**基于时间加权三部图的分众分类标签推荐算法**

李贵王爽 李征宇 韩子扬 孙平 孙焕良

沈阳建筑大学信息控制工程学院

**摘 要**分众分类作为一种公众对网络信息分类组织的方法，会随着新用户的增加生成大量的基于分众分类的标签元数据，传统的基于三部图的分众分类标签推荐算法没有考虑时间对用户偏好的影响，从而影响了标签推荐的准确性，本文提出一种基于时间加权三部图的分众分类标签推荐算法，将分众分类转化为一个加权无向三部图，应用BM25加权方案并结合标签的时间加权因子计算三部图的边权，并生成基于三部图的分众分类邻接矩阵。最后通过Katz值计算来预测标签节点与用户物品节点之间所有可能的路径权重之和，并依据路径权重来决定用户-物品对的个性化标签推荐排名。实验结果表明该方法较以往的标签推荐算法在准确率和召回率等性能指标上有明显提升。

**关键词** 分众分类，三部图，时间加权，标签推荐

**Folksonomy tag recommendation algorithm based on a time-weighted tripartite graph**

Li Gui Wang Shuang Li Zheng-yu Han Zhi-yang Sun Ping Sun Huan-liang

*（Faculty of Information & Control Engineering , Shenyang Jianzhu University, Shenyang 110168, China）*

**Abstract** Folksonomy as a way of classifying and organizing Web information, will generate a huge amount of user-generated metadata. The traditional folksonomy tag recommendation algorithm based on tripartite graph without considering the influence of time on users preferences, in turn affected the accuracy of tag recommendation. This paper proposed a folksonomy tag recommendation algorithm based on a time weighted tripartite graph. We transform folksonomy into a weighted , undirected tripartite graph. Firstly combine the BM25 weighing scheme and tag time-weighted factor to calculate the link weight of tripartite graph, and generate the adjacency matrix for the folksonomy graph. finally applying the Katz measure, a path-ensemble based proximity measure, to predict both the path weight between user and tag, and the path weight between item and tag. Thereafter, We calculates a ranking score of tag based on such predicted path weights and thus generates a list of ranked tags suited to user for item. Experiment results show that the proposed approach outperforms previous algorithms in terms of both precision and recall.

**Key words** Folksonomy ,Tripartite graph, Time-weighted, Tag recommendation

# 1 引言

近年来Last.fm、Delicious、Movielens等大型社会标签推荐网站越来越流行，标签服务系统已经为网络用户提供了强有力的应用平台，吸引了企业界以及研究机构的注意力，许多社会标签应用系统都具有一个共同的特征，允许用户自由的分配所选关键字，也就是自由的给物品添加“标签”[1]。而且社会标签服务系统灵活性很高，在内容上不受任何严格的等级制度限制，因此，社会标签成为一种分类、分享以及组织社会传媒内容(视频，影像，音乐等)的流行方式，进而导致生成大量用户元数据和分众分类(folksonomies) [2]。

分众分类是一种新兴的网络信息资源分类方法，在大众用户自发的为网络信息资源添加一个或多个标签的过程中根据标签被使用的频次选用高频标签作为该类信息资源类别的标识。相关研究工作提出一种基于图的分众分类法实现个性化的标签推荐，将分众分类转变成一个三部图[1][3]，图中的三个节点分别代表用户集、物品集、标签集。由于分众分类信息多变且增长迅速，当前需要解决的问题就是如何准确预测三部图的边权，因此本文将分众分类看作是一个加权无向三部图，提出一种基于时间加权的三部图边权预测标签推荐算法。将标签推荐问题转变成边权预测问题，在图的基础上，利用Katz值[4]测量(计算路径权重的测量方法)计算连接两个邻近节点所有可能路径的权重之和来量化两个节点的邻近程度。本文算法的基本前提是为一个给定的用户-物品节点对(u,i)推荐合适的标签，预测标签节点t如何既与用户节点u邻近又与物品节点i邻近。通过计算用户-物品对(u,i)与邻近标签节点的邻近程度来预测可能出现的新三部图 (用户-物品对(u,i)与邻近的标签节点构建的新三角关系)。

# 2相关工作

国内外很多学者将推荐算法应用于社会标签推荐系统并进行了相关研究，一些早期的研究学者提出一种形式化的分众分类模型，简称FolkRank[1]，是在PageRank算法基础上的一种扩展算法，该算法将用户、资源和标签看作三部图中的三个节点集合，不同集合中的节点之间的权重为与他们共现的另一集合中的对象数，并应用Folksonomy架构来计算所有对象的PageRank向量，从而进行标签推荐和搜索排名。文献[5]提出一种以用户为中心的标签模型，简称UCTM，该算法应用三阶张量来描述用户、物品、标签之间的关联，并将分众分类中相应被标注的物品映射成一个个性化的标签词汇列表，用于标签搜索以及标签推荐。文献[6]提出一种个性化的针对物品推荐的搜索框架，该框架是通过将用户输入的标签查询与用户的兴趣偏好相匹配来实现网页排名。文献[7]提出一种基于线性加权的混合算法用于实现基于标签的资源推荐。文献[8]是从机器学习的角度提出一种以双向二分图的形式来表示标签数据的标签推荐算法来解决标签推荐问题。文献[9]提出一种基于图的表示方法来进行音乐搜索，该方法将原始相似度与标签同时应用于一个概率框架用于实现音乐排名。

由于以上算法都没有考虑时间对用户偏好的影响，因此本文提出一种基于时间加权三部图的分众分类标签推荐算法，将分众分类转化为一个加权无向三部图，通过分析三部图中固有的三元关系，预测潜在的用户-物品对(u,i)与邻近的标签节点t之间的路径权重，利用Katz值测量方法分别计算用户节点u与标签节点t以及物品节点i与标签节点t之间所有可能的路径权重之和。最后依据路径权重来决定用户-物品对的个性化标签推荐排名。

# 3定义及注释

本节介绍本文相关注释及定义，在社会标签推荐系统中，用户用标签注释物品，建立用户、物品、标签三者之间的三元关系。并生成用户集U ={u1, u2,...,u|U|}，标签集T={t1, t2,...,t|T|}和物品集I={i1, i2,...,i|I|}。分众分类被定义为一个四元组F:={ U, T, I, Y}，其中Y ⊆ U × T × I代表U、T、I之间的三元关系，称作标签分配[1]。根据标签分配，能够得到由用户、物品、标签三者之间两两结合的三个矩阵，分别为|U|×|T|阶用户-标签矩阵MUT，|U|×|I|阶用户物品矩阵MUI，以及|T|×|I|阶标签-物品矩阵MTI，矩阵中的每个元素代表Y中相对应的行与列同时发生的次数[5]。通过应用这三个矩阵，分众分类可以转换成一个无向三部图G=(U∪T∪I,E)，三部图中的节点可以划分成三个独立的节点集合，分别为用户集U、标签集T、以及物品集I，其中E表示三部图中所有由U、T、I中节点组成的三部图边权(三部图的边)的集合。并且每个集合中的每个节点都至少与其他两集合中任一集合中的一个节点毗邻[1]。E中包含的每条边的权重都与矩阵MUT、MUI、MTI中相应的行与列所对应的非零元素相对应。另外，每条边权的权重都等于相应矩阵中元素的值。下面列举一个应用频率加权方案(权重为相邻的两个节点之间发生关联的次数)将分众分类转换成无向三部图的实例，如图1所示。



图1：将分众分类转换成加权无向三部图

事实表明分众分类可以转换成无向三部图，因此本文研究的问题可以看做是边权预测问题，根据三部图的边权来进行标签推荐。例如一个三部图，如图2所示，如果用户u有意给物品i打标签，那么用户u 与物品i之间就建立了新的关联权重也就是三部图的边权，为了给已建立的用户-物品对(u,i)推荐合适的标签t，本文的方法致力于预测用户-标签节点对(u,t)以及物品-标签节点对(i,t)之间所有可能的路径权重之和。然后基于已预测的路径权重计算标签t的排名得分，最后建立用户u 对物品i的top-N标签排名推荐列表。

在社会标签推荐系统中，本文的标签推荐定义如下：对应于分众分类的无向三部图G=(U T I, E)；一个正整数N；一个用户-物品节点对(u,i)，其中u ∈ U，i ∈ I，并确保由标签节点构成一个完全有序的标签集T(u,i)，使其满足|T(u, i)| ≤ N，且T(u,i) ⊆ T，其中标签集T(u,i)是指以三部图的形式伴随用户-物品对(u,i)出现的所有可能的标签节点集合。



图2：三部图边权预测标签推荐

# 4 基于时间加权三部图的分众分类边权预测

## 4.1 基于时间加权三部图的分众分类邻接矩阵

首先定义一个对应于分众分类的无向三部图G的邻接矩阵**A**，**A**中记录了三部图中所有u、i、t节点之间的边权。如图3所示，邻接矩阵中的元素表示该元素在矩阵中所对应行与列的两节点之间的边权。

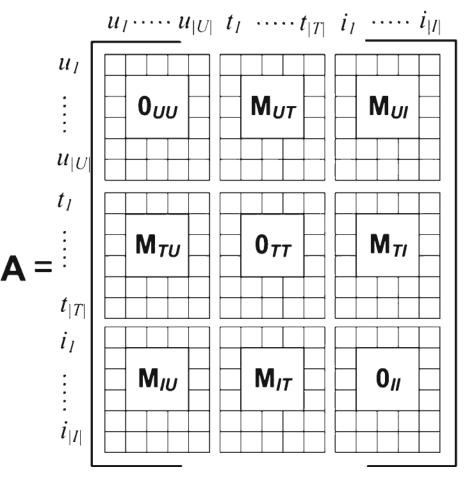


图3：分众分类的邻接矩阵A

设定三部图中的边权可以有不同的加权方案，加权方案不同，邻接矩阵也不同，本文列举三种加权方案并将其与本文提出的加权方案进行实验比较。

(1)Binary:权重设定为0或1，例如，如果用户u给物品i打标签t，那么相对应的节点对(u,t)、(u,i)、(t,i)之间的边权为1，否则为0；

(2)Frequency:其权重设定与应用于信息检索中的词频非常相似，节点对(u,t),(u,i),(t,i)之间的边权分别表示用户u打过标签t的物品数；用户u给物品i打过的标签数；以及物品i被打标签t的用户数；

(3)BM25：该方案具有以下两个特点，1)逆文档频率(文档频率越高，该标签越常见，文档频率越低，该标签越罕见)；2)该标签在特定文档(与文档长度相关联)中出现的频率[10]；(用户和物品同理)；

本文中文档指的是用户和物品，在BM25加权方案实例中，物品的长度表示只在该物品上被标注的标签数量，用户的长度表示该用户打过的标签数量。节点对(t,i)的BM25边权计算定义如下[6] [17]：

 (1)

其中：  (2)

 (3)

式(1)中描述的是逆物品频率，描述的是标签t与物品i的相关性权重，其中|I|是指物品总数量，N(it)是指被标注标签t的物品数量，tf(t,i)是指物品i被打标签t的次数，|i|是指物品i的长度，avg(I)是指物品的平均长度。参数k1和b为常数，一般情况下，取实验经验值k1=2，b=0.75。其中参数b的作用是调整文档(用户或物品)长度对相关性权重影响的大小，其中b越大，文档长度的影响对相关性权重就越大。同理, 节点对(u,t)的BM25边权计算定义如下：

 (4)

其中|U|是指用户总数量，N(ut)是指打标签t的用户数量，tf(u,t)是指用户u打标签t的次数，|u|是指用户u的长度。avg(U)是指用户u的平均长度。同理，k1=2，b=0.75。

以上三种加权方案都没有考虑标签的时间衰减问题，也就是用户受近期行为的影响比较大，简言之用户在近期给物品所打标签的(u,t)边权大于早期给物品所打标签的(u,t)边权，而且对于评价周期不同的用户，早期添加的标签对用户的影响力也不同，因此定义用户的网络活跃期如下所示：

 (5)

其中表示用户ui最后选择物品并添加标签的时刻，表示用户ui第一次选择物品并添加标签的时刻。然而，针对于被打标签周期较短的物品，早期标签对物品的影响力大于被打标签周期较长的物品。同理定义物品的活跃期如下所示：

 (6)

其中表示物品ij最后被添加标签的时刻， 表示物品ij第一次被添加标签的时刻。

由于用户和物品不同的网络活跃期对标签推荐的预测也不同，因此本文将用户和物品的网络活跃期与标签的时间衰减特性结合在一起，定义标签的时间加权因子 [18]如式(7)(8)所示：

 (7)

 (8)

其中t为当前时刻,为用户u添加标签t的时间，为物品i被打标签t的时间。式(7)(8)可理解为每个标签当前时刻在用户和物品上获得的一份推荐权重。

举一实例来更好地说明该算法，假设当前时间t为120，目标用户为U1，该用户使用过标签T2、T3，使用次数分别为3次、2次，使用T2的时间分别在t=30、40、50，使用T3的时间分别在t=90、100，由式(5)得到用户的网络活跃期为lifetime(U1)=70,由式(7)可得到标签T2的初始资源为：

 (9)

其中标签的初始资源可理解为目标用户U1在打标签T2的所有时间点上的推荐权重之和。由于BM25加权方案没有考虑标签的时间加权问题，因此，本文将BM25加权方案加以改进，将原始的BM25加权方案与标签的时间加权因子相结合,将改进后的加权方案定义为BM25-TA(BM-time attenuation)，并将BM25-TA加权方案节点对(t,i)的权重计算定义如下：

 (10)

同理，将节点对 (u,t)的BM25-TA权重的计算定义如下：

 (11)

本文应用改进后的BM25-TA加权方案来构建基于三部图的分众分类邻接矩阵A，A中的元素代表矩阵中该元素所对应行与列节点在三部图中的边权。如图3所示，本文将邻接矩阵分为9个模块，其中包括|U|×|U|、|T|×|T|、|I|×|I|零方阵**0**UU、**0**TT、**0**II，|U|×|T|用户-标签矩阵MUT， |U|×|I|用户-物品矩阵MUI，|T|×|I|标签-物品矩阵MTI，以及矩阵MUT、MUI、MTI的转置矩阵MTU、MIU、MIT，其中矩阵MUT、MUI、MTI及其转置矩阵均由分众分类法分别按各个加权方案将用户、物品、标签聚合所得。

## 4.2 路径权重Katz值计算

构建邻接矩阵A后，计算矩阵A中所有节点对(vk,vk+1)( k=1,,n)的Katz值。Katz值计算是一种计算路径权重的方法[4]，并已成功应用于不同的应用领域中，例如边权预测[11]，图形理论[14]，以及推荐系统[13]等等。本文应用该方法计算两个邻近节点之间所有可能路径的权重之和。在介绍如何计算三部图中两邻近节点之间的Katz值之前，首先来定义三部图G中的一条路径以及它的路径权重，三部图中的一条路径代表从起始节点开始穿越三部图中的边到达目标节点的一条路径，路径的长度n满足1,路径中包含的节点可以用v1,v2,vn,vn+1表示,因此连接两节点vk、vk+1的边可以由节点对(vk,vk+1)表示,其中k=1,,n。并将两节点之间的边权定义为W(vk,vk+1)。路径的长度是指路径中包含的边的数量。因此，如果一条路径的长度为1，则表示两节点之间只有一条边。本文将三部图G从节点x到节点y的所有长度为n的可能路径的集合定义为Pn(x,y)，一条路径p∈Pn(x,y)的路径权重被定义为：

 (12)

其中x代表起始节点v1,y代表终止节点vn+1，从节点x到节点y的Katz值计算如下：

 (13)

其中是一个递减参数(通常是一个很小的值(0.05或0.005)，因为路径长度的路径对Katz值总和的贡献较小)。Katz值计算的是连接两个节点间所有不同长度的路径权重之和。三部图中所有节点对之间的Katz值都可以用下面的矩阵形式表达出来[13]：

 (14)

其中I为单位矩阵，A为三部图的邻接矩阵，，其中表示邻接矩阵A的任意特征值绝对值的最大值，这一条件决定了的最大值。Katz矩阵中的每一个元素K代表矩阵中从相应的行节点到相应的列节点中所有不同长度的路径权重之和。最后本文通过求得的矩阵K来预测个性化的标签推荐。

## 4.3 个性化标签推荐

本节根据给定的用户-物品对，并以Katz值矩阵为基础，提出一种新的标签排名方法。为了鉴别前N个标签是否合适，需要借助Katz计算来评估一个确切的标签与邻近的用户和物品的邻近程度，通过将Katz值求和来计算推荐标签的排名得分。假设给定一个用户-物品对(u,i),对于推荐标签t的排名得分计算如下：

 (15)

其中Ku,t和Ki,t分别为从用户u到标签t和从物品i到标签t的Katz值，并用矩阵的形式简单的描述相关用户u和物品i的所有标签的排名分数，如下所示：

S=vTK (16)

其中vT是一个向量，向量中与用户u和物品i相对应的元素为1，其余为0。并将该向量与Katz值矩阵相结合得到邻接排名向量S，从而得到用户u 对物品i的前N个得分最高的标签排名推荐列表(top-N推荐列表)。图4描述了对于同一物品i1，不同用户u1、u2，标签t2的排名得分，以及同一用户u1，不同物品i1、i2，标签t2的排名得分。



图4 通过Katz值计算标签排名得分

# 5实验及结果

## 5.1数据集

如表1所示，实验中所用数据来源于Movielens和Last.fm。Movielens推荐系统的主要功能是应用协同过滤技术以及用户对电影的偏好，向用户推荐电影的网络平台，通过对一部分电影的评分，来预测对其他电影的评分，在此，本文将评分以标签的形式来进行电影推荐。为了提取数据集中比较稠密的数据，本文将数据进行以下筛选：

(1)用户至少给5个物品打过标签，且至少给物品打过5个标签；

(2)标签至少给5个物品标注过，且至少被5个用户标注过；

(3)物品至少被打过5个标签，且至少被5个用户标注过标签；

Last.fm是一个帮助用户发现音乐，分享音乐，标注音乐的社会化音乐服务网站。各数据集的属性如表1所示：

表1 各数据集的属性

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 数据集 | 用户 | 物品 | 标签 | 标签分配 |
| Movielens | 2652 | 4127 | 2496 | 161412 |
| Last.fm | 1923 | 12689 | 9810 | 187104 |

然后将两个数据集中得到的标签分配(如表1)分解为三个二维矩阵，分别为用户-物品矩阵MUI、标签-物品矩阵MTI、以及用户-标签矩阵MUT，各矩阵的属性如表2所示：

表2 由标签分配分解得到的二维矩阵的属性

|  |  |  |  |
| --- | --- | --- | --- |
| 数据集 | MUI | MTI | MUT |
| Movielens | 65894 | 72853 | 63574 |
| Last.fm | 71105 | 110056 | 38201 |

相对于Last.fm评估数据集来说，Movielens评分数据集更干净(噪声少)更稠密。说明Movielens数据集中的噪声标签比Last.fm数据集中的噪声标签少。另外，针对用户与标签、用户与物品、以及标签与物品之间的数据关联性而言，Movielens数据集优于Last.fm数据集。

## 5.2 实验设计及指标

为了评估该标签推荐算法的性能，本文应用交叉验证法来进行实验验证。该方法被广泛应用于评估标签推荐算法[14] [15] [16]，对于每一个用户，本文都随机的从训练集中删除一个物品以及被标注在该物品上的标签。例如，如果用户u1用两个标签t1和t2来注释物品i1，本文将从与用户u1相关的所有标签分配任务中删除物品i1以及标签t1、t2。然后在测试集里验证该算法是否能向用户u1针对于物品i1在前N个标签排名列表中推荐标签t1和t2。并将这一过程通过使用不同的训练集和测试集重复5次。以下的实验结果数据均为这五次实验的平均值。

本文首先应用评估指标MAP(准确率的平均值)：

 (17) 其中表示用户u给测试集中的物品所打标签的集合，表示前N个标签的准确值，Bn是一个二元随机变量，在推荐列表中，排名第n的标签出现在标签集合中则Bn取1，否则取0。

其次，本文还应用了评估指标MRR[17] (倒序排名均值)：

 (18)

其中rank(t)是指针对于用户u的推荐标签t在标签推荐列表中的排名。标签t在排名列表中的排名越高，MRR值越大。在实验过程中本文还应用准确率和召回率作为评估指标将本文算法与其他算法进行性能比较。

## 5.3 参数的选择

但是如何在计算Katz值的过程中恰当的选择值的大小以确保长路径的低效性仍是一个需要解决的问题，为了解决这一问题本文通过调整值来提高推荐系统的性能。

因此，本节将研究值的大小将如何决定推荐系统性能的高低。由于Katz值的大小取决于三部图中边权的大小，而不同的值将直接影响三部图中边权的大小，因此值的大小将间接影响推荐系统的性能。在Katz值计算过程中，随着Katz值的收敛，值应该满足，其中表示邻接矩阵A的任意特征值绝对值的最大值，根据加权方案的不同，值的约束条件也不同(如表3表示)。以BM25为例，为了确保Katz值的收敛，在数据集Movielens中值必须低于0.007，在数据集Last.fm中必须低于0.0059；因此，本文根据这些约束值来降低值。并将Binary、Frequency以及BM25加权方案应用于Katz值计算，进行系统性能的评估，实验结果表明，较小的值在性能上要略优于较大的值。值越小，长路径对于Katz值计算的影响越小。最短路径(如直连路径)对Katz值计算的影响最大。

本文将BM25-TA加权方案与Binary、 Frequency以及BM25加权方案进行实验比较，每一个加权方案均选择最佳的值，如表3所示, 表3中第一列为用于评估加权方案的两个真实数据集Movielens和Last.fm。第二列为用于计算Katz值的加权方案，第三列为邻接矩阵A的任意特征值绝对值的最大值，第四列为实验中验证的各加权方案下的最佳值，第五列为值在约束范围内能取到的最大值。然后应用最佳的值来分别计算评估指标MAP和MRR的值,如图5、6所示。

表3 不同数据集中针对不同加权方案的最佳值

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 数据集 | 加权方案 |  |  |  |
| Movielens | Frequency | 551.30 | 0.0005 | 0.0018 |
| Binary | 115.27 | 0.007 | 0.0088 |
| BM25 | 142.35 | 0.006 | 0.007 |
| BM25-TA | 153.91 | 0.005 | 0.006 |
| Last.fm | Frequency | 1158.37 | 0.0001 | 0.0009 |
| Binary | 135.23 | 0.006 | 0.0074 |
| BM25 | 166.98 | 0.005 | 0.0059 |
| BM25-TA | 177.35 | 0.004 | 0.0049 |

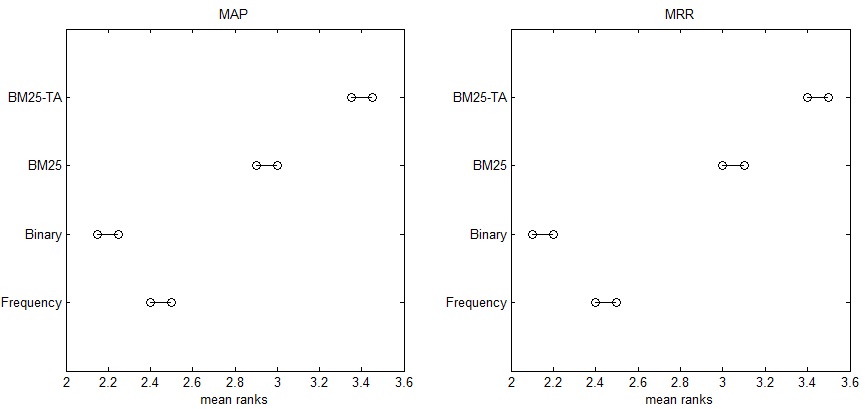


图5 数据集Movielens中各加权方案的实验比较

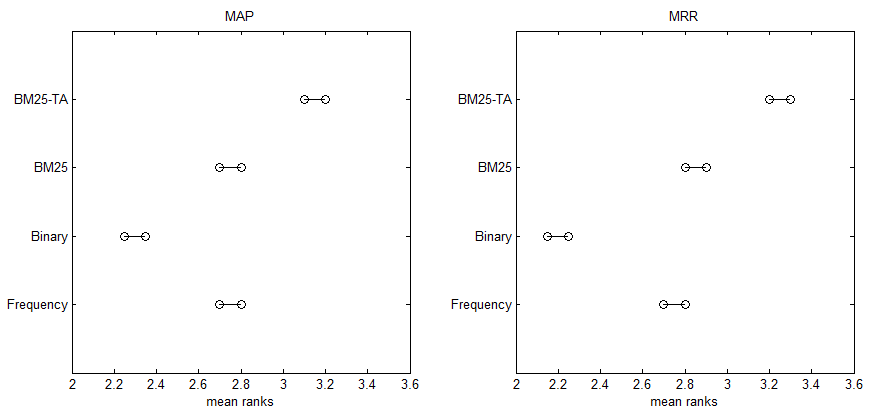
****

图6 数据集Last.fm中不同加权方案之间的实验比较

从图5、图6的实验结果可以看出在数据集Movielens、 Last.fm上应用BM25-TA加权方案均优于其他加权方案。

5.4 实验结果

本文应用数据集Movielens和Last.fm作为实验平台,将MAP、MRR以及准确率、召回率作为评估指标，将本文提出的方法(定义为KatzBM25-TA)与其他三种算法(FolkRank[1]、UCTM[5]、KatzBM25[17])进行实验比较，实验结果如图7、8、9、10所示。

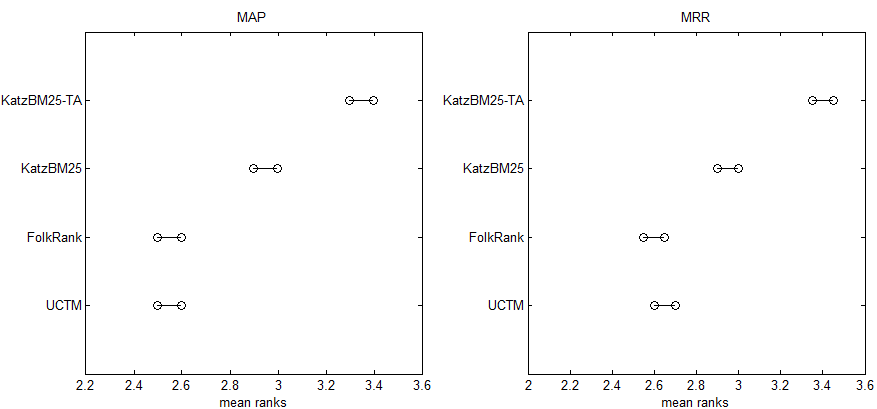


图7 数据集Movielens中不同算法之间的实验比较

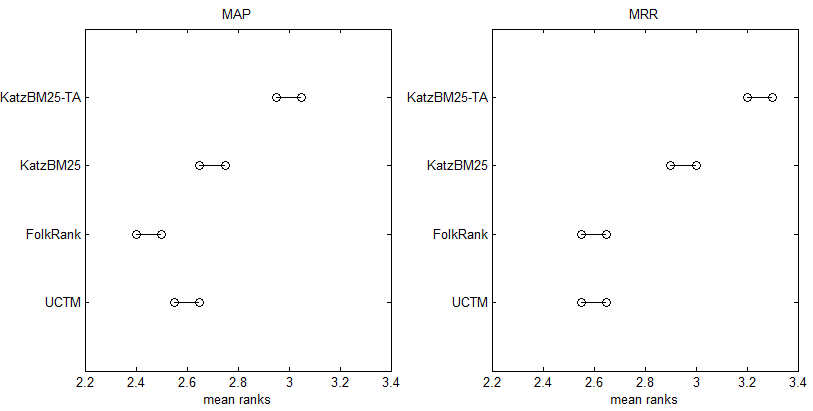


图8 数据集Last.fm中不同算法之间的实验比较

从图7、图8的实验结果可以看出在数据集Movielens、Last.fm中本文提出的KatzBM25-TA算法较其他三种算法在MAP、MRR评估指标上有明显提升。

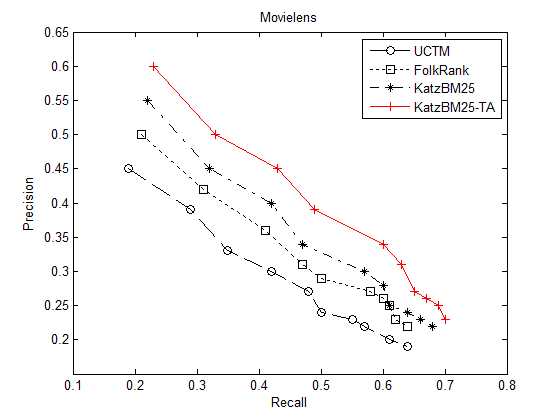


图9 数据集Movielens中各算法准确率与召回率实验比较

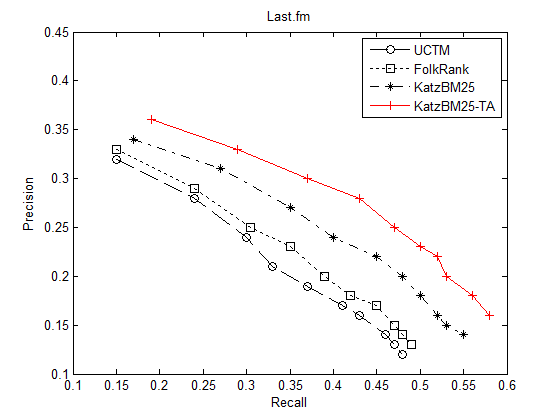


图10 数据集Last.fm中各算法准确率与召回率实验比较

从图9、图10的实验结果可以看出在数据集Movielens、Last.fm中本文提出的KatzBM25-TA算法较其他三种算法在准确率、召回率评估指标上有明显提升。

# 6 结论

本文将分众分类看作一个加权无向三部图，在三部图的基础上提出一种基于时间加权的三部图边权预测标签推荐算法，将标签进行时间加权，解决了用户偏好的改变对资源的即时需求问题，并利用Katz值计算求解预测的标签节点与给定的用户、物品节点组成的新三部图的路径权重之和。路径权重之和越小该预测标签所处的标签排名越高。该算法成功的实现了向三部图中存在的潜在用户-物品对推荐合适的标签，使其满足用户对物品的兴趣偏好，实现个性化的标签推荐。实验结果表明相对于UCTM、FolkRank、KatzBM25推荐算法，本文提出的KatzBM25-TA算法在MAP、MRR以及准确率、召回率评估指标上都有明显提升。

参考文献

[1] Hotho A, Jäschke R, Schmitz C, Stumme G. Information retrieval in folksonomies: Search and ranking. In Proceedings of the 3rd European semantic web conference. 2006:411–426

[2] Bischoff K, Firan C.S, Nejdl W, Paiu R. Can all tags be used for search? In Proceedings of the 17th ACM conference on information and knowledge management. New York: ACM,2008:193–202

[3] Ramezani M. Improving graph-based approaches for personalized tag recommendation. Journal of Emerging Technologies in Web Intelligence. 2011, 3(2): 168–176

[4] Kenichi Bannai, Guido Kings. P-adic elliptic polylogarithm, p-adic eisenstein series and Katz measure.arXiv,2007:1–4

[5] Wetzker R, Zimmermann C, Bauckhage C, Albayrak S. I tag, you tag: Translating tags for advanced user models. In Proceedings of the 3rd ACM international conference on web search and data mining. New York: ACM,2010: 71–80

[6] Xu S, Bao S, Fei B, Su Z, Yu Y. Exploring folksonomy for personalized search. In Proceedings of 31st annual international ACM SIGIR conference. New York: ACM, 2008:155–162

[7] Gemmell J, Schimoler T, Ramezani M et al. Tag-based resource recommendation in social annotation applications. In Proceedings of the 6th European semantic web conference on the semantic web: research and applications. Berlin: Springer-Verlag, 2011:195–206

[8] Song Y, Zhang L, Giles C. Automatic tag recommendation algorithms for social recommender systems. ACM Transaction on the Web, 2011,5(1):1–31

[9] Miotto R, & Orio N. A probabilistic model to combine tags and acoustic similarity for music retrieval. ACM Transactions on Information Systems, 2012,30(2): 1–29

[10] Sparck Jones K, Walker S, Robertson S.E. A probabilistic model of information retrieval: development and comparative experiments. Information Processing and Management, 2000,36(6):809–840

[11] Liben-Nowell, D., & Kleinberg, J. (2007). The link-prediction problem for social networks. Journal of the American Society for Information Science and Technology, 58(7): 1019–1031

[12]Foster K, Muth S, Potterat J, Rothenberg R. A faster Katz status score algorithm. Computational & Mathematical Organization Theory, 2001,7(4):275–285

[13]Huang Z ,Li X, Chen H. Link prediction approach to collaborative filtering. In Proceedings of the 5th ACM/IEEE-CS joint conference on digital libraries, New York: ACM,2005:141–142

[14] Jäschke R, Marinho L, Hotho A, Schmidt-Thieme L, Stumme G. Tag recommendations in social bookmarking systems. AI Communications, 2008,21(4):231–247

[15] Zanardi V, & Capra L. Social ranking: Uncovering relevant content using tag-based recommender systems. In Proceedings of 2nd ACM conference on recommender systems, New York: ACM, 2008:51–58

[16] Rendle S, & Schmidt-Thieme L. Pairwise interaction tensor factorization for personalized tag recommendation. In Proceedings of the 3rd international conference on web search and web data mining, New York: ACM, 2010: 81–90

[17] Vallet D, Cantador I, Joemon J. Personalizing web search with folksonomy-based user and document profiles. In Proceedings of the 32nd European conference on advances in information , Berlin: Springer-Verlag, 2010:420–431

[18] Gu Yiran, Chen Min. One tag time-weighted recommend approach on tripartite graphs networks[J] .Computer Science, 2012,39(8):1-4(in Chinese)(顾亦然，陈敏等.一种三部图网络中标签时间加权的推荐方法[J].计算机科学,2012,39(8):1-4)