

文章编号:1671-4229(2021)02-0077-09

MOOC 社区中基于移动情境感知的学伴推荐方法研究

廖宏建^a, 曲哲^b, 胡莹^a

(广州大学 a. 网络与现代教育技术中心; b. 图书馆, 广东 广州 510006)

摘要: 在大规模学员参与的 MOOC 社区中智能推荐适配的学习伙伴, 对促进学习投入和提升课程完成率具有重要意义. 基于探究社区理论模型分析 MOOC 社区知识地图结构, 梳理影响学伴推荐的用户情境(学习投入、评分、时间及位置)和社会情境(交互行为、强度及时效), 相应地构建用户情境信任度和社会情境信任度; 在社会情境信任度构建中, 引入信任奖励因子和时间衰减因子, 反映 MOOC 社交信任的动态性; 最后融合用户情境和社会情境计算动态综合信任度, 通过信任传播构建 MOOC 社区复杂信任网络, 基于此为学员生成 Top-N 推荐列表. 实验表明, 融合用户情境和社会情境的学伴推荐方法, 能充分挖掘学员兴趣偏好及交互行为变化所反映的动态信任关系, 在高动态性和移动性的 MOOC 社区中具有较好的推荐效果; 同时, 可以通过调整多维情境的权重值实现在不同推荐场景的灵活应用.

关键词: MOOC 社区; 情境感知; 信任网络; 知识地图; 学伴推荐系统

中图分类号: TP 391.1 **文献标志码:** A

A learning partner recommendation method for MOOC community based on mobile context-awareness

LIAO Hong-jian^a, QU Zhe^b, HU Ying^a

(a. Network and Modern Educational Technology Center; b. Library, Guangzhou University, Guangzhou 510006, China)

Abstract: Intelligent learning partner recommendation is of great significance in promoting learning engagement and improving course completion rate in a large-scale MOOC community. Based on Community of Inquiry (COI), the structure of knowledge map in MOOC community was analyzed. As user context and social context were important factors which influence learning partner recommendation, degree of trust based on user context and social context were constructed correspondingly. In terms of social context trust degree, trust reward factor were time attenuation factor were introduced to reflect the dynamic nature of MOOC social trust. Finally, the Top-N recommendation list was generated for learners through the construction of a complex trust network in the MOOC community based on trust propagation. According to the research results, learning partner recommendation method, which integrates user context and social context, can fully mine the dynamic trust relationship reflected by the changes in students' interests and interaction behavior. This method's effect is greatly recommended particularly in a high dynamic and mobile MOOC community. The method can also be applied to different recommendation scenarios flexibly by adjusting the weight value of the multi-dimensional context.

Key words: MOOC community; context-awareness; trust network; knowledge map; learning partner recommendation system

基金项目: 国家社科基金青年资助项目(15CTQ034)

作者简介: 廖宏建(1980—),男,副研究员. E-mail:337917539@qq.com

引文格式: 廖宏建, 曲哲, 胡莹. MOOC 社区中基于移动情境感知的学伴推荐方法研究[J]. 广州大学学报(自然科学版), 2021, 20(2): 77-85.

近年来,大规模在线开放课程(Massive Open Online Course, MOOC)的迅猛发展,为人们提供了前所未有的学习机会,每年有数以千万计的学习者参与 MOOC 学习,但 MOOC 的高辍学率也同样引起了关注(完成率仅约为 5%~15%)^[1]. 师生时空分离引起的孤独感、边缘参与和群体动力缺乏等是高辍学率的主要归因^[2-3]. 大量文献研究和教学实践已证明,学伴互动对降低孤独感,增加持续学习投入和促进深层次学习具有重要意义^[4-5],而学伴推荐(Learning Partner Recommendation, LPR)能促进学伴互动.

联通主义学习理论认为,学习即连接及形成网络^[6],学习者与其他同伴以学习为目的进行社交时,会伴随生成隐形人际关系,并推动学习社区形成. 随着大规模在线学习的兴起,MOOC 已不仅仅是一门视频课程,更是一个知识创造、交流和分享的学习社区,且具有以下特点:①学员的规模化与多样性. 一门课程的学员往往成千上万,来自不同的行业和年龄段,有着不同的知识水平、兴趣偏好和自我调节学习能力等;②学习与社交趋向融合. 学习社交化,社交学习化,MOOC 社区具有学习与社交的双重属性;③学习情境的移动化与泛在化. 基于移动媒介的微学习,形成移动社会网络,行为日志数据呈现自然浮现、多模态、碎片化和时空情境化等特点.

MOOC 社区的“规模化、社交化和移动化”等特征使得学习者情境更为复杂,如何在 MOOC 社区中智能、动态地推荐适配的学伴,成为 MOOC 发展中的重要研究问题. 网络分析被认为是研究在线学习大规模交互的有效方法^[7],它将行动者抽象为节点,将关系抽象为节点之间的边,形成一个社群图,从而描述网络关系. 本文在网络关系中引入信任度概念,使用信任度来统摄和描述 MOOC 社区节点间的兴趣、时间和位置等情境信息的相似度,构建动态信任网络,实现基于移动情境感知的学伴推荐.

1 相关工作

MOOC 社区中的学习同伴推荐(LPR)借鉴社交网络好友推荐的相关方法,主要有 3 类:①基于好友关系拓扑网络的推荐;②基于用户行为、兴趣等相关信息的推荐;③基于信任的推荐. 徐彬等^[8]分析了 MOOC 论坛区学习者的行为特征及形成的关系边类型,基于课程内容关键词创建固定主题模型,根据学生讨论内容主题分布的相似性来推荐学伴,实验验证该方法能有效挖掘出主题相关的学习者. 但其缺点在于需要事先为每门课程建立关键词词典. 蔡霞等^[9]构建了基于学习者特征(包含学习者属性、能力和偏好)的同伴推荐算法,推荐互补型或相似型的学习同伴. Hugues 团队^[10]使用对照研究分析了 MOOC 中“同伴推荐能否促进学习投入”及 3 种

推荐策略的效果,结果表明,有学伴推荐的实验组学员在参与度和课程完成率方面均优于无学伴推荐的对照组;在 3 种推荐策略上,基于人口统计信息相似性(包括地理位置、先前经验)的同伴推荐效果稍优于基于 MOOC 进度的推荐和随机推荐^[11]. 刘海鸥等^[12]提出了一种基于大数据深度画像学习伙伴匹配的理论模型,使用学习者教育背景、内容偏好和学习互动等情境计算学习风格,从而推荐风格相近或风格互补的学伴,但缺乏具体实践. 为了体现学习者行为和兴趣的动态性,Elghomary 等^[13]提出了一种基于信任管理系统(TMS)的动态学伴推荐系统(DPRS),该方法吸收了社交物联网(SIoT)的社会信任指标,如直接信任度、对象中心度、能量、社区兴趣、合作性和服务得分来计算节点的可信度,并定期进行信任更新.

综上所述,MOOC 学伴推荐研究尚处于起步阶段,主要借鉴了一般性社会网络中好友推荐的思路和方法. MOOC 作为一种在线知识社区,包含显性知识(课程资源)和隐性知识(学习者)节点,学习是特定节点之间建立连接并形成网络的过程,从本质上讲,学伴推荐即知识连接. 因此,在推荐情境和推荐目的上,MOOC 学伴推荐有其自身的特点:①学习者与课程内容之间的连接及交互行为是反映学习者知识水平、兴趣偏好的重要依据,也是影响学伴连接的重要情境. 已有推荐方法忽略了学习者与学习内容交互行为特征的分析及量化;② MOOC 社交网络具有高动态性和移动性,学伴连接的出现或消失是 MOOC 社交网络的链接拓扑随时间、空间、学习内容、兴趣偏好、知识水平和交互关系等诸多情境变化引起的,且根据推荐目标的侧重点不同,情境的取舍也会不同,如地理位置能为在线社区中学伴线下交流不紧密的问题提供情境线索. 选择适切的、促进学伴连接的情境,是实现有效学伴推荐的关键.

2 MOOC 社区情境分类

2.1 探究社区模型与 MOOC 知识地图

探究社区理论模型(Community of Inquiry, CoI)指通过发展“教学临场、社会临场和认知临场”3 个相互依存的要素,创造深层次和有意义的学习体验过程^[14]. 学习者与课程内容的交互等构成“教学临场”;学习者对知识的连接、批判性反思和深加工等活动构成“认知临场”;学习者之间互动(情感反应、交互式响应和凝聚力反应)构成“社会临场”^[15]. 其中,交互式响应是社会临场的核心,它是指引用学习同伴的话语信息来维持学习主题的讨论,或表达对他人的同意,或赞赏,或转发等. 3 种临场感的载体为“学习者、课程资源和交互”,3 者的连接关系及构建的知识地图可描述如图 1 所示.

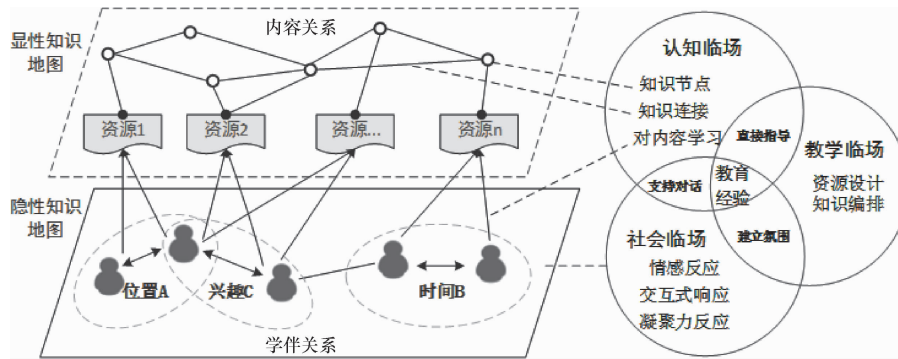


图 1 探究社区视角下的 MOOC 知识地图构成

Fig. 1 The construction of MOOC knowledge map from the perspective of COI

MOOC 社区中的知识构成包括显性知识和隐性知识.学习者本身作为一种隐性知识载体,是整个知识网络中的重要组成节点,节点之间的交互意向、信任关系创建等,无疑受到来自教学临场、认知临场中的时间、兴趣、位置和交互等情境影响.实现个性化学伴推荐(即构建隐性知识连接网络),需要计算和量化知识地图中的相关情境信息,并根据情境相似性构建学习者间的信任关系及动态信任网络.

2.2 面向学伴推荐的 MOOC 学习情境分析

依据图 1,学习者的交互行为可以分为两类:学习者—课程资源交互、学习者—学习者交互.第一类交互行为主要反映了学习者的兴趣、时间、地点和影响力等情境信息,跟用户自身相关,在网络中具有全局性,谓之用户情境;第二类交互行为主要反映学习者之间的交互关系,是点对点的行为,是“从虚拟或现实社会中获取的人际互动信息集”^[16],谓之社会情境.因此,面向学伴推荐的情境信息分类如表 1 所示.

表 1 面向学伴推荐的情境信息分类

Table 1 Context classification for peer recommendation

情境分类	情境子类	描述
用户情境	兴趣偏好	学习行为投入度
		[自主行为]课程登录次数,课件浏览数(视频、文档),在线学习时长
		[反思行为]学习笔记频次,学习笔记长度
		[知识水平]测验平均分,课程总成绩
时间情境	在线学习时间	[评价]对课程评分、评级
		位置情境 在线学习地理位置
社会情境	学伴交互	[交互行为]回帖、评论,互评(评价他人,被评),点赞,转发,分享...
		[交互强度]交互行为的频次,时长等
		[交互时效]交互随时序演化情况

对上述情境量化计算和融合,使用信任来统摄和度量学员间的情境相似度,并构建信任网络,实现基于情境感知的学伴推荐.

3 MOOC 社区信任度计算

在好友推荐中,定义信任为目标用户对所推荐用户的有效性、真实性和可靠性的一种信赖和肯定程度.MOOC 社区学员间综合信任度由用户情境信任度和社会情境信任度构成.

3.1 基于用户情境的信任度

3.1.1 兴趣偏好信任度

根据自我知觉理论(Self-perception theory),用户的态度和偏好在一定程度上可以通过用户的行为来预测^[17].根据该理论,学员过去的真实学习行为投入能在一定程度上反映其对该 MOOC 的偏好.学习行为投入是指学习者在学习活动中的行为表现,包含了投入时间、努力程度和活动强度等^[18].除了态度和偏好,在同一门课程或同一活动中学习行为投入的相似度也能在一定程度上反映学习者知识水平的相似性.借鉴文献[19]的学习行为投入评测框架,本研究中选取自主学习、行为反思和知识水平 3 个维度共 7 项指标,见表 1.此外,学员对课程的评分也是反映兴趣偏好的重要数据.因此,反映学习者兴趣偏好的信任度计算步骤如下.

(1)学习行为投入系数.为消除量纲对结果的影响,对 7 项指标使用极大值-极小值做归一化处理,对归一化后的值进行累加得到学员 u 对课程 i 的学习行为投入系数 $e_{u,i}$,然后使用经典算法皮尔逊相关系数计算学员间信任度,如式(1)所示.

$$eTrust_{u,v} = \frac{\sum_i (e_{u,i} - \bar{e}_u)(e_{v,i} - \bar{e}_v)}{\sqrt{\sum_i (e_{u,i} - \bar{e}_u)^2} \sqrt{\sum_i (e_{v,i} - \bar{e}_v)^2}} \quad (1)$$

其中, $e_{u,i}$ 和 $e_{v,i}$ 分别表示学员 u 和学员 v 对课程 i 的投入系数, \bar{e}_u 和 \bar{e}_v 表示学员 u 和学员 v 对参与课程的平均投

入系数.

(2)课程评分.课程评价一般采用 $[1,5]$ 区间的整数表示,相比学习投入矩阵,评分矩阵数据更为稀疏.假如2名学员仅有1门共同评分课程,且评分相同,使用皮尔逊相关系数计算得到相似度为1,这与实际情况不符.因此,使用均方偏差和Jaccard系数的乘积来增强度量学员评分相似性.假设学员 u 和学员 v 对课程均有评分,则

$$sTrust' = MSD_{u,v} * JAC_{u,v} = \left(1 - \frac{1}{I_{u,v}} \sum_{i=1}^{I_{u,v}} \left(\frac{r_{u,i} - r_{v,i}}{\max - \min}\right)^2\right) * \frac{I_u \cap I_v}{I_u \cup I_v} \quad (2)$$

课程评分具有一定的主观性,有些学员的评分习惯性偏高,有些学员则习惯性偏低,这种评分偏好使用偏差系数来表示,即用某学员与所有学员打分的差值来表示.将信任度乘以打分偏差系数,能够消除打分偏好的影响,得到更为客观的信任值,如式(3)所示.

$$sTrust = \frac{1}{1 + \sum_{i=1}^M (S_{ui} - \bar{S}_i) * I_{ui}} * sTrust' \quad (3)$$

综合式(1)和式(3)得到调和后的基于学习者兴趣偏好的信任度如式(4)所示.

$$preTrust_{u,v} = a * eTrust_{u,v} + (1 - a) * sTrust_{u,v} \quad (4)$$

其中, a 为权重参数,用于调节学习投入和评分在兴趣偏好信任度中的比重.

3.1.2 地理位置信任度

MOOC平台在选课日志、登录日志和每次学习行为日志中均记录了学习者的IP信息.通过IP信息可以获得学习者的地理位置信息,实现基于相同或相近位置的学伴推荐,促进学习互动从线上走向线下.任意学习者 u 和学习者 v 在空间上的相似性可以使用离散地理位置相似性度量.计算如式(5)所示.

$$locaTrust_{u,v} = \frac{\sum_{l \in P_{u,v} \cap N} P_{u,l} P_{v,l}}{\|P_u\| * \|P_v\|} \quad (5)$$

其中, N 表示学员序列集合; P 表示地理位置(城市)序列集合; $p_{u,l}$ 表示学员 u 在位置站点 l 出现的概率; $P_u = [P_{u,1}, P_{u,2}, \dots, P_{u,m}]$ (m 表示位置站点的总数)表示学员 u 在各个位置站点的出现概率向量; $\|P_u\|$ 表示 P_u 的范数.

3.1.3 同时在线时间信任度

在MOOC社区交互中,同步互动比异步互动更能提升社会临场感,能提高交互率和学习效果.任意学员 u 和 v 在时间上的相似度可以用同时在线时间相似度来表示,计算如式(6)所示.

$$timeTrust_{u,v} = \frac{\sum_{l \in T_{u,v} \cap N} P_{u,l} P_{v,l}}{\|P_u\| * \|P_v\|} \quad (6)$$

其中, N 表示所有学员序列集合; T 表示时间段序列集合,这里可根据需要,将一个自然日划分为不同的时段; $p_{u,l}$ 表示学员 u 在时间段 l 在线学习的概率;具体计算方式: $1/\text{该学员在某一自然日内总在线时间段的个数}$,所有 $p_{u,l}$ 之和为1. $P_u = [P_{u,1}, P_{u,2}, \dots, P_{u,k}]$ (k 表示时间段的总数)表示学员 u 在各个时间段在线学习的概率向量; $\|P_u\|$ 表示 P_u 的范数.

3.1.4 用户情境信任度综合计算

将上述基于兴趣偏好、地理位置和在线时间的信任值按权重综合,如式(7)所示,得到用户情境信任度.

$$UserTrust = \alpha * preTrust + \beta * locaTrust + \gamma * timeTrust \quad (7)$$

其中, α, β, γ 分别为这类信任度的权重,取值为 $[0,1]$ 区间,且 $\alpha + \beta + \gamma = 1$.

3.2 基于社会情境的信任度

交互行为体现了学习者之间的信任关系.在MOOC社区中,学员之间的回帖、点赞和转发等行为都代表了某种信任关系.尤其当学员频繁回复或者转发另一学员发布的内容时,表明该学员与该内容发布者之间具有较强的信任关系.

3.2.1 初始交互信任度计算

本文将MOOC交互行为集 $v = \{\text{回帖, 点赞, 转发}\}$ 作为交互信任度的计算指标.已有文献主要使用交互频次构建信任度,而MOOC交互作为一种知识学习和分享行为,交互时长能在一定程度上反映交流的深度.MOOC平台一般不记录某次交互行为的时长,因而这里使用回帖文本长度模拟交互时长.

(1)对不同交互行为赋予不同的权重

$$CF_{u \rightarrow v} = \lambda * \frac{u \text{ 回帖 } v}{u \text{ 回帖 } ALL} + \mu * \frac{u \text{ 转发 } v}{u \text{ 转发 } ALL} + \nu * \frac{u \text{ 点赞 } v}{u \text{ 点赞 } ALL} \quad (8)$$

$$CT_{u \rightarrow v} = \frac{Len(u \text{ 回帖 } v)}{Len(u \text{ 回帖 } ALL)} \quad (9)$$

其中, $CF_{u \rightarrow v}$ 表示学员 u 指向学员 v 的行为频次; $CT_{u \rightarrow v}$ 表示时长;同样,可以计算得到 v 指向 u 的行为频次 $CF_{v \rightarrow u}$ 和时长 $CT_{v \rightarrow u}$. λ, μ, ν 分别表示3种交互行为的权重.

(2)构建信任度

MOOC交互具有方向性,如果只是单方面主动发起联系,即便单方强度高,并不能代表双方熟识且信任度高,双向互动比单向互动构成好友的可能性更大,基于双向互动构建的信任度才有意义.因此,本文综合考虑双向互动来构建学员间的信任度,如式(10)所示, $CF_{u,v}$ 表示基于交互频次的信任度.

$$CF_{u,v} = \frac{2 \times \min(CF_{u \rightarrow v}, CF_{v \rightarrow u})}{2} + \frac{0.5 \times \max(CF_{u \rightarrow v}, CF_{v \rightarrow u})}{2} - \frac{0.5 \times \min(CF_{u \rightarrow v}, CF_{v \rightarrow u})}{2} \quad (10)$$

用同样的方法计算基于交互时长的信任度 $CT_{u,v}$ 。因学员交互时长和交互频次与信任度之间符合边际效应递减理论,故采用对数来度量交互与信任度之间的关系,得到初始交互信任度。

$$SociaTrust(u,v) = b \times \ln(CT_{u,v}) + (1-b) \times \ln(CF_{u,v}) \quad (11)$$

其中, $CT_{u,v}$ 和 $CF_{u,v}$ 分别表示学员 u 和 v 的交互时长和交互频次; b 表示调整交互频次和时长的权重。

3.2.2 交互信任的时序动态更新

在 MOOC 社区中,学员的兴趣与关注点是随着学习进程不断演化的,其学伴关系也会随着时间的变化而变化。具体而言,学伴的推荐受到两个方面的影响:①交互强度。当两个学员交互升温并进入专注期,且交互频次和深度明显高出其他人,这时应该予以奖励,使其脱颖而出;②交互时效。历史交互行为对信任值有远近效应,近期交互对当前信任值的影响大,远期交互对当前信任值的影响小。为了更关注近期交互的影响,引入时间衰减因子来削弱或遗忘历史行为记录的影响。相应地,本研究通过信任奖励因子和时间衰减因子来实现交互信任值的动态调整。将整个课程学习周期划分为若干个大小为 T 的时间窗口(按照 MOOC 交互规律,一般 $T=3d$),信任值每经过一个窗口 T 时更新一次。如图 2 所示,学员 u 和 v 、 u 和 w 的信任值依据课程中交互行为动态变化, t 时刻为信任评估和推荐的时间节点(可根据需要将课程周期划分为多个 t 节点)。

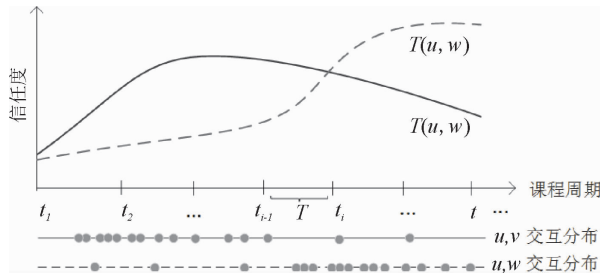


图2 奖励因子和时间衰减因子作用下的信任度演化示例
Fig.2 An example of trust evolution under reward factor and time decay factor

(1) 奖励因子

在第 i 个时间窗口 t_i 时刻,学员 u 对 v 的信任值可表示为 $ITrust(u,v,t_i)$, 因为 t_i 时刻的交互信任是 t_{i-1} 时刻交互信任的延续,也可以看作是对 t_{i-1} 窗口交互体验

的反馈。为量化交互反馈,定义反馈得分 S_i 为在 t_i 时刻,用户 u 对 v 的交互行为(包含频次和时长)占用户 u 对所有学员交互行为的比例。计算公式如下:

$$S_i = \frac{CF_{u,v}}{\sum_{k \in U_u} CF_{u,k}} + \frac{CT_{u,v}}{\sum_{k \in U_u} CT_{u,k}} \quad (12)$$

其中, U_u 表示与学员 u 有直接交互的学员集合, $\sum_{k \in U_u} CF_{u,k}$ 表示学员 u 对所有学员交互频次;频次 CF 包括回帖、点赞和转发 3 种行为,时长 CT 指回帖文本长度。 S_i 得分越高,说明学员 u, v 的当前交互越高于其他学员,交互更专注和深入。

比较 t_i 时刻和 t_{i-1} 时刻交互反馈得分,若 $S_i - S_{i-1} > 0$,则为正反馈;反之,若 $S_i - S_{i-1} < 0$,则为负反馈。引入信任奖励因子,对正反馈进行奖励。MOOC 社区交互不存在一般信任网络中的恶意攻击和伪装用户等,故不考虑施加惩罚因子,使得信任的自适应性特征改变为“快增慢减”,更吻合 MOOC 学习过程中学伴推荐的实际情况。

信任奖励因子 $r(k)$ 定义为两个连续时间窗口反馈得分的比值,即斜率 $r(k) = S_i/S_{i-1}$ 表示 k 时刻的奖励因子。交互结果为正反馈时,则 $r(k) \geq 1$;为负反馈时,则 $r(k) < 1$ 。

因此,经过 T 时间窗口后更新的信任值如式(13)所示。

$$SociaTrust(u,v,t_i)' = SociaTrust(u,v,t_i) * r(k) \quad (13)$$

经过奖励因子调节后,信任值在每个时间窗口 T 呈现了动态变化性。

(2) 时间衰减因子

在推荐时间节点 t ,为了给远近不同的历史交互行为施加差异权重,使用时间衰减因子更新 t 周期内的信任值。时间衰减因子定义如下:

$$\delta(t) = e^{-c \cdot [(t-t_i)/T]} \quad (14)$$

其中, t 表示信任评估与推荐时刻, t_i 表示周期 t 内第 i 个时间窗口时刻, T 为时间窗格, c 为衰减系数,则施加时间衰减因子后的交互信任值为

$$SociaTrust(u,v,t_i) = SociaTrust(u,v,t_i) * \delta(t) \quad (15)$$

从式(15)中可知, t_i 离评估 t 时刻越远,则经历的窗口越多,衰减幅度越大。

3.3 MOOC 社区信任网络构建

3.3.1 信任调和

将前面基于用户情境的信任度和基于社会情境的信任度通过式(16)调和,得到综合信任度。

$$MOOCTrust_{u,v} = \eta \cdot UserTrust_{u,v} + (1-\eta) \cdot SociaTrust_{u,v} \quad (16)$$

其中, η 为调和参数($0 \leq \eta \leq 1$), 若 $\eta < 0.5$, 则用户情境信任度占主导; 若 $\eta > 0.5$, 则社会情境信任度占主导; $\eta = 1$ 时, 则完全使用用户情境, 解决冷启动用户社会情境信任数据的稀疏性问题。

3.3.2 信任传递与信任网络构建

(1) 信任传递. 信任具有传递性, 用户往往会接受其信任的朋友推荐的朋友, 即朋友的朋友 (Friend of friend) 也可能会成为朋友. 信任传递可以缓解数据稀疏性, 为目标用户匹配更多的邻居用户. 在 MOOC 社区中, 信任是多链路传递的, 一个学员收到不同信任传递的链路越多, 其信任度就越客观. 本文采用广泛使用的有序加权平均聚合算法计算多链路传递^[20], 该方法为每条信任链赋予不同的权重, 能真实反映综合信任度. 首先将目标用户 U_u 到 U_v 的 N 条信任链按长短排序, 根据 OWA 算子定理, 信任链权重的计算公式为

$$w_j = F\left(\frac{j}{n}\right) - F\left(\frac{j-1}{n}\right) \quad (17)$$

其中, $F(x) = x^2$, w_j 表示第 j 条传递路径的权重值, 且 $\sum_{j=1}^n w_j = 1$.

(2) 信任网络. 得到所有学员的信任关系矩阵后, 可构建 MOOC 社区学员信任网络, 该网络是一种有向加权网络. 使用图论方法可形式化描述如下: 使用一个三元组 $G(U, E, W)$ 表示复杂信任网络, 其中, 集合 $U = \{U_1, U_2, \dots, U_n\}$ 表示学员节点集合; 集合 $E = \{E_1, E_2, \dots, E_n\}$ 表示图中边的集合, 即学员间的信任关系; $W = \{W_1, W_2, \dots, W_n\}$ 表示边集合 E 对应的信任权重集合; 基于该网络, 依据信任权重值实现对目标学员的 Top-N 推荐。

4 实验过程与结果分析

4.1 实验数据描述与预处理

在国内大型 MOOC 社区平台“好大学在线”随机抽取了课程“现代文学经典导读”某一期开课的真实日志数据作为实验对象. 该课程选课人数为 2 125 人, 课程周期为 15 周, 学习活动有观看视频、学习课件、文学创作、主观作业互评、议题讨论和自由讨论等. 用到的主要数据文件有 5 个: uc_action_log (含用户基本信息、登录时间、请求 IP、请求学习资源 ID 及访问页面等)、course_comments (课程多维评价)、uc_discuz (帖子基本信息、发帖、回帖、点赞、转发及标记等)、uc_exam (互评、测验、考试及分数等) 和 uc_notice (笔记时间、内容及所关联资源等). 用户原始行为日志数据约 37 万余条, 按照表 1 和式(1)计算学习行为投入系数, 剔除系数为 0 或过低 (偶尔或零星行为) 的非活跃学员, 得到 1 350 人及 32 万余

日志数据进入最终测试数据集。

为了从原始数据中获得推荐需要的情境信息, 对数据进行了预处理和转换。

(1) 学习行为投入. ①在线学习时长计算. MOOC 平台日志一般只记录某一个行为 (如看视频、参与讨论) 的开始时间戳, 学员总在线时长通过前后两次行为的时间戳差值的累计来计算. 如果用户直接关闭浏览器时系统无法记录学习结束时间, 则根据 MOOC 微学习特征, 该次学习时长取 30 min, 即前后相邻行为的时间戳差值的阈值设为 30 min. ②交互时长计算. 发帖和回帖时长 (这里忽略点赞和转发行为的时长) 通过其帖子的文本长度模拟转换为时长。

(2) 时间情境. 为计算在线学习时间段相似性, 根据 MOOC 学习特点, 将一个自然日按 30 min 划分为 48 个时段, 用序列集合 T 表示。

(3) 位置情境. 使用百度 IP 定位 API 接口, 将日志中每次请求的 IP 地址转换为城市信息, 用序列集合 P 表示。

(4) 信任更新周期. 原始数据包括 105 d 的学员行为, 时间窗格 T 分别取值为 3 d、5 d、7 d 计算时, 全样本中回帖、点赞及转发等交互行为频次的平均更新率为 3.21%、5.52%、9.85%, 为平衡推荐精度和计算性能, 令时间窗格 $T=5$, 课程全周期中信任评估和推荐次数设为 7 (即图 4 中 $t=7$, 每次推荐周期为 105 d/7 次 = 15 d)。

4.2 实验结果

4.2.1 学习者信任网络图谱

根据用户情境和社会情境, 计算测试集中学员的信任值, 并构建信任关系矩阵, 使用社会网络分析工具 UCINET 生成 MOOC 学习者的复杂信任网络图谱. 图 3 为随机抽取的某目标学员 (昵称为徜徉@塞纳河畔) 在 $t=5$ 时刻的信任网络图, 为简化节点和连线, 提高显示效果, 学伴跳数设为 2, 二级学伴信任阈值设为 0.2。

通过对 MOOC 社区相关情境信息进行系列量化和计算, 构建目标学员的信任网络, 实现了对图 3 中学伴关系的可视化分析和展示。

4.2.2 推荐结果分析

MOOC 学伴推荐的目的在于促进推荐后的交互行为发生, 文献[8]将所推荐的学伴在后期与目标用户确实产生交流作为推荐成功的衡量指标. 本文也基于这一思路, 分别统计课程 7 个信任评估和推荐周期内学员间的交互频次, 并将频次大于一定阈值的统计为本周期的真实好友。

(1) 评价指标. 本文使用的是 Top-N 推荐, 评价指标采用推荐系统中常用的准确率 (precision)、召回率 (recall) 和 F1 值作为实验效果的评价指标, 具体计算方法如式(18)~(20)所示。

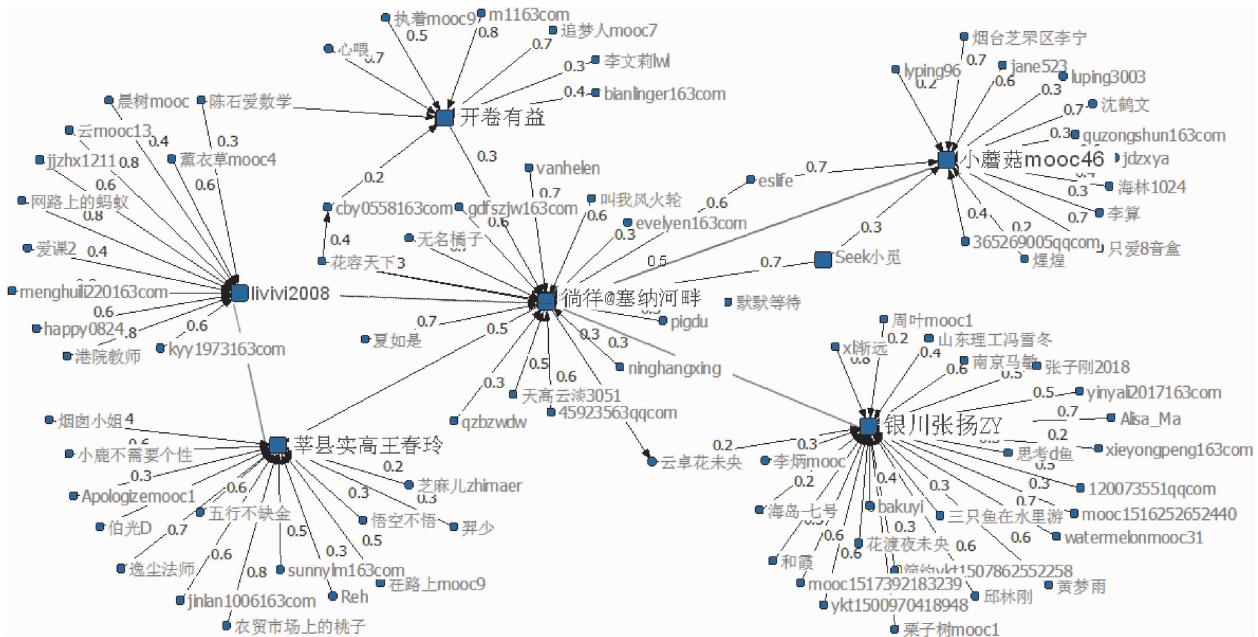


图 3 目标学员 t 时刻信任网络图谱

Fig. 3 Target student's trust network map at T -Time

$$precision = \frac{|R(u) \cap T(u)|}{|R(u)|} \quad (18)$$

$$recall = \frac{|R(u) \cap T(u)|}{|T(u)|} \quad (19)$$

$$F1 - measure = \frac{2 \times precision \times recall}{precision + recall} \quad (20)$$

其中, $R(u)$ 表示在 t 阶段 ($t = 1, 2, \dots, 7$) 系统推荐给目标学员 u 的好友数量, $T(u)$ 表示学员 u 在 t 之后所有阶段的真实好友数量. 准确率是指推荐命中的学伴数量占推荐总人数的比例, 反映推荐方法的精度. 召回率是指推荐命中的学伴数量占测试数据集中真实学伴总数的比例, 反映推荐方法的覆盖率. 这两个指标值越大, 推荐总体效果越好. 但这两个指标有时会矛盾, 故使用两者的加权调和平均指标 $F1-measure$ 表示综合推荐效果.

(2) 各个参数的设定

1) 在用户情境中, 根据前期相关研究^[19], 学习投入相比课程评分对兴趣偏好具有更强的解释力, 因此, 学习投入权重设为 0.6, 评分权重设为 0.4; 在用户情境信任度中, 兴趣偏好权重默认设为 1, 在线时段和位置权重根据实际需要可设为大于 0 的系数, 默认设为 0.

2) 在社会情境中, 一般地, 点赞的交互性较低, 而“转发”或“回帖”的交互性较高^[21], 代表对目标更为信任, 因此, 其权重也稍高, 令 $\lambda = 0.4$, $\mu = 0.4$, $\nu = 0.2$. 交互频次与时长权重的确定, 根据文献^[22]的经验值, 频次取 0.4, 时长取 0.6.

3) 时间衰减系数 c 取值范围为 $[0, 10]$ 之间的整数,

值越大, 时间衰减因子值越小. 对数据集进行多轮模拟计算, 本实验中取 $c = 4$.

4) 参数 η 的确定, 默认取值为 0.5, 即用户情境信任度和社会情境信任度各占一半比重.

(3) 推荐准确性的时序变化. 分别在 7 次信任评估和更新时间节点计算推荐准确率, 在推荐列表个数为 10 时, 结果如图 4 所示.

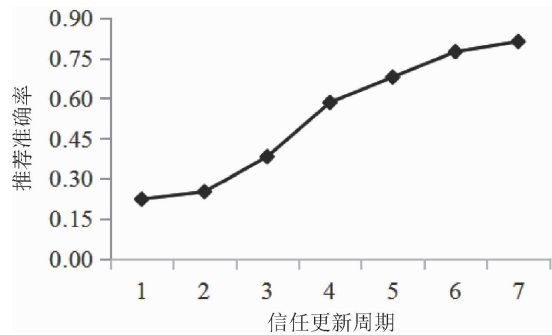


图 4 准确率时序变化分析

Fig. 4 Analysis of accuracy changes with time sequence

在课程的前 3 次更新周期中, 由于学生基数较大, 而学习投入和交互行为关系矩阵数据稀疏, 推荐准确率偏低; 从第 4 次开始, 推荐准确率明显得到提升, 这是因为随着学习投入和交互行为数据的丰富, 形成的关系边及信任网络逐步完善, 同时, 部分浅尝辄止的学员不断退出, 有效行为数据更加聚集, 算法性能得到体现.

(4) 不同情境的影响权重. 本文考察比较了混合法、只考虑用户情境和只考虑社会情境 3 种方法的推荐效

果,并计算推荐列表个数分别为 4、8、12、16 等多种情况下的综合推荐效果 $F1$ 值,结果如图 5 所示.

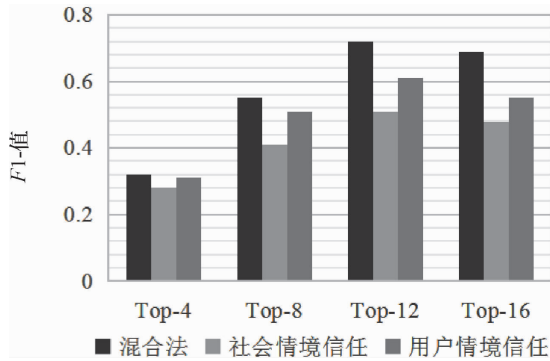


图 5 推荐个数及不同方法下的综合准确率

Fig. 5 Comprehensive accuracy rate under different recommend numbers and methods

实验发现,在推荐个数为 12 左右时具有最佳的推荐效果.综合用户情境和社会情境的混合法反映了学习投入和互动关系,其推荐效果要远高于只考虑互动关系的社会情境方法,这是因为 MOOC 是一种内容型社区,学伴间产生连接的主要前提是对学习内容的投入程度和兴趣偏好.与只考虑用户情境方法相比,混合法的推荐效果也要稍高,在 top 值增大时变得更为明显,这是因为用户情境法局限于依靠学习投入和兴趣相似度推荐学伴,无法在用户数据增加时利用交互关系构建连接拓扑网络来拓展学伴圈子,导致数据稀疏,从而影响其推

荐效果.

为进一步确定用户情境和社会情境对综合推荐效果的影响权重,按 Top-12 标准,对调和参数 η 从 0 到 1 按 0.1 递增计算综合指标 $F1$ 值,结果绘制如图 6 所示.从图 6 可知,当 $\eta = 0.6$ 时具有最优的推荐效果,也即在 MOOC 学伴推荐中,用户情境占有一定的主导作用.

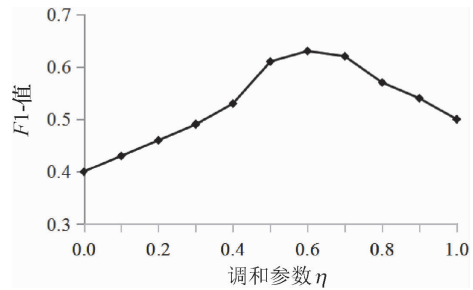


图 6 信任调和参数对推荐效果的影响

Fig. 6 Influence of trust reconciliation parameters on recommendation effect

(5) 不同应用场景的适配

本文提出的方法综合考虑了学习投入、时间、位置、交互行为、频次及时长等情境因素,通过对多维权重值的动态调整,可以实现具体应用场景的灵活应用.本文在上述 MOOC 数据集上进行了对照实验,实验组是本文提出融合多类情境的学伴推荐模型,对照组是文献[23]提出基于情境感知的移动社交网络好友推荐方法,实验结果如表 2 所示.

表 2 不同应用场景的参数设置及意义

Table 2 Parameter setting and significance of different application scenarios

方法	情境	参数	$F1$ 值	推荐场景
本文方法	Case1: 仅考虑位置	$\alpha = 0, \beta = 1, \gamma = 0$	0.221	促进线上走向线下交互
	Case2: 仅考虑时间	$\alpha = 0, \beta = 0, \gamma = 1$	0.316	促进异步转向同步交互
	Case3: 仅考虑学习行为与评分	$\alpha = 1, \beta = 0, \gamma = 0$	0.563	基于相似兴趣和知识水平的推荐
	Case4: 仅考虑交互频次	$b = 0$	0.421	基于交互活跃度的推荐
	Case5: 仅考虑交互时长	$b = 1$	0.446	基于交互深度的推荐
	融合上述所有情境因素		0.612	
文献[23]方法	Case1: 仅考虑位置	$\lambda = 0$	0.201	处于相近位置的推荐
	Case2: 仅考虑时间	$\lambda = 1$	0.312	基于通信频次和时长的推荐
	融合上述所有情境因素		0.435	

从表 2 可见,两种方法均可以实现学伴(好友)推荐,但在 $F1$ -measure 比较中,本文提出的推荐模型具有更好的效果.本文方法可以在不同类型的课程(或同一课程不同阶段)调整权重参数获得更优推荐.如针对交互要求较少的课程(或课程前阶段),可适当增大调和参数 η 值以增加学习投入行为对推荐的比重;针对交互行为较多的课程(或课程后阶段),可适当减少 η 值以增加互动行为和互动关系对推荐的比重.

5 结 语

本文基于探究社区模型,分析和梳理了 MOOC 社区中影响学伴推荐的情境因素,提出了一种结合用户情境和社会情境构建学员间信任网络并实现 Top-N 推荐的方法.本文的创新之处在于,在用户情境中引入学习者对学习内容的行为投入量化和计算,反映了学员的兴趣

偏好;社会情境在学员间交互关系中引入奖励因子和时间衰减因子,反映了信任关系的动态变化性.实验数据表明,该方法对具有高动态性和移动性的 MOOC 学伴推荐具有较好的效果,并且对多维参数权重值的动态调整可以实现不同推荐场景的灵活应用.学习投入包含行

为、认知和情感3个维度,本研究的不足之处在于未考虑认知和情感投入对学伴推荐的影响,下一步工作是挖掘 MOOC 社区中学习者的认知投入和情感投入及其量化方法,并纳入信任度构建模型,进一步提升 MOOC 社区学伴推荐的准确度.

参考文献:

- [1] Ardchir S, Talhaoui M A, Jihal H, et al. Predicting MOOC dropout based on learner's activity features[J]. *International Journal of Engineering and Technology*, 2018, 7(32):124-126.
- [2] Yeonji J, Jeongmin L. Learning engagement and persistence in Massive Open Online Courses(MOOCs)[J]. *Computers & Education*, 2018, 122:9-22.
- [3] 吴华君,葛文双,何聚厚.教师支持对 MOOC 课程持续学习意愿的影响研究——基于 S-O-R 和 TAM 的视角[J]. *现代远程教育*, 2020, 180(3):89-96.
- [4] Hew K F. Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCs[J]. *British Journal of Educational Technology*, 2016, 47(2):320-341.
- [5] Li Q J, Baker R. The different relationships between engagement and outcomes across participant subgroups in massive open online courses[J]. *Computers & Education*, 2018, 12(127):41-65.
- [6] Siemens G. Connectivism: Learning as network-creation[EB/OL]. [2019-05-23]. <http://www.elearnspace.org/Articles/networks.htm>.
- [7] 张婧婧,杨业宏,王焯宇,等.国际视野中的在线交互与网络分析:回顾与展望[J]. *电化教育研究*, 2019, 40(10):26-34.
- [8] 徐彬,杨丹,张昱,等.基于学习者行为特征的 MOOCs 学习伙伴推荐[J]. *计算机科学与探索*, 2015, 9(1):71-79.
- [9] 蔡霞,徐舒凡,敬懿,等.基于混合式学习的同伴推荐算法设计与实现[J]. *计算机时代*, 2020(4):38-40.
- [10] Hugues L, François B, Rémi B, et al. Does a peer recommender foster students' engagement in MOOCs'[C]//*International Educational Data Mining Society, Proceedings of the 9th International Conference on Educational Data Mining*. New York: ACM, 2016:418-423.
- [11] François B, Hugues L, Kalina Y, et al. Comparing peer recommendation strategies in a MOOC[C]//*Adaptation and Personalization, Proceedings of the 25th Conference on User Modeling*. New York: ACM, 2017:129-134.
- [12] 刘海鸥,刘旭,姚苏梅,等.基于大数据深度画像的个性化学习精准服务研究[J]. *图书馆学研究*, 2019, 15(21):68-74.
- [13] Elghomary K, Bouzidi D. Dynamic peer recommendation system based on trust model for sustainable social tutoring in MOOCs[C]//*IEEE, 2019 1st International Conference on Smart Systems and Data Science (ICSSD)*. Piscataway: IEEE, 2019:1-9.
- [14] Garrison D R, Arbaugh J B. Researching the community of inquiry framework: Review, issues and future directions[J]. *The Internet and Higher Education*, 2007, 4(10):157-172.
- [15] Garrison D R, Akyol Z. Toward the development of a metacognition construct for communities of inquiry[J]. *The Internet and Higher Education*, 2015, 1(24):66-71.
- [16] Sánchez-Rada J F, Iglesias C A. Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison[J]. *Information Fusion*, 2019, 12(52):344-356.
- [17] 俞国良. *社会心理学*[M]. 北京:北京师范大学出版社, 2015.
- [18] Pekrun R, Linnenbrink-Garcia L. Academic emotions and student engagement[M]. *Handbook of Research on Student Engagement*. Boston, MA:Springer, 2012.
- [19] 廖宏建,谢亮,曲哲.一种基于隐式信任感知的 MOOC 推荐方法[J]. *情报理论与实践*, 2021, 44(2):97-104.
- [20] 朱文强,徐军.融合社交信息的多属性元路径好友推荐方法[J]. *小型微型计算机系统*, 2020, 41(10):2036-2044.
- [21] 龙增艳,陈志刚,徐成林.基于用户交互的社交网络好友推荐算法[J]. *计算机工程*, 2019, 45(3):132-137.
- [22] 朱强,孙玉强.一种基于信任度的协同过滤推荐方法[J]. *清华大学学报(自然科学版)*, 2014, 54(3):360-365.
- [23] 孔聪聪,陈曙东.上下文感知的移动社交网络好友推荐算法[J]. *小型微型计算机系统*, 2018, 39(1):23-26.