



计算几何若干方法 及其在 空间数据挖掘中的应用

樊广仝 著



冶金工业出版社
Metallurgical Industry Press

计算几何若干方法及其在 空间数据挖掘中的应用

樊广全 著

北 京
冶 金 工 业 出 版 社
2010

内 容 提 要

计算几何作为计算机科学的一个分支,本书对其新发展和研究工作进行综述性的介绍。论述了 KDTIM 理论的内涵;通过对计算几何中的一些问题的研究,提出一些新的理论与算法;将计算几何的理论方法应用于空间数据挖掘中,用计算几何中的理论和方法解决知识发现中的一些问题。

本书适合从事计算几何、数据挖掘等计算机科学相关领域的工作人员阅读。

图书在版编目(CIP)数据

计算几何若干方法及其在空间数据挖掘中的应用/
樊广伶著. —北京:冶金工业出版社, 2010. 3
ISBN 978-7-5024-5158-5

I. ①计… II. ①樊… III. ①计算几何—计算机算法
②计算几何—应用—地理信息系统 IV. ①TP301.6 ②P208

中国版本图书馆 CIP 数据核字(2010)第 021708 号

出 版 人 曹胜利

地 址 北京北河沿大街嵩祝院北巷 39 号, 邮编 100009

电 话 (010)64027926 电子信箱 postmaster@cnmip.com.cn

责任编辑 杨盈园 美术编辑 李 新 版式设计 葛新霞

责任校对 栾雅谦 责任印制 牛晓波

ISBN 978-7-5024-5158-5

北京百善印刷厂印刷;冶金工业出版社发行;各地新华书店经销

2010 年 3 月第 1 版, 2010 年 3 月第 1 次印刷

850mm × 1168mm 1/32; 6.125 印张; 161 千字; 184 页; 1-1500 册

25.00 元

冶金工业出版社发行部 电话: (010)64044283 传真: (010)64027893

冶金书店 地址: 北京东四西大街 46 号(100711) 电话: (010)65289081

(本书如有印装质量问题, 本社发行部负责退换)

前 言

计算几何作为计算机科学的一个分支学科,自 20 世纪 70 年代诞生以来得到了迅猛的发展。该领域中的问题所带来的挑战性,使得一大批科研人员为之呕心沥血、辛勤耕耘,在比较短的时间内使这一崭新的研究领域取得了辉煌的成果,对许多问题有了一系列比较成熟的计算几何算法。但是,在该领域仍存在着一些问题没有得到解决或没有解决好,一些算法在运行效率等方面仍不能满足一些应用领域的要求,需要人们继续发展、完善它。知识发现也是一门年轻的学科,其他一些学科,如统计学、人工智能、模糊数学、粗糙集、图论与超图理论等多个学科的理论与方法都纷纷地运用到知识发现中来,为这一领域的发展带来了活力。如果能将计算几何中的一些方法运用到知识发现领域,解决知识发现过程中遇到的一些问题,无疑将会对两门学科都产生深远的影响。

本书是以 2003 年国家科技成果重点推广计划项目——集成化组合构件式知识发现软件系统 (ICCKDSS, 项目编号 2003EC000001) 为背景而编写的。ICCKDSS 是基于内在机理的知识发现理论 (Knowledge Discovery Theory based on Inner Mechanism, KDTIM) 在空间数据挖掘领域的进一步扩展。本书主要对计算几何中的一些理论与方法进行研究,并尝试将这些理论与方法应用于空间数据挖掘中,解决空间数据挖掘中的一些问题。

本书首先希望丰富 KDTIM 理论的内涵;其次,通过对计算几何中的一些问题的研究,提出一些新的理论与算法,从而对计算几何的完善贡献一份力量;再次,将计算几何的理论方法应用于空间数据挖掘中,可以用计算几何中的理论和

方法解决知识发现领域中的一些问题。

本书针对平面点集的凸壳，提出并证明了平面点集的城堡定理，设计了城墙快速搜索算法；提出了一种新的平面点集凸壳快速算法；提出了两种近似凸壳算法，即点集坐标旋转法（PSCR）和多方向极值近似凸壳算法（MDEV）；提出了一种任意简单多边形单调剖分算法；提出了对经典艺术画廊问题的两种解决方法，即基于可见传播规则的解决方法和基于顶点可见关系矩阵的解决方法；揭示了平面点集三角剖分的一个性质；在研究目前知识发现特别是空间数据挖掘中计算几何方法应用现状的基础上，提出了一种基于空间邻近关系的可视化空间数据聚类算法。

计算几何和空间数据挖掘都是比较新的领域。如何提出更好的计算几何算法、更多地将其运用于空间数据挖掘中，今后将需要不断进行更深层次的研究。

本书是作者在博士论文研究成果的基础上编撰而成的。在此，谨向我的导师杨炳儒教授致以最崇高的敬意和最真诚的感谢！在编写本书的过程中，得到了周培德教授、石纯一教授、钱旭教授、尹怡欣教授、郑雪峰教授、杨扬教授、穆志纯教授、闵乐泉教授、郑德玲教授、王志良教授、周显伟教授、王成耀教授和王昭顺教授的大力支持和悉心指导，在此向他们表示最诚挚的感谢！感谢所有曾给予我热心帮助和关心的老师和同学们。

由于水平和经验所限，书中如有不妥之处，恳请广大读者批评指正。

作 者

2009年10月

目 录

1 相关领域研究与发展现状	1
1.1 计算几何概述	1
1.1.1 计算几何简介	1
1.1.2 计算几何的研究内容	2
1.1.3 计算几何的发展与现状	3
1.1.4 计算几何与其他学科的关系	4
1.2 知识发现概述	5
1.2.1 KDD 的产生与发展	5
1.2.2 KDD 技术研究和应用存在的问题与发展趋势	6
1.2.3 基于内在认知机理的知识发现理论	8
1.3 空间数据挖掘概述	9
1.3.1 空间数据挖掘的研究现状与发展	9
1.3.2 空间数据结构和空间数据库	33
1.3.3 GIS 数据库的特点	45
1.3.4 空间数据库模型	46
1.3.5 空间查询与空间索引	50
2 平面点集的凸壳	57
2.1 凸壳问题简介	57
2.2 凸壳的应用	58
2.2.1 混合物勾兑	58
2.2.2 加速碰撞检测	58
2.3 平面点集凸壳的已有算法	59
2.3.1 卷包裹法	59
2.3.2 格雷厄姆算法	60

2.3.3	快速凸壳算法	62
2.3.4	分治算法	63
2.3.5	增量算法	65
2.3.6	周培德算法	67
2.3.7	实时凸壳算法	70
2.4	海量平面点集凸壳的解决方案	73
2.4.1	平面点集凸壳的城堡定理	73
2.4.2	城墙快速搜索算法	77
2.5	平面点集凸壳的一种高效算法	82
2.5.1	算法的基本思想	83
2.5.2	算法设计与实现	87
2.5.3	算法的效率分析与实验验证	88
2.5.4	快速凸壳算法的进一步优化	92
2.6	子凸壳的外直角三角定理	93
2.6.1	子凸壳的外直角三角定理	93
2.6.2	改进后的快速凸壳算法	95
2.6.3	实验结果	96
2.7	平面点集凸壳的两种近似算法	97
2.7.1	现有的近似凸壳算法	97
2.7.2	凸壳的近似度度量	99
2.7.3	点集坐标旋转法	100
2.7.4	多方向极值算法	105
3	平面点集的 Delaunay 三角剖分与 Voronoi 图	113
3.1	平面点集三角剖分简介	113
3.2	平面点集三角剖分的已有算法	113
3.2.1	逐点插入法	113
3.2.2	三角网生长法	114
3.2.3	分治方法	114
3.3	Delaunay 三角剖分	114

3.3.1	Delaunay 三角剖分简介	114
3.3.2	Delaunay 三角网构建方法	115
3.3.3	约束 Delaunay 三角网	118
3.3.4	一致性约束 Delaunay 三角网	119
3.4	Voronoi 图	120
3.4.1	Voronoi 图简介	120
3.4.2	Voronoi 图的构建	121
3.5	平面点集三角剖分的两条性质	121
4	多边形的三角剖分	124
4.1	多边形三角剖分简介	124
4.2	多边形三角剖分的已有算法	127
4.2.1	算法 1	127
4.2.2	算法 2	128
4.2.3	算法 3	128
4.2.4	算法 4	129
4.2.5	其他算法	130
4.3	简单多边形的快速单调剖分算法	130
4.3.1	算法相关概念与基本思想	131
4.3.2	算法描述	133
4.3.3	实现与验证	135
4.4	多边形的 Delaunay 三角剖分	136
5	艺术画廊问题	138
5.1	艺术画廊问题简介	138
5.1.1	艺术画廊问题及其数学模型	138
5.1.2	艺术画廊问题的诸多变种	139
5.1.3	艺术画廊问题解决现状综述	139
5.2	基于可见传播规则的艺术画廊问题的求解方法	143
5.2.1	可见传播规则	144

5.2.2	艺术画廊问题的求解步骤	145
5.2.3	实验验证	145
5.2.4	算法效率分析	147
5.3	基于顶点可见关系矩阵的艺术画廊问题解决方法 ...	147
5.3.1	可见关系矩阵及其性质	147
5.3.2	多边形两个顶点可见性的判定	150
5.3.3	多边形顶点可见关系矩阵的自动建立	152
5.3.4	监视顶点问题求解——初步解	153
5.3.5	艺术画廊问题解的判定——最终解	156
5.3.6	实验及结论	158
5.3.7	监视点可位于多边形内部或边界上的艺术 画廊问题	160
6	计算几何与空间数据挖掘	162
6.1	概述	162
6.2	凸壳与空间数据分类	162
6.3	基于 Delaunay 三角网的可视化空间数据聚类	163
6.3.1	算法的基本思想	164
6.3.2	算法的设计与实现	165
6.3.3	算法效率与特点分析	168
6.4	基于 Delaunay 三角网的高维空间数据聚类	169
6.4.1	算法的基本思想	170
6.4.2	该算法的相关定义	171
6.4.3	算法描述	173
参考文献		176

1 相关领域研究与发展现状

1.1 计算几何概述

1.1.1 计算几何简介

计算几何的历史，最早可以追溯到 17 世纪中叶。但人们普遍认为该研究领域出现于 20 世纪 70 年代。

1644 年，法国数学家 Descartes（笛卡儿）论及 Voronoi 图。1759 年，瑞士数学家、物理学家 Euler（欧拉）和 Vandermonde 一起讨论过欧几里德 TSP 问题。1975 年，Shamos（沙莫斯）和 Hoey（霍伊）利用计算机有效地计算平面点集的 Voronoi 图，并发表了一篇著名论文，从此计算几何诞生了^[1]。1978 年，Shamos 在他的论文中定义了现代计算几何。1985 年，Preparata（普雷帕拉塔）和 Shamos 撰写了第一本计算几何教科书。

由于大量的应用领域提供了特有的几何问题，这些问题的解决都要依靠设计精巧的几何算法，这就是孕育计算几何这门学科的基础。特别是 CAD 领域，提出的一系列问题均需要计算几何基础理论的支持。

从对问题的明确表述到得到高效而优雅的解决方法，往往需要经历漫长的过程，其间既要克服很多困难，也要积累一些次优的中间结果。由几何问题带来的挑战也吸引了众多的研究人员。因此，自计算几何这门学科诞生以来，该研究领域取得了辉煌的成果，使得计算几何成为理论计算机科学领域中一个新的极有生命力的子领域，并且，计算几何中的研究成果已在计算机图形学、图像分析、化学、统计分析、模式识别、地理数据库以及其他许多领域中得到了广泛的应用。

1.1.2 计算几何的研究内容

计算几何是计算机科学中的一个分支,是专门研究有关几何对象问题的。计算几何被定义为“针对处理几何对象的算法及数据结构的系统化研究”,其重点在于“渐进快速的精确算法”。计算几何研究的典型问题由几何基元、查找、优化等问题类组成。

首先,几何基元包括凸壳和 Voronoi 图、多边形的三角剖分、划分问题 (partition problems) 与相交问题。 E^{d+1} 中点集 S 的下凸壳在 E^d 中的投影恰好是点集 S 在 E^d 中投影点的 Delaunay 三角剖分,然后由 Delaunay 三角剖分可以容易地得到 Voronoi 图。换言之, Voronoi 图是凸壳的特例,因此,构造 E^{d+1} 中点集凸壳的算法也可以用于构造 E^d 中点集的 Voronoi 图。对多边形的三角剖分问题可以提出如下要求:设计复杂度低的算法构造多边形三角剖分以及设计三角形最小角最大化的三角剖分算法;分割线段长度之和最小的三角剖分算法。前者已有线性时间算法。划分问题是多边形三角剖分的推广,它要求把几何体划分成若干好的部分。好的部分通常是指下述两个目标之一:划分成尽量少的凸部分;各凸部分最小角最大化。另外在几何体中可以加入 Steiner 点 (新的顶点),然后再进行划分,使得划分线段长度之和最小化或者是提出其他要求。二维中的典型相交问题:给定平面上 n 条直线段,确定所有的相交线对。三维中的相交问题一般考虑两个凸多面体的交以及两个多面体的交。

其次,几何查找包括点定位、可视化、区域查找等问题。计算机图形学、数据库中的区域查找及地理图形中的点定位等都是几何查找中的典型例子。在平面细分 (planar subdivision) 中定位一个询问点或者在 E^d 中 ($d \geq 3$) 内由 n 个超平面构成的结构中定位询问点的问题是一个典型问题,现在不仅有解决这个问题的确定性算法,而且设计了动态随机增量算法。给定平面上 n 个顶点的简单多边形 P ,由点 q 向任一方向引射线 l ,确

定 l 和 p 相交的第一条边，这个问题的解决为可视化问题的求解提供了前提。 E^d 中给定点集 S 及区域集合 B , $b \in B$, 要求在 b 中查找 S 中的点，这就是区域查找问题。

最后，几何优化包括参数查找和线性规划。参数查找技术是将一个优化问题的检验算法变成寻找解的算法，它必须满足某些条件（检验算法是可以并行的），并且具有广泛的应用性。例如，可用它来求解平面中二维中心问题，还可以用来完成三维空间中射线的安置。众所周知，有确定变元数目的线性规划问题已有线性时间算法求解，但对于广义线性规划是否存在多项式时间算法还有待进一步研究。

此外，计算几何中各种问题的下界的确定、推导下界的方法以及求解各种几何问题的算法的复杂性分析等，也是计算几何研究中重要内容。

计算几何中引入随机化之后，已经设计出非常有效的概率算法求解诸多几何问题。随机化给几何算法设计带来了两种新的设计思想：基于随机抽样的分治方法；利用随机顺序插入产生随机增量结构。此外，随机几何算法的复杂性分析以及随机增量结构的非随机化也是重要的研究内容。

今天的计算几何学常被称为计算机科学中的一个新学科，它主要在概念上研究设计和分析有更丰富内容的算法。电子计算机对这个领域有强有力的影响，是赋予计算科学以技术翅膀的工具。计算科学的实质是结构性的（即计算的、算法的）数学，它独立于应用算法的机械技术，而算法产生于机械技术的应用。

现在的计算几何学是一个广大的领域，它包括几何探测、画廊算法和理论、计算机图形学、动态计算几何学、并行计算几何学、正交计算几何学、数字计算几何学、计算拓扑学等。

1.1.3 计算几何的发展与现状

20 世纪 70 年代末，计算几何从算法设计与分析中孕育而

生^[2]。该学科已经有了巨大的发展，不仅产生了一系列重要的理论成果，也在众多实际领域中得到了广泛的发展。

计算几何是随着计算机科学的发展而成长起来的一门新学科，一方面，计算机工具的引入大大拓展了几何学的传统研究方法和研究对象；另一方面，计算机在各行各业中的广泛应用又提出了一系列亟待解决的新的几何问题。因此，计算几何的内容日益丰富，发展十分迅速。信号与图像处理、图像与模式识别、可视化等领域遇到的许多问题可归结为计算几何问题。

目前，计算几何已经拥有自己的学术刊物和学术会议，形成了一个由众多活跃的研究人员组成的学术群体，成长为一个被广泛认同的学科。它涉及的问题及其解答本身所具有的美感吸引了众多的学者。在众多的应用领域中（诸如计算机图形学、地理信息系统、机器人学等）计算几何算法都发挥了重要的作用。有关计算几何的主要国际性学术会议有：SCG：ACM Symp on Computational Geometry；GMP：Geometry Modeling and Processing；CCCG：Canadian Conf on Computational Geometry；JCDCCG：Japan Conference on Discrete and Computational Geometry。

新近的计算几何包括几何抽样理论、计算实代数几何、计算拓扑、运动规划、并行计算几何、隐藏面的移动、结构和图形、网络生成以及计算机视觉中的几何问题等。计算机在各学科领域深层次的应用将为计算几何提出更多的研究问题，反之，计算几何的研究成果也将促进这些学科的进一步发展。

1.1.4 计算几何与其他学科的关系

作为计算机科学的一个分支，计算几何主要研究解决几何问题的算法。在现代工程和数学领域，计算几何在图形学、机器人技术、超大规模集成电路设计和统计学等诸多领域有着广泛的应用。实际上，计算几何应该是古典几何学与现代计算机科学相结合的产物。而 GIS、CAD/CAM、计算机图形学、机器人学乃至知识发现等则是计算几何的应用领域。没有古典几何

学成熟的理论基础，没有计算机科学的飞速发展，就不可能有计算几何这门学科的创立。

没有 GIS、CAD/CAM、计算机图形学、机器人学乃至知识发现等的发展，也就没有计算几何的迅猛发展，使之很快成为一门学科。例如，计算几何在 GIS 中应用比较多的有 Delaunay 三角网、Voronoi 图、凸壳、最短路径求解、拓扑计算等。这些模块在 GIS 中的应用范围很广。利用 Delaunay 三角网建立 DEM，利用 Delaunay 三角网、Voronoi 图协调目标间的拓扑关系、点状地物的凸壳、嵌套结构化与 Voronoi 图相结合进行自动选取，交通运输中的最短路径求解等。

值得一提的是，在我国的学科分类标准中，计算几何学隶属于几何学，而非计算机科学。

1.2 知识发现概述

1.2.1 KDD 的产生与发展

KDD (Knowledge Discovery in Databases) 是人工智能、机器学习与数据库技术等相结合的产物，是一门交叉性学科^[13,14]。KDD 一词是在 1989 年 8 月在美国底特律召开的第 11 届国际人工智能联合会议的专题讨论会上首次正式提出的。随着 KDD 在学术界和工业界的影响越来越大，于 1995 年 KDD 组委会将专题讨论会更名为国际会议，每年均召开一次，第一次 KDD 国际学术会议是在加拿大蒙特利尔市召开。1998 年在美国纽约举行的第四届知识发现与数据挖掘国际学术会议不仅进行了学术讨论，并且有 30 多家软件公司展示了它们的数据挖掘软件产品，不少软件已在北美、欧洲等国得到应用。

用数据库管理系统来存储数据，用机器学习的方法来分析数据，挖掘大量数据背后的知识，正是这两者的结合使得 KDD 技术应运而生。Usama M. fayyad^[15] 等人给出了 KDD 描述性定义：KDD 是从大量数据中提取出可信的、新颖的、有效的并能

被人理解的模式的高级处理过程。通过这一过程，感兴趣的知識或高层信息可以从数据库相关数据集中抽取出来并进行研究。KDD 也可以这样来描述：它是指对真实数据库（具有大数据量、不完全性、不确定性、结构性、稀疏性等特征）中数据所隐藏的、先前未知的及具有潜在应用价值的信息进行非平凡抽取，这些信息包括知识、规则、约束和正则性等。

近几年来，国际上对 KDD 的研究已取得了一定的成果，许多研究院和大型计算机公司开发出了一些较成型的产品或工具，不同的工具可以完成不同的数据挖掘任务，并可利用不同的方法达到同一目标。

与国外相比，国内对 KDD 的研究稍晚，没有形成整体力量。1993 年国家自然科学基金才首次支持对该领域的研究项目。目前，国内的许多科研单位和高等院校竞相开展知识发现的基础理论及其应用研究，这些单位包括清华大学、中科院计算技术研究所、空军第三研究所、海军装备论证中心等。北京系统工程研究所对模糊方法在知识发现中的应用进行了较深入的研究，北京大学也在开展对数据立方体代数的研究；华中理工大学、复旦大学、浙江大学、中国科技大学、中科院数学研究所、吉林大学等单位开展了对关联规则开采算法的优化和改造；南京大学、四川联合大学和上海交通大学等单位探讨、研究了非结构化数据的知识发现以及 Web 数据挖掘；北京科技大学知识工程研究所主要从事 KDD 内在机理的研究，在国内率先构建并逐步完善与拓展了基于内在机理的知识发现理论 KDTIM。

1.2.2 KDD 技术研究和应用存在的问题与发展趋势

目前，KDD 技术的研究还有待进一步完善，其应用还存在较大的局限性，具体表现在以下几方面：

(1) 数据挖掘算法的有效性可与测性问题。数据挖掘的对象向更大型的数据库、更高的维数和属性之间更复杂的关系方

向发展,从而导致组合爆炸,大大提高了知识搜索的代价。从一个大型数据库中抽取知识的算法必须高效、可测量,否则就不具有实用价值。

(2) 多种形式的输入数据问题。目前,数据挖掘工具能处理的数据形式有限。

(3) 与数据库的无缝连接问题。当前的数据分析工具倾向于离线存取数据库,这就会导致大量费时的重复性的 I/O 处理。

(4) 用户参与领域知识问题。目前的数据挖掘系统或工具很少能真正做到让用户参与到挖掘过程中。将相关领域的知识融入数据挖掘系统中是一个重要但没有很好解决的问题。

(5) 证实 (Validation) 技术的局限问题。数据挖掘使用特定的分析方法或逻辑形式发现知识,比如,归纳或演绎。但是系统却没有能力去证实发现的知识,使得发现的知识没有普适性而不能成为有用的知识。

(6) 不同技术的集成问题。任何算法都不是万能的,一个真正有用的工具必须为解决不同的问题提供不同的解决方法。因此,重要的是提出一种体系结构,使得新方法易于合成,已有方法便于运用。

(7) 知识的表达和解释机制问题。在许多实用系统中,最为重要的是用户能够理解发现的知识。只有当 KDD 系统能提供更好的解释机制,用户才能更有效地评价这些知识。

(8) 知识的维护和更新问题。目前研究采用增量更新的方法、数据快照和时间戳等方法对知识进行动态维护和及时更新。比如, D. W. Cheung 等人提出了维护关联规则的增量算法。

(9) 支持的局限、与其他系统的集成问题。目前的数据挖掘系统尚不能支持多种平台。数据挖掘系统和其他一些用户已经熟悉决策知识系统的有机集成,对于系统充分发挥作用是非常重要的。

(10) 数据私密性和安全性问题。可通过改进数据库安全方法, 阻止非法侵入, 以避免信息泄露和丢失。

1.2.3 基于内在认知机理的知识发现理论

综观知识发现的理论和技术方法的发展状况与存在的问题, 杨炳儒教授对知识发现的内在机理进行了全面和深入的研究。从知识发现、认知科学与智能系统等多学科交叉结合的角度, 以认知自主性为核心概念, 将知识发现视为一个开放的和不断进化的认知系统, 研究它的系统结构、方法、进化与运行机制。并于 1997 年提出了知识发现系统内在机理的研究方向^[16,17] (提出其涵盖的三个机制: 双库协同机制、双基融合机制^[18]、信息扩张机制), 揭示了其作为认知系统潜在的本质、规律复杂性; 在国内外率先独立构建并逐步完善与扩展了基于内在机理的知识发现理论 KDTIM (由 5 个层面组成)^[19,20]。

第一层面 (理论基础层): 由若干基础理论研究成果组成。如提出多层次结构逻辑、广义归纳逻辑因果模型; 提出因果关系能行可判定方法、因果关系定性推理模型与方法; 提出专家知识的归纳获取机制、语言场与语言值结构的知识表示方法等。

第二层面 (内在机理层): 由相关的 3 个原理 (即双库协同机制^[21,22]、双基融合机制、信息扩张机制) 构成的内在机理的内涵组成。其中包括: 结构对应定理、可达关系概率估计定理、启发与维护协调算法; 过程模型逻辑等价定理、RST 三类协调算法; 参数演化定理、“不动点原理”与“突变性原理”等。

第三层面 (结构模型层): 由内在机理研究诱导的 6 个新结构模型组成。其中包括: (1) 基于双库协同机制的 KDD* 结构模型 (用于处理结构化数据挖掘问题, 它区别于固有的 KDD 模型); (2) 基于双库协同机制的复杂类型数据挖掘模型 DFSSM; (3) 基于双基融合机制的 KDK* 结构模型; (4) KD (D&K) 结构模型 (它强调知识发现过程中的认知自主性, 突出知识的自动发现); (5) DKD (D&K) 总体结构模型; (6) KDD*E 结构