

优 化 与 决 策

最优结构图

——从复杂关系中发现规律

胡一凡 著

上海科学技术出版社

C934/74

2007

优 化 与 决

最优结构图

——从复杂关系中发现规律

胡一凡 著

上海科学技术出版社

图书在版编目(CIP)数据

最优结构图:从复杂关系中发现规律/胡一凡著. —上海:上海科学技术出版社, 2007. 12

(优化与决策)

ISBN 978—7—5323—9100—4

I. 最... II. 胡... III. 决策学—规律—普及读物
IV. C934—49

中国版本图书馆 CIP 数据核字(2007)第 152348 号

上海世纪出版股份有限公司 出版、发行
上海科学技术出版社

(上海钦州南路 71 号 邮政编码 200235)

新华书店上海发行所经销

苏州望电印刷有限公司印刷

开本 889×1194 1/32 印张 2.625

字数: 54 000

2007 年 12 月第 1 版 2007 年 12 月第 1 次印刷

印数: 1—3 000

定价: 12.00 元

本书如有缺页、错装或坏损等严重质量问题,
请向工厂联系调换

《优化与决策》丛书编委会

名誉主编 吴文俊 谷超豪

主 编 胡毓达

编 委 (以姓氏笔画为序)

王长钰 王兴华 王则柯 方伟武

石钟慈 史树中 刘源张 李 端

汪寿阳 张连生 陈光亚 茆诗松

姚恩瑜 袁亚湘 顾基发 徐利治

唐国春 章祥荪 越民义 韩继业

管梅谷 魏权龄

序



“人尽其能，物尽其用”，是人类进步的重要标志和社会发展的根本动力。

在现代社会中，小至个人事务的处理，大到国家政策的制定，无不需要人们进行关于“人”和“物”的选优抉择，以求取好的结果。在科技日新月异和经济快速发展的 21 世纪，人们要日益面临各种愈来愈复杂的决策问题，因此，现代优化思想和科学决策知识，已是当今人们普遍需要具备的基本素养。

现代教育提倡对学生进行创新精神和综合能力的素质培养。在我国大中学教育中，让学生们了解某些现代优化方法和进行决策能力培养，也正是素质教育的重要内容。

为了向广大读者普及最优化和科学决策的思想和方法，在中国运筹学会及其决策科学分会、数学规划分会和排序分会，中国系统工程学会，中国数学会计算数学分会以及上海运筹学会的倡议和支持下，我们邀请了在相应领域卓有成就的有关专家，撰写了这套《优化与决策》系列丛书。

这套丛书具有以下特点：

选题实用求新 本丛书的重要特色是内容的实用性。各选题在扩大知识的同时，均注重联系实际结合应用展开讨论。不论是定量或定性的决策问题，进行选优建模和效益分析一般要归为用数量刻画和作数值计算，因此，数学是这套丛书各选题的基本工具。但是，与以往作为中学数学教科书内容补充的科普性数学读

物或抽象的数学专著不同,本丛书强调综合运用数学和有关知识去解决现实中的应用问题。另外,丛书的选题既考虑其内容是具代表性的,同时也注意对新领域和某些发展中问题的介绍。

表述浅出深入 本丛书着力于用通俗易懂的方式引导读者掌握现代优化和决策知识。书中特配置了形象的图画以帮助加深读者对内容的理解。我们计划,具有高中数学基础的读者,即可读懂其中的基本内容。但是,为了每一选题的系统性和完整性,也不放弃对一些最基本和著名理论结果的介绍。因此,主要想了解思想方法和借鉴应用手段的读者,阅读时可以略去其中某些理论和证明部分,而不会影响对主要内容的理解。然而,对于具有相当数学素养并对理论结果同样有兴趣者,这些较深入的内容对他们是有价值的,其中有些结果即使对于同行学者也将具有重要的参考意义。因此,不同层次的读者,阅读本丛书后均会有所得益。

这套丛书的读者面是多层次和极广泛的,它既适用于各行业管理者,各级行政公务员,广大科技工作者,以及各专业大学生、研究生和教师们阅读,同时也可作为大专和高中学生的选读材料或课外读物。

写作这套既具科普性又基于一定理论分析的现代应用丛书,对于丛书作者是一种新的尝试。本丛书从酝酿组织、确定选题,直到现在与读者见面,曾经历了较长的时间。许国志院士和俞文鱿

教授生前曾积极参与出版本丛书的策划,并热情承担了写作任务,可惜未及如愿。值得一提的是,各位作者对分担撰写选题的内容都进行了精心选择和安排。特别是,许多作者专业造诣精深,但写作科普著作则是第一次,因此在可读性方面曾倾注了许多心血。对于他们的这种认真和奉献精神,谨致以衷心感谢和崇高的敬意!鉴于著述这套丛书对多数作者是一件新的工作,其中难免或有不足之处,期待读者们不吝指正。

最后,我们对吴文俊院士和谷超豪院士关心和支持本丛书的出版,并乐于担任名誉主编致以诚挚的谢忱!同时,感谢上海科学技术出版社对出版这套丛书所作的一系列努力。

胡競达

2006年9月19日

前言



随着网络的发展和信息量的日益增长,如何从这些信息中发现规律和知识,变得越来越重要了。比如,给出一个社会网络,如何自动找出每个人所属的社会集团;根据一个网上消费者的数据库,如何找出消费行为的特点,进而自动为消费者推荐相关产品;给出某个细胞内多种蛋白质相互作用的关系,如何用图像显示这些关系;以及给出多个计算机程序的关系,如何自动生成流程图,等等。这些问题都可用画出其最优结构图及其有关的数学技术来解答。

2006年5月初,美国一些新闻机构报道,美国国家安全局建立了一个巨大的数据库,这个数据库储存了美国全国每一个人打电话的记录。据称,该数据库只是记载了谁与谁通了电话,而不是具体的通话内容。即便如此,消息传出,舆论纷纷指责这一行为是侵犯了个人隐私。但是,国家安全局宣称建立电话数据库对于情报信息是重要的。

那么,通过这些电话记录,国家安全局究竟能获得些什么情报呢?可以想象,这么大的数据库,光靠肉眼和手工处理,是很难从中找出有用的信息的。必须使用计算机,利用算法和软件,才有可能处理这么大的数据库。

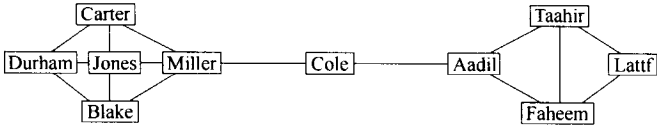
使用绘制最优结构图的软件,将数据库的有关信息显示出来,或者利用网页排行这样的网络分析算法对数据间的关系作分析,就是从庞大复杂的数据库中找出信息和规律的可行方法。

比如,设以下是一些人通电话的记录:

{Faheem→Lattf, Blake→Durham, Jones→Miller, Miller→Blake, Miller→Cole, Cole→Aadil, Lattf→Taahir, Taahir→Aadil, Aadil→Faheem, Jones→Blake, Durham→Jones, Carter→Miller, Faheem→Taahir, Durham→Carter, Carter→Jones }。

光是读这些记录,很难看出什么规律来。如果画出其最优结构图,就可以将以上记录清晰地表示出来。

从下图可以看出,这个电话通讯记录是有规律的,它由两个社会集团组成,左边的集团都有欧美名字,右边都是非欧美的名字。这两个集团通过 Cole 建立联系。



通话网络

使用本书最后部分讨论的网页排行的知识,可以对以上通电话的记录做出分析。下表给出上图中每一个人通电话的排行得分,可以看出,Miller, Aadil, Jones 和 Faheem 得分最高,表明他们在这个社会网络中比较重要。

通电话记录的网页排行分析表

Miller	Aadil	Jones	Faheem	Durham
0.109 8	0.109 8	0.108 3	0.108 3	0.084 1
Lattf	Taahir	Blake	Carter	Cole
0.084 1	0.083 8	0.083 8	0.083 8	0.060 3

本书第1节介绍要讨论的问题,第2节简要回顾了应用结构图表达事物关系的历史,第3和第4节是本书的主要内容,阐述了如何自动生成最优结构图的方法,最后一节则讨论了在构建最优结构图基础上的网页排行问题。

目 录

前言…… I

1. 最优结构图问题……1

2. 结构图应用简史……11

3. 最优结构图的力引导方法……21

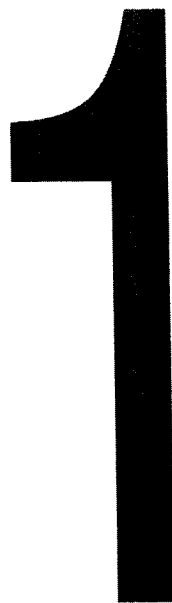
4. 有向最优结构图的分层方法……41

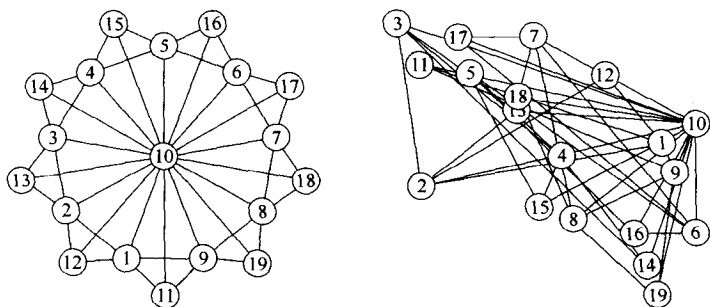
5. 最优结构图和网页排行……51

参考文献……67

致谢……69

最优 结构图问题





最优结构图问题,就是要把复杂的结构关系用“最优”的图像形式表达出来。因此,绘制某问题的最优结构图是从复杂关系中发现规律的一个有效方法。

但是怎样画图才是“最优”的,通常较难准确地用数值来计量。一个图是否画得合理,也取决于所画图的应用领域。一般认为,一个最优的结构图应该没有或有最少相交的边,边的长度应该尽量均匀,顶点在空间中的分布也要尽量均匀。上面两个图具有同一结构关系,显然,左边图的结构安排要比右边的图好,它是这类结构图中的最优结构图。

前言中提到的多个人的通电话记录实际上就是一个图(又称为网络)。从几何角度讲,很多复杂关系都可以用图来直观地表达。

比如说,下面的计算机网络,就是一个很简单的图。

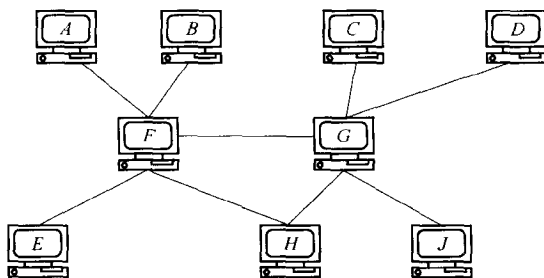


图 1.1 计算机网络

图与最优结构图 究竟什么是图呢?简单地讲,图是由顶点(又称为节点)和边组成的。在图 1.1 中,每一台计算机就是图的一个顶点,而计算机之间的连接是图的边。所以,上图有 9 个顶点,9 条边。图有时也称为网络,本书交替使用这两个词。

用 $A-B$ 表示两个顶点 A 和 B 之间有一条边相连。因此,以上的计算机网络可以用文字符号表述为 $\{A-F, B-F, C-G,$

$G-D, F-G, E-F, F-H, G-H, G-J\}$ 。

在这个例子中,计算机 A 和计算机 F 相连,那么显然 F 和 A 也相连。在这种情况下,每一条边是没有方向性的,这样的图称为无向图。

还可以定义顶点之间的图距离。比如,图 1.1 中, A 和 B 的距离是 2,因为从 A 到 B ,需要经过两个顶点 F 和 B , A 和 J 的距离是 3(即 A 到 F 到 G 到 J)。

顶点 A 和 B 之间的距离,是指从顶点 A 到顶点 B ,至少要经过几个顶点(包括 B)。

并非所有的图都是无向图。举个例子说,有一个十个人的小组。用 $A \rightarrow B$ 表示 A 是 B 的上级。这十个人的上下级关系如下:

$\{李 \rightarrow 安, 李 \rightarrow 张, 李 \rightarrow 江, 安 \rightarrow 马, 张 \rightarrow 马, 张 \rightarrow 唐, 江 \rightarrow 林, 江 \rightarrow 唐, 江 \rightarrow 杨, 安 \rightarrow 许, 江 \rightarrow 金\}$ 。

这又是一个图(见图 1.2)。这个图有 10 个顶点。图的 11 条边就是以上列出的。严格地讲,这是一个所谓的有向图,在这个图中,每条边都有一个方向,比如,“ $李 \rightarrow 安$ ”是一条边,但“ $安 \rightarrow 李$ ”不是一条边,因为“ $李$ ”是“ $安$ ”的上级,而“ $安$ ”不是“ $李$ ”的上级。

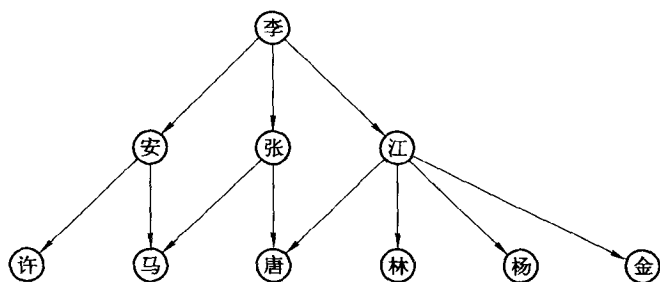


图 1.2 上下级关系组成的有向图

简洁地展示出来。比如,给出以上用文字表达的上下级关系,最优结构图问题就是要想办法把它变成图 1.2。如果所考虑的图只涉及少量的顶点,那么最优结构图问题用手工的办法就能解决。但实际问题中产生的图可以有很多顶点,对于这样的图,用手工的办法是很难画出的。因此,本书讨论怎样用自动的方法来画出问题的最优结构图。

那么怎样画出问题的最优结构图呢?这通常较难准确地用数值来计量。另外,一个图是否画得合理,也取决于所画的图的应用领域。一般认为,一个最优结构图应该是没有或有最少相交的边,边的长度应该尽量均匀,顶点在空间中的分布也要尽量均匀。比如,图 1.3 是同一个结构的图的两种画法,比较这两个图,显然,图 1.3(a)的结构安排要比图 1.3(b)好。

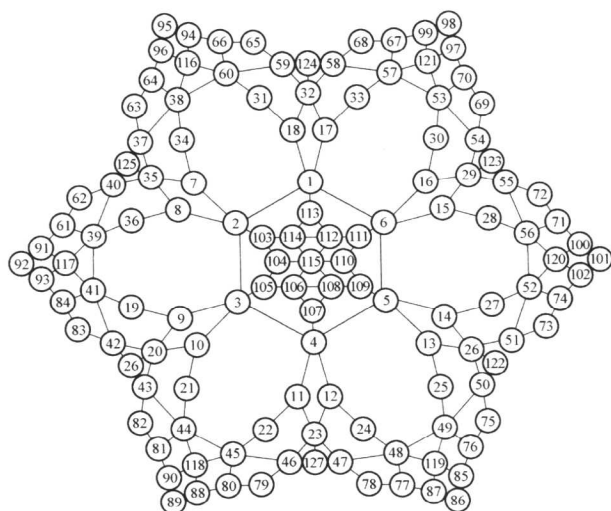
最优结构图问题的应用非常广泛。因为军事、运输、科技、管理、计算机等方面的大量问题都可以用图来表达,由此产生的图包括:组织机构图、计算机程序流程图、大规模集成电路图、电路图、社会网络、运输网络、蛋白质相互作用网络(protein interaction network)、分子结构图,等等。

最优结构图问题例子 比如说,人们通常用图画来展示一个单位的组织机构。图 1.4 用最优结构图表示了某个公司的组织机构。

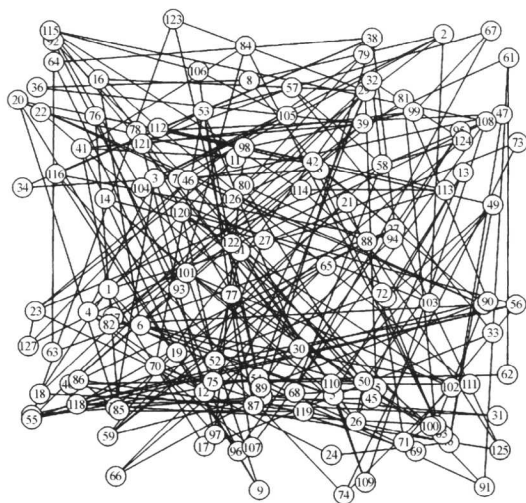
再比如说,考虑以下这个生物链:

{浣熊→鸟,浣熊→虫,猫→鸟,猫→鼠,狐狸→鸟,狐狸→蛇,狐狸→蝶螈,狐狸→兔,狼→兔,狼→狐狸,狼→鼠,狼→鼬,狼→鹿,熊→鹿,熊→鼠,熊→植物,鸟→植物,蛇→虫,蛇→蛙,蝶螈→虫,兔→植物,鼬→鼠,鼬→虫,鹿→植物,蛙→虫,虫→植物}。

用图 1.5 来表示,就要清楚得多了。



(a)



(b)

图 1.3 同一个图的两画法

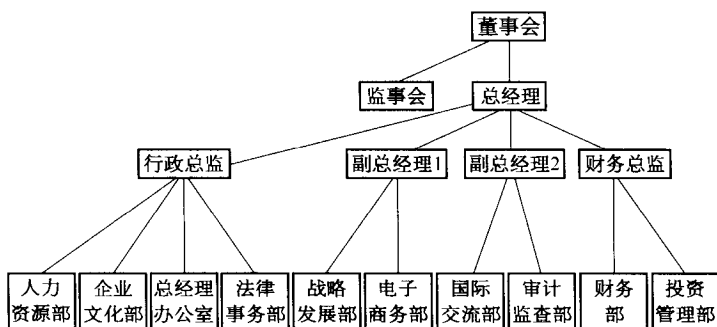


图 1.4 一个组织机构图

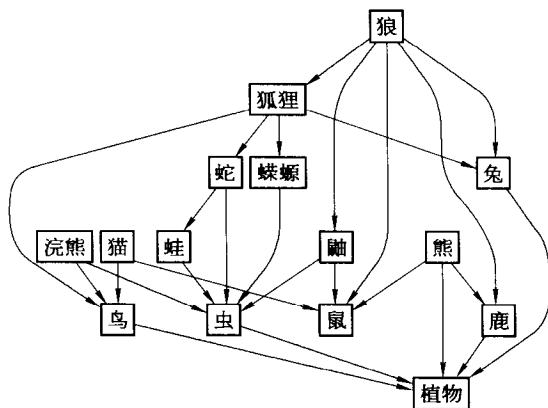


图 1.5 一个生物链示意图

最优结构图问题也是研究社会网络的有力工具。比如有以下这个社会网络：

{ 李→胡, Carol→Andre, Carol→Fernando, 王→胡, Carol→Diane, Andre→Diane, Andre→Fernando, Andre→Beverly, Fernando→Diane, Fernando→Garth, Fernando→Heather, Diane→Beverly, Diane→Garth, Diane→Ed, Beverly→Garth, 李→王,