Collabo-eNOTE: A Hybrid Recommender System for Group Learning Support

Xin Wan, Fumihiko Anma, Toshie Ninomiya, Toshio Okamoto

Graduate School of Information System, The University of Electro-Communications
1-5-1 Chofugaoka, Chofu-shi 182-8585, Tokyo, Japan

Summary
We report on the development of a WEB based intelligent e-NOTEBOOK system that provides an environment for group learning. That is, the system may recommend the useful notes automatically based on the item-to-item collaborative filtering and content-based filtering by using learners' reading histories and the contents of notes, which should improve the efficiency of the group learning. The performance of the proposed approach is compared with traditional collaborative filtering. Experimental results demonstrate that our approach is more effective and efficient to group learning.
Key words:
Collaborative Filtering, Content-based filtering, Hybrid filtering, Intelligent System, Collaborative learning.

1. Introduction

Collaborative learning is a learning method that contrasts with the traditional 'direct transmission' model in which learners are passive, isolated receivers of knowledge delivered by instructors[8]. Group learning is an important component of traditional teaching in all levels of the educational system and it is one type of collaborative learning. It is motivating, the learners become active and involved, and they enjoy meeting with other learners and doing work together. In this study, we have developed the e-NOTEBOOK system for group learning support which is a WEB based, multimedia database that is structured to support the learners through the inquiry processes and provides them with a mechanism for working cooperatively with others. Group learners may express their ideas, record their actions, and communicate with others by using this system. To help learners to find out each other's relevant contributions is a kind of advantageous way to make group learning work effectively when the number of learners increases. Therefore, making use of the recommender systems in the e-NOTEBOOK system can help the learners seek out their suitable notes, which can be taken as efficient reference.

We developed the system that can recommend the useful notes by collecting reading histories and the contents of learners' notes. In the mean time, the useful notes and learners with similar interest will be recommended automatically based on collaborative filtering and content-based filtering.

2. Related Work

Generally, recommender system can be divided into three categories[10]: Content-based filtering (CBF), Collaborative filtering (CF), and Hybrid approaches. Content-based recommendation systems use data about the requested items and the information regarding only the active user [11]. NewsWeeder [12] applied this method to build up a Netnews filtering system. Other applications are WebWatcher [13], InfoFinder[14], Mooney[15] and so on. Collaborative filtering was proposed to automate the process of "word-of-mouth" [16] by leveraging like-minded users' opinions. It is an information filtering technique that depends on human beings' evaluations of items. Many approaches and systems, such as Amazon.com [3], GroupLens [17] etc., adopt this technique. Hybrid filtering combines collaborative and content-based approach. Content-based and collaborative methods can be combined into the hybrid approach in several different ways [10]. Fab is an implementation of a hybrid content-based, collaborative Web-page recommendation system that eliminates many of the handicaps of the pure versions of either approach [18].

Recommender system in education needs to improve the "educational provision" [19]. The approach in educational recommender system is hybrid and it is obtained by two main mechanisms: the learners' learning processes and analysis of social interaction. A personal recommender system for e-learning therefore would have to search for potential learning activities and recommend the most suitable learning activities to the individual learner or learner group.

Manuscript received September 5, 2008.
Manuscript revised September 20, 2008.

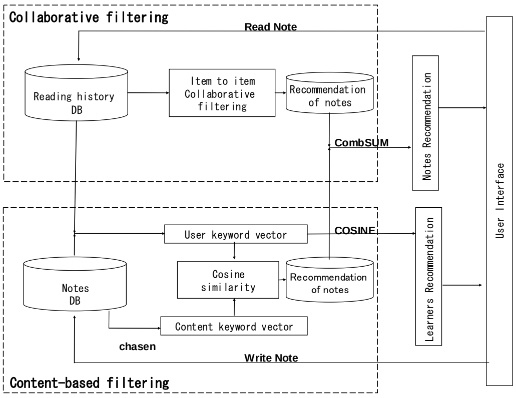


Fig.1 The brief architecture of system

As stated above, we focus on the hybrid filtering strategy in this paper. And we study and analysis the learners' recorded behaviors to predict the learners' preferences.

3. Collabo-eNOTE

The system collects reading histories and contents of every learner's notes by a database. It predicts recommendation notes based on content-based filtering and collaborative filtering. At the same time useful notes and the similar learners will be recommended (Fig.1). The system generates html pages such as the list of note recommendation or the list of learner recommendation dynamically. A learner can read a generated page freely.

3.1 Acquisition of learner profile

Generally, a collaborative filtering algorithm uses a collection of user profiles to identify interesting "information" for users. A particular user gets a recommendation based on the user profiles of other similar users. User profiles are commonly obtained by explicitly asking users to rate the items.

Previous research has shown that users are very unlikely to provide an explicit rating [7]. Alternatively, user profiles can also be obtained by implicitly observing user interactions with the system. And reading time spent on WEB page is an effective element to measure a user's interest [1].

All in all, this system measures acquisition of learner (user) profile by measuring time a learner spends on reading a note. In addition, we use the mechanism that removes the invalid reading time to improve the accuracy of the predication. And we predict recommendation note by using learners' reading histories with memory-based method.

The process is as follows:

Step1 While a learner reads the note the system records ID of the note which is read. The time of starting is recorded as a history temporarily.

Step2 When the learner finished reading the note, the system records the time of finishing and calculates length of reading time. On the same time, speed of reading is calculated. The reading speed is according to gamma distributions [9]. If confidence coefficient is under 95%, the reading time is deleted.

Step3 The valid reading time as the learner's preference is saved in learner profile.

3.2 Notes recommendation

Collaborative filtering can even recommend under the situation with a few data. However, it has often been reported as a cold-start problem [2]. Cold-start problem is related to the situation when a user enters the system and others express no ratings, which causes collaborative filtering not to be able to compute a recommendation. In contrast, it is possible to recommend if we understand the contents of a recommendation object by content-based filtering. Accordingly, it is necessary to adopt a new filtering algorithm that can put two kinds of filtering together well. We use collaborative filtering and content-based filtering together to recommend notes.

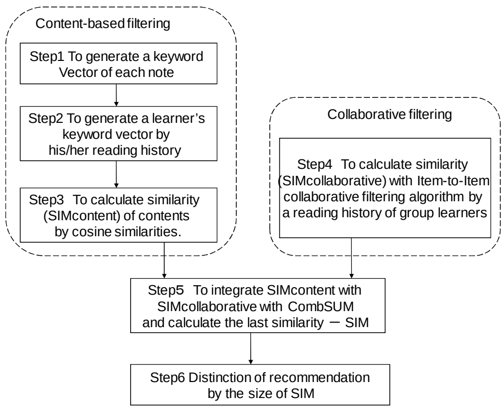


Fig.2 The process of notes recommendation

The process of notes recommendation is as follows (Fig.2).

Step1 To generate keyword vectors of each note.

The contents of every note is analyzed by a Japanese morphological analysis system automatically—chasen [4]. According to it, keywords are noun and unfamiliar words which are extracted from a note. The length of the keywords is more than 2 letters. Keywords list includes keywords and the word appearance frequency. When a learner writes the note, a keyword list is generated. The keyword vector of each note is expressed as equation (eq.1).

$$N = (w_{bk1}, w_{bk2}, \ldots, w_{bkn}) \qquad (1)$$

Step2 To generate a learner's keyword vector by his/her reading history.
Using the keyword vector which generated in Step1 and the reading history of each notebook, we can generate the keyword vector of the learner (eq.2).

$$U = (w_{uk1}, w_{uk2}, \ldots, w_{ukn}) \qquad (2)$$

Step3 To calculate similarity (SIMcontent) of contents by COSINE method.
Using equation (eq.1) and (eq.2) we can calculate similarities by using formula (eq.3).

$$SIMcontent(U, N_j) = \frac{\sum_{i=1}^{n} w_{uki} \times w_{bjki}}{\sqrt{\sum_{i=1}^{n} w_{uki}^2 \times \sum_{i=1}^{n} w_{bjki}^2}} \qquad (3)$$

Step4 To calculate similarity (SIMcollaborative) with Item-to-Item collaborative filtering algorithm by reading histories of group learners.
We take reading time as a pattern to predict preference of a learner. The system collects the learners' reading time by network automatically. With Weighted Slope One algorithm [6] of Item-to-Item collaborative filtering [3], calculate the ranks of recommendation.
Step5 To integrate SIMcontent with SIMcollaborative with CombSUM to calculate the last similarity (SIM).
The report [5] shows CombSUM method is effective in the text document search. First we use SUM method to normalize similarities by using equation (eq.4).

$$S_i^{'}(B_j) = \frac{S_i(B_j)}{\sum_{j=1}^{N} S_i(B_j)} \qquad (4)$$

Second we integrate SIMcontent with SIMcollaborative by using CombSUM method as equation (eq.5).

$$S^{''}(B_j) = \sum_{i=1}^{N} S_i^{'}(B_j) \qquad (5)$$

Step6 Distinction of recommendation by the size of SIM(Fig.3)

3.3 Learners recommendation

The system use reading histories to recommend learners who have the same interest and reading tendency (Fig.4). It can lead a recommended learner to a chat system by which learners can argue with each other.
The process is as follows:
Step1 To generate keyword vector of each learner (eq.6).

$$U = (w_{uk1}, w_{uk2}, \ldots, w_{ukn}) \qquad (6)$$

Step2 To calculate similarity of learners by COSINE method (eq.7).

$$sim(u_a, u_b) = \frac{\sum_{i=1}^{n} w_{uaki} \times w_{ubki}}{\sqrt{\sum_{i=1}^{n} w_{uaki}^2 \times \sum_{i=1}^{n} w_{ubki}^2}} \qquad (7)$$

Step3 Distinction of recommendation by the size of similarity.

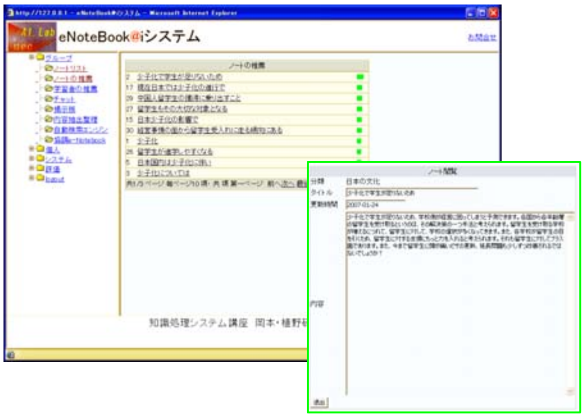


Fig.3 GUI of notes recommendation

Fig.4 GUI of learners recommendation

4. Experimental Evaluation

The theme, "The influence of Japanese declining birthrate on oversea students", was assumed and a total of 30 notes from 10 oversea students were collected. The note recommendation by two systems was done separately to the participants and they evaluated it. One is a traditional e-NOTEBOOK system developed only by the collaborative filtering and uses an explicit rating through peer assessment, and the other, a Collabo-eNOTE system by using hybrid filtering in this paper (Tab.1).

Tab.1 Comparison of the two systems

	The proposed system	Comparison system
Main technology	Morphological analysis Collaborative filtering Content-based filtering	Collaborative filtering
Algorithm of recommend	Collaborative filtering Content-based filtering	Collaborative filtering
Integration	SUM, CombSUM	
Acquisition of user profile	Implicit(reading history)	Explicit(peer assessment)
compute	online&offline	offline
Cold-start problem	no	yes

4.1 Evaluation of note recommendation

A number of metrics have been proposed in literature for evaluating the robustness of a recommender system. Breese et al. suggested a metric based on the expected utility of the recommendation list [20]. The utility of each item is calculated by the difference in vote for the item and a "neutral" weight. The metric is then calculated as the weighted sum of the utility of each item in the list where the weight signifies the probability that an item in the ranking list will be viewed. This probability was based on an exponential decay. We use this method as equation (eq.8) and compare the learning object recommendation result and evaluate this system.

$$R = 100\frac{\sum_a R_a}{\sum_a R_a^{\max}} \qquad (8)$$

$R_a^{\max}$ is the maximum achievable utility if all observed items had been shown at the top of the ranking list. In addition, R_a is an equation of (eq.9).

$$R_a = \sum_j \frac{\max(r_{a,j} - d, 0)}{2^{(j-1)/(\alpha-1)}} \qquad (9)$$

In the above equation, $r_{a,j}$ is user a's predicted vote to item j. d is the neutral vote and α is the viewing halflife. The halflife is the number of the item on the list, there is a 50-50 chance for the user to review that item. In this experiment, we used a halflife of 5 items.

Fig.5 is the conformity rate result. The conformity rate of note recommendation by 10 participants is higher than 60%. By using Breese's evaluation method, a result of 80.42 was obtained by the approach proposed by us, and a result of 79.43 by the traditional collaborative filtering approach. This shows obviously that the result obtained by our approach is higher. The reason why an excellent user evaluation was able to be obtained is that the content of notes becomes one of the patterns of recommendation. In addition, in comparison with the system using collaborative filtering only, it solved the cold start problem.

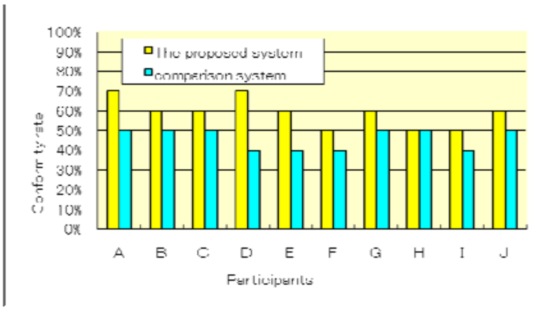


Fig.5 Result of notes recommendation conformity rate

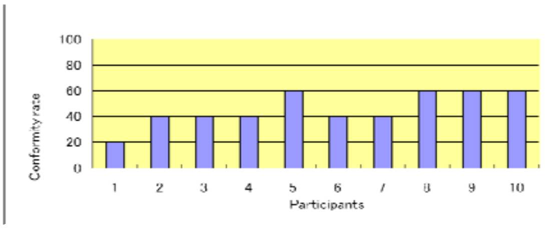


Fig.6 Result of learners recommendation conformity rate

4.2 Evaluation of learner recommendation

In order to evaluate the correctness of the recommendation learners, we use the equation (eq.10) to calculate the conformity rates of learners' recommendation.

$$conformity rate = \frac{total\ number\ of\ correct\ learners\ evaluated\ by\ learner}{total\ number\ of\ learners\ recommendation} \qquad (10)$$

The result from the 10 participants is as follows (Fig.6). Almost the group learners are satisfied with our recommendation. The average of the correctness of the learners' recommendation is 0.46.

We also evaluated the satisfaction of the recommendation of learners who are generated by our approach throughout questionnaire using five-stage evaluation (the most satisfaction is 5. the most dissatisfaction is 1). The result of the questionnaire is 4.33. That is to say, our learner recommendation is effective.

5. Conclusion and Future Work

With a view of helping to achieve efficient group learning for each learner, we have developed this Collabo-eNOTE system that can recommend useful notes from among a large amount of the notes.

The main contributions can be summarized as follows:

(1) We proposed a new hybrid filtering that combine collaborative filtering and content-based filtering.

(2) A new mechanism of learner profile's acquisition is proposed based on learner's reading time.

(3) The system makes it easier for group learners to study and communicate.

In future, we will work on adding patterns of measuring learners' preferences to improve accuracy of recommendation.

References

[1] Morita, M. and Shinoda,Y.(1994). Information Filtering Based on User Behavior Analysis and Best Match Text Retrieval, Proc. of the 17th Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval, pp.272-281.

[2] A.Schein, A.Popescul, L.Ungar, D.Pennock(2002). Methods and metrics for cold-start recommendations, 25th Annual ACM SIGIR Conference, pp. 253-260.

[3] G.Linden, B.Smith, J.York(2003). Amazon.com Recommendations: Item-to-Item Collaborative Filtering, Proc. 7, IEEE Internet Computing, Vol. No. 1s.

[4] "chasen" , http://chasen.aist-nara.ac.jp

[5] Montague,M.and Aslam, J.(2001). Relevance score normalization for metasearch, Proc. Of the ACM Tenth International Conference on Information and Knowledge Management(CIKM).

[6] Lemire Daniel, Maclachlan Anna(2005). Slope One Predictors for Online Rating-Based Collaborative Filtering, In Proc. of SIAM Data Mining.

[7] M.Claypool,M.W.P.Le,and D.C.Brown(2001). Implicit interest indicators, Proc. Of IUI'01, pp.31-36.

[8] De Corte, E.(1996). Instructional psychology: overview. In De Corte, E. & Weinert, F.E. (Eds.) Encyclopedia of Developmental and Instructional Psychology. London, Pergamon.

[9] Jansen, Margo G. H.(1997). Rasch's model for reading speed with manifest explanatory variables". Journal Psychometrika, pp.393-409.

[10] Adomavicius, G. and Tuzhilin, A.(2005), Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions, IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 6, pp. 734-749.

[11] Goldberg, D., Nichols, D., Oki, B. M., and Terry, D. (1992). Using collaborative filtering to weave an information tapestry. Communications of the ACM, vol.35, no.12, pp.61–70.

[12] Lang, K. (1995). NewsWeeder : Learning to filter netnews. In Proceedings of the 12th international conference on machine learning, pp. 331–339.

[13] Armstrong, R., Freitag, D., Joachims, T., and Mitchell, T. (1995). Webwatcher: A learning apprentice for the world wide web. AAAI Spring Symposium on Information Gathering, pp.6–12.

[14] Krulwich, B., and Burkey, C. (1997). The InfoFinder agent: learning user interests through heuristic phrase extraction. IEEE Expert-Intelligent Systems, Their Applications, vol.12, vol.5, pp. 22–27.

[15] Mooney, R. J., and Roy, L. (1999). Content-based book recommending using learning for text categorization. In ACM SIGIR '99 Workshop Recommender Systems: Algorithms and Evaluation, pp.195–240.

[16] Shardanand, U., and Maes, P.(1995). Social Information filtering: Algorithms for Automating 'Word of Mouth'", in Proc. CHI'95, pp. 210- 217.

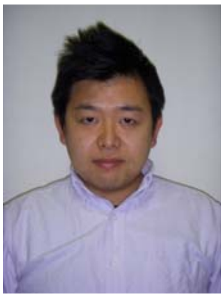
[17] Resnick, P., and Varian, H.R.(1997). Recommender systems, Communications of the ACM vol.40, pp.56-58

[18] Guido, P.(2005). Communion(Solidarity) and Power Conveyed by Social Relations: A Matter of Content or Structure?. ESCON TCK Conference, September 1st-4th 2005.

[19] Drachsler, H., Hummel, H.G.K., and Koper, R.(2008). Personal recommender system for learners in lifelong learning networks: requirements, techniques and model, International Journal of Learning Technology, Vol.3, No.4, pp.404-423.

[20] J. S. Breese, D. Heckerman, and C. M. Kadie(1998). Empirical analysis of predictive algorithms for collaborative filtering, in UAI, pp. 43- 52.

Xin WAN received the B.E. degree from Hubei University of Technology, Wuhan, China in 1996 and the M.E. degree from the University of Electro-Communications, Tokyo, Japan in 2007. She is currently a PhD candidate at Graduate school of Information Systems, the University of Electro-Communications, Tokyo, Japan. Her research interests include recommender system, data mining, artificial intelligence and collaborative learning.



Fumihiko ANMA received his B.E. from Tokyo Institute of Technology, Japan in 2000 and his Master Degree and PhD Degree from Shizuoka University, Japan in 2002 and 2005 respectively. He is currently an assistant professor with the Graduate school of Information Systems, The University of Electro-Communications, Japan. His research interests include Knowledge Computing, E-Learning, Artificial Intelligence, and Software Agent. He is a member of Japanese Society for Information and Systems in Education, Japanese Society for Artificial Intelligence.

Toshie NINOMIYA graduated from Department of Science and Technology in 1993, Keio University and worked in Toyota Motor Corporation as an engineer of industrial technology. She has an experience in education. She is currently a research associate at the University of Electro-Communications, Graduate School of Information Systems. Her research interests are in personalized adaptability of e-Learning environment including mentoring and monitoring system.

Toshio OKAMOTO obtained his PhD from Tokyo institute of Technology in 1988. He is currently a professor at the University of Electro-Communications (UEC), Graduate School of Information Systems and a director of the Center for e-Learning Research and Promotion in UEC. His research interests include theoretical and application studies/design of artificial intelligence, e-Learning, computer-supported collaborative learning systems and curriculum development in Information education. He is a convener of WG2 (Collaborative Technology) of LTSC/ISO SC36 (Learning Technologies Standards Committee).