

个性化习题路径推荐方法研究综述

冯旭光, 张 峰

(山东科技大学计算机科学与工程学院, 山东 青岛 266590)

✉1812553272@qq.com; zhangfengskd@163.com



摘 要: 个性化习题路径推荐技术能够综合考虑学习者的个性化特征, 为学习者量身定制习题路径。文章系统地梳理了个性化习题路径推荐研究工作。首先, 从推荐方式的角度, 介绍了全局最优习题路径推荐和局部迭代习题路径推荐方法, 总结了两类推荐方法的优势及其存在的问题。然后, 从目前个性化习题路径推荐工作使用较多的核心算法的角度, 介绍了基于协同过滤、认知诊断、知识追踪、深度学习和强化学习五类方法。最后, 探讨了该领域当前的研究难点, 并展望未来研究工作的方向。

关键词: 习题路径推荐; 个性化; 全局最优路径; 局部迭代路径; 推荐算法

中图分类号: TP391 **文献标识码:** A

A Research Overview of Personalized Exercise Path Recommendation Methods

FENG Xuguang, ZHANG Feng

(College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China)

✉1812553272@qq.com; zhangfengskd@163.com

Abstract: The personalized exercise path recommendation technology can take into account the individual characteristics of learners and customize the exercise path for them. In this paper, the research work on personalized exercise path recommendation is systematically reviewed. First of all, from the perspective of recommendation methods, the global optimal exercise path recommendation and local iterative exercise path recommendation methods are introduced, and the advantages and problems of the two types of recommendation methods are summarized. Then, from the perspective of the core algorithms that are widely used in the current personalized exercise path recommendation work, five methods are introduced, which are based on collaborative filtering, cognitive diagnosis, knowledge tracking, in-depth learning and reinforcement learning. Finally, current research difficulties in this field are discussed, and the future research directions are prospected.

Keywords: exercise path recommendation; personalized; global optimal path; local iterative path; recommendation algorithm

1 引言(Introduction)

在线题库和在线判题系统[如“力扣”(leetcode), 为全球程序员提供IT技术职业化提升的平台]因学习资源多且用户基数大, 采用人工方式无法为每位平台的学习者提供个性化指导。个性化习题路径推荐旨在根据学习者的学习能力、知识背景等个性化特征, 自动从题库中选取适合该学习者的习题, 并以合理的顺序推荐习题。本文对近年来开展的个性化

习题路径研究工作进行了梳理, 首先按照推荐形式的不同, 将现有工作分为全局最优路径推荐和局部迭代路径推荐两种类型, 并对其实现方法和优点、缺点进行了介绍; 然后对常用的五种个性化习题路径推荐工作中常用的五种核心算法及使用这些算法的代表性工作进行了总结和归纳; 最后从个性化特征的挖掘和习题路径的合理性两个角度阐述了当前该领域研究工作的难点及展望了未来的研究方向。

2 个性化习题路径推荐类型(Types of personalized exercise path recommendation methods)

个性化习题路径推荐的主要任务是根据学习者的知识水平、学习风格、学习习惯等个性化特征，从在线题库中选择适合学习者的题目，并按照习题特点和学习者的学习偏好对推荐的习题进行排序，最终以习题序列的形式推荐给学习者，个性化习题路径推荐流程如图1所示。

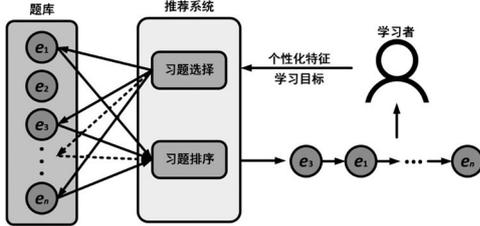


图1 个性化习题路径推荐流程

Fig.1 Recommendation process of personalized exercise path

不同的个性化习题路径推荐工作的实现方法和推荐目标各不相同。本文梳理了以往的个性化习题路径推荐研究工作，借鉴云岳等^[1]对个性化学习路径推荐的研究思路，将现有的推荐方法按推荐形式分为注重完整推荐内容的全局最优路径推荐和注重学习者学习过程的局部迭代路径推荐两种类型。本节将分别介绍两类推荐方法及其优点和缺点。

2.1 全局最优路径推荐

全局最优路径推荐工作是从推荐内容的完整性出发，在一次推荐中为学习者呈现能够满足其学习目标的完整习题路径，如图2所示。该类方法能够为学习者展示整个练习流程，方便学习者对将要练习的题目建立完整的认识和清晰的规划。此外，全局最优路径推荐中，每个推荐任务仅需运算一次，具有推荐速度快、运行效率高的特点。



图2 全局最优路径推荐

Fig.2 Global optimal path recommendation

在全局最优路径推荐工作中，XIA等^[2]首先将学习者答题结果划分为六种类型，用于计算习题难度。然后将学习者已经完成的习题作为当前状态，通过马尔可夫模型计算出的转移矩阵，得到后续每个习题的出现概率。最后根据转移矩阵、习题难度和学习者需求，为学习者推荐流行路径、挑战路径和渐进路径。SEGAL等^[3]提出了个性化习题难度排序算法(EduRank)，用来为学习者生成个性化习题难度递增序列。本研究根据相似学习者的已知难度排名，通过协同过滤投票机制，在不通过预测成绩的情况下，直接为不同的学习者生成个性化的习题难度排序。XIA等^[4]将习题路径的练习范围能够覆盖到所有指定的知识点作为推荐目标。本文首先定义了知识概念图，用于表示知识点之间的关系；然后根据知识概念图筛选出与目标知识点相关的所有题目；最后考虑习题难度，为学习者推荐符合知识点学习顺序且难度适宜的

习题路径。

全局最优路径推荐方法能够为学习者提供完整习题路径，但忽略了学习者在题目练习过程中知识水平等个性化特征的变化，导致路径的后段推荐精度降低。

2.2 局部迭代路径推荐

局部迭代路径推荐方法考虑到学习者在学习过程中个性化特征的变化，在一个推荐任务当中，每次向学习者推荐部分习题，并收集学习者的练习得分、知识水平变化等信息，然后根据这些信息推荐后续习题，如此反复迭代，直到完成学习目标，如图3所示。局部迭代路径推荐以更细的粒度为学习者规划习题路径。通过追踪学习者个性化特征的变化情况，局部迭代路径推荐能够实现更高的推荐精度。



图3 局部迭代路径推荐

Fig.3 The core algorithm used for personalized exercise path recommendation

针对局部迭代路径推荐方法，SAITO等^[5]将学习者的当前能力和目标能力用图表进行统计，并根据历史学习者答题记录使用长短期记忆神经网络模型(Long Short-Term Memory, LSTM)推荐下一道适合学习者的习题。YERA等^[6]使用协同过滤的方法，为学习者推荐下一道习题。本研究根据题库中每道习题的回答状态，将每位历史学习者的学习记录建模为学习者矩阵，对于每道没有被当前学习者解决的习题，根据前k个相似的用户计算出推荐分数，并在每次推荐中推荐得分最高的习题。DIAO等^[7]首先根据知识点的先决、包含和平行关系构造知识树；然后将学习目标表示为知识点集合，根据知识树表示的知识点之间的关系以及知识点掌握程度，为学习者逐步推荐未完全掌握知识点的相关习题。当学习者完成这些习题，则开始判断学习者对知识点是否掌握，如果掌握，就推荐下一个知识点的习题，否则继续推荐与该知识点相关的习题，直到学习者掌握最终目标知识点。

局部迭代路径推荐方法考虑了学习者在练习过程中个性化特征发生的变化，但是由于习题路径在学习过程中不断生成，导致学习者难以把握学习任务的整体内容，也很难做到对学习进度的控制。此外，由于需要在一个任务中进行多次计算和推荐，所以增大了推荐系统的运行压力。

3 个性化习题路径推荐所使用的核心算法(The main algorithm used for personalized exercise path recommendation)

推荐算法是一个推荐系统的核心，不同的推荐算法适用于不同的推荐目标和推荐场景。本节将介绍个性化习题路径推荐工作使用最多的五类核心算法，即协同过滤、认知诊断、知识追踪、深度学习和强化学习，并对使用各类算法的主要工作进行梳理和总结。

3.1 基于协同过滤的方法

协同过滤根据其他用户的偏好向目标用户推荐，它首先找出一组与目标用户偏好一致的邻居用户，然后分析该邻居用户，把邻居用户喜欢的项目推荐给目标用户^[8]。

在基于协同过滤的习题路径推荐工作中，LABAJ等^[9]发现当学习者在没有推荐的情况下自由选择习题时，会选择目前对他们来说比较容易和比较难的习题。为解决这个问题，本研究提出了一种基于相似学习者的难度反馈，为目标学习者推荐适当难度习题的推荐方法。TOLEDO等^[10]将模糊技术应用与协同过滤方法结合，根据其他用户的历史答题信息为当前用户推荐习题路径。本研究还提出了一种数据预处理方法，用于去除历史信息中可能影响推荐的异常行为，进一步提高推荐精度。

3.2 基于认知诊断的方法

认知诊断的目标是通过一批学习者在一批题目上的作答情况，分析每一位学习者对知识点的掌握程度。

在基于认知诊断的研究工作中，朱天宇等^[11]首先使用认知诊断模型对学习者的掌握程度进行建模，然后结合概率矩阵分解方法预测学习者正确回答习题的概率 p_i ，最后根据实际需求，设置推荐习题的难度范围，即 $[\beta_1, \beta_2]$ ，并通过 p_i 计算该习题对学习者的相对难度，筛选相对难度在推荐难度范围内的试题，从而向每一位学习者进行个性化习题序列推荐。LI等^[12]根据学习者选择的知识点目标和期望的分数范围，为学习者推荐与知识点目标高度契合的习题。首先，通过认知诊断判断学习者对知识的掌握程度，并为每一位学习者对每个知识点的掌握程度标定等级。然后，通过历史学习者的做题情况统计每道练习对它所包含的每个知识点的掌握程度的要求，并据此向学习者推荐习题。

3.3 基于知识追踪的方法

知识追踪(Knowledge Tracing, KT)旨在建立学习者知识状态随时间变化的模型，以判断学习者对知识的掌握程度。知识追踪和认知诊断一样，其目的都是为了获得学习者对知识点的掌握度；它们区别在于认知诊断需要一批学习者的一批答题记录，知识追踪则只需要一个学习者对应的答题序列，并动态更新学习者的实时知识点掌握度^[13]。

在基于知识追踪的工作当中，WU等^[14]从知识点覆盖度、知识点掌握程度、习题多样性三个角度为学习者规划习题路径。首先，本研究使用向量表示每道题目的知识点覆盖情况。然后，基于RNN神经网络，根据学习者从开始学习到现在所完成的练习，预测下一次出现的练习应该包括哪些知识点，并使用深度知识追踪(Deep Knowledge Tracing, DKT)预测学习者对这些知识点掌握程度。最后，通过学习者对不同知识点的掌握程度，为学习者推荐练习范围能够覆盖到掌握程度不足的所有知识点，且习题具有适当难度的习题路径。为保证推荐的习题路径具有新颖性和多样性，本研究还加入了两个过滤层用来过滤重复和相似的习题。DAVID等^[15]考虑到单一的知识追踪模型并不适合所有的学习场景，于是设

计了一种多模型选择算法，根据被推荐者过去的做题情况，从训练好的多个模型中选择最优模型，然后根据这个模型预测的知识点掌握程度和分数，为每一道待推荐的习题进行打分，每次向学习者推荐得分最高的习题。

3.4 基于深度学习的方法

近年来，深度学习被广泛应用于教育推荐领域。各类深度学习算法的训练本质上都是对自然界复杂映射关系进行抽象与拟合，而学习者与习题路径之间正对应一种复杂的映射关系，其中的规律从理论上可由深度神经网络进行表示和学习。

在基于深度学习方法的研究工作中，WANG等^[16]首先使用深度学习训练和生成学习者和练习习题的初始特征表示，然后使用因果推理干预算法对这些特征进行微调，最后再次使用深度学习预测学习者做题的得分情况，并从中挑选难度适宜的习题推荐给相应的学习者。FANTOZZI等^[17]基于自动编码神经网络(Auto-encoder Neural Network, ANN)构造了一种推荐方法，能够逐步为学习者推荐难度递增的习题序列。HUO等^[18]认为目前的知识追踪方法不足以对学习者的知识状态进行个性化建模，并提出了一种个性化的建模方法，通过LSTM网络为每一位学习者估计和维护一个独立的知识状态向量，并由此构造了一种新的个性化习题推荐模型，其具有个性化机制的长短期记忆神经网络模型(LSTMCQP)。实验证明，该模型推荐不同练习的能力优于现有方法。

3.5 基于强化学习的方法

强化学习(Reinforcement Learning, RL)是一个用于目标导向学习和顺序决策建模和自动化的计算框架。RL特别适用于涉及代理人需要学习在不同情况下做什么的策略——即如何将状态映射到行动上，以使长期利益最大化的设置。

在基于强化学习方法的研究工作中，HUANG等^[19]将学习者的做题过程抽象成一个马尔科夫过程，并使用深度强化学习推荐下一步的最优习题。在该过程中，作者设计了三个奖励函数用于同时优化三个推荐目标。CAI等^[20]将学习者的习题练习过程表示为马尔科夫决策过程，并使用深度知识追踪技术从学习者练习习题的过程中动态检索学习者的知识水平，并将其作为学习状态，在此基础上使用强化学习技术为学习者推荐最优的习题。

4 研究难点与未来发展方向(Research difficulties and future development direction)

根据对个性化习题路径推荐领域的分析和研究，其研究难点及未来的发展方向主要体现在以下两个方面。

(1)个性化特征的挖掘。在个性化习题路径推荐方法中，一个关键步骤是根据学习者的个性化特征对其进行建模。然而，不同学习者进行习题练习的场景和目的各不相同，如何根据不同的推荐场景和学习目标挖掘具有代表性的个性化特征，成为能否成功推荐个性化习题路径的关键，这也是个性化习题路径推荐领域目前研究的难点之一。

(2)习题路径的合理性。随着信息技术与传统教育方法的融合，许多优秀的教育理论被应用在电子学习之中。针对个

性化习题路径推荐领域,如何将更为科学的教育理论应用到推荐技术当中,使习题路径在满足学习者个性化要求的基础上符合教育规律,是目前面临的主要困难和未来重要研究方向。

5 结论(Conclusion)

个性化习题路径推荐作为在线学习的一个研究领域,有很多需要深入研究的问题。本文对近年来针对个性化习题路径推荐的相关研究工作,从推荐形式和核心方法两个角度对其进行了分类和整理,并对当前的研究难点和未来的研究方向进行了阐述。希望能够为个性化习题路径推荐研究领域构建一个较为完整的全景图,为相关研究人员提供参考和借鉴。

参考文献(References)

- [1] 云岳,代欢,张育培,等.个性化学习路径推荐综述[J].软件学报,2021,33(12):4590-4615.
- [2] XIA M, SUN M, WEI H, et al. Peerlens: peer-inspired interactive learning path planning in online question pool[C]// BREWSTERS, FITZPATRICK G, COX A, et al. Proceedings of The 2019 CHI Conference on Human Factors in Computing Systems. New York, USA: ACM, 2019:1-12.
- [3] SEGAL A, GAL K, SHANI G, et al. A difficulty ranking approach to personalization in E-learning[J]. International Journal of Human-Computer Studies, 2019, 130:261-272.
- [4] XIA J, LI G, CAO Z. Personalized exercise recommendation algorithm combining learning objective and assignment feedback[J]. Journal of Intelligent & Fuzzy Systems, 2018, 35(3):2965-2973.
- [5] SAITO T, WATANOBE Y. Learning path recommender system based on recurrent neural network[C]// ZHU Q, HAN Y, MIAO J, et al. 2018 9th International Conference on Awareness Science and Technology. Los Angeles, USA: IEEE, 2018:324-329.
- [6] YERA R, MARTÍNEZ L. A recommendation approach for programming online judges supported by data preprocessing techniques[J]. Applied Intelligence, 2017, 47(2):277-290.
- [7] DIAO X, ZENG Q, DUAN H, et al. Personalized exercise recommendation driven by learning objective within e-learning systems[J]. International Journal of Performability Engineering, 2018, 14(10):2532.
- [8] 冷亚军,陆青,梁昌勇.协同过滤推荐技术综述[J].模式识别与人工智能,2014,27(8):720-734.
- [9] LABAJ M, BIELIKOVÁ M. Utilization of exercise difficulty rating by students for recommendation[C]// POPESCU E, LAU R, PATA K, et al. International Conference on Web-Based Learning. Berlin, Germany: Springer, 2014:13-22.
- [10] TOLEDO R Y, MOTA Y C, MARTÍNEZ L. A recommender system for programming online judges using fuzzy information modeling[C]// BRYANT A. Informatics. Basel, Switzerland: MDPI, 2018: 17.
- [11] 朱天宇,黄振亚,陈恩红,等.基于认知诊断的个性化试题推荐方法[J].计算机学报,2017,40(1):176-191.
- [12] LI Z Z, XU F. A personalized exercise recommendation method for teaching objectives[C]// LU Y, CHENG C. International Conference on Computer Application and Information Security. Bellingham, USA: SPIE, 2022:450-455.
- [13] 李习华,曹云波.融入常态化教学的个性化学习模式探索[J].人工智能,2022,2:88-95.
- [14] WU Z, LI M, TANG Y, et al. Exercise recommendation based on knowledge concept prediction[J]. Knowledge-Based Systems, 2020, 210:106481.
- [15] DAVID Y B, SEGAL A, GAL Y. Sequencing educational content in classrooms using Bayesian knowledge tracing[C]// GAŠEVIĆ D, LYNCH G, DAWSON S, et al. Proceedings of The Sixth International Conference on Learning Analytics & Knowledge. New York, USA: ACM, 2016:354-363.
- [16] WANG S, MA Z, JI H, et al. Personalized exercise recommendation method based on causal deep learning: Experiments and implications[J]. STEM Education, 2022, 2(2):457.
- [17] FANTOZZI P, LAURA L. Collaborative recommendations in online judges using autoencoder neural networks[C]// DONG Y, VIEDMA E, MATSUI K, et al. International Symposium on Distributed Computing and Artificial Intelligence. Berlin, Germany: Springer, 2020:113-123.
- [18] HUO Y, WONG D F, NI L M, et al. Knowledge modeling via contextualized representations for LSTM-based personalized exercise recommendation[J]. Information Sciences, 2020, 523:266-278.
- [19] HUANG Z, LIU Q, ZHAI C, et al. Exploring multi-objective exercise recommendations in online education systems[C]// ZHU W, TAO D, CHENG X, et al. Proceedings of the 28th ACM International Conference on Information and Knowledge Management. New York, USA: ACM, 2019:1261-1270.
- [20] CAI D, ZHANG Y, DAI B. Learning path recommendation based on knowledge tracing model and reinforcement learning[C]// GUAN S, LI T, HE X, et al. 2019 IEEE 5th International Conference on Computer and Communications. Los Angeles, USA: IEEE, 2019:1881-1885.

作者简介:

冯旭光(1997-),男,硕士生.研究领域:教育数据挖掘。

张峰(1981-),男,博士,副教授.研究领域:教育数据挖掘,软件工程.本文通信作者。