Using Concept Lattice for Personalized Recommendation System Design

Caifeng Zou, *Student Member, IEEE*, Daqiang Zhang, *Member, IEEE*, Jiafu Wan*, *Member, IEEE*, and Jaime Lloret, *Senior Member, IEEE*

Abstract **— A novel personalized recommendation system based on concept lattice is presented. The system is divided into the offline part and the online part. In the offline part, the formal context and the concept lattice are constructed from the transaction database, and the association rules based on concept lattice are extracted. The obtained association rules are stored in the rule library. The concept lattice and the rule library are updated regularly using the new added data. In the online part, using the behavior data of target user, and the concept lattice and the rule library obtained in the offline part, the ordered recommendation result is calculated, and returned to the user. There are two recommendation methods in the online part, which are recommendation based on association rules and collaborative filtering recommendation. It is necessary to further study the efficient concept lattice construction algorithm, and association rule library update algorithm based on concept lattice in the future.**

Index Terms **— Association rule extraction, collaborative filtering recommendation, concept lattice, formal concept analysis, formal context, Hasse diagram, similarity degree.**

I. INTRODUCTION

With the rapid development of Internet and information technology, in the information sea of big data, the amount of information grows exponentially and the problems of information explosion and information overload have emerged. Because of heterogeneity, diversity, complexity and breadth of big data, people can get an increasing number of information, but it becomes increasingly difficult to quickly get the valuable information which people are interested in. How to get the latest useful data quickly from big data, how to improve the intelligence level of information retrieval, and how to meet the individual needs of users, are challenging tasks which the new information service system faces.

Manuscript received July 30, 2014.

This work was supported in part by the National Natural Science Foundation of China (No. 61262013).

C. Zou is with the School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China (e-mail: caifengzou@gmail.com).

D Zhang is with the School of Software Engineering, Tongji University, Shanghai 201804, China (e-mail: dqzhang@ieee.org).

J. Wan is with the School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510641, China (e-mail: jiafuwan_76@163.com).

J. Lloret is with the Department of Communications, Universidad Politecnica de Valencia, Valencia, Spain (jlloret@dcom.upv.es)

*Corresponding author: Jiafu Wan, Email: jiafuwan_76@163.com

As a novel intelligent information service mode, personalized recommendation can provide accurate information and services to users, and effectively solve the problems of information overload and information lost. Thus, personalized recommendation system attracts increasing attention and has been widely used in the field of the Internet industry, such as electronic commerce and social networks.

In the personalized recommendation system, collaborative filtering recommendation is one of the most successful technologies, whose basic idea is to generate recommendation result to the target user according to the rating data of the nearest similar neighbors [1]. In addition, with the development of data mining, association rules have been applied to the recommendation system. The personalized recommendation system based on association rules can find new interest points of customers, and does not rely on domain knowledge.

In collaborative filtering recommendation and association rule mining algorithms, the idea of Formal Concept Analysis (FCA) can be adopted, and concept lattice can be used as the fundamental data structure, in order to improve the efficiency of data mining and knowledge discovery algorithms, and obtain useful knowledge from massive information timely and efficiently.

The concept is the basic unit of human cognition and an important research object of artificial intelligence disciplines. German mathematician Wille proposed FCA in 1982 [2]. He systematically studied the hierarchies of concepts, properties of lattice algebra, and the isomorphic nature of concept lattice and formal context, which established foundations for the field of FCA.

FCA is a powerful tool for data analysis and rule extraction from the formal context. FCA expresses concepts, attributes, and relationships of the ontology with formal context. According to the context, concept lattice is constructed to show the structure of the ontology clearly, and describe the generalization and specialization of the concept. Concept Lattice, also known as Galois Lattice, is the core data structure of FCA. In concept lattice, each node is a formal concept. Formal concept consists of extension part and intension part [3]. Extension of the concept is considered as the set of all objects belonging to this concept, and intension is considered as the set of the common characteristics or attributes of all these objects [4]. Concept lattice essentially describes the affiliation between objects and features, and shows the relationships of generalization and specialization between the concepts. The corresponding Hasse diagram contributes to data visualization.

This paper studies the application of the concept lattice in personalized recommendation service, and designs a novel personalized recommendation system based on concept lattice. Take movie resource recommendation system for example, this paper uses association rule mining algorithm and collaborative filtering recommendation algorithm based on concept lattice to provide the most appropriate movie recommendation service to users.

This paper is organized as follows. Section II surveys related research work. Section III presents our design for the personalized recommendation system based on concept lattice. Section IV outlines the recommendation algorithm based on association rules. Section V describes the collaborative filtering recommendation method based on concept lattice. Section VI concludes this paper and discusses future research.

II. Related Work

A. Formal Concept Analysis and Concept Lattice

The basic concepts of FCA include formal context, concept lattice, Hasse diagram, and so on [5].

1) Formal Context

Formal context is defined as a triple *K (U, A, I)*, where *U* is a set of objects, *A* is a set of attributes, and *I* is a binary relation between object *U* and attribute *A*, i.e. $I \subseteq U \times A$. If there is $(x,a) \in I$, then xIa shows that object x has attribute a. The form of two-dimensional data table is also a type of formal context. The tuple represents object or instance, and the column represents attribute.

When there is
$$X^* = \{a \mid a \in A, \forall x \in X, xIa\}, X \subseteq U$$
$$B^* = \{x \mid x \in U, \forall a \in B, xIa\}, B \subseteq A$$

if $\exists X^* = B$ and $B^* = X$, then (X,B) is called a formal concept or simply a concept [6]. *X* is defined as the extension of concept (X,B). *B* is defined as the intension of concept (X,B). Extension of the concept indicates the set of all objects belonging to this concept, and intension of the concept indicates the set of the common attributes of all these objects. For example, $C((1,5),\{b,c,e\})$ indicates that concept *C* covers two objects 1 and 5. The common attribute of these two objects is $\{b,c,e\}$.

2) Concept Lattice

Concept lattice [7] is used to indicate the relationship between attributes and objects. There is a kind of partially ordered relationship between the nodes of concept lattice. Given $C_1(X_1,B_1)$, $C_2(X_2,B_2)$, then $C_1 < C_2 \Leftrightarrow B_1 < B_2$.

This partially ordered relationship means that $C_1(X_1,B_1)$ is a senior concept of $C_2(X_2,B_2)$, or $C_1(X_1,B_1)$ is a generalization of $C_2(X_2,B_2)$.

For formal context *(U, A, I)*, there is a unique partially ordered set in relationship *I*. This partially ordered set produces a lattice structure. Lattice *L* generated from the context *(U, A, I)*, is called the concept lattice. The concept lattice of a formal context is unique.

3) Hasse Diagram

Hasse diagram [7] expresses the partially ordered relationship in the concept lattice with a concise and effective diagram. It is a commonly used expression method of concept lattice, which can intuitively show the overall demonstration of generalization and specialization relationships among all concepts in the concept lattice. Specifically, for the partially ordered set $(S,\prec)$, each element of *S* is represented by the vertex of the plane. If *y* covers *x* (that is, *x* <*y* and $\neg\exists Z$ satisfies *x* <*z* <*y*), then the line from *x* up to *y* is drawn. These lines can be cross but can not touch any non-endpoint vertex. The diagram with labeled vertices can uniquely determine the partial order of the set.

B. Association Rule Extraction Based on Concept Lattice

Concept lattice is an effective tool for data analysis and association rule extraction. As a knowledge representation, concept lattice is used to discover potentially novel relationships between data items in the dataset, and has been widely applied in big data processing in Internet of Things, Cyber-Physical Systems, M2M Networks and other fields [8-10].

The most important application of concept lattice in data mining is to extract the association rules. Concept lattice is a natural platform for association rule mining. Association rule describes the relationship between intension sets, while concept lattice reflects the unification between the intension and extension of the concept. Each node of the concept lattice is actually a maximum itemset, and the relationship between nodes reflects generalization and specialization between the concepts. Thus, the concept lattice is very suitable as data structure of rule extraction. Compared to other methods, extracting rules using concept lattice can get very good effects [11].

Concept lattice can effectively describe the extraction of association, implication, equivalence and classification rules [12]. Hasse diagram of the concept lattice can be used to get frequent itemsets clearly and concisely. Extracting the interesting rules using Hasse diagram, can reduce the amount of redundant rules and improve the accuracy of data mining.

Godin and Missaoui [13] proposed association rules extracting algorithm using the concept lattice based on incremental concept formation approach. Missaoui, et al. [14] put forward the exact and approximate association rule extraction algorithm. Hu, et al. [15] proposed an integration algorithm to extract classification and association rule based on the concept lattice.

Research of association rule mining algorithm based on concept lattice has obtained some achievements.

These algorithms can be divided into two categories: 1) The association rules are extracted directly from the concept lattice according to the preset confidence and support threshold; 2) The frequent item sets are extracted from the concept lattice according to the preset support threshold, then the association rules are extracted from the frequent item sets according to the traditional methods.

1) Association rules are extracted directly from the concept

lattice. Xie and Liu [16] implemented an association rule mining algorithm based on concept lattice. The algorithm can extract rules directly according to the preset support threshold and confidence threshold.

2) Frequent item sets are extracted from the concept lattice according to the support threshold. In order to extract association rules from concept lattice, the transaction database *TD* can be understood as a formal context $K = (O, D, R)$. O is the set of transactions in the database *TD*, and it is corresponding to the objects in the concept node, namely the extension. D is the set of all possible features in the database, which corresponds to the attributes in the concept nodes, namely the intension. For $x \in O, d \in D, xRd$ if and only if d belongs to the item sets of transaction x. According to the concept lattice corresponding to the formal context K, all the frequent item sets can be calculated to extract association rules [17].

C. *Collaborative Filtering Recommendation*

Goldberg et al proposed the collaborative filtering algorithm in 1992 [18]. Its basic idea is to collect historical data of the user and perform statistical calculations to predict some similar users, and recommend the interested items of similar users to the target user.

In the collaborative filtering recommendation system, there is a user-item rating matrix, where the higher score corresponds with the stronger user interest. The unknown values are estimated based on the known value in the user-item rating matrix. The most commonly used models in the collaborative filtering recommendation system are neighbor relationship model and implicit vector model.

In the neighbor relationship model [19], the associations between users or between items are constructed. The unknown values are estimated according to scores made by neighbors of the current user, or scores obtained by neighbors of the current item.

In the implicit vector model [20], matrix factorization technique is used to analyze the rating matrix. The users and the items are mapped to the implicit vector space with same dimension. The implicit vectors are trained according to the existing scores. The inner product of implicit vector pair is used to predict the unknown ratings.

Compared with the neighbor relationship model, the implicit vector model describes the multidimensional nature of the data more fully. However, the neighbor relationship model is more flexible and easier to integrate with other models, and its recommendation results are more intuitive and comprehensible [19]. Therefore, collaborative filtering techniques based on neighbor relational model have been used widely.

III. Personalized Recommendation System Based on Concept Lattice

This work proposes a novel personalized recommendation system based on concept lattice, which includes offline part and online part, as shown in Fig. 1.

In the offline part, the formal context and the concept lattice

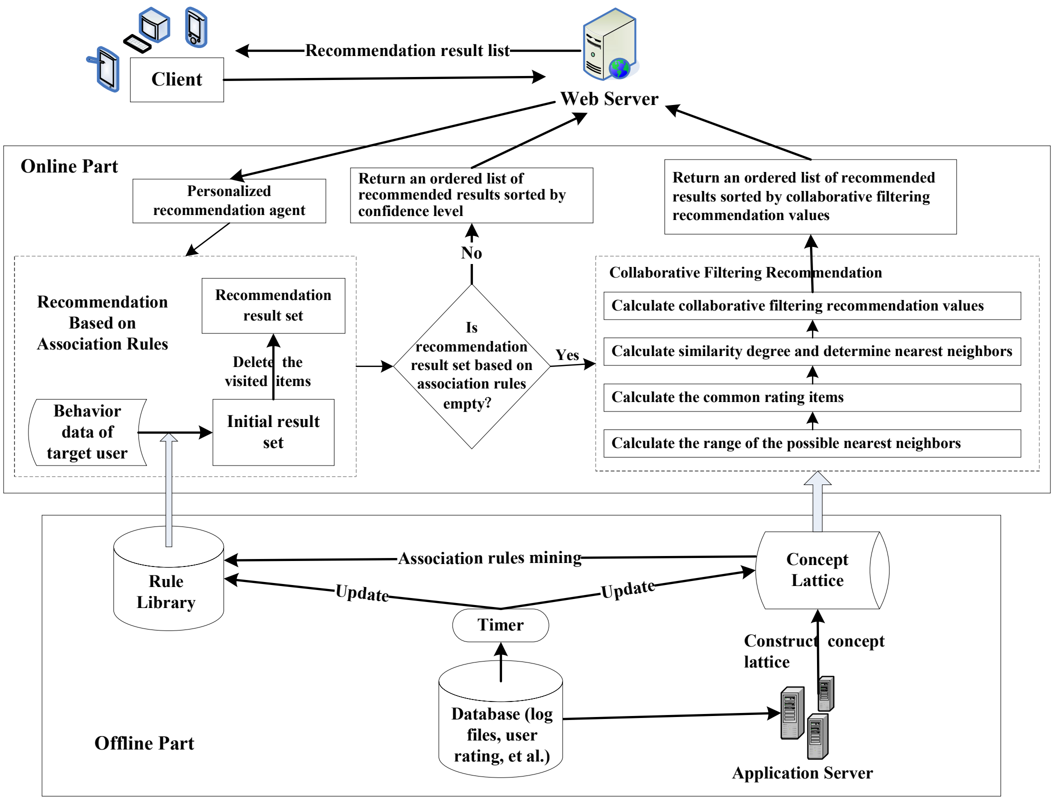


Fig. 1. Personalized recommendation system based on concept lattice.

are constructed from the transaction database (including log files, user-item rating, et al.), and the association rules based on concept lattice are extracted. The obtained association rules are stored in the rule library. The concept lattice and the rule library are updated regularly (such as every 24 hours) using the new added data.

In the online part, using the behavior data of target user, and the concept lattice and the rule library obtained in the offline part, the ordered recommendation result is calculated, and returned to the user. There are two recommendation methods in the online part, which are recommendation based on association rules and collaborative filtering recommendation.

According to the recommendation method based on association rules [21], the behavior data (including current and historical data) of the target user is used to search the rule library. If the user data can match with the antecedents of some association rules, the corresponding consequents of the rules are added into the initial result set. Then the items which the target user has visited are deleted from the initial result set, and the recommendations result set is obtained. If the result set contains multiple recommended items, they will be sorted according to the confidence level, and an ordered list of recommended results is generated to return to the user [21].

If the recommendation result set based on association rules is empty, the collaborative filtering method based on concept lattice can be used. The method uses the concept lattice generated in the offline part to calculate the range of possible nearest neighbors, and their common rating items with the target user. Then the similarity degrees are calculated to determine the nearest neighbors, and calculate the collaborative filtering recommendation values of the target user to items. The recommended items will be sorted according to the collaborative filtering recommendation values, and an ordered list of recommended results is generated to return to the user.

IV. Recommendation Algorithm Based on Association Rules

Take the movie personalized recommendation system for example, the recommendation algorithm based on association rules can be divided into offline part and online part. The offline part constructs the recommendation model using association rule mining algorithm based on concept lattice. The online part provides real-time recommendation service to users according to the established association rules recommendation model and the history records of users watching movies. Because the association rules recommendation model is established in the offline part, the real-time computing complex of the recommendation algorithm can be reduced to accelerate real-time response.

A. Construction of Formal Context and Concept Lattice of User-Movie Rating

The rating of each user to a movie constitutes a record in the database, which describes the detail information of the watched film, user ID and score. These records can reflect the interest of users.

By traversing and analyzing database, the information matrix can be created with the user ID as row, and the rating item as column. In this matrix, users, movies, and binary relations between them constitute the formal context of user-movie rating. Scores reflect the interest level of the user to the corresponding movies. The range of score is 0-10, and the maximum score is 10. The property or item in the formal context can be the name of movie or the type name of the classified movies. Take the rating of users 1-9 to comedy movie for example, the corresponding formal context is shown in Table I.

TABLE I
FORMAL CONTEXT OF USER-MOVIE RATINGS

User	Turkeys (a)	Kung Fu Panda (b)	The Croods (c)	Monsters University (d)	Ratatouille (e)
1	7.1	8.4	8.5		8.1
2	5.5			8.5	
3	6.1	8.5	8.6	8.2	
4		9.2	9.1	9.1	7.4
5		8.2			9.1
6		7.5	9.5	9.2	
7	6.1	9.5		8.1	8.5
8	6.5	8.5	9.2	7.9	7.8
9	5.6	8.1			8.4

If two users give the scores to the same multiple movies, it means that their interests are similar. Thus, the rating matrix of users can be binarization processed to get binary formal context of user-movie rating, as shown in Table II.

TABLE II
BINARY FORMAL CONTEXT OF USER-MOVIE RATING

User	Turkeys (a)	Kung Fu Panda (b)	The Croods (c)	Monsters University (d)	Ratatouille (e)
1	*	*	*		*
2	*			*	
3	*	*	*	*	
4		*	*	*	*
5		*			*
6		*	*	*	
7	*	*		*	*
8	*	*	*	*	*
9	*	*			*

Note: * indicates user rating items; Otherwise, the user does not give scores.

According to the binary formal context of user-movie rating in Table II, the concept lattice can be constructed and the Hasse diagram is generated, as shown in Fig. 2.

Hasse diagram can express the inheritance relationships between the various concept nodes in the concept lattice intuitively and rapidly. The concept node ({1,3,7,8,9}, {*a, b*}) means that the users numbered by 1,3,7,8,9 have watched the film *a* and film *b*, and they have given scores.

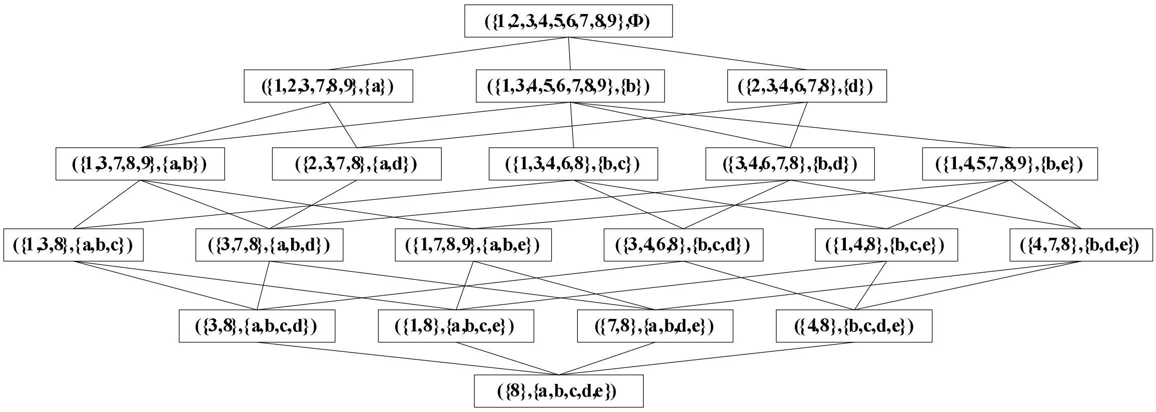


Fig. 2. Hasse diagram of the concept lattice for user-movie rating.

B. Association Rule Mining Based on Concept Lattice

Concept lattice is a natural association rule mining platform, because the rule is described by the relationship between intension sets, and concept lattice reflects the unity of intension and extension of concept. With the support of Hasse diagram of concept lattice, the frequent item sets can be obtained clearly and concisely, and the interesting rules can found according to Hasse diagram.

The transaction database *TD* about users' watching records can be considered as a formal context *K = (U, D, R)*, where *U* is the set of transactions in the database *TD*, *D* is the set of all possible attributes. When $x \in U, d \in D$, *xRd* if and only if *d* belongs to the item set of *x*. According to the corresponding concept lattice of formal context *K*, all the frequent item sets can be calculated.

The association rule $A \Rightarrow B$ corresponds to the unique node two-tuples (C_1, C_2) in the concept lattice, wherein $C_1 = (g(A), f(g(A))), C_2 = (g(A \cup B), f(g(A \cup B)))$. $g(A \cup B)$ is the support transaction set of $A \cup B$. The rule $A \Rightarrow B$ is generated by the node two-tuples (C_1, C_2), and (C_1, C_2) is the generation two-tuples of the rule $A \Rightarrow B$. The support level and confidence level of $A \Rightarrow B$ can be calculated according to its generation two-tuples.

The support level of $A \Rightarrow B$ is $supp(A \Rightarrow B) = |g(A \cup B)| / |U|$, wherein $|g(A \cup B)|$ is the amount of extensions of the concept node C_2. The confidence level of $A \Rightarrow B$ is $conf(A \Rightarrow B) = |g(A) \cap g(B)| / |g(A)|$, wherein $|g(A)|$ is the amount of extensions of the concept node C_1.

If $A \cup B$ is frequent, i.e. $supp(A \Rightarrow B) > \theta$, and $conf(A \Rightarrow B) > \varphi \Leftrightarrow |g(A \cup B)| / |g(A)| > \varphi$, i.e. $|g(A)| < |g(A \cup B)| / \varphi$, the rule $A \Rightarrow B$ is called (θ, φ)-association rule.

In the formal context, if $C_1 = (A_1, B_1)$ and $C_2 = (A_2, B_2)$ satisfy $A_1 \subset A_2$, the relationship between C_1 and C_2 is sub-concept and super concept. The direct sub-concept and super concept relationship between C_1 and C_2 is $C_1 < C_2$, and $\neg \exists C (C_1 < C < C_2)$.

The definition of the generalized super concept $sup(C)$ is as follows [22]:

3) The direct super concept of C is the generalized super concept of C.
4) The generalized super concept of the direct super concept of C is contained in the generalized super concept of C.

In the concept lattice, association rules can be obtained by the following method [22]:

1) If $C_1 = (A_1, B_1)$ and $C_2 = (A_2, B_2)$ satisfy $C_2 \in sup(C_1)$, the association rule $B_2 \Rightarrow B_1 - B_2$ can be obtained, whose confidence level is $|A_1| / |A_2|$. Conversely, the confidence level of $B_1 - B_2 \Rightarrow B_2$ also can be calculated from the concept lattice.
2) If the maximum common sub-concept $C = (A, B)$ of $C_1 = (A_1, B_1)$ and $C_2 = (A_2, B_2)$ is not empty, the association rule $B_1 \Rightarrow B_2, B_2 \Rightarrow B_1$ can be obtained, whose confidence level is $|A| / |A_1|, |A| / |A_2|$ respectively.

Take Fig. 2 for example, it is supposed that the minimum support threshold *MinSup* is 30% and the minimum confidence threshold *MinConf* is 79%, then association rules can be obtained according to the concept lattice of user-movie rating and the above method, as shown in Table III.

C. Personalized Recommendation Based on Association Rules

According to the recommendation method based on association rules, the behavior data (including current and historical data) of the target user is used to search the rule library. If the user data can match with the antecedents of some association rules, the corresponding consequents of the rules are added into the initial result set. Then the items which the target user has visited are deleted from the initial result set, and the recommendations result set is obtained.

TABLE III
ASSOCIATION RULES BASED ON CONCEPT LATTICE

Rule ID	Association Rule	Support Level	Confidence Level
1	$\phi \Rightarrow (b)$	88.9%	88.9%
2	$(a) \Rightarrow (b)$	55.6%	83.3%
3	$(d) \Rightarrow (b)$	55.6%	83.3%
4	$(c) \Rightarrow (b)$	55.6%	100%
5	$(e) \Rightarrow (b)$	66.7%	100%
6	$(a,b) \Rightarrow (e)$	44.4%	80%
7	$(b,c) \Rightarrow (d)$	44.4%	80%
8	$(b,d) \Rightarrow (c)$	44.4%	80%
9	$(c) \Rightarrow (d)$	44.4%	80%

Take the target user 2 for example, the visited items are (*a, d*), which can match with the association rules whose rule ID are (1, 2, 3), and the initial result set generated by the corresponding rule consequents is $\{b\}$. After the visited items of the target user 2 are deleted from the initial result set, the recommended result set $\{b\}$ is obtained. Therefore, the movie *Kung Fu Panda (b)* can be recommended to the target user 2. If the result set contains multiple recommended items, they will be sorted according to the confidence level, and an ordered list of recommended results is generated to return to the user.

Take the target user 9 for example, the visited items are (*a,b,e*), which can match with the association rules whose rule ID are (1, 2, 5, 6), and the initial result set generated by the corresponding rule consequents is $\{b,e\}$. After the visited items of the target user 9 are deleted from the initial result set, the obtained recommended result set is empty. Then the collaborative filtering recommendation method based on concept lattice can be used to calculate the recommendation results.

V. COLLABORATIVE FILTERING RECOMMENDATION METHOD BASED ON CONCEPT LATTICE

In the offline part, the formal context is constructed, which reflects the binary relation between users and items. The corresponding concept lattice is generated and updated regularly. In the online part, the collaborative filtering recommendation method based on concept lattice includes four steps:

1) Calculation of the range of the possible nearest neighbors based on the concept lattice.
2) Calculation of the common rating items of the possible nearest neighbors and the target user.
3) Calculation of similarity degree between the possible nearest neighbors and the target user based on the common rating items, and determination of the target user's nearest neighbors.
4) Calculation of the collaborative filtering recommendation values, and determination of items to be recommended to the target user and the corresponding recommendation strength.

A. Calculation of the Range of the Possible Nearest Neighbors

In order to recommend some movies to the target user, it is necessary to select some users with similar interests as nearest neighbors from those users who have watched these movies. Before the target user's nearest neighbors are determined, the range of the possible nearest neighbors must be calculated.

In the concept lattice, it is necessary to find the concept node with the maximum extension set whose intension set contains the item to be recommended (i.e. the movie to be rated) by the bottom to top search method. The users in its corresponding extension set are likely to recommend the item to the target user, because these users have watched the movie to be recommended and given the score. In these users, those who have larger similarity degree with the target user may become the nearest neighbors.

Since the traversal time complexity of concept lattice is great, it is necessary to adopt pruning strategy. According to the definition of concept lattice, the intension set of super concept is the sub-set of the intension set of sub-concept inevitably. If the item to be recommended does not exist in the intension set of a concept node, it must not exist in all the generalized super concept nodes of this concept node. Then the entire branch tree which consists of this concept node and all of its generalized super concept nodes can be cut off in order to reduce the amount of computation greatly.

Furthermore, when a concept node contains the item to be recommended, according to the depth-first search method, just along its first direct super concept node which includes the item to be recommended, the search towards the top of concept lattice does not be stopped until the concept node with maximum extension set whose intension set contains the item to be recommended is found. The search only needs to deepen towards the top of concept lattice, without backtracking to the bottom.

Take Fig. 2 for example, the steps of calculating the range of the possible nearest neighbors which are used to recommend the movie *Monsters University (d)* to the target user 9 are as follows: the depth-first search from bottom to top is carried out to find that the concept node with maximum extension set whose intension set contains the item *Monsters University (d)* is ($\{2, 3, 4, 6, 7, 8\}$, $\{d\}$). The users (2, 3, 4, 6, 7, 8) in its extension set have given the scores to the item *Monsters University (d),* so these users are possible to recommend the item *(d)* to the target user 9.

B. Calculation of the Common Rating Items

According to the range of the possible nearest neighbors, the common rating items of each possible nearest neighbor and the target user should be calculated. In the concept lattice, it is necessary to find the concept node with the maximum intension set whose extension set contains the possible nearest neighbor and the target user, by the top to bottom search method. Its corresponding intension set contains the common rating items of these two users.

Since the traversal time complexity of concept lattice is great, it is necessary to adopt pruning strategy. According to the definition of concept lattice, the extension set of sub-concept is the sub-set of the extension set of super concept inevitably. If the possible nearest neighbor and the target user do not coexist in the extension set of a concept node, they must not coexist in all the sub-concept nodes of this concept node. Then the entire branch tree which consists of this concept node and all of its sub-concept nodes can be cut off in order to reduce the amount of computation greatly.

Furthermore, when a concept node contains the possible nearest neighbor and the target user, according to the depth-first search method, just along its first sub-concept node which contains these two users, the search towards the bottom of concept lattice does not be stopped until the concept node with maximum intension set whose extension set contains these two users is found. The search only needs to deepen towards the bottom of concept lattice, without backtracking to the top.

Take Fig. 2 for example, the steps of calculating the common rating items of the possible nearest neighbor user 7 and the target user 9 are as follows: the depth-first search from top to bottom is carried out to find that the concept node with maximum intension set whose extension set contains these two user {7, 9} is ({1, 7, 8, 9}, {*a, b, e*}). Then the common rating items of the possible nearest neighbor user 7 and the target user 9 are (*a, b, e*). The common rating items of each possible nearest neighbor and the target user 9 can be obtained in a similar way, as shown in Table IV.

TABLE IV
THE COMMON RATING ITEMS OF EACH POSSIBLE NEAREST NEIGHBOR AND THE TARGET USER 9

User	Turkeys (a)	Kung Fu Panda (b)	Ratatouille (e)
2	*		
3	*	*	
4		*	*
6		*	
7	*	*	*
8	*	*	*

Note: * indicates the common rating item of this user and the target user 9; Otherwise, it is not the common rating item.

C. Calculation of Similarity Degree and Determination of the Nearest Neighbors

By calculating the similarity degree between users, the nearest neighbors can be selected. The similarity degree can be calculated according to the users' scores to attribution items. According to the common rating items in Table IV, the scores of the common rating items of each possible nearest neighbor and the target user 9 can be obtained, as shown in Table V.

According to the scores of the common rating items of each possible nearest neighbor and the target user 9 in Table V, the corresponding similarity degree can be calculated. There are many methods to calculate similarity degree. This work adopts the adjusted cosine similarity calculation method [23] proposed by Sarwar et al. In the collaborative filtering recommendation method based on users, the adjusted cosine similarity degree between user *u* and user *v* is calculated as follows:

$$sim(u,v)=\frac{\sum_{i\in I_{uv}}(R_{u,i}-\overline{R_u})(R_{v,i}-\overline{R_v})}{\sqrt{\sum_{i\in I_u}(R_{u,i}-\overline{R_u})^2}\sqrt{\sum_{i\in I_v}(R_{v,i}-\overline{R_v})^2}} \tag{1}$$

Wherein, i and j are attribution items; $R_{u,i}$ is the score of user u to item i; I_u is the rated item set of user u; I_{uv} is the common rating item set of user u and user v (I_{uv}=$I_u \cap I_v$); $\overline{R_u}$ is the average score of user u.

TABLE V
SCORES OF THE COMMON RATING ITEMS OF EACH POSSIBLE NEAREST NEIGHBOR AND THE TARGET USER 9

User	Turkeys (a)	Kung Fu Panda (b)	Ratatouille (e)
2	5.5, 5.6		
3	6.1, 5.6	8.5, 8.1	
4		9.2, 8.1	7.4, 8.4
6		7.5, 8.1	
7	6.1, 5.6	9.5, 8.1	8.5, 8.4
8	6.5, 5.6	8.5, 8.1	7.8, 8.4

Note: the first value in each score pair is the score of the possible nearest neighbor, and the second value is the score of the target user 9.

Take the user 7 for example, the adjusted cosine similarity degree between the user 7 and user 9 can be calculated by using (1). I_7={a,b,d,e} ; I_9={a,b,e} ; I_{79}=$I_7 \cap I_9$ = {a,b,e} ; $R_{7,a}=6.1$; $R_{7,b}=9.5$; $R_{7,d}=8.1$; $R_{7,e}=8.5$; $R_{9,a}=5.6; R_{9,b}=8.1; R_{9,e}=8.4$;

$\overline{R_7}=(6.1+9.5+8.1+8.5)/4=8.05$;

$\overline{R_9}=(5.6+8.1+8.4)/3=7.37$;

$$sim(7,9)=\frac{\sum_{i\in\{a,b,e\}}(R_{7,i}-\overline{R_7})(R_{9,i}-\overline{R_9})}{\sqrt{\sum_{i\in\{a,b,d,e\}}(R_{7,i}-\overline{R_7})^2}\sqrt{\sum_{i\in\{a,b,e\}}(R_{9,i}-\overline{R_9})^2}}=0.925$$

Likewise, $sim(2,9)=0.125$; $sim(3,9)=0.831$; $sim(4,9)=-0.539$; $sim(6,9)=-0.271$; $sim(8,9)=0.656$.

The similarity degrees between the users (4, 6) with the target user 9 are negative, so they are not chosen as the nearest neighbors. The users (2, 3, 7, 8) are selected as the nearest neighbors. Then according to the scores of the nearest neighbors to item *Monsters University (d)*, the possible score of the target user 9 to this item can be predicted, and the strength of recommending the movie *Monsters University (d)* to the target user 9 can be calculated.

D. Calculation of Collaborative Filtering Recommendation Values

After the target user's nearest neighbors and their similarity degrees are determined, the collaborative filtering recommendation values of the target user to some items can be calculated. The K-nearest neighbor method is often used to calculate the collaborative filtering recommendation value.

This method chooses K users with the highest similarity degrees as the target user's nearest neighbors.

Let N_u expresses the nearest neighbor set of user u, and $R(u,i)$ expresses the predicted score of user u to item i, then $R(u,i)$ can be calculated by using (2) [24].

$$R(u,i) = \overline{r_u} + \frac{\sum_{v \in N_u} \mathrm{sim}(u,v)(r_{v,i} - \overline{r_v})}{\sum_{v \in N_u} \mathrm{sim}(u,v)} \quad (2)$$

$sim(u,v)$ is the similarity degree between user u and user v. $r_{v,i}$ is the score of user v to item i. $\overline{r_u}$ is the average score of user u to items. $\overline{r_v}$ is the average score of user v to items.

The collaborative filtering recommendation value of the target user 9 to item *Monsters University (d)* can be calculated by using (2).

$$R(9,d) = \overline{r_9} + \frac{\sum_{v \in \{2,3,7,8\}} \mathrm{sim}(9,v)(r_{v,d} - \overline{r_v})}{\sum_{v \in \{2,3,7,8\}} \mathrm{sim}(9,v)} = 7.556$$

According to the above calculation result, the collaborative filtering recommendation value (i.e. the predicted score) of the target user 9 to item *Monsters University* (*d*) is approximately 7.556. The collaborative filtering recommendation values of the target user to other items can be calculated by using the same method. Then the items with the maximal collaborative filtering recommendation values should be recommended to the target user. This method can provide a more efficient collaborative filtering recommendation service.

VI. CONCLUSION

This work proposes a novel personalized recommendation system based on concept lattice, which contains offline part and online part.

In the offline part, the formal context and the concept lattice are constructed from the transaction database, and the association rules based on concept lattice are extracted. The obtained association rules are stored in the rule library. The concept lattice and the rule library are updated regularly using the new added data.

In the online part, by using the behavior data of target user, and the concept lattice and the rule library obtained in the offline part, the ordered recommendation result is calculated, and returned to the user. There are two recommendation methods in the online part, which are recommendation based on association rules and collaborative filtering recommendation.

According to the recommendation method based on association rules, the behavior data of the target user is used to search the rule library to obtain the recommendations result set. If the result set contains multiple recommended items, they will be sorted according to the confidence level, and an ordered list of recommended results is generated to return to the user.

If the recommendation result set based on association rules is empty, the collaborative filtering method based on concept lattice can be used. This method uses the concept lattice generated in the offline part to calculate the range of possible nearest neighbors, and their common rating items with the target user. Then the similarity degrees are calculated to determine the nearest neighbors, and calculate the collaborative filtering recommendation values of the target user to items. The recommended items will be sorted according to the collaborative filtering recommendation values, and an ordered list of recommended results is generated to return to the user.

The complexity of concept lattice construction and update is very high, so it is necessary to further study the efficient concept lattice construction and update algorithm in the future. The association rule library also need to constantly update and maintain, so it is also necessary to further study the association rule library update algorithm based on concept lattice.

REFERENCES

[1] G. Adomavicius and A. Tuzhilin, "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, pp. 734-749, 2005.

[2] R. Wille, "Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts," in *Ordered Sets*. vol. 83, I. Rival, Ed., ed: Springer Netherlands, 1982, pp. 445-470.

[3] B. Ganter and R. Wille, "Formal concept analysis: mathematical foundations," ed. Berlin: Springer, 1999.

[4] Q. Yang and M. Zhao, "Progress in concept lattice research," *Computer Engineering and Design,* vol. 29, pp. 5293-5296, 2008.

[5] Y. Cai, N. Cercone, and J. Han, "An attribute-oriented approach for learning classification rules from relational databases," in *Proceedings of Sixth International Conference on Data Engineering*, 1990, pp. 281-288.

[6] R. Brüggemann and G. P. Patil, "Formal Concept Analysis," in *Ranking and Prioritization for Multi-indicator Systems*, ed New York: Springer New York, 2011, pp. 117-133.

[7] L. Lambrechts, "Formal Concept Analysis," Master, Vrije Universtiteit, Brussel, 2012.

[8] C. Zou, J. Wan, M. Chen, and D. Li, "Simulation modeling of cyber-physical systems exemplified by unmanned vehicles with WSNs navigation," in *Proc. of the 7th Int. Conf. on Embedded and Multimedia Computing Technology and Service*, Gwangju, Korea, 2012, pp. 269-275.

[9] J. Wan, M. Chen, F. Xia, D. Li, and K. Zhou, "From Machine-to-Machine Communications towards Cyber-Physical Systems," *Computer Science & Information Systems,* vol. 10, pp. 1105-1128, 2013.

[10] M. Chen, J. Wan, S. Gonzalez, X. Liao, and V. C. M. Leung, "A Survey of Recent Developments in Home M2M Networks," *IEEE Communications Surveys & Tutorials,* vol. 16, pp. 98-114, 2014.

[11] K. Hu, Y. Lu, and C. Shi, "An Integrated Mining Approach for Classification and Association Rule Based on Concept Lattice," *Journal of Software,* vol. 11, pp. 1478-1484, 2000.

[12] C. Ma, X.-D. Gao, and Z.-W. Zeng, "A Spatial Data Mining Method Based on the Concept Lattice of Compact Dependencies," *International Review on Computers & Software,* vol. 8, pp. 341-346, 2013.

[13] R. Godin and R. Missaoui, "An incremental concept formation approach for learning from databases," *Theoretical computer science,* vol. 133, pp. 387-419, 1994.

[14] R. Missaoui, R. Godin, and A. Boujenoui, "Extracting exact and approximate rules from databases," in *Incompleteness and Uncertainty in Information Systems*, ed London: Springer-Verlag, 1994, pp. 209-222.

[15] K. Hu, Y. Lu, L. Zhou, and C. Shi, "Integrating classification and association rule mining: A concept lattice framework," in *New Directions in Rough Sets, Data Mining, and Granular-Soft Computing*, ed: Springer, 1999, pp. 443-447.

[16] Z. Xie and Z. Liu, "Concept lattice-based association rule discovery," *Mini-micro Systems,* vol. 21, pp. 1028-1031, 2000.

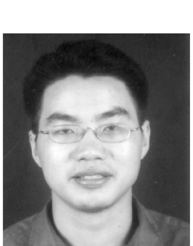
[17] Y. Li, T. Li, J. Cai, and L. Chen, "Extracting Succinct Association Rules Based on Concept Lattice," *Journal of Nanjing University of Posts and Telecommunications(Natural Science),* vol. 27, pp. 44-47, 2007.

[18] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry, "Using collaborative filtering to weave an information tapestry," *Communications of the ACM,* vol. 35, pp. 61-70, 1992.
[19] Y. Koren, "Factor in the neighbors: Scalable and accurate collaborative filtering," *ACM Transactions on Knowledge Discovery from Data,* vol. 4, pp. 1-24, 2010.
[20] M. Kurucz, A. A. Benczúr, and K. Csalogány, "Methods for large scale SVD with missing values," in *KDD Cup Workshop at Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, California, 2007, pp. 31-38.
[21] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Analysis of recommendation algorithms for e-commerce," presented at the Proceedings of the 2nd ACM conference on Electronic commerce, Minneapolis, Minnesota, USA, 2000.
[22] D. Wang, X. Hu, and H. Wang, "Algorithm of mining association rules based on quantitative concept lattice," *Journal of Hefei University of Technology (Natural Science),* vol. 25, pp. 678-682, 2002.
[23] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th international conference on World Wide Web*, Hong Kong, 2001, pp. 285-295.
[24] J. S. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," presented at the Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence, Madison, Wisconsin, 1998.

Caifeng Zou (M'13) received her MS in Computer Application Technology from Sun Yat-sen University in Guangzhou in 2008, and her BS in Computer Science and Technology from Shanghai University in Shanghai in 2006. She is currently pursuing her PhD in Computer Software and Theory from South China University of Technology in Guangzhou. She is currently a lecturer in Guangdong Mechanical & Electrical College in Guangzhou. Her current research interests include data mining, cloud computing, and internet of things.

This author became a Student Member (M) of IEEE in 2013.

Daqiang Zhang received his B.Sc. degree in management science and M.Sc. degree in computer science from Anhui University in 2003 and 2006, and his Ph.D. degree in computer science from Shanghai Jiao Tong University in 2010. He has published more than 60 papers in major journals and international conferences. As the first author, he has published papers in IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Computers, IEEE Transactions on Emerging Topics in Computing, IEEE Wireless Communications, IEEE Network, IEEE Systems Journal, IEEE ICPP, IEEE ICC and IEEE WCNC. His research includes mobile computing, distributed computing, and wireless sensor networks. Currently, he is an associate professor in the School of Software Engineering at Tongji University, China.



Jiafu Wan is an Associate Professor in School of Mechanical & Automotive Engineering, South China University of Technology (SCUT), China. He received the Ph.D. degree in Mechatronic Engineering from SCUT in Jun 2008. From Oct 2008 to Jun 2012, he was a Post-doctor of Computer Science and Engineering in SCUT. He is also managing editor for IJAACS (EI) and IJART (EI), and workshop chair of M2MC2012, M2MC2013, and MCC2013. He has authored/ co-authored one book and more than 60 scientific papers. His research interests include cyber-physical systems, internet of things, cloud computing, and embedded systems. He is a CCF senior member, and a member of IEEE.

Jaime Lloret received his M.Sc. in physics in 1997, his M.Sc. in electronic engineering in 2003, and his Ph.D. in telecommunication engineering (Dr.Ing.) in 2006. He is currently an associate professor in the Polytechnic University of Valencia. He is currently Chair of the Internet Technical Committee (IEEE Communications Society and Internet Society). He has authored 12 books, and has had more than 240 research papers published in national and international conferences, an international journals (more than 80 with ISI Thomson Impact Factor). He has been Co-Editor of 15 conference proceedings and Guest Editor of several international books and journals. He is Editor-in-Chief of the international journal Networks Protocols and Algorithms, IARIA Journals Board Chair (eight Journals), and an Associate Editor of several international journals. He is an IARIA Fellow.