

网络结构视角下用户异质性对知识协作效果的影响研究

石燕青¹, 李露¹, 是沁^{2*}

(1. 南京农业大学 信息管理学院, 南京 210095; 2. 南京中医药大学 卫生经济管理学院, 南京 210023)

摘要: [目的/意义]探讨如何优化在线知识协作平台网络结构和平衡团队内部的知识与经验, 为促进跨领域信息流通, 避免信息孤岛的形成, 通过集体知识协作, 推动知识的创造、传播和应用。[方法/过程]本研究利用编程问答网站 Stack Overflow 的数据, 筛选出含“Python”标签的问题帖和对应的回答帖, 结合负二项回归模型探讨了网络结构特征和用户异质性对知识协作质量和效率的影响。[结果/结论]网络结构特征方面, 节点中心性显著提升了协作的质量和效率, 较高的聚集系数和较大的连边跨度限制了信息流通, 不利于知识协作效率。用户异质性方面, 知识背景和注册时长的高异质性通常阻碍协作, 经验异质性中注册时长异质性对协作效果均产生负面影响, 回答被采纳率异质性仅对协作质量产生了负面影响, 而活动强度异质性则产生了积极影响。

关键词: 知识协作效果; 社会网络结构; 用户异质性; 信息行为

中图分类号: G252

文献标识码: A

文章编号: 1002-1248 (2024) 03-0072-11

引用本文: 石燕青, 李露, 是沁. 网络结构视角下用户异质性对知识协作效果的影响研究[J]. 农业图书情报学报, 2024, 36(3): 72-82.

0 引言

在数字化时代, 知识协作平台如在线问答社区、学术论坛以及各类专业网络平台已成为知识共享和集体智慧发挥作用的重要场所。平台聚集了不同领域、不同专业背景和经验水平的用户, 用户共同参与问题解答、观点交流, 形成了复杂且动态的社交网络。基于该环境, 知识的创建、传播和应用不再是个别专家

的孤立行为, 而是通过集体智慧的互动和协作实现。在线知识协作平台不仅提高了知识的可获取性, 而且通过集合不同个体的智慧和专长, 促进了知识的多样性和创新。这些平台成为了跨学科交流、问题解决和创新思维碰撞的孵化器。

在知识协作平台中, 多领域多背景用户一方面促进异质性知识融合、重组与创新; 另一方面, 由于参与者的背景、经验和观点的差异性, 如何有效地管理和协调

收稿日期: 2024-02-16

基金项目: 江苏省社会科学基金青年项目“社交网络中基于交互特征的学术知识扩散研究”(20TQC003); 江苏省社会科学基金青年项目“多主体协同下古籍数字化价值共创模式研究”(23TQC009)

作者简介: 石燕青, 女, 博士, 副教授, 研究方向为网络科学与科技数据挖掘。李露, 女, 硕士研究生, 研究方向为社交媒体数据挖掘

*通信作者: 是沁, 女, 博士, 讲师, 研究方向为数据管理。Email: 280305@njucm.edu.cn

成为了一个关键问题。此外, 知识协作的质量和效率还受到网络结构特征的影响, 例如网络中信息流动的路径、关键节点的作用以及小团体间的互动方式等。随着知识增长与创新逐渐趋于知识资源集成化和行为主体协同化, 社群关系网络已经成为社群知识协作的载体^[1]。社群关系结构不仅会影响社群知识协作效率和个体知识表现, 而且对社群个体之间的知识交互过程也会造成影响^[2]。比如研究发现度分布以及节点之间的知识势差是网络中知识扩散效率高低的决定性因素, 就证实了网络结构差异对知识扩散绩效差异的作用^[3]。

基于此, 在社会网络的框架下, 本研究将采用全球最大的编程问答网站 Stack Overflow 的真实数据, 基于用户协作关系构建知识协作网络, 探讨网络结构异质性和用户异质性如何作用于知识协作效果。

1 文献综述

1.1 社会网络中的弱关系和长连边

社会网络是由行动者和行动者间的关系构成的集合, 反映了处于网络中的成员信息交流、协同合作、知识学习等方面的动态行为, 对成员之间的交流促进、资源利用以及知识流动起到了关键的作用。美国著名社会学家 GRANOVETTER 根据人的互动频率、感情力量、亲密程度、互惠交换将社会关系分为强关系和弱关系^[4]。强关系维系着群体、组织内部的关系, 具有长期性、稳定性和高信用度; 而弱关系跨越不同的信息源, 能作为信息和资源的桥梁, 促进不同群体间的信息流动^[5]。研究表明虚拟社区的知识共享网络通常是稀疏的, 以弱联结为主, 成员间的互动属于非直接互惠的一般的交互^[6,7]。

近年来, 弱关系理论衍生出了新的社会网络分析视角——长连边关系, 这种跨越较长路径的关系在社交网络中被视为有价值的弱联系。当两节点间直连关系被切断时, 一个节点需要经过第二短路径到达另一个节点, 当第二短路径足够长时, 两节点间的关系将被称为“长连边关系”。当个体与不同社交圈子中的个

体都有联系时, 长连边关系能够提供更广泛的传播路径, 加速信息在不同社交圈子间的传播^[8]。长连边关系在网络中扮演着“社会桥”的角色, 促进异质性个体提供新的信息或不同的专业知识, 这些关系虽需时间维系, 却因其直接的互惠性而具有稳定性^[9]。同时, 长连边结构也可以避免信息层次的同质化和网络结构的冗余^[10]。

1.2 在线知识协作网络结构研究

在线知识协作平台中用户之间的行为与关系都可以被视为社会网络进行研究, 社会网络分析等理论应用到了对在线知识协作的结构分析上, 揭示了 CSDN、百度百科、Wikipedia 等在线知识协作社区中的小世界、无标度等复杂网络特性^[11,12]。学者们通过研究这种结构对群体行为产生的影响, 明确用户知识共享过程中的交互过程。在线知识协作社区内容的生成和完善是大量自愿参与者集体协作行为的结果, 许多研究表明主体之间的合作质量、合作绩效与网络结构密切相关^[13,14]; 基于 Wikipedia 的研究表明适当的联结与内聚, 以及少数核心成员的影响与领导, 都有助于提高网络的构建效率^[15]; 基于“内容-贡献者”构建的二模网络显示网络嵌入程度和用户人数的适当增加有利于提高内容的价值, 网络的度中心性和特征向量中心性都与文章质量成正比^[16]。

综上所述, 知识协作社区中会形成社团结构, 社团内部关系紧密, 而不同社团间关系较弱。长连边关系尽管在接触频率或强度方面看似较弱, 但常作为不同社区间的网络桥梁, 促进知识和资源的流动。此外, 知识贡献者在协作网络中的网络特征与其知识贡献具有相关性, 合理的网络结构和核心成员的活跃参与可以显著提高网络的效率和输出质量。通过研究社会网络结构对群体行为产生的影响, 有助于明确用户知识共享中的交互过程。

2 相关理论与研究假设

2.1 用户重要性与知识协作效果

社会资本理论认为, 个体在社会结构中所处的位

置会为其带来不同的社会资源^[17,18]。在线知识共享平台的各项研究均显示,在信息传递、知识协作和资源获取等社会资本的不同维度都对用户行为有显著影响^[19,20]。在以不同类型的用户为节点组成的社交网络中,占据更核心、更重要的位置,或与更多成员建立合作关系,能够增加异质性知识的获取机会^[21]。在网络分析中,节点重要性常常通过中心性指标来衡量,节点在网络中的位置能影响他们接触和控制信息的能力,以及他们对网络内其他节点的影响力。核心节点的作用至关重要,弱关系的传播主要集中在少数核心节点之间,他们往往起到“桥梁”的作用^[7]。已有多项研究发现,节点的中介中心性是复杂合作环境中绩效的主要影响因素,对于社区成长起到正向作用^[22],度数中心性在资源密集的平台型企业群中也能够正向影响创新绩效^[23]。因此,本研究提出假设:

H1a: 节点中心性会正向影响知识协作质量;

H1b: 节点中心性会正向影响知识协作效率。

2.2 用户关系密度与知识协作效果

用户之间的关系密度常常通过聚集系数指标来衡量,聚集系数可以衡量网络中的小世界特性,较高的聚集系数被认为可以提高网络成员间的信息传输效率,有助于合作创新活动的开展^[24]。在高密度的知识网络中,各类知识元素能更有效地整合重构,进而转化为创新成果^[21],促进知识的交流和共享。此外,有实证结果检验了较短的网络平均路径长度有利于企业创新绩效的提高,聚集系数与企业的创新绩效呈U型关系^[25]。因此提出假设:

H2a: 聚集系数会正向影响知识协作质量;

H2b: 聚集系数会正向影响知识协作效率。

2.3 用户关系跨度与知识协作效果

利用人口规模网络数据的研究发现长连边与更紧密社交圈子中的短连边几乎同样强大,表明这些远程联系在信息传播和社会融合过程中扮演着关键角色^[8]。另外,基于两年社交网络数据的实证分析和计算建模,研究揭示了联系强度(通过互动频率和时长衡量)与

连边跨度之间的“U型”关系,互动频率和时长随着联系范围的增加而减少,但之后会增加^[9]。高连边跨度可能连接更多异质的群体,从而带来新的信息和观点,增加知识协作的创新性,而高连边跨度可能导致知识协作效率下降,因为节点之间的连接路径较长,信息需要更长的时间在网络中传播。因此提出假设:

H3a: 连边跨度会正向影响知识协作质量;

H3b: 连边跨度会负向影响知识协作效率。

2.4 用户知识差异与知识协作效果

用户知识差异性指团队成员在知识背景、知识结构和认知方式上的差异性,包括他们所具备的专业知识和方法的差异^[26]。在线知识协作平台通常汇集了各学科背景的用户,形成以异质性为基础,资源共享为核心的弱关系网络^[27]。知识异质性有助于生成新的想法和解决方案,丰富团队的知识基础^[16],知识其带来的思维模式和技术路径上的多样性成为团队研发活动的重要驱动力^[28],对突破性创新具有正向影响^[29,30]。团队能够在决策时考虑更多的信息和观点,做出更全面的判断。但同时知识异质性可能会导致沟通障碍,由于成员需要将异质性知识加以整合,会激发其进行更深入的思考,阻止团队过早达成共识^[31]。因此提出假设:

H4a: 知识异质性会正向影响知识协作质量;

H4b: 知识异质性会负向影响知识协作效率。

2.5 用户经验差异与知识协作效果

用户经验差异性,即成员在专业技能、知识深度及解决问题经验上的差异^[32],对知识协作平台上的团队动态产生深远影响,它涉及了成员在特定任务中的经验水平和时间投入的差异。在知识协作中,经验的差异性既是一个优势也可能成为障碍。信息采纳理论(Information Adoption Model)表明,当信息质量不容易判定时,信息阅读者往往会参考信息来源的可信度,进而间接感知信息价值并决定是否接受该信息^[33,34]。对在线评论的研究结果显示评论者的经验能力会提升其答案的可信度,其发表的评论更易于被用户接受^[35],还有相关研究表明发帖数量与频率在一定程度上反映

了该用户发布信息对于其他用户的态度与观点更加具有正向影响^[6]。然而在某些情况下, 经验丰富的用户能够迅速识别和解决问题, 经验较少的用户可能要花更多的时间来理解问题, 则可能降低整体的效率。因此提出假设:

H5a: 经验异质性会正向影响知识协作质量;

H5b: 经验异质性会负向影响知识协作效率。

综上所述, 本文提出的研究模型如图 1 所示。

3 研究设计与方法

3.1 数据收集

本文选用 Stack Overflow 平台的开源数据集, 聚焦高活跃度的标签“Python”, 筛选了 2020 年内有被采纳回答且至少有 3 条回答的问题帖 3 862 条, 清理缺失数据后得到 3 726 条问题帖和 12 217 条回答帖, 涉及 9 589 位用户。

3.2 变量设计

3.2.1 在线知识协作效果

Stack Overflow 等平台通常依赖社区成员的互动协

作来解决技术问题, 因此知识协作的质量和效率均是需要关注的指标。质量方面, 社区对于问题回答的集体评价能够展现协作成员共同努力的效果, 反映集体智慧在解决问题中的作用, 故本文通过所有回答总得分来衡量知识协作质量, 总得分代表了针对特定问题的所有协作努力得到的社区认可的总量, 得分越高表明社区成员认为回答质量越高; 效率方面, 问题快速得到被采纳的回答通常意味着高效的知识协作, 表明社区能够迅速集中资源和专业知识来解决问题, 故本文通过问题与被采纳回答创建的时间差来衡量知识协作效率, 其时间差越大, 说明知识协作效率越低。

3.2.2 用户网络结构特征

网络结构特征主要体现在节点中心性、聚集系数和连边跨度 3 个方面。节点中心性选用度中心性, 同时以用户回答相同问题的次数作为权重。以问题作为研究单元, 计算该问题下所有回答者的度中心性、聚集系数和连边跨度的平均值。其中连边跨度 (Tie Range) 由两个相连节点之间的第二短路径的长度表征。在用户协作网络中切断两个节点之间的直接联系 (即他们之间的最短路径长度为 1), 则将这个最短路径长度设为无穷大 (Inf); 再计算新的用户网络中所有节点对之间的最短路径长度, 其结果即为连边跨度值。

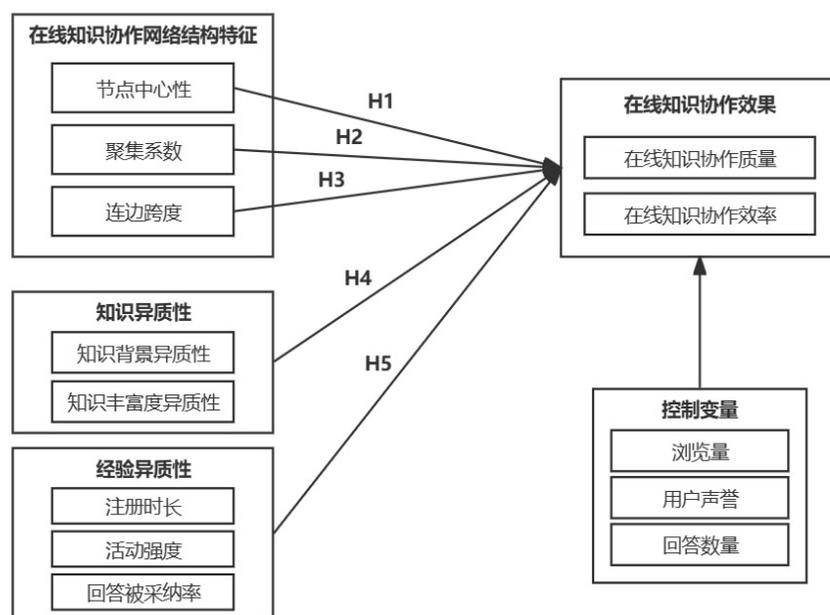


图 1 研究模型

Fig.1 Research model

3.2.3 用户异质性特征

3.2.3.1 用户知识异质性

Stack Overflow 提供了每个问题帖的标签, 这些标签是问题涉及知识领域的映射, 反映了用户的知识和兴趣点。首先, 本研究采集了 2020 年所有问题帖共 2 104 260 条, 并提取出与每个问题相关联的标签 (Tags); 然后以标签为节点, 根据标签之间的共现关系构建词共现网络; 接着采用 Louvain 算法识别共现频率高的标签群组, 对标签进行有效聚类。基于以上操作, 将上述问题帖中涉及的 44 649 个标签聚类成 12 个类, 代表最为活跃和显著的知识领域, 最后统计回答者分别参与 12 个类别标签的数量, 在此基础上计算以下指标。

(1) 知识背景异质性。知识背景异质性指参与回答的用户群体中, 每位用户所代表的知识背景的差异。汇总回答者所参与过的所有 Tag 记录, 去除重复标签, 统计 Tag 分别所属的类别数量, 生成 Tag 向量; 再计算每条问题帖下用户组中每对用户之间的 Tag 向量的异质性 (1-余弦相似度); 最后计算所有用户对异质性平均值。计算公式为:

$$Heterogeneity = 1 - \frac{1}{N} \sum_{i=1}^n (1 - \text{CosineSimilarity}(A_i, B_i)) \# \quad (1)$$

其中 N 是每条回答帖中用户对的数量, A_i 和 B_i 是每对用户的 Tag 向量。

(2) 知识丰富度异质性。知识丰富度异质性衡量回答者使用 Tag 类别的丰富程度。本研究采用香农多样性指数对其进行量化, 该指数考虑了用户参与 Tag 类别的丰富度及其分布的均匀性, 即一个用户的知识是否均匀分布在多个领域上, 或者是否集中在某几个领域。指数高意味着用户具有较广泛的知识面和较高的知识丰富度。计算公式为:

$$H' = - \sum_{i=1}^R (p_i \cdot \log(p_i)) \# \quad (2)$$

R 是用户参与的不同 Tag 类别的总数, p_i 是用户在第 i 个 Tag 类别中的参与频率与其在所有 Tag 类别中的总参与频率之比。

计算每条问题下所有回答者知识丰富度的离散系数作为用户组异质性的量化指标。

3.2.3.2 用户体验异质性

经验异质性通过回答者的注册时长、活动强度以及回答被采纳的频率来量化。首先, 注册时长以天为单位, 计算回答者从注册平台到回答特定问题时的时间跨度; 其次, 活动强度根据用户回答次数和评论次数总和来衡量; 再次, 回答被采纳率依据每位回答者被标记为 “Accepted Answer” 数量占其回答总数的比例。最后, 使用 3 项个体特征指标的离散系数来计算用户组的异质性。

与此同时, 本文将每条问题的浏览量、回答数量和用户声誉异质性作为控制变量。

4 实证分析

4.1 在线知识协作网络构建

根据回答者之间的关系构建在线知识协作网络, 以回答者为节点, 回答同一问题的关系作为边, 边的权重为共同回答的问题数量。在此基础上, 计算主要的整体性网络指标, 结果如表 1 所示。从网络密度 ($D=0.001$) 来看, 该网络较为稀疏, 说明回答者之间的合作较少; 从平均聚集系数 ($C=0.824$) 和平均最短路径 ($d=4.723$) 来看, 该网络具有较大的平均聚集系数和较小的平均最短路径, 具有明显的小世界特征; 较高的模块化程度 ($Q=0.746$) 也表明网络中具有明显

表 1 Stack Overflow 平台在线知识协作网络指标

Table 1 Indicators of online knowledge collaboration network on stack overflow platform

节点总数/个	边总数/条	平均度	平均加权度	平均聚集系数	平均最短路径	直径	密度	模块化程度
6 741	14 098	2.091	2.116	0.824	4.723	13	0.001	0.746

的社区结构,且社区内部节点联系密集。

在分析知识协作效果影响因素之前,对主要变量进行了相关性分析,结果显示因变量与各自变量之间具有较高的相关性。此外,对自变量进行方差膨胀因子检验,结果显示方差膨胀因子的最大值为 3.725,平均值为 1.94,表明变量间不存在多重共线性。

4.2 在线知识协作效果影响因素分析

为进一步检验知识协作效果和网络结构特征、用户异质性特征之间的关系,本文使用 R 软件对数据进行回归分析。由于回答得分和问答间隔时长(以分钟为单位)均为计数型变量,因此选择计数模型。在计数模型中,泊松模型的适用条件为因变量的均值和方差相等。基于此假设,对问题得分及回答数量分别进行 K-S 假设检验,检验结果表明显著性系数均小于 0.05,因而变量不符合泊松分布。同时,两个变量的数值分布离散程度较大,因此,本文最终使用负二项回归模型进行分析与验证,回归分析结果如表 2 和表 3 所示。

从结果中可以看出,AIC 随着变量的加入而逐渐

减小, $2 \times \log\text{-likelihood}$ 值(对数似然函数值的两倍)随着变量的加入而逐渐增加,说明随着变量维度加入,在线知识协作效果模型的拟合程度越来越好。如表所示,模型 4 和模型 8 具有最低的 AIC 值和最大的对数似然值,这表明纳入所有变量的模型 4 和模型 8 拟合效果最好,解释性最强。

4.2.1 在线知识协作质量影响因素分析

表 2 展示了知识协作质量的影响因素回归分析结果。模型 1 是仅引入控制变量的基线模型,问题浏览量($\beta=5.12E-05$, $p<0.001$)和回答数量($\beta=1.21E-01$, $p<0.001$)均对知识协作质量有显著的正向影响,声誉值异质性($\beta=-7.35E-02$, $p<0.001$)对其有显著的负向影响。

模型 2 加入知识异质性,知识背景异质性对知识协作质量具有显著的负向影响($\beta=-5.36E-01$, $p<0.001$),知识丰富度异质性对知识协作质量具有显著的负向影响($\beta=-4.67E-02$, $p<0.1$),假设 H4a 未得到验证。

模型 3 加入经验异质性,注册时长异质性($\beta=-6.23E-02$, $p<0.001$)和回答被采纳率异质性($\beta=-1.01E-01$, $p<0.001$)对知识协作质量有显著的负向

表 2 用户异质性和网络结构特征对在线知识协作质量的回归结果

Table 2 Regression results of user heterogeneity and network structure characteristics on online knowledge collaboration quality

项目	模型 1		模型 2		模型 3		模型 4	
	Coef	z	Coef	z	Coef	z	Coef	z
问题浏览量	5.12E-05***	16.83	5.19E-05***	17.55	5.28E-05***	18.06	5.32E-05***	18.40
回答数量	1.21E-01***	12.44	1.19E-01***	12.38	1.19E-01***	12.37	1.22E-01***	12.59
声誉值异质性	-7.35E-02***	-4.74	-1.24E-02	-0.58	-3.95E-02	-1.53	-6.44E-02*	-2.47
知识背景异质性			-5.36E-01***	-8.99	-3.09E-01***	-4.60	-2.43E-01***	-3.48
知识丰富度异质性			-4.67E-02	-1.86	-6.13E-02*	-2.11	-2.93E-02	-1.00
注册时长异质性					-6.23E-02***	-3.61	-6.28E-02***	-3.65
活动强度异质性					6.20E-02*	2.14	4.11E-02	1.41
回答被采纳率异质性					-1.01E-01***	-5.15	-7.99E-02***	-3.95
聚集系数							7.47E-02	1.09
度中心性							2.07E+01***	6.26
连边跨度							-9.42E-02*	-2.27
AIC		17 996		17 878		17 831		17 781
$2 \times \log\text{-likelihood}$		-17 985.64		-17 864.03		-17 810.6		-17 755.13

注: $P<0.1$ 为“.”, $P<0.05$ 为“”, $P<0.01$ 为“***”, $P<0.001$ 为“****”

表3 用户异质性和网络结构特征对在线知识协作效率的回归结果

Table 3 Regression results of user heterogeneity and network structure characteristics on online knowledge collaboration duration

项目	模型 5		模型 6		模型 7		模型 8	
	Coef	z	Coef	z	Coef	z	Coef	z
问题浏览量	1.20E-03***	37.96	1.35E-03***	43.16	1.34E-03***	42.67	9.94E-04***	31.93
回答数量	-8.47E-02	-1.35	-1.35E-02	-0.22	1.82E-02	0.29	1.84E-01**	2.97
声誉值异质性	2.08E-01*	2.24	2.54E-01*	1.97	3.73E-01*	2.42	4.99E-01**	3.26
知识背景异质性			3.96E+00***	11.54	3.75E+00***	9.66	2.16E+00***	5.39
知识丰富度异质性			-5.17E-01***	-3.48	-4.67E-01**	-2.71	-1.16E+00***	-6.72
注册时长异质性					2.01E-01*	1.97	1.58E-01	1.57
活动强度异质性					-2.96E-01	-1.71	2.07E-01	1.20
回答被采纳率异质性					1.53E-01	1.34	-2.33E-01*	-1.99
聚集系数							4.54E+00***	11.25
度中心性							-1.17E+01**	-0.58
连边跨度							1.35E+00***	5.64
AIC	55 357		55 238		55 234		55 087	
2×log-likelihood	-55 346.763		-55 223.841		-55 213.548		-55 061.045	

注: $P<0.1$ 为 “·”, $P<0.05$ 为 “”, $P<0.01$ 为 “***”, $P<0.001$ 为 “****”

作用。活动强度异质性则是起到 5%水平上的显著正向影响 ($\beta=6.20E-02$, $p<0.05$)。假设 H5a 部分得到验证。

模型 4 加入网络结构特征, 节点中心性 ($\beta=2.07E+01$, $p<0.001$) 对知识协作质量有显著的正向作用, 假设 H1a 得到验证; 聚集系数对知识协作质量的正向影响不显著, 假设 H2a 未得到验证; 连边跨度 ($\beta=-9.42E-02$, $p<0.05$) 对知识协作质量有显著的负向作用, 假设 H3a 未得到验证。

4.2.2 在线知识协作效率影响因素分析

表 3 展示了知识协作时长的影响因素回归分析结果。协作时长越长, 则协作效率越低, 自变量正向影响知识协作时长即意味着负向影响知识协作效率。

模型 5 仅引入控制变量的基线模型, 问题浏览量 ($\beta=1.20E-03$, $p<0.001$) 和声誉值异质性 ($\beta=2.08E-01$, $p<0.05$) 对知识协作效率有显著的负向影响。

模型 6 中加入知识异质性, 知识背景异质性对知识协作效率具有显著的负向影响 ($\beta=3.96E+00$, $p<0.001$), 知识丰富度异质性对知识协作效率具有显著的正向影响 ($\beta=-5.17E-01$, $p<0.001$), 假设 H4b 部分得到验证。

模型 7 加入经验异质性, 注册时长异质性 ($\beta=$

$2.01E-01$, $p<0.05$) 对知识协作效率有显著的负向作用, 活动强度异质性 ($\beta=-2.96E-01$, $p<0.1$) 对知识协作效率有显著的正向作用。

模型 8 加入网络结构特征, 节点中心性 ($\beta=-1.17E+01$, $p<0.01$) 对知识协作效率有显著的正向作用, 假设 H1b 得到验证。聚集系数 ($\beta=4.54E+00$, $p<0.001$) 对知识协作效率有显著的负向作用, 假设 H2b 未得到验证。连边跨度 ($\beta=1.35E+00$, $p<0.001$) 对知识协作效率有显著的负向作用, H3b 得到验证。另外可以观察到, 活动强度异质性不再显著正向影响知识协作效率, 相反回答被采纳率异质性显著正向影响知识协作效率 ($\beta=-2.33E-01$, $p<0.05$)。

综合上述结果来看, 对于知识协作质量, 知识异质性产生负向影响, 经验异质性中仅有活动强度异质性对其有正向影响; 网络结构特征中, 节点中心性对知识协作质量有正向影响, 聚集系数和连边跨度的影响不显著。对于知识协作效率, 知识异质性中知识背景异质性起负向作用, 而知识丰富度异质性起正向作用, 经验异质性中注册时长异质性起负向作用, 回答被采纳率异质性起正向作用; 网络结构特征中, 节点

中心性对其有正向影响, 聚集系数和连边跨度均对其有负向影响。

5 结论与展望

5.1 研究结论

在线知识协作平台作为知识共享与创新合作的重要场所, 集结了来自世界各地的专家和爱好者, 构成了一个涉及信息流通、知识创新和协作效率的复杂知识网络。特别是在编程和软件开发领域, 在线问答社区 Stack Overflow 为全球的开发者提供了即时、互助的知识交换平台。本文深入探讨了知识协作网络的结构特征以及团队用户之间的知识和经验异质性如何影响知识协作的效果, 研究结论如下:

在网络结构方面, 知识协作在紧密连接的网络中趋向于产生更高质量的输出, 这归因于高聚集系数的网络促进了信任和共同规范的建立。然而遗憾的是, 该类型网络可能不利于协作过程的快速进行。网络的中心节点通过有效获取和分配资源, 促进知识传播, 从而提升协作的质量和效率。跨越多节点的连边虽然可以连接不同知识领域的用户, 以增加知识多样性, 但也在一定程度上增加了沟通、认知、转化成本, 从而影响知识协作效率。

在用户异质性方面, 团队成员在知识背景差异较大时, 沟通和理解上的挑战抵消了多样性带来的潜在创新优势, 知识协作效率呈现下降趋势。不同的知识和技能一方面可以实现知识补充、知识融合、知识转化; 另一方面也会增加协调、整合的时间与成本。与此同时, 经验差异较大的成员在沟通和协作中遇到更多障碍, 导致效率和效果持续下降。此外, 被频繁采纳的回答增加了信息的信任度, 加快了知识的接受和应用速度。

针对本研究的结果, 可为知识协作平台提供服务优化建议。首先, 完善激励机制, 鼓励用户积极参与回答和讨论, 通过优化算法推荐系统, 提升中心用户的回答可见度, 从而加强知识传播的效率; 其次, 采用技术手段和政策引导, 鼓励用户分享多样化的知识并回答跨领域问题, 增强网络连通性和多样性, 避免

信息过度聚集, 促进网络结构的优化, 并简化专家贡献知识的流程; 最后, 鼓励用户间的互动与交流, 提供社交功能和对新用户的培训和指导, 帮助他们快速融入社区, 提升协作效率。

5.2 研究不足与展望

虽然本研究在探究网络结构异质性和用户异质性对知识协作效果的影响方面取得了一些有意义的发现, 但仍存在不足及值得进一步探讨之处。

(1) 本研究仅以 Stack Overflow 平台中 2020 年“Python”相关问题及回答者作为研究对象, 未来的研究可以考虑扩大样本范围, 涵盖更多不同主题和领域的问题, 以增加研究结果的可靠性和普适性。

(2) 未来的研究可关注网络结构和异质性的动态变化, 一方面分析用户在不同时间点的行为模式和网络连接情况, 以更好理解网络结构对知识协作的影响, 发现网络演化规律, 提高协作效果的预测能力; 另一方面可更深入地研究不同类型的异质性如何随时间影响协作动态, 例如通过追踪个别问题和答案的演变过程。

参考文献:

- [1] PODOLNY J. Networks as the pipes and prisms of the market[J]. American journal of sociology, 2001, 107(1): 33-60.
- [2] MORONE P, TAYLOR R. Knowledge diffusion and innovation[M]. Cheltenham: Edward Elgar Publishing, 2010.
- [3] MUELLER M, BOGNER K, BUCHMANN T, et al. The effect of structural disparities on knowledge diffusion in networks: An agent-based simulation model[J]. Journal of economic interaction and coordination, 2017, 12(3): 613-634.
- [4] GRANOVETTER M S. The strength of weak ties[M]//Networks in the knowledge economy. Oxford: Oxford University Press, 2003.
- [5] 荀亚玲, 毕慧敏, 张继福. 基于弱关系的异质社交网络推荐[J]. 计算机工程与设计, 2021, 42(6): 1526-1534.
- [6] XUN Y L, BI H M, ZHANG J F. Heterogeneous social network recommendation based on weak ties[J]. Computer engineering and design, 2021, 42(6): 1526-1534.
- [6] 雷静, 吴晓伟, 杨保安. 虚拟社区中的公共知识与知识共享网

- 络[J]. 情报杂志, 2012, 31(3): 145-150, 144.
- LEI J, WU X W, YANG B A. Public knowledge and knowledge sharing network in virtual communities[J]. Journal of intelligence, 2012, 31(3): 145-150, 144.
- [7] 李思佳, 郑德铭, 刘博. 考虑节点间强弱关系的突发事件信息传播网络分析[J]. 农业图书情报学报, 2024, 36(1): 83-96.
- LI S J, ZHENG D M, LIU B. Network analysis of emergency information dissemination considering the strength relationship between nodes[J]. Journal of library and information science in agriculture, 2024, 36(1): 83-96.
- [8] PARK P S, BLUMENSTOCK J E, MACY M W. The strength of long-range ties in population-scale social networks[J]. Science, 2018, 362(6421): 1410-1413.
- [9] LYU D, YUAN Y, WANG L, et al. Investigating and modeling the dynamics of long ties[J]. Communications physics, 2022, 5: 87.
- [10] NIAN F Z, REN J H. Higher-order spreading structure in social networks[J]. International journal of modern physics C, 2023, 34(7): 34.
- [11] 达一菲, 刘旭东, 孙海龙. 大数据驱动的开发者社区中知识交流网络的分析[J]. 计算机科学, 2018, 45(9): 113-118.
- DA Y F, LIU X D, SUN H L. Big data driven analysis of knowledge exchange network in developer community[J]. Computer science, 2018, 45(9): 113-118.
- [12] 关希望. 维基百科知识网络结构及编辑行为模式研究[D]. 大连: 大连理工大学, 2020.
- GUAN X W. On the structure of wikipedia knowledge link network and patterns of its editing behavior[D]. Dalian: Dalian University of Technology, 2020.
- [13] 李海林, 徐建宾, 林春培, 等. 合作网络结构特征对创新绩效影响研究[J]. 科学学研究, 2020, 38(8): 1498-1508.
- LI H L, XU J B, LIN C P, et al. Structural features of cooperative network impact on innovation performance[J]. Studies in science of science, 2020, 38(8): 1498-1508.
- [14] 赵健宇, 任子瑜, 裘希. 知识嵌入性对合作网络知识协同效应的影响: 吸收能力的调节作用[J]. 管理工程学报, 2019, 33(4): 49-60.
- ZHAO J Y, REN Z Y, XI X. Influence of knowledge embeddedness on knowledge synergistic effect in collaborative network: Absorptive capacity as a moderator[J]. Journal of industrial engineering and engineering management, 2019, 33(4): 49-60.
- [15] QIN X J, CUNNINGHAM P, SALTER-TOWNSHEND M. The influence of network structures of Wikipedia discussion pages on the efficiency of Wiki projects[J]. Social networks, 2015, 43: 1-15.
- [16] RANSBOTHAM S, KANE G C, LURIE N H. Network characteristics and the value of collaborative user-generated content[J]. Marketing science, 2012, 31(3): 387-405.
- [17] 贾平雷. 社交问答社区讨论线程的交流结构对问题解决的影响研究[D]. 济南: 山东大学, 2023.
- JIA P L. Research on the influence of communication structure of discussion threads on problem solving in social question-and-answer community[D]. Jinan: Shandong University, 2023.
- [18] OBSTFELD D. Social networks, the tertius iungens orientation, and involvement in innovation[J]. Administrative science quarterly, 2005, 50(1): 100-130.
- [19] 陆泉, 刘婷, 邓胜利. 基于社会资本理论的社交问答用户健康信息行为研究[J]. 图书情报工作, 2019, 63(17): 118-127.
- LU Q, LIU T, DENG S L. Research on user's health information behavior on social question and answer based on social capital theory[J]. Library and information service, 2019, 63(17): 118-127.
- [20] 赵杨, 袁析妮, 李露琪, 等. 基于社会资本理论的问答平台用户知识付费行为影响因素研究[J]. 图书情报知识, 2018(4): 15-23.
- ZHAO Y, YUAN X N, LI L Q, et al. The impact factors of users' paying behavior for knowledge on social Q & A platform based on social capital theory[J]. Documentation, information & knowledge, 2018(4): 15-23.
- [21] 程跃, 钟雨珊, 陈婷. 协同创新网络成员和知识多样性对区域创新绩效的影响研究——基于网络结构的调节作用[J]. 创新科技, 2023, 23(6): 66-78.
- CHENG Y, ZHONG Y S, CHEN T. Impact of collaborative innovation network members and knowledge diversity on regional innovation performance - The moderating effect of network structure[J]. Innovation science and technology, 2023, 23(6): 66-78.
- [22] ANTONACCI G, FRONZETTI COLLADON A, STEFANINI A, et al. It is rotating leaders who build the swarm: Social network determi-

- nants of growth for healthcare virtual communities of practice[J]. *Journal of knowledge management*, 2017, 21(5): 1218-1239.
- [23] 周文浩, 李海林. 合作网络异质性特征与企业创新绩效的关系[J]. *系统管理学报*, 2023, 32(2): 367-378.
- ZHOU W H, LI H L. Relationship between heterogeneous characteristics of collaborative network and innovation performance of enterprises[J]. *Journal of systems & management*, 2023, 32(2): 367-378.
- [24] WATTS D J, STROGATZ S H. Collective dynamics of "small - world" networks[J]. *Nature*, 1998, 393(6684): 440-442.
- [25] 曹洁琼, 其格其, 高霞. 合作网络“小世界性”对企业创新绩效的影响——基于中国 ICT 产业产学研合作网络的实证分析[J]. *中国管理科学*, 2015, 23(S1): 657-661.
- CAO J Q, QI G Q, GAO X. The influence of cooperation network's "small universality" on enterprise's innovation performance - An empirical analysis based on industry -university -research cooperation network of ICT industry in China[J]. *Chinese journal of management science*, 2015, 23(S1): 657-661.
- [26] 裘江南, 王婧贤. 在线知识社区中团队异质性对知识序化效率的影响[J]. *情报学报*, 2018, 37(4): 372-383.
- QIU J N, WANG J X. The impact of group heterogeneity on knowledge ordering efficiency in online knowledge communities[J]. *Journal of the China society for scientific and technical information*, 2018, 37(4): 372-383.
- [27] 张莉. 社会网络视域下的用户协同信息行为与图书馆信息服务新趋势[J]. *图书情报工作*, 2012, 56(7): 49-53.
- ZHANG L. Collaborative information behavior and new trend of library information services from the perspective of social network[J]. *Library and information service*, 2012, 56(7): 49-53.
- [28] 孟添天, 柴菁敏, 郑敏钰. 知识异质性对研发团队创造力的影响——知识整合能力的中介作用和主观关系体验的调节作用[J]. *技术经济与管理研究*, 2022(4): 41-45.
- MENG T T, CHAI J M, ZHENG M Y. The influence of knowledge heterogeneity on team creativity: The mediating role of knowledge integration ability and the moderating role of subjective relationship experience[J]. *Journal of technical economics & management*, 2022 (4): 41-45.
- [29] 赵君, 汪惠玉, 刘智强, 等. 高管团队异质性对突破性创新的影响机制研究[J]. *管理学报*, 2023, 20(9): 1303-1312.
- ZHAO J, WANG H Y, LIU Z Q, et al. The influencing mechanism of top management team heterogeneity on radical innovation[J]. *Chinese journal of management*, 2023, 20(9): 1303-1312.
- [30] 王秋博, 葛玉辉. 高管团队异质性与多样型多元化、创新绩效——基于适应能力的中介效应[J]. *经营与管理*, 2024: 1-12. <https://doi.org/10.16517/j.cnki.cn12-1034/f.20230804.003>.
- WANG Q B, GE Y H. Executive team heterogeneity and diverse diversification, innovation performance - Mediating effects based on adaptive capacity[J]. *Management and administration*, 2024: 1-12. <https://doi.org/10.16517/j.cnki.cn12-1034/f.20230804.003>.
- [31] VAN KNIPPENBERG D, SCHIPPERS M C. Work group diversity[J]. *Annual review of psychology*, 2007, 58: 515-541.
- [32] COLLINS R, BLAU P M. Inequality and heterogeneity: A primitive theory of social structure[J]. *Social forces*, 1979, 58(2): 677.
- [33] CHU S C, KAMAL S. The effect of perceived blogger credibility and argument quality on message elaboration and brand attitudes[J]. *Journal of interactive advertising*, 2008, 8(2): 26-37.
- [34] 邢飞, 刘彩华, 柴雪飞, 等. 基于社交平台的老年用户健康信息采纳行为影响因素研究——以微信为例[J]. *农业图书情报学报*, 2022, 34(7): 53-64.
- XING F, LIU C H, CHAI X F, et al. Influencing factors of elderly users' health information adoption behavior based on social platforms: Taking WeChat as an example[J]. *Journal of library and information science in agriculture*, 2022, 34(7): 53-64.
- [35] CHEUNG M Y, LUO C, SIA C L, et al. Credibility of electronic word-of-mouth: Informational and normative determinants of on-line consumer recommendations [J]. *International journal of electronic commerce*, 2009, 13(4): 9-38.
- [36] 丁汉青, 王亚萍. SNS 网络空间中“意见领袖”特征之分析——以豆瓣网为例[J]. *新闻与传播研究*, 2010, 17(3): 82-91, 111.
- DING H Q, WANG Y P. Analyzing "opinion leader" attributes in SNS cyberspace: An investigation of douban.com[J]. *Journalism & communication*, 2010, 17(3): 82-91, 111.

Impact of User Heterogeneity on Knowledge Collaboration Effectiveness from a Network Structure Perspective

SHI Yanqing¹, LI Lu¹, SHI Qin^{2*}

(1. College of Information Management, Nanjing Agriculture University, Nanjing 210095;

2. School of Health Economics and Management, Nanjing University of Chinese Medicine, Nanjing 210023)

Abstract: [[Purpose/Significance] In the context of the digital age, knowledge collaboration platforms such as online Q&A communities, academic forums, and various professional networking platforms have become important venues for knowledge sharing and collective wisdom. These platforms bring together users from different fields, with diverse professional backgrounds and levels of expertise. They actively engage in problem solving, exchange views, and form complex and dynamic social networks. Online knowledge collaboration platforms not only enhance the accessibility of knowledge but also serve as incubators for interdisciplinary communication, problem solving, and innovative thinking by harnessing the collective wisdom and expertise of individuals. This article explores how to optimize the network structure of online knowledge collaboration platforms and balance the internal knowledge and expertise within teams. The goal is to promote cross-domain information flow, prevent the formation of information silos, and promote the creation, dissemination, and application of knowledge through collective knowledge collaboration. [Methods/Process] Due to the diversity of participants' backgrounds, experiences, and viewpoints, effectively managing and coordinating this heterogeneity becomes a critical issue. Additionally, the quality and efficiency of knowledge collaboration is also influenced by the characteristics of the network structure, such as the flow of information paths, the role of key nodes, and the interaction patterns of small groups. This study is based on actual data from Stack Overflow, the world's largest programming Q&A website. It focuses specifically on the following aspects of influence: clustering coefficient, node centrality, edge span, user knowledge heterogeneity, and user experience heterogeneity. By constructing a negative binomial regression model, the study investigates how network structure characteristics and team user heterogeneity affect the quality and efficiency of knowledge collaboration. [Results/Conclusions] The results show that, with respect to network structural characteristics, node centrality significantly improves the quality and efficiency of collaboration, and higher aggregation coefficients and larger span of connecting edges restrict information flow and are detrimental to the efficiency of knowledge collaboration. In terms of user heterogeneity, high heterogeneity in knowledge background and registration duration usually hinders collaboration, heterogeneity in experience heterogeneity in registration duration negatively affects collaboration effectiveness in both cases, heterogeneity in response acceptance rate only negatively affects collaboration quality, while heterogeneity in activity intensity positively affects it. In addition, this study still has shortcomings that deserve further exploration. First, future research could consider expanding the sample to include more questions on different topics and domains to increase the reliability and generalizability of the findings. Second, future research could focus on the dynamic changes of network structure and heterogeneity in order to better understand the impact of network structure on knowledge collaboration and to improve the prediction ability of collaboration effects; it could explore more deeply how different types of heterogeneity affect collaboration dynamics over time.

Keywords: knowledge collaboration outcomes; social network structure; user heterogeneity; information behavior