

基于可信性理论的作者耦合关系强度分析方法

袁国强, 张颂飞

河北金融学院统计与数据科学学院, 河北 保定

收稿日期: 2025年11月11日; 录用日期: 2025年12月5日; 发布日期: 2025年12月10日

摘要

基于可信性理论, 本文研究作者耦合关系中的一些模糊因素, 并且通过期望值方法提出一种新的计算作者耦合强度的方法。最后, 通过2011年至2017年中国知网中部分作者参考文献的相关数据对本文提出的算法的有效性进行验证。

关键词

作者耦合, 可信性理论, 期望值方法, 模糊耦合强度

Analysis Method of Author Coupling Relationship Strength Based on Credibility Theory

Guoqiang Yuan, Songfei Zhang

College of Statistics and Data Science, Hebei Finance University, Baoding Hebei

Received: November 11, 2025; accepted: December 5, 2025; published: December 10, 2025

Abstract

Based on credibility theory, this paper studies some fuzzy factors in author coupling relationship, and proposes a new method of calculating author coupling strength by expect value method. Finally, the effectiveness of the proposed algorithm in this paper is verified by the relevant data of some authors' references in CNKI (China National Knowledge Infrastructure) from 2011 to 2017.

Keywords

Author Coupling, Credibility Theory, Expect Value Method, Fuzzy Coupling Strength

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 背景介绍

作者耦合关系是引文分析理论中的一个重要概念, 它是利用第三方文献建立的一种隐藏学术关系。目前, 一些学者的研究已经表明两位(或多位)作者共同引用的文献频次和数量越多, 说明他们的研究方向更加接近, 也就是他们之间的耦合强度越大。因此, 正确有效地挖掘作者之间的耦合关系可以较好地发现某学科领域的知识结构, 进而为学科的发展奠定良好的学术研究基础。

作者耦合是基于文献耦合[1]提出的, 但其又不同于文献耦合。作者耦合相比文献耦合的静态结构是动态结构, 并且作者之间的耦合强度是随时间而变化的。2008年, Zhao 和 Strotmann [2]首次针对作者耦合进行了实证研究, 并发现作者耦合的实证结论对作者同被引的研究进行了必要的完善。自此众多学者对作者耦合理论的相关研究工作进行了深入细致的研究。2010年, 刘志辉和张志强[3]基于作者耦合分析方法提出了作者关键词耦合概念。2011年, 国内学者陈远和王菲菲[4]首次利用作者耦合分析方法及其相关理论进行了实验应用研究。2012年, Ma [5]从多个角度研究了作者耦合分析的相关理论及其实证效用。2014年, 宋艳辉和武夷山[6]利用作者耦合分析方法研究了近年来情报学的知识结构。邱均平和刘国徽[7]对国内作者耦合分析方法的研究现状进行了分析。2016年, 赵文莲和张艺蔓[8]利用期刊耦合与作者耦合方法分析了学科的内部结构。Yang [9]等人利用作者关键词耦合分析方法对信息科学领域进行了可视化研究。

综上所述, 由于作者耦合分析方法及其相关理论的研究起步较晚, 从而相关的理论还不尽成熟, 并且许多知识还处于不断探索阶段。从现有的文献分析可以看出, 大多数的学者均是利用作者耦合分析方法或结合其它理论对某一领域的学科知识结构进行实证分析研究。然而, 对于作者耦合分析方法自身的理论研究比较少见。因此, 本文将从作者耦合分析方法的理论角度出发, 对作者耦合强度的理论计算方法进行更加深入的研究。目前, 作者耦合关系强度的算法主要以简单算法、最小值算法和组合算法为主[5]。这三种算法在计算单篇参考文献的耦合强度时均将参考文献的数量简单作为衡量的主要标准。然而, 在实际的研究中, 我们有时会遇到两位(或多位)作者引用了同一篇参考文献, 但是引用了文章中两个不同的知识点, 此时作者之间的耦合度就会偏低, 甚至也可能作者之间并没有耦合关系。研究表明, 上述的现象在目前的科学的研究中多以交叉学科比较常见。

本文针对以上的问题将基于可信性理论[10]在不确定环境中分析研究作者的耦合关系。首先, 在不确定环境下提出作者耦合强度的算法思想。其次, 利用期望值方法计算两位(或多位)作者的耦合强度。最后, 对于不确定环境下的作者耦合关系强度算法的可行性和有效性进行分析说明。

2. 预备知识

为了研究不确定环境下的作者耦合关系, 下面首先回顾可信性理论的一些基本内容。

设 $P(\Gamma)$ 是给定论域 Γ 的幂集, Pos 是定义在幂集 $P(\Gamma)$ 上的集函数。若集函数 Pos 满足:

- (1) $Pos(\emptyset) = 0$, $Pos(\Gamma) = 1$ 。
- (2) 对于幂集 $P(\Gamma)$ 上的任意子类 $\{A_i | i \in I\}$ 有 $Pos(\bigcup_{i \in I} A_i) = \sup_{i \in I} Pos(A_i)$ 。

称 Pos 是可能性测度[11] [12], 其中 I 是任意指标集。

基于可能性测度, Liu 和 Liu [9]定义了下面的自对偶集函数 Cr 。设三元组 $(\Gamma, P(\Gamma), Pos)$ 是一个可能空间, 若定义集函数 Cr 为

$$\text{Cr}(A) = \left(1 + \text{Pos}(A) - \text{Pos}(A^c)\right)$$

则称集函数 Cr 是事件 A 的可信性测度, 其中 A^c 是集合 A 的补集。

假设 $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ 是一个从 Γ 到实数空间 \Re^n 上的集函数, 则称 ξ 是定义在可信性空间 $(\Gamma, \mathcal{P}(\Gamma), \text{Cr})$ 上的模糊向量。特别地, 当 $n=1$ 时, ξ 称为模糊变量。

假设 ξ 是一个离散型的模糊变量, 且可能性分布为

$$\mu_\xi(x) = \begin{cases} \mu_1, & x = a_1 \\ \mu_2, & x = a_2 \\ \vdots \\ \mu_n, & x = a_n \end{cases}$$

不失一般性, 设 $a_1 \leq a_2 \leq \dots \leq a_n$, 则模糊变量的期望值为

$$E[\xi] = \sum_{i=1}^n p_i a_i, \quad p_i = \frac{1}{2} \left(\max_{j=1}^i \mu_j - \max_{j=0}^{i-1} \mu_j \right) + \frac{1}{2} \left(\max_{j=i}^n \mu_j - \max_{j=i+1}^{n+1} \mu_j \right),$$

并且 $\mu_0 = 0$, $\mu_{n+1} = 0$, $p_i \geq 0$, $\sum_{i=1}^n p_i = \max_{i=1}^n \mu_i = 1$ 。

关于可信性理论更多的知识及其应用可以参阅文献[13]-[15]。

3. 不确定环境下的作者耦合关系强度算法

3.1. 作者耦合关系简介

作者耦合是引文分析理论中的重要概念之一。它是基于文献耦合[1]而产生的一个分析作者之间动态结构的重要工具。利用作者耦合关系分析和探求学科的知识结构主要通过作者耦合关系的强度来衡量作者之间的学术联系强度。目前, 关于作者耦合关系强度的算法主要有简单算法、最小值算法和组合算法三种方法[5]。

为了研究的方便, 下面通过一个例子来说明原有作者耦合关系强度的算法原理。另外, 在示例中本文只研究第一作者的参考文献, 不考虑其他的作者。

Table 1. Examples of author coupling relationships

表 1. 作者耦合关系示例

学者 A	学者 B
参考文献 1 (引用 1 次)	参考文献 1 (引用 1 次)
参考文献 2 (引用 1 次)	参考文献 2 (引用 1 次)
参考文献 3 (引用 1 次)	参考文献 3 (引用 1 次)
参考文献 4 (引用 2 次)	参考文献 4 (引用 2 次)
参考文献 5 (引用 2 次)	参考文献 5 (引用 4 次)
参考文献 6 (引用 5 次)	参考文献 6 (引用 3 次)
参考文献 7 (引用 4 次)	参考文献 7 (引用 4 次)
参考文献 8 (引用 3 次)	参考文献 8 (引用 6 次)
参考文献 9	参考文献 10
参考文献 11	*
参考文献 12	*

根据表1中给出的数据,下面分别对三种作者耦合关系强度算法的原理进行说明。

(1) 简单算法

算法原理是将两位(或多位)作者共同的不重复参考文献的数目作为作者耦合强度。表1中作者A不重复参考文献数目是11,作者B不重复参考文献数目是9,作者A和作者B共同的参考文献数目是8,那么两位作者的耦合强度就是8。

(2) 最小值算法

算法原理是将两位(或多位)作者共同的不重复引用次数为1次的参考文献数目与各重复引用参考文献重复引用次数的最小值之和作为作者耦合强度。表1中两位作者共同引用1次的参考文献数目是3,参考文献4引用次数最小值是2,参考文献5引用次数最小值是2,参考文献6引用次数最小值是3,参考文献7引用次数最小值是4,参考文献8引用次数最小值是3,那么两位作者的耦合强度就是

$$3 + \min\{2, 2\} + \min\{2, 4\} + \min\{5, 3\} + \min\{4, 4\} + \min\{3, 6\} = 17$$

(3) 组合算法

算法原理是将两位(或多位)作者共同的不重复引用次数为1次的参考文献数目与各重复引用参考文献重复引用次数的乘积之和作为作者耦合强度。表1中两位作者共同引用1次的参考文献数目是3,参考文献4引用次数乘积是4,参考文献5引用次数乘积是8,参考文献6引用次数乘积是15,参考文献7引用次数乘积是16,参考文献8引用次数乘积是18,那么两位作者的耦合强度就是

$$3 + 2 \times 2 + 2 \times 4 + 5 \times 3 + 4 \times 4 + 3 \times 6 = 64$$

3.2. 不确定环境下作者耦合关系强度的计算方法

在3.1节中提到的三种计算作者耦合强度的方法已经得到一些学者的实证研究,并且结果显示出最小值算法相比较另外两种算法的效果更好。综合现有作者耦合的相关研究成果可以看出,大部分学者均是应用作者耦合关系强度进行了实证研究,而关于理论的研究并不多见。因此,本文将主要针对作者耦合进行一些理论方面的研究,从而对原有理论进行必要的补充和完善。从3.1节中提到的三种计算作者耦合关系强度的算法原理不难发现,两位(或多位)作者共同引用的参考文献数量越多,作者之间的耦合强度越大,并且对于共同引用的重复参考文献重复次数越多,作者之间的耦合强度越大。另外,在实际的引文分析理论研究中,虽然两位(或多位)作者共同引用了相同的参考文献,但是他们引用的论文关键点可能不同,这样的情况可能导致他们的研究兴趣并不一定相同,也就是说他们之间可能并没有很强的耦合关系。同时,作者之间的耦合关系强度简单利用参考文献引用次数表示也有些不尽合理。这样利用“次数”直接表示耦合关系“强”和“弱”的程度在一定意义上会影响研究的最终效果。从现有的文献看,以上提到的这些有关作者耦合强度计算的问题并未得到有效的分析和研究。

为了处理以上关于作者耦合关系强度的计算问题,本文将利用可信性理论的相关知识研究作者耦合关系强度。一般来说,两位(或多位)作者引用同一篇参考文献的关键点具有一定的不确定性,只有正确处理这种不确定性才能有效分析作者之间的耦合关系。事实上,作者之间的引用关系或多或少带有一定的主观性,这就会使得作者的耦合关系产生一定的模糊性。因此,为了更好地探究学科知识结构,本文将利用模糊数学知识度量作者耦合关系的强度。首先,假设两位(或多位)作者的共同参考文献数量是*i*(*i*=1,2,⋯,n);第*j*位作者引用第*i*篇参考文献的次数为*A_{ij}*(*j*=1,2,⋯,m)。根据以上的符号,下面给出第*i*篇共同参考文献的加权引用次数*n_i*的表达式为

$$n_i = \begin{cases} \min_{1 \leq j \leq m} A_{ij}, & \max_{1 \leq j, j' \leq m, j \neq j'} |A_{ij} - A_{ij'}| > \alpha \\ \frac{1}{m} \sum_{j=1}^m A_{ij}, & \max_{1 \leq j, j' \leq m, j \neq j'} |A_{ij} - A_{ij'}| \leq \alpha \end{cases}$$

其中, 参数 α 是决策者事先给定的一个确定引用次数。

其次, 为了建立两位(或多作者)的模糊耦合强度 ξ 的可能性分布, 我们要先对上述第 i 篇共同参考文献的加权引用次数 n_i 进行排序。假设排序后共同参考文献的加权引用次数是 a_t 的有 k_t 个($t=1, 2, \dots, T$), 并且假设具有相同引用次数 a_t 的共同参考文献的模糊耦合强度是相同的。另外, 对于大于等于 a_3 的共同参考文献的模糊耦合强度也相同。因此, 假设具有相同引用次数 a_t 的共同参考文献的模糊耦合强度为 μ_t , 并给出两位(或多作者)的模糊耦合强度 ξ 的可能性分布 μ_ξ 为

$$\mu_\xi(x) = \begin{cases} \mu_1 = \frac{a_1}{a_1 + a_2}, & x = a_1 \\ \mu_2 = \frac{a_2}{a_1 + a_2}, & x = a_2 \\ \mu_3 = 1, & x = a_{t \geq 3} \end{cases}$$

特别地, 当上面可能性分布 μ_ξ 中的 x 仅取 a_1 一个值时, 则定义

$$\mu_\xi(x) = \begin{cases} \mu_1 = 1, & x = a_0 \\ \mu_2 = \frac{a_1}{a_1 + 2}, & x = a_1 \end{cases}$$

这里规定 $a_0 = 0$, 且表示排序后共同参考文献的加权引用次数为 0。当上面可能性分布 μ_ξ 中的 x 取 a_1 和 a_2 两个值时, 则定义

$$\mu_\xi(x) = \begin{cases} \mu_1 = \frac{a_1}{a_1 + a_2}, & x = a_1 \\ \mu_2 = 1, & x = a_2 \end{cases}$$

以上部分详细给出了两位(或多作者)作者模糊耦合强度的计算方法。另一方面, 在实际的研究中往往通过作者之间的耦合关系强度来探究某一领域的学科知识结构。针对这样的研究特点, 要想合理地建立某一领域中任意两位(或多作者)作者之间的学科兴趣关联结构就需要比较他们之间的耦合强度大小。然而, 实际研究中单凭两位(或多作者)作者的模糊耦合强度 ξ 的可能性分布 μ_ξ 难以对某一领域的学科知识结构进行有效分析。因此, 本文将采用两位(或多作者)作者的模糊耦合强度的期望值对某一领域的学科知识结构进行迅速分析。期望值可以迅速分析出作者之间的耦合关系, 并且通过耦合关系也能真正反映出多位作者在学科领域中的重要性。

设 $E[\xi]$ 表示两位(或多作者)作者的模糊耦合强度 ξ 的期望值, 且

$$E[\xi] = (k_1 a_1) p_1 + (k_2 a_2) p_2 + \left(\sum_{t=3}^T k_t a_t \right) p_3,$$

$$p_t = \frac{1}{2} \left(\max_{j=1}^t \mu_j - \max_{j=0}^{t-1} \mu_j \right) + \frac{1}{2} \left(\max_{j=t}^3 \mu_j - \max_{j=t+1}^4 \mu_j \right), \quad t = 1, 2, 3,$$

并且 $\mu_0 = 0$, $\mu_4 = 0$, $p_t \geq 0$, $\sum_{t=1}^3 p_t = \max_{t=1}^3 \mu_t = 1$ 。

下面以表 1 中的数据为例, 计算作者 A 和作者 B 之间模糊耦合强度的期望值。设共同参考文献的加权引用次数 n_i 的表达式中的参数 $\alpha = 2$, 则表 1 中 8 篇共同参考文献的加权引用次数分别为

(1) 因为作者 A 和作者 B 引用参考文献 1-4 及参考文献 7 的次数是相同的, 所以

$$\max_{1 \leq j, j' \leq 2, j \neq j'} |A_{ij} - A_{ij'}| = 0 < 2, \quad i = 1, 2, 3, 4, 7$$

则 $n_i = 1$, $i = 1, 2, 3$, $n_4 = 2$, $n_7 = 4$ 。

(2) 因为作者 A 引用参考文献 5 为 2 次, 作者 B 引用参考文献 5 为 4 次; 作者 A 引用参考文献 6 为 5 次, 作者 B 引用参考文献 6 为 3 次, 所以

$$\max_{1 \leq j, j' \leq 2, j \neq j'} |A_{ij} - A_{ij'}| = 2 = 2, \quad i = 5, 6$$

$$\text{则 } n_5 = \frac{2+4}{2} = 3, \quad n_6 = \frac{5+3}{2} = 4.$$

(3) 因为作者 A 引用参考文献 8 为 3 次, 作者 B 引用参考文献 8 为 6 次, 所以

$$\max_{1 \leq j, j' \leq 2, j \neq j'} |A_{8j} - A_{8j'}| = 3 > 2,$$

则 $n_8 = \min\{6, 3\} = 3$ 。

根据作者模糊耦合强度 ξ 的可能性分布 μ_ξ 的定义, 可以得到

$$\mu_\xi(x) = \begin{cases} \mu_1 = \frac{1}{3}, & x = 1 \\ \mu_2 = \frac{2}{3}, & x = 2 \\ \mu_3 = 1, & x = 3, 4 \end{cases}$$

再根据作者模糊耦合强度 ξ 的期望值定义, 可以得到

$$\begin{aligned} p_1 &= \frac{1}{2}(\mu_1 - \mu_0) + \frac{1}{2}\left(\max_{j=1}^3 \mu_j - \max_{j=2}^4 \mu_j\right) = \frac{1}{2}\left(\frac{1}{3} - 0\right) + \frac{1}{2}(1 - 1) = \frac{1}{6} \\ p_2 &= \frac{1}{2}\left(\max_{j=1}^2 \mu_j - \max_{j=0}^1 \mu_j\right) + \frac{1}{2}\left(\max_{j=2}^3 \mu_j - \max_{j=3}^4 \mu_j\right) = \frac{1}{2}\left(\frac{2}{3} - \frac{1}{3}\right) + \frac{1}{2}(1 - 1) = \frac{1}{6} \\ p_3 &= \frac{1}{2}\left(\max_{j=1}^3 \mu_j - \max_{j=0}^2 \mu_j\right) + \frac{1}{2}(\mu_3 - \mu_4) = \frac{1}{2}\left(1 - \frac{2}{3}\right) + \frac{1}{2}(1 - 0) = \frac{2}{3} \end{aligned}$$

则有

$$\begin{aligned} E[\xi] &= (k_1 a_1) p_1 + (k_2 a_2) p_2 + \left(\sum_{t=3}^T k_t a_t\right) p_3 \\ &= 3 \times 1 \times \frac{1}{6} + 1 \times 2 \times \frac{1}{6} + (2 \times 3 + 2 \times 4) \times \frac{2}{3} = \frac{61}{6} \end{aligned}$$

上面的部分通过表 1 中的数据给出了作者模糊耦合强度 ξ 的期望值。通过与以往作者耦合强度的结果比较, 利用期望值的方法计算作者之间的耦合强度更能真实反映作者在学科领域中的作用。本文提出的期望值方法首先考虑了作者引用参考文献的不确定性, 从一定程度上利用模糊数学的方法处理了这些不确定性因素, 从而也说明了本文提出的计算作者耦合强度方法的理论可行性。其次, 为了更好地反映两位(或多位)作者共同引用文献数量越多, 作者之间的耦合强度越大, 并且对于共同引用文献的重复次数越多, 作者之间的耦合强度越大的现象, 本文通过设计参考文献的加权引用次数公式和作者模糊耦合强度的可能性分布函数合理反映了上面提到的耦合强度变化现象。因此, 本文提出的计算作者耦合强度方法具有一定的理论有效性。

4. 实证分析

4.1. 数据来源与研究对象

本文的数据来源于 2000~2017 年中国知网中收录的部分期刊数据, 并对数据进行了细致的筛选。这

里需要强调的是本文数据选用的时间间隔主要考虑该段时间的发文量、文献关联关系等因素。为了能很好地对本文提出的方法进行说明和验证, 下面对所选作者和发文的数量进行了必要的控制。首先, 利用主题词“引文耦合”和“作者耦合”在中国知网中仅筛选出 10 位相关领域的学者(仅限第一作者)作为本文的研究对象。其次, 利用主题词“耦合”在中国知网中筛选出以下十位作者所发表的期刊论文, 并且滤去所选论文中“导语”、“评论”以及“讲话”等形式的论文。本文在中国知网中所选作者和发文的数量见表 2。

Table 2. Authors and publication status
表 2. 作者和发文情况

作者(第一作者)	发文数量	作者(第一作者)	发文数量
邱均平	12	杨瑞仙	1
马瑞敏	2	李国俊	2
沈利华	1	靖培栋	1
赵文涟	1	肖明	4
许振亮	2	王菲菲	2

通过对以上 10 位作者之间的引用关系分析, 并且分别统计出了任意两位作者之间的参考文献引用关系。为了研究的方便, 本文对以上 10 位作者仅仅分析了任意两位作者之间的耦合关系, 不考虑两位以上作者的耦合关系。

4.2. 作者耦合关系的研究方法

首先, 利用在 3.2 节中设计的作者耦合关系强度的计算方法分别计算出 10 位作者中任意两位作者的模糊耦合强度期望值。其次, 建立 10 位作者之间的耦合关系矩阵, 矩阵中的数据分别代表任意两位作者的模糊耦合强度期望值, 如表 3 所示。矩阵很显然是一个对称矩阵, 对称性代表了作者之间的耦合关系是单向的。另外, 对于没有共同参考文献的两位作者之间的耦合强度以及作者本人与本人的耦合强度均定义为常数 0。最后, 在 UCINET 软件中利用画图工具绘制出 10 位作者之间的耦合关系图, 如图 1 所示。

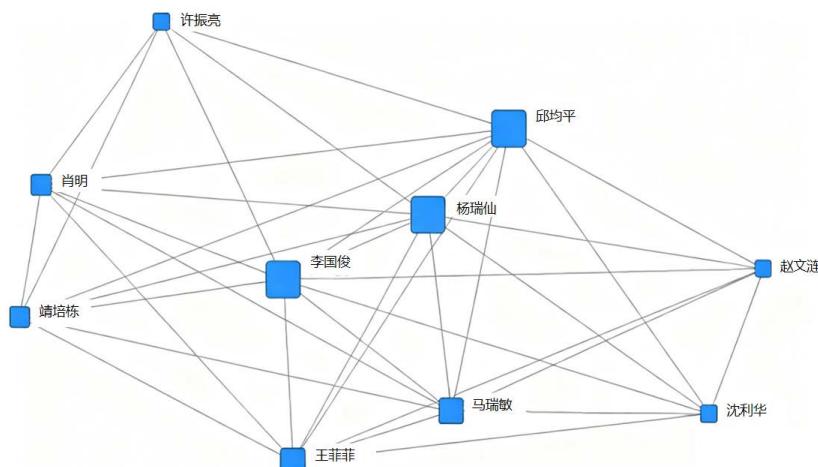


Figure 1. Network diagram of the author's fuzzy coupling strength relationship
图 1. 作者模糊耦合强度关系网络图

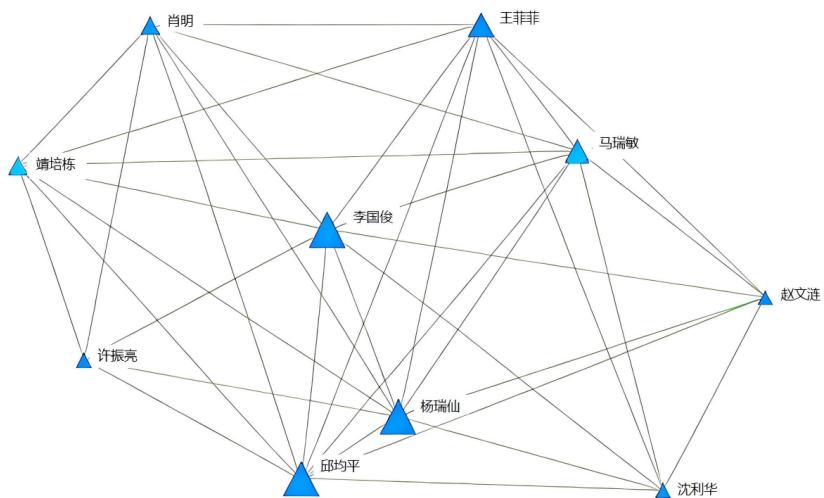


Figure 2. Network diagram of the author coupling strength relationship based on the minimum value algorithm

图 2. 基于最小值算法的作者耦合强度关系网络图

Table 3. Author coupling strength relationship matrix

表 3. 作者耦合强度关系矩阵

	邱均平	马瑞敏	沈利华	赵文涟	许振亮	杨瑞仙	李国俊	靖培栋	肖明	王菲菲
邱均平	0	$\frac{8}{3}$	2.3	3.25	$\frac{1}{3}$	5.4	2.35	1.4	2.1	$\frac{19}{6}$
马瑞敏	$\frac{8}{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{5}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1.8	$\frac{2}{3}$
沈利华	2.3	$\frac{1}{3}$	0	$\frac{2}{3}$	0	0.5	$\frac{1}{3}$	0	0	1.4
赵文涟	3.25	$\frac{1}{3}$	$\frac{2}{3}$	0	0	$\frac{5}{6}$	$\frac{1}{3}$	0	0	1.6
许振亮	$\frac{1}{3}$	0	0	0	0	0.25	$\frac{1}{3}$	0.25	$\frac{5}{12}$	0
杨瑞仙	5.4	$\frac{5}{6}$	0.5	$\frac{5}{6}$	0.25	0	1.6	$\frac{1}{6}$	$\frac{1}{3}$	1.6
李国俊	2.35	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1.6	0	0.25	2.7	1.4
靖培栋	1.4	$\frac{1}{6}$	0	0	0.25	$\frac{1}{6}$	0.25	0	$\frac{53}{28}$	$\frac{1}{6}$
肖明	2.1	1.8	0	0	$\frac{5}{12}$	$\frac{1}{3}$	2.7	$\frac{53}{28}$	0	$\frac{1}{6}$
王菲菲	$\frac{19}{6}$	$\frac{2}{3}$	1.4	1.6	0	1.6	1.4	$\frac{1}{6}$	$\frac{1}{6}$	0

4.3. 结果分析

利用 UCINET 软件对 10 位作者之间的耦合关系进行了可视化, 基于模糊期望值方法和最小值算法分别得到了图 1 和图 2 的结果。

首先, 仅仅从 10 位作者在耦合强度关系网络图中的学术地位而言, 从图 1 中可以明显分辨出学者邱均平、杨瑞仙和李国俊处于第一集团, 王菲菲和马瑞敏处于第二集团, 靖培栋和肖明处于第三集团, 许振亮、沈利华和赵文涟处于第四集团。同时, 从图 2 中的结果也可以得到图 1 中得到的结论。因此, 从

这一点上也就能在一定程度上说明本文提出的作者耦合强度关系的模糊期望值方法是有效和可行的。

其次, 利用本文提出的作者耦合强度关系的模糊期望值方法还可以有效区分相同耦合强度情形下的作者之间的耦合强弱关系。在 4.2 中计算作者之间耦合强度的过程中, 邱均平和肖明的共同参考文献有 3 篇, 引用次数分别是(2 次, 3 次), (1 次, 2 次)和(1 次, 1 次), 并且王菲菲和马瑞敏的共同参考文献有 4 篇, 引用次数分别是(1 次, 1 次), (1 次, 1 次), (1 次, 1 次)和(1 次, 1 次)。通过最小值算法计算可以得到邱均平和肖明以及王菲菲和马瑞敏之间的耦合强度均是 4, 但是通过模糊期望值方法计算可以得到他们之间的耦合强度分别是 2.1 和 2/3。从以上的分析可以看出, 虽然利用最小值方法计算部分作者之间的耦合强度得到了相同的结果, 但从实际的引用效果可以明显看出邱均平和肖明之间具有更高的耦合强度。另一方面, 利用模糊期望值方法有效区分了作者之间的耦合强度, 从而也说明了本文的方法具有一定的实际意义。

5. 结论

本文利用可信性理论处理了作者耦合关系中的一些不确定性因素, 并且通过期望值方法给出了作者耦合关系强度的一种新的计算方法。下面对本文的主要工作进行总结如下:

- (1) 建立了一种新的参考文献加权引用次数公式, 有效处理了两位(或多位)作者引用文献次数差别较大对作者耦合强度计算的影响。
- (2) 建立了新的作者模糊耦合强度可能性分布函数, 并以此合理分析了作者耦合强度中的不确定性因素。
- (3) 通过期望值方法给出了作者耦合强度的计算方法。

本文给出了一种新的作者耦合关系强度的计算方法, 但是在某些方面还存在一定的不合理性。其一, 本文研究的前提是只考虑第一作者引用情况, 因此他们的影响力只由作为第一作者的引用文献来决定, 这样就会造成一定的不合理性。换句话说, 本文未考虑同一篇论文不同作者的引用重要性以及共同引用文献的时效性等领域。其二, 虽然本文给出了作者耦合关系强度的计算方法有效性的数据验证, 但是数据的规模较小, 还不足以完全验证算法的可行性, 这方面还需要进一步实证研究。其三, 本文采用的评价方法仅利用期望值算子, 未尝试利用方差、置信区间等方法对作者耦合关系强度进行评价。

基金项目

河北省金融科技应用重点实验室项目“河北省城市商业银行潜在客户挖掘与系统开发研究”(项目编号: 2024004)。

参考文献

- [1] Kessler, M.M. (1963) Bibliographic Coupling between Scientific Papers. *American Documentation*, **14**, 10-25. <https://doi.org/10.1002/asi.5090140103>
- [2] Zhao, D. and Strotmann, A. (2008) Evolution of Research Activities and Intellectual Influences in Information Science 1996-2005: Introducing Author Bibliographic-Coupling Analysis. *Journal of the American Society for Information Science and Technology*, **59**, 2070-2086. <https://doi.org/10.1002/asi.20910>
- [3] 刘志辉, 张志强. 作者关键词耦合分析方法及实证研究[J]. 情报学报, 2010, 29(2): 268-275.
- [4] 陈远, 王菲菲. 基于 CSSCI 的国内情报学领域作者文献耦合分析[J]. 情报资料工作, 2011, 32(5): 6-12.
- [5] Ma, R. (2012) Author Bibliographic Coupling Analysis: A Test Based on a Chinese Academic Database. *Journal of Informetrics*, **6**, 532-542. <https://doi.org/10.1016/j.joi.2012.04.006>
- [6] 宋艳辉, 武夷山. 基于作者文献耦合分析的情报学知识结构研究[J]. 图书情报工作, 2014, 58(1): 117-123.
- [7] 邱均平, 刘国徽. 国内耦合分析方法研究现状与展望[J]. 图书情报工作, 2014, 58(7): 131-136.

- [8] 赵文涟, 张艺蔓. 基于期刊耦合和作者耦合的学科内部结构研究[J]. 情报科学, 2016, 34(12): 108-114.
- [9] Yang, S., Han, R., Wolfram, D. and Zhao, Y. (2016) Visualizing the Intellectual Structure of Information Science (2006-2015): Introducing Author Keyword Coupling Analysis. *Journal of Informetrics*, **10**, 132-150.
<https://doi.org/10.1016/j.joi.2015.12.003>
- [10] Liu, B. and Liu, Y.K. (2002) Expected Value of Fuzzy Variable and Fuzzy Expected Value Models. *IEEE Transactions on Fuzzy Systems*, **10**, 445-450. <https://doi.org/10.1109/tfuzz.2002.800692>
- [11] Wang, P. (1982) Fuzzy Contactibility and Fuzzy Variables. *Fuzzy Sets and Systems*, **8**, 81-92.
[https://doi.org/10.1016/0165-0114\(82\)90032-x](https://doi.org/10.1016/0165-0114(82)90032-x)
- [12] Klir, G.J. (1999) On Fuzzy-Set Interpretation of Possibility Theory. *Fuzzy Sets and Systems*, **108**, 263-273.
[https://doi.org/10.1016/s0165-0114\(97\)00371-0](https://doi.org/10.1016/s0165-0114(97)00371-0)
- [13] Zhai, H., Liu, Y. and Yang, K. (2016) Modeling Two-Stage UHL Problem with Uncertain Demands. *Applied Mathematical Modelling*, **40**, 3029-3048. <https://doi.org/10.1016/j.apm.2015.09.086>
- [14] Li, J.F., Liu, Y.-K. and Yang, G.Q. (2023) Two-Stage Distributionally Robust Optimization Model for a Pharmaceutical Cold Supply Chain Network Design Problem. *International Transactions in Operational Research*, **31**, 3459-3493.
<https://doi.org/10.1111/itor.13267>
- [15] Wang, J.P., Bai, X.J. and Liu, Y.-K. (2025) Building Sustainable Hazardous Products Supply Chain against Ambiguous Risk with Accelerated Benders Decomposition Algorithm. *Transportation Research Part E: Logistics and Transportation Review*, **194**, Article ID: 103941. <https://doi.org/10.1016/j.tre.2024.103941>