Hans Journal of Data Mining 数据挖掘, 2019, 9(2), 27-33
Published Online April 2019 in Hans. http://www.hanspub.org/journal/hjdm
https://doi.org/10.12677/hjdm.2019.92004

Hans汉斯

Analysis of the Co-Authorship of Papers Based on Social Networks

—Taking the Relevant Papers in the Field of Partial Differential Equations as an Example

De Sun

Shandong University of Science and Technology, Qingdao Shandong
Email: qingdaosd@126.com

Received: Mar. 8th, 2019; accepted: Mar. 22nd, 2019; published: Mar. 29th, 2019

Abstract

This paper combines the relevant data of the papers in the field of partial differential equations in the last five years of CNKI, establishes a network of papers, analyzes the network characteristics of the papers, and puts forward some problems in the co-authoring of the papers and the solutions. Firstly, this paper reflects the characteristics and laws of the co-existing phenomenon of the papers in the field of partial differential equations from the perspective of the co-authorship rate and the number of authors. Then, UCINET software is used to construct the paper co-authoring network and analyze the small world characteristics of the network and no scale characteristics. Finally, the problems and development countermeasures of the co-authorization of papers related to the field of partial differential equations in the past five years are proposed.

Keywords

Paper Co-Authoring Network, Field of Partial Differential Equations, Network Characteristics, Paper Co-Authorship

基于社交网络的论文合著情况分析

——以偏微分方程领域的相关论文为例

孙　得

山东科技大学，山东　青岛
Email: qingdaosd@126.com

文章引用: 孙得. 基于社交网络的论文合著情况分析[J]. 数据挖掘, 2019, 9(2): 27-33.
DOI: 10.12677/hjdm.2019.92004

收稿日期：2019年3月8日；录用日期：2019年3月22日；发布日期：2019年3月29日

摘　要

本文结合CNKI中近五年来偏微分方程领域论文合著的有关数据，建立论文合著网络，分析了论文合著网络的网络特征，提出该领域论文合著情况的一些问题，并给出了解决措施。首先，本文从合著率和论文著者数两个角度反映了偏微分方程领域相关论文合著现象的特征和规律；然后，介绍了社交网络的特征属性，利用UCINET软件构建论文合著网络并分析了该网络的小世界特性和无标度特性；最后，提出了近五年来偏微分方程领域相关论文的合著情况存在的问题以及一些建议。

关键词

论文合著网络，偏微分方程领域，网络特征，论文合著情况

Copyright © 2019 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

随着新学科和大量分支学科的产生，科研合作在现代学术发展中的作用显得尤为重要。经研究发现，20 世纪初论文的合著率仅有 10%，而到了 21 世纪末，论文的合著率超过了 30%，并且一直在稳步上升中。论文合著不仅能提高科学研究的质量，还能产生更多知识的碰撞，加快知识的扩散。总的来说，论文合作对于促进相关学科建设，高质量的科研成果产出以及学科整体科研水平的提高具有重要意义。

据了解，我国多采用文献计量法对论文合著情况进行研究。文献计量法从各种不同的角度对论文合著情况进行统计、分析，得到论文合著的相关规律。近年来，由于社交网络相关研究逐渐在各领域流行，社交网络成为论文合著现象研究的热门主题。某领域的作者之间相互合作，形成了该领域的论文合著网络，作者之间通过该社交网络进行知识的交流，从而促进相关学科建设，提高科研水平。而对于社交网络，非常值得研究的是它的小世界特性和无标度特性。

本文在当前研究的背景下，以 2014~2018 年这五年以来 CNKI 中偏微分方程领域的相关论文为研究对象，应用文献计量法和网络分析法，对偏微分方程领域的论文合著情况进行了分析和研究。文献计量方面，本文从合著率、论文著者数两个角度对该领域论文合著的相关规律进行了探讨；社交网络方面，本文构建了偏微分方程领域的论文合著网络，分析了该网络的小世界和无标度特性，并根据当前论文合著情况存在的问题提出了相关建议。

2. 偏微分方程领域论文的合著情况研究

数据来源与处理

本文研究的是近五年来偏微分方程领域论文的合著情况，因此在 CNKI 数据库搜索“偏微分方程”，选择 2014~2018 年近五年的数据，共得到 1243 篇论文。然后，使用 BICOMB 软件，抽取作者信息，得到 1243 篇论文及对应的作者，部分整理数据如表 1 所示。

Table 1. Partial raw data
表 1. 部分原始数据整理

文章	作者
类 Kuramoto-Sivashinsky 方程的输出反馈边界控制	廖钰靓；赵圣涛；谢成康；
基于 FLAC~(3D)的边坡生态防护根系作用数值模拟初探	彭栓栓；
常气压下非饱和土的一维瞬态渗流	李顺群；桂超；夏锦红；
《偏微分方程数值解》课程的探究式教学方法初探	李郴良；
Ca_(v1.2)离子通道马尔科夫随机过程模型的建立与应用	宋璟波；张镇西；
一类 Riesz 空间分数阶时滞扩散微分方程的隐-显差分格式	杨水平；刘红良；
三次 B-样条配点法定价欧式看跌期权	吴蓓蓓；
基于图像分割的效应靶自动报靶系统	杜红棉；葛竹；孟晓山；

借助 EXCEL 对 1243 篇偏微分方程领域的论文进行汇总、统计，发现 1243 篇偏微分方程论文中，两人及两人以上的合著论文有 469 篇，占总论文数的 37.73%。

3. 论文合著情况分析

目前，论文的合著情况主要通过作者合著度和合著率来表示。论文合著率即为某一领域内参与合著的论文数与论文总数的比值；而论文合著度则比较复杂，它与当前的论文合著规模、科研发展水平有关，由于论文合著度的相关研究比较少，目前还没有统一的定论，所以我们这里只对论文合著率和论文著者数进行研究。

1) 论文合著率研究

CNKI 中偏微分方程领域论文各年的合著率如表 2 所示。

Table 2. Cohesion rate of papers in the field of partial differential equations
表 2. 偏微分方程领域论文的合著率情况

	2014		2015		2016		2017		2018	
	论文数	合著数	论文数	合著数	论文数	合著数	论文数	合著数	论文数	合著数
总计	325	116	285	102	273	106	244	86	116	59

根据表 2 我们可以得到，该领域的论文中有 1/3 是合著论文。在这五年期间，CNKI 中偏微分方程领域论文的合著率总体上呈上升趋势，但上升速率较为缓慢，2017 年有略微波动。合著率最高为 2018 年的 0.5086，合著率最低为 2014 年的 0.3569，两者差值为 0.1517，说明在这五年期间偏微分方程领域论文的合著情况有较为明显的发展。

2) 论文著者数研究

统计出近 5 年偏微分方程领域不同著者数的论文所占的比例，如表 3 所示。

Table 3. The number of different authors in the paper
表 3. 论文不同著者数比率

著者数	1 人	2 人	3 人	4 人	5 人	5 人以上
比率	0.6235	0.1617	0.1158	0.0676	0.0169	0.0145

从表 3 可以看出，2014~2018 年间，偏微分方程领域共发表论文 1243 篇，其中单篇论文合作著者署名最高达到 8 人。通过对各层次著者数比率分析可见，随着著者数从 1 人增加到 5 人以上，论文数所占比率在逐渐下降，著者数为 5 人以上的论文所占比率最小为 0.0145；独著论文所占比率最高为 0.6235。在合著论文中，两人合著的论文所占比例最高为 0.1617。这说明偏微分方程领域相关论文的大部分合著仅限于两人到三人合著，尚未形成更大的合著网络。

4. 偏微分方程领域论文的合著网络及其分析

4.1. 论文合著网络的建立

下面需要通过构建论文的作者合著网络来反映论文合著现象。使用 BICOMB 软件，对附件二中 1243 篇论文的作者信息进行处理，共涉及 1869 个作者。

利用公式

$$a_{ij}=\begin{cases}1, \text{第}i\text{个作者和第}j\text{个作者合作过论文}\\ 0, \text{第}i\text{个作者和第}j\text{个作者没有合作过论文}\end{cases}$$

得到论文合著网络的邻接矩阵 $A=\left[a_{ij}\right]_{1869*1869}$。将上述矩阵导入 UCINET 软件，生成 1869 个作者的合著网络，如图 1 所示，该网络共有 1659 条边，是一个无向网络。

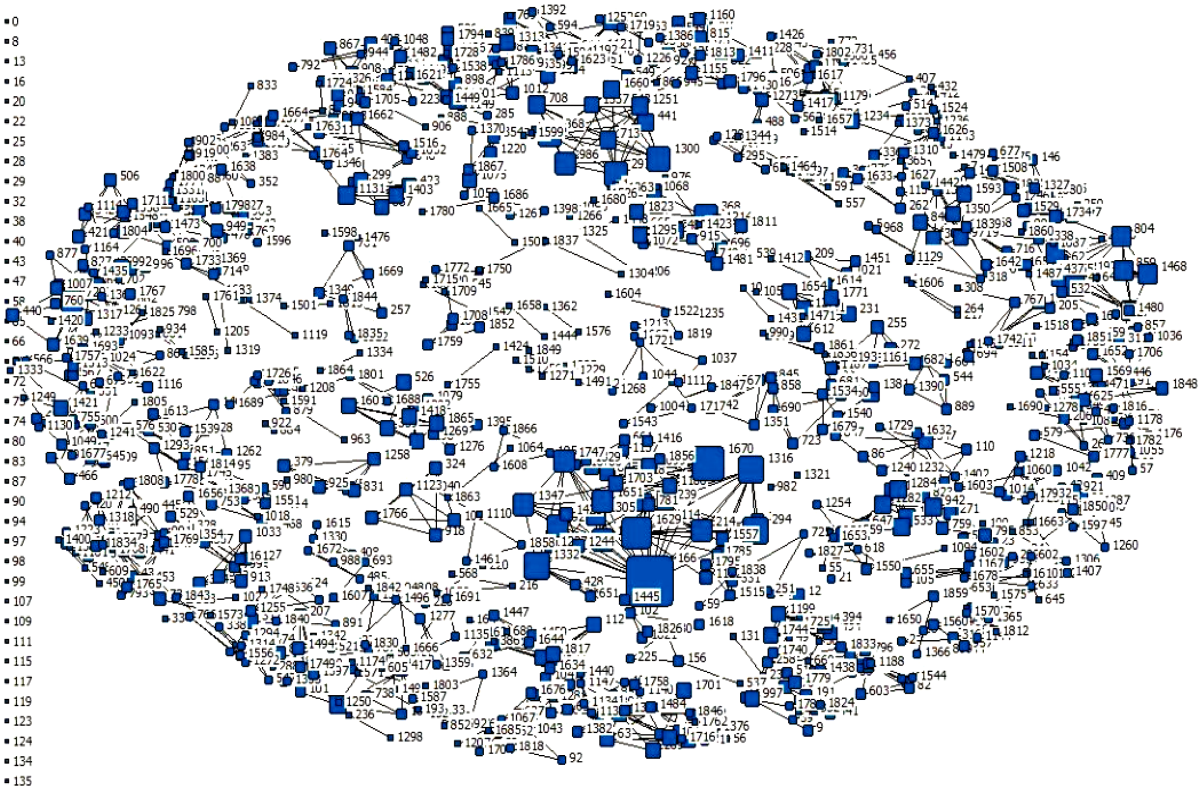

Figure 1. Co-authored network of papers in the field of partial differential equations
图 1. 偏微分方程领域论文的合著网络

由图 1 可见，该网络属于稀疏网络，密度很低，存在相当数量的作者没有与其他作者合作过，删除他们对网络的结构和网络的一些重要特征是没有影响的。

4.2. 网络特征属性的描述

1) 最短路径

无向网络中两节点之间路径的边数表示为两节点之间的距离 d_{ij}，当距离最小时，$\min\left(d_{ij}\right)$ 称为最短路径。对于一个有 N 个节点的网络，其平均最短路径长度为：

$$\mathrm{L}=\frac{1}{\frac{1}{2}N\left(N-1\right)}\sum_{i>j}d_{ij},i,j=1,\cdots,N. \tag{1}$$

2) 聚类系数

聚类系数可以用于表示网络的聚集化程度。在无向网络中，节点 v_{ij} 的聚类系数可用 C_i 来表示：

$$C_i=\frac{n}{C_k^2}=\frac{2n}{k\left(k-1\right)},i=1,2,\cdots,N. \tag{2}$$

其中，k 表示节点 v_i 的所有相邻节点个数；n 表示节点 v_i 的所有相邻节点之间相互连接的边的个数。而整个网络的聚类系数 $\mathrm{C}=\frac{1}{N}\sum_{i=1}^{N}C_i$.

平均最短路径和聚类系数是社交网络中两个重要的特性。

3) 度与度分布

度(degree)表示与一节点相连接的边数。在论文合著网络中，度代表与某作者合作的作者数目。

度分布(degree distribution)描述了网络中节点度的分布情况。通常采用度分布函 P(k)来体现任意网络节点的选择概率，它的值等于度为 k 的节点数占网络中节点总数的比例。

4.3. 论文合著网络的网络特征分析

1) 小世界性

小世界网络，其特征路径长度随着 log(n)增长中说明，在从规则网络向随机网络转换的过程中，实际上特征路径长度和聚合系数都会下降，到变成随机网络的时候，减少到最少。也就是说，如果一个网络点之间特征路径长度小，接近随机网络，而聚合系数依旧相当高，接近规则网络，则该网络具有小世界特性[1]。

对 UCINET 数据矩阵进行二值化处理，计算该论文合著网络的聚类系数和平均最短路径。计算得该微分方程领域的论文合著网络其平均最短路径为 2.120，可以看出该网络的平均最短路径较小，论文作者平均不超过 3 个作者就可以与其他的作者产生合作。此外，该合著网络的聚类系数为 0.327，表明论文合著网络中与一作者合作的两个作者之间有合作的概率为 0.327。

利用 UCINET 构造一个与该论文合著网络相同规模的随机网络，设置其密度与该合著网络密度相同为 0.001，通过计算得到随机网络的聚类系数 0.015，平均最短路径为 2.351。可以得到，该随机网络的聚类系数远小于论文合著网络的聚类系数，而其平均最短路径与论文合著网络差别不大。因此，可以认为该论文合著网络具有小世界特性。

2) 无标度性

无标度网络具有严重的异质性，其各节点之间的连接状况(度数)具有严重的不均匀分布性。在一个网络中，若极少数节点具有大量连接，而大多数节点具有极少的连接，则称该网络具有无标度特性[2]。

计算得到了网络中各节点的度，统计该论文合著网络的度分布情况，如表 4 所示。

Table 4. Degree distribution of network nodes
表 4. 网络节点的度分布

节点度	节点数量	度分布	节点度	节点数量	度分布
0	648	0.346709	7	9	0.004815
1	322	0.172285	8	8	0.00428
2	370	0.197967	9	4	0.00214
3	302	0.161584	10	3	0.001605
4	115	0.061530	12	1	0.000535
5	66	0.035313	13	1	0.000535
6	19	0.010166	20	1	0.000535

统计结果发现，在 1869 个节点的网络中度很小的节点有很多，有 1340 个节点的度小于 3；度较大的节点只有很少，有 6 个节点的度大于 9。这说明了偏微分方程相关论文的大部分合著仅限于少数人之间合著，尚未形成更大的合著网络。分别用幂率函数和指数函数对度分布进行逼近，如图 2 所示。

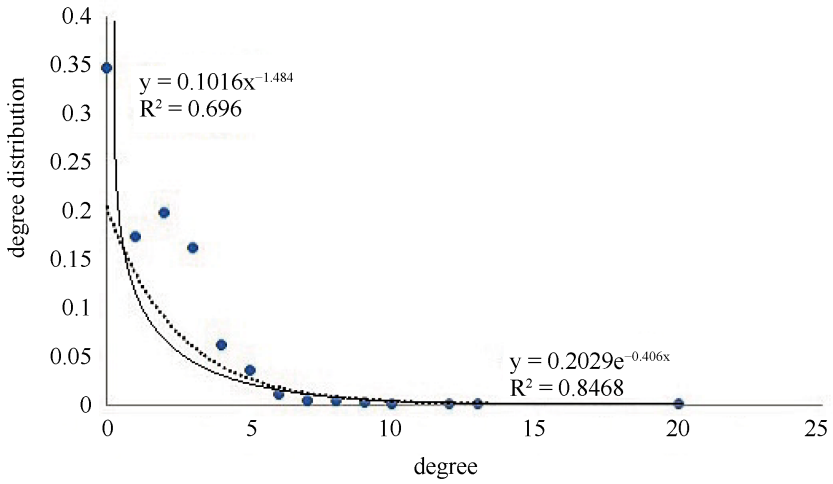

Figure 2. Degree distribution map of paper co-authorship network
图 2. 论文合著网络的度分布图

如果用幂率函数进行逼近，度分布函数为 $y = 0.1016x^{-1.484}, R^2 = 0.696$，逼近效果并不明显；用指数函数进行逼近，可以得到理想的度分布函数 $y = 0.2029e^{-0.406x}, R^2 = 0.8468$。可见该论文合著网络的度分布呈指数分布。根据以上度分布情况来看，该网络中极少数节点的度较大，而大多数节点的度很小，可以得出该论文合著网络具有无标度性[3]。

5. 结论及建议

5.1. 结论

本文利用文献计量法和网络分析法对常微分方程领域的论文合著现象进行分析，并得到了相关结论：

1) 2014~2018 年 5 年期间偏微分方程领域的论文合著率总体上呈上升趋势，但总体而言尚处于较低水平，且上升速率比较缓慢，这说明偏微分方程领域的论文合著情况尚不完善，还存在很大的发展空间。

2) 在偏微分方程领域的论文中，独著论文所占比例最高，且大部分合著论文仅限于两人到三人合著，

这说明偏微分方程领域相关论文的合著度较低，尚未形成更大的合著网络。

3) 从构建的偏微分方程领域论文合著网络来看，网络中有很多节点的度为 0，也就是说，存在相当数量的作者没有与其他作者合作过；而有一些节点的度大于 9，有很多合作过的作者，这说明不同作者之间的论文合著情况有较大差异。

4) 该论文合著网络具有无标度特性，就是说网络中大部分作者具有较少的合作，而只有一小部分作者有很多合作，且作者合著主要以 2~3 人合著为主，合著规模比较小。

5.2. 建议

随着偏微分方程学科领域的发展，对科研能力的要求越来越高，科研论文合作在学术发展中显得尤为重要。因此，科研人员要提高科研学术交流与合作，参考借鉴该领域其他作者的先进理论、创新思维；通过彼此启发、相互借鉴，拓展偏微分方程领域研究的深度和广度[4]。此外，应该扩大论文合作的规模，不仅限于两到三人合著，合著网络的规模的扩大对于促进资源共享、提高科研水平有着很大的积极意义。

参考文献

[1] 孟繁华. 基于复杂网络的供应链建模及其鲁棒性分析[D]: [硕士学位论文]. 北京: 华北电力大学, 2018.

[2] 世攀科, 徐锡, 程文聪. 大规模网络拓扑属性特征及模型分析[C]//中国计算机学会. 中国计算机大会论文集: 2008 年卷. 西安: 中国计算机学会, 2008: 464-472.

[3] 刘贵香, 蔡永明, 刘璐, 英玉超. 社会网络信息传播机制研究——以新浪微博为例[J]. 经贸实践, 2018(6X): 69-70.

[4] 施梅, 胡婷婷. 我国科技期刊论文合著现象分析——以图书情报学核心期刊为例[J]. 出版发行研究, 2018.

Hans 汉斯

知网检索的两种方式：

1. 打开知网页面 http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD
   下拉列表框选择：[ISSN]，输入期刊 ISSN：2163-145X，即可查询
2. 打开知网首页 http://cnki.net/
   左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：http://www.hanspub.org/Submission.aspx
期刊邮箱：hjdm@hanspub.org