

## 摘要

医院和患者作为医疗系统中重要的组成部分，对社会医疗保险体制的发展起着决定性的影响，而医疗资源作为患者与医疗卫生机构之间建立联系的纽带，直接关系到患者的就医需求能否得到满足以及医疗保险体制能否正常运行。目前，西安市的医疗资源处于资源不足与资源浪费并存的局面，因此有必要对该医疗系统从系统角度进行定量与定性相结合的具体分析，从而进一步把握该医疗系统中存在的问题，制定出相应的解决方案。

本文将西安市的医疗系统视为一个社会网络，并分别对该网络中的医院节点和患者节点进行了位置取向分析，对医院与患者之间的关系进行了关系取向分析，从而为我们进一步了解西安市医疗系统的特征和其中存在的问题提供了依据，并进而制定出相应的优化方案，在研究角度和方法上为我国社会医疗保险事业的研究与发展提供了一定的参考。

首先，本文论证了将西安市医疗系统视为社会网络的可行性，构建出西安市医疗系统的社会网络模型，定义了模型中的单元。其次，从位置取向角度分别对医疗系统中的医院节点和患者节点进行社会网络分析，并将聚类分析、频数分析、均值比较的方法运用于对医院、患者的“角色”分析。再次，从关系取向角度对医疗系统中医院和患者之间的关系进行社会网络分析。最后，对西安市医疗网络的特征和其中存在的问题进行总结，并相应提出优化方案，从而将定量分析和定性分析有机的结合到一起。

**关键词：**社会网络分析 数据挖掘 医疗网络 资源整合

## **Abstract**

Hospital and patient which are the very important parts of medical system have a decisive effect on the development of system of social medical insurance. Meanwhile, the medical resources which establish the ties between the hospitals with patients have a direct effect on meeting the patient's hospitalizing demand and the normal operation of medical insurance system. Presently, it is necessary to have a quantitative and qualitative analysis to the medical system from the system point of view, because there is a situation that coexists medical resource lacking and wasting in Xi'an medical system. The analysis to Xi'an medical system is propitious to find and solve problems.

This paper regards Xi'an medical system as a social network, makes a positional analysis to hospitals and patients, and makes a relational analysis to the relation between the hospitals with patients, which gives us a better understanding to character of Xi'an medical system and the basis of improvement program. The paper provides a certain reference to the social medical insurance for further development and study.

Firstly, this paper demonstrates the feasibility that Xi'an medical is a social network, designs the medical network model, and defines the model's units. Secondly, this paper makes a positional analysis to hospitals and patients, uses the methods of cluster、frequency、compare means to analyze the roles of hospitals and patients. Thirdly, this paper makes a relational analysis to the relation between the hospitals with patients. Finally, this paper summarizes the characters and problems in Xi'an medical system, makes a program to develop the system, which combines the quantitative analysis and qualitative analysis.

**Keywords:**   **Social Network Analysis**      **Data Mining**      **Medical Network**  
                  **Resource Integration**

## 西安电子科技大学

### 学位论文独创性（或创新性）声明

秉承学校严谨的学风和优良的科学道德，本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果；也不包含为获得西安电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中做了明确的说明并表示了谢意。

申请学位论文与资料若有不实之处，本人承担一切的法律责任。

本人签名：唐汉一

日期 2008.1.14

## 西安电子科技大学

### 关于论文使用授权的说明

本人完全了解西安电子科技大学有关保留和使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属西安电子科技大学。学校有权保留送交论文的复印件，允许查阅和借阅论文；学校可以公布论文的全部或部分内容，可以允许采用影印、缩印或其它复制手段保存论文。同时本人保证，毕业后结合学位论文研究课题再撰写的文章一律署名单位为西安电子科技大学。

（保密的论文在解密后遵守此规定）

本学位论文属于保密，在\_\_\_年解密后适用本授权书。

本人签名：唐汉一

日期 2008.1.14

导师签名：唐汉一

日期 2008.1.14

## 第一章 绪论

### 1.1 研究背景

随着我国经济的不断发展,建立、健全社会医疗保险体制是维护社会稳定、经济和社会全面可持续发展的必要条件。医疗保障制度关系国计民生,是社会经济发展的“安定网”和“稳定器”。

社会医疗保险是指通过国家立法,强制性地由国家、企事业单位和参保人三方共同集资建立社会医疗保险基金。参保人缴纳保费的多少取决于其经济收入,即按照其经济收入的一定比例缴纳,保险费率与参保者的年龄、性别及健康状况无关,参保人缴纳保费的多少也不影响享用医疗卫生服务和费用的平等性。当参保人因病需要获得必需的医疗服务时,由社会医疗保险机构提供医疗费用补偿,国家财力的大小和社会经济发展的水平决定了社会医疗保险的保障水平<sup>[1]</sup>。

1998年12月,国务院召开全国城镇职工基本医疗保险制度改革工作会议部署改革工作,国务院下发《关于建立城镇职工基本医疗保险制度的决定》(国发[1998]144),明确了改革的目标、原则和主要政策,这一决定标志着我国医疗保险制度的改革迈出了关键的一步,之后全国各地的城镇职工基本医疗保险制度进入全面实施阶段<sup>[2]</sup>。在目前的医保制度下,医保管理机构、医保服务机构、参保人、参保单位已共同构成了一个社会化的医保运作体系,并取得了一定的成就。

然而,现行的医疗保险制度仍然不能满足人们的要求,存在着许多的问题。随着经济的迅猛发展和人民生活水平的不断提高,居民对自己身体状况的关注意识不断加强,从而促进了医疗服务需求的不断增长,而我国的医疗资源虽然日益得到改善,但鉴于其医疗资源的配置不合理等原因,在医疗系统中仍存在医疗资源不足与资源浪费共存的现象<sup>[3]</sup>。医疗资源作为一种稀缺性资源,相比于医疗服务需求的迅猛增长态势,目前仍不能完全满足普通居民的医疗需求,进而导致了“看病难”、“看病贵”现象的发生,这不仅限制了我国居民健康水平的提高,也有碍于目前社会医疗保险体制的不断发展。

### 1.2 研究现状

对于医疗保险体制的研究,国内学者大多集中于对医疗卫生资源现状、资源利用效率、提高医疗卫生资源的途径等方面进行定性研究。例如,易红,胡祖斌等四人在文章《城市医疗资源调整及其作用的理论研究》一文中阐述了城市医疗卫生资源优化配置的理论,并对目前我国城市医疗资源配置的现状和存在的问题及其原因进行了分析<sup>[4]</sup>。

在我国学者对医疗系统中的医疗资源进行的定量研究中,有以下几种思路:1,侯本贞在《山东省医疗资源利用效率评价研究》一文中利用柯布一道格拉斯生产函数对山东省各地级市医疗资源的利用效率现状进行评价<sup>[5]</sup>。2,申一帆在文章《广州市医疗资源配置研究》中运用系统分析法对广州市医疗资源供给状况进行了分析<sup>[6]</sup>。3,李严在《关于医疗资源共享的研究》一文中采用 Web Services 技术对医疗信息资源共享进行系统设计<sup>[7]</sup>。4,扬起在文章《湖北省卫生资源现状分析》中采用灰色预警模型 GM (1.1) 对 1996~2001 年湖北省城市和农村主要卫生资源配置情况、医疗设备进行预测分析评价,预测医护、病床未来增加趋势和人均拥有卫生资源情况<sup>[8]</sup>。

在国外学者对医疗系统进行的分析中,有以下简要几个方面。Elliot B.Sloane 在《JEDI—An Executive Dashboard and Decision Support System for Lean Global Military Medical Resource and Logistics Management》一文中采用联合医疗资产仓库模型(JMAR)对军用医疗资源进行了分析<sup>[9]</sup>。Dave Sims 在《WWW Extends “Apprentice’s Assistant” to Global Medical Resource》一文分析了计算机和电子网络对全球医疗资源的影响<sup>[10]</sup>。

目前,社会网络分析理论多用于对产业集群、组织知识管理等方面的研究。例如,苑雅文在《环黄渤海合作区域产业集群的社会网络分析》一文中运用社会网络分析理论对在韩资产业集群的社会网络的形成、特点及作用进行了研究<sup>[11]</sup>。梁昌勇,陈发祥在文章《基于社会网络下的组织知识管理研究》中将组织内的知识管理从社会网络的角度构建了一个知识管理动态模型,探索出实施知识管理的有效途径<sup>[12]</sup>。本文运用社会网络分析理论对西安市医疗系统进行分析,试图探索如何有效地合理利用医疗资源,以解决看病难的现实问题。

### 1.3 研究意义

医院和患者作为医疗系统中重要的组成部分,对社会医疗保险体制的发展有着决定性的影响作用,而医疗资源作为患者与医疗卫生机构之间建立联系的纽带,直接关系到患者的就医需求能否得到满足以及医疗保险体制能否正常运行。目前,西安市的医疗资源处于资源不足与资源浪费并存的局面<sup>[13]</sup>,如果不能对医疗系统的现状进行充分了解,也就无从谈起对医疗资源进行合理配置和资源整合,进而有碍于医疗卫生事业和社会医疗保险体制的发展。

本文通过运用社会网络分析和数据挖掘相结合的方法,对西安市医疗系统现状进行了积极的探索,从而为我们进一步了解该医疗系统中存在的问题提供了依据,并进而制定出相应的优化方案。这为西安市医疗资源的合理配置和资源整合提供了定量和定性相结合的方法,进而有助于改善目前“看病难”、“看病贵”、

资源不足与浪费并存的现状，为我国社会医疗保险体制和医疗卫生事业的发展提供了可借鉴的方法。

## 1.4 研究思路及内容

### 1.4.1 研究思路与方法

首先，本文论证了将西安市医疗系统看作为一社会网络的可行性，并构建出西安市医疗系统的社会网络模型，定义了模型中的单元。

其次，从位置取向角度分别对医疗系统中的医院节点和患者节点进行社会网络分析，并将聚类分析、频数分析、均值比较的方法运用于对两节点的“角色”分析。

再次，从关系取向角度对医疗系统中医院和患者之间的关系进行社会网络分析。

最后，对西安市医疗网络的特征和其中存在的问题进行总结，并相应提出模型优化方案，从而将定量分析和定性分析有机的结合到一起。

### 1.4.2 研究内容

(1)第一章，绪论。包括医疗保险体制的现状与本文的研究意义，以及本文的研究思路、内容和结构框架。

(2)第二章，分别对社会网络分析和数据挖掘理论进行介绍，并剖析了将数据挖掘与社会网络分析理论进行交叉融合的必要性。

(3)第三章，论证了将西安市医疗系统看作为一社会网络的可行性，并构建出西安市医疗系统的社会网络模型，定义了模型中的单元。

(4)第四章，运用社会网络分析理论，对西安市医疗网络中的医院节点进行位置取向分析，其中，在角色分析中引入聚类分析方法对西安市 11 家定点医院从治疗效果和费用水平上分别进行了聚类。

(5)第五章，运用社会网络分析理论，对西安市医疗网络中的患者节点进行位置取向分析，其中，在角色分析中引入频数分析和均值比较的方法对患者的病种和入院情况进行统计描述，并应用系统聚类分析方法对 9 种高发疾病从患者特征角度进行了聚类。

(6)第六章，运用社会网络分析理论，从关系取向角度对西安市医疗网络中医院与患者之间的关系进行分析，并对西安市医疗网络的特征进行总结，同时针对模型中存在的问题相应提出优化方案。

(7)第七章，结论。对本文的主要研究方法与研究成果进行总结，提出在本文

研究基础上有待于进一步研究和完善的方面，并展望了西安市医疗系统和医疗保险体制的发展。

## 1.5 研究的结构框架

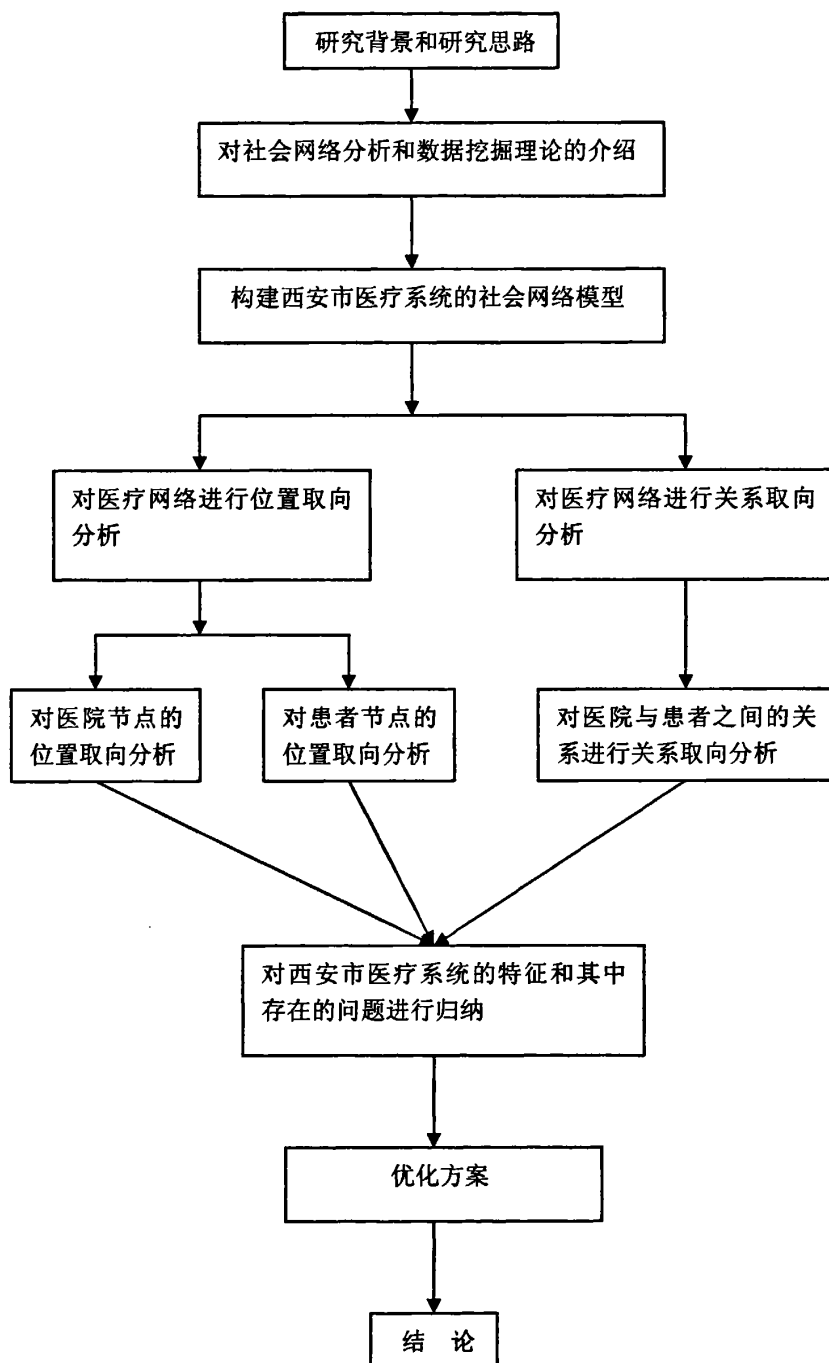


图 1.1 文章结构框架



## 第二章 社会网络分析理论

### 2.1 社会网络分析理论

#### 2.1.1 概念介绍

在社会网络分析中,“社会网络”被视为是联结行动者之间的一系列社会联系或社会关系,它们之间相对稳定的关系模式构成社会结构<sup>[14]</sup>。这说明行动者和社会关系(社会联系)是社会网络分析的两个基本要素。行动者一方面是具有独立意识的行为主体,另一方面又不得不要受到它所处社会地位的制约,表现出该社会地位所具有的属性,因此,可以说行动者既是“自主”的,又是“镶嵌”在网络中的;而社会关系则是指在行动者之间因某些特定的关系而发生互动基础上积累起来的联系模式<sup>[15]</sup>。

社会网络结构指网络内部各行动者之间相对稳定的关系模式,其中,“结构洞”模式是关系模式中一个非常重要的内容,它指的是网络中某些行动者之间关系(联系)缺乏的现象。伯特(Burt, 1992)是最早系统阐述“结构洞”概念的学者,按照伯特的观点,当网络中人们之间没有直接关系或关系缺失时,网络就会出现空洞,即所谓的“结构洞”。“结构洞”的重要性体现在,占据“结构洞”的行动者控制着网络中信息的折射和资源流动。因此,行动者在网络中占有的“结构洞”越多,其在结构上所占的优势就越大,通过这些优势获得回报的可能性就越高<sup>[16]</sup>。

根据分析的着眼点不同,社会网络分析可以分为两种基本视角:关系取向(relational approach)和位置取向(positional approach)<sup>[17]</sup>。关系取向关注行动者之间的社会性粘着关系,通过社会联结本身——如网络密度、关系强度、关系对称性、节点规模等——来说明特定的行为和过程。按照这种观点,那些强关系的、密集的且相对孤立的社会网络可以促进统一社会网络特征的形成。而位置取向则关注存在于行动者之间的、且在结构上相处于相等地位的社会关系的模式化,它讨论的是两个或两个以上的行动者和第三方之间的关系所折射出来的社会结构,强调用“结构等效”来理解人类行为。

#### (1) 关系取向中的主要分析内容

由于社会网络分析是以网络中的关系或通过关系流动的信息、资源等为主要研究对象的,这种取向中的主要分析内容大多集中在网络节点的“关系”上。其所包含的主要分析内容如下:

##### ① 规模

处于特定社会网络中的行动者都与其它行动者有着或多或少、或强或弱的关系,规模测量的是行动者之间关系的数量。如果把研究的焦点集中在某一特定行

动者上时,对关系数量的考察就变成了对网络集中性(centrality)的考察。所谓“集中性”是指特定行动者身上凝聚的关系的数量,用来衡量谁是该网络的主要节点。一般来说,某行动者凝聚的关系数量越多,它在网络中的地位就越重要<sup>[18]</sup>。衡量网络集中性的公式有如下两个:

$$C_D(n_i) = d(n_i) = \sum_j X_{ij} = \sum_j X_{ji} \quad \text{式 (2-1)}$$

$$C'_D = \frac{d(n_i)}{g-1} \quad \text{式 (2-2)}$$

式(2-1)用来衡量节点关系数量的绝对数值,即是把节点的关系数进行加总,其中, $X_{ij}$ 是0或1的数值,代表行动者*i*是否与行动者*j*存在关系;式(2-2)用来衡量节点关系数量的标准化数值,将节点关系数除以网络最多关系数,其中, $g$ 是此网络中的节点数。

## ② 强度

衡量关系强度的变量包括关系的时间量(包括频度和持续时间)、情感紧密性、熟识程度(相互信任)以及相互之间的交易金额。在医疗网络中,患者在某医院住院时间越长,医疗费用越多,则患者与此医院的关系就越强,反之则越弱。

## ③ 密度

网络中一组行动者之间关系的实际数量和其最大可能数量之间的比率称为“密度”。它用来表示行动者的关系是否紧密,可以用于测量社会网络中行动者之间的连结程度<sup>[19]</sup>。当实际的关系数量越接近于网络中的所有可能关系的总量,网络的整体密度就越大,反之则越小。一个团体可以有紧密关系,也可以有疏离关系,紧密团体的社会行为十分不同于疏离团体。一般来说,关系紧密的团体合作行为较多,信息流通较容易,团体工作绩效也会较好,而关系十分疏远的团体则常会出现信息不通、情感支持太少、工作满意程度低的问题。图形密度的公式如下,其中, $L$ 为网络中的关系数, $g$ 为网络中的节点数。

$$\Delta = \frac{2L}{g(g-1)} \quad \text{式 (2-3)}$$

## ④ 内容

所谓网络关系的“内容”,是指两行为者间关系产生的原因与关系建构基础,<sup>[20]</sup>即使在相同的网络中,行动者之间的关系也会具有不同的内容。任何可能将行动者联系起来的東西都能使行动者之间产生关系,因此内容的表现形式也是多种多样的,交换关系、亲属关系、信息交流关系、感情关系、工具关系、权力关系等都可以成为具体的内容。例如,在患者与医院建立的关系中可以有门诊关系、住院关系、健康检查关系等不同形式。

## ⑤ 不对称关系与对称关系

在不对称关系中,相关行动者的关系在规模、强度、密度和内容方面是不同的,而在对称关系中,行动者的关系在这些方面的表现都是相同的。例如:当信息只从患者流向医院,而医院不向患者提供信息时(由于双方信息不对称),两者之间的关系就是不对称关系。

#### ⑥ 直接性与间接性

网络关系的另一个内容就是直接性或间接性,前者指行动者之间直接发生的关系,后者则指必须通过第三者才能发生的关系。一般来说,直接关系联结的往往是相同或相似的行动者,他们往往彼此认同,具有相同的价值观,因此其关系通常为强关系;而间接关系中由于有中间人的存在,相互联系的行动者之间关系的强度受距离(中间人的数量)的影响很大,经历的中间环节越多,关系越弱,反之则可能越强<sup>[21]</sup>。例如,在医疗网络中,两家医院因“联合会诊”某一名患者而建立医疗合作,此时两医院之间的医疗合作关系正是通过患者(第三方)建立起来的,因此属于间接性关系。

#### ⑦ 主动关系或被动关系

关系产生时,因为行为者本身意向的主被动的不同,会产生不同的关系模式。例如,在医疗系统的医患关系中,患者主动寻找医院就医,医院没有选择患者的权力,因此,患者处于主动关系状态,而医院处于被动关系状态。

#### ⑧ 关系的方向

关系可分成有方向性及无方向性。例如,两患者在同一家医院就医,若两患者之间由于没有交流,则两患者之间的关系属于无方向性,但医院与患者的诊疗关系却完全不同,必定是由医院诊疗患者,因此可视为是具有方向性的关系。

#### (2) 位置取向中的主要分析内容

与关系取向不同,位置取向强调的是网络位置的结构性特征。如果说关系取向是以社会粘着为研究基点,以关系的各种特征为表现的话,那么,位置取向则以结构上的相似为基点,以关系的相似性为基本特征。在位置取向看来,位置所反映出来的结构性特征更加稳定和持久,更具有普遍性,因而对现实也更有解释力,且需要分析的内容也更为简单明了。其主要内容有:

##### ① 结构等效

当两组或两组以上的行动者(他们之间不一定具有关系)与第三个行动者具有相同的系统时,即为“结构等效”。这里强调的是在同一社会网络中所谓的“结构等效点”必须与同一个点保持相同的系统。网络中“结构等效点”的数量和质量将对网络的驱动力产生很大的影响。

##### ② 位置

作为位置取向的核心概念,“位置”在这里指的是结构上处于相同地位的一组行动者或节点,是被剥落了行动者而剩下的结构性特征,哪个行动者处在这个位

置并不重要，重要的是这个位置在网络本身中的处境。

### ③角色

与位置密切相关的另一项内容是角色，它是在结构上处于相同地位的行动者在面对其它行动者时表现出来的相对固定的行为模式。反过来说，具有相同社会角色的往往在社会网络结构或地位网络结构中处于相同的位置。因此，角色在某种程度上是位置的行为规范。例如，在医疗网络中，角色可以代表医院在面对患者时表现出来的行为模式，这种行为模式主要体现在对患者的治疗效果和医疗费用水平。

## 2.1.2 社会网络分析的单元

社会网络的分析单元共有点、关系与联系三种<sup>[23]</sup>，分别说明如下：

### (1)点

社会网络中的点是各个行动者，线是行动者之间的各种社会关系<sup>[22]</sup>。在社会网络研究领域，任何一个社会单位或者社会实体都可以看成点。例如，在医疗网络中，行动者可以是某一个患者实体，也可以是医院这样的社会单位。

### (2)关系

关于对关系的描述，与上文关系取向分析的内容相同，此处不再赘述。

### (3)联系

所谓的联系是指两行动者间的关系组合，是行动者之间的多元关系<sup>[23]</sup>。联系所含的关系可能只有一种，也有可能由多种的关系组合而成，例如两人彼此有商业合作、信息交换、学术交流等关系。又根据联系的“强弱”又可分成弱联系及强联系两种观念，及其定义如下（Wellman, 1992）：

强联系：行动者之间关系紧密，经常联系的视为强联系，具有强联系的行动者之间会自发性的维护联系，且常透过各种不同的关系来达到彼此间长时间的互动，从而产生更多的信息分享和相互协助<sup>[24]</sup>。

弱联系：行动者之间不紧密联系，彼此间接触较不频繁，或是通过间接联络的社会关系视为弱联系。

## 2.1.3 社会网络分析的基本特征

社会网络分析不能预测社会事实的发展，它更多的是一种分析视角。正是由于这种分析视角上的差异，不同立场的学者在研究主题、方法论、研究结论方面都会有所不同，但社会网络分析作为一个学术流派仍然具有一些基本特征。

第一，社会网络分析理论以关系模式为基础，用行动者的相对特征和与其它行动者的社会关系探讨社会网络的结构，而对社会网络中占有一定位置的个人所

具有的绝对特征不予考虑。

第二,经过社会网络理论对网络结构的重新定义,可以用“角色”、“地位”等概念来表明网络结构特征,这种对网络结构特征的定义并不是由各节点的特征简单加总而得,而是认为该网络的特征是基于节点的特征和节点间的社会关系所共同体现出来的。

第三,社会网络分析反对过分强调行动者的自主性行为,反对将社会过程视为行动者有意识行动的结果。在社会网络分析理论看来,社会结构一旦由行动者建立起来,它就会对行动者具有约束力,行动者的行为模式更多的是体现它所处社会地位所具有的属性,不可能完全突破社会网络结构的限制<sup>[25]</sup>。

第四,受图形论、概率论和几何学等数学方法的推动,兴起于 20 世纪 60 年代的社会网络分析把研究重点聚焦在社会系统的层面上,它关注的是整个互动领域和社会情境的作用。

不管社会网络中的行动者处于哪个层次和地位,社会网络分析的目标都是通过对行动者的分析获得对整个社会网络特征的认识。社会网络分析理论认为,群体的性质往往是由其内部行动者之间的互动决定的,而这些行动者的互动则是由他们所处的群体之间的互动决定的。因此,社会网络分析必须同时兼顾个体和群体两个层面,并在个体和群体之间建立沟通的桥梁。

#### 2.1.4 处理社会网络分析的软件

在处理社会网络分析的软件中最常见的有 Krackplot、STRUCTURE、NEGOPY、UCINET 四种,其中又以 UCINET 为最常使用的分析软件,在本文对西安市医疗网络模型所做的社会网络分析中,对于节点集中性、网络密度等指标的计算正是基于 UCINET 软件进行实现的。

UCINET 软件由 Borgatti, Everett, Freeman 制作而成。UCINET 网络分析集成软件,包括一维与二维数据分析的网络绘图,还有正在发展应用的三维展示分析软件 Mage 等,同时集成了 Pajek 用于大型网络分析的 Free 应用程序,该软件可以对关联性、嵌入性、自我中心网、中心性、密度等社会网络分析指标进行计算<sup>[27]</sup>。

## 2.2 社会网络分析与数据挖掘的交叉融合

### 2.2.1 社会网络分析与数据挖掘理论交叉融合的必要性的必要性

社会网络分析理论作为研究社会网络的一门社会学理论,首先是以大型理论的身份出现在世人面前的,然而,如果社会网络分析理论总停留在大型理论阶段,

则实证工作很难衔接,所以其必须由大型理论到中层理论,再发展到可检证的、指定清楚的因果模型(causal model),且环环相扣。为了解决或说明实际问题,社会网络分析理论必须不能止于概念架构的陈述以及用这种概念诠释所观察到的现象,而是应该可以指定因果模型的控制变量、变量测量方法、函数型态、以及因果结构,满足其可检证因果模型所需要的模型指定。即换句话说,社会网络分析理论需要中层理论在大型理论与因果模型之间建立一座桥,为一组概念找到具体可测量的被解释变量,也提供这组概念如何解释被解释变量的因果结构。少了中层理论发展的大型理论,在实证研究上不具有意义,相反,少了理论导引的因果模型,在过去十多年的社会网络分析理论方法论战中也被认为见树不见林,找出来的模型不必然具有社会学上的解释力,成为“无大脑的实证主义”(Mindless Empiricism)。因此,正是由于中层理论的出现,使得社会网络分析理论在概念架构(大型理论)和解释实际问题方面(因果模型)都可以施展作用<sup>[28]</sup>。

而数据挖掘理论作为一种中层理论,可以和社会网络分析理论相交叉融合,通过对社会网络随机抽样、确定可观测变量及函数模型、处理和分析数据,使社会网络分析理论在说明具体问题上得到了数据和模型的佐证,从而更有说服力。

同时社会网络分析理论中所涉及的样本数据都是大量纷繁的,而数据挖掘的功能就在于“从海量的数据中挖掘出可能有潜在价值的信息”,因此,在设立模型处理数据的过程中,也同样需要数据挖掘来为社会网络分析予以辅助。

### 2.2.2 数据挖掘理论的介绍

#### (1)数据挖掘的定义

随着科学技术的进步,新的数据采集和获取技术不断发展,使得数据库中所存储的数据量也随之急剧增长。然而,数据处理技术的发展却相对落后,数据库技术仍然停留在相对简单的录入、查询、统计、检索阶段,对数据库中的数据之间存在的关系和规则、数据的群体特征、数据集内部蕴涵的规律和趋势等,缺乏有效的技术手段将其提取出来,从而出现所谓的“被数据淹没,却饥渴于知识”(John Naisbett, 1997)的现象。为了解决这种现象,科学家们于20世纪80年代末期创立了一个新的研究领域,即数据挖掘(Data Mining)。

目前对于数据挖掘还没有一个完整的定义,各国研究学者提出了多种数据挖掘的定义。例如,SAS研究所(1997)认为数据挖掘是“在大量相关数据基础之上进行数据探索和建立相关模型的先进方法”。Bhavani(1999)认为数据挖掘是“使用模式识别技术、统计和数学技术,在大量的数据中发现有意义的新关系、模式和趋势的过程”。而Hand et al(2000)认为数据挖掘就是“在大型数据库中寻找有意义、有价值信息的过程”<sup>[29]</sup>。

综上所述,数据挖掘就是从海量的数据中挖掘出可能有潜在价值的信息的技术,用以支持决策,可以为企业带来利益,或者为科学研究寻找突破口。它是在数据库技术、机器学习、人工智能、统计分析等基础上发展起来的一个交叉性的学科<sup>[30]</sup>。

### (2)数据挖掘的功能

自从数据挖掘的概念于1989年8月首次出现在第11届国际联合人工智能学术会议以来,数据挖掘领域的研究和应用得到了长足的发展,形成了一些行之有效的理论和方法,并逐渐成为大量数据信息处理领域的研究热点。数据挖掘的主要功能有如下5个方面<sup>[31]</sup>:

①分类:通过分析对象属性值的异同,建立不同的组类来描述事物。

②聚类:对分析对象的内在规则进行识别,按照这些规则把对象分成若干类。

③关联规则和序列模式的发现:希望分析出当某一事物发生变化时,其它事物是否也会发生变化以及将发生何种变化。

④预测:把握分析对象发展规律,对未来的趋势作出预见。

⑤偏差的检测:对分析对象的少数的、极端的特例的描述,揭示内在的原因。

### (3)数据挖掘的方法

数据挖掘通过对数据的总结、分类、聚类、关联等分析,实现对数据内在结构特征的理解和对未知数据的预测。常见和应用最广泛的数据挖掘方法有传统的统计分析描述、可视化技术、联机分析处理(OLAP)、决策树、神经网络、遗传算法、关联规则挖掘算法<sup>[32]</sup>。

### (4)处理数据挖掘的软件

目前许多公司和研究机构都开发了一系列的软件工具用于数据挖掘,例如:SPSS公司的SPSS15.0和Clementine、SAS公司的Enterprise Miner、IBM公司的Intelligent Miner等。本文对11家定点医院和9种高发疾病的聚类分析正是基于SPSS15.0软件进行的实现。

## 2.2.3 数据挖掘与社会网络分析的交叉融合点

数据挖掘理论作为一种中层理论方法,与以大型理论角色出现的社会网络分析理论存在极其广泛的交叉融合点,两种理论之间互为补充,互为促进。在本文对西安市医疗网络模型从位置取向角度进行社会网络分析时,运用到聚类分析、频数分析、均值比较三种数据挖掘方法对医院和患者节点进行“角色”研究,从而更加量化的分析出各节点在网络中的属性和特征,希望通过这样的理论方法融合带给对社会网络分析更多积极的思考。其中,本文运用系统聚类分析方法对11家定点医院分别从治疗效果和费用水平进行了聚类,对9种西安市高发疾病从患者

特征角度进行了聚类。运用频数分析和均值比较的方法对各入院情况和病种下的患者年龄和性别特征进行统计描述。

### (1) 聚类分析

聚类分析根据“物以类聚”的原理,将本身没有类别的样本聚集成不同的组,并且对每一个这样的组进行描述的过程。其主要依据(即目的)是聚到同一个组中的样本应该彼此相似,而属于不同组的样本应该足够不相似。与分类分析不同,进行聚类前并不知道将要划分成几个组和什么样的组,也不知道根据哪些空间区分规则来定义组,其目的旨在发现空间实体的属性间的函数关系<sup>[34]</sup>。

聚类分析按其分类方法又分为系统聚类法、动态聚类法等。其中,系统聚类分析也称为分层聚类法(Hierarchical Cluster),它是聚类分析中应用最广泛的一种方法。在本文通过SPSS15.0对医院和病种进行聚类分析的过程中,也正是基于这一方法。系统聚类分析的思想是:开始将样品或者指标各视为一类,根据类与类之间的距离或相似程度将最相似的类加以合并,再计算新类与其它类之间的相似程度,并选择最相似的类加以合并,这样每合并一次就减少一类,不断继续这一过程,直到所有样品(或指标)合并为一类为止;最常用的系统聚类法又可以分为:最短距离法,最长距离法,中间距离法,重心法,类平均法,可变类平均法,可变法以及离差平方和法等多种方法<sup>[35]</sup>。

因此,我们可以运用聚类分析的方法来探寻社会网络中的节点构成和节点特征,将具有相同或相似属性的结点划分成一类,从而更进一步发掘在网络中众多节点的关联性,化简了结点个数,避免了因结点过多而引发的分析不便。

### (2) 频数分析和均值比较

频数分析是对符合条件的分析对象进行个数统计,进而比较出各种情况下的事件发生个数,从中寻找事物的内在规律<sup>[31]</sup>。通过一维频数分析,我们可以对各病种和各年龄段的患者数进行统计,从而分析出患者在各病种和各年龄段的患者人数构成,有助于我们了解患者的属性——易患病种、易患病年龄段。通过二维和三维频数分析,我们可以了解在各病种下和各入院情况下,各性别和各年龄段患者的人数构成。

频数分析是对个案数的统计,而均值比较是对各个案属性值的分析,它是指将属于某一情况下各个案的属性值进行平均,通过比较各种情况下的属性平均值发现事物内在的属性。通过均值比较,我们可以了解到各病种下患者的平均年龄等类似结论,可以说,频数分析和均值比较是从两个不同的角度对事物的属性进行了描述。



## 第三章 构建西安市医疗系统的社会网络模型

### 3.1 构建医疗系统网络模型的可行性

社会网络被认为是联结行动者的一系列社会联系或社会关系，它们之间相对稳定的关系模式构成社会结构。社会网络的两个基本要素是行动者和行动者之间的社会联系，其中，行动者是具有意识的行为主体，而社会联系则是在行动者之间因某些特定的关系而发生互动的基础上积累起来的联系模式。行动者的行为既是“自主”的，也是“镶嵌”在互动网络中的，因其所具备的社会结构属性而受到社会网络的制约。

在医疗系统中，主要的行为主体有以下四类：政府、社会医疗保险机构、医疗机构、患者。其中，政府可以通过制定、实施政策和财政拨款对社会医疗保险机构和医疗机构产生直接的影响，并对患者产生间接的影响；社会医疗保险机构通过与医疗机构和患者的资金往来产生联系；患者通过到医疗机构就医与其产生关系。政府、社会医疗保险机构、医疗机构、患者正是通过这些关系而有机的结合在一起，并形成医疗系统。因此，在医疗系统中也存在着社会网络模型定义中所要求的行动者和行动者之间的社会关系，从而把医疗系统也可以看成是一个社会网络模型。

通过将医疗系统构建成一个社会网络，我们可以从宏观的角度，运用社会网络分析理论对医疗系统进行研究，进而有助于我们明晰该医疗网络节点的结构属性和节点间的社会联系，为政府和社会医疗保险管理机构制定、实施政策提供科学、量化的依据。本文主要针对医疗机构和患者的结构属性和两者之间的诊疗关系进行研究，因此，本文在下面的研究中，将医疗系统网络模型简化成只包含医疗机构和患者两个节点。

### 3.2 西安市医疗网络模型中的单元

在由西安市医疗系统构建成的社会网络模型中，关于模型中的单元，本文将从点、关系、联系三个方面来进行定义：

#### 3.2.1 点

在社会网络分析中，最基本的元素就是点和线，点代表行动者，而线代表行动者之间的关系或是联系。在由西安市医疗系统构建成的社会网络模型中，本文只针对医疗机构和患者的结构属性和两者之间的诊疗关系进行研究，即该网络只包括医院和患者两个节点。

### 3.2.2 关系

本文将从内容、方向、强度、主动关系或被动关系四个方面对西安市医疗网络模型中的关系进行定义。

#### (1)内容

内容就是指两行为者间关系产生的原因与关系的建构基础，对于西安市的医疗网络模型，医院与患者之间产生的关系正是基于患者到医院就医，以及随之产生的医疗服务和费用。此外，本文在研究过程中，对医院之间以及患者之间的关系进行简化，认定医院与医院之间和患者与患者之间没有联系，进而不会相互影响。

#### (2)方向

在西安市医疗网络模型中，本文认定在同一家医院就医的两个患者之间的关系属于无方向性，同理，一个患者若先后在 A、B 两家医院就医，那么认定 A、B 两家医院之间的关系也是属于无方向性的；而医院与在该医院就医的患者之间的关系属于有方向性，且这种由医院向患者提供医疗服务、患者向医院缴纳医疗费用产生的关系属于双向性关系。

#### (3)强度

在衡量医院与在该医院就医患者之间这种双向性关系的强度时，我们可以选择住院天数和医疗费用金额两个指标进行评价，即住院时间长、发生的医疗费用多意味着该患者与医院之间的关系强度大。

#### (4)主动关系或被动关系

在医疗网络模型中，我们结合实际情况可以知道，医患关系的成立是由患者主动寻找医院就医而完成的，因此，在存在关系的医院与患者之间，患者属于主动关系状态；由于存在信息不对称，除非医院能够向患者提供有关治疗服务质量的信息，否则医院属于被动关系状态。

### 3.2.3 联系

从第二章的理论介绍中，我们了解到，所谓的联系是指两行动者间的关系组合，其联系所包含的关系可能只有一种，也有可能由多种的关系组合而成。在本文的医疗网络模型中，医院与患者之间发生的联系可以按患者的就医类型（门诊、住院、身体检查）进行关系的划分。鉴于本文所采集的是关于住院患者的数据，因此，本文只对发生住院关系的医院与患者进行研究，而对由门诊、身体检查所建立起来关系的医院与患者不予考虑。

在发生住院关系的医院与患者中，也存在“强联系”与“弱联系”的划分。其中，本文将住院次数多、住院时间长、发生的医疗费用大的患者定义为与该医

院具有“强联系”，反之，将住院次数少、住院时间短、发生的医疗费用小的患者定义为与该医院具有“弱联系”。

### 3.3 构建西安市医疗网络模型

在 3.1 节中，本文论证了将西安市医疗系统构建成社会网络的可行性；在 3.2 节中，本文给出由医院和患者两个节点构建成的西安市医疗系统网络模型单元的定义。下面，本文将以矩阵的形式给出描述西安市医疗系统网络模型的数据<sup>[38]</sup>。假设在西安市医疗网络模型中，医院节点共有  $N$  个，分别定义为  $A_1, A_2, \dots, A_N$ ；患者节点共有  $M$  个，分别定义为  $X_1, X_2, \dots, X_M$ ；那么，医疗网络模型中将由  $N+M$  个节点组成。假设用  $a_{ij}$  来表示网络中节点  $i$  和节点  $j$  之间的关系，其中， $1 \leq i, j \leq N+M$ 。于是，我们可以将西安市医疗网络模型的数据表示为矩阵  $A_{(N+M) \times (N+M)}$ 。

$$A_{(N+M) \times (N+M)} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{1N} & a_{1(N+1)} & \cdots & a_{1(N+M)} \\ a_{21} & a_{22} & \cdots & a_{2N} & a_{2(N+1)} & \cdots & a_{2(N+M)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} & a_{N(N+1)} & \cdots & a_{N(N+M)} \\ a_{(N+1)1} & a_{(N+1)2} & \cdots & a_{(N+1)N} & a_{(N+1)(N+1)} & \cdots & a_{(N+1)(N+M)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{(N+M)1} & a_{(N+M)2} & \cdots & a_{(N+M)N} & a_{(N+M)(N+1)} & \cdots & a_{(N+M)(N+M)} \end{pmatrix}$$

$$= \begin{pmatrix} A_{N \times N} & A_{N \times M} \\ A_{M \times N} & A_{M \times M} \end{pmatrix}$$

其中， $A_{N \times N}$ ， $A_{N \times M}$ ， $A_{M \times N}$ ， $A_{M \times M}$  为  $A_{(N+M) \times (N+M)}$  中的四个分块矩阵。假设将矩阵  $A_{(N+M) \times (N+M)}$  的节点  $1 \sim N$  定义为  $N$  家医院，节点  $N+1 \sim N+M$  定义为  $M$  名患者。结合 3.2 节关于节点关系的叙述，我们可以知道，对于医院之间和患者之间的节点关系定义为零，且医院与自身和患者与自身的关系也定义为零，那么，表明  $A_{N \times N}$  和  $A_{M \times M}$  为两个零矩阵，即我们可以将  $A_{(N+M) \times (N+M)}$  进一步表示为如下形式：

$$A_{(N+M) \times (N+M)} = \begin{pmatrix} 0 & A_{N \times M} \\ A_{M \times N} & 0 \end{pmatrix}$$

$$\text{其中， } A_{N \times M} = \begin{pmatrix} a_{1(N+1)} & \cdots & a_{1(N+M)} \\ a_{2(N+1)} & \cdots & a_{2(N+M)} \\ \vdots & \vdots & \vdots \\ a_{N(N+1)} & \cdots & a_{N(N+M)} \end{pmatrix}, \quad A_{M \times N} = \begin{pmatrix} a_{(N+1)1} & a_{(N+1)2} & \cdots & a_{(N+1)N} \\ \vdots & \vdots & \vdots & \vdots \\ a_{(N+M)1} & a_{(N+M)2} & \cdots & a_{(N+M)N} \end{pmatrix}$$

在矩阵  $A_{N \times M}$  和  $A_{M \times N}$  中，我们将  $a_{ij}$  用来定义医院与患者之间的关系，若患者  $i$  没有在医院  $j$  就医，则这两者之间不存在关系，即  $a_{ij}=0$ ；反之，若患者  $i$  在医院

$j$  就医并发生医疗费用, 则  $a_{ij} > 0$ , 且  $a_{ij}$  的大小反映了患者与医院之间的关系强度, 可以根据患者的住院天数和医疗费用进行衡量。此外, 需要说明的是, 由于患者与医院之间的关系属于双向性关系, 因此, 可以将  $A_{N \times M}$  看成是  $A_{M \times N}$  的转置矩阵, 即  $A_{N \times M} = A_{M \times N}^T$ 。

我们可以通过下面的例子来更进一步理解上面的数学定义。假设在一个医疗系统中, 只存在三家医院和三名患者, 且医院分别为 A、B、C; 患者分别为 X、Y、Z。它们之间具有如下关系: 患者 X 只在医院 A 就医, 且住院五天; 患者 Y 先后分别在医院 B 住院三天, 在医院 C 住院四天; 而患者 Z 只在医院 C 就医, 且住院一天。如果我们单从住院天数衡量医院与患者之间的关系强度 (即以住院天数作为反映关系强度的指标) 的话, 那么, 该医疗网络模型的数据可以通过如下矩阵表达:

|   | A | B | C | X | Y | Z |
|---|---|---|---|---|---|---|
| A | 0 | 0 | 0 | 5 | 0 | 0 |
| B | 0 | 0 | 0 | 0 | 3 | 0 |
| C | 0 | 0 | 0 | 0 | 4 | 1 |
| X | 5 | 0 | 0 | 0 | 0 | 0 |
| Y | 0 | 3 | 4 | 0 | 0 | 0 |
| Z | 0 | 0 | 1 | 0 | 0 | 0 |

## 第四章 对西安市医疗网络模型医院节点的位置取向分析

根据第二章对社会网络分析理论的介绍,我们知道,在社会网络分析中,根据分析的着眼点不同,社会网络分析可以分为两种基本视角:关系取向和位置取向。其中,位置取向关注的是存在于行动者之间的、且在结构上处于相等地位的社会关系的模式化,它以结构上的相似为基点,以关系的相似性为基本特征,讨论的是两个或两个以上的行动者和第三方之间的关系所折射出来的社会结构,强调用“结构等效”来理解网络节点的行为。从位置取向对社会网络进行分析主要基于以下三个方面的内容:结构等效、位置、角色。本章试图从位置取向角度对西安市医疗网络模型中的医院节点进行研究,从而进一步明晰医院在与患者产生关系的过程中所表现出的结构特征和属性。

### 4.1 对医院节点的“结构等效”分析

当两组或两组以上的行动者(他们之间不一定具有关系)与第三个行动者具有相同的关系时,即为“结构等效”。根据社会网络分析理论,在同一社会网络中所谓的“结构等效点”必须与同一个点保持相同的关系。

针对本文所研究的医疗网络模型,从总体上看,如果将全部患者看成一个节点,那么,可以将各个医院节点看成是所谓的“结构等效点”,因为它们与患者之间都是“患者就医——医院诊疗”的关系,正是基于这些相同的关系做为纽带(即使各个医院之间不存在关系),我们可以认定各个医院之间具有“结构等效”。从具体上看,即将每个患者都看成是一个单独的节点,那么,只有与某患者 $i$ 都发生“患者就医——医院诊疗”关系的医院才可以称之为“结构等效点”,而那些与患者 $i$ 不存在“患者就医——医院诊疗”关系的医院节点并不纳入该“结构等效点”的定义范畴。

根据社会网络分析理论,网络中“结构等效点”的数量和质量将对网络的驱动力产生很大的影响。针对本文所研究的医疗网络模型,这一点也是适用的。在医疗网络模型中,医院的数量直接影响到患者的病情是否能够得到及时的医治,目前,患者在就医时感觉到的“看病难”问题正是缘由医院数量少,特别是随着人民生活水平的提高,对医疗服务的需求增加,而大医院医疗资源的发展速度明显落后于居民的医疗需求,从而导致目前在某些情况下,医院不能完全满足(或不能及时满足)居民对医疗服务的需求。因此,在医疗网络模型中,医院的数量和质量直接影响到与患者的医患关系,进而对医疗保障体制的进一步发展起到制约/促进的作用。目前,国家为舒缓大医院的接诊压力、改善居民“看病难”的问题,正在大力发展社区医院,从而使病情较轻的患者在社区医院就能得到及时的

医治,病情严重的患者再由社区医院转向大医院,而这种就医模式也正是基于“将社区医院与大医院结成结构等效点”的思想。

## 4.2 对医院节点的“位置”分析

“位置”在这里指的是结构上处于相同地位的一组行动者或节点,是被剥落了行动者而剩下的结构性特征。在本文所研究的西安市医疗网络模型中,由于各等级医院在医疗设备、医疗服务、诊疗技术方面具有相同或相似的水平,进而会反映出相似的结构特征,也即可以认为同等级的医院在医疗网络中具有相同的“位置”。因此,我们将医院等级用于对医院节点“位置”的定义,即拥有同一等级的医院是在结构上处于相同地位的一组行动者或节点,那么,从“位置”的角度可以将医疗网络模型中的医院节点划分成9组节点,即三级甲等、三级乙等、三级丙等、二级甲等、二级乙等、二级丙等、一级甲等、一级乙等、一级丙等。在以位置取向对社会网络的分析中,“哪个节点处在这个位置并不重要,重要的是这个位置在网络本身中的处境”,因此,我们在对西安市医疗网络模型的研究中更注重对相同位置节点结构性特征的探讨。

## 4.3 对医院节点的“角色”分析

在社会网络分析理论中,“角色”是指在结构上处于相同地位的行动者在面对其它行动者时表现出来的相对固定的行为模式。反过来说,具有相同社会角色的往往在社会网络结构或地位网络结构中处于相同的位置。因此,“角色”在某种程度上是“位置”的行为规范。在本节对西安市医疗网络模型医院节点的角色分析中,“角色”指的是处于相同“位置”(相同医院等级)的医院在面对患者时表现出来的行为模式,这种行为模式主要体现在对患者的治疗效果和医疗费用水平。下面,本文通过系统聚类分析对各等级医院的“角色”进行实证研究。

### 4.3.1 从治疗效果角度对各医院进行聚类分析

#### (1)数据来源

本文从西安市11家定点医院中分别提取了共838名冠心病患者的住院数据,其中,数据涵盖4家三级甲等医院(分别为H3、H4、H5、H6)3家三级乙等医院(分别为H2、H7、H11)、2家二级甲等医院(分别为H1、H8)、2家一级医院(分别为H9、H10)。由于数据采集有限,并未能涵盖所有等级医院的冠心病患者数据,但由于所采集到的各等级医院在各“位置”中具有代表性,因此,所做的聚类分析依然具有结论意义。此外,鉴于数据涉密等方面的原因,为保证数据采集样本中各医院的权益,本文以H1~H11分别代表各医院的名称。

## (2)评价指标体系的建立

医院的治疗效果正如“商品”的质量和性能一样，是评价医院所提供服务这一“商品”的重要指标。本文从患者出院情况的三个方面（治愈率，好转率，死亡率）对医院的治疗效果进行描述，意图通过上述三个指标来量化医院的治疗效果，从而更好的对各医院的“角色”进行分析<sup>[39]</sup>。

其中，本文所指的治愈率（ $x_1$ ）、好转率（ $x_2$ ）、死亡率（ $x_3$ ）是就冠心病而言，病情治愈、好转、死亡的患者人数占该病种就医患者总数的比例。

## (3)聚类分析的实现

该样本的原始数据见表 4.1:

表 4.1 11 家医院治疗效果的原始数据

|     | $x_1$ | $x_2$ | $x_3$ |
|-----|-------|-------|-------|
| H1  | 63.6  | 36.4  | 0.0   |
| H2  | 11.3  | 88.7  | 0.0   |
| H3  | 3.8   | 94.2  | 1.9   |
| H4  | 2.8   | 91.7  | 5.6   |
| H5  | 8.7   | 89.1  | 2.2   |
| H6  | 0.0   | 100.0 | 0.0   |
| H7  | 8.3   | 79.2  | 8.3   |
| H8  | 0.0   | 98.0  | 2.0   |
| H9  | 2.8   | 94.4  | 0.0   |
| H10 | 2.2   | 89.1  | 0.1   |
| H11 | 7.2   | 81.9  | 0.0   |

由于本数据各指标值不涉及量纲和数量级上的差异，因此，此处无需对数据进行标准化处理，进而可以直接运用 SPSS15.0 的 Analyze→Hierarchical Cluster 命令来进行对系统聚类分析的实现。SPSS15.0 的系统聚类分析实现界面见图 4.1:

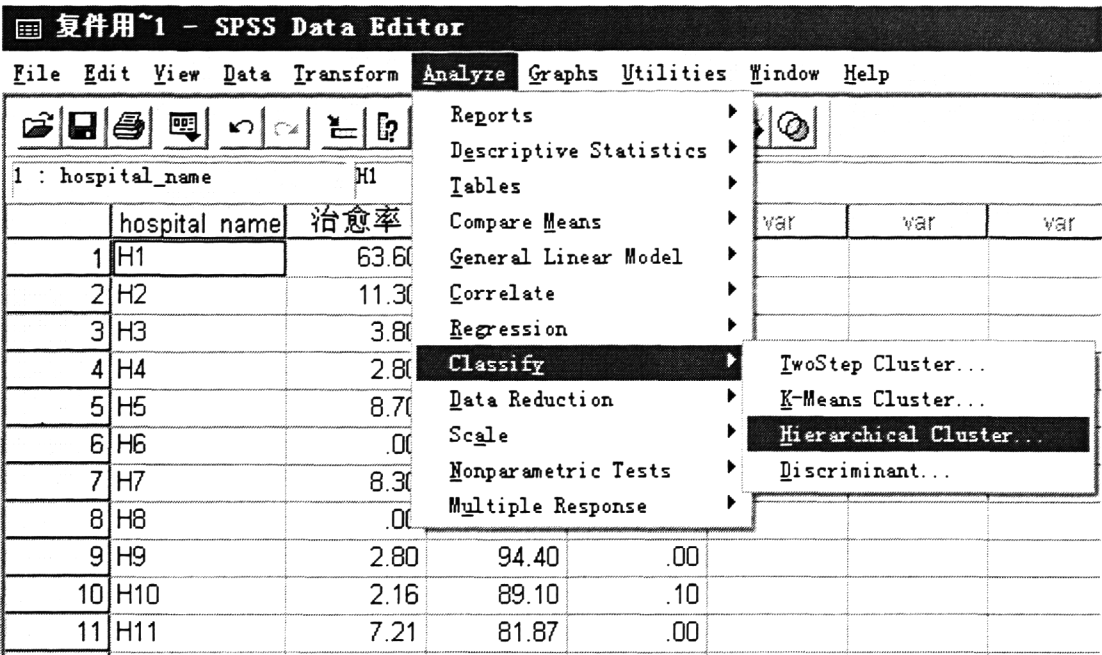


图 4.1 系统聚类分析实现界面一

(4)聚类分析的结果图

表 4.2 凝聚过程表

Agglomeration Schedule

| Stage | Cluster Combined |           | Coefficients | Stage Cluster First Appears |           | Next Stage |
|-------|------------------|-----------|--------------|-----------------------------|-----------|------------|
|       | Cluster 1        | Cluster 2 |              | Cluster 1                   | Cluster 2 |            |
| 1     | 9                | 10        | .007         | 0                           | 0         | 4          |
| 2     | 3                | 8         | .007         | 0                           | 0         | 5          |
| 3     | 2                | 11        | .016         | 0                           | 0         | 6          |
| 4     | 6                | 9         | .020         | 0                           | 1         | 6          |
| 5     | 3                | 5         | .026         | 2                           | 0         | 7          |
| 6     | 2                | 6         | .044         | 3                           | 4         | 7          |
| 7     | 2                | 3         | .081         | 6                           | 5         | 9          |
| 8     | 4                | 7         | .152         | 0                           | 0         | 9          |
| 9     | 2                | 4         | .627         | 7                           | 8         | 10         |
| 10    | 1                | 2         | 1.761        | 0                           | 9         | 0          |

表 4.2 是系统聚类分析的凝聚过程表，Stage 这一列给出聚类分析的步骤，本例共有 10 个步骤；Cluster Combined 这一列表示聚类过程；Coefficients 这一列表示样本或类之间的距离，距离最短的样本或类先聚成一类<sup>[40]</sup>。



|                    |  | Vertical Icicle |      |      |      |      |        |      |      |        |      |      |
|--------------------|--|-----------------|------|------|------|------|--------|------|------|--------|------|------|
| Number of clusters |  | Case            |      |      |      |      |        |      |      |        |      |      |
|                    |  | 7:H7            | 4:H4 | 5:H5 | 8:H8 | 3:H3 | 10:H10 | 9:H9 | 6:H6 | 11:H11 | 2:H2 | 1:H1 |
| 1                  |  | X               | X    | X    | X    | X    | X      | X    | X    | X      | X    | X    |
| 2                  |  | X               | X    | X    | X    | X    | X      | X    | X    | X      | X    | X    |
| 3                  |  | X               | X    | X    | X    | X    | X      | X    | X    | X      | X    | X    |
| 4                  |  | X               | X    | X    | X    | X    | X      | X    | X    | X      | X    | X    |
| 5                  |  | X               | X    | X    | X    | X    | X      | X    | X    | X      | X    | X    |
| 6                  |  | X               | X    | X    | X    | X    | X      | X    | X    | X      | X    | X    |
| 7                  |  | X               | X    | X    | X    | X    | X      | X    | X    | X      | X    | X    |
| 8                  |  | X               | X    | X    | X    | X    | X      | X    | X    | X      | X    | X    |
| 9                  |  | X               | X    | X    | X    | X    | X      | X    | X    | X      | X    | X    |
| 10                 |  | X               | X    | X    | X    | X    | X      | X    | X    | X      | X    | X    |

图 4.2 冰柱图

图 4.2 是系统聚类分析结果的冰柱图，从图中我们可以看出每一步的凝聚状态及结点的凝聚顺序，即哪些结点先聚为一类，之后又与哪些结点聚为另一类。例如，节点从 10 类变为 9 类时，H3 与 H8 两家医院之间的距离最小，率先聚成新的一类。

Dendrogram using Average Linkage (Between Groups)

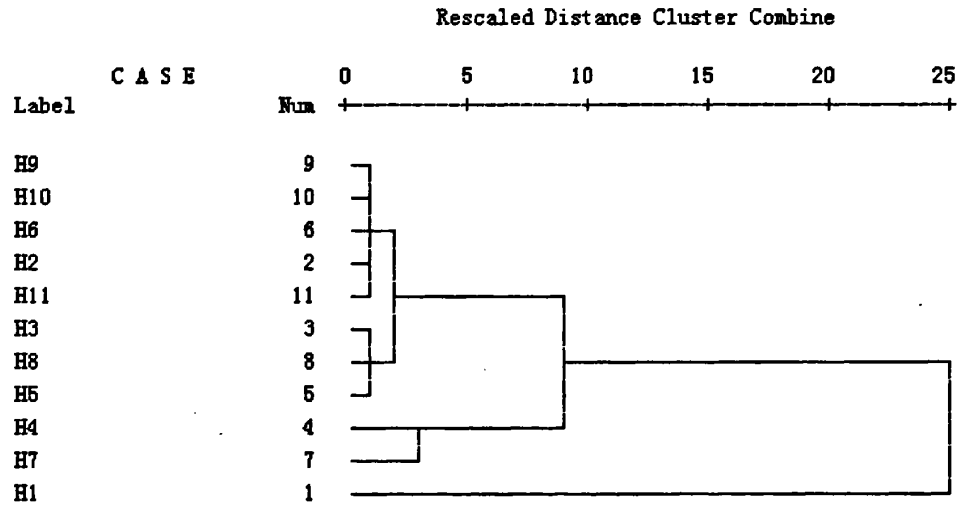


图 4.3 树形图

图 4.3 是反映系统聚类分析全过程的树形图，垂线表示连接的聚类，横线的长短表示被连接的两类之间的距离大小。可以在此图上用一把尺子垂直方向放在图上左右移动，与尺子相交的每一根横线就是一类。每根横线左端与之联系的各观测量就是分到该类的成员。

(5)聚类结果的分析

从树形图（图 4.3）中我们可以看出将这 11 家定点医院分为 3 类时，类间距离比较大，说明各类的特点比较突出，从而能更好的对各类医院的治疗效果进行

定义。

第一类是 H2, H3, H5, H6, H8, H9, H10, H11 这八家医院, 第二类包括 H4, H7 两家医院, 而第三类只包括 H1 这一家医院。

在将这 11 家定点医院分为 3 类以后, 我们对这 3 类医院的各项“治疗效果”评价指标值进行统计描述, 见表 4.3, 并进而可以对这 3 类定点医院的诊疗水平进行比较分析。

表 4.3 三类医院的治疗效果值

|     | $x_1$ | $x_2$ | $x_3$ |
|-----|-------|-------|-------|
| 第一类 | 4.5   | 91.9  | 0.8   |
| 第二类 | 5.6   | 85.5  | 6.9   |
| 第三类 | 63.6  | 36.4  | 0.0   |

通过对表 4.3 中各项指标值的纵向对比, 我们发现第一类医院的特点突出体现在好转率 ( $x_2$ ) 很高, 达到 91.9%, 比第二类和第三类的好转率分别高出 6.4% 和 55.5%, 但与此同时, 该类医院冠心病病情痊愈的患者却只占到 4.5%, 并有少量的医治无效现象 (死亡率达到 0.8%)。综上所述, 该类医院在冠心病方面基本能够使患者得到有效的治疗, 使病情得到好转。

第二类医院的特点突出体现在死亡率 ( $x_3$ ) 较高, 达到 6.9%, 为三类医院中死亡率最高的医院, 此外, 治愈率与好转率两项指标分别三类医院中的中间值, 说明该类医院在冠心病方面治疗效果一般, 并不具备治疗专长。

第三类的特点突出体现在治愈率 ( $x_1$ ) 在三类医院同项指标中最高, 达到 63.6%, 分别为第一类和第二类治愈率的 14.13 倍和 11.36 倍, 且该类医院在我们提取的患者样本中没有死亡现象的发生 (死亡率为 0)。说明该类医院在治疗冠心病患者方面具有一定的治疗专长。

根据表 4.3 中的数据, 第三类医院在冠心病方面的治愈率远高于第一类医院, 但这并不能被简单的理解为第三类医院的诊疗技术和医疗设施远优于第一类医院。因为我们在结合样本患者的实际病例后发现, 在第一类医院就医的患者多为病情较严重 (入院情况为病危或病急), 从而造成病情完全治愈的难度较大, 而在第三类医院就医的患者多为病情较轻 (入院情况为一般), 靠药物治疗就能够达到治愈的效果。但与此同时, 也说明第三类医院在治疗冠心病病情较轻的患者时, 能够满足患者的医疗需求, 也即在采用药物治疗方面具有一定的技术专长, 并能够使患者的病情得到好转或治愈<sup>[41]</sup>。

通过从治疗效果角度对各医院的聚类, 可以将西安市医疗网络中治疗效果相似的医院认定是处于“结构等效”关系的地位, 正是因为各类“结构等效点”医院的治疗效果存在差异性, 使得患者在选择医院进行就医时更愿意与治疗效果好

的医院产生“强联系”，而与哪些治疗效果一般的医院产生“弱联系”，从而影响了患者的行为模式。

#### 4.3.2 从费用水平角度对各医院进行聚类分析

各医院在某一病种上的费用水平是患者和社会医疗保险管理机构对其评价的主要参考指标。如果将患者的治疗效果看成医院所提供医疗服务这一“商品”的质量的话，那么费用水平将是医院所提供医疗服务这一“商品”的价格。患者在选择医院就医时，将根据自身的经济水平、医保待遇选择适合自身条件的医院进行就医，所以治疗费用必定是患者对医院评价的一个关键因素。同时，各医院的费用水平也是社会医疗保险管理机构对其考核的重要评价指标。

此外，在医疗网络模型中，费用水平也是医院节点的一个重要结构特征，各等级医院在费用水平上的差异体现出在与患者行为关系模式与结构特征上的不同，因此，有必要对这一“角色”进行分析。

##### (1) 评价指标体系的建立

本文将从以下几个方面对费用水平进行衡量：实际住院天数（ $x_4$ ）、统筹支付费用（ $x_5$ ）、日平均费用（ $x_6$ ）、床位费（ $x_7$ ）、护理费（ $x_8$ ）、西药费（ $x_9$ ）、化验费（ $x_{10}$ ）、诊疗费（ $x_{11}$ ）、手术费（ $x_{12}$ ）、检查费（ $x_{13}$ ）和手术患者占总患者的比例（ $x_{14}$ ）。

##### (2) 数据处理

该样本的原始数据见表 4.4:

表 4.4 各医院费用水平的原始数据

|     | $x_4$ | $x_5$   | $x_6$  | $x_7$ | $x_8$ | $x_9$  | $x_{10}$ | $x_{11}$ | $x_{12}$ | $x_{13}$ | $x_{14}$ |
|-----|-------|---------|--------|-------|-------|--------|----------|----------|----------|----------|----------|
| H1  | 14.0  | 2035.2  | 217.9  | 171.3 | 100.3 | 1459.1 | 198.1    | 91.8     | 265.0    | 850.2    | 4.6      |
| H2  | 13.6  | 2596.0  | 410.6  | 376.6 | 105.0 | 1960.4 | 318.2    | 382.6    | 0.0      | 796.3    | 0.0      |
| H3  | 11.5  | 7286.5  | 917.1  | 254.2 | 276.9 | 3348.0 | 596.5    | 421.5    | 2965.0   | 1363.9   | 15.4     |
| H4  | 15.2  | 3048.5  | 351.9  | 436.7 | 91.7  | 2574.5 | 602.7    | 817.2    | 45.5     | 428.2    | 8.3      |
| H5  | 11.4  | 11275.1 | 1180.0 | 623.8 | 219.9 | 1817.2 | 539.1    | 221.3    | 5927.3   | 2110.1   | 32.6     |
| H6  | 12.7  | 10379.9 | 1365.9 | 371.1 | 261.4 | 1983.4 | 395.4    | 261.2    | 3166.7   | 1242.1   | 22.2     |
| H7  | 12.5  | 7359.1  | 1396.9 | 291.0 | 131.0 | 2059.0 | 27.9     | 412.7    | 1866.7   | 1209.2   | 23.1     |
| H8  | 12.0  | 2434.9  | 317.9  | 13.1  | 84.6  | 2007.8 | 180.7    | 182.3    | 8.0      | 515.9    | 3.9      |
| H9  | 11.5  | 1495.6  | 188.4  | 111.2 | 90.8  | 1596.9 | 76.0     | 264.0    | 0.0      | 25.4     | 0.0      |
| H10 | 11.4  | 1302.5  | 178.5  | 121.0 | 91.9  | 1247.6 | 85.0     | 255.0    | 0.0      | 20.0     | 0.0      |
| H11 | 12.5  | 3072.4  | 617.9  | 396.7 | 108.0 | 1840.9 | 420.0    | 279.5    | 0.0      | 720.4    | 0.0      |

为了消除各指标值量纲和数量级上的差异，本文选用级差法对数据进行标准

化处理,即  $X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$ , 其中  $X$  为原始变量值,  $X_{\min}$  为该列变量中的最小值,  $X_{\max}$  为该列变量中的最大值。本文通过 SPSS15.0 的 Analyze→Hierarchical Cluster→Standardize (Range 0 to 1) 命令来实现这一数据处理过程<sup>[42]</sup>。

(3)聚类分析的结果图

关于聚类分析的SPSS软件实现过程与上文从治疗效果对医院进行聚类相同, 所以此处不再对实现界面进行赘述。

表 4.5 凝聚过程表

Agglomeration Schedule

| Stage | Cluster Combined |           | Coefficients | Stage Cluster First Appears |           | Next Stage |
|-------|------------------|-----------|--------------|-----------------------------|-----------|------------|
|       | Cluster 1        | Cluster 2 |              | Cluster 1                   | Cluster 2 |            |
| 1     | 9                | 10        | .029         | 0                           | 0         | 3          |
| 2     | 2                | 11        | .162         | 0                           | 0         | 4          |
| 3     | 8                | 9         | .275         | 0                           | 1         | 5          |
| 4     | 1                | 2         | .553         | 0                           | 2         | 5          |
| 5     | 1                | 8         | .798         | 4                           | 3         | 9          |
| 6     | 5                | 6         | .931         | 0                           | 0         | 7          |
| 7     | 3                | 5         | 1.476        | 0                           | 6         | 8          |
| 8     | 3                | 7         | 1.906        | 7                           | 0         | 10         |
| 9     | 1                | 4         | 2.230        | 5                           | 0         | 10         |
| 10    | 1                | 3         | 3.675        | 9                           | 8         | 0          |

Vertical Icicle

| Number of clusters | Case |      |      |      |      |       |      |      |       |      |      |     |     |     |     |     |      |     |
|--------------------|------|------|------|------|------|-------|------|------|-------|------|------|-----|-----|-----|-----|-----|------|-----|
|                    | 7:10 | 6:10 | 5:10 | 3:10 | 4:10 | 10:10 | 9:10 | 8:10 | 11:10 | 2:10 | 1:10 | 7:9 | 6:9 | 5:9 | 3:9 | 4:9 | 10:9 | 9:9 |
| 1                  | X    | X    | X    | X    | X    | X     | X    | X    | X     | X    | X    | X   | X   | X   | X   | X   | X    | X   |
| 2                  | X    | X    | X    | X    | X    | X     | X    | X    | X     | X    | X    | X   | X   | X   | X   | X   | X    | X   |
| 3                  | X    | X    | X    | X    | X    | X     | X    | X    | X     | X    | X    | X   | X   | X   | X   | X   | X    | X   |
| 4                  | X    | X    | X    | X    | X    | X     | X    | X    | X     | X    | X    | X   | X   | X   | X   | X   | X    | X   |
| 5                  | X    | X    | X    | X    | X    | X     | X    | X    | X     | X    | X    | X   | X   | X   | X   | X   | X    | X   |
| 6                  | X    | X    | X    | X    | X    | X     | X    | X    | X     | X    | X    | X   | X   | X   | X   | X   | X    | X   |
| 7                  | X    | X    | X    | X    | X    | X     | X    | X    | X     | X    | X    | X   | X   | X   | X   | X   | X    | X   |
| 8                  | X    | X    | X    | X    | X    | X     | X    | X    | X     | X    | X    | X   | X   | X   | X   | X   | X    | X   |
| 9                  | X    | X    | X    | X    | X    | X     | X    | X    | X     | X    | X    | X   | X   | X   | X   | X   | X    | X   |
| 10                 | X    | X    | X    | X    | X    | X     | X    | X    | X     | X    | X    | X   | X   | X   | X   | X   | X    | X   |

图 4.4 冰柱图

Dendrogram using Average Linkage (Between Groups)

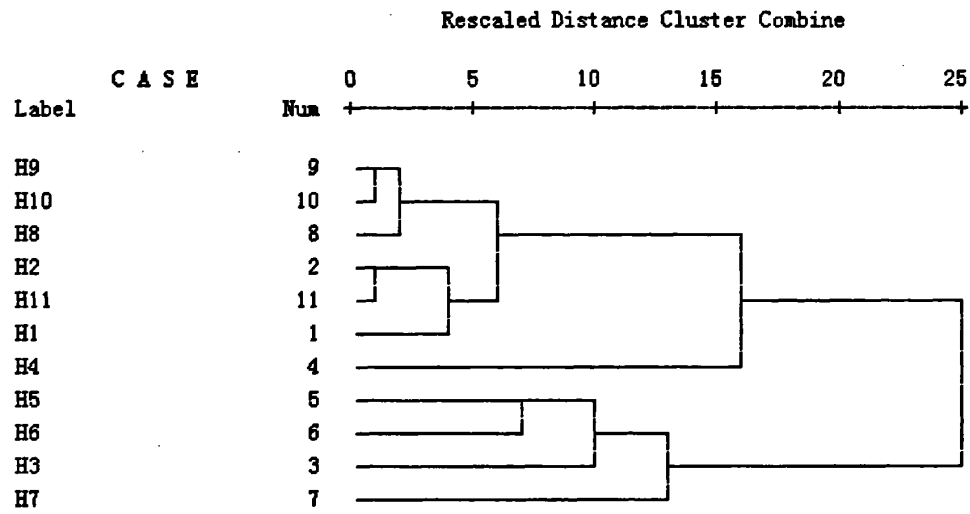


图 4.5 树形图

(4)聚类结果的分析

从树形图（图 4.5）中我们可以看出将这 11 家定点医院分为 5 类时，类间距离比较大，说明各类的特点比较突出，从而能更好的对各类医院的费用水平进行定义。

第一类包括 H1, H2, H8, H9, H10, H11 这六家医院，第二类只有 H4 一家医院，第三类包括 H5, H6 两家医院，第四类包括 H3 一家医院，第五类只包括 H7 一家医院。

在将这 11 家定点医院分为 5 类以后，我们对这 5 类医院的各项“费用水平”评价指标进行统计描述，见表 4.6，并进而可以对这 5 类定点医院的医疗费用水平进行比较分析。

表 4.6 五类医院费用水平值

|     | $x_4$ | $x_5$   | $x_6$  | $x_7$ | $x_8$ | $x_9$  | $x_{10}$ | $x_{11}$ | $x_{12}$ | $x_{13}$ | $x_{14}$ |
|-----|-------|---------|--------|-------|-------|--------|----------|----------|----------|----------|----------|
| 第一类 | 12.5  | 2156.1  | 321.9  | 198.3 | 96.8  | 1685.5 | 212.9    | 242.5    | 45.5     | 488.0    | 1.4      |
| 第二类 | 15.2  | 3048.5  | 351.9  | 436.7 | 91.7  | 2574.5 | 602.7    | 817.2    | 45.5     | 428.2    | 8.3      |
| 第三类 | 12.0  | 10827.5 | 1272.9 | 497.5 | 240.7 | 1900.3 | 467.2    | 241.2    | 4547.0   | 1676.1   | 27.4     |
| 第四类 | 11.5  | 7286.5  | 917.1  | 254.2 | 276.9 | 3348.0 | 596.5    | 421.5    | 2965.0   | 1363.9   | 15.4     |
| 第五类 | 12.5  | 7359.1  | 1396.9 | 291.0 | 131.0 | 2059.0 | 27.9     | 412.7    | 1866.7   | 1209.2   | 23.1     |

通过对表 4.6 中各项指标值的纵向对比，我们发现第一类医院的特点突出体现在医疗费用在五类医院中最低，并且手术患者比例（ $x_{14}$ ）最低。其中，个人统筹支付费用（ $x_5$ ）平均值为 2156.1 元，较第三类医院同类指标值低了近 8000 元，且手术患者的比例仅为 1.8%，较第三类医院同类指标值低了 26%，说明该医院在

治疗冠心病患者时多采用非手术治疗的方式,且各单项费用也低于其它类医院。通过对第一类医院构成分析,我们发现该类医院主要由三级乙等(H2和H11)、二级甲等(H1和H8)和一级医院(H9和H10)组成。此外,通过树形图(图4.5),我们进一步发现,H8、H9、H10医院先聚为一类,H1、H2、H11医院先聚为一类,说明H8、H9、H10在费用水平上更相似,同理,H1、H2、H11在费用水平上也更为相似。说明如果从费用水平上进一步划分该类医院节点,那么可以将二级甲等医院和一级医院聚为一类,而三级乙等单独作为一类。

第二类医院由医院H4代表,它的特点主要体现在住院天数( $x_4$ )在五类医院中最高,为15.2天,个人统筹支付费用( $x_5$ )平均值虽略高于第一类医院,但仍在五类医院中属于费用较低水平。通过对第一类和第二类医院日平均费用( $x_6$ )的对比发现,日平均费用分别为321.9元和351.9元,说明两类医院的医疗费用标准基本相同,只是由于第二类医院住院天数较多,从而引发了更多的医疗费用,致使床位费( $x_7$ )、护理费( $x_8$ )、西药费( $x_9$ )等各单项费用的增长,最终导致了第二类医院的总费用略高于第一类医院。此外,需要说明的两个问题是:一、该类医院的诊疗费用( $x_{11}$ )在五类医院同类指标中最高,为817.2元,较第一类、第三类医院和第四类医院、第五类医院分别高出600元和400元。二、该类医院的手术患者比例( $x_{14}$ )也较低,只有8.3%,说明该类医院在治疗冠心病方面也多采用非手术治疗的方式。

第三类医院由医院H5、H6组成,这类医院的特点主要体现在个人统筹支付费用( $x_5$ )最高,为10827.5元;且手术患者比例( $x_{14}$ )最高,为27.4%,因此而引发的手术费用( $x_{12}$ )也是在五类医院中最高,为4547.0元。说明该类医院在冠心病方面多采用手术治疗方式。通过对该类医院医疗费用的横向比较发现,手术费用占到个人统筹支付费用的近50%,因此,手术费用是构成个人统筹支付费用的主要因素。通过与第一类和第二类医院费用水平的纵向对比发现,床位费( $x_7$ )、护理费( $x_8$ )等各单项费用也高于第一类和第二类医院。从该类医院的组成来看,该类医院是由两家三级甲等医院构成,因此,从第三类与第一类、第二类医院各单项费用水平的纵向对比中,我们可以量化出三级甲等医院医疗费用在冠心病方面较三级乙等、二级甲等、一级医院高出的程度。

第四类医院由H3代表,这类医院的特点主要体现在个人统筹支付费用( $x_5$ )较高,为7286.5元,但住院天数( $x_4$ )在五类医院中最低,为11.5天。此外,该类医院H3为三级乙等医院,在与其它三级甲等医院医疗费用的纵向对比中,手术患者比例( $x_{14}$ )这一指标属于较低值,为15.4%,但西药费用( $x_9$ )却在五类医院中最高,为3348.0元,说明该类医院多采用手术与药物治疗相结合的方式,且所采用的药物价格较高。此外,日平均费用( $x_6$ )为917.1元,相较于其它三级医院,属于较低值,说明该医院在冠心病方面的费用标准在三级医院中属于较

低水平。

第五类医院由 H7 代表,这类医院的特点主要体现在个人统筹支付费用 ( $x_5$ ) 较高,为 7359.1 元,且手术患者比例 ( $x_{14}$ ) 较高,为 23.1%,这也是导致其总体费用较高的主要原因。H7 作为一家三级乙等医院,手术患者比例大从一个侧面说明该家医院在冠心病方面具有手术治疗的能力。此外值得一提的是,该类医院的日平均费用 ( $x_6$ ) 在五类医院同类指标中最高,为 1396.9 元,说明在冠心病方面,该类医院的医疗费用水平较高。

通过从费用水平角度对各医院进行聚类,可以将西安市医疗网络中治疗效果相似的医院认定是处于“结构等效”关系的地位,从而对医院在医疗网络中的角色有了更进一步的了解,有助于我们理解医院在面对患者时表现出来的固定行为模式。此外,正是因为各类“结构等效点”医院的费用水平存在差异性,使得患者在选择医院进行就医时更愿意与费用水平低(若不考虑医院治疗效果)的医院建立“强联系”,而只与费用水平高的医院建立“弱联系”,从而影响了患者的行为模式。

#### 4.4 本章小结

本章运用社会网络分析理论,从位置取向角度对西安市医疗网络模型中的医院节点进行分析,主要分析内容主要包括三个方面:结构等效分析、位置分析和角色分析。

其中,在结构等效分析方面,从宏观角度可以将医疗网络中的各医院看作是具有结构等效的节点,从微观角度,认定只有与某一单个患者发生诊疗关系的医院才具有结构等效;在位置分析方面,将具有相同等级的医院认定具有相同的“位置”;在角色分析中,通过运用系统聚类分析的方法,分别对西安市 11 家定点医院在冠心病方面的治疗效果和费用水平进行实证研究,并据聚类分析的结果刻画出西安市各等级医院在冠心病治疗效果和费用水平的结构性特征,从而为把握西安市医疗网络中各医院节点的“角色”提供了参考。

## 第五章 对西安市医疗网络模型患者节点的位置取向分析

在社会网络分析理论中,位置取向是以结构上的相似为基点,以关系的相似性为基本特征。在位置取向看来,位置所反映出来的结构性特征更加稳定和持久,更具有普遍性,因而对现实也更有解释力,且需要分析的内容也就更为简单明了。本章试图从位置取向角度对西安市医疗网络模型中患者节点的结构等效、位置、角色进行研究,从而进一步明晰患者在医院就医过程中所表现出的结构特征和属性。

### 5.1 对患者节点的“结构等效”分析

根据上文社会网络分析理论中对“结构等效”概念的介绍,针对本文所研究的医疗网络模型,从总体上看,如果将全部医院看成为一个节点,那么,可以将各个患者节点看成是所谓的“结构等效点”,因为它们与医院之间都是“患者就医——医院诊疗”的关系,正是基于这些相同的关系作为纽带(即使各个患者之间不存在关系),我们可以认为各个患者之间具有“结构等效”。从具体上看,即将每个医院都看成是一个单独的节点,那么,只有在同一医院就医的患者之间才可以称之为“结构等效”,而那些在不同医院就医的患者节点之间不属于“结构等效”的定义范畴。

根据社会网络分析理论,网络中“结构等效点”的数量和质量将对网络的驱动力产生很大的影响。针对本文医疗网络模型所研究的患者节点,这一点也是适用的。在医疗网络模型中,患者的数量直接影响到医院的接诊人数和收入来源,如果患者数量多,则医院将产生更多的医疗收入,可以进一步用于对医疗设备的改善和医疗技术的提高,从而推动医院诊疗水平的不断发展,因此,患者的数量对医疗网络具有驱动力。此外,患者的“质量”间接帮助了医院诊疗水平的提升,如果患者病情严重,那么医生为了诊疗成功将努力钻研,进而有助于其提高诊疗技术,推动医院整体诊疗技术的不断提升。

### 5.2 对患者节点的“位置”分析

根据上文社会网络分析理论中对“位置”概念的介绍,针对本文的医疗网络模型,本文采用年龄和性别这两项指标对患者的“位置”进行定义。“位置”是指剥落了行动者而剩下的结构性特征,哪个行动者处在这个位置并不重要,重要的是这个位置在网络本身中的处境。在医疗网络模型中,本文认为具有相同性别或年龄段的患者具有相同的“位置”,即这一年龄段的人群容易患有某些相同的高发



疾病,处于相同的“处境”。从性别对患者节点的“位置”进行定义,我们可以将患者节点划分成男、女两个节点;从年龄段进行划分,由于在我们提取到的患者样本数据中,最小的患者 23 岁,最大的 95 岁,因此,我们可以将患者根据年龄段分为 20~29、30~39、40~49、50~59、60~69、70~79、80~89、90~99 这八个节点,每个年龄段的患者具有相同的“位置”。

本文从性别和年龄段两个角度对患者节点的“位置”进行定义,试图分析具有相同“位置”的节点在与医院发生“患者就医——医院诊疗”关系的过程中表现出哪些固定的行为模式和相同的结构属性,也即相同性别、相同年龄段的患者容易患有哪些疾病,入院情况如何。下面,本章将对处于同一“位置”的患者在与医院发生“患者就医——医院诊疗”关系的过程中所表现出的“角色”进行更进一步的分析。

### 5.3 角色分析

根据上文对网络节点“角色”的定义,应用在本文所研究的医疗网络模型中,对患者节点的“角色”分析是指患者在与医院建立“患者就医——医院诊疗”关系的过程中所表现出来的固定行为模式,这种固定的行为模式主要体现在患者的患病种类和入院情况。由于医疗费用为患者被动接受,因此,在考虑患者主动采取的“行为模式”中对医疗费用不予考虑。下面,本章将运用频数分析、均值比较和系统聚类分析的方法对西安市医疗网络模型中患者节点的“角色”进行实证研究。

#### 5.3.1 一维频数分析

通过对病种、患者入院情况、性别、年龄段、住院次数进行一维频数分析<sup>[43]</sup>,我们可以从宏观上了解西安市医疗网络中患者节点的各项属性,进而有助于我们了解该医疗网络中患者的“角色”,以及其相应与医院在发生“患者就医——医院诊疗”关系过程中的固定行为模式。

##### (1)对病种的频数分析

在本文对西安市各医院患者提取的样本数据中,通过应用 SPSS15.0 软件进行频数分析,结果见图 5.2,实现界面见图 5.1。

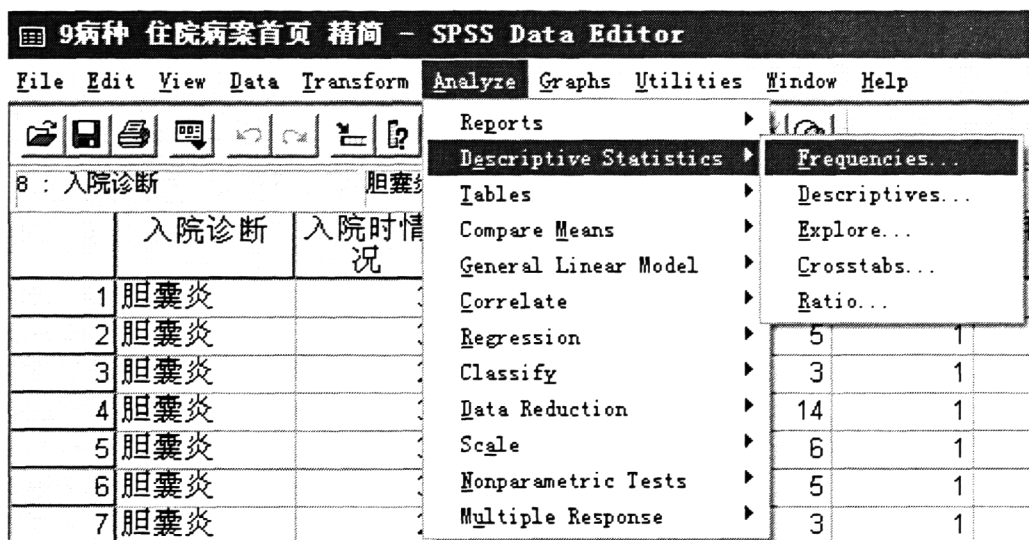


图 5.1 频数分析实现界面一

| 入院诊断  |       |           |         |               |                    |
|-------|-------|-----------|---------|---------------|--------------------|
|       |       | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | 胆囊炎   | 46        | 9.4     | 9.4           | 9.4                |
|       | 肺癌    | 52        | 10.6    | 10.6          | 20.0               |
|       | 高血压   | 72        | 14.7    | 14.7          | 34.8               |
|       | 冠心病   | 131       | 26.8    | 26.8          | 61.6               |
|       | 脑梗塞   | 55        | 11.2    | 11.2          | 72.8               |
|       | 乳癌    | 29        | 5.9     | 5.9           | 78.7               |
|       | 糖尿病   | 43        | 8.8     | 8.8           | 87.5               |
|       | 眩晕    | 32        | 6.5     | 6.5           | 94.1               |
|       | 支气管炎  | 29        | 5.9     | 5.9           | 100.0              |
|       | Total | 489       | 100.0   | 100.0         |                    |

图 5.2 病种频数分析结果图

由于患有胆囊炎、肺癌、高血压、冠心病、脑梗塞、乳癌、糖尿病、眩晕、支气管炎这九种疾病的患者占到本次实证研究所提取患者样本总数的 61.9%，因此，可以认为这九种疾病在西安市患者当中属于高发疾病。根据图 5.2，我们发现在这九种高发疾病中，又尤以冠心病的患病率最高，在九种疾病患者中占到 26.8%。因此，可以认为在西安市医疗网络模型中，患者节点最有可能患有冠心病，即最有可能通过冠心病与医院发生“患者就医——医院诊疗”关系。

#### (2) 对入院情况的频数分析

我们根据患者病例，将入院情况分为三种情况：病危、病急、一般。其中，在上述九种高发疾病的患者样本中，入院情况为“病危”的患者占到患者样本总数的 5.3%，“病急”的患者占到 27.8%，“一般”的患者占到 66.9%，说明西安市医疗网络中患者的入院情况多为“一般”。

入院时情况

|          | Frequency | Percent | Valid Percent | Cumulative Percent |
|----------|-----------|---------|---------------|--------------------|
| Valid 病危 | 26        | 5.3     | 5.3           | 5.3                |
| 病急       | 136       | 27.8    | 27.8          | 33.1               |
| 一般       | 327       | 66.9    | 66.9          | 100.0              |
| Total    | 489       | 100.0   | 100.0         |                    |

图 5.3 患者入院情况频数分析图

## (3)对患者性别的频数分析

在上述九种高发疾病的患者样本中, 男性患者占到 55.0%, 女性患者占到 45.0%, 从中我们可以明晰西安市医疗网络模型中各患者节点“位置”的规模。

性别

|         | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|-----------|---------|---------------|--------------------|
| Valid 男 | 269       | 55.0    | 55.0          | 55.0               |
| 女       | 220       | 45.0    | 45.0          | 100.0              |
| Total   | 489       | 100.0   | 100.0         |                    |

图 5.4 患者性别频数分析图

## (4)对患者年龄段的频数分析

根据上文对患者年龄段的划分, 我们将 20~29、30~39、40~49、50~59、60~69、70~79、80~89、90~99 八个年龄段分别由 2~9 这八个数字来代表。通过图 5.5, 我们可以看出在西安市患者中, 60~69、70~79 这两个年龄段最容易患病, 分别占到患者样本总数的 29.0%和 29.7%。

年龄段

|         | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|-----------|---------|---------------|--------------------|
| Valid 2 | 2         | .4      | .4            | .4                 |
| 3       | 21        | 4.3     | 4.3           | 4.7                |
| 4       | 70        | 14.3    | 14.3          | 19.0               |
| 5       | 94        | 19.2    | 19.2          | 38.2               |
| 6       | 142       | 29.0    | 29.0          | 67.3               |
| 7       | 145       | 29.7    | 29.7          | 96.9               |
| 8       | 14        | 2.9     | 2.9           | 99.8               |
| 9       | 1         | .2      | .2            | 100.0              |
| Total   | 489       | 100.0   | 100.0         |                    |

图 5.5 患者年龄段频数分析图

## (5)对患者住院次数的频数分析

通过图 5.6, 我们发现在西安市医疗网络中, 第一次入院诊疗的患者占到 59.1%, 说明患者多为首次住院接受诊疗。

|         |        | 住院次数      |         |               |                    |
|---------|--------|-----------|---------|---------------|--------------------|
|         |        | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid   | 1      | 289       | 59.1    | 59.5          | 59.5               |
|         | 2      | 103       | 21.1    | 21.2          | 80.7               |
|         | 3      | 39        | 8.0     | 8.0           | 88.7               |
|         | 4      | 26        | 5.3     | 5.3           | 94.0               |
|         | 5      | 9         | 1.8     | 1.9           | 95.9               |
|         | 6      | 9         | 1.8     | 1.9           | 97.7               |
|         | 7      | 6         | 1.2     | 1.2           | 99.0               |
|         | 8      | 2         | .4      | .4            | 99.4               |
|         | 9      | 3         | .6      | .6            | 100.0              |
| Total   |        | 486       | 99.4    | 100.0         |                    |
| Missing | System | 3         | .6      |               |                    |
| Total   |        | 489       | 100.0   |               |                    |

图 5.6 患者住院次数的频数分析图

5.3.2 二维频数分析和均值比较

本小节从入院情况和病种两个角度分别对患者性别和年龄段进行二维的频数分析和均值比较。其中，通过从入院情况角度对患者性别和年龄段进行二维频数分析，可以明晰入院情况分别为病危、病急和一般的患者数量在各性别和年龄段上的分布，并对各入院情况的患者性别和年龄进行均值比较；同理，从病种角度对患者性别和年龄段进行二维频数分析可以明晰各病种的患者数量在各性别和年龄段上的分布，并对各病种的患者性别和年龄进行均值比较。

(1)从入院情况角度对患者性别和年龄段的二维频数分析和均值比较

| 性别·年龄段·入院时情况 Cross tabulation |       |     |    |    |    |    |    |    |   |       |
|-------------------------------|-------|-----|----|----|----|----|----|----|---|-------|
| Count                         |       | 年龄段 |    |    |    |    |    |    |   | Total |
| 入院时情况                         |       | 2   | 3  | 4  | 5  | 6  | 7  | 8  | 9 |       |
| 病危                            | 性别 男  |     |    |    | 3  | 7  | 6  | 0  |   | 16    |
|                               | 女     |     |    |    | 0  | 4  | 5  | 1  |   | 10    |
|                               | Total |     |    |    | 3  | 11 | 11 | 1  |   | 26    |
| 病急                            | 性别 男  |     | 3  | 5  | 15 | 24 | 29 | 2  |   | 78    |
|                               | 女     |     | 2  | 8  | 8  | 20 | 21 | 1  |   | 58    |
|                               | Total |     | 5  | 13 | 21 | 44 | 50 | 3  |   | 136   |
| 一般                            | 性别 男  | 1   | 7  | 24 | 39 | 54 | 43 | 6  | 1 | 175   |
|                               | 女     | 1   | 9  | 33 | 31 | 33 | 41 | 4  | 0 | 152   |
|                               | Total | 2   | 16 | 57 | 70 | 87 | 84 | 10 | 1 | 327   |

图 5.7 各入院情况下患者在各性别和年龄段上的频数分析图

从图 5.7 中，我们可以发现在西安市医疗网络中的病危和病急患者中，多为 60~79 年龄段，说明这一年龄段的患者最容易患上上述九种高发疾病，且患病程度较为严重。同时，本文发现从 40~49 年龄段到 70~79 年龄段的一般病情患者数持续快速增长，但 30~39 年龄段的患者数与 40~49 年龄段的患者数差距却很大，说明西安市居民从 40 岁开始患上上述九种高发疾病的概率逐渐增大，因此，西安市居民从 40 岁开始有必要更加重视对自己的身体检查，预防上述九种疾病。此外，

通过图 5.7，我们发现各种入院情况下的男性患者都高于女性患者，说明男性患者较之女性患者更容易患上上述九种高发疾病。

| Report |                |        |      |
|--------|----------------|--------|------|
| 入院时情况  |                | 年龄     | 性别   |
| 病危     | Mean           | 68.58  | 1.38 |
|        | Std. Deviation | 6.500  | .496 |
| 病急     | Mean           | 63.95  | 1.43 |
|        | Std. Deviation | 10.895 | .496 |
| 一般     | Mean           | 60.16  | 1.46 |
|        | Std. Deviation | 12.600 | .500 |
| Total  | Mean           | 61.66  | 1.45 |
|        | Std. Deviation | 12.107 | .498 |

图 5.8 各入院情况下患者年龄和性别的均值比较

通过图 5.8，入院情况为一般、病急和病危患者平均年龄分别为 61.66、63.95、68.58 岁，呈增长趋势，说明西安市居民随着年龄的增长，其病情的严重程度也逐渐加重。此外，需要说明的是，本文在对性别进行定量分析时将男性患者定义为 1，女性患者定义为 2，因此，病危患者的性别平均值为 1.38 说明在病危患者中男性占较大比重。

(2)从病种角度对患者性别和年龄段的二维频数分析和均值比较

| 性别·年龄段·入院诊断 Cross-tabulation |       |     |   |    |    |    |    |   |   |       |
|------------------------------|-------|-----|---|----|----|----|----|---|---|-------|
| Count                        |       | 年龄段 |   |    |    |    |    |   |   | Total |
| 入院诊断                         | 性别    | 2   | 3 | 4  | 5  | 6  | 7  | 8 | 9 |       |
| 胆囊炎                          | 男性    | 1   | 6 | 8  | 3  | 5  | 3  |   |   | 26    |
|                              | 女性    | 0   | 3 | 4  | 6  | 5  | 2  |   |   | 20    |
|                              | Total | 1   | 9 | 12 | 9  | 10 | 5  |   |   | 46    |
| 肺癌                           | 男性    |     |   | 0  | 6  | 18 | 12 | 2 |   | 36    |
|                              | 女性    |     |   | 1  | 2  | 4  | 7  | 2 |   | 16    |
|                              | Total |     |   | 1  | 8  | 20 | 19 | 4 |   | 52    |
| 高血压                          | 男性    |     | 2 | 4  | 11 | 11 | 11 | 1 |   | 40    |
|                              | 女性    |     | 0 | 7  | 7  | 6  | 12 | 0 |   | 32    |
|                              | Total |     | 2 | 11 | 18 | 17 | 23 | 1 |   | 72    |
| 冠心病                          | 男性    |     |   | 3  | 15 | 24 | 24 | 3 | 1 | 70    |
|                              | 女性    |     |   | 6  | 9  | 25 | 19 | 2 | 0 | 61    |
|                              | Total |     |   | 9  | 24 | 49 | 43 | 5 | 1 | 131   |
| 脑梗塞                          | 男性    |     |   | 5  | 4  | 13 | 13 | 1 |   | 36    |
|                              | 女性    |     |   | 1  | 1  | 5  | 10 | 2 |   | 19    |
|                              | Total |     |   | 6  | 5  | 18 | 23 | 3 |   | 55    |
| 乳癌                           | 男性    | 0   | 1 | 0  | 0  | 0  | 0  |   |   | 1     |
|                              | 女性    | 1   | 3 | 14 | 5  | 2  | 3  |   |   | 28    |
|                              | Total | 1   | 4 | 14 | 5  | 2  | 3  |   |   | 29    |
| 糖尿病                          | 男性    |     | 0 | 4  | 8  | 7  | 6  |   |   | 25    |
|                              | 女性    |     | 1 | 4  | 4  | 3  | 6  |   |   | 18    |
|                              | Total |     | 1 | 8  | 12 | 10 | 12 |   |   | 43    |
| 眩晕                           | 男性    |     | 1 | 3  | 4  | 3  | 3  |   |   | 14    |
|                              | 女性    |     | 2 | 4  | 3  | 5  | 4  |   |   | 18    |
|                              | Total |     | 3 | 7  | 7  | 8  | 7  |   |   | 32    |
| 支气管炎                         | 男性    |     | 0 | 2  | 6  | 6  | 6  | 1 |   | 21    |
|                              | 女性    |     | 2 | 0  | 0  | 2  | 4  | 0 |   | 8     |
|                              | Total |     | 2 | 2  | 6  | 8  | 10 | 1 |   | 29    |

图 5.9 各病种患者在各性别和年龄段上的频数分析图

从图 5.9 中，我们可以总结出各病种的高发年龄段，例如，胆囊炎的高发年龄段为 30~69，跨度较大；乳癌的高发年龄段为 40~49；眩晕的高发年龄段为 40~

79, 年龄跨度较大; 高血压、糖尿病、支气管炎的高发年龄段为 50~79; 肺癌、冠心病、脑梗塞的高发年龄段为 60~79; 总结起来, 患有胆囊炎和眩晕的患者年龄跨度较大, 乳癌属于“中年疾病”, 高血压、糖尿病、支气管炎、肺癌、冠心病、脑梗塞属于“老年疾病”。此外, 在各种疾病的患者人数上, 男性患者基本都多于女性, 但乳癌除外 (因为乳癌是妇科疾病)。

Report

| 入院诊断  |                | 年龄     | 性别   |
|-------|----------------|--------|------|
| 胆囊炎   | Mean           | 51.39  | 1.43 |
|       | Std. Deviation | 13.247 | .501 |
| 肺癌    | Mean           | 66.83  | 1.31 |
|       | Std. Deviation | 8.554  | .486 |
| 高血压   | Mean           | 61.53  | 1.44 |
|       | Std. Deviation | 11.545 | .500 |
| 冠心病   | Mean           | 65.45  | 1.47 |
|       | Std. Deviation | 9.481  | .501 |
| 脑梗塞   | Mean           | 66.44  | 1.35 |
|       | Std. Deviation | 10.169 | .480 |
| 乳癌    | Mean           | 48.14  | 1.97 |
|       | Std. Deviation | 10.680 | .186 |
| 糖尿病   | Mean           | 59.95  | 1.42 |
|       | Std. Deviation | 10.488 | .499 |
| 眩晕    | Mean           | 57.44  | 1.56 |
|       | Std. Deviation | 13.176 | .504 |
| 支气管炎  | Mean           | 63.55  | 1.28 |
|       | Std. Deviation | 12.673 | .455 |
| Total | Mean           | 61.66  | 1.45 |
|       | Std. Deviation | 12.107 | .498 |

图 5.10 各病种患者年龄和性别的均值比较

从图 5.10 中, 我们可以看出各种疾病患者的年龄和性别均值, 这从另外一个角度印证了上文对各疾病患者在年龄和性别上的频数分析结论。

### 5.3.3 三维频数分析和均值比较

本小节以冠心病患者为例, 对该病种下各入院情况的患者性别和年龄段进行三维频数分析。对于其它疾病各入院情况的患者性别和年龄段的三维频数分析与此同理, 本文不多作赘述。

| 性别·年龄段·入院时情况·入院诊断 Crosstabulation |       |       |   |    |    |    |   |       |    |
|-----------------------------------|-------|-------|---|----|----|----|---|-------|----|
| Count                             |       | 年龄段   |   |    |    |    |   | Total |    |
| 入院诊断                              | 入院时情况 | 性别    | 4 | 5  | 6  | 7  | 8 |       | 9  |
| 冠心病                               | 病危    | 男     |   | 1  | 4  | 4  |   |       | 9  |
|                                   |       | 女     |   | 0  | 3  | 1  |   |       | 4  |
|                                   |       | Total |   | 1  | 7  | 5  |   |       | 13 |
|                                   | 病急    | 男     | 0 | 4  | 12 | 10 | 1 |       | 27 |
|                                   |       | 女     | 2 | 3  | 9  | 10 | 1 |       | 25 |
|                                   |       | Total | 2 | 7  | 21 | 20 | 2 |       | 52 |
|                                   | 一般    | 男     | 3 | 10 | 8  | 10 | 2 | 1     | 34 |
|                                   |       | 女     | 4 | 6  | 13 | 8  | 1 | 0     | 32 |
|                                   |       | Total | 7 | 16 | 21 | 18 | 3 | 1     | 66 |

图 5.11 冠心病下各入院情况患者性别和年龄段的频数分析图

| Report |       |                |        |      |
|--------|-------|----------------|--------|------|
| 入院诊断   | 入院时情况 |                | 年龄     | 性别   |
| 冠心病    | 病危    | Mean           | 69.08  | 1.31 |
|        |       | Std. Deviation | 6.614  | .480 |
|        | 病急    | Mean           | 66.79  | 1.48 |
|        |       | Std. Deviation | 7.915  | .505 |
|        | 一般    | Mean           | 63.68  | 1.48 |
|        |       | Std. Deviation | 10.739 | .504 |
|        | Total | Mean           | 65.45  | 1.47 |
|        |       | Std. Deviation | 9.481  | .501 |

图 5.12 冠心病下各入院情况患者性别和年龄段的均值比较图

从图 5.11 中，我们可以看出冠心病患者在入院时病情多为病急和一般，且病危和病急患者的年龄段多集中在 60~79，而在入院病情一般的患者中，年龄段多集中在 50~79。此外，在各入院情况下和年龄段下，男性患者始终多于女性，说明较之女性，男性更容易患冠心病。

从图 5.12 中，我们可以看出对于冠心病，病危患者的平均年龄为 69.08，病急患者的平均年龄为 66.79，病情一般的患者平均年龄为 63.68，说明高年龄段的冠心病患者病情较严重。此外，病危患者中性别平均值为 1.31，小于其它两种入院情况的性别平均值，说明在病情危急的患者中，男性所占比例更高。

综上所述，具有相同性别、年龄段的患者为具有相同“位置”的节点，而“角色”分析主要指患者在与医院发生关系的过程中所表现出的固定行为模式，即可以从患者所患病种和入院情况两个角度进行对患者节点的“角色”分析。因此，通过 5.3.1、5.3.2、5.3.3 节对各“位置”患者所患病种和入院情况的频数分析和均值比较，我们可以对西安市医疗网络中患者节点的“角色”（即固定的行为模式）有进一步更为量化的把握，并有助于明晰西安市患者的结构性特征。

5.3.4 对病种的聚类分析

(1)指标体系及数据来源

本文在对病种进行聚类分析时,采用各病种患者的年龄、性别、入院情况、住院次数作为对病种评价的指标变量,试图通过系统聚类方法分析患各种疾病的患者特点,将具有相同患者特点的疾病聚为一类,从而为我们进一步了解西安市医疗网络中患者的属性提供了参考。

本节从西安市 11 家定点医院中提取了九种高发疾病的患者数据,运用 SPSS15.0 中的 Analyze→Compare Means 命令对各病种的患者年龄、性别、入院情况、住院次数进行均值比较,得到下图(图 5.13),并以此作为对各病种进行系统聚类分析的数据。此外,需要说明的是,我们对入院情况进行了量化的处理,分别将入院情况为病危、病急和一般定义为数值 1、2、3。

**Report**

| Mean  |       |      |       |      |
|-------|-------|------|-------|------|
| 入院诊断  | 年龄    | 性别   | 入院时情况 | 住院次数 |
| 胆囊炎   | 51.39 | 1.43 | 2.74  | 1.15 |
| 肺癌    | 66.63 | 1.31 | 2.87  | 4.00 |
| 高血压   | 61.53 | 1.44 | 2.67  | 1.41 |
| 冠心病   | 65.45 | 1.47 | 2.40  | 1.64 |
| 脑梗塞   | 66.44 | 1.35 | 2.64  | 1.45 |
| 乳癌    | 48.14 | 1.97 | 2.97  | 3.41 |
| 糖尿病   | 59.95 | 1.42 | 2.77  | 1.37 |
| 眩晕    | 57.44 | 1.56 | 2.59  | 1.31 |
| 支气管炎  | 63.55 | 1.28 | 2.21  | 1.62 |
| Total | 61.66 | 1.45 | 2.62  | 1.85 |

图 5.13 各病种患者年龄、性别、入院情况、住院次数的均值比较图

## (2)数据处理

为了消除各指标量纲和数量级上的差异,本文选用级差法对数据进行标准化处理,即  $X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$ , 其中  $X$  为原始变量值。 $X_{\min}$  为该列变量中最小值,  $X_{\max}$  为该列变量中的最大值。

这一数据处理过程,本文通过 SPSS15.0 的 Analyze→Hierarchical Cluster→Standardize (Range 0 to 1) 命令来实现。

## (3)聚类分析结果

此聚类分析的实现界面与第四章对医院进行聚类分析的实现界面相同,所以本文此处不再赘述。下面是对病种进行系统聚类分析后的三个结果图:凝聚过程表、冰柱图和树形图。



表 5.1 凝聚过程表

| Agglomeration Schedule |                  |           |              |                             |           |            |
|------------------------|------------------|-----------|--------------|-----------------------------|-----------|------------|
| Stage                  | Cluster Combined |           | Coefficients | Stage Cluster First Appears |           | Next Stage |
|                        | Cluster 1        | Cluster 2 |              | Cluster 1                   | Cluster 2 |            |
| 1                      | 3                | 7         | .025         | 0                           | 0         | 2          |
| 2                      | 3                | 8         | .103         | 1                           | 0         | 4          |
| 3                      | 4                | 5         | .137         | 0                           | 0         | 4          |
| 4                      | 3                | 4         | .229         | 2                           | 3         | 5          |
| 5                      | 3                | 9         | .421         | 4                           | 0         | 6          |
| 6                      | 1                | 3         | .531         | 0                           | 5         | 7          |
| 7                      | 1                | 2         | 1.243        | 6                           | 0         | 8          |
| 8                      | 1                | 6         | 1.992        | 7                           | 0         | 0          |

表 5.1 是系统聚类分析的凝聚过程表，Stage 这一列给出聚类分析的步骤，本例共 8 个步骤；Cluster Combined 这一列表示聚类过程；Coefficients 这一列表示样本或类之间的距离，距离最短的样本或类先聚成一类。

|                    | Case  |       |        |       |       |        |        |          |        |        |         |         |         |        |        |
|--------------------|-------|-------|--------|-------|-------|--------|--------|----------|--------|--------|---------|---------|---------|--------|--------|
|                    | 9.乳腺癌 | 10.肺癌 | 11.食管癌 | 12.胃癌 | 13.肝癌 | 14.胰腺癌 | 15.宫颈癌 | 16.子宫内膜癌 | 17.卵巢癌 | 18.鼻咽癌 | 19.甲状腺癌 | 20.前列腺癌 | 21.结直肠癌 | 22.膀胱癌 | 23.胆管炎 |
| Number of clusters |       |       |        |       |       |        |        |          |        |        |         |         |         |        |        |
| 1                  | X     | X     | X      | X     | X     | X      | X      | X        | X      | X      | X       | X       | X       | X      | X      |
| 2                  | X     |       | X      | X     | X     | X      | X      | X        | X      | X      | X       | X       | X       | X      | X      |
| 3                  | X     |       | X      |       | X     | X      | X      | X        | X      | X      | X       | X       | X       | X      | X      |
| 4                  | X     |       | X      |       | X     | X      | X      | X        | X      | X      | X       | X       | X       | X      | X      |
| 5                  | X     |       | X      |       | X     | X      | X      | X        | X      | X      | X       | X       | X       | X      | X      |
| 6                  | X     |       | X      |       | X     | X      | X      | X        | X      | X      | X       | X       | X       | X      | X      |
| 7                  | X     |       | X      |       | X     | X      | X      | X        | X      | X      | X       | X       | X       | X      | X      |
| 8                  | X     |       | X      |       | X     | X      | X      | X        | X      | X      | X       | X       | X       | X      | X      |

图 5.14 冰柱图

图 5.14 是系统聚类分析结果的冰柱图，从图中我们可以看出，每一步的凝聚状态及结点的凝聚顺序，即哪些结点先聚为一类，之后又与哪些结点聚为另一类。例如，节点从 7 类变为 6 类时，脑梗塞与冠心病之间的距离最小，率先聚成新的 一类。

Dendrogram using Average Linkage (Between Groups)

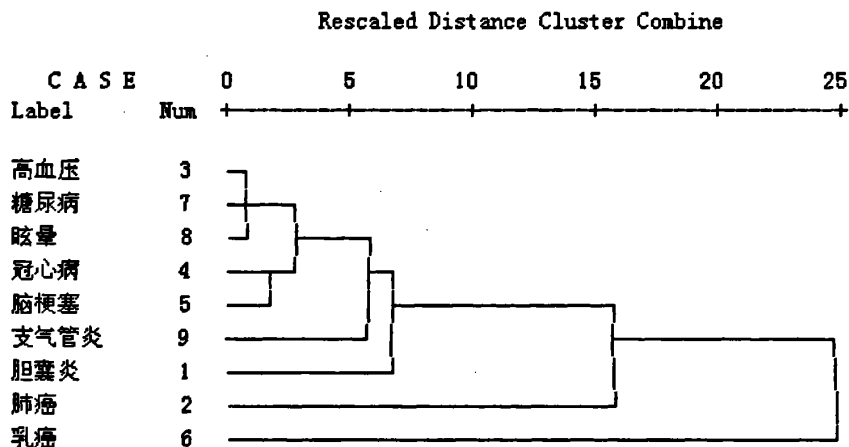


图 5.15 树形图

图 5.15 是反映聚类分析全过程的树形图，垂线表示连接的聚类，横线的长短表示被连接的两类之间的距离大小。可以在此图上用一把尺子垂直方向放在图上左右移动，与尺子相交的每一根横线就是一类。每根横线左端与之联系的各观测量就是分到该类的成员。

#### (4) 聚类结果的分析

从树形图（图 5.15）中我们可以看出将这 9 种疾病分为 3 类时，类间距离比较大，说明各类的特点比较突出，从而能更好的对各种疾病患者的特征进行定义。

第一类包括高血压、糖尿病、眩晕、冠心病、脑梗塞、支气管炎、胆囊炎这 7 种疾病，第二类只包括肺癌，第三类也只包括乳癌这种疾病。

在这 9 种疾病分为 3 类以后，我们对这 3 类疾病的各项评价指标进行统计描述，得出表 5.2，并进而可以对患有这 3 类疾病的患者特征进行比较分析。

表 5.2 三类疾病的患者指标值

|     | 年龄    | 性别   | 入院情况 | 住院次数 |
|-----|-------|------|------|------|
| 第一类 | 60.82 | 1.42 | 2.57 | 1.42 |
| 第二类 | 66.83 | 1.31 | 2.87 | 4.00 |
| 第三类 | 48.14 | 1.97 | 2.97 | 3.41 |

通过对表 5.2 中各项指标值的纵向对比，我们发现第一类疾病的患者年龄、性别值处于三类疾病的中间水平，而该类疾病患者的入院情况和住院次数平均值最低。其中，患有该类疾病患者的平均年龄为 60.82 岁，在三类疾病中处于中间水平，但此种疾病仍属于“老年疾病”。性别平均值为 1.42，说明在该类疾病中女性患者所占比重较大，但也处于三种疾病同类指标的中间水平。由于将入院情况为病危、病急、一般定义为数值 1、2、3，因此，该类疾病的患者入院情况平均

值为 2.57, 说明该类疾病的就医患者多为病急和一般情况, 而病危的就医患者较少。该类疾病患者的住院次数在三类疾病中最低, 为 1.42 次, 说明该类疾病的病情通过医院治疗能够基本得到好转, 从而降低了该类疾病屡次就医的患者比例。

第二类疾病的患者年龄和住院次数在三类疾病中最高, 性别平均值最低, 患者的入院情况平均值处于中间水平。其中, 该类疾病的患者年龄平均值为 66.83 岁, 在三类疾病中最高, 属于“老年疾病”范畴。该类疾病患者的住院次数为 4 次, 说明该类疾病通过在医院的治疗, 并不能使病情得到较大的好转, 即该病情的治疗难度较大, 致使反复就医的患者比例增加。该类疾病的患者性别平均值为 1.31, 说明在该类疾病的患者中, 男性患者所占比例较大。该类疾病患者的入院情况平均值为 2.87, 处于三类疾病中的中间水平, 说明该类疾病的患者在住院时多为病情一般情况。

第三类疾病的特点在于该类疾病的患者年龄在三类疾病中最小, 且多为女性, 入院情况平均值接近于 3 (即病情一般), 患者住院次数较多, 在三类疾病中属于中间水平。由于第三类疾病由乳癌代表, 而乳癌是一种妇科疾病, 所以, 患者绝大多数都是女性。此外, 该类疾病的患者平均年龄在三类疾病中最低, 只有 48.14 岁, 因此可以认为该疾病属于一种“中年疾病”。患者入院情况的平均值为 2.97, 说明该类疾病的患者在入院时多为病情一般的情况。该类疾病患者的住院次数平均值为 3.41 次, 说明该类疾病的治疗难度较大。

综上所述, 本小节以患者的年龄、性别、入院情况、住院次数作为对各病种聚类分析的评价指标, 根据系统聚类分析的结果, 我们可以了解到患有各类疾病的患者特征。疾病作为患者与医院联系在一起的纽带, 是患者在与医院发生关系的过程中所表现出的行为模式, 即患者的“角色”, 因此, 对各病种的聚类分析也即是对患者节点的“角色”分析, 有助于我们进一步明晰各“位置”患者节点在与医院发生关系中的“角色”特征。

## 5.4 本章小结

本章运用社会网络分析理论, 从位置取向角度对西安市医疗网络模型中的患者节点进行分析, 主要分析内容包括三个方面: 结构等效分析、位置分析和角色分析。

其中, 在结构等效分析方面, 从宏观角度可以将医疗网络中的各患者看作是具有“结构等效”的节点, 从微观角度, 认定只有在同一医院就医的患者之间才可以称之为“结构等效”; 在位置分析方面, 本文认为具有相同性别或年龄段的患者在医疗网络中具有相同的“位置”。在角色分析中, 本章分别从入院情况和病种两个角度, 运用频数分析、均值比较方法对具有相同“位置”的患者进行了“角

色”分析，并运用系统聚类分析的方法对九种高发疾病进行了聚类，聚类结果为第一类包括高血压、糖尿病、眩晕、冠心病、脑梗塞、支气管炎、胆囊炎，第二类包括肺癌，第三类包括乳癌，由于聚类分析的评价指标为患者的各项特征，因此根据对疾病的聚类分析结果可以进一步挖掘出西安市患者的各项结构性特征。

## 第六章 对西安市医疗网络模型的关系取向分析

根据第二章对社会网络分析理论的介绍,我们知道,在社会网络分析中,根据分析的着眼点不同,社会网络分析可以分为两种基本视角:关系取向和位置取向。其中,关系取向关注行动者之间的社会性粘着关系,通过社会联结本身——如密度、强度、对称性、规模等——来说明特定的行为和过程。按照这种观点,那些强关系的、密集的且相对孤立的社会网络可以促进统一的社会网络特征的形成。从关系取向对社会网络进行分析主要基于六个方面的内容:节点规模、关系强度、网络密度、关系的内容、不对称关系与对称关系、直接性与间接性。本章试图从关系取向角度对西安市医疗网络模型进行研究,从而进一步明晰医院与患者之间的社会关系。

### 6.1 西安市冠心病医疗网络模型的数据

根据第三章的分析结论,西安市医疗系统符合社会网络的特点,可以构建成一个社会网络,并且在 3.3 节中构建了西安市医疗网络模型,给出了该模型用于社会网络分析的数据类型。通过第五章的分析,我们知道各个病种的患者与医院之间的关系模式相差很大,因此,本章在对西安市医疗网络模型从关系取向进行社会网络分析时,只是基于冠心病这一病种对社会网络数据进行构建,这种简化方法虽然具有一定的缺陷,但是由于冠心病的发病率很高,因此,基于冠心病的社会网络数据也能够一定程度上反映西安市医疗网络中患者与医院之间关系的特征。

根据第四章对各等级医院系统聚类分析的结果,我们结合实际情况可以把西安市医疗网络中的医院节点划分成三类,即三级甲等医院、三级乙等医院、二级和一级医院。同时,根据第五章对患者节点的“角色”分析,我们可以根据患者入院情况将患者划分成三类,即病危患者、病急患者、病情一般患者。因此,西安市医疗网络模型中的节点可以被归纳为六个,即病危患者、病急患者、病情一般患者、三级甲等医院、三级乙等医院、二级和一级医院。

此外,在这六个模型节点中,如果我们以患者的住院天数来衡量患者与就医医院之间关系强度的话,那么,针对本文在西安市冠心病患者中提取的样本数据,入院情况为病危的冠心病患者三等甲级医院、三等乙级医院、二级和一级医院的平均住院天数分别为 13.83 天、14.65 天、10.82 天;病急患者在上述三类医院的平均住院天数为 11.78 天、12.27 天、12.33 天;入院情况为一般的患者在上述三类医院的平均住院天数为 12.78 天、10.85 天、11.75 天。结合本文 3.3 节给出的西安市医疗网络的数据类型,我们可以给出西安市医疗网络模型在冠心病下的社会

网络数据为如下形式:

$$A_{6 \times 6} = \begin{matrix} & \begin{matrix} P1 & P2 & P3 & H1 & H2 & H3 \end{matrix} \\ \begin{matrix} P1 \\ P2 \\ P3 \\ H1 \\ H2 \\ H3 \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 13.84 & 14.65 & 10.82 \\ 0 & 0 & 0 & 11.78 & 12.27 & 12.33 \\ 0 & 0 & 0 & 12.78 & 10.85 & 11.75 \\ 13.84 & 11.78 & 12.78 & 0 & 0 & 0 \\ 14.65 & 12.27 & 10.85 & 0 & 0 & 0 \\ 10.82 & 12.33 & 11.75 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

其中, P1、P2、P3 分别代表入院情况为病危、病急、一般的患者; H1、H2、H3 分别代表三级甲等医院、三级乙等医院、二级和一级医院; 矩阵  $A_{6 \times 6}$  中的元素  $a_{ij}$  ( $1 \leq i, j \leq 6$ ) 代表各类入院情况的冠心病患者在各等级医院的平均住院天数, 用以表述患者与医院之间的关系强度。例如,  $a_{55} = 14.65 = \max\{a_{ij}\}$ , 说明若以住院天数衡量节点关系, 那么, 在医疗网络节点间中关系最强的是病危冠心病患者与三级乙等医院之间的关系。

同理, 如果我们以各入院情况患者在各等级医院的住院人数来衡量患者与医院之间关系强度的话, 那么, 针对本文对西安市冠心病患者提取的样本数据, 在入院情况为病危的冠心病患者中, 到三级甲等医院、三级乙等医院、二级和一级医院的住院人数分别为 9 名、25 名、11 名; 病急患者中到上述三类医院的住院人数分别 41 名、83 名、57 名; 而入院情况为一般的患者中到上述三类医院的住院人数分别为 102 名、15 名、42 名。结合本文 3.3 节给出的西安市医疗网络的数据类型, 我们可以给出西安市医疗网络模型在冠心病下的社会网络数据为如下形式:

$$A'_{6 \times 6} = \begin{matrix} & \begin{matrix} P1 & P2 & P3 & H1 & H2 & H3 \end{matrix} \\ \begin{matrix} P1 \\ P2 \\ P3 \\ H1 \\ H2 \\ H3 \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 9 & 25 & 11 \\ 0 & 0 & 0 & 41 & 83 & 57 \\ 0 & 0 & 0 & 102 & 15 & 42 \\ 9 & 41 & 102 & 0 & 0 & 0 \\ 25 & 83 & 15 & 0 & 0 & 0 \\ 11 & 57 & 42 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

其中, 矩阵  $A'_{6 \times 6}$  中的元素  $a'_{ij}$  ( $1 \leq i, j \leq 6$ ) 代表各类入院情况的冠心病患者在各等级医院的住院人数, 用以表述患者与医院之间的关系强度。例如,  $a'_{34} = 102 = \max\{a'_{ij}\}$ , 说明若以住院人数衡量节点关系, 那么, 在医疗网络节点间中关系最强的一般冠心病患者与三级甲等医院之间的关系。

## 6.2 以住院天数衡量关系强度的关系取向分析

### 6.2.1 节点规模

处于特定社会网络中的行动者都与其它行动者有着或多或少、或强或弱的关系，节点规模测量的是行动者与其它行动者之间关系的数量。根据式 (2-1)，西安市医疗网络模型中的各节点都与（且只与）其它三个节点产生关系，即各节点的规模都是 3，其中，对于各患者节点，只与三类医院产生关系，同样，对于各医院节点，只与三类患者节点产生关系。

如果把研究的焦点集中在某一特定行动者（节点）上时，对关系数量的考察就变成了对网络集中性的考察。所谓“集中性”是指特定行动者身上凝聚的关系的数量，用来衡量谁是该网络的主要节点。一般来说，特定行动者凝聚的关系数量越多，它在网络中的地位就越重要。本节通过 UCINET 软件对西安市冠心病下的社会网络数据进行“集中性”的分析，结果如图 6.2 所示，其中，图 6.1 为 UCINET 软件的实现界面。由于本文所用社会网络数据中的元素不是表示关系存在与否的二值变量，而是用住院天数  $a_{ij}$  ( $1 \leq i, j \leq 6$ ) 代表患者与医院之间的关系强度，因此，从图 6.2 中我们可以知道在西安市医疗网络中，病危患者和三级甲等医院身上所凝聚的关系数量最多，“集中性”指标分别为 53.666 和 52.423，也即其在网络中所处的地位更为重要。

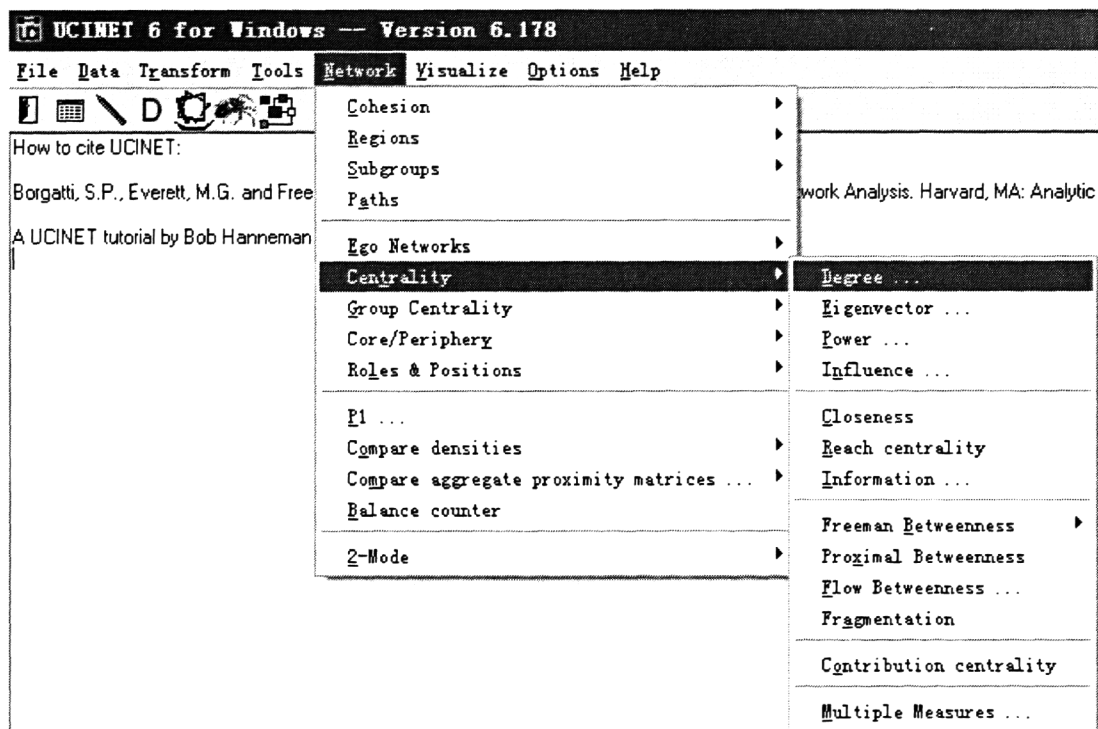


图 6.1 UCINET 软件计算“集中性”的实现界面

|   |       | 1      | 2         | 3     |
|---|-------|--------|-----------|-------|
|   |       | Degree | NrmDegree | Share |
| 1 | 病危    | 39.310 | 53.666    | 0.177 |
| 4 | 三甲    | 38.400 | 52.423    | 0.173 |
| 5 | 三乙    | 37.770 | 51.563    | 0.170 |
| 2 | 病急    | 36.380 | 49.666    | 0.164 |
| 3 | 一般    | 35.380 | 48.300    | 0.159 |
| 6 | 二级和一级 | 34.900 | 47.645    | 0.157 |

图 6.2 以住院天数定义关系强度的网络集中性结果图

### 6.2.2 关系强度

在衡量患者与医院这种双向性关系的强度时，本节选用患者的住院天数作为评价指标，即患者住院天数长，意味着该患者与医院之间的关系紧密，反之，认为关系薄弱。在 6.1 节对冠心病下患者与医院之间建立的社会网络数据中， $a_{ij}$  正是每类患者在各等级医院的平均住院天数。其中，病危患者在三类医院就医的平均住院天数定义为  $a_1$ ，用于反映这类节点的在网络中的关系强度，

$$a_1 = \sum_{j=4}^6 a_{1j} / 3 = 13.10, \text{ 同理, 病急患者的关系强度 } a_2 = 12.13, \text{ 入院情况为“一般”}$$

患者的关系强度  $a_3 = 11.79$ ，因此在患者节点中，其关系强度的比较为：  $a_1 > a_2 >$

$$a_3; \text{ 三级甲等医院的关系强度 } a_4 = \sum_{i=1}^3 a_{i4} / 3 = 12.80, a_5 = 12.59, a_6 = 11.63, \text{ 且有 } a_4$$

$> a_5 > a_6$ 。在整个冠心病医疗网络中，有  $a_1 > a_4 > a_5 > a_2 > a_6 > a_3$ 。说明病危患者和三级甲等医院在网络中的关系强度最大，即与其它节点的关系更为紧密，这也从一个侧面印证了“病危患者和三级甲等医院在网络中处于重要地位”的结论。

### 6.2.3 网络密度

网络中一组行动者之间关系的实际数量和其最大可能数量之间的比率称为“密度”。当实际的关系数量越接近于网络中的所有可能关系的总量，网络的整体密度就越大，反之则越小。运用 UCINET 软件计算本章所研究的冠心病医疗网络密度<sup>[44]</sup>，得出其密度值为 0.60，结果如图 6.4 所示，图 6.3 为 UCINET 软件计算网络密度的实现界面。根据第二章的理论介绍，关系紧密的团体合作行为较多，信息流通较容易，而关系十分疏远的团体信息较少，效率较低。目前，西安市冠



心病医疗网络的密度较低，为提高西安市医疗网络的密度、加强网络中的联系应通过媒体和政府来促进患者与患者之间、医院与医院之间、患者与医院之间的合作联系，以促进网络内的信息流通。

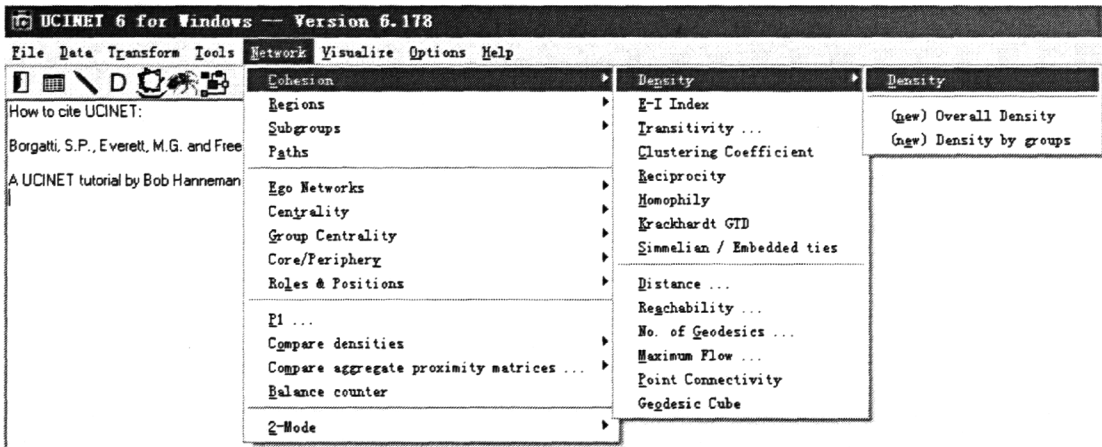


图 6.3 UCINET 软件计算“密度”的实现界面

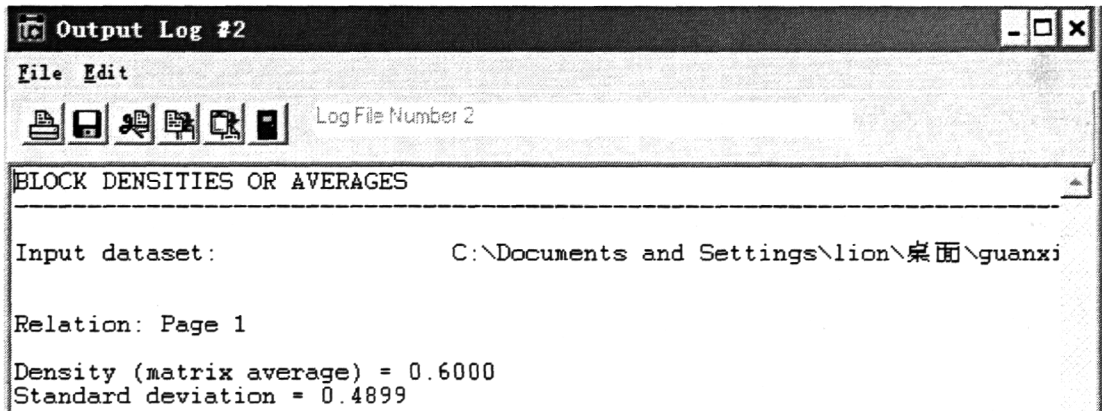


图 6.4 以住院天数定义关系强度的网络密度结果图

### 6.2.4 关系的内容

关系的内容就是指两行为者间关系产生的原因与关系的建构基础，即使在相同的网络中，行动者之间的关系也会具有不同的内容。任何可能将行动者联系起来的東西都能使行动者之间产生关系，因此关系内容的表现形式也是多种多样的。对于在医疗网络中的患者和医院节点，本文认定它们之间只存在“患者就医——医院诊疗”的关系模式，而对患者与患者之间、医院与医院之间的信息交流和合作关系不做考虑。根据患者就医方式的不同可以将患者与医院“患者就医——医院诊疗”的关系内容分为门诊关系、住院关系、健康体检关系等几种形式，本章西安市冠心病医疗网络数据正是基于患者与医院之间的住院关系所进行构建。

### 6.2.5 不对称关系与对称关系

在不对称关系中，相关行动者的关系在节点规模、关系强度、网络密度和关

系内容方面是不同的,而在对称关系中,行动者的关系在这些方面的表现都是相同的。在医疗网络模型中,对于在同一医院就医、入院情况相同的冠心病患者,他们之间在节点规模、关系强度、网络密度、关系内容方面的相似性导致其之间存在对称关系,同理,相同等级的医院之间也是对称关系。此外,在患者诊疗过程中,患者与医院之间存在信息不对称,主要体现反映在对医疗费用的认知不同,这从而导致在医疗网络中患者与医院节点的不对称关系。

#### 6.2.6 直接性与间接性

网络节点的直接性是指行动者之间直接发生的关系,而间接性则指必须通过第三者才能发生的关系。在医疗网络中,这种直接性主要体现在患者对三级甲等医院的肯定度较大,在患病时只愿去三级甲等医院就医,即与其发生直接性的“强联系”。随着居民生活水平的提高,使越来越多的患者都对三级甲等医院具有很高的认可度,但是三级甲等医院的医疗资源属于稀缺性资源,不能完全满足广大患者的就医需求,因此导致了目前“看病难”现象的发生。为解决医疗网络中“看病难”的问题,我们可以引用社会网络分析中的“间接性”概念。例如,三级甲等医院可以将部分病情较轻和术后只需简单护理的患者分流到其指定的低级别医院,在这样的医疗合作模式中,患者与低级别医院之间属于间接性关系,而三级甲等医院处于患者与低级别医院之间的“介绍人”的位置。三级甲等医院通过这种将患者“外包”给低级别医院的方式,一方面有助于自身高效使用医疗资源、缓解接诊压力,另一方面,也有助于改善低级别医院浪费医疗资源——“病床闲置”的局面,同时,也能保证病情严重的患者能够得到及时的治疗,进而提高了整个医疗网络的运转效率。

#### 6.2.7 绘制医疗网络节点图

根据社会网络分析的理论,针对本章构建的西安市冠心病医疗网络数据运用 UCINET 软件绘制出其网络节点图,结果如图 6.5 所示。通过图 6.5 的网络节点图,我们可以直观的看出各患者节点与各医院节点的联系与否与关系强度。其中,节点间的距离越远(即线的距离越长),意味着节点间的关系强度越弱,反之,关系强度越大,从而反映出入院情况不同的各冠心病患者与各等级医院的关系强度。例如,病危患者与三等甲级医院之间的距离较近,而其与三等乙级医院的距离较远,说明病危患者与三等甲级医院的关系强度更大,更愿意与之建立“强联系”。此外,根据图 6.5,各患者之间(各医院之间与之同理)没有连线,说明其之间没有直接联系,属于“结构洞”关系,但是各患者可以通过某一医院建立间接联系,但此种联系属于一种“弱联系”。

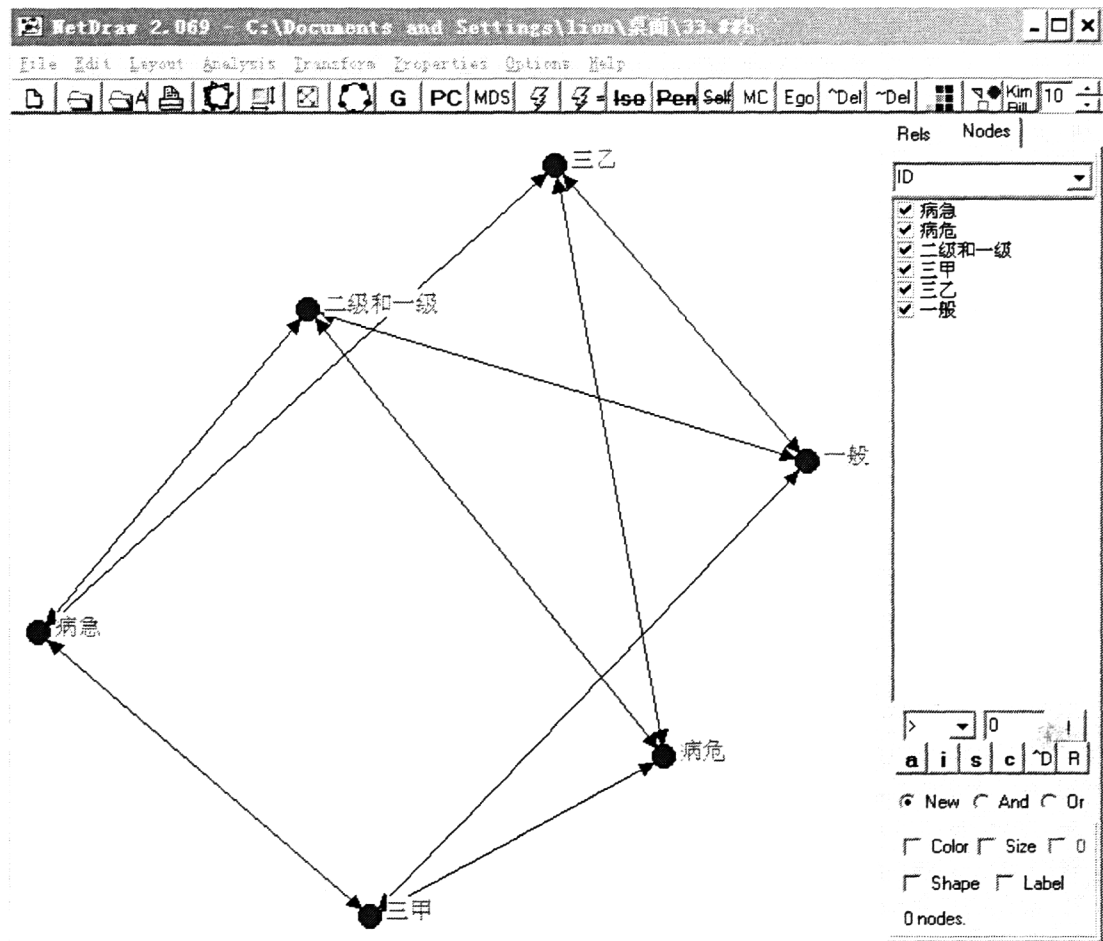


图 6.5 以住院天数定义关系强度的医疗网络节点图

6.3 以住院人数衡量关系强度的关系取向分析

在以住院人数衡量患者与医院关系强度的关系取向分析中，关于节点规模、网络密度、关系内容、对称关系与不对称关系、直接性与间接性方面的分析与 6.2 节相同，本小节不再赘述。

6.3.1 网络集中性

本节通过 UCINET 软件对西安市冠心病下的社会网络数据进行网络“集中性”的分析，结果如图 6.6 所示。从图 6.6 中我们可以知道在以住院人数定义患者与医院关系强度时，西安市医疗网络中病危和入院情况为一般的患者节点身上所凝聚的关系数量最多，“集中性”指标分别为 181.000 和 159.000，也即其在网络中所处的地位更为重要。由于病危患者较少，所以与医疗网络中的各医院节点联系也即很少，所以，较之病危患者和一般患者，其在西安市医疗网络中的地位并不重要，这与 6.2 节从住院天数定义关系强度所做出的集中性结论有所不同。

|   |       | 1       | 2         | 3     |
|---|-------|---------|-----------|-------|
|   |       | Degree  | NrmDegree | Share |
| 2 | 病急    | 181.000 | 35.490    | 0.235 |
| 3 | 一般    | 159.000 | 31.176    | 0.206 |
| 4 | 三甲    | 152.000 | 29.804    | 0.197 |
| 5 | 三乙    | 123.000 | 24.118    | 0.160 |
| 6 | 二级和一级 | 110.000 | 21.569    | 0.143 |
| 1 | 病危    | 45.000  | 8.824     | 0.058 |

图 6.6 以住院人数定义关系强度的网络集中性结果图

### 6.3.2 关系强度

若以住院人数定义患者与医院的关系强度，在 6.1 节对冠心病下患者与医院之间建立的社会网络数据中， $a'_{ij}$  正是每类患者在各等级医院的住院人数。其中，本文所提取的冠心病患者总数为 385 名，若以  $a'_1$  表示病危患者节点的关系强度，则  $a'_1 = \sum_{j=4}^6 a_{1j} / 385 = 11.69\%$ ，同理，病急患者的关系强度  $a'_2 = 47.01\%$ ，入院情况为

“一般”患者的关系强度  $a'_3 = 41.30\%$ ，因此在患者节点中，其关系强度的比较为：

$a'_2 > a'_3 > a'_1$ ；三级甲等医院的关系强度  $a'_4 = \sum_{i=1}^3 a_{i4} / 385 = 39.48\%$ ， $a'_5 = 31.95\%$ ，

$a'_6 = 28.57\%$ ，且有  $a'_4 > a'_5 > a'_6$ 。在整个冠心病医疗网络中，有  $a'_2 > a'_3 > a'_4 > a'_5$

$> a'_6 > a'_1$ 。说明病急、一般患者和三级甲等医院在网络中的关系强度较大，即与其它节点的关系更为紧密，这也从一个侧面印证了上文关于“病急、一般患者和三级甲等医院在网络中处于重要地位”的结论。

### 6.3.3 绘制医疗网络节点图

根据社会网络分析的理论从住院人数角度定义患者与医院的关系强度，针对 6.1 节所构建的西安市冠心病医疗网络数据，运用 UCINET 软件绘制出其网络节点图，结果如图 6.7 所示。通过图 6.7 的网络节点图，我们可以直观的看出各患者节点与各医院节点的联系与否与关系强度。其中，病情一般的患者与三等甲级医院之间的距离较近，而与三等乙级医院的距离较远，说明病情一般的患者与三等甲级医院的关系强度更大，而病危患者与三级甲等医院的距离最远，关系强度最小，这说明在目前西安市冠心病医疗网络中，三级甲等医院接诊更多的是病情一

般的患者,同时,病情危急的患者更多的只能在三级乙等和二级、一级医院就医,这种资源分配的不合理直接导致了病危患者不能得到最有效的医治以及三级甲等医院的资源浪费。

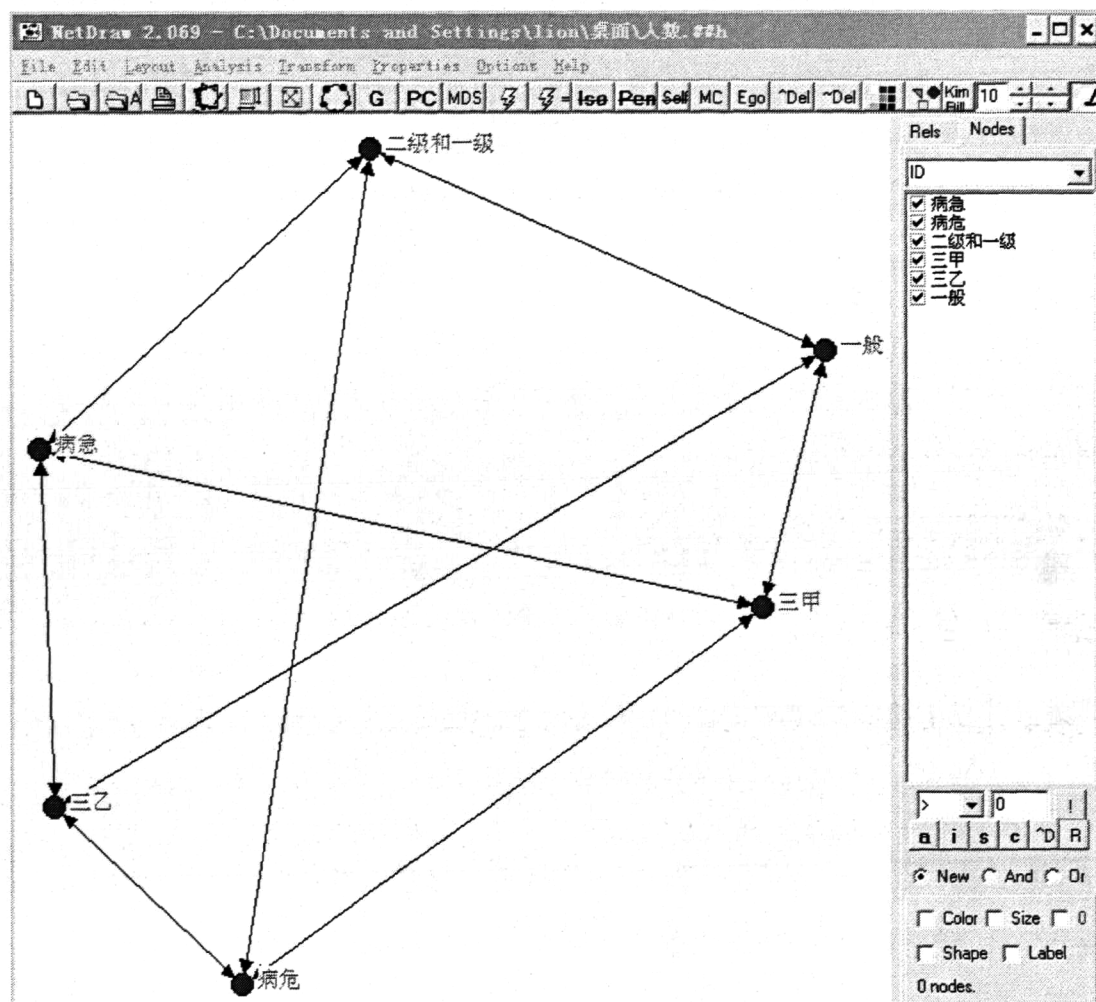


图 6.7 以住院人数定义关系强度的医疗网络节点图

## 6.4 西安市医疗网络的特征及优化方案

### 6.4.1 西安市医疗网络的特征

经过第四章和第五章对西安市医疗网络中医院和患者节点的位置取向分析以及 6.2 和 6.3 节针对西安市冠心病医疗网络的关系取向分析,我们现在可以初步把握西安市医疗网络的特征和其中存在的问题<sup>[45]</sup>,现总结如下:

(1)特征一:各等级医院在医疗合作方面存在互补性。

根据 4.3.1 节从治疗效果对各等级医院进行的聚类分析结果,并结合具体的实际病例,我们可以发现,二级和一级医院在治疗病情较轻的患者时,通过药物治疗具有使其病情治愈的能力。这使得三级医院可以将病情较轻,可以靠药物治疗

病情的患者分流到指定的二级和一级医院中,从而为三级医院与二级、一级医院的互补性医疗合作奠定了可行性。

(2)特征二:各等级医院的费用水平差距较大<sup>[47]</sup>。

本文在 4.3.2 节运用系统聚类分析对各等级医院进行研究时,通过提取冠心病患者的数据作为样本,从医疗费用水平的角度对各等级的医院进行聚类。虽然,得出的结果具有一定的片面性(因为仅仅是以冠心病的患者数据作为样本),但是,我们还是能够得到一些积极的结论。其中,通过表 4.6 对各类医院的费用水平进行纵向对比分析,我们可以知道,三级甲等医院与三级乙等、二级、一级医院在费用水平上存在着较大差异。在人均个人统筹支付费用方面,三级甲等医院比三级乙等医院多出 3655 元,比二级和一级医院多出 6180.4 元。而三级甲等医院中手术患者比例高,进而产生高额的手术费用,是造成这一巨大费用水平差距的主要原因。

(3)特征三:胆囊炎、肺癌、高血压、冠心病、脑梗塞、乳癌、糖尿病、眩晕、支气管炎为西安市居民的九种高发疾病。

由于患有胆囊炎、肺癌、高血压、冠心病、脑梗塞、乳癌、糖尿病、眩晕、支气管炎这九种疾病的患者占到本次实证研究所提取患者样本总数的 61.9%,因此,可以认为上述这九种疾病在西安市患者当中属于高发疾病。此外,在九种高发疾病中,又尤以冠心病的患病率最高,该疾病患者人数在这九种高发疾病患者中占到 26.8%。

(4)特征四:在高发疾病中(乳癌除外),男性患者居多。

通过 5.3.1 节对患者性别的一维频数分析,发现在西安市居民九种高发疾病中男性患者居多,占到 55%。此外,通过 5.3.2 节从入院情况对患者性别和年龄段的二维频数分析中,也发现在病危患者中男性也居多,占到 61.54%。

(5)特征五:可以将九种高发疾病从患者特征角度聚为三类。

通过 5.3.4 节运用系统聚类分析方法,根据西安市患者特征——年龄、性别、入院情况、住院次数对上述九种高发疾病进行聚类,可以将其分为三类,第一类:高血压、糖尿病、眩晕、冠心病、脑梗塞、支气管炎、胆囊炎;第二类:肺癌;第三类:乳癌。

(6)特征六:医疗网络资源配置不合理<sup>[46]</sup>。

从 6.2 节对西安市医疗网络的关系取向分析中,我们知道病危患者和三级甲等医院具有更重要的地位,但是通过 6.3 节从以住院人数衡量节点关系强度的关系取向分析中,我们知道由于三级甲等医院接诊的患者更多的是一般病情患者,这使得病危患者并不能及时在三级甲等医院得到医治,这说明西安市医疗网络的资源存在配置不合理性。毋庸置疑,三级医院在诊疗技术和高科技医疗设备上具有较二级、一级医院更强的能力,这也正是大多患者看重到三级医院就医的原因,

但在过多病情一般的患者涌向三级医院就医的同时,不仅使得二级和一级医院的病床大量闲置,浪费了医疗卫生资源,也使得三级医院的床位却非常紧张,病危患者不能得到及时的医治。

#### 6.4.2 优化方案

根据 6.4.1 节对西安市医疗网络特征的总结,我们可以基本了解该网络的结构属性。下面,本节根据西安市医疗网络的属性和其中存在的问题,提出以下几条建议用于对西安市医疗系统网络进行优化。

(1)建议一:对各等级医院的医疗卫生资源进行整合<sup>[52]</sup>。

根据第三章对各等级医院的聚类分析结果,并结合具体的实际病例,我们可以发现,二级和一级医院在治疗病情较轻的患者时,通过药物治疗具有使其病情治愈的能力(治愈率达到 17.2%,好转率达到 79.5%)。这使得三级医院可以将病情较轻、可以靠药物治疗病情的患者分流到指定的二级和一级医院中,进行后续的简单治疗和护理。

各三级医院可以选择某些二级或一级医院作为自己的指定医院,与之建立医疗合作关系,将病情较轻或是只需要简单护理(但仍需住院)的患者“分流”到该指定医院中去,患者可以定期到三级医院进行复查(或是将病情化验报告送于三级医院的主管大夫)。这样的医疗合作模式其优点在于:1,使三级医院减轻接诊压力,从而使其有更多精力去医治病情较重的患者,同时也节约出更多的医疗卫生资源;2,使二级和一级医院产生更多的业务量,提高了医疗卫生资源的利用率,同时,通过与三级医院的医疗合作和联系,可以弥补其在医疗卫生资源上的贫乏劣势,并且提高了自身的医疗技术和管理经验;3,通过第三章对各等级医院医疗费用的对比分析,我们了解到,三级医院的各单项医疗费用都明显高于二级和一级医院,所以,采用这样的医疗合作模式,既保证了患者在二级和一级医院的治疗和护理需求得到满足,同时也降低了患者承担的医疗费用。

(2)建议二:对患者就医观念进行正确引导,培育良好的就医观念。

目前,广大患者在就医时都感觉到“看病难”,其实,这主要体现在到三级医院就医的“看病难”,具体形式表现在“挂号难”、床位紧张、手术预约时间长等几个方面。这种现象的发生一方面是由于目前的医疗卫生资源分配不合理,但其中也存在着患者盲目就医心理的因素。很多病情较轻的患者在发病后都只选择到三级医院进行就医,而不相信二级和一级医院的诊疗效果,从而造成目前“就医拥堵”的现象。因此,有必要通过政府、医疗机构、媒体对患者的就医观念进行正确引导,培育良好的就医观念。在进一步加强二级和一级医院医疗卫生资源状况提升的基础上,培育患者对二级和一级医院治疗效果的信心,并且,通过建议

一中提到的三级医院和二级、一级医院的医疗合作模式,使患者可以在二级和一级医院进行联合就医。目前,北京的许多三级医院已经采用了这种模式,纷纷与各社区医院进行医疗合作、建立联系,使得患者可以在社区医院完成在三级医院的就诊预约,同时,也满足了患者简单病情的治疗需求。可以说,对患者就医观念的正确引导可以促进病情较轻的患者“主动分流”到二级和一级医院,从而整合了医疗卫生资源,避免了资源浪费。

(3)建议三:鼓励各种性质医疗卫生机构发展<sup>[48]</sup>。

目前,许多营利性的私营医院如雨后春笋般出现在陕西省的医疗卫生市场中,并得到了蓬勃的发展。根据中国卫生部颁布的《2007 中国卫生年鉴》数据:2006 年陕西省共有营利性医院 198 家,其中,综合医院 122 家,拥有床位 4487 张;中医医院 25 家,拥有床位 463 张;专科医院 47 家,1667 张。这些不同性质医院的设立改变原有人人们对医院“事业单位”的角色认定,为医疗卫生市场注入了新的活力,为居民的就医提供了更多的选择。

营利性医院的迅速崛起首先丰富了医疗卫生资源,缓解了各大医院的接诊压力,对医疗卫生资源水平的提升有着积极的作用,但同时,这类医院也以其先进的医院管理模式、优越的医疗服务、灵活的资源配置为优势,向原有的各公立医院发起了竞争的号角。

此时,政府应该在鼓励各种性质医疗卫生机构发展的同时,又要努力加强监管,避免公立医院与营利性医院无序竞争进而造成医疗卫生资源浪费局面的出现。

(4)建议四:在不与现有医院医疗卫生资源冲突的情况下,大力发展社区医院<sup>[50]</sup>。

社区医院在目前医疗网络中的作用日益得到体现,首先,患者的病情可以就近通过社区医院得到基本的治疗,从而丰富了医疗卫生资源;其次,社区医院为居民提供的定期体检也方便患者及时发现病情;此外,社区医院提供的日常保健服务也有利于提高居民的身体状况。因此,应该着力发展社区医院,改善社区医院的医疗卫生条件,提高城乡社区医院的覆盖率,并且按照小区居民的特点丰富其医疗卫生服务功能<sup>[51]</sup>。但与此同时,我们应该首先对各社区的医疗卫生服务供给进行调研分析,例如,某社区已有一家二级医院,并且能够满足该社区居民的日常医疗需求,那么,为避免医疗卫生资源的浪费,不应再设立社区医院,而是应该在该二级医院中增设社区医院的特有服务功能,将社区医院的角色融入进该二级医院中。

(5)建议五:加强对西安市居民高发疾病的预防。

针对上述九种高发疾病,加强对西安市居民的预防工作,并结合各年龄段的高发疾病和其患者的性别、住院次数等特征做到有的放矢的针对性预防,从而提高西安市居民的整体健康水平。



## 6.4 本章小结

本章从患者住院天数和住院人数两个角度对医疗网络中患者与医院的关系强度进行衡量, 分别对西安市冠心病医疗网络进行关系取向分析, 分析内容包括该网络的节点规模、关系强度、网络密度、关系内容、不对称关系与对称关系、直接性与间接性, 并得出相关结论。此外, 结合对西安市医疗网络的位置取向和关系取向分析结论, 总结出该网络的特征和其中存在的问题, 并相应提出对西安市医疗网络的优化方案, 用于改善西安市医疗网络的资源配置, 提高网络效率。

## 第七章 结论

### 7.1 成果总结

本文从系统、科学的角度出发,将西安市的医疗系统视为一个社会网络,构建出该医疗系统的社会网络模型,并分别对该网络中的医院节点和患者节点进行了位置取向分析,对医院与患者之间的关系进行了关系取向分析,据此总结出西安市医疗网络的特征和其中存在的问题,提出了一个模型优化方案,从而将定量分析和定性分析有机的结合在一起。其中,在对医院和患者进行位置取向分析中,主要涉及结构等效分析、位置分析、角色分析三个方面;在对医院与患者之间的关系进行关系取向分析时,主要涉及节点规模、关系强度、网络密度、关系内容、不对称关系与对称关系、直接性与间接性六个方面,并应用 UCINET 软件绘制出该网络的节点图。同时,本文运用聚类分析理论分别对该网络中的医院和患者进行“角色”研究,通过 SPSS15.0 软件的实现,分别对 11 家医院和 9 种高发疾病进行了聚类。此外,在对患者节点进行位置取向分析时,运用频数分析和均值比较方法量化的对患者的“角色”进行了探讨。

本文对西安市医疗网络所做的社会网络分析,通过定量和定性分析相结合,可以科学、清晰的反映出西安市医疗网络的现状,其分析方法和分析过程严谨科学。本文 6.4 节提供的模型优化方案在改进西安市医疗网络现状方面进行了有益的探索,其思路清晰、方法可行,对于政府和社会医疗保险管理机构有一定的参考价值。具体来说,有如下几个优点:

(1)本文从宏观和网络的角度进行分析,创新的将社会网络分析理论运用到研究医疗系统方面,对政府和社会医疗保险管理机构了解西安市医疗网络现状和结构特征有一定的帮助,实现了对医疗保险体制研究的新探索。

(2)从位置取向和关系取向两个角度对西安市医疗网络进行分析,使我们了解西安市医疗系统现状的角度更加多元化,分析更加全面。

(3)将聚类分析、频数分析、均值比较的方法引入到对医院和患者的“角色”分析中,其中,对西安市 11 家定点医院和 9 种高发疾病进行了聚类,从性别和年龄段角度对各入院情况和各病种下的患者数进行了频数分析和均值比较,并结合对西安市医疗网络的关系取向分析,总结出该网络的特征和其中存在的问题,并通过模型优化方案将定量分析和定性分析有机的结合在一起。

## 7.2 问题与展望

### 7.2.1 存在的不足

(1)对医疗系统和医疗服务市场的研究还不够深入,对相关理论、方法和模型的学习还不够全面,有待于进一步的完善。

(2)本文在社会网络分析中所涉及的样本数据有限,并未涵盖西安市所有的医院和病种数据,在一定程度上限制了对西安市医疗网络模型的实证研究结论。此外,在本文医疗网络模型中,如果能够得到各等级医院之间进行交流合作的具体数据,那么运用社会网络分析理论将能得到更进一步的分析。

(3)在本文运用聚类分析对医院和病种进行研究时,由于时间、精力所限未能对聚类分析中所用到的评价指标体系和最优类数做进一步具体的论证。

(4)在本文所构建的西安市医疗网络模型中,只是对医院与患者节点之间的关系进行了分析,尚缺乏对网络模型中其它节点的研究,譬如缺乏对政府、医疗保险管理机构与医院之间的关系强度研究。

### 7.2.2 对西安市医疗系统的展望

随着经济发展水平的迅速提升,西安市医疗系统也必将随之不断发展,并且在国家加大医疗卫生资源投入和建立社区医院的政策指引下,西安市医疗系统的发展注定迎来新的美好篇章。在医疗卫生资源水平提升和进行医疗资源整合的背景下,所谓的“看病难”,“看病贵”现象也会逐渐得以改善,广大居民的医疗服务需求也将得到满足。

此外,随着数据挖掘、社会网络分析等新的、更为科学的量化分析方法引入到社会医疗保险体制研究中来,我国的社会医疗保险体制研究必将有新的突破与创新,从而带动我国社会医疗保险事业的不断发展!

## 致谢

随着毕业论文的完成，我的研究生生活也即将结束。时光荏苒，回首连续六年多的大学校园生活，我对我的导师、父母和同学们怀有深深的感激之情。

首先向我的导师温小霓教授表示深深的谢意。我从本科就是温老师的学生，在六年多的大学学习和生活中，温老师在治学和人生观方面教会了我很多东西。她治学严谨，对我的小论文进行了认真的阅读，仔细考证文中的每个词、每句话，并提出了合理的修改意见。在毕业论文的撰写过程中，她总是询问论文的进度和遇到的困难，并时常提醒我们保证论文的质量。同时，在生活中温老师非常的平易近人，十分关心我们的生活和思想动态，如慈母般的呵护着我不断成长。温老师认真的工作态度、严谨的治学作风，温文尔雅的人格魅力，深深的感动了我。在此论文完成之际，特向温老师表示衷心的感谢！

其次，我要感谢我的父母。在这么多年的学习生活中，父母为我倾注了毕生的心血，不辞劳苦的为我筹集学杂费，无微不至的关心着我在学校的生活。在此，即将毕业的我要对我的父母表示万分的感谢，我将用我以后兢兢业业地工作表现来报答你们对我的养育之恩。

再次，感谢我的同窗同学：管兴亮、蔡汝骏、郭丽红、宇文佳子、曾嵘欣、曹海涛、李泉、刘栋等等，大家在一起生活、学习这么多年，相互之间形成了融洽的关系，为我们的生活共创了轻松愉快的环境。即将各奔东西的朋友们，我祝福大家在以后的工作中，百尺竿头，更进一步！

这里还要感谢我的女朋友席燕，感谢她在我的论文完成过程中给予我的帮助和一直以来对我的支持！

最后，向审阅本文的老师们表示感谢。谢谢你们在百忙之中审阅我的论文！

## 参考文献

- [1] 长思锋,温海红,赵文龙主编.《社会保障概论》[M]. 武汉:科学出版社. 2003. 8. 31-42
- [2] 彭涛,方少红. 数据仓库技术在医保系统中的应用[D]. 华中科技大学. 硕士学位论文. 2006
- [3] 高先兵. 浅谈我国城镇职工基本医疗保险制度的运行与完善[J]. 皖西学院学报. 2006 年第 12 期. 61-64
- [4] 易红,胡祖斌,彭想等. 城市医疗资源调整及其作用的理论研究[J]. 中国社会医学杂志. 2006 年第 12 期. 209-211
- [5] 侯本贞,徐凌中. 山东省医疗资源利用效率评价研究[D]. 山东大学. 硕士学位论文. 2004
- [6] 申一帆,胡善联. 广州市医疗资源配置研究[D]. 复旦大学. 硕士学位论文. 2004
- [7] 李严,邹宽城. 关于医疗资源共享的研究[D]. 长春工业大学. 硕士学位论文. 2004
- [8] 杨起,卢祖洵. 湖北省卫生资源现状分析[D]. 华中科技大学同济医学院. 硕士学位论文. 2002
- [9] Elliot B. Sloane, JEDI—An Executive Dashboard and Decision Support System for Lean Global Military Medical Resource and Logistics Management [J], Engineering in Medicine and Biology Society, Aug.2006. 5440-5443
- [10] Dave Sims, WWW Extends “Apprentice’s Assistant” to Global Medical Resource [J], Volume 16, 1996. 14-15
- [11] 苑雅文. 环黄渤海合作区域产业集群的社会网络分析[J]. 天津社会科学. 2003 年第 2 期. 84-87
- [12] 梁昌勇,陈发祥. 基于社会网络下的组织知识管理研究[J]. 情报杂志. 2007 年第 7 期. 24-27
- [13] 西安市医疗资源卫生现状的统计数据  
<http://www.xawsj.gov.cn/Class/XXGS/index.html>
- [14] 刘军.《社会网络分析导论》[M]. 社会科学文献出版社. 2004. 12. 1-27
- [15] Kim Klyver, Kevin Hindle, Denny Meyer, Influence of social network structure on entrepreneurship participation[J], The International Entrepreneurship and Management Journal, 2005
- [16] 张存刚,李明,陆德梅. 社会网络分析——一种重要的社会学研究方法[J]. 甘

肃社会科学. 2004 年第 2 期. 109-111

[17] 周涛, 张际平. Wiki 社群的社会网络分析[D]. 华东师范大学. 硕士学位论文. 2005

[18] 侯贇慧, 刘洪. 基于社会网络的城市群结构定量化分析[J]. 复杂系统与复杂性科学. 2006 年第 6 期. 35-42

[19] 王永固, 陈品德. 社会网络分析在虚拟学习社区中的应用[R]. 第十屆全球华人计算机教育应用会议论文集. 2006. 189-195

[20] 王霄宁, 王轶. 新经济社会学视角下基于社会网络分析的产业集群定量化研究[J]. 探索. 2005 年第 3 期. 95-99

[21] 王伟, 靖继鹏. 公共危机信息传播的社会网络机制研究[J]. 情报科学. 2007 年第 7 期. 979-982

[22] 约翰·斯科特著, 刘军译. 社会网络分析法[M]. 重庆大学出版社. 2007. 1. 53-68

[23] 陈成文, 肖卫宏. 农民养老: 一个社会网络的分析框架[J]. 湖北社会科学. 2007 年第 4 期. 57-62

[24] 叶新东, 张际平. 网络环境下的社会网络分析[D]. 华东师范大学. 2004

[25] Chung-kwan Shin, A hybrid approach of neural network and memory-based learning to data mining[J], Neural Networks, Volume 11, 637-646

[26] Ichiro Kawachi, Lisa F. Berkman, Social Ties and Mental Health[J], The New York Academy of Medicine, Volume 78, 458-467

[27] UCINET 软件的使用说明

<http://faculty.ucr.edu/~hanneman/nettext/>

[28] 罗家德. 《社会网络分析讲义》[M]. 社会科学文献出版社. 2005.4. 1-27

[29] 王斌会. 数据挖掘技术及其应用现状[J]. 统计与决策. 2006 年第 5 期. 122-124

[30] 范明, 孟小峰等. 《数据挖掘: 概念与技术》[M]. 北京: 机械工业出版社. 2001

[31] 别荣芳, 尹静, 邓六爱译. 《数据挖掘技术: 市场营销、销售与客户关系管理领域应用》[M]. 北京: 机械工业出版社. 2006.07

[32] 朱世武, 崔嵬, 张尧庭. 数据挖掘运用的理论与技术[J]. 统计研究. 2003.08

[33] R.Agrawl, H.Mannila. Fast discovery of association rules. In: MF Usana ed. LAdvances in knowledge Discovery and Data Mining. Manlo Yark, CA: AAAI/ MIT Press, 1996, 307-328

[34] 陈学进. 数据挖掘中聚类分析的研究[J]. 计算机技术与发展. 2006 年第 9 期. 44-49

[35] 杨桦. 聚类分析在投资决策中的应用[J]. 当代经理人. 2006 年第 21 期. 1547-1549

- [36] Everitt B, Landau S, Leese M. Cluster Analysis. Fourth Edition [M]. London : Arnold New York : Oxford University Press, 2001, 21-25
- [37] Chiu T, Fang D, Chen J, et al. A robust and scalable clustering algorithm for mixed type attributes in large database environment. Proceedings of the Seven ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2001, 263-268
- [38] 王志亮, 党延忠. 社会网络分析方法在科研协作网中的应用研究[D]. 大连理工大学. 2005
- [39] 唐国滨, 温小霓. 陕西省医疗资源分布的聚类分析研究[J]. 中国医院管理. 2006 年第 11 期. 18-20
- [40] 林杰斌, 林川雄等. 《SPSS13.0 统计建模与应用实务》[M]. 北京: 中国铁道出版社. 2006
- [41] 惠俊刚, 卫海燕, 郎玉泉. 陕西区域经济发展水平的聚类分析[J]. 兰州商学院学报. 2007 年第 2 期. 55-70
- [42] 余建英, 何旭宏. 《数据统计分析与 SPSS 应用》[M]. 北京: 人民邮电出版, 2003. 4 251-252
- [43] 数据挖掘技术及研究现状, [http: //www. cioworld. net](http://www.cioworld.net)
- [44] 罗家德. 《社会网分析讲义》[M]. 社会科学文献出版社. 2005. 4. 133-146
- [45] 赵林海, 李绍华. 现阶段医疗保险评价的定位[J]. 中国卫生事业管理. 2006 年第 4 期. 214-215
- [46] 熊光前. 中国医疗保障制度模式选择[D]. 天津大学管理学院. 2004
- [47] 杨倩. 医疗费用测算研究[J]. 西安工业学院学报. 2002. 3. 83-86
- [48] 邓宏. 整合社会医疗保险与商业健康保险资源[J]. 医药世界. 2006 年第 6 期. 40-41
- [49] Peter Diamond, Organizing the Health Insurance Market [J], Econometrica, 1992.11. Volume 60. 1233-1254
- [50] 朱一龙, 赵云鹏, 陈虾. 社区医院构建院前急救绿色通道的体会[J]. 岭南急诊医学杂志. 2006 年第 12 期. 451
- [51] 于瑞英, 杜慧敏, 杨晓蓉. 城市 社区居民护理服务需求调查分析及研究[J]. 现代护理. 2006 年第 24 期. 2349-2350
- [52] 魏东海. 医疗资源应该怎样配置[J]. 中国医院管理. 2001 年第 5 期. 57-59

## 攻读硕士学位期间论文发表情况

发表和录用的论著:

[1]唐汉一, 温小霓. 《聚类分析在社会医疗保险研究中的应用》. 财经界. 2007年第10期. 87-88