

推荐系统冷启动问题解决策略研究

乔 雨 李玲娟

(南京邮电大学 计算机学院 江苏 南京 210003)

摘 要: 推荐系统利用机器学习技术进行信息过滤,快速准确地定位用户需要的信息,并且能够预测用户对目标项目的喜好程度。由于新用户与新项目的存在,传统的推荐系统在缺少数据信息的情况下面临着冷启动问题的挑战,导致系统无法为用户产生准确的推荐。分析冷启动产生的原因,阐述解决冷启动问题的意义,从是否考虑冷启动类型等方面对目前推荐系统冷启动问题的研究成果进行分类总结,并尝试给出冷启动问题未来的研究重点与难点。目前较为普遍的处理方式是将多种数据源与多种推荐方法进行混合使用,从而提高系统推荐的准确度与效率,但是仍然存在着如在收集用户各类信息的同时如何保护个人隐私、如何建立推荐系统的效用评价等难点问题。

关键词: 推荐系统; 协同过滤; 用户冷启动; 项目冷启动; 解决策略

中图分类号: TP311

文献标识码: A

文章编号: 1673-629X(2018)02-0083-05

doi: 10.3969/j.issn.1673-629X.2018.02.019

Research on Solution of Solving Cold Start Problem in Recommender Systems

QIAO Yu LI Ling-juan

(School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: Recommendation systems apply machine learning techniques to filter and locate information accurately, and can predict whether a user would like a given resource. Traditional collaborative filtering systems have to deal with the cold-start problems as new users and items are always present, which fail to produce the accurate recommendation for users. In this paper we first illustrate the causes and the significances of solving the cold-start problems according to the current achievements in research, and then summarize the existing algorithms and compare the performance of them. Finally, we try to give the difficulties and future directions of recommender system. It was found that the most popular way is to have mixed data sources and algorithms to improve the accuracy and efficiency of recommender system at present, but there still have been some difficulty like how to protect personal privacy during getting users' information or how to establish the performance evaluation of recommendation systems.

Key words: recommender system; collaborative filtering; new-user cold-start; new-item cold-start; solving methods

0 引 言

互联网的快速发展使人们能够快速、便捷、低成本地制造和接收信息。虽然网络带来了极大的便利,但随之而来的信息超载问题使得如何从这些海量的数据中快速、准确地获取有用信息成为研究热点^[1]。目前解决信息超载的方法主要分为两种:一种是利用用户主动检索的关键词进行筛选,过滤掉多余的信息,这种方法即为搜索引擎,它在帮助用户获取所需的网络信息方面发挥着极其重要的作用,如 Google、Baidu 等。另一种则是从系统本身出发,根据用户已有的信息分

析出用户的需求,采用适当算法进行计算处理,并将用户可能感兴趣的信息推荐给用户,此方法即为个性化推荐系统。一个好的推荐系统不仅能为用户提供个性化服务,还能与用户建立密切关系,让用户对推荐产生依赖^[2],这也成为解决信息超载一种非常有潜力的方式。这两种方式在底层技术设计上有大量类似的地方,但是在相应的用户需求和应用场景上,推荐系统离用户更进一步。

推荐系统的概念提出后,在电子商务、电影音乐推荐等方面得到了广泛的应用,学术界也成立了关于推

收稿日期: 2017-02-17

修回日期: 2017-06-21

网络出版时间: 2017-10-19

基金项目: 国家自然科学基金(61302158, 61571238)

作者简介: 乔 雨(1993-),女,硕士,CCF 会员(E200052171G),研究方向为数据挖掘与推荐系统;李玲娟,教授,研究方向为数据挖掘、信息安全、分布式计算。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20171019.1559.002.html>

荐系统的主题会议,其应用与研究的核心都围绕推荐算法在应用过程中的性能优化。根据推荐算法的不同,目前推荐系统分为基于规则(rule-based)的推荐系统、基于内容(content-based)的推荐系统、协同过滤(collaborative filtering)推荐系统^[3]。其中协同过滤推荐系统是应用最广泛、最成功的推荐系统^[4],然而该算法需要使用大量的用户行为信息作为计算依据,因此对于新用户、新物品以及新系统该类算法存在明显的不足之处,因此需要研究出针对冷启动状态下的算法。文中对该领域的研究进展和发展趋势进行了总结与分析,为后续研究提供参考。

1 冷启动问题描述

协同过滤推荐系统的流程如图 1 所示。根据用户的基本信息与历史行为信息,分析出用户的行为偏好;根据物品的基本信息与被评价的信息提取物品的特征,再将用户的偏好信息与物品的特征信息进行适配,选取处理结果匹配度最高的 Top-N 个物品,并将其推荐给用户。

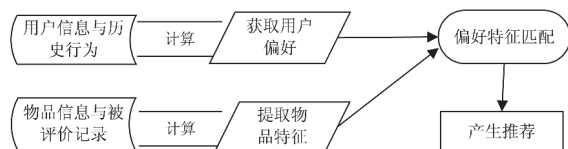


图 1 推荐系统模型流程

1.1 产生的原因

协同过滤推荐算法在缺少数据或者数据稀疏的情况下推荐效果急剧下降,这是由于算法无法利用足够的信息进行用户行为分析,因此无法产生准确的推荐效果。表 1 为用户-项目评分事例(评分范围为 1~5 分),“-”表示没有评分。

表 1 用户对项目的评分事例

	Item ₁	Item ₂	Item ₃	Item ₄	Item ₅
User ₁	3	5	-	5	4
User ₂	-	-	-	-	-
User ₃	2	3	-	-	-
User ₄	-	-	-	4	3
User ₅	4	5	-	5	5
User ₆	3	4	-	5	-

该表中描述了 6 个用户(User₁, User₂, User₃, User₄, User₅, User₆)对 4 个物品(Item₁, Item₂, Item₃, Item₄)的评分信息。系统根据用户已评分信息,可以分析出 User₁, User₅与 User₆具有相似的喜好,因为他们对 Item₁, Item₂和 Item₄有着极为相似的评分,那么系统会推荐 Item₅给 User₆,因为与 User₆偏好相似的 User₁与 User₅均对 Item₅的评分较高。同时, User₂作为新用户

进入系统而不具有历史评分信息,因此就无法根据协同过滤推荐算法的原理分析出 User₂的行为偏好;而从物品的角度看, Item₃作为新项目存在于系统中,因缺少评分信息导致系统无法感知它的存在,所以 Item₃也就无法被推荐出去。这就形成了协同过滤推荐系统中的冷启动问题。

1.2 分类

(1) 新用户冷启动:假设新用户 u₁进入该模型,由于 u₁不完善的个人信息与行为信息,因此系统不能通过模型分析 u₁的偏好,推荐系统也就无法利用推荐算法向该用户推荐其可能喜欢的物品。这种现象被称为新用户的冷启动问题。

(2) 新物品冷启动:假设新物品 i₁进入该模型,由于 i₁没有被用户评价的信息,导致评分数据的稀少,系统不能建立更详细的特征信息表,也就不能将其与用户的喜好进行很好的匹配,因此新物品 i₁被推荐的概率就大大减少。这种现象称为新物品的冷启动问题。

(3) 新系统冷启动:可以将其理解为(1)、(2)的一种特殊情况,即所有的用户、项目对于系统来说都是新的,一个全新的系统无法产生相应的推荐即为新系统的冷启动问题。

2 冷启动问题研究现状

在基于协同过滤和基于矩阵分解的推荐算法中,冷启动问题表现得尤为明显,而这两种算法在推荐系统中应用最为广泛,因此解决推荐系统的冷启动问题也就更有必要。目前解决冷启动问题效果较好的有混合推荐、融和其他数据源、动态情景敏感策略等方式。以下就是是否考虑冷启动问题类型的解决策略进行描述与分析。

2.1 不考虑冷启动问题类型的解决策略

在推荐系统的发展过程中,随着推荐精度的不断提高,算法的复杂程度也越来越高。在不考虑冷启动类型的情况下,通常采用的策略有随机推荐策略、数值策略、基于上下文推荐等。

2.1.1 随机推荐策略

随机推荐策略是实际推荐系统应用中最简单直观的方法,即无论是否为新用户或新物品,系统都采取随机推荐的方式,再根据推荐后的反馈来了解用户的兴趣偏好以及新项目被哪些用户群所喜欢。这种策略的覆盖率较广,但准确率较低。长期来看,随机推荐策略的准确率不会超过 50%,这会使得用户对推荐系统不满意,而导致信任度的降低^[4]。

2.1.2 常用数值策略

推荐系统中常用的数值类方法包括平均值策略、流行度策略、信息熵法等。

(1) 平均值策略: 选取所有项目的评分均值对原始评分矩阵进行填充, 然后再根据填充后的评分矩阵利用协同过滤的算法产生推荐。该方法采用统一的数据对矩阵进行填充, 即认为用户对物品的喜爱程度均一致, 这种情况在实际情况中的概率是非常小的, 不符合个性化推荐的原则。因此有文献提出以平均喜好为轴, 个人的实际兴趣偏好围绕此轴上下波动的预测模式^[5], 这种方式虽然考虑了个体评分的差异性, 但是绕均值上下波动的幅度在进行矩阵填充时是很难把握的, 因此这种方法并不能够很好地解决均值法在推荐应用中的个体差异性问题。

(2) 流行度策略: 指采用所有用户评价过的最多的评分值作为对未评分项目的预测评分值。这种策略是出于考虑目标用户可能喜欢绝大多数用户都喜欢的项目, 但是该思路只能是从统计学的角度说明预测准确的概率高于不准确的概率^[6]。存在这样的情况: 某些用户具有较小众的喜好, 其兴趣偏好与大多数人的兴趣爱好存在较大的差异性, 那么对于这类用户来说, 采用众数法进行预测时不仅误差较大, 有时甚至是相反的预测, 因此众数法在某种程度上存在着较大的漏洞。

(3) 信息熵策略: 信息熵方法实际上也是一种均值策略, 它是利用系统中已被评价的项目信息, 利用一定的计算方法求出这些项目的信息熵, 并对这些信息熵的值进行排序, 采用信息熵大的项目评价作为对待评分项目的预测依据, 最后将目标用户在这些物品上的平均预测值作为该用户对新物品的预测评分^[7]。文献[8]是以整个数据集的评分均值划分好评与差评, 计算出的结果具有一定的绝对性, 忽略了个体特征的差异性; 文献[9]是以各项评分均值作为划分依据, 这类类似于决策树学习中通过信息的增益来选择分类属性, 虽然考虑到了特征属性的影响因子, 但是仍然没有达到推荐的个性化要求。在实际研究中也会对熵策略进行改进, 比如文献[10]对缺失值的改进熵进行预测评分。

2.2 考虑冷启动问题类型的解决策略

上述方法只是在整体上缓解了冷启动问题, 并且以牺牲用户的个性化需求为代价。为了提高推荐的准确性, 很多文献给出了针对冷启动问题类型的解决策略, 主要分为混合使用多种算法、融合多种数据源、考虑动态的情景推荐等策略。

2.2.1 混合推荐算法

混合算法通过考虑各种推荐算法的特点, 根据各自优劣, 融合不同的思想形成新的算法^[11], 在实际应用中证明混合使用多种推荐技术对提升推荐效果、改善推荐系统的性能等方面都有着重要意义。

在新用户冷启动方面, 文献[12]通过对已有用户的新闻点击行为进行信息记录, 从而提取用户在不同环境中的上下文情景信息, 并且利用兴趣分类的方法构建出决策树分类模型; 即当新用户到达时, 获取该用户当前环境中所带有的上下文信息, 并将其与决策树模型进行匹配, 以此来预测新用户的新闻浏览兴趣, 进而完成新闻推荐。文献[13-14]提出了基于原始评分矩阵扩充的方法, 该方法是将用户的人口统计信息和项目的内容特征信息扩充到原始的用户-项目评分矩阵中, 使得矩阵的行和列有所增加, 从而形成新的评分矩阵, 并在新的评分矩阵上应用协同过滤推荐算法。这样即使新用户或新项目不存在历史行为信息, 其新的扩充矩阵中数据也是存在的, 因此可以继续利用新的评分矩阵计算相似邻居并产生推荐, 从而避免了新用户或者新项目带来的冷启动问题; 该方法是目前解决冷启动问题研究的方向性方法。

在新项目冷启动方面, 文献[15]提出了 LCE (local collective embeddings) 算法, 融合项目内容信息与用户历史行为信息进行统一矩阵分解, 结合了矩阵分解的预测准确性与项目内容的属性信息来克服项目冷启动问题。文献[16]混合了内容过滤方法和协同过滤方法, 为项目信息文件进行虚拟分配, 从而对内容进行筛选, 在此基础上利用协同过滤产生推荐。文献[17]提出一种基于两阶段聚类的推荐算法, 将基于图摘要的方法与基于内容相似度的算法相结合, 实现以用户兴趣为主题的推荐; 实验证明该方法在稀疏数据和冷启动的情况下具有更好的推荐效果, 并且利用离线处理大量数据集的方式, 使其具有更好的在线推荐效率。

2.2.2 融合其他数据源算法

常用的推荐算法都有各自依赖的数据源, 例如协同过滤推荐算法需要用户对项目的评分信息, 基于内容的推荐算法需要项目的特征信息等。从混合使用算法的角度出发, 不难理解融合了多种数据源的方式也有助于解决冷启动问题。

因此通过引入用户社会关系信息, 帮助建立新用户特征模型, 在一定程度上缓解新用户的冷启动问题。例如, 文献[18-19]融合了用户评论与评分数据, 利用主题模型生成评论主题的分布, 再根据评分数据生成评论态度影响因子, 通过评论态度影响因子来放大评论主题分布中的突出特征, 建立更为准确的用户偏好与物品特征, 进而进行评分预测与物品推荐。文献[20]在协同过滤算法中引入用户的人口统计学信息以及社会网络信息 (包括性别、年龄、社交软件中的好友等), 结果证明了社交网络内容的巨大预测能力, 及其在冷启动用户的推荐方面显著的作用。文献

[21]则利用了 Web 日志信息中的用户浏览路径,针对用户的浏览行为,特别是与来自其他领域的访问相关的行为(这在之前的研究中已经被忽略),该文献使用从 Yahoo News 的大型视图中提取出群体浏览记录,跟踪构建出一个 BrowseGraph(浏览视图),并将 Referrer-Graph(指向视图)定义为由具有相同引用域的会话引发的子图,该子图的结构和时间属性展示了用户浏览的内容及浏览的时间等行为是高度依赖于会话的引用路径的,因此在这个观察的基础上,提出了一个解决新闻系统中冷启动问题的推荐器。

2.2.3 动态情景敏感策略

情景敏感推荐系统(CARS)是近年来推荐系统研究的热点^[22],情景也称为上下文(context),指的是利用用户模型中的额外信息(如位置、时间、天气等)来提高推荐系统的准确度与实时性^[23]。

动态情景策略通过算法感知当前状态(如新用户或新物品),从而选择适合此情景的最佳项目进行推

荐。文献[24]提出了构建“元情景”(meta-context)的策略,即对不同情景策略进行组合形成预设情景,在此基础上划分出“元情景”。这种思路类似于选取具有一定特征的项目组成种子集,通过让新用户对这些种子集的评分分析用户的喜好;而“元情景”是根据新用户与推荐系统交互产生的实时反馈信息,选择相对合适的情景作为新用户可能感兴趣的推荐,此方法与评价种子集的方法相比能更主动地了解用户、服务用户。文献[25]开发了一个交互式推荐系统,系统根据用户行为的变化能够动态地自适应,从而去匹配与用户最相近的情景,产生合适的推荐结果。算法的独特之处是包含了一个“情景变化侦测器”,用来感知用户行为的变化,一旦检测到用户行为发生重大改变时,如从某一分类的查看切换到另一种类的搜索,系统就会优先采用用户最新行为信息来重建情景模型,产生新的推荐,这样使得推荐更加即时化、智能化。

算法比较如表 2 所示。

表 2 算法比较

推荐方法	核心原理、算法	优点	缺点
数值类推荐 ^[5-10]	根据数值方法如随机数、均值、众数、信息熵等产生推荐,利用用户的反馈更新偏好信息	推荐结果具有一定的依据,且无需大规模的训练数据	个性化低,推荐覆盖率低
混合类推荐 ^[11-17]	混合使用多种算法,基于内容推荐与协同过滤推荐相结合等	产生较为可靠的推荐结果,具有针对性,可缓解冷启动问题与数据稀疏性问题	需要大规模的训练数据,稳定性差,推荐结果解释性差
融合多数据源推荐 ^[18-21]	引入用户的社会关系信息、评分、评论、位置信息等	围绕用户为中心快速产生推荐结果,缓解冷启动问题	需要学习的参数多,推荐效果受限于历史数据
动态情景推荐 ^[22-25]	获取实时动态的信息(如天气、位置、心情等)产生最新的推荐	实时分析用户偏好,产生即时的推荐结果,个性化强,准确率高	需要多种实时数据,用户推荐覆盖率低

3 结束语

推荐系统是大数据背景下最能满足用户个性化需求的信息服务工具,吸引了众多科技公司及研究学者的研究,国内外也专门成立了推荐系统方面的学术交流会议,如 ACM RecSys 等,其会议内容反映了推荐系统的前沿和热点,但同时也指出了推荐系统发展的重点与难点。综合来看,通过机器学习等方法对推荐系统中的新用户和新项目进行集中预处理,研究出既适用于冷启动问题情景又适用于一般用户、项目的推荐方法是当前研究的主体思路;而基于多种方法混合、多种数据源融合的混合推荐将是未来推荐系统研究的主要方向。

在不断提高推荐精度的同时,推荐技术也面临着很多挑战,主要包括在融合用户个人信息数据源的过程中如何确定和量化用户属性信息,如何做好个人信息的隐私保护工作以及推荐系统的效用评价,这也是未来推荐系统健康发展过程中需要着重关注的地方。

参考文献:

- [1] 朱郁筱,吕琳媛.推荐系统评价指标综述[J].电子科技大学学报,2012,41(2):163-175.
- [2] 王国霞,刘贺平.个性化推荐系统综述[J].计算机工程与应用,2012,48(7):66-76.
- [3] 孟祥武,刘树栋,张玉洁,等.社会化推荐系统研究[J].软件学报,2015,26(6):1356-1372.
- [4] GHAZANFAR M A, PRUGELBENNETT A. A scalable, accurate hybrid recommender system[C]//Third international conference on knowledge discovery and data mining. Phuket: IEEE Computer Society, 2010: 94-98.
- [5] GRECO G, GRECO S, ZUMPANO E. Collaborative filtering supporting Web site navigation[J]. AI Communications, 2004, 17(3): 155-166.
- [6] 郭弘毅,刘功申,苏波,等.融合社区结构和兴趣聚类的协同过滤推荐算法[J].计算机研究与发展,2016,53(8):1664-1672.
- [7] 孙少华.协同过滤系统的稀疏性与冷启动问题研究[D].杭州:浙江大学,2005.

- [8] BOBADILLA J ,ORTEGA F ,HERNANDO A ,et al.Recom-
mender systems survey [J]. Knowledge - Based Systems ,
2013 ,46(1) : 109-132.
- [9] LEMPEL R ,KOREN Y.On bootstrapping recommender sys-
tems[C]//Proceedings of ACM international conference on
information & knowledge management. New York: ACM ,
2010: 1805-1808.
- [10] 景民昌,张 芹,唐第官.协同过滤系统中基于种子集评分
的新用户冷启动推荐研究[J].图书情报工作,2013,57
(5): 124-128.
- [11] RASHID A M ,ALBERT I ,COSLEY D ,et al.Getting to
know you: learning new user preferences in recommender
systems[C]//Proceedings of the 7th international conference
on intelligent user interfaces. New York: ACM ,2002: 127-
134.
- [12] 杨秀梅,孙 咏,王美吉,等.新闻推荐系统中用户冷启动
问题的研究[J].小型微型计算机系统,2016,37(3): 479-
482.
- [13] 杨 阳,向 阳,熊 磊.基于矩阵分解与用户近邻模型的
协同过滤推荐算法[J].计算机应用,2012,32(2): 395-
398.
- [14] 张 川.基于矩阵分解的协同过滤推荐算法研究[D].长
春: 吉林大学,2013.
- [15] SAVESKI M ,MANTRACH A.Item cold-start recommenda-
tions: learning local collective embeddings [C]//Proceed-
ings of the 8th ACM conference on recommender systems.
New York: ACM ,2014: 89-96.
- [16] LIU H S ,GOYAL A ,WALKER T ,et al.Improving the dis-
criminative power of inferred content information using seg-
mented virtual profile [C]//Proceedings of the 8th ACM
conference on recommender systems. New York: ACM ,
2014: 97-104.
- [17] 陈克寒,韩盼盼,吴 健.基于用户聚类的异构社交网络推
荐算法[J].计算机学报,2013,36(2): 349-359.
- [18] LING G ,LYU M R ,KING I.Ratings meet reviews ,a com-
bined approach to recommend [C]//Proceedings of the 8th
ACM conference on recommender systems. New York:
ACM ,2014: 105-112.
- [19] 杨兴耀,于 炯,吐尔根·依布拉音,等.融合奇异性和扩
散过程的协同过滤模型[J].软件学报,2013,24(8): 1868-
1884.
- [20] SEDHAIN S ,SANNER S ,BRAZIUNAS D ,et al.Social col-
laborative filtering for cold-start recommendations [C]//
Proceedings of the 8th ACM conference on recommender
systems. New York: ACM ,2014: 345-348.
- [21] TREVISIOL M ,AIELLO L M ,SCHIFANELLA R ,et al.
Cold-start news recommendation with domain-d pendent
browse graph [C]//Proceedings of the 8th ACM conference
on recommender systems. New York: ACM ,2014: 81-88.
- [22] ADOMAVICIUS G ,TUZHILIN A.Context-aware recom-
mender systems [C]//Proceedings of the 8th ACM confer-
ence on recommender systems. New York: ACM ,2014: 335-
336.
- [23] 景民昌.从 ACM RecSys 2014 国际会议看推荐系统的热点
和发展[J].现代情报,2015,35(4): 41-45.
- [24] TANG Liang ,JIANG Yexi ,LI Lei ,et al.Ensemble contextual
bandits for personalized recommendation [C]//Proceedings
of the 8th ACM conference on recommender systems. New
York: ACM ,2014: 73-80.
- [25] HARIRI N ,MOBASHER B ,BURKE R.Context adaptation
in interactive recommender systems [C]//Proceedings of the
8th ACM conference on recommender systems. New York:
ACM ,2014: 40-48.

(上接第82页)

- [7] 王 旭.基于能耗约束的虚拟机调度策略研究[D].昆明:
云南大学,2011.
- [8] RAGHAVENDRA R ,RANGANATHAN P ,TALWAR V ,et
al.No power struggles: coordinated multi-level power man-
agement for the data center [J].ACM SIGARCH Computer
Architecture News ,2008,36(1): 48-59.
- [9] PINHEIRO E ,BIANCHINI R ,CARRERA E V ,et al.Load
balancing and unbalancing for power and performance in
cluster-based systems [C]//Workshop on compilers and op-
erating systems for low power. [s.l.]: [s.n.],2003: 182-
195.
- [10] CLARK C ,FRASER K ,HAND S ,et al.Live migration of
virtual machines [C]//Proceedings of the 2nd conference on
symposium on networked systems design & implementation.
Berkeley ,CA ,USA: USENIX Association ,2005: 273-286.
- [11] SVÄRD P ,HUDZIA B ,WALSH S ,et al.Principles and per-
formance characteristics of algorithms for live VM migration
[J]. ACM SIGOPS Operating Systems Review ,2015,49
(1): 142-155.
- [12] HINES M R ,GOPALAN K.Post-copy based live virtual ma-
chine migration using adaptive pre-paging and dynamic self-
ballooning [C]//Proceedings of the 2009 ACM SIGPLAN/
SIGOPS international conference on virtual execution envi-
ronments. Washington ,DC: ACM ,2009: 51-60.
- [13] 陈伯龙,程志鹏.云计算与 OpenStack (虚拟机 Nova 篇)
[M].北京: 电子工业出版社,2013.
- [14] 戴 友.OpenStack 开源云-王者归来[M].北京: 清华大学
出版社,2014.
- [15] OpenStack. Install guides [EB/OL].2016. <http://docs.openstack.org/kilo/>.
- [16] BELOGLAZOV A ,BUYA R.Adaptive threshold-based ap-
proach for energy-efficient consolidation of virtual machines
in cloud data centers [C]//Proceedings of the 8th interna-
tional workshop on middleware for grids ,clouds and e-sci-
ence. New York: ACM ,2010: 1-6.