

基于传热学理论的文献时空分布与知识传播规律研究

赵志远 郑彦宁 赵筱媛 贾亚敏

(中国科学技术信息研究所 北京 100038)

摘要 引入了传热学中的理论研究方法,建立了文献时空分布和知识传播模型。基于传热学的理论在年份、热度、传播方式、衰减老化、影响能力、研究领域、研究内容等方面进行了创新性研究。提出了文献热度、文献热度累积量、文献热流量、文献热扩散系数等概念,用传热学中热源扩散以及能量守恒等观点研究文献的分布、知识传播能力以及生命周期的分析判断,并进行了实证研究。

关键词 文献时空分布,知识传播,传热学,数学建模,MATLAB,技术竞争情报

中图分类号 O236 **文献标识码** A

Research on Temporal and Spatial Distribution of Literatures and Knowledge Dissemination Based on Heat Transfer Science

ZHAO Zhi-yuan ZHENG Yan-ning ZHAO Xiao-yuan JIA Ya-min

(Institute of Scientific and Technical Information of China, Beijing 100038, China)

Abstract This paper introduced the theory and research method in heat transfer science and established the model of the temporal and spatial distribution of literatures and knowledge dissemination. And based on the theory of heat transfer science, we made the innovation research in the year, heat degree, transmission model, attenuation or aging, the ability of influence, research field, research content, and industry division. Then, this paper proposed a series of new concepts, which are literature heat degree, amount of literature heat degree accumulation, amount of literature heat flow, literature thermal diffusion coefficients and so on. Furthermore, we used the heat diffusion, energy conservation and other point of view in heat transfer science to research literature distribution, capability of knowledge dissemination, the analysis and judgment of life cycle. And finally, we carried out CTI empirical study.

Keywords Temporal and spatial distribution of literatures, Knowledge dissemination, Heat transfer science, Mathematical modeling, MATLAB, CTI

1 引言

在学术研究中,每个研究领域都是由一个或几个研究主题开始逐步向四周扩散,不断有学者开拓研究,不断有新的学术文献产生,最终拓展成一个研究领域。那么,文献的时空分布能否直观表示?知识传播的快慢能否度量?研究主题或文献的扩散能力能否定量表征?知识传播的快慢与研究主题或文献的影响力有何种关系?关系能否定量表示?不同或相似的研究主题、文献之间能否定量比较?这些都是值得情报学界学者深入思考的问题。在情报学领域,已经有学者研究知识的传播与扩散问题,也有学者应用物理学中的热力学领域知识来研究情报学问题的成功案例,比如熵的引入^[1]。

受热物理领域“熵”的概念的启发,本文作者认为文献的时空分布与知识传播规律可以使用物理学中的热量传递的众多规律解释。因此,本研究尝试用传热学中的理论方法和研究思路来研究在信息情报领域的文献时空分布与知识传播的规律,并以对技术竞争情报的研究为例,建立基于传热学理论

的文献时空分布与知识传播模型,从而为信息情报与文献计量领域提供一种新的研究方法和思路。

2 国内外相关研究现状

针对文献的时间分布和空间分布规律,前人已经建立了很多模型进行研究,如普赖斯指数函数模型、逻辑增长模型、Brookes 积累指数模型、布拉德福德定律、文献老化规律等^[1],这些大都是通过总结科技文献的分布规律进而提炼归纳出较理想的数学模型,这些原始模型只针对空间分布规律或者时间变化规律中的一种进行研究^[2]。后来的学者们大都是基于前人已提出的模型进行改进和创新完成分析研究,这也是研究主题向外扩散和知识向外传播的表现。当表征科学家影响的情况 H 指数提出后,一些其他指数例如 G 指数^[3]、AR 指数^[4]等基于 H 指数相继被提出。化柏林利用五要素构成的 TAIPPO 模型对文献计量分析研究的发展趋势进行探讨^[2];刘志辉、郑彦宁在作者关键词耦合分析方法的基础上提出利用研究专业演化图谱对研究领域演化进行分析的方

赵志远(1992—),男,硕士,主要研究方向为竞争情报、情报技术与方法, E-mail: hit_zzy@163.com; 郑彦宁(1965—),男,博士生导师,主要研究方向为竞争情报、情报技术与方法; 赵筱媛(1978—),女,研究员,硕士生导师,主要研究方向为竞争情报、知识服务、公共管理; 贾亚敏(1991—),女,硕士,主要研究方向为情报技术与方法、未来学。

法^[5];吴夙慧等人提出了一种基于学术文献同被引分析的初始聚类中心和K值选取的K-means改进算法^[6],并提出了基于引用上下文、文献同被引和文献耦合的相关反馈改进算法^[7];俞立平等人基于结构方程进行了学术期刊评价研究,为科技评价提供了一种新的思路^[8]。

把数学和物理学规律应用到情报学领域,国外开始得较早。1948年,信息论的创始人申农提出了信息的度量这一问题,把哈特利公式扩大到概率不同的情况,提出了著名的计算信息熵的公式^[1]。几乎在同一时间,维纳出版了学术巨著《控制论:或关于在动物和机器中的通讯和控制的科学》,在这一著作中,维纳类似地用熵定义了连续信号的信息量,把消息看成可测事件的时间序列,建立了维纳滤波理论,提出了度量信息量的申农-维纳公式^[10]。

在我国,使用数学物理方法定量研究信息情报领域开始得较晚。1988年,文学领域的学者邱均平在文献信息系统中尝试使用定量的研究方法^[11],从而引发了国内以数理统计和物理理论为研究工具和方法的学术潮流。吴国恩发表的《情报信息熵研究》和《情报信息的测度》对信息的物理描述^[12]和物理测度^[13]进行了深入研究。陈颖把申农的信息论中信息熵的概念和数学表示式引入文献信息这一信息科学下的分支学科来描述文献信息^[14]。进入21世纪以来,国内学者更多地把物理扩散和空间场论的知识引入到科技创新、网络信息扩散等研究领域。康凯把场论和系统论的方法理论迁移到技术创新扩散的领域,研究了技术创新空间扩散的散度、梯度及吸引力等方面的规律^[15]。近几年,邓忆瑞把物理学中场的概念迁移过来,使用场论的观点研究信息扩散的机理,尝试从场的角度描述信息扩散的本质和规律^[16],并且研究在网络环境下的信息扩散的微观运动形式^[17]。随着跨领域研究的进一步深入,许智宏等人应用蚁群算法在技术创新领域进行了扩散研究^[18]。

3 基于传热学理论的文献时空分布与知识传播模型

传热学就是研究由温差引起的热能传递规律的科学。热能的传递有3种基本方式:热传导、热对流与热辐射。传热学中称单位时间内通过某一给定面积的热量为热流量。与重力场、速度场等一样,物体中存在温度的场,称为温度场,它是各个时刻物体中各点温度所组成的集合,又称为温度分布。一般地说,物体的温度场是空间坐标与时间的函数^[19]。

在传热分析中,热扩散率 a 是热导率 λ 与比热容 c 和密度 ρ 的乘积之比。热扩散率又叫导热系数,它表示物体在加热或冷却中温度趋于均匀一致的能力;在 a 高的物质中热能扩散得很快,而 a 低的物质中热能则扩散得较慢^[20,21]。

在传热学领域中,决定热扩散率的根源是传热介质本身的物理属性。从上面的介绍可以看出,热扩散率 a 由热导率 λ 、比热容 c 、密度 ρ 决定,而热扩散率不同反映出来的结果就是不同层温度变化的快慢。总结来说,如图1所示,热扩散率由自身物性决定,表现出来是温度。

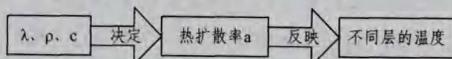


图1

正如热能的传递需要温差提供动能,本文认为知识的传播也需要势能差来提供能量。这种势能差是由影响力的不同

产生的,因此有必要研究文献或知识点之间的势能差,即应该提出体现各文献和知识的影响力不同的物理量,这有助于从事物的内因出发来研究外在的规律。

温度场即在物理空间中,不同的点具有不同的温度,空间中所有点的温度就构成了该空间的温度场。参照传热学用温度场来描述物理空间,本文使用热度场来描述文献的时空分布。本文把一篇文章看成一个点,领域内的所有文章构成了一个领域空间,这个空间存在一个热度场,而热度场中的每一点都有各自的热度,每一点的热度即这篇文章在当年的被引次数,用 T 表示,热度场是由场中的各点的热度组成的。热度 T 与温度相似,是一个瞬态量,会随着年份发生变化,从而热度场也是会随着时间变化的。热度场除了可以描述文献这种介质外,还可以描述专利、知识等介质。至此,就建立了一个描述科技文献时间和空间分布的热度场。使用MATLAB软件,编写MATLAB代码,绘制初始的文献时空分布与知识传播图,如图2所示。为方便观察,仅展示2年的热度场图,假如内部球层为2006年,外面透明球层为2007年,把2006年球层标为深色以便于观察。该热度场的球层由里向外,年份由远及近变化,2006年内部是早于2006年的年份球层,透明的2007年球层外面分别是2008年球层、2009年球层等。如果该领域开始于1995年,则球心起始于1995年一点,向外扩散生成完整的时空分布图。如图2所示,2006年有2篇文章A和B,2007年有文献A'、B'、B''、C、C',其中A'引用A和B,B'和B''引用B,C'引用C。通过观察2007年的热度图谱可以看出文献的热度情况,热度越高则越靠近上面的红色极点,热度越低则越靠近下面的蓝色极点,即 $T_C > T_{C'} > T_{A'} > T_B > T_{B'}$ 。

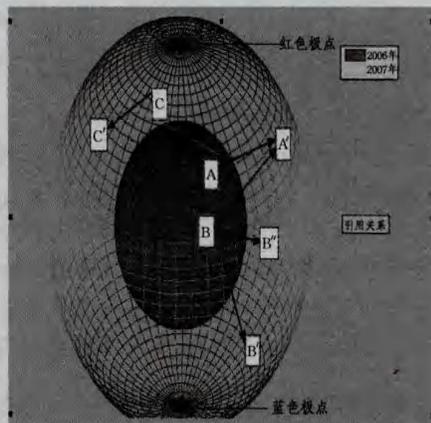


图2 文献时空分布与知识传播的热度场图(简单的基于年份的热度分布与知识传播关系)

传热学中热能通过热传导、热对流、热辐射3种基本方式传递。在某一领域中,任意两篇文献之间的关系同样可以有3种:第一种就是与热传导类似的较强关系,一篇文献是另一篇文献的施引文献或被引文献;第二种与热对流类似,两篇文献不直接接触,即两篇文献是同被引文献或共引文献;第三种与热辐射类似,两篇文献之间关联较弱或者没有关联,是除以上两种关系的其他任意关系,例如:二次印证或二次参考等。

与热量的传递会衰减一样,知识的传播过程也会出现老化现象。热量通过热源的扩散,熵值增加,周围环境逐渐达到平衡状态,这样势差就会降低,传递随之衰减;知识在传播的过程中也会从原来的较专业、较新颖的知识变为被大家广泛

接受和普遍认同的知识,这样传播的势差也会降低,随之所在领域中研究该知识点的热度就会逐渐降低。热量的传播过程有时会遇到奇点,即该点把传播过程中的热量不断吸走,使得热量传递过程与原始的均匀扩散的传播路径不同;而在知识的传播过程中也会遇到这样的情况,即当大家发现某个正在传播的知识点是错误的或没有研究的价值时,该知识的传播过程就遇到了奇点,该传播过程也随之改变,传播路径也发生变化。热量的传播过程还会遇到激发态的情况,即当热量传至某点时,该点的介质会发生化学变化被激发,从而变成热源或其他新的辐射源,使原来的传播环境变为新的传播环境;在知识的传播过程中,这种情况也较普遍,原有的知识在经过某个传播点时被完善、升华、改变,然后影响了传播环境,甚至会变成传播的起点而创建新的研究领域。

借鉴传热学理论中的热量累积和热流量,在研究主题或文献扩散领域引入文献热度累积量和文献热流量的概念。

文献热度累积量表示通过所研究年份的所有被引文献的数量,数值计算上即从该研究年份开始的所有被引文献数量之和,用 Q 表示,则 N_0 年发表的某篇文献在第 N_i 年的文献

热度累积量 Q_i 用公式表示为 $Q_i = \sum_{t=N_0}^{i=N_i} T_t$ 。文献热度累积量

Q_i 的意义是每一年的 Q_i 都会随着时间的推移而增加,例如,研究范围从 2006 年至 2014 年的 Q_{2006} 大于研究范围从 2006 年至 2010 年的 Q_{2006} ,很明显,这是热量累加的结果;而同一研究范围内,每年文献热度累积量 Q_i 会随着时间轴向后推移而逐渐减少,例如, $Q_{2006} > Q_{2007} > Q_{2008} > \dots$,文献从发表之时起,发出的热量总是小于内热源产生的热量,这也符合能量守恒定律,即:传递的能量+损耗的能量=总能量。

与文献热度累积量不同,文献热流量是一个有向向量,表示从文献发表年份传递到文献被引用年份的热量的流量大小和方向,用 \vec{Q} 表示,则 N_0 年发表的某篇文献在第 N_i 年的文献热度累积量 \vec{Q}_i 用公式表示为 $\vec{Q}_i = (N_i - N_0 + 1) \times T_i$,文献热度累积量 \vec{Q}_i 与传递的时间、传递的热量成正比。文献热度累积量 \vec{Q}_i 的意义是文献传递至该年份的热量强度,体现了热源影响力的大小。

借鉴传热学理论中的热扩散系数,引入文献热扩散系数的概念,用 α 表示。文献热扩散系数 α 表征的是文献的整体扩散能力, N_0 年发表的某篇文献在第 N_i 年的文献热扩散系数 α_i 用公式表示为 $\alpha_i = \frac{\vec{Q}_i}{Q_i} = \frac{(N_i - N_0 + 1) \times T_i}{\sum_{t=N_0}^{i=N_i} T_t}$ 。通过观察 α_i

随年份的变化,可以判断文献所处的生命时期以及文献半衰期的长短,并客观评价文献的生命力和总体扩散能力。

在研究主题或文献的分布空间后,可以按不同类型对其进行划分。

按照印证文献的重要性可以将文献的时空分布模型划分为若干层级。被引次数是不能准确衡量文献的影响力的,因为不同的文献影响力不同,所以被不同的文献引用的影响力是不同的,被高 H 指数的学者引用与被低 H 指数的学者引用的重要性是不同的,这就是所谓的引用认同问题^[22]。而基于传热学理论的文献时空分布与知识传播模型可以解决这一问题,即按照印证文献的重要性对模型进行划分,从而体现该文献能够影响到哪一层级的文献,以及研究不同层级之间及层级内的关系。由图 2 可以看出,不同的文献所处的热度层

是不同的,因此被不同热度层的文献引用的效果也是不同的,即同样是被引用了 10 次,但是被较高热度层引用次数多的点的文献热度累积量更高,从而解决了引用认同问题。

另外,可以按照研究领域的不同进行划分,把文献的时空分布和知识传播环境以领域为界限划分为块领域进行研究。整个球层被若干经线划分,两相邻经线之间为一个研究领域,编写 MATLAB 代码,绘制基于领域划分的文献时空分布与知识传播模型图,如图 3 所示,图中的 3 条斜线指示的 3 个领域分别是:文献计量领域、竞争情报领域、数据挖掘领域。每一个领域都被若干纬线划分,形成不同的热度层,通过该模型,可以看出不同领域的不同热度文献之间的关系,从而清晰描绘领域内和领域间的文献分布和知识传播,并展示出是哪些领域联系最为密切,领域之间哪个层次交流最为密切。

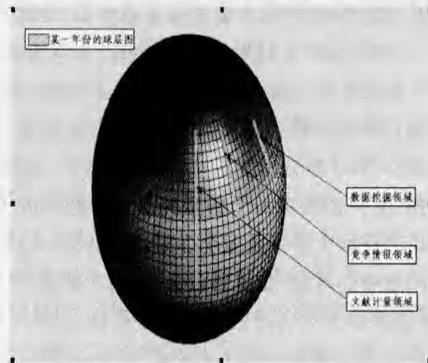


图 3 文献时空分布与知识传播模型图(简单的基于领域划分的热度分布关系)

按照文献的研究内容划分为概念理论、方法工具、实证研究 3 个方面进一步研究。在一个领域中,首先会有一些概念和理论被学者提出,随着研究的深入,有很多学者会开始方法工具和应用实证层面的研究,而与此同时,概念理论的研究也会随着方法的研究继续深入,并有新的理论概念被提出,就这样循环向前推进,一个研究领域慢慢形成,该研究领域的概念理论、方法工具、实证研究不断成熟。划分好科技文献的分布空间后,就可以按年份添加热度点,对每一年该领域进行概念理论、方法工具、实证研究三层次的划分,并把点放入所在的空间层,并且根据点的自身热度大小进行颜色的填充选择。当把各个年份的空间分布依次做好后,就建立了动态的基于研究内容的时间空间分布模型。

还可以按照显性知识的传播和隐性知识的传播把文献和知识的时空分布模型划分成不同的层次,并通过模型图展示出来。

通过该时空分布模型首先可以进行自身的纵向对比,可以比较得出自身研究领域随时间热度的变化以及研究层面的变化。另外还可以进行行业和行业之间的横向对比,通过比较得出行业间热度的大小,还可以对多个行业进行跨时间维度的对比,如图 4 所示。

该知识传播模型还可以比较领域间的知识传播能力,并研究比较知识传播趋势。研究领域间的知识传播能力,可以选择每个领域的期刊进行研究。总的传播能力用被引用期刊的种数与被引用次数的乘积表示,即被引用期刊的种数越多,被引用次数越高,传播能力越强。扩散系数以研究的第一年为基数,分别被每一年的总扩散能力除,得到基于年份的扩散系数变化图,从而可以研究扩散能力的变化趋势。

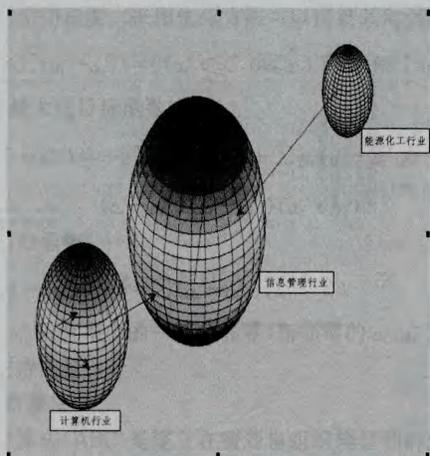


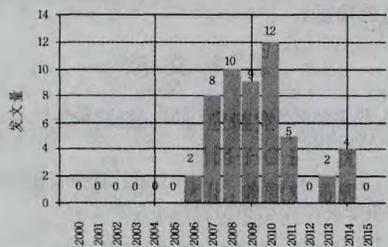
图4 文献时空分布与知识传播模型图(简单的跨行业的热度分布与知识传播关系)

该模型的优点就是比以前描述文献分布的模型更具综合性、直观性和可比较性,原来的文献分布模型只针对时间或空间一方面进行分析,而本模型可以进行跨时间维度的空间之间的比较分析以及同一空间不同空间层或不同空间同一空间层或不同空间的比较,并且可以定量计算层级之间扩散率的大小,从而进行不同方面的比较。模型也存在缺点和局限性,即原始的物理模型的温度分布是连续无离散分布,而新模型是离散点的分布,因此只能进行数值大小或普通的数值运算,而不能像原模型那样可以进行数值求导或积分等高级数值分析和计算。

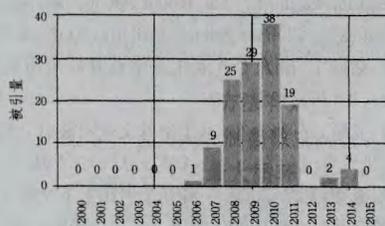
4 实验研究

本文利用文献时空分布与知识传播模型先对技术竞争情报领域进行简单的小领域实证对比研究。本文通过维普期刊资源整合服务平台进行文献印证追踪查询,其中使用“题名或关键词=技术竞争情报,时间:1989-2015”进行精确检索,共检出52篇文献。统计引证文献年代分布:2000-2015,引证文献量:87,被引次数:127,H指数=7。

检出文献年代分布和引证文献年代分布如图5所示。



(a) 所选文献年代分布图



(b) 引证文献年代分布图

图5

对以上数据应用传热学模型进行分析,首先绘制总的文献热量变化,如图6所示。

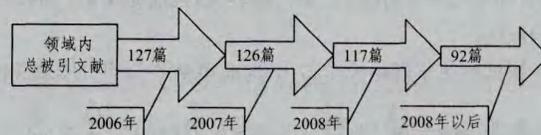


图6 技术竞争情报领域总的文献热量变化

首先选取了一篇领域内被引用次数达81的高被引论文进行研究,即以《技术竞争情报的现状分析》为例。该文献发表于2006年,从2006年至2014年被引用的数量如图7所示,使用热力学模型进行分析。

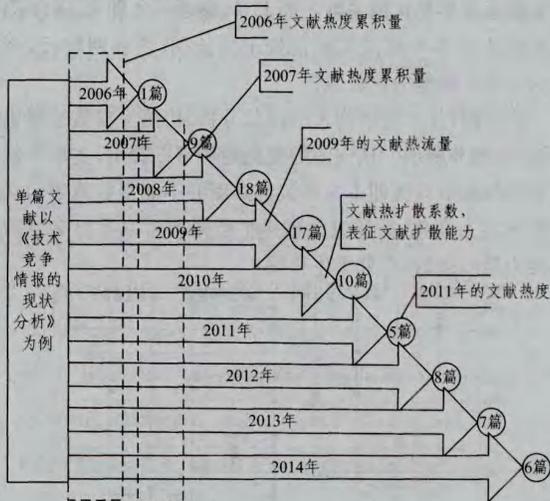


图7 《技术竞争情报的现状分析》模型参数解释

该文献在2006年的热度为 $T_{2006} = 1$,在2007年的热度为 $T_{2007} = 9$,2008年至2014年的数据见表1。通过研究每年的热度,可以得出文献热度变化的一般规律,即先增加后减少,这也是很好理解的。

表1 文献《技术竞争情报的现状分析》传热学参数

参数	2006	2007	2008	2009	2010	2011	2012	2013	2014
热度	1	9	18	17	10	5	8	7	6
文献热度累积量	81	80	71	53	36	26	21	13	6
文献热流量	1	18	54	68	50	30	56	56	54
文献热扩散系数	0.012	0.225	0.760	1.283	1.389	1.154	2.667	4.308	9

根据传热学理论模型得出2006年至2014年每年文献热度累积量: $Q_{2006} = \sum_{i=2006}^{i=2014} T_i = 81$, $Q_{2007} = \sum_{i=2007}^{i=2014} T_i = 80$,同理,2008年至2014年的数据见表1。

根据传热学理论模型得出2006年至2014年每年的文献热流量: $\bar{Q}_{2006} = (N_i - N_0 + 1) \times T_i = (2006 - 2006 + 1) \times 1 = 1$,同理,2007年至2014年数据见表1。通过研究每年的文献热流量可以看出,文献刚发表的前两年,由于距离热源较近,辐射较强,流动紊乱性强,文献热流量 \bar{Q} 的数值变化明显,然后从2008年至2014年,文献热流量 \bar{Q} 的数值趋于稳定,尤其是最近几年,数值几乎不变,说明热源在近几年有比较稳定均匀流量,随着时间的推移,文献热流量 \bar{Q} 的数值会逐渐衰减,但不排除中途有大幅波动的情况。如有大幅波动,说明内

热源热度有所变化,即可能出现距 N_0 若干年后理论的推演变化,如果几十年后 \bar{Q} 的数值仍然较大,说明该文献中的理论影响深远。

根据传热学理论模型得出 2006 年至 2014 年的文献热扩

$$\text{散系数为 } \alpha_{2006} = \frac{\bar{Q}_i}{Q_i} = \frac{\bar{Q}_{2006}}{Q_{2006}} = \frac{1}{81} \approx 0.01, \alpha_{2007} = \frac{\bar{Q}_i}{Q_i} = \frac{\bar{Q}_{2007}}{Q_{2007}} =$$

$$\frac{18}{80} \approx 0.23, \text{同理,2008 年至 2014 年的数据见表 1。通过研究}$$

该文献的热扩散系数可以看出,该文献第一年热扩散系数较低,比较符合常理,接下来几年的热扩散系数基本保持不变,虽然每篇文献随着时间的推移热扩散系数都会逐渐降低,但该文献近几年热扩散系数不断上升,最近一年即 2014 年热扩散系数达到 9,可见该文献正处于生长期,半衰期较长,该文献的总体扩散能力比较强。

将文献《技术竞争情报的现状分析》传热学参数绘制成趋势图,如图 8 所示。从该图的变化趋势可以看出,文献计量领域常用的被引频次即本文所说的热度并不能很好地表示文献的扩散能力变化,而通过文献热扩散系数可以看出文献的扩散能力是增强的,正处于生长期。

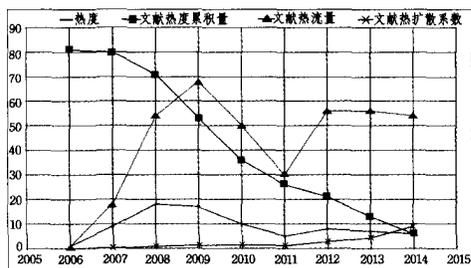


图 8 文献《技术竞争情报的现状分析》传热学参数变化

传统的文献指标评价存在的问题就是仅以数学模型或基于数学模型的改进模型来评价文献,描述知识传播和信息扩散,但这样带来的问题就是割裂了文献或知识的内在影响力。例如,一篇发表于 2000 年的文献在 2001 年至 2014 年中除了 2009 年均被引用 10 篇以上,而 2009 年仅被引用了 3 篇,那么,文献在 2009 年是否影响力很低呢?用原始模型会有一个很大的波谷式波动,但是应用本模型后的波动就不会很强烈,因为本模型应用了综合累计影响的思路。例如,一篇发表于 2000 年的文献,在 2005 年和 2050 年均被引用了 2 次,在本模型中,认为这两年的 2 次是不同的,2050 年的 2 次被引体现文献具有更大的影响力。

然后应用知识传播模型比较领域间的知识传播能力,并研究比较知识传播趋势。在众多学科领域中,管理学作为较为综合的学科,引入其他学科的知识较多,也对其他学科有较大影响,因此,研究管理学的传播扩散情况比较典型。由于最近几年的数据有不全的情况,因此选取 2003 年至 2007 年进行研究。本文使用张青学者论文中的数据,该论文统计的是管理学被各领域学科的期刊引用的情况^[23]。在此,仅宏观使用知识传播模型计算管理学向各领域传播的能力,通过图 9 可以看出,管理学向经济学传播的非常多,而向其他领域传播的较少;其次,管理学对图书情报学领域也有一定影响。通过图 10 可以看出,管理学向图书政治学领域传播的能力呈较强趋势,向法学领域传播的能力有所上升,而向经济学、教育学、图书情报学领域传播的能力基本呈现不变或略有降低的趋势。

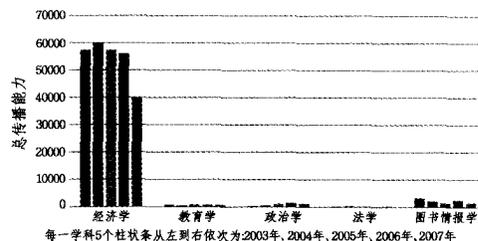


图 9 管理学在各领域间的传播情况比较

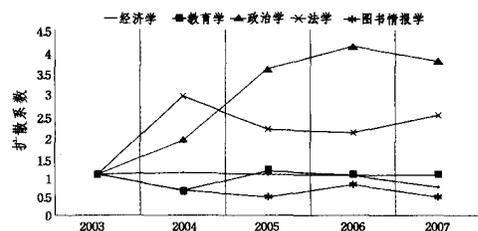


图 10 管理学在各领域内的传播情况比较

结束语 本研究主要是尝试用传热学中的理论研究方法来解释文献的分布和扩散规律,创新性地提出了文献热度、文献热度累积量、文献热流量、文献热扩散系数等概念,用传热学中热源扩散以及能量守恒等观点研究文献的分布、扩散能力以及生命周期的分析判断。

利用情报学领域外的方法来研究情报学问题,并通过这项研究为情报学领域提供一种新的定量研究方法,本研究尝试突破研究领域,并且是基于方法模型层面的研究,在适用性和宏观性上要高于工具层面的跨领域研究。基于传热学理论构建的模型的定量分析方法有助于清晰了解和比较研究主题或文献的发展情况。该模型可以基于时间序列定量预测未来的研究发展情况,可以为后续的学术研究提供指导和方向。该模型可以扩展升级为情报领域一个新的范式,对开拓科学哲学视野中的思维方式有一定帮助。但是,由于篇幅有限,本文不能进行详细复杂的跨行业、多领域、基于热度的综合多指数模型实证分析,详细的综合模型实证分析有待日后进一步研究。本文最大的意义在于提供了具有物理意义的可视化的文献分布综合表达与分析方式,并提供了一个看待文献引用与知识传播的角度。

参考文献

- [1] 庞景安. 科学计量研究方法论[M]. 北京: 科学技术文献出版社, 2002
- [2] 化柏林. 基于 TAIPO 模型探析文献计量分析研究的发展趋势[J]. 情报学报, 2010, 29(3): 479-485
- [3] Egghe L. Theory and practise of the g-index [J]. Scientometrics, 2006, 69(1): 131-152
- [4] Rousseau R, Egghe L. The R-and AR-indices; Complementing the h-index[J]. Chinese Science Bulletin, 2007(6): 855-863
- [5] 刘志辉, 郑彦宁. 研究专业演化图谱及其应用研究[J]. 情报学报, 2011, 30(11): 1178-1186
- [6] 吴凤慧, 成颖, 郑彦宁, 等. 基于学术文献同被引分析的 K-means 算法改进研究[J]. 情报学报, 2012, 31(1): 82-94
- [7] 吴凤慧, 成颖, 郑彦宁, 等. 一种基于引用上下文和引文网络的相关反馈算法[J]. 情报学报, 2012, 31(10): 1052-1061
- [8] 俞立平, 潘云涛, 武夷山, 等. 基于结构方程的学术期刊评价研究[J]. 情报学报, 2010, 29(1): 136-141
- [9] Shannon C E. The Mathematical Theory of Communication

[10] Wiener N. Cybernetics, or Control and Communication in the Animal and the Machine [M]. Cambridge, MA: MIT Press, 1948
 [11] 邱均平. 文献计量学[M]. 北京: 科技文献出版社, 1988
 [12] 吴国恩. 情报信息熵研究[J]. 情报理论与实践, 1993(1): 4
 [13] 吴国恩. 情报信息的测度[J]. 情报学刊, 1993(5): 343
 [14] 陈颖. 以信息论为基础的文献计量方法[J]. 河北大学学报(自然科学版), 1994(4): 106
 [15] 康凯, 苏建旭, 张会云. 技术创新扩散场——技术创新空间扩散研究的一种新方法[J]. 河北工业大学学报, 2000(2): 27-31
 [16] 邓忆瑞. 基于场论的信息扩散研究[J]. 情报杂志, 2008(8): 31-34
 [17] 邓忆瑞. 网络环境下信息扩散的模式研究[J]. 图书情报工作,

[18] 许智宏, 高静静, 胡浩. 蚁群算法在技术创新扩散模型中的应用研究[J]. 河北工业大学学报, 2012(1): 9-13
 [19] 杨世铭, 陶文铨. 传热学[M]. 北京: 高等教育出版社, 2006
 [20] Jakob M. Heat transfer [M]. New York: John Wiley & Sons, 1949
 [21] Eckert E R G, Goldstein R J. Measurements in heat transfer [M]. Washington D C: Hemisphere Publishing Corp, 1976
 [22] Howard D W. Authors as citers over time [J]. Journal of the American Society for Information Science and Technology, 2001, 52(2): 87-108
 [23] 张青. 我国管理学期刊引用网络分析[J]. 西南民族大学学报(人文社科版), 2010, 31(11): 274-276

(上接第 6 页)

所选评价准则是:

c_1 : 使用的需要; c_2 : 分析研究; c_3 : 新仪器和新技术的开发。

6 个专家对 5 个实验、3 个指标进行的判断排序矩阵如下:

$$A^1 = \begin{matrix} & c_1 & c_2 & c_3 \\ a_2 & \begin{bmatrix} 5 & 3 & 3 \end{bmatrix} \\ a_3 & \begin{bmatrix} 2 & 1 & 2 \end{bmatrix} \\ a_4 & \begin{bmatrix} 3 & 4 & 4 \end{bmatrix} \\ a_5 & \begin{bmatrix} 4 & 5 & 5 \end{bmatrix} \\ a_6 & \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \end{matrix} \quad A^2 = \begin{matrix} & c_1 & c_2 & c_3 \\ a_2 & \begin{bmatrix} 3 & 4 & 4 \end{bmatrix} \\ a_3 & \begin{bmatrix} 2 & 2 & 1 \end{bmatrix} \\ a_4 & \begin{bmatrix} 5 & 3 & 5 \end{bmatrix} \\ a_5 & \begin{bmatrix} 4 & 5 & 2 \end{bmatrix} \\ a_6 & \begin{bmatrix} 1 & 1 & 3 \end{bmatrix} \end{matrix}$$

$$A^3 = \begin{matrix} & c_1 & c_2 & c_3 \\ a_2 & \begin{bmatrix} 3 & 4 & 4 \end{bmatrix} \\ a_3 & \begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \\ a_4 & \begin{bmatrix} 5 & 3 & 5 \end{bmatrix} \\ a_5 & \begin{bmatrix} 4 & 5 & 1 \end{bmatrix} \\ a_6 & \begin{bmatrix} 2 & 2 & 3 \end{bmatrix} \end{matrix} \quad A^4 = \begin{matrix} & c_1 & c_2 & c_3 \\ a_2 & \begin{bmatrix} 4 & 1 & 3 \end{bmatrix} \\ a_3 & \begin{bmatrix} 2 & 3 & 1 \end{bmatrix} \\ a_4 & \begin{bmatrix} 5 & 4 & 5 \end{bmatrix} \\ a_5 & \begin{bmatrix} 3 & 2 & 4 \end{bmatrix} \\ a_6 & \begin{bmatrix} 1 & 5 & 2 \end{bmatrix} \end{matrix}$$

$$A^5 = \begin{matrix} & c_1 & c_2 & c_3 \\ a_2 & \begin{bmatrix} 4 & 4 & 4 \end{bmatrix} \\ a_3 & \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \\ a_4 & \begin{bmatrix} 5 & 5 & 5 \end{bmatrix} \\ a_5 & \begin{bmatrix} 3 & 3 & 2 \end{bmatrix} \\ a_6 & \begin{bmatrix} 2 & 1 & 3 \end{bmatrix} \end{matrix} \quad A^6 = \begin{matrix} & c_1 & c_2 & c_3 \\ a_2 & \begin{bmatrix} 1 & 5 & 5 \end{bmatrix} \\ a_3 & \begin{bmatrix} 3 & 1 & 2 \end{bmatrix} \\ a_4 & \begin{bmatrix} 5 & 4 & 4 \end{bmatrix} \\ a_5 & \begin{bmatrix} 4 & 3 & 3 \end{bmatrix} \\ a_6 & \begin{bmatrix} 2 & 2 & 1 \end{bmatrix} \end{matrix}$$

(1) 根据准则 c_1 可以得出 Borda 分矩阵 B_1' :

$$B_1' = \begin{matrix} & 1 & 2 & 3 & 4 & 5 & 6 \\ a_2 & \begin{bmatrix} 0 & 2 & 2 & 1 & 1 & 4 \end{bmatrix} \\ a_3 & \begin{bmatrix} 3 & 3 & 4 & 3 & 4 & 2 \end{bmatrix} \\ a_4 & \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ a_5 & \begin{bmatrix} 1 & 1 & 1 & 2 & 2 & 1 \end{bmatrix} \\ a_6 & \begin{bmatrix} 4 & 4 & 3 & 4 & 3 & 3 \end{bmatrix} \end{matrix}$$

(2) 根据算法修正 B_1' , 得到修正的 Borda 分 C_1' :

$$C_1' = \begin{matrix} & 1 & 2 & 3 & 4 & 5 & 6 \\ a_2 & \begin{bmatrix} 0 & 3.06 & 3.06 & 2.25 & 2 & 5.23 \end{bmatrix} \\ a_3 & \begin{bmatrix} 4.83 & 5.26 & 6.46 & 5.08 & 6.16 & 3.23 \end{bmatrix} \\ a_4 & \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ a_5 & \begin{bmatrix} 1 & 1.4 & 1.4 & 3.25 & 3 & 1.4 \end{bmatrix} \\ a_6 & \begin{bmatrix} 5.83 & 6.26 & 5.46 & 6.08 & 5.16 & 4.23 \end{bmatrix} \end{matrix}$$

同理可以得到 C_2' 和 C_3' 。

(3) 根据 $W=(0.28, 0.32, 0.4)$, 计算得出各方案的综合值:

$$G = [G_2, G_3, G_4, G_5, G_6] \\ = [12.37, 32.62, 3.96, 14.22, 28.89]$$

所以, 最终各方案之间的优序关系为 $a_3 > a_6 > a_5 > a_2 > a_4$ 。

将修正算法计算结果与模糊 Borda 法计算结果进行对比, 如表 1 所列。可以发现方案 a_2 和 a_5 的次序颠倒。需要说明的是, 模糊 Borda 法综合值一栏不等式中的前一个数表示方案作为当前排序的比重, 后一个数表示第二个可能排当前序的方案比重。所以模糊 Borda 中 a_2 和 a_5 的次序都是不够占绝对优势的, 特别是 a_5 的次序 (0.44), 这可能是模糊 Borda 法进行线性变换导致的。所以, 本文给出的结果应更为合理。

表 1 方法结果对比

次序	模糊 Borda 法		本文算法	
	方案	综合值	方案	综合值
1	a_3	$0.72 > 0.28$	a_3	32.62
2	a_6	$0.72 > 0.28$	a_6	28.89
3	a_2	$0.6 > 0.4$	a_5	14.22
4	a_5	$0.44 > 0.4$	a_2	12.37
5	a_4	$0.84 > 0.16$	a_4	3.96

结束语 本文在传统 Borda 法的基础上, 设计了一种 Borda 分修正算法。算法借鉴 SPAN 算法思想, 在修正成员关于方案的 Borda 分矩阵时, 参考其余成员的一致的评价意见; 修正后的 Borda 分矩阵能够反映出方案之间的优劣程度, 且方法具有一定的防操纵性; 采用加权和方法集结方案的综合值并排序择优, 通过算例将本文方法与模糊 Borda 法进行了对比, 凸显了本文算法的准确性与合理性。

参 考 文 献

[1] Banks J S. Sophisticated Voting Outcomes and Agenda Control [J]. Social Choice and Welfare, 1986, 1(4): 295-306
 [2] Arrow K J. Social choice and individual values (2nd edition) [M]. New York: Wiley, 1963
 [3] 杨季美, 史本山. 群体评价中的并和方法[J]. 系统工程理论与实践, 1992(1): 49-53
 [4] 苏为华, 陈骥. 模糊 Borda 法的缺陷分析及其改进思路[J]. 统计研究, 2007, 24(7): 58-64
 [5] 许晓东, 吴松, 路小刚. 简单多数票法和 Borda 法的防策略性分析[J]. 华中科技大学学报(自然科学版), 2005, 33(11): 86-89
 [6] 岳超源. 决策理论与方法[M]. 北京: 科学出版社, 2003