

地理信息系统国际发展形势文献计量分析

王雪梅, 肖仙桃

(中国科学院资源环境科学信息中心, 730000 兰州市天水路中路8号)

摘要 2005年7月6日在美国科学情报研究所(Institute for Scientific Information, ISI)网络数据库(Web of Knowledge)上检索1900年至今的关于地理信息系统(Geographic Information System, GIS)的文献6988条记录。运用德文特分析软件(Derwent Analytics 3.2a)进行数据挖掘和可视化分析,从宏观的全局概括到关键词的详细剖析,用文献计量学的方法对这一学科的国际发展现状进行科学定量的分析研究,以期对该学科研究现状获得较全面客观的认识。

关键词 地理信息系统 德文特分析 文献计量学 词频 共生矩阵

1 引言

地理信息系统始于20世纪60年代的加拿大和美国,之后各国相继投入大量研究,自80年代后期繁荣至今,且不断发展变化。人们用它来进行空间数据的分析和综合处理并通过信息高速公路实现全球的数据交换和共享,强大的空间分析能力和辅助决策功能使其在社会经济、自然科学等各个领域得到广泛应用^[1-5]。21世纪,随着计算机软硬件技术的发展和数据处理与数据建库技术的进步,地球观测全球数字化、信息高速公路、网络和虚拟现实技术等将促进地理信息系统在理论研究和应用开发方面得到更大的发展,在知识经济社会中占有愈来愈重要的地位,成为解决资源、环境与人口问题的必要手段和有效方法^[6]。发达国家加快了本国基础地理信息数据库的建设,通过网络为用户提供快捷方便的信息数据服务。我国在上世纪八九十年代经历了地理信息系统的快速发展期,GIS作为一种现代化的技术迅速在各领域普及,目前虽然基本建立起数字化技术体系,但对这一领域本身深入的研究还相对滞后,今后在那些方向加强纵深研究,缩小国际差距,跟上国际发展步伐,是我们亟须了解的问题。因此有必要对这一新兴学科的国际发展现状和中国对该学科的研究在国际上所处的地位进行分析研究,在综合比较的基础上,找出我们的差距和不足,明确未来的发展方向,为我国GIS技术战略目标和制定提供科学的决策依据。

2 计量方法

通过美国科学情报研究所ISI (Institute for Scientific Information) Web of Knowledge 网络数据库^[7],以文章标题、关键词和摘要为范围,geograph* information system* 为主题词,在1900-2005年7月期间检索到关于地理信息系统的收录共6988条。ISI收录的论文既有研究成果原创性的价值体现,又反映了科学前沿的发展动态,用它所包含的科研论文来进行某学科发展形势的分析评价具有一定的说服力。应用美国汤姆森公司生产的德文特分析软件(Derwent Analytics 3.2a)^[8],进行数据挖掘和可视化分析。

自相关图(Auto-correlation Map)能显示一系列要素间的关系。例如,作者相关图能够显示共同写作的研究人员小组的关系,关键词相关图显示在同一篇文献中被同时引用的关键词具有高度的相关性。共生矩阵(Co-occurrence matrix)能显示两列要素之间的共现特性。关键词是表达文献主题概念的自然语言词汇,它们能够反映研究成果的核心内容,某学科领域研究成果的关键词的变化是该学科研究发展趋势的直接反映,通过分析关键词的变化可以全面地把握该学科发展的动态过程、特点和规律^[9]。本研究利用自相关图分析了关键词之间的相关关系,应用共生矩阵分析了国家与关键词之间的共现特性。

3 结果与分析

3.1 国际发展概况

根据主题词检索到的最早的地理信息系统文献于 1977 年出版。从图 1 可见, 文献数快速增长, 特别是 20 世纪 90 年代增长尤其迅速, 从 1983-1988 年的 99 篇跃升到 1995-2000 年的 2702 篇, 12 年间增长了 27.3 倍, 且仍在持续上升。该数据库收集的文献以论文为主, 占总文献数的 90.94%, 其次是书评、综述、社论、会议文摘等。

地理信息系统文献中被引用次数最高的是 Daly C、Neilson RP 和 phillips DI 在 1994 年 *Journal Of Applied Meteorology* 第 33 卷第 2 期上发表的 *A Statistical Topographic Model For Mapping Climatological Precipitation Over Mountainous Terrain* 一文, 到下载日期为止, 已被引用 368 次。本领域被引用达 50 次以上 (包括 50 次) 的论文 71 篇, 占 1.02%; 从未被引用的文章 2413 篇, 占总数的 34.53%。

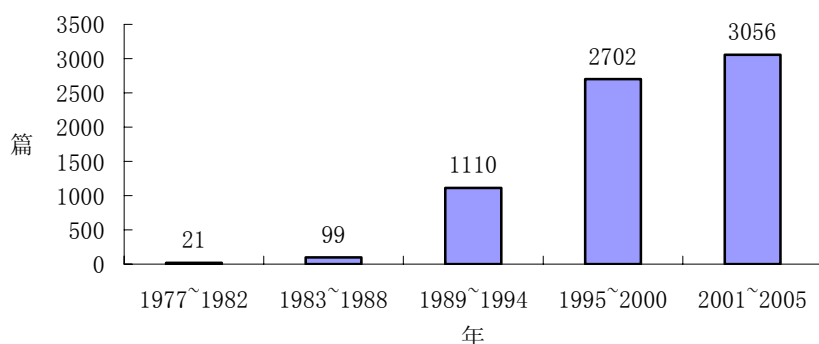


图 1 地理信息系统的 SCI 论文数量

3.2 科学影响力分析

统计得到该学科 SCI 论文数位居前 20 位的国家、机构和科研单位。论文产出量排名前十位的国家依次是美国、英国、加拿大、澳大利亚、中国、意大利、德国、荷兰、法国、印度 (表 1)。中国虽然位居第五, 但发表文章数只是美国的 1/10。

SCI 论文作者所属科研机构最多的依次是美国卡罗来纳州大学、中科院、美国德州大学、美国地质调查局、美国农业部农业研究服务中心、美国威斯康星大学、美国科罗拉多大学等 (表 2)。在前 10 个科研机构中美国有 8 个, 中国有 2 个, 中科院位居第 2。

作者所在科研单位出现次数最多的是美国俄亥俄州立大学地理系、苏格兰 Macaulay 土地利用研究所、美国圣地亚哥大学地理系、英国利兹大学地理学院、美国加利福尼亚大学圣巴巴拉分校地理系等 (表 3), 其中美国在前 20 名的单位最多, 达到 11 个, 其次是英国为 5 个 (英格兰 3 个、苏格兰 2 个)。中国有两个单位在前 20 之列, 香港大学土木工程学系排名第 11, 中科院地理科学与资源研究所排名第 14。

表 1 论文数居前 20 位的国家

排名	1	2	3	4	5	6	7	8	9	10
国家	USA	UK	Canada	Australia	Peoples R China	Italy	Germany	Netherlands	France	India
排名	11	12	13	14	15	16	17	18	19	20
国家	Spain	South Africa	Japan	Switzerland	Greece	Brazil	Sweden	New Zealand	South Korea	Finland

表 2 论文数居前 20 位的机构

排名	1	2	3	4	5	6	7	8	9	10
机构	Univ Carolina	Chinese Acad Sci	Univ Texas	US Geol Survey	USDA ARS Univ Wisconsin	Colorado State Univ	Florida	Univ S Florida	US Forest Serv	Univ Hong Kong
排名	11	12	13	14	15	16	17	18	19	20
机构	Univ Arizona	Western Oregon State Univ	Penn State Univ	Univ Calif Santa Barbara	Univ Calif Davis	Univ Minnesota	Iowa State Univ Sci & Technol	Purdue Univ	Univ Georgia	Univ Illinois

表 3 论文数居前 20 位的科研单位

排序	科研单位
1	Ohio State Univ, Dept Geog, Columbus, OH 43210 USA
2	Macaulay Land Use Res Inst, Aberdeen AB15 8QH, Scotland
3	San Diego State Univ, Dept Geog, San Diego, CA 92182 USA
4	Univ Leeds, Sch Geog, Leeds LS2 9JT, W Yorkshire, England
5	Univ Calif Santa Barbara, Dept Geog, Santa Barbara, CA 93106, USA
6	Univ E Anglia, Sch Environm Sci, Norwich NR4 7TJ, Norfolk, England
7	Purdue Univ, Dept Agr & Biol Engn, W Lafayette, IN 47907 USA
8	Georgia Inst Technol, Sch Civil & Environm Engn, Atlanta, GA 30332 USA
9	Univ Arizona, Dept Geog & Reg Dev, Tucson, AZ 85721 USA
10	Oregon State Univ, Dept Geosci, Corvallis, OR 97331 USA
11	Univ Hong Kong, Dept Civil Engn, Hong Kong, Hong Kong, Peoples R China
12	Univ Stirling, Dept Environm Sci, Stirling FK9 4LA, Scotland
13	Univ Wisconsin, Dept Forest Ecol & Management, Madison, WI 53706 USA
14	Chinese Acad Sci, Inst Geog Sci & Nat Resources, Beijing 100101, Peoples R China
15	Natl Univ Singapore, Dept Civil Engn, Singapore 119260, Singapore
16	SUNY Buffalo, Dept Geog, Buffalo, NY 14261 USA
17	Univ Arkansas, Dept Anthropol, Fayetteville, AR 72701 USA
18	Univ Calgary, Dept Geog, Calgary, AB T2N 1N4, Canada
19	Univ N Texas, Dept Geog, Denton, TX 76203 USA
20	Univ Southampton, Dept Geog, Southampton SO17 1BJ, Hants, England

3.3 研究热点分析

运用文献计量学的词频分析方法，对地理信息系统国际论文的关键词进行统计分析。把 keywords author's 和 keywords plus 合并，得到 16025 条关键词，进行同义词、近义词合并。对出现次数较多的关键词共现频次的多少进行聚类 and 映射分析，根据其亲疏关系发现该学科领域的核心选题^[10]。

从发表论文最多的前 10 位国家和前 20 个关键词的共生矩阵（表 4）可以发现，美国对该领域的研究具有明显优势，特别是对模型的研究占主导地位。英国、加拿大、中国、澳大利亚在模型研究方面也有一定的优势；美国、英国、印度对遥感方面的研究较多；在管理方面应用较多的是美国、澳大利亚、英国；在植被方面应用较多的是美国、澳大利亚、加拿大；美国和英国对分类、土地利用、土壤、栖息地、区域、流行病、景观等方向研究也较多。

对出现频次较多的前 22 个关键词（它们都在 100 篇以上的文章中出现）进行自相关分

析，从图 2 可见 geographic information systems、model、remote sensing、spatial analysis 较紧密地聚为一类，conservation、management、population、landscape、habitat、ecology、patterns、scale、dynamics、vegetation、classification 相互交错聚为一类，system、area、land-use、soil、epidemiology、simulation 等与 geographic information systems 之外的其他关键词共现的次数较低。

表 4 SCI 论文 TOP10 国家和 TOP20 关键词的共生矩阵

排 序	记录 关键词	国家	1	2	3	4	5	6	7	8	9	10
			2041 美国	540 英国	314 加拿 大	228 澳大 利亚	217 中国	157 意大 利	137 德国	128 荷兰	108 法国	94 印度
1	473	Model	161	38	24	20	23	13	3	12	8	4
2	290	Remote Sensing	78	17	12	11	11	7	4	7	4	15
3	288	Vegetation	97	12	20	23	6	7	6	2	4	3
4	203	Management	84	21	9	16	9	1	3	2	2	
5	186	Patterns	85	12	5	11	5	4	1	2	4	2
6	175	System	60	17	3	7	7	4	3	8	4	3
7	144	Conservation	45	18	8	16	2		4	1	2	5
8	137	Classification	37	17	7	5	7	4	2	3		1
9	131	Land-Use	57	15	2	2	4	2	2	1	3	3
10	130	Soil	39	13	6	6	4	4	8	5	1	2
11	124	Habitat	44	17	4	4	1	1	3			1
12	123	Area	42	16	4	3	7	2	2	2	5	
13	120	Epidemiology	40	15	5	3		4	4	1	8	
14	117	Simulation	25	6	9	4	6	5	6	2	4	3
15	116	Population	56	10	10	3		2	3		3	
16	113	Spatial Analysis	47	11	6	4	2	3	3	2	3	
17	113	Ecology	36	3	3	13	3	3	2	2	3	1
18	112	Landscape	47	12	5	3	2	3	1	2	1	
19	110	Dynamics	44	7	10	3	2	3	6	2	4	1
20	106	Scale	47	5	5	4	2	3		1	3	

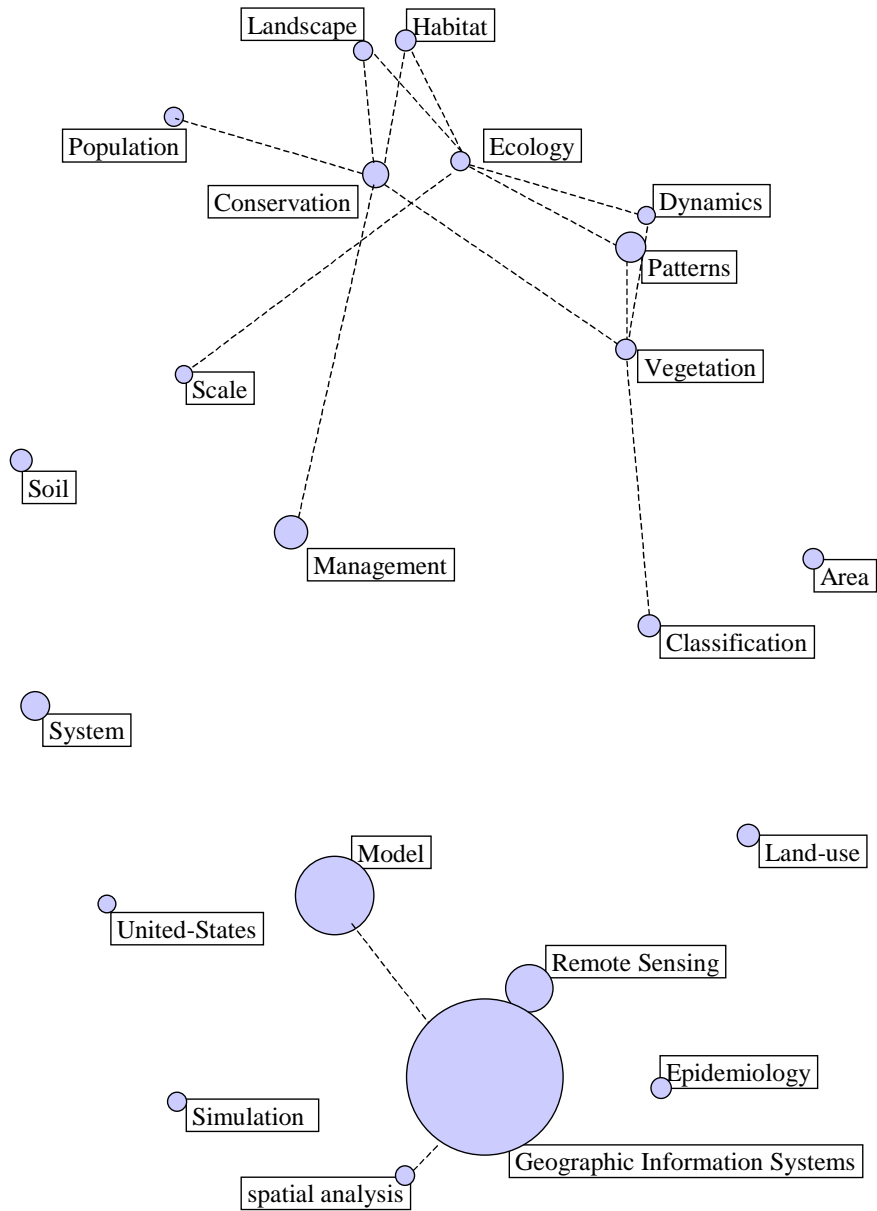


图 2 出现次数 TOP20 的关键词自相关分析图

4 结论与讨论

本文运用文献计量学的方法，对地理信息系统的 SCI 论文数、引文和被引情况做了统计分析，归纳出在该学科发表论文数位于前 20 位的国家、机构和科研单位，对关键词进行统计分析，了解该学科当代国际研究热点和中国在国际研究中所处的位置，为我国相关领域的科研人员提供有益的参考。

美国对地理信息系统的研究应用在世界上遥遥领先，有诸多科研机构、大学和科研工作者在各个领域从事相关研究，特别是在模型、管理、遥感、土地利用、人口等方面取得了比较多的研究成果。中国科学院也有很多的学者在进行地理信息系统及其应用的研究，但研究相对分散，只有地理科学与资源研究所的科研力量比较集中，发表论文数进入前 20 个科研单位的行列。从发表论文最多的前 10 位国家和前 20 个关键词的共生矩阵中可以看出，中国在模型、遥感、管理、分类方面的研究具有一定的优势。

对出现次数较高的关键词进行自相关分析，发现地理信息系统、模型、遥感、空间分

析之间的研究较紧密的结合在一起, 守恒、人口、管理、景观、生态、尺度、动力、植被、分类等相互交错聚为一类。通过对地理信息系统关键词共现分析, 展示了国际地理信息系统科技论文关键词之间的网络联系, 大体勾勒出该学科研究领域的发展现状及趋势。

研究中只考虑 ISI 数据库所录的文献, 虽然具有较强的代表性, 但没有对其他数据库数据进行搜集整理, 因此可能会失之片面。以 geograph* information system* 为主题词进行检索, 在减少误录的同时也导致了相关文献的收录缺失, 比如文章的标题、关键词和摘要中只有 GIS 简称而不用 geograph* information system* 的文献就可能被漏掉, 如何寻找一条收录全面又简单易行的方法仍有待进一步研究。

另外, 对关键词的处理还有待开展更深入的研究, 建立一套科学的关键词分类系统非常重要。信息论一般认为一个词的信息量与它的频率成反比, 一个低频词可能包含更多的信息量, 这些词往往阐明了作者所要探讨的主题^[11]。由于受到运算量的限制本文对出现次数较少的关键词没做分析, 因此可能丢失部分有用信息, 特别是一些新兴的、有潜力的研究方向的关键词。

时间序列的纵向分析将是下一步工作的重点, 以了解不同时期该学科的研究方向、重点, 和国家、机构研究核心领域的转移变化, 把握国际发展动态, 预测未来发展方向。

参考文献

- [1] Abed J, Kaysi I. Identifying urban boundaries: application of remote sensing and geographic information system technologies. CANADIAN JOURNAL OF CIVIL ENGINEERING, 2003, 30 (6): 992-999.
- [2] Tsai YC, Gao B, Lai JS. Multiyear pavement-rehabilitation planning enabled by geographic information system - Network analysis linked to projects. TRANSPORTATION RESEARCH RECORD, 2004 ,(1889): 21-30.
- [3] Zhang WC, Chen J, Ogawa K, et al. An approach to estimating evapotranspiration in the Urumqi River basin, Tianshan, China, by means of remote sensing and a geographical information system technique. HYDROLOGICAL PROCESSES, 2005, 19 (9): 1839-1854.
- [4] Brabyn L. Solutions for characterising natural landscapes in New Zealand using geographical information systems. JOURNAL OF ENVIRONMENTAL MANAGEMENT, 2005,76 (1): 23-34.
- [5] Castella JC, Trung TN, Boissau S. Participatory simulation of land-use changes in the northern mountains of Vietnam: the combined use of an agent-based model, a role-playing game, and a geographic information system. ECOLOGY AND SOCIETY , 2005,10 (1).
- [6] 赵俊三,赵耀龙. GIS发展的最新趋势及其应用前景. 测绘工程, 2000,9(2):21~25.
- [7] <http://isi01.isiknowledge.com/>
- [8] <http://thomsonderwent.com/products/dapt/derwentanalytics/>
- [9] 李文兰,杨祖国. 从关键词的变化看中国图书馆学研究主题的发展. 图书情报工作, 2004,48(12):115~118.
- [10] 谢彩霞,梁立明,王文辉. 我国纳米科技论文关键词共现分析. 情报杂志,2005,3:69~73.
- [11] 李旒. 关键词词频统计法的评价. 江苏图书馆学报, 1999(6): 20~21.