

Recognizing User Interest and Document Value from Reading and Organizing Activities in Document Triage

Rajiv Badi, Soonil Bae, J. Michael Moore,
Konstantinos Meintanis, Anna Zacchi,
Haowei Hsieh, Frank Shipman
Department of Computer Science
Texas A&M University
College Station, TX 77843-3112 USA
1 979 845 5534
shipman@cSDL.tamu.edu

Cathy Marshall
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052 USA
cathymar@microsoft.com

ABSTRACT

People frequently must sort through and identify relevant materials from a large set of documents, such as looking through the results of a web search. During this process of document triage there is reading and organizing activity. Moreover, these tasks can occur in different applications. A user's interests can be identified from reading and organizing activity and used as a basis for providing cues to other potential documents of interest in the set. To most effectively identify related documents of interest, activity data must be collected from all applications used in document triage. In this paper we present a common framework (the Interest Profile Manager) for collecting and analyzing user interest. We also present models for identifying user interest based only on reading activity, only on organizing activity, and models incorporating both reading and organizing activity. A study comparing document values calculated using the different models shows that incorporating interest information from both reading and organizing activity more accurately estimates users' valuation of documents than using either type of activity alone.

Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval – *search process, selection process.*

General Terms

Design, Experimentation, Human Factors.

Keywords

Document Triage, Information Triage, Sensemaking, User Interest Recognition, User Interest Modeling, Visual Knowledge Builder.

1. INTRODUCTION

With the growth of the internet and other sources of information, locating data about a topic of interest is less of a problem than determining which particular documents are valuable. Frequently, the desired information is obscured within a long list of resources. Users are inundated with so much information that they spend the majority of their time sifting through documents rather than on the information task itself. New systems are needed that help users cull through related documents to gain the information they need.

Document triage is the practice of quickly determining the usefulness and relevance of documents in a vast collection of documents (e.g. from a search engine). In triage, the attention of the user shifts from document to document to contextual overview (e.g. the list of search results, bookmarks, desktop or (another) visual overview of documents). Hence, document triage involves extensive reading (engagement with multiple documents at once) and hyper-extensive reading (engagement with subdocument components and fragmentary information), as opposed to intensive reading (engagement with a single document).

As users sort through a set of documents, they begin to organize relevant documents. They bookmark websites, place papers into piles according to topic, or even create categorized lists of documents. Alternatively, documents may be deleted or ignored. With extremely large sets of documents, users will frequently stop when what they have is good enough. Consequently, potentially better documents are left having never been reviewed.

As users engage in document triage, their activities indicate user interest. Tools that can interpret these activities and infer interest can provide cues indicating documents that are more likely to be of interest as the task proceeds.

1.1 Supporting Document Triage

We are exploring ways to support users in their triage tasks based on inferred user interest. We identified four steps for providing this support: recognizing user interest and document value, representing user interest, recognizing documents of interest, and visualizing interest information.

In describing these four steps, we will refer to three different types of applications involved in document triage: *document*

overview software for listing results from explicit and implicit searches, document reading software for presenting the content of a document, and document organization software for collecting the results of the triage and sensemaking activity.

1.1.1 Recognizing User Interest & Document Value

Systems can gain an understanding of user interest either explicitly or implicitly. Explicit interest indicators require the user to rate documents after reading/skimming them. It has been noted that with explicit ratings, users read a lot more articles than they rate [10]. Recognizing user interest through implicit means requires software specific techniques – implicit expression of document interest in overview software is different than implicit expression of document interest in reading software.

1.1.2 Representing User Interest

Multiple software applications are involved during a triage task. For example someone may be preparing a presentation on the effects of rising oil prices. First a web browser is used to do a search. A variety of reading applications are used including a web browser, Acrobat Reader, media players, etc. Finally, links to relative resources are copied and pasted into categorized lists in a word processor. Each application can have its own methods for recognizing user interest, but the representation of inferred interest must be shared across all applications for this to be useful to the entire triage process. User interest can be represented as a set of documents or document components, abstractions of these elements (e.g. term vectors and metadata), or actions on a document (e.g. scrolling patterns and time spent on a document).

1.1.3 Recognizing Documents of Interest

As applications receive interest information from their own use and the use of other applications, they can begin to identify documents related to this interest (among search results in overview software) or sub-components of a document that match the interest (in reading software). The method for determining when a document or a sub-component of a document is related to an area of interest depends on the representation of interest.

1.1.4 Visualizing Interest Information

Once information relevant to user interests is identified, users must be notified. There are a variety of approaches for notifying users, including suggestion mechanisms and visualization techniques. For example, in overview software, relevant documents can be identified by visually distinguishing proposed documents of interest. In a text reading application, sections of relevant text within the document can be highlighted with a yellow background.

This paper addresses the first two steps, recognition and representation of user interest. Specifically, the focus is on recognizing user interest based on a user’s behavior across multiple applications involved in triage activity. All applications collaborate to recognize user interest during triage. Next we discuss approaches for recognizing user interest. That is followed by an introduction to our system for facilitating the sharing of interest information across applications. We then discuss four models for identifying user interest and compare these models against user-specified document values.

2. RECOGNIZING USER INTEREST

Much research has gone into gathering and recognizing user interest. Systems built to recognize user interest employ implicit indicators, explicit indicators or a combination of both to recognize user interest.

2.1 Explicit Interest Indicators

With explicit indicators (e.g. ratings), users tell the system how interesting or uninteresting a given document is. Explicit indicators are well-understood, easy to implement and fairly precise. However, stopping to enter explicit ratings can alter normal patterns of browsing and reading [3], and can lead to increased cognitive load on the user. Users may stop providing ratings when they perceive that there is no benefit in providing explicit ratings [5]. Research on the GroupLens system [10] found with explicit ratings that users rated many fewer documents than they read. Thus, even though explicit ratings are fairly precise in recognizing interest, their efficacy is limited.

2.2 Implicit Interest Indicators

Implicit indicators are a viable alternative. Nichols [9] identifies some implicit interest indicators and discusses the potential of implicit ratings. He states that “the limited evidence available suggests that implicit ratings have great potential, but their effectiveness remains unproven”.

Overall, time spent reading a document is a good indicator of interest [2, 3, 7, 8]. Conversely, Kelly and Belkin [6] find no general, direct relationship between display time and usefulness. They do note, however, that display times varied according the specific task and user. So this may not be problematic in a more focused task such as document triage.

User events such as scrolling and mouse events are informative as well [3, 4]. These studies do conflict on the efficacy of mouse movement to user interest.

User annotations are a final class of implicit indicators. The XLibris system aids users as they re-read documents by providing visualizations based on user annotations [13]. The system identifies area highlights, area circling, underlining, the use of margin bars, highlighting, comments, etc. A higher emphasis value is assigned to interpretive marks (comments, symbols and callouts). Even though recall and precision rates are low, the paper differentiates interpretive tasks from normal reading and navigational tasks.

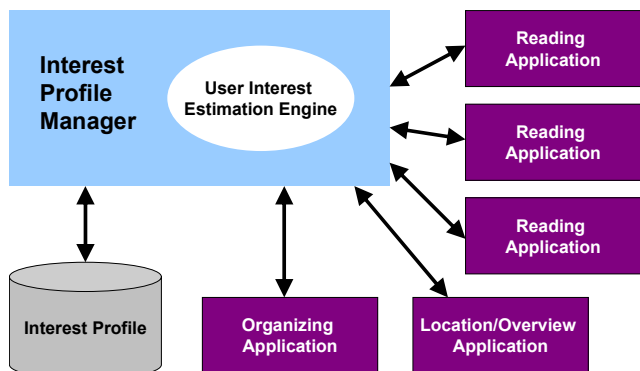


Figure 1: Interest Profile Manager

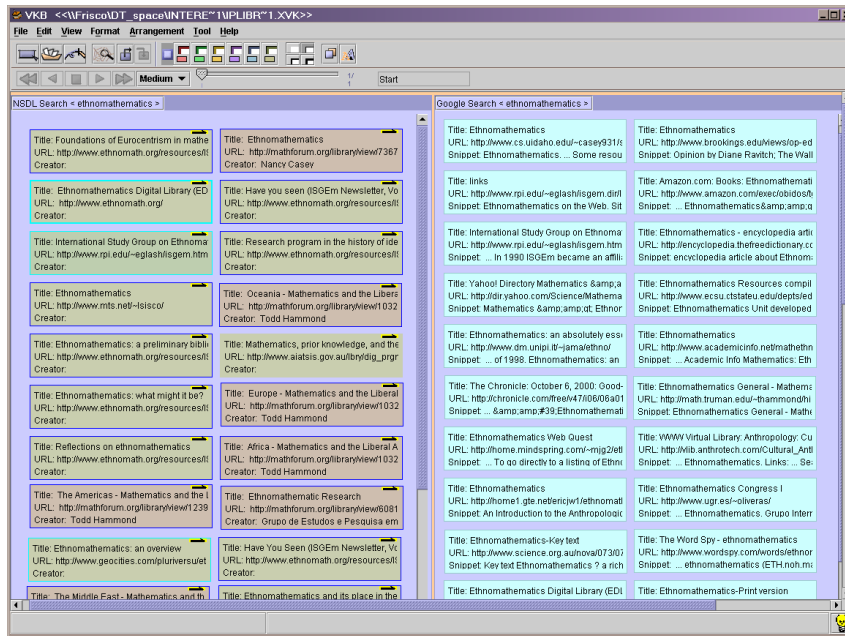


Figure 2: VKB workspace with search results

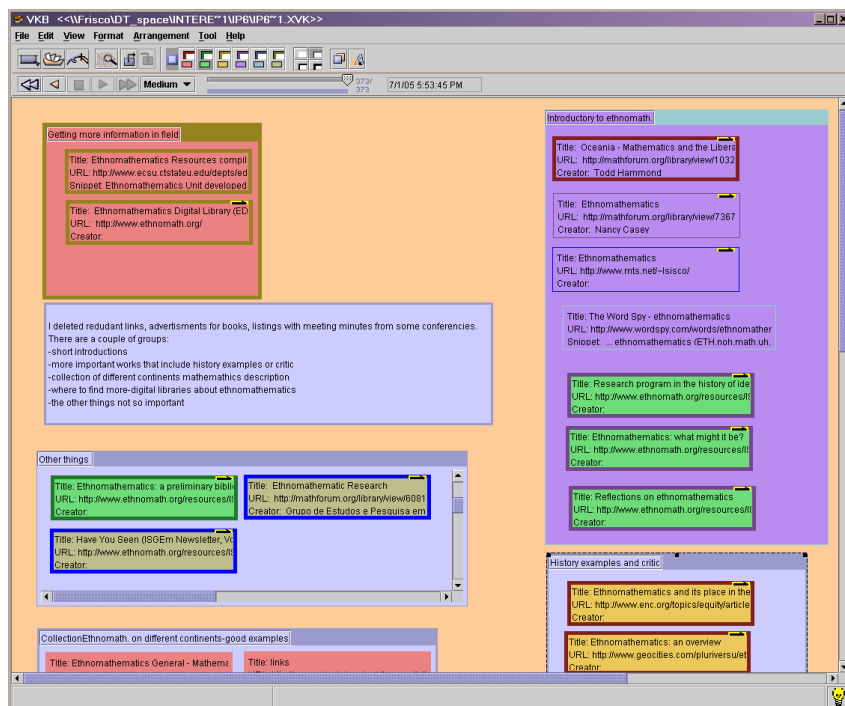


Figure 3: VKB workspace after organizing activity

3. INTEREST PROFILE MANAGER

Prior work has focused on a single reading application as the source for interest indicators. However, as indicated earlier, triage occurs in the context of multiple applications. We have created the Interest Profile Manager as the basis for determining, sharing and storing user interest based on interest indicators from multiple applications (Figure 1). The Interest Profile Manager acts as an independent server. A library is provided that applications can use to connect and then store and retrieve interest information. The

library can be used to modify existing programs to interface with the Interest Profile Manager or to create connector applications that interface the Interest Profile Manager to an existing application with its own interface library such as Internet Explorer. In this way reading, organizing, and location/overview applications can connect to the Interest Profile Manager.

As users work, client applications send interest-related activity information to the Interest Profile Manager as a set of attributes and values and this information is stored in the Interest Profile. Changes to the Interest Profile are propagated to all clients. The Interest Profile Manager includes a User Interest Estimation Engine that evaluates user interest based on information in the Interest Profile and its interest models. The interest profile is augmented with the results calculated from the interest models, and these calculations are communicated to all clients.

For our initial implementation, we modified the Visual Knowledge Builder (VKB) [11] to act as both the location/overview application and the organizing application. In VKB, search results are returned as information objects inside of a two-dimensional space called a *collection* (Figure 2). Each information object refers to a search result and double clicking on an object opens the corresponding URL. VKB also allows users to add text and associate metadata to both collections and objects. Moreover, visual attributes including border width, object size, and background and border color can be edited for objects and collections. The addition and manipulation of text, metadata, and visual attributes can act as implicit interest indicators. Figure 3 shows the workspace after organizing activity.

The reading application in our implementation is an instrumented version of Microsoft's Internet Explorer (IE). IE was augmented to gather and log user activity and communicate with the Interest Profile Manager. The next section describes how user interest models were developed based on use of these applications.

4. INTEREST MODELS

Due to the shortcomings of explicit interest ratings, we currently rely on only implicit indicators. This information is gathered unobtrusively from a reading interface (currently IE) and an organizing interface (currently VKB). Data gathered includes characteristics of the document (i.e. document attributes), data coming from the document reading activity, and data coming from the document organizing activity (Table 1).

Table 1: Data collected

Category	Parameter
Document Attributes	Number of characters (words/page)
	Number of links
	Number of images
	File size of a document (Bytes)
Document Reading Activity (IE)	Time spent in a document
	Number of mouse clicks
	Number of text selections
	Number of scrolls
	Number of scrolling direction changes
	Time spent scrolling
	Scroll offset
	Total number of scroll groups
	Number of document accesses
	Document Organizing Activity (VKB)
Number of object moves	
Number of object resizes	
Number of object deletions	
Number of content changes	
Number of background color changes	
Number of border color changes	
Number of border width changes	
Number of font changes	
Number of font color changes	
Number of canvas color changes	
Number of changes in z-order	
Number of transparency changes	

Document attributes are inherent to the documents themselves, rather than on users' interactions with the documents. We primarily consider document length (e.g. number of pages, number of characters and number of words), number of links or images in a document, and file size.

Document reading activity includes user actions during passive reading in the reading application. This consists of time spent on a document, number of mouse clicks, number of text selections, characteristics of the users scrolling behavior, and frequency of document access.

Document organizing activity refers to user actions on objects representing documents in the VKB workspace (i.e. the organizing application). Creating categories, assigning document objects to categories or changing spatial or visual attributes of categories or document objects are included in this activity. Through these actions, users express their interpretation of the documents and their interrelations.

A prior study [1] showed a correlation between user interest with user events such as reading time and a few document attributes. A more comprehensive correlation analysis revealed additional user events for estimating document value from VKB (organizing application), IE (reading application) and document attributes. Table 2 shows events correlated to explicit ratings given to documents by users. With one exception, all p-values are less than 0.005. "Number of characters" has a p-value of 0.01 but is included due to interviews with subjects in a Fall 2004 document triage study strongly indicating that users consider the amount of content in a document to be an important criterion for classifying documents. The only negative correlation, "Number of object deletions" makes sense since removing a document should correspond to little or no interest in the document. As has been shown in prior studies, time spent in a document positively

Table 2: Results from correlation analysis

Parameter	Pearson coefficient	p-value
Number of characters	0.431	0.010
Time spent in a document	0.527	0.001
Number of scrolls	0.630	<0.001
Scroll offset	0.641	<0.001
Number of scrolling direction changes	0.589	<0.001
Total number of scroll groups	0.590	<0.001
Number of document accesses	0.476	0.004
Number of object moves	0.711	<0.001
Number of object resizes	0.622	<0.001
Number of object deletions	-0.495	0.003
Number of background color changes	0.597	<0.001
Number of border color changes	0.628	<0.001
Number of border width changes	0.525	0.001

correlates to user interest. Note that four different types of scrolling events correlate to user interest.

The correlation analysis provides a set of parameters for estimating users' interest from user activities in IE (reading application), VKB (organizing application) and document attributes. These are the basis for building mathematical and qualitative models to estimate user interest. Since there are significant correlations from both the reading and organizing applications, models for user interest can take advantage of events from both classes of applications.

4.1 Mathematical Models

Three mathematically-derived models are presented that calculate user interest. Data gathered from the Fall 2004 [1] study was used to derive these models. This data includes the parameters identified in Table 2 and explicit user assessment of documents. In this case, the explicit user assessment involved each user selecting the five most interesting documents and the five least interesting documents out of the corpus of 34 documents after performing a document triage activity. To generate a quantitative value of interest, documents rated as being of least interest were assigned an interest value of 0, documents rated as being of high interest were assigned an interest value of 2, and the remainder were given an interest value of 1.

Rather than classifying documents as being of low, average, or high interest, all three mathematical models were developed to produce floating point values between 0 (low interest) and 2 (high interest). A value below 1 indicates the document is of less than average interest while a value greater than 1 indicates a greater than average interest.

The first model is limited to reading activity in IE and document

Table 3: Factor score coefficients for the reading activity model

Parameter	Factor1	Factor2
Number of characters	-0.199	0.537
Number of scrolls	0.123	0.187
Number of document accesses	0.294	-0.229
Time spent in document	0.265	-0.069
Scroll offset	-0.072	0.442
Number of scrolling direction changes	0.227	0.010
Total number of scroll groups	0.258	-0.050

Standardized parameter =
(parameter – mean) / standard deviation.

Factor1 =

- 0.199 * (standardized number of characters)
- + 0.123 * (standardized number of scrolls)
- + 0.294 * (standardized number of visits)
- + 0.265 * (standardize time spent in document)
- 0.072 * (standardized scroll offset)
- + 0.227 * (standardized number of scrolling direction changes)
- + 0.258 * (standardized total number of scroll groups)

Factor2 =

- 0.537 * (standardized number of characters)
- + 0.185 * (standardized number of scrolls)
- 0.229 * (standardized number of visits)
- 0.069 * (standardized time spent in IE)
- + 0.442 * (standardized scroll offset)
- + 0.010 * (standardized number of scrolling direction changes)
- 0.050 * (standardized total number of scroll groups)

Figure 4: Reading-Activity equations to calculate factors

attributes. The second model is limited to organizing activity in VKB. The final model combines reading activity in IE, organizing activity in VKB, and document attributes to estimate user interest. All three models were developed using aggregated user activity and averaged user evaluation of the documents' value.

4.1.1 Reading-Activity Model

The reading-activity model uses document attributes, such as number of characters, images and links; and user events generated during reading, such as number of scroll events and time spent in reading. Correlation analysis between these document attributes/user events and the explicit user ratings on documents (document score) identified seven variables. Many of these variables are correlated to each other, and regression modeling requires independent variables. Factor analysis produced two independent factors based on the seven variables. These factors can be calculated from the factor coefficients generated for each parameter as shown in Table 3. Figure 4 demonstrates how factors are calculated. The multiple regression model based on those factors is:

$$\text{Document Score} = 0.877 + 0.133 * \text{Factor1} + 0.120 * \text{Factor2}$$

4.1.1.1 Organizing-Activity Model

The organizing-activity model takes user events generated during organizing activities such as moving/resizing/deleting objects and changing the background/border color of objects. The correlation analysis between these user events and explicit user ratings for documents identified six variables. We have extracted two independent factors through factor analysis. The coefficients for calculating the factors in this model are found in Table 4. Factors are calculated in a similar fashion as Figure 4. The multiple regression model based on those factors is:

$$\text{Document Score} = 0.877 + 0.185 * \text{Factor1} - 0.092 * \text{Factor2}$$

4.1.1.2 Combined Model

The combined-activity model is based on document attributes and user events generated during reading and organizing (i.e. the combination of those used in the reading-activity and organizing activity models). In this case, factor analysis identified four independent factors based on the 13 variables. Table 5 shows the factor coefficients for each parameter used in calculating the four factors. The multiple regression model based on those factors is:

$$\text{Document Score} = 0.877 + 0.125 * \text{Factor1} + 0.152 * \text{Factor2} + 0.0662 * \text{Factor3} + 0.0653 * \text{Factor4}$$

4.2 Qualitative Model

The previous three models are based on the statistical analysis of prior study data. They were derived so as to minimize the error with respect to the explicit user evaluation of documents in the earlier study. Due to the limitation of subjects to identify five documents as high value and five as low value, with the remaining 25 considered to be of average value, this mathematical approach awards being conservative in modeling above- or below-average interest. As such, the above models tend to be conservative in their estimates of documents varying from average interest.

A fourth model was defined based on a combination of qualitative and quantitative assessment of previous study data. Qualitative data included interviews with subjects about what made documents valuable and analyzing videotapes of user activity during the performance of the task. Table 6 summarizes this model. The weight field in the table is the degree to which the particular parameter contributes to the overall document score. Currently, document attributes constitute 10% of the qualitative interest model, reading activity is used for 37.5% of the interest value, and organizing activity is the remaining 52.5%.

Table 4: Factor score coefficients for the organizing activity model

Parameter	Factor1	Factor2
Number of object moves	0.331	0.202
Number of object resizes	0.325	0.327
Number of object deletions	0.164	0.820
Number of background color changes	0.210	-0.232
Number of border color changes	0.234	-0.174
Number of border width changes	0.164	-0.105

Table 5: Factor score coefficients for the combined activity model

Parameter	Factor1	Factor2	Factor3	Factor4
Number of characters	-0.143	-0.038	0.549	0.174
Number of scrolls	0.164	-0.142	0.204	-0.091
Number of visits	0.197	0.117	-0.246	0.250
Time spent in a document	0.268	-0.112	-0.057	-0.119
Scroll offset	0.005	-0.140	0.472	0.044
Number of scrolling direction changes	0.244	-0.134	0.035	-0.079
Number of scroll groups	0.266	0.113	0.047	0.137
Number of object moves	0.026	0.291	-0.123	0.149
Number of object resizes	0.066	0.195	-0.070	0.317
Number of object deletions	-0.036	-0.055	0.170	0.826
Number of background color changes	-0.109	0.319	0.040	-0.073
Number of border color changes	0.101	0.339	0.009	0.023
Number of border width changes	-0.091	0.338	-0.139	-0.132

Table 6: Weights for qualitative model

Category	Parameter	Weight
Document Attributes (10%)	Number of characters	5
	Number of links	5
Document Reading Activity (37.5%)	Number of scrolls	5
	Number of document accesses	10
	Time spent in a document	10
	Scroll offset	5
	Number of scrolling direction changes	5
	Number of scroll groups	2.5
Document Organizing Activity (52.5%)	Number of object moves	17.5
	Number of object resizes	15
	Number of object deletions	5
	Number of background color changes	7.5
	Number of border color changes	2.5
	Number of border width changes	5

This model was developed to generate a wider range of interest values and to be the starting point for a model that is not tied to one particular document set and user activity – which is a danger when developing models from a user study involving a single task and corpus, no matter how many subjects are included. It will need to be refined based on data from follow on studies that includes other tasks and document sets.

4.3 Comparison of Models

The capability of the three statistical models to predict user interest on documents has been compared in terms of R^2 and the adjusted R^2 of the models in Table 7. A value of R^2 is a measure of how much of the variability in the outcome of the models is accounted for by the predictors. For the first model, the value is 0.477, which indicates that parameters of reading activity account for 47.7% of the expressed user interest for documents. Similarly, parameters of organizing activity account for 63.6% of user interest. Finally, parameters of both reading and organizing activity account for 70.8% of user interest. The adjusted R^2 shows how the model generalizes. Ideally adjusted R^2 should be the same, or very close to, the value of R^2 .

Table 7: Comparison of mathematical models

Model	R	R^2	Adjusted R^2
Reading-activity model	0.690	0.477	0.444
Organizing-activity model	0.797	0.636	0.613
Combined model	0.841	0.708	0.669

5. MODEL EVALUATION

The prior comparison of models shows the error is reduced when using a combined model. This is guaranteed due to the increased number of variables available to model the same data. It remains to be seen whether the combination will improve the model's predictions for other users. Thus, a study was conducted to evaluate the effectiveness of the interest models presented in the previous section. The following subsections discuss the study and the results.

5.1 Study

The study took place in the Center for the Study of Digital Libraries at Texas A&M University. Sixteen graduate students

and research associates from the university participated. Subject ages ranged from 22 to 32. All subjects had basic familiarity with using a computer and browsing the web and had used a computer regularly for five or more years.

This is the same task and topic as in some of our prior studies [1, 12]. Subjects were placed in the role of a research librarian that had to select and organize documents for a high school teacher preparing a class on ethnomathematics, which is the study of a group's culturally-specific mathematical practices as they go about their everyday activities.

Since VKB is the organizing application, all subjects were given a brief training on VKB emphasizing features considered relevant for the task. VKB allows users to organize information objects (links to websites in this study) in a hierarchy of two-dimensional workspaces. Subjects were also limited to using the augmented version of Internet Explorer as the reading application.

Subjects started with 20 documents returned from the National Science Digital Library (NSDL) and 20 documents returned from Google placed in lists in a VKB workspace (Figure 2). All subjects received the same 40 links. The documents varied in their level of difficulty, relatedness to the topic and volume of information. Although no time limit was set, all subjects took less than one and a half hours for the task. Figure 2 shows the VKB document (with 40 links) as provided to subjects.

Subjects had to organize the 40 links into visual structures for the high school teacher. They had to determine their own criteria for organizing the links, and were free to add, modify and delete the links. They were also free to modify other attributes in the VKB space such as background color, border thickness, font color, etc., as well as adding text or annotations to the VKB space as they deemed necessary.

Following the task, subjects provided explicit document usefulness ratings for all documents. Documents were rated on a scale of 1 (Not Useful) to 5 (Very Useful). Five of the 40 links were not included in the following analysis as their activity data was not available (e.g. the augmented IE did not record activity within an embedded viewer such as Adobe's PDF plugin). Subjects were also interviewed to gain a deeper understanding of their task practices, document ratings, usage of metadata in evaluating document relevance, etc. The main focus of the interview was gaining an understanding of the subject's triage practice and the criteria the subject used to rate documents.

User actions in VKB and IE were logged to collect events and times for model calculation. The Interest Profile Manager was used to store user behavior details from web pages.

5.2 Results

The 16 subject's evaluations for each document were averaged and scaled to a continuous value between 0 (indicating least useful) and 2 (indicating most useful). To investigate the accuracy of each model's prediction of user interest, we calculated the residue for each document. The residue is the absolute value of the difference between the explicit user rating and a model's predicted rating. A perfect predictive model would have an average residue of zero. The results in **Error! Reference source not found.** are averaged across all documents for each model. This shows the overall prediction accuracy for each model.

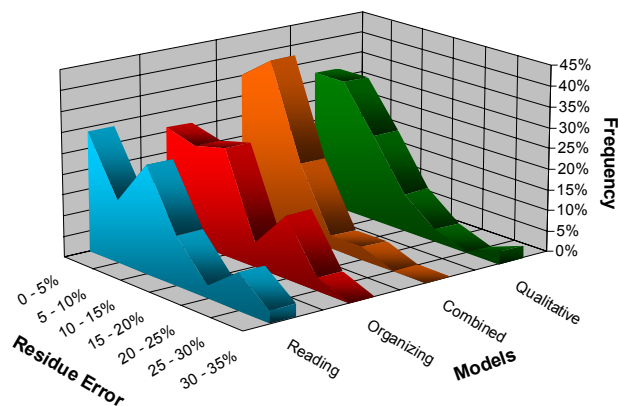


Figure 5: Distribution of residue errors

Both models relying on a combination of reading and organizing activity have substantially lower residue values than either of the models based solely on either reading or organizing activity. ANOVA can identify which models produced residues that are significantly different from each other. Based on the results of ANOVA, only the residues from the reading-activity model were significantly different from those of the combined model ($p=0.025$). The organizing-activity model compared with the combined model resulted in a p -value of 0.281, and the organizing-activity model compared with the reading-activity model resulted in a p -value of 0.238.

Moreover, paired t -test analysis between the different models also showed a significant difference between the residues of the reading-activity model and the combined model ($p=0.02$). While, not significant at the 0.05 level, paired t -test analysis between the residues of the organizing-activity model and the combined model resulted in a p -value of 0.123.

As previously mentioned, residues greater than zero represent errors in a model's prediction of user interest. Since document interest values are scaled from 0 (low interest) to 2 (high interest), each 0.1 increment in residue represents a 5% error of predicted values from user assessed values. Figure 5 shows a distribution of errors for each model. Ideally, 100% of all errors would fall in the 0-5% range. So, keeping a large proportion of errors towards 0% is desirable. All models have a majority of their errors distributed in the 0-15% range. The combined model has 94% of its error contained within the 0-15% error range. However, both the reading and organizing models have more errors spread beyond 15% making these models' predictions more random. Note the peak in the 20-25% range for the organizing model and the peak in the 25-30% range for the reading model.

5.3 Discussion

The results consistently show that the combined activity model performs the best of the three mathematical models while the organizing activity model performs better than the reading activity model. The qualitative model performed reasonably well – better than either the reading or organizing activity models.

Table 8: Residue comparison of models

Model	Average Residue	Standard Deviation
Reading-activity model	0.258	0.192
Organizing-activity model	0.216	0.146
Combined model	0.175	0.138
Qualitative Model	0.197	0.134

The prior analyses are based on the aggregated user activity and averages of document valuations. We are currently investigating each model's predictive capability when applied to individual data. As the mathematical models were developed based on aggregated data, they are not likely to perform well for individuals. On the other hand, the qualitative model has been developed with an understanding of how triage practices vary between users.

To understand these differences in individual document triage practices, we describe some results from our qualitative analyses. In interviews that followed the task, the users were asked why they rated documents lower and higher, about revisiting documents, and about their overall organizational strategy. The following discussion is based on both the quantitative results of the study as well as the qualitative assessment based on the task and the interviews.

Subjects claimed to rate documents mainly based on their content. Some subjects considered web pages with more information as useful, e.g. rating academic journals and some comprehensive sources as more useful, while others rated them low as they felt that the teacher would prefer shorter or introductory material for his/her class.

Metadata was also used to evaluate documents. Most subjects relied on the domain names of the documents for authenticity. Therefore they rated documents from .edu domains and digital libraries higher than other documents. It is interesting to note that a couple of subjects relied on the update frequency of the web pages as a criterion in rating them evidenced by rating recently updated documents higher. They relied on the "Last updated" information in a web page for this criterion. Some subjects rated a few documents without visiting them, basing their decisions on metadata alone. For example, one of the subjects moved the Amazon link to a collection based on the URL and title information provided in the VKB symbol, and subsequently assigned a rating of 4 (out of 5) to the URL after the task.

Subjects revisited documents for different reasons. One subject visited most links cursorily in the first pass "to see what was out there", classified most documents during the second pass, and then classified the remaining documents in the third pass. Another subject revisited some of the documents since he did not recall if he had already visited them.

Another aspect for consideration is the deletion of links from the VKB space. Subjects had not been given any specific instructions regarding discarding links from the space. Some users chose to delete links which they felt were not useful. Some others chose to place such links in separate collections giving those collections labels such as "Other". Some subjects chose not to delete links and left them in their original collections without moving them to any new collections they created. One of the subjects deleted most of the links leaving only 10 links in the VKB space. The subject mentioned that the 10 links would suffice for the teacher,

and that many of the other links could be reached from these 10 links. This subject provided explicit ratings for 34 documents. One of the hypotheses going into the study was that subjects deleted links of documents that were not useful. In this case some useful documents were removed. However, in the context of this task, the remaining documents were relatively more valuable than those removed.

The usage of color in VKB was also idiosyncratic based on each subject's strategy for completing the task. One subject used elaborate color coding for web pages; red for good examples, purple for introductory material, and thick red borders for good examples which were already assigned a different color. The subject mentioned that she would set the color and border thickness of a symbol after reading the corresponding web page and in the second pass drop the matching symbols into collections. Another subject who came up with a chapter based classification, used color to show the relationship between the chapter headings and their content.

6. CONCLUSIONS

Document triage is the process of identifying valuable documents within a set of resources. We are investigating methods for actively supporting document triage.

Triage tasks involve searching or accessing lists of resources, skimming or otherwise reading bits of these resources, and selecting and organizing desired resources for further use. Thus, the triage process involves multiple applications, including searching or browsing tools, reading tools, and document organization and management tools.

User activity in these applications often corresponds to user interest in resources and can be the basis for supporting later triage activity. In the past, each application involved in the triage task would have to try to infer the user's interest. The Interest Profile Manager allows applications to communicate information about user activity during the triage task and to estimate user interest in documents based on that activity.

A reading-only activity model, an organizing-only activity model, and a combined activity model were developed based on a prior study of document triage practice. These three models were then compared to explicit user statements of interest in a subsequent study. The evaluation showed that the combined activity model was better at predicting user interest than either single-application model and that the organizing activity model performed better than the reading activity model.

All of these models were developed from aggregated user activity and average document assessments from the earlier study. They are unlikely to perform as well for individuals due to the idiosyncratic nature of document triage practice. In order to cope with such variation, models estimating user interest need to increase the number of triage activity and document characteristics included in the model and better balance their contributions to the inferred value. The qualitative model is our first step towards such a model.

7. REFERENCES

- [1] Bae, S., Badi, R., Meintanis, K., Moore, J.M., Zacchi, A., Hsieh, H., Marshall, C., Shipman, F. "Effects of Display Configurations on Document Triage," *Proceedings of IFIP Interact Conference*, 2005.
- [2] Chan, P. "A Non-Invasive Learning Approach to Building Web User Profiles," *Workshop on Web Usage Analysis and User Profiling*, 1999, pp. 7-12.
- [3] Claypool, M., Le, P., Waseda, M., Brown, D. "Implicit interest indicators," *Proceedings of ACM Intelligent User Interfaces*, 2001, pp. 33-40.
- [4] Goecks, J., Shavlik, J. "Learning Users Interests by Unobtrusively Observing Their Normal Behavior," *Proceedings of ACM Intelligent User Interfaces*, 2000.
- [5] Grudin, J. "Groupware and Social Dynamics: Eight Challenges for Developers," *Communications of the ACM*, 35:92 – 105, 1994.
- [6] Kelly, D., Belkin, N. "Display time as implicit feedback: Understanding task effects," *Proceedings of the 27th Annual ACM International Conference on Research and Development in Information Retrieval (SIGIR 04)*, 2004, pp. 377-384.
- [7] Kim, J., Oard, D.W., Romanik, K. "Using implicit feedback for user modeling in internet and intranet searching," University of Maryland CLIS Technical Report, 2000.
- [8] Morita, M., Shinoda, Y. "Information filtering based on user behaviour analysis and best match text retrieval," *Proceedings of the 17th Annual ACM International Conference on Research and Development in Information Retrieval (SIGIR '94)*, Springer-Verlag, 272-81.
- [9] Nichols, D. "Implicit Rating and Filtering," *Proceedings of the 5th DELOS Workshop on Filtering and Collaborative Filtering*, Budapest, Hungary, 10-12, November 1997.
- [10] Sarwar, B., Konstan, J., Borchers, A., Herlocker, J., Miller, B., Reidl, J. "Using Filtering Agents to Improve Prediction Quality in the GroupLens Research Collaborative Filtering System," *Proceedings of ACM Conference on Computer Supported Collaborative Work (CSCW)*, 1998.
- [11] Shipman, F., Hsieh, H., Maloor, P., Moore, J. M. "The Visual Knowledge Builder: A Second Generation Spatial Hypertext," *Proceedings of ACM Hypertext*, 2001.
- [12] Shipman, F., Hsieh, H., Moore, J.M., Zacchi, J.M. "Supporting Personal Collections across Digital Libraries in Spatial Hypertext," *Proceedings of the ACM and IEEE Joint Conference on Digital Libraries*, 2004.
- [13] Shipman, F., Price, M., Marshall, C., Golovchinsky, G., Schilit, B. "Identifying Useful Passages in Documents Based on Annotation Patterns," *Proceedings of European Conference on Digital Libraries*, Springer Verlag, 2003, pp. 101-112.