

1 Parallel sets: Interactive exploration and visual analysis of categorical data[6]

用 Parallel Set 方法可视分析分类数据

Parallel Set 是用来探索分类数据的可视化方法，它展示的是数据出现的频率而不是通常的单个数据点的特征。方法基于平行坐标的坐标轴布局，轴上的盒子（boxes）代表分类，轴之间的四边形代表分类之间的关系。通过交互用户可以重新定义数据和分类的映射，分析考虑更多的数据维度。

Parallel Set 是 Parallel Coordinates 的变种。Parallel Coordinates 是一种多维数据的可视化方法，多个坐标轴平行放置，每个坐标轴代表一个维度，一条数据对应一条折线，折线与坐标轴的焦点代表数据在这个维度上的值。因此一般每个维度为数值数据。

应用到分类数据的时候，分类到数值的映射可以是任意的，因此需要定义类别到坐标轴的映射。最基本的是把分类映射到离散的几个值上 (图1a)，如果引入数据出现的频率，可以根据某一分类数据量的多少把分类映射到坐标轴上的某一区间 (图1b)。

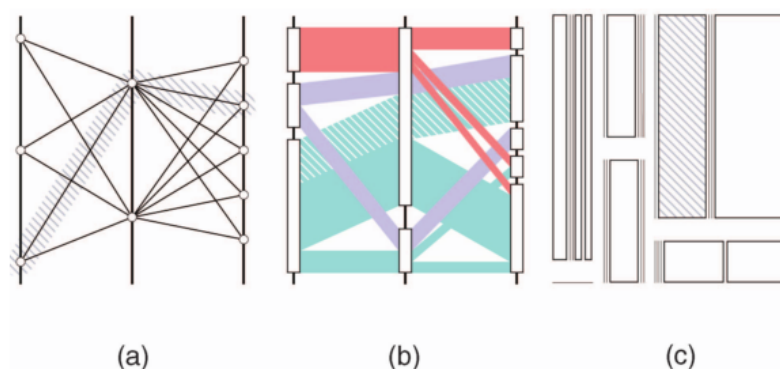


Fig. 3. Three different visualization techniques displaying the same data: (a) The categories are represented by points on continuous axes in parallel coordinates. (b) Parallel sets show the frequencies of categories and relations. (c) A mosaic display provides a compressed view of the data (the hatched parts represent the same subset).

Figure 1: Parallel Coordinates

Parallel Set 是一种可视化技术，还是一种交互框架。它能够很自然的将类别变量映射到可视实体上，这使得交互分析更有效率。

1.1 交互 (图2)

Shneiderman's visual information seeking mantra: **overview first, zoom and filter, then details-on-demand**

- 选择高亮
- 交互查询
- 过滤
- 维度和分类重排

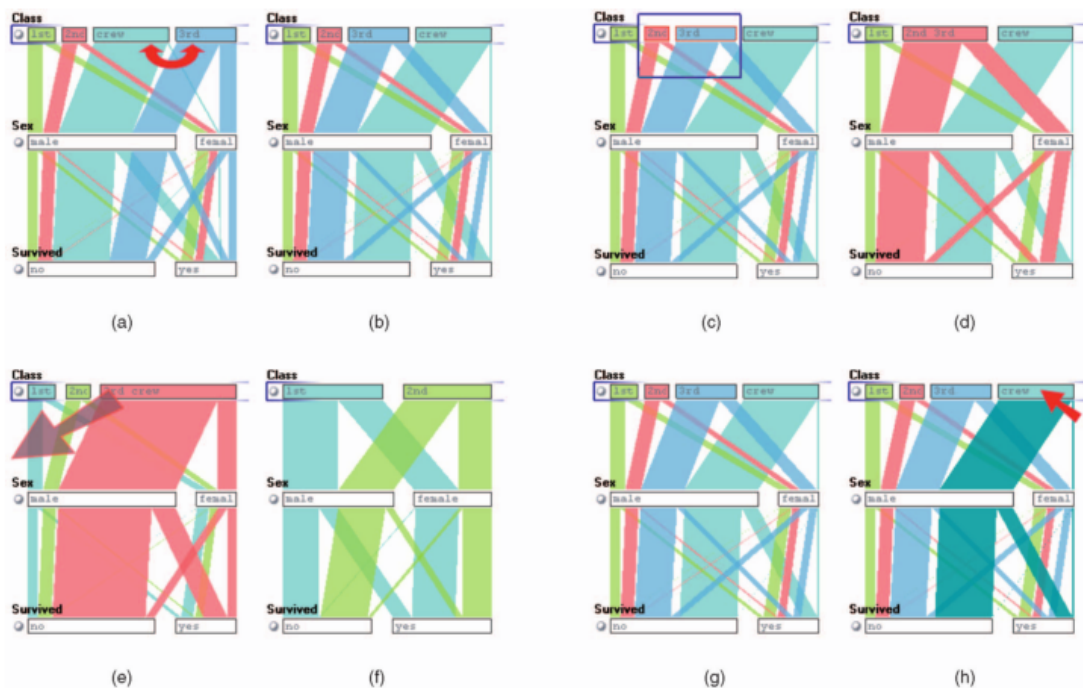


Figure 2: a 分类重排, b 重新布局 layout, c 合并分类, d 层次分析, e 去除分类, f 过滤, g 高亮, h 多维关系

用户选择感兴趣的数据 → 探索数据间的关系 → 选择数据创建新的 View(Dimension)

1.2 其他

对于不是分类属性的维度，比如说连续属性，需要将其分段划分。
坐标轴的顺序可能会影响数据的探索。

对于层次数据，每一层可以作为一个坐标轴来进行分类，每层的不同类型的节点作为不同的分类。

2 Ordered and Quantum Treemaps: Making Effective Use of 2D Space to Display Hierarchies[3]

文章提出了两种 treemap 算法，ordered-treemap 和 quantum-treemap。

- ordered-treemap 算法用来解决一般 treemap 算法中节点位置突变的问题。在普通的算法中，当数据发生细小的变化时往往造成 treemap 结构的很大变化。ordered-treemap 算法使得数据中相邻的节点在生成的 treemap 中依旧相邻（或不会相隔很远），所以当数据属性发生变化时，整个图结构不会发生很大的变化。
- quantum-treemap 用来显示 fixed-size objects，如图片。在显示图片时需要保证图片的 aspect-ratio, squarified-treemap 可以作为一个候选，但是它能够在一定程度上保证横纵比，但是无法保证 fixed-size object 的固横纵比。

3 Multitrees: Enriching and Reusing Hierarchical Structure[5]

文章引入了 multitree 这种结构来展示信息。Multitrees 是一类带有 unusual properties 的 DAG，把树作为其子结构增加了它的可视别性。在层次上下文交换或者在一个层次重用的模型中，这些子树具有自然的语义解释 (These subtrees have a natural semantic interpretation providing alternate hierarchical contexts for information, as well as providing a natural model for hierarchical reuse) Multitrees 中的子树提供基于树的图交互。

DAG 可以表示抽象的语义概念：比如父类子类 (class above subclass)、聚合部分 (aggregations above subparts)。DAG 的问题在于其结构很难布局并理解，即使很简单的邻居关系就可能造成非平面的布局，导致边的交错。

3.1 Motivation

- 寻找一种介于 DAG 和 tree 的结构，在展现信息上提供一种新的选择。
- reuse of hierarchical structure. 对于公司具有的文档树，每个职员可能需要自定义符合自己需求的结构，职员大多不需要这么一个大的聚合文档结构，而是需要在小视图中寻找自己需要的部分。这就需要一种能够把树的片段重构成新的树结构的途径。3

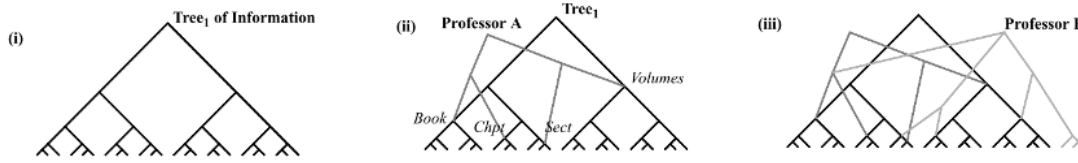


Figure 3: Example for hierarchy reuse

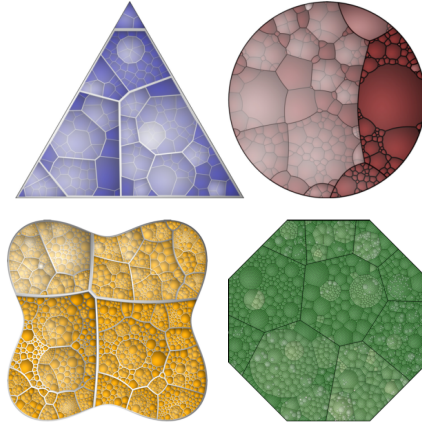


Figure 4: Voronoi Treemap Examples

4 Voronoi Treemaps[1]

该论文提出的方法将 treemap 构成的基本元素从矩形拓展到了任形状，从而消除了一般 treemap 采用矩形时出现的横纵比不均衡以及不能清楚的展示数据的层次结构等问题。所以 Voronoi Treemap 更具灵活性，能够在更多的应用中使用。层次结构经常用作数据的分类、排序、组织的一种抽象，被广泛应用在多种数据中。(Hierarchical structures are an often used abstraction for the classification, sorting, and organization of a broad variety of data.) Treemap 具有的 space-efficiency 特性，使得它被广泛应用在系谱图 (family tree)、文件系统的组织结构、软件架构、金融分析、体育报告等多种领域。

文章中的 Voronoi Treemap 的基础是 centroidal voronoid tessellation[4] 的使用。这个技术被广泛应用在能量最小化的各个领域，比如数据压缩，图像处理，网格细化 (mesh refinement)，资源分配以及科学可视化中。

5 Cheops: A Compact Explorer For Complex Hierarchies[2]

数据量的增多，加大了层次数据可视化的难度，一方面数据越来越复杂，另一方面数据的节点一层一层呈指数增长。一般的可视化技术能够有效的处理 1000 个几点左右的层次数据，一旦

数据量超过 5000, 就需要一些特殊的技术来弥补数据量过大造成的节点重叠等问题, 比如通过 cluster 来减少节点数目, 再通过交互方式显示被隐藏的节点。但是 cluster 技术会影响可视结果的稳定性。

文章提出了一个浏览大量复杂层次数据的可视技术, Cheops, 可以处理 1000 万到 10 亿的节点数目。它能够在保持节点上下文环境的同时浏览节点的细节信息。

传统的方法一般是在基本的表现形式上通过交互来加强信息的显示, 最常用的是 focus+context 和 fisheye。当这些方法用在比较复杂的数据上时会存在一些问题, 比如因为使用了 DOI (Degree-of-Interest) 函数使得周围的节点扭曲。当焦点改变的时候有可能需要重绘整个视图, 效率就会有所下降。

5.1 论文大意

Cheops 方法使用三角形作为可视的基本元素, 通过元素重叠的方式节省空间, 如图5。然后通过交互手段根据需要进行选择性的显示重叠的节点。在实际中节点并不是完全重叠。

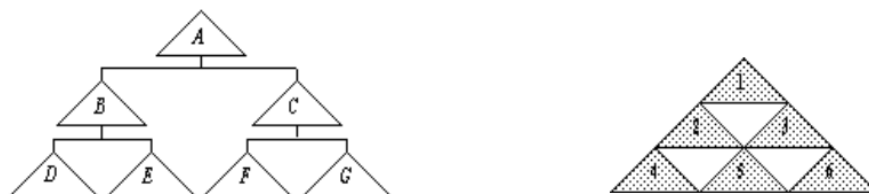


Figure 5: 从左图到右图的转换, 左图中的 EF 两个节点在右图中重叠为节点 5

为了能够在动态的上下文中浏览节点, 实现中使用了 *selection* 和 *pre-selection* 技术。

- selection 操作是说当用户选中一个节点 (logical node) 时, 高亮显示以整个节点为根的树, 其余的节点只显示一个轮廓如图6左。旁边面板显示每个层次中选中的点的信息。
- pre-selection 是说用户选中一个节点后, 随机的选择一个用户可能选择的子节点, 用户通过鼠标操作可以在子节点中进行切换, 如图6右。

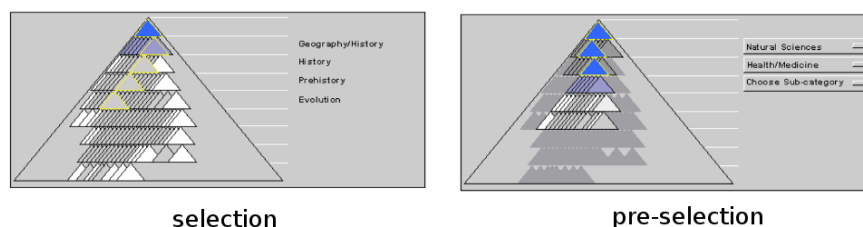


Figure 6: selection 和 pre-selection

这两种选择方式能够让用户在复杂的层次结构中快速的定位到节点。

另为除了在图上选择点以外，旁边面板提供的类似 comb select 的方式来选择节点，图的下方也有 navigation button 来帮助用户选择。

其他的交互 slider、bookmarks、histogram。bookmarks 允许用户保存某个视图，histogram 用户显示节点或者当前选中结构的信息。如图7

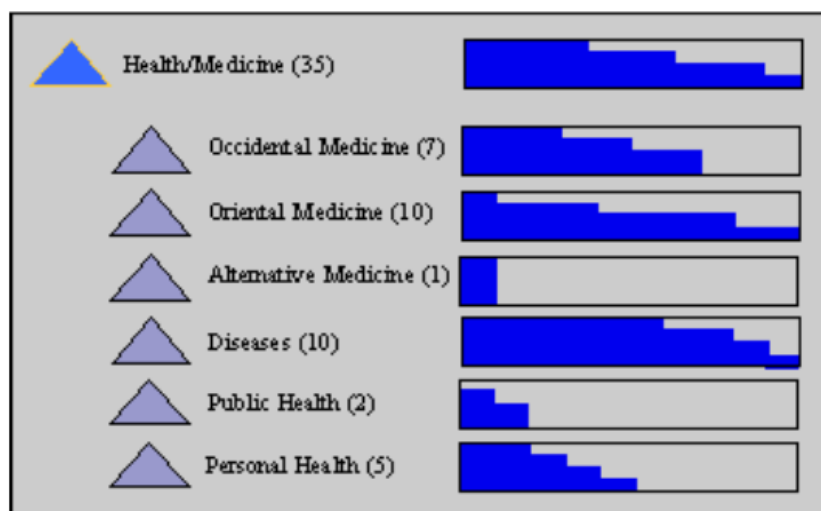


Figure 7: 图中可以看出，通过比较，disease 节点应该是用户关注的下一个节点

5.2 缺点

这个方法适合在大量数据中查找信息或者查看数据的结构特性。在对数据属性的显示上呈现的效果不是很丰富，所以不适合用于数据的比较或者查看数据属性之间的关系。

5.3 启发

1. 交互中，同一个功能可能需要多种方式实现来方便不同用户或者在不同的状态下进行选择。
2. 交互不是目的，交互是为了方便用户提取图中的信息。所以可以以一种显示 -> 交互 -> 信息提取 -> 显示 -> 交互 -> 信息提取的循环模式逐步让用户提取信息。

5.4 思考

overlap 的方式在什么情况下适用？缺点？优点？

References

- [1] Michael Balzer and Oliver Deussen. Voronoi Treemaps. In *Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization*, INFOVIS '05, pages 7—, Washington, DC, USA, 2005. IEEE Computer Society.
- [2] Luc Beaudoin, Marc-Antoine Parent, and Louis C Vroomen. Cheops: A Compact Explorer for Complex Hierarchies. In *Proceedings of the 7th Conference on Visualization '96*, VIS '96, pages 87—ff., Los Alamitos, CA, USA, 1996. IEEE Computer Society Press.
- [3] Benjamin B Bederson, Ben Shneiderman, and Martin Wattenberg. Ordered and Quantum Treemaps: Making Effective Use of 2D Space to Display Hierarchies. *ACM Trans. Graph.*, 21(4):833–854, October 2002.
- [4] Q. Du, V. Faber, and M. Gunzburger. Centroidal voronoi tessellations: Applications and algorithms. *SIAM Review*, 41(4):637–676, 1999.
- [5] George W Furnas and Jeff Zacks. Multitrees: Enriching and Reusing Hierarchical Structure. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '94, pages 330–336, New York, NY, USA, 1994. ACM.
- [6] Robert Kosara, Fabian Bendix, and Helwig Hauser. Parallel sets: Interactive exploration and visual analysis of categorical data. *Visualization and Computer Graphics, IEEE Transactions on*, 12(4):558–568, 2006.