

国家科学图书馆青年人才领域前沿项目

社会网络分析在专利权人网络分析中的
方法与应用研究

项目负责人：陈云伟
课题组成员：王春，平昭，梁田

2011年6月

目 录

1	研究目标.....	4
2	国内外发展现状调研.....	4
2.1	社会网络分析的概念和特征.....	4
2.1.1	社会网络分析的定义.....	4
2.1.2	社会网络分析的特征.....	5
2.1.3	社会网络的一些基本概念（图论）.....	5
	(1) 点、边和密度.....	5
	(2) 中心度和中心势.....	6
	(3) 成分、核与派系.....	6
2.1.4	社会网络的一些基本概念（矩阵）.....	8
	(1) 邻接矩阵（adjacency matrix）.....	8
	(2) 关联矩阵（incidence matrix）.....	8
	(3) 距离矩阵（distance matrix）.....	8
	(4) 数值矩阵（valued matrix）.....	8
	(5) 单位矩阵（identify matrix）.....	8
2.2	社会网络分析在专利分析中的应用现状.....	8
2.2.1	社会网络分析在专利引文分析中的应用.....	9
	(1) 专利引用与社会网络分析的关系 [□]	9
	(2) 社会网络分析在专利引文分析中的应用.....	9
2.2.2	社会网络分析在专利合作研究中的应用.....	10
2.2.3	社会网络分析在专利技术主题关联分析中的应用.....	11
2.2.4	专利社会网络分析中的科学地图可视化作图方法.....	11
	(1) 专利社会网络分析可视化“工作流”.....	12
	(2) 专利社会网络分析可视化图形的主要类型.....	12
2.2.5	研究意义与发展前景.....	12
3	研究方法与实践路线.....	13
3.1	研究方法.....	13
3.2	解决的关键问题.....	14
3.3	实施路线图.....	14
4	社会网络分析在专利权人网络分析中的应用方法.....	15
4.1	基于合作的专利权人合作网络分析方法.....	17
4.1.1	基本原理.....	17
4.1.2	分析方法研究与内容.....	18
	(1) 全球规模的专利权人合作网络分析方法.....	18
	(2) 特定学科领域（技术主题）的专利权人合作网络分析方法.....	19
	(3) 特定专利权人的合作网络分析方法.....	23
	(4) 自我中心的合作网络分析方法.....	26
4.2	基于引用的专利权人引用网络分析方法.....	28
4.2.1	基本原理.....	28
4.2.2	分析方法与内容.....	29
	(1) 专利权人直接引用网络.....	29

(2)	专利权人共引网络.....	31
(3)	基于专利文献耦合专利权人网络.....	31
4.2.3	功能与作用.....	31
4.3	基于技术主题的专利权人网络分析方法.....	32
4.3.1	基本原理.....	32
4.3.2	分析方法设计与内容.....	33
(1)	技术主题的设计依据.....	33
(2)	数据集的构建.....	33
(3)	分析指标.....	33
(4)	分析实例.....	33
4.3.3	功能与作用.....	38
5	专利权人网络分析案例研究-以中国科学院为例.....	38
5.1	中科院专利权人ego合作网络.....	38
5.1.1	中科院发明专利申请数量增长趋势.....	38
5.1.2	合作者的增长.....	39
5.1.3	合作网络的稠化与增长.....	40
5.1.4	网络直径.....	40
5.2	中科院专利权人global合作网络.....	40
5.2.1	研究所与合作者的增长.....	40
5.2.2	Global合作网络的稠化与增长.....	41
5.2.3	网络直径.....	42
5.2.4	中科院global合作网络的特征.....	42
6	讨论.....	44
7	参考文献.....	44

1 研究目标

本课题研究思想源于社会网络分析在社会学研究中的广泛应用及研究成果,在研究方法上借鉴近几年来图书情报学领域对社会网络分析方法的引用实践,特别是在基于论文的社会网络分析方法实践经验的基础上,在专利分析中进行社会网络分析方法的研究与实践。课题将研究社会网络分析在专利权人网络分析中的方法,并对中国科学院专利权人合作网络进行案例应用分析,挖掘中国科学院专利权人合作网络的属性,网络的演变特征,在 1985-2009 年的时间里,随和中国科学院研发活动的开展,特别是知识创新工程实施以来,合作行为的涌现及演化特征,在网络增长过程中合作行为的产生是否具有优先情节?合作网络是否呈现小世界现象?专利权人合作是否符合幂律分布和无标度特征?研究影响合作行为的因素包括哪些?如地域因素?发现的合作网络特征能为提升中国科学院专利创新活动带来哪些值得参考的信息?希望研究成果能对我馆专利情报研究工作者更加深入地开展专利分析工作有所帮助。

2 国内外发展现状调研

2.1 社会网络分析的概念和特征

2.1.1 社会网络分析的定义

社会网络分析(Social Network Analysis, SNA)是对社会关系结构及其属性加以分析的一套规范和方法,它主要分析的是不同社会单位(个体、群体或社会)所构成的关系的结构及其属性^[1]。

对社会网络分析的定义,需要突出以下四点:第一,社会网络分析首要的是研究节点之间的相互关系,其次才关注个体的特性,然而,在理解社会现象时,个体特征与相互关系都非常重要;第二,社会网络分析研究的一个重要方面是网络结构如何影响个体的行为;第三,社会网络分析的两种主要类型是Ego Network 和Global Network,前者研究个体为中心的网络,后者研究网络中所有参与者之间的相互关系;第四,社会网络分析是一跨学科领域,数学和计算科学是促进其发展的两个重要学科^[2]。

社会网络分析为研究人们的相互关系提供了可视化的和数学的分析。通过社会关系网络图及社会关系矩阵两种形式化描述方式,可以评估网络中节点的位置,发现节点的中心性,以及网络的中心与边缘,对网络中各种节点和簇进行深入的分析,发现簇在哪里,簇中包含哪些节点,谁是连接节点,谁是领导者,谁又是孤立节点等信息。社会网络分析法是在人类学、心理学、社会学、数学以及统计学等领域中发展起来的,已经经历了 70 多年的历史,与图论和复杂网络理论一起称为网络研究三部曲,如果说结构-功能主义统治了 20 世纪 40—60 年代的欧美社会学界,那么从 20 世纪 70 年代初期至今占据着欧美社会学特别是美国社会学主流地位的则是社会网络分析,其领军人物包括:伯特(Ronald Burt)、格兰诺维特(Mark Granovetter)、诺科(David Knoke)、马斯登(Peter Marsden)、维尔曼(Barry Wellman)、怀特(Harrison white)等学者^[3]。至今,社会网络分析法已经形成了一系列专有术语和概念,被广泛应用于社会学研究中,成为社会科学研究的一种新的范式。

社会网络分析近十几年的重要进展包括：二方关系、三方关系、一般化的块模型、动态网络模型、P*模型、社会影响模型、社会选择模型、动态分析等。

如何解释社会网络研究的结果？刘军等^[4]认为，有时候仅仅从社会网络的角度给出的解释也是不够的。在数学领域中有一个著名的定理——哥德尔不完全性定理。该定理告诉我们，任何公设系统都不是完备的，其中必然存在着既不能被肯定也不能被否定的命题，任何完备系统都不能仅仅利用自身的定理来解释。在社会网络分析中又何尝不是如此呢！在一定意义上可以说，“网络只是管道”，管道的结构固然重要，但是管道内流通的内容也是我们不能忽视的。因此社会网络研究的结果需要结合抽象的“文化”、“地方性知识”等来研究。只有如此，我们才能理解、解释社会网络现象。

2.1.2 社会网络分析的特征

社会网络分析又称结构分析，或者说，社会网络分析不仅是对关系或结构加以分析的一套技术，还是一种理论方法——结构分析思想。社会网络分析关注的是行动者之间的关系，而非行动者的属性。社会网络分析家巴里·韦尔曼指出，“网络分析探究的是深层结构——隐藏在复杂的社会系统表面之下的一定的网络模式。”网络分析者特别关注特定网络中的关联模式如何通过提供不同的机会或限制，从而影响到人们的行动。行动者之间相互连接而形成的关系是社会网络分析的基础，这也是社会网络区别于其他分析方法的重要特征。

当代社会网络分析中都可以发现四点特性，这四点特性也共同定义了这个研究领域^[5]：

- 社会网络分析源自于联系社会行动者的关系基础之上的结构性思想；
- 它以系统的经验数据为基础；
- 它非常重视关系图形的绘制；
- 它依赖于数学或计算模型的使用。

2.1.3 社会网络的一些基本概念（图论）

社会网络分析借用图论和矩阵法等表现社会关系及其结构，已发展出了自己的一套概念体系。为便于对社会网络分析并不熟悉的读者阅读本报告，下面对社会网络的一些基本概念进行介绍。

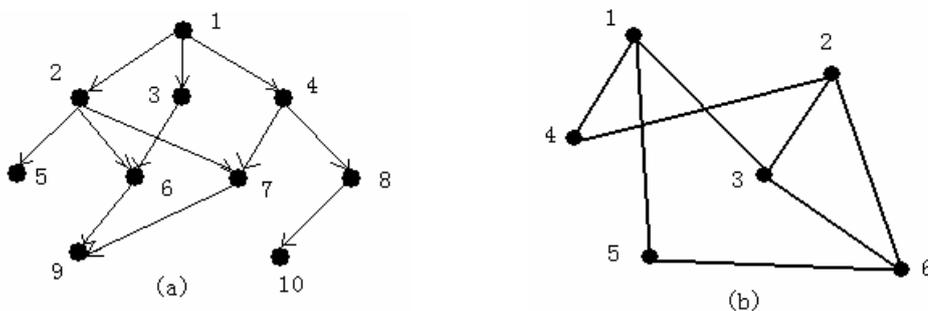


图1. 网络拓扑结构示意图

说明：(a) 是 10 阶有向图；(b) 是 6 阶无向图

(1) 点、边和密度

点 (Node) ——也称“行动者 (Actor)”，社会网络中的一切个体、社会实体或事件都称为行动者，每个行动者在网络中的位置被称为“点”。

边 (edge) ——也称“关系纽带 (relation tie)”，行动者之间相互的关联。

度数 (degree) ——与某特定点相邻的那些点称为该点的邻域 (neighborhood), 邻域中的总点数, 称为度数。

线路 (walk) ——各个点可以通过一条直线连接, 也可以通过一系列线间接连接, 这一系列线叫做一条线路。

途径 (path) ——如果线路中的每个点和每条线都不同, 则称该线路为途径。

长度 (length) ——一个途径的长度是构成该途径的边的条数来测量。

距离 (distance) ——指连接两点之间的最短途径的长度。

内含度 (inclusiveness) ——图中关联的点数与总点数的比值。

密度 (density) ——图中实际拥有的连线数与最多可能拥有的连线数的比值。

(2) 中心度和中心势

局部中心度 (local centrality) ——以度数为基础的对点中心度的测量。即与某点直接相连的点数, 称为该点的局部中心度, 也可以测量距离为 1 和 2 或更大的点数。对局部中心度进行测量的时候, 不涉及整个网络是否有独一无二的“核心”点这样的问题, 认识到这一点很重要。局部中心度的局限性在于, 中心度数仅仅在同一个图的成员之间或者在同等规模的图之间进行比较才有意义。

相对中心度——点的实际度数与可能联络的最多度数之比。

整体中心度 (global centrality) ——如果一个点与其他许多点的距离都很短, 则称该点是整体中心点, 在图中, 这样的点与许多其他点都“接近”。即该点与其他各点的距离之和, 和越小, 整体中心度越大。

中间度 (betweenness) ——一个度数相对比较低的点可能起到重要的“中介”作用, 因而处于网络的中心。在针对点中心度的各种测度中, 中间度可能最难计算。

结构洞 (structure holes) ——当两个点以距离 2 相连的时候, 就说二者之间存在一个结构洞。结构洞的存在使第三者扮演经纪人或中间人的角色。

中心势和图的中心——密度和中心势 (centralization) 这两个概念代表的是一个图的总体“紧凑性” (compactness) 的不同方面。密度指的是一个图的凝聚力的总体水平; “中心势”描述的则是这种内聚性能够在多大程度上围绕某些特定点组织起来。二者是两个重要的、相互补充的量度。中心势测度的一般性程序: 找出最核心点的中心度和其他点的中心度之差。中心势就是实际的差值总和与最大的差值总和之比。一个图的结构中心 (structural center) 是单个点或者一个点集, 它像圆心和球心一样, 是图结构的枢纽。

(3) 成分、核与派系

试图发现网络的各个“派系”以及网络可以分成多少个凝聚子群 (cohesive sub-groups), 这是投身于社会网络分析的学者们持续关注的问题之一。各种关于子群的理论模型相继涌现, 如把子群描述成为“派系 (cliques)”、“聚类 (clusters)”、“成分 (components)”、“核 (cores)”、“圈子 (circles)”。

成分——是在各种子图概念中最简单的, 它的正式定义: “最大关联的子图”。在子图中所有的点都通过一条线或多条线相互连接, 但是它们与子图外的点无关联。成分是一个点集, 这些点通过连续的关系链连在一起。原则上讲, 一个成分的各个成员都可以相互交往, 无论这种交往是直接的还是通过一条中介链, 孤立点则无此机会。因此, 在一个图中发现的各个成分的模式——其数目和规模——可以看成是标志了其成员进行交往的机会和限制性, 或者表征了相应的网络中的资源流动。因此, 就此来说, 他们体现了早期的场域理论中的“拓扑式区域 (topological regions)”背后的观念。因此, 对一个网络进行结构描述的第一步便是确定其成分的数目和规模。

环成分 (cyclic components) ——环成分的概念依赖循环一词, 一个循环就是一个途径, 只不过它返回到其初始点, 并且与一个途径类似, 其长度也是可以是任意的。一个图中的循

环可以用其长度来描述。

桥线 (bridge) ——它本身不处于环中，但是却把两个或者更多的环连在一起。

悬挂点 (hangers) ——这是一些与环成分相连之点，但他们本身不在环上，这些点仅仅“挂”在环成分上。

桥点——这些点处在两个或者多个环成分之间，起到中介者 (intermediaries) 或者摆动者 (wavers) 的作用，但它们不是任意环成分的成员。这样，一个桥点挂在两个或者多个环成分上。

孤立树 (isolated trees) ——这是一些点链 (包括二人链 dyads)，它们与任何环成分都无关联。这些树的成员之间以非环 (non-cyclic) 的方式相连。

孤立点——度数为 0，与任何点都不相连的点。

有时候很难给长的关联途径以有实质意义的社会学解释，当长环把众多点连在一起的时候，问题就更加明显。一般考察 3-环或 4-环就可以了。如果研究者希望运用长度超过 4 的环，那么他赋予这种数学结构的社会学解释必须既清楚又有意义。

对有向图来说，也可以分析其环成分，最简单的做法就是不考虑线的方向。这种分析的基础是“半环 semi-cycles”，在这些环中都不考虑线的方向，当然这样做会失去一些信息，但是可以得到弱环成分 (weak cyclic components)。为了分析强环成分 (strong cyclic components)，方向性信息必须保留。还有一种称为可接受的半环。一个有向图中的所有环都可以分为有向环、可接受的环和不可接受的环三类。

在悬点 (hangers-on) ——指的是这样的一些悬点，即连着它的线指向强环成分的某一成员。未悬点 (hangers-off) ——指的是这样一些悬点，即存在一条从成分的一个成员出发并指向该点的线。

切割点 (cut-point) ——指的是这样的点，即如果去掉它，就会增加成分的数目，没有它，子图就会分成两个或多个独立的子群，并且各个子群之间无关联。可以认为切割点代表的行动者处于局部中心地位。

成分的轮廓——一种被称为“嵌套 (nesting)”的程序就可以揭示成分的轮廓。检验成分的各种算法把所有的关系都看成是二元数据，即简单地看关系的有无，对一个多值图的分析必须把真实值转换为二元值，即 1 和 0。人们提出两类相互替代的嵌套方法：一类是根据点的度数作为测量凝聚力的标准，另一类是根据线的多元性作为测量紧密度的标准。以度数为基础 (degree-based) 的测量确定的是“K 核 (k-cores)”，基于多元性的测量确定的是“m-核 (m-cores)”。

Seidman 对成分结构的研究可以运用最小度标准，以便区分高、低凝聚力的领域。他论证到，对一个图的“k-核”结构分析是对密度测度的一个重要补充。一个 k-核是一个最大子图，其中的每个点都至少与其他 k 个点连接：k-核中的每个点的度数都至少为 k。这样，一个简单的成分就是一个“1k-核”，其中所有点都相连，因而其度数至少为 1。为了确定“2k-核”，需要忽略所有度数为 1 的点，进而考察剩余各点之间的关联结构，2k-核由那些度数为 2 的剩余关联点组成。同理，确定一个 3k-核要去掉度数为 2 和 1 的点，以此类推。可以推出，一个 k-核便是在整个图中的一个凝聚力相对较高的区域，但是它不一定是最大的凝聚子图，因为有可能存在一些相互之间联系松散，却有很高凝聚力的区域。

核塌缩序列 (core collapse sequence) ——一个 k-核中的点可以分为 2 个集合：在 K+1 核中的点和不在该核中的点。在任何核中，剩余集合都是由那些当 k 增加 1 后在分析时将会消失的点组成。

核塌缩序列汇总了成分的密集性 (clumpiness)。Seidman 认为，如果核塌缩是缓慢的、逐渐的，则表示网络结构在总体上具有一致性。否则一个不规则的取值序列则说明存在着相对来说比较紧密的区域，该区域被比较多的边缘点包围着。

派系——派系的本质含义看成是“最大的完全子图 (maximal complete sub-graph)”。一个派系就是一个点的子集 (sub-set)，其中任何一对点都有一条线直接相连，并且该派系不被其他任何派系所包含。一个派系是一个最大的完全关联图。每个点都与其他点直接相连。最大完全子图这个概念对于社会网络来说是相当严格的，这种紧密联系的群体在实际中很少见。因此学者对此概念进行了推广。最早的是 n -派系 (n -clique)， n 指的是派系成员之间联络的最长途径之长。该思想的缺陷：1) 当 n 大于 2 时，很难对这样的松散派系进行说明；2) n -派系中的一些途径终节点并不是派系的成员。这里有一个派系的直径的概念——最长距离之长度。

K -从——一个 K -从就是满足如下条件的一系列点，即其中每一个点都与除了 k 个点之外的其他点直接相连。

2.1.4 社会网络的一些基本概念 (矩阵)

把社会网络中的每一个节点分别按行和列的方式排列即可形成网络矩阵，网络的节点数称为矩阵的阶数。社群图有不同的类型，与之相对应，其矩阵表达形式也有不同的类型，社会网络分析常用的矩阵有邻接矩阵、关系矩阵等。

(1) 邻接矩阵 (adjacency matrix)

邻接矩阵表示的是有 n 个点的矩阵中的元素从点 i 到点 j 的关系的条数，又被称为社群矩阵 (sociamatrix)。例如，有 n 个点的图 G 的邻接矩阵 $A = (a_{ij})$ 是一个 $n \times n$ 矩阵，其中若 n_i 邻接 n_j ，那么 $a_{ij} = 1$ ，否则 $a_{ij} = 0$ 。

若关系无方向，则成为对称矩阵。对称矩阵表示行动者之间关系的有无 (1,0)，但不能表示其关系强弱。若关系有方向是，则矩阵数值的分布就不对称了。

(2) 关联矩阵 (incidence matrix)

关联矩阵是关于无圈图 G 的点与线的矩阵，记作 $B = (b_{ij})$ 。具体来说，对于一个图，指的是一个 0, 1 矩阵，其中 i, j 元素取 1 当且仅当点 i 与 j 关联。

(3) 距离矩阵 (distance matrix)

距离矩阵是根据节点间的距离形成的矩阵，距离矩阵是对称矩阵。

(4) 数值矩阵 (valued matrix)

根据数值图可形成数值矩阵，其中的数值代表某种关联性 or 关联程度。

(5) 单位矩阵 (identify matrix)

单位矩阵，简称单矩阵，是指主对角线上的元素全是 1，其余元素全为 0 的正方阵，记作 I 。

2.2 社会网络分析在专利分析中的应用现状

作为一种非常有用的方法，社会网络分析法早已突破了社会学领域的范围，为其他领域的学者所采用。情报学领域也不例外，经过近年的发展，主要应用在合著网络与引文网络分析领域，并取得了一些成果。

专利是技术创新和科学技术发明的产物。国外自 20 世纪 60 年代末就开始研究专利产出对科学技术进步与经济发展的影响，创建了相应的专利评价指标体系，一些著名机构定期出版有用于对国家、区域、机构的技术创新能力评价的报告。如 OECD、ipIQ 公司等，在国家、区域层面上对技术创新力进行评价。

将社会网络分析方法引入专利分析,目前此方面的研究国内外均较少,且大多停留在基于专利引文分析的层面上。据相关研究报道,专利网络与论文网络存在不同的特征^[6],下面对国内外主要的研究成果进行简要的总结与分析。

2.2.1 社会网络分析在专利引文分析中的应用

(1) 专利引用与社会网络分析的关系^[7]

专利引用分析是一种重要的专利管理方法,并已在许多领域得到广泛应用。通过专利引用分析,可以对企业决策和技术研发、运用提供一种客观的评价方法,并根据竞争者的专利部署,分析自身在市场中的地位。如通过将每篇专利的引证情况作连线式的链接,即多级引用分析,可以得到专利引用图。通过图中专利引用的线索,可以找到最早被引证的专利文献,该文献很可能是一篇基础专利,包含重要的技术信息。此外,根据专利引用图的线索,还可以得到专利的承继性和发展历程,以帮助科研人员把握科技创新脉络,了解前沿技术,避免选题重复和人、财、物的浪费,推动科技进步。专利引用分析主要基于两个方面:引用分析和被引分析。引用分析,即对某专利在研发过程中引用的其他专利情况的分析。被引分析,即对某专利被审查或之后的其他专利所引用情况的分析。一方面,通过对专利引用次数的统计,分析各主体专利的经济或技术输入情况,寻找其技术基础;另一方面,通过对专利被引次数的统计,则可了解其知识溢出的情况,并可以在一定程度上反映在后专利的发明者对在先专利价值的评价。

与个人社会网络关系相似,各专利在相互引用中也具有社会网络关系。社会网络中的个体,即专利引用网络中引用或被引用的各件专利;联系即在专利引用网络中各专利自引和他引的关系;连接即专利之间相互引用形成的连接,包括引用连接和被引连接。从微观层次看,专利引用的社会网络关系是某技术领域或某组织内部各专利相互引用的关系。这些关系对技术的发展和资源配置效率具有极大的影响,随着各种技术的变迁,这些网络关系也会随之发生改变,从而判断各专利的质量和重要性。从宏观层次看,专利引用的社会网络关系是各组织、各技术之间的专利引用关系,从而形成各技术领域之间或地理区域之间的信息流。相对于微观层次的社会关系,宏观层次的社会网络关系对技术的发展、组织的壮大具有更大的影响力。各组织可通过这种宏观的社会网络关系,发现各交叉技术中的技术空白点,研发具有高价值的技术,或选择具有相同目标、相似背景的组织形成相对稳定的网络组织。

(2) 社会网络分析在专利引文分析中的应用

在专利引用网络中,使用向心度分析是为了衡量某专利在整个网络中的作用或影响,试图从引用网络中区分最重要的专利,尤其是可利用被引次数来衡量专利的重要性。

在专利引用网络中,不仅可以通过专利之间的引用关系,找出关键的重要技术,而且可以通过技术的发展脉络、引用网络的结构,判别竞争对手的专利竞争地位,将其划分为技术潜在竞争者、技术领导者、技术积极者和技术落后者,并结合组织自身的情况实行不同的战略,或进行合作开发,或进行兼并等。某机构或某专利为了保持其在行业领域或技术领域中的优势,则必须掌握处在结构洞中的关键技术或成为结构洞中关键的竞争对手,保持信息优势和控制优势。它们需关注于与其他独立或未联系的组织建立联系,开发结构洞,以便于将两个组织的成员连接在一起。如果节点拥有了结构洞的位置,它们可以较容易争取信息和控制优势,从其他组织中获利。

专利引用的社会网络分析为挖掘信息流的发展方向、发现技术主题发展历程中的核心专利以及在技术演进中处于关键过渡地位的专利、发明人或机构等,提供了强有力的手段。例如,德国汉堡大学的Wartburg等^[8]在2005年通过多级专利引用分析(multi-stage patent citation

analysis) 方法对技术演进进行了研究 (图 2-a); 北京大学的Gao等^[9]在 2009 年采用社会网络分析方法分析了美国专利数据库中的中国专利与科技期刊的文献耦合网络; 美国Fannie Mae公司的Gress^[10]在 2007 年通过 1963-2002 美国专利的引文分析研究了随着时间维度的专利引用网络, 探讨了信息流的发展方向, 图 2-b展示了基于 1975 年的 6 件专利的引用网络, 图 2-c展示的是美国专利 3858382 被引网络立体图。德国Sternitzke等^[11]对专利权人之间的引用关系进行可视化研究, 图 2-d展示的是ROHM公司专利的引用网络图。

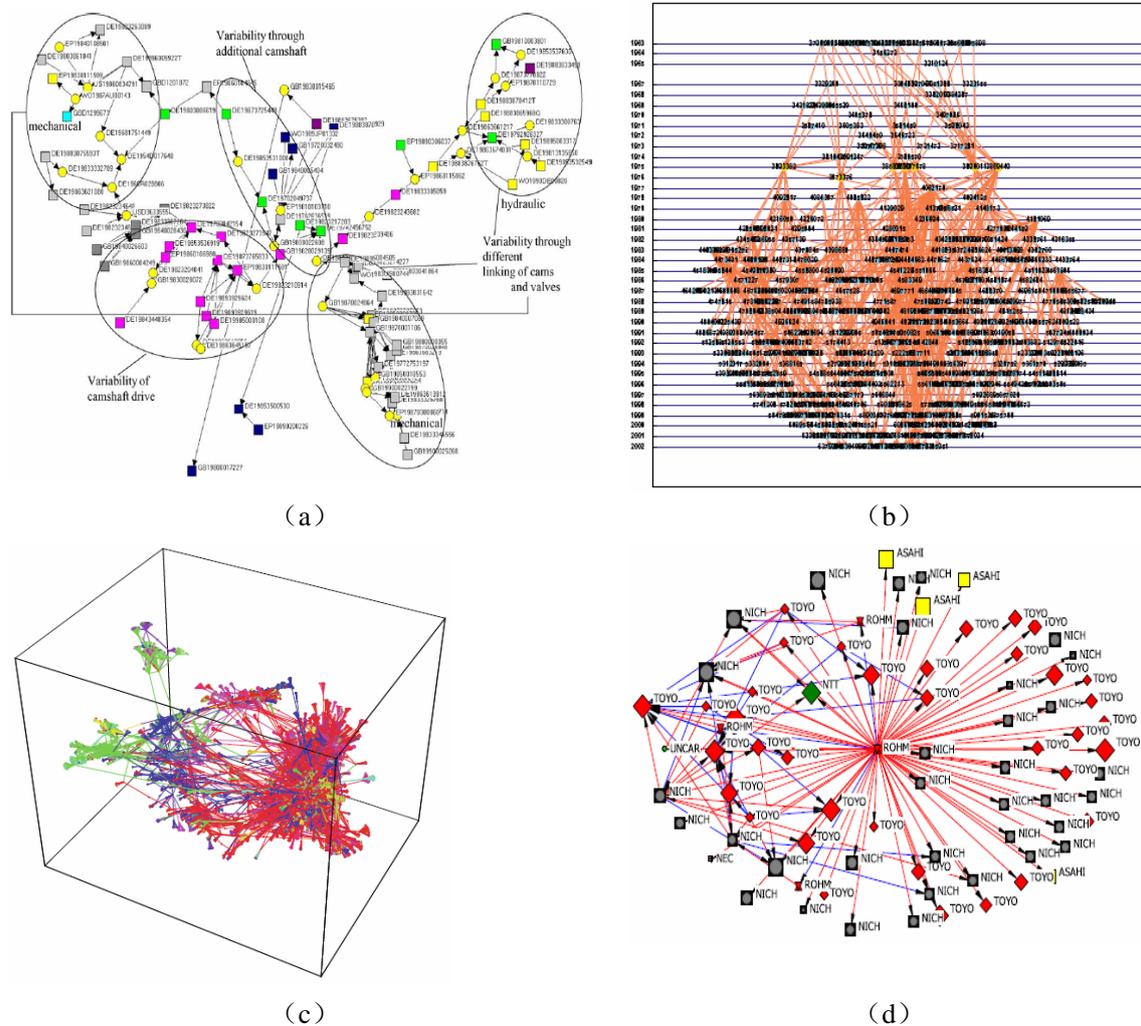


图2. 专利引用的社会网络分析图

2.2.2 社会网络分析在专利合作研究中的应用

引用网络分析仅仅是社会网络分析在专利分析中应用的一个方向, 有的学者把引用网络与社会网络看作是并列的两个概念。除了专利引用网络, 可以以国家、专利权人、发明人为节点, 从专利合作的角度开展社会网络分析。

专利合作网络可视化主要分为国家、专利权人、发明人三个层面, 主要用于发现他们之间的相互合作强度以及竞争合作关系。通过合作网络研究可以发现国家、机构的竞争合作关系, 挖掘研发团队成员之间关系和团队规模。例如, Dou等^[12]在 2007 年基于欧洲专利研究禽流感相关研究的研发战略和国家比较分析时, 采用国家和专利权人的合作网络可视化对他们的竞争合作关系进行研究, 图 3-a对禽流感相关专利的国家合作网络进行可视化分析。图

3-b是Sternitzke等^[13]在2008年做的专利权人的合作网络研究。

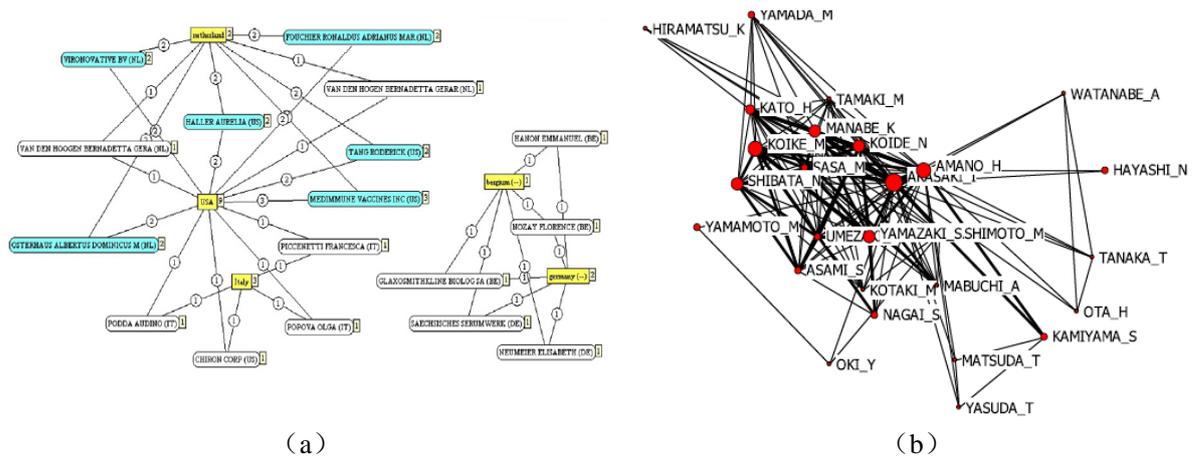


图3. 专利合作网络

2.2.3 社会网络分析在专利技术主题关联分析中的应用

专利技术主题信息可视化的研究可以从主题词和专利分类等角度进行研究。例如，图4-a是禽流感相关欧洲专利IPC相互关系图，反映各种相关IPC小类之间的关联强度。图4-b反映了德国马普学会发明专利申请数量最多的前二十个IPC小类的关联网络。

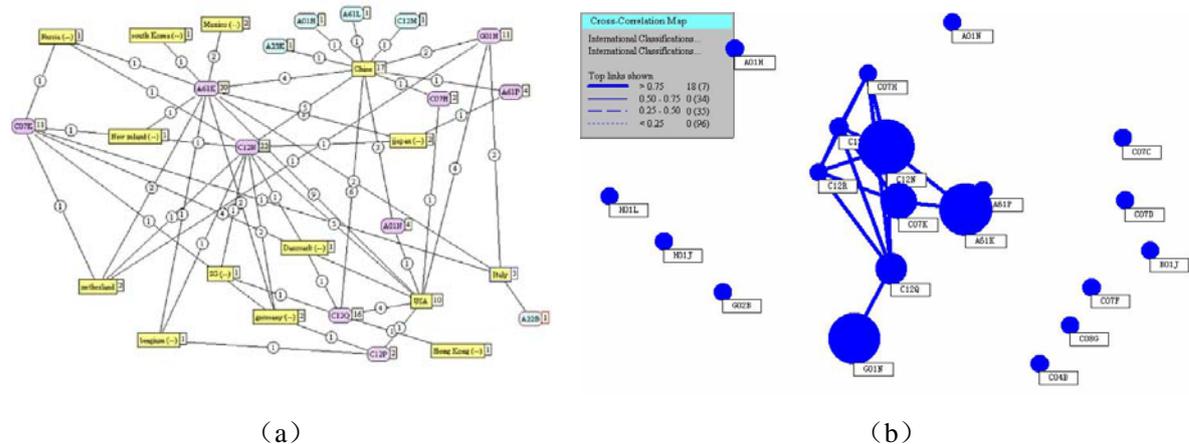


图4. 专利技术主题关联网络

2.2.4 专利社会网络分析中的科学地图可视化作图方法

专利情报可视化研究目前尤其值得关注的内容包括，通过基于发明人、专利权人合作的网络可视化分析，分析个人和机构在创新过程中所处的位置以及发挥的价值；通过领域和主题的网络可视化分析揭示不同机构的竞争或合作关系；通过引文网络可视化挖掘信息流的发展方向、发现技术主题发展历程中的核心专利以及在技术演进中处于关键过渡地位的专利、发明人或机构；还应重视可视化图谱在确定研发重点、制定专利研发战略等方面所应发挥的作用。在专利可视化技术的选择方面，应针对不同的需求，根据各种技术之间的表现力差异选择最适当的可视化技术。此外，专利数据与论文数据在数据结构上存在差异，如其特有的专利家族现象、转让与授权行为等都将成专利可视化分析的对象，如何面向这些特殊对象生成能反映真实相互关系的可视化图形、避免生成误导性的图形，这将是值得深入研究的内

容。

(1) 专利社会网络分析可视化“ workflow ”

专利情报可视化研究与文献计量等其他信息可视化技术相同,在许多情况下需要多个步骤,历经一系列的过程最后生成可视化图形,这个过程被称为 workflow。通常的 workflow 是专利数据经导入、抽样和模型化后,再经预处理、分析后实现可视化。一般的工作流包括 6 个步骤^[14]:专利数据准备、分析单元的定义、方法选择、各单元之间相似性的计算、各单元的关联与定位、对专利可视化结果的分析与说明,其中“各单元之间相似性的计算”与“各单元的关联与定位”两步又可同时定义为专利数据布局。若再广义地加以分类,专利情报可视化的整个过程又可以分为两大部分,第一部分是专利数据的收集、整理、分析与管理,此部分是专利情报可视化研究的基础,将决定所构建的可视化图形的科学性与准确性。此步骤的关键技术包括:专利引文分析、因子分析、聚类分析、多维尺度分析等。这些技术同时也是科学计量学、文献计量学所面临的最重要的难点之一。针对更多个性化的、复杂化的专利情报可视化研究工作,则需要根据具体的功能设计个性化的可视化方案,通过计算机编程实现。需要指出的是,随着该学科的不断发展和完善,工具也在不断的发展和完善。

(2) 专利社会网络分析可视化图形的主要类型

根据不同的分类标准,专利情报可视化可分为不同的类型,如基于介质可以将可视化图形分为时间图、地理图、人物图和主题图等;根据反映的客体,还可以将其分为微观(个人)、中观(机构)和宏观(国家)三个层面。根据陈云伟等^[15]对科学地图的分类,目前有关专利社会网络分析的地图多属于领域地图的范畴,包括专利引用网络地图、专利合作网络地图和专利技术主题地图。

2.2.5 研究意义与发展前景

可以预见,将社会网络分析方法应用于专利分析中具有重要意义,2008 年,德国学者 Christian Sternitzke 等就指出,社会网络分析方法才刚刚开始进入专利分析领域,应用前景广阔^[16]。社会网络分析在专利分析中的应用是对社会网络分析方法学的有益拓展,可以预见,将社会网络分析方法应用于专利分析中具有重要意义,未来专利社会网络分析的研究重点将至少包括以下几个方面:

首先,用于鉴别关键发明人或专利权人、发掘竞争对手。通过专利合作、技术主题关联和引用网络分析,可以发现处于网络重要位置的专利权人或发明人,这些重要的节点可以是中心(度数中心或中间中心)节点,也可以是桥点。综合这些信息,可以描绘发明人或专利权人的竞争力,用于构建更有创新能力的研发团队。特别是对专利文献所承载的专利权的所有者——专利权人的分析目前还相对较少,这将是非常广阔的研究领域,通过专利权人之间的关联网络分析,可以用于发现竞争对手^[17]进行竞争分析,或为创建联合研发项目寻找合作伙伴提供参考。此外,高被引专利可以用来评价专利丛林,引文网络中关系近的节点说明其技术相关,若在引文网络中相连的两个节点在合作网络中没有合作,说明二者存在高度竞争关系,若有合作则说明二者在联合研发新技术。

其次,专利社会网络分析在企业创新与管理研究中得到越来越多的关注,包括企业之间的竞争合作以及企业内部发明人之间的合作网络结构对创新的影响等。例如 Paruchuri 等(2010)通过企业之间的专利合作网络以及企业内部的发明人合作网络分析发现,公司内部发明人创新合作网络的发明人结构中心性与其对公司创新性的作用呈现倒“U”型关联关系,这种关系同时又受到公司所处的合作网络的中心性及结构洞的负调节^[18]。此外, Su 等(2009)采用专利家族的优先权专利网络对公司的专利组合(patent portfolio)进行了分析,服务于

企业技术保护与实现战略目标^[19]。

第三，基于论文和专利之间的引用关系探讨科学向技术的转移仍将是一大研究热点。现有研究大多数尚停留在计量和统计水平，而缺少在网络结构和演进动力学水平考察科学向技术的知识转移的研究，这将是社会网络分析在专利分析中的重要潜在应用领域之一。

第四，更多的专利数据来源将作为专利社会网络分析的数据基础。现有有关专利网络分析工作大多数都是基于美国专利商标局（USPTO）的专利数据开展的，基于欧洲专利、日本专利、中国专利以及德温特专利的专利社会网络分析工作还比较少，属于待发掘领域。

最后，专利网络动力学及演进分析将是今后发展的重点之一，特别是专利引用网络以及专利权人合作网络动力学及演进分析将是今后的研究热点之一。近期已经发表的几篇文章中也正体现了这种趋势，包括Barberá等（2011）基于专利引用网络对技术演进（主路径）进行的研究^[20]，Lee等（2009）基于ICT领域专利数据开展的协同进化研究^[21]，Kegler等（2010）对机构合作网络演进的分析^[22]，等等。技术演进路径识别的算法研究与应用也将是专利演进分析的一个热点方向，如常见的搜寻路径连接数目（SPLC）算法和搜寻路径节点对数（SPNP）算法等，刘倩楠等（2010年）利用SPNP算法对以太网技术演进路径进行了研究^[23]。

目前，我馆在专利分析上已经有数年的积累，在国内处于领先行列，包括利用计量学的方法从专利量化角度探讨国家和区域技术创新力、核心专利评价、专利地图等方面做了大量研究工作。在分析工具上也引进了TDA和Aureka等商业化软件，并自主开发专利在线分析系统，然而目前的研究工作以及所拥有的工具中尚缺乏对社会网络分析在专利中应用的关注和功能设计。将社会网络分析方法应用于专利分析中，将进一步完善我馆在专利情报研究领域的方法体系，对现有的指标体系、专利地图、核心专利等分析是有益的结合和补充，同时也是对社会网络分析应用方法的良性扩展应用。

立足于我馆专利分析工作的现实需求，鉴于专利权人在专利中的重要性，由于目前国内外对专利权人网络分析较少，缺乏可以引入到实际工作中的值得借鉴的直接经验和方法，本课题将在专利权人分析中引入社会网络分析的理念，进行方法和应用研究，并为日后专利网络分析工作积累经验。在未来的工作中，将逐渐扩展到包括国家、发明人、引文、主题、专利家族、资助项目等不同方面的社会网络研究工作。

3 研究方法与实践路线

对不同的社会网络有不同的社会测量和分析技术，这些都要视研究对象、目的的不同而有所不同。但是无论它们有何不同，通过社会网络资料回答的基本问题应是相同的：哪些行动者属于网络的成员？他们之间具有哪些形式的关系？如何对他们进行具体测量分析？

3.1 研究方法

1. 通过文献调研分析和专家咨询法，借鉴当前国内外社会网络分析在论文分析以及专利分析中的应用方法，研究出专利权人网络分析方法。

2. 结合专家咨询，设计针对中国科学院专利权人合作网络分析方案，具体步骤包括：下载1985-2009年中国科学院发明专利；基于方法研究的结果设计关联方案，对专利权人进行清洗和加工等预处理；通过软件对数据进行可视化，生成网络关联图谱；对网络关联图谱进行比较分析与解读。

3. 通过软件的组合应用及数据格式转变，实现专利权人关联网络的可视化处理，主要的工具包括 TDA、NWB 等。

4. Derwent Innovations IndexSM数据库是本课题案例研究的主要数据来源。

3.2 解决的关键问题

(1) 适用于专利数据的社会网络分析方法的研究

专利数据有其自身的结构特征，目前针对专利社会网络分析的工作较少，缺少直接借鉴的参考资料。

(2) 数据的关联方案

数据集中项之间的关联方案是决定专利社会网络分析结果是否具有科学价值的重要因素，因此必需根据专利数据自身特点建立科学实用的关联方案，形成科学严谨的关系网络。

(3) 案例研究中的数据加工与处理

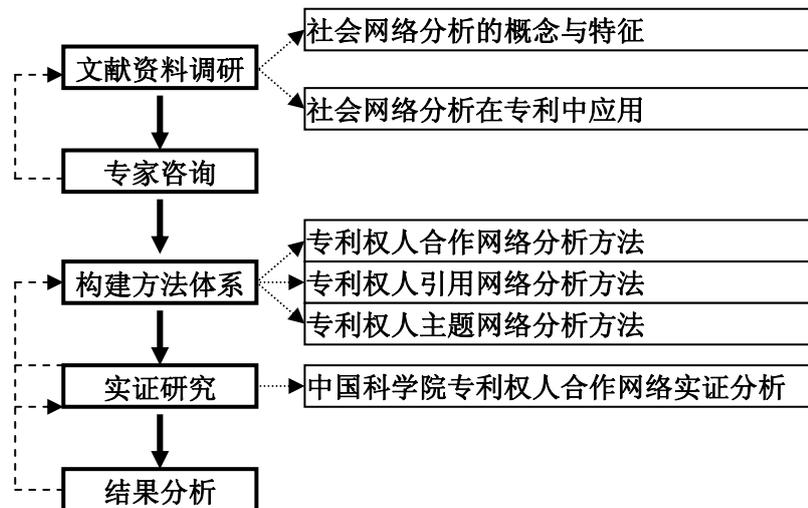
科学准确的数据是社会网络分析的基础，在网络关联图构建过程中数据的预处理占据重要的地位，如清洗噪音数据、数据集成和变换等。同时，在面对大量存在内在复杂联系的数据时，需要借助计算机分析工具，根据需要，将对数据格式进行加工或处理，或对现有软件进行功能改良，比如在现有软件的开放端口添加新的算法程序等。

(4) 专利权人合作网络演进图谱的实现

(5) 网络关联图谱的深入解析

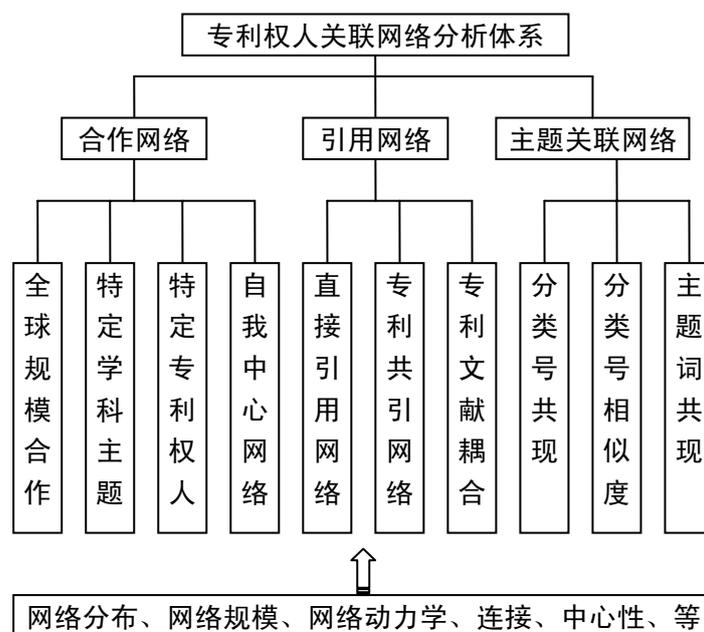
这一点也是目前研究的难点和需要深入研究的方向之一，对中心性、连接性、边缘性、此处分析是挖掘网络关联图谱所反映问题的重要工作

3.3 实施路线图



4 社会网络分析在专利权人网络分析中的应用方法

本课题基于社会网络分析方法现有的理论与方法体系及在社会学领域特别是在论文分析中的应用，结合专利文献的数据结构特征，在方法层面研究适用于专利权人网络分析的方法，包括基于合作、基于引用和基于技术主题三个方面。



由于不同的专利数据库著录结构略有不同，为了提高本课题研究结果的广泛适用性，本课题以Derwent Innovations IndexSM数据库（以下简称DII）收录的发明专利申请（以下简称“专利”）为基础，以基本专利年（DII数据库首次收录的专利家族成员专利的公开年）为年度划分依据，基于DII的数据结构特征，探讨社会网络分析在专利权人网络分析中的应用方法。表1列出了DII数据库所收录的发明专利的著录结构特征。

表1. DII收录的发明专利著录结构示例¹

PN（公开号）	WO2007044462-A1; US2007110358-A1; US7599584-B2	
TI（名称）	Planar mid-infrared integrated microphtonic platform for use in mid-infrared telecommunication, has laser to perform lasing functions on optical signal, where laser comprises chalcogenide glass.	
AU（发明人）	HU J AGARWAL A M KIMERLING L C	
AE（专利权人）	MASSACHUSETTS INST TECHNOLOGY (MASI) CENT NAT RECH SCI (CNRS)	A
	HU J (HUJJ-Individual) AGARWAL A M (AGAR-Individual) KIMERLING L C (KIME-Individual)	B

¹表中部分著录信息对原专利进行了修改，此处仅作为示例使用，目的是为了说明专利权人网络的情况。

GA (入藏号)	2007--533703	
AB (摘要)	(略)	
TF (技术焦点)	TECHNOLOGY FOCUS – 略.	
DC (德温特分类代码)	P81 (Optics); U12 (Discrete Devices, e.g. LEDs, photovoltaic cells); U13 (Integrated Circuits); W02 (Broadcasting, Radio and Line Transmission Systems)	C
MC (德温特手工代码)	U12-A01B; U12-A01C; U12-A02B2A; U12-A02B5D; U12-E01A4; U13-D04A; W02-C04B1	
IP(国际专利分类)	G02B-006/10; G02B-006/12; G02B-006/13; G02B-006/26	
PD (专利详细信息)	WO2007044462-A1 19 Apr 2007 G02B-006/12 200752 Pages: 23 English US2007110358-A1 17 May 2007 G02B-006/12 200752 US7599584-B2 06 Oct 2009 G02B-006/12 200966 English	
AD (申请详细信息)	WO2007044462-A1 WOUS038973 05 Oct 2006 US2007110358-A1 US543333 05 Oct 2006 US7599584-B2 US543333 05 Oct 2006	
FD (更多申请详细信息)	US2007110358-A1 Provisional Application US723869P US7599584-B2 Provisional Application US723869P	
PI(优先权申请信息和日期)	US723869P 05 Oct 2005 US543333 05 Oct 2006	
DS	略	
CP	US2003128922-A1 MASSACHUSETTS INST TECHNOLOGY (MASI) PETRICH G S, KOLODZIEJSKI L A	D
CR	略	

表 1 中:

A 区域是课题分析专利权人合作网络、基于技术主题的关联网络的节点基础。

B 区域是专利权著录项目中为个人的情况，本课题的专利权人不考虑个人，理由如下：美国的专利申请必须是以原始的发明人提出，然后再转让给公司（所以公司与员工劳务合同中都会有关于知识产权的事项），所以美国的专利记录用 assignee（受让人）来代表专利权人，当美国专利记录里没有申请人 applicants 的著录时，都用 inventors 来表示。因此，当美国专利公开文件记录中有 assignee 字段后，DII 就识别这个字段为专利权人，当没有 assignee 字段时（此时属于单位没有把和发明人签署的转让协议出示给 USPTO），DII 就识别 inventors 为 applicants，并著录到 assignee 中。为了避免因为美国专利法带来的 DII 专利权人著录项中增加了发明人的信息带来的影响，本课题清洗掉专利权人为个人的情况。然而，这样处理会漏掉个人切实作为共同专利权人的情况。故本课题的研究方法更适用于机构法人的专利权人网络分析。

C 区域是基于技术主题的专利权人网络分析的技术主题依据基础，包括国际专利分类号 IPC、德温特分类代码 DPC 和德温特手工代码 MC。其中：①IPC 适用于功能分类分析，因为该体系以功能分类和应用分类相结合，侧重功能分类。②DPC 适用于主题分类分析，德温特分类代码是以主题范畴分类为基础的分类体系。③MC 适用于应用分类分析，MC 体系与现有技术有很好的对应。课题重点研究基于 IPC 和 DPC 的专利权人网络。

D 区域是示例专利所引用的在先专利列表，其中包含被引专利的专利权人信息，这也是

构建专利权人引用网络的重要数据基础。

4.1 基于合作的专利权人合作网络分析方法

4.1.1 基本原理

专利权人合作网络分析主要用于发现专利权人之间的相互合作强度以及竞争合作关系，包括研发合作、专利权转移所产生的关联关系，进而分析机构与外界在技术创新方面的合作与交流情况。通过分析共同申请人（如基于研发合作）而产生的专利权人合作网络，可以发现机构在技术研发阶段整合或利用外界技术优势或资金优势的力度，这种合作网络属于无向网络（如图 5-a）；通过分析基于专利权转移而产生的专利权人合作网络，可以发现机构专利的产业化价值与市场潜力，这种合作网络属于有向网络（如图 5-b）。将分别基于技术研发阶段的无向网络图和基于专利权转移的有向合作图进行整合，则可以分析两种网络图的中心点是否存在重合或某种关联（如图 5-c）。然而，由于专利的特殊性，其专利权人也会因专利权的转移而发生增加或变更，专利数据库在收录专利信息时也缺乏对由于不同原因产生的专利权人进行明确标识，因此，本项目研究则忽略上述无向图和有向图之间专利权人性质的差异，对因专利权转移产生的合作关系和基于研发合作的共同专利权人的合作网络不作区分，均纳入专利权人合作网络（如图 5-d）。

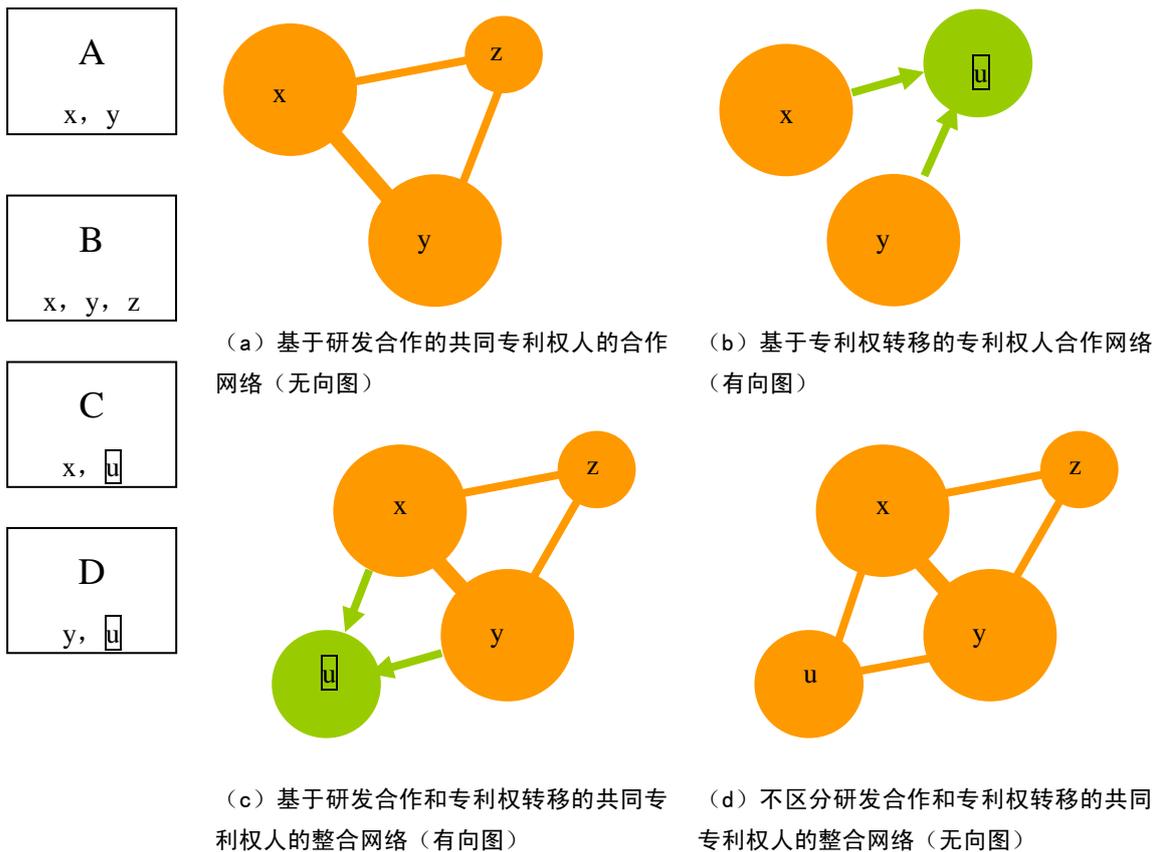


图5. 专利权人合作网络示意图

（说明：A、B、C、D——发明专利；x、y、z——专利的基于研发合作的共同专利权人；u——基于专利权转移产生的专利权人）

鉴于本项目研究以 DII 数据库作为主要数据来源，而该数据库在专利权人著录项没有对专利权人的性质进行区分，因此，本项目在下述研究内容中主要研究不区分研发合作和专利

权转移的共同专利权人的整合网络（无向图）。

在后续的研究中，可以针对小样本，例如最近几年中国科学院专利权人合作的关系，区分是在研发阶段的研发合作、企业对研究所的投资合作或者专利申请后或授权后的技术转移转化等建立的合作关系现象。

4.1.2 分析方法研究与内容

（1）全球规模的专利权人合作网络分析方法

在社会网络分析中，最典型的例子是巴恩斯对挪威一个渔村的考察研究，关注在社区整合的生产过程中亲属、朋友或邻里所起的作用^[24]。在社会网络分析发展的数十年内，更多的研究都是针对特定范围内的关系的研究，很少见到在全球范围内以某类事物为节点的研究，因为在全球范围内多数事物的节点数量巨大，很难对庞大的数据进行加工和分析，即使采用超级计算的方法获得了分析结果，面对庞大的网络，受到地理、文化、风俗等等诸多因素的影响，也很难针对这样的结果总结和归纳出一般性的规律。因此，社会网络分析更倾向于对有限的、相对封闭的小群体进行分析，归纳出广泛适用的规律，然后在其他群体中进行验证。然而，随着网络分析的一步步发展，面对复杂性网络的研究，面对一些特定的问题，为了理解网络复杂性，了解网络背后的组织原则，我们需要构建完整的网络。例如，在互联网领域，绘制完整的互联网图具有重要的现实意义，如果不知道互联网的拓扑结构，就不可能设计出更好的工具和服务；在医学领域，p53网络（与p53细胞周期调控因子相互作用的分子和基因的集合，50%的人类癌症都和p53基因有关）的研究具有重大意义，人们需要破解这个网络的拓扑结构，彻底了解里面所有的相互作用，最终找到修复p53细胞周期调控因子的药物；在经济领域，随着经济全球化的一步步增强，研究全球范围的完整企业关系网络，对挖掘企业的竞争合作行为、能力及发展潜力是不可或缺的^[25]。

可喜的是，在图书情报领域，论文和专利数据与地球上的大多数其他事物相比，相对“单一化”，具有数据易获取、易处理的特点。因此，基于现有数据库，我们可以获取特定时间段内的全球所有（数据库收录范围内）的论文或专利数据，进而对其进行分析。针对专利而言，DII数据库收录了全球41个专利授权组织的1400多万发明专利申请数据，对其专利权人进行合作网络研究，研究其增长动力学模型，可以发现全球创新合作的动力所在；通过网络分布特征分析，挖掘以专利权人为节点的合作是否符合幂律分布？通过分析专利权人合作网络的节点中心度或紧密中心度，挖掘出处于全球研发合作中心地位的中心节点；还可以通过间距中心度的测量，发现那些处于“中介”位置的专利权人，这些间距中心度高的专利权人具有沟通和桥梁的作用，特别是对连接跨学科机构之间的中介点、连接不同国家机构的中介点的研究，对专利权人合作网络研究具有重要意义。

然而，虽然美国已有学者基于全球规模的自然科学论文数据，对整个自然科学的关联网络进行了全面的描绘和分析^[26]，毕竟面向全球规模的合作网络研究还是太大了，虽然全球规模的专利权人合作网络可以呈现出明确的合作信息及其演变历程，然而，从社会网络分析所关注的问题焦点而言，社会网络分析更关心的是专利权人之间的关系，以及这些关系深处所隐含的内在动力。为此，可以采用抽样的方法在全球规模的专利数据中抽取样本，构建一个小群体，进行社会网络分析，抽样的方法可以采取随机抽样，也可以基于一定的标准有选择性地抽样。本课题则采取后者的方法，进行有选择地抽样研究，重点阐述特定小群体的专利权人合作网络，包括以学科领域为界定群体的依据或以国家和专利权人的性质为群体划分依据，开展专利权人合作网络的社会网络分析。以学科领域为划分依据在论文研究中已较为常见，并获得了相当多的成果，近年来颇具代表性的一份研究成果是文献^[27]，对化学学科的发展进行了详尽的网络分析。

(2) 特定学科领域(技术主题)的专利权人合作网络分析方法

① 局部网的构建方法

第一,基于数据库自身的分类体系,下载并分析某特定领域的数据库。如陈云伟等(2008)基于SSCI收录的Information Science & Library Science(情报学及图书馆学)64种期刊,对其收录的2001—2007年的70022篇论文进行科学计量分析^[28]。研究世界图书情报学科的年度发展趋势、国家和机构布局以及图书情报科学与其他学科的交叉布局。Leydesdorff与Rafols(2009)基于ISI的学科分类构建了全球科学地图^[29]。

第二,以某个期刊或会议为基础,对一定时期内的论文进行作图研究。如Börner, Penumathy和Meiss等(2006)采用科学地图的方法描绘了1982-2001年间在PNAS国际会议上发表论文最多的500个研究机构的信息^[30]。Chen, McCain和White等(2002)^[31]、Hou, Kretschmer^[32]和Liu(2008)、Dutt, Garg和Bali(2003)^[33]分别基于*Scientometrics*期刊进行了作者共引、机构合作和国家分布等科学地图研究。

第三,以关键词进行检索。如Judit(2008)采用关键词检索的策略对计量情报学在21世纪的发展进行了综述研究^[34]。

第四,利用国际专利分类号或美国专利分类号进行专利检索,用于构建专利数据集。例如,OECD^[35]采用国际专利分类号IPC对生物技术发明专利进行了定义,英国SPRU^[36]采用美国专利分类号USPC对生物技术及其所涵盖的8个主要方向进行了定义,Linton^[37]等采用USPC对工业生物技术专利进行了定义。本项目组在开展有关纤维素乙醇发明专利分析时,也采用IPC对发酵法生产纤维素乙醇和纤维素酶的发明专利进行了定义,在《中国科学院专利分析报告》的撰写中,借鉴了Linton的定义,构建了用于分析的相关群体基础。

② 存在的问题

从学科出发获取的某专利权人的专利,一般并非是该专利权人的所有专利,因此,此时的合作网络规模小于专利权人的实际合作网络,而且会“漏”掉专利权人跨学科的合作关系,但是,这种方法对分析某学科内部的专利权人合作网络而言是非常有效的。即,此时的网络是一个“局部网”(partial network),是在全网(所有专利权人网络)中专利权人基于所分析的学科而形成的一部分网络。同时,为了弥补“漏”掉的关系,可以对群体的构建方法进行改良,具体方法是,基于领域获取专利权人列表后,再以此列表重新检索专利,获得与出发领域相关的专利权人所有的专利数据(包含非出发领域的专利数据),再基于新获得的数据进行社会网络分析,然而此时的群体节点数会增加,新增加的专利权人恰好为与出发领域相关的专利权人,那些带来这种新增的领域内专利权人,则起到了重要的中介作用,一般拥有跨领域的研究主题。

③ 分析指标

网络研究,特别是复杂网络研究正在经历从发现不同网络共性到深入挖掘具有代表性的典型网络的特性这一过程,在开展复杂网络研究的最初几年,推动研究的主要力量是隐藏在各种各样网络中的统计共性,例如小世界现象、无标度效应、社团结构等等;最近三四年,大量的研究开始关注一些有代表性的网络具有的不同于其他网络的独特的性质,例如公路网的强几何约束条件^[38],计算机互联网独特的外围增长激励机制^[39],科学家合作网的富者俱乐部^[40]现象,等等。正是受这种从一般到特殊,从共性到个性的研究趋势的影响,本课题也将针对专利权人的合作网络特征进行研究。具体分析指标包括:

网络分布特征分析——分析网络的分布特性,节点是否符合幂律分布还是钟形分布,或其他分布特征,是否存在中心节点和中介点的数量与分布情况。

网络规模分析——网络密度大小分析,网络直径分析。

动力学分析——网络动力学性质的基本研究对象是动力学模型在不同网络上的性质与相应网络的静态统计性质的联系,包括已知和未知的静态几何量。如果我们发现了某个模型

在某一网络上有某种特殊的表现,那么可以认为是这一网络的某种特征影响了这个模型的表现。这种特征有可能是已经得到研究的这种网络的几何特性,也有可能是没有被发现的几何特征,那么前者将印证网络上这些几何量的重要性,而后者将会推动网络本身研究的发展。

弱连接分析——专利权人节点之间的连带强度在直观上可以分为强连接、弱连接和无连接。假定下列情境:任意选定的两个专利权人 A 和 B、以及一个特定的集合 S(S=C, D, E……),集合中的人与 A 或 B 的任一人有连带。当 A 与 B 之间的连带愈强,则在 S 当中的个人同时与 A 和 B 产生连接——不管是强连接还是弱连接——的几率将会增大。A 与 B 合作者的重叠成度,在 A 与 B 没有连接的时候重叠最少,是强连接的时候重叠最多,而是弱连接的时候,重叠程度则是中等的。基于这种假设,还有一个重要的结论:如果 A 和 B 很紧密地连接在一起,且 A 和 C 之间也有强连接,此时 B 和 C 之间没有连接的可能性是非常小的。即,三角关系的形成会随着专利权人之间相互了解的时间的增长而增加,并逐步发展成强连接。

桥分析——桥即在一个网络当中,提供给两点之间唯一路径的一条线。介于 A 与 B 之间的桥便是提供这样一条路径,是信息或影响力能够从 A 的任何一个合作专利权人,流动到 B 的任何一个合作专利权人。因此,从任何一个和 A 有合作的专利权人,到任何一个和 B 有合作的专利权人,都可能依靠这条桥。所以,在网络传播的研究中,我们可以预期这些桥均扮演者重要的角色。一般而言,社会学研究发现,除了在极少数不太可能发生的情况下,没有强连接能够成为桥,只在一个小团体中较有可能。然而弱连接却不受这些限制,虽然它们并非自然能成为桥,但重要的是,所有的桥都是弱连接。然而在专利权人合作网络中,桥的分布特征是否与社会学研究结构一致,这是进行专利权人合作网络桥分析的重大魅力所在。

弱连接的重要性在于当它们成为这些区间桥梁,可以创造更多,而且更短的路径。一般而言,弱连接的消失,对于信息传递的可能性所造成的损害,比起一般强连接的消失将会更大。

中心、中间专利权人分析——已有研究发现,发明人合作网络中,一个完整的合作网络一般会含有超过一个中心发明人,围绕多个中心发明人形成多个簇,簇与簇之间一般通过少数几个发明人进行连接。这种位于不同发明人合作团队之间或研发部门与机构之间的界面处的中间发明人具有更好的专利产出数及被引频次,说明处于中间位置的发明人受益于信息流,并对他们的专利数量和质量有证明影响,网络的中心发明人也有类似的特征。那么,专利权人合作网络中的中间专利权人是否具有类似的特征,这将是本课题实证研究部分需要研究的问题。本课题还将研究中心专利权人与中间专利权人是否拥有更多的 IPC 分类主题,进而揭示中心性与技术主题的关系。

④ 特定学科主题的专利权人合作网络分析实例

本课题以DII数据库收录的基本专利年(Basic Patent Year)为1996-2009的939件纤维素酶发明专利申请数据为基础,构建专利权人合作网络。数据获取基于Linton等^[41]对纤维素酶相关专利的定义,以DII数据库为基础,分析纤维素酶相关专利1996-2009年的发展情况。DII数据下载日期:2009年12月1日。

1996-2009年期间,平均每件专利仅有0.2位合作专利权人(不考虑专利权人为个人的情况,否则该值为1.8),表明从事纤维素酶研发的专利权人在研发和技术转移方面的行动活跃度较低。

表2. 纤维素发明专利申请前10位专利权人(TOP10)

专利权人		专利数
Novozymes AS	诺维信	101
Danisco/Genencor	丹尼斯克/杰能科	75
Dokuritsu Gyosei Hojin Sangyo Gijutsu So	宝神产业株式会社	30
Kao Corp	花王公司	29

Meiji Seika Kaisha Ltd	明治制果株式会社	27
Procter & Gamble Co	宝洁公司	21
Chinese Acad Sci	中国科学院	17
Univ Jiangnan	江南大学	17
Diversa Corp	DIVERSA 公司	14
Univ China Agric	中国农业大学	14

从 TOP10 专利权人角度分析，在纤维素酶领域，诺维信牢牢占据着全球领先的地位，其次是丹尼斯克/杰能科。2008-2009 年，中国的相关机构专利数明显增加，TOP10 专利权人中有 6 个来自中国，其中以江南大学专利数最多。从 1996-2009 年的排名分析，中国科学院、江南大学和中国农业大学进入全球 TOP10 中，预期未来数年内，在 TOP10 专利权人中，中国的相关机构所占的份额将进一步提升。

那么，社会网络分析可以在对深入挖掘纤维素酶发明专利专利权人的竞争行为上发挥什么样的功能呢？我们尝试利用社会网络分析方法，借助社会网络分析工具 NWB^[42] 进行分析与挖掘。

首先，在未清洗发明人著录在专利权人项的情况时，939 件纤维素酶发明专利申请总计拥有 821 个专利权人，其中有 177 个专利权人属于孤立节点，未与其他任何专利权人之间存在合作关系，这其中就包括中国的江南大学，也是 TOP10 专利权人中唯一没有与其他专利权人进行合作的机构。合作网络中形成一个相对超大的成分，由 362 个专利权人组成，相互之间存在 1409 条连线，见图 6，其中专利权人之间的最大合作次数 14 次，平均合作次数为 1.2 次，网络密度为 0.02，平均每个专利权人与 7.8 个专利权人存在合作关系。

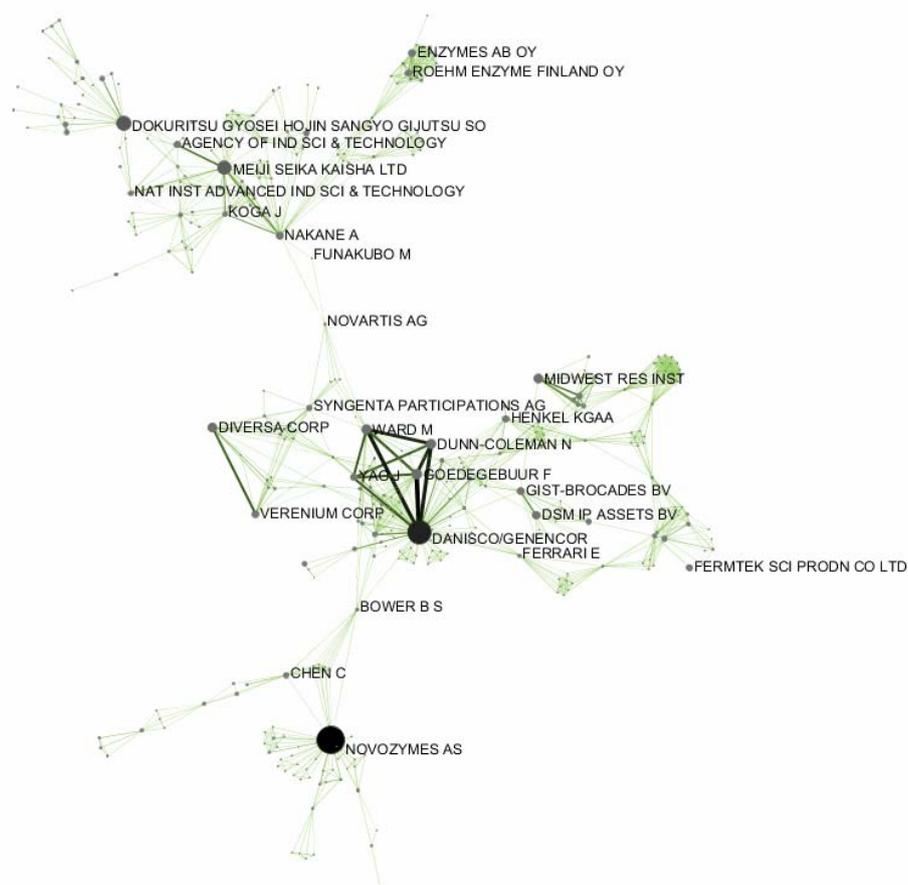


图6. 纤维素酶专利权人合作网络最大关联子图（未清洗发明人著录在专利权人项的情况）

图 6 清晰显示，在纤维素酶专利申请数 TOP10 专利权人中，有 5 个专利权人均未出现在合作网络最大关联子图中，它们是中国的一家机构——中国科学院、江南大学和中国农业大学，宝洁公司和花王公司。其他 5 个 TOP10 专利权人均出现在图 6 中，并且有 4 个还是网络的局部中心节点，它们是诺维信、丹尼斯克/杰能科、宝神产业株式会社和明治制果株式会社。

此外，从图 6 中不难发现，诺华公司（Novartis AG）和 Bower BS 是两个明显的中间专利权人，在最大关联子图的形成上起到了重要的连接作用。

然而，这里的两个发现与已有的研究并不完全相符^[43]，在特定学科领域的专利权人合作网络中，中心节点虽然具有了拥有较多专利数的特征，然后两个中间节点的专利数则非常少。有必要对此现象进行深入剖析，发现在学科领域专利权人合作特定网络中，中心节点与中间节点的特征。本报告暂不做深入分析。

另一方面，4 个主要的中心节点的周边分布着大量的发明人著录为专利权人的噪音影响，阻碍了对真实专利权人合作网络的呈现。同时，两个重要的中间节点之一 Bower BS 是一发明人著录为专利权人的噪音，而且该发明人仅属于丹尼斯克/杰能科公司，与诺维信的弱连接也是一种错误的构建。因此，有必要对发明人著录为专利权人的噪音数据进行清洗，构建更加准确的合作网络。

其次，我们研究将专利权人著录项中的发明人进行清洗后的专利权人合作网络，经清洗后，有 46 件专利因为专利权人著录项只有发明人而被清除，进而获得剩余 893 件专利进行合作网络分析。此时的网络包含 372 个专利权人，其中有 200 个专利权人为孤立节点，未与其他任何专利权人存在合作关系，其余 172 个专利权人相互之间存在 186 条连线，由 47 个子成分构成，其中最大成分包含 29 个节点。

数据表明，经清洗发明人后的专利权人合作网络已变得非常松散。专利产出数最多的 TOP10 专利权人之间合作强度非常低，仅宝神产业株式会社（Dokuritsu Gyosei Hojin Sangyo Gijutsu So）和明治制果株式会社（Meiji Seika Kaisha Ltd）之间存在微弱的直接关系，因此也形成了包含 29 个专利权人的最大成分，此时这两个专利权人之间的链接则担当重要的桥线的作用。此外，在 TOP10 专利权人中，诺维信公司、丹尼斯克/杰能科、中国科学院和 Diversa 公司都与其他专利权人存在不同程度的合作，而花王公司、宝洁公司、江南大学和中国农业大学则没有任何合作行为。网络分析的结果提示我们，在纤维素酶的国家研发行为中，存在非常强的竞争关系，特别是研发实力较强的机构之间的竞争关系更加激烈，在研发活动上呈现出“不相往来”的现状；另一个重要特征为，虽然重要的专利权人之间“不相往来”，但他们却热衷于与其他专利权人开展合作，也许是为了整合多方面的力量，用于对抗竞争对手。值得一提的是 Diversa 公司，其与专利权人 Verenium 公司的关系实际是由公司重组合并而带来的，2007 年，Diversa 公司和 Celunol 公司合并为 Verenium 公司，所以，从这个角度考虑，这两个节点应合二为一，其与另外一个重要的专利权人先正达制药公司存在微弱的关联。中国科学院则与山东大学之间存在微弱的关联。

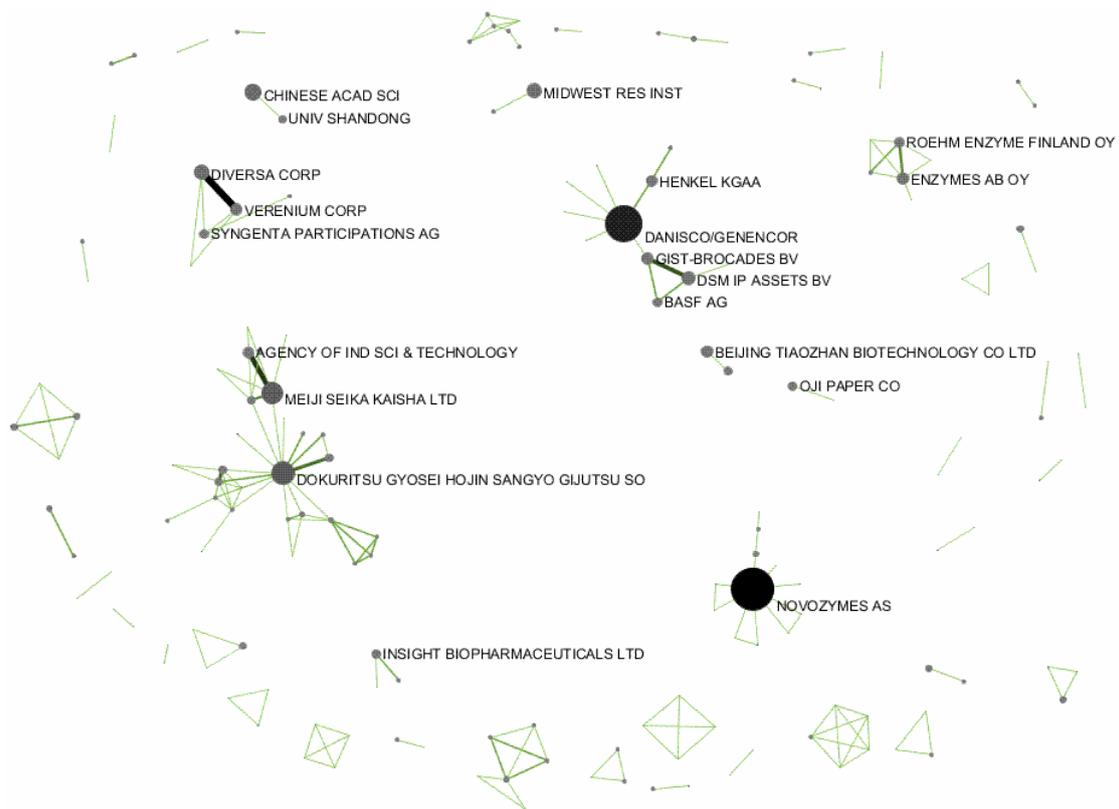


图7. 纤维素酶专利权人合作网络最大关联子图（已清洗发明人）

在社会网络分析中，还可以深入分析网络的演化动力学，并对未来的演化进行预测，预测新连接的出现，网络的演进速率及发展情况。本节仅作方法的研究以及所能实现的功能，具体动力学研究在此不作深入分析，在实证部分将对中国科学院专利权人合作网络的发展动力学进行分析。

表3. 专利权人著录项中发明人清洗前后专利权人合作网络参数比较

网络参数（统计信息）		发明人清洗前	发明人清洗后
节点数（专利权人数）		821	372
孤立节点数（无合作专利权人数）		177	200
边数（合作关系数）		1918	186
边的最大权重值（最大合作次数）		14	8
边的平均权重值（平均合作次数）		1.2	1.3
删除孤立节点后的网络参数	平均度数（各机构的平均合作系数）	6	1
	成分数/最大成分节点数	70/362	47/29
	网络密度	0.01	0.003

从表3的统计不难看出，对著录在专利权人项中的发明人进行清洗后，整个专利权人的合作网络结构发生了重大变化，基于前文对发明人著录在专利权人项中的原因的分析，可以得出在专利权人合作网络分析中，清洗发明人是非常关键的一个步骤，否则得到的网络无法真实反映专利权人的实际合作关系。

（3）特定专利权人的合作网络分析方法

现有的合作网络研究，不论是基于论文的作者合著网络还是基于专利的专利权人合作网

络，多数研究都集中在针对特定主题领域的研究，然而，除了从学科领域出发外，还可以分析包括不同性质的专利权人（如高校、企业、科研院所）之间的合作关系。然而，在专利分析方面有关研究较少。

①局部网的构建方法（数据集的构建方法）

第一，以行政地区为依据，分析全球、洲、国家或省市的专利权人合作网络，可以对特定国家进行比较研究。

第二，以专利权人性质为划分依据，例如，分析高校、科研院所或公司企业的专利合作网络。

②分析指标

同 4.1.3 (2) ② 章节。

③分析实例

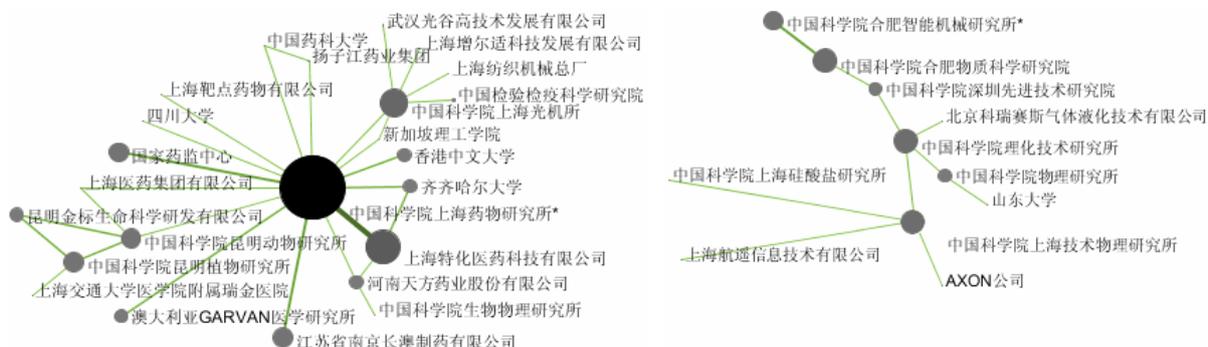
本课题以 DII 数据库收录的最早优先权年（Priority Years Earlist）为 2005-2008 的 124 件合作专利申请数据为基础，构建专利权人合作网络。DII 数据下载日期：2009 年 10 月 30 日。

2005-2008 年，以中国科学院院属各研究所及其合作机构为分析节点，院内共有 48 个研究所有专利合作行为，它们共有 67 个院外合作专利权人，合作专利数总计 124 件，其中院内研究所之间的合作 25 件，与院外合作 99 件。

表4. 2005-2008 年中国科学院院属研究所合作专利的合作网络参数

网络参数（统计信息）	2005-2008 年
节点数（专利权人数）	115
边数（合作关系数）	103
边的最大权重值（最大合作次数）	7
边的平均权重值（平均合作次数）	1.42
平均度数（各机构的平均合作关系数）	1.79
网络密度	0.02
网络直径	6
网络平均最短路径	2.48

表 4 给出了 2005-2008 年中国科学院院属研究所合作专利的合作网络参数，除了院外的 67 个合作专利权人外，院内 48 个研究所有专利合作行为。总计 115 个合作专利权人之间的最大合作次数为 7，平均合作次数仅为 1.42 次，平均每个机构仅与 1.79 个机构存在合作关系，网络密度值仅为 0.02。



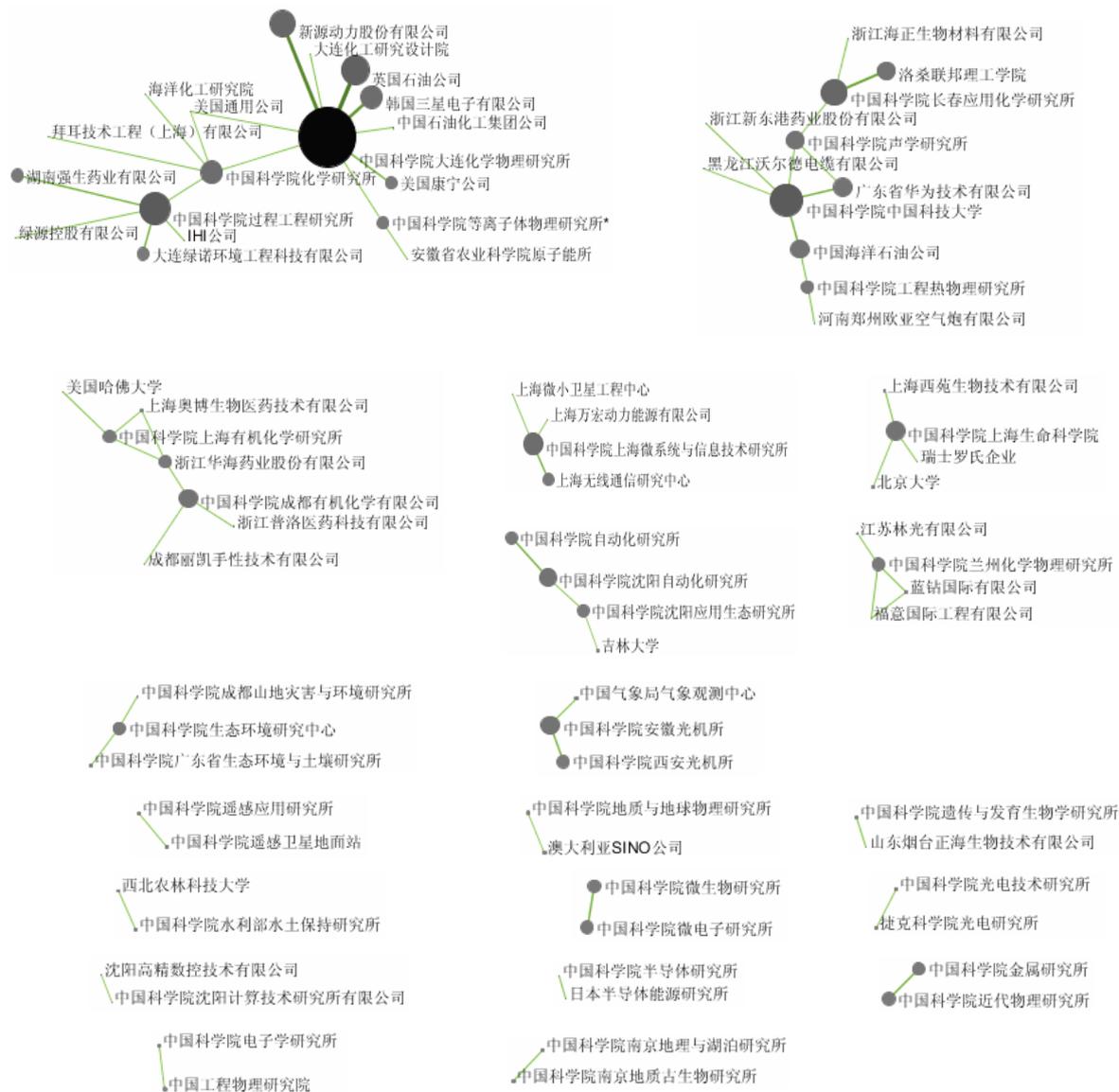


图8. 2005-2008 年中国科学院院属研究所合作专利的合作网络图

图 8 展示的中国科学院院属研究所合作网络关系图显示，上海药物研究所和大连化学物理研究所的网络中心性最强，其次上海光机所、化学研究所、过程工程研究所和中国科技大学拥有相对较高的合作系数。其中，最大的簇由 24 个专利权人组成，以中国科学院上海药物研究所为中心，大多为药物研发及相关的研究机构或企业；其次是一个由 18 个专利权人组成的簇，以中国科学院大连化学物理研究所为中心，大多为化工类研究机构或企业；有两个簇分别由 11 个专利权人组成，一个以中国科学院上海技术物理研究所和中国科学院理化技术研究所为中心，大多为理化类研究机构，另一个以中国科技大学为中心，以物理、化学类机构为主。其次依次为含有 7 个专利权人的簇 1 个、4 个专利权人的簇 4 个、3 个专利权人的簇 2 个、2 个专利权人的簇 11 个。

相互之间合作次数最多的机构是上海药物研究所与上海特化医药科技有限公司之间的合作，以及大连化学物理研究所与新能源动力股份有限公司之间的合作。

数据表明，在中国科学院院属研究所中，上海药物研究所、大连化学物理研究所、上海光机所、化学研究所、过程工程研究所和中国科技大学，拥有相对较高的合作系数，合作活跃度较高。然而，机构间合作次数多的关系多发生于院属研究所与院外机构之间，相比之下，院属研究所之间的相互合作强度则相对较弱。

(4) 自我中心的合作网络分析方法

① 自我中心网的构建方法（数据集的构建方法）

最近几年，有大量的文献已经开始分析关于个人所镶嵌的社会网对个人行为的影响。自我中心网络的不同部分或许会有不同的密度。他将一个人与之“互动非常紧密有相当固定”的朋友，以及由此一关系而互相认识所形成的网络称为“有效网络”（effective network），其余则构成延伸性网络（extended network）。目前有一个论点一直未获得普遍的同意，亦即自我中心网络是否只包括那些与自身有直接连接的人即可，或是应该包括与他有连接的人的朋友，或者是其他人。强调简化个人网络的，倾向于采用前者，而后者则强调网络的实际运作。鉴于此，本节实例研究中国科学院自我中心的专利权人合作网络分析方法，仅考虑与中科院有直接连接的专利权人与中科院的合作关系，在后续实证研究中，则以中国科学院为中心，研究其专利权人合作网络。

② 分析指标

演进动力学分析——通过网络演进动力学分析网络演进过程中新增节点的增加特点及规律，网络的增长对中心节点的贡献是什么？分析二级间接连接点转化成直接连接点的平均时间、比例以及转化条件。分析新增节点的主题分布特征，特别是结合中心节点的历史研发战略、规划、计划和具体举措分析伴随这些举措的同时其网络演进的特征，分析两者之间的关系，为以后的研发提供决策依据，也为从事竞争情报分析人员提供剖析竞争对手研发行为提供客观依据。

中间专利权人分析——这里所述的中间专利权人是一种借用中间中间性并加以改良的一种概念，重点挖掘在网络演进过程中和中心节点的合作强度并未随着时间的推移发生显著增强，但却通过该中间节点为网络引入重要的间接连接节点、并最终发展成为重要直接连接者的这样处于“中介”位置的专利权人。

③ 分析实例

本实例将整个中国科学院作为一整体单元（节点），研究其与国内外专利权人（大学、公司企业、科研院所等）的合作网络。以DII数据库收录的最早优先权年（Priority Years Earlist）为2005-2008的99件合作专利申请数据为基础，构建专利权人合作网络。DII数据下载日期：2009年10月30日。

经统计，除中国科学院属研究机构外，2005-2008年间，中国科学院共有67个合作专利权人，合作专利数总计99件。

表5. 2005-2008年中国科学院专利合作网络参数

网络参数（统计信息）	05-08
节点数（专利权人数）	68
边数（合作系数）	72
边的最大权重值（最大合作次数）	7
边的平均权重值（平均合作次数）	1.54
平均度数（各机构的平均合作系数）	2.12
网络密度	0.03
网络直径	2
网络平均最短路径	1.97

表5统计了2005-2008年99件合作专利的合作网络参数，结果反映，2005-2008年间仅占中国科学院发明专利申请总数不到1%的合作专利，其最大合作次数为7次，99件专利的平均合作次数仅为1.54次，主要贡献来自于合作数量排在前5位的上海特化医药科技有限公司、英国石油公司、新源动力有限公司、韩国三星电子有限公司和国家药监中心（见表

3-9)。每个机构平均只与 2.12 个机构存在合作关系，而且该平均值的主要贡献是中国科学院作为网络的中心，其度数为 67。

表6. 2005-2008 年中国科学院专利 TOP5 合作专利权人

TOP	合作专利权人	专利数
1	上海特化医药科技有限公司	7
2	英国石油公司	6
3	新源动力有限公司	5
4	韩国三星电子有限公司	4
5	国家药监中心	3

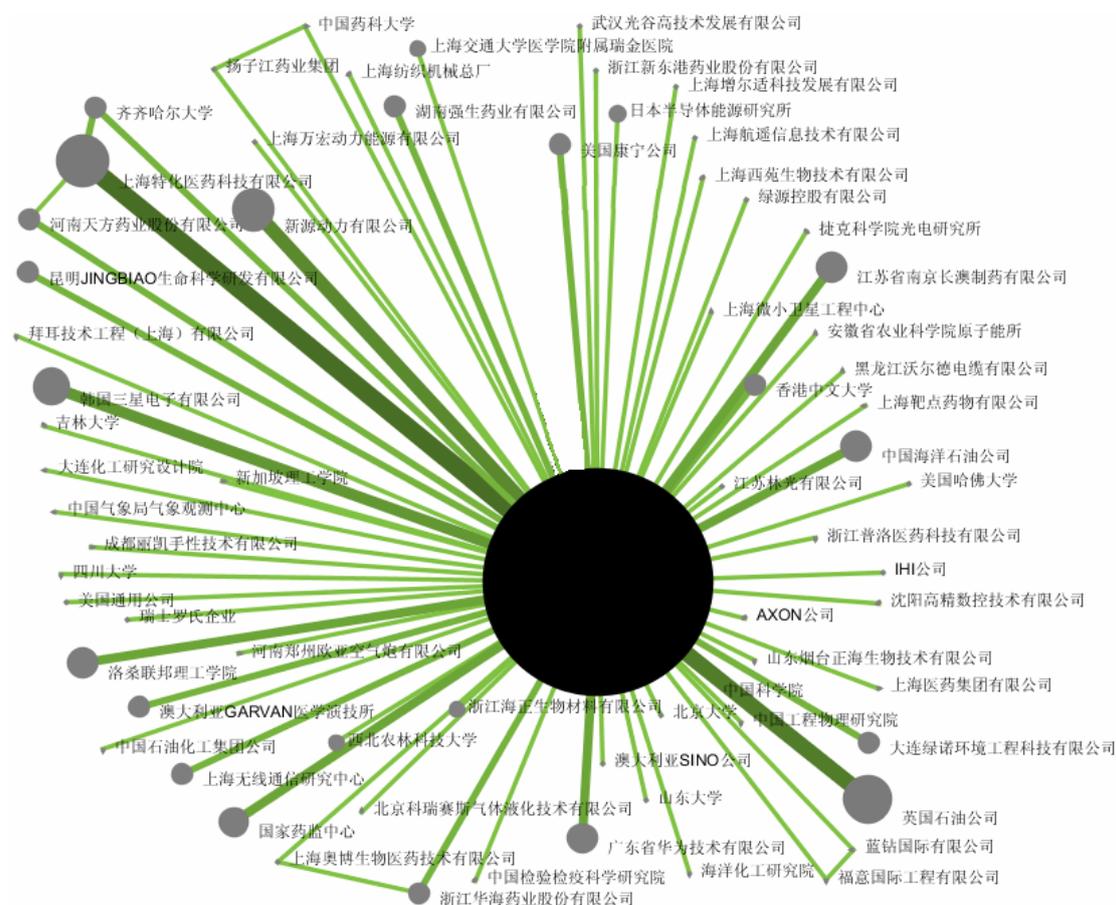


图9. 2005-2008 年中国科学院 99 件合作专利的合作网络图

从图 9 可见，67 个合作专利权人中，中科院仅与一家机构合作的专利权人有 58 个，与两家机构以上合作的专利权人有 9 个，其中上海特化医药科技有限公司分别与河南天方药业股份有限公司（1 件）和齐齐哈尔大学（2 件）一起作为中国科学院的专利合作者，中国药科大学与扬子江药业集团（1 件）、蓝钻国际有限公司与福意国际有限公司（1 件）、上海奥博生物医药技术有限公司与浙江华海药业股份有限公司（1 件）分别一起作为中国科学院的专利合作者。

数据显示，中国科学院在专利合作行为上，与国内外机构的合作力度还相对较弱，虽然这在一定程度上能反映出中国科学院研发拥有相对较强的独立性，但也应看到，中国科学院还有与国内外相关机构加强合作的空间，尤其是加强与国内外一流研发机构和企业开展相关

合作研究的同时，更应重视专利技术向公司企业等产业应用的转移转化工作。

4.2 基于引用的专利权人引用网络分析方法

4.2.1 基本原理

引文分析是科学计量学领域研究的重要内容，特别是针对论文的引文分析，被认为可以用来衡量真实环境中作者的信任关系，进而引文网络也可以被看做是一种信任网络^[44]。Judit在2008年发表在Journal of Informetrics期刊上发表的一篇综述把二十一世纪初期几年有关引文分析的内容分为了六个部分，分别为基于引文分析对期刊或研究领域进行描述、引文分析的理论和方法研究、自引分析、自我中心的引文分析、质量评估、引文分析与统计的错误研究^[45]。

从引文分析的基本原理角度，可以分为直接引用分析和共引分析，出发的角度可以包括论文本身、作者、机构、国家。具体包括三个大的方面，第一是定向连接网络，包括文章-文章的引用网络和作者-文章网络——一篇文章通过以参考文献的形式引用其他文章形成一个未加权的、定向论文引用图表，利用此类图表可以按照论文发表的顺序通过箭头指示信息流的方向。一篇文章的参考文献和引文可以及时地反映对先发论文的引用及被引图表，同时也是作为一篇文章的重要指标。第二是参考文献共现网络——拥有相同参考文献的文章、专利或其他学术产品被称为具有文献著录连接关系。学术文章的文献耦合（bibliographic coupling, BC）强度可以根据它们共同引用的参考文献的数目进行计算，连接强度被认为反映主题的相似性。第三是共引网络，包括文献共引网络和作者共引网络——文献共引网络最早于1973年提出，也是引文分析中最为重要的领域。

目前专利引用分析主要集中在对专利家族之间的引用关系方面，例如，Sternitzke等（2008）分析了LED领域特定年份内专利之间的引用关系^[46]，德国汉堡大学的Iwan von Wartburg等在2005年通过多级专利引用分析（multi-stage patent citation analysis）方法对技术演进进行了研究^[47]。此外也有学者对专利权人之间的引用网络进行了研究^[48]，但均未涉及基于共引和专利文献耦合的引用网络分析。

基于上述的论文引文分析内容与方法，本项目提出基于专利的专利权人直接引用网络（图10：x→z→v）、专利权人共引网络（图10：x-y）、以及基于专利文献耦合的专利权人网络（图10：z-w）三种主要的专利权人引用网络分析方法。本节将从这三方面对专利权人引用网络的方法和功能进行研究。

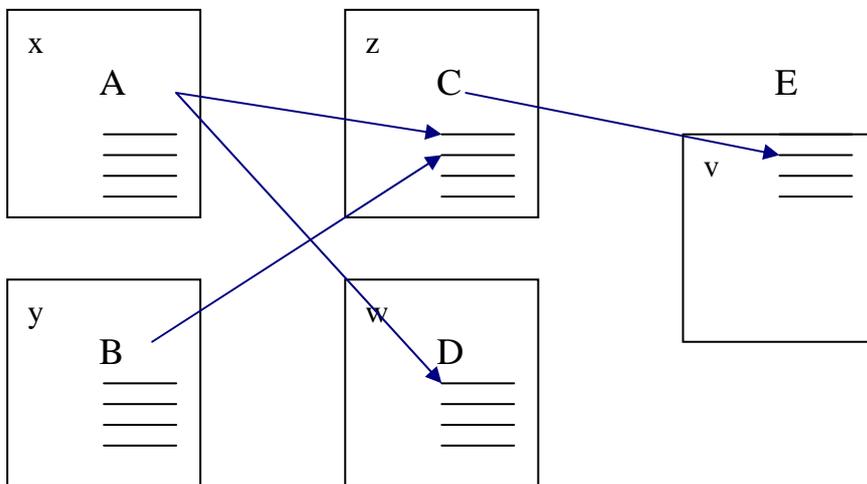


图10. 专利权人引用网络示意图

(说明: A、B、C、D、E——发明专利; x、y、z、w、v——专利权人, 当专利权人为多个时, 本节专利权人引用网络分析部分仅研究第一专利权人)

4.2.2 分析方法与内容

(1) 专利权人直接引用网络

根据不同的研究目的, 专利权人直接引用网络也可以分为不同的类型:

①从数据集的构建角度, 可以研究:

首先, 全球范围的专利权人引用网络, 虽然这是一项非常庞大的工程, 而且面临无法对数据进行有效清洗的弊端, 但此项研究仍具价值, 可以从宏观上反映出国际科技创新的关联关系, 发现学科关联与相互贡献, 迄今还未见此方面的研究成果。

其次, 研究不同学科主题之间的跨学科引用网络或学科内部的引用网络, 数据集的构建方法同 4.1.2 (2)。如Sternitzke等 (2008) 分析了LED领域特定年份内专利权人之间的引用网络^[49], 图 11 演示了他们构建的专利权人引用网络, 节点大小代表专利数目的多少, 节点之间的连线代表引用关系, 线的粗细代表引用关系的强弱。由于网络中节点和连线数目较多, 为了更加清晰地挖掘出重要专利权人之间的引用关系, 该文作者又对引用网络进行了挖掘, 只演示最活跃的专利权人之间的引用网络 (图 12)。

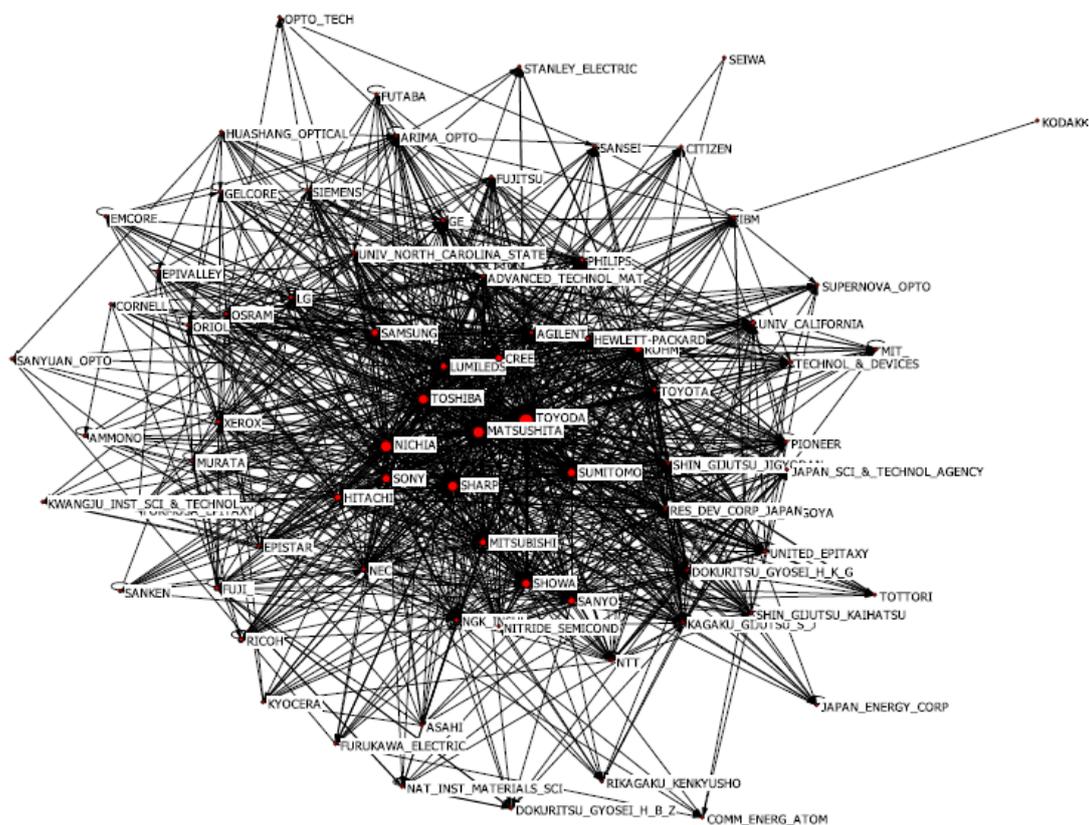


图 11. 专利权人引用网络[LEDs]

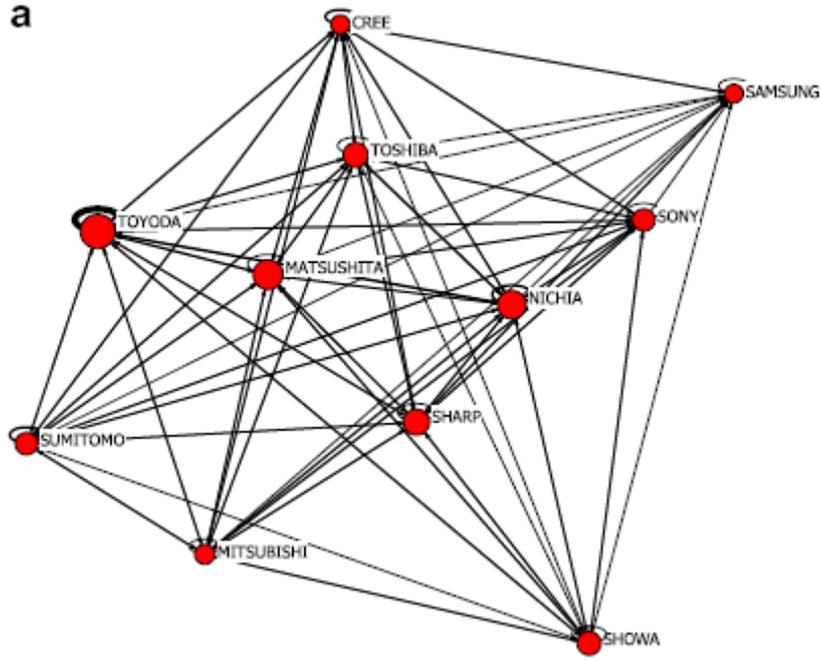


图12. 最活跃的专利权人引用网络[LEDs]

第三，研究特定专利权人的引文网络，如不同性质的专利权人（如高校、企业、科研院所）之间的合作关系。

最后，研究专利权人自引网络分析，严格意义上讲，专利权人自引网络并不属于专利权人引用网络，而是某特定专利权人所拥有的专利家族之间的引用关系。如图 13 是 Stermitzke 等（2008）分析的 CREE 公司的自引网络，网络中的每个节点代表 CREE 公司的专利。

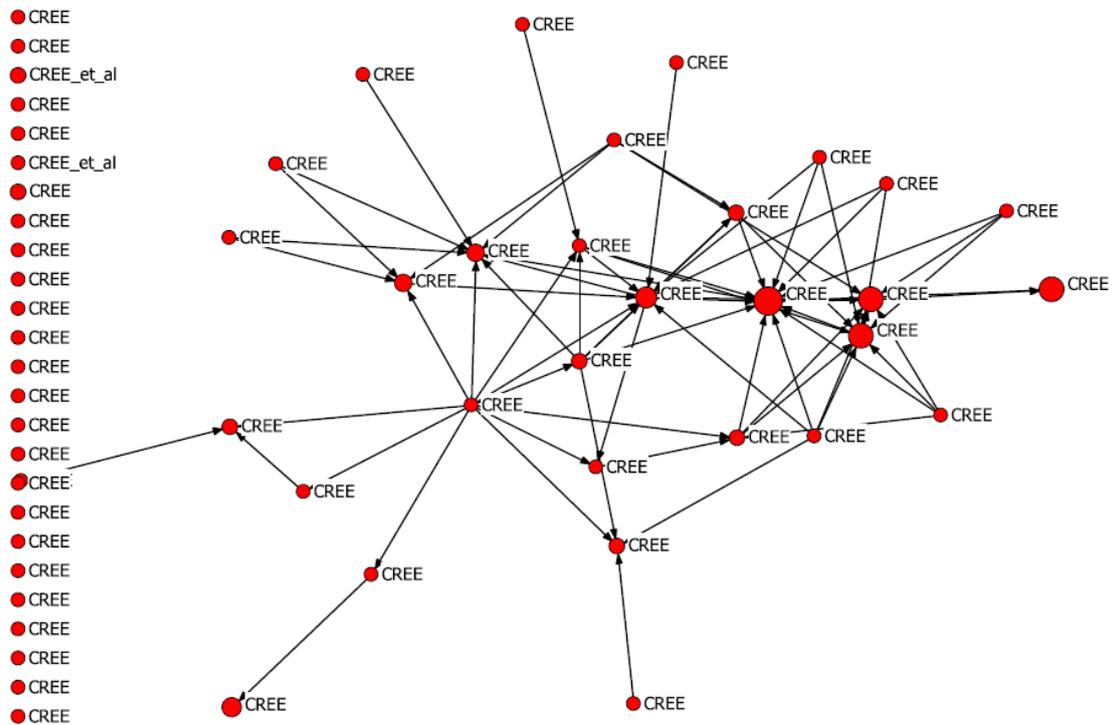


图13. CREE 公司自引网络（LEDs）

②从引文网络的层次角度分析，目前关注的一个重点是专利的多级引用网络，进行专利

的引文流分析。图 14 是多级引用分析的示意图。

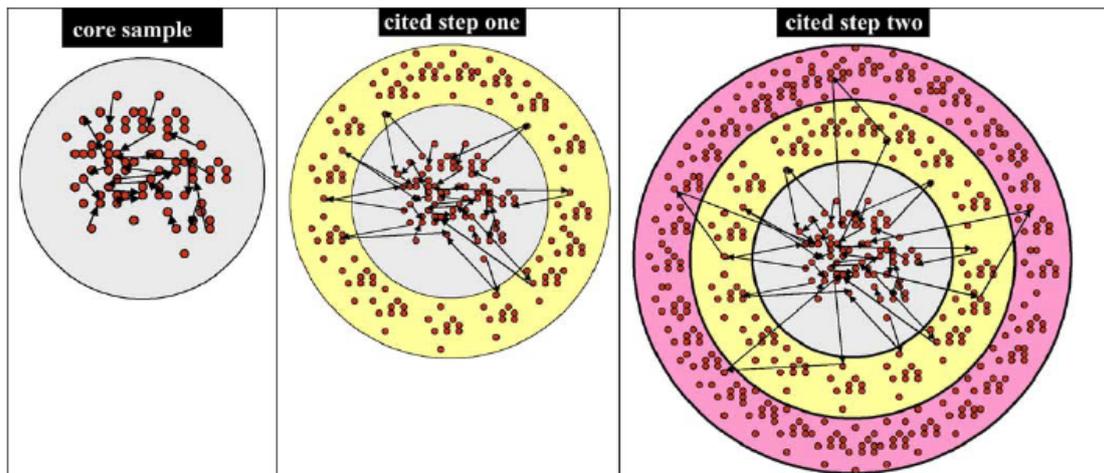


图14. 多级引用分析的示意图^[50]

(2) 专利权人共引网络

当两件专利文献同时被后来的一篇或多篇专利文献引用,同时把共同引用这两篇专利文献的数量称为共引强度。共引分析是引文分析中最具影响力的分析方法之一,利用共引分析可以得出分析对象之间由引文关系形成的关联关系,对这些信息用学科专业知识加以解释和分析判断,可以揭示研究对象的规律、预测其发展趋势。两篇专利文献被后来专利文献同时引用的次数越多,表明这两篇专利文献的关联程度越大,即说明这两篇专利文献在内容上相似性较大,也即技术内容(公司的研发领域)相似性较大。

(3) 基于专利文献耦合专利权人网络

从专利文献耦合角度研究专利权人的关联网络的研究报道还比较少见,基本原理与基于论文的文献耦合理论完全一致。特别是对技术情报价格更高的专利而言,分析专利权人的耦合网络,对分析专利权人之间的竞争关系与挖掘竞争对手而言,具有非常重要的作用。此方面的研究在在日后工作中一步步开展。

4.2.3 功能与作用

需要特别指出的是,尽管上述专利引用分析方法在一定程度上能反映出专利权人的技术影响力、研发竞争行为等特征,但如果要借助这些信息直接对机构的技术影响力进行判断,还需慎重。首先,尽管一些研究结果证明专利引用可以用于表征知识流^[51],但也有针对专利引用作为专利分析指标的批评^[52]。其次,Alcacer & Gittelman (2006)强调,美国专利商标局和欧洲专利局的审查员在审查发明专利申请时的方法完全不同,从而对每件专利的相对引文数量产生了强烈的影响^[53]。而且,美国专利商标局与欧洲专利局所依据的不同法律也影响了专利引文的数量。Michel & Bettels (2001)指出,美国专利商标局的引用数量是欧洲专利局的3倍^[54]。最后,研究人员还发现,发明人更倾向于引用地理位置相对接近的发明人的专利,这也会影响到引用数据的变化^[55]。

此外,若针对中国专利引用网络进行研究时还应注意,DII数据库收录我国国家知识产权局的发明专利申请数据中不包含引用信息。因此在实证研究中,本课题组也发现中国科学院的发明专利被引数量相对于国外机构的差距相当明显^[56],造成中国科学院专利被引数量低的可能原因之一是DII数据库收录我国国家知识产权局的发明专利申请数据中不包含引用信息,从而导致我国专利权人对中国科学院专利的引用无法体现在DII数据库中,致使数值相对偏低。本报告的对比数据仅以数据库著录为依据,请谨慎参考。

4.3 基于技术主题的专利权人网络分析方法

4.3.1 基本原理

基于技术主题对专利权人之间的关系进行研究，可以在全球范围内、不同层面挖掘技术主题相近或关系紧密的专利权人，可以用于特定专利权人的竞争对手分析与挖掘，或进行预警分析。专利文献的分类体系如国际专利分类号（IPC）、美国专利分类号（USPC）、德文特专利分类号（DPC）等是分析专利技术主题的主要切入点。例如，杨璧嘉等（2008）利用网络分析方法基于IPC计算出专利之间的Pearson相关系数，再利用Pajek进行社会网络可视化，研究结果用于技术路线图研究^[57]。

本报告的研究主要基于 IPC 研究专利权人之间的关联关系。由于一件专利一般都含有一个以上的 IPC，且一个专利权人一般会有多个专利都包含某个或多个相同的 IPC，因此，本课题的研究重点考察不同专利权人拥有多少个不同的 IPC，从而分析专利权人之间的技术关联广度，而不考虑相同的某个 IPC 所涉及的专利件数。具体原理如表 7、表 8、图 15 所示。

表7. 专利权人及其所有专利的 IPC 统计矩阵

IPC	专利权人 1	专利权人 2	专利权人 3	……	专利权人 n
IPC1	30	28	24		1
IPC2	26	25	23		0
IPC3	25	0	5		1
……					
IPCn	1	0	0		1

根据研究需要，将表 7 中的矩阵值转换为 1-0 型矩阵（表 8），只分析某专利权人是否涉及到某 IPC 号，而忽略涉及到某 IPC 号有多少件专利。

表8. 专利权人及其所有专利的 IPC 统计 1-0 型矩阵

IPC	专利权人 1	专利权人 2	专利权人 3	……	专利权人 n
IPC1	1	1	1		1
IPC2	1	1	1		0
IPC3	1	0	1		1
……					
IPCn	1	0	0		1

假如分析对象仅包括表 8 中的 3 个专利权人和 3 个 IPC，那么所获得的基于 IPC 的专利权人关联网络则如图 15 所示。

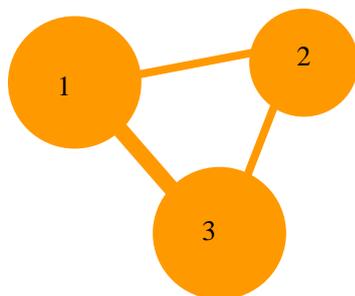


图15. 基于 IPC 的专利权人关联网络示意图

4.3.2 分析方法设计与内容

(1) 技术主题的设计依据

技术主题可以依据 IPC、DPC、USPC 以及日本专利局、欧洲专利局等对专利文献进行技术主题分类的专利分类号进行研究,本课题研究依托于 DII 数据库,可以选择 IPC 和 DPC 进行分析。IPC 又可以选择小类、大组或小组进行研究,对同一数据集而言,基于小类的专利权人关联网络的网络规模大,但专利权人之间的关联权重值低,而基于小组的关联网络网络规模相对较小,但专利权人之间的关联权重值高,更适于挖掘学科重合度较高的专利权人,适用于竞争对手发觉分析。

还可以根据专利权人 IPC 的相似度计算,通过设定阈值,决定不同专利权人之间的相似度,构建关联网络。

此外还可以从主题词的角度出发进行聚类或相似度测算,进而构建基于主题词的关联网络。

(2) 数据集的构建

首先,全球范围的基于技术主题的专利权人关联网络,然而,这是一项非常庞大的工程,但此项研究仍具价值,可以从整体上反映出国际技术创新的机构之间的关联关系,迄今还未见此方面的研究成果。

其次,研究特定专利权人的基于技术主题的关联网络,如不同性质的专利权人(如高校、企业、科研院所)之间的基于技术主题的关联关系。

最后,此方法不太适合特定学科主题分析,理由是特定的学科主题必然在专利分类上集中于某些专利分类号,必然形成一个关系紧密的关联网络,挖掘价值难以体现。

(3) 分析指标

同 4.1.3 (2) ② 章节。

(4) 分析实例

① 基于小类的关联网络

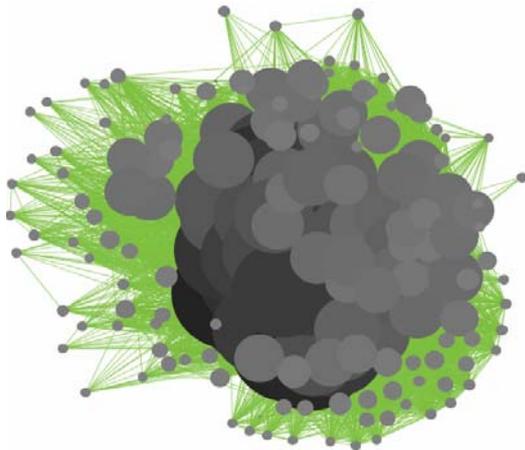
本课题以分析中国科学院院属各研究所之间专利产出主题上的相近性为实例,首先基于院属研究所专利所涉及的 IPC 小类构建研究所之间的关联网络,从而反映中国科学院整体的研发布局情况。表 9 统计了 2005-2008 年中国科学院研究所基于 IPC 小类的相关关联网络参数,对应的关联网络图见图 16。

表9. 2005-2008 年中国科学院研究所基于 IPC 小类的相关关联网络参数

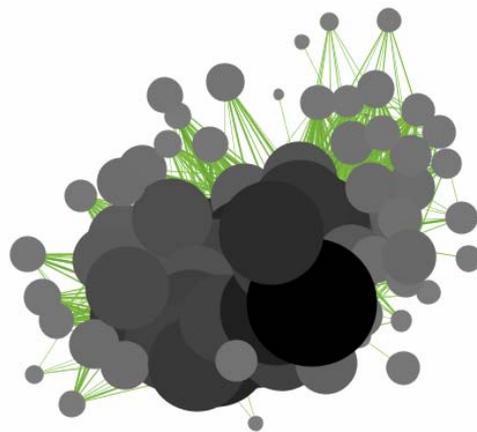
网络参数(统计信息)	相关 IPC 小类数(边的权重)					
	>=1	>=10	>=20	>=30	>=40	>=50
节点数(专利权人数)	194	86	55	39	27	19
边数(相关关系数)	9078	1463	544	269	129	50
边的最大权重值(最大相关 IPC 小类数)	70	70	70	70	70	70
边的平均权重值(平均相关 IPC 小类数)	5.50	20.54	32.30	40.96	48.75	57.08
平均度数(各机构的平均相关系数)	93.6	34.0	19.8	13.8	9.6	5.3
网络密度	0.48	0.40	0.37	0.36	0.37	0.30
网络直径	3	4	3	3	4	4
网络平均最短路径	1.51	1.64	1.71	1.66	1.76	1.86

分析发现,在中国科学院 2005-2008 年间的所有发明专利申请所涉及的 IPC 小类中,至少都涵盖两个及两个以上的专利权人,所有的研究所相互形成一个紧密联系的簇(边权重

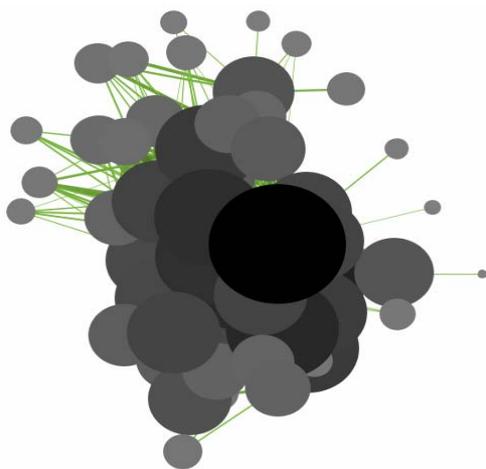
≥ 1), 网络密度高达 0.48, 研究所之间 IPC 小类相同数目最多的为 70 次, 位于中国科技大学与上海光机所之间。为了通过可视化手段揭示相互关联强度较强的研究所之间的关联关系, 本节分别对研究所之间存在相同 IPC 小类的数目 (边的权重) 在 10、20、30、40 和 50 次以上的关联网络进行可视化分析。当边权重值在 10 次以上时, 即只有当两个研究所之间存在相同 IPC 小类数目超过 10 个以上 (含 10 个) 时, 它们才呈现在关联网络中 (20、30、40 和 50 同)。从网络参数分析, 整个网络即时取到边权重值在 50 以上时, 网络仍然保持着较高的密度 (0.30), 网络直径为 4。



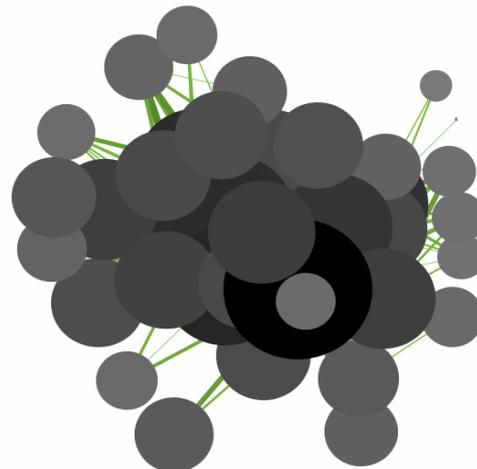
边权重 ≥ 1 (a)



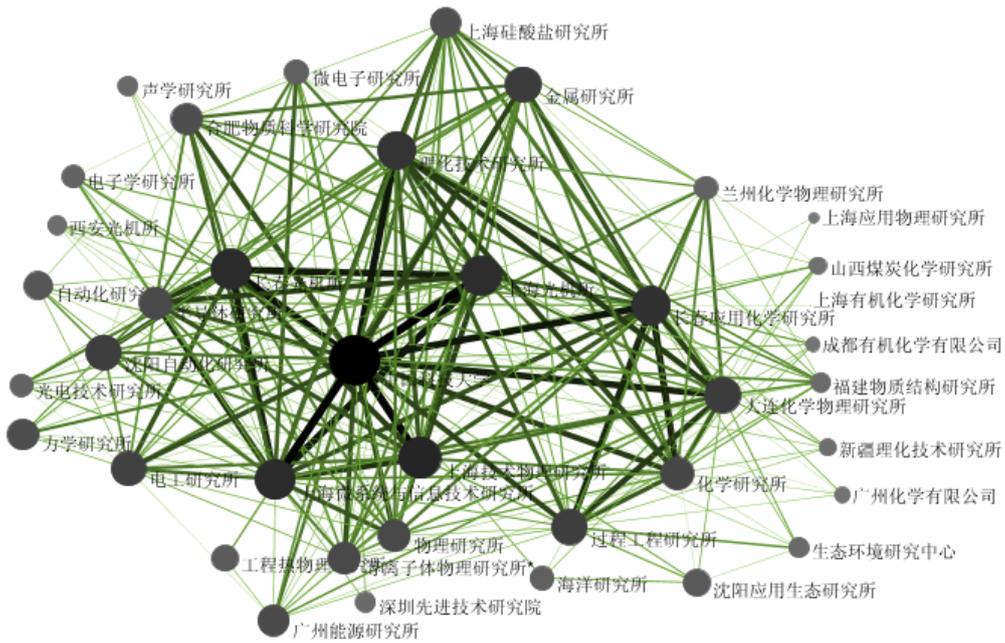
边权重 ≥ 10 (b)



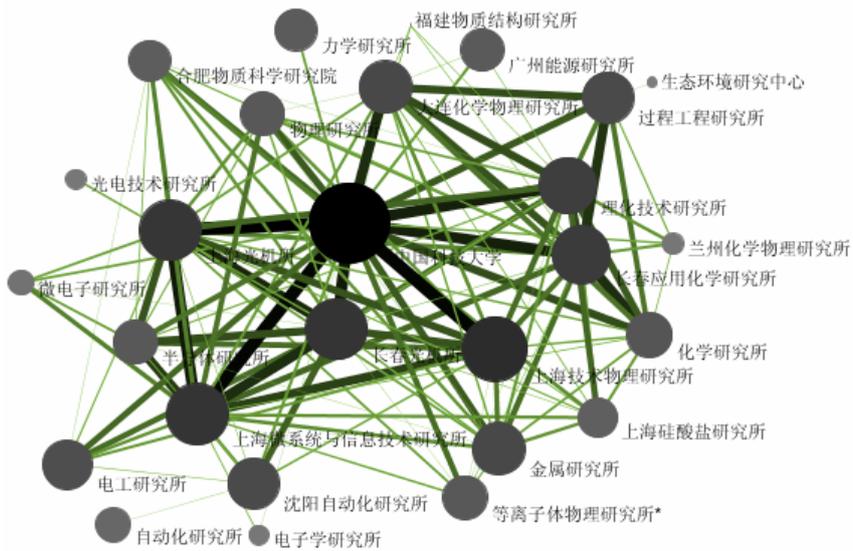
边权重 ≥ 20 (c)



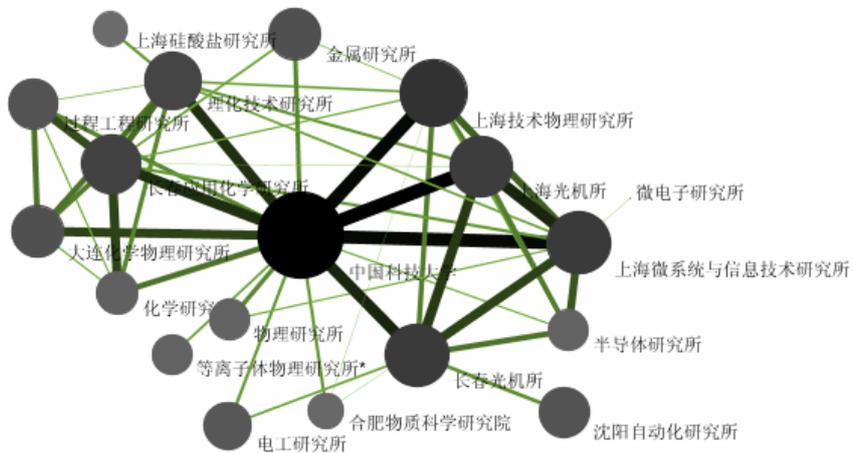
边权重 ≥ 30 (d)



边权重 ≥ 30 (e)



边权重 ≥ 40 (f)



边权重 ≥ 50 (g)

图16. 2005-2008年中国科学院研究所基于IPC小类的相关关联网

图16(e)、(f)和(g)逐渐清晰显示出,基于IPC小类的关联网中,中国科学院的信息技术、物理与化学类研究所形成相对更为集中的研发布局,从中国科学院整体水平考虑,中国科学院在上述领域形成了具有较强研发基础的能力。

从整个网络结构分析,中国科技大学研究涉及领域最广,是整个网络的核心,也是一个中间连接点。与其关联强度较大的研究所形成两个主要的簇,其一是包括上海技术物理研究所、上海光机所、上海微系统与信息技术研究所、长春光机所和半导体研究所在内的簇,主要涉及信息技术领域;另一个是包括理化技术研究所、长春应用化学研究所、过程工程研究所、大连化学物理研究所和化学研究所在内的簇,主要涉及物理和化学领域。

图16(e)、(f)和(g)揭示了网络中度数较大(连线较多,即与相对更多的研究所存在相同的IPC小类)的研究所的分布情况,中国科技大学、上海光机所、长春光机所、上海微系统与信息技术研究所、上海技术物理研究所、长春应用化学研究所、大连化学物理研究所、化学研究所、过程工程研究所成为网络的几个中心点。说明在中国科学院的各研究所内,这些研究所的研究领域涉及面相对较广。

② 基于大组的关联网

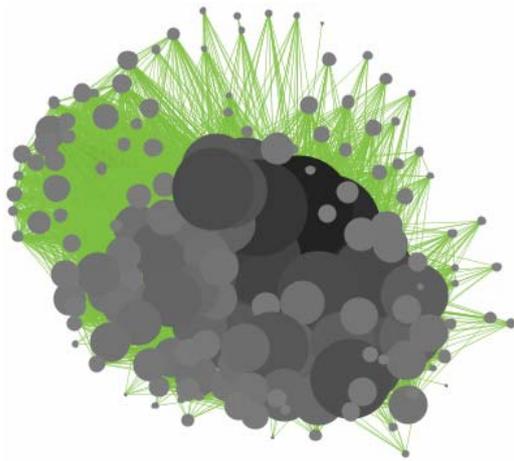
由于IPC小类相对分类较为宏观,为了从具体的技术方向考察中国科学院院属各研究所之间的相互关联关系,本节基于IPC大组构建中国科学院院属研究所的关联网。根据机构间拥有相同的IPC大组数目(边的权重值)的多少,本节分别构建相关IPC大组数在1、10、20、50、100次以上的各研究所之间的关联网。

表10. 2005-2008年中国科学院研究所基于IPC大组的相关关联网参数

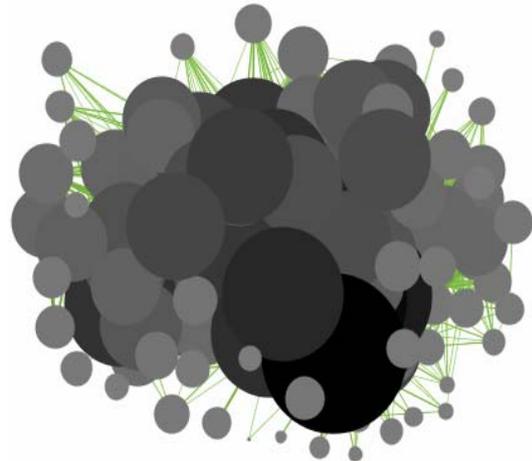
网络参数(统计信息)	相关IPC大组数(边的权重)				
	>=1	>=10	>=20	>=50	>=100
节点数(专利权人数)	187	97	72	37	10
边数(相关关系数)	7190	1554	739	174	22
边的最大权重值(最大相关IPC大组数)	196	196	196	196	196
边的平均权重值(平均相关IPC大组数)	7.98	26.44	40.61	73.61	121.5
平均度数(各机构的平均相关关系数)	76.90	32.04	20.53	9.41	4.4
网络密度	0.41	0.33	0.29	0.26	0.49
网络直径	3	4	4	4	2
网络平均最短路径	1.60	1.73	1.86	1.95	1.51

表10统计的网络参数显示,中国科学院院属研究所基于IPC大组的关联网密度值为0.41,机构间的最大相关IPC大组数为196,发生于化学研究所和长春应用化学研究所之间。当边权重值逐渐提高后,网络密度依旧维持较高的数值,尤其当边权重值超过100的关联网的密度达到0.49。

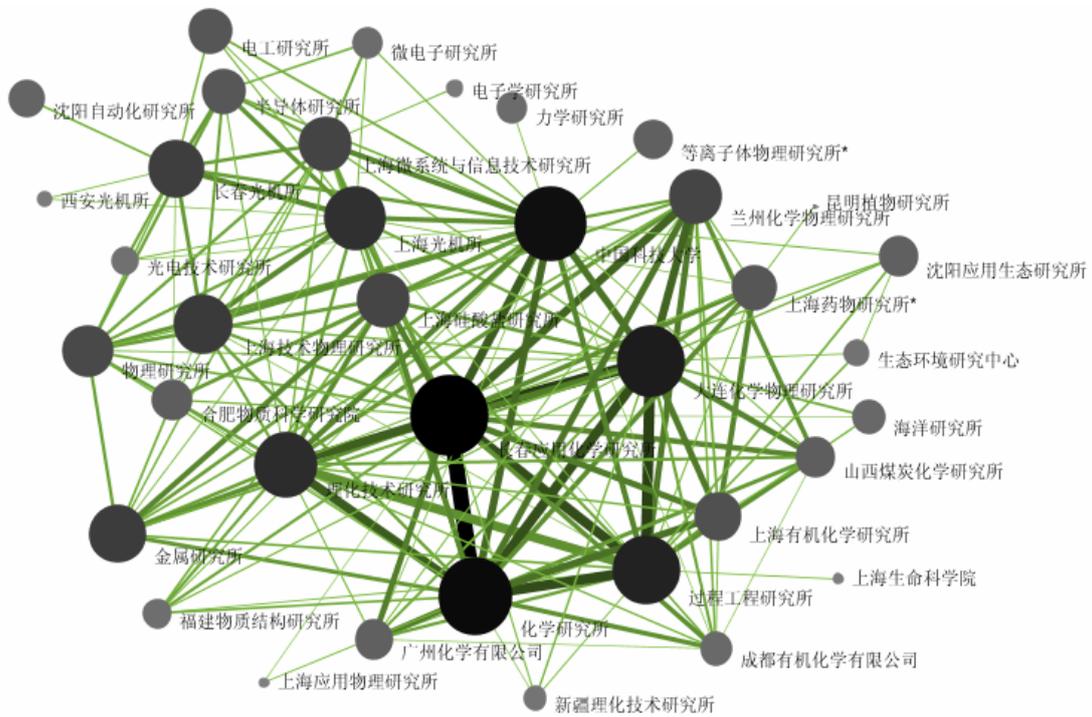
图17显示,从边的权重值大于或等于1、10以及20的关联网分析,与基于IPC小类的分析相似,中国科学院各研究所之间基于IPC大组的相似性构建成一个紧密相连的网络,网络密度达到0.41(边权重大于或等于1时)。边权重值大于或等于50的关联网反映的特征与基于IPC小类的关联网类似,关联强度较高的研究所主要分布在信息技术、物理和化学领域的研究所。然而边权重值大于或等于100的关联网反映的特征与基于IPC小类的关联网有差别,相互间关联最强的10个研究所都是与化学相关的研究所(含中国科技大学)。其中化学研究所、长春应用化学研究所、大连化学物理研究所和过程工程研究所形成了关联网的最强核心。



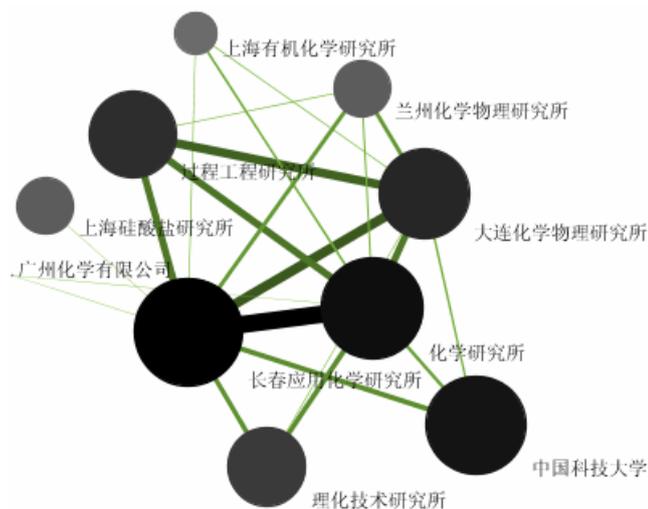
边权重 ≥ 1 (a)



边权重 ≥ 10 (b)



边权重 ≥ 50 (e)



边权重 ≥ 100 (f)

图17. 2005-2008年中国科学院研究所基于IPC大组的相关关联网络

4.3.3 功能与作用

基于本节针对基于技术主题的专利权人关联网络的构建原理,结合对中国科学院院属研究所的基于IPC的关联网络分析,中国科学院院属各研究所在技术研发主题上存在较强的关联关系,即说明它们在技术研发主题上存在较大的相似性。然而,通过机构合作网络分析(4.1.2(3)和4.1.2(4))却反映出院属研究所之间以及与院外机构的合作强度较弱,表明我院各研究所在一些相同或相似的技术主题上各自相对独立地在进行研发工作,院内研究所之间,以及与院外研发机构之间开展的合作较少。

由此可见,基于技术主题的专利权人关联网络分析和专利权人合作网络分析的联合应用,对分析特定专利权人集合的潜在竞争合作关系,具有直接的参考借鉴价值。

5 专利权人网络分析案例研究-以中国科学院为例

案例研究以Derwent Innovations IndexSM数据库收录的中国科学院发明专利数据为数据基础,从以下两个层面上研究从1985年到2009年的逐年专利权人网络,并对各年网络进行纵向比较研究:①将整个中国科学院作为一个节点,研究其与国内专利权人(大学、公司企业、科研院所等,统称“合作者”)的ego合作网络;研究其与国外专利权人的合作网络;②以院属各研究所为节点,研究各研究所之间及与国内外专利权人的global合作网络。

5.1 中科院专利权人 ego 合作网络

中科院 Ego 合作网络以整个中国科学院作为一个节点,研究其与院外专利权人之间的合作网络及演进情况。

5.1.1 中科院发明专利申请数量增长趋势

图18(a)展示了中科院逐年发明专利申请数的发展情况,1985-2009年间,中科院各年发明专利申请数量总体上呈现出逐年递增的态势,年均增长率达到23%,特别是进入1998年以后,增长趋势更为显著,年均增长率达到28%,而1985-1997年之间的年均增长率仅为17%。1998年以后中科院发明专利数量迅速增长的原因可能有两个:首先,我国机构近年来对专利申请的重视度逐渐提高,第二,受到国外机构在我国进行专利布局的竞争冲击,最后,中科院在1998年开始实施的知识创新工程极大地推进了其创新能力的提升与科研产出

的增加，这也许是 1998 年以后中科院发明专利申请数量快速增长的最重要的原因。

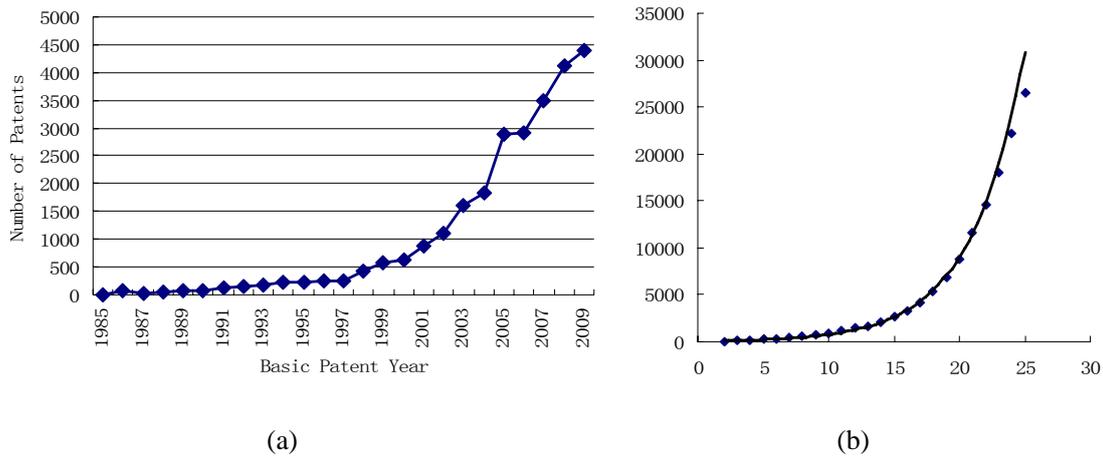


图 18. 中科院专利申请数增长趋势：1985-2009，a-逐年；b-累积

图 18 (b) 显示了中科院逐年累积的发明专利数量，采用非线性回归得到其累积增长曲线的方程为：

$$F(t)=67.7e^{0.25t} \quad (1)$$

其中 $F(t)$ 代表中科院从 1985 至 2009 年间的发明专利累积数量 $t=2$ 代表 1986 年 ($r^2=0.9942, t>1$)。由于中科院在 1985 年没有专利公开，因此曲线中不含 1985 年，曲线 b 表明中科院的发明专利逐年累积演进情况符合指数增长。

5.1.2 合作者的增长

研究发现，发明专利数量的增长特征与合作者的增长特征不同，最典型的一个特点是，合作者可以在首次出现后的年份里再次出现，但在演进网络里则不重复计算。

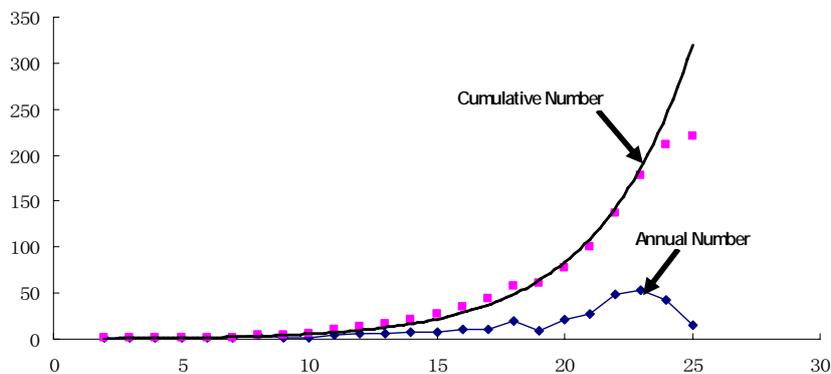


图 19. 中科院发明专利合作者数量增长趋势图：1985-2009

图 19 演示了 1985 至 2009 年中科院专利合作者的年度数量 (annual) 及逐年累积 (cumulative) 数量的变化情况，利用非线性回归得到合作者累积增长曲线的方程为：

$$C(t)=0.3847e^{0.27t} \quad (2)$$

其中 $C(t)$ 代表中科院从 1985 至 2009 年间的发明专利合作者逐年累积数量， $t=2$ 代表 1986 年 ($r^2=0.9718, t>1$)。由于中科院在 1985 年没有专利公开，因此曲线中不含 1985 年，cumulative 曲线表明中科院的发明专利合作者逐年累积演进情况符合指数增长。同时我们发现，2008 和 2009 年合作者的数量呈现不规律地减少，并偏离了指数曲线。导致这种非正常偏离的最可能原因是一些潜在的合作者并非是参与到专利研发过程的合作者，而是接受中科院发明专

利技术转移转化的权利受让人，理论上讲，只要专利权有效，就可能被转移或转化。然而，我们查阅了一些发生专利权转移的专利文献发现，专利权的转移大多发生在发明专利公开与授权之间、或在授权后的最初几年内，这意味着 2008 年和 2009 年的一些发明专利由于尚未授权，一些潜在的权利受让人还未出现，进而导致这两年合作者数量的减少。

5.1.3 合作网络的稠化与增长

Bettencourt Kaiser和Kaur（2009）指出，随着学科领域的发展，其研发合作网络也会变得越来越稠密，这意味着随着时间的发展，平均每个节点（作者）拥有的连接数逐渐增加，节点数与边的关系符合下面简单的标度律，其中标度指数 $\alpha>1$ ^[58]。

$$\text{edges} = A(\text{nodes})^\alpha, \quad (3)$$

其中 A 和 α 是常数，标度指数 α 与网络规模（节点数目）无关。

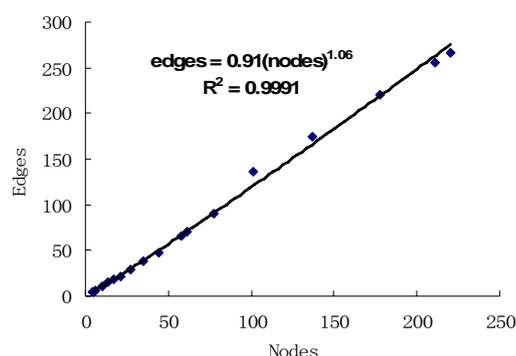


图20. 中科院 ego 合作网络的稠化标度律

在本报告实例研究中的数据（图 19）显示，随着时间的发展，合作者的数目逐渐增加，那么，是否中科院专利权人 ego 合作网络也表现出类似的标度律（ $\alpha>1$ ）？为了回答这问题，我们分析了中科院 ego 合作网络中节点数与边的关系（如图 20 所示）。其中节点代表整个中科院及其合作者，标度指数 $\alpha=1.06$ 。该值较低，说明中科院 ego 合作网络稠化进程较为缓慢，即随着合作者加入 ego 合作网络，我们并未检测到边数的更快速增长。由于在 1991 年之前没有合作者，数据起始年份为 1992 年。

5.1.4 网络直径

在 ego 合作网络中，中科院位于网络的中央并且与所有合作者都存在直接连接关系，网络没有孤立点，所以网络的直径（指任意两节点间的最长距离边数）始终为 2。

5.2 中科院专利权人 global 合作网络

中科院自 1949 年创建以来，历经 60 年的发展，许多研究所发生过改名、重组或新建等情况，目前已拥有 120 余家研究所，本报告为了分析研究所合作网络的演进情况，特对改名、重组的情况统一根据目前名称进行清洗，最终获得 125 个研究所。

5.2.1 研究所与合作者的增长

图 21 显示了中科院发明专利 global 合作网络中专利权人（包含院属研究所和合作者）数量增长趋势图，期与图 19 的中科院 ego 合作网络合作者数量增长特征总体一致。利用非线性回归得到专利权人累积增长曲线的方程为：

$$C(t)=26.47e^{0.11t} \quad (4)$$

其中 $C(t)$ 代表中科院从 1985 至 2009 年间的专利权人逐年累积数量， $t=2$ 代表 1986 年

($r^2=0.9803$, $t>1$)。Cumulative曲线表明中科院的发明专利专利权人逐年累积演进情况符合指数增长。

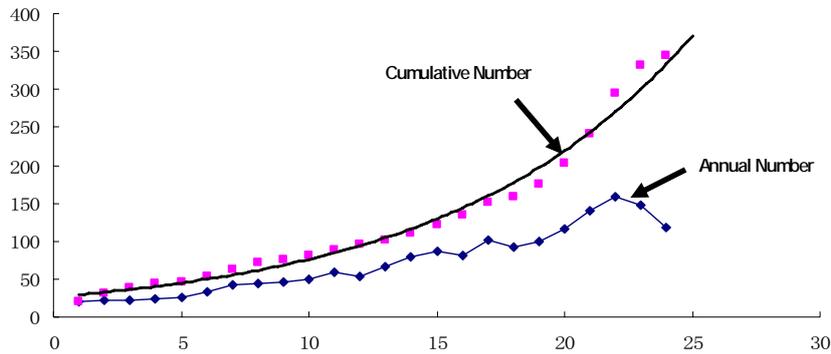


图21. 中科院发明专利专利权人（研究所+合作者）数量增长趋势图：1985-2009

然而，global 合作网络的专利权人数增长曲线与 ego 合作网络的合作者增长曲线相比则显得更加平缓，这种差异源自 1985 年以来有专利产出的研究所数量的增长的影响。通过分析院属研究所的数量增长曲线发现（图 22），院属研究所逐年累积数量增长曲线是非线性的对数增长曲线，方程为：

$$C(t)=40.115\ln(t)-20.035 \quad (5)$$

其中 $C(t)$ 代表中科院从 1985 至 2009 年间的研究所逐年累积数量， $t=2$ 代表 1986 年 ($r^2=0.9479$, $t>1$)。这与合作者累积数增长曲线完全相反。

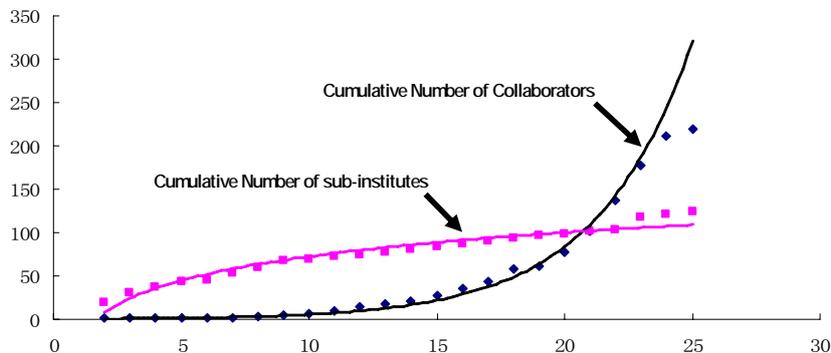


图22. 中科院发明专利院属研究所与合作者数量增长趋势比较图：1985-2009

图 22 清晰显示，1986 年中科院仅有 20 家研究所有专利申请，这 20 家也是中科院最早申请专利的 20 家研究所，在随后的几年内这个数字迅速增加到 1991 年的 53 家。然而，1992 年以后，拥有专利产出的研究所的数量则呈现出缓慢增长的态势。事实上，研究所的数量也不可能保持始终增长，迄今为止，几乎中科院所有的研究所都有专利产出，以后这个数量的增长将完全依赖于新建研究所，例如在 2007-2009 年间，研究所累积数量均超过了对数曲线，主要原因都是新建研究所造成的，如深圳先进技术研究院、苏州纳米技术与纳米仿生研究所等。可以预见，新建研究所也不可能不停增长，因此，如果 global 合作网络依旧想保持专利权人数量的累积增长曲线呈现指数增长态势，将依赖于不断有新的合作者的不断增加。

5.2.2 Global 合作网络的稠化与增长

利用公式 (3) 的标度律测度中科院 global 合作网络中节点与边的关系（图 23），节点

代表研究所和合作者，数据起始年份为 1992 年。

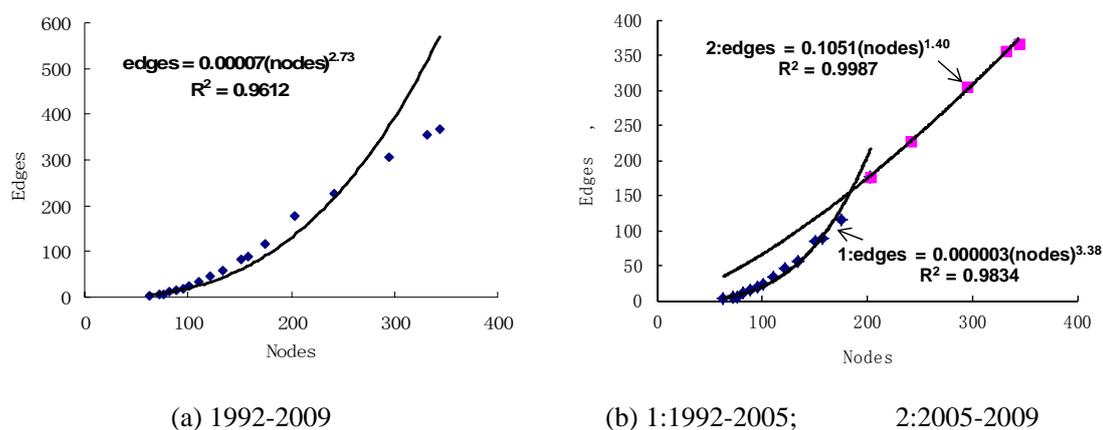


图23. 中科院 global 合作网络的稠化标度律

图 23(a)显示 global 合作网络的标度指数高达 $\alpha=2.73$ ，说明在中科院发明专利 global 合作网络中，随着有专利产出的研究所数量以及合作者数量的不断增加，专利权人之间的合作行为增长的相对更为快速。Global 合作网络的标度指数($\alpha=2.73$)显著高于 ego 合作网络的标度指数($\alpha=1.06$)，这意味着研究所之间的合作增长速率高于研究所与合作者之间的合作。

然而，图 23(a)的指数回归曲线与真实数据拟合度较低 ($R^2=0.9612$)，分析发现在 2005 年存在一个明显的拐点，为了观察 2005 年前后的网络增长特征的差异，图 23(b)分别对 1992-2005 与 2005-2009 的标度律进行了分析，标度指数分别为 $\alpha=3.38$ 和 $\alpha=1.40$ 。数据说明在 2005 年以前中科院 global 发明专利合作网络稠化的速率非常快，而 2005 年以后这种稠化现象则明显放缓。原因可能包括：首先，研究所的数量只有 125 个，研究所之间的合作边数不可能总是持续增长，尽管某些边可以变得越来越强；其次，最近几年节点数的增加主要依赖于新建研究所和新的合作者，因此在缺乏研究所之间合作边增加的支撑下，主要由新建研究所和新合作者的增加所带主导的节点的增加则相对抑制了 α 值。

5.2.3 网络直径

Leskovec 等 (2005) 的研究发现^[59]，随着网络的增长网络直径倾向于降低，直径的降低意味着合作网络越来越稠密。Bettencourt 等 (2009) 发现科技领域的合作网络在初期直径快速增长，然后稳定在 12-14 左右。然而，在中科院 global 发明专利合作网络中，网络直径从 1994 年的 1 增加到 2009 年的 16，迄今尚没有明显迹象表明直径是否会继续增长或保持在现有水平维持不动，或逐渐降低。基于网络直径保持稳定不变或降低将代表网络的转型这一理论假设，那么我们尚无法确定中科院的 global 发明专利合作网络什么时候会发生网络转型。

5.2.4 中科院 global 合作网络的特征

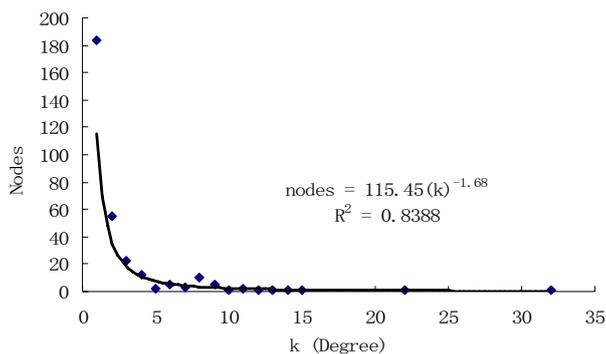


图24. 中科院 global 发明专利合作网络节点度分布特征（1985-2009）

历经 25 年的发展,中科院有越来越多的研究所与其他研究所或合作者建立了合作关系,截止 2009 年,总计有 125 家研究所拥有发明专利申请,总计有 219 个合作者。这 344 个专利权人构建了最终现有的中科院发明专利 global 合作网络,其中包括一个主成分保护 252 个节点,2 个包含 6 个节点的弱成分,2 个包含 4 个节点的弱成分,7 个包含 3 个节点的弱成分,7 个包含 2 个节点的弱成分,以及 37 个孤立节点(均是研究所)。数据表明中科院有 30%的研究所虽然申请了专利,但在专利活动中与院内其他研究所及院外研究机构均无合作关系。最终合作网络节点与度数分布满足幂律分布特征(见图 24)。

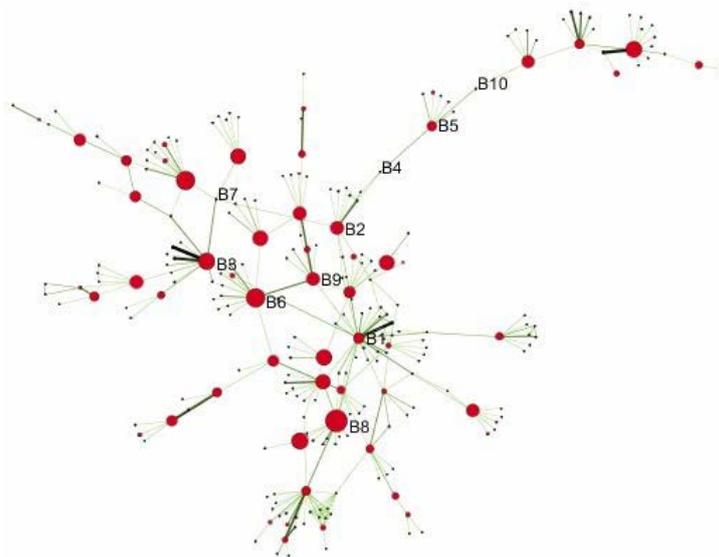


图25. Global 合作网络主成分网络图（252 节点）

图 25 显示了 global 合作网络的最大的主成分,包含 252 个节点,其中 60 个(占 125 个研究所数量的 48%)是院属研究所,192 个为院外合作者。通过计算网络的中间度(betweenness),在网络中标注了中间度最大的 10 个节点 B1-B10,其中有三个节点为院外合作者,它们分别是 B4、B7 和 B10,其他 7 个则是专利产出数量也位居前列的院属研究所。这些高中间度的节点对网络的演化起到了重要的调控作用,特别是院外的 3 家合作者对我院发明专利合作网络的贡献度不容忽视。在未来的工作中,我们将深入剖析这些节点的促进网络演化中所发挥的作用。

6 讨论

社会网络分析在专利分析中的应用是对社会网络分析方法学的有益拓展,可以预见,将社会网络分析方法应用于专利分析中具有重要意义,根据本文的综述,下面提出未来专利社会网络分析的研究重点将至少包括以下几个方面,供讨论:

首先,用于鉴别关键发明人或专利权人、发掘竞争对手。通过专利合作、技术主题关联和引用网络分析,可以发现处于网络重要位置的专利权人或发明人,这些重要的节点可以是中心(度数中心或中间中心)节点,也可以是桥点。综合这些信息,可以描绘发明人或专利权人的竞争力,用于构建更有创新能力的研发团队。特别是对专利文献所承载的专利权的所有者——专利权人的分析目前还相对较少,这将是非常广阔的研究领域,通过专利权人之间的关联网络分析,可以用于发现竞争对手^[60]进行竞争分析,或为创建联合研发项目寻找合作伙伴提供参考。此外,高被引专利可以用来评价专利丛林,引文网络中关系近节点说明其技术相关,若在引文网络中相连的两个节点在合作网络中没有合作,说明二者存在高度竞争关系,若有合作则说明二者在联合研发新技术。

其次,专利社会网络分析在企业创新与管理研究中得到越来越多的关注,包括企业之间的竞争合作以及企业内部发明人之间的合作网络结构对创新的影响等。例如Paruchuri等(2010)通过企业之间的专利合作网络以及企业内部的发明人合作网络分析发现,公司内部发明人创新合作网络的发明人结构中心性与其对公司创新性的作用呈现倒“U”型关联关系,这种关系同时又受到公司所处的合作网络的中心性及结构洞的负调节^[61]。此外,Su等(2009)采用专利家族的优先权专利网络对公司的专利组合(patent portfolio)进行了分析,服务于企业技术保护与实现战略目标^[62]。

第三,基于论文和专利之间的引用关系探讨科学向技术的转移仍将是一大研究热点。现有研究大多数尚停留在计量和统计水平,而缺少在网络结构和演进动力学水平考察科学向技术的知识转移的研究,这将是社会网络分析在专利分析中的重要潜在应用领域之一。

第四,更多的专利数据来源将作为专利社会网络分析的数据基础。现有有关专利网络分析工作大多数都是基于美国专利商标局(USPTO)的专利数据开展的,基于欧洲专利、日本专利、中国专利以及德温特专利的专利社会网络分析工作还比较少,属于待发掘领域。

最后,专利网络动力学及演进分析将是今后发展的重点之一,特别是专利引用网络以及专利权人合作网络动力学及演进分析将是今后的研究热点之一。近期已经发表的几篇文章中也正体现了这种趋势,包括Barberá等(2011)基于专利引用网络对技术演进(主路径)进行的研究^[63],Lee等(2009)基于ICT领域专利数据开展的协同进化研究^[64],Kegler等(2010)对机构合作网络演进的分析^[65],等等。技术演进路径识别的算法研究与应用也将是专利演进分析的一个热点方向,如常见的搜寻路径连接数目(SPLC)算法和搜寻路径节点对数(SPNP)算法等,刘倩楠等(2010年)利用SPNP算法对以太网技术演进路径进行了研究^[66]。

7 参考文献

- ¹ 林聚任, *社会网络分析: 理论与应用*. 第1版 ed.; 北京师范大学出版社: 北京, 2009.
- ² Otte, E., Rousseau, R. Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*, 2002, 28(6): 441-453.
- ³ 李金华. 网络研究三部曲:图论、社会网络分析与复杂网络理论. *华南师范大学学报(社会科学版)*, 2009, (02):136-138.
- ⁴ 原著John Scott. 刘军 译. 社会网络分析. 重庆大学出版社, 重庆, 2007.
- ⁵ 原著John Scott. 刘军 译. 社会网络分析. 重庆大学出版社, 重庆, 2007.

-
- ⁶ Otte, E., Rousseau, R. Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*, 2002, 28(6): 441-453.
- ⁷ 王俨, 郭婕婷, 肖国华. 社会网络理论在专利引用中的应用. *情报理论与实践*, 2008, (03):364-366.
- ⁸ Wartburg I, Teichert T, Rost K. Inventive progress measured by multi-stage patent citation analysis[J]. *Research Policy*, 2005, 34:1591-1607.
- ⁹ Gao X, Guan J. Networks of scientific journals: An exploration of Chinese patent data[J]. *Scientometrics*, 2009, 80(1):283-302.
- ¹⁰ Gress B. Properties of the USPTO patent citation network 1963-2002[J]. *World Patent Information*, 2009: In press.
- ¹¹ Sternitzke C, Bartkowski A, Schramm R. Visualizing patent statistics by means of social network analysis tools[J]. *World Patent Information*, 2008, 30(2):115-131.
- ¹² Dou H, Bai Y. A rapid analysis of Avian Influenza patents in the Esp@cenet database- R&D strategies and country comparisons[J]. *World Patent Information*, 2007, 29(1):26-32.
- ¹³ Sternitzke C, Bartkowski A, Schramm R. Visualizing patent statistics by means of social network analysis tools[J]. *World Patent Information*, 2008, 30(2):115-131.
- ¹⁴ 陈云伟, Katy, Borner. 科学地图的科学价值[J]. *图书情报知识*, 2009, (6):27-33;74.
- ¹⁵ 陈云伟, Katy, Borner. 科学地图的科学价值[J]. *图书情报知识*, 2009, (6):27-33;74.
- ¹⁶ Sternitzke C, Bartkowski A, Schramm R. Visualizing patent statistics by means of social network analysis tools[J]. *World Patent Information*, 2008, 30(2):115-131.
- ¹⁷ Sternitzke C, Bartkowski A, Schramm R. Visualizing patent statistics by means of social network analysis tools[J]. *World Patent Information*, 2008, 30(2):115-131.
- ¹⁸ Paruchuri S. Intraorganizational Networks, Interorganizational Networks, and the Impact of Central Inventors: A Longitudinal Study of Pharmaceutical Firms[J]. *Organization Science*, 2010, 21(1):63-80.
- ¹⁹ Su FP, Lai KK, Sharma RRR, et al. Patent Priority Network: Linking Patent Portfolio to Strategic Goals[J]. *Journal of the American Society for Information Science and Technology*, 2009, 60(11):2353-2361.
- ²⁰ Barberá-Tomás D, Jiménez-Sáez F, Castelló-Molina I. Mapping the importance of the real world: The validity of connectivity analysis of patent citations networks[J]. *Research Policy*, In Press, Corrected Proof.
- ²¹ Lee S, Kim M-S, Park Y. ICT Co-evolution and Korean ICT strategy--An analysis based on patent data[J]. *Telecommunications Policy*, 2009, 33(5-6):253-271.
- ²² Kegler MC, Rigler J, Ravani MK. Using network analysis to assess the evolution of organizational collaboration in response to a major environmental health threat[J]. *Health Education Research*, 2010, 25(3):413-424.
- ²³ 刘倩楠, 基于专利引文网络的技术演进路径识别研究[D]. 硕士, 大连理工大学, 2010.
- ²⁴ 约翰·斯科特 著. 刘军 译. 社会网络分析法(万卷方法), 第二版, 重庆大学出版社, 重庆, 2007, 23 页.
- ²⁵ 艾伯特·巴拉巴西 著. 徐彬 译. 连接-网络新科学. 湖南科学技术出版社, 2007, 长沙.
- ²⁶ <http://mapofscience.com/>
- ²⁷ Boyack K W, Börner K, Klavans R. Mapping the Structure and Evolution of Chemistry Research[J]. *Scientometrics*, 2009, 79: In press.
- ²⁸ 陈云伟, 方曙, 邓勇. 基于SSCI研究世界图书情报学科布局[J]. *情报杂志*, 2008, 27(12): 97-95, 114
- ²⁹ Leydesdorff L, Rafols I. A Global Map of Science Based on the ISI Subject Categories[J]. *Journal of the American Society for Information Science and Technology*, 2009, 60(2):348-362.
- ³⁰ Börner K. Mapping the structure and evolution of science[M]. Symposium on Knowledge Discovery and Knowledge Management Tools at NIH Natcher Conference Center, Bethesda. February 6, 2006.
- ³¹ Chen C M, McCain K, White H, Lin X. Mapping Scientometrics (1981 -2001)[M]. ASIST 2002: PROCEEDINGS OF THE 65TH ASIST ANNUAL MEETING, 2002, 39: 25-34.
- ³² Hou H Y, Kretschmer H, Liu Z Y. The structure of scientific collaboration networks in *Scientometrics*[J]. *Scientometrics*, 2008, 75(2):189-202.
- ³³ Dutt B, Garg K C, Bali A. Scientometrics of the international journal *Scientometrics*[J].

Scientometrics, 2003,56(1):81-93.

³⁴ Judit B I. Informetrics at the beginning of the 21st century—A review[J]. *Journal of Informetrics*, 2008,2:1-52.

³⁵ OECD *OECD biotechnology statistic 2006*; 2006; p 157.

³⁶ Patel P *SPRU-UK performance in biotechnology-related innovation an analysis of patent data*; 2003; p 36.

³⁷ Linton K, Stone P, Wise J. Patenting trends & innovation in industrial biotechnology. *INDUSTRIAL BIOTECHNOLOGY*, 2008, 4(4):367-390.

³⁸ Barthélemy M, Flammini A. Modeling urban street patterns. *Phys Rev Lett*, 2008, 100: 138702.

³⁹ Zhang G Q, Yang Q F, Cheng S Q, et al. Evolution of the internet and its cores. *New J Phys*, 2008, 10: 123027.

⁴⁰ Colizza V, Flammini A, Serrano M A, et al. Detecting rich-club ordering in complex networks. *Nat Phys*, 2006, 2: 110-115.

⁴¹ Linton K, Stone P, Wise J. Patenting trends & innovation in industrial biotechnology. *INDUSTRIAL BIOTECHNOLOGY*, 2008, 4(4):367-390.

⁴² NWB Team. (2006). Network Workbench Tool. Indiana University, Northeastern University, and University of Michigan, <http://nwb.slis.indiana.edu>.

⁴³ Sternitzke C, Bartkowski A, Schramm R. Visualizing patent statistics by means of social network analysis tools[J]. *World Patent Information*, 2008, 30(2):115-131.

⁴⁴ Judit B, J. Informetrics at the beginning of the 21st century - A review. *Journal of Informetrics*, 2008, 2(1): 1-52.

⁴⁵ Judit B, J. Informetrics at the beginning of the 21st century - A review. *Journal of Informetrics*, 2008, 2(1): 1-52.

⁴⁶ Sternitzke C, Bartkowski A, Schramm R. Visualizing patent statistics by means of social network analysis tools[J]. *World Patent Information*, 2008, 30(2):115-131.

⁴⁷ Wartburg I, Teichert T, Rost K. Inventive progress measured by multi-stage patent citation analysis[J]. *Research Policy*, 2005, 34:1591-1607.

⁴⁸ Sternitzke C, Bartkowski A, Schramm R. Visualizing patent statistics by means of social network analysis tools[J]. *World Patent Information*, 2008, 30(2):115-131.

⁴⁹ Sternitzke C, Bartkowski A, Schramm R. Visualizing patent statistics by means of social network analysis tools[J]. *World Patent Information*, 2008, 30(2):115-131.

⁵⁰ Wartburg I, Teichert T, Rost K. Inventive progress measured by multi-stage patent citation analysis[J]. *Research Policy*, 2005, 34:1591-1607.

^[51] Criscuolo, P. & Verspagen, B. (2008). Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. *Research Policy*, 37, 1892-1908.

⁵² Jaffe, A.B., Fogarty, M.S. & Banks, B.A. (1998). Evidence from patents and patent citations on the impact of NASA and other federal labs on commercial innovation. *Journal of Industrial Economics*, 46, 183-205.

⁵³ Alcacer, J. & Gittelman, M. (2006). Patent citations as a measure of knowledge flows: the influence of examiner citations. *Review of Economics and Statistics*, 88, 774-779.

⁵⁴ Michel, J. & Bettels, B. (2001). Patent citation analysis. A closer look at the basic input data from patent search reports. *Scientometrics*, 51, 795-816.

⁵⁵ Criscuolo, P. & Verspagen, B. (2008). Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. *Research Policy*, 37, 1892-1908.

⁵⁶ 中国科学院专利分析报告 2009.

⁵⁷ 杨璧嘉, 张旭. 专利网络分析在技术路线图中的应用. *现代图书情报技术*, 2008, (05): 61-66.

⁵⁸ Bettencourt, L. M. A., Kaiser, D. I., & Kaur, J. (2009). "Scientific discovery and topological transitions in collaboration networks." *Journal of Informetrics*. 3(3): 210-221.

⁵⁹ Leskovec, J., Kleinberg, J. & Faloutsos, C. (2005). "Graphs over Time: Densification Laws, Shrinking Diameters and Possible Explanations" ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD).

⁶⁰ Sternitzke C, Bartkowski A, Schramm R. Visualizing patent statistics by means of social network analysis tools[J]. *World Patent Information*, 2008, 30(2):115-131.

⁶¹ Paruchuri S. Intraorganizational Networks, Interorganizational Networks, and the Impact of Central Inventors: A Longitudinal Study of Pharmaceutical Firms[J]. *Organization Science*, 2010,

21(1):63-80.

⁶² Su FP, Lai KK, Sharma RRK, et al. Patent Priority Network: Linking Patent Portfolio to Strategic Goals[J]. Journal of the American Society for Information Science and Technology, 2009, 60(11):2353-2361.

⁶³ Barberá-Tomás D, Jiménez-Sáez F, Castelló-Molina I. Mapping the importance of the real world: The validity of connectivity analysis of patent citations networks[J]. Research Policy, In Press, Corrected Proof.

⁶⁴ Lee S, Kim M-S, Park Y. ICT Co-evolution and Korean ICT strategy--An analysis based on patent data[J]. Telecommunications Policy, 2009, 33(5-6):253-271.

⁶⁵ Kegler MC, Rigler J, Ravani MK. Using network analysis to assess the evolution of organizational collaboration in response to a major environmental health threat[J]. Health Education Research, 2010, 25(3):413-424.

⁶⁶ 刘倩楠, 基于专利引文网络的技术演进路径识别研究[D]. 硕士, 大连理工大学, 2010.