

# Serendipity-Oriented Recommender System Considering Product Awareness in Communities

Junya Yamazaki, and Shinsuke Nakajima

**Abstract**—Recently, recommender systems are expected to achieve not only high accuracy but also high serendipitous recommendation results. Thus, we are trying to develop a serendipity-oriented recommender system that can recommend products whose awareness are not so high against general users but quite high against maniac users in the community of a target topic. In addition, we perform an experimental evaluation comparing our proposed method to a conventional method.

**Index Terms**—recommender system, Serendipity-oriented, product awareness, community

## I. INTRODUCTION

IN recent years, the amount of information which people can obtain is getting bigger with development of the Internet. Therefore, a position of recommender systems is getting more significant. Because they can provide useful information from Big Data to users. By the way, the novelty and serendipity of recommended items may become worse if researchers try to improve only its accuracy when developing recommender system. Namely, it becomes easy to fall into the situation of already knowing or having them. It is necessary to develop a recommendation method whose novelty and serendipity of recommended items become better, in order to build a really useful recommender system.

As conventional approaches, Sawaizumi et al[1] have reported that verification of the effectiveness of serendipity-oriented recommendation. Hijikata et al[2] have proposed a novel method for estimating unknown items using similarities between users.

In this paper, We propose a serendipity-oriented recommender system that can consider product awareness in the community of a target topic, in order to recommend even if unknown items. Moreover, we perform experimental evaluations for our proposed method.

The remainder of this paper is organized as follows. The related work is given in Section II. Then Section III describes the method for Serendipity-oriented recommendation considering product awareness in communities. In Section IV, we show the experimental evaluation. We conclude the paper in Section V.

## II. RELATED WORK

As a research regarding high unexpected recommendation method in information recommendation, Hijikata [3] how evaluation from the historical development of the history of the recommendation of the study, describes a typical

Manuscript received Jan. 8, 2017; revised February 2, 2017. This work was supported in part by the MEXT Grant-in-Aid for Scientific Research(C) (#26330351).

J. Yamazaki and S. Nakajima are with Graduate School of Frontier Informatics, Kyoto Sangyo University, Motoyama, Kamigamo, Kita-ku, Kyoto, Kyoto, 603-8555, Japan. e-mail: ({i1558183, nakajima}@cse.kyoto-su.ac.jp).

problem. Murakami et al[4] is in the content the user prefer, feel the surprise to content that does not habitually access, based on the assumption that leads to the satisfaction of users, it has proposed a recommendation system. Akiyama et al[5] is of Serendipity and content to feel the useful information that the user is divided into two is not aware of the information that the information and the user to be aware of the user useful information to the user not aware of the Serendipity It has proposed a suggestion of a recommender system. Besides, Oku et al.[6] have proposed a system that produces a chance artificially by mixing the features of any of the two products. The user in this system can mix up to convincing the two products, has proposed a system that does not think that the time and effort the need to search.

## III. SERENDIPITY-ORIENTED RECOMMENDATION CONSIDERING PRODUCT AWARENESS IN COMMUNITIES

This section describes the Serendipity-oriented recommendation in consideration of the user's degree of preference in Community to item.

### A. Outline of the proposed method

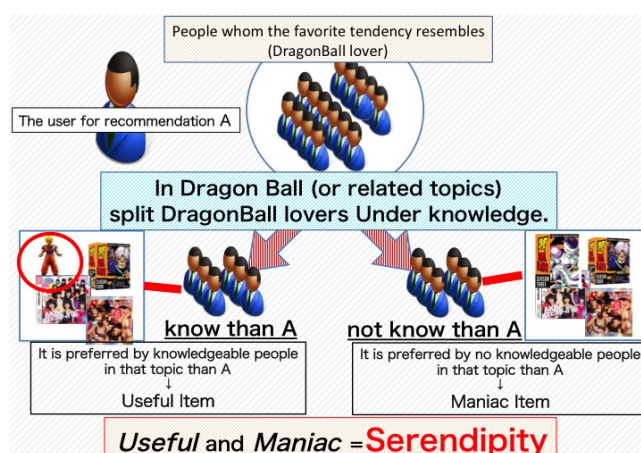


Fig. 1. Conceptual diagram of our proposed method

The items recommended using our proposed method should be unexpected and useful for users, and difficult to discover by themselves. Fig.1 shows a conceptual diagram of our proposed method. It describes how to judge whether a target user is interested in the Dragon Ball or not, and how to select recommendation items by the system. At first, our system try to identify a user group who are interested in same favorite topic with a target user A and also have a detailed knowledge of it. Then, the system recommend items that members of the user group browsed or purchased before, because the items are likely to be useful and valuable for the

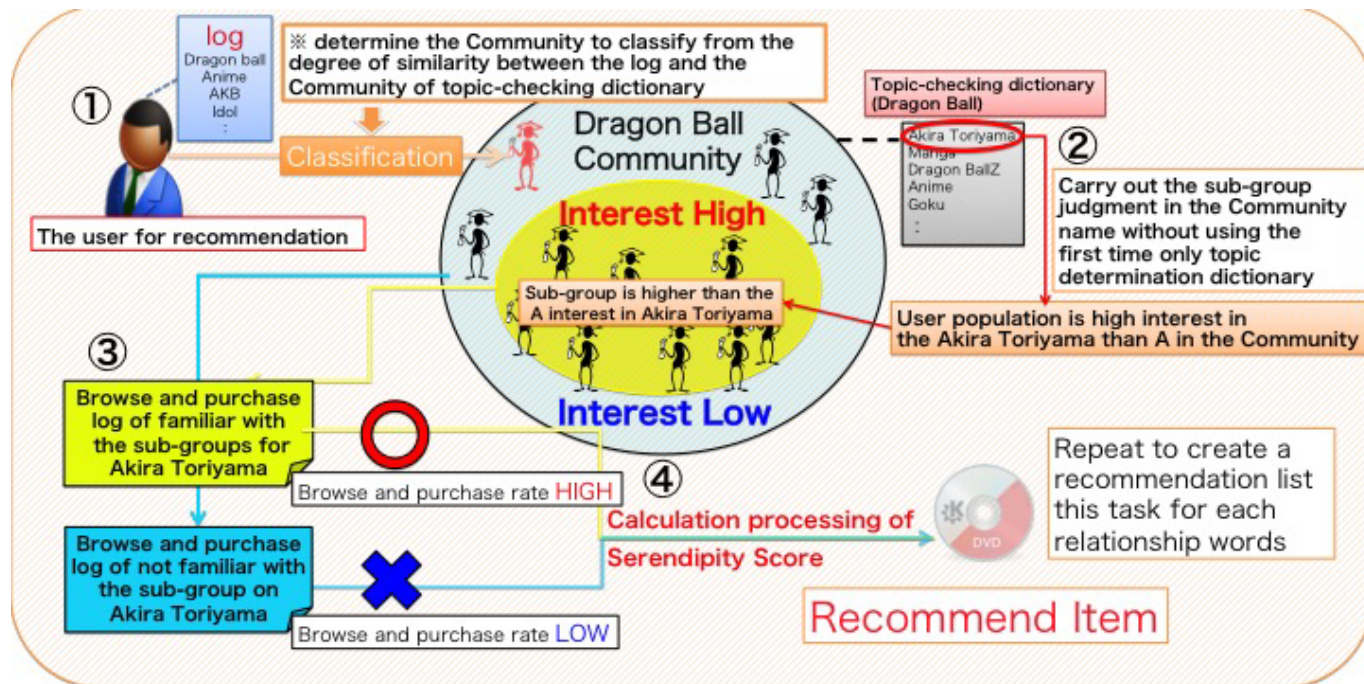


Fig. 2. Processing procedure of proposed method

target user. Moreover, it is possible that User A feels items, which unfamiliar user group with Dragon Ball browse or purchase, as too common. By using the browse and purchase log of the two user groups, user groups not familiar from the user A to the preference of the topic of topics to user A more detailed user group products and user A is viewing and purchase preferences of the user A view and the trade is surprising and useful for the user by not not recommend products to buy, can recommend a further item is not high self-discovery possibilities. Although it also has an element of contents base type recommendation by extracting to an interested topic by this method, it is distinguishable from the difference that a contents base is not what is recommended from the log of the user who is only alike.

In addition, Iaquinta et al[7] also proposed content-base recommender system considering Serendipity.

**B. Procedure of the proposed method**

Fig.2 is procedure and shows the example when the user A is presumed to be interested in a topic called DragonBall from an inspection and a purchase log as well as the outline like the point. This section explains the flow of the proposed method systematically. By this recommendation method, it roughly divides and three procedure occurs.

Community classification of the user for recommendation  
The subgroup judging in Community by a user's preference degree judging

Calculation and recommendation of Serendipity Score of each Community related product  
Each explanation is given below.

1) *Community classification of the user for recommendation:* Fig.3 explains the acquisition method of the candidate user's candidate user's log. First, the topic which expresses the feature of that product called a related topic altogether is attached to the item used by this research. About the

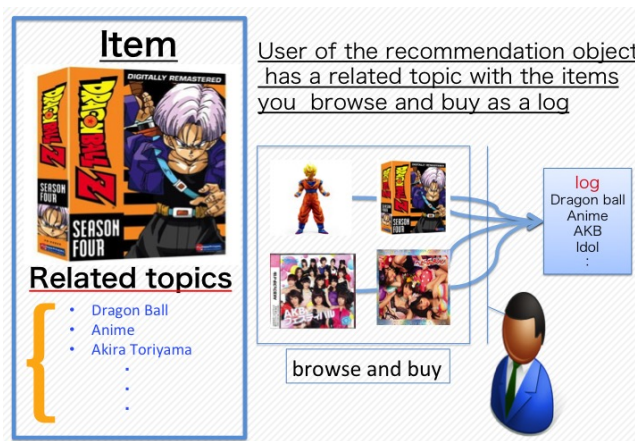


Fig. 3. How to get a target user's log through online shopping history

grant process of the related topic to item, it is automatically given by the item related topic automatic grant system as Fig. 4. A dictionary called the dictionary for a topic judging used here is a dictionary which there is a dictionary of the coincidence word to a certain topic obtained by a blog analysis called the brogram dictionary drawn up by research of the laboratory, and was processed except for words and phrases other than a noun etc. based on it. A user has as a log a set of the related topic given to the item which carried out inspection purchase, and recommends based on the log. This operation is automatically applied also to new item. Hence, reflection of a quick trend is attained. In Fig.5, a candidate user's Community classification is performed using the log created by Fig.3. The user for recommendation presumes the preference of the user for recommendation to an user group which is called Community and which is interested to a certain topic based on a log, and is classified into it. This Community has a dictionary for a topic judging to the

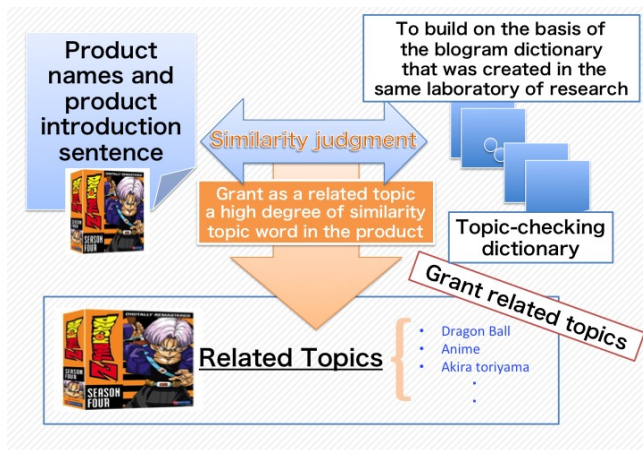


Fig. 4. Automatic finding method for topics related to target item

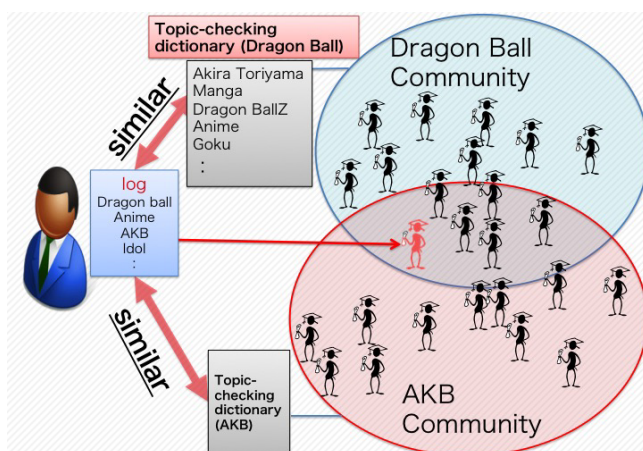


Fig. 5. Community classification method for a target user

topic word of Community. Whether he is interested in the Community sees and judges the similarity of the dictionary for a topic judging Community's, and a user's log created by Fig. 3. At this time, the user for recommendation is classified into two or more communities according to it being judged from a log that he is interested.

2) *The subgroup judging in Community by a user's preference degree judging:* The degree of preference performs the subgroup judging using the word of the dictionary for a topic judging explained as the point. About the word taken out first, the high user of the concern about a relational term is taken out from the user for recommendation as a detailed group to the word out of Community. And an inspection within each subgroup and a purchase log are created the group for a detailed subgroup and the other user group in Community to the word as a subgroup which is not detailed to the word.

3) *Calculating SerendipityScore for each Community related product:* Fig.6 explains calculation and recommendation of SerendipityScore of each Community related product. Serendipity Score is computed from an inspection and purchase log of a detailed subgroup, the subgroup which is not detailed, and these two subgroups to a certain related term word created, and it recommends from what has a high score. Serendipity Score is calculated by the formula in the Fig.6. Calculation processing of this Serendipity Score is the item

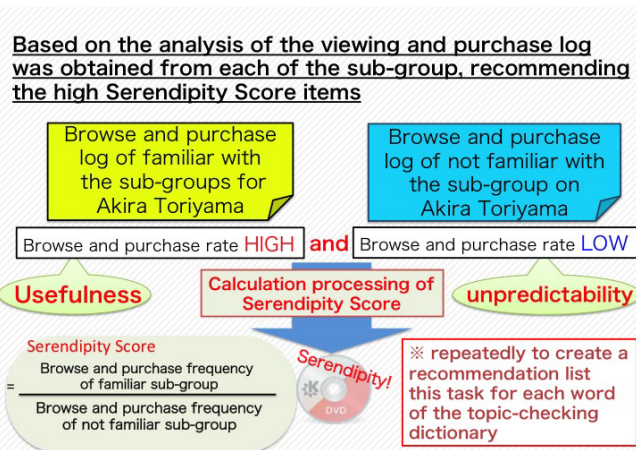


Fig. 6. Calculating of SerendipityScore for each Community-related product

in an inspection and purchase log of a detailed subgroup, and is processing in which the item which are not in an inspection and purchase log of the subgroup which is not detailed are determined as recommendation item.

The maniac item known by only the user detailed in the topic which the user for recommendation likes by this operation can be collected. Furthermore, since the possibility of the item which are related to the interested topic of the user for recommendation is high, this product is considered that usefulness is also high.

#### IV. EXPERIMENTAL EVALUATIONS

##### A. Comparative experiment for proposed method

Here, the Experimental evaluation of this proposed method is described. In this experiment, the reviewers 15 people Amazon wearing a review products related to Dragon Ball as a user with the same preference as the subjects, an experiment was conducted the commodity 145 to the reviewers who gave a review as a pseudo purchase log. The purpose of this experiment is to perform the check of the validity of the proposed method in an experiment.

Data for the experiment

- 1) log 15 servings of Dragon Ball favorite amazon reviewers
- 2) 145 reviewed item
- 3) 1162 dictionaries for a topic judging

It experimented to the one 20th generation man using the above data. As the procedure of an experiment Out of the item of 145 collected in advance, I have multiple selection of the item to purchase to a subject made, and it creates a false use log. Recommended the proposed method as from the review history of the Amazon reviewers of usage history and the 15 people who created. Item base collaborative filtering was used about the comparative approach. About evaluation, I recommended five item at a time from the proposed method and the conventional method, and had you reply to three questions which used five steps of scale methods about each item.

- 1) Were item known or not?
- 2) Did you regard item as wanting or not?

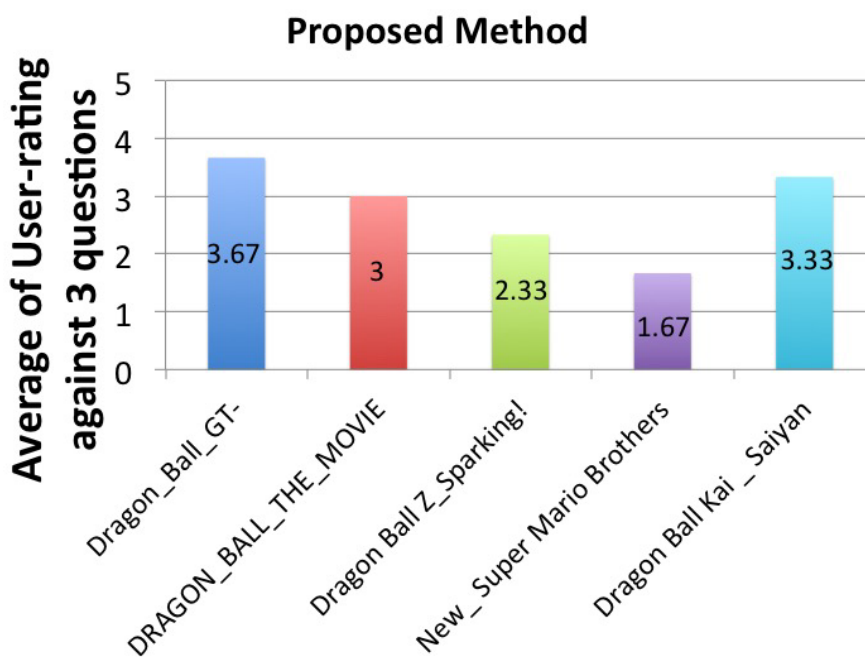


Fig. 7. verification experiment Result

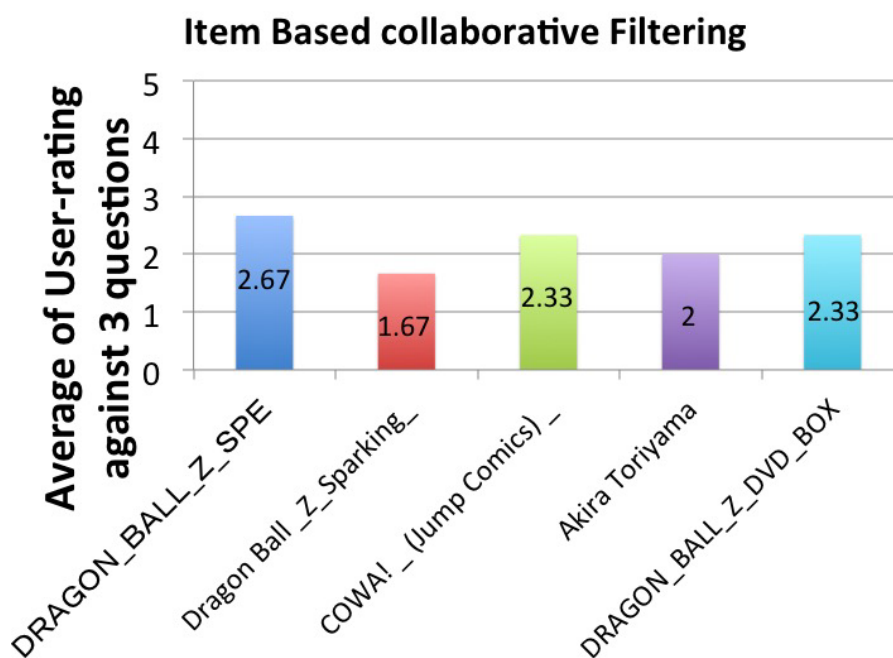


Fig. 8. verification experiment Result

3) Did you regard item as the ability to discover by oneself or not?

*B. experimental result*

Fig.7 corresponds to the proposed method result, and Fig.8 corresponds to the result of the comparative method, which is concerned with the evaluation each of the products in the 5-point scale. As a result, an average of 2.8 points in the proposed method, whereas the maximum was 3.67 points, an average of 2.2 points in the comparison method, the maximum is 2.67 points,Shimese effectiveness. When each evaluation of three questions was compared, the difference

had come out most by evaluation of (2). It seems that the proposed method attached evaluation of (2) to the item which were not known highly, and the difference came out of it from variation having appeared in name recognition while the item of the comparative approach had equal name recognition. As an unexpected result, the difference seldom arose with the question of (3). I think that a difference will come out notably when they both increase, since there are little number of subjects and number of times of trial under the present circumstances, and it is visible why it became like this. Furthermore, in the process to recommend, since there were many item in which the direction of the proposed

method serves as a candidate of recommendation when the same log as a comparative approach was given, it is surmised that the proposed method is a thing strong against a cold start.

We believe that our proposed method effectively performs against a cold start problem. In addition, we can improve its performance by adopting a method for presuming users' favorite categories, methods for detecting buzzwords[8], [9], [10]. They are going to work through them as our future works.

Moreover, it is necessary to also consider data and the acquisition method of using it for topic acquisition of Community, and to also take the difficulty of acquisition of a relational term into consideration simultaneously.

Word of the associated likely word and item from the co-occurrence word of the Community of topics in addition to, but simply co-occurrence word is expected to be a relationship between words with respect to relations language, there are many products, for example, co-occurrence is low prevalent topic expanding the range of products by taking into account such as, I believe it is possible to recommend a more Serendipity products.

In addition, as the future direction, performs a verification experiment in an environment closer to implementation, evaluation of the evaluation method of Murakami et al[11] with respect to evaluation of serendipity, with respect to the system Herlocker et al[12] how is the plan to make the evaluation as a reference.

It can be told including the data in Rakuten data Release [13] presentation that the restriction on use is also large to the open data which can be used for a verification experiment.

Therefore, selection and the usage of suitable data are due to inquire also including and the other data currently released, and to be examined.

## V. CONCLUSION

In this paper, We proposed a serendipity-oriented recommender system that can consider product awareness in the community of a target topic, in order to recommend even if unknown items. Based on our prototype system, we perform experimental evaluations. The experimental result indicated that our proposed method could provide better recommendation candidates than a conventional method.

## ACKNOWLEDGEMENT

This work was supported in part by Rakuten,Inc for providing the resources of Rakuten Data Release.

## REFERENCES

- [1] Shigekazu Sawaizumi, Osamu Katai, Hiroshi Kawakami, and Takayuki Shiose. Use of Serendipity Power for Discoveries and Inventions. *Intelligent and Evolutionary Systems, Studies in Computational Intelligence*, Vol. 187/2009, pp. 163-169, 2009.
- [2] Yoshinori Hijikata, Takuya Shimizu, and Shogo Nishida. Discovery-oriented collaborative filtering for improving user satisfaction. In *IUI2009:Proceedings of the 14th international conference on Intelligent user interfaces*, p. 67, New York, New York, USA, 2009. ACM Press.
- [3] Yoshinori Hijikata, methods of Preference Extraction for Information Recommendation, *IPSI Journal*, Vol.48, No.9, pp963-964, 2007.
- [4] Tomoko Murakami, Koichiro Mori, and Ryohei Orihara, A Method to Enhance Serendipity in Recommendation and Evaluation, *Journal of the Japanese Society for Artificial Intelligence* Vol.24 No.5, pp428-436, 2009.

- [5] Akiyama Takayuki, Obara Kiyohiro, and Tanizaki Masaaki. proposed and Evaluation of Serendipitous Recommendation by Introducing Metric Independent of User Profiles, *FIT2010*, RO-007, p157-164, 2010.
- [6] Oku Kenta and Hattori Fumio. User Evaluation of Fusion-based Recommender System for Serendipity-oriented Recommender System. *Proceedings of the Workshop on Recommendation Utility Evaluation: Beyond RMSE (RUE 2012)*, at the 6th ACM International Conference on Recommender Systems (RecSys 2012) (Web), pp.(6 pages), Dublin, Ireland, 2012.
- [7] Leo Iaquina, Marco De Gemmis, Pasquale Lops, Giovanni Semeraro, Michele Filannino, and Piero Molino. Introducing Serendipity in a Content-Based Recommender System. In *2008 Eighth International Conference on Hybrid Intelligent Systems*, pp. 168-173. Ieee, September 2008.
- [8] Shinsuke Nakajima, Jianwei Zhang, Yoichi Inagaki and Reyn Nakamoto. Early Detection of Buzzwords Based on Large-scale Time-Series Analysis of Blog Entries, *23rd ACM Conference on Hypertext and Social Media (ACM Hypertext 2012)*, pp.275-284, June 2012.
- [9] Jianwei Zhang, Seiya Tomonaga, Shinsuke Nakajima, Yoichi Inagaki, and Reyn Nakamoto, "Finding Prophets in the Blogosphere: Bloggers Who Predicted Buzzwords Before They Become Popular", *Proc. The 17th International Conference on Information Integration and Web-based Applications & Services (iiWAS 2015)*, pp. 100-109, Brussels, Belgium, December 2015.
- [10] Jianwei Zhang, Seiya Tomonaga, Shinsuke Nakajima, Yoichi Inagaki, and Reyn Nakamoto, "Prophetic Blogger Identification Based on Buzzword Prediction Ability", *International Journal of Web Information Systems*, Vol. 12, No. 3, pp.267-291, August 2016.
- [11] Tomoko Murakami, Koichiro Mori, and Ryohei Orihara. Metrics for evaluating the serendipity of recommendation lists. *New Frontiers In Artificial Intelligence, Lecture Notes in Computer Science*, Vol. 4914/2008, pp. 40-46, 2008.
- [12] Jonathan L. Herlocker, Joseph A. Konstan, Loren G. Terveen, and John T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, Vol. 22, No. 1, pp. 5-53, 2004.
- [13] Rakuten Data Release, <http://rit.rakuten.co.jp/rdr/>