Improvement of Attribute Correlation Method and Proposal of Collaborative Attribute Method in Text Recommender Systems for E-Learners

Yuji Wada*†, Takuya Segawa*, Jun Sawamoto**†† and Hiroyuki Sato**††

*Tokyo Denki University, Japan
**Iwate Prefectural University, Japan
†yujiwada@mail.dendai.ac.jp
††{sawamoto, sato_h}@iwate-pu.ac.jp

Abstract -E-learning is used in various places. However, many systems do not show advantages, such as online exams, and simply enumerate the teaching material, etc. In our An Individual Reviewing System (abbreviated AIRS), contents of each user are optimized according to recommendations using Collaborative Filtering (what we call CF). This system multiplies the load to the user by smoothly improving study efficiency. However, this CF method has disadvantages in that if insufficient data is available, recommendations may show poor accuracy. This is what we call Cold-Start problem. In this paper, to solve this Cold-Start problem, firstly we provided a solution of Attribute Correlation Method that uses metadata which are belonged to users. And, we experimented with this Attribute Correlation Method, but the good results were not obtained. Secondly, in order to improve this Attribute Correlation Method, we proposed a new approach (called Collaborative Attribute Method) is to address this Cold-Start problem and showed the experimental results.

Keywords: Recommender System, Web Digital Texts, E-Learning, Cold-Start Problem.

1 INTRODUCTION

E-learning, in which students can learn anywhere, at any time, has been coming into broader use in universities, corporate training and other settings. However, many existing systems simply make teaching materials available and conduct online testing, without providing the full range of unique learning advantages available through e-Learning.

One example of the existing systems is an individualized reviewing system (called AIRS). With AIRS, provision of content is tailored to the specific learner. This system uses an algorithm that helps students learn efficiently, based on the student's own historical data and the historical data of other learners, as described in [1].

The first other example is a bidirectional recommendation system. This system extracts the relationship among the learning web digital texts with historical logs and recommends an effective web digital text for learners, as discussed in [2].

The second other example is a recommendation system that recommends the optimum learning texts based on data mining of learning historical data. This system is called a collaborative learning recommendation system that mines the data of similar users sharing non-favorite subjects using historical logs and user attribute data, as discussed in [3].

However, among the existing systems mentioned above, there is a common disadvantage that the systems cannot handle recommendation before any historical data have been accumulated. This is a so-called Cold-Start problem.

To solve the Cold-Start problem, firstly we proposed Attribute Correlation Method using the background data of the user, and evaluated the usefulness of this approach, as mentioned in [4]. However, the results did not show this method to be particularly useful.

Secondly, in order to improve this Attribute Correlation Method, we proposed another solution to the Cold-Start problem in our research. So, we adopted the following approach;

· We proposed Attribute Correlation Method using the background data of the user.

· We tested subjects using Attribute Correlation Method, and evaluated the results.

· We examined whether this Attribute Correlation Method is effective or not.

· We proposed a new method (what we call Collaborative Attribute Method, described later), after considering improvements to this Attribute Correlation Method.

2 RELATED RESEARCH

Research in systems that anticipate user preferences and recommend contents is currently advancing, with a number of Web services using this approach. For instance, with the EC services used by Amazon [5], products are recommended that are likely to appeal to the user, based on the user's product page viewing history y, purchasing history and other data. Many of these systems use collaborative filtering (CF), as shown in [6]. In terms of education, however, research in the use of CF as opposed to education based on classroom lectures and other realistic environments is being conducted, as discussed in [7], but there are few cases in which this has actually been incorporated into e-Learning systems. With AIRS, learning content is recommended to the learner. With CF, however, a Cold-Start problem exists, in which the user has to use the

This research was supported by one Grant-in-Aid for Scientific Research C (Subject No. 21500908), and is currently supported by the other Grant-in-Aid for Scientific Research C (Subject No. 24500122).

contents to some extent, or no history can be obtained, and this makes it impossible to provide recommendations with a high level of accuracy, as described in [8]. This poses a drawback for users who want to use the system to solve questions in content learned through lectures and other means, or to review content already acquired. The research presented here proposes Attribute Correlation Method, which focuses on the Cold-Start problem.

3 COLLABORATIVE LEARNING RECOMMENDATIONS

Collaborative learning recommendations are recommendations carried out through the same procedure as CF. Hereafter the user will be referred to as the “learner”, and the historical data as "learning history ".The procedure for making collaborative learning recommendations comprises the following sequence of steps.

3.1 Extraction of Similar Learners

Other learners who have preferences similar to those of the learner for whom contents are to be recommended are extracted as “similar learners”. A database of the learning histories of learners is compiled, and correlations are drawn between learners based on that database, with learners being sorted in sequence based on the size of the correlation coefficient. Higher-order learners with a particularly large correlation are extracted as similar learners.

3.2 Extraction of Recommendation Contents

The actual content to be recommended is extracted from among the learners extracted as similar learners. The learning histories of similar learners are used to identify difficulties encountered by those persons, and analogies are drawn based on the way that those difficulties were overcome in order to extract relevant content.

3.3 Presentation of Recommendation Results

The extracted content is presented to the user via the system. This involves the system interface, and will not be addressed here.

4 ATTRIBUTE CORRELATION METHOD

As described above, collaborative learning recommendations are formulated by selecting recommended content based on the history of the learner. For this reason, similar learners cannot easily be extracted for learners who do not already have a learning history, or learners for whom a certain level of learning history has not been compiled (hereafter, we will call these “new learners”). As a result, it will not be possible to present highly accurate recommendation results. Given this, we propose a method of extraction in which background data for new learners is compiled and treated as attribute data, and learners with attribute data similar to that of the learner for whom recommendations are being provided are extracted as persons with similar attributes.

4.1 Overview

A primary reason for the Cold-Start problem that occurs in the collaborative learning recommendation method is that new learners do not have extensive histories, making it difficult to identify similar learners, as described in Section 3.2. In other words, this problem could possibly be solved if correlations between new learners and existing learners could be evaluated by other means. Figure 1 shows an overall flowchart incorporating the proposed method.

4.2 Attribute Data

Attribute data are acquired from meta-data, for example, age, sex, hobbies and preferences, strong subjects, weak subjects, and other personal data. This data is certain to be available for new learners, even if they do not have a learning history.

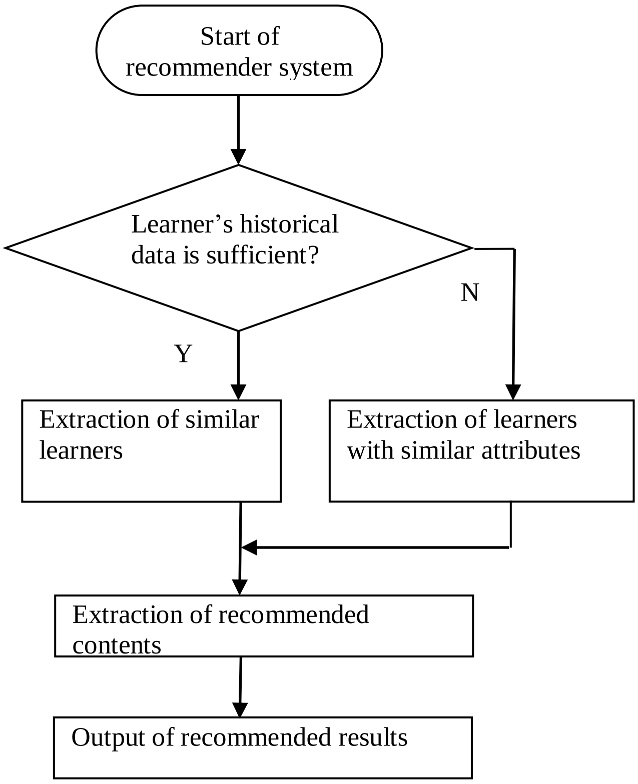

Figure 1: Attribute correlation method flow chart.

4.3 Systematization of Attribute Data

In attribute data, there is relevance among data items. For example, no relevance can be identified in a high school education between writing and physics, but a certain degree of relevance can be found between items that are both in a science curriculum, such as mathematics IA and physics. Systematizing attribute data within itself and expressing it is believed to be a necessary step, the reason being that one can envision that there will be little attribute data that can be compared to the learning history and used.

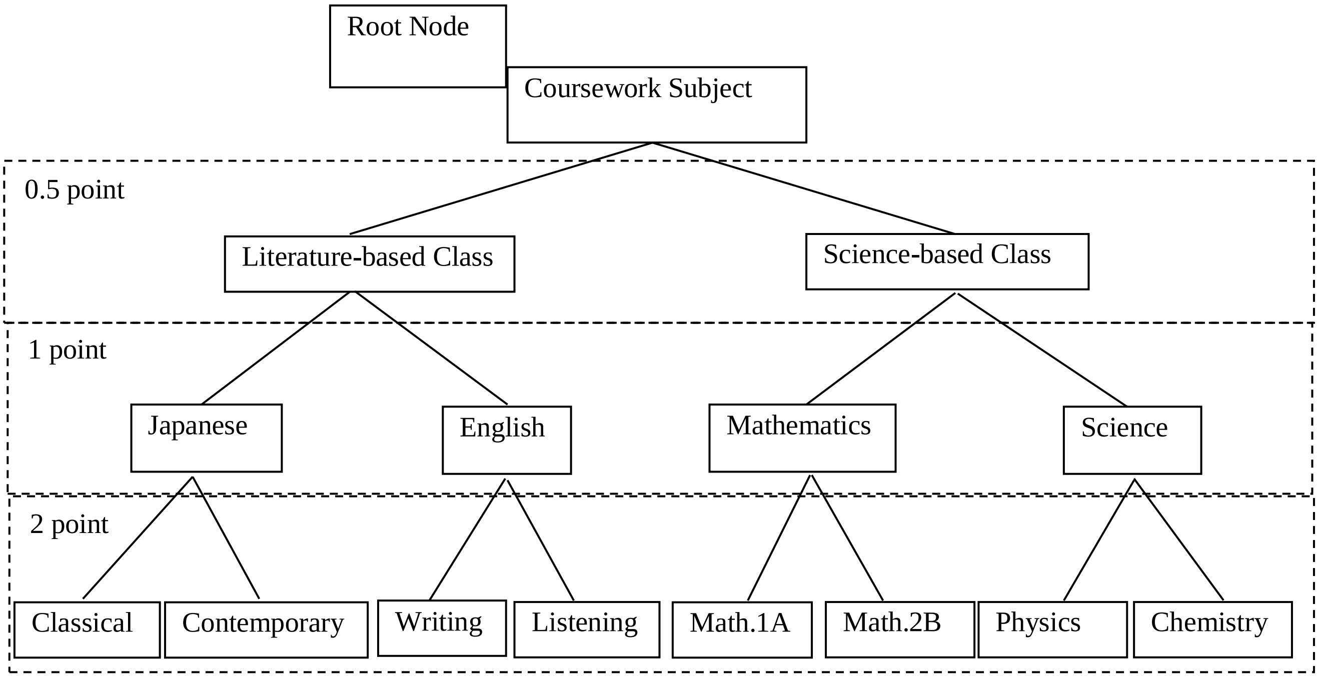

Figure 2: Hierarchy of attribute data.

With learning attributes, taking, for instance, a high school education as an example, coursework subjects are classified into root nodes, with science-based classes and literature-based classes as sub-nodes. These sub-nodes are further classified into generalized coursework classifications. Even more detailed names and definitions of classes are provided at the next layer, and a hierarchical structure is created. Moving further down the hierarchical layers, data become more specific, and thus carry greater weight as information. This weight can be expressed in terms of points: the first layer directly beneath the root node is counted as 0.5 points; and underlying layers are counted as 1, 2, and 4 points respectively, so that each layer has double the weight of the layer just above. This is done to increase the estimated value of the deeper layers. Figure 2 shows an example of the systematization of attribute data pertaining to learning. Here, only those types of attributes necessary for the evaluation, such as "learning" and "occupation", are created.

4.4 Extraction of the Degree of Attribute Data Similarity and Users with Similar Attributes

Conformances of attributes between new learners and all other learners are compared, and scores of all of the attributes are added together. A ranking is then created, with the highest scores at the top, and learners with particularly high conformance values are taken as learners with similar attributes. In the example shown in Table 1, Learner N is strong in the subject of physics, and thus has information in science and in science-category classes, which are upper-level nodes. Learner X matches completely, so 2 points are assigned, while Learner Y matches only in science-category subjects, and is thus assigned 0.5 points. Consequently, at this stage, Learner X will be a learner with similar attributes. The available attributes continue to be added up in this way. Ultimately, learners with the highest scores are extracted as learners with similar attributes.

4.5 Relationship between Similar Learners and Users with Similar Attributes

Table 1: Attribute table

	Science-category class	Natural Science	Physics	Chemistry	Math-ematics
N	0.5	1	2	0	0
X	0.5	1	2	0	0
Y	0.5	0	0	0	1

The flowchart in Figure 3 shows that when a sufficient learning history is available, Attribute Correlation Method is bypassed and recommendations are based on the normal algorithm for collaborative learning recommendations. This is because it can be surmised that Attribute Correlation Method will not produce better results by extracting similar learners based on learning history. This is because the recommended content itself is used as the history when extracting similar learners. In comparison, the background information of the learner, which has no direct relation, is used with attribute correlation. When these two approaches are compared, the learning history

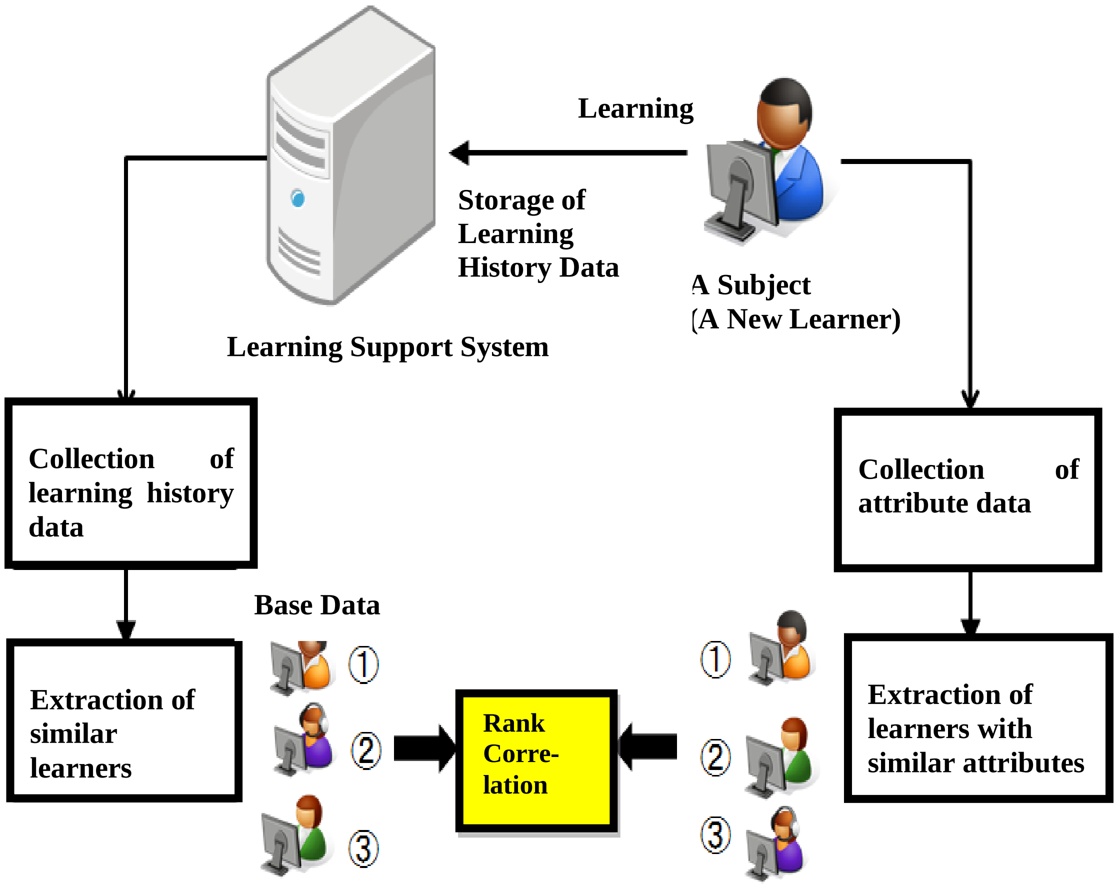

Figure 3： Outline of experiment.

clearly constitutes pure information in terms of the system. For example, in order to recommend books to a person who has not read any books to date, the thinking is incorporated that books will be recommended that may appeal to that person's preferences, based on elements such as other interests and skills. The primary aim of this method is to solve the Cold-Start problem.

4.6 Testing

Testing was conducted on subjects to clarify the outcomes of the proposed method. The following two items were evaluated.

· Is the proposed method effective?

· Was the hypothesis pertaining to attribute data selection proven?

4.6.1 Hypothesis Pertaining to the Selection of Attribute Data

As described in Section 3.1, attribute data serve as the meta-data for learners. However, not all of the personal data of learners is necessarily required. For example, if one were recommending exercises to help a person stay fit, physical information such as height and weight would be important, but this type of information is not necessary when recommending novels. In other words, it was theorized that attributes that are relevant to the content being recommended will probably demonstrate a high correlation. Here, because we are creating a recommendation system to be used in an education support system, information relating to learning will demonstrate a high correlation compared to attributes that are not particularly related to learning.

4.6.2 Test Method

***Advance preparation*:** To prepare for testing, courses from a high school curriculum were systematized as attributes related to learning, and hobbies were systematized as attributes other than learning-related attributes. The reason for choosing hobbies as attributes was that learners acquire and actively choose hobbies, as opposed to inherent information such as height, so these were assumed to closely reflect learner preferences. High school courses were selected as learning attributes in order to eliminate differences based on school year, since the students taking part in the testing were university students. As no models existed that were systematized with respect to hobbies, systematization was done based on speculation. For high school courses, however, we referred to the “Senior High School Education Guidelines” issued by the Ministry of Education, Culture, Sports, Science and Technology, as shown in [9]. Attribute hierarchies were each organized into three layers, with the objective of suppressing any bias created by differences in scores occurring as a result of changes in the weight of scores based on the depth of the hierarchy layer. Attribute data obtained as a result consisted of two attributes and three hierarchical layers.

***Subjects*:** Subjects were grouped into two groups comprising a total of 18 students, and a questionnaire was conducted prior to the testing. Participants answered the following two questions.

· What were your strong subjects when you were in high school? (Learning attributes)

· What are your current hobbies? (Hobby attributes)

Attributes of subjects were compiled based on the questionnaire. As a large number of attributes could be selected, the questionnaire was conducted in a self-reporting format, but in cases where the student did not respond correctly, that student was asked the question again by the tester, for the purpose of normalizing the attribute information. Subsequently, the following three items pertaining to the contest of the test were explained to the subjects, and testing was conducted.

· Learning time would be 15 minutes.
· Content would be in the form of a database.
· An achievement test would be performed after the study time had ended.

Moreover, the database comprising the content was something that could not be learned in its totality in 15 min, so subjects were asked to select portions that they did not understand, and to focus on those items when learning. This was done in order to avoid having subjects start at the beginning and study the contents in sequential order. The achievement test was also designed to increase the motivation of subjects to study efficiently in a short period of time, and would not affect the test evaluation itself.

***Test content*:** For this test, we used the text content to study the relational algebra operations of database technology with AIRS. The relational algebra operations cover the nine topics listed below.

· Selection Operation
· Projection Operation
· Summation Operation
· Intersection Operation
· Difference Operation
· Division Operation
· Cartesian Product Operation
· Join Operation
· Natural Join Operation

Figure 4 shows the examples of contents on Projection Operation and Join Operation.

***Analysis method*:** Figure 3 shows a schematic for the testing. The degree of similarity (similar learners) was calculated based on the learning history obtained from the 15-min period of learning, and the degree of similarity (learners with similar attributes) was calculated based on the compiled attribute data. The two were then compared and evaluated. Specifically, the same number of rank correlations was acquired as the number of subjects, and correlations were acquired in relation to the rankings of similar learners and learners with similar attributes obtained from each of the two similarity scales noted above. The Jaccard coefficient was used to calculate the degree of similarity based on learning history, as described in [10], and Kendall's rank coefficient correlation was used to calculate the rank correlation, as shown in [11]. Attribute Correlation Method is designed only to address new learners. The degree of similarity based on learning history shows a high degree of reliability, and so was used as the reference. In other words, the aim was to obtain the rank correlation between the ranking for the degree of similarity based on learning history (similar learners) and the degree of similarity calculated based on the proposed method (learners with similar attributes), so if the average of all subjects was high, reliability in terms of the extraction of similar learners would be seen as high, and the approach could be considered effective.

4.6.3 Test Results

Table 2 and Table 3 show test results for the two groups. The figures represent mean and standard deviation for the group as a whole, calculated based on the rank correlation between the ranks of learners with similar attributes and those of learners with similar learning histories. As the rank correlation is a correlation coefficient, values were taken from between -1 and 1. The closer the value is to 1, the stronger the correlation. The closer the value is to -1, the stronger the inverse correlation. The closer the value is to 0, the weaker the correlation. As can be seen from the two tables, the average was |0.1| or less for both, so no correlation was demonstrated, and no significant results were obtained. Moreover, with respect to learning attributes and hobby attributes, the only differences were due to error, so the hypothesis was negated. Except for one item, standard deviations were all≤0.2 as well, indicating that this conclusion is appropriate.

Table 2: Experimental results for group 1

	Learning Attribute	Hobby Attribute	Whole Attributes
Average	-0.077	-0.044	-0.095

Table 3: Experimental results for group 2

	Learning Attribute	Hobby Attribute	Whole Attributes
Average	0.1032	0.0238	-0.0397

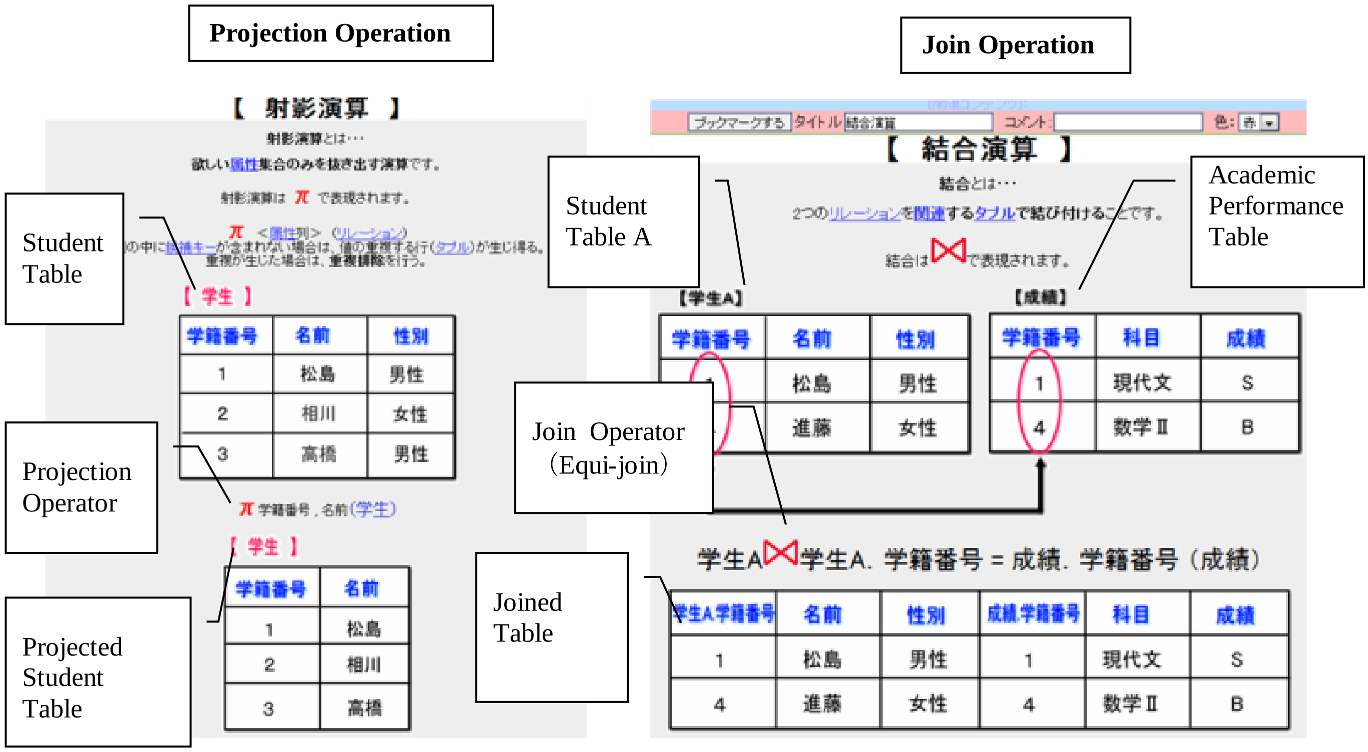

Figure 4: Examples of test content.

4.6.4 Discussion

Considering the causes of the results produced, the possibility arises that the amount of attribute data was insufficient. In that light, looking at the individual data for each subject, in the rankings based on attribute correlation, it was seen that rankings at the same ratio occurred for many subjects. Among these, there were a number of cases in which hobby attributes ended up being the same numeric values as those for other subjects as a whole, and no ranking correlations could be determined. However, despite the small volume of sample data, the fact that the average value for correlation coefficients was close to zero cannot be ignored. One other problem was that the relationship between the content being recommended and the attribute data was not clear. As indicated in Section 4.5, the reliability of attribute data is unclear, from an objective standpoint.

5 COLLABORATIVE ATTRIBUTE METHOD

In Collaborative Attribute Method of testing described in Section 4.6, usefulness could not be confirmed, for the reasons described in Section 4.6.4. Given that, we used the background information as attribute data. Collaborative Attribute Method is proposed here as a method for extracting new learners

.

5.1 Overview

Using the background information of the learner as attribute data is the same approach used in Attribute Correlation Method. This Attribute Correlation Method consisted of systematizing this data before use, but the data are not systematized in the method proposed here, but rather used as is. The degree of similarity between learners is surmised with reference to the degree of similarity between learners based on learning history, and to the attribute data.

5.2 Degree of Similarity between Attribute Data

The degree of similarity between attribute data was calculated in advance. Here, we take Learner I and Learner J, for whom a certain amount of learning history has been compiled. Attribute data for these two users were acquired when they were new learners, so we already have degrees of similarity in learning histories and respective attribute data at this stage. Amounts of attribute data are not determined in particular, but let us assume in this example that we have two attribute data: A and B. Taking the degree of similarity in learning histories between these two persons as X, we can say that the combination of attributes for these two persons, for some reason, has similarity X. If this combination is also seen among other learners, we take the average. These degrees of similarity are then accumulated in a database. Figure 5 shows a schematic diagram of this.

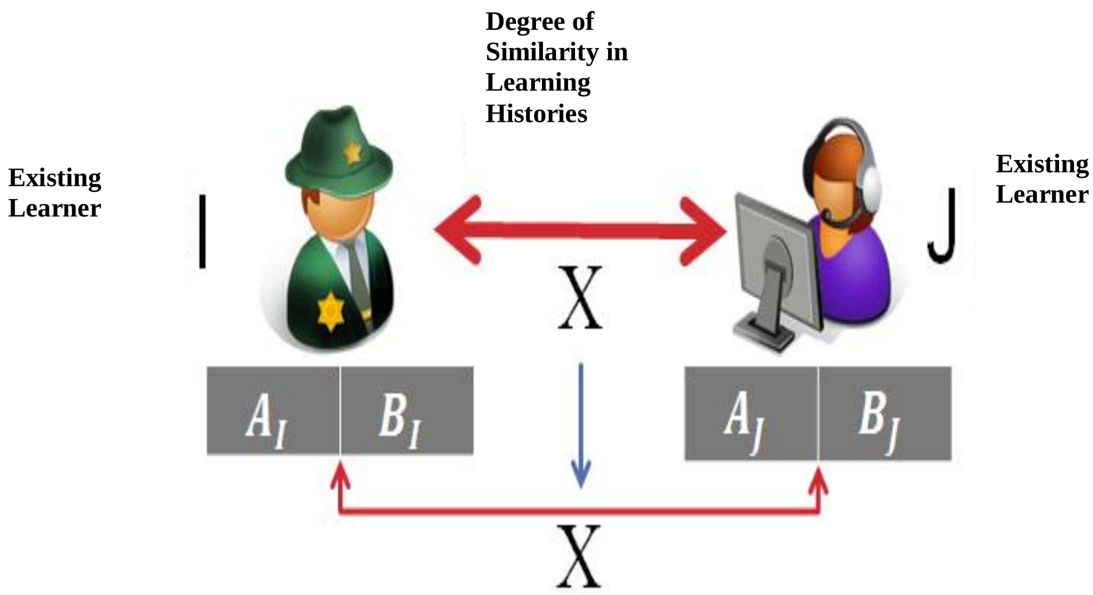

Figure 5: Calculating the similarities between attribute data.

5.3 Deriving the Degree of Similarity

When actually making recommendations for new learners for whom no degree of learning history similarity exists, we refer to similarities between attribute data that have been accumulated, and extract learners having combinations with the highest degrees of similarity between attribute data as learners with similar attributes, as shown in Figure 6. Content is then recommended based on these users.

5.4 Differences between This Method and the Attribute Correlation Method

In Attribute Correlation Method, attributes are systematized and the number of points is totaled. In Collaborative Attribute Method, however, similarities between attributes are measured using similarities in learning histories, which are reliable, as a resource. As a result, the data can be expected to be more reliable. Conversely, because the approach taken is similar to that in CF, recommendations will similarly be less accurate if only small amounts of data have been accumulated.

5.5 If the Amount of Attribute Data Accumulated Is Insufficient

As indicated in Section 5.4, this method also involves accumulated attribute data, and there are concerns that the extraction of persons with similar attributes will be less accurate if insufficient information is available. If the amount of attribute data accumulated by forming combinations of attributes of a learner for whom recommendations are being made is smaller than a stipulated amount, attributes A and B are split and calculated, as shown in Figure 7.

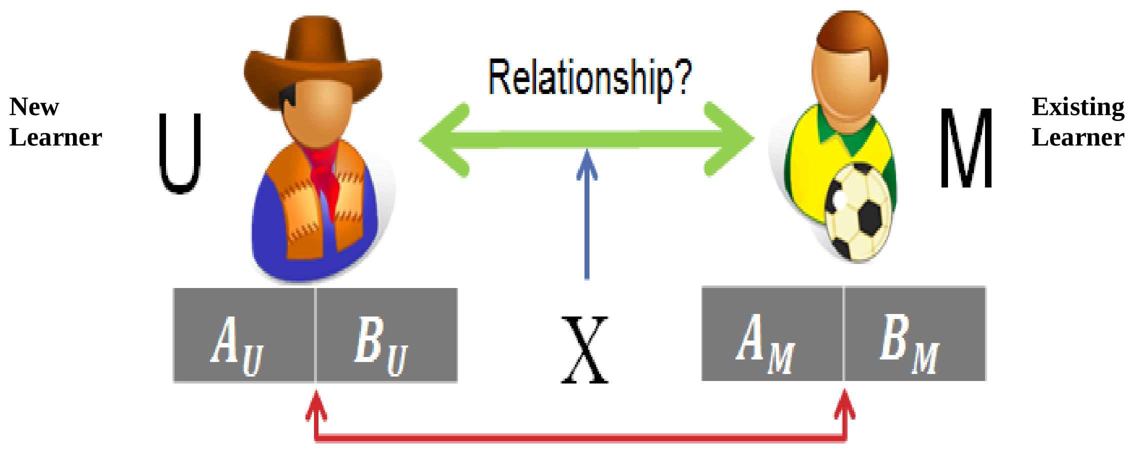

Figure 6: Extracting the attribute analogy.

Now, assume that we want to find the similarity of A_X and A1. We load combinations that include and from a table of attribute data similarities that have been accumulated, and we take the similarity of each of these and divide the number of points by the ratio of the number of elements. For example, if the ratio of the number of elements of A and the number of elements of B is 1:2, and the similarity between $A_XB_{a'}$ and A_1B_n is 0.6, this result of 0.6 would be divided by 1/3 to obtain a result of 0.2. This would be carried out for the number of combinations A_X of and A1, and the average of all values would be taken. This would be done as many times as there are combinations of the attributes of A and B, and recommendation content would be extracted from users having the combinations with the highest values.

5.6 Experimental Results

We experimented with Collaborative Attribute Method using the same data as those in Section 4.6. Table 4 shows the results of rank correlations. We can see that this Collaborative Attribute Method provides better results than Attribute Correlation Method in Table 4, but they are not so high values. Some values of the rank correlations which are above 0.4 exist among the results before averaging. So, we can expect the averaged rank correlation will be higher if we can collect more data.

Table 4: Experimental results of rank correlation with Collaborative Attribute Method

	Rank Correlation
Average	0.188
Average	0.237

6 CONCLUSION

In the testing described in Section 4.6, the usefulness of Attribute Correlation Method was not able to be proved. This was attributed to the fact that the relationship between attribute data and learning history is not understood, and Collaborative Attribute Method is proposed in which similarities in learning history are referenced and attributes are used. At the same time, however, this method has not yet been perfected and still has scope for improvement. In addition, it may simply be that not enough testing has been conducted on Attribute Correlation Method. In the future, we intend to continue conducting testing on Attribute Correlation Method, and to develop, implement and test Collaborative Attribute Method.

7 ACKNOWLEDGMENT

We appreciate the technical support on AIRS to Dr. Shinichi Dohi who is a Professor in Tokyo Denki University and also thank to all subjects who participated in our experiments.

This research was supported by one Grant-in-Aid for Scientific Research C (Subject No. 21500908), and is currently supported by the other Grant-in-Aid for Scientific Research C (Subject No. 24500122).

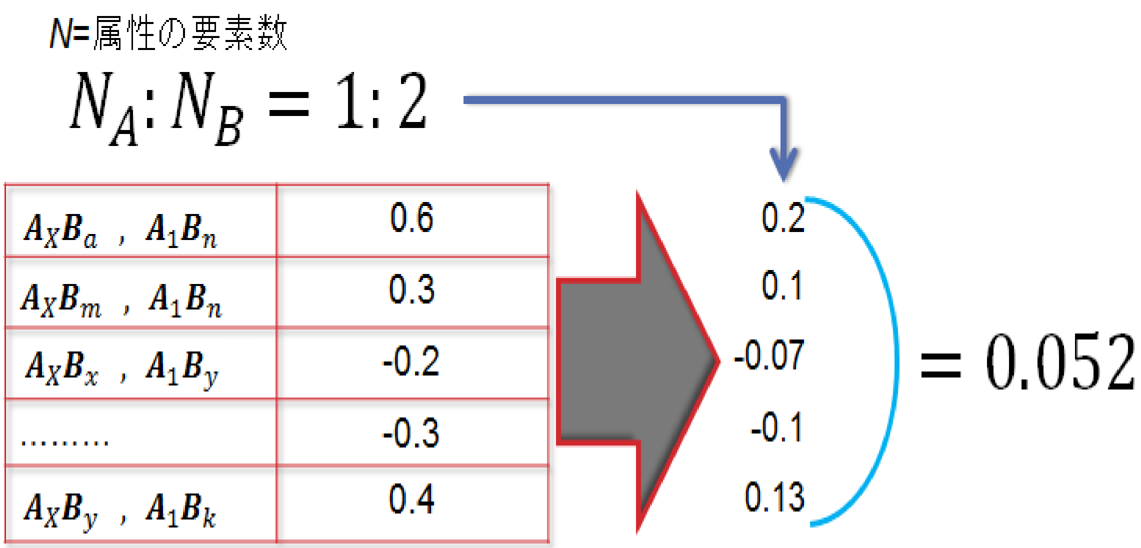

Figure 7: Algorithm for split attributes.

REFERENCES

[1] M. Yamaguchi, Y. Takahashi, T. Matsuzawa, T. Takahashi, and S. Dohi, “Development and Evaluation of e-Learning System AIRS with Multi-Expression Contents,” IEICE-ET2006-118, Vol.106, No.583, pp.69-74 (2007).

[2] Y. Wada, S. Matsuzawa, M. Yamaguchi, and S. Dohi, “Bidirectional Recommendation Technology for Web Digital Texts, Journal of Digital Information Management,” vol. 8, no. 4, pp. 240–246 (2010).

[3] Y, Wada, Y. Hamadume, S. Dohi, and J. Sawamoto, “Technology for Recommending Optimum Learning Texts Based on Data Mining of Learning Historical Data,” International Journal of Information Society IJIS., vol. 2, no.3, pp.78–87 (2011).

[4] Y. Wada, T. Segawa, J. Sawamoto, and H. Sato, “RESEARCH ON RECOMMENDER FUNCTIONS FOR LEARNING SUPPORT SYSTEMS,” 2nd International Conference on Applied and Theoretical Information Systems Research (2012).

[5] Amazon: http://www.amazon.com

[6] P. Resnick, N. Iacovou, M. Suchak, P. Bergstorm, and J. Riedl, “GroupLens: An Open Architecture for Collaborative Filtering of Netnews,” Proc. ACM Conf. on computer Suspend Coparative Work, pp.175-186 (1994).

[7] S. Kitamura, “Consistency between Theoretical Interests in Collaborative Learning Studies and Methods of Statistical Analysis: A Review of Statistics for Hierarchical Data,” Japan Society for Educational Technology (ISSN 1349-8290), vol.33 , no.3 , pp.342-352 (2010).

[8] A. Schein, A. Popescul, L. Unger, and D. Pennock, “Methods and Metrics for Cold-Start Recommendations,” 25th Annual ACM SIGIR Conf., pp253-260 (2002).

[9]http://www.mext.go.jp/b_menu/shuppan/sonota/990301d.htm

[10]T. Segaran, “Programming Collective Intelligence Building Smart Web 2.0 Applications,” O’REILLY (ISBN 978-4-87311-364-7), (2008).

[11]H. Iwasa and H. Yadohisa, ISBN 978-4-7980-2396-0, pp.352 (2009-10).

(Received November 29, 2013)

(Revised November 10, 2014)

Yuji Wada received the B.E. and the M.E. in electrical engineering from Waseda University in 1974 and 1976, respectively. He joined Mitsubishi Electric Corporation in 1976. He received the PhD degree in computer science from Shizuoka University of Japan in 1997. He is currently a Professor in the Department of Information Environment, Tokyo Denki University. His research interests include database systems, data mining, and recommendation. He is a member of the IPSJ, the IEICE, the JSAI, the JSSST and the DBSJ.

Takuya Segawa received the B.E. and the M.E. in information environment engineering from Tokyo Denki University of Japan in 2010 and 2012, respectively. He is currently working on a Recommend Software Development at DWANGO Co., Ltd. His research interests include dispersion treatment, recommend, and clustering.

Jun Sawamoto is currently a Professor of Faculty of Software and Information Science, Iwate Prefectural University, Japan. He received the B.E. and M.E. in mechanical engineering from Kyoto University in 1973 and 1975. He joined Mitsubishi Electric Corporation in 1975. He received his PhD degree from Tokyo Denki University in 2004. His research interests include ubiquitous computing, human-interface system, multi-agent systems, and cooperative problem solving. He is a member of IPSJ, IEEE-CS, ACM.

Hiroyuki Sato is currently a Professor of Faculty of Software and Information Science, Iwate Prefectural University, Japan. He received the B.E. in information engineering from Tsukuba University in 1982. He joined Mitsubishi Electric Corporation in 1982. He received his PhD degree from Tsukuba University in 2003. His research interests include parallel processing, and high performance computing. He is a member of IPSJ, IEICE and IEEE-CS.