Cross-Domain Recommendation via Clustering on Multi-Layer Graphs

Aleksandr Farseev*, Ivan Samborskii** *, Andrey Filchenkov**, Tat-Seng Chua*
*National University of Singapore, **ITMO University
farseev@u.nus.edu,samborskiy.ivan@gmail.com,afilchenkov@niuitmo.ru,chuats@nus.edu.sg

ABSTRACT

Venue category recommendation is an essential application for the tourism and advertisement industries, wherein it may suggest attractive localities within close proximity to users' current location. Considering that many adults use more than three social networks simultaneously, it is reasonable to leverage on this rapidly growing multi-source social media data to boost venue recommendation performance. Another approach to achieve higher recommendation results is to utilize group knowledge, which is able to diversify recommendation output. Taking into account these two aspects, we introduce a novel cross-network collaborative recommendation framework C^3R, which utilizes both individual and group knowledge, while being trained on data from multiple social media sources. Group knowledge is derived based on new cross-source user community detection approach, which utilizes both inter-source relationship and the ability of sources to complement each other. To fully utilize multi-source multi-view data, we process user-generated content by employing state-of-the-art text, image, and location processing techniques. Our experimental results demonstrate the superiority of our multi-source framework over state-of-the-art baselines and different data source combinations. In addition, we suggest a new approach for automatic construction of inter-network relationship graph based on the data, which eliminates the necessity of having pre-defined domain knowledge.

ACM Reference format:
Aleksandr Farseev*, Ivan Samborskii** *, Andrey Filchenkov**, Tat-Seng Chua*. 2017. Cross-Domain Recommendation via Clustering on Multi-Layer Graphs. In *Proceedings of SIGIR '17, Shinjuku, Tokyo, Japan, August 07-11, 2017,* 10 pages.
https://doi.org/10.1145/3077136.3080774

1 INTRODUCTION

There has been an exponential growth of publicly available data on the Web. This growth may be linked to users using multiple online social networks, such as that seen among many adult users, who were found to use three or more online social networks daily [8]. These social networks are inter-connected through the deployment of the so-called cross-linking functionality [10, 33], which offers information about the same users from different perspectives [9]. In an attempt to make sense of the overwhelming amount of multi-source data, various recommender systems were designed to extract relevant information from individual users. However, most of them utilize data from a single-source or of a single modality, while disregarding the potential of cross-source multi-modal recommendation. Another aspect which lacks research, and is worth exploring, is the role of personal and group knowledge integration in cross-source multi-modal recommendation. It has been established in an earlier study that effective recommendation can be achieved based on the combination of group and individual knowledge [7, 53]. While individual knowledge is useful for the incorporation of personal experience [48], group knowledge helps to improve the recommendation diversity [26].

Aiming to bridge these two research gaps, in this study, we focus on the problem of **recommendation across data sources** (also known as cross-domain recommendation) based on multi-view social media data. Specifically, this work is devoted to the emerging topic of venue category recommendation [4, 18, 49], where we recommend a ranked list of Foursquare venue categories to users with accounts in three social networks, namely Twitter, Instagram, and Foursquare [19]. By performing filtering based on venue categories of places around user's current location, our recommender system does not depend on any particular geographical area. This means that it can be used for many real-world scenarios, such as tourist route planning [32], facility arrangement [46], or interactive mobile assistance [20]. For example, based on one's venue category preferences, we could further recommend a particular venue (i.e. Chinese Restaurant, Movie Theatre, etc.) near his/her current location, rather than a similar place located far away, even if the venue located farther away may slightly better fit his/her preferences. Last but not least, the venue category recommendation (but not venue recommendation) also helps to overcome evaluation difficulty, which often arises due to location datasets' sparsity.

Despite its potential, we recognize the challenges of integrating cross-source venue category recommendation with individual and group knowledge. One challenge is **data integration**. Multiple data sources often describe distinct sides of users' activities. For example, Twitter may casually reveal users' daily life updates, while Instagram may uncover their visual preferences. At the same time, social media content comes in many forms (modalities), such as text, images or videos, which also exhibit different facets of users' online experiences. Such multi-source multi-modal data must be integrated into one recommendation framework in a mutually-consistent fashion, which is an open research problem. Another challenge pertains to **group representation learning**. Group representation can be naturally expressed in the form of user communities, which must be extracted from heterogeneous multi-source data. However, the detection of such user communities from multi-source multi-modal

SIGIR '17, August 07-11, 2017, Shinjuku, Tokyo, Japan

ACM ISBN 978-1-4503-5022-8/17/08…$15.00
https://doi.org/10.1145/3077136.3080774

data is not an easy task due to the necessity of proper inter-source and intra-source relationship modeling.

Inspired by previous studies and the challenges above, we seek to address three research questions. First, to support the assumptions behind this study, it is important to answer: **(RQ1) Is it possible to improve the recommendation performance by integrating individual and group knowledge?** Second, even though the topic has been discussed with respect to some problems, it is still unclear if: **(RQ2) Inter-source relationship information enable us to find better user communities.** Finally, for further recommendation improvement, it is important to understand: **(RQ3) What the contribution is of each data source (modality) towards venue category recommendation.**

To answer these research questions, we introduce a novel recommender framework $\mathbf{C^3R}$ that utilizes group and individual knowledge to perform **C**ross-Source User **C**ommunity-Based **C**ollaborative venue category **R**ecommendation. Individual knowledge is obtained from user's experience, which is modeled as the distribution among venue categories that a user has visited in past. To incorporate group knowledge, we detected cross-source user communities in a latent space, where the relationship between users is modeled as a multi-layer graph. The community detection approach incorporates inter-source relationships during the process of learning individual source representations and preserves inter-source consistency at the stage of learning latent sources' representation. The inter-source relationship graph is computed automatically from the data and further utilized via novel graph-constrained regularization. The experimental results show that our framework can achieve better recommendation performance in three geographical regions, as compared to state-of-the-art baselines and different data source combinations.

The main contributions of this paper are threefold: First, we present a **cross-source venue recommendation framework** that utilizes both individual and group knowledge; second, we propose a **novel cross-source user community detection approach** that utilizes both inter-source relationship and sources' ability to complement each other via efficient regularization; third, we suggest a **new approach for automatic construction of inter-source relationship graph** based on the data, which eliminates the necessity of having expert knowledge.

2 RELATED WORKS

Probably one of the first studies towards improving recommendation performance based on multi-source data was conducted by Abel et al. [1]. The authors aggregated user profiles from Flickr, Twitter, and Delicious to demonstrate that their cross-network user modeling strategies have a large impact on the recommendation quality in cold-start settings. At the same time, Tiroshi et al. [47] utilized network-related and domain-related features to perform user interest recommendation. Later on, Yan et al. [56] proposed a two-stage solution of cross-source video recommendation problem: first, user preferences were transferred from an auxiliary network by learning cross-network behavior correlations; next, the transferred preferences were integrated with the observed behaviors on target network in an adaptive fashion. Concurrently, Qian et al. [37] introduced a probabilistic framework that solved the problem of cross-domain recommendation by utilizing shared domain priors and modality priors for collaborative learning of a latent representation. Recently, Farseev et al. [18] performed cross-source venue category recommendation by implementing a recommender system based on their proposed Multi-Source re-Ranking approach, where the ranks of individual sources were obtained by performing nearest neighbor collaborative filtering. These works are related to our study regarding the cross-source approaches utilized. Finally, Farseev et al. [17] and Wang et al. [52] proposed cross-domain recommender systems, where inter-domain linking was implemented via the so-called "bridge" users (social media users who have accounts on two or more social networks). However, they do not incorporate both group and individual knowledge into recommendation, which our study is recommending, and is an essential aspect of our study.

At the same time, several works highlighted the potential of multi-source data to find better user communities. For example, Su et al. [45] demonstrated the usefulness of multi-source community discovery for various applications; Rhouma and Romdhane [40] proposed an approach for multi-source user community detection in partially-overlapping social network graphs; while Dong et al. [12] introduced a multi-source clustering approach, where the distance between different data sources was measured on Grassmann Manifolds. The works mentioned above are supportive of utilizing multi-source community detection strategies; however, they are limited because inter-source relationship was not incorporated during the community discovery process. This could potentially lead to sub-optimal results in real world settings.

There were also research efforts in studying the contribution of multiple data sources for different applications. For example, Farseev et al. [15, 16] proposed a multi-source multi-task learning frameworks that aim to combine multi-source multimodal data and data from wearable sensors for Body Mass Index inference; while Song et al. [44] and Akbari et al. [3] developed multi-task learning frameworks for user interests inference and wellness events categorization, respectively. In these four studies, the inference category relationship "weights" were automatically inferred from the data and used to guide the learning model, which can also be applied to solve our problem.

3 CROSS-SOURCE RECOMMENDATION

Strategic decision making is known to be influenced by external factors like personal experiences and public opinions [5]. The public opinion can be expressed by explicit and implicit user communities that are formed based on social relations and their data similarity, respectively. This phenomenon can be leveraged to enhance recommendation performance. To do so, we perform venue category recommendation based on both personal and group knowledge, which naturally models the impact of society on an individual's behavior during the selection of a new place to go. Formally, our proposed C^3R recommendation approach is defined as follows:

$$\text{rec(u)} = \text{sort}\left(\gamma \cdot vec_u + \theta \frac{\sum_{v \in C_u} vec_v}{|C_u|}\right) \tag{1}$$

where vec_u is the distribution of the user u among items (venue categories) **in u's past**, and $\frac{\sum_{v \in C_u} vec_v}{|C_u|}$ is the normalized (by the number of members in user community C_u) distribution of all community members among venue categories **in C_u's past**, γ controls the personal aspect of recommendation, while θ regulates the group experience impact of the user community C_u. Besides the benefits described earlier, incorporation of user communities reduces the search space during the recommendation process and provides better candidates to compare for further collaborative filtering [42]. While personal information (i.e. users' previous activity) is often available, in most of the cases user communities are not explicitly indicated, which gives rise to user community detection challenge.

4 USER COMMUNITY DETECTION

4.1 Similarity Graph Construction

The first step in finding representative user groups is the modeling of users' relationships in the form of a graph so that dense subgraphs of such graph can be treated as user communities. The graph can be constructed based on: (a) social connections between users (i.e. follower/followee relationship) that are often hidden behind the privacy settings; or (b) user-generated content (UGC), where similarity between users is estimated as a distance between data representations of users and each data source (modality) modeled as a layer in a multi-layer graph. In our work, we adopt the graph construction based on UGC to avoid privacy concerns and limitations in mining explicit user social relations. Specifically, for every graph node pair (i, j) from the m-th graph layer, the corresponding distance can be computed by applying *Heat* kernel:

$$d_{m_{i,j}} = e^{-\frac{||x_{m_i} - x_{m_j}||^2}{\sigma}},$$

where $||x_{m_i} - x_{m_j}||$ is the Euclidean norm, and σ is a graph sparsity-related parameter that could be found by grid search. There are certain benefits of such graph construction approach [14]. Firstly, it does not suffer from the lack of information about users' relationship, since the relations between users are simply modeled as distances between users based on UGC similarity. Secondly, it naturally solves the problem of cross-region recommendation, where the condition for related users to be explicitly connected in social networks is relaxed.

4.2 Problem Formulation

To simplify the reading process, we summarize all defined notations in Table 1.

One of the commonly used formulations of the community detection problem is its representation in a form of NCut formulation, which conditions the sum of graph edges' weights in each community to be minimized among all communities [43]. This simply means that all communities are formed by users that are most "similar" to each other. Such a definition is naturally applicable to our task of group representation learning based on users' interests. We thus adopt the NCut formulation in our study. The NCut definition is given below:

$$NCut(C_1, ..., C_k) = \sum_{i=1}^{k} \frac{W(C_i, \overline{C}_i)}{vol(C_i)} = \sum_{i=1}^{k} \frac{cut(C_i, \overline{C}_i)}{vol(C_i)},$$

where $vol(C_i)$ is the sum of weights of all edges attached to vertices in C_i.

Table 1: Notations summary

Symb.	Description
vec_u	Distribution of the user u among items (venue categories) in u's past
C_u	Community of the user u
γ	Parameter that controls personal aspect of rec-n
θ	Parameter that controls group aspect of rec-n
N	Number of users
M	Number of data sources (graph layers)
L_i	Laplacian matrix of the i-th layer
U_i	Eigendecomposition matrix of the i-th layer
$\hat{L_i}$	Inter-layer relationship regularized Laplacian of the i-th layer
$\hat{U_i}$	Inter-layer relationship regularized eigendecomposition matrix of the i-th layer
$\hat{L}_{mod}$	Sub-space regularized Laplacian matrix
W_R	Adjacency matrix of inter-layer similarity graph
k	Parameter that controls the number of clusters
α	Parameter that controls sub-space regularization
β_i	Parameter that controls inter-layer regularization for the layer i

However, the NCut problem is proven to be $\mathcal{NP}$-hard [51]. Fortunately, there exists a state-of-the-art approximation that is defined as a standard trace minimization (also known as *spectral clustering*) [50]:

$$\min_{U \in R^{n \times k}} \operatorname{tr}(U^\intercal L_{sym} U), \ \ s.t. \ U^\intercal U = I. \tag{2}$$

By the Rayleigh-Ritz theorem, the solution of the problem in Equation (2) is given by the first k eigenvectors of the normalized graph Laplacian $L_{sym} = I - D^{-\frac{1}{2}} W D^{-\frac{1}{2}}$, where W is adjacency matrix, and D is degree matrix [30]. The mapping of each user to a cluster can be further obtained by i.e. k-means clustering over the eigenvector space [50].

4.3 Spectral clustering on multi-layer graph

The well-known disadvantages of early fusion and late fusion data aggregation strategies [3] encourage us to perform joint clustering from all graph layers simultaneously. The final data representation (latent representation) must be consistent with all graph layers so that clustering will not be biased towards individual graph layers. One way to keep the latent representation consistent to all layers of the multi-layer graph is to apply regularization during the clustering process, which will keep graph layers "close" to the target representation. A well-adopted approach for measuring "closeness" between target latent space U and the space of each layer U_i is based on Grassman Manifolds [25], where the projected distance between two spaces S_1 and S_2 on Grassman Manifold is equal to [12]:

$$d^2_{Proj}(S_1, S_2) = \frac{1}{2} ||S_1 S_1^\intercal - S_2 S_2^\intercal||_F^2, \tag{3}$$

where $||A||_F$ is the Frobenius norm [22] of the matrix A.

Since mappings $S_1 S_1^\intercal$ and $S_2 S_2^\intercal$ preserve distinctness [12], we consider the projection distance as a distance measure between subspaces. The distance between target subspace S and all the other individual subspaces $\{S_i\}_{i=1}^M$ can thus be defined as the sum of squared projection distances between S and all individual subspaces $\{S_i\}_{i=1}^M$ [12]:

$$d^2_{Proj}(S, \{S_i\}_{i=1}^M) = kM - \sum_{i=1}^{M} \mathrm{tr}(SS^\top S_i S_i^\top). \tag{4}$$

By utilizing the Equation (4), the Equation (2) can be extended to introduce the subspace-regularized objective:

$$\min_{U \in \mathbb{R}^{n \times k}} \sum_{i=1}^{M} \mathrm{tr}(U^\top L_i U) + \alpha(kM - \sum_{i=1}^{M} \mathrm{tr}(UU^\top U_i U_i^\top)), \quad s.t.\ U^\top U = I, \tag{5}$$

where α controls sub-space regularization. We can rearrange Equation (4) to present it in a form of standard trace minimization problem [12]:

$$\begin{aligned} &\min_{U \in \mathbb{R}^{n \times k}} \sum_{i=1}^{M} \mathrm{tr}(U^\top L_i U) + \alpha(kM - \sum_{i=1}^{M} \mathrm{tr}(UU^\top U_i U_i^\top)) \\ &= \min_{U \in \mathbb{R}^{n \times k}} \mathrm{tr}(U^\top \sum_{i=1}^{M} (L_i - \alpha U_i U_i^\top) U), \end{aligned} \tag{6}$$

which, by the Rayleigh-Ritz theorem [30], can be solved as the first k eigenvectors of the modified Laplacian $L_{mod} = \sum_{i=1}^{M} (L_i - \alpha U_i U_i^\top)$. The eigen decomposition of the matrix L_{mod} is not in the scope of the paper, but can be efficiently computed by well-known algorithms [28].

4.4 Incorporating inter-layer relationship

Even though the subspace-regularized spectral clustering approach performed well on synthetic dataset [12], real world problems often require a consideration of inter-layer relationship (similarity) during the learning process. Such a requirement is caused by the difference between data modalities and the way they describe the users [18]. Specifically, for particular applications some data sources or modalities may be more informative than others. For example, Foursquare check-ins could be more useful for venue category recommendation than textual posts, while Twitter text data is of crucial importance for demographic profiling [19].

Inspired by previous works [6, 18, 54] and based on our observations, we take a step further in multi-layer clustering by introducing a novel model regularization schema that guides clustering process based on predefined inter-layer relationship graph. Given the multi-layer user relationship graph G, which is constructed as described in previous sections, let's assume that there exists a complete undirected inter-layer similarity graph R with the adjacency matrix W_R, where each value of the matrix $w_{i,j}$ represent the similarity between i-th and j-th layer of G. While W_R can be automatically computed from the data (see next section for details), we first describe our inter-layer regularized spectral clustering approach.

Essentially, our goal is to regularize the conventional spectral clustering in such a way that the representation of each layer $\hat{U}_i$ would be computed with respect to the inter-layer similarities $w_{i,j}$, which are taken from adjacency matrix of the inter-layer relationship graph R. Let's also note that the approach in Equation (6) relies on pre-computed layers' Laplacians $\{L_i\}_{i=1}^M$ and their spectral spaces $\{U_i\}_{i=1}^M$. In our work, we apply the inter-layer regularization on the process of computing the layers' spectral spaces, so that the final multi-layer spectral space is computed as in Equation (6). By using previously defined distance on Grassman manifold, we define the new objective function for the i-th layer as follows:

$$\min_{\hat{U}_i \in \mathbb{R}^{n \times k}} \mathrm{tr}(\hat{U}_i^\top L_i \hat{U}_i) + \beta_i (kM - \sum_{j=1, j \neq i}^{M} w_{i,j}\, \mathrm{tr}(\hat{U}_i \hat{U}_i^\top U_j U_j^\top)), \quad s.t.\ \hat{U}_i^\top \hat{U}_i = I, \tag{7}$$

where $\hat{U}_i^\top$ is the new spectral space of the i-th layer, β_i — parameter that controls inter-layer regularization for the layer i, $\{U_j\}_{j=1}^M$ — spectral spaces of all layers after standard spectral clustering, $w_{i,j}$ — similarity between layer i and layer j. The problem in Equation (7) can be then presented as a standard trace minimization:

$$\begin{aligned} &\min_{\hat{U}_i \in \mathbb{R}^{n \times k}} \mathrm{tr}(\hat{U}_i^\top L_i \hat{U}_i) + \beta_i (kM - \sum_{j=1, j \neq i}^{M} w_{i,j}\, \mathrm{tr}(\hat{U}_i \hat{U}_i^\top U_j U_j^\top)) \\ &= \min_{\hat{U}_i \in \mathbb{R}^{n \times k}} \mathrm{tr}(\hat{U}_i^\top (L_i - \beta_i \sum_{j=1, j \neq i}^{M} w_{i,j} U_j U_j^\top) \hat{U}_i), \end{aligned}$$

thus, by the Rayleigh-Ritz theorem, it can be solved as the first k eigenvectors of the regularized Laplacian $\hat{L}$:

$$\hat{L}_i := L_i - \beta_i \sum_{j=1, j \neq i}^{M} w_{i,j} U_j U_j^\top.$$

We now presented all necessary components to define our multi-layer clustering approach with the following objective function:

$$\begin{aligned} &\min_{U \in \mathbb{R}^{n \times k}} \sum_{i=1}^{M} \mathrm{tr}(U^\top \hat{L}_i U) + \alpha(kM - \sum_{i=1}^{M} \mathrm{tr}(UU^\top \hat{U}_i \hat{U}_i^\top)) \\ &= \min_{U \in \mathbb{R}^{n \times k}} \mathrm{tr}(U^\top \sum_{i=1}^{M} (\hat{L}_i - \alpha \hat{U}_i \hat{U}_i^\top) U), \\ &s.t.\ U^\top U = I. \end{aligned} \tag{8}$$

To make the clustering procedure clear, we present the pseudocode as shown in Algorithm 1.

From the pseudocode, it can be seen that optimization of the Equation (8) and further clustering consists of four main steps: 1) Perform conventional spectral clustering on each layer to obtain L_i and U_i; 2) By incorporating inter-layer relationship graph R, perform inter-layer relationship regularized spectral clustering on each layer to obtain $\hat{L}_i$ and $\hat{U}_i$; 3) Execute subspace-regularized spectral clustering on each layer to obtain $\hat{L}_{mod}$ and U; 4) Normalize U to obtain U_{norm} and execute the x-means clustering over it [35].

The value of the subspace regularization parameter α and the inter-layer regularization parameters β_i can be found by grid search. In next section, we outline the construction of inter-layer similarity graph G.

4.5 Computing Inter-Layer Relationship

Intuitively, the inter-layer relationship graph R must represent the similarity between layers in terms of clustering results, summarized for different values of k. We thus define the similarity $sim(w, q)$ between graph layers q and w as a normalized difference between $N \times N$ k-clustering co-occurrence matrices $M_{q,k}$, $M_{w,k}$, in which each value $m_{i,j}$ is equal to 1 if user i is assigned to the same cluster as user j in both layers w and q, and 0 otherwise. The clustering co-occurrence matrices are obtained by performing single-layer spectral clustering on each layer of the multi-layer graph and for

Algorithm 1 C^3R clustering

function CLUSTER($\{W_i\}_{i=1}^M$, W_R, k, α, $\{\beta_i\}_{i=1}^M$)
  ▷ $\{W_i\}_{i=1}^M$ are weighted adjacency matrices of layers $\{G_i\}_{i=1}^M$, W_R is adjacency matrix of inter-layer similarity graph, k is target number of clusters, α and $\{\beta_i\}_{i=1}^M$ are regularization parameters

  for $i \leftarrow [0; M-1]$ **do**
    Compute L_i and U_i for G_i [50]
    ▷ L_i is the normalized Laplacian matrix of the layer i, U_i is subspace representation of the layer i, G_i is ith layer graph
    Compute $\hat{L}_i \leftarrow L_i - \beta_i \sum_{j=1, j \neq i}^{M} w_{i,j} U_j U_j^\intercal$
    ▷ $\hat{L}_i$ is the regularized Laplacian matrix of ith layer
    Compute $\hat{U}_i \in \mathbb{R}^{N \times k}$
    ▷ $\hat{U}_i$ is the the matrix of first k eigenvectors of $\hat{L}_i$[28]
  end for
  Compute $\hat{L}_{mod} \leftarrow \sum_{i=1}^{M} (\hat{L}_i - \alpha \hat{U}_i \hat{U}_i^\intercal)$
  ▷ $\hat{L}_{mod}$ is the modified Laplacian matrix [12]
  Compute $U \in \mathbb{R}^{N \times k}$
  ▷ U is the matrix of first k eigenvectors [28] of $\hat{L}_{mod}$
  Normalize rows of U to get U_{norm}
  $\{C\}_{i=1}^k \leftarrow$ FINALCLUSTERING(U_{norm})
  ▷ FINALCLUSTERING() is k-means or x-means clustering

  return $\{C\}_{i=1}^k$
  ▷ $C_1, ..., C_k$ are cluster assignment
end function

different values of k ($k = 2..K$, where $K = \sqrt{N}$ [23]). We then take an average among relation values obtained for different k:

$$sim(w, q) = \left(\sum_{k=2}^{K} \left(1 - \frac{||M_{w,k} - M_{q,k}||}{\sqrt{N(N-1)}} \right) \right) \; / \; (K-1).$$

The above formulation is the modified and normalized version of the Partition Difference measurement [29]. Being averaged over different values of k, it is able to serve as a reliable indicator of the similarity between different social networks in terms of clustering results. We explicitly would like to mention that our-proposed inter-layer relationship graph construction approach is purely automated and does not require any expert knowledge. This suggests its further usage for other graph-constraint unsupervised learning approaches.

4.6 Computational time complexity analysis

To analyze the complexity of C^3R clustering, we need to estimate the complexity of each step of Algorithm 1. If N is the number of users, M the number of graph layers (data modalities), and k is the number of first eigenvectors to compute, C^3R time complexity can be estimated as $O(N^2(M^2k + MN + k^2))$. Below, we discuss the complexity in more details.

First, each Laplacian (L_i) and eigenvector matrix (U_i) computational complexity is $O(N^3)$, which sums up to $O(MN^3)$ for computing them for all graph layers. The computation of each $\hat{L}_i$ costs $O(MN^2k)$, which gives the computational complexity of $O(M^2N^2k)$. The total cost of $\hat{U}_i$ computation is $O(MN^3)$. $\hat{L}_{mod}$ can be computed in $O(MN^2k)$ time. The computation of matrix U takes $O(N^3)$ time, while the complexity of x-means clustering in space U_{norm} is $O(N^2k^2)$ [35]. The total C^3R time complexity, thus, is $O(M^2N^2k) + O(MN^3) + O(N^2k^2) = O(N^2(M^2k + MN + k^2))$, where $M, k \ll N$.

5 EVALUATION

5.1 On Community Detection Evaluation

There are two main approaches for evaluating community detection algorithms: direct evaluation and indirect evaluation. Direct evaluation uses a quality measure (e.g. Modularity) to compare community detection results achieved by different algorithms explicitly. However, there is no any widely accepted measure to quantify community detection results in the case of multi-source community discovery. Moreover, many of such quality estimation measures were found to be weakly related to the actual quality of the detected communities [21]. Indirect evaluation, in turn, compares results achieved by approaches from other application domains (e.g. Recommendation, Classification, etc.). Such approaches must be created based on earlier obtained communities. The latter conforms well with our study and allows for evaluating both our proposed cross-source recommendation approach and its backbone — multi-layer community detection approach. In this work, we thus perform the indirect evaluation.

5.2 Dataset

To answer our research questions, we evaluate the C^3R recommendation framework based on largest available multi-source multi-modal cross-region social dataset NUS-MSS [19]. The dataset is provided for three social networks (Twitter, Foursquare, and Instagram), and was collected during the period of *10 July 2014 – 20 Dec 2014* in Singapore, London, and New York [19]. Farseev et al. [19] first collected a set of active users, who have recently posted tweets through the cross-linking functionality of Instagram or Swarm mobile apps. Further, authors utilized Twitter REST API to perform the location-dependent tweets search in three geographical regions. Based on the active user list, Farseev et al. [19] crawled user generated contents for those users, who posted their activities from other social networks on Twitter. For example, each sampled Twitter cross-linking post (e.g. Foursquare check-in) contains a short link to the original check-in page, where the check-in details are available [15].

For every geographical region the following five feature types are provided:

Textual features (3 feature types combined in one vector): 70 LIWC distribution features [36], distribution over 50 LDA topics, and 14 writing style features [19] from Twitter posts, Instagram image captions, and Foursquare "shouts".

Location (Venue Category) Features: distribution over 764 venue categories.

Visual features: distribution over 1000 ImageNet [11] concepts, extracted from Instagram images.

Based on the data collection time frame, Farseev et al. [19] split NUS-MSS dataset into fixed training and test sets. In other words, **check-ins posted by the same users but in different time intervals were used to form training and testing sets**. The *training set* consists of the first 3 months of data, while the *testing set* consists of the last 2 months. Only users who contributed content to all three social networks during both train and test set time frames were included in the evaluation process[1]. This gives rise to 1801 users from Singapore, 813 users from London, and 1602 users from New York. The number of recommendation items (venue categories) in all three geographical regions equals to 764.

5.3 Additionally-Extracted Features

In addition to the features provided by NUS-MSS dataset, we extracted a 48-dimensional temporal feature and 4 mobility features. It is known that online activity of social-media users is tightly knit to temporal and mobility aspects [34], which makes it reasonable to incorporate such data into community detection process. Intuitively, users with similar mobility patterns (i.e. often co-located) and similar temporal patterns (i.e. often perform activities at similar time intervals) may have similar interests and can form an interests-based community. Below, we give a brief description of the mobility and temporal data features that we used to form mobility and temporal layers of user relationship graph.

The mobility features were computed based on users' *areas of interest (AOIs)* [38], which are geographical regions of user's high geo-location density (regardless the geo-location semantic meaning, which could be a geo-located tweet, geo-located Instagram image, or Foursquare check-in). AOIs were obtained by performing density-based clustering[2] [41] over the geo locations of each user and considering the convex hull of each cluster as a new AOI. ε was computed by analyzing the average distance between neighbors in MinPts-distance graph ($MinPts = 3$ was selected empirically) [13]. User's AOIs represent his/her geographical mobility patterns [34, 38] that can be related to user's lifestyle and interests. We extracted the following **Mobility And Temporal** features:

Average number of posts during each of the 8 daytime durations, where each time duration is 3 hours long. The temporal features were computed for each data source with respect to weekday/weekend factor. In total, there were $8 \times 3 \times 2 = 48$ temporal data dimensions computed. The temporal features indicate users' temporal online activity and related to their urban mobility [34].

Number of areas of interest (AOI). This feature reflects the mobility side of user's physical activity. For example, users with the higher number of AOIs may have a physically intense lifestyle. AOIs also represent users' frequent areas of activity (i.e. home, office, university/school) [38] and may indicate how far users are willing to travel on a daily basis.

Median size of user's AOIs[3], which indicates users' mobility inside each AOI. The feature is an indicator of users' traveling habits inside their main activity areas.

Normalized number of AOI outliers. The feature indicates how often users visit places that are not located inside their AOI, which may show how often users deviate from their regular mobility patterns.

[1]The requirement of conducting evaluation based on users with data from all three social networks is dictated by the necessity to make a fair comparison of clustering performance on different data source combinations. Such comparison is only possible in case when the results were obtained based on fixed training and testing sets that consist of the same users.

[2]Most of NUS-MSS's check-ins belong to three geographical regions (Singapore, New York, London), which makes it possible to compute DBScan clusters for most of the NUS-MSS users

[3]Where AOI size is defined as the median distance between the center of mass and all points inside AOI.

Median distance between AOIs. This feature reflects users' mobility at intra-city/inter-city/international level. Specifically, it shows how often and how far users travel between their activity zones and can be useful to infer travel-related user communities.

5.4 Evaluation Measures

To evaluate our framework against competing systems, we chose the following two widely accepted measures:

Normalized Discounted Cumulative Gain ($NDCG$) measure, which is defined as:

$$NDCG@p = \frac{DCG@p}{IDCG@p}, \; DCG@p = \sum_{i=1}^{p} \frac{2^{rel_i}}{\log_2(i+1)}, \; rel_i = \frac{Cat_i}{N_{Cat}},$$

where $IDCG$ is the maximum possible (**I**deal) DCG for a given set of queries, rel_i is the graded relevance of the result at position i, Cat_i is number of times user checked-in at venue of category i, and $N_{Cat} = 764$ is the total number of Foursquare venue categories in the dataset.

Average Precision (AP), which is defined as:

$$AP@p = \frac{1}{\sum_{i=1}^{p} r_i} \sum_{i=1}^{p} r_i \left(\frac{\sum_{j=1}^{i} r_j}{i} \right), \; r_i = \begin{cases} 1, & \text{i is in top p visited cat.)} \\ 0, & \text{otherwise.} \end{cases}$$

In this section, we briefly describe competing recommendation approaches, and different C^3R modifications.

5.4.1 *Recommender System Baselines.*

Popular (POP) — performs recommendation only based on user's past experience (distribution on user's check-ins among 764 Location (Venue Category) features in past). To note, it is the special case of C^3R recommendation (Equation (1)), where $\theta = 0$.

Popular$_{All}$ (POP$_{All}$) — performs recommendation based on aggregated experience of all users, which produces user-independent recommendation output (764 Venue Category features were utilized).

Multi-Source Re-Ranking (MSRR) [18] — linearly combines recommendation results from all data modalities and determines source weights via Stochastic Hill Climbing With Random Restart (SHCR) (See Farseev et al. [19] for details): $w_{tw} = 0.42$, $w_{fsq} = 0.95$, $w_{isnt} = 0.36$, $w_{temp} = 0.65$, $w_{mob} = 0.02$.

Nearest Neighbor Collaborative Filtering (CF) [2] — produces recommendation based on top $k = 20$ (determined by SHCR) most similar users from the recommendation source (Foursquare, in our case) using a similarity measure (in our case, $Heat$ distance) over 764 Venue Category features.

Early Fusion (EF) [55] — fuses multi-source data into a single feature vector (of dimension 1952) and performs recommendation via CF.

Implicit Feedback-Enhanced Singular Value Decomposition (SVD++) [27] — makes use of the "implicit feedback" information in factorization model (trained on 764 Venue Category features), where $\lambda = 0.67$, $\gamma = 0.06$, $k = 277$, $iter = 55$ are obtained by SHCR [19].

Factorization Machines (FM) [39] — brings together the advantages of different factorization-based models via regularization. In our study, we trained FM based on 764 Venue Category features, utilized the $MCMC$ optimization technique [39] and the regression loss, where $k = 30$, $iter = 140$ are obtained by SHCR [19].

5.4.2 $\mathbf{C^3R}$ *And Its Modifications.*

In four $\mathbf{C^3R}$ modifications below, recommendation was performed according to Equation 1, and community detection was performed over all 1952 features.

$\mathbf{C^3R}$ — our proposed C^3R recommendation approach (Equation 1), where user communities are detected as minimization of Equation (8) via the Algorithm 1. The estimated parameters are: $k = 10$, $\alpha = 0.922$, $\beta_i = 1$ ($i = 1..M$), $\gamma = 1$, $\theta = 0.248$, which are inferred by SHCR. The adjacency matrix W_R of the automatically constructed (See Section 4.5) inter-layer similarity graph is given below:

$$W_R = \left(\begin{array}{c|ccccc} & \textbf{txt} & \textbf{loc} & \textbf{vis} & \textbf{tmp} & \textbf{mob} \\ \hline \textbf{txt} & 1 & 0.632 & 0.621 & 0.643 & 0.561 \\ \textbf{loc} & 0.632 & 1 & 0.614 & 0.631 & 0.570 \\ \textbf{vis} & 0.621 & 0.614 & 1 & 0.621 & 0.551 \\ \textbf{tmp} & 0.643 & 0.631 & 0.621 & 1 & 0.560 \\ \textbf{mob} & 0.561 & 0.570 & 0.551 & 0.560 & 1 \end{array}\right)$$

$\mathbf{C^3R_{-\hat{L}}}$ — C^3R recommendation without inter-layer regularization ($\beta_i = 0, i = 1..M$), where $k = 10$, $\alpha = 0.521$, $\gamma = 1$, $\theta = 0.1$ are obtained by SHCR [19].

$\mathbf{C^3R_{-\hat{L}-L_{mod}}}$ — C^3R recommendation without inter-layer regularization ($\beta_i = 0, i = 1..M$) and sub-space regularization ($\alpha = 0$), where $k = 22$, $\gamma = 1$, $\theta = 0.147$ are obtained by SHCR [19].

$\mathbf{C^3R_{-Comm}}$ — C^3R recommendation without user community extraction (all users are considered to be in the same community).

$\mathbf{C^3R}$ **(DBScan)** — C^3R recommendation, where user communities are detected by Density-Based clustering (DBScan) [41]. The DBScan's $\varepsilon = 0.9$ and $MinPts = 6$ were obtained by grid search.

$\mathbf{C^3R}$ **(x-means)** — C^3R recommendation, where user communities are detected by x-means clustering [35]

$\mathbf{C^3R}$ **(Hierarchical)** — C^3R recommendation, where user communities are detected by Hierarchical Clustering. Parameter $k = 10$ and "single linkage" inter-cluster distance measure were obtained by grid search.

5.5 Comparing With Baselines

To answer **RQ1**[4], we evaluate C^3R framework against other recommender systems. The evaluation results are presented in Figure 1.

First, an interesting observation is the poor performance of the EF approach, regarding both *AP* and *NDCG* evaluation metrics, which, once again [18], suggests the necessity of using proper data fusion approaches for multi-source learning. On the other hand, the recommendation approach based on user's past items distribution (POP) produces comparably good recommendation output, especially for the small values of p. It means that Foursquare users tend to re-visit venue categories that they have already visited in past, which highlights the importance of individual knowledge for venue category recommendation and allows *POP* for achieving high recommendation performance. At the same time, the POP_{All} baseline (non-personalized popularity-based recommendation) outperforms FM and EF approaches concerning *NDCG* metric, which suggests the usefulness of group knowledge for recommendation purposes. Finally, it can be seen that the combination of individual and group knowledge in C^3R significantly outperforms other baselines in terms of both *AP* and *NDCG* evaluation metrics. This confidently confirms the necessity of individual and group knowledge integration and **positively answers RQ1**.

Let us also highlight the weak performance of the factorization-based models. For example, even after proper negative sampling [24], FM is not able to perform relevant recommendation in the head of the recommendation list, which explains its poor performance in terms of *NDCG* measure. At the same time, SVD++ fails to outperform others in terms of *Precision*, which could be explained by the limited applicability of the "implicit feedback" concept to the task of venue category recommendation. Specifically, the fact that some users did not visit a venue of a particular category does not necessarily mean that they do not appreciate such venue type. Alternatively, such behavior could be a result of the geographical [34] or social [31] constraints. Another interesting observation is the advance of popularity-based baseline (POP) over other recommendation approaches in New York region (regarding *AP*). The possible explanation is that Foursquare users often re-visit venues of the same category, which means that their corresponding training and test sets significantly overlap in many cases (regarding item distributions). The last is not typical for recommendation systems and may lead to the sub-optimal performance.

5.6 Comparing $\mathbf{C^3R}$ Modifications

To answer **RQ2**[5], we compare different C^3R modifications, which include different variants of our proposed community detection approach as well as other state-of-the-art clustering approaches. The evaluation results are presented in Figure 2.

First, we note that the non-regularized version of the framework ($C^3R_{-\hat{L}-L_{mod}}$) performs the worst against its relatives and other clustering algorithms. This is due to the inability of such non-regularized clustering to build a consistent latent representation, which does not consider inter-network relationship and leads to poor recommendation performance. At the same time, the inter-layer regularized C^3R framework beat all the baselines in all three cities, which could not be achieved by its non-regularized versions ($C^3R_{-\hat{L}-L_{mod}}, C^3R_{-\hat{L}}, C^3R_{-Comm}$). The above observation suggests the importance of subspace regularization in combination with inter-source regularization for balanced user-community detection and, as a result, better recommendation performance. It also allows us to give a **positive response to RQ2**.

Let us also outline the good performance of other clustering approaches, where the recommendation based on x-means clustering (C^3R (x-means)) performs slightly better than other clustering baselines. The above is consistent with the previous study [17] and can be explained by algorithm's ability to automatically infer the parameter k (number of clusters), which allows for finding more distinctive (regarding Bayesian Information Criterion) user communities and, eventually, provide more relevant recommendation.

5.7 Comparing Data Source Combinations

To gain insight into the role of different data sources (modalities) for venue category recommendation and answer **RQ3**[6], we evaluate our framework trained on all data modalities against different

[4]Is it possible to improve the recommendation performance by integrating individual and group knowledge?

[5]Does inter-source relationship information enable us to find better user communities?

[6]What the contribution is of each data source (modality) towards venue category recommendation?

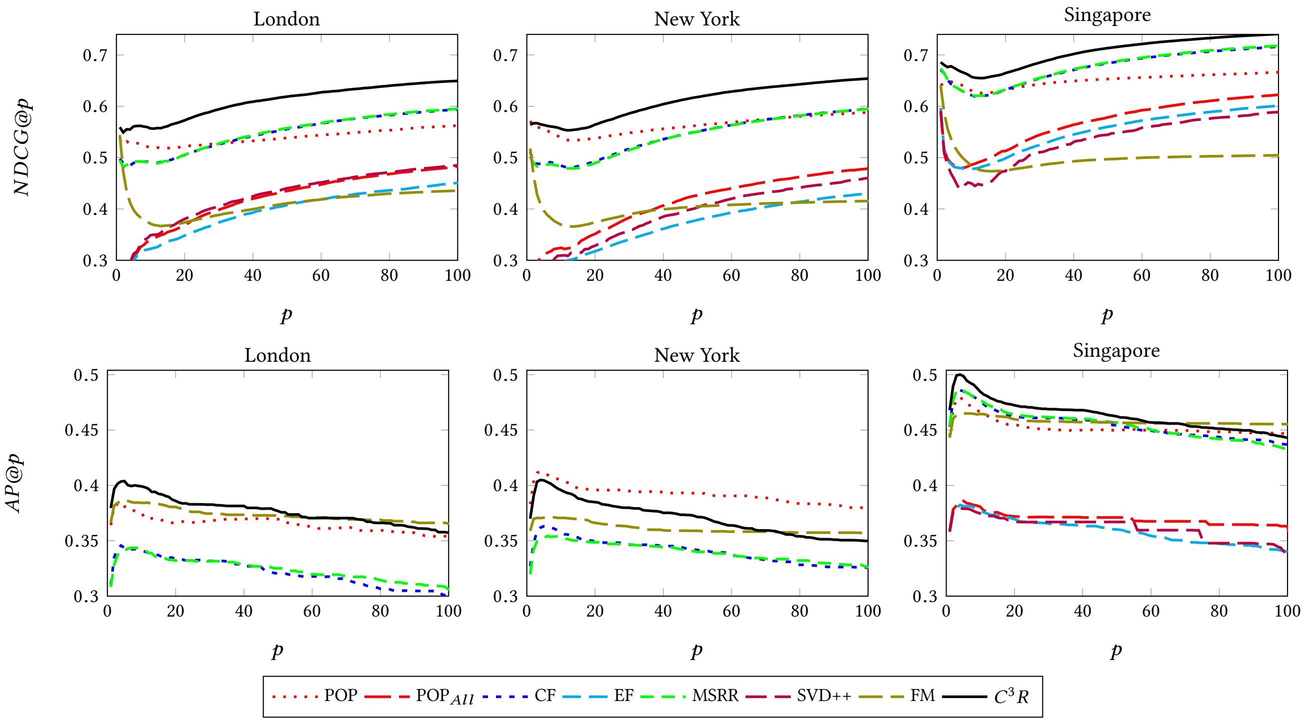

Figure 1: Evaluation of C^3R framework against other recommendation approaches ($NDCG@p$, $AP@p$)

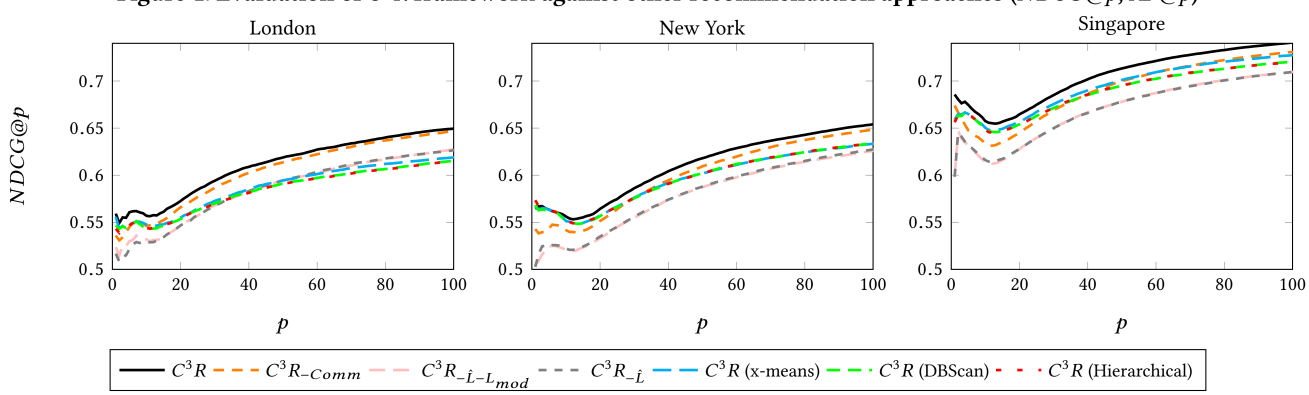

Figure 2: Recommendation performance of different modifications of C^3R framework (different clustering algorithms)

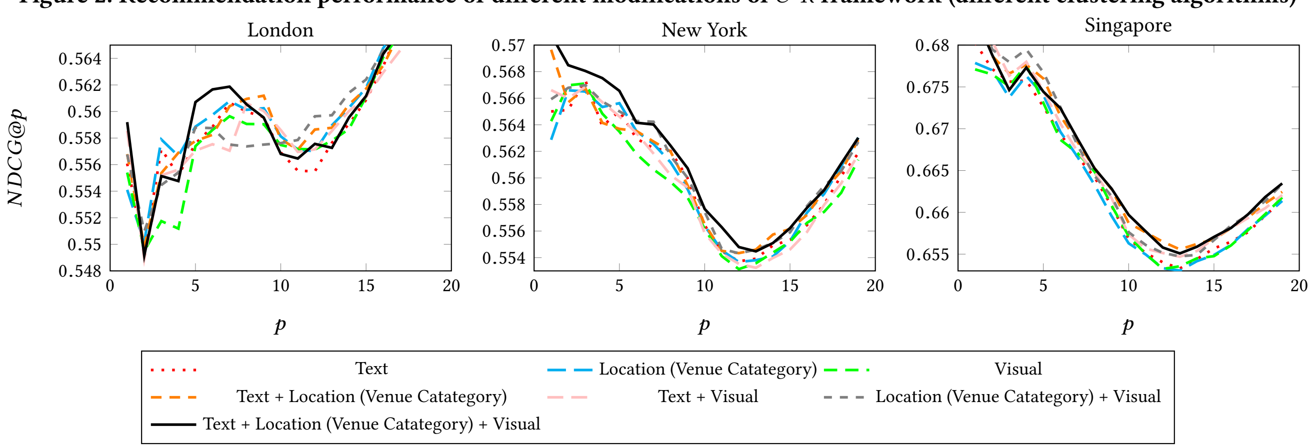

Figure 3: Recommendation performance of C^3R framework based on different data source combinations

data modality combinations. Due to space constraints, in Figure 3 we do not present all feature combinations (31 different combinations), but only show the combinations that include Temporal and Mobility features. For example, the combination "Text" includes Textual, Temporal, and Mobility features, while the combination "Text + Location (Venue Category)" includes Text, Location (Venue Category), Temporal, and Mobility features.

It is not surprising that the recommendation purely based on venue category distribution performs the best among all single-modal baselines in all three cities [18]. The reason is that the venue category data from Foursquare contains explicit knowledge about the distribution of recommendation items in the training set, and thus can forecast the distribution of items in the test set accurately. Bi-source combination results also provide for interesting observations. For example, the Location (Venue Category) + Text combination achieves the best recommendation performance in London and New York regions, but not in Singapore (where text from Twitter gives way to Instagram images). The possible reason is the differences in daily social media usage patterns in different geographical regions: London and New York users mainly post interest-related messages on Twitter, while Singapore users (where Twitter is not widely used) upload pictures of their interests on Instagram (i.e. pictures of food [31]). The above is also supported by the previous study [19], where image data plays a crucial role for the task of multi-source demographic profiling in Singapore. Lastly, we also notice that the combination of all data sources performs the best for all three geographical regions, where the maximum recommendation performance is achieved in Singapore. This could be possibly because of the fact that Singapore has the largest amount of available multi-source data, in comparison with New York and London [19]. Summarizing the above, we **answer the RQ3 by highlighting the importance of venue category, temporal and location-based data** as a major contributors towards recommendation performance. At the same time, depending on geographical region and users' posting behavior, **visual data and textual data may impact differently**, while the **combination of all data sources allows for achieving the best recommendation performance** in most of the cases.

5.8 Qualitative Evaluation

The key idea of our recommendation approach is the incorporation of group knowledge into recommendation via detecting relevant user communities from multiple social multimedia data sources. The evaluation against the baselines indirectly shows the ability of C^3R framework to detect important user communities. To support our answer to **RQ2** and demonstrate the community detection performance explicitly, in Table 2, we have listed the profiles of the 3 largest user communities detected in Singapore. User profiles are constructed as a "bag-of-words" over textual, visual, and location data modalities. From the table, it can be seen that most popular data representations ("words") in all data modalities are consistent with each other and represent distinct user communities. For example, the user community "Com1" is represented by words: "device", "launcher", "android"; visual concepts: "mouse", "digital clock", "hard disk"; and venue categories: "electronics store", "startup", "technology building". We thus named this community as "Gadgets". The distinctness of detected user communities allows C^3R to perform collaborative filtering based on those social media users, who are semantically "close" to the user of the recommendation system, which helps to boost the recommendation performance. The inter-modal consistency of the community profiles shows that C^3R can perform balanced clustering on a multi-layer graph, which is achieved via inter-layer regularization. The overall results suggest the applicability of the proposed techniques for a task of multi-source user community detection and encourage further research along these directions.

Table 2: User communities profiles

Community	Bag of Words for different modalities		
	Text	Visual	Location
(Com1) "Gadgets" 832 users	device, launcher, android	mouse, digital clock, hard disc	electronics store, tech. building
(Com2) "Arts" 538 users	painting, landscape, reflection	obelisk, paint-brush, pencil box	arts & crafts store, arts & museum
(Com3) "Food" 446 users	dining, coffee, cooking	pineapple, microwave, frying pan	italian restaurant, pizzeria, restaurant

6 LIMITATIONS

Although C^3R outperforms the baselines, there are several limitations that we would like to highlight. First, our proposed multi-source clustering approach requires simultaneous availability of data from all three social networks. This does not allow for its usage in a cold-start settings. Such problem could be addressed by introducing multi-source data completion prior to clustering process, or by performing late model fusion [18]. Another limitation is the time complexity of the framework, which is introduced by the graph Laplacians computation. The issue could be potentially resolved via reformulating the spectral clustering problem in a smaller space [10]. Particularly, by replacing $N \times N$ graph Laplacians with $P \times P$ matrices ($P \ll N$), a solution of mathematically-equivalent problem in significantly reduced space could be obtained. The final clustering could be then recovered from the obtained solution, while P can be chosen to be much smaller than N so that the running time could grow almost linearly to N [10]. We plan to address the above limitations in our future work.

7 CONCLUSION

In this paper, we presented a study on cross-source collaborative recommendation. The proposed recommendation framework C^3R utilizes both individual and group knowledge to solve a task of venue category recommendation. Individual knowledge is modeled as the distribution among venue categories that user has visited in past, while group knowledge is modeled as the distribution of Foursquare venue categories among user community members. The user communities are detected based on novel regularized spectral clustering approach that is able to perform an efficient partitioning of multi-layer user relations graph. By performing comprehensive evaluation against the state-of-the-art baselines and different data source combinations, we demonstrated that our proposed user-community-empowered venue category recommendation framework can achieve superior recommendation performance. Additionally, we contributed a new fully-automated method for inter-network relationship graph construction, which eliminates the necessity of involving a pre-defined expert knowledge.

8 ACKNOWLEDGMENTS

NExT++ research center is supported by the National Research Foundation, Prime Minister's Office, Singapore under its IRC@SG Funding Initiative.

REFERENCES

[1] F. Abel, S. Araújo, Q. Gao, and G.-J. Houben. Analyzing cross-system user modeling on the social web. In *International Conference on Web Engineering*. Springer, 2011.

[2] 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*, 2005.

[3] M. Akbari, X. Hu, N. Liqiang, and T. Chua. From tweets to wellness: Wellness event detection from twitter streams. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

[4] M.-D. Albakour, R. Deveaud, C. Macdonald, I. Ounis, et al. Diversifying contextual suggestions from location-based social networks. In *Proceedings of the 5th Information Interaction in Context Symposium*, pages 125–134. ACM, 2014.

[5] A. Ambrus, B. Greiner, P. Pathak, et al. Group versus individual decision-making: Is there a shift. *W. Paper 91*, 2009.

[6] K. Buraya, A. Farseev, A. Filchenkov, and T. Chua. Towards user personality profiling from multiple social networks. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

[7] R. Burke. Integrating knowledge-based and collaborative-filtering recommender systems. In *Proceedings of the Workshop on AI and Electronic Commerce*, 1999.

[8] P. R. Center. Social networking fact sheet. http://www.pewinternet.org/fact-sheet/social-media.

[9] J. Chen, H. Zhang, X. He, L. Nie, W. Liu, and T. Chua. Attentive collaborative filtering: Multimedia recommendation with item- and component-level attention. In *Proceedings of the 40th ACM SIGIR Conference*. ACM, 2017.

[10] Y. Chen, C. Zhuang, Q. Cao, and P. Hui. Understanding cross-site linking in online social networks. In *Proceedings of the 8th Workshop on Soc. Net. Mining and Analysis*, 2014.

[11] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009.

[12] X. Dong, P. Frossard, P. Vandergheynst, and N. Nefedov. Clustering on multi-layer graphs via subspace analysis on grassmann manifolds. *IEEE Transactions on Signal Processing*, 2014.

[13] D. Eppstein, M. S. Paterson, and F. F. Yao. On nearest-neighbor graphs. *Disc. & Comp. Geometry*, 1997.

[14] A. Farseev, M. Akbari, I. Samborskii, and T. Chua. 360° user profiling: past, future, and applications. *ACM SIGWEB Newsletter*, (Summer), 2016.

[15] A. Farseev and T. Chua. Tweet can be fit: Integrating data from wearable sensors and multiple social networks for wellness profile learning. *ACM Transactions on Information Systems (TOIS)*, 2017.

[16] A. Farseev and T. Chua. Tweetfit: Fusing multiple social media and sensor data for wellness profile learning. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. AAAI*, 2017.

[17] A. Farseev, N. Gukov, I. Gossoudarev, and U. Zarichnyak. Cross-platform online venue and user community recommendation based upon social networks data mining. *Computer Instruments in Education*, 6:28–38, 2014.

[18] A. Farseev, D. Kotkov, A. Semenov, J. Veijalainen, and T. Chua. Cross-social network collaborative recommendation. *ACM International Conference on Web Science*, 2015.

[19] A. Farseev, L. Nie, M. Akbari, and T. Chua. Harvesting multiple sources for user profile learning: a big data study. In *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, 2015.

[20] A. Farseev, I. Samborskii, and T. Chua. bbridge: A big data platform for social multimedia analytics. In *Proceedings of the 2016 ACM on Multimedia Conference*, pages 759–761. ACM, 2016.

[21] S. Fortunato. *Community detection in graphs.* Physics Reports, 2010.

[22] G. H. Golub and C. F. Van Loan. *Matrix computations.* 2012.

[23] J. A. Hartigan and M. A. Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 1979.

[24] X. He, H. Zhang, M.-Y. Kan, and T. Chua. Fast matrix factorization for online recommendation with implicit feedback. In *Proceedings of the 39th ACM SIGIR Conference*, 2016.

[25] S. Helgason. The radon transform on euclidean spaces, compact two-point homogeneous spaces and grassmann manifolds. *Acta Mathematica*, 1965.

[26] R. Hu and P. Pu. Helping users perceive recommendation diversity. In *Proceedings of the Workshop on novelty and diversity in recommender systems.*, 2011.

[27] Y. Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2008.

[28] R. B. Lehoucq and D. C. Sorensen. Deflation techniques for an implicitly restarted arnoldi iteration. *SIAM Journal on Matrix Analysis and Applications*, 1996.

[29] T. Li, M. Ogihara, and S. Ma. On combining multiple clusterings. In *Proceedings of the 13th ACM international conference on Information and knowledge management*, 2004.

[30] H. Lutkepohl. Handbook of matrices. *Computational Statistics and Data Analysis*, 1997.

[31] Y. Mejova, H. Haddadi, A. Noulas, and I. Weber. # foodporn: Obesity patterns in culinary interactions. In *Proceedings of the 5th International Conference on Digital Health*, 2015.

[32] A. S. Niaraki and K. Kim. Ontology based personalized route planning system using a multi-criteria decision making approach. *Expert Systems with Applications*, 2009.

[33] L. Nie, L. Zhang, M. Wang, R. Hong, A. Farseev, and T.-S. Chua. Learning user attributes via mobile social multimedia analytics. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 8(3):36, 2017.

[34] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo. Mining user mobility features for next place prediction in location-based services. In *Proceedings of the Data mining (ICDM), 2012 IEEE 12th international conference on*, 2012.

[35] D. Pelleg, A. W. Moore, et al. X-means: Extending k-means with efficient estimation of the number of clusters. In *ICML*, volume 1, 2000.

[36] J. W. Pennebaker, M. E. Francis, and R. J. Booth. Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 2001.

[37] S. Qian, T. Zhang, R. Hong, and C. Xu. Cross-domain collaborative learning in social multimedia. In *Proceedings of the 23rd Annual ACM Conference on Multimedia*, 2015.

[38] Y. Qu and J. Zhang. Trade area analysis using user generated mobile location data. In *Proceedings of the 22nd international conference on World Wide Web*, 2013.

[39] S. Rendle. Factorization machines with libFM. *ACM Trans. Intell. Syst. Technol.*, 3(3).

[40] D. Rhouma and L. B. Romdhane. An efficient algorithm for community mining with overlap in social networks. *Expert Systems with Applications*, 41(9), 2014.

[41] J. Sander, M. Ester, H.-P. Kriegel, and X. Xu. Density-based clustering in spatial databases: The algorithm gdbscan and its applications. *Data mining and knowledge discovery*, 1998.

[42] 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*, 2001.

[43] J. Shi and J. Malik. Normalized cuts and image segmentation. *Transactions on Patt. Anal. and M. Intelligence*, 2000.

[44] X. Song, L. Nie, L. Zhang, M. Liu, and T. Chua. Interest inference via structure-constrained multi-source multi-task learning. In *Proceedings of the International Joint Conference on Artificial Intelligence*, 2015.

[45] W.-C. Su. Integrating and mining virtual communities across multiple online social networks: Concepts, approaches and challenges. In *4th International Conference on digital Inf. and comm-n Tech. and it's App.* IEEE, 2014.

[46] J. C. Teixeira and A. P. Antunes. A hierarchical location model for public facility planning. *European Journal of Operational Research*, 2008.

[47] A. Tiroshi, S. Berkovsky, M. A. Kaafar, T. Chen, and T. Kuflik. Cross social networks interests predictions based ongraph features. In *Proceedings of the 7th ACM conference on Recommender systems*. ACM, 2013.

[48] S. Trewin. Knowledge-based recommender systems. *Encyclopedia of Library and Information Science: Volume 69-Supplement 32*, 2000.

[49] M. H. Veiga and C. Eickhoff. Privacy leakage through innocent content sharing in online social networks. *arXiv preprint arXiv:1607.02714*, 2016.

[50] U. Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 2007.

[51] D. Wagner and F. Wagner. *Between min cut and graph bisection.* Springer, 1993.

[52] X. Wang, X. He, L. Nie, and T. Chua. Item silk road: Recommending items from information domains to social users. In *Proceedings of the 40th ACM SIGIR Conference*. ACM, 2017.

[53] X. Wang, L. Nie, X. Song, D. Zhang, and T. Chua. Unifying virtual and physical worlds: Learning toward local and global consistency. *ACM Transactions on Information Systems (TOIS)*, 36(1):4, 2017.

[54] X. Wang, Y.-L. Zhao, L. Nie, Y. Gao, W. Nie, Z.-J. Zha, and T. Chua. Semantic-based location recommendation with multimodal venue semantics. *Transactions on MM*, 2015.

[55] P. Winoto and T. Tang. If you like the devil wears prada the book, will you also enjoy the devil wears prada the movie? a study of cross-domain recommendations. *New Generation Computing*, 2008.

[56] M. Yan, J. Sang, and C. Xu. Unified youtube video recommendation via cross-network collaboration. In *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, 2015.