

Eye-tracking Analysis of User Behaviors in Document Similarity Judgment

Minghuang Chen^{*1}Seiji Yamada^{*2}Yasufumi Takama^{*1}^{*1}Tokyo Metropolitan University^{*2}National Institute of Informatics

This paper investigates the behaviors of users judging similarity of documents by eye-tracking analysis. This investigation relates with recently-proposed concept of Minimal User Feedback (MUF), which aims at decreasing the cost of a user providing feedback information. In order to achieve this goal, this paper focuses on minimizing the cost of judging similarity of documents, which is supposed to be fundamental task for a user using an interactive document clustering system. In the experiment, 21 test participants were asked to judge the similarity of documents. As the clue for the judgment, 3 types of information; original text, snippet, and term, are investigated. An eye-tracking device is used to record the participants' viewing behaviors. The areas of interesting (AOI) are analyzed, and the result shows that the areas at which the participants more frequently looked are different between the conditions of judging documents of same topics and different topics. It also shows that participants most frequently switched AOIs between documents when terms are presented. The obtained results will contribute to the design of interface that can minimize the user's feedback cost.

1. Introduction

This paper investigates the behaviors of users judging similarity of documents by eye-tracking analysis. The concept of Minimal User Feedback (MUF) [Okabe 07] has been proposed, which aims at decreasing the cost of a user providing feedback information. The MUF employs two approaches; minimizing the quantity of feedback information and minimizing the cost of generating each of feedback information (i.e. relevance judgment for a single document). A related work with the former approach is active learning [Onoda 07], in which a learner actively gathers training data effective for learning.

This paper addresses the latter approach of minimizing the cost of generating each of feedback information. A user usually generates feedback information by judging target objects. For example, a user judges the relevance of a document to a query in the case of document retrieval. In order to provide must-links and cannot-links for constrained clustering [Basu 08], a user has to judge the similarity between target objects. Therefore, decreasing the cost of judgment is important.

This paper investigates the behaviors of users judging similarity of documents. In particular, we focus on the effect of the type of providing information on users' judgment behaviors. The behaviors of users judging relevance of objects have been investigated. However, to our best knowledge, the behaviors of users judging similarity of objects have not been investigated.

In this paper, test participants are asked to judge the similarity of two documents. Given a pair of news articles, a participant judges whether those articles relate with the same topic or not. As the clue for judging similarity, three kinds of information; original text, snippets, and terms, is mutually provided. We suppose that information identifying the difference and commonality of documents is effective. Therefore, common and specific terms / snippets are presented to test participants in a separate manner.

An eye-tracking system was used to record user's eye movement on the computer screen. By analyzing AOI (area of

interest) and focusing time, typical viewing behaviors are investigated. The obtained results will contribute to the design of interface that can minimize the user's feedback cost.

2. Related works

Compared with similarity judgment, much work has been done for studying users behaviors in judging relevance of documents, which include users' viewing behaviors in search result pages and web pages [Lorigo 06] [Rodden 07] [Sun 05], and study on the effect of snippet on relevance judgment [Cutrell 07] [Chen 09]. Among them, Chen et al compared accuracy of relevance judgment and judgment time between the condition of providing snippet and that of providing original text [Chen 09]. Most of the studies have not only investigated judgment accuracy and judgment time, but also analyzed users' behaviors by using eye-tracking systems.

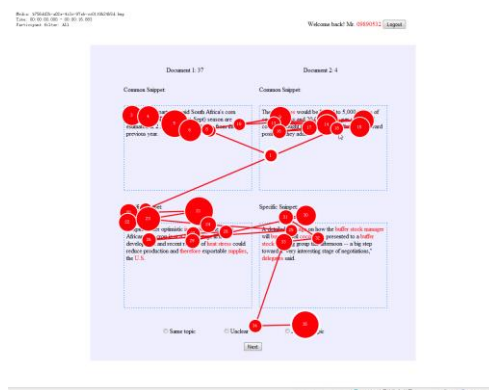


Fig 1. A screenshot of gaze plot by T60 Eye-Tracking System

An eye-tracking system can record user's eye movement on a computer screen. This paper uses T60 from Tobii technology to investigate the user behaviors in document similarity judgment. It can record eyes movement and focusing area (AOI) on the screen. The focusing time is also recorded. By using the Tobii's accessory software Tobii Studio, the recorded data can be visualized. Figure 1 shows a gaze plot, which shows the

trajectory of focusing point. In the gaze plot, the size of node represents the focusing time. Figure 2 is called a heat map that is also visualized by the Tobii Studio. A heat map is a graphical representation of data where the accumulated focusing time is represented as colors. The heat map can be used for displaying areas of a Web page that are frequently scanned by readers [Cutrell 07]. The most frequently focused areas are highlighted by red colors.



Fig 2. A screenshot of heat map by T60 Eye-Tracking System

Such eye-tracking systems have been widely used for studying users' behaviors in viewing Web search results. Cutrell [Cutrell 07] has investigated the effect of task type, snippet length, and the position of the best result in MSN search result page on users' viewing behaviors. Interesting results were obtained, such as that snippet length has different effect between navigational and informational tasks. That is, for informational task, longer snippet improved click accuracy while reducing the task time, but opposite effect was observed for navigational task. This result was explained based on the analysis of eye-tracking record that user performing informational task tended to rely on snippets.

Rodden [Rodden 07] has explored the relationship between mouse movements, and eye movements when performing a search task with using Google. Various interesting patterns are observed, such as keeping the mouse still while reading and using the mouse as a reading aid.

Lorigo et al. [Lorigo 06] has investigated users' search and evaluation behaviors based on the analysis of scan path recorded by using eye-tracking system. The results have shown that users tended to make decision on performing new search before viewing entire page of retrieved result, and that they tended to reexamine top 1 and 2 results frequently.

3. Outline of experiment

3.1 Experimental setting

This paper investigates users' behaviors in similarity judgment. The task of test participants is to judge the similarity of two documents. Given a pair of documents, they are asked to judge whether those documents relate with the same topic or not.

For the experiment, we implement the experiment system that is written in VB language as ASP pages. Figure 3 shows the screenshot of the experiment system, which can be accessed with ordinary web browsers. These pages are displayed using T60

eye-tracking system, and eye movements of the participants during the experiment are recorded. In each session, a document pair is displayed by one of the three types of information: topic terms, snippets and original text. In these figures, information about documents is arranged in two columns; we call left-hand document area and right-hand document area hereinafter.

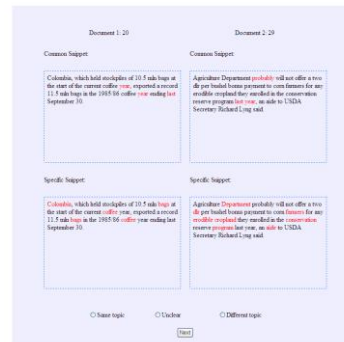


Fig 3. A screenshot of the experiment system

Two kinds of topic terms (snippet): common and specific terms (snippets) are presented. The common terms (snippet) are displayed in the upper part of the screen, and specific one is displayed in the lower part. Topic terms are highlighted with red when snippet is displayed. The documents and topics are selected from Reuter Test Collection¹. It includes 21578 documents with 135 topics. The document pair is extracted randomly from 42 documents of 3 topics (each topic contains 14 documents).

3.2 Extraction of topic terms

Terms that represent the topic of the document are supposed to work as a clue for judging similarity of documents. In particular, the terms indicating the difference and commonality between documents should be presented to a user. Based on this consideration, we classify the topic terms into common and specific terms, which are extracted with the following two steps.

In the first step, given a set of documents D (42 documents used in the experiment), terms that have high TF-IDF values are extracted as topic terms.

In the second step, for a pair of documents that are to be compared, the topic terms that occur in both of the documents are selected as common terms, whereas the terms exclusively occur in either of the documents are selected as specific terms.

3.3 Snippet generation

Based on the same consideration as noted in Sec. 3.2, two types of snippets, common and specific snippets, are employed in this paper. The snippets are generated by the following steps:

- Step 1: Extraction of topic terms (Sec. 3.2).
- Step 2: Score calculation for each sentence.
- Step 3: Extraction of a set of sentences as a snippet.

In step 2, the score of a sentence is calculated based on the TF-IDF values of specific / common terms that are contained in the sentence.

¹<http://www.daviddlewis.com/resources/testcollections/>

In step 3, a set of sentences with the highest score is selected as a snippet. The snippet that consists of the sentences containing specific (common) terms is called specific (common) snippet.

4. Experimental results

By analyzing the eye-tracking data, we found that the position and topic of documents affected participants' viewing behaviors. We also found the viewing behavior that is specific to the case when terms are presented. These findings are described in the rest of this subsection.

4.1 Effect of document position on participants' viewing behaviors

Figure 4-6 show the distribution of AOI per participant, in which focusing time is accumulated for right-hand and left-hand document areas, respectively. It is noted that among 21 participants, 6 participants were excluded from these analysis because of stability problems with the eye tracking, leaving us with 15 participants.

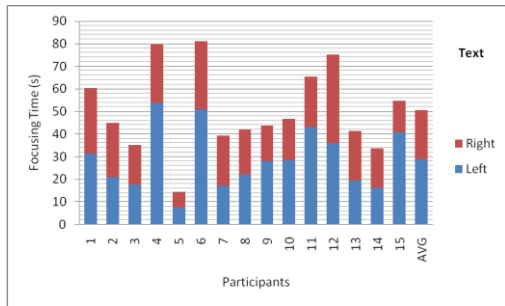


Fig 4. AOI distribution per participant (Text)

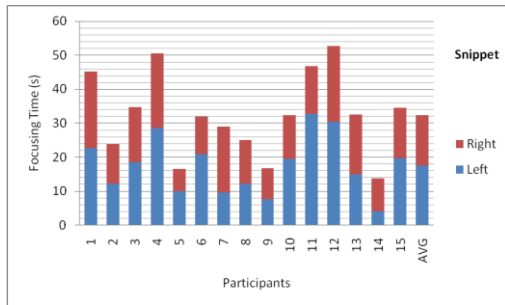


Fig 5. AOI distribution per participant (Snippet)

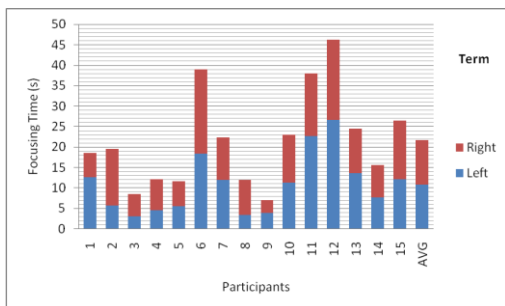


Fig 6. AOI distribution per participant (Term)

These figures show that participants spent more time on looking at left-hand document area when either text or snippet is presented. This tendency is stronger when text is presented than snippet.

We think this tendency can be described by two assumptions about our behaviors when reading documents. First, it is supposed that we usually read a document from left to right. Second, when we compare two things, we use one as a basis and try to find the difference from another one. As a result, we suppose participants read the left document first and used it as a basis.

In order to examine this assumption, we analyzed which document area (right-hand or left-hand) the participants gazed first. Table 1 shows the frequency of gazing at right-hand / left-hand document areas per type of information. From the table, we can see the same tendency as Fig. 4-6, i.e., participants viewing text were the most likely to gaze at left-hand document area first, and those viewing snippet were second most likely to gaze at it first.

Table 1. Summary of firstly-gazed area

	Text	Snippet	Term	Sum
Left	42	39	32	113
Right	3	6	13	22
Sum	45	45	45	135

4.2 Effect of topic on participants' viewing behaviors

Table 2 shows the number of participants who focused on the corresponding (common or specific) area more than another areas. It is counted in two categories: when documents of the same topic are presented and those of different topic are presented.

Table 2. Effect of topic on viewing behaviors

Topic	Snippet		Term	
	Common	Specific	Common	Specific
Same	22	5	19	12
Different	7	11	4	10

The table shows that in the same topic condition participants more frequently looked at common snippet / terms than specific ones. On the contrary, in the different topic condition, specific snippet / terms were more frequently looked at by participants than common ones. From the result, it is supposed that participants first assume whether given document pair relate with same topic or not, then examine the assumption by reading the corresponding information. That is, participants would carefully read specific terms/snippet when they assume the documents relate with different topic from each other. This result suggests participants need different kind of information according to assigned task. In that sense, separately providing common / specific information (snippets and terms) as employed in the paper is effective for supporting the task of similarity judgment.

4.3 Viewing behavior specific to terms condition

Figure 7 shows a gaze plot of viewing behavior that is specific to the case when terms are presented. The figure shows that the participant frequently switches AOIs between left-hand and right-hand document areas. This behavior was frequently observed when terms are presented. On the contrary, as shown in Fig. 8, such switching behavior was less observed when text or snippet is presented.

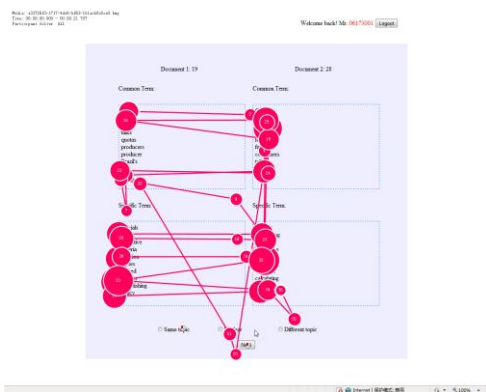


Fig 7. Frequent switches of AOIs when viewing terms

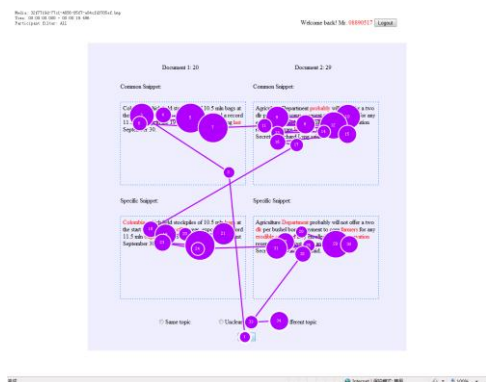


Fig 8. Fewer switches when viewing snippet

In order to investigate this tendency in more detail, we calculated the number of “switches”; a switch occurs when participants changed AOI from left-hand to right-hand documents or vice versa. Table 3 shows the average switching frequency per a participant that is counted per type of information. It shows that participants most frequently switched when terms are presented.

Table 3. Average switching frequency

	Text	Snippet	Term
Average frequency	4.49	7.06	12.34
STDEV	3.28	3.54	6.82

It seems that each term can be examined independently, whereas participants have to read a sequence of terms when snippets or original text is displayed. As the difference is expected to affect the user's feedback cost, the difference

between the type of providing information in terms of judgment time and accuracy should be investigated, which will be one of our future works.

5. Conclusions

This paper investigates the behaviors of users judging similarity of documents by eye-tracking analysis. The aim of the investigation is to obtain the hint for minimizing the cost of users judging similarity of documents, which is an essential task for users when performing interactive document clustering.

An experiment system was implemented, with using which 21 test participants were asked to judge the similarity of given pair of documents. As for the clue for the judgment, 3 types of information; original text, snippet, and term, are investigated. The behaviors of participants were recorded by T60 eye-tracking system. Analyzing the recorded data found typical viewing behaviors. The result shows that the areas at which the participants more frequently looked are different between the conditions of judging documents of same topics and different topics. It also shows that participants most frequently switched AOIs between documents when terms are presented.

Our future work include the design of interface that supports interactive document clustering based on constrained clustering method. The obtained result will contribute to realization of interface that can minimize the user's feedback cost.

References

[Okabe 07] M. Okabe and S. Yamada, “Semi-supervised Query Expansion with Minimal Feedback,” IEEE Trans. Knowledge and Data Engineering, Vol.19, No.11, pp.1585-1589, 2007.

[Basu 08] Basu, I. Davidson, and K. Wagstaff eds., Constrained Clustering: Advances in Algorithms, Theory, and Applications, Chapman & Hall, 2008.

[Onoda 07] T. Onoda, H. Murata, and S. Yamada, “SVM-based Interactive Document Retrieval with Active Learning,” New Generation Computing, Vol. 26, No. 1, pp. 49-61, 2007.

[Lorigo 06] L. Lorigo, B. Pan, H. Hembrooke, T. Joachims, L. Granka, and G. Gay, “The influence of task and gender on search and evaluation behavior using Google,” Information Processing and Management Vol. 42, No. 4, pp. 1123-1131, 2006.

[Rodden 07] K. Rodden and X. Fu, “Exploring How Mouse Movements Relate to Eye Movements on Web Search Results Pages,” SIGIR 2007 Workshop on Web Information Seeking and Interaction (WISI), pp.29-32, 2007.

[Sun 05] Q. Li, L. Sun, and J. Duan, “Web Page Viewing Behavior of Users: An Eye-Tracking Study,” Int'l Conf. on Services Systems and Services Management (ICSSSM'05), pp. 244-249, 2005.

[Cutrell 07] E. Cutrell, Z. Guan, “An eye-tracking study of information usage in Web search: Variations in target position and contextual snippet length,” CHI'07, pp. 407-416, 2007.

[Chen 09] M. Chen, S. Yamada, and Y. Takama, “An Investigation of Snippet Generation for Minimum User Feedback,” 23rd Annual Conference of the Japanese Society for Artificial Intelligence (JSAI2009), 2B2-1, 2009.