Motivation
Social Collaborative

Our Approach

Related Work

Learning Parameters

Expectation-Maximization Algorithm
Maximization of posterior is equivalent to maximizing the complete log-likelihood of $U, V, S, D, P$ and $Q$ given $\lambda_u, \lambda_D, \lambda_Q$ and $\beta$

Update Equations

Optimize this function by gradient ascent approach by iteratively optimizing the collaborative filtering and social network variables $u_i, v_j, s_{ij}$ and topic proportions $\beta_j$

Given U and V, we can learn topic proportions $\beta$. We cannot optimize $\beta$ analytically, so we use projection gradient approaches

Experiments

Dataset Description

Results

Evaluation

In-matrix prediction: Item rated by at least one user $e_{ij}(D) = e_{ij}[D] = e_{ij}(D_i) + e_{ij}(D)$

Out-matrix prediction: New item (no ratings)

We consider ‘recall’ as evaluation metric

Fig 1: In-matrix recall comparisons (Left: Lastfm, Right: Delicious)

Fig 2: In-matrix recall prediction. Left: Comparing our model with state-of-the-art approaches. Right: Impact of social network parameter $\lambda$ by varying content parameter $\lambda_c$

Fig 3: In-matrix prediction recall by varying content parameter $k$ and social network parameter $k_m$ at $K=250$. Top: Lastfm, Bottom: Delicious dataset

Complexity Analysis

Conclusions

Contributions

Future work

Using final ‘static’ social network in recommendation systems could lead to potential information leak. Our model could be making better predictions before social network information. Fig. 4 confirms information leak. We observe that smaller training dataset (sparser social graph) results in larger information leak.

Fig 4: Recall of proposed model by varying social network structure. Dataset used: Delicious

Social Information Leak

Introduction

Motivation

Social Networks have become important platform for content, opinion sharing and provide rich information to study social circle’s influence on user’s decision process

Problem: How can social networks help improve the social recommendation systems? Does information leak occur in social recommendation systems?

Solution: We propose a novel hierarchical Bayesian model which jointly incorporates content topic modeling and matrix factorization of social networks. Empirical experiments on two large-scale datasets show that our algorithm outperforms the state-of-the-art approaches. Our results reveal interesting insight that the social circles have more influence on people’s decisions about the usefulness of information (e.g., bookmarking preference on Delicious) than personal taste (e.g., music preference on Lastfm)

Collaborative Topic Regression with Social Network Factorization for Recommendation Systems
Sanjay Purushotham, Yan Liu, C.-C. Jay Kuo

Related Work

Collaborative Filtering (CF) approaches predict user preferences based on collective rating records of similar users or items. However, CF-based methods suffer from sparsity problem and imbalance of rating data, especially for new and infrequent users

Latent factor methods such as Probabilistic Matrix Factorization (PMF) incorporate user interests into the CF-model. However, users are assumed to be i.i.d and additional information is ignored

Collaborative Topic Regression (CTR) incorporates content information via LDA into collaborative filtering framework. CTR represents users with topic interests and assumes that items are generated by a topic model

Social recommendation (SoRec) based, on matrix factorization, uses social network information and user ratings to recommend items

Current approaches cannot predict ratings for new/unseen items and new/infrequent users of social network

Our Approach

Our model is generalized hierarchical Bayesian model which jointly learns the user, item and social factor latent spaces. We use LDA to capture item’s content information in latent topic space, and we use matrix factorization to derive latent feature space of user from his social network graph.

Conditional distribution over social network relationships:

Through Bayesian inference, combining LDA with social matrix factorization, we get:

$\approx p(U_i, V_j, S, D, P | Q, U, V, S, D, P)$

Learning Parameters

Fig 2: Future implementation of our algorithms and capturing social network dynamics into our model

Fig 1: Future implementation of our algorithms and capturing social network dynamics into our model


Contact information: Sanjay Purushotham, Email: spurusho@usc.edu