Context-Aware Personalized Recommendation Methods for Mobile Digital Library Resources: Postprint

Authors: Hong Liang, Qian Chen, Fan Xing

Date: 2017-11-08T00:00:00+00:00

Abstract

【Purpose】Integrating the resource layout and push characteristics of mobile digital libraries, this study designs a context-aware personalized resource recommendation approach.

【Methodology】Based on the consideration that similar users have similar choices, we introduce the concept of role to simulate user interest selection, design an effective WSSQ algorithm to construct a user trust network, and thereby propose an improved context-aware recommendation method. Simulation experiments are conducted on an extended Epinion dataset.

【Results】Experiments demonstrate the feasibility of the proposed method, which outperforms other recommendation methods under metrics such as MAE and RMSE, exhibiting good recommendation accuracy.

【Limitations】When the user sample is sufficiently large, the method faces the problem of context and role sparsity.

【Conclusion】This study provides an approach for resource recommendation in mobile digital libraries, which is conducive to the improvement and refinement of their recommendation systems.

Full Text

Preamble

Research on Context-Aware Personalized Recommendation Methods for Mobile Digital Library Resources

Hong Liang, Qian Chen, Fan Xing
(School of Information Management, Wuhan University, Wuhan 430072, China)

Abstract

[**Objective**] This study designs a context-aware personalized resource recommendation approach for mobile digital libraries, considering their resource distribution and push notification characteristics. [**Methods**] Based on the principle that similar users make similar choices, we introduce the concept of "roles" to model user interest selection, design an effective WSSQ algorithm to construct a user trust network, and propose an improved context-aware recommendation method. We conduct simulation experiments on an extended Epinions dataset. [**Results**] Experiments demonstrate the feasibility of our method, which outperforms other recommendation approaches in terms of MAE, RMSE, and other metrics, showing good recommendation accuracy. [**Limitations**] When user samples are sufficiently large, the method faces sparsity issues in contexts and roles. [**Conclusions**] This research provides a new perspective for mobile digital library resource recommendation, contributing to the improvement and refinement of their recommendation systems.

Keywords: Mobile digital library; Context-aware recommendation; WSSQ algorithm

1. Introduction

The deepening of social informatization and digitization has created an explosive information environment, while the proliferation of mobile terminals exposes users to countless fragmented and disordered information streams every moment. The contradiction between rapidly expanding data volume and limited user capacity for data acquisition has reduced the efficiency of active search and retrieval, representing a concentrated manifestation of the current information overload problem. To adapt to this environment of information overload and better serve users, libraries—as social information hubs—have begun evolving toward mobile digital libraries. Mobile digital libraries represent a new development of libraries in the mobile internet environment, primarily enabling unified search functionality for various library resources on handheld devices and satisfying users' needs to access and read literature anytime and anywhere.

Resource availability and authority serve as key elements in evaluating mobile digital library resources, with user utilization rates being the most 朴素 yet practical manifestation of these evaluation methods. Whether users can accurately obtain resource access paths, whether their reading intentions are satisfied, and whether they have a superior user experience are all important metrics for evaluating mobile digital library services. In an environment where handheld mobile terminals have become the primary source of information reception, library users are constantly influenced by rich and continuously changing contextual information, leading to different resource needs or reading preferences. Therefore, how to dynamically meet users' changing needs constitutes a major challenge for mobile digital libraries in the current situation.

Personalized recommendation systems offer an effective solution to these problems. Recommendation systems input user information and employ recommendation algorithms to produce reasonable recommendations. By eliminating the need for users to search independently, personalized recommendations focus on collecting user interests and making targeted recommendations, thereby better satisfying user needs. If we consider how user interests change under the influence of contextual factors—that is, context-aware recommendations—we can provide an answer to the problem of changing user needs in mobile digital libraries.

The academic community generally agrees that mobile digital libraries should emphasize innovative development of traditional service models rather than simply implementing networked information services. How to achieve innovation depends on examining the actual environment and needs. Mobile digital libraries rely on the rich and orderly collection resources of physical libraries, which are of relatively high quality. Sheng Xiaoping considers resource availability, authority, and user utilization rates as essential elements for evaluating mobile digital library resources, with user utilization being the most 朴素 and practical manifestation of these evaluation approaches.

2. Related Research Progress

Research on context-aware recommendation and user trust networks has always been an important component of recommendation system studies, attracting significant scholarly attention and yielding numerous achievements.

Forestier et al.' s survey results show that user roles widely exist in social networks and continuously evolve with changing contexts. Adomavicius et al. discuss how to simulate contexts in recommendation systems, while Anand et al. propose new ways to incorporate user contextual information into recommendation methods. Bao et al. present an unsupervised approach to model personalized context information for mobile users. Sen et al.' s recommendation system incorporates contextual information for context-aware recommendation. Karatzoglou et al. propose a Multiverse Recommendation (MV) method that combines context-aware and collaborative filtering approaches, while Rendle et al. propose using Factorization Machines (FM) to simulate contextual information, thereby improving recommendation accuracy and reducing algorithmic complexity compared to the MV method. However, these existing methods do not consider the context-aware interests and selection tendencies shared by user groups, nor the trust relationships among them.

Trust-based recommendation often assumes the existence of a trust network among users. Massa et al. study trust propagation in trust networks and propose a trust-aware recommendation approach. Jamali et al. propose a random walk model called TrustWalker to combine trust-based and collaborative filtering methods for recommendation. They also present an innovative Top-N item recommendation method that uses a weighted combination of results from both trust-based and collaborative filtering approaches. Dubois et al. employ trust paths to cluster users into trust groups, thereby improving recommendation accuracy. Ma et al. propose a recommendation framework that integrates personal preferences with those of trusted friends.

In summary, current research on context-aware recommendation primarily suffers from two problems: user interest models that are divorced from reality, and the neglect of user trust relationships. Specifically, existing context-aware recommendation systems examine users' common interests based on the current context, but in reality, users often share the same information preferences across interconnected contexts. Typical context-aware recommendation systems assume user decision-making independence while ignoring relationships and interactions among users, though Massa et al.' s research has demonstrated that user decisions are significantly influenced by their trusted counterparts. Based on this, we introduce the concept of roles to represent users' common interests across multiple contexts and use roles to find similar users and build trust models. On the other hand, trust-based recommendation research also contains unrealistic assumptions, such as believing that trust between users does not change with context, or that all trust relationships in social trust networks are explicitly expressed. Existing methods do not consider how trust relationships manifest in different contexts, which reduces recommendation quality. Therefore, this paper uses a trust model to calculate implicit trust relationships among users to enrich the trust network and incorporates context as a parameter in the calculation of user trust values and similarity to reflect the impact of contextual changes on user trust relationships.

3. Overview of the Recommendation Scheme

The proposed recommendation scheme consists of online and offline components, as shown in [Figure 1: see original paper], and includes three specific steps.

[Figure 1: see original paper]

Step 1: Use a role mining algorithm to identify roles from the UCB matrix and employ a random walk model to measure the association strength between each role and different contexts. This work will be completed offline. The UCB matrix is the "User-Context-Behavior" matrix, where rows represent users and columns represent contexts, with each element indicating a user' s selection behavior in the corresponding context. This paper uses users' preferences and choices for items as manifestations of user behavior. These items are described by content tags such as "economic commentary" and "economics," which are derived from the resource classification tags adopted by mobile digital libraries. The recommendation results are mobile digital library collection resources labeled with corresponding classification tags. In this paper, context is defined as a collection of spatiotemporal situations, including factors such as location and time that influence user behavior.

Step 2: Construct a role-based trust model to calculate similarity between users. An effective WSSQ algorithm will be used to find all similar users for user u. This process will be completed online.

Step 3: Based on the role-based trust model established for each user, predict user ratings for different items and recommend items with high ratings to users. This process will also be completed online.

Research by Golbeck and McPherson et al. demonstrates that identity is universal in social networks and can enhance mutual trust among users; that is, the higher the similarity between users, the stronger the trust relationship between them. This paper constructs a role-based trust network by calculating implicit trust values. This trust network is context-aware and personalized, with user trust values and trusted counterparts changing with context. Then, based on such a trust network under context c and the UIR matrix, recommendations are completed through two steps: determining each item as a user behavior choice, and predicting unknown ratings for item v through ratings given by users in the role-based trust network. Items are sorted in descending order of rating results, and the top k items are recommended to users. The UIR matrix is the "User-Item-Rating" matrix that records users' ratings for each item, reflecting user preferences for those items.

4.1. Role Mining

In the role mining method, a role is defined as the common behavior set of a user group in a specific context. Therefore, using roles to represent users' common interests is feasible, and discussing user similarity based on roles can also align with reality. For example, in [Figure 2: see original paper], users 1, 2, and 5 make the same choices in the two contexts of "subway" and "study room," while users 2, 3, and 4 share the same behavioral tendencies in the two contexts of "dormitory/home" and "CD store." Thus, we can understand that users 1, 2, and 5 enact role r_1 in the "subway" and "study room" contexts, while users 2, 3, and 4 enact role r_2 in the "dormitory/home" and "CD store" contexts.

More precisely, we introduce the definition of a *tile* to represent any submatrix in the UCB matrix where elements in the same column are identical, thereby defining a role as a tile containing at least m rows and n columns. Users' decisions are often influenced by their roles. For example, the choices of users 2, 3, and 4 in both "dormitory/home" and "CD store" contexts are related to pop music, so r_2 can be regarded as pop music enthusiasts. Therefore, users 2, 3, and 4 may be interested in pop music-related content when in "dormitory/home" and "CD store" contexts. The recommendation method proposed in this paper will focus on behaviors shared by at least m users in at least n contexts, where the values of n and m depend on specific applications.

In existing Role-Based Access Control (RBAC) systems, role mining algorithms have been designed for Boolean matrices, but these algorithms cannot be directly applied to the non-Boolean role matrix in this paper. In fact, the role mining problem in the UCB matrix can be equivalent to the Minimal Tiling problem, which has been proven to be NP-complete. The Minimal Tiling problem aims to find the minimum number of tile elements that cover all non-empty elements in the UCB matrix. This paper proposes a learning heuristic role mining algorithm using Conditional Database.

This algorithm initializes the role set ROLES as empty, then constructs a c -conditional database for each column c of the UCB matrix. For each c , it considers all c satisfying $1 \leq j \leq n$ and finds all tiles with at least m rows (denoted as t) in each pair (c , c). For each found t, it checks whether there exists t' ROLES such that t and t' can be merged (two tiles can be merged if and only if they have the same rows). If such t' exists, t and t' are merged; otherwise, t is directly added to ROLES. The final result is the set ROLES composed of tiles with at least n columns.

For example, consider the UCB matrix in [Figure 3: see original paper], extracted from [Figure 2: see original paper], where the first column represents 5 users, the first row represents 4 contexts, and m to n represent "economic commentary," "economics," "advanced guitar," and "modern metal rock," respectively. Assuming both m and n are 2, we construct c -conditional databases, such as c-cdb, remove all tiles with fewer than 2 columns, and obtain ROLES that meet the requirements. Based on ROLES, the UCB matrix is decomposed into a Role-Context-Behavior (RCB) matrix and a User-Role (UR) matrix. The RCB matrix has roles as rows and contexts as columns, with each element representing the behavior of the corresponding role in the corresponding context. The UR matrix has users as rows and roles as columns, with each element indicating whether the corresponding user enacts the corresponding role (Boolean value).

After role mining, each user can be represented by their role set $R(u,\ c) = \{r_1,\ \cdots,\ r\}$, where c is the current context and r_1 to r represent roles enacted by user u in context c. In the RCB matrix, each role corresponds to different context-aware behaviors. Accordingly, if a new user u' exhibits the same context-aware behavior as role r in the RCB matrix, then u' enacts role r. Moreover, if u' ' s behavior changes over time, their role set will be updated promptly. Additionally, there may exist users who do not enact any roles; these users have interests and tendencies in specific contexts that do not align with any user group, but they still possess a set of behaviors across different contexts. For such users, we can find similar user groups based on common context-aware behaviors and measure similarity between this user and others using behavioral set similarity.

After role mining, we calculate role weights in corresponding contexts, i.e., measuring the association strength between role r and context c, using a random walk model and denoted as $W(r,\ c)$. Obviously, if user u enacts role r in context c, then $W(r,\ c)$ should be larger than the weight values of their other roles.

[Figure 4: see original paper]

As shown in [Figure 4: see original paper], we construct a bipartite graph G_1 between the user set and role set, connecting a user and a role if and only if the user' s role set contains that role. We construct another bipartite graph G_2 between the role set and context set, connecting a role and a context if and only if the role covers that context in the UCB matrix. G_1 and G_2 combine to form graph G. *W(r, c)* is measured by a random walk with restart model, i.e., the steady-state probability that a particle starting from node r finally stops at node c.

4.2. Role-Based Trust Model

In reality, trust relationships between users manifest in both explicit and implicit ways, but explicit trust is mostly reflected in communication, comments, etc., and such data is relatively sparse, making it impossible to accurately measure user trust. Therefore, we construct a user trust model to solve the problem of calculating implicit trust values between user u and user *u*' , which is not Boolean and is denoted as *T(u, u' , c)*. User similarity and user interaction are key to model construction.

This paper proposes a context-aware user similarity measurement based on users' role sets. Literature [19] proposes the Jaccard similarity formula, defining the similarity between two elements s and r as the ratio of the L1-norm of their intersection to the L1-norm of their union. This paper improves upon the Jaccard similarity formula to simulate user similarity. Let user u and user *u*' in context c have role sets *R(u, c)* and *R(u' , c)*, respectively. Then user similarity can be expressed as:

$$S(u, u', c) = \frac{\sum_{r \in R(u,c) \cap R(u',c)} W(r, c)}{\sum_{r \in R(u,c) \cup R(u',c)} W(r, c)}$$

Next, we consider user interaction. Liu et al. [20] argue that user interactions in social networks, such as chatting and commenting, affect mutual trust. A user is often more willing to trust information provided by another user with whom they have communicated. In our model, we use *I_{uu' }* to represent the number of interactions between user u and user *u*' . We simulate the trust value between users u and *u*' as the conditional probability that *u*' becomes *u*' s trust object, proposing the following formula to express trust:

$$T(u, u', c) = P(F(u, u', c) = 1 \mid S(u, u', c), I_{uu'})$$

F(u, u' , c) = 1 means that user *u*' is user *u*' s trust object in context c.

Due to its advantages of accommodating both continuous and discrete variables and higher efficiency in large datasets, this paper selects the Logistic regression model to simulate trust relationships. Expanding formula (2) according to the Logistic regression model, we examine the influence of *S(u, u' , c)*, *I_{uu' }*, and *I_{u' u}* on *T(u, u' , c)*:

$$T(u,u',c) = P(F(u,u',c)=1 \mid S(u,u',c), I_{uu'}) = \frac{1}{1+e^{-\beta_0-\beta_1\cdot S(u,u',c)-\beta_2\cdot I_{uu'}-\beta_3\cdot I_{u'u}}}$$

β_0, β_1, β_2, β_3 are trained parameters. This paper uses explicit trust relationships as training data to obtain parameter values. First, we propose the following expectation function based on the binomial distribution concept:

$$P(u,u',c) = F(u,u',c)\cdot(1-P(u,u',c))^{1-F(u,u',c)}$$

P(u, u' , c) represents the probability that *F(u, u' , c) = 1*. Taking the logarithm of both sides of formula (4), we obtain the log-likelihood function:

$$L = \sum_{u,u'\in D}[F(u,u',c)\ln P(u,u',c) + (1-F(u,u',c))\ln(1-P(u,u',c))]$$

$$= \sum_{u,u'\in D}[(F(u,u',c)-1)(\beta_0+\beta_1\cdot S(u,u',c)+\beta_2\cdot I_{uu'}+\beta_3\cdot I_{u'u})-\ln(1+e^{-\beta_0-\beta_1\cdot S(u,u',c)-\beta_2\cdot I_{uu'}-\beta_3\cdot I_{u'u}})]$$

Next, we use a gradient-based algorithm to estimate β_0, β_1, β_2, β_3 to maximize *L*. Setting the derivatives of *ln L* with respect to the parameters to zero, we then apply the Newton-Raphson iterative method to estimate the parameters.

4.3. Finding Similar Users

As can be seen from formula (3), user similarity is a key factor in the trust model. By finding similar users *u'* for user *u* in context *c* and calculating similarity *S(u, u' , c)*, we can obtain the trust value *T(u, u' , c)*. In the online recommendation process, the first problem to solve is finding similar users. Specifically, given context *c*, user *u*, and similarity threshold δ, we need to find similar users *u'* for user *u* such that the similarity between *u* and *u'* satisfies *S(u, u' , c)* $\geq \delta$. This paper adopts the WSSQ (Weighted Set Similarity Query) algorithm to solve this problem.

(1) WSSQ Algorithm

The WSSQ algorithm takes an ordered role set and similarity threshold δ as input, and outputs all similar users meeting the condition of similarity greater than the threshold along with their corresponding similarity values. Roles in the set are arranged in reverse order according to context-aware weights, while the similarity threshold δ is trained from the dataset, with its optimal value being 0.6. Experiments prove that when this default value is input, trust measurement accuracy reaches a peak of 89.1%. The algorithm is completed in two parts: offline and online, as illustrated in [Figure 5: see original paper].

[Figure 5: see original paper]

In the offline part, we first establish an inverted index for all roles of user u according to role weights, as shown in the Query part of Figure 5: see original paper. As mentioned earlier, roles represent user interest tendencies and continuously change with context. Therefore, role weight values are also context-dependent. As shown in Figure 5: see original paper, under contexts c_1 and c_2, each role has different weight values. Thus, there are as many reverse-ordered role sequences as there are contexts. When the system generates k contexts c (i = 1, ···, k), for each context, k reverse-ordered user lists *L (r)* are retained under each role, and each user u in *L (r)* corresponds to their respective role set *R(u, c)*, allowing users and role sets to be interchangeably expressed in the analysis. In *L (r)*, all roles in each role set *R(u, c)* are arranged in reverse order according to their weights under context c , and then the role sets in each reverse-ordered list L_1*(r)* are arranged in ascending order by their L1-norm. As shown in Figure 5: see original paper, role B generates reverse-ordered list L_1*(r)* under context c_1, containing 5 users who enact role B. User 1' s complete role set is {B, E, F}, and these three roles are presented in reverse order according to their weight values under context c_1.

In the online part, when given context c and user u, we use two filtering techniques—Prefix filtering and L1-norm pruning—to find all similar users *u'* . In the Prefix filtering process, we identify inverted sequences that meet the conditions. In the L1-norm pruning process, we eliminate some unnecessary role sets from the identified inverted sequences to save search space. Through these steps, we can calculate similarity *S(u, u' , c)* for each similar user *u'* of user *u*.

(2) Prefix Filtering

Given context c and user u in this context, all roles of u are arranged in reverse order of weight to form a Ranked Role Set, denoted as *R(u, c)*. The *p*-prefix is the first p roles in *R(u, c)*, denoted as *R(u, c, p)*. For users u and *u'* , if *u'* does not possess any role in *R(u, c, p)*, we propose the following formula to represent the Upper Bound of *S(u, u' , c)*:

$$UB\text{-}P(u, u', c) = 1 - \frac{\sum_{r \in R(u,c,p)} W(r, c)}{\sum_{r \in R(u,c)} W(r, c)}$$

The principle of Prefix filtering is: given context c and user u, for another user *u'* , if *R(u' , c)* does not cover *R(u, c)* on the *p*-prefix, then the Upper Bound of *S(u, u' , c)* can be calculated through formula (6). Based on this, we can find the maximum *p*-prefix of *R(u, c)*. The maximum *p*-prefix means: if user *u'* ' s roles cover *R(u, c, p-1)*, then *UB-P(u, u' , c)* $\geq \delta$; if not covered, then *UB-P(u, u' , c)* $< \delta$. Similarly, for any role in *R(u, c, p)*, if *u'* ' s role set is not included in any *L(r)*, then *UB-P(u, u' , c)* must be less than δ. Therefore, to find all similar role sets, we only need all inverted sequences *L(r)* that satisfy *r R(u, c, p)*.

(3) L1-Norm Pruning

Given user u in context c, we can find its maximum p-prefix $\{r_1, \cdots, r\}$ according to Prefix filtering. Then we input inverted sequences *L(r)* one by one. Since a user may appear in different inverted sequences, we adopt the following measure: if a user' s role set *R(u, c)* in *L(r)* has a role r satisfying $1 \leq j \leq$ *i-1*, then *R(u, c)* will not be input because it was already input when accessing *L(r)*. Additionally, when accessing *L(r)*, some role sets can be eliminated using the Upper Bound method of similarity.

Given ordered role set *R(u, c) = {*r_1*, ···, r }* and a user *u*'' s role set *R(u' , c)* in *L(r)*, satisfying $1 \leq p \leq p$, if all other roles (r_1 to $r\ _1$) besides r do not belong to *R(u' , c)*, we propose the following formula to define the Upper Bound of the L1-norm of *S(u, u' , c)*:

$$UB\text{-}L(u, u', c) = \frac{\sum_{r \in R(u,c) \cap R(u',c)} W(r, c)}{\sum_{r \in R(u,c)} W(r, c) + \sum_{r \in R(u',c) \setminus R(u,c)} W(r, c)}$$

4.4. Top-k Recommendation

First, we identify all roles enacted by user u in context c. For each role r in *R(u, c)*, we check whether it corresponds to context c in the UCB matrix. If so, r is a role enacted by u in c. As mentioned earlier, a role is a collection of context-based behaviors, which are the content that users tend to select. When u enacts a certain role, u may be interested in the corresponding content in context c. For each selection tendency, the recommendation system predicts the rating that user u would give. Users often refer to their trust objects' choices, so the system finds all trust objects of u and calculates *u*' s rating for option v in context c, denoted as *RT(u, v, c)*. We propose the following formula:

$$RT(u, v, c) = \frac{\sum_{u' \in U(u,c)} UIR[u'][v] \cdot T(u, u', c)}{\sum_{u' \in U(u,c)} T(u, u', c)}$$

UIR[u'][v] represents the rating that user *u*' would give to option v. *U(u, c)* represents the set of trusted users of user u who have already rated option v in context c. The top k options with the highest final ratings will be recommended to user u in context c.

Additionally, there are three types of special users. The first type exhibits very few behaviors, and the second type has unique behaviors different from all other users. Both types hardly enact any roles. For these users, we find similar users based on common context-aware behaviors. The third type enacts many roles in other contexts but has no behavior in the current context. For these users, we find similar users based on role set similarity. Based on user similarity, we can construct a trust network for each special user in the current context and similarly recommend the top k options with the highest predicted ratings.

4.5. Recommendation System Experiments and Quality Evaluation

(1) Experimental Data Description

The experimental data is derived from the Extended Epinions dataset. Epinions is a user review website that allows users to rate items on a scale from 1 to 5, where users also express explicit trust relationships. Our experiments primarily use book review data from this website; users' reading and rating of books on mobile terminals form a mobile digital library. The extended Epinions dataset includes 132,000 users who provided 717,667 trust expressions and a total of 13,668,319 ratings for 1,560,144 articles. Since the original dataset is not context-dependent, we apply a context modification method [8] to add two artificial contextual factors to the original data: selecting 90% of items and randomly extracting 50% of ratings for modification, making context affect the rating items in the modified dataset.

(2) Method Comparison Analysis

In the experimental results analysis, we compare the proposed Role-based Trust Network recommendation method (RRTN) with the following methods:

1. **RRTN method**: This is the context-independent version of RRTN, which calculates similarity based only on common behaviors between users u and u' , i.e., using *T(u, u')* instead of *T(u, u' , c)* for recommendation.
2. **RRTN method**: This method does not build a role-based trust network for all users but makes recommendations for user u based on roles enacted by u in context c. Comparing with RRTN evaluates the context-aware module of the RRTN method.
3. **RSTE method**: This method [14] is based on social trust recommendation, with the core idea that each user' s choice reflects both personal preferences and recommendations from trusted users.
4. **MV method**: This method [8] is designed for context-aware recommendation, using a tensor factorization-based collaborative filtering approach that models data as an N-dimensional tensor of user-item-context instead of the traditional user-item binary state.
5. **FM method**: This method [21] is also a context-aware recommendation approach that outperforms other methods such as MV in terms of prediction accuracy and recommendation time. The FM method incorporates both user trust relationships and contextual information, thus considering both context and user trust for recommendation.

(3) Recommendation Quality

Evaluation Metrics This paper uses MAE and RMSE methods to evaluate the accuracy of user-item rating predictions, and Precision at k and NDCG at k methods to assess the quality of the recommended k items.

In MAE and RMSE methods, smaller MAE or RMSE values indicate higher prediction accuracy and thus better recommendation quality. Based on the metrics in this paper, we adjust them and define them as:

$$MAE = \frac{\sum_{i,j} |RT_{ij} - \hat{RT}_{ij}|}{N}, \quad RMSE = \sqrt{\frac{\sum_{i,j}(RT_{ij} - \hat{RT}_{ij})^2}{N}}$$

RT_{ij} and $\hat{RT}_{ij}$ represent the actual and predicted ratings given by user i to item j, respectively, and N represents the total number of test samples.

For each user u in the test set, Precision at k represents the proportion of items that appear in both the top k items with highest predicted ratings and user u' s experimental Top k list. We sort user u' s rated items in reverse order of actual ratings and select the top k items to form user u' s experimental Top k list.

Recommendation quality is sensitive to the positions of the top k recommended items with highest ratings. We use the Normalized Discounted Cumulative Gain (NDCG) metric to measure the quality of the top k recommended items. NDCG is defined as:

$$NDCG@k = \frac{1}{Z_k} \sum_{j=1}^{k} \frac{RT(j)}{\log_2(1+j)}$$

RT(j) represents the rating of the item at position j in the ordered item list, and constant *Z_k* is set so that the NDCG value equals 1 in the ideal state.

Impact of Parameters *m* and *n* In our method, parameters m and n balance the number of roles and the level of role abstraction. If m and n values are large, the number of mined roles will decrease, resulting in a large portion of users not enacting any or very few roles. In this case, finding similar users for a user based on common context-aware behaviors rather than using the WSSQ algorithm based on role sets for similarity calculation and similar user selection will harm recommendation quality. On the other hand, if m and n values are too small, the role scale becomes smaller, leading to a situation of "almost no similarity" when calculating user similarity, which affects the system' s final determination. To find appropriate m and n values that improve recommendation quality, we conduct experiments comparing MAE and RMSE changes under different m and n values.

[Figure 6: see original paper] shows the impact of parameters m and n on MAE and RMSE in the experimental dataset. It can be seen that when m and n are 750 and 2, respectively, MAE reaches its minimum value of 0.605 and RMSE reaches its minimum value of 0.805. Therefore, the optimal values for m and n should be 750 and 2, respectively. In subsequent comparisons and evaluations, m and n will also take these optimal values to ensure the recommendation algorithm achieves the best quality.

(4) Quality Evaluation of Role-based Trust Network Recommendation (RRTN) We compare the final prediction accuracy using training data proportions ranging from 20% to 99%, measured by MAE and RMSE. For example, results at the 90% proportion mean using randomly selected 90% training data to predict the remaining 10% rating data. Rating prediction quality is measured by Precision at k and NDCG at k values under different k values.

[Figure 7: see original paper] shows the MAE performance of different recommendation methods. [Figure 8: see original paper] shows the RMSE performance. [Figure 9: see original paper] shows the Precision at k values. [Figure 10: see original paper] shows the NDCG at k values.

From [Figure 7: see original paper] to [Figure 10: see original paper], we can see that RRTN achieves more accurate recommendations than RSTE for two main reasons: RSTE is not a context-aware recommendation method and ignores contextual factors; RRTN makes recommendations based on a dense role-based trust network, where many implicit trust relationships can be quantified through the trust model, while RSTE relies only on sparse explicit trust relationships. The results also show that RRTN' s recommendation quality is significantly better than RRTN and RRTN . RRTN is the context-independent version of RRTN, calculating similarity between users u and u' based only on their common behaviors without considering context c. RRTN does not build a role-based trust network for all users but makes recommendations based on roles enacted by user u in context c. RRTN' s superior quality over RRTN is mainly due to its consideration of contextual factors. RRTN ignores the impact of trust relationships between users on their choices in reality, as users tend to select items highly rated by their trust objects. These results demonstrate that trust relationships and contextual information need to be combined in recommendation.

Although the FM method considers both trust relationships and contextual information, its performance is still inferior to RRTN, mainly because RRTN introduces roles to represent user groups' common context-aware behaviors. Roles help the system determine users' context-aware interest preferences and trust relationships. The FM method ignores implicit trust expressions among users.

Taking the FM method as an example: in [Figure 7: see original paper] and [Figure 8: see original paper], when using 20% of the dataset, RRTN is 34% lower than FM in MAE and 48% lower in RMSE; when using 99% of the dataset, RRTN is 28% lower than FM in MAE and 51% lower in RMSE. All methods' MAE and RMSE values in [Figure 7: see original paper] and [Figure 8: see original paper] show a decreasing trend, confirming that more training data yields better recommendation results. [Figure 9: see original paper] shows that when $k = 10$, RRTN is at least 2.8% better than other methods, and when $k = 50$, it is at least 2.9% better. [Figure 10: see original paper] shows that when $k = 10$, RRTN is at least 2% better than other methods, and when $k = 50$, it is at least 3% better.

5. Conclusion

Based on the real-world scenario of mobile digital libraries providing personalized recommendation services to users in mobile environments, and grounded in the theoretical foundation that users enact different roles in different contexts and that users with the same roles share preferences for certain recommended items, this paper proposes a context-aware personalized recommendation approach for mobile digital library users. First, we use a role mining algorithm to mine roles in the UCB matrix. Then, we construct a role-based trust model to measure trust values between any two users. Subsequently, we employ an effective WSSQ algorithm to build a role-based trust network for user u, ensuring that u' s trust objects and trust values change with context and roles. This recommendation method finds trusted users through role set similarity, effectively solving the sparsity problem of explicit trust data among users. Experimental results also show that our proposed method achieves better recommendation quality compared to other methods.

However, in this paper, contexts are predefined and role mining is implemented offline. If user samples are sufficiently large, new contexts and roles become sparse. Therefore, future research will design an effective and scalable role mining algorithm that can adapt to new contexts and users.

The national mobile digital library system aims to establish a knowledge kingdom with outlets across the country—a nationwide cross-regional, cross-departmental, and cross-industry network of scientific, technological, and cultural information resources. In the current situation where major commercial information service platforms have developed personalized recommendations relying on the internet environment, simple resource content pushing is insufficient to highlight the characteristics of libraries themselves, nor can it maximize the value of collections and fulfill the responsibility of social information communication. The service tenet of mobile digital libraries inherits the library' s user philosophy of "every book has its reader." Under new circumstances, it should further expand its in-depth and professional reader service spirit with the help of rapidly developing technological theories. Future research on resource recommendation systems should also be integrated into this development. We believe the following points can serve as references for in-depth development:

1. Establish personal user accounts to collect more user information and subdivide existing item tags to make recommended content more aligned with users' actual needs. Users' resource needs are sometimes implicit in their personal information such as education level and social position. If we only consider user interests while ignoring their characteristics as social individuals, the recommendations may be biased. For example, economics can be subdivided into economic life common sense and economics academic papers, thus serving non-economics professionals, business people, and academics differently.

2. Combine users' social identities and work needs to integrate recommendation services, virtual reference services, and selective dissemination of information services, providing professional consultation pushes and special project follow-ups. This will further inherit and strengthen the advantageous projects of traditional libraries in serving professionals. When discussing the development of physical libraries, research libraries and public libraries are often distinguished, with the former' s key service project being reference consultation. When the situation shifts to mobile digital libraries, the consultation needs of researchers and professionals still exist. Mobile digital libraries, relying on internet platforms, can utilize more resources and serve more diverse users in terms of industry, professional level, and spatial distribution. Integrating recommendation systems into consultation platforms and project service platforms is a requirement of economic and social development for mobile digital libraries.

References

[1] Wang Zhijin, Hou Yanxiang. On the Service Pattern of the Sustainable Digital Library in China [J]. Journal of Academic Libraries, 2005, 23(5): 2-6.

[2] Sheng Xiaoping. Collection Evaluation of Digital Libraries [J]. Library and Information Service, 2003, 47(5): 40-43.

[3] Forestier M, Stavrianou A, Velcin J, et al. Roles in Social Networks: Methodologies and Research Issues [J]. Web Intelligence & Agent Systems, 2012, 10(1): 117-133.

[4] Adomavicius G, Tuzhilin A. Context-Aware Recommender Systems [C]. In: Proceedings of ACM Conference on Recommender Systems. ACM, 2008: 2175-2178.

[5] Anand S S, Mobasher B. Contextual Recommendation [A]. // From Web to Social Web: Discovering and Deploying User and Content Profiles [M]. Springer Berlin Heidelberg, 2006: 287-313.

[6] Bao T, Cao H, Chen E, et al. An Unsupervised Approach to Modeling Personalized Contexts of Mobile Users [J]. Knowledge & Information Systems, 2010, 31(2): 345-370.

[7] Sen S, Adomavicius G, Sankaranarayanan R, et al. Incorporating Contextual Information in Recommender Systems Using a Multidimensional Approach [J]. ACM Transactions on Information Systems, 2005, 23(1): 103-145.

[8] Karatzoglou A, Amatriain X, Baltrunas L, et al. Multiverse Recommendation: N-dimensional Tensor Factorization for Context-aware Collaborative Filtering [C]. In: Proceedings of ACM Conference on Recommender Systems (RecSys2010), Barcelona, Spain. 2010: 79-86.

[9] Rendle S, Gantner Z, Freudenthaler C, et al. Fast Context-aware Recommendations with Factorization Machines [C]. In: Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2011), Beijing, China. 2011: 635-644.

[10] Massa P, Avesani P. Trust-aware Recommender Systems [C]. In: Proceedings of Conference on Recommender Systems. 2007: 17-24.

[11] Jamali M, Ester M. TrustWalker: A Random Walk Model for Combining Trust-based and Item-based Recommendation [C]. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2009: 397-406.

[12] Jamali M, Ester M. Using a Trust Network to Improve Top-N Recommendation [C]. In: Proceedings of ACM Conference on Recommender Systems. ACM, 2009: 181-188.

[13] Dubois T, Golbeck J, Kleint J, et al. Improving Recommendation Accuracy by Clustering Social Networks with Trust [C]. In: Proceedings of ACM RecSys09 Workshop on Recommender Systems & the Social Web. 2009.

[14] Ma H, King I, Lyu M R. Learning to Recommend with Social Trust Ensemble [C]. In: Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2009: 203-210.

[15] Golbeck J. Trust and Nuanced Profile Similarity in Online Social Networks [J]. ACM Transactions on the Web, 2009, 3(4): 3451-3459.

[16] McPherson M, Smithlovin L, Cook J M. Birds of a Feather: Homophily in Social Networks [J]. Annual Review of Sociology, 2001, 15(4): 344-349.

[17] Vaidya J, Atluri V, Guo Q. The Role Mining Problem: Finding a Minimal Descriptive Set of Roles [C]. In: Proceedings of the 12th ACM Symposium on Access Control Models and Technologies. 2007: 175-184.

[18] Tong H, Faloutsos C, Pan J Y. Fast Random Walk with Restart and Its Applications [C]. In: Proceedings of the 6th International Conference on Data Mining. IEEE Computer Society, 2006: 613-622.

[19] Hadjieleftheriou M, Srivastava D. Weighted Set-Based String Similarity [J]. Bulletin of the Technical Committee on Data Engineering, 2010, 33(1): 25-36.

[20] Liu H, Lim E P, Lauw H W, et al. Predicting Trusts Among Users of Online Communities: An Epinions Case Study [C]. In: Proceedings of the 9th ACM Conference on Electronic Commerce (EC-2008), Chicago, IL, USA. 2008: 310-319.

[21] Rendle S. Factorization Machines with libFM [J]. ACM Transactions on Intelligent Systems & Technology, 2012, 3(3): 219-224.

[22] Meng Liansheng. Some Thinking about How to Develop Digital Libraries in China [J]. Documentation, Information & Knowledge, 2002(1): 9-12.

Author Contributions: Hong Liang: proposed research ideas, designed research plan, completed experiments and analysis; Qian Chen: implemented the scheme, drafted the paper; Hong Liang, Qian Chen, Fan Xing: revised the final version of the paper.

Conflict of Interest Statement: All authors declare no conflict of interest.

Received Date: 2016-04-22
Revised Date: 2016-05-20

Context-aware Recommendation System for Mobile Digital Libraries

Hong Liang, Qian Chen, Fan Xing
(School of Information Management, Wuhan University, Wuhan 430072, China)

Abstract: [Objective] This paper proposes a context-aware recommendation system for mobile digital libraries, with the help of the latter' s collection features and users' behaviors. [Methods] Based on the theory that similar users make similar choices, we first modeled users' interests by introducing the concept of roles. Second, we designed an effective Weighted Set Similarity Query (WSSQ) algorithm to build a role-based trust network for the users. Finally, we modified the existing context-aware recommendation system, which was then evaluated with an Extended Epinions dataset. [Results] The proposed new recommendation system was feasible, and had better performance than other methods. [Limitations] The contexts and roles were not rich enough to process large user samples. [Conclusions] This study could help us improve the mobile digital library' s resource recommendation system.

Keywords: Mobile digital library; Context-aware recommendation; WSSQ algorithm

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.