

# 大型层次信息可视化方法研究

景民昌<sup>1,2</sup>, 孙洁丽<sup>1</sup>

(1. 中国科学院 文献情报中心, 北京 100080; 2. 中国石油大学 图书馆, 北京 102249)

**摘要:** 层次信息的可视化是信息可视化领域的研究热点。本文介绍了几种常见的大型层次信息的可视化方法, 并对它们在显示空间、信息节点组织与显示、空间利用率等方面的优缺点作了对比分析。

**关键词:** 信息可视化; 层次信息; 节点连线图; 空间填充

**中图分类号:** G350      **文献标识码:** A      **文章编号:** 1007 - 7634(2008)04 - 0541 - 05

## Study on Methods of Large Hierarchical Information Visualization

JING Min - chang<sup>1,2</sup>, SUN Jie - li<sup>1</sup>

(1. Library of Chinese Academy of Sciences, Beijing 100080, China;

2. Library of China University of Petroleum, Beijing 102249, China)

**Abstract:** The visualization of hierarchical information is a hot topic in the information visualization community. This paper introduces several important methods of large hierarchical information visualization, and compares them in display space, organizing and layout of node, efficiency of space using.

**Key words:** information visualization; hierarchical information; node - link diagram; space - filling

## 1 引言

信息可视化是情报学研究的新领域<sup>[1]</sup>。信息可视化将信息或数据转换为人们可以直观、形象理解的图形或图像表达方式, 从而大大加快数据的处理速度, 使时刻产生的海量数据得到有效利用。在“信息爆炸”的今天, 信息可视化在人和信息之间架起了一座桥梁, 为日益严重的“信息迷航”问题提供了一种可行的解决方法<sup>[2]</sup>。

信息可视化的信息类型包括一维信息、二维信息、三维信息、时态信息、多维信息、层次信息、网状信息<sup>[3]</sup>。其中, 层次信息是一种较为普遍的信息类型, 如文件系统、单位组织结构、家谱图、各

种分类法等都是层次信息。层次信息的特点是除根节点外, 每一节点都有一个链接指向其父节点, 父子节点及其链接还可以包含有多个属性。

层次信息的可视化是信息可视化领域的研究热点之一<sup>[3-4]</sup>。针对大型层次信息的可视化问题, 国内外作了大量研究工作, 提出了许多新的可视化方法, 构建了一些可视化系统。本文分析了几种主要的大型层次信息可视化技术, 并对其作了比较研究。

## 2 大型层次信息可视化方法

传统的描述层次信息最直观的方法是树型结构。然而, 对于大型层次信息来说, 传统的方法很

收稿日期: 2007 - 09 - 03

作者简介: 景民昌 (1972 -), 男, 山西省永济人, 工程师, 博士生, 从事可视化技术与数据挖掘研究; 孙洁丽 (1969 -), 女, 山东省文登人, 副教授, 博士生, 从事文本挖掘、可视化与竞争情报系统研究。

难满足要求。这主要是因为大型层次信息节点多、层次多,需要占据大量的可视化显示空间,而一般的显示空间(主要是计算机屏幕)是十分有限的,无法用一个视图表示整个结构。用户经常需要通过拖拉滚动条来浏览层次节点,很难一次性概览整个层次结构;另一方面由于层次信息的结构在横向(每层节点的个数)和纵向(层次结构的层数)扩展上不成比例,分支很快就会拥挤交织在一起,导致树型结构变得混乱不堪<sup>[4]</sup>。如何在有限的计算机屏幕空间上,显示层次结构尽可能多的节点,并将焦点部分突出显示,同时又能够把整个层次结构显示出来,是层次信息可视化研究的重点和难点。

根据实现方法的不同,大型层次信息的可视化基本可以分为两种:一种是节点连线图法(mode-link),一种是空间填充法(space-filling)。

### 2.1 节点连线图法

节点连线图法是在一定空间中,用线段将具有关系(主要是父子关系)的节点彼此联接起来,形成树型结构的一种方法,它可以直观地表示层次关系。由于大型层次信息具有节点多、层次多的特点,需要一定的技术来实现树型结构。常用的技术有空间压缩、约束树、交互技术、焦点+上下文(focus+context)技术等<sup>[5]</sup>。其中焦点+上下文技术应用较多,它将一个信息集合特定部分的细节视图,通过某种方式和该信息集合的总体结构视图混合在一起,呈现给读者,使读者在浏览一个大的信息空间(context)的同时,还能够关注部分细节(focus)。

在信息可视化中,常用的空间模式有平面空间、双曲空间、三维空间等。与平面空间相比,双曲空间和三维空间可利用的空间范围更大,故常被用于大型层次信息的可视化。在不同的空间模式上都可以应用焦点+上下文技术,如Hyperbolic Tree<sup>[6]</sup>是在双曲空间上应用了焦点+上下文技术,Cone Tree<sup>[7]</sup>是在三维空间上应用了焦点+上下文技术。下面对这两种技术作一介绍。

(1) Hyperbolic Tree。Hyperbolic Tree是John Lamping等提出的一种基于双曲几何的可视化和操作大型层次信息的focus+context技术<sup>[6]</sup>,称为双曲树。这种技术将更多的可视化空间用于层次结构中当前被关注的部分,而同时又能把整个层次结构显示出来。该技术通过一种规范的算法将层次信息显示在一个双曲平面上,然后将这个双曲平面映射到

显示区域中。所选择的映射方式提供了一种鱼眼变形来支持focus和context之间的平滑过渡(图1)。

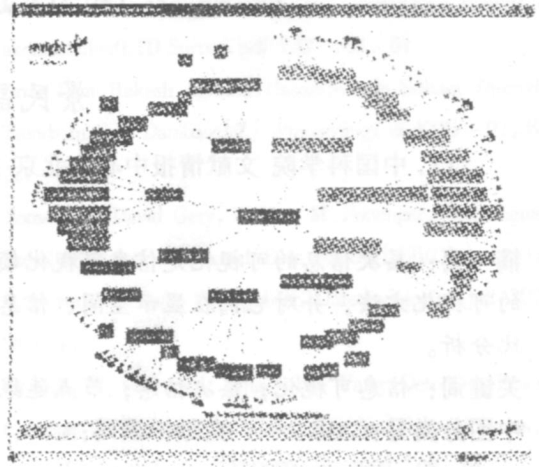


图1 Hyperbolic Tree

Hyperbolic Tree以新颖、动态、易于实现等特点,引起许多关注。如,美国国家科学数字图书馆(NSDL)在其主页上利用Hyperbolic Tree实现学科导航功能<sup>[8]</sup>;其分馆计算与信息技术数字图书馆(CITIDEL)在开发的交互式检索系统Citiviz中也应用了Hyperbolic Tree对检索结果进行可视化显示<sup>[9]</sup>;靖培栋等利用Hyperbolic Tree实现了《中图法》分类体系的可视化<sup>[10]</sup>。

(2) Cone Tree。Cone Tree是Robertson等提出的一种利用三维图形技术对层次结构进行可视化的方法,称为锥形树<sup>[7]</sup>。其基本思想是利用三维图形技术将传统的二维树形表示法扩展到三维空间。在Cone Tree中,表示层次结构的整棵树以三维的形式进行组织和显示。利用三维锥形体来实现层次结构中父子节点之间的连接,层次结构的顶部放置在可视化空间的顶端或左端,每个锥体的顶点表示该层结构的顶点,其子节点(三维)均匀排列在该锥体的底部(图2)。锥体的底面直径随着层次结构的深入逐渐减小,以保证最低层的结构也能在可视化空间中有效表示。每个锥体之间透明遮拦,从而可以保证每个锥体能够很容易被感知,还不会妨碍后面的锥体显示。同时辅以旋转、拖动等交互技术,可以很容易的实现复杂层次关系的把握。

### 2.2 空间填充法

空间填充法的基本实现思想是将屏幕空间划分成一系列的形状单元(如矩形、圆形等),每个形状单元代表层次结构中的一个信息节点,形状的面

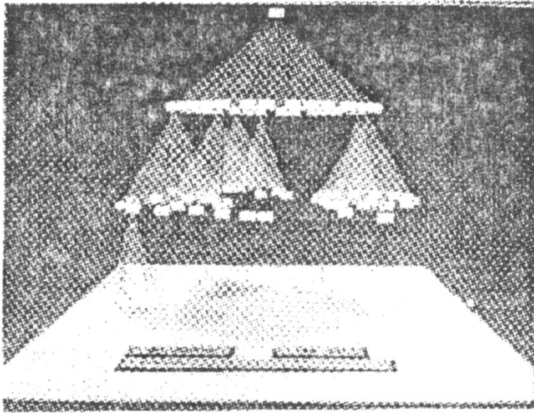


图 2 Cone Tree

积表示相应节点的权重。与节点连线图法相比，空间填充法一般更能有效地利用显示空间。

根据填充方式的不同，主要有矩形填充法、圆形填充法等。各种方法的关键区别在于对屏幕空间的划分方式不同，最具代表性的矩形填充算法是 Tree - maps (标准树图)<sup>[11]</sup>。由 Tree - maps 也衍生了一些改进算法，如 Cushion Treemaps (柔型树状图)<sup>[12]</sup>、Beamtrees (柱型树状图)<sup>[13]</sup>等；圆形填充算法是用圆、圆环或分割圆环的扇形体代表层次信息的节点。与矩形填充算法相比，圆形填充算法更能展示出层次信息的整体结构，可使用户对层次结构有一个概览，但其空间利用率一般要比矩形填充算法低。圆形填充算法有用放射圆 (radial circles) 来代表节点，如 John Stasko 等开发的 Sunburst 可视化系统<sup>[14]</sup>；也有用嵌套圆 (nested circles) 来代表节点，如最近引人注目的新一代搜索引擎 Grokker<sup>[15]</sup>。

(1) 矩形填充算法。Tree - maps 是 Shneiderman 等提出的一种层次信息可视化模型<sup>[11]</sup>。它用矩形表示层次结构中的每个节点，矩形的面积表示相应节点的权重。父节点的所有子节点矩形都嵌套在父节点矩形之内，被该父节点矩形所包围。Tree - map 可视化结构充分利用了显示空间，将层次结构映射为一个矩形，使大型层次结构能够在有限的空间中进行显示。显示空间被分割成互相包围的一些矩形，这些矩形表示树形结构，包围在某个矩形中的节点的画法完全依赖于节点的内容，每个节点的显示尺寸基于它相对于整棵树的比重。

Tree - maps 由于其实现简单，显示简洁，应用较多。如，新浪网开发的“沪深 300 地图”<sup>[16]</sup>，通过行业对构成沪深 300 指数的 300 只沪深两市股票

进行组织，每一个小矩形块代表一家公司，矩形的面积代表该公司流通市值在整个沪深 300 成分股流通市值中所占比重，其面积会根据每天涨跌自动调整；矩形颜色代表该股票当日涨跌情况，其中红色为上涨，绿色为下跌，其颜色深浅根据涨跌幅度而变化；当鼠标指向某一小块儿时，会自动显示该公司基础资料，点击则可查看其实时行情。同时，还可在查询框中输入股票代码、名称或拼音缩写对股票进行查询 (图 3)。

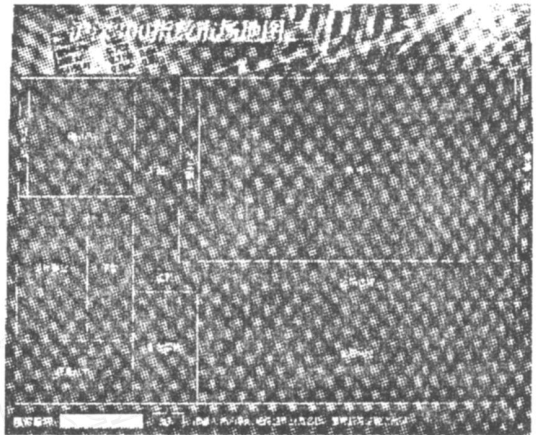


图 3 Tree - maps

Cushion Treemaps (柔型树状图)<sup>[12]</sup>是对标准树状图的一种改进算法。它通过给每个矩形表面填充阴影，利用光线的明暗变化，来隐喻每个节点的信息。从侧面看，好像每个矩形都有不同的隆起高度。柔型树状图充分利用了每一个小矩形的表面来承载更多的信息，改善了树状图的外观。

Beam Trees (柱型树状图)<sup>[13]</sup>也是标准树状图的一种改进，它利用矩形的堆叠、阴影、和 3D 效果来显示层次信息树的深度结构。与标准树状图不同的是，分配给叶子节点的空间减少了，但是其层次深度关系更明确了，并且提高了空间利用率。

另外还有其他的一些改进算法，如 squarified treemap, ordered treemap、strip treemap、quantum treemap 等。

(2) 圆形填充算法。Sunburst 是 John Stasko 等开发的基于放射圆填充的一种可视化层次信息的工具<sup>[14]</sup>。它将根节点置于中心，圆环向外辐射，层次越低的节点距离中心圆环越远，相同层次的节点位于同一圆环上，代表子节点的扇体分布在父节点扇体的外围，扇体的角度与节点的权重大小成正比 (图 4)。

Grokker 实际上是一个对检索结果可视化处理

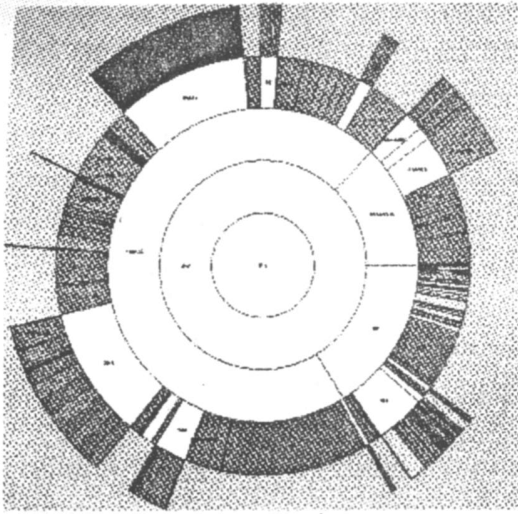


图4 放射圆填充法

的搜索引擎<sup>[15]</sup>。它首先将检索引擎得到的检索结果进行分类归档,每个分类目录被显示为一个彩色的圆,圆内可以嵌套子分类或叶节点。当点击相应的圆时,相关的区域就会被放大,里面就会显示更多的子内容(图5)。Grokker 提供了一种崭新的检索结果显示方式,引起了广泛关注,如 IBM 公司、Stanford 大学、Michigan 公司应用了 Grokker 进行信息管理<sup>[17]</sup>;著名的数据库商 EBSCO 也在其检索系统中引进了 Grokker 插件,形成了其独具特色的“视觉检索”模块<sup>[18]</sup>。

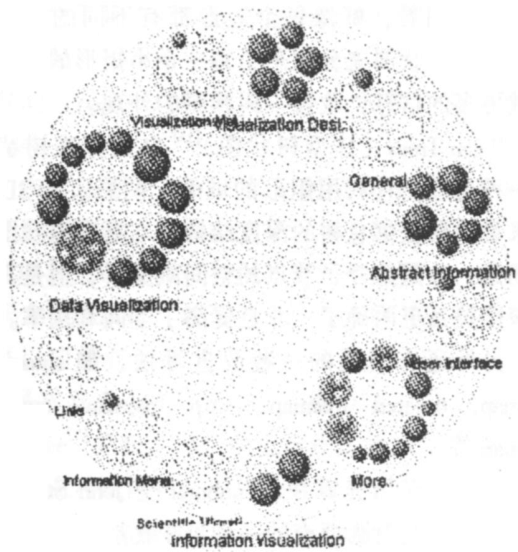


图5 Grokker

Weixin Wang 等借鉴了 TreeMap 和 Grokker 的特点,利用圆填充(Circle Packing)技术,在二维空间和三维空间上实现了基于嵌套圆填充的层次信息

可视化方法(图6)。

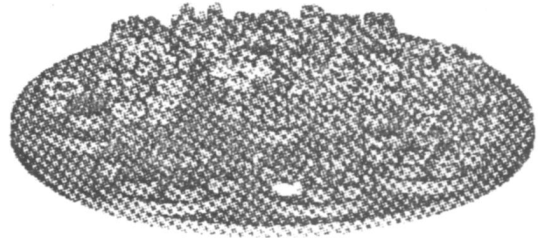


图6 基于 Circle packing 的填充法

### 3 几种方法的比较

大多数大型层次信息的可视化研究主要基于三点考虑:增加可利用的显示空间,如利用三维空间或双曲空间;对信息节点进行合理组织,如分组或隐藏部分节点,从而减少显示的节点数量;充分利用有限的显示空间,改进空间的利用效率。本文中的这几种可视化方法,在显示空间、信息节点的组织 and 显示、空间利用率等方面各不相同,下面对它们的优缺点作一比较。

(1) Hyperbolic Tree 采用双曲空间作为信息的显示空间,增加了有限空间上可容纳的信息节点;利用焦点+上下文技术,隐藏非焦点的信息节点,从而突出显示焦点节点。同时通过交互手段,解决了层次结构中节点和上下文之间的平滑过渡问题;但是,与空间填充法(如 Tree - maps)相比,其空间利用效率不高。另外,节点之间的关系有时较难把握,比如,当把当前焦点节点移到显示区域的中心时,其父节点则可能被推移到周围,这时较难分辨出父子关系;对处于同一层次上的不同父节点的子节点关系(准兄弟关系)体现较难。

(2) Cone Tree 采用了三维空间对信息进行显示,可以比较容易地体现出树形结构的整体信息,在有限的屏幕上也可以显示更多的节点;在信息节点的组织方面,可以利用的手段较多,如,除了几何结构外,还可以通过圆锥在平面上的阴影映射现实节点的分布情况;但是,其空间利用率有限,空间分布极度不均匀,有的区域信息过于密集,而有的空间则完全没有得到利用,尤其是在层次结构本身不均衡的情况下尤为突出;另外,节点之间相互遮挡的情况也较为严重。

(3) Tree - maps 的最大优点在于其显示空间利用率最高,它充分地利用了计算机屏幕空间,而且显示简洁,易于实现,使得大型层次信息能够在有

限的屏幕空间中进行显示,并且使语义信息的表示变得相对容易;每个节点的显示尺寸是基于它相对于整棵树的比重的,因而很容易找到信息树中的最大尺寸属性的节点。但是,由于 Tree - maps 将绝大多数空间分配给了叶子节点,而枝节点是编码隐含的,因而丧失了层次结构的直观性,树的层次关系很难分清。最坏的情况是平衡树的显示,当每个双亲节点都有相同的子节点,而且每个叶子节点的尺寸属性都相同,这时,树状图就退化成为了标准的网格。柔型树状图是 Tree - maps 的改进,保留了空间利用率高的优点。以光线的明暗变化来隐喻每个节点的信息,改善了树状图的外观,通过矩形表面的明暗变化来克服标准树状图无法显示信息树结构的缺陷;但是恶化了布局的稳定性,还需要观察者有强大的认知能力去从图形中提取重构信息。

(4) 放射圆填充法在信息节点的组织 and 显示上,比 Tree - maps 更能够揭示出整个层次结构,但是其空间利用效率不如 Tree - maps 高,因为放射圆外围的空间不能被有效利用。另外,在层次节点较多,深度较大时,在放射圆外围的低纵横比的扇形有时不能够区分清楚。

(5) Grotter 是一种新颖的层次信息可视化方法,但是, Grotter 不能展示整个层次信息结构的概览。因为当浏览父节点时,子节点不可见,只有点击该节点才能显示该节点的子分类。Grotter 的空间利用率虽比 Tree - maps 等矩形填充算法低,但要比 Hyperbolic Tree、Cone Tree 等节点连线图法高。

(6) Circle Packing 法使用嵌套圆代替矩形,容易看清楚节点的分组以及结构间的关系,层次结构的总体关系更为清晰。其节点分布比较紧凑,能够在有限的空间显示较多的节点。其空间利用的有效性虽不如 Tree - maps,但要比 Grotter 高。

## 4 结 语

信息可视化虽然还是一个较新的研究领域,但已引起人们的广泛注意。大型层次信息的可视化是可视化领域的研究热点之一,本文从大型层次信息可视化的两种实现方式出发,介绍了几个常见的可视化技术或系统,并对它们的优缺点作了对比和分析。

### 参考文献

1 靖培栋. 信息可视化——情报学研究的新领域[J]. 情报

科学,2003,21(7):685-687.

- 2 周宁,张会平,陈勇跃. 信息可视化与知识组织[J]. 现代图书情报技术,2006,(7):62-65.
- 3 Ben Shneiderman. The Eyes Have It: A task by Data Type Taxonomy for Information Visualizations[C]. Proc. IEEE Workshop Visual Language'96,1996:336-343.
- 4 周宁,刘玮,赵丹. 信息提供的可视化研究[J]. 情报科学,2004,22(3):257-260.
- 5 Stuart K. Card,David Nation. Degree - of - Interest Trees: A Component of an Attention - Reactive User Interface[EB/OL]. <http://www2.sims.berkeley.edu/courses/is247/f05/readings/Card.DOITrees.AVI02.pdf>. ,2007-07-25.
- 6 J. Lamping, R. Rao. The hyperbolic browser: A focus + context technique for visualizing large hierarchies[J]. J. Visual languages and Computing,1996,7(1):33-55.
- 7 George G. Robertson, Jock D. A4ctckinlay, Stuart K. Card. Cone trees: animated 2d visualizations of hierarchical information[EB/OL]. <http://portal.acm.org/citation.cfm?doid=108844.108883>. ,2007-07-25.
- 8 Nithiwat kampanya, Rao Shen, Seonho Kim, et al. Citiviz: A visual user interface to the CITIDEL System[EB/OL]. <http://infovis.cs.vt.edu/papers/ecdl2004-citviz.pdf>. ,2007-07-25.
- 9 靖培栋. 基于双曲几何的《中图法》分类体系的可视化[J]. 现代图书情报技术,2005,(1):40-42.
- 10 B. Shneiderman. Tree visualization with tree - maps: a 2d space - filling approach[J]. ACM Transactions on Graphics,1990,11(1):92-99.
- 11 TreeMap Java Library[EB/OL]. <http://treemap.sourceforge.net/>. ,2007-07-25.
- 12 Frank van Ham, Jarke J. van Wijk. Beamtrees: compact visualization of large hierarchies[J]. Information Visualization,2003,(2):31-39.
- 13 John Stasko, Eugene Zhang. Focus Context Display and Navigation Techniques for Enhancing Radial, Space - Filling Hierarchy Visualizations[EB/OL]. <http://www.ifs.tuwien.ac.at/~silvia/wien/gwa/ws05/articles/Stasko-InfoVis2000.pdf>. ,2007-07-25.
- 14 Grotter[EB/OL]. <http://www.groxis.com>. ,2007-07-25.
- 15 沪深300指数市场地图[EB/OL]. <http://finance.sina.com.cn/stock/stockmap/hs300.shtml>. ,2007-07-25.
- 16 Client Case Studies[EB/OL]. <http://www.groxis.com/service/grotter/resources.html>. ,2007-07-25.
- 17 EBSCO[EB/OL]. <http://www.ebsco.com>. ,2007-07-25.
- 18 Weixin Wang, Hui Wang, Guozhong Dai, et al. Visualization of Large Hierarchical Data by Circle Packing[C]. ACM Conference on Human Factors in Computing Systems (CHI '06), Canada,2006.

(责任编辑:孙晓明)