Collaborative Filtering by User-Item Clustering Based on Structural Balancing Approach

Katsuhiro Honda, Akira Notsu and Hidetomo Ichihashi,

Osaka Prefecture University, Gakuencho 1-1, Nakaku, Sakai, Osaka, 599-8531 JAPAN

Summary
Collaborative filtering is a technique for reducing information overload and is achieved by predicting the applicability of items to users. In neighborhood-based algorithms, the applicability is predicted by the weighted averages of ratings of neighbors. This paper considers a new approach to user-item clustering in collaborative filtering. The new clustering method plays a role for selecting the user-item neighbors based on a structural balance theory used in social science, in which users and items are partitioned into two clusters by balancing a general signed graph composed of alternative evaluations on items by users.

Key words:
Collaborative filtering, Clustering, Signed graph, Perceptual balance.

1. Introduction

Automated collaborative filtering is computational realization of ``word-of-mouth'' in network community and also achieves reduction of information overload. Predicting the applicability of items to an active user based on a database of ratings given by users, personalized recommendation is performed by selecting the items that are in his/her interest. Then, the problem space is given as a matrix of users versus items, in which each element represents a user's rating of a specific item. So, the performance of a recommendation system depends on the ability of predicting missing values in the data matrix [1]. Neighborhood-based algorithms are built on the assumption that items to be recommended to a user are the items preferred by other users who have similar interests to the active user. The original GroupLens algorithm [2] generates predictions for the active user by calculating the weighted averages of ratings given by the ``neighbors'' and the subset of neighbors is chosen considering similarities (Pearson correlation coefficients) to the active user.

Clustering is a fundamental technique for grouping similar objects into ``clusters'' so that objects belonging to same clusters are mutually similar. Fuzzy *c*-Varieties (FCV) [3] is the FCM-type linear fuzzy clustering method, in which prototypes of clusters are given by linear varieties, and can be used for local principal component analysis because the prototypes are identified with local principal sub-spaces in local PCA [4]. Then, the local PCA technique was applied to the missing value estimation problem in collaborative filtering, in which missing values are estimated assuming that data points including missing values should exist on the nearest points to the prototypical linear varieties spanned by the local principal component vectors [5, 6]. In this sense, the clustering-based prediction method is a kind of neighborhood-based approach, in which the clustering part and the prototype estimation part are responsible for neighborhood selection and weighted average calculation, respectively.

This paper considers an approach to user-item clustering that is an extended relational data clustering. In the new approach, a structural balance theory used in social science is applied, in which the balancing process is identified with clustering of vertices in multi-vertex network models.

2. Structural Balance and Minimal Balancing Processes

2.1 Perceptual Balance

P-O-X Theory introduced by Heider [7] is a theory of perceptual balance or cognitive balance based on the naive psychology (or common-sense psychology). In the theory, a person's situation is expressed by words based on word analysis as well as the situation itself based on situation analysis, and various situations are categorized into the 8 notions of relation types. Then, either of positive (+) or negative (−) categories are given for all these relations. Heider considered the consistency of the relations based on the perceptual or cognitive balance of a person (P) with another person (O) with respect to an entity (X) in a localized situation setting. Assume that we have a situation shown in the right sub-figure of Fig. 1 where

- P likes O (+ : positive sentiment relation).
- O has a positive feeling for X (+ : positive unit formation relation).
- P has a negative feeling for X (−: negative sentiment relation).

Manuscript received December 5, 2008
Manuscript revised December 20, 2008

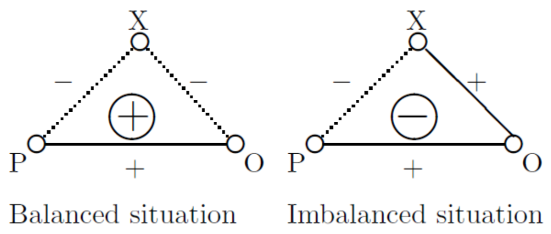

Fig. 1 Perceptually balanced and imbalanced situations in P-O-X theory.

In the figure, positive and negative relations are depicted by solid and broken lines, respectively. In the Heider's consideration, a balanced situation can be accepted by P without stress, while an ``imbalanced'' or ``unbalanced'' situation makes P feel stressful and uncomfortable. The balance of a triangular system is given by the sign of the product of the categories presented on three arcs, and the positive sign $(+)$ means a balanced situation. Then, the right sub-figure in Fig. 1 is regarded as ``imbalanced'' or ``unbalanced'' because $(+)\times(+)\times(-) = (-)$, and the imbalanced state (situation) should be replaced with a balanced one as is shown in the left sub-figure in Fig. 1. We also have other balanced situations among the three entities (P, O and X) [7].

2.2 General Signed Graph and Minimal Balancing Processes

Although Heider considered a localized situation setting with three vertices, we can also extend the P-O-X theory to more realistic social problems represented by a general signed graph with no restrictions on the number of members or items. In the following, we treat a general signed graph with no distinction between members and items used in social science [8, 9, 10, 11].

A (balanced or imbalanced) graph G is defined by the adjacency matrix $A(G) = [a_{ij}]$ whose entities are given by

$$a_{ij} \begin{cases} 1 & \text{if the relation between vertices } i \text{ and } j \text{ is positive,} \\ -1 & \text{if the relation between vertices } i \text{ and } j \text{ is negative,} \\ 0 & \text{if the relation of vertices } i \text{ and } j \text{ does not exist.} \end{cases} \quad (1)$$

The structural balance theorem proposed by Cartwright and Harary [9] says that the set of vertices (notions) is partitioned into two subgroups (one of which may be empty) in a balanced system and the relations among the vertices of the same subgroup have positive signs while relations between the vertices of different subgroups have negative ones as is shown in Fig. 2.

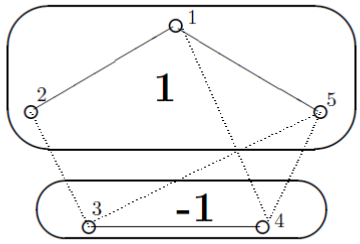

Fig. 2 A graph in balanced situation.

When we have n vertices in a graph G and give +1 to vertices of one subgroup and –1 to those of the other subgroup, the vertices of the signed graph for a balanced system is represented by a sign vector $s = (s_1, \quad s_2, \quad \cdots, \quad s_n)^T$, where

$$s_i \in \{1, \quad -1\}, \quad \text{for } i = 1,2,\cdots,n. \quad (2)$$

T denotes the transpose of the vector. For example, the balanced graph in Fig. 2 is represented as $s = (1, \quad 1, \quad -1, \quad -1, \quad 1)^T$.

We can calculate the number of different sign relations between a graph G with the adjacency matrix $A(G) = [a_{ij}]$ and a balanced graph G' whose subgroup of G' is represented by a sign vector $s = (s_1, \quad s_2, \quad \cdots, \quad s_n)^T$ as follows:

$$\begin{aligned} l &= \sum_{i,j=1}^{n} \frac{\left| a_{ij} - s_i \cdot s_j \cdot |a_{ij}| \right|}{4} \\ &= \frac{1}{4}\left(\left(\sum_{i,j=1}^{n} |a_{ij}| \right) - \left(\sum_{i,j=1}^{n} s_i \cdot s_j \cdot a_{ij} \right) \right) \\ &= \frac{1}{4}\left(\left(\sum_{i,j=1}^{n} |a_{ij}| \right) - \left(s^T A(G) s \right) \right). \end{aligned} \quad (3)$$

When the number of different sign relations l is minimum, the balanced graph G' is called the minimum balanced situation from graph G and the sign vector s^* satisfies the following equation:

$$s^{*T} A(G) s^* = \max_s s^T A(G) s \quad (4)$$

The balanced situation is achieved by transiting unbalanced vertices and a balancing process for deriving a minimum balanced situation is called a minimum balancing process. Here, the minimal balancing process is not necessarily unique.

Katai and Iwai [12, 13] discussed how to derive a minimal balancing process. In a minimum balanced situation, we have

$$\max_{s} s^T A(G) s \le e^{*T} A(G) e^*, \tag{5}$$

and $\| e^* \| = n$, where e^* is the eigenvector of $A(G)$ corresponding to the maximum eigenvalue λ^* and $\| \bullet \|$ is the Euclidean norm. Then, an approximately optimum sign vector s^* can be given by:

$$s^* = \left(s_1^*, \quad s_2^*, \quad \cdots, \quad s_n^*\right)^T,$$

$$s_i^* = \begin{cases} 1 & \text{if } e_i^* \ge 0 \\ -1 & \text{if } e_i^* < 0 \end{cases} \quad \text{for } i = 1, 2, \cdots, n. \tag{6}$$

3. User-item Clustering in Collaborative Filtering Based on Structural Balancing Approach

3.1 Problem Space of Collaborative Filtering

Assume that $X = [x_{ij}]$ is an ($n \times m$) data matrix consisting of m dimensional ratings of n users and its element of x_{ij} is the rating for item j given by user i.

GroupLens [2] is the most famous neighborhood-based algorithm that calculates the applicability of item j for active user i as a weighted sum of the ratings of other users:

$$y_{ij} = \bar{x}_i + \frac{\sum_{u=1}^{n} \left(x_{uj} - \bar{x}_u\right) \times r_{ij}}{\sum_{u=1}^{n} r_{ij}}, \tag{7}$$

where $\bar{x}_i$ is the average of the ratings voted by user i. The weight r_{iu} is a similarity measure between user i and user u, and the original GroupLens used Pearson correlation coefficients.

Honda *et al*. [5, 6] introduced the user clustering approach and proposed a linear fuzzy clustering-based prediction algorithm, in which users are partitioned into several linear fuzzy clusters and prediction models are estimated in clusters based on local sub-space learning (local PCA). In this paper, a new approach to user-item clustering in collaborative filtering is considered.

3.2 A Structural Balancing Approach to User-Item Clustering

The problem of predicting the applicability of items to an active user in collaborative filtering can be regarded as an alternative recommendation process if the filtering system recommends items whose prediction values are larger than a pre-defined threshold. In the alternative process, the evaluation value in matrix X can be identified with the alternative relation of ``positive'' or ``negative'' in the social systems modeling. In the remaining part of this section, user-item clustering in collaborative filtering is performed based on a structural balancing approach.

Assume that the evaluation matrix X is transformed into a partial adjacency matrix $Z = [z_{ij}]$ where z_{ij} is given as:

$$z_{ij} = \begin{cases} 1 & \text{if } x_{ij} \ge \varepsilon, \\ -1 & \text{if } x_{ij} < \varepsilon, \\ 0 & \text{if } i \text{ did not evaluate } j, \end{cases} \tag{8}$$

where ε is a pre-defined threshold. In order to generate a graph representing the mutual relation among users and items, the adjacency matrix A is defined as the following block matrix:

$$A = \begin{pmatrix} O & Z \\ Z^T & O \end{pmatrix}. \tag{9}$$

Note that the first n (or the remaining m) vertices correspond to n users (or m items), and users and items are not distinct as in the social systems modeling by Cartwright and Harary. Mutual relation among users (or items), however, is not known. So, the diagonal blocks of A are given by O (zero matrix).

Then, a balanced system is given by partitioning the set of vertices (users and items) into two subgroups in such a way that the relations between the vertices of the same subgroup have positive signs and relations between the vertices of different subgroups have negative signs, i.e., users and items are simultaneously partitioned without distinction between users and items.

3.3 Implementation Strategy

After use-item clustering, we can expect that users have positive relation with items included in the same cluster. Then, a simple strategy is to recommend all non-evaluated items in the cluster to the user. This strategy will be useful for improving the recall ratio of the system because almost half of items are to be recommended. However, user may be confused by too many recommendations when we have enormous items.

Table 2: Derived eigenvector and approximately optimum sign vector

	user						item									
	a	b	c	d	e	f	1	2	3	4	5	6	7	8	9	10
eigenvector e^*	0.35	0.23	0.35	-0.30	-0.09	-0.32	0.22	0.10	0.18	0.22	0.30	-0.30	-0.10	-0.22	-0.17	-0.30
sign vector s^*	1	1	1	-1	-1	-1	1	1	1	1	1	-1	-1	-1	-1	-1

Table 1: An artificially generated matrix with alternative evaluation

user	item									
	1	2	3	4	5	6	7	8	9	10
a	1	-1	1	1	1	-1	-1	-1	-1	-1
b	0	1	1	0	1	-1	1	0	-1	-1
c	1	1	0	1	1	-1	-1	-1	-1	-1
d	-1	-1	0	-1	-1	1	0	1	0	1
e	1	0	-1	1	-1	1	1	-1	0	1
f	-1	0	-1	-1	-1	1	0	1	0	1

Then, the second strategy is to use the clustering result for further selection in the conventional recommendation systems in order to improve the precision ratio of the system. For example, GroupLens recommendation system sometime picks up many candidate items to be recommended. In such cases, further selection from the candidates should be done and the result of use-item clustering can be applied.

4. Numerical experiments

4.1 Artificially generated data

A numerical experiment was performed by using an artificially generated evaluation matrix shown in Table 1. In the table, evaluation values on items by users are given in an alternative manner, in which ``1'' and `` -1 '' represent a positive and negative evaluations while ``0'' means that the user have not evaluated the item. The numbers of users and items are $n=6$ and $m=10$ , respectively. Note that users ``a-c'' mainly gave positive evaluations to the first 5 items while users ``c-f'' mainly gave negative ones to them. The (16×16) adjacency matrix A was generated. Then, the eigenvector corresponding to the largest eigenvalue of A and the approximately optimum sign vector s^* (16 dimensional vectors) were given as Table 2. From the table, we can see that the users form the two clusters of ``a-c'' and ``c-f'', and items ``1-5'' (or ``6-10'') were associated to the first (or second) cluster. In this way, the users were properly partitioned into two subgroups and the items were also assigned to the subgroups.

Table 3: Average recommendation ability with strategy 1

	GroupLens	Strategy 1
Precision	0.736544	0.674856
Recall	0.756660	0.811967

Table 4: Average recommendation ability with strategy 2

	GroupLens	Strategy 1
Precision	0.736544	0.750391
Recall	0.756660	0.698329

4.2 Movie evaluation data

Next, the proposed approach was applied to a ratings data set collected for purposes of anonymous review from the MovieLens movie recommendation site [14]. The data set is originally composed of 100,000 ratings from 943 users, with every user having evaluated at least 20 ratings on a scale from 1 to 5 based on the semantic differential (SD) method. Only 82,275 ratings from 874 users ($n=874$) for 598 movies ($m=598$) were used in this experiment so that each movie had been evaluated by at least 50 users and each user had evaluated at least 20 movies because other movie ratings were difficult to predict from correlations among the users. In this experiment, the pre-defined threshold ε was set as 3.5. The evaluation data matrix was transformed into the $(874+598)\times(874+598)$ adjacency matrix A whose elements are given in an alternative manner. Then, the users and movies were partitioned into two clusters by calculating the eigenvector and its associated sign vector.

The recommendation ability of the system was validated by using the two implementation strategies. In the strategy 1, users were recommended all movies included in the same cluster from the system. Table 3 shows the average recommendation ability given by 5-hold cross validation. In this experiment, ``recall ratio'' and ``precision ratio'' were used for validating the recommendation ability. Recall (or also called sensitivity) is the fraction of the movies that are relevant to the user that are successfully retrieved. Precision is the percent of retrieved movies that are relevant to the search. The larger the values, the higher the recommendation ability of the system. As is shown in Table 3, the strategy 1 derived a higher recall ratio than

GroupLens because the strategy recommends more movies, i.e., the strategy can dig up the movies that were missed by GroupLens while more movies are recommended by the system. On the other hand, in the strategy 2, the user-item clustering result was used for further selection in GroupLens. Table 4 shows the result and indicates that the strategy improved the precision ratio of GroupLens, i.e., the strategy is useful for further selecting high quality subset after GroupLens selection. In this way, the proposed user-item clustering approach can be used in conjunction with the conventional recommendation systems.

5. Conclusions

This paper proposed a new user-item clustering method in collaborative filtering based on structural balancing approach. Experimental results demonstrated that the two implementation strategies of simply all selection and further selection can be used in conjunction with the conventional recommendation systems for improving the recommendation quality.

Although the proposed method partitioned users and items into two clusters based on the consideration given by Cartwright and Harary, it was also pointed out that the decomposition into just two subgroups is not necessarily enough [15,16]. Discussion on the further clusterability is remained in future work.

Another potential future work is a comparative study with other clustering algorithms based on eigen decomposition. For example, PCA-guided *k*-Means [17] partitions data through eigen decomposition of a matrix composed of inner product similarity measure while the proposed method used the eigenvector of an adjacency matrix given by alternative evaluation. In [18], clusters are sequentially extracted by calculating the principal eigenvector of a matrix of non-negative similarity measures. The theoretical comparison will be useful for further improvement of the proposed method.

Acknowledgments

This work was funded in part by the Ministry of Internal Affairs and Communications, Japan, under Strategic Information and Communications R&D Promotion Programme (SCOPE).

References

[1] J. L. Herlocker, J. A. Konstan, A. Borchers and J. Riedl: An algorithmic framework for performing collaborative filtering, Proc. of Conference on Research and Development in Information Retrieval, 1999.

[2] J. A. Konstan, B. N. Miller, D. Maltz, J. L. Herlocker, L. R. Gardon and J. Riedl: Grouplens: applying collaborative filtering to usenet news, Communications of the ACM, vol.40, no.3, pp.77-87, 1997.

[3] J. C. Bezdek: Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, 1981.

[4] K. Honda and H. Ichihashi: Regularized linear fuzzy clustering and probabilistic PCA mixture models, IEEE Trans. Fuzzy Systems, vol.13, no.4, pp.508-516, 2005.

[5] K. Honda, N. Sugiura, H. Ichihashi and S. Araki: Collaborative filtering using principal component analysis and fuzzy clustering, Web Intelligence: Research and Development, Lecture Notes in Artificial Intelligence 2198, Springer, pp.394-402, 2001.

[6] K. Honda and H. Ichihashi: Component-wise robust linear fuzzy clustering for collaborative filtering, Int'l J. Approximate Reasoning, vol.37, no.2 pp.127-144, 2004.

[7] F. Heider: The Psychology of Interpersonal Relations, John Wiley, 1958.

[8] F. Harary: A structural analysis of situation in the middle east in 1961, J. Conflict Resolution, vol.5, no.2, pp.167-178, 1961.

[9] D. Cartwright and F. Harary: Structural balance, a generalization of Heider's theory, Psychological Review, vol.63, pp.167-293, 1956.

[10] T. M. Newcomb: An approach to the study of communicative acts, Psychological Review, vol.60, pp.393-404, 1953.

[11] A. Notsu, H. Ichihashi and K. Honda: Agent simulation based on perceptual balance, Int'l J. Computer Science and Network Security, vol.6, no.12, pp.50-54, 2006.

[12] O. Katai and S. Iwai: Studies on the balancing, the minimal balancing, and the minimum balancing processes for social groups with planar and nonplanar graph structures, J. Math. Psychol., vol.18, no.2, pp.141-176, 1978.

[13] O. Katai and S. Iwai: On the characterization of balancing processes of social systems and the derivation of the minimal balancing processes, IEEE Trans. Systems, Man, and Cybernetics, vol.SMC-8, no.5, pp.337-348, 1978.

[14] MovieLens Web Page; http://www.movielens.org/

[15] J. A. Davis: Clustering and structural balance in graph, Human Relation, vol.20, pp.181-187, 1967.

[16] S. Wasserman and K. Faust: Social Network Analysis, Cambridge University Press, 1994.

[17] C. Ding and X. He: K-means clustering via principal component analysis, Proc. of Int'l Conf. Machine Learning (ICML 2004), 225-232, 2004.

[18] K. Inoue and K. Urahama: Sequential fuzzy cluster extraction by a graph spectral method, Pattern Recognition Letters, vol. 20, pp.699-705, 1999.

Katsuhiro Honda received the B.E., M.E. and D.Eng. degrees in industrial engineering from Osaka Prefecture University, Osaka, Japan, in 1997, 1999 and 2004, respectively.

He is currently an Assistant Professor, Department of Computer Sciences and Intelligent Systems, Osaka Prefecture University. His research interests include hybrid techniques of fuzzy clustering and multivariate analysis, data mining with fuzzy data analysis and neural networks. He received paper award and young investigator award from Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT) in 2002 and 2005, respectively, and gave a tutorial on "Introduction to Clustering Techniques" at 2004 IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE 2004).

Akira Notsu received the B.E., M.I. and D. Informatics degrees from Kyoto University in 2000, 2002 and 2005, respectively.

He is currently an Assistant Professor, Department of Computer Sciences and Intelligent Systems, Osaka Prefecture University. His research interests include agent-based social simulation, communication networks, game theory, human-machine interface, and cognitive engineering.

Hidetomo Ichihashi received the B.E. and D.Eng. degrees in industrial engineering from Osaka Prefecture University, Osaka, Japan, in 1971 and 1986, respectively.

From 1971 to 1981, he was with the Information System Center of Matsushita Electric Industrial Co., Ltd., Tokyo, Japan. From 1981 to 1993, he was a Research Associate, Assistant Professor, and Associate Professor at Osaka Prefecture University, where he is currently a Professor in the Department of Computer Sciences and Intelligent Systems. His fields of interest are adaptive modeling of GMDH-type neural networks, fuzzy C-means clustering and classifier, data mining with fuzzy data analysis, human-machine interface, and cognitive engineering.