

中国信息服务业空间集聚研究

惠洁, 李晓琳

(成都理工大学管理科学学院, 成都 610059)

摘要:通过产业集聚的 EG 指数对中国各省份信息服务业的空间集聚程度进行研究, 得出集聚程度比较显著。在此基础上, 找到这一服务业的两个相关指标, 利用产业集中度 CR_n 判断出信息服务业主要集中于北京、广东等六个地区, 然后对两指标数据进行无量纲化处理, 加权求和得到新的综合指标, 最后对这一新指标进行了系统聚类分析, 得知这一产业主要集中在北京和广东两个地区。

关键词:空间集聚; EG 指数; 产业集中度 CR_n ; 系统聚类

中图分类号:C81

文献标志码:A

引言

目前, 由于科技创新和社会变革在空间上趋于集中, 我国各地区的产业发展差距在较长时间内呈扩大趋势, 学者们认为产业发展具有区域集中性的特征, 虽然近几年有所逆转但尚不稳定^[1]。随着社会的快速发展, 各个产业领域的发展已成为我们关注的焦点, 近几年来人们也开始关注生活质量问题, 随之迎来了各种新兴产业的发展, 生产性服务业的空间集聚特征也快速崛起, 不断影响着城市的发展, 人们的需求, 而信息服务业的空间集聚特征更是给高速发展的社会带来了推动作用^[2-4]。因此研究信息服务业的空间集聚特征可以更全面的促进社会的发展, 目前, 研究这一产业的学者较少, 本文在前人的研究基础上, 结合传统产业集聚理论、区域经济学和产业经济学的相关知识对我国各个省份的这一服务业进行整体研究。

1 Ellison - Glaeser 产业集聚指数

在求取产业集聚指数的时候, 一般采用的传统方法

为 EG 指数法^[5], 其数学模型为:

空间基尼系数:

$$G = \sum_i (x_i - s_i)^2 \quad (1)$$

其中, G 为空间基尼系数, 是度量行业区域集聚程度的常用系数; x_i 为 i 地区全部就业人数占全国总就业人数的比例; s_i 为 i 地区某产业就业人数占该产业全国就业人数的比例。

由式(1)可以看出, 用空间基尼系数考察产业集聚程度时, 由于没有考虑到具体的产业组织状况及区域差异, 从而不能区分是来自产业结构还是本身存在的地理集中, 最终导致结果失真。

赫芬达尔系数:

$$H = \sum_{j=1}^n z_j^2 \quad (2)$$

其中, H 为赫芬达尔系数, 是衡量市场集中度的常用系数; z_j 为 j 企业就业规模占总企业就业规模的比例。

同样, 用赫芬达尔系数进行衡量某产业的地理集中程度上存在一定的难度, 要求有每个企业就业规模数据, 而此数据不易得到。

综合以上方法的优缺点, 得出了测量产业空间集聚程度更合适的指数— γ_{EG} 。

$$\gamma_{EG} = \frac{G/(1 - \sum_i x_i^2) - H}{1 - H} = \frac{G - (1 - \sum_i x_i^2)H}{(1 - \sum_i x_i^2)(1 - H)} \quad (3)$$

其中, γ_{EG} 为地理集中指数。

用 γ_{EG} 衡量产业空间集聚程度来研究产业集聚,既考虑到了产业组织规模及区域差异,使得产业在不同区域能够真实反映集中程度,又弥补了空间基尼系数的不足。为解决 H 系数中数据不易得到的问题,对 γ_{EG} 指数进行修正。修正后的 γ_{EG} :

$$\gamma_{EG} = \frac{G/(1 - \sum_i x_i^2) - 1/n}{1 - 1/n} = \frac{G - (1 - \sum_i x_i^2)/n}{(1 - \sum_i x_i^2)(1 - 1/n)} \quad (4)$$

其中, n 表示这一产业的企业个数,通过做差验证可以知道 $H - \frac{1}{n} \geq 0$, 从而得到修正模型 EG 指数方差与原

模型 EG 指数方差之比小于 1,也就是修正模型的 EG 指数方差小于原模型 EG 指数方差。而当企业所有规模相同时,两种模型 EG 指数就相同,但是事实不然,就统计学方差检验而言,修正后的模型更优,因此,本文选用修正后的 EG 模型来衡量产业的集聚程度。

本文数据来源于 2010 - 2012 年中国统计年鉴数据,由产业集聚指数 γ_{EG} 找到 2010 - 2012 年信息服务业就业人数和单位数及全部行业就业人数和单位数两类指标,通过计算得到 γ_{EG} 值(表 1)。

表 1 2010 - 2012 年信息服务业空间集聚值

| G | | | γ_{EG} | | |
|---------|---------|----------|---------------|---------|---------|
| 2010 | 2011 | 2012 | 2010 | 2011 | 2012 |
| 0.03450 | 0.03719 | 0.039010 | 0.03752 | 0.04055 | 0.04270 |

由已有的研究成果,知道如果当 $\gamma_{EG} < 0$ 时,表示厂商存在分散趋势;当 $\gamma_{EG} = 0$ 时,不能说明厂商空间分布

表 2 2010 - 2012 年信息服务业排名前 6 位集聚地区及比重

| 2010 年集聚度 | | | | 2011 年集聚度 | | | | 2012 年集聚度 | | | |
|-----------------|----------|-----------------|----------|-----------------|----------|-----------------|----------|-----------------|----------|-----------------|----------|
| 就业人数 | | 单位数 | | 就业人数 | | 单位数 | | 就业人数 | | 单位数 | |
| $CR_6 = 49.8\%$ | | $CR_6 = 50.9\%$ | | $CR_6 = 50.7\%$ | | $CR_6 = 52.0\%$ | | $CR_6 = 51.1\%$ | | $CR_6 = 50.0\%$ | |
| 北京 | 0.224435 | 北京 | 0.109205 | 北京 | 0.230708 | 广东 | 0.104657 | 北京 | 0.236206 | 广东 | 0.101918 |
| 广东 | 0.094726 | 广东 | 0.097405 | 广东 | 0.092772 | 北京 | 0.102196 | 广东 | 0.083454 | 江苏 | 0.093707 |
| 浙江 | 0.06028 | 江苏 | 0.083842 | 浙江 | 0.059028 | 江苏 | 0.088294 | 浙江 | 0.063152 | 北京 | 0.090166 |
| 江苏 | 0.046825 | 山东 | 0.079181 | 江苏 | 0.046621 | 浙江 | 0.084073 | 江苏 | 0.048412 | 浙江 | 0.081203 |
| 上海 | 0.03606 | 浙江 | 0.078287 | 山东 | 0.039336 | 山东 | 0.079383 | 山东 | 0.040979 | 山东 | 0.075797 |
| 山东 | 0.035522 | 上海 | 0.063034 | 上海 | 0.038537 | 上海 | 0.061144 | 上海 | 0.038388 | 上海 | 0.057105 |

是完全均匀的,而是在没有任何自然优势和外溢存在时进行完全随机、独立的区域选择所呈现的一种分布状态;当 $\gamma_{EG} < 0.02$, $0.02 \leq \gamma_{EG} \leq 0.05$, $\gamma_{EG} > 0.05$ 时分别表示低、中、高三种集聚水平。

由表 1 可知,信息传输,计算机服务和软件业在 2010 - 2012 年呈现出一种中等集聚水平,但是从 γ_{EG} 可以看出 2010 年的集聚程度最低,并且集聚程度依次递增,到 2012 年集聚程度在慢慢向高集聚靠近,如果全国信息服务业发展模式保持稳步改革,在未来的几年里,这一产业会呈现出一种高集聚状态。因此有必要对此服务产业集聚状况进行研究,这样有助于国家对这一产业进行更好的规划,使之发展的越来越好。集聚是体现在某些地区性的一种发展状态,因此有必要对每个地区的集聚水平做出一定判断,从而可以对此产业的集聚做出综合分析,以便于未来信息传输,计算机服务和软件业产业有更好的发展。

2 产业集聚度

信息服务业虽然呈现的是一种中度集聚状态,但是从 2010 - 2012 年的集聚程度变化看,这一产业在未来有变为高度集聚的趋势,因此,现在需要进一步对这一产业每个地区的集聚水平^[6-8]进行判断,本文选用产业集聚集中度对每个地区进行分析,

$$CR_n = \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^N x_i} \quad (5)$$

其中, CR_n 为业中规模前 n 个地区所占的比重, x_i 为第 i 个地区产业指标值, N 为地区总数,当所取 n 值相对地区个数很小,但 CR_n 值也很高时,说明这一产业主要集中在某些地区,相应集聚水平就越高。

通过计算得到前 6 个地区的产业集聚度值见表 2。

从表2可以看出,当选取 n 为6时,2010年得到两个指标的集中度值为49.8%和50.9%;2011年分别为50.7%和52.0%;2012年分别为51.1%和50.0%。从这些集聚度知这6个地区占了31个地区信息服务业的一半,由此,可以得出这一产业主要集中在这6个地区,而且集中程度较高。就业人数的集聚度是不断增大,也就是从2010年到2012年就业人数是不断趋于集中,而单位数集聚度无明显变化趋势。

3 系统聚类

为了进一步确定这一产业在各地区的发展情况,将这一产业的单位数和就业人数作为判断指标^[9],并用R软件进行系统聚类分析,得出城市发展相似的类,并结

合相应结论情况判断出集聚情况。

聚类步骤:

(1)将该产业单位数和就业人数进行无量纲化处理,这里采用的是对数归一化:

$$x^* = \frac{\log_{10}(x)}{\log_{10}(\max)} \quad (6)$$

其中, x 为指标值, \max 为 x 中的最大值, $x^* \in [0,1]$ 。经过变换后的数据便于不同量级指标之间加权比较。

(2)由于两个指标对这一产业研究的影响都非常重要,因此,分别对两个指标赋予0.5的权重,然后得到一个综合性指标 $y = 0.5x_1 + 0.5x_2$,结果见表3,其中, x_1 , x_2 为标准化后的两指标值。

表3 各省份综合指标值

| 2010年 | | | 2011年 | | | 2012年 | | | | | |
|-------|----------|-----|----------|-----|----------|-------|----------|-----|----------|----|----------|
| 地区 | y值 | 地区名 | y值 | 地区名 | y值 | 地区名 | y值 | 地区名 | y值 | 地区 | y值 |
| 北京 | 1 | 湖北 | 0.647166 | 北京 | 0.998808 | 湖北 | 0.667936 | 北京 | 0.993952 | 湖北 | 0.668237 |
| 天津 | 0.495099 | 湖南 | 0.687393 | 天津 | 0.489155 | 湖南 | 0.702618 | 天津 | 0.55157 | 湖南 | 0.703646 |
| 河北 | 0.673569 | 广东 | 0.878637 | 河北 | 0.650552 | 广东 | 0.883014 | 河北 | 0.65769 | 广东 | 0.868742 |
| 山西 | 0.587579 | 广西 | 0.608976 | 山西 | 0.59933 | 广西 | 0.6244 | 山西 | 0.615276 | 广西 | 0.624473 |
| 内蒙 | 0.564729 | 海南 | 0.294867 | 内蒙 | 0.566986 | 海南 | 0.346076 | 内蒙 | 0.566943 | 海南 | 0.354282 |
| 辽宁 | 0.716928 | 重庆 | 0.560275 | 辽宁 | 0.731727 | 重庆 | 0.604795 | 辽宁 | 0.741469 | 重庆 | 0.608155 |
| 吉林 | 0.594416 | 四川 | 0.705286 | 吉林 | 0.600331 | 四川 | 0.697923 | 吉林 | 0.576248 | 四川 | 0.690295 |
| 黑龙江 | 0.633494 | 贵州 | 0.516483 | 黑龙江 | 0.637706 | 贵州 | 0.5106 | 黑龙江 | 0.629594 | 贵州 | 0.498518 |
| 上海 | 0.727315 | 云南 | 0.581182 | 上海 | 0.74328 | 云南 | 0.599532 | 上海 | 0.742175 | 云南 | 0.604917 |
| 江苏 | 0.776665 | 西藏 | 0.131682 | 江苏 | 0.786137 | 西藏 | 0.059455 | 江苏 | 0.795896 | 西藏 | 0.104821 |
| 浙江 | 0.807074 | 陕西 | 0.669425 | 浙江 | 0.813984 | 陕西 | 0.678852 | 浙江 | 0.822357 | 陕西 | 0.680684 |
| 安徽 | 0.611865 | 甘肃 | 0.427758 | 安徽 | 0.626348 | 甘肃 | 0.42715 | 安徽 | 0.632055 | 甘肃 | 0.454886 |
| 福建 | 0.646164 | 青海 | 0.271654 | 福建 | 0.666296 | 青海 | 0.286484 | 福建 | 0.65991 | 青海 | 0.295391 |
| 江西 | 0.539109 | 宁夏 | 0.260635 | 江西 | 0.534755 | 宁夏 | 0.265315 | 江西 | 0.583478 | 宁夏 | 0.272466 |
| 山东 | 0.736763 | 新疆 | 0.461596 | 山东 | 0.758991 | 新疆 | 0.147105 | 山东 | 0.764396 | 新疆 | 0.502637 |
| 河南 | 0.643764 | | | 河南 | 0.658308 | | | 河南 | 0.665838 | | |

(3)计算样品两两间的距离 d_{ij} ,本文选用欧氏距离进行计算,记 $D_{(0)} = (d_{ij})$ 其中, i, j 为样本序号。

(4)将距离最近的两类合并为一个新类,本文选择系统聚类法中最远距离法进行聚类,最远距离克服了链接聚合的缺陷,两类合并后加大了其它类的距离,更容易进行分类。

用R软件对这一产业2010-2012年综合指标值进行系统聚类,结果如图1-图3所示。

由2010年聚类图(图1)可知,中国各省份根据其发展情况从上至下依次定义为4类,结合标准化后的综合指标值及两指标的集中程度值(表2,表3),由于,北京和广东信息服务业标准化后的综合值为1和0.878637位于31个地区的前两个地区,且就业人数的比重分别

为0.224435和0.094726,单位数的比重值为0.109205和0.097405,这三种衡量值都位于第一和第二,因此,信息服务业这一产业主要集中在北京和广东两个地区。同样,结合2011年聚类图(图2)及表2、表3,将各地区也依次分为了4类,且每项值第一和第二都是北京或广东,因此,2011年的信息服务业也主要是集中在北京和广东两个地区。而2012年聚类图(图3),北京、广东、江苏和浙江归为一类,从表2和表3可知这4个地区的各项集聚值相对其它地区高,因此,结合这几种结果,可以知道2012年信息服务业也主要集中在北京、广东、江苏和浙江这4个地区。

由表2、表3值及聚类结果,可以看出信息服务业主要是集中在北京和广东两个地区,沿海地区发展仅次于

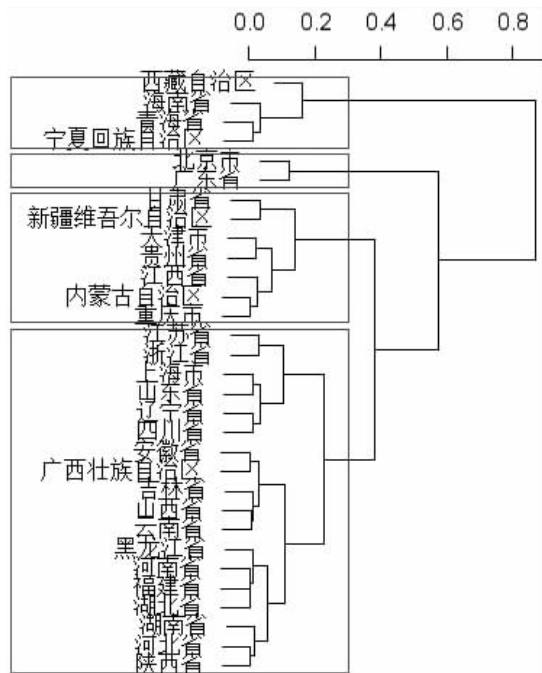


图 1 2010 年各省份综合值聚类图

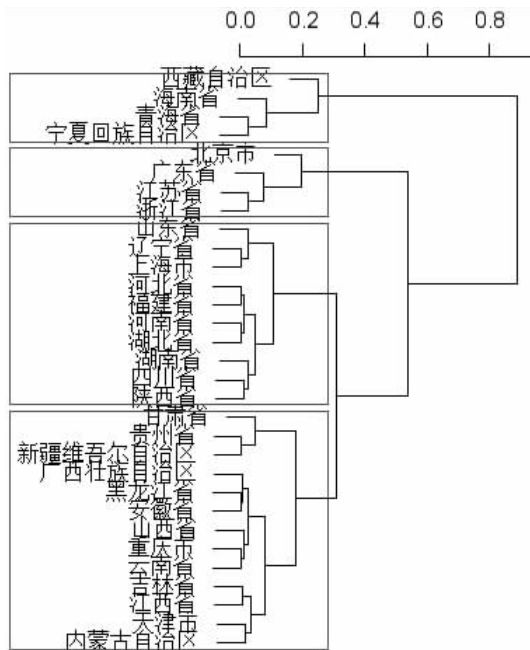


图 3 2012 年各省份综合值聚类图

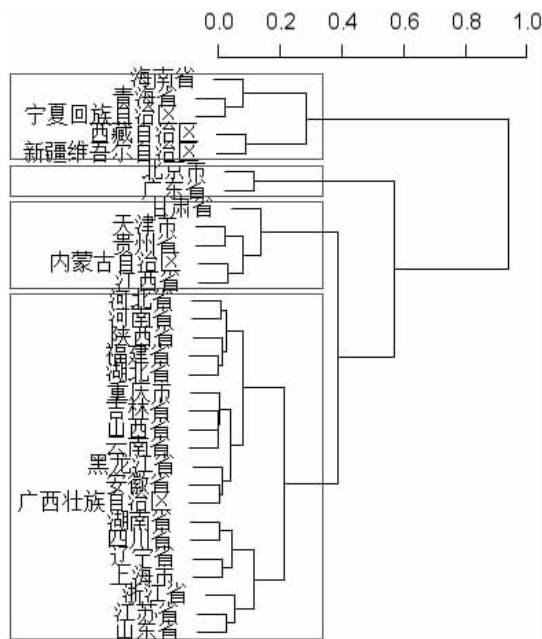


图 2 2011 年各省份综合值聚类图

北京和广东,中西部地区发展相对迟缓。2012 年集聚情势有了江苏和浙江的步入,可以知道这一产业在不断地扩大市场,从而带动更多地区这一产业的发展,总体看来是非常好的发展趋势。

4 结 论

在研究了各省份的信息服务业的 EG 指数值后,分析出这一服务产业在产业间和地理上都呈现出显

著的集聚状态,在这一产业呈现集聚显著的基础上,通过这一产业单位数和就业人数两个指标得到新的综合测量值,并对此进行聚类得出这一产业主要集中在北京和广东这两个省份,从而得出以下结论:

- (1) 在服务产业间,信息服务业这一产业首先属于行业区域集聚。
- (2) 从 EG 指数可以判断出信息服务业在地理上也处于高度集聚状态。

通过聚类分析可以知道哪些省份的发展类似,便于得出这一产业主要集中在哪些省份。

基于以上结论,针对不同区域,本文提出以下建议:加快中部地区发展,因为中部地区毗邻沿海地区,有较好的资源和市场,如果发挥这部分地区的优势,加强与东部、国际合作,不断加快城市发展进而辐射农村发展,此外,中部地区有辐射面积很广的优势,在中部地区信息服务业发展加快的情况下,辐射到周围也相对会比较容易。因此,有必要加强中部地区信息服务业的发展。再则,要不断加强北京,广东等地区的高集聚优势,可以巩固加强沿海地区,不断推动新型服务产业的发展,进而不断扩大这些地区经济圈的辐射作用和范围,以正向影响周边地区的信息服务业发展。最后,不能放松西部大开发政策,政府要不断加强落实实施情况,全面推进发展机遇,积极同东部地区建立长期合作,还要不断引进国际投资,进而发挥自身优势,以促进西部地区信息传输,计算机服务和软件服务业的发展。

参考文献:

- 集聚水平研究[J].科技进步与对策,2012,29(8):43-46.
- [1] 张华,梁进社.产业空间集聚及其效应的研究进展[J].地理科学进展,2007,26(2):14-24.
- [2] 赵黎明,邢雅楠.基于EG指数的中国旅游产业集聚研究[J].西安电子科技大学学报:社会科学版,2011,21(2):43-48.
- [3] 张翔.工业空间集聚指数及其在上海的应用[J].上海大学学报:自然科学版,2003,9(5):467-455.
- [4] Illeris S, Philippe J. Introduction: the role of services in regional economic growth[J]. Service Industries Journal, 1993,13(2):3-10.
- [5] 类骁,韩伯棠.基于EG指数模型的我国制造业产业
- [6] 张文君.基于EG指数的我国制造业产业集聚度比较[J].华北电力大学学报:社会科学版,2012,2(4):32-35.
- [7] Amott R. Congestion tolling and urban spatial structure [J]. Journal of Regional Science, 1998,38(3):495-504.
- [8] 邓宏兵,刘芬,庄军.中国旅游业空间集聚与集群化发展研究[J].长江流域资源与环境,2007,16(3):289-292.
- [9] Han Shengsun, Qin Bo. The spatial distribution of producer services in Shanghai [J]. Urban Studies, 2009,46(4): 877-896.

Spatial Agglomeration Study of China's Information Service Industry

HUI Jie, LI Xiaolin

(School of Management Sciences, Chengdu University of Technology, Chengdu 610059, China)

Abstract: The industrial concentration index of information service industry of China's each province is used to study the spatial agglomeration degree, and it is reached that the degree of agglomeration is great. On this basis, two related indicators of the service industry are found, using industrial concentration to determine that information services are mainly concentrated in six areas such as Beijing, Guangdong and so on, then the two indicators data are dimensionless processed, and the new comprehensive index is got by weighted summation, finally, the new index is system clustering analysed, the result shows that the industry is mainly concentrated in Beijing and Guangdong.

Key words: spatial agglomeration; EG index; industrial concentration CR_n ; system clustering