

我国工会维权领域研究的文献分析

——基于文献计量分析视角

杨宗典 1403408006

摘要：以 CNKI 为数据来源,以"工会维权"为主题检索,从年代、期刊、作者、学科等方面进行了文献计量分析。研究发现，工会维权领域研究较为缺乏核心作者，多数重要成果为硕博毕业论文，学科交叉情况十分明显。

关键词：工会维权 文献计量 多学科交叉

一、前言

在劳动关系失衡的大背景下，中国工会作为法定的职工权益代表者和维护者，对促进我国当代社会和谐发展，维护社会稳定，协调企业劳资关系中有着至关重要的作用。工会维权作为工会的重要职能，作为职工的重要维权手段，具有十分重要的研究意义。本文将通过文件计量的方法对工会维权领域已有的研究成果进行分析，对研究现状进行梳理。

二、数据来源与研究方法

本文使用基于 Python 语言的自动化脚本，使用检索式为“SU = ‘工会维权’”，采集 CNKI 数据库中的全量数据，即将 CNKI 数据库中主题为工会维权的所有文献采集为样本数据，采集属性包括文献的题名、第一作者、第二作者、关键词、分类号、文献出处、参考文献和引用文献。在此基础上，本文将所有没有参考文献的样本视为无效样本，以剔除报道性文献（如报纸等）和低质量文献，保证样本基本为具有一定质量的研究性文献。经清洗，样本总量为 819 篇。

出于效率考虑，本文所涉及的参考文献和引用文献的分类号主要使用题名匹配的方式，采自于万方数据库。对于第一次采集缺失的数据，使用题名匹配的方式，在百度学术平台上进行了二次采集。两次采集后，缺失的数据约占总量的 5%，基本不影响分析结果。

三、我国工会维权研究的基本情况

1、年代分布

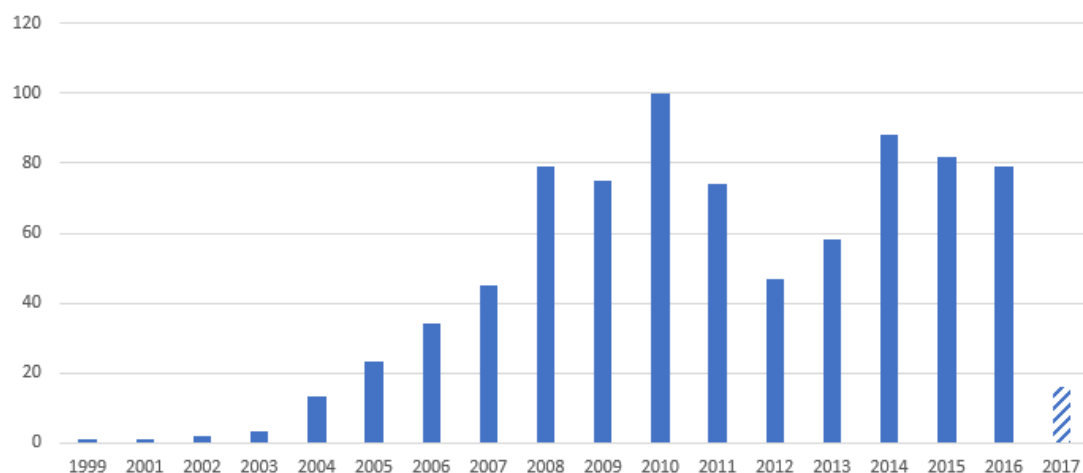


图 1 样本文献的年代分布

图 1 中的柱状图为该年的文献样本量，由于 2017 年度的数据不完整，以斜纹填充，仅供参考。折线图为趋势线，由 Excel 中 4 周期移动平均方法得出。移动平均法是一种简单平滑预测技术，它的基本思想是：根据时间序列资料、逐项推移，依次计算包含一定项数的序时平均值，以反映长期趋势。4 周期移动平均数公式为 $M = \sum_{t-4}^t \frac{x_t}{4}$ ，其中 t 为观测期数， x_t 为第 t 期的观测数^[1]。

根据图 1，我们可以发现，1999-2010 年期间处于工会维权研究的快速上升阶段，2011-2013 年为工会维权研究的研究的低谷期，随后每年的研究产出量又保持在一个较为稳定的水平上。

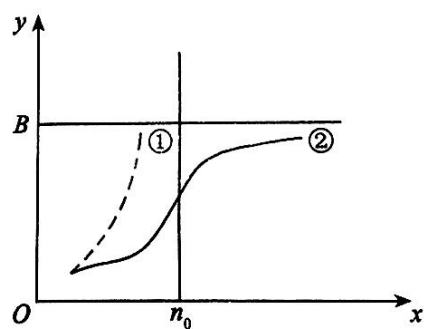


图 2

1 李云刚.基于移动平均法的改进[J].决策与统计,2009(19)

如果将 2011-2013 年的研究产出“低谷”填平，那么工会维权领域的文献增长情况将极为符合出现“饱和现象”的文献指数增长率，如图 2 曲线②^[2]。

为试探究 2011-2013 年出现文献产出“低谷”的原因，笔者通过查阅国家统计局的统计数据（<http://data.stats.gov.cn> 只提供了 2006-2015 年的规模以上工业企业数和亏损数），发现规模以上工业企业的亏损率与文献样本量有一定相关性，如果将亏损率滞后两年（如图 3），两者的相关系数较高，约达 0.648。于是，笔者提出一个猜测：工业企业的亏损会影响到工人权利的保障，导致工人通过工会等渠道进行维权，进而使学者更加关注这一领域的研究，而这两次因果传递的时间大致为 2 年。

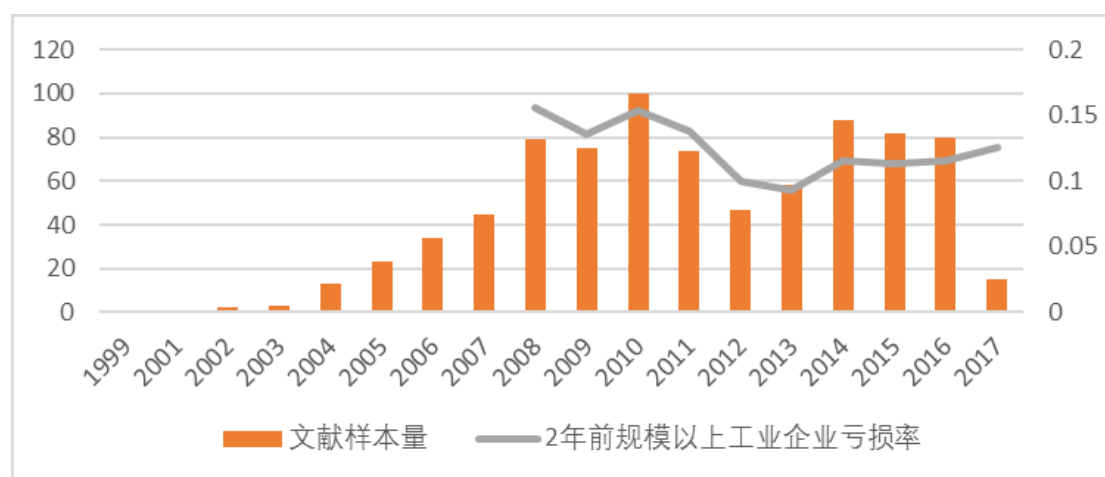


图 3 文献样本量与 2 年前规模以上工业企业亏损率

2、期刊分布

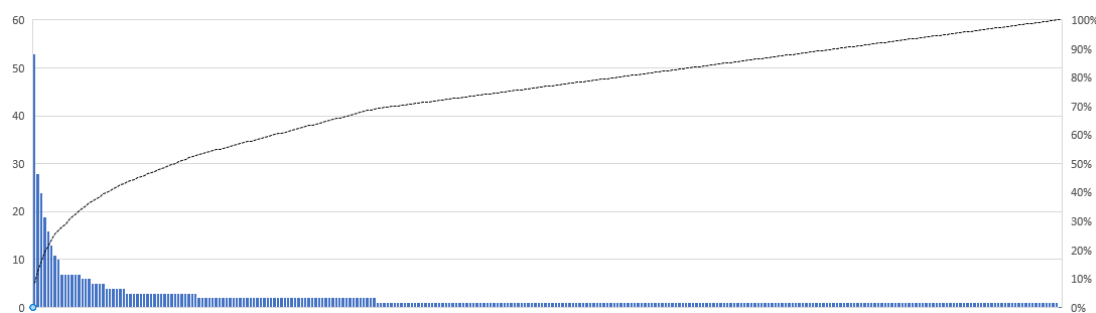


图 4 期刊分布图

通过统计样本（除硕博论文样本）的来源期刊，得到图 4，可以发现工会维权领域各期刊载文量呈现很强的长尾现象，符合文献信息集中和分散的一般规律。为此，笔者按照布拉德福定律的描述对样本文献统计，以进一步讨论工会维权领域的文献分布情况。统计结果见表 1。

表 1

| 区 | 期刊数量 | 样本文献数量 |
|---|------|--------|
| 1 | 14 | 214 |
| 2 | 78 | 214 |
| 3 | 206 | 215 |

布拉德福定律可以基本描述为：如果按一个领域中期刊的载文数量降序排列，分成三组，每组期刊的总载文数约占所有文章的三分之一，那么每组中的期刊数将形成 $1:n:n^2$ 的比例，而 n 一般约为 5^[3]。本文中三区比例约为 1:5.57:14.71，基本符合布拉德福定律的描述。其中核心区（即 1 区）与非相关区（即 3 区）期刊数之比较低、非相关区期刊数相对较少的情况，应该与剔除了大量报道性文献（如报纸等）和低质量文献有关。核心区期刊详细情况见下表 2。

表 2 核心区期刊一览

| 序号 | 期刊名 | 载文数 | 序号 | 期刊名 | 载文数 |
|----|---------------------|-----|----|----------------------|-----|
| 1 | 中国劳动关系学院学报 | 53 | 8 | 才智 | 10 |
| 2 | 东方企业文化 | 28 | 9 | 工会理论研究(上海工会管理职业学院学报) | 7 |
| 3 | 工会论坛(山东省工会管理干部学院学报) | 24 | 10 | 法制与社会 | 7 |
| 4 | 山东工会论坛 | 19 | 11 | 企业改革与管理 | 7 |
| 5 | 天津市工会管理干部学院学报 | 16 | 12 | 工会理论与实践 | 7 |
| 6 | 经营管理者 | 13 | 13 | 新西部(理论版) | 7 |
| 7 | 北京市工会干部学院学报 | 11 | 14 | 学理论 | 7 |

3 Bradford, Samuel C. Sources of information on specific subjects, Journal of Information Science, 10:4, 1985 (October), pp. 173–180

3、基金分布

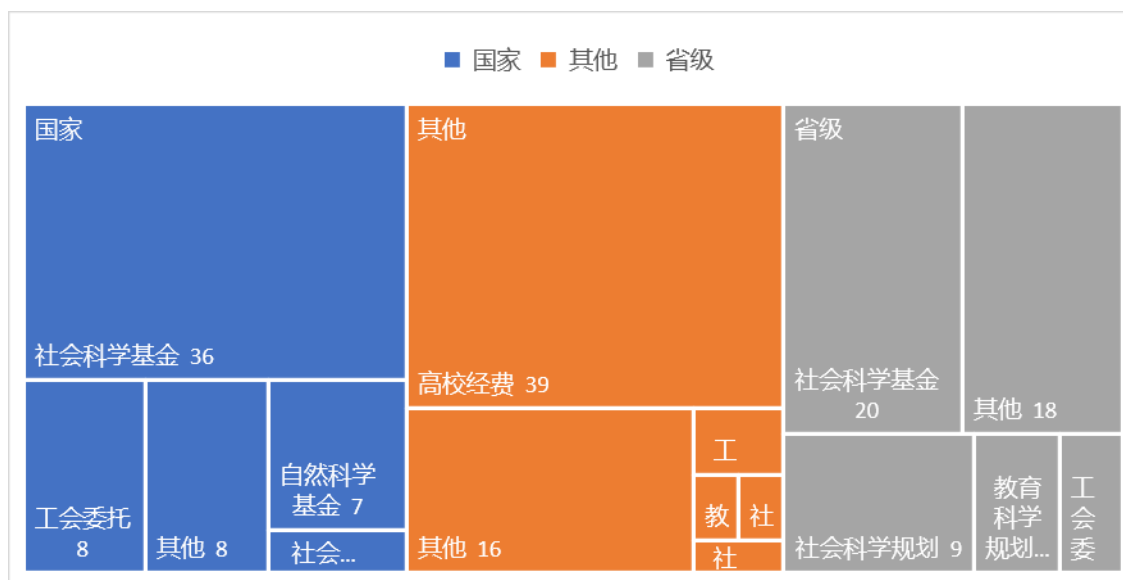


图 5 基金等级与类别层次结构图

总样本 819 篇文献中，共涉及各类基金 175 次。在数据清洗后，从等级上分为国家级、省级和其他 3 类，从类别上分为社会科学基金、高校经费、工会委托、社会科学规划、教育科学规划、自然科学基金和其他 6 类。

其中，国家级 61 次，省级 54 次，其他 60 次，等级分布十分均匀；各级社会科学基金 57 次，高校经费 39 次，各级工会委托 13 次，社会科学规划 12 次，自然科学基金 7 次，教育科学规划 5 次，其他 42 次。

从图 5 可以发现，涉及到的基金主要由国家级、省级社科基金以及高校科研经费组成，值得关注的是，全国总工会以及其他各级工会通过委托研究的方式，为工会维权领域的研究提供了一定的资金支持。

4、作者分布与关系

表 3 处理后的作者分布表

| | | | | |
|---------|---|---|----|-----|
| 处理后的发文数 | 5 | 3 | 2 | 1 |
| 作者数 | 2 | 8 | 53 | 991 |

由于第一作者与第二作者对研究的参与程度和对研究成果的贡献程度有所差异，所以按第一作者 1.0 权重，第二作者 0.7 权重对样本文献的作者数据进行

加权处理，为方便进一步统计，对每位作者的加权发文数四舍五入，形成表 3。

表 4 实际情况与洛特卡定律下的正常值相比较

| 处理后的发文数 | 5 | 3 | 2 | 1 |
|------------------------|------|-------|------|---|
| 与发文数为 1 的作者数之比 | 0.2% | 0.8% | 5.3% | - |
| 洛特卡定律下的正常值 ($1/n^2$) | 4% | 11.1% | 25% | - |

著名的洛特卡定律认为写 N 篇论文的作者数量约为写一篇论文作者数量的 $1/n^2$ ，而写一篇论文作者的数量约占所有作者数量的 60%^[4]。但是通过表 4 对比可以发现，工会维权领域的作者分布情况与洛特卡的描述差异巨大。具体有以下两点原因：1、洛特卡定律是对物理、化学两学科抽样研究出的理论，且非精确的统计分布，所以必然会产生一定误差；2、洛特卡定律要求研究的学科要相对稳定，研究的论文时间区间要足够长，作者人数要足够大，精确性才会相对提高^[5]，而我国工会维权领域的研究是近十年才逐步发展起来的，学科较不稳定，作者人数也相对较少。此外这种差异也在一定程度上反映出工会维权领域的研究门槛相对较低，缺乏核心作者，研究相对分散的情况。

多位作者的合作发文是作者关系最明显、最直接、最重要的体现，但除了合作发文这种相对显性的关系，同时还存在相对隐性的引用关系。作者通过大量阅读前人发表的文献，并在此基础上开展进一步的研究，在某种程度上也是一种跨域时空的单项合作。

为发现这种隐形的合作关系，本文采用引文网络分析的方发。引文，又称参考文献，通过正文后附参考文献列表以及在正文中进行引用标注，表明对前人研究成果的承认、借鉴、继承、修正、反驳、批判或者是向读者提供更进一步研究的参考线索等，从而在著文与引文之间建立起一种引证关系^[6]。

4 Lotka, Alfred J. (1926). "The frequency distribution of scientific productivity". Journal of the Washington Academy of Sciences. 16 (12): 317–324. 转引自维基百科

https://en.wikipedia.org/wiki/Lotka%27s_law

5 Egghe, Leo (2005). "Relations between the continuous and the discrete Lotka power function". Journal of the American Society for Information Science and Technology. 56 (7): 664–668 转引自维基百科 https://en.wikipedia.org/wiki/Lotka%27s_law

6 周晓英. 基于引文网络的知识链接框架研究[J].情报杂志,2010, (10): 21-22.

表 5 引文关系权重设计表

| | 期刊文献或硕博论文 | 书籍或其他文献 |
|--------------|-----------|---------|
| 引用 | 0.1 | 0.25 |
| 被引 | 0.5 | 无 |
| 注：合作发文的权重为 1 | | |

为使能够如实反映工会维权领域文献的作者网络，对引文关系进行加权处理是十分必要的。绝大多数文献的参考文献量远大于引证文献量；硕博论文的参考文献量是期刊文献的参考文献量的数倍到十余倍，而其中绝大多数为期刊文献，据此对引文关系进行了权重设计，详见表 5。

在对引用关系进行分级加权处理后，为更好得对作者网络进行可视化展示，减少次要引用关系对视觉的干扰，对数据进行两次过滤：第一次使用“K-核心”过滤方法，过滤度小于 2 的节点，即过滤连入度与连出度之和小于 2 的作者节点，即过滤掉作者网络中与其他作者关系过弱的作者；第二次过滤作者关系总权重小于 1 的节点，从而使核心作者网络更为清晰。

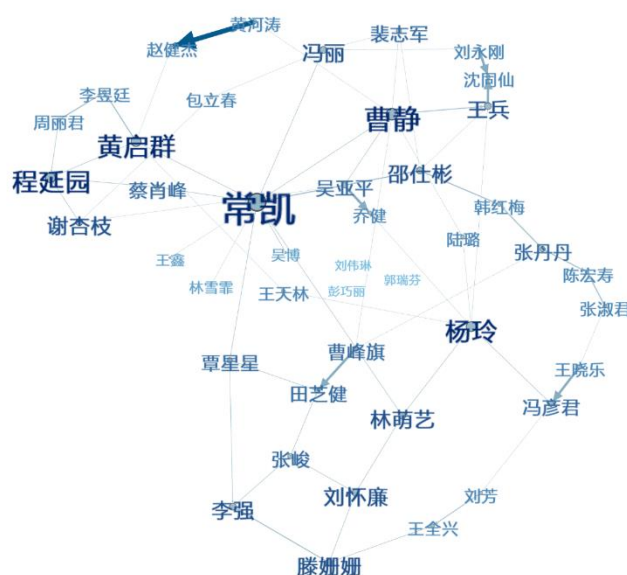


图 6 作者网络图

可视化方面，使用基于 Java 语言的 Gephi 进行处理，使用节点和节点字体大小及颜色的深浅来表示度的量，使用边的粗细和颜色的深浅来表示加权后

两位作者间引文关系的强弱程度，使用内置的 ForceAtlas 2 布局，开启劝阻 Hubs、LinLog 模式和防止重叠，并在此基础上对布局进行手动微调，使引文网络更为清晰，由此生成的作者网络图见图。

表 6 作者网络中关键节点简介

| 作者 | 简介 |
|-----|--|
| 常凯 | 中国人民大学劳动人事学院教授、中国人民大学劳动关系研究所所长，代表作品有《工会法通论》等 |
| 程延园 | 中国人民大学劳动人事学院教授、中国人民大学劳动关系研究所副所长，主要著作包括《劳动关系》、《劳动争议案例》等 |
| 黄启群 | 复旦大学公共管理硕士 |
| 曹静 | 天津大学公共管理硕士 |
| 杨玲 | 中南大学公共管理硕士 |

引文网络中关键节点简介详见表 6，其中曹静和杨玲各有一到两名同名作者，但由于同名作者的影响力较小，不影响引文网络的整体情况，故未做剔除。

通过对作者网络中关键节点的探索发现，中国人民大学在工会维权领域可能具有较强的影响力；部分著名学者通过本人著作的被引，对工会维权领域的研究产生影响；由于研究领域的体量相对较小，也并非热点研究领域，高质量的硕博士毕业论文成为本领域的关键文献，被大量引用。

5、学科分布

学者和期刊编辑在标注文献分类号时，会将最能总结本文内容的分类号放在首位，将次重要的分类号置于其后。同时，绝大多数文献仅使用一到两个分类号，拥有三个或三个以上分类号的文献极为少见，所以本文使用表 7 的方式对分类号进行加权处理。将采集到的中图分类号按第一分类号权重为 1，第二分类号权重 0.7 的方式进行统计，制成图 5，学科分布旭日图。旭日图中由内到外第一环为中图分类号一级类目，如 D 政治、法律、F 经济；第二环主要为中图分类号三级类目，如 D41 工人运动与组织；第三环则为四级类目，如 C913 社会生活与社会问题，第四环则为样本文献所标注的原分类号，如 F249.26 劳动关系。

表 7 学科分布权重设计表

| | 第一分类号 | 第二分类号 | 第三分类号 | |
|----|-------|-------|-------|-------|
| 权重 | 1 | 0.7 | 0 | 0 |

根据图 7，可以发现 D-政治、法律大类占据了近 3/4 的比例，继续细分可以发现，其中 D412.6-工会工作一类占据了近 60% 的比重，是工会维权领域研究中最常用的分类，即本领域研究的核心内容。

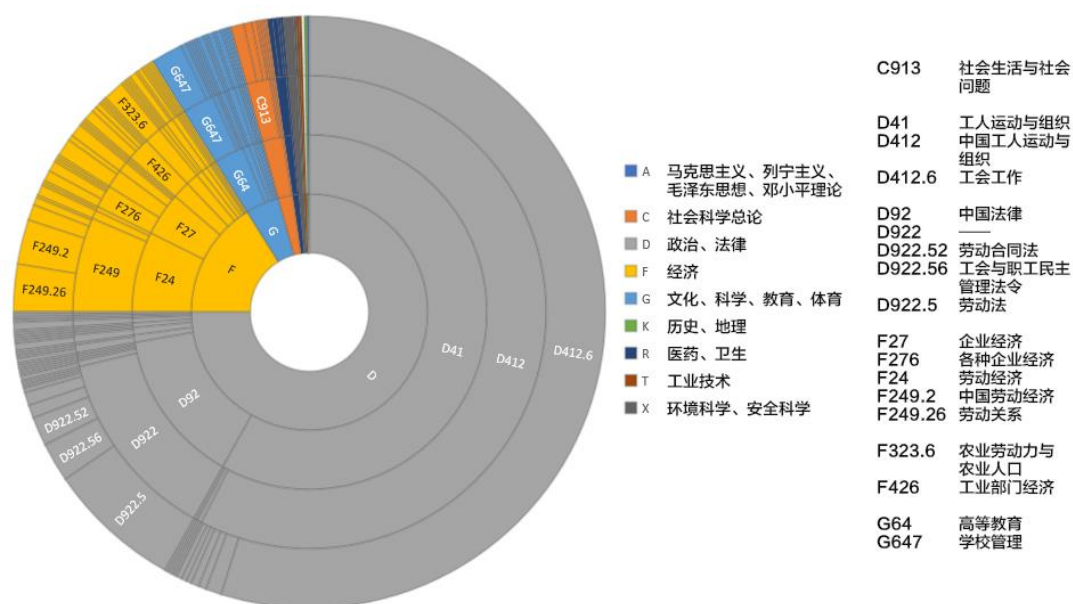


图 7 学科分布旭日图

此外，三级类目 D92-中国法律占据了整体约 20% 的比重，而其下的细分类目约 70% 为 D922.5-劳动法。可见有众多学者从法律视角研究工会维权问题，其中主要是从劳动法的角度进行研究，此外也有少数学者从行政法等领域展开讨论。

F-经济大类占比与 D92-中国法律相当，基本可细分为三类：F2-经济计划与管理、F3-农业经济、F4-工业经济，其中 F2-经济计划与管理下的 F249.2-中国劳动经济约占本大类 30% 的比重。值得一提的是，F323.6-农业劳动力与农业人口占据了约 10% 的比重，其中主要研究的研究主体为农民工，可见农民工相关的工会维权问题较为突出，受到部分学者的重视。

G-文化、科学、教育、体育大类约占总体的 10%。这一类中绝大多数为学

校学者在高校教职工工会方面的研究。经抽样调查发现，使用该分类发文的很多学者并不是人文社科学者，而是自然科学学者。他们的发文多数是自身在教学研究过程中遇到了工会维权的相关问题，进而有所研究、思考产生的成果。这一方面说明工会维权问题有很广泛的涉及面；另一方面也侧面说明该领域对知识背景的要求不高，研究门槛相对较低。

亦有较少学者使用 C-社会科学总论大类，从社会学角度进行研究。此外也有极少数学者从工会的历史发展（K 大类）、劳动卫生（R 大类）、工会维权的相关信息系统开发（T 大类）、劳动安全工程（X 大类）等角度进行研究。

从以上分析，可以很明显的发现，工会维权领域研究成果所使用的分类号以 D412.6-工会工作为主，以经管、法律相关分类号为辅，同时有部分非主流分类存在。由此，可以得出一个设想：工会维权领域研究存在较为明显的学科交叉现象。

6、引文关系

为进一步证实多学科交叉的猜想，笔者对样本文献中引证关系进一步处理，将文献间的引用关系转变为文献分类间的关系。由于第一和第二分类号体现文献关键所属学科的能力不同，所以在利用参考文献和引证文献研究引文关系时应对分类号进行一定程度的加权，本文的权重设计见表 3。

表 3 引文关系权重设计表

| | 被引文献第一分类号 | 被引文献第二分类号 |
|-----------|-----------|-----------|
| 引用文献第一分类号 | 1 | 0.7 |
| 引用文献第二分类号 | 0.7 | 0.4 |

由于各作者和编辑标记分类号的标准不同，但绝大多数都采用三级分类或更细的分类，所以为了统一统计标准，并使图像表意更为清晰，多数使用二级分类，少数使用三级分类。在此基础上，使用基于 JavaScript 的 Echarts 插件绘制图 8，图中自左到右分别为参考文献、样本文献和印证文献。

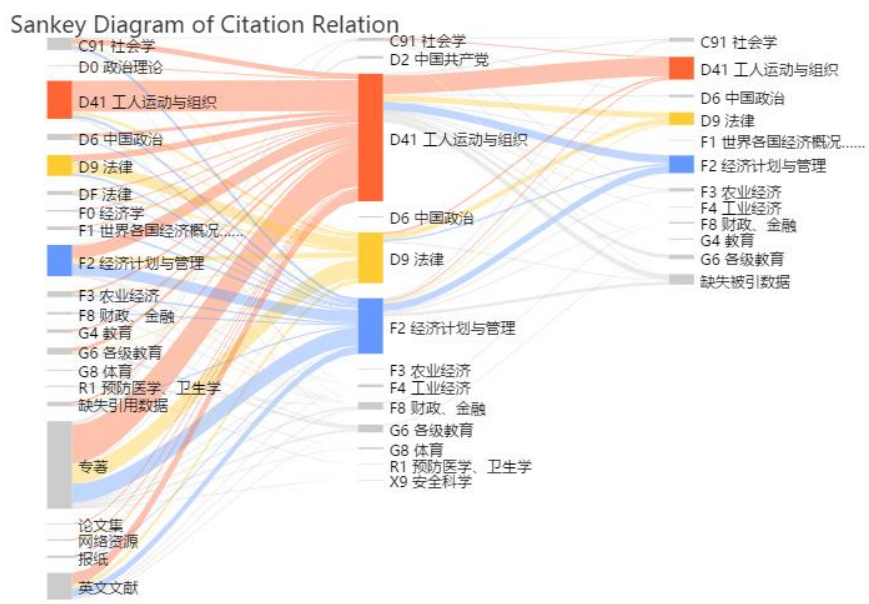


图 8 学科关系桑基图

由图可发现，样本文献中占大多数的 D41-工人运动与组织、D9-法律和 F2-经济计划与管理分类下的样本文献大量使用非本分类的参考文献，同时其成果也被其他领域的文献大量引证，呈现出非常强的学科交叉的特点。

四、结语

根据一系列的文献计量分析，可以发现工会维权领域的研究呈现两大特点：1、分散有余而集中不足，期刊分布中相关区期刊的数量相对不足，而作者分布中集中程度严重不足，同时缺乏长期在本领域研究做出可观贡献的核心作者；2、学科交叉现象明显，工会维权研究中政党及群众组织、劳动法和经济管理三大学科相互交叉、融合，在一定程度上推动了本领域研究的发展、深入。

此外在年代分布研究过程中提出的猜测也值得继续求证，即工业企业的亏损是否会影响到工人权利的保障，如果会，工会如何在其中发挥更大的作用，工人如何更好地保障自身的权利。