Capturing User Reading Behaviors for Personalized Document Summarization

Hao Jiang[*]
Department of Computer Science
The University of Hong Kong
Hong Kong S.A.R., P.R. China

Songhua Xu[*,⋆]
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831,
USA

Francis C.M. Lau
Department of Computer Science
The University of Hong Kong
Hong Kong S.A.R., P.R. China

ABSTRACT

We propose a new personalized document summarization method, which observes a user's reading behaviors, including user facial expressions, gaze positions, and reading durations, during his or her past reading activities to infer the user's personal reading preferences. Once a user's personal reading preferences are derived, our algorithm can then automatically generate document summarization in a personalized way. We compare the performance of our algorithm with that of a few peer algorithms and software packages. The result of our comparative study shows that our algorithm can produce superior personalized document summaries than those peer methods in that the automatic document summarization generated by our algorithm can better satisfy a user's personal preferences.

General Terms

Human Factors, Measurement, Performance

Author Keywords

Personalized document summarization, reading preference, personal preferences, facial expressions, gaze positions, reading durations

ACM Classification Keywords

H.5.2 User Interfaces: input devices and strategies, interaction styles; I.7.5 Document Capture: document analysis

INTRODUCTION

To cope with today's information explosion, automatic document summarization has increasingly become important, as heatedly pursued by many information science researchers. However, despite many fruitful research advances in the area, only very limited efforts have been dedicated to generating personalized summaries to suit for individual readers' preferences. Such a lack of research emphasis on personal document summarization leads to one of the major differences between human summaries and machine summaries generated by existing automatic algorithms. To address this problem, in this paper, we study how to generate personalized document summaries that observe the preferences of individual readers. To attain this goal, when generating a personalized document summary for a particular user, our algorithm attempts to best accommodate both the user's reading preferences and the document author's writing interests by selecting a few key sentences from the input document which can maximally include content words reflecting intentions of both sides.

One of the most closely related work to our study here is the automatic document summarization algorithm proposed by Nenkova et al. [8]. In their approach, a few sentences from the input document are selected to maximally cover the high frequency content words in the input document. Different from their approach, our method attempts to include key words from the document that can best satisfy a reader's personal reading interests and preferences, the latter of which are acquired from the reader's previous reading activities. In this way, the document summaries produced by our algorithm are customized for individual readers. The design of our algorithm is also influenced by the multi-document summarization algorithm proposed in [13] which maximizes the inclusion of informative content words. Similarly, the main difference between their algorithm and our method is that besides considering word level statistics in the original document when deciding whether to include a certain content word in a document's summary, our method additionally considers a reader's personal reading interests and preferences when making algorithmic decisions regarding document content selection during our extractive document summarization process.

As a specialized line of research, eye-tracking is recognized as a promising approach for obtaining implicit user feedbacks [3] to estimate the user's personal reading interests, for building personalized online recommendation systems, e.g. [10], and generating personalized document summaries [11]. Both commercial eye-trackers such as electro-oculographic (EOG) based systems [2] and web camera based commodity approaches such as [6] are available for detecting user gaze positions. Using web cameras, we can also simultaneously

[*] To be considered as equal first authors.
[⋆] Corresponding author; contact him at xus1 *at* ornl *dot* gov.

IUI'11, February 13–16, 2011, Palo Alto, California, USA.

detect user facial expressions as another type of implicit user feedbacks [12]. Witnessing such an advantage, in this paper, we adopt a web camera based approach to capture user facial expressions and gaze positions to understand a user's personal reading interests for generating personalized document summaries.

OBSERVING WORD LEVEL PERSONAL READING INTERESTS OF A USER

For each content word w_i, we introduce a user reading interest property, $\phi_{u_k}(w_i)$, to represent an arbitrary user u_k's reading interest over the word w_i. We rely on this property to represent the user's personal reading interests and preferences. Such a property will play a critical role in our personalized document summarization process, which will be discussed later in this paper. To acquire the reading interests of a user over individual words, in this work, we observe and analyze a user's facial expressions, gaze positions, and reading time durations exhibited in the user's previous reading activities.

Detecting User Facial Expressions for Estimating Reading Concentration

To detect user facial expressions during one's reading or browsing activities, we adopt a recent facial emotion analyzer software package, called "eMotion" [1]. The output of eMotion is a series of probabilistic values that represent the likelihood of a user displaying certain facial expressions. In our approach, we use eMotion's facial expression detection output to estimate the reading concentration of a user, i.e. the user's reading attentiveness.

Let $rc(u_k, t)$ be the *reading concentration* of the user u_k at the moment t. For a specific user u_k, a larger reading concentration value indicates more attentively the user reads a webpage or a document, which further implies that the contents of the reading materials more strongly attract the user. As mentioned above, we measure a user's reading concentration at any moment according to the output from the facial emotion analyzer. In our current algorithm implementation, $rc(u_k, t)$ is estimated as follows:

$$rc(u_k, t) = 1 - F_{neutral}, \quad (1)$$

where $F_{neutral}$'s value range is $[0, 1]$, representing the probability of having a neutral facial emotion at the moment t, as detected by the human facial emotion capturing software eMotion [1]. Here we assume if a piece of information appears interesting to a user, the user tends to display some non-neutral expressions when encountering the information.

Detecting User Gaze Positions for Estimating Reading Zone

As mentioned earlier, our approach uses a web camera as the basic input device, which is coupled with computer vision techniques to track a user's eye movement. We did not choose commercial eye-trackers, mainly because of their high cost. In our system, we detect a user's gaze positions using an off-the-shelf software package, called "Enable Viacam" [7].

Deriving Observed Word Level User Reading Interest

Once a user's reading concentration over a document page or a webpage is captured, we then uniformly assign the captured reading concentration samples to all the visible individual words at that moment. In the following, we will no longer differ a document from a webpage, as we apply the same procedure to process user reading concentration data obtained over both types of reading materials.

Let $[t_0, t_1]$ be the duration captured of the user u_k when he or she reads an actively displayed reading zone Ω. We assume there are n distinct content words, i.e. verbs, nouns, adjectives, and adverbs, appearing in Ω, which are denoted as $w_1, w_2, \cdots, w_n$ respectively. Given these notations, the amount of reading concentration samples assigned to the word w_i, denoted as $\phi_{u_k}(w_i)$, is computed as a weighted fraction of the integral of $rc(u_k, t)$ over the time period of $[t_0, t_1]$, where the weight is determined according to w_i's occurrence times versus the total number of occurrences of all the content words displayed in the period of $[t_0, t_1]$.

ESTIMATING PERSONAL READING INTERESTS OF A USER OVER NEW WORDS AND SENTENCES

Estimating Personal Reading Interests Over New Words

For an arbitrary word w_s which is previously unseen by a user u_k, we first identify all the words that the user has read in the past, with which the pairwise word semantic similarities are above a certain threshold. Let $Sim(w_i, w_j)$ be the semantic similarity between a pair of words w_i and w_j, where $Sim(w_i, w_j) \in [0, 1]$. In our current algorithm implementation, we calculate $Sim(w_i, w_j)$ using the semantic similarity estimation algorithm proposed in [5] due to its relative ease of implementation and the method's satisfying performance. We empirically tune the minimum word similarity threshold to be 0.1. Under such a threshold, assume we find a total of n words which the user u_k has read previously. Without loss of generality, we denote these words as $w_{1,u_k}, \cdots, w_{n,u_k}$ respectively, and call them the *sample words*. We then use the following equation to estimate user u_k's personal reading interest over the word w_s as follows:

$$\phi_{u_k}(w_s) = \frac{\sum_{j=1}^{n} \left(Sim(w_s, w_{j,u_k}) \phi_{u_k}(w_{j,u_k}) \right)}{\sum_{j=1}^{n} Sim(w_s, w_{j,u_k})}. \quad (2)$$

To understand the meanings of the above equation, basically, the more similar a sample word w_{j,u_k} is to the target word w_s, the more influential the sample word is when estimating personal reading interests over the target word.

Estimating Personal Reading Interests Over a Sentence

We assume the more user interested content words appear in a sentence, the more intriguing the sentence appears to the user as the sentence either delivers more personally interested information to the user or better satisfies the user's personal reading preferences. Closely related to our study here is Nenkova et al. [8]'s work on combining frequencies of individual words to derive a sentence's importance for document summarization. In their study, they examined three combination forms — multiplication, summation,

and average. Under the three combination forms, the importance of a sentence is determined respectively according to the product, sum, and average of the frequencies of all the content words in the sentence. Inspired by their work, in this paper, we also assume the overall importance of a sentence, i.e. the value to include the sentence in the document's summary, is a function of the properties of content words in the sentence. Given a sentence S, we first select the sentence's top five words which hold the highest reading interest values for a given user u_k. We denote the user's personal reading interest values over these five words as $\phi_{u_k}(w_1), \phi_{u_k}(w_2), \cdots, \phi_{u_k}(w_5)$ respectively. And then, we take the sum of the algebraic and arithmetic averages of these five words' personal reading interest values as the sentence's overall personal reading interest value for the user. The reason why we adopt both types of averaging operators when deriving a sentence's overall reading interest is due to the respective advantages demonstrated by the multiplication and summation methods for synthesizing a sentence's overall importance based on the importance of individual words, as revealed by the prior study [8].

GENERATING PERSONALIZED DOCUMENT SUMMARY

Once a user's personal reading interest value over each sentence in a document is known, we can then generate a personalized summary of the document. In our document summarization process, we first implement the algorithm presented in [4] to quantify the importance of individual sentences in a document. We denote the importance of a sentence S as estimated by their method as $\chi(S)$. For each sentence S in the input document, we will then estimate user u_k's personal reading interest over the sentence as $\phi_{u_k}(S)$ according to the method presented in the previous section. After we compute the respective $\chi(S)$ and $\phi_{u_k}(S)$ values for all the sentences in the input document, we normalize the two values respectively to ensure that the highest $\chi(S)$ and $\phi_{u_k}(S)$ values for a sentence are always 1. Assuming for a specific document summarization task, the specified compression rate is $c\%$. Under that circumstance, we will select the top $c\%$ sentences, which carry the largest values of $\chi + \alpha\phi_{u_k}$ in the document, to compose a personalized summary for the document. Here, α is a user tunable parameter, whose default value is 1.

EVALUATION AND EXPERIMENT RESULTS

We collected five sets of articles for evaluating the performance of our algorithm, including science/technology research papers, and readings for entertainment or leisure. More concretely, we randomly selected 60 articles from each of the following sources to form an article set respectively: 1) "ACM Digital Library"; 2) "Science" magazine website; 3) "National Geographic"; 4) novel chapters in "Free Online Novels"; 5) "New York Times". We expect these article sets cover a diverse range of topics appealing to readers of different knowledge background, preferences, and reading behaviors. A few key statistics of the five article sets are reported in Table 1.

For each article set, twelve people were invited in our evaluation study. Each of them was asked to read ten uniformly

Article Set	I	II	III	IV	V	Overall
Articles in the set	60	60	60	60	60	300
Words per article	6684	979	1111	3906	942	2724
Sentences per article	304.1	37.6	51.6	236.8	53.2	136.7
Paragraphs per article	52.8	9.1	21.1	69.4	11.3	32.8
Pages per article	8.0	1.2	1.0	7.4	1.4	3.8
Sentences per summary	20.8	12.4	15.5	32.9	14.7	19.3
Manual compression rate	6.9%	33.0%	30.0%	13.9%	27.6%	22.3%

Table 1. Key statistics of the five sets of articles used in our experiment. Set "i" means the i-th article set ($i = I, \cdots, V$). The column titled "Overall" reports the statistics over all the five article sets. "Manual compression rate" is the average ratio between word lengths of user manually produced document summaries and word lengths of the original documents in an article set.

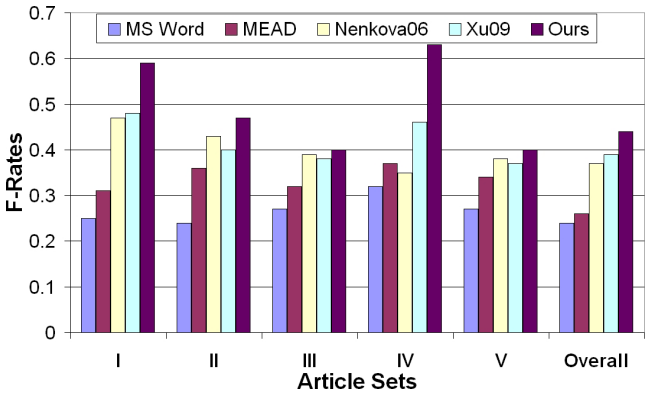

Figure 1. F-rates of automatic document summarization by our personalized document summarization algorithm and four peer methods respectively over the five article sets separately and as a whole.

randomly selected articles from the article set. Therefore, on average, each article in a set will be read by two people. After reading each article, the participant was asked to provide a summary for the article by selecting a subset of the sentences in the article that best describe the contents of the article. All the manually produced summaries composed by a participant in this way were treated as the groundtruth document summaries for the participant. During our evaluation experiment, we use the subject's reading behaviors captured during his or her readings over nine of the ten articles to infer the person's reading interests using our algorithm introduced at this paper. A person's reading behaviors captured from these nine of the ten reading sessions were used to generate a machine summary of the tenth article via our algorithm. When generating automatic document summaries, the objective compression rate is set the same as that of the corresponding article's groundtruth summary for the corresponding participant. In our experiments, we measured the quality of an article summary produced by our algorithm

against the corresponding manually produced human summary result using the F-rate evaluation metric. The overall performance of our algorithm was measured as the average performance of our algorithm for all the articles read and summarized by the twelve participants. In Figure 1, we report the measured F-rates of our algorithm in performing automatic document summarization over the five article sets respectively as well as our algorithm's overall performance for document summarization over the entire five article sets as a whole.

In our evaluation, we also compared the performance of our algorithm with four other peer summarization methods, including "NenKova06" [8]—a recent summarization algorithm which outperforms many earlier summarization algorithms, "Xu09" [11]—a recent personalized summarization algorithm which captures personal reading interests through eye-tracking for automatic document summarization, and two popular document summarization software packages—"Microsoft Word AutoSummarize" as provided in Microsoft Office Professional Edition 2003 and the MEAD summarizer system [9]. Figure 1 reports the averaged F-rate measurements for each of the four peer methods for automatic document summarization over the five article sets respectively as well as their respective overall performance of automatic document summarization over the entire five article sets as a whole. As revealed by these comparative experiment results, our algorithm consistently outperforms the other four peer methods in all the document summarization tasks.

CONCLUSION

In this paper, we proposed a new algorithm for personalized document summarization, which respects both the author's writing purpose and the readers' reading interests during the automatic document summarization process. In contrast to traditional document summarization algorithms, our new algorithm observes readers' individual reading interests during automatic document summarization process, an issue often overlooked by existing automatic summarization algorithms. Our algorithm has yielded very positive results in a series of experiments, in comparison with several recently proposed summarization algorithms and popular software packages. These experiment results confirm the effectiveness and advantages of our newly proposed algorithm for personalized automatic document summarization.

Acknowledgement

Songhua Xu performed this research partially as a Eugene P. Wigner Fellow and staff member at the Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. Department of Energy under Contract DE-AC05-00OR22725.

REFERENCES

1. eMotion, Visual Recognition. http://www.visual-recognition.nl, 2006-2008.
2. A. Bulling, J. A.Ward, H. Gellersen, and G. Tröster. Robust recognition of reading activity in transit using wearable electrooculography. In *Pervasive '08: Proceedings of the 6th International Conference on Pervasive Computing*, pages 19–37, 2008.
3. E. H. Chi, M. Gumbrecht, and L. Hong. Visual foraging of highlighted text: An eye-tracking study. In *HCII '07: Proceedings of HCI International Conference*, pages 589–598, 2007.
4. J. Leskovec, N. Milic-frayling, M. Grobelnik, and J. Leskovec. Extracting summary sentences based on the document semantic graph. Technical report, Microsoft Research (MSR-TR-2005-07), 2005.
5. Y. Li, Z. A. Bandar, and D. Mclean. An approach for measuring semantic similarity between words using multiple information sources. *IEEE Transactions on Knowledge and Data Engineering*, 15(4):871–882, 2003.
6. Y.-P. Lin, Y.-P. Chao, C.-C. Lin, and J.-H. Chen. Webcam mouse using face and eye tracking in various illumination environments. *EMBS '05: Proceedings of 27th IEEE Annual International Conference of Engineering in Medicine and Biology Society*, pages 3738–3741, 2005.
7. C. M. Loba. Enable Viacam, CREA Software Systems. http://eviacam.sourceforge.net, 2008-2009.
8. A. Nenkova, L. Vanderwende, and K. McKeown. A compositional context sensitive multi-document summarizer: exploring the factors that influence summarization. In *SIGIR '06: Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 573–580, New York, NY, USA, 2006. ACM.
9. D. R. Radev, S. Teufel, H. Saggion, W. Lam, J. Blitzer, H. Qi, A. Çelebi, D. Liu, and E. Drabek. Evaluation challenges in large-scale multi-document summarization: the mead project. In *ACL 2003: Proceedings of the Annual Meeting of the Association of Computational Linguistics*, pages 375–382, Sapporo, Japan, 2003.
10. S. Xu, H. Jiang, and F. C. Lau. Personalized online document, image and video recommendation via commodity eye-tracking. In *RecSys '08: Proceedings of the 2008 ACM Conference on Recommender systems*, pages 83–90, New York, NY, USA, 2008. ACM.
11. S. Xu, H. Jiang, and F. C. Lau. User-oriented document summarization through vision-based eye-tracking. In *IUI '09: Proceedings of the 13th International Conference on Intelligent User Interfaces*, pages 7–16, New York, NY, USA, 2009. ACM.
12. S. Xu, H. Jiang, and F. C. Lau. Observing facial expressions and gaze positions for personalized webpage recommendation. In *ICEC '10: Proceedings of the 12th International Conference on Electronic Commerce*, pages 77–86, 2010.
13. W. T. Yih, J. Goodman, L. Vanderwende, and H. Suzuki. Multi-document summarization by maximizing informative content-words. In *IJCAI'07: Proceedings of the 20th International Joint Conference on Artifical Intelligence*, pages 1776–1782, 2007.