

# 基于统一超图的音乐推荐:结合社交媒体信息和音乐内容

卜家军 1 号, 谭树龙 1 号, 陈春 1 号, 王灿 1 号, 郝武 1 号, 张立军 1 号, 河 2 号  
北京, 老挝, 1984, 中国, 加拿大, zljzju}@zju.edu.cn, 吴昊, xiaofeihe@cad.zju.edu.cn

1 浙江省服务机器人重点实验室 2 浙江大学计算机科学学院 CAD&CG 国家重点实验室 浙江大学计算机科学学院  
中国杭州, 310027

## 摘要

近年来, 基于声学的音乐推荐系统受到了越来越多的关注。由于低级声学特征和高级音乐概念之间的语义鸿沟, 许多研究人员已经探索了音乐推荐系统中的协同过滤技术。传统的协同过滤音乐推荐方法只关注用户评分信息。然而, 在 Last.fm 和 Pandora 等音乐社交社区中, 存在各种各样的社交媒体信息, 包括不同类型的对象以及这些对象之间的关系。这些信息对于音乐推荐很有价值。然而, 利用这种丰富的社交媒体信息有两个挑战: (a) 在音乐社交社区中有许多不同类型的对象和关系, 这使得很难开发考虑所有对象和关系的统一框架。 (b) 在这些社区中, 一些关系比成对关系复杂得多, 因此不能简单地用图来建模。本文提出了一种新的音乐推荐算法, 该算法综合利用了多种社交媒体信息和基于音乐声学的内容。我们用超图代替图来建模各种对象和关系, 把音乐推荐看成是这个超图上的一个排序问题。普通图的一条边只连接两个对象, 而超边代表一组对象。通过这种方式, 超图可以自然地用于建模高阶关系。在音乐社交社区 Last.fm 数据集上的实验证明了该算法的有效性。

## 类别和主题描述符

H. 3. 3 【信息存储与检索】: 信息搜索与检索——信息过滤; h . 5. 5 [信息接口和表示]: 声音和音乐计算——方法和技术

允许免费制作本作品全部或部分的数字或硬拷贝供个人或课堂使用, 前提是不得以盈利或商业利益为目的制作或分发拷贝, 并且拷贝第一页带有本声明和完整引用。以其他方式复制、重新发布、在服务器上发布或重新发布到列表, 需要事先获得特定许可和/或付费。

2010 年 10 月 25 日至 29 日, 意大利费伦泽。

版权所有 2010 美国计算机学会 978-1-60558-933-6/10/10... \$10.00.

## 泛称

算法, 实验。

## 关键词

推荐系统、音乐推荐、超图、社交媒体信息

## 1. 介绍

随着万维网成为各种数字音乐的来源和分发渠道, 人们可以访问大量的音乐曲目。由于用户查找和选择他/她想要的音乐通常是困难和费时的, 音乐推荐系统成为一个不可缺少的工具[26]。音乐推荐在许多真实世界的应用中是有价值的, 例如社交音乐社区、在线音乐商店和一些音乐设备(例如, PC 和 MP3 播放器), 其中音乐推荐可以用于生成音乐播放列表。

对于音乐推荐任务, 最常见的方法是直接分析音频信号。这些方法被称为基于声学的音乐推荐[22, 9, 8, 30]。由于低级声学特征和高级音乐概念之间的语义差距[10], 基于声学的音乐推荐的结果并不令人满意。有必要在推荐系统中考虑更多的信息[11]。一些研究人员试图通过应用协同过滤方法来利用用户评级信息[41, 20, 38, 40]。还存在利用与音乐曲目相关联的元数据(例如, 流派)中的信息的工作[4, 28, 27]。然而, 所有这些方法只利用了有限种类的信息, 没有考虑丰富的社交媒体信息。

在典型的音乐社交社区中, 如 Last.fm 和 Pandora2, 有丰富的社交媒体信息, 包括各种类型的对象和这些对象之间的关系。图 1 示出了 Last.fm 的示例。在 Last.fm 中, 每个用户可以与其他用户交朋友, 加入群组, 收听音乐曲目, 并使用标签来标记音乐曲目、专辑和艺术家等资源。资源之间也有一些关系, 比如曲目和专辑之间的包含关系。另外, 可以基于音频内容来计算音乐轨道之间的相似性关系。

上面提到的各种社交媒体信息是

<sup>1</sup><http://www.last.fm2http://www.pandora.com>

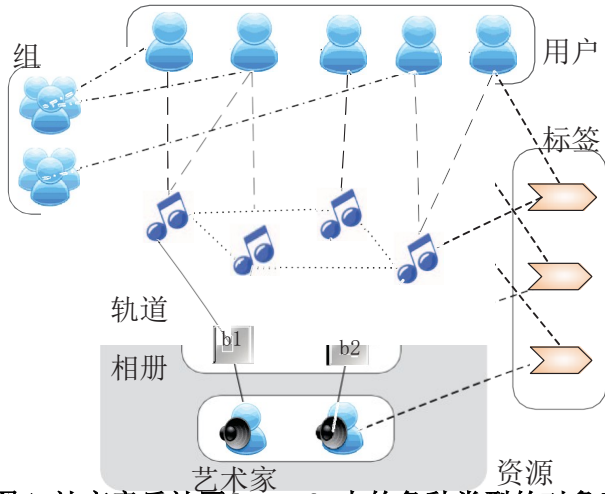


图 1: 社交音乐社区 Last.fm 中的各种类型的对象和关系, 这些关系包括友谊关系、成员关系、收听关系、标记关系、资源(例如, 曲目和专辑)之间的包含关系以及音乐曲目之间的相似性关系。

对音乐推荐很有用。推荐系统中的一个关键步骤是建立用户的偏好简档

[42] 这可以从他们对资源的多类型动作(例如评级和标记)中推断出来。通过向每个用户推荐相似的音乐, 可以在音乐推荐中使用音乐相似性关系。此外, 用户的兴趣可能受到他/她的朋友的影响[19, 24]。

然而, 利用所有这些信息有两个主要挑战。首先, 社交媒体社区中有许多不同类型的对象和关系, 这使得很难开发一个同时考虑所有对象和关系的统一框架。第二, 在社交媒体社区中, 有些关系不仅仅是成对的。例如, 两个以上的用户加入同一个群组, 或者一个用户通过标签将音乐曲目加入书签。我们称这种关系为高阶关系。处理成对关系的传统方法[45, 2]不能正确地模拟这些高阶关系。

传统的推荐算法, 如协同过滤(CF) [29, 16, 21], 只考虑用户项目评分矩阵, 而没有利用其他类型的社交媒体信息。最近, 人们对利用社交媒体信息来提高推荐性能产生了相当大的兴趣[39, 36, 19, 24, 33, 43]。例如, 一些先前的工作使用普通的图来建模推荐问题的标签数据[19, 43]。图 2(a) 示出了使用普通图来模拟标记关系的简单示例。有三种标记关系:  $u1$  分别用标记  $t1$  和  $t2$  标记资源  $r1$  和  $r2$ ,  $u2$  用标记  $t2$  标记资源  $r1$ 。图 2(b) 示出了我们的用于建模标记关系的统一超图方法。在我们的统一超图模型中, 三类对象之间的高阶关系可以自然地表示为三元组:  $(u1, t1, r1)$ ,  $(u1, t2, r2)$ ,  $(u2, t2, r1)$ 。显然, 普通的图模型不能捕捉标记关系

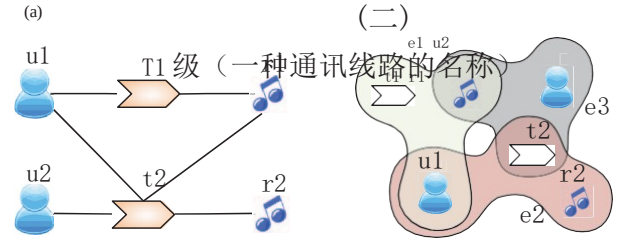


图 2: 用两种模型表示的标记关系: (a) 普通图模型, 和 (b) 我们的统一超图模型。这个超图包含六个顶点和三个超边, 即  $(u1, t1, r1)$ 、 $(u1, t2, r2)$  和  $(u2, t2, r1)$ 。

没错。例如, 从图 2(a) 中, 不清楚是否  $u2$  书签  $r1$  和/或  $r2$ 。

在本文中, 我们使用统一超图来建模音乐社交社区中的多类型对象和关系。基于声音信号的音轨之间的相似性被视为一种关系。这样, 我们在一个统一的框架中结合了基于声学的和协同过滤的推荐。超图是普通图的推广, 其中的边称为超边, 是顶点集的任意非空子集[3]。超图的每个顶点对应于任何类型的对象。超边用于模拟高阶关系, 如图 2(b) 所示。通过使用统一超图模型, 我们可以准确地捕捉各种类型对象之间的高阶关系, 而不会丢失任何信息。我们进一步将音乐推荐视为超图上的排序问题, 以找到每个用户想要的音乐曲目。

以下几点突出了本文的贡献:

1. 多源媒体融合。我们在音乐推荐中整合了多源媒体信息, 包括多种社交媒体信息和音乐声学信号, 以提高推荐性能。
2. 我们提出用超图代替传统的图来建模社会媒体信息中的高阶关系。这样, 在表示各种类型的关系时没有信息损失。
3. 我们实证研究了不同类型的社交媒体信息对推荐绩效的贡献。我们的结果对实际的音乐推荐系统是有帮助的。

本文的其余部分组织如下。第 2 节回顾了相关工作。在第三节中, 我们介绍了问题的形式化定义, 并描述了如何在超图上进行排序。在第四节中, 我们讨论如何在音乐推荐中应用超图排序。第五节介绍了大量的实验结果。在第 6 节中, 我们总结了我们的论文, 并对未来的工作提出了建议。

## 2. 相关著作

在本文中, 我们结合基于声学 and 协同过滤的音乐推荐方法来开发丰富的

使用基于超图的学习技术的社交媒体信息和基于声音的内容信息。我们的工作涉及混合音乐推荐、使用社交媒体信息的推荐和基于图形/超图的学习。在这一节中，我们将对这些作品做一个简要的回顾。

## 2.1 混合音乐推荐

有几种混合方法将基于声学的和协同过滤的音乐推荐结合起来，以提高预测的整体准确性[41, 20, 38, 14, 40]。Yoshii等人[41, 40]通过使用概率模型整合了评级和音乐内容信息。不可观测的用户偏好直接通过引入潜在变量来表示。Li等人[20]提出了一种基于项目的概率模型，该模型利用音频特征来捕捉项目(即音乐)之间的准确相似性。Tiemann等人[38]研究混合音乐推荐的集成学习方法。他们应用集成学习方法来组合基于项目的协同过滤和基于声学的推荐的输出。Donaldson[14]通过统一频谱图和声学特征向量，为混合音乐推荐系统开发了播放列表和声学信号中的音乐同现信息。所有上述工作使用传统的协同过滤方法，并且仅利用有限种类的信息，而没有考虑更复杂的社交媒体信息。

## 2.2 使用社交媒体信息的推荐

事实证明，社交媒体信息，如标签关系和友谊关系，对推荐很有价值。Tso-Sutter等人[39]通过将标签视为用户或资源，将标签关系中的三类对象(用户、资源和标签)简化为两类，然后分别应用传统的基于项目或基于用户的协作过滤算法[1]。Diederich等人[13]介绍了用户的TF-IDF标签简档，并在基于使用的CF算法中使用这些简档向量来测量用户-用户相似性。Zhang等人[43]提出了一种通过在用户标签项目三向图上整合扩散的推荐算法。马等人[24]提出了一个概率因子分析框架，该框架自然地将用户的喜好和他们信任的朋友的偏好融合在一起。为了利用友谊和标签关系，Konstas等人[19]创建了一个协作推荐系统，该系统构建了一个关于用户、标签和资源的社交图。Sen等人[33]通过首先推断用户标签偏好，然后基于标签偏好计算资源项目偏好来解决资源推荐。他们提出了一些启发式方法，以在标签偏好生成步骤中利用各种社交媒体信息，例如点进和搜索信息。膝盖等人[18]利用基于网络的音乐艺术家相似性信息来减少必要的基于声学的音乐相似性计算的数量，然后在音乐播放列表生成的任务中使用音乐相似性。尽管上述方法在资源推荐应用中取得了巨大的成功，但它们未能充分利用社交媒体社区中的高阶关系。在这项工作中，我们提出使用超图，而不是普通的图，来精确地捕捉高阶关系，从而提高推荐质量数据性能。

## 2.3 基于图的排序和超图

我们的工作也与基于图的排序和hypergraph学习相关[45, 2, 3, 44, 12, 35, 7]。

周等人提出了一种流形排序算法，该算法根据数据中的内在几何结构对数据对象进行排序[45]。他们首先构造一个加权图并设置查询点，然后让所有数据点通过加权图将它们的排名分数传播给它们附近的邻居。重复扩展过程，直到达到全局稳定状态。Agarwal[2]提出将数据对象建模为加权图，并将这种图结构作为正则项纳入排名函数。以这种方式，所获得的排序函数在图形上平滑地变化。为了生成个性化标签推荐，Guan等人提出了一种基于图的多类型相关对象排序算法[15]。

最近，人们对用超图学习产生了浓厚的兴趣[3, 44, 12, 35, 7]。Bul`o等人引入了hypergraph聚类算法，使用高阶(而不是成对)相似性从一组对象中提取最大相干组[7]。周等人开发了一个适用于超图数据的分类、聚类和嵌入的通用框架[44]。这些研究只关注超图的分类、聚类和嵌入。然而，通过将多种类型的社交媒体对象及其关系建模为统一超图，我们将音乐推荐视为统一超图上的排序问题。

## 3. 统一超图上的排序

在本节中，我们将讨论如何在统一hypergraph模型中建模各种类型的对象及其关系，以及如何在统一hypergraph上执行排名。我们从问题的描述和符号开始。

### 3.1 符号和问题定义

设 $G(V, E, w)$ 表示一个超图，其中 $V$ 是顶点的集合， $E$ 是超边的集合， $w$ 是定义为 $w: E \rightarrow \mathbb{R}$ 的权函数。每个超边 $e \in E$ 是 $V$ 的子集。超边 $E$ 的度由 $\delta(e) = |e|$ 定义，即 $E$ 的基数。如果每个超边的度为2，则超图退化为普通图。顶点 $v$ 的度 $d(v) = \sum_{e \in E, v \in e} w(e)$ 。我们假设在顶点 $v_l$ 和 $v_k$ 之间有一条超路径

如果存在不同顶点和超边 $v_1, e_1, v_2, e_2, \dots, e_k, v_k$ ，使得 $v_i, v_{i+1} \in e_i$ ， $i=1, \dots, k$ 。我们定义超图是连通的，如果每对顶点都有一个超路径。

一个顶点超边关联矩阵 $H \in \mathbb{R}^{|V| \times |E|}$ ，如果 $v \in e$ ，其元素 $h(v, e)$ 为1，否则为0。然后我们有：

$$d(v) = \sum_{e \in E, v \in e} w(e) h(v, e), \quad (1)$$

$$\delta(e) = \sum_{v \in V, v \in e} h(v, e). \quad (2)$$

设 $D_e$ 和 $D_v$ 分别是由超边和顶点度组成的两个对角矩阵。设 $W$ 是包含超边权的 $|E| \times |E|$ 对角矩阵。

在下文中，我们定义了统一超图，其将用于对不同类型的对象之间的高阶关系进行建模。统一超图是具有多类型顶点和超边的超图。假设一个统一



我们定义  $\alpha = 1/(1 + \mu)$ 。注意到  $\mu/(1 + \mu)$  是一个常数，不会改变排名结果，我们可以将  $f$  改写如下：

$$f f = I \alpha A y。$$

(8)

可以证明矩阵  $I - \alpha A$  是可逆的。由于篇幅所限，省略了证明。注意，矩阵  $I - \alpha A$  是高度稀疏的。因此，计算可以非常高效。

## 4. 超图上的音乐推荐

在这一节中，我们将介绍我们在超图 (MRH) 上的音乐推荐方法。

### 4.1 数据收集

为了评估我们的算法，我们在 2009 年 12 月从 Last.fm 收集了数据。首先，我们收集了前 340 位最受欢迎的艺术家，以及对这些艺术家感兴趣的用户。将所有这些用户的朋友相加，我们得到了用户的候选集。然后，我们通过限制每个用户在集合中至少有一个朋友来减少候选用户集合。最终用户集由表示

表 1:数据集中的对象。

目标	记号	数数
用户	$U$	2596
组	$G$	1124
标签	$Tr$	3255
轨道	$Tr$	16055
相册	$Ar$	4694
艺术家	$Ar$	371

表 2:数据集中的关系。

关系	记号	数数
友谊关系	$R1$	4503
成员关系	$R2$	1124
倾听关系	$R3$	304860
标记轨道上的关系	$R4$	10936
相册上的标记关系	$R5$	730
艺术家的标签关系	$R6$	36812
音轨-专辑包含关系	$R7$	4694
专辑-艺术家包含关系	$R8$	371
轨道之间的相似性	$R9$	-

你。我们收集了基于这个用户集的其他对象和关系。我们下载了这些用户加入的所有组，并通过确保每个组在最终用户集中至少有五个成员来减少组集。对于资源对象和关系，我们抓取了每个用户的前 500 首经常播放的音乐曲目，以形成曲目的候选集。为了获得资源之间的包含关系，我们下载候选曲目集中所有曲目的所有对应艺术家和专辑，并移除候选曲目集中包含少于 5 首曲目的专辑。之后，我们获得了最终的资源集，即曲目集、专辑集和艺术家集，分别用  $Tr$ 、 $al$  和  $Ar$  表示。我们收集了本质上是三元组的标记关系，

(用户、标签、艺术家)。对于每个用户，我们下载了他/她的所有标签关系。我们只保留资源在  $Tr$  中的那些关系，或者先前获得的那些关系。最后一组标签由  $Ta$  表示。最后，我们从网上下载了音乐文件 (mp3 或 wma 格式)。总结了我们的实验中使用的对象和关系。

分别在表 1 和表 2 中。基于音乐内容计算音乐曲目之间的相似性。

#### 4.2 基于声学测量的相似性

音乐相似性的声学测量在最近几年得到了广泛的研究 [23, 37, 5, 25]。这些算法主要集中在几个中心问题上:1)提取什么样的代表特征;2)如何建立有限元模型

音乐的真实分布;3)如何度量分布模型之间的相似性。

为了简洁地表示音乐内容，在本文中，我们从 Mel 频率倒谱系数 (MFCCs) 中导出特征 [5]。MFCCs 在音频分类中很普遍。给定的音乐轨道被分割成短帧，并且为每一帧计算 MFCC。类似于 [23]，我们使用 K-means 将每个轨迹的所有帧分组到几个簇中。对于所有的聚类，平均值、协方差和权重被计算为音乐轨道的签名。到

比较两个不同轨道的特征，我们使用推土机的距离 (EMD) [31]。

### 4.3 统一超图构造

我们考虑了上述数据集中的六种类型的对象和九种类型的关系。这些对象包括用户、组、标签和三种类型的资源 (即曲目、专辑和艺术家)。这些关系分为四类:社会关系、对资源的作用、资源之间的包含关系和基于声学的音乐相似性关系。社会关系包括友谊关系和成员关系 (例如，兴趣小组)，分别由  $R1$  和  $R2$  表示。对资源的动作涉及四种类型的关系，即，收听关系 ( $R3$ )，以及对曲目、专辑和艺术家的标记关系 ( $R4$ 、 $R5$  和  $R6$ )。资源之间的包含关系是曲目和专辑、专辑和艺术家 ( $R7$  和  $R8$ ) 之间的包含关系。基于声学的音乐相似性关系由  $R9$  表示。

这六种类型的对象构成了 uni- 的顶点集

fied 超图。所以  $V = U \cup G \cup Ta \cup Tr \cup Al \cup Ar$ 。在统一超图中有九种类型的超边，每一种对应于一种特定类型的关系，如表 2 中所列。我们将超边集表示为  $E(i)$  对应于  $R_i$ ,  $i = 1, \dots, 9$ 。九种超高压的结构如下:

$e$ : 我们建立对应于每个成对友谊的超边，并将超边权重设置为 1。

$e$ : 对于每个组，我们建立一个超边，它包含对应于该组中所有用户的顶点，以及该组本身。注意，组本身也是一个对象。我们将超边权重设置为 1。

$e$ : 对于每个用户-音轨收听关系，我们建立一个包含用户和音乐音轨的超边。权重  $w(e(3))$  ( $e(3) \in E(3)$ ) 被设置为  $fre$

用户  $ui$  收听音轨  $tr_j$  的频率

$$w(e(3)) = \frac{1}{|Tr|} \sum_{ui \in U \text{ 和 } tr_j \in Tr} (ui, tr_j)$$

其中  $Q$  表示集合  $Q$  中包含的元素数量。为了消除偏差，我们将权重归一化为

$$w(e(3)) = \frac{w(e(3))}{\sum_{(3) \in E(3)} w(e(3))} \quad (9)$$

此外，为了处理不同类型的关系

(除了轨道之间的相似性关系) 同样，权重被进一步归一化如下:

$$w(e(3)) = \frac{w(e(3))}{\text{平均}(\text{西}(\text{东}), \text{西}(\text{西}))} \quad (10)$$

其中  $\text{ave}(w(e(3)))$  是用户  $ui$  的归一化权重的平均值。

$E/E/E$ : 如图 2(b) 所示，我们构建了超边缘来标记三种类型资源的关系。每个超边包含三个顶点 (对应于一个用户、一个标签和一个资源)，并且权重被设置为 1。

表 3: 统一超图及其子矩阵的关联矩阵 H。

(1)E(2)E(3)E(4)E(5)E(6)E(7)E(8)E(9)

U	用户设备(1)	用户设备(2)	用户设备(3)	用户设备(4)	用户设备(5)	用户设备(6)	0	0	0
G	0	通用电气(2)	0	0	0	0	0	0	0
Ta	0	0	0	TaE(4)	TaE(5)	TaE(6)	0	0	0
Tr	0	0	TrE(4)	TrE(5)	0	0	TrE(6)	0	TrE(9)
Al									
Ar									

$E/E$ : 我们为每张专辑建立了一个超级边缘, 其中包含了这张专辑中的所有曲目和专辑本身。同样, 艺术家的超边缘包含属于艺术家和艺术家本人的所有专辑。对应于专辑和艺术家的超边缘的权重被设置为 1。

$e$ : 我们基于基于声学的音乐相似性来构建  $k$  最近邻 ( $knn$ ) 图, 并且为我们的统一超图构建对应于  $knn$  图的边的超边。重量  $w(e(9))$  是  $sim$ —

第 4.2 节中计算的轨道长度  $tri$  和  $trj$ 。为了消除偏差, 我们将权重归一化为

$$f = \frac{w(e(9))}{\max(w(e(9)))} \quad (11)$$

其中  $\max(w(e(9)))$  是所有音乐模拟中的最大值  $ilarities$ 。我们引入参数  $c$  来控制音乐曲目的声学内容和其他社交媒体信息之间的相对重要性。最后, 重量是

$$w(e(9)) f = c f w(e(9)) f。$$

(12)

最后, 我们得到顶点-超边关联矩阵 H, 如表 3 所示, 以及权重矩阵  $w$ 。

#### 4.4 方法学

我们的音乐推荐算法 MRH 有两个阶段, 离线训练和在线推荐。在离线训练阶段, 我们首先构造如上所述的统一超图, 并获得顶点超边关联矩阵 H 和权重矩阵  $w$

gree 矩阵  $D_v$  和超边度矩阵  $D_e$  根据 H 和  $w$  计算得出。最后, 我们计算  $(I \alpha D_v 1/2 hwd 1HT D_v 1/2) 1$ , 记为  $(I \alpha A) 1$ , 其中  $\alpha$  设置正确。在在线推荐阶段, 我们需要首先构建查询向量  $y$ 。然后可以计算排名结果  $f f$ 。

我们的方法也可以应用于其他应用通过选择不同的顶点作为查询, 并考虑不同顶点类型的排序结果。例如, 如果我们选择一个用户作为查询, 音乐的排名结果

曲目可以用于音乐曲目推荐(即本文的主要关注点), 用户的排名结果

可用于好友推荐, 群组排名结果可用于兴趣群组推荐。对于标签推荐问题[34, 15], 我们

应该将目标用户和目标资源设置为查询, 并考虑标签的排名结果。

设置音乐曲目推荐的查询向量  $y_0$  有三种方法: (1) 将目标用户  $u$  对应的  $y$  项设置为 1, 其他项设置为 0。

(2) 将与目标用户  $u$  相对应的  $y$  的条目, 以及通过某种边缘连接到  $u$  的所有其他对象设置为 1。(3) 将目标用户  $u$  对应的  $y$  的条目设置为 1。此外, 如果  $u$  连接到对象  $v$ , 则将对对应于  $v$  的  $y$  的条目设置为  $Au, v$ 。注意,  $Au, v$  是  $u$  和之间的相关度的度量

动词 (verb 的缩写) 第一种方法没有考虑密切相关的对象, 这些对象也反映了用户的兴趣。第二种方法可能不是一个好的选择, 因为直觉上不同

ent 对象在不同程度上反映了用户的兴趣。因此, 在我们的实验中我们采用第三种方法。设置查询向量后, 排序结果  $f f$  可以是

计算出来的。对于音乐曲目推荐问题,

我们只考虑如上所述的音乐曲目的排序结果。最后, 我们可以向用户推荐他/她以前没有听过的排名靠前的曲目。

## 5. 实验

在这一节中, 我们研究了我们的音乐曲目推荐方法的使用。

### 5.1 比较算法

我们将我们的 MRH 算法与五种推荐算法进行了比较。第一种是基于用户的协同过滤 (CF) 方法[29, 19], 其仅使用收听关系。我们选择基于用户的 CF 算法, 因为与 CF 的传统数据集不同, 我们的数据集比用户有更多的音乐曲目。给定一个目标用户  $ui$ , 设  $r_{ui}, trp$  为音乐曲目  $trp$  的用户  $ui$  的预测排名分数, 由[19]给出

$$r_{ui}, trp = \frac{\sum_{j=1}^k (3) *_{ij} - (3)}{k}, \quad (13)$$

$w(ei.)$  + 哪里

$$\frac{\sum_{trp=0}^{\sigma} (3) *_{trp} - (3) *_{trp} w(eip)}{trp \in}$$

以及  $s_{ui}, u_j$  是用户  $s_{ui}$  和  $u_j$  之间的相似性权重。  $k$  是用户  $ui$  的最近邻居的数量。我们采用基于余弦的方法[6, 32]来计算相似性

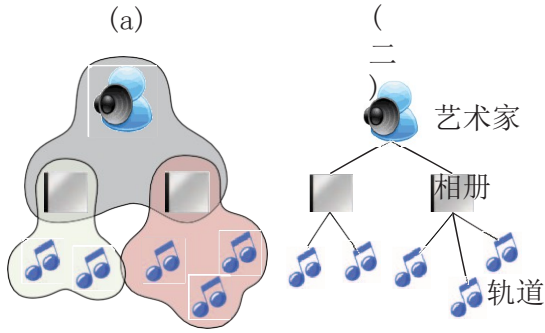


图 3:在两个模型中表示的包含关系: (a) 我们的统一超图模型, 和 (b) 普通图模型。

用户之间:

$$sui, u_j = \sigma \left( \frac{Tr_{p=1} w(e(3))}{\sum_{p=1}^2 w(e(3))} \right) \quad (15)$$

基于所获得的相似性, 我们使用显著性 [17] 中提出的加权方法, 用于提高推荐性能。具体来说, 如果两个用户之间共同收听的音乐曲目的数量 (用  $N$  表示) 小于阈值数量  $N$ , 那么我们将它们的相似性乘以  $n/N$ 。在我们的实验中, 我们根据经验将  $N$  的值设置为 20, 将最近邻居的数量  $k$  设置为 5, 以实现最佳性能。

第二种比较算法是基于声学的音乐推荐方法 [38], 其使用收听关系和音乐相似性关系。它用 AB 表示。

第三种比较算法使用我们下载的数据集中的所有信息。与 MRH 不同, 我们使用普通图表来模拟社交媒体信息。具体来说, 我们通过图结构对标记关系进行建模, 如

图 2(a), 并通过如图 3 所示的树结构对成员和包含关系进行建模。[45] 中描述的图排序算法用于计算最优

排名分数。我们把这种算法称为统一图上的推荐 (RUG)。

第四种比较算法是我们的 MRH 方法, 但是仅使用收听关系和音乐相似性关系 (即,  $R3$  和  $R9$ )。这种方法称为 MRH 混合法。第五种比较的算法是我们的 MRH 方法, 但是没有使用音乐相似关系。它使用所有其他八种类型的关系。这种方法被称为 MRH 社会。

## 5.2 估价

为了评估我们的 MRH 算法和其他比较算法的性能, 对于每个用户, 我们随机选择 20% 的收听关系作为评估目的的测试数据。如果用户可以访问测试装置中的某个轨道  $tr$ , 我们要求他/她不能访问训练装置。为了实现这一点, 我们移除所有对应的标记关系, 留给我们最终的训练集。

对于评价指标, 我们使用精确度、召回率、F1、平均精确度 (MAP) 和归一化折扣累积增益 (NDCG) 来衡量不同

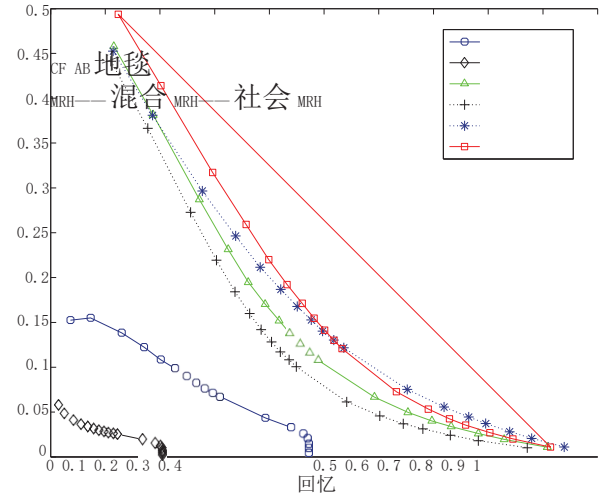


图 4:所有六种人工智能的召回精度曲线

算法。

推荐项目的数量。召回的定义是

正确推荐的项目数除以应该推荐的项目总数 (即目标用户实际收听的项目)。F1 是精确度和召回率的调和平均值。Average Precision (AP) 是在推荐列表中每个正确推荐项目的点上计算的平均精度:

$$AP = \frac{\text{普通}}{\text{正确推荐的项目数量}}, \quad (16)$$

其中  $Precision@i$  是排名位置  $i$  的精度,  $N$  是推荐项目的数量, 如果位置  $i$  的项目被正确推荐, 则  $corri = 1$ , 否则  $corri = 0$ 。MAP 是所有用户的平均精度分数的平均值。n 位的 NDCG 定义为:

$$NDCG@n = \frac{IDCG}{IDCG} \quad (17)$$

推荐算法的性能。精确度定义为正确推荐的项目数除以 to-



其中  $r_i$  是等级为  $i$  的项目的相关性评级。在我们的情况下，如果用户已经听过该推荐的音乐，则  $r_i$  为 1，否则为 0。选择 IDCG 使得完美排名的 NDCG 值为 1。

### 5.3 性能比较

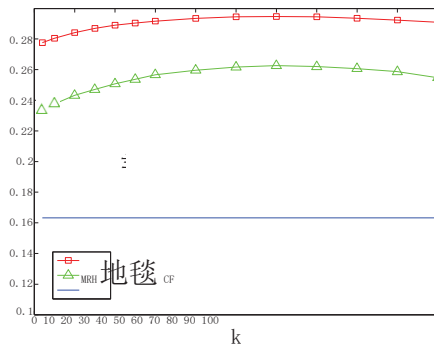
我们使用 5.2 节中提到的所有评估指标来衡量每个推荐算法的性能。图 4 显示了所有六种算法的召回精度曲线。我们在表 4 (MAP 和 F1) 和表 5 (NDCG) 中报告了所有六种算法在 MAP、F1 和 NDCG 方面的性能。很明显，我们提出的算法在大多数情况下明显优于其他推荐算法，尤其是在低秩的情况下。注意，我们提出的 MRH 算法通过超边对高阶关系建模，而 RUG 使用普通图来近似这些高阶关系。MRH 相对于 RUG 的优势表明超图确实是在社交媒体社区中建模复杂关系的更好选择。基于声学的 (AB) 方法效果最差。这是因为基于声学的方法导致了语义差距，并且基于声学内容的相似性

表 4: 推荐算法在 MAP 和 F1 方面的比较。(粗体排版表示性能最佳。\*表示与第二好的相比,  $p < 0.001$  的统计显著性。)

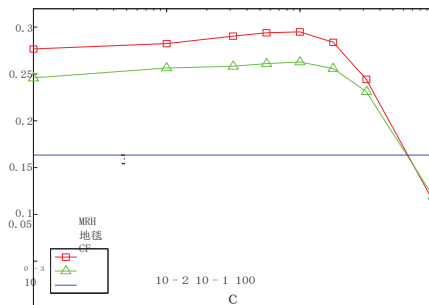
	地图	F1@5	F1@10	F1@20	F1@30	F1@50	F1@70	F1@100	F1@200
随军牧师	0.1632	0.0557	0.0929	0.1243	0.1329	0.1294	0.1197	0.1064	0.0765
AB 型血	0.0762	0.0226	0.0303	0.0377	0.0403	0.0421	0.0415	0.0401	0.0334
小地毯	0.2626	0.1729	0.2323	0.2587	0.2516	0.2237	0.1988	0.1701	0.1169
MRH 混合动力车	0.2470	0.1653	0.2224	0.2451	0.2377	0.2099	0.1855	0.1581	0.1076
MRH-社会	0.2755	0.1705	0.2311	0.2654	0.2660	0.2440	0.2202	<b>0.1906</b>	<b>0.1318*</b>
MRH	<b>0.2948*</b>	<b>0.1855*</b>	<b>0.2510*</b>	<b>0.2839*</b>	<b>0.2799*</b>	<b>0.2509*</b>	<b>0.2227</b>	0.1892	0.1270

表 5: 根据 NDCG 比较推荐算法。(粗体排版表示性能最佳。\*表示与第二好的相比,  $p < 0.001$  的统计显著性。)

	NDCG@5	NDCG@10	NDCG@30	NDCG@50	NDCG@70	NDCG@100	NDCG@200
随军牧师	0.1522	0.1713	0.2519	0.2987	0.3278	0.3579	0.4120
AB 型血	0.0733	0.0820	0.1241	0.1532	0.1749	0.2027	0.2556
小地毯	0.4849	0.4318	0.3826	0.4109	0.4345	0.4587	0.5037
MRH 混合动力车	0.4587	0.4091	0.3640	0.3911	0.4124	0.4346	0.4753
MRH-社会	0.4759	0.4268	0.3866	0.4197	0.4480	0.4763	0.5264
MRH	<b>0.5192*</b>	<b>0.4650*</b>	<b>0.4174*</b>	<b>0.4484*</b>	<b>0.4740*</b>	<b>0.4987*</b>	<b>0.5419*</b>



(a)  $k$  与地图



(b)  $c$  与 MAP

图 5: 相似关系中  $k$  和  $c$  的参数设置。首先, 我们把  $c$  固定在 0.1, 让  $k$  变化。(a) 显示了通过 MAP 测量的性能。然后我们把  $k$  固定在 60, 让  $c$  变化。(b) 显示了通过 MAP 测量的性能。

并不总是与人类的知识相一致 [10]。CF 算法也不好。这可能是由于我们的数据集中的用户跟踪矩阵非常稀疏, 只有大约 0.6% 的非零条目。MRH 混合算法只使用了音轨和收听关系之间的相似性关系, 但是它比 AB 和 CF 效果好得多。

与 MRH 社交相比, MRH 额外使用了音乐曲目之间的相似关系。我们发现, 使用这种基于声学的信息可以改善推荐结果, 尤其是当召回率较低时。这是因为基于声学信息可以缓解协同推荐系统中与数据稀疏性相关的一些众所周知的问题, 例如, 用户偏差、非关联和冷启动问题 [20]。

## 5.4 探索参数设置

在我们的算法中有三个参数, 即 4.3 节中提到的最近邻的数量  $k$ , 等式中的  $c$ 。等式 (12) 和  $\alpha$ 。(8)。

为了探索参数  $k$  和  $c$  的影响, 我们使用 MAP 作为评估度量。图 5 显示了结果。首先, 我们根据经验将  $c$  固定在 0.1, 让  $k$  变化。图 5(a) 显示了作为  $k$  的函数测得的性能。当  $k$  在 60 左右时获得最佳结果。然后我们把  $k$  固定在 60, 让  $c$  变化。图 5(b) 显示了作为  $c$  的函数测量的性能。获得了最佳结果。

当  $c = 0.1$  时。可以看出, 我们的算法在大范围的参数变化中始终优于其他两个比较的算法。在我们的实验中, 对于 MRH、MRH 混合动力车和 RUG, 我们将  $k$  设为 60,  $c$  设为 0.1。  $\alpha$  是我们的 MRH 算法和 RUG [45] 共有的一个参数。在我们的实验中, 我们将  $\alpha$  设置为

0.98。MRH, MRH-混合, MRH-社会, 从经验上看是崎岖不平的。

## 5.5 社会信息贡献

为了探索不同类型的社交媒体信息对推荐性能的贡献, 我们研究了 MRH 在四个不同社交媒体信息子集上的性能。第一个子集仅包含监听关系 (即 R3), 它被视为基础

表 6:根据 MAP 和 F1 比较不同社会信息子集的 MRH。(粗体排版表示性能优于单独使用监听关系 (R3) 的性能。\*表示与仅使用收听关系的算法相比, 在  $p < 0.001$  时的统计显著性。)

	地图	F1@5	F1@10	F1@30	F1@40	F1@60	F1@70	F1@100	F1@200
R3 上的 MRH	0.2303	0.1430	0.1996	0.2332	0.2143	0.1945	0.1772	0.1695	0.1184
MRH 在 R1, R2, R3	<b>0.2308</b>	<b>0.1444</b>	<b>0.1998</b>	<b>0.2337</b>	<b>0.2146</b>	0.1943	0.1772	0.1695	0.1181
R4 R3 上的 MRH	0.2303	<b>0.1432</b>	<b>0.1997</b>	0.2332	0.2143	0.1945	<b>0.1773</b>	0.1695	0.1184
R3、R7 和 R8 上的 MRH	<b>0.2757*</b>	<b>0.1748*</b>	<b>0.2339*</b>	<b>0.2642*</b>	<b>0.2413*</b>	<b>0.2176*</b>	<b>0.1970*</b>	<b>0.1878*</b>	<b>0.1299*</b>

关系。第二个子集包含倾听关系和社会关系(即 R1、R2)。第三个子集包含曲目上的列表关系和标记关系(即 R4)。第四子集包含倾听关系和包含关系(即, R7, R8)。从表 6 中, 我们可以看到包含关系显著提高了推荐性能。通过利用资源之间的包含关系, 我们可以推荐相同或相似唱片中的音乐曲目, 以及由相同或相似艺术家演奏的曲目。可以看出, 通过使用社会关系, 在低回忆区域有轻微的改善。直觉上, 用户的喜好可以从友谊和会员关系中推断出来。标记关系不会提高性能。这是因为人们通常会把他们已经听过的音乐收藏起来。因此, 听力关系和标记关系之间有很强的相关性, 因此标记关系的使用受到限制。

## 6. 结论和未来工作

本文针对音乐社交社区中的音乐推荐问题, 着重结合各类社交媒体信息和音乐听觉信号。我们将推荐问题建模为统一超图上的排序问题, 并提出了超图(MRH)上的音乐推荐算法。MRH 构建了一个超图, 将音乐社交社区中的多类型对象建模为顶点, 将这些对象之间的关系建模为超边。基于声音信号的音乐轨道之间的相似性被视为一种关系。这样就可以自然地捕捉到社会信息中的高阶关系。此外, 协同过滤和基于声学音乐推荐被结合在一个统一的框架中。基于所构建的超图, 我们然后使用正则化框架来导出查询主题的排名结果。我们将用户视为查询, 并向用户推荐排名靠前的音乐曲目。在音乐社交社区 Last.fm 数据集上的实验表明, 该算法明显优于传统推荐算法, 丰富的社交媒体信息对音乐推荐非常有用。

MRH 还可以用于其他类型的社交媒体社区的推荐系统, 例如电影和图片。在这项工作中, 我们平等地对待所有类型的社会关系(除了音乐相似性关系)。然而, 在实际应用中, 不同类型的关系可能具有不同的重要性。例如, 在一些纯社交网络中, 如 Facebook<sup>3</sup> 和 LinkedIn<sup>4</sup>, 人们的口味

<sup>3</sup><http://www.facebook.com> 四 <http://www.linkedin.com>

用户会受到他们朋友的显著影响。在这种情况下, 我们应该赋予社会关系相对较高的权重, 如友谊和成员关系。另一方面, 对于特殊兴趣社交媒体社区(例如 Last.fm 和 YouTube<sup>5</sup>), 统一超图模型应该更加强调用户对资源的行为(例如评级和标记)以及资源之间的关系(例如包含关系)。

此外, 如 4.4 节所述, 我们的方法不限于音乐曲目推荐。我们可以在不同的应用中利用它, 例如朋友推荐和个性化标签推荐。这些问题留给我们今后的工作。

## 感谢

本工作得到了国家重点科技 R&D 计划(2008BAH26B00, 2007BAH11B06)和国家自然科学基金(60875044, 90920303)的资助。

## 7. 参考

- [1] G. Adomavicius 和 A. Tuzhilin. 走向下一个推荐系统的产生: 一个综述  
最先进的和可能的扩展。IEEE 知识与数据工程汇刊, 17(6):734 - 749, 2005。
- [2] 南阿加瓦尔. 图表数据的排名。进行中。第 23 届国际机器学习会议, 匹兹堡, 宾夕法尼亚州, 2006 年。
- [3] 南阿加瓦尔, k. 布兰森和 s. 贝隆吉. 用图进行高阶学习。进行中。第 23 届国际机器学习会议, 匹兹堡, 宾夕法尼亚州, 2006 年。
- [4] J. -J. 奥库图里尔和 f. 帕切特. 扩大音乐播放列表的生成。进行中。2002 年瑞士卢桑纳 IEEE 多媒体与博览会国际会议。
- [5] A. Berenzweig, B. Logan, D. P. W. Ellis 和 B. Whitman.  
声学 and 主观音乐相似性测量的大规模评估。计算机音乐杂志, 28(2):63 - 76, 2004。
- [6] J. Breese, D. Heckerman 和 C. Kadie. 协同过滤预测算法的实证分析。进行中。第 14 届人工智能不确定性会议, 旧金山, 加州, 1998 年。
- [7] 南 R. Bul`o 和 M. Pelillo. 超图聚类的博弈论方法。神经信息处理系统进展 22, 加拿大温哥华, 2009。
- [8] R. 蔡, 张, 张, 马伟英. 通过搜索进行可扩展的音乐推荐。进行中。第 15 届国际多媒体会议, 德国奥格斯堡, 2007。
- [9] 页 (page 的缩写) 卡诺, m. 科彭伯格和 n. 瓦克. 基于内容的音乐音频推荐。进行中。第 13 届 ACM 多媒体国际会议, 新加坡, 2005。
- [10] o`. 塞尔马. 跨越音乐推荐中的语义鸿沟。进行中。第五届国际语义网会议, 雅典, 格鲁吉亚, 2006。

<sup>5</sup><http://www.youtube.com>

- [11] o`。塞尔玛和p.拉梅尔。如果你喜欢披头士,你可能会喜欢...:音乐推荐教程。进行中。第16届ACM多媒体国际会议,温哥华,加拿大不列颠哥伦比亚省,2008。
- [12] 南陈,王,张。基于高阶关系的同步异构数据聚类。进行中。2007年在内布拉斯加州奥马哈举行的第7届IEEE国际数据挖掘研讨会。
- [13] J.迪耶德里奇和t.约弗丘。基于大众分类法从用户档案中发现实践社区。进行中。2006年在希腊克里特岛举行的第一届建筑技术实践社区强化学习解决方案国际研讨会。
- [14] J.唐纳森。面向流行音乐的混合社会声学推荐系统。进行中。2007年纽约ACM推荐系统会议。
- [15] Z.关、步军、梅青、陈、王。在多类型相关对象上使用基于图的排序的个性化标签推荐。进行中。第32届ACM SIGIR信息检索研究与发展会议,波士顿,马萨诸塞州,2009。
- [16] A.哈帕莱和杨。面向协同过滤的个性化主动学习。进行中。2008年第31届SIGIR ACM信息检索研究与发展会议。
- [17] J.L. Herlocker, J. A. Konstan, A. Borchers and J. Riedl。用于执行协同过滤的算法框架。进行中。第22届美国计算机学会SIGIR信息检索研究与发展会议,1999年。
- [18] 页(page的缩写)膝盖, T. Pohle, M. Schedl 和 G. Widmer。将基于音频的相似性与基于网络的数据相结合,以加速自动音乐播放列表生成。进行中。第八届ACM多媒体信息检索国际研讨会,加州圣巴巴拉,2006。
- [19] I. Konstantas、V. Stathopoulos 和 J. M. Jose。社交网络与合作推荐。进行中。第32届ACM SIGIR信息检索研究与发展会议,波士顿,马萨诸塞州,2009。
- [20] Q. Li、S. H. Myaeng 和 B. M. Kim。一种考虑用户意见和音频特征的概率音乐推荐器。信息处理与管理, 43(2):473 - 487, 2007。
- [21] 名词(noun的缩写)刘。还有q.杨。一种面向排序的协同过滤方法。进行中。2008年第31届SIGIR ACM信息检索研究与发展会议。
- [22] B.洛根。歌曲集的音乐推荐。进行中。第五届国际音乐信息检索会议,西班牙巴塞罗那,2004年。
- [23] B.洛根和a.所罗门。基于信号分析的音乐相似度函数。进行中。2001年日本东京IEEE国际多媒体与博览会会议。
- [24] H.马,金和吕先生。用社会信任集合学习推荐。进行中。第32届ACM SIGIR信息检索研究与发展会议,波士顿,马萨诸塞州,2009。
- [25] C.麦凯和我藤永。组合从音频、符号和文化来源提取的特征。进行中。第九届国际音乐信息检索会议,费城,宾夕法尼亚州,2008。
- [26] H.-s. Park j. o. Yoo 和 s. b. Cho。基于效用理论的模糊贝叶斯网络的上下文感知音乐推荐系统。模糊系统与知识发现, 4223:970 - 979, 2006。
- [27] 南鲍斯、w.维海格和m.沃森。使用本地搜索快速生成最佳音乐播放列表。进行中。第七届国际音乐信息检索会议,加拿大维多利亚,2006。
- [28] R.拉格诺、伯格和赫利。推断音乐对象之间的相似性并应用于播放列表
- 代。进行中。第七届ACM SIGMM多媒体信息检索研讨会,新加坡希尔顿,2005。
- [29] 页(page的缩写)雷斯尼克, n.亚科武, m.苏查克, p.博格斯特伦和J.里德尔。Grouplens:网络新闻协同过滤的开放架构。进行中。美国计算机学会计算机支持的协同工作会议,教堂山,北卡罗来纳州,1994。
- [30] 南Rho, b.韩军和E. Hwang。基于Svr的音乐情绪分类和基于上下文的音乐推荐。进行中。第17届ACM多媒体国际会议,中国北京,2009。
- [31] Y. Rubner、C. Tomasi 和 L. J. Guibas。推土机的距离作为图像检索的度量。国际计算机视觉杂志, 40(2):99 - 121, 2000。
- [32] B.萨瓦尔、g.卡里皮斯、j.康斯坦和j.雷德尔。基于项目的协同过滤推荐算法。进行中。2001年,第十届国际互联网大会,中国香港。
- [33] 南森、j.维格和j.里德尔。Tagommenders:通过标签将用户与项目联系起来。进行中。2009年在西班牙马德里举行的第18届国际互联网大会。
- [34] Y.宋,庄,李,赵,李俊杰,李伟中,和C. l.贾尔斯。实时自动标签推荐。进行中。2008年第31届SIGIR ACM信息检索研究与发展会议。
- [35] 长度孙,季,叶。多标签分类的超图谱学习。进行中。第14届ACM SIGKDD知识发现和数据挖掘国际会议,内华达州拉斯维加斯,2008。
- [36] 页(page的缩写)Symeonidis, M. Ruxanda, A. Nanopoulos 和Y.马诺洛普洛斯。个性化音乐推荐的社会标签三元语义分析。进行中。2008年美国宾夕法尼亚州第九届国际音乐信息检索会议。
- [37] D.陶,刘,唐。K-box:一个基于歌曲查询的音乐检索系统。进行中。第12届ACM多媒体国际会议,纽约,纽约州,2004年。
- [38] 米(meter的缩写)蒂尔曼和s.鲍斯。面向混合音乐推荐的集成学习。进行中。2007年纽约ACM推荐系统会议。
- [39] K.H. L. TsoSutter、L. B. Marinho 和 L. Schmidt-Thieme。融合协同过滤算法的标签感知推荐系统。进行中。2008年ACM应用计算研讨会,巴西福塔莱萨。
- [40] K.吉井和后藤先生。用于混合音乐推荐的连续pLSI和平滑技术。进行中。第十届国际音乐信息检索学会会议,日本神户,2009。
- [41] K.Yoshii, M. Goto, K. Komatani, T. Ogata 和 H. G. Okuno。使用具有潜在用户偏好的概率模型的混合协作和基于内容的音乐推荐。进行中。第七届国际音乐信息检索会议,加拿大维多利亚,2006。
- [42] Y.张和j.科伦。推荐系统中高效的贝叶斯层次用户建模。进行中。第30届ACM SIGIR信息检索研究与发展会议,荷兰阿姆斯特丹,2007。
- [43] Z.张,周,张。基于用户-商品-标签三分图的集成扩散个性化推荐。物理A:统计力学及其应用, 389(1):179 - 186, 2009。
- [44] D.周,黄军,和B. Scho lkopf。用超图学习:聚类、分类和嵌入。神经信息处理系统进展19,加拿大温哥华,2006。
- [45] D.周、j.韦斯顿、a.格雷顿、o.布斯凯和B. Sch olkopf。数据流形上的排名。神经信息处理系统进展16,剑桥,麻省,2003。