

中国图书情报领域专利研究的计量分析

刘玉琴 刘晶 张勇斌

北京印刷学院绿色印刷包装产业技术研究院 北京 102600

摘要 为发现国内图书情报领域专利研究的现状与趋势,论文基于文本分析与可视化工具 ITGInsight 对中国图书情报领域 22 种核心学术期刊 2000 年 1 月-2018 年 4 月专利研究相关论文进行计量分析,采用数据统计、文本挖掘、信息可视化方法,分析该领域研究趋势、主题分布和演化、期刊关联和演化、机构关联和演化、作者分布和演化,挖掘国内图书情报领域专利研究的特征。研究发现:中国图书情报领域专利研究的热度经过 2009-2014 年快速发展后,在最近 5 年有所减弱;早期侧重专利检索、知识产权,中期、后期侧重专利分析、专利地图,近期侧重高校图书馆知识服务;整体研究重点以专利分析方法与领域专利分析为主;情报类期刊研究偏重于专利分析、专利地图,图书馆类期刊偏重于知识产权、高校图书馆知识服务、专利检索;文献主要集中在情报类期刊;突出的研究机构包括中国科学技术信息研究所、北京工业大学经济与管理学院、中科院成都文献情报中心等单位。

关键词: ITGInsight; 专利分析; 可视化; 演化分析

中图分类号: TP391 G35

开放科学(资源服务)标识码(OSID)



Quantitative analysis of patent research in China's library and information field

LIU Yuqin LIU Jing ZHANG Yongbin

Academic of Printing and Packaging Industrial Technology, Beijing Institute of Graphic Communication, Beijing 102600, China

基金项目: 国家重点研发计划重点专项“知识产权信息共享与运营服务应用示范”(2017YFB1401900)。

作者简介: 刘玉琴(1979-),高级工程师,博士后,研究方向:专利分析、信息可视化, E-mail: liuyuqin2004@126.com; 刘晶(1979-),讲师,研究方向:科研管理、创新管理; 张勇斌(1974-),副教授,研究方向:大数据、人工智能。

Abstract To discover the situation and trend of patent research in library and information field of China, this paper analyzed the papers published from 2000.1 to 2018.4 related to the patent research of twenty two journals. The study adopted data statistics, text mining, and information visualization methods to analyze the research trends, topic distribution and evolution, journals distribution and evolution, institutional relevance and evolution, author distribution and evolution. The results showed that the popularity of patent research in the field of Chinese library and information technology has weakened in the last five years after the rapid development of 2009–2014; the early stage focus on patent search, intellectual property, patent analysis, patent maps in the middle and later stages, and recent emphasis on university libraries knowledge services; The overall research focus on patent analysis methods and field patent analysis; intelligence journal research focuses on patent analysis, patent maps, library journals focus on intellectual property, university library knowledge services, patent search. The outstanding research institutions include the China Institute of Scientific and Computational Information, the School of Economics and Management of Beijing University of Technology, and the Chengdu Document and Information Center of the Chinese Academy of Sciences.

Keywords: ITGInsight; patent analysis; visualization; evolution analysis

引言

现代社会进入了信息社会，信息资源已成为现代社会中最重要的战略资源之一。专利作为一种特殊的信息和战略资源，在国家信息资源建设开发与利用中有着特殊的地位和作用^[1]。不同领域研究者对专利展开分析和讨论，如企业管理、创新管理、技术预测、知识产权、情报分析等领域。专利作为文献计量学研究的主要对象之一，同样得到图书情报领域专家学者的高度重视。研究内容即涉及情报方法^[2-5]，又包括具体领域的情报分析^[6-10]，以及专利数据库、软件工具的使用^[11-13]等诸多方面内容。国内著名学者邱均平于2010年对比分析CSSCI与SCIE中1998-2008年间国内外图书情报领域的专利研究^[14]，得出我国专利分析起步晚，合作度不高，研究领域差别大的结论。谭晓基于Web of Science数据对图书情报领域专利分析文献进

行计量分析^[15]，数据检索仅限于专利分析相关的文献，挖掘专利分析方法、技术以及应用等方面的最新进展。邱均平与谭晓对图书情报领域专利分析计量所依赖的数据距今较远，研究范围一个面向国内外，一个仅限于专利分析。

本文选择中国图书情报领域2000年以后发表的核心期刊论文作为数据样本，从不同角度分析国内图书情报领域在专利研究上的主要特征，发现持续跟踪国内图书情报领域专利研究的现状与趋势，研究内容对选择合作伙伴、识别竞争对手、发现研究空白、科研评价与科研管理具有一定的借鉴和参考意义。

1 数据来源与分析工具

1.1 数据来源

本文以中国知网图书情报领域核心期刊作为数据源，采用高级检索，检索时间为2018年

5月15日,限定文献时间范围2010年1月-2018年4月,采用精确匹配,在主题、关键词、摘要中检索含有关键词“专利”,期刊名称含有关键词“图书”或“情报”的期刊文献,自定义格式下载到本地后依据论文名称去重,最后获得22个期刊1713篇文献^①。

1.2 分析工具

随着信息技术的发展,文本挖掘和信息可视化技术已被应用到情报分析领域,产生了众多的情报分析工具。本文应用文献分析工具ITGInsight进行文献计量分析^[16],采用数据统计、文本挖掘、信息可视化方法,从时间、主题、期刊、机构、作者角度,挖掘中国图书情报领域专利研究特征。ITGInsight是作者开发的一款高级的科技文本挖掘与可视化分析工具,主要针对科技文本,如专利、论文、报告、报刊等进行可视化分析与挖掘,可视化挖掘方法有合著关系可视化、同现关系可视化、耦合关系可视化、关联关系可视化、引证关系可视化、演化分析可视化,可视化输出包括网络图、热力图、

世界地图、矩阵图、演化图。该工具增强了对大规模数据的处理,将聚类分析、技术热力图、技术地形图、技术气象图整合到系统中。目前,ITGInsight在众多图书情报机构展开应用。

2 研究论文的计量分析

2.1 年度趋势

论文发表数量和发表时间趋势一定程度上可以反映出某学科或研究领域的发展状态、热度和趋势。图1为中国图书情报领域核心期刊论文中专利研究的文献数量趋势变化。从图1可以看出,相关论文累计数量呈递增趋势,2000-2013年间每一年度论文数量高低起伏变化,2000-2009年各年论文数量总量不大,均在50篇左右,2010-2014年快速增加,均在100篇左右;2014、2015年高峰均达到182篇;之后论文有递减的趋势。进一步查看论文作者和机构的数量趋势情况,发现与之具有相似的趋势变化。专利研究的热度经过2009-2014年的快速发展后,在最近5年有所减弱。

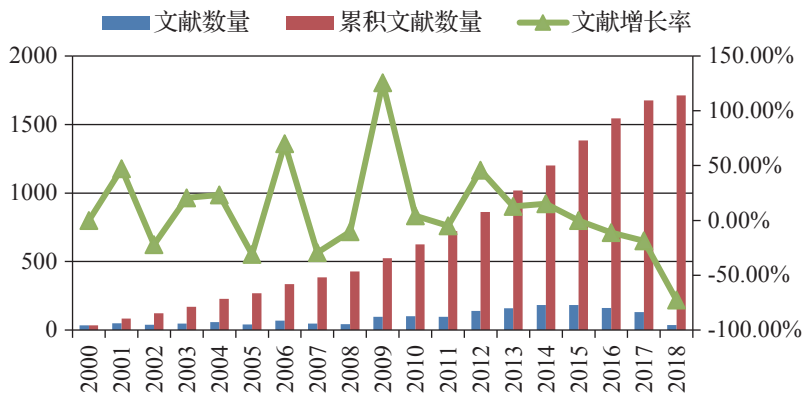


图1 中国图书情报领域专利研究论文趋势变化

①.TXT格式下载数据地址为: <https://pan.baidu.com/s/1wkPvUBxW3wzuPzK4V75MPQ>密码: a096; 去重后数据下载地址为: <https://pan.baidu.com/s/1RaTbkJSJNobNCPbLbDfCRQ>密码: 95Jw, 去重后数据为dataset格式, 需要用ITGInsightV1.3.7打开。

2.2 主题分布和演化

技术主题图通过类似于地理信息系统中的等高线图实现对科技文本数据的可视化,通过颜色的深浅区别数据的多少以及数据之间的关系^[17]。技术主题图是进行技术主题布局分析的典型计量学方法之一。基于 ITGInsight 的主题图绘制功能,绘制中国图书情报领域专利研究的主题分布,如图 2 所示。国内图书情报领域专利研究可以分为以下几方面:

1) 专利分析方法与领域专利分析研究,主题词涵盖“专利分析、专利计量、专利地图、数据挖掘、文本挖掘、文献计量、专利引用、知识图谱、可视化”等;其中领域分析较为分散,主题词并不显著、也不聚焦,查询文献后发现,关于具体领域的专利分析涉及能源、医药、海洋工程、智能手机、碳纤维、网络认证等技术。

2) 专利法律属性研究,主题词涵盖“核心专利、专利诉讼、专利预警”等。

3) 高校专利利用,主题词涵盖“高校、专利申请、发明专利”等。

4) 专利检索,主题词涵盖“专利数据库、专利检索、专利引文、文献检索”等。

5) 高校图书馆专利信息服务,主题词涵盖“高校图书馆、信息服务”等。

6) 中国、日本、美国专利数据的使用。此外,还有科技创新,本体,知识产权管理等小众化研究主题。

主题演化分析作为学科新兴趋势探测方法之一,有助于了解领域主题产生、消亡、增强、减弱、聚合和裂变的过程,在情报分析中广泛

应用,也得到众多研究者的关注。国内著名学者王晓光设计主题演化工具 NEViewer,对学科主题演化过程进行计量分析和可视化展示^[18]。唐果媛应用共词分析构建学科知识结构的主题演化模式分析方法^[19]。本文应用 ITGInsight 进行主题演化跟踪,以主题词演化反应领域的主题变化情况。其基本过程:提取每个时间段的主题词(或主题词组),统计各主题词(或主题词组)频数,列出前 N 个主题词,依据数量多少排序;计算各个时间段主题词(或主题词组)之间的同现关系强度,上一阶段的同现关系作为上一阶段主题词(或主题词组)与下一阶段主题词(或主题词组)之间的关系强度。基于相同的计算方法,ITGInsight 将主题演化拓展到作者、机构、地区、关键词、技术类别演化。

筛选出每一年度数量最多的前 10 个主题词,应用 ITGInsight 绘制主题词演化图,如图 3 所示。从图 3 可以看出,2000-2003 年主要是对专利文献、专利检索、专利数据库进行的研究分析;2004 年转到知识产权、企业管理、专利战略方向,之后知识产权角度进行的专利研究一直持续,但排序逐渐被专利分析、专利地图、专利计量取代;专利分析、专利地图、专利计量从 2005 年开始排序一直较为靠前;2012 年开始人们在分析专利情报时,对专利质量逐步重视,不断提到核心专利;2015、2016、2017 高校图书馆的知识服务排序逐年靠前。整体上看,早期侧重专利检索、知识产权;中期、后期侧重专利分析、专利地图;近期侧重高校图书馆知识服务。

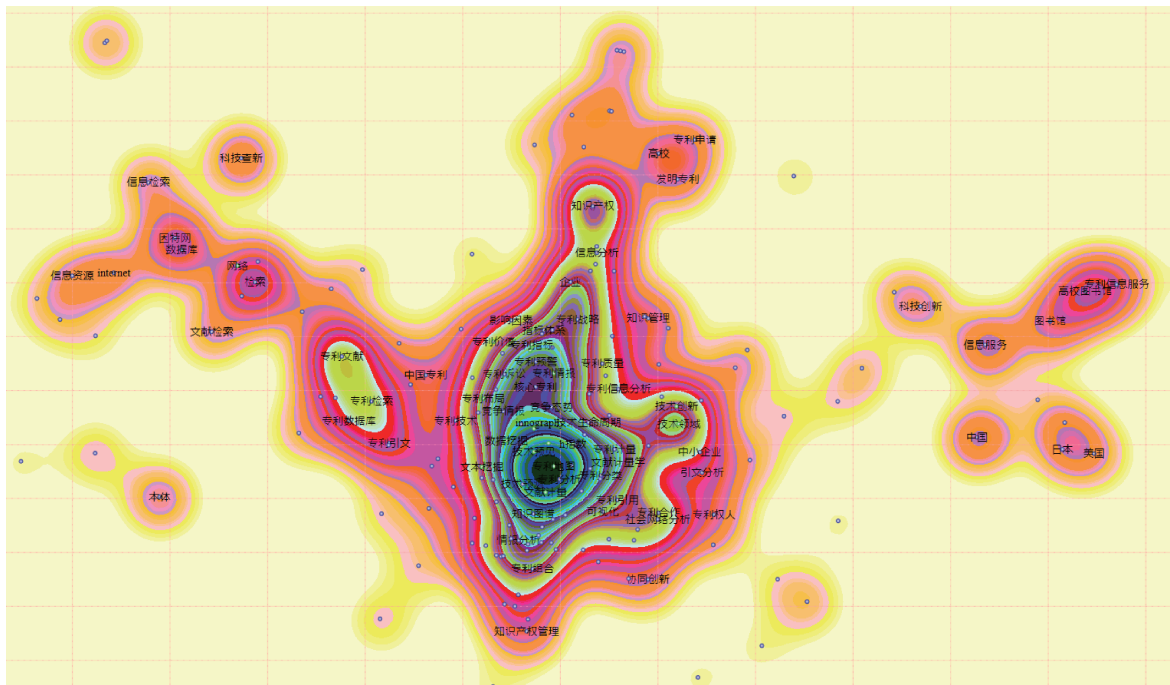


图 2 中国图书情报领域专利研究主题词分布

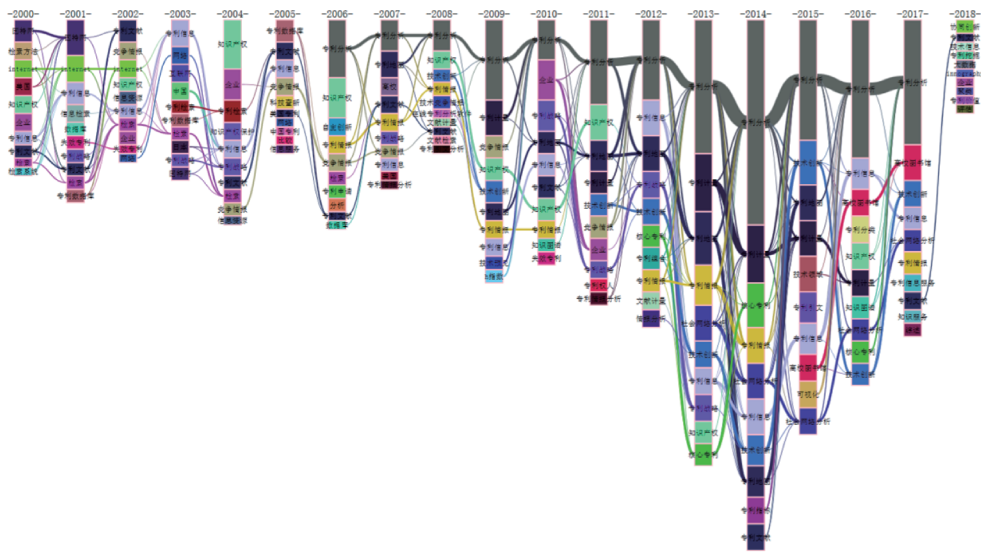


图3 中国图书情报领域专利研究主题词演化

2.3 期刊关联和演化

本次分析的图书情报领域专利研究核心期刊包括 22 种期刊, 图 4 为这 22 种期刊之间的研究主题关联性分析。图中节点数字为该期刊相关论文数量, 节点大小与之成正比。连线粗

细表示期刊研究主题的相关性强度,节点备注为期刊使用最多的关键词及其数量。计算方法为 ITGInsight 的关联分析^[20],依据论文的主题词相似性度量期刊之间的关联程度。从图 4 可以看出,情报杂志、图书情报工作、情报理

论与实践这3种期刊相关论文最多,占比超过60%,分别达到591、256、193篇。情报类期刊研究偏重于专利分析、专利地图,图书馆类期刊偏重于知识产权、高校图书馆知识服务、专利检索,情报类期刊的研究在数量上占绝对优势。这与两类期刊的办刊宗旨有一定的关系。从图5的期刊演化趋势看,2007年以前情报杂志、情报科学、现代情报期刊的排序较为靠前;

2008 年以后情报杂志仍然保持靠前的位置, 图书情报工作、情报理论与实践则替换情报科学、现代情报, 在专利研究领域迅速发展。现代情报在中断一段时间后, 相关研究论文于 2015 年又逐渐增加, 并且有超过图书情报工作和情报理论与实践的趋势。在整个研究中, 情报学报在专利研究方面的文献数量并不显著, 这与该期刊的出版周期有一定关系。

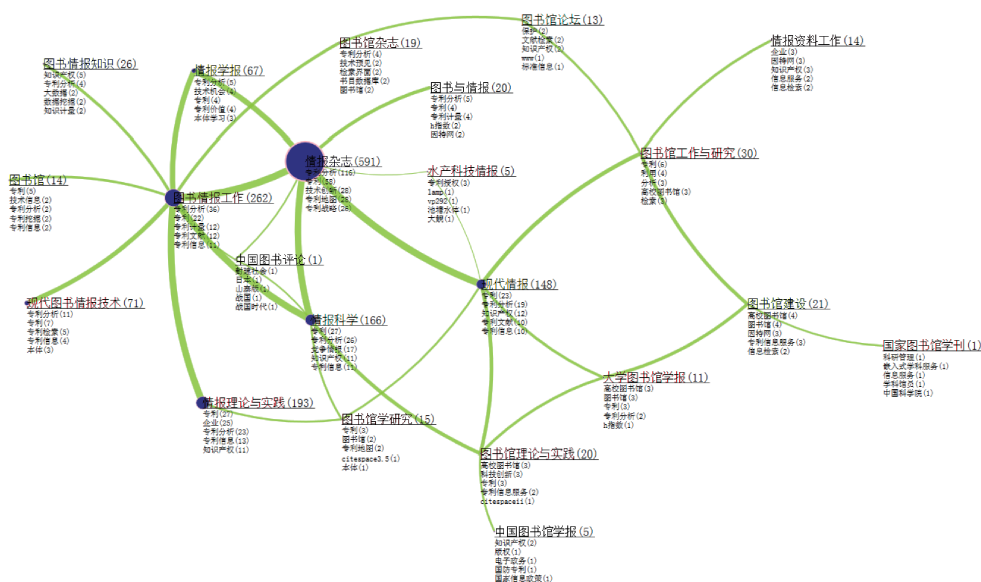


图 4 中国图书情报领域专利研究期刊关联

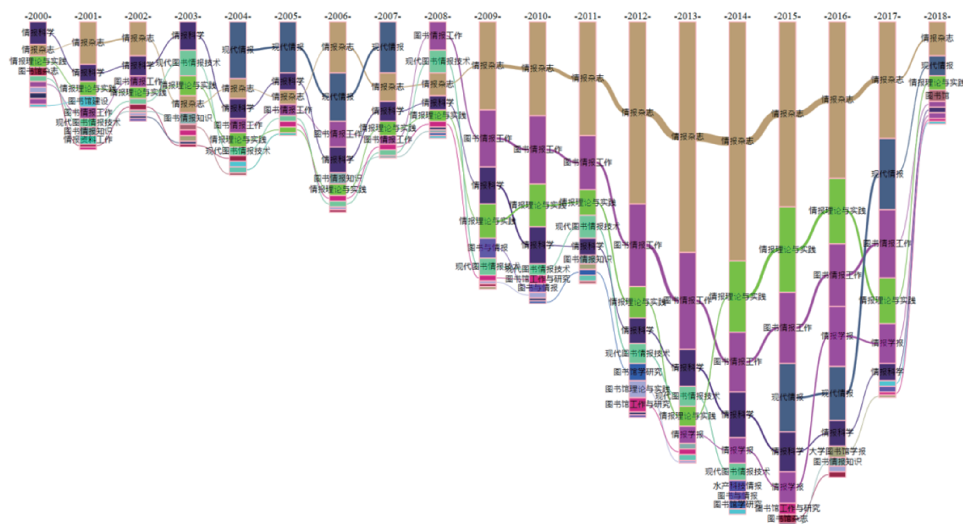


图5 中国图书情报领域专利研究期刊演化

2.4 机构关联和演化

基于文献文本特征的关联分析是挖掘领域研究主体研究内容关联性的有效方法之一, ITGInsight 采用文献 [21] 中提出的学术主体关联算法构建机构、作者、区域的关联关系。本文利用 ITGInsight 的机构关联分析, 挖掘发文机构的研究主题词侧重, 计算发文机构之间的主题关联强度, 揭示发文机构之间的竞争关系, 如图 6 所示。图中节点大小与专利文献数量多少成正比, 节点红、绿、黄色分别表示署名第一、第二、第三及以后的专利文献数量。图中连线粗细与发文机构之间的主题关联强度成正比。节点标注文字为该发文机构名称。从研究内容上看, 中国科学技术信息研究所研究主题广泛, 与众多科研机构表现出了一定的相关性, 且其论文数量最多, 达到 112 篇, 主要发表在情报杂志、图书情报工作、情报理论与实践、情报学报和情报科学上。北京工业大学经济与管理学院排名第二, 达到 59 篇, 主要发表在情报杂志和情报学报, 且与其他机构的主题关联性较弱, 仅与北京理工大学管理与经济学院、上海理工大学管理学院表现出较弱的关联性。中国科学院大学 54 篇, 但绝大部分为第二、三作者的合著论文, 与中国科学院国家科学图书馆、中国科学院成都文献情报中心、中国科学院国家科学图书馆成都分馆、中国科学院文献情报中心合著较多(图 6 椭圆部分)。华中科技大学管理学院 53 篇, 位列第 4, 主要发表在情报杂志上, 研究主题涉及专利质量、专利诉讼、专利池, 与其它发文机构的关联性不显著; 江苏大学科技信息研究所 48 篇, 位列第 5, 主要发表在情报杂志、情报理论与实践。此外,

吉林大学管理学院、南京大学信息管理学院、山东理工大学科技信息研究所、北京理工大学管理与经济学院、湘潭大学知识产权学院、南京工业大学图书馆、华中科技大学知识产权战略研究院、哈尔滨工程大学经济管理学院、上海图书馆上海科学技术情报研究所、四川大学公共管理学院、南京大学信息管理系的文献数量均在 10 篇以上。

选择每一年度文献数量排序前 10 位的机构绘制机构演化图, 如图 7 所示。从图 7 的机构演化来看, 南京大学 2002 年、2003 年发文数量排序靠前, 是较早从事相关研究的机构, 但后续研究跟进不足。中国科学技术信息研究所 2009 年以后在专利研究上一直位列前三。北京工业大学经济与管理学院在 2013、2014、2015 三年发文迅速增加, 在 2015 年达到峰值, 2016、2017 年有减弱的趋势。江苏大学科技信息研究所的发文排序呈倒 U 形趋势, 在 2014 年触底反弹。

2.5 作者合著与演化

提取发文数量排序前 100 名的论文作者, 构建作者合著关系网络图, 如图 8 所示。图中节点大小与文献数量多少成正比, 节点红、绿、黄色分别表示署名第一、第二、第三及以后的文献数量。图中连线粗细与合著数量多少成正比。在中国图书情报领域, 专利研究第一群体是以翟东升、黄鲁成为中心的北京工业大学经济与管理学院团队; 研究主题偏向关联分析、专利引用、专利共类、MapReduce 等; 发表期刊主要为情报杂志、情报科学。第二群体是以方曙、张娴、许海云、陈云伟等为中心的中科

院成都文献情报中心团队；研究主题偏向引文分析、合作网络、技术演化、指标分析等；发表期刊主要为图书情报工作。第三群体是以刘桂锋为中心的江苏大学科技信息研究所团队；研究强调大数据、专利组合、核心专利、专利指标等；发表期刊主要为情报杂志。此外山东

理工大学科技信息研究所的吴红团队、武汉东湖新技术开发区管委会的周磊团队、湘潭大学知识产权学院的邱红华、四川大学管理学院的李睿、河南科技大学管理学院温芳芳、北京理工大学的汪雪锋、朱东华团队等，也是专利研究领域比较突出的研究团队或个人。

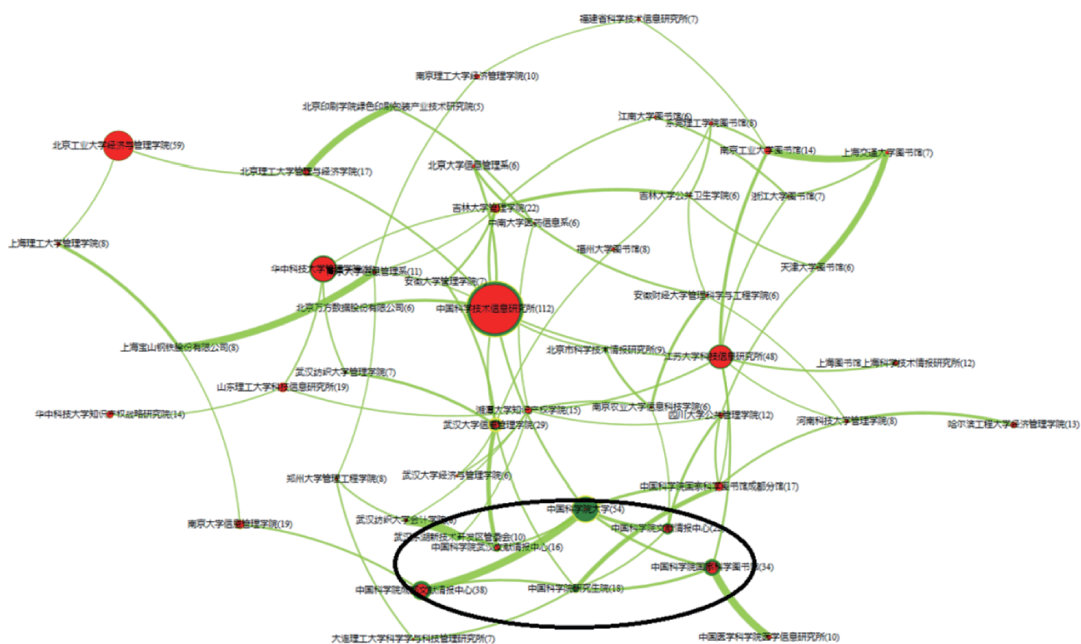


图6 中国图书情报领域专利研究机构关联

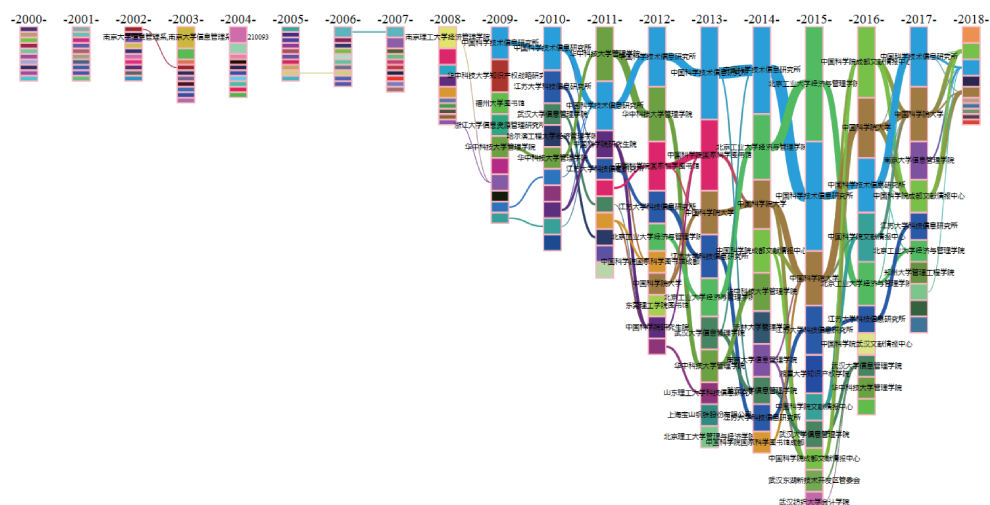


图7 中国图书情报领域专利研究机构演化

在作者合著网络中,分析作者单位属性,会发现作者之间的合著呈现地区性、系统性、内部性特点。如武汉地区的高校、研究机构合著,中国科学院系统的研究机构合著,高校内部研

究者之间的合著相对显著。中国科学技术信息研究所与众多科研机构 and 高校之间具有合著关系, 但合著数量较少。除此之外, 机构之间的合著总体不显著。

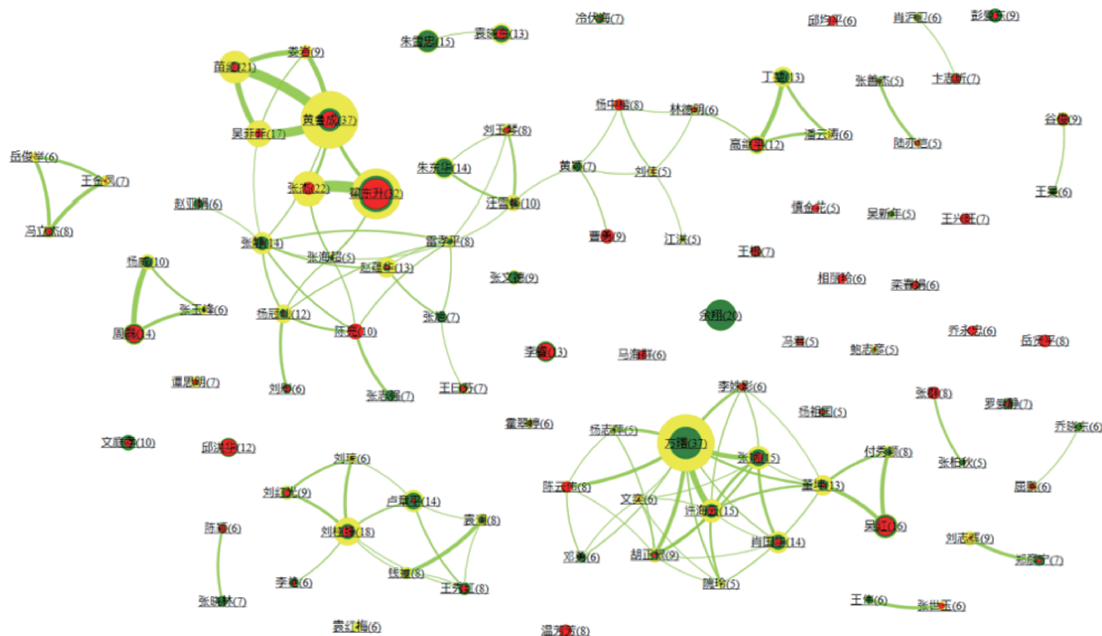


图 8 中国图书情报领域专利研究作者合著

选择每一年度文献数量排序前 10 位的作者绘制机构演化图,如图 9 所示。从图 9 历年作者演化趋势图看,北京工业大学经济与管理学院翟东升、黄鲁成团队在 2012-2015 年一直较为靠前;其团队成员吴菲菲、张杰、苗红也是较为突出的研究者。中科院成都文献情报中心方曙、许海云团队在 2016、2017 年排序超过北京工业大学经济与管理学院,方曙是该领域持续研究的重要学者之一。江苏大学科技信息研究所刘桂锋团队在 2015、2016 年较为突出。北京理工大学朱东华团队是早期该领域的重要研究团队之一。

3 总结

对中国图书情报领域专利研究特征进行分析, 分析结果总结如下:

(1) 专利研究的热度经过 2009-2014 年的快速发展后,在最近 5 年有所减弱。

(2) 研究主题主要涉及: 1) 专利分析方法与领域专利分析研究, 领域专利分析涉及能源、医药、海洋工程、智能手机、碳纤维、网络认证等技术。2) 专利法律属性研究。3) 高校专利利用。4) 专利检索。5) 高校图书馆专利信息服务。6) 中国、日本、美国专利数据的使用。

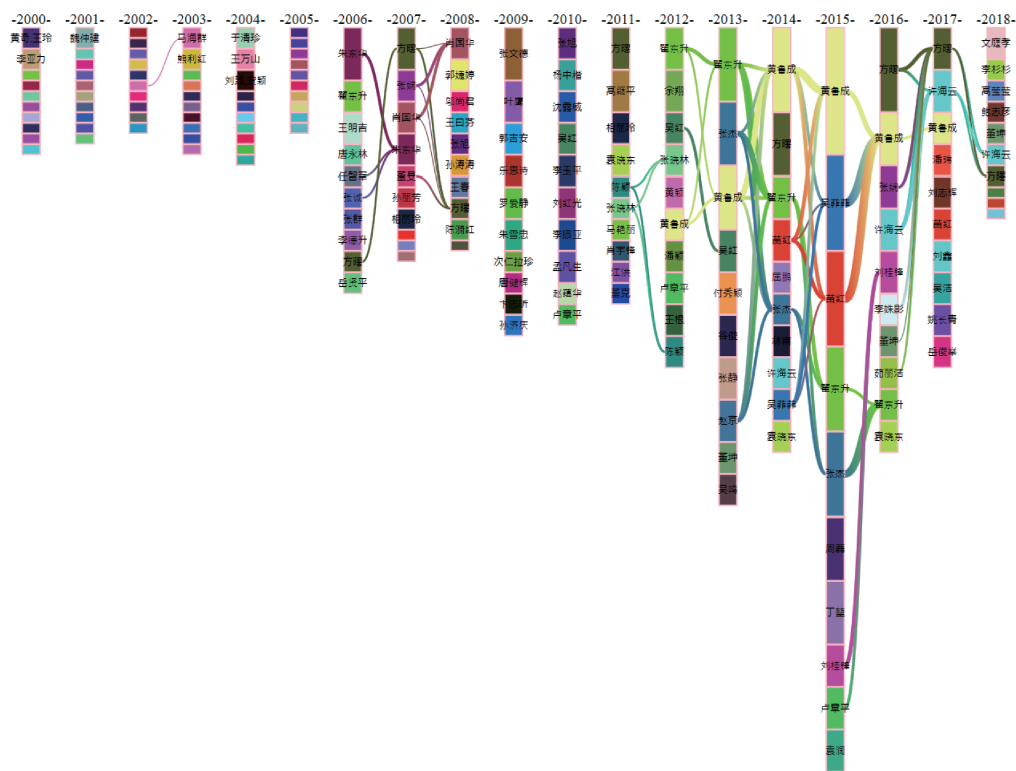


图9 中国图书情报领域专利研究作者演化

(3) 研究主题变化, 早期侧重专利检索、知识产权; 中期、后期侧重专利分析、专利地图; 近期侧重高校图书馆知识服务。

(4) 情报杂志、图书情报工作、情报理论与实践这3种期刊相关论文最多, 占比超过60%。情报类期刊研究偏重于专利分析、专利地图, 图书馆类期刊偏重于知识产权、高校图书馆知识服务、专利检索, 情报类期刊的研究在文献数量上占绝对优势。

(5) 中国科学技术信息研究所论文数量最多, 研究主题广泛, 与众多科研机构表现出了一定的相关性。北京工业大学经济与管理学院排名第二, 与其他机构的主题关联性较弱。中国科学院大学大部分为第二、三作者的合著论文, 与中国科学院国家科学图书馆、中国科学

院成都文献情报中心、中国科学院国家科学图书馆成都分馆、中国科学院文献情报中心合著较多。

(6) 南京大学是较早从事相关研究的机构, 但后续研究跟进不足。中国科学技术信息研究所2009年以后在专利研究上一直位列前三。北京工业大学经济与管理学院在2013-2015年发文迅速增加, 近两年有减弱的趋势。江苏大学科技信息研究所的发文排序呈倒U形趋势。

(7) 作者合著网络呈现地区性、系统性、内部性特点。武汉地区的高校、研究机构合著, 中国科学院系统的研究机构合著, 高校内部研究者之间的合著相对显著。

专利蕴含了丰富的技术信息、法律信息, 是图书情报领域重要的研究对象。本文对国内

图书情报领域核心期刊专利研究相关文献进行计量分析,发现该领域研究趋势、主题分布和演化、期刊关联和演化、机构关联和演化、作者分布和演化,挖掘国内图书情报领域专利研究的特征,有利于客观认识领域发展趋势、研究侧重,对于合作伙伴选择、竞争对手识别、寻找研究空白点具有重要意义。下一步研究中,将进一步对比分析国内图书情报领域与科研管理领域的专利研究,国内与国外图书情报领域的研究,发现研究差异,寻找科研机会。

参考文献

- [1] 刘玉琴,彭茂祥.国内外专利分析工具比较研究[J].情报理论与实践,2012,35(9):120-124.
- [2] 文庭孝.专利信息计量研究综述[J].图书情报知识,2014(5):72-80.
- [3] 谢萍,袁润,钱过.基于TOPSIS方法的核心专利识别研究[J].情报理论与实践,2015,38(6):88-92+108.
- [4] 王博,刘盛博,丁莹,等.基于LDA主题模型的专利内容分析方法[J].科研管理,2015,36(3):111-117.
- [5] 孙涛涛,唐小利,李越.核心专利的识别方法及其实证研究[J].图书情报工作,2012,56(4):80-84.
- [6] 许琦,林晗.基于专利引证网络的专利产业化评价方法研究——以燃料电池领域为例[J].数字图书馆论坛,2016(9):49-56.
- [7] 吴云龙,赵蕴华.基于专利的中日美质子交换膜燃料电池技术发展对比研究[J].数字图书馆论坛,2016(8):16-23.
- [8] 贡强,彭喆.全球锂离子动力电池失效专利分析[J].数字图书馆论坛,2016(8):24-28.
- [9] 葛亮.基于专利信息分析的技术生命周期研究——以我国石墨烯制备技术为例[J].情报工程,2015,1(4):58-64.
- [10] 吴日欣,凌锋,刘会景.我国空气净化器技术现状专利分析[J].情报工程,2015,1(5):77-85.
- [11] 战玉华,潘乐影,程爱平.利用Innography进行专利情报分析——以OLED为例[J].图书情报工作,2013,57(18):104-109.
- [12] 王春,杨志萍,方曙.利用德温特分析家软件开展专利情报研究[J].现代情报,2005(12):14-17.
- [13] 冯浩然,方曙.Thomson Data Analyzer在机构专利情报分析中的应用研究[J].情报科学,2008,26(12):1833-1836+1843.
- [14] 邱均平,李慧.国内外图书情报领域专利计量研究的对比分析[J].图书情报工作,2010,54(10):83-87+91.
- [15] 谭晓,张志强.图情领域中专利分析主题的研究进展——基于WOS的文献分析[J].图书情报工作,2012,56(20):85-91.
- [16] 刘玉琴,汪雪锋,雷孝平.科研关系构建与可视化系统设计与实现[J].图书情报工作,2015,59(8):103-110+125.
- [17] 刘玉琴,逢金辉,崔志成,等.一种简易的技术主题图绘制方法[J].图书情报工作,2017,61(13):125-132.
- [18] 王晓光,程齐凯.基于NEViewer的学科主题演化可视化分析[J].情报学报,2013,32(9):900-911.
- [19] 唐果媛.基于共词分析法的学科主题演化研究方法的构建[J].图书情报工作,2017,61(23):100-107.
- [20] 王有国,刘玉琴,汪雪锋.面向科研评价的学术关系可视化方法研究[J].科学学与科学技术管理,2014,35(5):13-18.
- [21] 刘玉琴.学术关系可视化方法研究[D].北京:中国科学技术信息研究所,2013.