**基于情境的社交网络推荐**

**一． 引言**

随着社交网络的发展，极度爆炸的信息使得用户越来越不容易获得所需的数据。因此社交网络中的个性化推荐也已成为其一项重要的主动服务机制。它需要从大量的信息中主动选择符合用户兴趣和需要的内容，以满足用户的个性化需求。

传统的个性化推荐技术[1]主要有基于内容的推荐和协同过滤，以及其混合推荐技术。在社交网络中，协同过滤的方法已经不能用来完全刻画复杂的社交网络关系。而在社交网络中的推荐中，研究人员引入了基于用户间信任(Trust-base)和基于用户间影响(Influence-base)的方法来利用社交网络中用户之间的关系，从而使社交网络中的推荐更为精确。

无论是传统的推荐方法和现有社交网络推荐都是对观测到的用户评分矩阵，用户特征矩阵，项目特征矩阵等二维矩阵进行处理，从而得到预测的用户评分矩阵来描述用户对项目的偏好。这些二维的推荐机制忽略了情境对用户行为的影响，只考虑了用户和项目两个维度，没有考虑情境这一维度。

在现实生活中的社交场合，人们除了考虑交流的人和话题本身之外，会根据当时的情境来选择交流的对象和话题。也就是说，在动态环境中，人对人的影响，以及人对事物的喜爱度会随着情境的变化而变化。如，当一个人有烦恼或困扰时，他会选择同沉稳的或者年长有经验的人交流，而交流的内容可能从他平时喜欢的体育等项目转为人生哲理类；而当他处于工作期间，他交流的对象就会转向与他工作相关的同事，交流的内容也与工作相关；然而同样是工作期间，时间却是即将到来的周末的前一天周五，他和同事交流的内容就可能会更多的与周末活动相关。

在与真实的社会生活相近的社交网络中，影响用户行为的因素主要有如图1中三个方面。图1描述了一个社交网络中用户选择和过滤信息的完整情境。用户会根据自己所处的情境，信息的发送者和其的内容，对社交网络中的信息进行筛选，选择自己当时情境下感兴趣的内容。所以在建立社交网络中用户的行为模型时，既要考虑到(1)用户的个人喜好，要考虑(2)用户之间的影响，同时也要考虑(3)用户所处的情境。

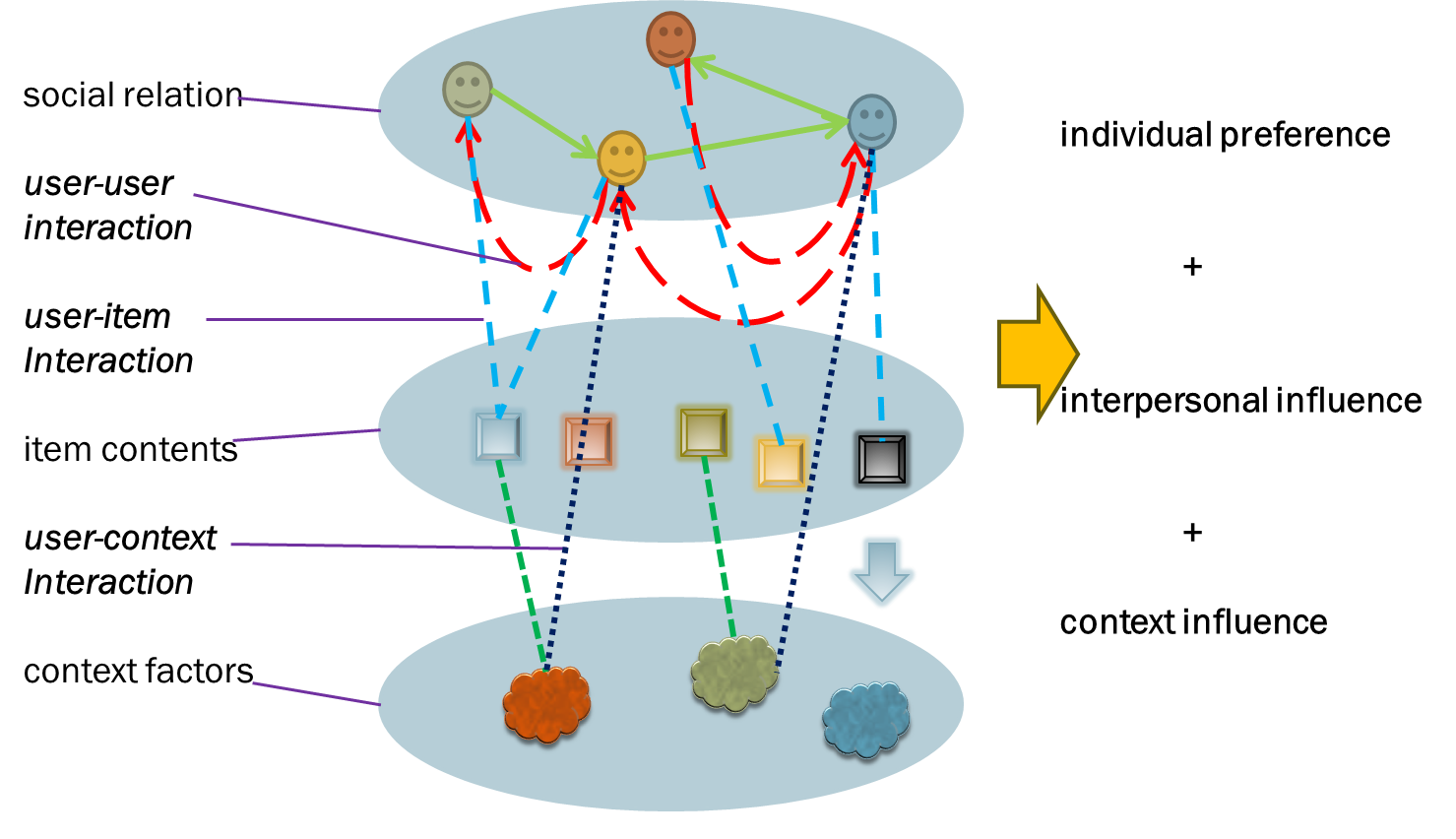


图1

在社交网络推荐中，为了提高推荐的精确性，我们提出了一个基于情境的三维社交网络推荐模型，将情境这一第三维度引入到原有的用户项目的二维模型中。

我们利用已有的社交网络新浪微博的数据进行实验和推荐模型的验证。首先我们爬取新浪微博的数据，然后分析识别数据中的情境数据，发送者和项目内容。由于现在的社交网络对情境感知的渠道和内容比较局限，情境数据主要有包括通过移动设备感知的地点数据，时间数据，以及用户通过新浪心情接口发送的心情，及分析用户表情得到的情境数据。我们将实验数据分为训练集和测试集，采用最常用的推荐度量方法MAE平均绝对误差作为评价标准。

本文组织如下。章节2介绍相关工作。章节3具体介绍如何对用户，项目及情境进行抽象和建模。章节4将对我们的基于情境的社交网络推荐模型公式化。章节5介绍我们的实验设计和实验结果。章节6进行总结。

**二．相关工作**

传统的个性化推荐技术[1]主要有基于内容的推荐和协同过滤，以及其混合推荐技术。协同过滤技术是应用最为广泛和成熟的推荐方法。协同过滤技术过又可分为基于邻居（基于用户或基于项目）的协同过滤和基于模型的协同过滤。基于用户的协同过滤技术[?]主要建立在相似用户具有相似兴趣这一假设的基础上。根据用户对项目的历史评价，找到与当前用户相似的邻居用户，然后根据邻居用户对项目的评价，预测用户对其评价。同样，基于项目的协同过滤[]的基本原理就是用所有用户对项目的偏好，发现物品和物品之间的相似度，然后根据用户的历史信息，将类似的物品推荐给用户。基于邻居的协同过滤主要会用Pearson Correlation Coefficient [13]和Vector Space Similarity[]算法和奇异值分解SVD/LSI[3]。基于模型的协同过滤方法主要是用已有的用户评分矩阵，来训练一个推荐模型，然后根据实时的用户喜好的信息进行预测，计算推荐。用到的算法包括Personality Diagnosis[17]，Bayesian Network Model[8]和Clustering Model[]。

在社交网络中，协同过滤的方法已经不能用来完全刻画复杂的社交网络关系。而在社交网络中的推荐中，研究人员引入了基于用户间信任(Trust-base)[2][9][11][22]和基于用户间影响(influence-base)[4][10][16][18]的方法来利用社交网络中用户之间的关系，从而使社交网络中的推荐更为精确。但是Trust-base的推荐是基于同质原则(即用户同他信任的人偏好相似)的假设，该假设并不完全符合现实，这种方法只适用于trust机制的网络，而不是真正的社交网络。而Huang [10]引入了社交网络中用户间相互的影响因素，却忽略了用户本身的影响因素。Jiang[4]综合考虑了社交网络的情境，提出了一个新的社交网络中的用户行为模型，即用户的行为受到用户间的影响和用户的偏好两者的共同作用，较全面的刻画了用户在社交网络中的行为，但是忽略了在不同情境下，用户间的影响和用户的偏好是变化的。在推荐方法的计算效率方面，所有的方法所以矩阵分解[6][7]和矩阵正规化[5]等的方法也被提出用来计算海量的社交网络数据，以提高计算效率。但是无论是传统的推荐方法和现有社交网络推荐都是对观测到的用户评分矩阵，用户特征矩阵，项目特征矩阵等二维矩阵进行处理，从而得到预测的用户评分矩阵来描述用户对项目的偏好。这些二维的推荐机制忽略了情境对用户行为的影响，只考虑了用户和项目两个维度，没有考虑情境这一维度。在本文中我将情境对用户间影响和用户偏好的作用应用的我们的社交网络推荐中。

**三．基于情境的社交网络推荐模型**

**情境定义**

情境的定义在不同的领域会有所不同，本文采用Schmidt的定义[20]：情境描述一个设备或用户所处的态势和环境，一个情境可以用一个唯一的名字来标识，每个情境都拥有一组相关的特征和描述该特征的值来确定。

在社交网络中，情境描述用户在浏览社交网络时所处的状态和环境，包括外在环境因素和用户本身状态的因素。由于网络的限制，用户浏览社交网络时的外在环境因素具有局限性主要包括时间和场所，而用户本身因素可以描述为心情等。例如，用户可能在家里，工作地点或公共场所浏览社交网络信息，可以在一天的不同时间段浏览信息，浏览社交网络时用户或开心或不开心或无感觉。

**基于情境的三维模型**

维度是用来表示推荐空间的范围。除了用户维U和项目维I，还有情境维C。推荐空间S就是用户维，项目维和情境维的笛卡尔乘积：

传统的推荐系统中，通过用户对项目评分来表示对项目的喜爱。通过不同方法预测用户的评分二维矩阵R来进行推荐。

在考虑了情境这一维度的多维的推荐模型中，我们通过计算预测推荐空间S的值来进行。如图多维空间的评分矩阵可以直观的表现出在某一情境下用户的项目的评分R(I,U,C)。在社交网络中R为采纳举证，即值1表示用户u采纳了项目i，值0表示没有采纳。故：。

推荐空间S中用户的各种评分结果可以直观地用如下图3的三维立体表示，每个小方块表示一个用户在一个情境下对一个项目的采纳情况，在本文中我们成为采纳矩阵R。

Rijk=

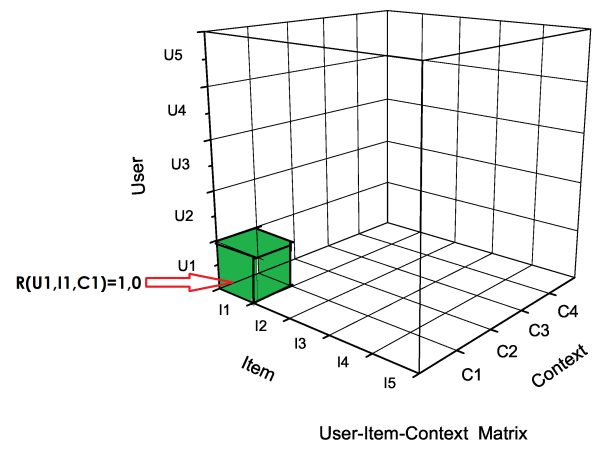
****

图2 用户-项目-情境 采纳矩阵

**社交网络推荐算法**

基于情境的社交网络推荐就是要预测出用户对项目的采纳矩阵R，然后根据R的值进行推荐。

根据图1对用户在社交网络中行为的分析，我们提出了一个如图3所示的基于情境的多维概率矩阵社交网络推荐模型。该模型综合考虑了两种主要因素：1不同情境下用户之间的影响度，和2不同情境下用户的个人偏好变化。从而使推荐结果更具准确性和精确性。这样原有的评分矩阵R(User-Item)就需要从用户偏好和用户间影响两方面计算得到(分为User-Item Influence矩阵和User-Item Preference矩阵)，并将二维的评分矩阵R扩展到三维，由User-Item 变为User-Item-Context三维矩阵，再将其分解为User-Item-Context Influence矩阵和User-Item-Context Preference矩阵。

根据历史数据，我们用LDA[]来计算项目Item的计算其话题分布，从而得到Item在不同情境下的Item Latent Feature矩阵，根据用户采纳项目的历史数据，可以计算得到在不同情境下的用户对不同Latent的偏好度(User Latent Feature Matrix)，同时，从而可以计算得到User-Item-Context Preference矩阵(1)。另一方面，根据采纳项目的发送者信息，计算得到User-User-Context Influence 矩阵，然后结合项目Item的实际发送者情况(Item Sender Matrix)得到User-Item-Context Influence矩阵(2)。最后由(1)(2)计算得到User-Item-Context评分矩阵。这样就可以根据该矩阵和用户当前的情境为用户进行推荐。

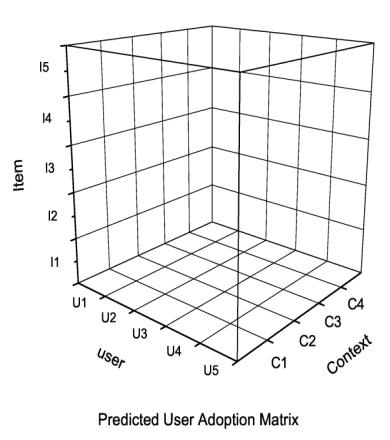
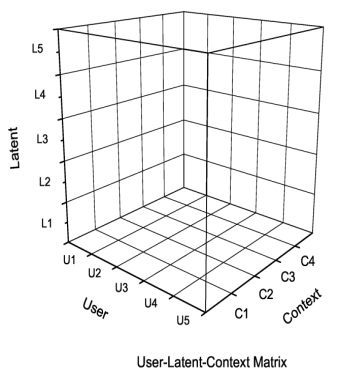
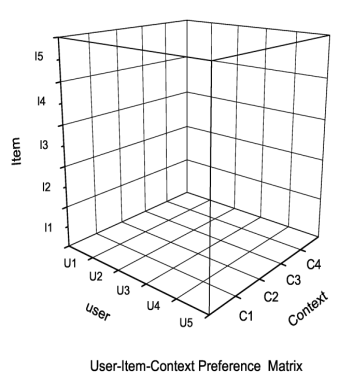
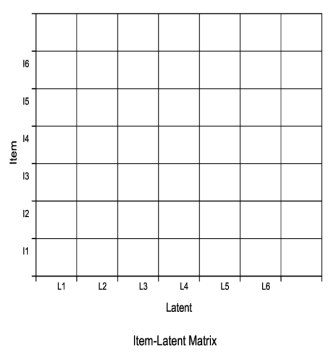
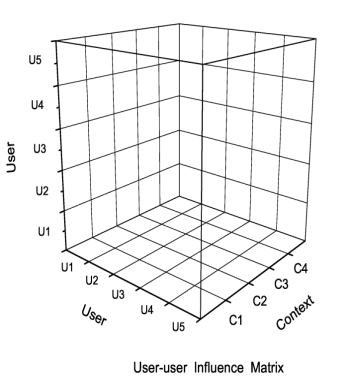
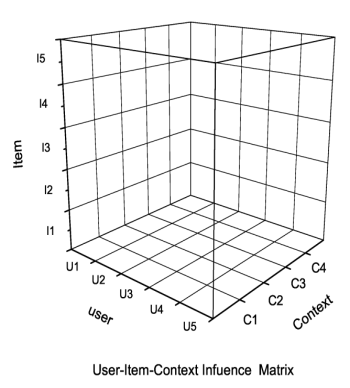
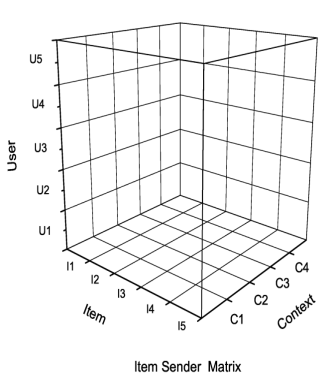


图3：基于情境的社交网络推荐模型

在我们的模型中，我们假设用户是否采纳一个项目取决于4个因素：1项目本身内容，2项目的发送者，3用户对项目偏好，和4当时情境。

首先，我们定义有m个用户，n条项目，k种情境。用矩阵M表示项目发送者(Item Sender Matrix)，Mij=1表示用户uj发送了项目pi；矩阵N为用户间影响矩阵(User-User-Context Influence Matrix)，矩阵中Nijk表示ui对用户uj在情境ck的影响力。用矩阵G表示用户在不同情境下对项目不同latent的偏好(User Latent Context Matrix)，用矩阵S表示项目在话题上的分布(Item Latent Matrix)。这样我们就可以通过Hadamard乘积来近似计算用户采纳矩阵R了。

我们用情境向量Cx(c1,c2,…,ck)表示情境x，其中c1,c2,…,ck 分别表示情境x在k种情境中的取值。

对于项目latent分布矩阵S，我们用LDA模型对每个项目内容进行话题建模，向量Ta表示项目a的话题分布，所有的项目的话题分布向量就构成了矩阵S。

在不同情境下用户v用户u的影响向量H为：

其中S(u,v)是用户v发送给u的项目集合，A(u)用户采纳的项目集合，Ca’为用户采纳项目a’时的情境向量。所有这些向量构成三维矩阵N。

在不同情境下用户对项目偏爱的话题分布二维矩阵为：

其中A(u)用户采纳的项目集合，Ca’为用户采纳项目a’时的情境向量，Ta’为话题分布向量。这些所有的二维举证就构成了三维矩阵G。

G表示用户在不同情境下对项目不同latent的偏好：已知用户u,项目a，当时他情境x。求x情境下用户对项目的偏爱度：

用户影响矩阵I：

其中V(u,a)是所有转发项目a给用户u的用户集合，A(u)是用户u采纳项目的集合，S(u,v)是用户v发送给u的项目集合，Ca’为用户采纳项目a’时的情境向量。

偏爱度矩阵P：

其中为项目a的话题分布向量，为项目a评测的情境向量值，同样为项目a’采纳时的情景向量的值，为用户发送的所有除a外的项目集合， 为集合中元素的个数。

基于情境的社交网络推荐算法：

输入：用户u，待推荐项目a，用户情境x。

输出：目标用户对待评测项目a的评分ruax。

1V(u,j)发送项目j给用户u的用户集合，A(u)用户采纳的项目集合，S(u,v)用户v向用户u发送的项目集合。

2

3

情境相似度算法：

**四．实验**

在这一章节中，我们用真实的社交网络数据进行了实验，并与其它推荐方法进行了比较。

**实验描述**

我们利用已有的社交网络新浪微博的数据进行实验和推荐模型的验证。新浪微博是在中国类似于Twitter的著名社交网络，人们可以关注自己感兴趣的人，然后收到这些人所发的信息，用户可以收藏，转发和赞这些微博信息。信息发送时可以附带表情，地理位置信息，新浪微博还提供了发送心情的接口。首先我们爬取新浪微博的数据，然后分析识别数据中的情境数据，发送者和项目内容。由于现在的社交网络对情境感知的渠道和内容比较局限，情境数据主要有包括通过移动设备感知的地点数据，时间数据，以及用户通过新浪心情接口发送的心情，及分析用户表情得到的情境数据。实验数据统计如下表：

|  |  |
| --- | --- |
| Statistic | Value |
| User Number |  |
| Tweet Item Number |  |
| Forward Behavior Number |  |

实验目标：我们的实验需要对用户的正在阅读的项目进行预测。实验设置：与其他网络用户浏览静态信息的不同，在社交网络中，用户所见的信息是实时更新的，用户不一定能看到所有发给他的信息。同时我们不能获得用户在线的时间，也不能获得用户阅读微博时的实时情境信息，而只能根据用户的已有的行为判断用户的在线时间和情境信息。所以我们根据用户的有效行为（如发表微博，转发、收藏微博来）来判断用户在线，对于每个有效行为，我们定义一个有效时间段online session（该行为前后一段时间），表示在这个时间段内用户在线。我们从新浪微博中提取三种情境信息（时间，地点，心情），因为用户行为并不一定包含所有情境信息，我们将包含两种或以上的行为作为情境有效行为，如图4，在该行为有效时间段内的其他用户行为视为有效行为，同时从用户该行为得到情境信息持续这个时间段。这个有效时间段信息的设置，对于其他对比算法同样有效。

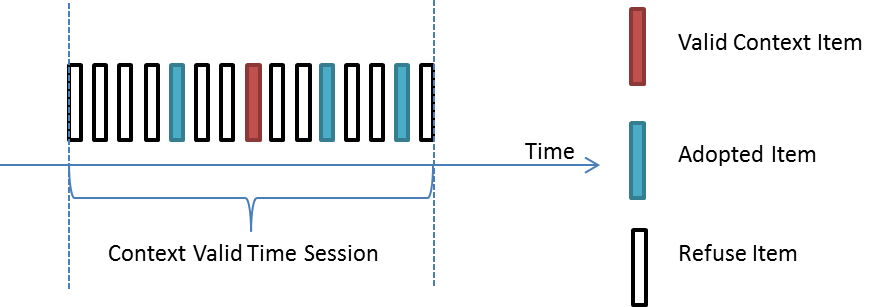


图4：情境数据的时效性

我们将实验数据分为训练集和测试集，从训练集中用我们的模型计算出用户间在不同情境下的用户间影响和项目特征矩阵，不同情境下用户的特征矩阵。同时对于我们从测试集中提取有效的时间段。然后用户对比算法和我们的模型进行计算。

//图

**对比算法：**

**度量方法：**

采用最常用的推荐度量方法Mean Absolute Error(MAE)平均绝对误差和Root Mean Square Error (RMSE)作为评价标准，通过计算预测评分和真实评分之间的偏差来度量推测的准确性，MAE和RMSE值越小则推荐的质量越高。

MAE方法定义如下：

RMSE方法定义如下：

其中表示用户*i*对项目*j*的评分，表示预测的用户评分，T表示测试的评分数量。

**实验结果和分析：**

**五．总结和未来工作**

**参考文献**

[1] Emmanouil Vozalis, Konstantinos G. Margaritis, Analysis of Recommender Systems’ Algorithms, 2003

[2] P. Bedi, H. Kaur, and S. Marwaha. Trust based recommender system for semantic web. In Proc. of

IJCAI’07, 2007

[3] Bardul M. Sarwar, Sparsity, Scalability, and Distribution in Recommender Systems, Ph.D. thesis, University of Minnesota, 2001

[4] Meng Jiang,Peng Cui, Rui Liu，Social Contextual Recommendation,2012

[5] Hao Ma,Denglong Zhou, Recommender Systems with Social Regularization,2011

[6] Y. Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In Proceeding of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), pages 426–434, 2008.

[7] Y. Koren. Collaborative filtering with temporal dynamics. In Proceeding of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2009.

[8] John S. Breese, David Heckerman, and Carl Kadie, Empirical analysis of predictive algorithms for collaborative filtering, 1998

[9] Samaneh Moghaddam, Mohsen Jamali, Martin Ester. FeedbackTrust: Using Feedback Effects in Trust-based Recommendation Systems, 2011

[10] J. Huang, X. Cheng, J. Guo, H. Shen, and K. Yang. Social recommendation with interpersonal influence.In Proceedings of the 19th European Conference on Artificial Intelligence (ECAI), pages 601–606, 2010

[11] M. Jamali and M. Ester. Trustwalker: a random walk model for combining trust-based and item-based recommendation.In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2009

[12] HAO MA, IRWIN KING, MICHAEL R. LYU. Thomas Kriechbaum, Learning to Recommend with Explicit and Implicit Social Relations, 2011

[13] Paul Resnick, Neophytos Iacovou, Mitesh Sushak, Peter Bergstrom, and John Riedl, Grouplens: An open architecture for collaborative filtering of netnews , 1994

[14] Shuang Hong Yang, Bo Long, Alex Smola, Like like alike — Joint Friendship and Interest Propagation in Social Networks, 2011

[15] Yelong Shen,Ruoming Jin, Learning Personal+Social Latent Factor Model for Social Recommendation,2012

[16] J. Leskovec, A. Singh, and J. Kleinberg. Patterns of influence in a recommendation network. Lecture Notes in Computer Science, 2006

[17] David M. Pennock, Eric Horvitz, Steve Lawrence, and C. LeeGiles, Collaborative filtering by personality diagnosis: A hybric memory- and model-based approach, 2000

[18] L. Liu, J. Tang, J. Han, M. Jiang, and S. Yang. Mining topic-level influence in heterogeneous networks. In Proceedings of CIKM, 2010

[19] Ke Sun,Yunbo Cao,Xinying Song , Learning to Recommend Questions Based on User Ratings, 2009

[22] P. Massa and P. Avesani. Trust-aware recommender systems. In Proceedings of the2007 ACM Conference on Recommender Systems (RecSys), 2007