nternational Journal of Fuzzy Logic and Intelligent Systems, vol. 6, no. 2, June 2006 pp. 110-115

Collaborative Filtering based Context Information for Real-time Recommendation Service in Ubiquitous Computing

Se-Il Lee* and Sang-Yong Lee**

***Dept. of Computer Engineering, Kongju National University**

****Division of Computer Science & Engineering, Kongju National University**

Abstract

In pure P2P environment, it is possible to provide service by using a little real-time information without using accumulated information. But in case of using only a little information that was locally collected, quality of recommendation service can be fallen-off. Therefore, it is necessary to study a method to improve quality of recommendation service by using users' context information. But because a great volume of users' context information can be recognized in a moment, there can be a scalability problem and there are limitations in supporting differentiated services according to fields and items.

In this paper, we solved the scalability problem by clustering context information per each service field and classifying it per each user, using SOM. In addition, we could recommend proper services for users by quantifying the context information of the users belonging to the similar classification to the service requester among classified data and then using collaborative filtering.

Key Words : Collaborative Filtering, Context, Recommendation System, Pure P2P, SOM

1. Introduction

In order to receive recommendation service in ubiquitous environment by means of mobile devices, recommendation systems provide users with recommendation service by service history information being stored in a server[1]. However, because of excessive contents recommended by various people, the server of the systems requires enormous memory space. One method to solve this problem is to recommend service according to real-time information by pure P2P. But this also has the following problems. That is, although it is necessary to recommend service by comparing service requester's own service information with other similar users', accumulated information can not be used in the pure P2P. So only a little service information is used and compared with other similar users', which can have the possibility of resulting in inaccurate information. In order to solve these problems, we can use collaborative filtering method for context information, which is most frequently used in recommendation systems. Thereby we can upgrade quality of recommendation information. But because there is limitation for users' context information to support differentiated services according to areas and items, there can be also the scalability problems caused by collection of too much information at a certain area.

In this paper, in order to solve these problems, we intend to have context information clustered per each service area by using SOM(Self-Organizing Maps) and then to classify it per user; then to do collaborative filtering by quantifying context information included in the similar classification to the user's among classified data. In such a way, the researcher solved the scalability problem by classifying only similar context information to the service requester's. In addition, the method is not to use only single service information but to have collaborative filtering of other context information by quantifying it, so it is possible to support proper service for service requestors.

As composition of this paper, we describe relevant researches in chapter 2; analyze context information using SOM in chapter 3; describe experimented and evaluated result in chapter 4; and mention about conclusion finally in chapter 5.

2. Related Works

2.1 P2P

P2P is an abbreviation of peer to peer, meaning a distributed computing structure, which enables direct transmission and management of information between clients without server or through a single management server. P2P technology can be divided into hybrid P2P with central management server and pure P2P without central server. The first system of this pure P2P method is Gnutella. Gnutella is not a routing based on physical connection centering on circuits, but logical routing where node itself has router's movement and delivers passing through data to adjacent nodes. Pure P2P has good capability of connection and extension, but lacks in managemental aspects and security[3].

2.2 Context

Definition of context mentioned in ubiquitous computing is

Manuscript received Jun. 1, 2006; revised Jun. 13, 2006.
Corresponding Author : Sang-Yong Lee

information about the users such as their present location, behaviors and works, etc, and continuous change of such information under users' environment[4]. In addition, Schilt of Xerox PARC is defined as identity and information related to users and objects. Abowd of GATECH[5] is defined as every information related to individual such as users, space, object and etc.

2.3 Collaborative Filtering

Collaborative filtering is method to recommend only necessary information among enormous amount of information and it is most frequently used. In order to predict items users might like, collaborative filtering is to recommend items based on evaluation of other customer items with similar preference[6,7].

For example, if we give points from 1 to 5 as for 5 users and 4 items respectively, we can get a table like Table 1.

Table 1. Evaluation Item

Item / User	Item 1	Item 2	Item 3	Item 4
User 1	2	5	3	1
User 2	1	2	4	
User 3	4		5	4
User 4	5	1	3	1
User 5	2	5	2	?

In order to have the item 1 of the user 5 recommended, if we consider relation, we can see that referring to the user 1 having the most relation with the user 5 would be good.

3. Analysis of Context Information using SOM

Order of analyzing context information and applying such information for collaborative filtering under pure P2P environment is as follows (Fig. 1).

① To receive and integrate context information from sensors and then to analyze and classify only necessary context information.

② To classify classified context information per each service area by using SOM.

③ To recommend service information to users by means of collaborative filtering after quantifying it in order to make a certain value.

3.1 Recognition of Context

In order for a service requester to receive desired service using mobile device, users' real-time context information is necessary in pure P2P environment. For recommendation of service, first of all, the context integration agent integrates its own profile information and users' context information drawn from sensors and the local P2P incorporates only basic information. Integrated context information classifies only context information necessary for service. In this paper, the researcher applied it to movies in which people have much interest, and necessary context information is as Table 2.

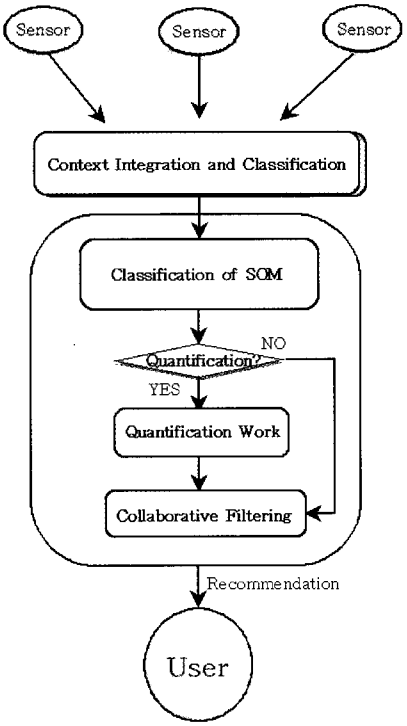

Fig 1. Conceptual Diagram

Table 2. Context Information for Service

Type	Contents
Users' ID	Distinction of user
Sex	Male/Female
Place	Service place
Time	Service time
Genre	Genre of movies(Clustering)
Age	Age is classified by 10 units.
Companions	Intension (friend, Lover, ...)
Title	Tile of movies (Recommendation)

In order for clustering of classified context information per genre, we move to the step of classification by means of SOM.

3.2 Clustering by Means of SOM and Quantifying Context

If amount of context information collected from users increases, there can be the scalability problem. As the first step to solve this problem, it is necessary to cluster service information per genre in order to obtain service information with similar degree of interest to user's by using SOM.

SOM is an algorithm based on a nerve network being

proposed by Kohonen where the connecting line between input and output layers indicates connection strength and both layers are completely connected. In addition, for connection strength vector and input vector, normalized value from 0 to 1 is used in general and only new run with closest distance between connection strength vector and input vector can have output [8, 9]. Performance algorithm for classification per genre is as follows.

① To initialize weight. Then to designate a small number to the line connecting N numbers' input and M numbers' output as weight.

② To suggest new input (per genre).

③ To calculate similarity between input and output.

$$d_b = \sum_{a=0}^{n-1} (R_a(t) - V_{ab}(t))^2 \quad (1)$$

Ra(t) is the a-th input of node in time t and Vab(t) has weight from a to b in time t.

④ To select node b with the highest similarity.

⑤ To renew weight between the node with high similarity and the adjacent node.

$$V_{ab}(t+1) = V_{ab}(t) + \eta(t)(R_a(t) - V_{ab}(t)) \quad (2)$$

$\eta(t)$ is studying ratio, having value from 0 to 1, which becomes smaller by means of more study.

⑥ To repeat again from the 2nd step until you has desired studying result.

Classification using SOM is as Fig. 2.

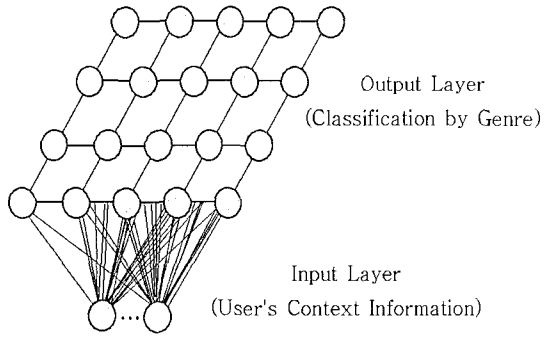

Fig. 2. Classification by genre using SOM

Every context information is not used directly but used after being classified, so we can improve the problem of scalability.

In order to use this classified context information directly for collaborative filtering, quantification is necessary. Context information to be used for quantification includes age, sex, companions, service time and etc. and quantified value is from 1 to 5. In order to quantify age among context information, it is necessary to allocate high point to the users with similar age to service requester's and lower point to users according to age difference. For according to age, there is difference in interest in movies. The following formula(3) is the quantifying formula per age where the minimum point 1 is allocated for items with the biggest difference in age and the highest point 5 for items showing similarity in age. Followings are quantification method of age.

$M_{age} = | S_{Age} - U_{Age} |$
if Mage>4 then Mage=4
if Mage<0 then Mage=0
Mage=Mage+1 //For quantification from 1 to 5 (3)

Sage is the share calculated by dividing service requester's age by 10 and Uage is the share calculated by dividing compared user's age by 10. Table 3 is quantified value obtained by classifying service requester's and users' age.

Table 3. Measure of Age

		Users' Age				
		~ 19	20	30	40	50~
Service Requester's Age	~ 19	5	4	3	2	1
	20	4	5	4	3	2
	30	3	4	3	4	3
	40	2	3	2	5	4
	50~	1	2	1	4	5

In addition, because distinction from sex has shown similar taste as the result of collecting data, point 4 and 5 is allocated respectively. Then using information of mobile group, the researcher allocated points from 1 to 5 using the graph in Figure 4based on whether a person has come to a movie theater together with companions or not. And because there is also difference in the taste of watching movies according to serviced time, the researcher allocated value from 1 to 5.

3.3 Collaborative Filtering Using Context Information

If composition of a table is like Table 1, there aren't any problems in recommending information using collaborative filtering. But because there aren't any accumulated data in pure P2P, there can be problems of mutual irrelevance as Table 4and it is difficult to indicate customers' satisfaction with service by means of numbers.

Table 4. Comparative table of pure P2P

Item / User	Service 1	Service 2	Service 3	Service 4
User 1	Firewall			
User 2				Flight plan
User 3		Munich		
Service Requester			Curious George	

However, in the method of collaborative filtering using context, it is possible to provide service requesters with

recommendation service by using even a single serviced item of similar type together with other context information. Table 5 is an example of context information required for collaborative filtering.

Table 5. Context Information

Item / User	Age	Sex	Intention	Time	Title
User 1	19	M	Single	Afternoon	Panther
User 2	59	M	Family	Evening	King
User 3	26	F	Friend	Afternoon	Flight plan
User 4	28	F	Lover	Evening	Geisha
Service Requester	33	M	Lover	Evening	?

With the above context information, quantification like Table 6 is made.

Table 6. Measured Value

Item / User	Age	Sex	Intention	Time	Title
User 1	3	5	4	4	Panther
User 2	3	5	4	5	King
User 3	4	4	3	4	Flight Plan
User 4	4	4	5	5	Geisha
Service Requester	5	5	5	5	?

If collaborative filtering is applied to quantified value, service information can be recommended by reference to the user 4 with the most similarity to service requester's. Therefore, service requester can receive services with better quality.

4. Evaluation

For this paper, the researcher designed and made experiment using C#, J2ME and WIPI under environment of Pentium IV, 2.8Ghz and 512MB. As examples of data used for experiment are as follows:

"0111234567\M\22\Drama\Movie\200602131830\Friend\King"

"0111234567" indicates user's ID and "M" is distinction between male and female; "22" indicates age; "Drama", genre, "Movie", service place; "200602131830", date and time; "Friend", intention for coming to see a movie; "King", serviced items. Table 7 shows examples of context information.

Among the above context information, user's ID, service place and genre are used for classification works and as context information requiring quantification, there are age, companions, time, and sex (Table 8). In addition, titles of movies are items to be used for service recommendation.

Because age has much relation, its quantification is easy, but it is not easy to quantify relation of companions, time and sex, so the researcher made up questions to 427 persons and investigated relation. Among respondents to questionnaires, males were 241 and females were 186 as the Fig. 3.

Table 7. Example of Context Information

Order	Context Information
1	0111234567\M\22\Drama\Movie\200602131830\Friend\King
2	0162343456\M\35\Romance\Movie\200602131830\Lover\Love
3	0111011222\F\23\Comedy\Movie\200602131900\Friend\Panther
4	0172324563\F\25\Romance\Movie\200602131830\Lover\Love
5	0161529535\M\37\Drama\Movie\200602131900\Friend\King
6	0182245932\M\29\Drama\Movie\200602131900\Family\George
7	0111594234\F\53\Drama\Movie\200602131900\Family\My brother
8	0167112134\M\26\Comedy\Movie\200602131900\Friend\Panther
9	0119364743\M\25\Action\Movie\200602131800\Friend\Kungfu
10	0114132872\F\32\Drama\Movie\200602131830\Lover\King
...	...

Table 8. Classified Context Information

Age	~ 9	20~ 9	30~ 9	40~ 9	50~
Sex	Male	Female			
Time	Early Morning	Morning	Afternoon	Evening	Midnight
Companions	Companion	Friend	Single	Lover	Family

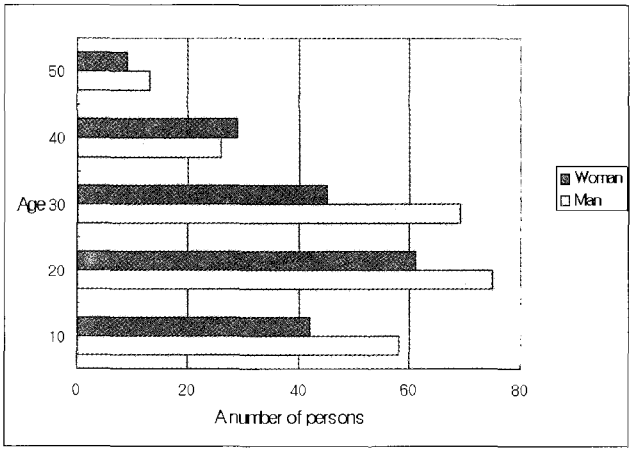

Fig. 3. User Groups classified by Age

In case of companions, the researcher asked them to select the genre they want to see the most and then compared relation with high rank item and obtained the result. The result was obtained by comparing relation with high rank item

while moving toward 'companions' and 'friends' on the basis of 'single'. Fields of action and comedy showed closer relation. 'Lover' seemed to be more intimate to the fields of 'romance' and 'drama'. And 'family' was intimate to 'family drama' or 'drama'. Fig. 4 is a graph expressing co-relation of companions by using preferred high rank genre.

Although relation was weaker than companions, time also showed some relation. The researcher could find that in the earlier time before noon, people preferred 'Action' and 'Comedy' movies, but in the later time at midnight, people preferred 'Romance' or 'Drama' movies.

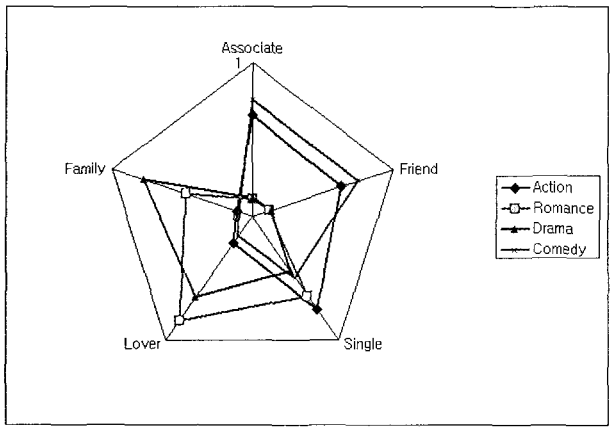

Fig. 4. Association Graph of Companion

$$|E| = \frac{\sum |v_i - r_i|}{N} \quad (4)$$

As the result of experiment, the researcher could find that the collaborative filtering using SOM(CFS) on the basis of context information had far less errors than the existing GroupLen method(CF) and it had about 0.0044 less errors in average than the collaborative filtering applying Naive Bayesian Algorithm(CFN)[2].

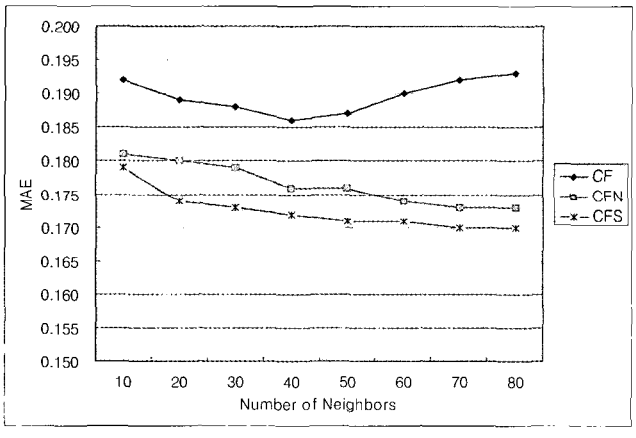

Fig. 5. MAE according to the Number of Neighbors

5. Conclusion

In the case of collaborative filtering using only a little information in the pure P2P environment, there can be problems of falling-off in quality of service, so the researchers used context information. But in case of using context information, there can be the scalability problems caused by collection of excessive information. Therefore, by clustering information per each service area by SOM, we improved the scalability problems.

In addition, it is difficult to provide differentiated service according to areas and items with only users' context information. So we could improve quality of services by quantifying context information such as age, companions, time, sex and etc. and then by using collaborative filtering.

As the result of experiment using WIPI emulation, the proposed method was evaluated to be 9.9% better than CF and 2.3% better than CFN in the aspect of evaluation of service supporting function.

Reference

[1] Mi-Sug Ku, Jeong-Hee Hwang, Nam-Kyu Choi, Doo Young Jung, Keun Ho Ryu, "*Context-based Incremental Preference Analysis Method in Ubiquitous Commerce*", Korea Information Processing Society, Vol.11-D, No. 07, pp.1417-1426, 2004.12

[2] Se-Il Lee, Sang-Yong Lee, "*Collaborative Filtering Method Using Context of P2P Mobile Agents*", Korea Fuzzy Logic and Intelligent Systems Society, Vol.15, No.5, pp.643-648, 2005

[3] Eung-Gon Kim, Dong-Hyun Kim, Sung-Ju Lee, "*A Study on P2P Business Model Analysis*", Korea Fuzzy Logic and Intelligent Systems Society, Vol.11, NO.01, pp.203-206, 2001

[4] S.Jang, W.Woo, "ubi-UCAM: A Unified Context-Aware Application Model.", LNAI(Context03), pp.178-189, 2003

[5] D.Salber, A.K.Dey and G.D.Abowd, "*The Context Toolkit:Aiding the Development of Context-Aware Application*", In the Workshop on Software Engineering for Wearable and Pervasive Computing (Limerick Ireland), June, 2000

[6] B. M. Sarwar, J. A. Konstan, Al Borches, J. Herlocker, B. Miller, and J. Riedl. "*Using Filtering Agents to Improve Prediction Quality in the GroupLens Research Collaborative Filtering System*", Proceedings of the 1998 Conference on Computer Supported Cooperative Work, 1998

[7] N.Good, B. Schafer, J.Konstan, A. Borchers, B.Sarwar, J. Riedl, "*Combining Collaborative filtering with personal Agents for Better Recommendation*", AAAI/IAAI, pp.439-446, 1999

[8] T. Kohonen, *Self-Organizing Maps*, Springer-Verlag, Berlin, 1995

[9] Hee-Seung Kim, *Pattern Recognition*, Saing-Reung, pp.323-346, 1993

Se-Il Lee

1993 : Dept. of Computer sciences, Taejon National University of Technology (B.S)

2001 : Dept. of Computer Science and Electronics Engineering, Chungwoon University (M.S)

2004～ : Dept of Computer Engineering, Kongju National University (Doctoral Student)

Research Interests : Ubiquitous Computing, Context-Awareness, Collaborative Filtering, Game Algorithm

E-mail : lsilhr@kongju.ac.kr

Sang-Yong Lee

1984 : Dept. of Computer Science, Chungang University (B.S)

1988 : Dept. of System Science, Tokyo Institute of Technology (M.S)

1988～ 989 : Researcher of NEC Research Center in Japan

1993 : Dept. of Computer sciences, Chungang University (Ph.D)

1996～ 997 : Visiting Professor of University of Central Florida in U.S.A.

1993～ : Professor of Division of Computer Science & Engineering, Kongju National University

Research Interests : Artificial Intelligent, Agent System, Ubiquitous Computing, Context-Awareness

E-mail : sylee@kongju.ac.kr