experience [2], when you instruct participants to read and rate articles, they actually spend time reading them even if they do not find them interesting. This may make the time/interest correlation even weaker.

Konstan et al [9] describe how the GroupLens system for filtering Usenet News studied the correlation between time spent reading an article and the explicit ratings. They could obtain substantially more ratings by using implicit ratings, and predictions based on time spent reading are nearly as accurate as predictions based on explicit ratings. They also provide confirmation of the results of Morita and Shinoda [12]. Our work seeks to extend their experiments into alternative domains, as well as to greatly expand the number of implicit ratings examined.

Goecks and Shavlik [3] measure browsing activity in an attempt to predict the future activity of the user. They modify Microsoft's Internet Explorer to measure the amount of mouse and scrolling activity. A single user browsed the web looking for specific documents while their modified browser collected data. A neural network was trained on the data, to see if they could accurately predict user activity on other documents the user did not read. While they were able to accurately predict user behavior for some unread documents, their evaluation did not ascertain how well the user activity

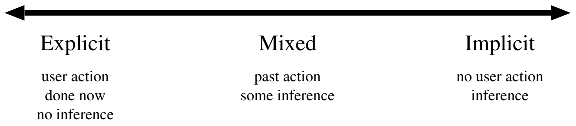

Figure 1: Explicit/Implicit Dimension of Interest.

correlates with user interest. Their methodology for gathering Implicit Interest Indicators may prove valuable for our experiments, and our proposed work will similarly analyze mouse movement and scrolling. In addition, we will analyze additional user activities, while correlating the data to explicit interest.

2.3 Experiments on Marking

Hill et al [8] monitor "read" and "edit" actions on a document. The amount of time spent reading or editing an item is termed the "wear" on the item, and is implicitly assumed to indicate interest. However, these implicit ratings were not analyzed to determine how accurately they correlated with interest, but were merely displayed in a scrollbar so that users can infer interest themselves by the "wear" provided by other users. In addition to their time study, Morita and Shinoda [12] record the actions (marks) on the Usenet News articles: posted, saved or followed-up. They hypothesize that this data could be useful for predicting interest. However, they do not analyze the correlation with user interest. Our work provides a methodology for doing this.

Siteseer [15] uses the overlap between bookmark files to determine similarity among individuals. A user's bookmarks are assumed to imply interest. The correlation among bookmarks, is similar to the Fab system described above. Our research proposes to study to what degree implicit interest indicators do, in fact, indicate interest. This would allow systems such as Siteseer and Fab to adjust their prediction algorithms accordingly.

Letizia [11] uses different levels of marking to imply different amounts of interest. Letizia, which works in a web-based environment, infers that saving a reference to an item implies a strong amount of interest, following a link implies a tentative amount of interest, repeated visits indicate an increasing amount of interest, and passing over a link indicates no interest unless the item is selected later. Our work proposes to explicitly measure the level of interest for similar interest indicators.

3. INTEREST INDICATOR CATEGORIES

Implicit interest indicators can be categorized in a variety of ways. The most basic is to consider them on an Implicit/Explicit dimension, as depicted in Figure 1.

This dimension is based on the time at which the user the provided input (i.e., an action), and on whether, and how much, inference is needed. The time might be "now", at the time of viewing the page (e.g., explicit rating) or earlier (e.g., user provided keywords). By "user action" we mean an action that is 'intended' to indicate interest. An example of Explicit is "providing a rating", of Mixed is "keyword match", and of Implicit is "time spent reading". While this dimension clearly needs some additional study and refinement (e.g., as it mixes action, intent and inference), another beneficial view is to consider 'what' the user's input is.

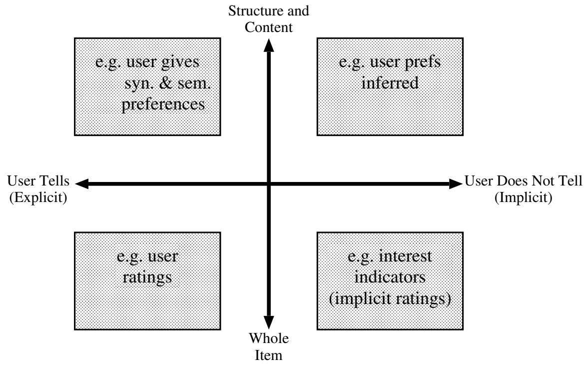

Figure 2: Categorizing Interest Indicators.

Figure 2 extends Figure 1 by depicting a two-dimensional representation of all interest indicators. The horizontal axis represents how explicit or implicit the interest indicator is. The vertical axis represents whether the interest indication comes from the structure or content of the item or from the whole item. Explicit interest ratings are at the bottom left of the Figure. The implicit interest indicators we propose to measure are in the bottom middle to bottom right of the Figure. Another categorization:

- *Explicit Interest Indicators.* To explicitly indicate interest, a user might select an interest value from a 'scale' that provides continuous levels. Alternatively, they can be asked to select from 'degree of interest' buttons, representing a fixed scale.

- *Marking Interest Indicators.* Various user actions might be considered as a form of marking, and can be interpreted as interest. These include bookmarking a Web page, deleting a bookmark, saving the page as a file, emailing the page, or printing it.

- *Manipulation Interest Indicators.* Some actions, such as cutting and pasting, can be considered as 'manipulation'. Others include opening a new browser window (i.e., perhaps the user is keeping the current browser window open to its current page because it is interesting), searching in the page for text, or scrolling.

- *Navigation Interest Indicators.* If the user spends time with the page open, follows, or doesn't follow a link, then we can consider these to be forms of 'navigation' indicators.

- *External Interest Indicators.* External indicators are concerned with the user's 'physical' responses to information, such as heart-rate, perspiration, temperature, emotions and eye movements. While clearly difficult to obtain directly without special instrumentation, some physical responses might be inferred from user actions. For example, eye movements might be indicated by the user 'following along' through the text with the cursor, or circling text with the cursor, while emotional response might be indicated by rapid changes in the rate of interaction.

- *Repetition Interest Indicators.* In general, we can hypothesize that doing 'more' of something means more interest. Thus inferences might be made from the user spending more time on a page, doing lots of scrolling through a page, and repeatedly visits to the same page.

- *Negative Interest Indicators.* Absence of an indicator might be considered to be a "negative" indicator. We suspect that there are some negative indicators that are worth including. The problem with this approach is that it is very difficult to distinguish between, for example, deliberately not visiting a page, and merely just not visiting it. However, one could accumulate evidence in order to increase the reliability of the indicator. For example, if a user is 'touring' a web site, and on many occasions is only one link (i.e., one click) away from visiting a web page, then we can assume with some confidence that this web page is not of interest.

It is worth noting that some indicators may be context sensitive, depending on the user's task/goal (e.g., browsing versus searching), or the "category" of the page: i.e., whether it is a page of links in a menu-style, or just plain text with embedded links. This might effect the importance of links 'not' taken. In general, layout has an effect on page function, which affects the user's behavior.

In addition, different combinations of indicators might mean different things. For example, if a user does not read a document for very long, but they do bookmark it, the short time might suggest that they do not like the page, while the bookmark might suggest that they do. In this case, they probably bookmarked it for later reading and we do not yet know if they like it or not.

4. APPROACH

Our approach is to experimentally measure and analyze several promising indicators presented in the previous section (Section 3), in order to ascertain their effectiveness in predicting explicit interest. We used the following methodology:

- Implement a browser to capture gather data on as many Implicit Interest Indicators as possible.
- Conduct a user study with many participants browsing the Web with our custom browser.
- Analyze correlation between implicit interest indicators gathered and explicit interest.

This section details the Web browser we implemented, called *The Curious Browser*, to capture some implicit interest indicators from user actions as they browsed the Web. The Curious Browser provides a Graphical User Interface (GUI) that also captures mouse and the keyboard actions as the user browses the Web. The first time each Web page is visited, the Curious Browser stores the user name, the URL, the time and date, the explicit rating and all implicit interest indicators. Subsequent returns to the same page are not recorded.

4.1 Graphical User Interface

The graphical user interface is written with Microsoft's Internet Explorer (version 5.0) in mind, with additional buttons for evaluation, user study instructions, and exiting. Figure 3 shows the main interface of the Curious Browser.

Figure 3: The Curious Browser. This is a screen shot of the main interface.

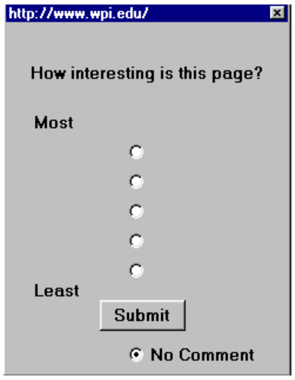

Figure 4: Evaluation Window. This is a screen capture of the window that pops up for users to give their explicit rating of the current Web page.

As in normal Web browsing, clicking on a link will load the appropriate Web page. However, before the current Web page is closed, the user is presented with an evaluation window that prompts the user for their explicit rating on the page just visited (see Section 4.5). Figure 4 shows a screen-capture of the evaluation window. The explicit rating is indicated by checking one of five unlabeled radio buttons presented with a scale labeled "least" to "most" interest. There is a sixth button labeled "no comment" that is the default button selected.

4.2 Mouse Activities

The Curious Browser captures two mouse activities: the number of mouse clicks and the time spent moving the mouse, in milliseconds. Mouse activities are only captured when the mouse is inside the bowser window and the browser is in focus. The mouse is out of the browser window when the mouse cursor is out of the main HTML page, the vertical scroll bar, and the horizontal scroll bar. The browser window is not focused when a user activates another application. The mouse activities are accumulated for each user while on the page.

4.3 Scrollbar Activities

The Curious Browser captures two kinds of scrollbar activities: the number of mouse events (clicks) on the horizontal and vertical scroll bars and time spent scrolling. Similar to the mouse activities, scrolling activities are only captured when the mouse is inside the browser window and the browser is in focus.

4.4 Keyboard Activities

As some people prefer using a keyboard to scroll instead of the mouse, the Curious Browser captures action on 4 keys: Page Up, Page Down, Up Arrow and Down Arrow. There are two different keyboard activities: the number of times that a user holds down these keys; and the other is the amount of time, in milliseconds, that these keys were held down. We store the data separately for each key.

4.5 Explicit Ratings

The Curious Browser explicitly asks for ratings (using the window shown in Figure 4) whenever the user changes from one page to another. This is typically done by following a link, but can also be done by pushing the Evaluation button. There are also several ways to change a page to another: push the Back button, push the Forward button, or type a URL address directly into the Address Bar and hit the Enter key. In addition, the user can select the Evaluation button at any time to enter an explicit rating.

5. EXPERIMENTS

We installed the Curious Browser on about 40 PC's running Microsoft Windows 98 on a computer lab open to all WPI students and a private computer lab open to only computer science students enrolled in our *Webware* (cs4241) course.

Students from a Human-Computer Interaction class (cs3041) as well as students from *Webware* were encouraged to participate in the user study experiments. Students were instructed to open up the Curious Browser and browse the Web for 20-30 minutes, but were not told the purpose of the experiments.

The Curious Browser was available from March 20, 2000 to March 31, 2000. During this time, 75 students visited a total of 2267 Web pages. 72 of the students visited all their Web pages in one session while 3 students had 2 sessions. They provided explicit ratings on only 1823 (80%) of the URL's (the others were "no comment"). Figure 5 depicts a histogram of the rating breakdown. The mean explicit rating was 3.3.

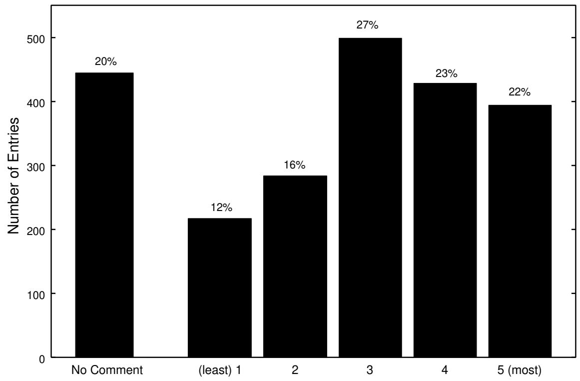

Figure 5: Explicit Rating Histogram. This figure shows the number of each explicit rating, along with its percentage of all ratings.

6. ANALYSIS

The implicit interest indicators we analyze in this section are:

1. The time spent on a page (Section 6.1).
2. The time spent moving the mouse (Section 6.2).
3. The number of mouse clicks (Section 6.3).
4. The time spent scrolling (Section 6.4).

In addition, we analyze the coverage and accuracy of different types of implicit interest indicators (Section 6.5).

Initially, we analyzed the mean of each implicit interest indicator versus the explicit rating. However, because of some extreme outliers, the mean of the implicit indicator proved to be a poor indicator of explicit interest. Thus, we focus on the median and distribution of each indicator using a Kruskal-Wallis test[3] (based on .05 level of significance) to examine the degree of independence of the medians among each explicit rating groups for each implicit interest indicator. Details on the test results can be found in [10][4], but are only summarized here due to lack of space.

We present the results below showing a box-plot, where the box represents the range of values from the bottom quartile (25%) to the top quartile (75%) and the median is depicted by a line in the middle. Although typical box-plots are extended on the top and bottom by two "whiskers" that extend to the full range of values, most of the whiskers are cropped in the below figures.

6.1 Time on Page versus Explicit Rating

The time spent on a page is captured immediately after loading the page until right before the page is exited. It includes all the actions and the actual reading time for the page, but does not include the time that the Curious Browser is not in focus. Thus, factors that influence its accuracy include loading time (which, in turn, depends upon speed of connection, CPU speed and the amount of Internet traffic) and how much of the active window time the user actually spends looking at the Web page (as opposed to going out for coffee). Before running the test, we filtered out 91 outliers: 4 data points that have more than 1,200,000 milliseconds (about 20 minutes) spent on a page as the users had likely stopped reading the page, and 87 data points that had less than 1000 milliseconds (1 second) spent on a page as we believe users cannot accurately assess interest in a page in less than 1 second.

Figure 6 depicts a box-plot of the time spent on a page versus the explicit rating. The Kruskal-Wallis rejected the null hypothesis (that the median values are the same), mean-

[3]Details on the Kruskal-Wallis test can be found in typical statistics books.

[4]On the Web at: `http://perform.wpi.edu/`

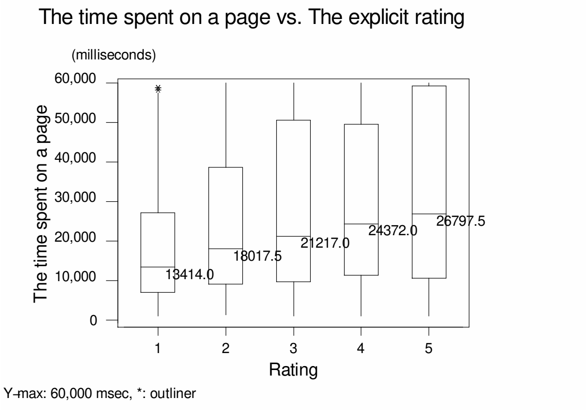

Figure 6: Time versus Explicit Rating.

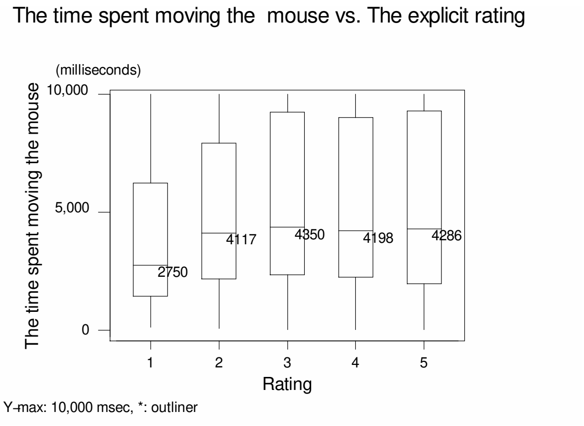

Figure 7: Time Moving Mouse versus Explicit Rating.

ing that the median values for each explicit rating group differed. Our conclusion is that the total time spent on a Web page is a good indicator of interest. This is a more general result than found in [12] and [16] which showed the correlation between time spent reading a News article and interest.

6.2 Time Moving Mouse versus Explicit Rating

The time spent moving the mouse is measured as the total time the mouse position is changing inside the active browser. Some users move the mouse while reading the window text or looking at interesting objects on the page, while others move the mouse only to click on interesting links. Either way, we hypothesized that the more mouse movement, the more interesting a user would find the page.

Figure 7 depicts a box-plot of the time spent on a page versus the explicit rating. The results from the Kruskal-Wallis test rejected the null hypothesis, meaning that the median values for each explicit rating group differed.

The median for a rating of 1 is significantly less than the median for the other explicit rating groups. The other explicit rating groups (2-5) have only small differences in the median and distribution. Thus, we can observe that the time spent moving the mouse is directly proportional to the

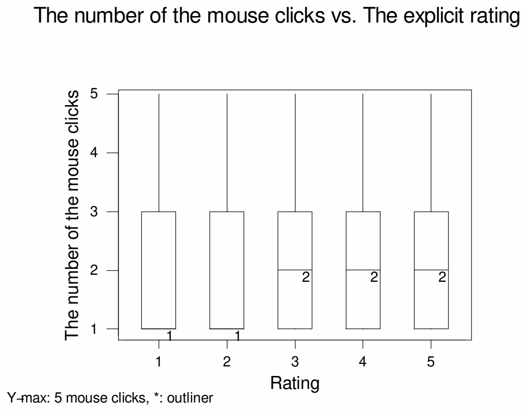

Figure 8: Number of Mouse Clicks versus Explicit Rating.

explicit rating. However, they are not linearly proportional to the explicit rating.

Our conclusion is that there is a positive relationship between the time spent moving the mouse and the explicit rating, but mouse movements alone appear only useful for determining which pages receive have the least amount of interest but are not accurate for distinguishing amongst higher levels of interest.

6.3 Number of Mouse Clicks versus Explicit Rating

Mouse clicking may be a useful interest indicator, too, as users click on links they find interesting (suggesting the current page is a good gateway to interesting sites) and may click on items on the page that look appealing.

Figure 8 depicts a box-plot of the number of mouse clicks versus the explicit rating. The Kruskal-Wallis test failed to reject the null hypothesis, meaning that the median values for each explicit rating group may be the same. Our conclusion is that for this experiment the number of mouse clicks is not a good indicator of interest.

6.4 Scrolling versus Explicit Rating

We hypothesized that users scroll down a page that they find interesting, most likely as they read the material or occasionally as they search the page for interesting links to follow. Users may scroll in a variety of ways: clicking on the scroll bar, clicking and dragging the scrollbar, hitting page up/down keys or hitting up/down arrow keys. Early analysis of each scrolling method by itself revealed them to be poor indicators of interest. We then attempted to combine some of scrolling methods by adding the time spent in each in an attempt to capture a the "total" scrolling amount.

Figure 9 depicts a box-plot of the time spent scrolling by the mouse and the keyboard versus the explicit rating. The Kruskal-Wallis test rejected the null hypothesis, meaning that the median values for each explicit rating group are different. We conclude that the total time spent scrolling by the mouse and the keyboard is a good indicator of interest.

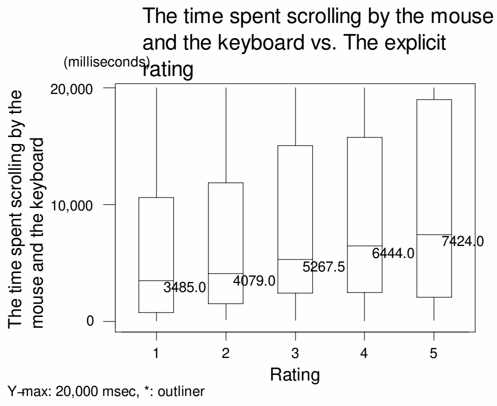

Figure 9: Combined Scrolling versus Explicit Rating.

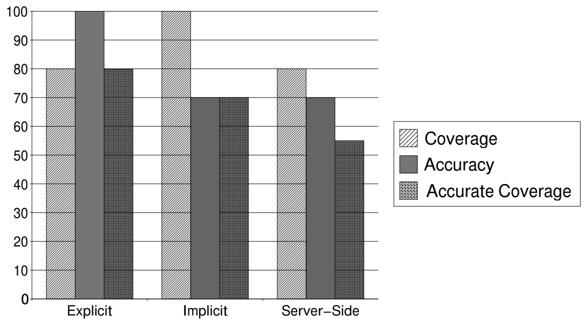

Figure 10: Coverage and Accuracy of Interest Indicator Methods. *Coverage* refers to the percentage of indicators that can be obtained. *Accuracy* is how likely they are to reflect true interest. *Accurate coverage* is a combination of the accuracy and coverage.

6.5 Summary

In this work, we developed a user interface in the form of a customized Web browser in order to capture implicit interest indicators. However, implicit detection of interest can be deployed at the server or even at a proxy as well as at the interface of a client. There are numerous advantages to having server-side detection of implicit interest, notably the ability of users to run any non-customized Web browser they wish. Server-side detection also allows flexibility in the back-end processing that may accompany interest detection, including storage in a database or updating a user profile.

If we assume that a Web server uses an established method for detecting Web sessions from the server logs [6], then, within a session, the time a spent on a page can be obtained by subtracting the access time for the previous page. However, this method is only effective for the current Web server. Thus, if a user jumps to another server, the time spent on the last page of the current server cannot be used as an implicit interest indicator.

Using this method of server-side implicit interest indicators, based on our data server-side implicit interest detection could only be used in about 70% of the Web pages visited, compared with client-side implicit interest detection that could be used in 100% of the Web pages visited. However, server-side detection is comparable to explicit interest indication in which users provided ratings for only 80% of the Web pages visited.

We can extend this analysis to the accuracy of the interest indicators. We assume that the explicit interest indicators are 100% accurate. We can measure the accuracy of the implicit indicators we studied using the graphs shown in this paper and measuring how many "false" predictions would be made for each type of indicator. We assume a "false" prediction is one that is off by more than 2 in terms of explicit interest, as this difference is enough to allow an implicit prediction of "like" (1 or 2) when the explicit interest could actually be a "dislike" (4 or 5) and vice versa. In doing this accuracy analysis, we find time and scrolling to be equally effective, providing about a 70% accuracy each.

Combining these results with the coverage results presented above, we find that explicit interest indicators provide about 80% accurate coverage and client-side implicit interest indicators provide about 70% accurate coverage. While the difference of 10% between them is nontrivial, it is probably an acceptable difference for practical purposes, suggesting that implicit interest indicators can provide the same effectiveness as explicit interest indicators without the user cost. Server-side only implicit interest indicators provide only about about 50% accurate coverage, significantly less than either implicit interest indicators or explicit interest indicators.

The relationship between coverage, accuracy and accurate coverage for the different types of interest detection are depicted in Figure 10. We note that combinations of interest detectors, such as time spent on a Web page and the amount of scrolling, may prove more accurate than any indicator alone. Doing this analysis is an area of future work (see Section 8).

7. CONCLUSIONS

One way a user-interface can be "intelligent" is to understand the interest of the user in the current document. Explicit methods, such as asking users to rate the documents they read, intrude upon the normal browsing process and often are ignored by users. Implicit methods, while requiring more sophisticated intelligent user interfaces, promise to provide more interest indicators without the "cost" to the users.

In this research we have categorized and experimentally evaluated the effectiveness of several implicit interest indicators in determining the explicit interest in a Web page. Based on over 40 hours of Web browsing by over 70 students, we find that time is good implicit indicator of interest mouse movement and mouse clicks by themselves are ineffective implicit interest indicators. However, in using mouse clicks and keyboard actions to infer the level of scrolling, we obtain an means of determining the "amount" of scrolling that also provides an effective indicator of interest.

The techniques used in this research provide a means of gathering implicit interest indicators at the client through a customized browser. However, implicit interest indicators can be gathered at a Web server, too, primarily through server logs. Although server-side indicators do not require custom client software, they provide less accurate results than do client-side implicit interest indicators.

The results presented promise to strengthen the predictions by today's recommender systems and provide insight into other intelligent user interfaces that must infer user interest in order to be effective.

8. FUTURE WORK

In this work, we have considered only the implicit interest indicators alone, such as time versus interest or scrolling versus interest. Combinations of interest detectors, such as time spent on a Web page *and* the amount of scrolling, may prove to be more accurate than any indicator alone. Implicit interest indication may be combined with more explicit indicators, such as ratings or even purchase history, to provide even more effective interest indication.

Future work also suggests searching for a prediction function that accurately predicts explicit interest for a large percentage of users on a large percentage of pages tested. Similarly, there may be a personalized prediction function that can be tailored to an individual user, resulting in a more accurate means of predicting explicit interest.

While our intent here was to establish the relationship between implicit interest indicators and any kind of Web browsing, it may be possible to come up with more accuracy if the test domain is limited to specific types of pages or a specific task. For instance, the correlation between time spent reading a page and a user's interest may be stronger if it is known that the user will not be doing tasks other than browsing. In addition to browsing the Web at large that we present here, we have considered casually reading an online newspaper, looking up a topic in an online encyclopedia, and searching for information using a search engine.

There are many more implicit interest indicators present in other literature [13, 14], such as bookmarking or printing, that need to be empirically evaluated as we have begun to do for time and mouse activity.

9. ACKNOWLEDGMENTS

We would like to acknowledge the help of Professor Isabel Cruz for encouraging her students to participate in our user study, as well as the students who participated.

The Curious Browser and the data gathered from our experiments can be downloaded from `http://perform.wpi.edu/`.

10. REFERENCES

[1] C. Avery and R. Zeckhauser. Recommender Systems for Evaluating Computer Messages. *Communications of the ACM*, 40:88 – 89, Mar. 1997.

[2] M. Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes, and M. Sartin. Combining Content-Based and Collaborative Filters in an Online Newspaper. In *Proceedings of ACM SIGIR Workshop on Recommender Systems*, August 19 1999.

[3] J. Goecks and J. W. Shavlik. Learning Users' Interests by Unobtrusively Observing Their Normal Behavior. In *Procedings of the ACM Intelligent User Interfaces Conference (IUI)*, Jan. 2000.

[4] J. Grundin. Groupware and Social Dynamics: Eight Challenges for Developers. *Communications of the ACM*, 35:92 – 105, 1994.

[5] D. Gupta, M. Digiovanni, H. Narita, and K. Goldberg. Jester 2.0: A New Linear-Time Collaborative Filtering Algorithm Applied to Jokes. In *Proceedings of Workshop on Recommender Systems: Algorithms and Evaluation*, Aug. 1999.

[6] D. He and A. Goker. Detecting Session Boundaries from Web User Logs. In *Proceedingsof of the BCS-IRSG 22nd Annual Colloquium on Information Retrieval Research*, 2000.

[7] W. Hill, L. Stead, M. Rosenstein, and G. Furnas. Recommending and Evaluating Choices in a Virtual Community of Use. In *Proceedings of ACM CHI'95*, pages 194 – 201, 1995.

[8] W. C. Hill, J. D. Hollan, D. Wroblewski, and T. McCandless. Edit Wear and Read Wear. In *Procedings of ACM CHI Conference on Human Factors in Computing Systems*, pages 3 – 9, 1992.

[9] J. Konstan, B. Miller, D. Maltz, J. Herlocker, L. Gordon, and J. Riedl. GroupLens: Applying Collaborative Filtering to Usenet News. *Communications of the ACM*, 40(3):77 – 87, 1997.

[10] P. Le and M. Waseda. Curious Browsers. Technical Report Major Qualifying Project MQP-DCB-9906, Worcester Polytechnic Institute, May 2000. Advisors Dave Brown and Mark Claypool.

[11] H. Lieberman. Autonomous Interface Agents. In *Proceedings of ACM Conference on Human Factors in Computing Systems (CHI)*, 1997.

[12] M. Morita and Y. Shinoda. Information Filtering Based on User Behavior Analysis and Best Match Text Retrieval. In *Proceedings of SIGIR Conference on Research and Development*, pages 272 – 281, 1994.

[13] D. M. Nichols. Implicit Rating and Filtering. In *Proceedings of the Fifth DELOS Workshop on Filtering and Collaborative Filtering*, Nov. 1997.

[14] D. Oard and J. Kim. Implicit Feedback for Recommender Systems. In *Proceedings of the AAAI Workshop on Recommender Systems*, July 1998.

[15] J. Rucker and M. Polanco. Siteseer: Personalized Navigation for the Web. *Communications of the ACM*, 40(3):73 – 76, Mar. 1997.

[16] B. Sarwar, J. Konstan, A. Borchers, J. Herlocker, B. Miller, and J. Riedl. Using Filtering Agents to Improve Prediction Quality in the GroupLens Research Collaborative Filtering System. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW)*, 1998.

[17] U. Shardanand and P. Maes. Social Information Filtering: Algorithms for Automating 'Word of Mouth'. In *Proceedings of ACM CHI'95*, 1995.

[18] A. Watson and M. A. Sasse. Measuring Perceived Quality of Speech and Video in Multimedia Conferencing Applications. In *Proceedings of ACM Multimedia Conference*, pages 55 – 60, Bristol, UK, Sept. 1998.