

在线教育环境中学习共同体研究综述^{*}

张俊涛¹, 杨先娣¹, 宋伟¹, 张雪龙¹, 贺宇阳¹, 彭智勇^{1,2}

¹(武汉大学 计算机学院, 湖北 武汉 430072)

²(武汉大学 大数据研究院, 湖北 武汉 430072)

通信作者: 彭智勇, E-mail: peng@whu.edu.cn



摘 要: 随着信息技术与教育的深度融合, 蓬勃发展的在线教育已成为教育信息化进程的新常态, 并产生了海量的教育数据, 但也面临辍学率高、课程完成率低、监管不足等问题, 因此如何对海量教育数据进行挖掘和分析是解决这些问题的关键. 学习共同体是以学习者为核心要素的学习组织, 强调学习过程中学习者之间互动交流、资源共享以及协作学习等行为, 从而完成共同的学习任务或目标. 对在线教育环境中学习共同体的研究进行回顾、分析和展望. 首先, 介绍在线教育环境中学习共同体的背景与重要性. 其次, 介绍不同学科中学习共同体的定义. 然后, 总结同质、异质和混合 3 种类型学习共同体的构建方法. 接着, 从共享、协作和激励 3 个方面讨论学习共同体的管理机制. 最后, 探讨和展望学习共同体未来的研究方向.

关键词: 在线教育; 学习共同体; 构建方法; 管理机制

中图法分类号: TP391

中文引用格式: 张俊涛, 杨先娣, 宋伟, 张雪龙, 贺宇阳, 彭智勇. 在线教育环境中学习共同体研究综述. 软件学报, 2023, 34(11): 5058–5083. <http://www.jos.org.cn/1000-9825/6735.htm>

英文引用格式: Zhang JT, Yang XD, Song W, Zhang XL, He YY, Peng ZY. Survey on Learning Communities in Online Education Environment. Ruan Jian Xue Bao/Journal of Software, 2023, 34(11): 5058–5083 (in Chinese). <http://www.jos.org.cn/1000-9825/6735.htm>

Survey on Learning Communities in Online Education Environment

ZHANG Jun-Tao¹, YANG Xian-Di¹, SONG Wei¹, ZHANG Xue-Long¹, HE Yu-Yang¹, PENG Zhi-Yong^{1,2}

¹(School of Computer Science, Wuhan University, Wuhan 430072, China)

²(Big Data Institute, Wuhan University, Wuhan 430072, China)

Abstract: With the in-depth integration of information technology and education, booming online education has become the new normal of the education informatization process and generated massive amounts of education data. However, online education also faces high dropout rates, low course completion rates, insufficient supervision, and other problems. How to mine and analyze the massive education data is the key to solving these problems. A learning community is a learning organization with learners as its core element, and it emphasizes the interactive communication, resource sharing, and collaborative learning among the learners in the learning process so that common learning tasks or goals can be completed. This study reviews, analyzes, and discusses the prospect of the research on learning communities in the online education environment. Firstly, the background and importance of learning communities in the online education environment are outlined. Secondly, the definitions of a learning community in different disciplines are presented. Thirdly, the construction methods for three types of learning communities, namely, homogeneous, heterogeneous, and hybrid learning communities, are summarized. Fourthly, the management mechanism for learning communities is discussed from the three aspects of sharing, collaboration, and incentive. Last but not least, directions for future research on learning communities are suggested.

Key words: online education; learning community; construction methods; management mechanism

随着现代化信息技术与教育的深度融合, 在线教育或在线学习 (online education or e-learning) 的发展与应用打破了课堂式教学环境的时空限制, 颠覆了传统的教学方式, 并成为教育信息化的发展趋势. 在线教育的蓬勃发展

* 基金项目: 国家自然科学基金 (U1811263, 62072349); 国家重点研发计划 (2020YFC1522602); 湖北省科技重大专项 (2021BEE057)
收稿时间: 2021-12-23; 修改时间: 2022-03-22, 2022-05-12; 采用时间: 2022-06-05; jos 在线出版时间: 2023-05-18
CNKI 网络首发时间: 2023-05-19

致使教育数据以前所未有的速度增长与积累, 如何对教育大数据进行挖掘与分析是教育领域亟需解决的任务^[1]. 由于在线教育是基于互联网的学习模式, 学习者在学习过程可以时空不同步且不受任何监督, 很容易导致学习者之间出现互动交流少、课程完成率低与辍学率高等问题, 因此面向教育大数据对学习者进行群体建模具有重要的研究价值与意义. 学习共同体或学习社区 (learning community) 是将学习者和助学者建模为一个学习群体或组 (learning group), 强调他们在学习过程中的互动交流、资源分享和协作学习等行为, 从而促进他们完成共同的任务或目标. 学习共同体在协作学习、同伴学习等方向具有重要的研究价值与应用前景^[2-7].

学习共同体很早已经在教育领域得到了深入研究, 它综合应用了教育学、计算机科学、社会学等多学科理论与技术来解决学习过程中所遇到的问题. 近年来, 随着互联网、云计算、大数据、人工智能等技术与教育的深度融合, 在线教育环境中学习共同体的研究掀起了热潮. 尤其是, 以 Coursera, Udacity, edX, XuetangX 等为代表的大规模开放式在线课程 (massive open online courses, MOOCs)^[8] 的兴起与发展, 使得在线教育环境中学习共同体的研究变得更加具有趣味性与挑战性. 例如, NovoEd 在线教育平台已开发了基于团队学习的功能模块, 它要求学习者学习课程前必须创建或加入学习共同体. 在传统的课堂教育环境中, 教师将学习过程中相似的学习者划分为一个学习共同体或者采用随机分组的方式, 以帮助改善教学效果、提升学习者的协作学习效率与探究能力. 然而, 随着在线教育的发展, 学习者在学习过程中不再受时空限制, 学习方式更加灵活、学习内容更加多样化, 并且一门课程可能有成千上万个学习者, 因此不能简单地通过教师分组或随机分组等方式解决.

本文对学习共同体的研究进展进行跟踪分析, 深入研究学习共同体在学习过程中的重要性, 并对学习共同体的相关研究工作进行归纳与总结. 本文第 1 节介绍在线教育环境中学习共同体的重要性, 本文综述与现有综述之间的区别以及本文贡献. 第 2 节从教育学和计算机科学两个领域介绍学习共同体的定义, 并给出本文综述对动态学习共同体的形式化定义. 第 3 节介绍 3 种类型学习共同体的构建方法, 并对现有研究工作的方法进行总结与分析. 第 4 节分别从共享、协作以及激励 3 个方面讨论学习共同体管理机制的研究进展. 第 5 节探讨学习共同体未来的研究方向. 第 6 节对本文综述的内容进行总结.

1 引言

1.1 研究背景

随着计算机、互联网、云计算、大数据与人工智能等技术与教育的深度融合, 在线教育环境打破了传统课堂教学环境的时空限制, 使教育信息化进入了一个全新的阶段, 呈现出智能化、泛在化、个性化、开放化、协同化的发展趋势^[9]. 尤其是, MOOCs 于 2012 年的兴起与发展, 以 Coursera, Udacity, edX, XuetangX 等为代表的在线教育平台使静态封闭的教育环境向互动开放的在线教育环境转变, MOOCs 作为一种新的教育方式受到了广泛关注. 如图 1 所示, 根据 2021 年 class central^[10] 的报告 (不包括中国), 全球 950 所高校共开设了 1.94 万门课程 (图 1 右半部分显示了 MOOCs 上开设课程数量变化趋势), 提供了 1670 个微证书和 70 个在线学位, 共有 2.2 亿位学习者注册 MOOCs 进行学习, 在线教育已经成为教育信息化进程的新常态. 由于学习者可以通过互联网与其他的学习者互动交流或访问课程资料等, 因此在线教育平台拥有了空前规模和分辨率的学习者活动记录^[11]. 面对在线教育蓬勃发展所产生的海量教育数据, 如何挖掘学习者之间的关联关系以及建模学习者群体以促进学习者之间的互动与协作, 使所有学习者的整体学习收益最大化是科学家和政策制定者应关注的问题^[4].

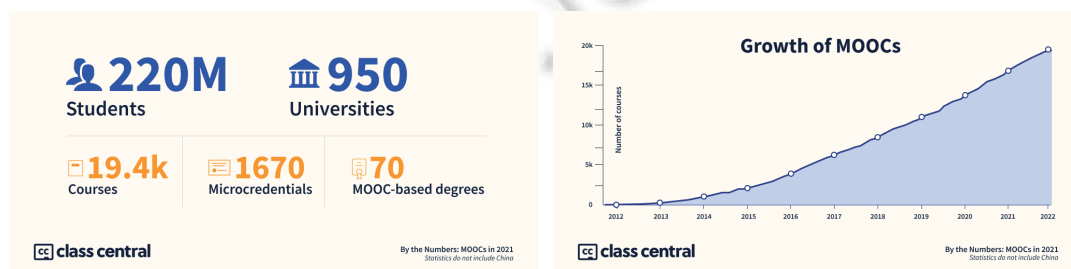


图 1 2021 年 MOOCs 统计和趋势综述

另一方面,虽然在线教育为教育信息化进程与跨时空学习提供了新的机遇,但是与传统的课堂教育环境相比,拥有大规模学习者的在线教育平台却面临着辍学率高^[12,13]、课程完成率低^[14]等问题.根据文献[15,16]的研究显示 40%–80% 的在线学习者会从在线课程中辍学,这使得在线教育中课程保留率比传统课堂学习方式低了约 10%–20%.Rovai 指出在线教育过程中学习者与学习者、学习者与教师的物理分离是导致辍学率上升的主要原因^[17].除此之外,导致在线教育辍学率高、课程完成率低的原因还有很多,例如缺乏有效的监督机制、缺乏对学习时间的严格约束、学习者之间缺乏沟通交流、协作互动等.尽管在线教育平台的任何一门课程都拥有数以千计的学习者在同时学习,但是这些在线学习者在时间和空间上可能是不统一的,这就意味着在线学习者实际上是独自学习,缺乏学习伙伴或学习小组的相互支持,从而导致了他们在学习过程中产生孤独感与厌学情绪.由于在线学习者无须任何的申请或说明就可以随时地退出当前的课程学习且不会受到任何的惩罚^[16],这也导致了学习者毫无顾忌地放弃课程学习.在线教育已经使学习活动成为以信息资源和人际交互为基础的在线社会化协作学习环境.从社会互动和协作学习的角度,构建学习共同体是促进在线教育持久性发展的关键特征^[18].同时,Rovai 认为学习者参与在线教育环境中的学习共同体对完成课程学习有积极影响^[17].因此,对于在线教育的开发者或教学者来讲,设计行之有效的算法或模型构建学习共同体来降低学习者的辍学率、缓解学习者的学习孤独感以及提高课程完成率等是一项值得深思的任务.

教育环境是开展教学和学习活动的场所和载体,它可以是传统的学校和教室,互联网上的远程教育系统,也可以是安装在智能手机上的教学软件,或者是微博、微信等社交媒体^[19].事实上,无论是传统的班级式教学环境,还是在线教育平台或教学软件,它们的目的都是将学习者归聚为一个学习群体,然后通过相互协作、共享资源等方式完成共同的学习任务或目标,而这个学习群体被视为学习共同体.受新型冠状病毒 (COVID-19) 影响,2020 年国内的教育教学任务被迁移到在线教育环境中完成,例如腾讯的 QQ 与微信、钉钉等社交+教育类型 APP 软件.这些在线教育软件的群组或分组功能成为构建学习共同体的载体,尤其是钉钉能根据不同的学习任务或目标手动将学习者分到不同层次的学习组中,促进他们在学习过程中的互动和协作以提高他们完成任务或目标的效率.此外,国外的 NovoEd 是一个基于学习团队的 MOOCs 平台,学习者开始学习课程时,必须发起或加入一个学习团队共同学习,通过团队合作提高学习过程中的学习效率.然而,仅将基于团队的学习纳入 MOOCs 并不会自动地带来成功的团队^[20].因此,在线教育环境中,开发自动构建的学习共同体系统并进行高效地管理,有助于提高学习者之间的互动与协作.

面对在线教育产生的海量教育数据以及所遇到的问题,有目的地根据学习者的属性特征对他们进行分组或组建学习者团队^[1],以及将具有共同学习目标或任务的学习者形成学习小组^[21]是非常必要的,这些团队或小组被称为“学习共同体”.学习共同体以学习者为核心,其目的是将学习者进行分组或形成不同的学习者团队,以便于他们通过资源共享、互动交流、相互补充等来提高学习过程的效率,从而保证能够完成学习目标或任务.在在线教育环境中,学习共同体综合运用了教育学、计算机科学、社会学等多学科的理论与技术对学习过程中的学习者进行建模,而基于学习共同体的 MOOCs 平台则有助于促进合作学习和终身学习的发展.当前,已有研究工作通过学习者的属性特征来构建学习共同体,然后通过共享、协作与激励等进行管理来保障学习共同体的有效运作,从而缓解在线教育过程所面临的问题.学习共同体已经在协作学习^[2,3]、同伴学习^[4–7]等方向产生了重要的研究价值与应用前景.

1.2 本文与现有综述的区别

当前,有关学习共同体的研究已经取得了一定的进展,MOOCs 的发展也使学习共同体受到前所未有的关注.已有综述分别从教育学和计算机科学两个领域对学习共同体的研究工作进行总结,本文综述则结合教育学与计算机科学等学科对学习共同体的研究进展进行总结,并与现有研究综述进行区别分析.

在教育领域,学习共同体的相关综述有文献[22–24],这些综述主要是说明学习共同体的含义以及它在学习活动中的重要性.文献[22]梳理了在线教育领域中学习共同体的内涵、构成要素、测量评价、不同类型、促进策略、参与者类型、发展阶段以及构建策略等内容,旨在推动国内学习共同体的研究.文献[23]分析了当前学习

共同体的研究状况, 对其概念进行溯源, 然后阐释了不同学者对学习共同体内涵的认知, 最后指出它的发展趋势与特征. 文献 [24] 首先介绍了什么是专业学习共同体, 然后从共同价值与愿景、群体责任、合作等说明专业学习共同体不同阶段的成长, 接着从学习过程、领导力、信任和积极的关系等说明创建专业学习共同体的过程, 并从个体变化、校园环境、专业学习基础设施等分析了发展专业学习共同体的影响因素, 最后讨论了专业学习共同体是否可持续. 与此不同的是, 本文遵从教育学领域对学习共同体的概念描述, 然后从计算机科学视角对学习共同体的形式化定义、构建方法与管理机制等进行总结与分析, 旨在提高学习者学习效率以及协作学习的能力.

据我们所知, 与计算机科学相关的学习共同体综述有文献 [25–27], 然而这些综述只是在构建学习共同体的方法上进行总结. 文献 [25] 从遗传算法的角度对学习共同体的构建进行总结, 并对现有工作中所提到的遗传算法在学习者属性特征、编码与评估方法、遗传因子、初始化与终止条件、算法评估以及分组结果等进行讨论与分析. 文献 [26] 对计算机支持的协作学习环境中学习共同体构建方法进行总结, 然后分别从学习者特征、算法 (包括遗传算法、粒子群算法、聚类算法以及分组标准等)、教师或系统决定、构建学习共同体类型 (同质、异质或混合)、算法有无改进可能、有无实验结果等方面对现有文献进行分析对比. 文献 [27] 从学习共同体的类型 (同质或异质)、学习者特征、规模大小以及算法等问题对研究工作进行分类梳理. 这些综述已经对学习共同体构建的多个内容进行较为全面总结, 本文在这些综述的基础上, 补充学习共同体构建的最新研究; 同时添加学习共同体的形式化定义描述, 并对学习共同体构建之后的管理机制研究进行总结与讨论.

本综述与现有综述的区别是: 我们从学习共同体的研究领域、定义描述、构建方法、管理机制 4 个维度分类并总结当前研究工作. 研究领域侧重计算机科学的研究工作, 并结合教育学、社会学等使综述内容更丰富; 定义描述则重点介绍计算机科学中学习共同体的形式化定义, 并给出本文综述对动态学习共同体的形式化定义, 这是对现有综述工作的补充; 构建方法是在现有综述基础上加入最新研究进展, 使本文综述的内容更全面; 管理机制从共享、协作与激励 3 个方面对构建后的学习共同体进行有效管理, 以保证学习者的积极性与学习共同体的有效运作, 这项工作是对现有综述内容的补充. 本综述详细介绍学习共同体每个维度的研究进展, 表 1 显示了与现有综述在各个维度的覆盖对比.

表 1 本文综述与现有综述的比较

文献	研究领域		定义描述		构建方法	管理机制		
	教育学	计算机科学	概念定义	形式化定义		共享机制	协作机制	激励机制
[22]	√	—	√	—	—	—	—	—
[23]	√	—	√	—	—	—	—	—
[24]	√	—	√	—	—	—	—	—
[25]	—	√	—	—	√	—	—	—
[26]	—	√	—	—	√	—	—	—
[27]	—	√	—	—	√	—	—	—
本文	√	√	√	√	√	√	√	√

1.3 本文贡献

本综述对学习共同体的现有研究工作和当前研究进展提供了一个深入的研究梳理与总结. 如第 1.2 节所描述的内容, 我们分别检索计算机科学、教育学等领域的期刊与会议, 从学习共同体的研究领域、定义描述、构建方法与管理机制等多个方面梳理与总结, 最后指出学习共同体的未来研究方向. 与现有综述相比, 本综述内容更加丰富、系统. 本文具体的贡献有以下 4 个方面.

(1) 分析当前在线教育中辍学率高、课程完成度低、监管不足等问题, 讨论学习共同体解决这些问题的意义; 与现有的学习共同体相关综述相比, 本综述涉及多个学科、内容丰富、讨论层次深.

(2) 对在线教育环境中学习共同体的构建方法进行梳理, 分别对学习共同体的类型、构建算法、规模大小、关键属性特征与测量指标等方面进行详细地总结, 并讨论其不足.

- (3) 为保障学习共同体的活跃性与持久性, 分别从共享、协作以及激励 3 个方面介绍学习共同体管理机制的研究工作, 并对研究工作进行归纳与总结.
- (4) 对比了现有学习共同体的研究进展, 从学习共同体的动态构建、多学习共同体的管理以及学习共同体的精准教学评价等 3 个方面探讨未来的研究发展.

2 学习共同体的定义描述

学习共同体是源于教育学领域的问题, 但是随着在线教育的深入发展, 它与计算机科学的联系变得越来越紧密. 实际上, 在计算机科学领域, 学习共同体被当作是如何根据学习者在学习过程中的属性特征划分学习小组或构建学习者团队等, 这与分组学生 (grouping student)、学习团队形成 (learning team formation) 等问题有相似之处, 因此有些研究会将它们相互转化和比较. 本节首先介绍教育学和计算机科学中学习共同体的定义, 然后结合这两学科领域的研究进展与需求, 给出动态学习共同体的形式化定义.

2.1 教育学中学习共同体的定义

在教育学领域, 目前较为广泛认可的学习共同体定义是: 指一个由学习者及其助学者 (包括教师、专家和辅导者等) 共同构成的团体, 他们彼此之间经常在学习过程中进行沟通、交流, 分享各种学习资源, 共同完成一定的学习任务, 因而在成员之间形成了相互影响、相互促进的人际联系^[22]. 该定义的内涵包括 3 个方面: (1) 学习者和助学者两个要素; (2) 强调学习过程中学习者之间的沟通交流、资源共享与相互协作; (3) 共同的学习任务.

学习共同体实例在教育学领域是很常见的, 如几个学习者组成学习组共同地学习数据挖掘课程或几个拥有数据库、C++程序设计、软件工程等技能的学习者一起开发一个图书管理系统而形成的团队等. 事实上, 学习者是学习共同体的核心要素, 而学习者之间的互动与协作是学习共同体的灵魂, 也是学习者完成共同学习任务或目标的基础. 基于上述定义和相关例子的描述, 在线教育环境中学习共同体构建过程如图 2 所示. 在图 2 中, 我们提取在线学习过程中学习者的学习记录、学习偏好等行为特征, 然后设计策略或算法 (如聚类算法、元启发式算法等) 构建学习共同体, 从而促进学习者与学习者、学习者与助学者之间相互交流、资源共享、相互协作等以完成学习任务. 随后, 学习共同体的群体学习模式会在在线教育系统中不断地改善学习者的知识水平、认知能力、沟通能力等, 并更新学习共同体中学习者的知识结构和知识状态, 从而改变整个学习共同体的结构状态. 因此, 我们需要重新构建学习共同体以满足在线教育过程中学习者之间教育资源的共享与传递、协作学习等行为. 在在线教育环境中, 由于助学者与学习者、学习者与学习者都处于时空分离的状态, 因此在线学习空间中的学习共同体也被称为“在线学习共同体 (online learning community)”^[21]. 在本综述中, 我们不区分学习共同体与在线学习共同体之间的不同.

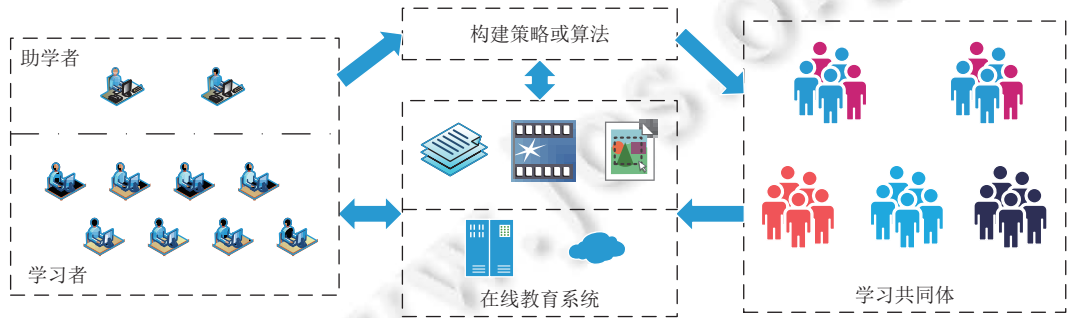


图 2 在线教育环境中学习共同体构建过程示意图

2.2 计算机科学中学习共同体的定义

根据教育学中学习共同体定义的描述, 在计算机科学中对学习共同体的形式化定义尚未有统一的标准. 目前, 有些研究从分组学生^[4-7]和学习团队形成^[28]两个角度对学习群体进行形式化定义, 目的是让学习者以群体学习

形式共同完成学习任务或目标, 从而提高学习者的学习效果. 本节介绍这两类问题的定义作为学习共同体的形式化定义.

● 分组学生

将学生划分为不同学习小组, 使学生之间便于互动与相互合作来完成学习任务, 然后最大限度地从同伴学习中获得收益. Agrawal 等人^[4]给出分组学生问题的正式定义, 并首次从计算的角度解决该问题.

假设给定 n 个学生 $X = \{X_1, X_2, \dots, X_n\}$ 学习特定的课程, 每个学生 X_i 具有相应的学习能力 $\theta_i \in [0, 1]$ (用项目反应理论等技术来确定), 学习能力按照递减排序 $\theta_1 > \theta_2 > \dots > \theta_n$. 定义一个学生分组 $T \subseteq X$ 的群体能力 $\hat{\Theta}_T$ 表示该组为其成员提供的收益, 学生通过参与组 T 活动获取平均收益, 而组 T 的群体能力 $\hat{\Theta}_T$ 是其成员的平均能力水平, 即 $\hat{\Theta}_T = 1/|T| \sum_{X_i \in T} \theta_i$. 一个学生分组 T 根据群体能力 $\hat{\Theta}_T$ 将该组 T 划分为两个集合: 领导者 (leaders) 和追随者 (followers), 领导者集合 $L_T = \{X_i \in T | \theta_i \geq \hat{\Theta}_T\}$ 是组 T 中学习能力在 $\hat{\Theta}_T$ 之上的成员; 追随者集合 $F_T = \{X_i \in T | \theta_i < \hat{\Theta}_T\}$ 是组 T 中学习能力在 $\hat{\Theta}_T$ 之下的成员. 用函数 $A_f(X_i, T)$ 和 $A_l(X_j, T)$ 分别表示追随者 X_i 和领导者 X_j 通过参与组 T 的增益, 则组 T 整体的增益定义为其所有成员的增益之和, 由公式 (1) 计算:

$$A(T) = \sum_{X_i \in F_T} A_f(X_i, T) + \sum_{X_j \in L_T} A_l(X_j, T) \quad (1)$$

追随者能够通过与更有能力的同伴互动和合作来提高自己的能力; 领导者则通过教导他人和给予帮助来从中获益. 然而, 文献 [4] 只关注追随者的增益, 则公式 (1) 修改为公式 (2).

$$A(T) = \sum_{X_i \in F_T} A_f(X_i, T) \quad (2)$$

Agrawal 等人^[4]描述的分组学生包括: 1-GROUP 和 ℓ -GROUP 两种问题. 这两种问题的定义描述如下.

1-GROUP 问题: 给定 n 个学生 $X = \{X_1, X_2, \dots, X_n\}$, 确定最多由 k 个学生组成的组 $T \subseteq X$, 使得该组 T 的收益 $A(T)$ 最大化.

ℓ -GROUP 问题: 给定一个整数 k 和 n 个学生 $X = \{X_1, X_2, \dots, X_n\}$ (满足 $n = k\ell$) 找出 X 的分组 $T_1, T_2, \dots, T_\ell$, 其中每个组的大小为 k , 使得所有组的收益之和 $\sum_{i=1}^{\ell} A(T_i)$ 最大化.

Agrawal 等人在文献 [5] 中进一步对 ℓ -GROUP 问题进行研究, 使学生可以通过与组内其他的学习伙伴互动来提高他们的收益, 从而使学习小组的整体收益最大化. 除了以学生的学习能力构建学习共同体外, 其他的研究利用学习者的学习风格、知识水平、学习行为等特征来构建学习共同体.

● 学习团队形成

在学习过程中根据学习者的学习特征组建学习者团队以覆盖学习任务或目标的需求, 从而促进学习者之间的协作与互动. Liu 等人^[28]通过建模学习者的技能掌握程度给出了学习团队形成的形式化定义.

假设有 n 个学习者 $X = \{X_1, X_2, \dots, X_n\}$ 学习同一个课程 (如数学), 用 $S = \{S_1, S_2, \dots, S_V\}$ 表示课程中 V 个技能集合. 学习者 X_i 将与技能熟练度的向量 $\alpha_i = \{\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{iV}\}$ 相关联, 其中 α_{ij} 表示学习者 X_i 对 S_j 的熟练程度, α_{ij} 值的范围是 0 到 1, 而所有学习者在所有技能上的熟练程度概述为 $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_V\}$. 因此, 根据熟练程度 α 形成学习团队, 令 $G = \{G_1, G_2, \dots, G_M\}$ 表示为 M 个学习团队集合. 形成学习团队是需要满足两个约束: (1) 保证每个学习者只属于一个学习团队中; (2) 学习团队规模最好是平等的, 规模差异不能超过一个学习者以保证学习团队的平衡. 则约束条件的形式化表示如公式 (3) 所示:

$$G_{k_1} \cap G_{k_2} = \emptyset, -1 \leq |G_{k_1}| - |G_{k_2}| \leq 1, k_1, k_2 = 1, 2, \dots, M, k_1 \neq k_2 \quad (3)$$

通过公式 (3) 保证 n 个学习者 X 均衡地分配到 M 个学习团队集合 G 中. 根据上述约束定义, Liu 等人^[28]通过基于不同的目标与基于增益的目标设计有效的学习团队形成算法, 以最大限度提高每个学习者在技能集合 S 中的熟练度. 例如, 基于不同的目标, 根据学习者对不同技能熟悉度的差异计算所有学习团队的异质性平均值作为集合 G 的异质性; 然后再考虑学习团队之间的平衡, 即学习团队之间异质性的差异. 最后, 考虑异质性和平衡性, 利用聚类算法形成学习团队. 基于增益的目标, 则是受到文献 [4] 启发, 首先选出每个学习团队中每项技能的领导者, 其

次所有跟随者根据技能增益加入相应的学习团队,然后定义增益函数并根据技能的熟练程度来衡量学习者的提升.最后计算团队集合 G 的平均增益和平衡增益,并提出一种考虑学习者收益和团队平均水平的算法,以形成平衡的学习团队.

2.3 动态学习共同体的形式化定义

根据图 2 的描述,我们知道在学习过程中学习共同体的结构状态与学习者的知识状态和知识结构是相互影响、动态变化的.因此,基于分组学生和学习团队形成的形式化定义,以及文献 [7] 预先定义完成任务的轮数动态分组学生,我们设定学习过程中多个阶段性的学习任务或目标以提取学习者的知识状态,然后给出计算机科学中动态学习共同体的形式化定义.

假设给定 n 个学习者 $X = \{X_1, X_2, \dots, X_n\}$ 在在线教育过程中共同学习某个课程,该课程包含有 L 个连续且不同的学习任务或目标.当学习者 X_i 开始学习该课程时,他在学习过程中的学习能力特征为 θ_i 、学习偏好特征(包括由点击知识点、浏览资源、知识结构等形成的向量)为 F_{X_i} .我们给定一个整数 K ,目标是设计算法或模型基于 n 个学习者的学习能力特征 θ_i 和学习偏好特征 F_{X_i} 形成 K 个规模大致相同的学习共同体 $C = \{C_1, C_2, \dots, C_K\}$,并满足 $C_i \cap C_j = \emptyset$, $-1 \leq |C_i| - |C_j| \leq 1$, $1 \leq i \neq j \leq K$.在学习共同体 C_k 中,我们既保证学习者在学习偏好特征上是同质的,又保证学习者之间的学习能力具有差异化(即将学习者之间的学习能力差距尽量最大化)以促进学习者之间相互学习,从而保证学习共同体中学习者的学习收益.然后,我们通过余弦相似性来计算学习者之间学习偏好的相似度 $\cos(F_{X_i}, F_{X_j}) = \frac{F_{X_i} \cdot F_{X_j}}{|F_{X_i}| \cdot |F_{X_j}|}$;而学习者之间学习能力的差值为: $\text{diff}(X_i, X_j) = \theta_j - \theta_i$,其中学习者 X_j 的学习能力 θ_j 高于学习者 X_i 的学习能力 θ_i .所有学习共同体 C 在 L 个连续不同的学习任务或目标中学习收益表示如下:

$$\text{gain}(C)_{T_i} = \sum_{k=1}^K \sum_{X_i, X_j \in C_k} (\lambda + \eta) \times \text{diff}(X_i, X_j) \quad (4)$$

$$\max \sum_{j=1}^L \text{gain}(C)_{T_j} \quad (5)$$

其中,公式 (4) 中的 λ 和 η 为 0 到 1 之间的值,分别表示学习者 X_j 和 X_i 从学习能力差中获得收益率,是完成第 T_i 个学习任务或目标以后所有学习共同体的学习收益之和.当完成学习任务或目标 T_i 后,我们需要更新学习者的学习能力和学习偏好,然后重新构建 K 个学习共同体以完成学习任务或目标 T_{i+1} .公式 (5) 表示完成整个课程 L 个学习任务或目标后的总体学习收益.在我们的定义中,我们需要设定相应的算法或模型构建学习共同体优化公式 (5) 以保证在学习过程中所有学习者的整体收益最大.

2.4 小结

在教育领域,学习共同体根据不同的学习需求,它被界定为一种学习环境、一种学习方式、一个学习团队或一个完成学习目标的生态系统等.在本综述中,我们将学习共同体视为学习过程中由学习者形成的学习小组或学习团队,以作为学习者之间协作学习、资源共享等行为的基础.在计算机科学领域,通过对学习者在学习过程中属性特征的量化,从分组学生和学习团队形成两个角度描述学习共同体的形式化定义来保证学习者的学习收益.文献 [6] 根据学习者的学习能力与亲和力两个特征将分组学生定义转化为目标约束问题,这是学习共同体形式化定义的延伸.文献 [7] 预先定义完成学习任务的轮数,然后基于每轮的学习任务动态分组学生以获取最大学习收益,这使静态学习共同体向动态学习共同体转变.基于现有学习共同体定义的理解,我们给出计算机科学领域中动态学习共同体的形式化定义.无论是在教育学领域还是计算机科学领域,学习共同体的核心思想是将学习者聚在“同一学习空间”,从而保证学习者之间能够相互协作、分享资源以完成共同的学习任务或目标,促进学习者的学习效率与学习绩效的提升.事实上,学习共同体的研究任务分为两部分,一是如何设计算法或模型构建学习共同体;二是如何对学习共同体进行管理以维持其有效地运作与稳定.

3 学习共同体的构建方法

构建学习共同体的目的是形成学习者群体基础,让学习者之间能进行良好地互动、资源共享以及协作学习等,提高学习者的学习效率,从而使学习者的学习绩效最大化.学习共同体构建方法需要基于学习者的属性特征形

成学习者群体,其核心问题是如何准确高效地构建学习共同体。目前,学习共同体的构建方法有4种:(1)随机构建;(2)让学习者自己选择学习伙伴;(3)教师基于设定的标准构建;(4)基于算法的构建。与第4种方式相比,前3种构建方法都存在一定的不足,如随机构建无法保证学习共同体的属性平衡;让学习者自己选择则要求保证良好的社会关系;而教师构建方式无法处理大规模学习者以及复杂分组标准^[25,26]。随着大数据、数据挖掘与机器学习等技术的发展,通过算法处理学习者的属性特征来构建学习共同体进入了快速发展时期,并在在线教育环境中显得也尤为重要。现有研究工作将学习共同体类型分为:同质、异质和混合3种类型^[29],本节对3种类型学习共同体的构建方法进行总结,并讨论每种类型构建方法中2-3个典型研究工作。

假设 n 和 k 分别表示学习者的个数和构建的学习共同体个数,然后我们从以下几个方面对不同类型的学习共同体构建方法进行分析与总结。

- (1) 构建不同类型学习共同体的具体算法有哪些?
- (2) 学习共同体成员规模多大时更有助于促进学习者之间的互动与协作?
- (3) 学习者哪些属性特征是构建学习共同体的关键因素?
- (4) 在学习共同体构建过程,哪些测量指标被用于判断学习者属性特征之间的距离或相似性?
- (5) 现有学习共同体构建方法所适用的应用场景描述是什么?

3.1 同质学习共同体的构建方法

同质学习共同体是将具有共同特征的学习者划分到同一学习共同体,以便于学习者之间资源共享、相互合作等,从而实现更加高效的学习。例如,将数据挖掘课程中学习分类、聚类、决策树等不同任务的学习者分别划分成不同的小组进行学习,这些学习组就是同质类型的学习共同体。表2列出了一些同质学习共同体构建方法的研究工作,这些研究从优化算法和聚类算法的角度将学习者聚集在同一学习空间而形成学习共同体。

表2 同质学习共同体构建的相关研究工作

文献	具体算法	规模大小	属性特征	测量指标	应用场景
[3]	K-means算法	[8, 15]	投票事件数量, Wiki页面编辑数量等14个特征	欧氏距离	在线学习环境TRAC
[4]	优化算法	n 或 n/k	学习能力	学习收益	学生能力测试系统
[5]	Percentile Partitions算法	n/k	学习能力	学习收益	
[30]	谱聚类算法	n/k	学习行为	相似度量	远程教育系统 SkyClass
[31]	K-means算法	$[1, n-k]$	学习者模型	欧氏距离	个性化在线学习系统
[32]	模糊聚类算法	$[1, n-k]$	学习者模型	相似系数	
[33]	基于矩阵的聚类算法	$[1, n-k]$	学习者模型	欧氏距离	
[34]	深度神经网络与K-means算法	n/k	学习力度	欧氏距离	智能导学系统 ASSISTments
[35]	稀疏自编码器与K-means算法	$[1, n-k]$	学习者间的无向权重图	欧氏距离	MOOCs平台
[36]	图注意自编码器	$[1, n-k]$	学习者间的属性图	—	
[37]	元路径与谱聚类算法	$[1, n-k]$	学习者间的异构关系图	相似度量	
[38]	图嵌入与谱聚类算法	$[1, n-k]$	学习者间的多类型关系图	相似度量	

由于同质学习共同体中的学习者具有一项或多项的共同属性特征,从而使不同学习空间中的学习者便于交流与互动,提高他们的学习效率;也使助学者识别学习者属于不同特征的学习共同体,便于助学者对不同的学习共同体因材施教。典型的研究如Agrawal等人^[4]通过量化学习者在在学习过程中的学习能力并进行由高到低排序,然后利用count-based和value-based两种增益函数对1-GROUP与 ℓ -GROUP两类学习共同体问题进行建模(具体定义见第2.2节)。事实上, count-based增益函数通过计算学习共同体中追随者的数量,学习者通过与学习能力强的学习者合作来提高自己的收益。然而,这个函数的局限是所有的追随者通过参与学习共同体都能获得相同的收益。value-based增益函数既考虑追随者能力的实际值,又关注学习共同体的集体能力,认为能力距离集体能力 $\hat{\Theta}_T$ 远的学习者通过参与学习共同体活动能获得更多的收益。Agrawal等人提出多个优化算法对两类学习共同体问题进行

优化, 例如通过 L&F 算法优化 count-based 增益函数与 1-GROUP 问题所形成的学习共同体, L&F 算法的伪代码如算法 1 所示. L&F 算法例子描述: 给定由 5 个学习者的能力值为 $\theta_1 = 0.9$ 、 $\theta_2 = 0.8$ 、 $\theta_3 = 0.6$ 、 $\theta_4 = 0.5$ 、 $\theta_5 = 0$, 通过算法考虑哪 3 个学习者组成的学习共同体收益最大. 如果学习者 1、2、3 组成学习共同体时, 他们的集体能力是 0.767, 只有学习者 3 才能从合作中收益, 这显然不合适. 当学习者 2、3、4 组成学习共同体时他们的集体能力是 0.633, 学习者 3 和 4 都可以互动中获得收益, 此时学习共同体的整体收益才最大.

算法 1. L&F Algorithm.

输入: Set of Learners $L = \{1, 2, \dots, n\}$ with sorted abilities $\theta_1 > \theta_2 > \dots > \theta_n$, Learning community size k ;

输出: Learning community T with the maximum $Ac(T)$.

```

1:  $T = \emptyset$ 
2: for  $i=1, \dots, k$  do
3:    $L = \text{top-}i$  ability learners
4:   for  $j = i+1, \dots, n-(k-i)+1$  do
5:      $F = \text{learners with abilities } \theta_j, \theta_{j+1}, \dots, \theta_{j+k-i-1}$ 
6: if  $F$  and  $L$  satisfy feasibility constraint then return  $T = L \cup F$ 

```

基于增益函数, 学习者在学习共同体中可以与同伴互动合作提高自己的收益^[39]. 随后, Agrawal 等人^[5]根据 ℓ -GROUP 问题将 n 个学习者设定为 k 个不同的学习共同体, 通过对目标函数进行优化使课程中 n 个学习者的学习收益最大. Agrawal 等人^[5]提出 p -Percentile 划分问题设定学习共同体, 其目标函数与增益函数的定义如下:

$$\max_G \sum_{g \in G} \Gamma^p(g), \Gamma^p(g) = \sum_{i=1}^{|g|} (p_i^g - \theta_i^g) \quad (6)$$

其中, 增益函数中 θ_i^g 与 p_i^g 分别表示学习共同体 g 中第 i 个学习者的学习能力与高于学习者 i 的学习能力在 p -Percentile 的学习者, p -Percentile 的指数值设定为: $\theta_{\lceil n \times p / 100 \rceil}$. 然后 Agrawal 等人提出 Percentile_Partitions 算法优化 k 个学习共同体, 使所有学习共同体中学习者的学习能力都达到 p -Percentile 以上.

聚类算法^[3,30-33]也是构建同质学习共同体的有效方法, 具有应用广泛、速度快、鲁棒性强等特点. 此外, K-means 算法^[3,31]处理大规模数据集时具有可伸缩性与高效性的优点, 模糊聚类算法^[32]则需要依赖于 K-means 算法, 基于矩阵的聚类算法^[33]不需要设定聚类簇、距离或规则的相似度更容易定义, 谱聚类算法^[30]具有计算复杂度小、能收敛到全局最优解和能有效聚类稀疏数据等特点. 典型的研究如, Xie 等人^[30]首先分析学习系统中学习者的行为主体、行为结果、行为被触发的时间以及主题关注的材料等 4 个基本组件, 并定义学习者的学习行为、行为序列等. 然后从数据模型、数据形状描述、结构相似性、微观相似性与混合相似性计算等测量学习者行为序列的相似性. 其中数据模型用于理解数据的特征与含义; 数据形状描述用于直观展现数据, 以便监管学习者的数据来源; 结构相似性表示学习者的行为序列具有相似结构特征; 微观相似性表示数据形状中片段的变化趋势与斜率变化的特征信息; 混合相似性是根据结构特征和微观特征来计算得到.

$$S_{\text{final}}(L_1, L_2) = \lambda S_{\text{structural}}(L_1, L_2) + (1 - \lambda) S_{\text{micro}}(L_1, L_2) \quad (7)$$

公式 (7) 中 L_1 和 L_2 表示两个学习者的行为序列轨迹, λ 是一个用于调整结构特征与微观特征比例的参数. 最后, 将行为序列视为图的顶点、行为序列比较视为图的边而相似性视为权值, 构建一个加权图 $G = (V, E, W)$, Xie 等人^[30]采用鲁棒性强、计算复杂度小的谱聚类算法将行为序列聚类任务转为图分割任务, 分割加权无向图 G 为 k 个不同子图, 即形成 k 个不同的学习共同体. 该算法主要步骤描述如下.

(1) 根据学习者的行为序列构建加权无向图 $G = (V, E, W)$ 和相应的非负对称相似矩阵 SM , 其中 V 为顶点集合、 E 为连接边集合、 W 为边的权重集合.

(2) 计算图拉普拉斯矩阵 $L = D - N$, D 是 G 的度矩阵而 N 是 G 的邻接矩阵.

(3) 计算 top- k 个特征值和特征向量, 形成一个特征矩阵.

(4) 通过 K-means 找到最佳的 k 个分割点.

目前, 深度学习、人工智能等技术用于构建同质学习共同体的研究较少, 但是一些研究工作已经开始探索. 如文献 [34,35] 采用堆栈自编码器构建深度神经网络对学习者的 (用户) 的多样化特征进行去噪与降维, 然后运行 K-means 算法构建学习共同体. 文献 [36] 基于学习者 (用户) 之间的属性与关系构建为属性图, 提出注意力自编码器最小化图重构损失来学习学习者 (用户) 的潜在嵌入表示, 然后利用自训练聚类模块基于潜在学习表示进行聚类, 从而形成聚簇 (学习共同体). 文献 [37,38] 将学习者 (用户) 与其他类型实体之间不同的类型关系构建异构信息网络或多关系图, 并利用深度学习技术对图嵌入表示, 然后通过谱聚类算法构建学习共同体. 如 Li 等人 [37] 利用元路径 (meta-path) 度量两个学习者 (用户) x_u 和 x_v 之间的相似性, 计算方式如下所示:

$$Sp(x_u, x_v) = \frac{2 \times \{p_{x_u \rightsquigarrow x_v} : p_{x_u \rightsquigarrow x_v} \vdash P\}}{\{p_{x_u \rightsquigarrow x_u} : p_{x_u \rightsquigarrow x_u} \vdash P\} + \{p_{x_v \rightsquigarrow x_v} : p_{x_v \rightsquigarrow x_v} \vdash P\}} \quad (8)$$

公式 (8) 中 $p_{x_u \rightsquigarrow x_v}$ 表示 x_u 和 x_v 之间的路径实例, $p_{x_u \rightsquigarrow x_v} \vdash P$ 表示 x_u 和 x_v 之间在元路径 P 上所有的路径实例. 给定一个元路径集合 PS , 则每条元路径 $P_i \in PS$ 基于公式 (8) 得到相似性矩阵 Sp_i , 然后构建所有学习者 (用户) 形成的矩阵 W 作为他们之间的权重加权和, 如下所示:

$$W = \sum_{i=1}^{|PS|} \lambda_i Sp_i \quad (9)$$

公式 (9) 的 λ_i 表示元路径权重向量, Li 等人 [37] 的目标就是通过 W 提供的信息推导出一个高质量的相似矩阵 S 显示清晰的聚类结构, 然后通过求解目标函数 $\Gamma(S, \lambda)$ 优化问题确定 S 和 λ 的值, 如下所示:

$$\begin{cases} \min \Gamma(S, \lambda) = \|S - W\|_F^2 + \alpha \|S\|_F^2 + \beta \|\lambda\|_2^2 \\ \text{s.t. } \sum_{j=1}^n S_{ij} = 1; S_{ij} \geq 0; \sum_{i=1}^{|PS|} \lambda_i = 1; \lambda_i \geq 0; \text{rank}(L_S) = n - k \end{cases} \quad (10)$$

公式 (10) 包含 3 项, 第 1 项是用于学习 S 接近 W , 后两项是正则项表示用于防止过拟合. 前 4 个约束项归一化 S 和元路径的权重值, 并防止它们为 0; 最后一个约束项是强制要求 S 中 k 个连通分量. 随后的目标是从一组基于元路径相似矩阵 W 中学习具有 k 个连通分量的相似矩阵 S .

同质学习共同体的构建以学习者的属性特征作为基础, 使学习者之间存在天然的亲密感, 不会因个别学习者的某个特征不同而不能融入学习共同体, 影响学习者之间的协作学习和互动沟通. 事实上, 构建同质学习共同体的相关算法仍存在一些不足, 如文献 [4,5] 中的优化算法虽然能简单、高效地通过学习者的特征划分为不同的学习共同体, 但面临大规模且多样化特征的在线学习者时计算代价大. 在聚类算法中, K-means 算法的不足是聚类好坏依赖初始聚类中心的选择、对异常样本数据较为敏感、 K 值的选择确定困难等; 基于矩阵的聚类算法改进基于矩阵的层次算法 [33], 其不足是计算复杂高、聚类结果依赖于合并点和分裂点的选择等; 谱聚类算法的不足是对相似图的变化与聚类参数的选择非常敏感等. 而深度学习、人工智能等技术还处于探索阶段, 它们仍需借助聚类算法完成. 此外, 构建同质学习共同体还需考虑公平性问题. (1) 学习积极性高低不同的学习者会形成各自不同的学习共同体, 前者的学习积极性高、活跃性好会有效地促进学习效果提升, 而后者中学习者的互动交流少导致学习效果差, 因此在学习过程中构建的学习共同体很容易出现两极分化及“富人俱乐部现象”[40], 从而使学习过程中出现不公平问题. (2) 研究表明学习共同体的规模大小影响学习者之间的互动频率及质量[41], 而聚类算法难以控制聚簇的规模大小, 导致构建的学习共同体之间的规模不统一, 使学习共同体之间难以实现公平学习.

3.2 异质学习共同体的构建方法

为避免同质学习共同体的问题, 研究者提出构建异质学习共同体. 异质学习共同体: 是将特征不相同的学习者划分到同一学习共同体中, 使学习者完成共同目标或任务时能够相互补充, 通过促进学习共同体内部互动交流来提高协作学习效果. 例如要完成一项图书管理系统的开发任务, 给定 3 个学习者 A、B 和 C 分别拥有不同的知识技能, 如数据库、C++程序设计、软件工程, 则他们组成的开发团队就是异质学习共同体. 表 3 列出了一些异质学习共同体构建方法的相关研究工作.

表 3 异质学习共同体构建的相关研究工作

文献	具体算法	规模大小	属性特征	测量指标	应用场景
[2]	动态分组策略算法	3	概念学习成功、失败和部分学习所形成的知识结构	学习成绩水平	课程的在线协作学习过程
[6]	目标约束优化算法	n/k	学习潜力和亲和力	技能提升	
[42]	遗传算法	3	知识水平、沟通技巧和领导能力	最优适应度值	智能协作学习系统
[43]	数学模型与蚁群优化算法	4	主题兴趣、个性特征、学业表现和自信度等	欧氏距离和异质性度	
[44]	动态分组与遗传算法	n/k	概念学习成功、失败和部分学习所形成的知识结构	最优适应度值	
[45]	遗传算法	5	英语听力、口语、阅读的成绩	最优适应度值	
[46]	遗传算法	4	多元智能、学习风格、领导能力和自信力	最优适应度值	课程的在线协作学习过程
[47]	基于动态分组策略的教学优化 (TLBO) 算法	5	教学过程中的知识分配与师生之间的交互行为	优化函数	
[48]	遗传算法	55或56	知识结构与社会关系的互补度	最优适应度值	基于Web的在线学习系统

构建异质学习共同体的核心旨在确保学习者之间的属性特征尽可能地多而不同,使学习者之间形成优势互补,从而促进他们互动交流、资源共享等来完成学习任务或目标.如 Esfandiari 等人^[6]将基于学习者的学习潜力和亲和力两个特征的异质学习共同体构建问题转化为目标约束优化问题,从而提出目标约束优化算法优化学习者的异质划分.给定 n 个学习者 $W = \{w_1, w_2, \dots, w_n\}$, 他们拥有的技能值表示为 w_i^s , 而优化目标则是将 n 个学习者优化成 k 个学习共同体 $G = \{g_1, g_2, \dots, g_k\}$, 同时保证学习潜力和亲和力两个目标最优,形式化如下:

$$\max_G \sum_{i=1}^k LP(g_i), \sum_{i=1}^k Aff(g_i) \text{ s.t. } |G| = k, |g_i| = n/k \tag{11}$$

其中,学习潜力 LP 表示为最熟练与最不熟练学习者之间的技能差值 (LpD) 或每个学习者与更高技能学习者之间的技能差值和 (LpA),例如在一个学习共同体 $g_1 = \{2, 3, 1, 5\}$ 中,学习潜力 $LP(g_i) = 4$ 或 13.亲和力 Aff 反映学习者之间基于社会人口统计特征的协作效率,而学习共同体亲和力是所有学习者之间最小的 ($AffD$) 或指定学习者与所有学习者之间最小的 ($AffC$).由于 LP 和 Aff 是两个互不相容的目标,目标优化的复杂性在于亲和力与学习共同体的规模,因此多目标优化问题转为约束学习潜力优化亲和力的约束优化问题,形式化表示如下:

$$\max_G \sum_{i=1}^k Aff(g_i) \text{ s.t. } \sum_i LP(g_i) > OptLP(g_i) \quad |G| = k, |g_i| = n/k \tag{12}$$

其中, $OptLP$ 表示学习潜力的最优值.基于公式 (12),Esfandiari 等人^[6]提出 ($AffC LpD$)、($AffC LpA$)、($AffD LpD$)、($AffD LpA$) 这 4 种学习共同体算法来约束学习者之间的学习潜力并优化学习者之间的亲和力,从而优化学习者在不同学习共同体中的异质划分以及保证学习者之间的公平性.

在线教育环境中,构建异质学习共同体需要处理学习者多样化的属性特征,而元启发式算法通过编码机制能处理多个变量并且能快速找到问题的最优解,因此它成为构建异质学习共同体广泛使用的算法,其包括遗传算法^[42,44-46,48]、蚁群优化算法^[43]、教学优化算法^[47]等.这些元启发式算法是一个通用的算法框架或黑盒优化器,几乎可以用于所有优化问题.如 Zou 等人^[47]通过教学优化算法 (teaching learning-based optimization algorithm, TLBO) 模拟教师教授知识给所有学习者和学习者之间通过相互帮助共同学习的学习过程,然后引入动态分组策略 (dynamic group strategy, DGS),提出一种基于动态分组策略的教学优化 (DGSTLBO) 算法构建异质学习共同体用于改善学习过程的学习效率.在 DGSTLBO 算法中学习者在教师阶段,分别从教师和自己所在的学习共同体中学习,则学习者 X 的知识更新 $newX$ 表示如下:

$$newX = X + r \times (Teacher - TF \times LCmean), TF = round(1 + rand[0, 1]) \tag{13}$$

其中, $Teacher$ 表示最优的知识水平, 也是所有学习者要达到的水平, $LCmean$ 表示学习者 X 所在学习共同体的平均知识水平, TF 是决定 $LCmean$ 变化的因素, r 是 0 和 1 之间的随机数. 在学习阶段, DGSTLBO 算法设置随机学习策略和量子行为学习策略进行学习, 学习者 X 的量子行为学习策略更新如下表示:

$$tempX = \alpha \times LCteacher + (1 - \alpha) \times Teacher \quad (14)$$

$$\begin{cases} newX = tempX + \beta |LCmean - X| \times \ln(1/u), & k < 0.5 \\ newX = tempX - \beta |LCmean - X| \times \ln(1/u), & k \geq 0.5 \end{cases} \quad (15)$$

其中, $LCteacher$ 表示学习共同体中学习能力最优的学习者, $tempX$ 表示学习 X 从教师和学习共同体中获得知识值. 公式 (15) 由 u 和 k (值为 0 到 1 之间) 控制动态地改变学习共同体中学习者的知识水平, 促进学习者之间进行知识交互. 总之, DGSTLBO 算法在教学阶段使用学习共同体的学习均值而不再是传统的班级; 在学习阶段, 通过设定随机学习策略或量子行为学习策略进行互动学习. 最后, DGSTLBO 算法设置初始化参数和随机生成初始学习共同体, 通过不断迭代和优化构建不同的异质学习共同体.

由于学习者的知识状态会随着学习深入而发生改变, 因此学习共同体的状态也需要动态地变化才能更有效地提高学习者的学习效果, 否则学习共同体中容易出现搭便车现象、角色固化以及合作倦怠等问题. 一些研究工作^[2,44,48]以学习者的知识结构、社会关系等作为互补, 动态地构建不同学习阶段的学习共同体. 典型研究如 Jong 等人^[2]通过概念图 (由概念点及其先序关系组成) 和序列概率比测试 (sequential probability ratio test, SPRT) 设计概念图诊断系统对不同学习阶段中学习者所学习概念的情况 (成功、失败、部分成功) 进行诊断, 生成学习者知识结构的概念图. 然后计算学习者的辅助分数来测量他们之间知识结构的异质程度 (表示某一个学习者在协作学习活动中帮助另一个学习者的程度). 最后设定动态调整阈值将学习过程分设为多个阶段, 基于每个阶段结束后学习者的互补分数提出动态分组策略算法, 从而在不同学习阶段动态地构建不同的异质学习共同体, 以提高学习者的学习成绩水平. 随后, Chan 等人^[44]基于文献^[2]的研究, 采用遗传算法和动态分组策略对学习者的知识结构和概念权重进行编码, 以知识之间的互补作为最优适应度值, 通过不断地迭代和设置策略 (选择、交叉或变异等) 直到算法终止或产生有效的近似最优解, 从而动态优化在不同学习阶段中的异质学习共同体.

文献^[6]目标约束优化算法通过约束优化学习者的两个属性特征能够实现异构学习共同体的构建以及学习者的公平划分, 当面对学习者多个的属性特征时, 可以将约束目标转为带约束的目标优化问题, 从而能高效快速地实现学习共同体的构建. 然而, 由于目标约束优化问题是 NP-hard 问题, 当面临大规模在线学习者的多样化特征时, 约束优化算法同时处理学习共同体规模、学习者属性特征会导致很高的复杂度与计算代价. 以蚁群优化算法、遗传算法等为代表的元启发式算法所构建的异质学习共同体规模相对较小, 能基于学习者的属性特征实现他们之间的互补, 从而有效促进学习者之间的交流互动、提高学习者的合作学习效率. 然而, 由于元启发式算法是依赖于问题的算法^[49], 在构建异质学习共同体时也存在一些不足: (1) 算法对初始化种群有一定的依赖性, 需要根据学习者的特征设计合理的初始化编码, 以保证形成的学习共同体快速收敛; (2) 缺乏有效的动态调节策略, 算法很容易陷入局部最优, 从而导致所构建的学习共同体收敛精度低或无法收敛; (3) 搜索空间、参数设置等都会影响算法的优化性能. 除此之外, 构建异质学习共同体还需要考虑以下 3 个问题: 一是设置动态改变学习共同体的合理阈值, 如果设置的阈值太小则学习共同体调整过于频繁而导致学习效果不佳, 如果设置的阈值太大则不能及时地改变学习共同体的学习伙伴, 从而让学习者感到疲惫或无聊^[50]. 二是学习共同体中学习者属性特征之间的平衡性, 如果学习共同体中学习者之间的属性特征平衡性较差, 则无法保证学习者之间优势互补、相互协作等, 从而导致学习共同体难以凝聚学习者而无法发挥群体智能. 三是现有的构建方法缺乏与深度学习、人工智能等技术相结合, 无法提取大规模在线学习者属性特征的潜在嵌入表示信息.

3.3 混合学习共同体的构建方法

同质和异质学习共同体的构建方法都是将 n 个学习者构建成为 k 个相似或相异的学习共同体, 然而所构建的学习共同体要么缺乏公平性, 要么平衡性不足. 与这两种类型学习共同体不同, 混合学习共同体既要保证学习者的有些特征是相似的, 又保证学习者的有些特征是不同的, 以便于学习者之间相处融洽且相互协作. 假设要完成一个

分布式图书管理系统, 现有 5 个学习者 {A, B, C, D, E}, 他们分别拥有的技能是: A={分布式系统, 数据库}, B={数据库, C++程序设计}, C={Web 编程, 算法设计}, D={软件工程, 算法设计}, E={软件项目管理}. 事实上, 只有他们 5 个组成学习共同体才能完成任务, 有些技能相似便于他们沟通交流, 而不同的技能则有助于他们合作与互补, 这就是混合类型学习共同体. 混合学习共同体既避免了出现同质学习共同体的两极分化, 又提高了异质学习共同体因学习者之间的差异而缺乏的亲和力, 从而保证学习者之间的公平性与平衡性. 混合学习共同体构建的研究相对较少, 表 4 列出了几个混合学习共同体构建方法的研究成果.

表 4 混合学习共同体构建的相关研究工作

文献	具体算法	规模大小	属性特征	测量指标	应用场景
[51]	模糊C-均值算法	$[2, 2\ln(n)]$	学习风格、知识水平、学习目标、兴趣爱好、性别和活跃度	划分系数和划分熵	课程的在线协作学习过程
[52]	粒子群优化算法和遗传算法	n/k	性别、认知水平、学习风格、学习时间兴趣爱好和学习偏好等	公平性、互补性和亲密性	
[53]	改进的遗传算法	3	编程技巧和优先选择同伴	平均值和标准偏差	cMOOCs
[54]	粒子群优化算法	5	知识水平和学习兴趣	最优适应度值	
[55]	增强粒子群优化算法	$[[n/k], [n/k]]$	主题的理解水平和兴趣	最优适应度值	基于网络的协作学习系统

构建混合学习共同体的思想是从学习者自身多项特征出发, 基于学习任务或目标的需求, 既保证有些学习者是同质的, 促进相互合作与交流; 又要保证学习者之间有异质性, 促进他们的互动与协作、形成优势互补, 从而便于完成任务. 典型研究如 Zheng 等人^[53]将在线教育环境中的学习共同体构建问题表述为一个完整的数学模型, 该模型首先融合同质和异质学习共同体的特征形成混合学习共同体的特征 (F_{mixed}), 然后在结合学习共同体的平衡性 (F_{balan}), 学习者对同伴的选择偏好 (F_{pref}) 以及学习者的聚集或分布 (F_{dist}) 等特征. 为满足在线教育环境中混合学习共同体构建需求, 根据这些特征标准建模为数学模型如下所示.

$$\min F = \frac{w_{\text{intra}} F_{\text{mixed}} + w_{\text{inter}} F_{\text{balan}} + w_{\text{pref}} F_{\text{pref}} + w_{\text{dist}} F_{\text{dist}}}{w_{\text{intra}} + w_{\text{inter}} + w_{\text{pref}} + w_{\text{dist}}} \quad (16)$$

其中, 公式 (16) 约束每个学习者只在一个学习共同体且任何两个学习共同体中人数差不能超过一个, 公式中每个 w 项表示相应特征的权重信息. 为求得该数学模型的最优解, Zheng 等人提出一种改进的遗传算法, 该算法结合随机全域抽样选择与基于适应度的重插入实现精英策略, 从而构建最优学习共同体, 以满足不同教育背景下学习共同体的需求.

Lin 等人^[55]以学习者对给定主题的理解水平和兴趣作为构建学习共同体的标准, 并根据两个标准考虑任意两个学习共同体中学习者对每个主题理解水平的差异、两个学习共同体之间每个主题理解水平的最大差异以及每个学习共同体对每个主题的兴趣差异最大等内容. 为构建最优混合学习共同体, 两个约束条件被要求: 一是每个学习者只能参与一个学习共同体, 二是学习共同体中人数尽可能相等或差异不超过一个. 从问题的描述可知, 这是一个 NP-hard 问题, 因此 Lin 等人提出一种增强粒子群优化算法用于对混合学习共同体的构建优化. 为满足两个约束条件, 提出了一个粒子编码规则:

$$P_y = [p_{11} p_{12} \dots p_{21} p_{22} \dots p_{k1} \dots p_{kx}] \quad (17)$$

其中, P_y 表示第 y 个粒子, p_{kx} 表示第 k 个学习共同体中第 x 个学习者. 为解决上述问题和约束, 一个适应度函数被设计用来作为增强粒子群优化算法的优化目标, 如下所示:

$$\min Z(P_y) = f + C_1 + C_2 \quad (18)$$

$$f = \frac{\sum_{l=1}^t \left\{ \max_{1 \leq j \leq k} \left\{ \sum_{x=1}^c w_l U_{xjl} \right\} - \min_{1 \leq j \leq k} \left\{ \sum_{x=1}^c w_l U_{xjl} \right\} \right\}}{ct}, \quad 0 < f < 1 \quad (19)$$

$$C_1 = \frac{\max_{1 \leq l \leq t} \left\{ \max_{1 \leq j \leq k} \left\{ \sum_{x=1}^c w_l U_{xjl} \right\} - \min_{1 \leq j \leq k} \left\{ \sum_{x=1}^c w_l U_{xjl} \right\} \right\}}{c}, 0 < C_1 < 1 \quad (20)$$

$$C_2 = \max_{1 \leq l \leq t} \left\{ 1 - \frac{\sum_{j=1}^k \min \left\{ 1, \sum_{x=1}^c L_{xjl} \right\}}{k} \right\}, 0 < C_2 < 1 \quad (21)$$

公式 (18) 是优化混合学习共同体构建的适应度函数, 它由公式 (19)–(21) 这 3 个函数组成. 其中公式 (19) 表示对于 t 个主题理解水平最高与理解水平最低的学习共同体之间的平均差异, 每个学习者对 t 个主题的理解度由开始前测试成绩推断. 公式 (20) 表示任意两个学习共同体对每个主题的理解水平的最大差异. 公式 (21) 表示计算学习共同体中每个学习者感兴趣的平均数量, 每个学习者对主题的兴趣是基于学习者的学习经验推断而来. 其中 t 是主题数量、 k 是学习共同体的数量、 c 是学习共同体的规模, U_{xjl} 和 L_{xjl} 分别表示第 j 个学习共同体中第 x 个学习者对第 l 个主题的理解水平与学习兴趣. 通过增强粒子群优化算法和适应度函数优化混合学习共同体的构建, 使学习共同体对每个主题有相同的理解水平, 同时至少有一个学习者对每个主题感兴趣, 从而促进学习者之间的交流和合作帮助教师规划不同的协作学习任务.

混合学习共同体的构建方法仍依赖于聚类算法与元启发式算法, 故不再赘述其优势与不足. 虽然混合学习共同体要通过学习者之间的不同属性特征来维持学习过程的公平性和平衡性, 保证了学习共同体中学习者之间的互动与协作, 但现有工作仍有一些问题: (1) 混合学习共同体构建的特征选择. 由于学习者不同的属性特征决定学习者不同状态, 尤其是动态特征具有实时变化的特性会影响混合学习共同体的状态, 因此设计动态构建算法是非常必要的. (2) 构建算法存在弊端. 聚类算法与元启发式算法的弊端已经表明, 需要从目标约束角度^[4-6]设计新算法构建混合学习共同体. (3) 缺乏与人工智能的结合, 当处理大规模在线学习者的多样化特征时捉襟见肘.

4 学习共同体的管理机制

在在线教育环境中, 学习共同体作为一种重要的交互式学习群体, 它能够帮助学习者克服在线教育场景中时空分离的影响, 改善在线教育所面临的问题. 然而, 随着学习任务或目标的不断深入, 如何保障整个学习共同体的稳定性与活跃性和学习者的学习积极性是一个非常关键的问题. 本节从共享、协作、激励等内容讨论学习共同体管理机制的研究进展, 以保障学习共同体的有效运作及稳定.

4.1 学习共同体的共享机制

学习共同体中学习者之间的共享是维持其运作的关键因素, 共享内容包括教育资源与知识, 教育资源是指教学过程中所使用的课程材料、视频、参考资料等; 知识则是对教育资源的认知、理解与判断等. 然而, 由于海量的教育数据具有大数据“5V”的特点, 使得学习者在搜索与学习海量教育资源的过程中引发“资源过载和信息迷航”的问题, 导致“教育资源越来越多、知识学习越来越困难”的尴尬局面. 因此, 在在线教育环境中, 实现学习共同体的共享机制显得尤为重要.

学习共同体的教育资源共享是学习者在学习过程中互动交流的保障, 也是知识共享与传播的基础. 知识共享是个体与个体、个体与群体之间进行的知识交换过程, 它已被认为是一个组织生存的积极力量, 并在组织、群体或个人之间传递或传播知识, 从而产生新想法或知识^[56]. 近年来, 有关学习共同体知识共享的研究主要运用社会学、心理学、系统学等理论, 其中以社会资本理论^[57]、社会认知理论^[58,59]、社会交换理论^[60,61]等社会学理论的应用最为广泛. 这些工作通过结构资本 (个体之间存在的结构联系)、关系资本 (个体具有理解和应用知识的认知能力)、认知资本 (个体的关系具有强烈、积极的特点)、自我效能、结果预期与个体互动等维度研究学习共同体中学习者知识共享的影响因素, 促进学习共同体中知识的双向交流、反复反馈、相互调整, 从而将知识从个体层面认知上升到群体层面认知, 以保证学习共同体的生存与发展.

面对一群学习风格、学习偏好、学习行为等特征相似或相异的学习者所构成学习共同体, 向学习共同体共享合适的学习课程或视频等资源是实现教育资源共享与知识传播的关键, 也是解决“资源过载和信息迷航”问题的重

要措施. 群组推荐 (group recommendation) 通过群体活动向群组推荐合适项目或资源, 它将推荐对象从单个用户扩展为一个群组^[62], 为群组决策过程提供强有力的建议以满足用户在群体活动中的需求^[63]. 在学习过程中, 通过群组推荐向学习共同体 (群组) 推荐合适的教育资源 (项目) 是实现在线教育中资源共享的重要机制. 群组推荐的核心问题是如何在学习共同体 (群组) 决策过程中聚合成员 (学习者) 之间的偏好. 从技术角度来说, 现有的群组推荐工作可分为两类: 基于 Memory 和基于 Model 的方法. 基于 Memory 方法是采用预定义策略来聚合群组 (学习共同体) 内成员 (学习者) 的偏好, 其进一步可分为偏好聚合 (preferences aggregation)^[64]和分数聚合 (score aggregation)^[65-67]. 偏好聚合首先聚合群组内每个成员的偏好成为群组偏好, 然后利用个体推荐技术对群组进行推荐; 分数聚合首先通过个体推荐获得群组内每个成员的预测分数, 然后利用预定义聚合策略 (如平均满意度、最小痛苦、最大满足度) 对群组内所有成员分数进行聚合以获得群组的分数. 然而, 由于每个学习者在学习共同体中的活跃性与积极性不同, 他们对学习共同体选择教育资源的贡献度也不同, 因此基于 Memory 方法不能很好地模拟学习共同体的决策过程. 与基于 Memory 方法不同, 基于 Model 方法则尝试对群组 (学习共同体) 的决策过程进行建模, 用于融合群组中不同的成员 (学习者) 对项目 (教育资源) 的偏好. 近年来, 基于 Model 方法利用注意力机制作为动态调整群组的聚合策略, 以捕获群组决策的复杂过程, 主要包括 AGREE (attentive group recommendation) 模型^[68]、MoSAN (medley of sub-attention networks) 模型^[69]、Sigr (social influence-based group recommender) 模型^[70]、GroupSA (group self-attention) 模型^[71]、KGAG (knowledge graph-based attentive group recommendation)^[72]模型. 在本节, 我们将基于 Model 方法的群组推荐作为学习共同体资源共享的机制, 以实现在学习共同体中教育资源共享, 并详细讨论几个典型的研究工作. 基于 Model 方法实现学习共同体中教育资源共享的群组推荐过程如图 3 所示.

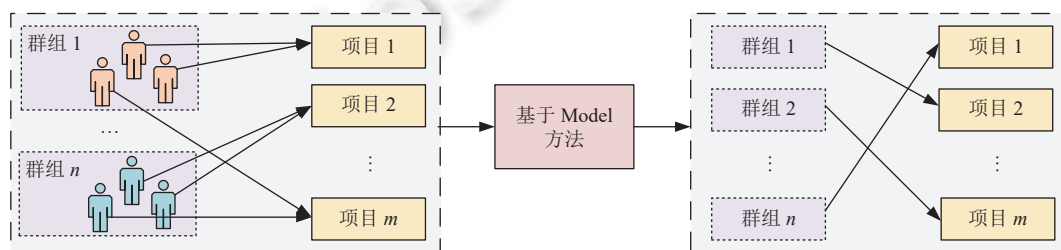


图 3 基于 Model 方法的群组推荐过程

在在线教育环境中, 学习共同体实际上是基于某些学习任务或行为特征而形成的临时群组, 学习者在学习共同体中彼此之间可能不熟悉且产生的影响也不同. 另一方面, 学习共同体与教育资源在学习过程中的可用交互是有限的, 因此在学习共同体中实现教育资源共享面临着学习者偏好聚合与数据稀疏两个问题. 为解决这些问题, 一些典型的研究工作被提出. 如 Cao 等人^[68]基于注意力网络和神经协同过滤的最新进展, 提出一种注意力群组推荐 (AGREE) 模型, 该模型能从历史数据中学习聚合策略来解决学习者的偏好聚合问题. 在 AGREE 中, 注意力网络能够在学习共同体与不同教育资源交互时为学习者分配不同的权重信息, 并动态调整学习共同体的聚合策略以建模决策过程; 神经协同过滤框架则捕获学习者与教育资源和学习共同体与教育资源的交互用于提高推荐的性能. 然而, 这项工作既没有考虑学习共同体与教育资源交互数据的稀疏性问题, 也没有考虑学习者在学习共同体中的社交影响力.

为解决上述问题, Guo 等人^[71]提出一种群组自我注意 (GroupSA) 模型建模学习共同体的决策过程, 该模型利用基于社交的自我注意聚合策略解决学习共同体内学习者的偏好聚合. 在 GroupSA 中, 学习共同体的决策过程被建模为一个多重投票过程, 并通过一个栈式社交自我注意网络模拟学习者的投票权重以达成整体共识; 对于数据稀疏性问题, 通过聚合学习者与学习者、学习者与教育资源的历史数据来丰富学习共同体与教育资源的特征表示. Deng 等人^[72]将知识图谱作为边信息引入群组推荐, 提出一种基于知识图谱的注意力群组推荐 (KGAG) 模型获得学习共同体的知识感知来解决数据稀疏性和偏好聚合问题. KGAG 模型包括信息传播模块、偏好聚合模块以及优化模块 3 部分, 信息传播模块获取知识图谱的信息传播以及聚合学习者与教育资源之间的高阶表示; 随后, 这

些高阶表示输入到偏好聚合模块, 通过注意力机制捕获学习者在学习共同体中对教育资源的自我坚持与同伴评分, 结合这两部分内容作为他在学习共同体中对教育资源的重要性, 并计算出学习共同体对教育资源的偏好; 最后, 优化模块对学习共同体对教育资源的偏好信息进行优化, 实现推荐合适的教育资源到学习共同体中. 文献 [71,72] 通过社交网络和知识图谱捕获学习者、学习共同体与教育资源之间丰富的特征信息来缓解数据稀疏性问题, 从而实现学习共同体与教育资源之间的高效共享. 通过注意力机制聚合学习共同体中所有学习者偏好的一般过程如下所示, 令 u_i 表示学习者在学习共同体 lc_k 中的嵌入表示、 v_j 表示教育资源的嵌入表示.

$$\tilde{\sigma}(u_i, v_j) = h^T \text{ReLU}(P_u u_i + P_v v_j + b) \quad (22)$$

$$\sigma(u_i, v_j) = \frac{\exp(\tilde{\sigma}(u_i, v_j))}{\sum_{u_i \in lc_k} \exp(\tilde{\sigma}(u_i, v_j))} \quad (23)$$

$$\eta(lc_k, v_j) = \sum_{u_i \in lc_k} \sigma(u_i, v_j) \times x_{u_i}^{v_j} \quad (24)$$

公式 (22) 表示教育资源 v_j 对学习共同体中学习者 u_i 的重要性, 公式 (23) 则表示学习共同体中学习者 u_i 对教育资源 v_j 重要性的归一化表示, 公式 (24) 表示学习共同体 lc_k 对教育资源 v_j 的整体偏好, 其中 $x_{u_i}^{v_j}$ 表示学习者 u_i 对教育资源 v_j 的偏好程度. 除上述研究外, 一些研究基于学习共同体的特征, 然后设计群组推荐系统来实现教育资源在学习共同体中共享. 如 Rodríguez 等人 [73] 根据 VARK 模型建模学习者的学习风格, 然后提出自适应群组推荐的多智能体系统. 该系统能够融合所有学习者的学习风格形成学习共同体, 并向学习共同体推荐合适的教育资源. Papamitsiou 等人提出一种基于非合作博弈论的教育资源群组推荐方法, 通过将推荐策略建模为纳什均衡问题, 即学习者无法独自偏离纳什均衡而改变自己的满意度, 从而优化所有学习者的偏好以及解决他们之间的利益冲突, 并在协作学习环境中实现学习共同体中教育资源共享 [74].

尽管通过群组推荐能够缓解学习者在学习过程中所遇到的“资源过载和信息迷航”问题, 并在学习共同体中实现教育资源的共享, 但是现有研究仍存在一些不足. (1) 忽视了学习者在不同学习共同体的影响. 在在线教育环境中, 学习者可能在多个学习共同体中学习, 现有工作没有考虑学习者在不同学习共同体中对不同教育资源的影响力, 没有挖掘不同教育资源之间的相关性以及学习者之间的潜在关联, 不能准确计算学习者对教育资源的重要性, 从而影响学习共同体的教育资源共享效果. (2) 忽视了学习共同体的动态变化. 由于学习共同体是临时形成的学习群组, 学习者在学习过程中存在不确定性且不受监督, 因此学习者可能会随时加入或离开学习共同体, 从而影响学习共同体对教育资源的最终决策, 使得向学习共同体共享的教育资源不是最佳的. 事实上, 资源共享研究的核心挑战是缺乏基于学习共同体动态演化的教育资源共享机制. 在知识共享方面, 虽然社会学理论有效地挖掘出了影响知识交换与共享的关键因素, 促进了学习共同体中学习者之间的知识共享和传播, 但是也存在一些问题, 如忽视学习共同体的知识共享是一个反复反馈、不断调整的过程, 现有研究只适合于某个阶段而无法满足长期动态演化的需求; 缺乏算法或模型量化学习者的知识共享行为. 由于在线教育环境存在诸多不确定性, 所以需要准确把握学习者的共享行为与意图以保证学习共同体的知识共享的稳定性和持久性.

4.2 学习共同体的协作机制

协作学习 (collaborative learning) 是教学过程中运用群体学习模式使学习者互动工作, 并以最大限度地促进自己及他人的学习 [75]. 事实上, 学习共同体与协作学习之间是相辅相成的, 学习共同体为师生之间、学习者之间的互动与协作提供群体基础与意识; 协作学习为学习共同体提供了一种学习模式, 它通过学习者之间的协作活动不断地优化学习共同体的状态. 与个体学习相比, 协作学习则强调学习共同体中学习者之间的互动与合作, 它跨越个体认知并向群体认知转变, 是一个复杂的学习过程.

学习活动是以学习者为中心且以学习任务或目标为驱动的协作式社会活动. 在在线教育环境中, 社区问答 (论坛) 系统可视为学习共同体, 问答协作则是一种促进学习者之间交互协作与提高学习者参与度的协作机制. 基于在线教育的问答系统, 学习者与助学者、学习者与学习者通过相互的问答, 能够高效地解决所遇到的学习问题或完成学习任务. 一些研究工作从教育问答系统 [76,77] 中学习者与助学者之间的协作问答, 不断地完善在线教育问答系

统, 并提高学习共同体中成员之间协作和信息分享的效率. 另一些研究工作为了能及时高效地解决在线问答的问题, 提出问题路由 (question routing) 机制^[78-83], 它将问题推送给那些可能回答的专家 (学习者或助学者), 从而提高学习共同体中成员之间的协作效率与问题回答的准确率. 总体来说, 问题路由是学习共同体中一种主动的协作策略, 它通过学习者或助学者历史回答记录或兴趣与当前问题构建联系, 使当前问题能及时高效地被回答, 促进学习共同体中学习者的协作交互, 提高学习过程的学习效率与质量. 本节根据问题路由的相关典型研究来介绍学习共同体的协作机制, 并讨论其不足.

虽然在线教育问答系统为学习者提供了一个解决问题的平台, 学习者可以根据自己的问题寻求问答, 但是学习者不能在短时间内获得问答结果, 并且还会陷入漫长等待. 这使得问题发布者和潜在答案回答者之间存在严重的差距, 降低了在线教育问答系统的性能以及影响学习者的学习积极性. 为了提高在线教育环境中学习共同体的问答效率, 问题路由研究被提出用于及时地将问题推动可能回答的专家. Zhou 等人^[79]定义问题路由: 给定一个问题及基于学习共同体的问答, 将问题路由视为一个分类问题, 判断学习者或助学者是否会基于他们的知识回答该问题. 然后构造问题、学习者 (助学者) 以及问题-学习者 (助学者) 的局部特征与全局特征, 并通过支持向量机预测能够回答问题的学习者 (助学者) 的权重值, 从而将问题推送给权重高的学习者 (助学者) 来提高学习共同体中成员之间的协作效率与问答质量. 然而, 该研究侧重于评估学习者 (助学者) 的权威级别以推荐他们回答问题, 忽视了问题发布者背景与兴趣爱好而无法满足不同学习者的个性化需求.

为实现学习共同体中不同学习者对问题路由的个性化需求, Li 等人^[80]将在线社区问答系统中的提问者、问题与回答者构成一个异构信息网络, 提出一个网络嵌入增强排名模型 (NeRank) 用于解决学习共同体中个性化问题路由. 该模型通过元路径嵌入算法和 LSTM (long short-term memory) 分别提取提问者与回答者的潜在表示和学习问题的内容表示, 这种实体训练表示既保留实体之间的邻近信息, 又保留提出者的特定知识背景与偏好信息. 最后, NeRank 将建模问题路由任务为排序问题, 并建模提问者、回答者、问题 3 个实体之间的相关性, 利用卷积神经网络计算回答者与给定查询 (提问者, 问题内容) 的排序得分, 根据排序得分将问题推送给合适的回答者, 从而提高在线社区中学习者之间问答协作的效率与质量.

由于文献 [78-81] 都是将问题推送给较为合适的学习者或助学者, 因此很容易导致这些学习者或助学者出现问答负载, 同时也导致其他的学习者参与度低而不利于学习过程的协作, 因此文献 [82,83] 考虑将问题路由给 top-k 个学习者或助学者. 如 Zhou 等人^[82]提出一个问题路由框架, 该框架首先根据学习共同体中历史问题的回答情况来计算学习者或助学者成为当前问题专家的概率; 然后利用论坛系统中学习者之间的结构关系, 通过概率对学习者的专业知识进行重新排名; 最后将这两个结果聚合到一个概率模型中计算所有学习者的最终排名, 并将问题推送给最有潜力的 top-k 个学习者, 促进学习者之间的协作效率和学习者的参与度.

虽然通过在线教育的问答系统和问题路由能够有效地促进学习共同体中成员之间的协作, 使得学习者能够高效地解决学习过程遇到的问题, 但是现有研究工作仍有不足. (1) 缺乏对所提出问题的词义与语义解析来解决问答的冷启动. 如果提出的问题是新问题, 以往的学习者或助学者没有任何回答记录, 则现有研究是无法匹配到合适的回答者. 事实上, 通过自然语言处理技术对问题进行词义与语义解析, 然后捕获历史回答相似或相近问题的回答者 (学习者或助学者), 并将问题推动给其中的 top-k 个回答者, 这样可能有效地提高问答协作的效率、减少等待反馈的时间. (2) 问答协作缺乏特定的知识库. 现有研究都是通过问题与学习者的历史回答数据进行关联匹配, 面对大规模在线教育环境的问答很显然代价会非常高, 因此现有研究缺乏面向特定领域的知识库快速匹配常见问题并反馈回答.

4.3 学习共同体的激励机制

在学习共同体中, 既面临学习者活跃度不足而导致协作与共享不积极, 又面临学习者随意地共享资源、问答互动等而导致信息过载与质量低下, 因此规范学习者的行为非常重要. 激励机制是一种以期望的方式影响或改善个体行为的策略^[84], 它在 P2P 网络^[85]、无线网络^[86]等方向已经存在大量的研究工作. 由于动机是人类一切行为活动的出发点与驱动力^[87], 因此学习共同体的激励机制也需要强调学习者的学习动机与驱动力, 目的是提高学习共

共同体中学习者的参与度、知识共享、协作互动以及维持学习共同体的活跃度等. 现有学习共同体的激励机制研究主要是通过增加学习者的虚拟积分或声誉值^[88-91]、计算学习者(用户)排名^[92-95]等策略鼓励学习者积极参与学习共同体活动以贡献学习资源、传递知识、提高协作等. 在本节, 我们将针对学习共同体激励机制的典型研究工作进行讨论与分析.

Cheng 等人^[89]提出一种自适应的激励机制, 该机制的基本思想是根据学习者的声誉与贡献质量和学习共同体的需求对学习者的参与行为进行奖励, 从而促进学习共同体的活跃度与激励学习者贡献高质量的学习资源. 自适应激励机制包括学习共同体模型和个体模型. 学习共同体模型描述当前学习阶段中学习共同体的实际需求, 包括所有学习者当前主题(Q_C)贡献的资源数量和学习共同体激励因子(F_C), Q_C 由学习共同体管理者根据主题的特征和学习者设定, F_C 反映新加资源对学习共同体的有用程度. 个体模型包括学习者的资源声誉(C_I)和学习者的评分声誉(R_I), C_I 表示学习者共享所有资源的平均值, R_I 表示学习者所有评分之间平均差值的倒数, 其公式表示如下:

$$R_I = \frac{N}{\sum_{i=1}^N |r_i - \bar{r}|} \quad (25)$$

其中, r_i 是学习者对资源的评分, $\bar{r}$ 是各个资源的平均评分. 学习者的贡献资源数量 Q_I 是 Q_C 的一部分, 学习者具有较高的资源声誉 C_I 则获得 Q_I 就越大, 这意味着具有良好资源声誉的学习者会做出更多贡献. Q_I 由公式(26)所示.

$$Q_I \approx Q_C \cdot C_I \left/ \sum_{i=1}^n C_I \right. \quad (26)$$

其中, n 表示学习者的数量.

综上所述, 自适应的激励机制通过考虑学习共同体的 Q_C 与 F_C 和个体模型的 C_I 、 R_I 与 Q_I 等设定学习者分享和评级资源的自适应权重, 然后计算分享和评级资源的虚拟积分作为奖励, 虚拟积分越高则表示学习者的声誉、关注度越高, 从而提高学习者的积极性和维持学习共同体的活跃度.

文献^[92-95]通过提出算法或模型对学习共同体(或学习论坛)中学习者(用户)在问答过程中提供的答案进行计算与排名, 以激励学习者(用户)贡献高质量的答案, 从而提高他们的学习绩效. 如 Shi 等人^[92]提出一个学习者(用户)反馈增加的重排名模型, 用于改善学习者(用户)体验和改变收集信息时他们的反馈. Narang 等人^[95]提出多关系图卷积模型对学习共同体(学习论坛)问答系统中学习者(用户)所提供的问题答案进行排名, 从而激励学习者(用户)贡献高质量的答案. 典型的研究: Cai 等人^[93]提出 ExpertRank 算法衡量学习者的专业知识并对他们进行排序, 该算法计算学习者 u_i 在第 $k+1$ 回答时的专业排名如下所示:

$$ER_{k+1}(u_i) = \sum_{j=1}^{|Q(u_i)|} QA_Quality(Q_j(u_i), u_i) + \alpha \sum_{j=1}^{|Q(u_i)|} ER_k(Q_j(u_i)) \quad (27)$$

其中, $|Q(u_i)|$ 表示学习者 u_i 在学习共同体中回答问题的数量, $ER_k(Q_j(u_i))$ 是学习者 u_i 在 k 次回答中的专业知识排名, $ER_0(u_i) = 0$. $QA_Quality(Q_j(u_i), u_i)$ 表示捕获学习者 u_i 回答第 j 个问题(学习者 u_j 的问题)的质量分, 它通过不同学习者的投票分数和答案排名质量分来描述学习者回答问题的质量. Cai 等人通过公式(27)对学习共同体中学习者问答行为进行排序, 激励学习者成为学习共同体中的知识专家.

Omari 等人^[94]在问题与候选答案之间的相关性基础上, 关注候选答案的多样性与新颖性, 在学习共同体(学习社区)的问答系统中提出一种答案排序方案来提升学习者(用户)的答案排名. 该方案先选择命题(传递学习者回答问题的连贯信息), 然后通过命题过滤将与问题无关的命题过滤掉, 并使用两个命题之间的相关性函数(包括 TF-IDF、Word2Vec、ESA 和 CQA-ESA)作为答案多样性和重要性的衡量标准. Omari 等人以 Word2Vec 为例, 计算命题之间的相似性 $sim(p, o)$ 如公式(29)所示.

$$Coverage(p, o) = \frac{1}{|p|} \sum_{t_p \in p} \max_{t_o \in o} [\cos(w2v(t_p), w2v(t_o))] \quad (28)$$

$$sim(p, o) = \sqrt{Coverage(p, o) \cdot Coverage(o, p)} \quad (29)$$

公式 (28) 中的 p 和 o 是命题, t_p 和 t_o 是 p 和 o 的项, $w2v(t)$ 是 t 项的 Word2Vec 表示. 该方案提出基于相似性的答案选择和基于层次聚类的答案选择的两种方法, 用于更新命题新颖性和选择下一个答案. 我们着重介绍基于相似性的答案选择方法, 该方法直接利用命题之间的相似性来选择答案. $Support(p, a)$ 表示基于答案与命题的相似度来衡量答案 a 支持命题中给出信息的支持度, 命题的新颖性评估由 $Novelty[p]$ 属性来维持, 根据它代表方面在更高排序答案中被其他命题所代表的程度. 答案 a 的分数 ($Score(a)$) 通过 a 对所有命题支持的总和决定, 并与它们的新颖性加权, 其计算方式如下所示:

$$Support(p, a) = 1 - \prod_{p_a \in a} (1 - sim(p, p_a)) \quad (30)$$

$$Score(a) = \sum_{p \in P} Novelty[p] \cdot Support(p, a) \quad (31)$$

$$Novelty[p] = Novelty[p] \cdot (1 - Support(p, selectedAnswer)) \quad (32)$$

公式 (30) 中 $Support(p, a)$ 是一种相似性表达, 只有当所有答案的命题 p 零相似时它才是 0, 并且当 a 命题至少一个与 p 相似时才支持 p . 公式 (31) 的 $Score(a)$ 只考虑未排序答案中命题对答案分数的贡献. 当选择更高分数的答案 $selectedAnswer$ 时, 每个命题的 $Novelty[p]$ 将由公式 (32) 更新. 最后, 根据学习者 (用户) 的答案所包含命题数量以及它们在所有答案中的重要性, 考虑它们与问题之间的相关性, 通过贪心算法对学习者的答案进行排序, 从而激励学习者 (用户) 贡献高质量的问题答案.

虽然关注学习者在学习共同体的个体排名能够激发学习者的积极性, 但是这种情况忽略了整个学习共同体的学习绩效、群体效能以及活跃度. Chang 等人^[91]提出一种群体激励机制, 该机制结合学习者的个体绩效和学习共同体中学习者之间的相互依赖计算学习共同体的激励分值 S_n , 如公式 (33) 所示.

$$S_n = 100 \times \frac{P_n}{T} - (T - P_n) \sum_{k=n-2}^{n-1} \frac{200R_k}{(n-k)T} \quad (33)$$

其中, P_n 是学习共同体中通过第 n 个学习阶段的学习者数量, T 表示是学习共同体中学习者的总数量, R_k 是未通过第 k 个学习阶段的学习者数量. 公式 (33) 前半部分表示学习者个体绩效对学习共同体激励分值的影响, 也是对学习者的自我激励; 后半部分表示学习共同体中学习者的学习进度是存在相互依赖的, 学习进度慢的学习者会受到惩罚, 从而影响整个学习共同体的排名. Chang 等人通过惩罚项来激励学习者之间相互合作完成学习任务以获得高的值 S_n , 从而使得学习共同体具有更高的群体效能和凝聚力.

学习共同体的激励机制在促进学习者积极参与、知识共享、互动以及维持学习共同体稳定性与持久性等起到重要作用. 然而, 学习共同体激励机制的研究仍存在一些不足: (1) 过于关注学习共同体中学习者的贡献排名, 缺乏互动协作行为的激励, 容易导致学习共同体中学习者的贡献排名两极分化. (2) 缺乏对学习者的资源评论、问题回答等内容自然语义处理与理解, 无法准确判断与学习共同体的主题或任务是否匹配, 故计算学习者声誉时容易存在偏差. (3) 缺乏整体到个体的激励策略, 应设计算法或模型计算学习共同体的整体行为, 并制定相应的激励策略, 然后基于对学习共同体的贡献设计分配策略奖励学习者, 从而促进学习者、学习共同体之间的竞争与合作.

5 学习共同体的研究展望

学习共同体的目的在于实现学习者之间的资源共享、互动交流、协作学习等, 调动学习者在学习过程中的自觉性与积极性, 解决在线教育环境面临辍学率高、课程完成率低、监管不足等问题. 然而, 面对复杂的在线教育环境, 现有的研究工作仍然处于起步阶段, 本节将从学习共同体的动态构建、多学习共同体的管理以及基于学习共同体的精准教学评价 3 个方面探讨学习共同体未来的研究.

5.1 学习共同体的动态构建

学习共同体是以学习者为中心的学习群体, 群体的成功发展是随时间序列不断演化^[96]. 已有一些研究工作通过概念图诊断系统生成学习者的知识结构来动态构建学习共同体, 从而提高学习者的学习成绩. 然而, 由于学习者在学习过程中具有一定的知识背景、学习方式、学习能力等差异, 因此他们的知识状态更新程度会有所不同, 从

而导致学习共同体演化的状态不同. 为保证学习共同体在在线学习过程中动态变化时的有效性、活跃性与可演化性, 研究学习共同体随学习过程不断演化的动态构建是一个极具挑战的问题. 学习共同体的动态构建存在以下挑战: (1) 如何设定学习共同体合理的规模约束 (如规模大小为 5 或 10) 以及学习者知识状态的阈值范围 (如 $[0.6, 0.9]$), 建模学习者在学习共同体中历史知识状态与当前学习行为的关联更新机制、设计函数或模型获得学习者知识的动态变化值, 实现学习者在学习共同体中的加入或退出以保证学习共同体状态的实时更新. (2) 为表达学习共同体结构的层次变化, 研究学习共同体中学习者的知识状态、交互关系等与实际任务需求变化 (如任务分割或合并) 的动态关联, 建模学习共同体与实际任务之间的需求机制, 实现当前学习共同体的结构演化 (如分割或合并) 以推动学习共同体的多样性与有效性发展, 是学习共同体结构演化的关键. (3) 如何与人工智能技术结合, 建模智能感知机制差异化捕获学习过程中学习者的知识更新度, 实现学习共同体的动态感知构建机制.

5.2 多学习共同体的管理

在学习过程中, 一个学习者根据不同的学习任务可能参与多个学习共同体, 并且这些学习共同体可能存在某些结构上的联系. 学习者作为学习共同体的核心要素, 是多个学习共同体之间资源共享、知识传播与协作学习等行为的纽带. 多学习共同体的管理既能帮助学者把握学习者在不同学习共同体中的学习状态、激励学习者之间互动、组织学习共同体之间的协作, 还能使学习者超越个体的认知而上升到群体认知层面实现自我价值, 因此多学习共同体之间的有效管理是一个值得深思的任务. 然而, 要实现多学习共同体的管理需要解决以下关键问题. (1) 设计学习者参与学习共同体的合理个数, 如何有效地建模学习者知识结构与学习共同体主题之间的关联机制, 并确定学习者在学习共同体中迁移的约束规则 (如学习知识的先后、任务的重要程度) 和参与学习共同体个数. (2) 在学习共同体的主题 (或任务) 演化时, 如何建模学习共同体之间主题 (或任务) 演化程度与关系, 实现遵循学习者之间知识结构和学习共同体之间主题依赖的资源流动、知识共享与协作学习. (3) 如何根据多学习共同体间的迁移约束规则与多个学习任务间的重要性和关联性, 建模学习者在多学习共同体中的迁移机制.

5.3 学习共同体的精准教学评价

教学评价一直是教育领域的研究重点, 它是整个教学过程中必不可少的重要环节, 也是呈现教学效果的有力手段. 传统教学评价方式是对教学过程的数据信息进行定量分析, 缺乏对整个教学过程多维度、细粒度、全面化、精准化的评价. 在在线教育环境中, 以学习共同体为中心的学习单位记录着授课内容、学习行为、社交行为、评论等信息, 因此基于学习共同体实现教学过程的精准评价是一个关键的任务. 通过学习共同体实现精准教学评价需要考虑以下内容: (1) 如何建模学习者在学习过程中贡献资源、问题评论、协作互动等行为的关联方法, 设计声誉机制评价学习者在学习共同体中的贡献值与排名, 保证学习者的积极性与活跃性. (2) 如何建模学习共同体中学习活动的动态迭代评价机制、设计众包激励机制、融合学习者的实时学习行为, 实现对学习共同体及其要素的动态评价与反馈, 这是一个极具挑战性的问题. (3) 如何设计学习者的声誉与学习共同体动态迭代评价之间的融合权重, 建模个体行为与群体特征的评价关联度, 提供自适应评价机制实现对教学过程全面且精准的评价, 并激励学习者和学习共同体不断地改善教学过程.

6 总 结

本文分析了挖掘教育大数据任务和解决在线教育面临辍学率高、课程完成率低、参与度低以及监管不足等问题的重要性, 指出了学习共同体解决这些问题的研究意义与价值, 给出了学习共同体未来的研究方向. 我们首先从计算机科学和教育学两个领域介绍了学习共同体的定义以及本文的动态定义, 总结了学习共同体的研究任务. 然后, 我们从构建算法、规模大小、属性特征、测量指标、应用场景等多方面对同质、异质和混合 3 种类型的学习共同体研究进行总结, 这 3 种学习共同体能够在课程学习、科研合作、系统开发等情景描述下适用. 虽然现有的研究工作在构建学习共同体方面取得了一定成果, 但是面对动态变化的在线教育环境, 学习共同体的构建方法仍有很大的探究空间. 在学习共同体的管理方面, 我们通过现有工作总结共享、协作、激励等机制研究促进学习者之间的资源共享、知识共享、协作互动、积极参与以及维持学习共同体的持久性与活跃性等. 此外, 我们指出

了现有研究工作仍存在的问题,以及缺乏多学习共同体管理的研究,无法发挥群体智能的优势.最后,根据现有研究在学习共同体构建与管理两个方向上存在的问题与不足,我们讨论了动态构建学习共同体在规模约束、阈值设定、结构层次变化以及与人工智能结合的问题和建模挑战;多学习共同体管理在迁移约束规则、主题演化、迁移机制的问题与建模挑战;基于学习共同体的精准教学评价在声誉评价机制、动态迭代评价和自适应全面评价机制的问题与关键技术,并简明地描述相应的研究内容,以便更好地推动学习共同体研究与发展.

尽管学习共同体在在线教育环境中能够有效地推动学习者之间的合作交流、资源共享等,但学习共同体的动态构建与多学习共同体的管理还处于萌芽阶段,没有形成完整的研究体系,相关研究工作仍需研究者不断地探究.本文工作是基于作者及其团队成员之间的理解与讨论完成的,未来还将在学习共同体方向上深入挖掘.

References:

- [1] Bakhshinategh B, Zaiane OR, ElAtia S, Ipperciel D. Educational data mining applications and tasks: A survey of the last 10 years. *Education and Information Technologies*, 2018, 23(1): 537–553. [doi: [10.1007/s10639-017-9616-z](https://doi.org/10.1007/s10639-017-9616-z)]
- [2] Jong BS, Wu LY, Chan TY. Dynamic grouping strategies based on a conceptual graph for cooperative learning. *IEEE Trans. on Knowledge and Data Engineering*, 2006, 18(6): 738–747. [doi: [10.1109/TKDE.2006.93](https://doi.org/10.1109/TKDE.2006.93)]
- [3] Perera D, Kay J, Koprinska I, Yacef K, Zaiane OR. Clustering and sequential pattern mining of online collaborative learning data. *IEEE Trans. on Knowledge and Data Engineering*, 2009, 21(6): 759–772. [doi: [10.1109/TKDE.2008.138](https://doi.org/10.1109/TKDE.2008.138)]
- [4] Agrawal R, Golshan B, Terzi E. Grouping students in educational settings. In: *Proc. of the 20th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. New York: Association for Computing Machinery, 2014. 1017–1026. [doi: [10.1145/2623330.2623748](https://doi.org/10.1145/2623330.2623748)]
- [5] Agrawal R, Nandanwar S, Musti MN. Grouping students for maximizing learning from peers. In: *Proc. of the 10th Int'l Conf. on Educational Data Mining*. Wuhan: EDM, 2017.
- [6] Esfandiari M, Wei D, Amer-Yahia S, Roy SB. Optimizing peer learning in online groups with affinities. In: *Proc. of the 25th ACM SIGKDD Int'l Conf. on Knowledge Discovery & Data Mining*. Anchorage: Association for Computing Machinery, 2019. 1216–1226. [doi: [10.1145/3292500.3330945](https://doi.org/10.1145/3292500.3330945)]
- [7] Wei D, Koutis I, Roy SB. Peer learning through targeted dynamic groups formation. In: *Proc. of the 37th IEEE Int'l Conf. on Data Engineering*. Chania: IEEE, 2021. 121–132. [doi: [10.1109/ICDE51399.2021.00018](https://doi.org/10.1109/ICDE51399.2021.00018)]
- [8] Pappano L. The year of the MOOC. *The New York Times*, 2012, 2(12): 1–7.
- [9] Zheng QH, Dong B, Qian BY, Tian F, Wei BF, Zhang WZ, Liu J. The state of the art and future tendency of smart education. *Journal of Computer Research and Development*, 2019, 56(1): 209–224 (in Chinese with English abstract). [doi: [10.7544/jssn1000-1239.2019.20180758](https://doi.org/10.7544/jssn1000-1239.2019.20180758)]
- [10] Shah D. A decade of MOOCs: A review of MOOC stats and trends in 2021. 2021. <https://www.classcentral.com/report/moocs-stats-and-trends-2021/>
- [11] Anderson A, Huttenlocher D, Kleinberg J, Leskovec J. Engaging with massive online courses. In: *Proc. of the 23rd Int'l Conf. on World Wide Web*. Seoul: Association for Computing Machinery, 2014. 687–698. [doi: [10.1145/2566486.2568042](https://doi.org/10.1145/2566486.2568042)]
- [12] Kizilcec RF, Piech C, Schneider E. Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In: *Proc. of the 3rd Int'l Conf. on Learning Analytics and Knowledge*. Leuven: Association for Computing Machinery, 2013. 170–179. [doi: [10.1145/2460296.2460330](https://doi.org/10.1145/2460296.2460330)]
- [13] Onah DFO, Sinclair J, Boyatt R. Dropout rates of massive open online courses: Behavioural patterns. In: *Proc. of the 6th Int'l Conf. on Education and New Learning Technologies*. Barcelona: IATED Academy, 2014. 5825–5834.
- [14] He JZ, Bailey J, Rubinstein BIP, Zhang R. Identifying at-risk students in massive open online courses. In: *Proc. of the 29th AAAI Conf. on Artificial Intelligence*. Austin: AAAI Press, 2015. 1749–1755.
- [15] Herbert M. Staying the course: A study in online student satisfaction and retention. *Online Journal of Distance Learning Administration*, 2006, 9(4): 300–317.
- [16] Prenkaj B, Velardi P, Stilo G, Distanto D, Faralli S. A survey of machine learning approaches for student dropout prediction in online courses. *ACM Computing Surveys*, 2020, 53(3): 57. [doi: [10.1145/3388792](https://doi.org/10.1145/3388792)]
- [17] Rovai AP. Building sense of community at a distance. *The Int'l Review of Research in Open and Distributed Learning*, 2002, 3(1): 1–16. [doi: [10.19173/irrodl.v3i1.79](https://doi.org/10.19173/irrodl.v3i1.79)]
- [18] Stanford-Bowers DE. Persistence in online classes: A study of perceptions among community college stakeholders. *Journal of Online Learning and Teaching*, 2008, 4(1): 37–50.

- [19] Zhou Q, Mou C, Yang D. Research progress on educational data mining: A survey. *Ruan Jian Xue Bao/Journal of Software*, 2015, 26(11): 3026–3042 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4887.htm> [doi: 10.13328/j.cnki.jos.004887]
- [20] Kulkarni C, Cambre J, Kotturi Y, Bernstein MS, Klemmer SR. Talkabout: Making distance matter with small groups in massive classes. In: *Proc. of the 18th ACM Conf. on Computer Supported Cooperative Work & Social Computing*. Vancouver: ACM, 2015. 1116–1128. [doi: 10.1145/2675133.2675166]
- [21] Liu IF, Chen MC, Sun YS, Wible D, Kuo CH. Extending the TAM model to explore the factors that affect intention to use an online learning community. *Computers & Education*, 2010, 54(2): 600–610. [doi: 10.1016/j.compedu.2009.09.009]
- [22] Hu Y, Li YF. The connotation and related researches reviews of virtual learning communities. *Modern Distance Education Research*, 2012, (3): 32–42 (in Chinese with English abstract). [doi: 10.3969/j.issn.1009-5195.2012.03.005]
- [23] Ji H, Zhu YF. Inheritance and innovation: From community to learning community. *Distance Education in China*, 2019, (10): 74–79 (in Chinese with English abstract). [doi: 10.13541/j.cnki.chinade.2019.10.007]
- [24] Stoll L, Bolam R, McMahon A, Wallace M, Thomas S. Professional learning communities: A review of the literature. *Journal of Educational Change*, 2006, 7(4): 221–258. [doi: 10.1007/s10833-006-0001-8]
- [25] Krouska A, Troussas C, Virvou M. Applying genetic algorithms for student grouping in collaborative learning: A synthetic literature review. *Intelligent Decision Technologies*, 2019, 13(4): 395–406. [doi: 10.3233/IDT-190184]
- [26] Costaguta R. Algorithms and machine learning techniques in collaborative group formation. In: *Proc. of the 14th Mexican Int'l Conf. on Artificial Intelligence*. Morelos: Springer, 2015. 249–258. [doi: 10.1007/978-3-319-27101-9_18]
- [27] Odo C, Masthoff J, Beacham N. Group formation for collaborative learning—A systematic literature review. In: *Proc. of the 20th Int'l Conf. on Artificial Intelligence in Education*. Chicago: Springer, 2019. 206–212. [doi: 10.1007/978-3-030-23207-8_39]
- [28] Liu YP, Liu Q, Wu RZ, Chen EH, Su Y, Chen ZG, Hu GP. Collaborative learning team formation: A cognitive modeling perspective. In: *Proc. of the 21st Int'l Conf. on Database Systems for Advanced Applications*. Dallas: Springer, 2016. 383–400. [doi: 10.1007/978-3-319-32049-6_24]
- [29] Gogoulou A, Gouli E, Boas G, Liakou E, Grigoriadou M. Forming homogeneous, heterogeneous and mixed groups of learners. In: *Proc. of the 2007 Workshop on Personalisation in E-learning Environments at Individual and Group Level, the 11th Int'l Conf. on User Modeling*. 2007. 33–40.
- [30] Xie T, Zheng QH, Zhang WZ. A behavioral sequence analyzing framework for grouping students in an e-learning system. *Knowledge-based Systems*, 2016, 111: 36–50. [doi: 10.1016/j.knsys.2016.08.001]
- [31] Du J, Zheng QH, Ding J, Gong ZY. A method for learner grouping based on personality clustering. In: *Proc. of the 10th Int'l Conf. on Computer Supported Cooperative Work in Design*. Nanjing: IEEE, 2006. 1–6. [doi: 10.1109/CSCWD.2006.253206]
- [32] Tian F, Wang SB, Zheng C, Zheng QH. Research on E-learner personality grouping based on fuzzy clustering analysis. In: *Proc. of the 12th Int'l Conf. on Computer Supported Cooperative Work in Design*. Xi'an: IEEE, 2008. 1035–1040. [doi: 10.1109/CSCWD.2008.4537122]
- [33] Zhang K, Cui LZ, Wang HY, Sui Q. An improvement of matrix-based clustering method for grouping learners in e-learning. In: *Proc. of the 11th Int'l Conf. on Computer Supported Cooperative Work in Design*. Melbourne: IEEE, 2007. 1010–1015. [doi: 10.1109/CSCWD.2007.4281577]
- [34] Zhang JT, Li B, Song W, Lin NZ, Yang XD, Peng ZY. Learning ability community for personalized knowledge tracing. In: *Proc. of the 14th Asia-Pacific Web (APWeb) and Web-age Information Management (WAIM) Joint Int'l Conf. on Web and Big Data*. Tianjin: Springer, 2020. 176–192. [doi: 10.1007/978-3-030-60290-1_14]
- [35] Tian F, Gao B, Cui Q, Chen EH, Liu TY. Learning deep representations for graph clustering. In: *Proc. of the 28th AAAI Conf. on Artificial Intelligence*. Québec City: AAAI, 2014. 1293–1299. [doi: 10.1609/aaai.v28i1.8916]
- [36] Wang C, Pan SR, Hu RQ, Long GD, Jiang J, Zhang CQ. Attributed graph clustering: A deep attentional embedding approach. In: *Proc. of the 28th Int'l Joint Conf. on Artificial Intelligence*. Macao: AAAI Press, 2019. 3670–3676. [doi: 10.24963/ijcai.2019/509]
- [37] Li X, Kao B, Ren ZC, Yin DW. Spectral clustering in heterogeneous information networks. In: *Proc. of the 33rd AAAI Conf. on Artificial Intelligence*. Honolulu: AAAI, 2019. 4221–4228. [doi: 10.1609/aaai.v33i01.33014221]
- [38] Sadikaj Y, Velaj Y, Behzadi S, Plant C. Spectral clustering of attributed multi-relational graphs. In: *Proc. of the 27th ACM SIGKDD Conf. on Knowledge Discovery & Data Mining*. Association for Computing Machinery, 2021. 1431–1440. [doi: 10.1145/3447548.3467381]
- [39] Webb NM. Peer interaction and learning in small groups. *Int'l Journal of Educational Research*, 1989, 13(1): 21–39. [doi: 10.1016/0883-0355(89)90014-1]
- [40] Vaquero LM, Cebrian M. The rich club phenomenon in the classroom. *Scientific Reports*, 2013, 3(1): 1174. [doi: 10.1038/srep01174]

- [41] Kim J. Influence of group size on students' participation in online discussion forums. *Computers & Education*, 2013, 62: 123–129. [doi: [10.1016/j.compedu.2012.10.025](https://doi.org/10.1016/j.compedu.2012.10.025)]
- [42] Moreno J, Ovalle DA, Vicari RM. A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics. *Computers & Education*, 2012, 58(1): 560–569. [doi: [10.1016/j.compedu.2011.09.011](https://doi.org/10.1016/j.compedu.2011.09.011)]
- [43] Graf S, Bekele R. Forming heterogeneous groups for intelligent collaborative learning systems with ant colony optimization. In: *Proc. of the 8th Int'l Conf. on Intelligent Tutoring Systems*. Jhongli: Springer, 2006. 217–226. [doi: [10.1007/11774303_22](https://doi.org/10.1007/11774303_22)]
- [44] Chan TY, Chen CM, Wu YL, Jong BS, Hsia YT, Lin TW. Applying the genetic encoded conceptual graph to grouping learning. *Expert Systems with Applications*, 2010, 37(6): 4103–4118. [doi: [10.1016/j.eswa.2009.11.014](https://doi.org/10.1016/j.eswa.2009.11.014)]
- [45] Wang YH, Li YC, Liao HC. Using a genetic algorithm to determine optimal complementary learning clusters for ESL in Taiwan. *Expert Systems with Applications*, 2011, 38(12): 14832–14837. [doi: [10.1016/j.eswa.2011.05.065](https://doi.org/10.1016/j.eswa.2011.05.065)]
- [46] Pinninghoff JMA, Contreras AR, Salcedo LP, Contreras AR. Genetic algorithms as a tool for structuring collaborative groups. *Natural Computing*, 2017, 16(2): 231–239. [doi: [10.1007/s11047-016-9574-1](https://doi.org/10.1007/s11047-016-9574-1)]
- [47] Zou F, Wang L, Hei XH, Chen DB, Yang DD. Teaching-learning-based optimization with dynamic group strategy for global optimization. *Information Sciences*, 2014, 273: 112–131. [doi: [10.1016/j.ins.2014.03.038](https://doi.org/10.1016/j.ins.2014.03.038)]
- [48] Shih TK, Gunaratne WKT, Ochirbat A, Su HM. Grouping peers based on complementary degree and social relationship using genetic algorithm. *ACM Trans. on Internet Technology*, 2018, 19(1): 2. [doi: [10.1145/3193180](https://doi.org/10.1145/3193180)]
- [49] Abdel-Basset M, Abdel-Fatah L, Sangaiah AK. Chapter 10—Metaheuristic algorithms: A comprehensive review. In: Sangaiah AK, Sheng M, Zhang ZY, eds. *Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications*. London: Academic Press, 2018. 185–231. [doi: [10.1016/B978-0-12-813314-9.00010-4](https://doi.org/10.1016/B978-0-12-813314-9.00010-4)]
- [50] Chen BY, Hwang GH, Lin TS. Impacts of a dynamic grouping strategy on students' learning effectiveness and experience value in an item bank-based collaborative practice system. *British Journal of Educational Technology*, 2020, 51(1): 36–52. [doi: [10.1111/bjet.12794](https://doi.org/10.1111/bjet.12794)]
- [51] Luo L, Yang Y, Ma Y. Research on online collaborative learning group division based on fuzzy C-means algorithm. *Computer Engineering and Applications*, 2017, 53(16): 68–73 (in Chinese with English abstract). [doi: [10.3778/j.issn.1002-8331.1612-0327](https://doi.org/10.3778/j.issn.1002-8331.1612-0327)]
- [52] Li HJ, Yue L, Zhang PW, Yang L. Online learning group formation method from the perspective of multi-objective optimization. *Journal of Chinese Computer Systems*, 2022, 43(4): 712–722 (in Chinese with English abstract). [doi: [10.20009/j.cnki.21-1106/TP.2020-0985](https://doi.org/10.20009/j.cnki.21-1106/TP.2020-0985)]
- [53] Zheng YQ, Li CR, Liu SY, Lu WG. An improved genetic approach for composing optimal collaborative learning groups. *Knowledge-based Systems*, 2018, 139: 214–225. [doi: [10.1016/j.knosys.2017.10.022](https://doi.org/10.1016/j.knosys.2017.10.022)]
- [54] Ullmann MRD, Ferreira DJ, Camilo CG, Caetano SS, De Assis L. Formation of learning groups in cMOOCs using particle swarm optimization. In: *Proc. of the 2015 IEEE Congress on Evolutionary Computation*. Sendai: IEEE, 2015. 3296–3304. [doi: [10.1109/CEC.2015.7257302](https://doi.org/10.1109/CEC.2015.7257302)]
- [55] Lin YT, Huang YM, Cheng SC. An automatic group composition system for composing collaborative learning groups using enhanced particle swarm optimization. *Computers & Education*, 2010, 55(4): 1483–1493. [doi: [10.1016/j.compedu.2010.06.014](https://doi.org/10.1016/j.compedu.2010.06.014)]
- [56] Charband Y, Navimipour NJ. Knowledge sharing mechanisms in the education: A systematic review of the state of the art literature and recommendations for future research. *Kybernetes*, 2018, 47(7): 1456–1490. [doi: [10.1108/K-06-2017-0227](https://doi.org/10.1108/K-06-2017-0227)]
- [57] Chang HH, Chuang SS. Social capital and individual motivations on knowledge sharing: Participant involvement as a moderator. *Information & Management*, 2011, 48(1): 9–18. [doi: [10.1016/j.im.2010.11.001](https://doi.org/10.1016/j.im.2010.11.001)]
- [58] Chiu CM, Hsu MH, Wang ETG. Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision Support Systems*, 2006, 42(3): 1872–1888. [doi: [10.1016/j.dss.2006.04.001](https://doi.org/10.1016/j.dss.2006.04.001)]
- [59] Shi CC, Hu P, Fan WG, Qiu LF. How learning effects influence knowledge contribution in online Q&A community? A social cognitive perspective. *Decision Support Systems*, 2021, 149: 113610. [doi: [10.1016/j.dss.2021.113610](https://doi.org/10.1016/j.dss.2021.113610)]
- [60] Yan ZJ, Wang TM, Chen Y, Zhang H. Knowledge sharing in online health communities: A social exchange theory perspective. *Information & Management*, 2016, 53(5): 643–653. [doi: [10.1016/j.im.2016.02.001](https://doi.org/10.1016/j.im.2016.02.001)]
- [61] Chen CJ, Hung SW. To give or to receive? Factors influencing members' knowledge sharing and community promotion in professional virtual communities. *Information & Management*, 2010, 47(4): 226–236. [doi: [10.1016/j.im.2010.03.001](https://doi.org/10.1016/j.im.2010.03.001)]
- [62] Zhang YJ, Du YL, Meng XW. Research on group recommender systems and their applications. *Chinese Journal of Computers*, 2016, 39(4): 745–764 (in Chinese with English abstract). [doi: [10.11897/SP.J.1016.2016.00745](https://doi.org/10.11897/SP.J.1016.2016.00745)]
- [63] Chen YL, Cheng LC, Chuang CN. A group recommendation system with consideration of interactions among group members. *Expert Systems with Applications*, 2008, 34(3): 2082–2090. [doi: [10.1016/j.eswa.2007.02.008](https://doi.org/10.1016/j.eswa.2007.02.008)]
- [64] Seo YD, Kim YG, Lee E, Seol KS, Baik DK. An enhanced aggregation method considering deviations for a group recommendation. *Expert Systems with Applications*, 2018, 93: 299–312. [doi: [10.1016/j.eswa.2017.10.027](https://doi.org/10.1016/j.eswa.2017.10.027)]

- [65] Baltrunas L, Makcinskas T, Ricci F. Group recommendations with rank aggregation and collaborative filtering. In: Proc. of the 4th ACM Conf. on Recommender Systems. Barcelona: Association for Computing Machinery, 2010. 119–126. [doi: [10.1145/1864708.1864733](https://doi.org/10.1145/1864708.1864733)]
- [66] Amer-Yahia S, Roy SB, Chawlat A, Das G, Yu C. Group recommendation: Semantics and efficiency. Proc. of the VLDB Endowment, 2009, 2(1): 754–765. [doi: [10.14778/1687627.1687713](https://doi.org/10.14778/1687627.1687713)]
- [67] Yuan Q, Cong G, Lin CY. COM: A generative model for group recommendation. In: Proc. of the 20th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Long Beach: Association for Computing Machinery, 2014. 163–172. [doi: [10.1145/2623330.2623616](https://doi.org/10.1145/2623330.2623616)]
- [68] Cao D, He XN, Miao LH, An YH, Yang C, Hong RC. Attentive group recommendation. In: Proc. of the 41st Int'l ACM SIGIR Conf. on Research & Development in Information Retrieval. Ann Arbor: Association for Computing Machinery, 2018. 645–654. [doi: [10.1145/3209978.3209998](https://doi.org/10.1145/3209978.3209998)]
- [69] Tran LV, Pham TAN, Tay Y, Liu YD, Cong C, Li XL. Interact and decide: Medley of sub-attention networks for effective group recommendation. In: Proc. of the 42nd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. Paris: Association for Computing Machinery, 2019. 255–264. [doi: [10.1145/3331184.3331251](https://doi.org/10.1145/3331184.3331251)]
- [70] Yin HZ, Wang QY, Zheng K, Li ZX, Yang JL, Zhou XF. Social influence-based group representation learning for group recommendation. In: Proc. of the 35th IEEE Int'l Conf. on Data Engineering. Macao: IEEE, 2019. 566–577. [doi: [10.1109/ICDE.2019.00057](https://doi.org/10.1109/ICDE.2019.00057)]
- [71] Guo L, Yin HZ, Wang QY, Cui B, Huang Z, Cui LZ. Group recommendation with latent voting mechanism. In: Proc. of the 36th IEEE Int'l Conf. on Data Engineering. Dallas: IEEE, 2020. 121–132. [doi: [10.1109/ICDE48307.2020.00018](https://doi.org/10.1109/ICDE48307.2020.00018)]
- [72] Deng ZY, Li CY, Liu SJ, Ali W, Shao J. Knowledge-aware group representation learning for group recommendation. In: Proc. of the 37th IEEE Int'l Conf. on Data Engineering. Chania: IEEE, 2021. 1571–1582. [doi: [10.1109/ICDE51399.2021.00139](https://doi.org/10.1109/ICDE51399.2021.00139)]
- [73] Rodríguez P, Giraldo M, Tabares V, Duque N, Ovalle D. Recommendation system of educational resources for a student group. In: Proc. of the 2016 Int'l Conf. on Practical Applications of Agents and Multi-agent Systems. Sevilla: Springer, 2016. 419–427. [doi: [10.1007/978-3-319-39387-2_35](https://doi.org/10.1007/978-3-319-39387-2_35)]
- [74] Papamitsiou Z, Economides AA. Motivating students in collaborative activities with game-theoretic group recommendations. IEEE Trans. on Learning Technologies, 2020, 13(2): 374–386. [doi: [10.1109/TLT.2018.2869582](https://doi.org/10.1109/TLT.2018.2869582)]
- [75] Johnson DW, Johnson RT, Smith KA. Cooperative learning: Improving university instruction by basing practice on validated theory. Journal on Excellence in College Teaching, 2014, 25(3–4): 85–118.
- [76] Wang CC, Hung JC, Yang CY, Chang HP. A question answering system approach for collaborative learning. In: Proc. of the 10th Int'l Conf. on Computer Supported Cooperative Work in Design. Nanjing: IEEE, 2006. 1–5. [doi: [10.1109/CSCWD.2006.253045](https://doi.org/10.1109/CSCWD.2006.253045)]
- [77] Agarwal A, Sachdeva N, Yadav RK, Udandara V, Mittal V, Gupta A, Mathur A. EDUQA: Educational domain question answering system using conceptual network mapping. In: Proc. of the 2019 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing. Brighton: IEEE, 2019. 8137–8141. [doi: [10.1109/ICASSP.2019.8683538](https://doi.org/10.1109/ICASSP.2019.8683538)]
- [78] Li BC, King I. Routing questions to appropriate answerers in community question answering services. In: Proc. of the 19th ACM Int'l Conf. on Information and Knowledge Management. Toronto: Association for Computing Machinery, 2010. 1585–1588. [doi: [10.1145/1871437.1871678](https://doi.org/10.1145/1871437.1871678)]
- [79] Zhou TC, Lyu MR, King I. A classification-based approach to question routing in community question answering. In: Proc. of the 21st Int'l Conf. on World Wide Web. Lyon: Association for Computing Machinery, 2012. 783–790. [doi: [10.1145/2187980.2188201](https://doi.org/10.1145/2187980.2188201)]
- [80] Li ZY, Jiang JY, Sun YZ, Wang W. Personalized question routing via heterogeneous network embedding. In: Proc. of the 33rd AAAI Conf. on Artificial Intelligence. Honolulu: IEEE, 2019. 192–199. [doi: [10.1609/aaai.v33i01.3301192](https://doi.org/10.1609/aaai.v33i01.3301192)]
- [81] Cheng X, Zhu SG, Su S, Chen G. A multi-objective optimization approach for question routing in community question answering services. IEEE Trans. on Knowledge and Data Engineering, 2017, 29(9): 1779–1792. [doi: [10.1109/TKDE.2017.2696008](https://doi.org/10.1109/TKDE.2017.2696008)]
- [82] Zhou YH, Cong G, Cui B, Jensen CS, Yao JJ. Routing questions to the right users in online communities. In: Proc. of the 25th IEEE Int'l Conf. on Data Engineering. Shanghai: IEEE, 2009. 700–711. [doi: [10.1109/ICDE.2009.44](https://doi.org/10.1109/ICDE.2009.44)]
- [83] Macina J, Srba I, Williams JJ, Bielikova M. Educational question routing in online student communities. In: Proc. of the 11th ACM Conf. on Recommender Systems. Como: Association for Computing Machinery, 2017. 47–55. [doi: [10.1145/3109859.3109886](https://doi.org/10.1145/3109859.3109886)]
- [84] Bezerra JM, Hirata C, Randall D. A conceptual framework to define incentive mechanisms for virtual communities. Journal of Universal Computer Science, 2015, 21(9): 1107–1135.
- [85] Tian CQ, Yang BJ, Zhong JD, Liu XJ. Trust-based incentive mechanism to motivate cooperation in hybrid P2P networks. Computer Networks, 2014, 73: 244–255. [doi: [10.1016/j.comnet.2014.08.010](https://doi.org/10.1016/j.comnet.2014.08.010)]
- [86] Wu NM, Zhou XW, Sun MX. Incentive mechanisms and impacts of negotiation power and information availability in multi-relay

- cooperative wireless networks. *IEEE Trans. on Wireless Communications*, 2019, 18(7): 3752–3765. [doi: [10.1109/TWC.2019.2918538](https://doi.org/10.1109/TWC.2019.2918538)]
- [87] Hou M, Han J, Zhang JS, Liu J. Research on incentive mechanism in network education platform. In: *Proc. of the 12th Int'l Conf. on Natural Computation, Fuzzy Systems and Knowledge Discovery*. Changsha: IEEE, 2016. 2288–2293. [doi: [10.1109/FSKD.2016.7603538](https://doi.org/10.1109/FSKD.2016.7603538)]
- [88] Hummel HGK, Burgos D, Tattersall C, Brouns F, Kurvers H, Koper R. Encouraging contributions in learning networks using incentive mechanisms. *Journal of Computer Assisted Learning*, 2005, 21(5): 355–365. [doi: [10.1111/j.1365-2729.2005.00140.x](https://doi.org/10.1111/j.1365-2729.2005.00140.x)]
- [89] Cheng R, Vassileva J. Design and evaluation of an adaptive incentive mechanism for sustained educational online communities. *User Modeling and User-adapted Interaction*, 2006, 16(3): 321–348. [doi: [10.1007/s11257-006-9013-6](https://doi.org/10.1007/s11257-006-9013-6)]
- [90] Jin FS, Niu ZD, Zhang QX, Lang HY, Qin K. A user reputation model for DLDE learning 2.0 community. In: *Proc. of the 11th Int'l Conf. on Asian Digital Libraries*. Bali: Springer, 2008. 61–70. [doi: [10.1007/978-3-540-89533-6_7](https://doi.org/10.1007/978-3-540-89533-6_7)]
- [91] Chang CH, Chen CM, Zhao RH. An effective group incentive mechanism in a collaborative problem-based learning system for enhancing positive peer interaction and learning performance. In: *Proc. of the 7th Int'l Congress on Advanced Applied Informatics*. Yonago: IEEE, 2018. 170–175. [doi: [10.1109/IIAI-AAI.2018.00041](https://doi.org/10.1109/IIAI-AAI.2018.00041)]
- [92] Shi C, Hu YX, Zhang ZM, Shao L, Jiang FJ. User feedback and ranking in-a-loop: Towards self-adaptive dialogue systems. In: *Proc. of the 44th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval*. Association for Computing Machinery, 2021. 2046–2050. [doi: [10.1145/3404835.3463079](https://doi.org/10.1145/3404835.3463079)]
- [93] Cai YZ, Chakravarthy S. Expertise ranking of users in QA community. In: *Proc. of the 18th Int'l Conf. on Database Systems for Advanced Applications*. Wuhan: Springer, 2013. 25–40. [doi: [10.1007/978-3-642-37487-6_5](https://doi.org/10.1007/978-3-642-37487-6_5)]
- [94] Omari A, Carmel D, Rokhlenko O, Szpektor I. Novelty based ranking of human answers for community questions. In: *Proc. of the 39th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval*. Pisa: Association for Computing Machinery, 2016. 215–224. [doi: [10.1145/2911451.2911506](https://doi.org/10.1145/2911451.2911506)]
- [95] Narang K, Krishnan A, Wang JT, Yang CQ, Sundaram H, Sutter C. Ranking user-generated content via multi-relational graph convolution. In: *Proc. of the 44th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval*. Association for Computing Machinery, 2021. 470–480. [doi: [10.1145/3404835.3462857](https://doi.org/10.1145/3404835.3462857)]
- [96] Backstrom L, Huttenlocher D, Kleinberg J, Lan XY. Group formation in large social networks: Membership, growth, and evolution. In: *Proc. of the 12th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. Philadelphia: Association for Computing Machinery, 2006. 44–54. [doi: [10.1145/1150402.1150412](https://doi.org/10.1145/1150402.1150412)]

附中文参考文献:

- [9] 郑庆华, 董博, 钱步月, 田锋, 魏笔凡, 张未展, 刘均. 智慧教育研究现状与发展趋势. *计算机研究与发展*, 2019, 56(1): 209–224. [doi: [10.7544/issn1000-1239.2019.20180758](https://doi.org/10.7544/issn1000-1239.2019.20180758)]
- [19] 周庆, 牟超, 杨丹. 教育数据挖掘研究进展综述. *软件学报*, 2015, 26(11): 3026–3042. <http://www.jos.org.cn/1000-9825/4887.htm> [doi: [10.13328/j.cnki.jos.004887](https://doi.org/10.13328/j.cnki.jos.004887)]
- [22] 胡勇, 李宇峰. 虚拟学习社区的内涵及其相关研究概述. *现代远程教育研究*, 2012, (3): 32–42. [doi: [10.3969/j.issn.1009-5195.2012.03.005](https://doi.org/10.3969/j.issn.1009-5195.2012.03.005)]
- [23] 纪河, 朱燕菲. 继承与创新: 由共同体走向学习共同体. *中国远程教育*, 2019, (10): 74–79. [doi: [10.13541/j.cnki.chinade.2019.10.007](https://doi.org/10.13541/j.cnki.chinade.2019.10.007)]
- [51] 罗凌, 杨有, 马燕. 基于模糊C均值的在线协作学习混合分组研究. *计算机工程与应用*, 2017, 53(16): 68–73. [doi: [10.3778/j.issn.1002-8331.1612-0327](https://doi.org/10.3778/j.issn.1002-8331.1612-0327)]
- [52] 李浩君, 岳磊, 张鹏威, 杨琳. 多目标优化视角下在线学习群体形成方法. *小型微型计算机系统*, 2022, 43(4): 712–722. [doi: [10.20009/j.cnki.21-1106/TP.2020-0985](https://doi.org/10.20009/j.cnki.21-1106/TP.2020-0985)]
- [62] 张玉洁, 杜雨露, 孟祥武. 组推荐系统及其应用研究. *计算机学报*, 2016, 39(4): 745–764. [doi: [10.11897/SP.J.1016.2016.00745](https://doi.org/10.11897/SP.J.1016.2016.00745)]



张俊涛(1989—), 男, 博士生, CCF 学生会员, 主要研究领域为教育大数据, 学习共同体, 数据挖掘, 机器学习.



张雪龙(1997—), 男, 硕士, 主要研究领域为教育大数据, 团队形成, 数据挖掘.



杨先娣(1974—), 女, 博士, 副教授, CCF 专业会员, 主要研究领域为教育大数据, 大数据管理与分析, 可信数据管理.



贺宇阳(1999—), 男, 硕士生, CCF 学生会员, 主要研究领域为教育大数据, 学习共同体.



宋伟(1978—), 男, 博士, 副教授, CCF 高级会员, 主要研究领域为大数据管理与分析, 云计算, 网络安全, 应用密码学.



彭智勇(1963—), 男, 博士, 教授, 博士生导师, CCF 会士, 主要研究领域为数据库, 大数据管理与分析, 可信数据管理, 复杂数据管理.