Recommendation System Keeping Both Precision and Recall Based on Uninteresting Information

Tsukasa Kondo
Graduate School of Science and Engineering Ritsumeikan University, Shiga, Japan
Email: tsukasa@de.is.ritsumei.ac.jp

Fumiko Harada and Hiromitsu Shimakawa
College of Information Science and Engineering, Shiga, Japan
Email: {harada, simakawa}@cs.ritsumei.ac.jp

***Abstract*—The recommendation system which recommend the interesting information for the target user must keep the high value of the precision and recall. However there is a trade-off relationship between precision and recall. In this paper, we propose the recommendation method keeping both precision and recall. The proposed method extracts the indicator from not only interesting information but also uninteresting information for the target user in the user-preferred-genre. Proposed method can keep the precision and recall by exclude the uninteresting information based on extracted indicator. From result of an experiment, we have verified the proposed method can improve the precision and recall.**

***Index Terms*—recommendation system, uninterest indicator, precision, recall, trade-off**

I. Introduction

Recently, there are many web pages in the World Wide Web. Since it is difficult for users to pick up the web pages that is interesting for themselves, recommendation systems have been developed to support users to find interesting web pages. Almost all existing recommendation systems judge the web page that is interesting to a user from him behavior. For example, user preference for a specific web pages is judged from the browsing time of the web page [1] or mouse operations [2]. In this way, recommendation systems judge the web page interested by a user. They extract the indicators to decide the web page recommended to him.

II. Need for Uninterest Indicator for Recommendation

A. Trade-off Between Precision and Recall

Recommendation systems must keep the high precision and high recall. However, there is a trade-off between them [3]. Let us consider Fig. 1, when the recommendation system recommends the web pages including the word "Keisuke Honda", the name of a famous Japanese football player, for the target user for viewing the web page whose content is "Keisuke Honda scored the goal in the international game". The indictor is only "Keisuke Honda", which is a weak constraint. Suppose that three web pages A, B, and C in Fig. 1 are interesting to the target user. The two web pages, D and E, though they match the indicator "Keisuke Honda", are supposed to be uninteresting to the target user, because they are not related to any excitement of the target user. This example indicates, the information matching the target user indicator is not always interesting for him. On the other hand, we can consider extracting stronger indicator in order to recommend only web pages that is interesting to the target user. Suppose the indicators "Keisuke Honda scored the goal" and "Keisuke Honda, a member of the national team" as indicator 2 in Fig. 1. With indicator 2, the recommendation system can recommend only the web pages A and B. However, it cannot recommend web page C, if no web pages in browsing history of the target user can be extracted with indicator 2. Therefore, if a recommendation system uses indicator 2, it cannot recommend C. In this way, any recommendation system cannot cover the all web pages that are interesting to the target user, using strong indicators.

Figure 1. Recommendation example, cannot guarantee the precision and recall.

Manuscript received Nov 30, 2012; revised Jan 28, 2013.

doi: 10.12720/joace.1.3.202-207

B. Interest Indicator and Uuninterest Indicator

There are two kinds of the information matching the indicator that is interesting to the target user. One appears in the user choices. In Fig. 1, the target user is aware that the information "Keisuke Honda participated in the Japan national team" is interesting to him. Because of this fact, the web page relevant to "Keisuke Honda participated in the Japan national team" has appeared in browsing history of the target user. Another one does not appear in the target user choices. It is the information that the target user finds to be interesting to him after make a browse it actually. The target user is not aware of such information is interesting before the browse. It implies that such information does not appear in the target user choices. In Fig. 1, suppose the target user does not know the information "Why Keisuke Honda achieves good performance in the games? ". In this case, target user cannot be aware of that it is interesting to himself. No matter how the recommendation indicator is set to be strong, the recommendation system cannot give priority to web pages whose similar pages do not appear in the target user choices. In Fig. 1, even if the indicator implies "Keisuke Honda" the recommendation system cannot generate the indicator leading to specific web pages about "Keisuke Honda" whose content is not similar to any pages which appear in the browsing history. If we can predict uninteresting web pages relevant to indicator 1 can make novel indicator from it, which enables the recommendation system to excludes the uninteresting ones. The recommendation system could cover the all web pages interesting to the target user, including ones which content is not similar to any pages appearing in the browsing history. The web pages D and E are not interesting for the target user in Fig. 1. Suppose a novel indicator which prevents the information relevant to D and E from the recommendation on "Keisuke Honda". With such indicator, a recommendation system can recommend the all web pages that are interesting to the target user in such as A,B and C, while D and E can be excluded in recommendation.

If the recommendation system can extract the indicator not only from the target user choices but also from uninteresting information, it can exclude only uninteresting ones from recommendation candidates. It can guarantee high precision and high recall.

C. Related Work

There are some researches to use information uninteresting to the target user to improve performance of the search engine and recommendation system. The methods proposed in [4] and [5] improve the result of search engine use the word the target user wants to exclude from result.

The method in [6] could be more accurate for recommendation using an indicator extracted from news articles which match this condition, "When the target user does not view a news article even though it is presented in the first time in online news service, we can assume it is uninteresting to him". However, target users have to choose words that he want to exclude from search results by himself in [4], [5]. The method in [6] can extract the indictor only to online news subscribers. Furthermore, the method in [6] uses the all words appearing in titles of articles uninteresting to the target user. It might use the word the target user has an interest as an indicator. If we want to use the uninteresting information as an indicator for recommendation, we not only have to extract the word uninteresting to the target user, but also we should automate the extraction of an indicator corresponding to the uninteresting word.

III. Proposed Method

A. Method over View

In this research, the web page interesting to the target user is referred to as an interesting web page. The web page uninteresting to the target user is referred to as an uninteresting web page. The indicator extracted from the former and the letter are referred to as an interest indicator and respectively an uninterest indicator.

In this paper, we propose the recommendation system based on interest indicator and uninterest indicator to keep high values in both of the precision and the recall. The proposed method automatically extracts the interest indicator and uninterest indicator. The method adds positive weights to web pages matching to the interest indicator, while negative weight to web pages matching to the uninterest indicator. Fig. 2 shows the flow of the proposed method, which has four steps.

- Step1: In the first step, target user chooses the word for recommendation, so that our method recommends the web pages which have contents relevant to the chosen word. The method extracts the web pages which have contents relevant to the chosen word from target user browsing history. The extracted web page is referred to as an related web page.
- Step2: Using browsing time and bookmark of related web pages, the method extracts the interest indicator.
- Step3: In the third step, the method extracts the uninterest indicator same way.
- Step4: In the final step, the proposed method decides the web pages recommend to the target user based on interest indicator and uninterest indicator.

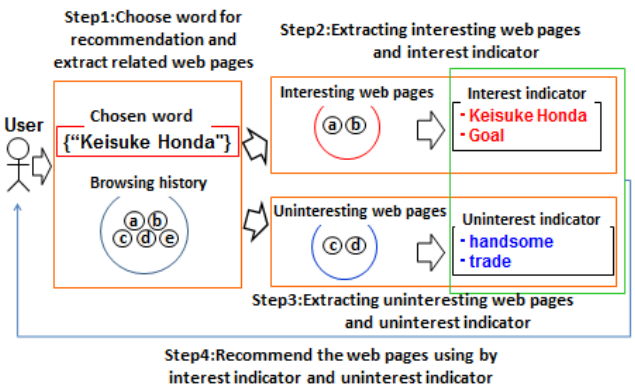


Figure 2. Method over view.

B. Related Web Pages

The word chosen by the target user for recommendation is referred to as a chosen word. Proposed method extracts the related web pages from target user browsing history. We suppose that the web page which includes the chosen word in the title or body text has probability of related web page. The proposed method extracts those web pages as related web pages.

C. Extracting Interest Indicator

Interesting web pages are picked up from related web pages using browsing time and bookmark. If the web page to satisfy at least one of the following conditions, proposed method extracts it as an interesting web page.

- A web page is browsed by the target user more than threshold θ_1 seconds
- A web page is bookmarked by the target user

The first one comes from an article interested by the target user is browsed for a long time [1]. It is reasonable to assume that bookmarking a web page is the user declaration to visit the web page again in the future. It seems to contain some information interesting to for him.

A user judges whether a web page is interesting or not based on its contents. Interest indicator has to be characteristic word can represent the contents of the interesting web page. Therefore, the method extracts the characteristic word of the interesting web pages based on TFIDF. We can suppose that characteristic words of the interesting web page have high TFIDF values with high probability. Extracts the five words which have top five TFIDF value in the each interesting web pages as an interest indicator. The chosen word is included into interest indicator regardless of its TFIDF value.

D. Extracting Uninterest Indicator

Uninteresting web pages are picked up from related web pages using browsing time and bookmark. If the web page to satisfy both of the following conditions, the proposed method extracts it as an uninteresting web page.

- A web page is browsed by the target user less than θ_2 seconds.
- A web page is not bookmarked by the target user.

A user stops browsing the web pages when the moment he has no interest to it [1]. Because of this fact, uninteresting web page has short browsing time. Therefore, we regard the web page browsed by the target user less than θ_2 seconds as an uninteresting web page. Any bookmarked web page is out of candidate of uninteresting web page. Since the target user bookmarks the web page to read it lately. We should not include the web page into uninteresting web pages, even if the browsing time of it is short.

Any uninteresting web page has a reason why target user judged it uninteresting for him at that time. A user judges whether a web page is interesting or not for himself based on its contents. We can suppose that uninteresting web page has a characteristic word which does not appear in the interesting web pages. Comparing the words including in the interesting web pages and uninteresting web pages, the method extracts which only appear in uninteresting web pages the words as an uninterest indicator.

E. Recommendation of the Web Pages

The proposed method decides the web pages to recommend to the target user with following (1).

$$\begin{cases} S(W) = \alpha \times \sum_{n=1}^{N} i(n) - (1-\alpha) \times \sum_{m=1}^{M} u(m) \\ \alpha > 0 \end{cases} \quad (1)$$

Let W, M and N stand for the web pages target user has not browsed, total number of the interest indicator include in the W, the total number of the uninterest indicator include the W, respectively. The score of interest indicator about the k-th word is represented with $i(n)(1 \le k \le N)$ the score of uninterest indicator about the l-th word is represented with $u(m)(1 \le l \le M)$. Parameter α is constant to control the weight for the interest indicator and the uninterest indicator. Equation (1) calculates the importance of W for the target user. If W include a word in the interest indicator, add the *IDF* value of it to *S(W)*. If W include the word in the uninterest indicator, subtract the *IDF* value of it from *S(W)*. In this way, the more interest indicator W has, the more *S(W)* is similarity, the more uninterest indicator W has, the less *S(W)* is. Proposed method uses the *IDF* value because TFIDF value is changed by length of body text. Therefore, TFIDF value cannot evaluate the each web page fairly. The proposed method recommends the web pages to the target user in the descending order of *S(W)*.

IV. Experiment to Evaluate the Proposed Method

A. Verification Items

To evaluate the validity of proposed method we have to investigate the following two verification items.

- Investigate the threshold θ_1 and θ_2 to judge the interesting web page and uninteresting web page and its accuracy.

  The proposed method judges the interesting web page and uninteresting web page to extract the interest indicator and uninterest indicator. Therefore, we have investigated the browsing time of interesting web page and uninteresting web page to decide the threshold θ_1 and θ_2. Furthermore, we have investigated the accuracy of judging interesting web page and uninteresting web page using threshold θ_1 and θ_2.
- precision and recall

  We have investigated the proposed method can improve the precision and recall more than existing method or not. We have investigated the precision and recall the case change the parameter α.

B. Collecting Data for Verification

The examinees of this experiment are six university students in IT department. The examinees have browsed the web pages and evaluated it. The total number of the

evaluated web pages by examinees varies from 81 to 199. We have prepared 99554 for candidate of browsed and evaluated by examinees. Data for verification have collected by following steps.

1). Each examinee chooses a word for recommendation. We have recommended the web pages which include the chosen word at random from prepared web pages.
2). The examinee browses to evaluate recommended web pages.
Browsed and evaluated recommended web pages by examinees. There is no need to browse web page completely. The examinees could finish the browsing by himself at any time. There are two evaluation items.
A) "Are there any interesting or useful information .in the recommended web page?"
The first evaluation item judges the whether recommended web page is interesting or not for the examinee and evaluated by the following four-grade evaluation. 4: It is very interesting web page. 3: It is interesting web page. 2: It is uninteresting web page. 1: I hate this web page.
B) "Do you want bookmark?"
The second evaluation item judges whether recommended web page is bookmarked or not by the examinee. 1: Yes, I want to bookmark it. 0: No, do not bookmark it.
3). We have measured the each recommended web pages.

The data we have got in this section called verification data. In the verification data, the web page evaluated as 4 or 3 in the first evaluation item A is referred to as an interesting web page, the web page evaluated as 2 or 1 is referred to as an uninteresting web page.

C. Interesting Web Page and Uninteresting Web Page

Table I shows the total number of interesting web page, bookmark, uninteresting web page and all web pages evaluated by examinee and Table. II shows the average browsing time of interesting web page, uninteresting web page and all web pages. The average browsing time of interesting web page is longer than that of all web pages, and the average browsing time of uninteresting web page is shorter than that of all web pages. All of examinees have this feature. Examinee D was always browse the web pages completely. If the browsed web page was uninteresting to D, he browses roughly. Because of this fact, his browsing time is longer than other examinees. Since the browsing time has individual difference, we have normalized the browsing time of examinee by the deviation. Table III shows the deviation of the interesting web page and the uninteresting web page by each examinee. The deviation of the interesting web page has few variance, which the deviation of the uninteresting web page has almost no variance. The deviation may be caused by the difference in the length of the body text. The web page which has long body text makes examinee use the long time for browsing. On the other hand, examinees stopped browsing of the uninteresting web page when it judged uninteresting for them. We consider

TABLE I. TOTAL NUMBER OF THE EACH WEB PAGES

Examinee	Interesting	Bookmark	Uninteresting	All
A	108	62	89	197
B	74	11	125	199
C	41	18	40	81
D	79	75	71	150
E	93	4	65	158
F	42	9	106	148

TABLE II. AVERAGE BROWSING TIME

Examinee	Interesting	Uninteresting	All
A	32sec	20sec	27sec
B	40sec	15sec	24sec
C	88sec	43sec	66sec
D	110sec	75sec	93sec
E	31sec	17sec	25sec
F	47sec	9sec	20sec

TABLE III. COMPARISON OF DEVIATION

Examinee	Interesting	Uninteresting
A	52.7	46.8
B	57.8	45.3
C	53.4	46.5
D	52.3	47.5
E	53	45.7
F	59.6	46.2
AVE	54.8	46.3

that body text of the uninteresting web page was not browsed completely therefore it has almost no variance. As a result, we use the deviation value 54.8 for set the threshold θ_1 and deviation value 46.3 for set the threshold θ_2. We calculate the threshold θ_1 and θ_2 with (2) and (3). The standard deviation of the browsing time and the, average browsing time referred to as an *SD* and *AB*. *SD* and *AB* is an individual value of the each examinees.

$$\theta_1 = \frac{54.8 - 50}{10} \times SD + AB \quad (2)$$

$$\theta_2 = \frac{46.3 - 50}{10} \times SD + AB \quad (3)$$

We use θ_1 to the condition of the interesting web page for the judgment of the interesting web page from verification data.

Table IV shows the precision and recall, the case we use only browsing time, bookmark and both of them for the judgment of the interesting web page. In the case only use browsing time for judgment, the average precision is 82% and the average recall is 39%. When we use bookmark, average precision is 94% and the average recall is 37%. The case only use bookmark, we could judge the interesting web page with high precision however recall has individual difference. Examinees A, C and D often make the bookmark, while B, E and F have few bookmark. The case using only the browsing time shows less accuracy than the case using only bookmark. However, we could guarantee at least 29% recall. The case we use both of the factors for judgment, the average

precision is 87%, average recall is 63%. Table V shows the precision and recall in case we use the browsing time and bookmark for the judgment of the uninteresting web page. The average precision is 82%,

TABLE IV. JUDGMENT OF INTERESTING WEB PAGES.

Examinee	Browsing time		Bookmark		Both	
	Precision	Recall	Precision	Recall	Precision	Recall
A	74%	30%	98%	56%	87%	77%
B	83%	51%	100%	15%	84%	57%
C	80%	29%	100%	44%	90%	63%
D	70%	33%	88%	84%	79%	86%
E	91%	33%	75%	3%	89%	35%
F	92%	57%	100%	21%	92%	57%
AVE	82%	39%	94%	37%	87%	63%

TABLE V.JUDGMENT OF UNINTERESTING WEB PAGES.

Examinee	Precision	Recall
A	84%	55%
B	92%	75%
C	77%	57%
D	82%	52%
E	63%	69%
F	95%	68%
AVE	82%	63%

and the average recall is 63%. The precision of the examinee E is less than other examinees in accuracy. The proposed method extracts the uninterest indicator comparing by interesting web pages and uninteresting web pages. Because of this, we suppose that if proposed method can judge the interesting web pages with high precision, there is less influence for extracting the uninterest indicator.

As a result, we should use the deviation value 54.8 in the browsing time of each examinees for calculate the threshold θ_1. We should use the deviation value 46.3 for calculate the threshold θ_2.

D. Evaluation by Rankscore

If proposed method can make many interesting web pages are placed in higher rank, many uninteresting web pages are placed in lower rank than exiting method, we can consider that propose method keep the both precision and recall regardless of the number of recommended web pages. We have evaluated the proposed method by how recommendation rank of the test data set is good. We considered that the case where α =1.0 in the (1) as an existing method, the case where α ={0.9, 0.8, ⋯, 0.2, 0.1} as the proposed method. Based on result of subsection 4.C, we have extracted 15 web pages for each of interesting web pages and uninteresting ones. The remainder of the verification data used for test data set. In this verification, we use the evaluation indicator Rankscore proposed by Breese [7]. The test data set ranked for recommendation by the proposed method is referred to as a L. Ranksocre evaluates how L is good or not. If interesting web pages are placed in higher ranks, the value of Rankscore gets higher. On the hand, if uninteresting web pages are placed in higher ranks, the value of Rankscore gets lower. Proposed method make the web page which include the interest indicator placed in high rank relatively by reduce the *S(W)* of the web page which include the uninterest indicator. If the proposed method works effectively the Ranksore of its recommendation result better than that of existing methods. Ransksore is calculated with (4), (5), (6). The Rankscore is referred to as a *RS* for L. β is the parameter to control the *RS* by the rank of the interesting web page. Method $r(W_i)$ gets the rank of the interesting web page W_i in the L. Method $idx(W_i)$ gets the highest rank of the interesting web page W_i in the L.

$$RS_p = \sum_{W_i \in L} \frac{1}{2^{\frac{r(W_i)-1}{\beta}}} \tag{4}$$

$$RS_m = \sum_{W_i \in L} \frac{1}{2^{\frac{idx(W_i)-1}{\beta}}} \tag{5}$$

$$RS = \frac{RS_p}{RS_m} \tag{6}$$

RS_p is the one kind of the *RS* calculated by the rank of the interesting web pages in the L. RS_m is the one kind of the *RS* if interesting web pages can get the most highest rank in the L. The division of RS_p by RS_m can investigate how L is close to the most effective L. For example, suppose that there are ten web pages in the test data set. Let three web pages be interesting web page, while other seven web pages are uninteresting web pages. The interesting web pages are supposed to be ranked in first, sixth and tenth rank in the L. In case β = 10, RS_p = 2.2429. RS_m is the value in the case of the highest rank in the L. In this case, the interesting web pages are ranked in first, second and third rank in the L. RS_m = 2.8035, finally the *RS* = 0.80003.

Let us compare the existing method and proposed method using *RS*. Fig. 3 shows the value subtracted the *RS* the case α = 1.0 from the *RS* the case α = {0.9, 0.8, ⋯, 0.2, 0.1}. According to the Fig. 3, examinees A and E improve the *RS* reduce the α. However examinees B, C and F get the highest *RS* in case α = 0.5 or 0.6. In case α = 1.0, *RS* is the lowest. The *RS* of examinee D is gradually reduce as α gets lose to 0.1. There is no tendency that stronger weights for the uninterest indicator make *RS* improve. This is because uninterest indicator includes noise words. If the proposed method extracts the uninterest indicator perfectly, the propose method should obtain the highest *RS* in the case α = 1.0. However, synonym words of the interest indicator and general words include as an uninterest indicator when the proposed method extract the uninterest indicator. Because of this fact, the weight for the uninterest indicator is too strong when the proposed method works when α = 1.0. In that case, interesting web pages are ranked lower in the L. Actually, in case α = {0.8, 0.7, 0.6, 0.5} where the

weight for interest indicator is stronger than or equal to that of the uninterest indicator, all examinees except examinee D improve *RS*.

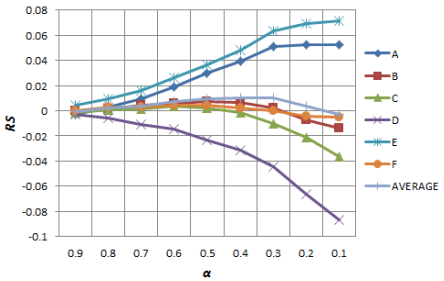


Figure 3 Relationship between *RS* and α

In case α = 1.0, we have achieved the highest average *RS*. We can conclude the most effective value for α in 0.5.

The *RS* of the examinee D calculated with the proposed method is lower than existing method ones. This is because the uninterest indicator is not extracted effectively for examinee D. The proposed method extracts the uninterest indicator by comparing the interesting web pages and uninteresting web pages. Therefore the proposed method cannot extract the uninterest indicator which appears in the web pages, if the method misjudges it as interesting web page. Since the browsing time of examinee D depends on the length of the body text, the proposed method have misjudged the uninteresting web page as an interesting web page. Therefore we fail to confirm the improvement of *RS*.

The, uninterest indicator can improve the *RS*. However, *RS* of the examinee D calculated with the proposed method is lower than that calculated with the existing method ones. We have two future works. First, we should decide more effective threshold θ_1 and θ_2 by considering the length of the body text. It would make the proposed method cover more examinees to improve *RS*. Second, we should prepare a method for noise words the proposed method extracts as the uninterest indicator. We have to accomplish the proposed method avoids extracting the synonym of the interest indicator and general word with low IDF value as the uninterest indicator.

V. Conclusion

In this paper, we have proposed a method to recommend the web page keeping both precision and recall based on interest indicator uninterest indicator. The proposed method extracts the interesting web pages and uninteresting web pages automatically to determine the interest indicator and uninterest indicator.

We have compared the existing method and the proposed method in the experiment from the view point of *RS*. The proposed method improves *RS* of 5 of 6 examinees. This result shows that proposed method can recommend the web page with high precision and high recall regardless of the number of the recommended web pages. However, the proposed method could not improve *RS* when the judgment precision of the interesting web pages is low. To improve *RS* of all examinees, we will try to improve the judgment precision of the interesting web pages.

References

[1] M. Morita and Y. Shinoda, "Information filtering based on user behavior analysis and best match text retrieval," in *Proc. 17th Annual International ACM-SiGiR Conference on Research and Development in Information Retrieval*, Dublin, 1994, pp. 272-281.

[2] H. Sakagami and T. Kamba, "Learning personal preferences on online newspaper articles from user behaviors," in *Proc. Sixth International World Wide Web Conference, In Computer Networks and ISDN Systems*, Sep 1997, vol. 29, pp. 1447-1455.

[3] M. Buckland and F. Gey, "The relationship between recall and precision," *Journal of the American Society for Information Science*, vol. 45, pp. 12-19, Jan 1994.

[4] Google. [Online]. Available: https://www.google.co.jp/

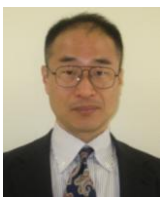
[5] T. Yamamoto, S. Nakamura, and K. Tanaka, "Rerank everything: A reranking interface for exploring search results," in *Proc. 20th ACM Conference on Information and Knowledge Management*, Glasgow, 2011, pp.1913-1916.

[6] K. Otsuki, G. Hattori, H. Haruno, K. Matsumoto, and F. Sugaya, "A preference cluster management method based on user access/non-viewed logs for online news delivery toward portable terminals," *DBSJ letters*, vol. 6, pp. 37-40, June 2007.

[7] J. S. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," in *Proc. Fourteenth Annual Conference on Uncertainly in Artificial Intelligence*, Madison, 1998, pp. 43-52.

Tsukasa Kondo received B.E from Ritsumeikan University in 2011. He advanced Graduate School of Ritsumeikan University. He engages in the research on data engineering. He is member of IPSJ.

Fumiko Harada received B.E. and M.E, and Ph.D degrees from Osaka University in 2003, 2004, and 2007,respectively. She joined Ritsumeikan University as an assistant professor in Ritsumeikan University in 2007, and is currently a lecturer. She engages in the research on real-time systems and data engineering. She is a member of IEEE.



Prof. Hiromitsu Shimakawa received Ph.D degree from Kyoto Univ. in 1999. He joined Ritsumeikan Univ. in 2002. Currently, he is a professor in Ritsumeikan Univ. His research interests include data engineering, usability, and integration of psychology with IT. He is a member of IEEE and ACM