

# 基于专利共被引方法的研究前沿识别 ——以脑机接口领域为例\*

高楠, 傅俊英, 赵蕴华

(中国科学技术信息研究所, 北京 100038)

**摘要:** 探索基于专利共被引分析来识别领域研究前沿的方法。采用基于原始观测值和余弦距离两种相似度算法, 建立专利相似度矩阵, 利用社会网络分析, 获得研究前沿簇, 再进行簇类命名, 得到研究前沿。利用Innography数据库的脑机接口领域专利数据对此方法进行验证, 并对两种算法获得的结果进行比较, 发现相对于基于原始观测值的相似度算法, 余弦距离相似度算法能识别出更多的研究前沿, 且识别结果的内容更为丰富。

**关键词:** 研究前沿; 专利; 共被引分析; 脑机接口; 相似度算法

**中图分类号:** G255.53; R318.6

**DOI:** 10.3772/j.issn.1673-2286.2016.1.006

对科技信息实时追踪、识别领域研究前沿, 是科技情报工作的重要内容之一。目前文献计量学领域对于研究前沿的识别分析, 多以论文作为数据源, 较少涉及专利数据。世界知识产权组织公布的一组数据显示<sup>[1]</sup>, 专利说明书中含有90%~95%的研发成果, 其中85%以上的技术将不再出现于其他技术文献中, 且要比其他载体早公开1-2年。专利数据涵盖了更为广博丰富的技术信息, 适于领域的研究前沿分析。但专利数据具有其特殊性如专利家族, 在实际分析中要针对具体情况加以研究。

## 1 研究进展

1965年Price<sup>[2]</sup>首次基于文献计量提出研究前沿的概念及识别方法, 认为研究前沿由30~50篇最近发表的高被引文献组成, 涉及的方法是被引频次统计。1974年, Small和Griffith<sup>[3]</sup>首先提出共被引分析(Co-Citation Analysis, CCA)方法, 认为共被引聚类得到的联系紧密的簇就是研究前沿, 同时利用不同时间段内共被引网络的差异分析研究前沿的产生、发展和

衍化情况, 并预测未来走向, 进而得到研究前沿演进图。Garfield<sup>[4]</sup>对SCI数据库收录的高被引文献的前100,000篇进行共被引聚类, 提出研究前沿是由共被引文献簇和这些簇类的引文组成。同样采用共被引分析法, Gonzalex和Castro<sup>[5]</sup>识别了管理科学领域的研究前沿, 侯海燕<sup>[6]</sup>对科学计量学领域进行了研究前沿分析。另外, 有学者将其他分析方法和共被引分析结合, 对研究前沿进行识别。如崔雷等<sup>[7]</sup>按时间排序筛选出不同时段的高被引文献, 然后结合主题词聚类分析, 对支气管肺泡灌洗术领域的研究前沿进行识别和预测; 刘则渊等<sup>[8]</sup>则结合知识图谱分析对管理学理论前沿、国际创新管理领域前沿和战略管理学领域前沿进行了识别。共被引分析方法的缺点是具有一定的时滞性, 因为文献被其他文献引用需要时间。

与论文共被引相似, 专利的共被引是指两条专利同时被其他专利引用的现象, 反映了两条专利所代表的技术在理论或方法上的相似性或同源性<sup>[9]</sup>。但是, 专利共被引分析的时滞性问题会更突出, 因为专利从申请到公开有最多18个月的滞后。本文将探讨以专利为数据源, 对其做相应处理后, 识别研究前沿的方法。

\* 本研究得到中央级公益性科研院所基本科研业务费专项基金“未来学”(编号: XK2015-2)资助。

## 2 研究方法

### 2.1 数据源的选择

专利数据存在专利家族的特殊性,一个专利家族的同族专利是指拥有相同优先权专利,在不同国家或组织多次申请、多次公布或审核批准的一系列内容相同或相似的专利文献<sup>[10]</sup>,表征相同或相似的技术。故本文对专利家族,而非单个专利进行研究。ProQuest Dialog公司的Innography专利信息检索和分析平台有专利家族号(Family ID, FID)字段,同族专利的家族成员拥有相同的FID,代表一个技术点,因而本研究采用该数据库作为数据源进行研究前沿分析。数据库中的专利前向引用指标(Forward Citations)即专利的被引用情况,反映了专利的技术重要性和社会价值<sup>[11]</sup>,在本研究中用于共被引分析。

### 2.2 时间窗划分

时间窗的划分一般通过技术周期(Technology Cycle Time, TCT)设置或人为划分,但采用TCT设置滑窗宽度在实际操作中难度较大,对于大样本数据的可操作性不强;而人为划分主观性太强。专利计量学中已有的专利技术生命周期测度方法<sup>[12]</sup>,是以专利申请数量为纵坐标,专利权人数量为横坐标,绘制散点图得到的。专利权人数反映了研发机构或个人的数量,专利申请量反映了该领域的科技产出情况,数量越多则该领域的科技活动越频繁。通过观察二者之间的关系,可以初步判断某技术领域的技术成熟度,即该技术正处于萌芽阶段、发展阶段、成熟阶段或衰退阶段。本文创新性地研究前沿识别方法中利用领域的专利技术生命周期图来划分时间窗。

### 2.3 相似度矩阵建立

将同一FID所有家族成员的前引数据首先合并,使得每个FID仅出现一次。但是,在合并前引数据时,由于有两种对相同前引专利的计数方式,本研究采用两种相似度计算方法,进而衍生出两种共被引相似度矩阵生成方法。

第一种基于原始观测值(Observed Value, OV)的相似度计算方法,是在合并前引数据时,对于专利家族

各成员的相同前引专利全部归一,只计数1次,专利对间的相似度即二者相同前引专利的数量。该方法反映的是样本间绝对的共现次数,可以真实地反映共现情况,但无法反映集合的引用偏好,对于两个集合的相似度判断存在缺陷。

第二种基于余弦距离(Cosine Distance, CD)的相似度计算方法,是将专利家族各成员的前引专利叠加,转化成空间向量,然后计算两两向量间夹角的余弦值,得到余弦距离(也称为余弦相似度<sup>[13]</sup>)。该方法反映的是样本间相对的共现次数,可以保留集合的引用偏好情况,但缺点是无法还原真实的共现次数。

两种方法各有利弊,互为补充,因此本文采用两种相似度计算方法分别建立共被引相似度矩阵。相似度矩阵的行与列均为FID,根据专利前向引用指标和两种相似度算法分别计算专利家族对间的相似度,遍历所有专利家族对后,即可分别建立基于原始观测值的共被引相似度矩阵(OV-CCA)和基于余弦距离的共被引相似度矩阵(CD-CCA)。

### 2.4 研究前沿识别

利用UCINET社会网络分析法对引文相似度矩阵进行聚类,连续调节阈值直到所得簇类结果最多为止,该阈值即此时间窗的阈值,所得簇类即研究前沿。统计结果时,根据张嘉彬<sup>[14]</sup>的研究,将包含三个及以上的专利家族聚类算作一个研究前沿。最后,对识别出的研究前沿进行命名,依据其共同的联合专利分类号(Cooperative Patent Classification, CPC)以及专利内容,综合运用人工解读、切词技术和专家智慧对研究前沿簇进行命名。

对命名结果进行汇总时,对各个聚类簇中专利家族所含家族成员的平均优先权年和平均公布年均进行理论计算。前者反映前沿簇最早出现的平均年份,后者是公众可查询的平均年份。另外,还对聚类簇所含的专利家族和专利成员的数量和所占比例进行统计。因为在同一个时间窗内,可以认为数量最多且占比最大的簇类是该时间阶段内技术发展的重点。

## 3 脑机接口领域实证分析

脑科学领域的脑机接口技术(Brain Computer Interface, BCI)指在无外周神经系统和肌肉组织参与

的条件下,通过计算机等电子设备采集转化大脑活动信号为输出控制信号,进而与外界环境进行交流,涉及计算机科学、神经生理学和康复医学等领域<sup>[15-17]</sup>,被国际多个权威机构选为21世纪的研究前沿和热点之一<sup>[18]</sup>,通过科学计量方法识别脑机接口领域的研究前沿,并与文献调研结果比较,有重要的理论和实践意义。

### 3.1 数据源

在Innography专利信息检索和分析平台检索脑机接口技术领域相关专利,检索时间范围为1986年-2015年5月30日,得到6,374件专利,进行数据清洗剔除无关专利,最终得到相关专利6,243件。

### 3.2 基于技术生命周期划分时间窗

依据2.2节方法对下载数据的专利权人字段进行预处理,绘制脑机接口领域的专利技术生命周期曲线(见图1),从图中可以看出该领域能细分为三个发展阶段:

第一阶段:1986-1995年,萌芽阶段。该阶段由于技术市场还不明确,研发风险较大,只有少数几个机构参与相关的技术与市场开发,专利权人数量和专利申请数量均增长缓慢。

第二阶段:1996-2005年,小幅增长阶段。该阶段技术生命周期图上的点间距离较密集,说明专利研发者和专利数量都有增长,但增长幅度并不太大。

第三阶段:2006-2015年,大幅增长阶段。该阶段技术生命周期图上的点间距离明显加大。表明该技术有了较大突破,其隐含的经济价值开始显现出来,有较多的机构在这一期间加入该领域的研究,专利数量也增加明显。

统计各个时间窗内的专利数量和专利家族数量如表1所示。

表1 各时间窗内专利数量和专利家族数量

时间窗	优先权年	专利数量(件)	专利家族数量(个)
第一阶段	1986-1995	374	226
第二阶段	1996-2005	2323	900
第三阶段	2006-2015	3546	1848

### 3.3 基于原始观测值相似度算法

在按优先权年划分的1986-1995、1996-2005、

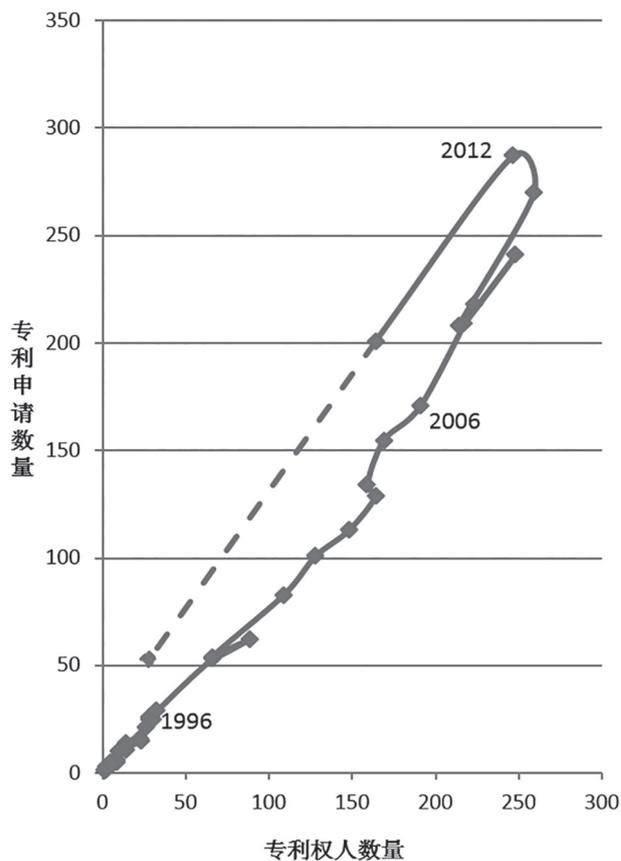


图1 脑机接口领域专利技术生命周期图

注:由于专利申请日和公开日之间存在18个月的时间间隔,因此近两年专利申请数据的统计不完整,不能真实反映脑机接口技术的发展趋势,图中用虚线表示。

2006-2015三个时间窗内,根据2.3节所述方法,利用VBA编程计算专利对间相似度,分别得到“226×226”、“900×900”和“1848×1848”的脑机接口领域的共被引相似度矩阵。识别各时间窗的聚类簇得到图2-图4所示结果,图中每个圆圈代表一个技术点,以专利家族号表示,线的粗细表示FID间联系的紧密程度。基于OV-CCA识别研究前沿,共得到三个时间窗内11个研究前沿簇,涵盖81个专利家族,共235件专利。其中2006-2015年间,聚类出的研究前沿簇最多,达到4个。最后,对研究前沿簇进行命名,得到表2中的命名结果。

#### (1) 1986-1995年间的研究前沿

调节观测值的阈值为23,最终得到1986-1995年间的共4个研究前沿簇,涵盖24个专利家族,共39件专利(见图2)。图中最右侧的研究前沿簇所含专利家族号最多,达到18个,占总有效聚类结果的75%,但进一步分析发现,该簇反映了两个研究前沿方向,分别是基于

诱发响应(如P300)和自发响应(如事件相关电位)的脑电图,以及通过神经元冲动控制电子游戏,以序号A1和A4分别表示。

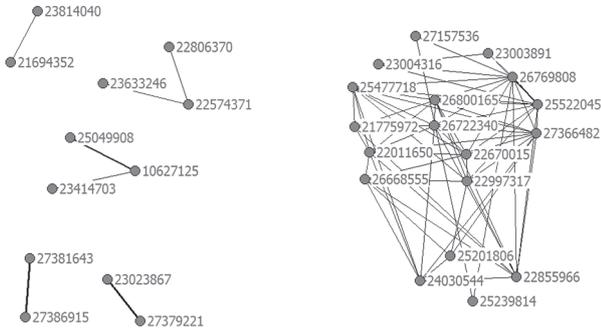


图2 基于OV-CCA识别1986-1995年间BCI研究前沿

(2) 1996-2005年间的研究前沿

调节观测值的阈值为60,最终得到1996-2005年间的共3个研究前沿簇,涵盖18个专利家族,共58件专利(见图3)。该聚类结果以研究前沿A6所占的比例最大,

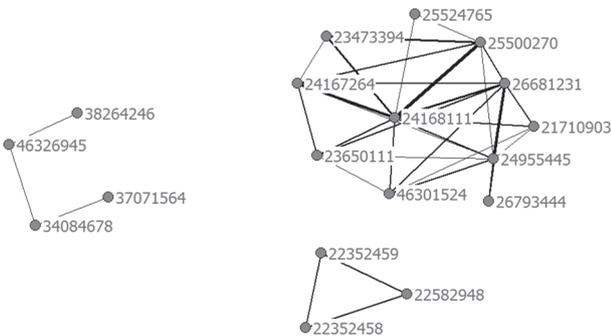


图3 基于OV-CCA识别1996-2005年间BCI研究前沿

达到61.1%,反映了这个时间窗内技术发展重点是评估神经系统的检测、测量或记录,属于脑机接口的信号采集部分。

(3) 2006-2015年间的研究前沿

调节观测值的阈值为10,最终得到2006-2015年间的共4个研究前沿簇,涵盖39个专利家族,共138件专利(见图4)。图中下侧的聚类簇所含专利家族号最多,达到33个,进一步可细化为A8利用诱发响应对脑电信号分类、定量评价精神状态和A11利用植入式神经刺激器治疗癫痫等神经系统疾病两个簇,反映了近十年脑机接口的研究从信号采集,深入到进行信号处理和利用脑机接口辅助诊断和治疗。国际脑机接口会议每三年举行一次,其所展示的研究前沿内容具备权威性,且最能反映近十年,乃至未来一到三年的重要发展方向。2013年举行的第五届国际脑机接口会议也提到与本文识别的研究前沿相同的BCI未来发展方向<sup>[19]</sup>:相比于非植入式BCI,植入式BCI展现出其对大脑活动信号控制更好的质量和速度,而成为新的研究聚焦点。另外,还可利用BCI进行意识检测和评估认知精神状态。

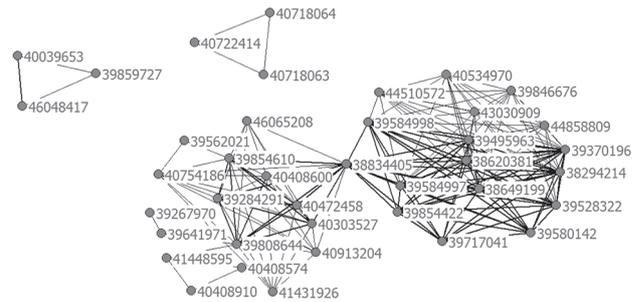


图4 基于OV-CCA识别2006-2015年间BCI研究前沿

表2 基于OV-CCA的研究前沿识别结果

时间窗	阈值	序号	研究前沿	聚类中专利		数量		所占比例%	
				平均 优先 权年	平均 公布 年	专利 家族	专利	专利 家族	专利
1986- 1995	23	A1	基于诱发响应(如P300)、自发响应(如事件相关电位)的脑电图	1986	1989	14	23	58.3	59
		A2	基于脑电信号的癫痫发作检测、预警系统	1989	1990	3	4	12.5	10.3
		A3	监测、记录大脑活动结合计算机分析,辅助医疗诊断或治疗,如计算机化处方或递送药物或饮食、医疗器械的计算机局部控制、医学专家系统或远程医疗	1986	1990	3	5	12.5	12.8
		A4	通过神经元冲动控制电子游戏	1990	1994	4	7	16.7	17.9

续表

时间窗	阈值	序号	研究前沿	聚类中专利		数量		所占比例%	
				平均优先权年	平均公布年	专利家族	专利	专利家族	专利
1996-2005	60	A5	利用电极检测、测量或记录脑电信号, 如脑电图	1999	2001	3	7	16.7	12.1
		A6	为评估神经系统进行的检测、测量或记录, 如脑电图、智能型脑起搏器、体腔内无线电测压器、双向脑机接口	1999	2004	11	42	61.1	72.4
		A7	治疗认知或精神病学(如痴呆、阿尔茨海默氏症, 抑郁症)的植入式神经刺激器	1998	2009	4	9	22.2	15.5
2006-2015	10	A8	利用诱发响应对脑电信号进行采集、分类	2007	2010	17	68	43.6	49.3
		A9	脑电信号获取, 如基于心理作业的脑电信号、皮层脑电	2008	2011	3	8	7.7	5.8
		A10	由身体位置、姿势或运动等产生的生理信号作为控制信号进行神经系统调制, 如电调制、磁调制	2008	2011	3	8	7.7	5.8
		A11	利用植入式神经刺激器治疗癫痫等神经系统疾病	2007	2011	16	54	41	39.1

### 3.4 基于余弦距离相似度算法

在按优先权年划分的1986-1995、1996-2005、2006-2015三个时间窗内, 根据2.3节所述方法, 利用汤森路透数据分析(Thomson Data Analyze, TDA)工具计算专利对间余弦相似度, 分别得到“226×226”、“900×90”和“1848×1848”的脑机接口领域的共被引相似度矩阵。识别各时间窗的聚类簇得到图5-图7所示结果, 基于CD-CCA识别研究前沿, 共得到三个时间窗内20个研究前沿簇, 涵盖119个专利家族, 304件专利。其中2006-2015年, 聚类出的研究前沿簇最多, 达到9个。最后, 对研究前沿簇进行命名, 得到表3中的命名结果。

#### (1) 1986-1995年间的研究前沿

调节余弦距离阈值为0.27, 最终得到1986-1995年间的共3个研究前沿簇, 涵盖25个专利家族, 共39件专利(见图5)。该聚类结果中a1所占比例最大, 达到52%, 反映了该时间窗内技术发展重点是利用视觉诱发响应检测、测量或记录脑电信号, 属于脑机接口的信号采集部分, 且信号源属于视觉诱发电位。视觉诱发电位是最常见也是最易引发和记录的脑电信号, 包括瞬态视觉诱发电位和稳态视觉诱发电位, 其中又以后者应用最多。高信噪比, 使得即使在头皮也能记录到较强的信号<sup>[20]</sup>; 它属于人对外界视觉刺激(如闪光、荧光刺激和颜色交替等)的正常反应, 所以几乎不需要用户训练, 这都促

使了稳态视觉诱发电位在早期阶段大量应用。



图5 基于CD-CCA识别1986-1995年间BCI研究前沿

#### (2) 1996-2005年间的研究前沿

调节余弦距离阈值为0.54, 最终得到1996-2005年间的共8个研究前沿簇, 涵盖57个专利家族, 共204件专利(见图6)。图中研究前沿a5所占的比例最大, 达到31.6%, 反映了该时间窗内技术发展重点是诊断或监视神经系统的活动及异常病变, 属于利用脑机接口辅助诊断或治疗。从初步的信号采集逐渐深化到有关神经系统疾病的诊断和治疗, 尤以癫痫的诊断研发为最。

#### (3) 2006-2015年间的研究前沿

调节余弦距离阈值为0.53, 最终得到1996-2005年间的共9个研究前沿簇, 涵盖37个专利家族, 共61件专利(见图7)。该时间窗内聚类出许多小的团簇, 一定程度上反映了近十年脑机接口发展的多元化。其中以a19神经响应测量的市场化应用所占比例相对较高,

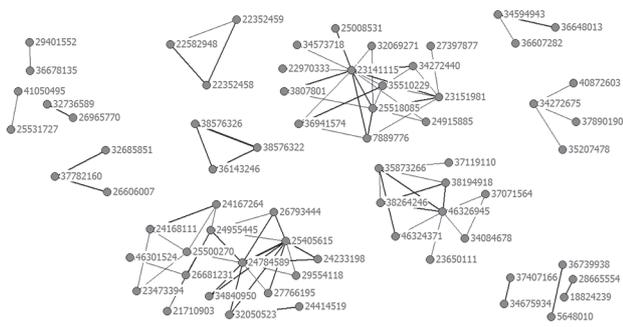


图6 基于CD-CCA识别1996-2005年间BCI研究前沿

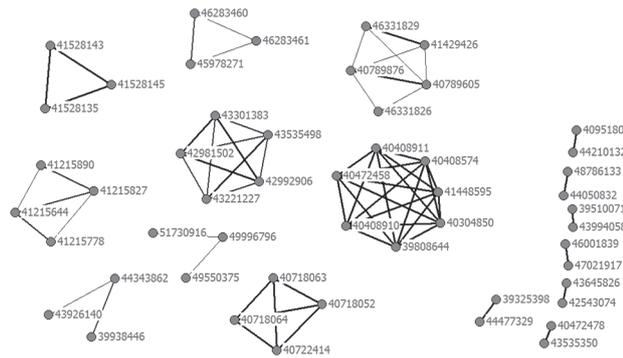


图7 基于CD-CCA识别2006-2015年间BCI研究前沿

达到18.9%。随着脑机接口的不断发展，市场化是必然的趋势，如神经刺激材料的研发等。另外，随着各种成像技术的发展使得作为脑机接口的输入信号源种类也

大为丰富，除了传统的脑电图，还有脑磁图、功能近红外光谱成像 (Functional Near-Infrared Spectroscopy, fNIRS)、功能性磁共振成像等脑部成像，甚至眼动电图、面部情感等都可以作为人机交互的输入端。这些对于脑机接口的市场化应用都起着重要的推动作用。第五届国际脑机接口会议上也提到了以上相同内容<sup>[19]</sup>：不断追求用户体验友好性、易用性，便于BCI的市场化推广。

另外，a6中还涉及基于近红外光谱的脑机接口 (fNIRS-BCI)，Sitaram等<sup>[21]</sup>于2007年才第一次提出fNIRS-BCI的相关研究，该技术对头动不敏感，设备轻便，收集信号过程无噪音<sup>[22]</sup>，是目前最新的脑机接口研究方向之一。这是OV-CCA识别结果中未涉及的重要的BCI研究前沿。a1和a7还提到了目前BCI技术最常用的信号处理方法，小波变换方法；在辅助诊断和治疗方面，除了与OV-CCA类似的利用植入式BCI，还涉及a3基于脑电传感数据的医疗保健设备和a5基于生物活性检测的医疗诊断系统，利用计算机辅助开具药物处方或药物输送。a7中提到的多模脑电图，在第五届脑机接口国际会议上也有重点强调<sup>[19]</sup>。由此可见，CD-CCA的识别结果除了在数量上占据明显优势外，在识别内容上也比OV-CCA所涉及的面更广、更丰富。

表3 基于CD-CCA的研究前沿识别结果

时间窗	阈值	序号	研究前沿	聚类中专利		数量		所占比例%	
				平均优先权年	平均公布年	专利家族	专利	专利家族	专利
1986-1995	0.27	a1	利用视觉诱发响应检测、测量或记录脑电信号，绘制脑电地形图	1985	1989	13	22	52	56.4
		a2	对脑电图记录的脑电数据进行统计学分析，如因子分析、聚类分析等	1986	1990	3	4	12	10.3
		a3	利用诱发电位的脑电信号检测、测量或记录，如稳态视觉诱发电位、P300	1991	1995	9	13	36	33.3
1996-2005	0.54	a4	利用电极检测、测量或记录脑电信号，如脑电图	1999	2001	3	7	5.3	3.4
		a5	诊断或监视神经系统的活动（如脑电波）及异常病变（如癫痫）	2000	2004	18	73	31.6	35.8
		a6	依据大脑活动进行测谎、脑指纹识别等	2001	2005	14	55	24.5	27
		a7	有关大脑活动的电流或磁场信号的检测、测量及记录装置，如脑电图、脑磁图	2003	2006	3	6	5.3	2.9
		a8	神经系统活动监测与输入设备，用于控制病人行动、康复训练	2004	2007	3	13	5.3	6.4
		a9	含有杂环化合物的药物制剂	2004	2007	4	24	7	11.7
		a10	治疗认知或精神病学（如痴呆、阿尔茨海默氏症、抑郁症）的植入式神经刺激器	2001	2008	9	21	15.7	10.3

续表

时间窗	阈值	序号	研究前沿	聚类中专利		数量		所占比例%	
				平均 优先 权年	平均 公布 年	专利 家族	专利	专利 家族	专利
1996-2005	0.54	a11	用于心肺系统与神经系统信号监测的植入式神经刺激器	2005	2009	3	5	5.3	2.5
2006-2015	0.53	a12	视觉诱发响应的脑电信号处理方法, 如小波变换方法	2008	2009	5	6	13.5	9.8
		a13	神经系统活动检测与输入设备, 例如脑电波 (EEG) 检测	2008	2010	3	3	8.1	4.9
		a14	基于脑电传感数据的医疗保健设备	2009	2010	5	5	13.5	8.2
		a15	用于治疗植入式神经刺激器	2008	2010	4	10	10.8	16.4
		a16	基于生物活性检测的医疗诊断系统, 利用计算机辅助开具药物处方或药物输送	2008	2011	4	7	10.8	11.5
		a17	与视觉刺激相关的神经系统活动检测与分析设备, 如视觉刺激生成器、解码视觉诱发电位的装置、基于近红外光的视觉驱动脑机接口	2009	2012	3	6	8.1	9.8
		a18	基于小波分析的术中运动区功能定位系统的多模脑电图	2012	2012	3	3	8.1	4.9
		a19	将神经响应测量市场化, 如神经刺激材料研发、脑电图 (EEG)、皮肤电反应 (GSR)、心电图 (ECG)、眼动电图 (EOG)、眼跟踪和面部情感编码测量等	2007	2012	7	18	18.9	29.6
		a20	用于测量脑部活动的光学传感器	2012	2015	3	3	8.1	4.9

### 3.5 结果对比

汇总基于OV-CCA和基于CD-CCA的研究前沿识别结果, 得到表4。从表中可以看出单个时间窗内, 无论是前沿的数量还是前沿所含的FID总数, CD-CCA能识别出的数量更多。同时, 由3.4节可知CD-CCA能识别出OV-CCA识别不出的前沿簇, 且识别结果所涵盖的内容范围更广, 具有更强的研究前沿识别能力。因而, 采用基于余弦距离相似度法更适合于建立共被引相似度矩阵及用于研究前沿的识别。

## 4 结论

本文采用共被引分析法, 利用专利数据源, 分析了

脑机接口领域近30年的研究前沿演变情况。对于共被引相似度矩阵的建立采用了基于原始观测值和基于余弦距离的两种相似度算法, 并基于研究前沿识别结果进行了对比分析。

共被引分析法是当前最常用的识别研究前沿的方法之一, 而共被引相似度矩阵的建立是识别研究前沿最关键的一步, 之后便可通过聚类识别出所需前沿结果。选择两种相似度算法建立引文相似度矩阵, 是因为专利数据的特殊性。专利具有专利家族的概念, 一个专利家族对应一个技术点, 因而相比于单篇专利更适合于研究前沿中技术点的识别。但两个专利家族间相似度的计算, 是两类专利文献集间相似度的计算, 在合并同一FID所有家族成员的前引数据时, 对相同前引数据的计数有两种方式, 因此, 必然导致两种相似度计算

表4 基于OV-CCA和CD-CCA所得研究前沿汇总

阈值	阈值		前沿个数 (个)		前沿所包含的FID总数 (个)	
	OV-CCA	CD-CCA	OV-CCA	CD-CCA	OV-CCA	CD-CCA
1986-1995	23	0.27	4	3	24	25
1996-2005	60	0.54	3	8	18	57
2006-2015	10	0.53	4	9	37	39

方式的使用。两种相似度算法各有利弊,但通过比较发现,这两种方法都可以识别出脑机接口领域的研究前沿发展情况;但是,基于余弦距离算法的前沿识别能力更强,所得前沿数量更多,且包含的技术点更全面。

因为专利的重要引用关系除了共被引,还有耦合关系。下一步将在本文共被引方法研究以及已开展的耦合方法研究的基础上,进一步探索将共被引和耦合方法进行融合,构建更完整的信息网络,进行研究前沿识别的方法研究,并比较该融合方法与单个引文信息分析识别研究前沿方法的差异,以期对研究前沿识别这一关键问题获得较为全面、科学的解答。

## 参考文献

- [1] 张强.基于专利计量的专利实施许可实证研究[D].重庆:西南政法大学,2012.
- [2] Price D.J. Networks of scientific papers[J]. Science,1965,149(3683):510-515.
- [3] Small H G, Griffith B C. The structure of scientific literatures I: Identifying and graphing specialties[J]. Science Studies,1974(4):17-40.
- [4] Garfield E. Research fronts[J]. Current Contents,1994,41(10):3-7.
- [5] Gonzalez F J, Castro B C. Dominant approaches in the field of management[J]. Organizational Analysis,2001,9(4):327-353.
- [6] 侯海燕.科学计量学知识图谱[M].大连:大连理工大学出版社,2008.
- [7] 崔雷,沈锡宾,方丽,等.内科学主要分支学科研究前沿的文献计量学分析[J].中华内科杂志,2013,52(2):144-150.
- [8] 刘则渊,王贤文.生态经济学研究前沿及其演进的可视化分析[J].西南林学院学报,2008,28(4):3-11.
- [9] 李睿,张玲玲,郭世月,等.专利同被引聚类与专利引用耦合聚类的对比分析[J].图书情报工作,2012,56(8):91-95.
- [10] 吴琳,魏星,霍翠婷,等.基于Web的专利双语语料自动获取研究及实现——以esp@cenet数据库为例[J].现代图书情报技术,2009(9):57-63.
- [11] 经济合作与发展组织.专利统计手册[M].高昌林.北京:科学技术文献出版社,2013:3.
- [12] 牟萍.专利情报检索与分析[M].北京:知识产权出版社,2010:141.
- [13] 刘妍.基于Lucene的余弦距离检测文档相似度方法的研究[J].信息系统工程,2014,(4):129-130,142.
- [14] 张嘉彬.以书目耦合及共被引探讨不同引用区间之研究前沿:以OLED领域为例[D].台北:国立台湾大学,2011:20-22.
- [15] Boyu W, Chi M, Feng W. Comparison of Different Classification Methods for EEG-Based Brain Computer Interfaces: A Case Study[C]//2009 IEEE International Conference on Information and Automation.
- [16] 杨建,刘进,吴明曦,等.脑机技术发展及其对军事领域的影响[J].国防科技,2013,34(6):19-23.
- [17] 尧德中,刘铁军,雷旭,等.基于脑电的脑-机接口:关键技术和应用前景[J].电子科技大学学报,2009,38(5):550-554.
- [18] Mason S G, Bashashati A, Fatuorechi M, et al. A comprehensive survey of brain interface technology designs[J]. Ann Biomed Eng,2007,35(2):137-169.
- [19] Huggins J E, Guger C, Allison B, et al. Workshops of the fifth international brain-computer interface meeting: defining the future[J]. Brain-Computer Interfaces,2014,1(1):27-49.
- [20] 程明,任宇鹏,高小榕,等.脑电信号控制康复机器人的关键技术[J].机器人技术与应用,2003(4):45-48.
- [21] Sitaram R, Zhang H, Guan C, et al. Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface[J]. Neuroimage,2007,34(4):1416-1427.
- [22] 尧德中.脑机接口:从神奇到现实转变[J].中国生物医学工程学报,2014,33(6):641-643.

## 作者简介

高楠,女,1992年生,硕士研究生,研究方向:情报学分析,E-mail:gaonan2013@istic.ac.cn.

傅俊英,女,1972年生,博士,研究员,研究方向:科技情报研究、生物技术研究,通讯作者,E-mail:fujunying@istic.ac.cn.

## Recognition of Research Fronts Based on Patent Co-Citation Analysis in the field of Brain-Computer Interface

GAO Nan, FU JunYing, ZHAO YunHua  
(Institute of Scientific and Technical Information of China, 100038 Beijing, China)

Abstract: Patent Co-Citation method issued to identify research front (RF) in a field in this paper. Based on two similarity algorithms - observed value and cosine distance, two kinds of patent similarity matrixes are then established, social network analysis is applied to get RF clusters which are then named, and we finally get the research fronts. Brain-computer interface (BCI) is selected to perform empirical analysis in terms of the above method, and results from two similarity algorithms are also compared. This study finds that RF can be got by Co-Citation method, while cosine distance algorithm can reveal more and detailed research fronts than observed value algorithm.

Keywords: Research Front; Patent; Co-Citation Analysis; Brain Computer Interface; Similarity Algorithm

(收稿日期: 2015-10-10)