# 面向隐式反馈的推荐系统研究现状与趋势

陆 艺 曹 健
(上海交通大学计算机科学与工程系 上海 200240)

**摘 要** 推荐系统作为解决信息过载的一种有效手段,已成为工业界和学术界的研究热点,它依据用户的显式或隐式反馈信息推测其需求、兴趣等,将其偏好的信息、产品等推荐给他们。面向显式反馈信息的推荐方法是目前的主流,而隐式反馈信息的普遍性使得基于此类的推荐方法具有更广的适用性,但是,隐式反馈信息并不能直接反映用户的偏好,因而利用它进行推荐具有很大的挑战。首先阐述了隐式反馈的特性以及基于此类信息进行推荐的必要性和所面临的问题;然后对面向隐式反馈的推荐算法给出了全面的、系统的分类,在此基础上比较了各类隐式反馈的推荐方法的优、缺点,并进一步分析了适用于隐式反馈推荐方法的多种评价指标;最后讨论了面向隐式反馈推荐方法的未来发展方向。

**关键词** 推荐算法,隐式反馈,推荐评估指标
**中图法分类号** TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.4.002

## Research Status and Future Trends of Recommender Systems for Implicit Feedback

LU Yi CAO Jian
(Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

**Abstract** As an effective approach addressing information overloading problem, recommender system has been a hotspot in both industry and academia, which infers users' potential requirements and interests by utilizing their explicit/implicit feedbacks, and then recommends them with preferable information or products. The recommendation methods based on explicit feedbacks are the mainstream approaches in this area, however, because of the prevalence of implicit feedbacks, the recommendation methods for implicit feedbacks can be more widely applied. Because the implicit feedbacks cannot reflect users' preferences directly, to recommend products relying on implicit feedbacks is a more challenging task. The characteristics of implicit feedback, necessity and problems of recommendation for implicit feedbacks were illustrated firstly. Then a systematic taxonomy for various recommendation methods on implicit feedback was proposed. On this basis, the strength/weakness of these approaches and evaluation metrics for implicit feedback oriented recommendation were analyzed. Lastly, the possible directions of implicit feedback oriented recommendation in the future were discussed.

**Keywords** Recommendation algorithms, Implicit feedback, Recommendation evaluation metrics

## 1 引言

随着信息技术的发展,人们面临着日益严重的"信息过载现象"(Information Overload),如何帮助用户迅速发现其所需的信息成为了一个重要挑战。在此背景下,推荐系统应运而生。无论是在电子商务中为用户寻找其可能喜欢的商品,还是在社交应用中向用户推荐朋友,个性化推荐都展示出了不可忽视的价值。在应用需求的推动下,学术界对此也产生了巨大的热情。因而,个性化推荐技术现在已成为学术界和工业界的研究和应用热点。

推荐系统通过对用户偏好的理解和建模,评估待推荐产品或事物对于用户的效用,并最终决定如何呈现给用户。用户偏好可以通过如评论、评分以及个人主动提供的相关反馈等来获得,它们直接表达了用户的喜好,因而这类反馈信息被称作显式反馈。与之相对,不能直接表达用户对事物喜恶程度的信息如用户收藏夹、购买历史、搜索模式以及点击方式等则被称作隐式反馈。

目前,许多推荐系统仅仅采用显式反馈作为输入,而忽视了隐式反馈信息,其主要原因在于该类信息对偏好无法完全表达(例如缺少用户对某产品不喜欢的反馈),据此难以理解用户行为并对其建模。尽管隐式反馈不如显式反馈精确,但在实际中隐式反馈远比显式反馈丰富,在一些场景中甚至只能获取到隐式反馈;而许多研究也发现[1-3],隐式反馈在交互环境中可以成为显式反馈的替代。因此,近年来许多学者将目光转到了基于隐式反馈的个性化推荐的研究上,并提出了不同的算法和策略。本文研究了基于隐式反馈的推荐方法,并对其做了系统的分类,结合隐式反馈的特点与推荐方法的不同侧重进行了分析和总结。

到稿日期:2015-04-11 返修日期:2015-08-14
陆 艺(1990—),男,硕士生,主要研究领域为推荐系统算法、数据分析,E-mail:lythesia@sjtu.edu.cn;曹 健(1972—),男,博士,教授,CCF高级会员,主要研究领域为服务计算、数据分析。

本文第 2 节阐述显式与隐式反馈的不同特点，强调了隐式反馈对个性化推荐带来的挑战；第 3 节分类介绍和分析当前主要的基于隐式反馈的推荐算法；第 4 节总结隐式反馈推荐适用的评价标准；最后对未来的研究方向进行展望。

## 2 隐式反馈推荐的关键问题

### 2.1 隐式反馈信息的特征

在推荐系统中，显式反馈往往被作为反映用户偏好的真值标准(Truth Ground)，因为它们同时表示了用户的喜好，并细化了喜好的程度和因素，信息更加完整与丰富。显式反馈最直观的类型便是数值评分，例如 Amazon[1] 商城与 IMDb[2] 都用星级评分来获取用户的反馈；此外，标签数据也是其中的一种，例如 Xiami[3] 在用户收藏音乐作品或艺人的同时提供打标签功能，这一类反馈更细致地反映了用户偏好。但是显式反馈在现实中并不总是可用的：一方面是来源问题，即提供某些反馈涉及到用户的隐私并且多数用户不愿承担提供反馈的额外负担；另一方面是可靠性问题，即用户对事物的评价可能受到其它因素的影响而缺乏客观性，或其兴趣变化导致旧反馈数据一定程度上的失效等。

相较之下，隐式反馈在一定程度上缓解了上述问题，不同于显式反馈，前者在实际的信息系统中大量存在，因此易于获取。例如之前提到的浏览记录和购买历史等，都是在用户没有察觉的情况下由系统自动记录的，因而避免了获取隐私的敏感以及用户操作负担；此外，该类反馈信息是连续产生的，无需主动提供，同样降低了信息滞后失效的可能。然而，隐式信息自身也存在着显著缺陷，Hu 等人[4]将其归纳如下：

1)缺少负反馈

通过用户的行为记录，可以推测其可能喜欢某件商品，但无法知道是否不喜欢某物品。例如，一个用户没有购买某物品的原因可能是对其不感兴趣，也可能仅仅是不知道它的存在。我们只能获取正反馈信息，而缺少负反馈信息——用户不喜欢的物品信息。而显式反馈中同时包含了正、负反馈(如高低评分)[5]。因此，对于数据中缺失负反馈信息的处理是隐式反馈推荐方法需要解决的重点问题。

2)数据噪音

仅依靠用户行为对用户偏好进行推测是很困难的。例如，用户的购买行为并非出于自己的目的(公司采购或者购买礼物等)，或者购买后而感到失望等，这导致了数据中的正反馈并不一定反映用户的真实喜好。尽管 Amatriain 等人[6]提出显式反馈中存在着数据噪音，但隐式反馈由于不是用户的主观反馈，其反馈可信度较显式更低、更缺乏规律性，因此也更难处理。

3)偏好表达与可信度

显式评分系统提供了用户直接表达喜好程度的能力，例如在五分制的评分系统中，1 表示完全不喜欢，5 则表示非常喜欢。隐式反馈的数值中通常包含用户行为的频率，比如播放次数、点击率和购买量等。更高的数值并不一定表示用户更喜欢该物品；但另一方面，它对偏好的表达又具有更高的可信度，因为一次性事件代表了偶然性，而高频事件则较为可信地反映了用户的喜好。

4)评估标准的有效性

在基于显式反馈的推荐系统中，传统的评估标准包括使用均方误差来衡量预测与实际评分之间的距离等，以此表达推荐结果的准确度。然而在隐式反馈的情形下，我们并不知道真正的负样本，因此基于准确度的评估准则无法适用；相较于显式反馈可以利用预测评分与实际评分的绝对差值来比较和排序，隐式反馈的预测则表现为二元性——非正样本即负样本，这些比较标准显然无法适用，而比较样本为正(负)样本的可能性高低这样的相对比较方式可能更为合适。因此，对隐式反馈推荐方法的评估标准的合理设计也是隐式反馈推荐方法中需要研究的问题。

Jawaheer 等人[7]将隐式和显式反馈的相关特点进行了比较，在此基础上结合上述分析将其归纳在表 1 中。

表 1 显式与隐式反馈的特征

| | 隐式反馈 | 显式反馈 |
|---|---|---|
| 准确度 | 低 | 高 |
| 丰富度 | 高 | 低 |
| 获取难度 | 容易 | 困难 |
| 数据噪音 | 较难识别 | 较易识别 |
| 上下文敏感 | 是 | 是 |
| 用户偏好的表达能力 | 只有正样本 | 包含正负样本 |
| 评估比较标准 | 相对比较 | 绝对比较 |

### 2.2 隐式反馈数据的获取

虽然隐式数据在信息系统中大量存在，但真实可用于研究和应用的数据集的获取仍然不是易事。一方面，由于数据产生的速率极快、数量极大，对数据的搜集和规整具有相当的挑战；另一方面，因为推荐系统兴起较晚，尤其是人们对隐式数据推荐的认识较晚，导致目前隐式数据虽然存在很多，但是积累很少。目前，在学术界和工业界的关注与努力下，越来越多的隐式数据集经由研究者整理并公布出来，以下是较为权威或常用的隐式数据集和数据源。

Last. fm[4]：世界最大的社交音乐平台，拥有超过 1 亿首歌曲曲目和超过 1000 万名歌手艺术家的收藏记录，它还提供了公开的 API 接口供使用者获取乐曲、艺术家以及用户的在线播放信息等，其中播放信息隐式地记录了用户播放的曲目以及时间戳。

KDD Cup 竞赛：由 SIGKDD 主办的国际知识发现和数据挖掘竞赛，该赛事历年的题目和数据取自不同的挖掘领域，且都具有很强的应用背景。虽然并不是针对隐式数据和推荐算法，但其仍有可用于研究的数据，比如 2012 届[5]的基于微博数据预测用户感兴趣的人物或话题，以及根据用户查询和广告信息来预测用户对广告的点击情况等。

Delicious. com[6]：世界最大的书签共享社交平台，拥有

1) http://amazon. com

2) http://www. imdb. com

3) http://www. xiami. com

4) http://www. last. fm/

5) http://www. kddcup2012. org

6) http://delicious. com

530 万条用户及 1.8 亿条 URL 书签，许多第三方组织和学者都曾收集整理该站点的用户收藏信息并加以研究。数据不仅包括用户对 URL 的收藏记录，也包括用户之间的好友和共享关系，在隐式数据的基础上补充了社交信息。

此外还有 RecSys 会议收录文章所整理的一些公开数据集，包括 MovieLens[1)] 和 Flixster[2)] 的电影评分数据（可以通过预处理转化为隐式数据集进行研究），以及 Million Song Chanllenge[3)] 竞赛根据用户收听历史预测其喜欢乐曲的公开数据等。

相比于显式反馈，虽然可用的隐式反馈数据显得较少，但是鉴于后者的推荐越来越受到学术界及业界的重视，相信更多的企业组织和研究者们会更加注重隐式反馈数据的积累和整理。希望本文研究能为后来的研究者们提供有益帮助。

## 3 面向隐式反馈的推荐技术

在介绍基于隐式反馈的推荐技术之前，首先简要回顾一下传统的推荐技术，它们主要包括：基于内容的推荐、基于协同过滤的推荐以及混合推荐[8]。基于内容的推荐的思想是：对用户评价过的对象抽取特征，构建用户的兴趣模型（用户资料，Profile），再依据用户资料与待预测项目的匹配程度进行推荐。基于协同过滤的推荐是目前被研究得最多的推荐技术，它基于邻居用户的信息得到目标用户的推荐，其方式主要分为两类：1）基于近邻的协同过滤，典型做法是通过相似计算得到目标用户（或者物品）的近邻集合，再以该集合中的信息进行预测推荐；2）基于模型的协同过滤，广泛使用的技术包括潜在因子模型（Latent Factor Model，LFM）[9] 和基于贝叶斯网络（Bayesian Network）[10] 的模型等。

然而，上述的推荐方法并不直接适用于隐式反馈：对于基于内容的推荐，虽然隐式反馈的数量很多，但是由于其数据噪声和稀疏性的缺陷，传统的方法（比如向量空间模型）难以进行特征抽取；对于基于协同过滤的推荐，一方面，加权评分的方式不适用于隐式反馈（因为其本身不包含评分信息），另一方面，由于缺少负样本，基于模型的方法并不能学习到关于负样本的信息，会使模型偏置。

针对隐式反馈中缺少负反馈和信息量较少等问题，不少学者提出了基于隐式反馈的推荐方法，这些方法可以分为以下几类：基于单类协同过滤的推荐、引入辅助信息的推荐以及基于排序的推荐。

### 3.1 基于单类协同过滤的推荐

如上所述，隐式反馈存在难以区分负样本和潜在正样本的缺陷，Pan 等人[11] 将只存在正样本的协同过滤称为单类协同过滤问题（One-Class Collaborative Filtering，OCCF）。

针对缺少负反馈的问题，存在两种比较直观的解决策略：1）把所有缺失数据视为负样本（All Missing As Negative，AMAN），尽管实际中用户不喜欢的物品可能确实占大多数，但这种方式使得单个用户喜好分布失衡，导致推荐结果出现偏斜；2）把所有缺失数据视为未知（All Missing As Unknown，AMAU），这种做法同样无法产生好的推荐结果，因为缺失部分的信息无法被利用，导致其结果近乎随机。

单类协同过滤问题的关键在于如何处理缺失数据中的负样本，因此产生了两类不同的处理方式：1）对每个样本实例基于不同的权重，将缺失数据也一并纳入基于权值的模型进行训练；2）对缺失数据的负样本进行分布假设，通过建立概率模型来避免正负样本偏斜的问题。

在基于权重的单类问题中，Hu[4] 提出了样本实例的置信度概念，他认为用户对一个物品的行为次数越多，则对应喜好的置信程度越高，因此给出如下的置信度公式：

$$c_{ui}=1+\alpha r_{ui}$$

其中，$r_{ui}$ 表示行为频度，$\alpha$ 表示控制系数，因此该信任指数可以直接嵌入协同过滤的计算模型中。例如基本的带权 LFM 的目标函数：

$$\min \sum_{u,i} c_{ui}(p_{ui}-x_u^{\mathrm{T}} y_i)^2+\lambda\left(\sum_u \| x_u \|^2+\sum_i \| y_i \|^2\right)$$

其中，$p_{ui}$ 为二元值，0 表示负样本，1 表示正样本，$x_u$ 和 $y_i$ 分别为用户和物品特征因子。但是该模型仅适用于反馈包含数值型数据的场景，并且仅仅使权重正相关于频数，并不一定能正确反映用户的偏好（尤其是对负样本的表示）。

Pan[11] 提出了 3 种基本的权重方案，如表 2 所列。首先，所有正样本的权值固定为 1，对于负样本则有不同的策略：1）从[0,1]中随机选取一个数值作为权值；2）构造与用户活跃度正相关的数值作为权值，因为如果一个用户拥有较多的正样本，则有较大的可能表示他不喜欢其他的物品，所以他的缺失数据很有可能是负样本；3）构造与物品流行度负相关的数值作为权值，类似地，一个物品拥有更多的正样本意味着该物品很受欢迎，不太可能是一个负样本。

表 2 权重取值策略

| | 正样本 | 负样本 |
|---|---|---|
| 标准 | $w_{i,j}=1$ | $w_{i,j}=\delta$ |
| 偏向用户 | $w_{i,j}=1$ | $w_{i,j}\propto\sum_j R_{i,j}$ |
| 偏向物品 | $w_{i,j}=1$ | $w_{i,j}\propto m-\sum_i R_{i,j}$ |

这 3 种方案看起来比之前的权重构造更加合理，同时它们也适用于只有二元数值的隐式反馈推荐系统。

确定加权矩阵的构成后便可以进行传统的协同过滤的推荐，带权重的最小二乘矩阵分解（Weighted Alternative Least Square，wALS）及其变型是基于 LFM 推荐的常用方法，Pan 在文献[12]中提出了将加权矩阵也进行低秩分解来优化 wALS 执行效率的方法。

另一方面，从概率或采样的角度来处理单类问题的核心思想是通过采样或概率分布假设来产生负样本以对观测数据进行补充，之后再运用标准的优化方法（如 ALS）求解潜在因子。类似地，对于负样本的采样策略，Pan 也提出了 3 种基本策略：1）以均匀的概率抽取负样本；2）根据用户活跃度采样，如果一个用户有较多的正样本，则他未接触的物品很有可能是负样本；3）根据物品流行度采样，如果一个物品很受欢迎，则它对其他未知用户来说是负样本的概率很低。在得到负样本后，将其与原正样本合成得到完整的反馈数据，使用

1) https://movielens.org

2) http://www.flixster.com

3) https://www.kaggle.com/c/msdchallenge

ALS进行还原，如此重复多次并取平均结果以避免单次采样的误差。

此外，Sindhwani 等人[13]从缺失数据中正负样本的分布概率出发，将缺失样本替换为其可能是正样本的概率值 $p_{ij}$，由此构成概率矩阵 $P$，参与 LFM 模型的优化目标函数如下：

$$\min \lambda \| W \|_F^2 + \gamma \| H \|_F^2 + \| C \otimes (X + P - WH) \|_F^2$$

其中，$W$、$H$ 为潜在因子（即潜在用户和物品特征向量构成的矩阵），$C$、$X$ 分别为权值矩阵和原始反馈数据（即初始 0-1 矩阵）。$P$ 同样作为待训练参数，在计算过程中与 $W$、$H$ 交替固定迭代，最终收敛得到模型参数。作者还对该模型进行了优化，在计算中给定选择缺失样本的比例上限参数，所以概率矩阵 $P$ 是低秩的，因此其相比传统的矩阵分解大大降低了计算复杂度。

Paquet 等人[14]则将用户与物品的关系矩阵视作一张二部图，并假设用户感兴趣的物品一定是其考虑过的，因此得到结论：对于用户与其曾考虑过的物品构成图 $H(h \in H)$ 以及用户与其感兴趣的物品构成图 $G(g \in G)$，前者是后者的子图。通过构造用户可能喜欢的物品的概率，建立如下关于用户和物品特征分布的二部图 $H$ 模型：

$$p(g|u,v,b,h=1) = \sigma(u^{\mathrm{T}} v + b)^g [1 - \sigma(u^{\mathrm{T}} v + b)]^{1-g}$$

其中，$u$、$v$ 为用户和物品特征因子，$b$ 为偏差参数，$\sigma$ 为 sigmoid 函数（以下皆同）。然后通过最大似然来求解其参数。其中，类似地，通过给定图 $H$ 的大小上限参数以及图中物品端点的入度分布，从观测图 $G$ 中进行采样得到 $H$。作者同时给出了在大规模计算下优化效率的接口来进行分块计算以提高该算法的性能。

不管是从加权还是采样角度来看，单类协同过滤问题的核心都在于如何对未知的负样本构造合理的假设，例如表 2 中偏向物品的权重取值策略假设流行物品为负样本的概率小，而反过来，正因为流行物品对某用户不是正样本，所以其是负样本的概率较高，这样的解释在一定程度上也是合理的。从这一点上看，不对未知负样本进行过强的假设，而通过概率模型训练获得其可能的分布则是更加合适的方法，甚至为了提高个性化推荐的程度，可以为每个用户（或者每一类用户）假设单独的正负样本分布，以此提高推荐的准确性。

### 3.2 引入辅助信息的推荐方法

如前所述，仅仅依靠隐式反馈进行推荐的效果不佳的根本原因在于缺乏直接反映用户偏好的信息。因此不少学者提出了引入外部信息辅助隐式反馈来进行推荐的方法。将此类方法划分为如下 4 类：结合内容的推荐、结合上下文的推荐、结合社交信息的推荐以及引入跨域知识的推荐。

#### 3.2.1 结合内容的推荐

其基本思想是通过项目或物品相关的属性特征的定义，基于用户（正）反馈对象的特征学习用户的兴趣，形成用户资料，然后据此计算与待测物品的匹配度来辅助隐式反馈的推荐。然而属性特征的定义没有固定的标准，通常根据应用场景的不同而变化；它们往往作为模型参数直接参与预测评分，也称伪评分（Pseudo Rating）的构成。

Li 等人[15]对 OCCF 进行扩展，引入了用户的购买、搜索和浏览记录信息，并且以此构造了基于近邻模型和 LFM 的两个基本方法：对于前者，他采用传统的基于用户的预测评分与在各个特征上用户与待测物品的相似度两个部分的线性模型构成预测评分；对于后者，他采用了基于加权矩阵的方法，将正样本的权值赋为 1，而未知样本的权值负相关于用户与该物品的相似度。在该文中，作者展示了多种不同方案下的对比结果，发现加入用户搜索特征与物品的内容相似度对推荐结果的准确性有一定的提高。

Pacula[16]和 Parra[17]都提出了针对音乐推荐的方法。前者在音乐推荐中将归一化的个人音乐播放频率加入到样本置信度的构成，建立基于该置信度加权的 LFM 模型，通过随机梯度下降（Stochastic Gradient Descent，SGD）进行求解。后者则将单曲播放数、总播放数和上次播放时间同时作为构成伪评分的元素，并采用逻辑回归调整评分以适应（单）分类场景，通过这种方式将隐式反馈转化为显式反馈，同样利用了置信度加权 LFM 的方法进行推荐。相比其他推荐方法，这类方法往往缺少通用性，并且在构造模型之前，特征需要根据能否反映用户偏好来合理选取，而在训练中也要考虑选取的因素对模型的影响程度（如伪评分的权重）等。

#### 3.2.2 结合上下文的推荐

传统的推荐算法主要研究如何发掘用户兴趣与物品的联系，将符合兴趣的物品推荐给用户，忽略了用户所处的上下文。上下文包括用户访问推荐系统的时间、地点、心情[18]等，这些因素对于用户选择有很大的影响，因此将上下文信息作为第三维（甚至是多维）引入，可以提高推荐的准确性。

Rendle 等人[19]利用购物历史数据连续的特点，将其视作与时间相关的马氏链（Markov Chain）模型，提出了一种购物篮推荐的方法。为了降低模型的复杂性和提升计算效率，他们将用户的购物篮历史建模为长度为 1 的齐次马氏链，由此用每个用户购物篮中购买物品的概率构成转移矩阵，转移矩阵可以通过 LFM 模型训练学习获得。作者指出该做法的好处在于：1）矩阵分解模型保持了用户的全局偏好；2）马氏链模型增加了用户偏好的（时间）局部性和个性化。在计算中，由于每个用户的转移矩阵构成一个转移张量（Tensor），作者采用了张量分解（Tucker Decomposition）[20]的方法来训练得到模型参数。Hidasi 等人[21]也提出了结合时令信息（Seasonality）推荐的 LFM 方法，类似地，他们把该信息作为第三维构成三维张量，并提出了基于 ALS 的快速张量分解算法。

除时间信息外，Levandoski 等人[22]认为地点上下文信息可以被引入推荐系统，并在研究中发现用户兴趣和地点相关的两种特征：1）兴趣本地化，即不同地区的用户兴趣存在较大差异；2）活动本地化，即用户往往在固定的区域附近活动。作者以基于物品的协同过滤算法为基础，将物品和用户之间的距离作为惩罚值构成伪评分，推荐距离用户较近且他会感兴趣的物品。Sarwt[23]在此基础上对该算法进行了优化，以提高其性能和可伸缩性。

而 Yao 等人[24]则提出了基于图的上下文感知的推荐模型，他们将上下文信息分为 3 类：用户上下文、物品上下文以及选择上下文。上下文信息中的实体（比如用户的性别、年龄属性，物品的种类属性以及选择上下文中的时间、地点等）成为图中的节点，这些节点被对应地划分为 3 层：用户层、物品层和选择层。作者通过定义同一层内和不同层之间节点的边的权值，采用随机游走（Random Walk）的方法构造了从用户到不同物品的排名分数，该分数即为用户对物品的偏好程度。该模型的优点在于通过图模型建立了上下文感知的一般表

达，同时通过定义图中不同类型边的权值，将异构的多层图转化为了同构图模型。

#### 3.2.3 结合社交信息的推荐

社交信息的引入为传统的推荐系统提供了更为复杂和全面的用户关系，比如好友关系、群组和社区等。一个用户的偏好可以借助社交关系网进行传播，借此可以挖掘该用户可能感兴趣的事物，这为单纯的用户隐式反馈提供了有效的辅助信息。

Yang 等人[25]提出了 FIP(Friendship-Interest Propagation)模型，分别对用户-商品和用户-用户的关系建立了概率模型。作者假设用户和物品分别存在可视特征(如用户注册档案和物品描述信息)和潜在特征(如用户兴趣和物品语义信息)，且后者是关于前者的概率分布，则用户与物品的关系是同时依赖于可视和潜在特征的分布：

$$\phi_i \sim p(\phi_i \mid x_i), \phi_j \sim p(\phi_j \mid x_j)$$

$$y_{i,j} \sim p(y_{i,j} \mid \phi_i, \phi_j x_i, x_j, \Theta)$$

其中，$\phi_i$、$\phi_j$ 分别表示用户和物品的特征，$y_{i,j}$ 表示关于用户-物品关系的分布，$\theta$ 为模型超参数；而在建模朋友关系方面，对于不同的两个用户，他们的关系服从依赖于两者可视和潜在特征的分布：

$$s_{i,i'} \sim p(s_{i,i'} \mid \phi_i, \phi_{i'}, x_i, x_{i'}, \Theta)$$

其中，$i$、$i'$ 表示两个用户，$s_{i,i'}$ 表示关于用户-用户关系的分布。至此，作者通过定义两个模型上各自的损失函数来构成 FIP 模型的后验，在求解最大后验的过程中得到模型参数。为了提高模型的健壮性，还引入了干扰参数来缓解未知样本中的负样本分布偏斜问题。

Delporte 等人[26]在 LFM 模型的基础上引入了相关用户对目标用户的影响作为预测评分的一部分：

$$F_{i,j} = U_i M_j + \sum_{k \in \mathcal{F}_i} \frac{\alpha_{ik}}{|\mathcal{F}_i|} U_k M_j$$

其中，$i$、$j$ 分别表示目标用户和物品，$k$ 表示目标用户的朋友用户，$\frac{\alpha_{ik}}{|\mathcal{F}_i|}$ 则是他们对目标用户归一化后的影响权值，作者假设 $\alpha_{ik}$ 也是模型需要学习的参数之一。之后建立基于置信度加权的 LFM 损失函数，通过类似交替最小二乘的方式不断更新用户物品特征矩阵以及表示用户关系影响的权重矩阵，最终得到这 3 个潜在因子。

结合社交信息的优点在于利用用户偏好的群组特征即同一群组(如朋友圈)的用户在偏好上具有相似性，以及用户偏好的传播，来挖掘潜在的感兴趣物品；而缺点是社交信息依然无法弥补隐式反馈负样本的缺失。

#### 3.2.4 引入跨域知识的推荐

跨域推荐是迁移学习的一种可行方法，它的思想是从一个领域(环境)中学习知识来辅助新领域(环境)的学习任务，在 LFM 模型中表现为在其他领域中学习得到的潜在因子可以辅助目标领域的特征因子的学习。

Singh 等人[27,28]提出的联合矩阵分解框架(Collective Matrix Factorization，CMF)将不同域的信息以矩阵形式引入，其中每个域上的潜在因子可能作用于多个域，因此将这些矩阵赋予不同的权值进行联合分解的同时，就让目标域的潜在因子吸收学习了其他域的知识。同时作者引入了广义线性模式的概念，允许定义不同的损失函数，使得 CMF 成为通用的框架，可以根据矩阵分解方法的不同细化为不同的模型。

而 Hu 等人[29]则借助张量的表示，将领域信息作为除用户和物品之外的第三维，利用张量分解来学习其他领域上的知识，同时结合置信度矩阵作为权值来调节模型以适应隐式反馈的单类问题。

引入跨域知识的优点在于可以将用户的偏好进行迁移以缓解用户行为数据过于稀疏甚至冷启动的问题；而其缺点在于不同领域之间的联系是未知的，即不能确定其他领域的知识对目标领域的影响程度，因而需要通过交叉验证或遗传算法等寻找最优的影响权值分配。

从本质上而言，引入辅助信息是一种信息补充的手法，它虽然能够在一定程度上弥补隐式数据的系数和丰富度不足的问题，但引入信息本身的合理性也是问题之一。例如 Parra 在之前的工作[30]中充分研究了各个因素与人们选择音乐偏好的关系后，才尝试并提出了由隐式向显式转化的回归模型；Fang[31]和 Hu[4]也在各自的工作中对不同数据集(或领域)在矩阵分解的权值设置策略中提出了相应的方法。因此，外部信息本身是不是对推荐目标有促进作用，以及引入信息与原隐式信息整合的方式是否有效，仍然是需要研究的问题。

### 3.3 基于排序的推荐

推荐系统的目的是为用户提供个性化的推荐列表，而推荐列表的排序直接体现了推荐结果的质量——用户更偏好的物品排名越高，则推荐结果的价值越高。传统基于显式反馈的推荐算法往往采用针对预测损失的优化模型，但第 2 节和本节开头就已说明这并不适用于隐式反馈，而针对排序的优化则提供了一个新的视角。因此针对优化排序的推荐方法得到了学者越来越多的关注和研究。

Karatzoglou 等人[32]对基于学习排序(Learning to Rank)的推荐方法进行了详尽的研究，将其分为以下 3 类：逐点排序(point-wise)、逐对排序(pair-wise)和逐表排序(list-wise)。

#### 3.3.1 逐点排序

逐点的方式是将排序问题归约为传统的多分类或回归问题。以 LFM 模型为例，当模型参数——用户和物品的潜在因子学习完毕之后，即可通过模型的评分预测函数得到具体数值，据此可以对推荐物品进行排序。

Ostuni 等人[33]基于隐式的公开互联数据(Linked Open Data，LOD)提出了一种 top-N 的推荐方法。LOD 数据集可以被看作包含语义信息的图模型，关于图中实体(比如一部电影、一个歌手等)的知识通过连接不同节点的边来表示(如一部电影属于喜剧、一个歌手属于流行乐派等)。作者定义了基于路径的语义相关排序 SPrank(Semantic Path-based rank)来表达图中物品关于特定用户的相关度(或该用户对物品的偏好程度)，并以 SPrank 作为每一个(用户，物品)关联向量的一个特征分量，特征数取决于不同路径的数量。作者通过回归树(Gradient Boosting Regression Tree，GBRT)[34]逐点训练得到关于关联向量的排序函数(Ranking Function)，利用该函数即可计算得到满足全序关系的排序分数。

#### 3.3.2 逐对排序

不同于直接获得物品相对于特定用户的相关度，逐对的方式只考虑任意两个物品之间的相对相关度(相较于 A 物品，用户对 B 物品具有更强的偏好)，通过物品对之间的偏序关系将排序问题转化为二分类问题。

基于贝叶斯后验优化的个性化排序(Bayesian Personalized Ranking，BPR)推荐是 Rendle 等人[35]提出的基于隐式

反馈的排序推荐框架。作者首先定义了偏序关系：如果用户 $u$ 购买过物品 $I_i$ 但未购买过物品 $I_j$，则有 $I_i >_u I_j$，即用户更偏向于物品 $I_i$；否则两者的偏序关系无法确定。假设不同用户的偏好相互独立且同一用户对不同物品的偏好相互独立，以此可以建立如下贝叶斯模型：

$$\prod_{u,i,j} p(\theta | I_i >_u I_j) \propto \prod_{u,i,j} p(I_i >_u I_j | \theta) p(\theta)$$

其中，$\theta$ 为所求模型参数，而用户偏好物品 $I_i$ 的概率表达可以细化为 $p(I_i >_u I_j | \theta) = \sigma(\hat{x}_{ui} - \hat{x}_{uj})$，使用 sigmoid 函数对 $(\hat{x}_{ui} - \hat{x}_{uj})$ 平滑处理以适应二分类，$\hat{x}_{ui}$ 为关于模型参数的排名函数（比如伪评分的构造函数）。对于不同的评分模型，$\hat{x}_{ui}$ 的形式也不尽相同（比如可以是近邻模型或者 LFM 模型等），因此 BPR 作为泛用的推荐框架，可以演化为不同的子模型。作者通过对当前 AUC（详见第 4 节）采用逐对的优化得到评分模型参数 $\theta$，即排序函数。

Fang[36] 也提出了相似的想法，首先定义了当给定了用户 $u$ 参照的物品 $c$ 时，推荐物品 $i$ 比 $j$ 更相关的表达 $l_{i,j}^{u,c} = 1$，否则 $l_{i,j}^{u,c} = -1$。作者借鉴了概率潜语义分析（Probabilistic Latent Semantic Indexing，pLSI）[37] 的思想，引入隐式参数 $z$ 来表示用户的潜在组特征（group feature），由此产生了需要最大化似然的模型：

$$P(l_{i,j}^{u,c} | c, i, j, z) = \sigma\left(\sum_{v=1}^{V} \lambda_{zv} (f_v(c,i) - f_v(c,j))\right)$$

其中，$f_v$ 为在特征 $v$ 上的物品相似度，$\lambda_{zv}$ 为其在对应组上的权重。文中作者设计了一系列相似度衡量指标，且通过 EM（Expectation-Maximization）算法来估计潜在参数 $\lambda_{zv}$。最终通过后验公式得到预测评分：

$$r_i^{u,c} = \sum_{z=1}^{K} P(z|u) \sigma\left(\sum_{v=1}^{V} \lambda_{zv} f_v(c,i)\right)$$

逐对模型的优点在于考虑了不同用户物品相关度之间的相对关系，更符合用户在选择物品时的比较心理，但该方法也有缺点：1）对不同级别的区分统一对待，没有考虑到排序靠前的结果对整个排序的重要性；2）相关物品集的大小不同导致模型偏置，例如某用户的相关物品很多，在训练排名函数时占据主导作用，使得对相关物品少的用户的推荐效果变差。

### 3.3.3 逐表排序

相比于前两种方法，逐表方法直接对推荐的结果列表进行优化，主要有两种：1）直接针对排序指标进行优化，比如常用的 MAP、NDCG（详见第 4 节）；2）优化损失函数的构造。

TFMAP[38] 是关于 MAP 优化的一个逐列表优化模型。作者在隐式反馈的基础上引入了上下文信息，将数据集建模为张量，通过 CP 分解（CANDECOMP/PARAFAC）[20] 建模用户、物品和上下文的潜在因子 $u$、$v$、$c$，然后以最大化 MAP 为目标学习潜在因子。其中 MAP 定义如下：

$$AP_{mk} = \frac{1}{\sum_{i=1}^{N} y_{mik}} \sum_{i=1}^{N} \frac{y_{mik}}{r_{mik}} \sum_{i=1}^{N} y_{mjk} \Pi(r_{mjk} \leqslant r_{mik})$$

$$MAP = \frac{1}{MK} \sum_{m=1}^{M} \sum_{k=1}^{K} AP_{mk}$$

其中，$m$、$k$、$i$ 分别表示用户、上下文和物品，$y_{mik}$ 表示对应隐式反馈的观测值，$r_{mik}$ 表示 CP 模型得到的预测排名分数，$\Pi$ 为二值指示函数（参数真时为 1，非真时为 0）。MAP 指标越高，表示列表排序越符合用户预期，推荐效果越好，且排名越靠前的物品 AP 对 MAP 影响越大。而 MAP 指标是非连续函数，难以进行传统梯度的优化，作者引入 sigmoid 函数来平滑该指标：

$$\Pi(r_{mjk} \leqslant r_{mik}) \approx \sigma(f_{mjk} - f_{mik}) = \sigma(\langle U_m, V_j - V_i, C_k \rangle)$$

$$\frac{1}{r_{mik}} \approx \sigma(f_{mik}) = \sigma(\langle U_m, V_i, C_k \rangle)$$

其中 $f_{mik} = \langle U_m, V_i, C_k \rangle = r_{mik}$ 为 CP 模型的排名函数。至此可以通过平滑后的 MAP 指标构造损失函数，采用传统的优化方法来训练模型的潜在因子。在文中作者还设计了采样策略来减少迭代优化的计算量，使得算法具有良好的伸缩性。

逐表方法的优点是目标明确，直接对整体排序优化克服了逐对方法的缺陷；缺点在于比前两种方法需要更多的计算量。

### 3.3.4 各推荐技术比较

将上述各类推荐技术的优缺点进行了比较，如表 3 所列。

表 3 推荐技术优缺点的比较

| 类目 | 推荐技术 | 优点 | 缺点 |
|---|---|---|---|
| 单类协同过滤 | 基于权重[4,11,39,40] | 不需要领域知识和特征选取；加权模型简单，容易解释 | 加权假设过强，可能导致模型偏置；考虑所有样本的计算复杂度较高 |
| | 基于采样[13,14] | 训练概率模型使算法不易过拟合；负样本采样的方式降低了计算复杂度 | 采样策略问题；多次采样训练结果的集成（ensemble）问题 |
| 辅助信息引入 | 结合内容[15-17,41] | 评分函数直观，容易解释 | 特征选取没有统一标准；复杂属性难以处理 |
| | 结合上下文[19,21-24] | 对用户偏好理解更符合实际情况 | 关于用户、物品或两者交互的上下文数据不易获得 |
| | 结合社交信息[25,26,42] | 用户偏好传播提供新的用户兴趣挖掘方式；社交信息容易获得 | 缺乏负样本数据的弥补；难以对社交圈外的事物进行推荐 |
| | 引入跨域知识[27-29,43] | 解决了当前领域的数据稀疏问题 | 跨域信息相关度问题；多领域知识的影响不同 |
| 学习排序 | 逐点排序[33,44] | 样本复杂度低，为 O(n) | 忽略了样本对之间的相关性 |
| | 逐对排序[35,36] | 用户偏好关于样本对的偏序关系较符合实际情况 | 样本复杂度较高为，$O(n^2)$；忽略了高排名样本间的差异；相关样本集不均衡导致的模型偏置 |
| | 逐表排序[38,45,46] | 序列优化更接近用户的实际期望序列 | 样本复杂度很高，为 O(!n) |

## 4 评价指标

如第 2 节讨论过的，因为隐式反馈缺少负样本的信息，所以传统的基于准确率的评价标准并不适用于此类推荐系统。因此选取一个合理的标准对隐式反馈推荐系统的评估是十分重要的。我们将在此类推荐系统中广泛使用的评估标准[47]筛选整理如下，并加以详细的说明和分析。

### 4.1 Precision/Recall

准确率（Precision）和召回率（Recall）是推荐系统中最常用的分类指标。准确率表示用户对系统推荐物品感兴趣的概率。研究者通常给定一长度为 $N$（Top$N$）的推荐列表，根据推荐算法的预测对待推荐物品排序，将前 $N$ 位推荐给用户，

并以此计算推荐的准确率，该测度通常称为 *Precision@N*。

而对一个用户未接触过的物品，其结果包含 4 种情况：系统推荐且用户喜欢(True Positive，TP)，系统推荐但用户不喜欢(False Positive，FP)，系统未推荐但用户喜欢(False Negative，FN)，以及系统未推荐且用户不喜欢(True Negative，TN)。表 4 直观地总结了这 4 种情况，其中 $N_{TP}$ 分别代表了各自情况的物品数目，不难发现 $N=N_{TP}+N_{FP}$，$B=N_{TP}+N_{FN}$，后者表示用户真实喜欢的物品数。

表 4　待预测商品的 4 种情况

| 是否喜欢 | 系统推荐 | 系统未推荐 |
|---|---|---|
| 喜欢 | $N_{TP}$ | $N_{FN}$ |
| 不喜欢 | $N_{FP}$ | $N_{TN}$ |

有了上述 4 类物品的表示，不难得到对于单个用户的 *Presion@N* 的定义：$P_u(N)=\frac{N_{TP}}{N_{TP}+N_{FP}}$，对所有用户的准确率求平均作为整个系统的推荐准确率：$P(N)=\frac{1}{M}\sum_u P_u(N)$，其中 $M$ 表示接受评估的人数。在离线测试的情况下，准确率很容易受到数据稀疏性的影响，比如对一个很少有购买或点击的低活跃用户的推荐准确率往往是低的，但并不意味着推荐效果不佳，可能存在用户未接触但确实喜欢的物品；而对于在线测试，$N_{TP}$ 的数值可以通过实时反馈统计得到，这部分数据是用户的真实反馈，往往更能够准确地体现推荐效果。

召回率表示一个用户喜欢的物品被系统推荐的概率，该指标同样可以通过表 4 中的数值指标来定义：$R_u(N)=\frac{N_{TP}}{N_{TP}+N_{FN}}$。而实际上由于系统无法知道未被推荐的物品是否是用户喜欢的，因此该指标很难用于在线评估。同样，系统召回率可以表示为所有用户的平均值：$R(N)=\frac{1}{M}\sum_u R_u(N)$。

### 4.2　F-measure

在实际中，准确率和召回率往往是负相关的，且二者都依赖于推荐列表的大小，一般来说，推荐列表增大，准确率会下降而召回率会上升。为了综合考虑二者对系统推荐结果的有效程度的贡献，有学者定义了 $F_1$ 测度：$F_1=\frac{2\cdot Precision\cdot Recall}{Precision+Recall}$，即准确率与召回率的调和均值，其取值范围在 0 到 1 之间，越接近 1 则效果越好。$F_1$ 测度的好处在于统一了两个指标，使得不同算法之间可以进行比较。而在某些场景下，不同算法对准确率和召回率有不同的侧重，因此提出了更加一般化的 F 测度定义：$F_\beta=\frac{(1+\beta^2)\cdot Precision\cdot Recall}{\beta^2\cdot Precision+Recall}$，其中 $\beta$ 表示召回率相对于准确率的权重比值，当 $\beta>1$ 时，给予召回率更高的权重，反之为准确率。

### 4.3　AUC

在推荐与否的阈值不确定的情况下，上述指标无法适用，此时往往采用 AUC 指标来衡量推荐结果的准确性，其好处在于不依赖推荐列表的长度，且不受喜好阈值的影响。AUC 指标表示 ROC[48](Receiver Operator Curve)曲线下的面积(Area Under Curve)，它衡量一个推荐系统能够在多大程度上将用户喜欢与不喜欢的物品区分出来。通常 ROC 曲线以 FPR(False Positive Rate，假命中率，即$\frac{N_{FP}}{N_{TN}+N_{FP}}$)为横轴，以 TPR(True Positive Rate，真命中率，即$\frac{N_{TP}}{N_{TP}+N_{FN}}$)为纵轴，依据同一模型中的不同阈值(即高于阈值的认为是正样本，反之为负样本)将($FPR$，$TPR$)坐标画在 ROC 空间里就形成了该模型的 ROC 曲线，该曲线下的面积就是 AUC 指标。由于绘制 ROC 曲线较繁琐，通常使用近似方法计算 AUC：每次从推荐列表中选择一个物品与另一个随机的不喜欢物品进行比较，若前者的预测评分(或正样本概率等，依据模型而定)大于后者，则加 1 分；若相等，则加 0.5 分。如此独立操作 $n$ 次，其中 $n'$ 次加 1，$n''$ 次加 0.5 分，则可以得到近似表达：$AUC=\frac{n'+0.5n''}{n}$。显然，随机预测的 AUC 为 0.5，因此 AUC 大于 0.5 的程度衡量了算法在多大程度上比随机推荐精确。但是 AUC 指标没有考虑具体排序的影响，因此在 ROC 曲线面积相同的情况下难以比较算法的好坏。

### 4.4　MAP

对于推荐排序要求严格的系统，需要用排序准确度指标来度量算法得到的有序列表与用户对物品排序的统一程度。Pan 在其工作[11]中使用了平均准确率均值(mean average precision)[49]，它计算了对所有用户推荐列表的平均准确率(average precision，AP)的均值，其中 AP 定义为：$AP_u=\frac{\sum_i r_{ui}\times prec(i)}{\#\ of\ preferred\ items}$，其中 $prec(i)$ 为物品相对特定用户的百分位排序，该数值越小，表明推荐物品的排序越符合用户的心理预期，推荐效果越好。Hu[4]和 Fang[31]也采用了相似的评价指标百分位排序均值(Mean Percentage Ranking，MPR)[50]。类似地，MAP 数值小于 50%意味着算法优于随机推荐。

### 4.5　HLU

衰期效用指标(Half-Life Utility)[51]是在用户浏览物品的概率与该物品在推荐列表中的排序值呈指数递减的假设下提出的，它度量的是推荐系统对用户的实用性程度，即用户真实评分与系统默认的差别。用户 $u$ 的效用期望定义为：$R_u=\sum_j\frac{\max(r_{uj}-d,0)}{2^{(j-1)/(h-1)}}$，其中 $r_{uj}$ 为实际评分，$d$ 为系统默认评分，而在隐式反馈系统中分子部分可以表示为 $\delta(j)$，其值为 1 意味排位在第 $j$ 位是正样本，为 0 则是负样本；$h$ 为半衰期，即用户有 50%的概率会浏览推荐列表的位置。当用户喜欢的物品都被放在列表的前端时，该指标达到最大值。基于此，整个系统的 HLU 指标定义为：$R=100\frac{\sum_u R_u}{\sum_u R_u^{\max}}$。半衰期参数的选择没有统一标准，不同学者可能会选择不同参数(如根据数据集的特点而定)，因此难以放在一起对比；此外，浏览概率的指数衰减假设也并不总是适用。

### 4.6　DCG

折扣累计利润(Discounted Cumulative Gain，DCG)[52]的主要思想是用户喜欢的物品被排列在前面比排在后面会更大程度上提高用户体验，在推荐中排 $n$ 位的 DCG 定义为：$DCG_n=\sum_{j=1}^{n}\frac{2^{r(j)}-1}{\log(1+j)}$，$r(j)$ 为相关度分级，分子可以理解为推荐列表中物品的 gain，该值随用户心理预期的排位递减(这里为指数

递减)，分母可以理解为排位不同物品的 gain 应当赋予不同的折算因子，排位越后折算越多(即最终 gain 越小)。同样，该数值越大则意味着排位推荐的效果越好。在此基础上，学者也定义了 NDCG(Normalized DCG)[53]，即对 DCG 进行归一化：$nDCG_r = \frac{DCG_r}{DCG_r^{max}}$，使得不同分级上的 DCG 值易于比较。类似地，推荐物品的 gain 取决于对用户喜好的分级，而不同的研究工作中对分级也会有不同的假设，导致相互间的算法结果无法直接比较。

为了给出更直观的比较，对上述几种评价指标进行总结，如表 5 所列。

表 5　评价指标比较

| 评价指标 | 名称 | 符号 | 依赖推荐列表长度 | 说明 |
|---|---|---|---|---|
| 分类准确度 | 准确率/召回率 | Precsion/Recall | 是 | 只适用于明确二分喜好(阈值)的推荐系统 |
| | F 测度 | $F_1(F_\beta)$ | 是 | |
| | ROC 曲线面积 | AUC | 否 | 适用于缺少负样本的系统；没有考虑排序质量 |
| 排序准确度 | 平均准确率均值 | MAP | 是 | 适用于对推荐排序要求严格的系统 |
| | 百分位排序均值 | MPR | 是 | |
| 基于排序加权的指标 | 半衰期效用指标 | HLU | 是 | 衰减加权模型更符合实际情况 |
| | 折扣累计利润(及其归一化) | DCG(NDCG) | 是 | 考虑了排序中不同的分级 |

## 5　研究展望

虽然基于隐式反馈的推荐方法在近几年得到了不错的发展，但现有的方法仍然存在一些不完善的地方。对这一课题的研究应有以下发展热点与方向：

1)深入算法背后的细节问题。虽然推荐算法被一个接一个地提出，但算法中的细节和假设往往欠缺坚实合理的说明，这也是推荐算法的共同问题——可解释性问题。比如对于在单类协同过滤中提到的几种加权方法，其本身并没有足够的理论和实际证据支持假设，因此难以判定其是否真的有效。好的解释理论不仅可以提高用户的推荐体验，更有助于算法本身的设计和完善。

2)对目前优化算法效率的改进。如第 3 节中介绍的诸多推荐算法，尤其是基于排序和张量分解的最优化算法存在计算复杂度大的缺陷。面对目前现实中海量的数据，提高这些方法的运算性能和实时性也是主要目标之一。许多学者致力于研究针对大数据的推荐算法，以及借助并行计算和分布式框架的力量来提高算法效率。

3)融合常识知识。在隐式反馈的推荐技术中需要各种假设来完成模型的建立，包括对负样本的分布以及伪评分或排序函数的前提假设等，这些假设需要融入常识性知识(Common Knowledge)来建立，目的在于：①简化模型的同时使算法更加有效；②使模型更符合实际情况，容易对推荐结果进行解释。

4)结合辅助信息。在现实的信息系统中，反馈信息中除了本身与用户的关联信息，往往还包含了诸如物品、用户属性等元信息，以及事件发生的上下文信息，如时间、地点等。前者可以在通过聚类等方法筛选数据来对用户、物品本身建模时起到重要作用；而后者更为推荐的多维度化提供了数据支持，可以预见未来的推荐系统应该具备根据用户不同的使用场景产生动态推荐结果的能力。

5)从推荐系统出发，结合多领域的研究成果。推荐系统是一个开放的研究领域，其本身也是结合了机器学习、数据挖掘的方法，以及复杂网络、异构信息网络、社交网络等不同领域的信息模型的具有实际应用价值的产物；不再局限于推荐行为本身，而考虑多类研究的方法，吸收多类数据的特性，例如借鉴深度学习的方法、小世界理论以及研究用户兴趣的深度层次结构等，才能不断完善推荐系统的基础，在未来迎来新的发展。

**结束语**　本文对现有的基于隐式反馈的推荐方法进行了系统的研究和分析，总结了隐式反馈推荐系统的分类以及各自的特点；特别是对单类问题和排序问题的研究，强调了从隐式反馈本身进行推荐的可能性和有效性。之后介绍了适合隐式反馈推荐系统的评价指标，在详细阐述指标含义的同时对各自的优缺点和使用场景进行了分析，特别强调了在推荐列表中商品排序对推荐评价的影响。文中总结了已有的一些研究成果，同时针对其不足之处提出了可能的解决方向，进一步的问题及研究需要研究者们发现和努力。

## 参 考 文 献


[1] Pitkow J, Schtze H, Cass T, et al. Personalized search[J]. Commun. ACM, 2002, 45(9): 50-55

[2] Teevan J, Dumais S T, Horvitz E. Personalizing search via automated analysis of interests and activities [C]// Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2005: 449-456

[3] Peška L, Vojtáš P. Estimating importance of implicit factors in e-commerce recommender systems [C]// Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics. ACM, 2012: 62

[4] Hu Y, Koren Y, Volinsky C. Collaborative filtering for implicit feedback datasets [C]// Eighth IEEE International Conference on Data Mining, 2008(ICDM'08). IEEE, 2008: 263-272

[5] Chao D L, Balthrop J, Forrest S. Adaptive radio: achieving consensus using negative preferences [C]// Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting group Work. ACM, 2005: 120-123

[6] Amatriain X, Pujol J M, Oliver N. I like it…i like it not: Evaluating user ratings noise in recommender systems [M]// User Modeling, Adaptation, and Personalization. Springer, 2009: 247-258

[7] Jawaheer G, Szomszor M, Kostkova P. Comparison of implicit and explicit feedback from an online music recommendation service [C]// Proceedings of the 1st International Workshop on

Information Heterogeneity and Fusion in Recommender Systems. ACM,2010:47-51

[8] Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions[J]. IEEE Transactions on Knowledge and Data Engineering,2005,17(6):734-749

[9] Schafer J B,Konstan J A,Riedl J. E-commerce recommendation applications [M]// Applications of Data Mining to Electronic Commerce. Springer,2001:115-153

[10] Heckerman D, Geiger D, Chickering D M. Learning Bayesian networks: The combination of knowledge and statistical data [J]. Machine Learning,1995,20(3):197-243

[11] Pan R, Zhou Y, Cao B, et al. One-class collaborative filtering [C]// Eighth IEEE International Conference on Data Mining, 2008 (ICDM'08). IEEE,2008:502-511

[12] Pan R, Scholz M. Mind the gaps: weighting the unknown in large-scale one-class collaborative filtering [C]// Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM,2009:667-676

[13] Sindhwani V,Bucak S S,Hu J,et al. One-class matrix completion with low-density factorizations [C]// 2010 IEEE 10th International Conference on Data Mining (ICDM). IEEE, 2010: 1055-1060

[14] Paquet U,Koenigstein N. One-class collaborative filtering with random graphs [C] // Proceedings of the 22nd International Conference on World Wide Web. International World Wide Web Conferences Steering Committee,2013:999-1008

[15] Li Y,Hu J,Zhai C,et al. Improving one-class collaborative filtering by incorporating rich user information [C]// Proceedings of the 19th ACM International Conference on Information and Knowledge Management. ACM,2010:959-968

[16] Pacula M. A Matrix Factorization Algorithm for Music Recommendation using Implicit User Feedback[OL]. http://mpacula.com/publications/lastfm. pdf

[17] Parra D, Karatzoglou A, Amatriain X, et al. Implicit feedback recommendation via implicit-to-explicit ordinal logistic regression mapping[C]// Proceedings of the CARS-2011. 2011

[18] Adomavicius G, Tuzhilin A. Context-aware recommender systems [M]// Recommender Systems Handbook. Springer,2011: 217-253

[19] Rendle S,Freudenthaler C,Schmidt-Thieme L. Factorizing personalized markov chains for next-basket recommendation [C]// Proceedings of the 19th International Conference on World Wide Web. ACM,2010:811-820

[20] Kolda T G,Bader B W. Tensor decompositions and applications [J]. SIAM Review,2009,51(3):455-500

[21] Hidasi B,Tikk D. Fast ALS-based tensor factorization for context-aware recommendation from implicit feedback [M]// Machine Learning and Knowledge Discovery in Databases. Springer,2012:67-82

[22] Levandoski J J, Sarwat M, Eldawy A, et al. Lars: A location-aware recommender system [C]// 2012 IEEE 28th International Conference Data Engineering (ICDE). IEEE,2012:450-461

[23] Sarwt M,Levandoski J J,Eldawy A,et al. LARS* :An Efficient and Scalable Location-Aware Recommender System[J]. IEEE Transactions on Knowledge and Data Engineering,2013,26(6): 1384-1399

[24] Yao W L,He J,Huang G Y,et al. A Graph-based model for context-aware recommendation using implicit feedback data[J]. World Wide Web,2015,18(5):1351-1371

[25] Yang S-H,Long B,Smola A,et al. Like like alike: joint friendship and interest propagation in social networks [C]// Proceedings of the 20th International Conference on World Wide Web. ACM,2011:537-546

[26] Delporte J,Karatzoglou A,Matuszczyk T,et al. Socially enabled preference learning from implicit feedback data [M]// Machine Learning and Knowledge Discovery in Databases. Springer, 2013:145-160

[27] Singh A P,Gordon G J. Relational learning via collective matrix factorization [C]// Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM,2008:650-658

[28] Singh A P,Gordon G J. A unified view of matrix factorization models [M]// Machine Learning and Knowledge Discovery in Databases. Springer,2008:358-373

[29] Hu L,Cao J,Xu G,et al. Personalized recommendation via cross-domain triadic factorization [C]// Proceedings of the 22nd International Conference on World Wide Web. International World Wide Web Conferences Steering Committee,2013:595-606

[30] Parra D,Amatriain X. Walk the Talk [M]// User Modeling, Adaption and Personalization. Springer,2011:255-268

[31] Fang Y,Si L. Matrix co-factorization for recommendation with rich side information and implicit feedback [C]// Proceedings of the 2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems. ACM,2011:65-69

[32] Karatzoglou A,Baltrunas L,Shi Y. Learning to rank for recommender systems [C]// Proceedings of the 7th ACM Conference on Recommender Systems. ACM,2013:493-494

[33] Ostuni V C,Noia T D,Sciascio E D,et al. Top-N recommendations from implicit feedback leveraging linked open data [C]// Proceedings of the 7th ACM Conference on Recommender systems. ACM,2013:85-92

[34] Elith J,Leathwick J R,Hastie T. A working guide to boosted regression trees[J]. Journal of Animal Ecology,2008,77(4):802-813

[35] Rendle S,Freudenthaler C,Gantner Z,et al. BPR: Bayesian personalized ranking from implicit feedback [C] // Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence. AUAI Press,2009:452-461

[36] Fang Y,Si L. A latent pairwise preference learning approach for recommendation from implicit feedback [C]// Proceedings of the 21st ACM International Conference on Information and Knowledge Management. ACM,2012:2567-2570

[37] Hofmann T. Probabilistic latent semantic indexing [C] // Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM,1999:50-57

(下转第 49 页)

tic Relatedness Measure Based on Wikipedia Structure[J]. Chinese Journal of Computers, 2012, 35(11): 2361-2370(in Chinese)
孙琛琛,申德荣,单菁,等. WSR:一种基于维基百科结构信息的语义关联度计算算法[J]. 计算机学报,2012,35(11):2361-2370
[12] Liu Xiao-liang. Research on Computation of Lexical Semantic Relatedness Based on Wikipedia Semantic Graph[J]. Journal of the China Society for Scientific and Technical Information, 2014,33(11):1124-1132(in Chinese)
刘晓亮. 基于维基语义图的词语语义相关度计算研究[J]. 情报学报,2014,33(11):1124-1132
[13] Bellomi F, Bonato R. Network analysis for Wikipedia[C]//Proceedings of Wikimania. 2005
[14] Lizorkin D, Medelyan O, Grineva M. Analysis of community structure in wikipedia[C]//Proceedings of the 18th International Conference on World Wide Web. ACM, 2009:1221-1222
[15] Li Yun. Research about semantic knowledge mining based on the Chinese Wikipedia[D]. Beijing: Beijing University of Posts and Telecommunications, 2009
[16] Kleinberg J M. Authoritative sources in a hyperlinked environment[J]. Journal of the ACM (JACM), 1999, 46(5):604-632
[17] Jarmasz M. Roget's thesaurus as a lexical resource for natural language processing[J]. arXiv:1204.0140, 2012
[18] Landauer T K, Foltz P W, Laham D. An introduction to latent semantic analysis[J]. Discourse Processes, 1998, 25(2/3): 259-284
[19] Levenshtein V I. Binary codes capable of correcting deletions, insertions, and reversals[C]//Soviet Physics Doklady, 1966, 10(10):707-710
[20] 维基百科数据集[OL]. http://dumps.wikimedia.org/
[21] DataMachine[OL]. http://search.maven.org/#search|ga|1|tudarm stadt. ukp

(上接第 15 页)

[38] Shi Y, Karatzoglou A, Baltrunas L, et al. TFMAP: optimizing MAP for top-n context-aware recommendation [C]//Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2012: 155-164
[39] Wang Peng, Jing Li-ping. Improved One-class Collaborative Filtering for Recommendtion System[J]. Journal of Frontiers of Computer Science and Technology, 2014, 8(10): 1231-1238(in Chinese)
王鹏,景丽萍. 改进的单类协同过滤推荐方法[J]. 计算机科学与探索,2014,8(10):1231-1238
[40] Li Gai, Li Lei. One-class collaborative filtering based on matrix factorization[J]. Application Research of Computers, 2012, 29(5):1662-1665(in Chinese)
李改,李磊. 基于矩阵分解的单类协同过滤推荐算法[J]. 计算机应用研究,2012,29(5):1662-1665
[41] Núñez-Valdéz E R, Lovelle J M C, Martínez O S, et al. Implicit feedback techniques on recommender systems applied to electronic books[J]. Computers in Human Behavior, 2012, 28(4): 1186-1193
[42] Li Gai, Li Lei. One-class collaborative filtering algorithm based on social network[J]. Journal of Hubei University(Natural Science), 2014, 36(4):333-338(in Chinese)
李改,李磊. 基于社交网络的单类协同过滤算法[J]. 湖北大学学报(自然科学版),2014,36(4):333-338
[43] Luo Sheng-mei, Lin Yun-zhen, Ye Xiao-long, et al. One Class Collaborative Filtering Algorithm Based on Transfer Learning [J]. Hans Journal of Data Mining, 2013, 3(1): 12-17(in Chinese)
罗圣美,林运祯,叶小伟,等. 基于迁移学习的单类协同过滤算法[J]. 汉斯出版社-数据挖掘,2013,3(1):12-17
[44] Rendle S. Factorization machines with libFM[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2012, 3(3):57
[45] Shi Y, Karatzoglou A, Baltrunas L, et al. Climf: learning to maximize reciprocal rank with collaborative less-is-more filtering [C] // Proceedings of the Sixth ACM Conference on Recommender Systems. ACM, 2012: 139-146
[46] Shi Y, Karatzoglou A, Baltrunas L, et al. xCLiMF: optimizing expected reciprocal rank for data with multiple levels of relevance [C]//Proceedings of the 7th ACM Conference on Recommender Systems. ACM, 2013: 431-434
[47] Zhu Yu-xiao, Lv Lin-yuan. Evaluation metrics for recommender systems[J]. Journal of Electronic Science and Technology of China, 2012, 41(2): 163-175
[48] Herschtal A, Raskutti B. Optimising area under the ROC curve using gradient descent [C]//Proceedings of the Twenty-first International Conference on Machine Learning. ACM, 2004: 49
[49] Manning C D, Raghavan P, Schütze H. Introduction to information retrieval [M]. Cambridge: Cambridge University Press, 2008
[50] Craswell N. Mean Reciprocal Rank [M]//Encyclopedia of Database Systems. Springer, 2009: 1703-1703
[51] Breese J S, Heckerman D, Kadie C. Empirical analysis of predictive algorithms for collaborative filtering [C]//Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence. Morgan Kaufmann Publishers Inc., 1998: 43-52
[52] Järvelin K, Kekäläinen J. IR evaluation methods for retrieving highly relevant documents [C]//Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2000: 41-48
[53] Järvelin K, Kekäläinen J. Cumulated gain-based evaluation of IR techniques [J]. ACM Transactions on Information Systems (TOIS), 2002, 20(4): 422-446