

网络可视化研究综述

孙 扬¹ 蒋远翔² 赵 翔¹ 肖卫东¹

(国防科技大学信息系统与管理学院 长沙 410073)¹ (新华社湖南分社 长沙 410013)²

摘 要 传统的文字+表格的展现形式已无法满足规模日益庞大的网络数据的分析管理需求,网络可视化技术作为辅助用户理解网络结构并从中挖掘隐含信息的有效工具而得到广泛应用。首先归纳了网络可视化的任务,在回顾部分绘图美学相关工作后,基于网络可视化技术的分类体系,详细阐述了各类技术主要代表方法的基本原理和特征,随后概括了网络可视化扩展问题的解决策略,最后比较分析了各类技术的特点,展望了网络可视化未来的研究方向及其面临的挑战。

关键词 网络可视化,绘图美学,网络可视化任务,信息可视化

中图分类号 TP391 **文献标识码** A

Survey on the Research of Network Visualization

SUN Yang¹ JIANG Yuan-xiang² ZHAO Xiang¹ XIAO Wei-dong¹

(School of Information System and Management, National University of Defense Technology, Changsha 410073, China)¹

(Xinhua News Agency, Hunan Branch, Changsha 410013, China)²

Abstract The traditional way of depicting network data, text + data sheet, can not meet the requirement of analyzing and managing the increasingly-larger network dataset. As an effective tool of assisting users' understanding the network structure, as well as mining the connotative information, network visualization techniques are widely applied nowadays. In the beginning of this paper, we summarized the tasks of network visualization. After the review of graph drawing aesthetics related works, the rationales and characters of representative methods were expatiated on each category of a taxonomy framework of network visualization. Afterwards, the strategy of solution to the scalability of network visualization was recapitulated. And a comparison of techniques on each category was presented in the end, followed by an expectation of future work and challenges of network visualization faced.

Keywords Network (Graph) visualization, Drawing aesthetics, Network (Graph) visualization task, Information visualization

信息技术在不同领域的大规模应用使用户能够获取的网络结构数据越来越多,如社会关系网络、生物食物链网络、学术论文互引网络等^[1]。但是,随着网络节点和边的增加,仅用数据表格+文字展现网络结构的方式已经无法满足用户分析、管理、利用及处理网络结构数据的需求。如何高效浏览、理解、挖掘、操作、导航大量的网络结构数据,成为信息领域一个迫切需要解决的问题^[2]。

网络可视化技术作为一类重要的信息可视化技术,充分利用人类视觉感知系统,将网络数据以图形化方式展示出来,快速直观地解释及概览网络结构数据,一方面可以辅助用户认识网络的内部结构,另一方面有助于挖掘隐藏在网络内部的有价值信息^[3],因此得到了各国学者的高度重视,并被广泛应用于各类网络数据分析管理领域中,从 20 世纪 90 年代中期开始,在 Graph Drawing, InfoVis, IV 等重要国际会议中成为一个越来越受关注的议题^[4]。网络可视化与网络拓扑可视

化^[5]虽然共同使用了很多相似的技术方法,但二者并不完全相同,前者包含的内容更广泛,后者只是前者的子集。网络拓扑可视化大都应用于具有物理结构的网络,如因特网、电力网等,通过对网络拓扑结构的可视化展现可辅助管理人员对网络进行管理、分析和评估等。而网络可视化不仅可以实现上述功能,还可以作用于抽象的具有网络关联结构的数据集,如社会关系网络、论文互引网络等,通过对其进行可视化并设计友好的人机交互操作可以辅助用户充分发掘、分析网络结构数据中隐含的语义知识信息等。当前网络可视化的研究大部分集中于两个方面:一是制定美学标准,使绘制出的网络结构能够得到最佳感知,从而得到更好的理解;二是根据任务主题,建立不同的可视化系统,辅助用户高效直观地挖掘分析网络结构数据。其中,设计尽量符合美学标准的网络节点布局方法又是一个重点研究方向。

本文第 1 节首先给出了网络的相关定义,如节点、边、度

到稿日期:2009-03-13 返修日期:2009-05-24 本文受国家自然科学基金(60172012),湖南省自然科学基金重点项目(03JJY3110)资助。

孙 扬(1983-),男,博士生,主要研究方向为信息可视化、人机交互技术等,E-mail:victor_830514@yahoo.com.cn;蒋远翔(1964-),男,高级工程师,主要研究方向为信息系统开发及维护;赵 翔(1986-),男,硕士生,主要研究方向为信息可视化等;肖卫东(1968-),男,教授,主要研究方向为信息资源管理等。

及路径等,然后介绍了网络可视化的任务;第2节回顾了与绘图美学标准相关的部分工作;第3节归纳了不同的网络可视化技术;第4节概括了解决网络可视化扩展问题的策略;最后总结、比较了不同网络可视化方法的特点并展望了未来发展的方向。

1 相关定义及网络可视化的任务

1.1 网络的相关定义

网络是一种常见的用于描述诸对象(节点)及其相互间联系的数据结构模型。从数学层面讲,网络就是由节点和连线构成的图(G),因此在信息可视化相关文献中网络(network)和图(graph)的概念是等同的,可相互替代。网络中的对象表示为节点,两点间的关系表示为连接(也称为边)。节点一般表示为一个圆形、正方形或长方形,而边一般表示为连接两点的直线或曲线。网络可形式化表述为一有序偶 $\langle V, E \rangle$,即 $G = \langle V, E \rangle$,其中 V 是一非空集合, E 是 V 中元素的无序偶或有序偶的集合。 V 和 E 分别称为 G 的节点的集合和(无向或有向)边的集合,并用 $|V|$ 、 $|E|$ 表示网络节点的个数及边的条数。

在网络 $G = \langle V, E \rangle$ 中,如果节点 u 和节点 v 之间存在一条边,即若存在 $e \in E$ 且 $e = (u, v)$,则称节点 u 和节点 v 是邻接的, e 关联于节点 u 和节点 v ,并称 u 和 v 是边 e 的端点。称关联于同一节点的不同边为相邻的。与节点 u 相连接的边的条数称为 u 的度,表示为 $d(u)$ 。

对于网络 $G = \langle V, E \rangle$ 和网络 $G_s = \langle V_s, E_s \rangle$,若 $V_s \subseteq V$ 并且 $E_s \subseteq E$,则称 G_s 为 G 的子网络,并表示为 $G_s \subseteq G$ 。

在网络 $G = \langle V, E \rangle$ 中,把边的一个序列 (e_1, e_2, \dots, e_t) 称为网络的一条路径,其中 $e_i \in E$ 并且 $e_i = (u_i, u_{i+1})$ ($u_i \in V, i = 1, 2, \dots, t$)。

1.2 网络可视化的任务

网络可视化作为信息可视化的一个重要分支,涵盖了其涉及的所有常见任务,如检索值、筛选、计算派生值、查找极值、排序、确定属性值范围、刻画分布、发现、揭示关联、查找相邻节点、扫视浏览和集合操作等。马里兰大学的Plaisant根据用户执行网络分析理解任务的需求及目的,针对网络数据结构的特性,对前述的常见任务进行组合,将网络可视化的任务归纳为4大类20小项^[6]。

1) 基于拓扑(Topology)的任务

1.1) 邻接(直接连接)

- T1. 查找一个节点的所有邻接节点的集合;
- T2. 确定一个节点的邻接节点数量;
- T3. 确定哪个节点拥有最多的邻接节点;

1.2) 可达(直接或间接连接)

- T4. 查找一个节点的所有可达节点的集合;
- T5. 确定一个节点的可达节点数量;
- T6. 查找距离给定节点不大于 n 的可达节点集合;
- T7. 确定距离给定节点不大于 n 的可达节点数量;

1.3) 公共连接

T8. 在给定节点中,查找与其中所有节点均连接的节点集合;

1.4) 连通性

- T9. 查找两个节点之间的最短路径;

T10. 识别聚类;

T11. 识别连通分量;

T12. 查找桥;

T13. 查找连接点;

2) 基于属性(Attribute)的任务

2.1) 节点

T14. 查找具有特定属性值的节点;

T15. 查看给定节点集合;

2.2) 链接

T16. 给定一个节点,查找通过特定关联与之相连接的节点;

T17. 给定一个节点,在所有相连节点中查找具有属性极值的节点;

3) 浏览(Browsing)任务

3.1) 追寻路径

T18. 跟踪一条给定路径;

3.2) 二次访问

T19. 重新回溯到原先访问过的节点;

4) 概览(Overview)任务

T20. 快速获得关于网络的一些概要信息。

除上述常见任务以外,网络可视化过程中还可能遇到其他一些高级任务,如比较两个网络的异同点、网络随时间的变化情况。但这些任务并不常见,且它们都可进一步分解为前述若干个任务的组合和重复。因此,T1—T20较完整地定义了网络可视化过程中可能涉及的任务集合。

2 绘图美学标准

早期研究人员就认识到,即使对于小型网络,较差的网络布局也会增加用户理解网络结构的困难程度。如图1所示,虽然该网络的节点及边的数量都不多,但图中展示的网络结构用户难以理解。而将布局优化为图2时,网络结构则变得非常美观而且容易理解。

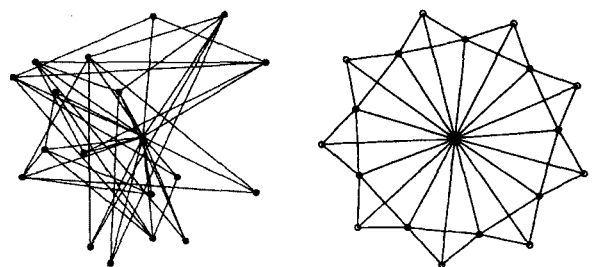


图1 不符合美学标准的网络图 图2 符合美学标准的网络图

为使用户可以最大程度发挥其视觉感知能力,从而能直观、深入地理解网络结构,研究人员提出了很多美学标准作为绘制网络结构的理论基础及原则。

绘图美学标准鼻祖 Sugiyama 在文献[7]中定义了5项美学标准,以改进网络图的易读性,包括:

- 边交叉数量最小原则。为能清晰展现网络结构,绘图时应尽量减少相互交叉边的数量。
- 邻接点空间位置接近原则。将相连接的节点尽量配置在相近的位置上,以减小边的长度。
- 直线边原则。网络中的边尽量使用直线,避免曲线。
- 边平衡布局原则。相同节点的多条边尽量以该节点为

中心平衡布局。

· 节点层次布局原则。引入层的概念,将节点尽量布局在水平或竖直的不同层上。

后来,Sugiyama 等又给出了规律性和可溯性的定义^[8]:

· 规律性是指在整个绘图过程中布局节点的参考原则应该始终保持一致。

· 可溯性是指网络边应该能够形成网络中的路径(尽量将边相连)。

Sindre 等也提出了一组美学标准^[9],其新增标准包括:

· 区域最小原则。绘制网络图应尽量节省屏幕空间。

· 高度节点居中原原则。高度节点应布局在绘图中心。

· 节点密度均匀原则。尽量使每一个单位区域的布局节点数量相等。

· 凸多边形原则。尽量使相接边形成的多边形为凸多边形。

Sindre 还提到了其他的标准,但这部分标准要么只是针对某一类网络图(如层次结构图中的垂直性原则),要么是 Sugiyama 标准的相似阐述(如最长边尽可能短原则就是邻接点空间位置接近原则的另一种说法),在此不予赘述。

为了验证可视化方法的有效性,Purchase 等也提出了 5 种美学标准^[10]:对称性、正交性、最小角最大化原则、边交叉数量最小原则和直线边原则。后两者前面已经给出了定义,前三者定义如下:

· 对称性。绘制网络时应将具有相同结构的子图围绕绘图中心平衡布局。

· 正交性。绘制网络时应以网格为背景,并尽量将节点布局在网格交叉处,使边可以沿着网格线绘制。

· 连接角最大化原则。使同一节点任意两条边形成的角度尽量大。

Purchase 在后续研究中发现最重要的美学标准是边交叉数量最小原则^[11]。但是,Ware 等在提出连通性原则之后(连通性原则是指尽量使相连边形成的网络路径保持为直线)指出,好的连通性甚至比边交叉数量最小原则还要重要,尤其是在用户需要确定网络中的最短路径时。他们研究发现,即使网络中存在少量的交叉边,如果能形成线形网络路径,依然可以增加用户对网络结构的理解力。而且,与减少整个网络的交叉边相比,减少网络路径交叉能得到更清晰的网络结构^[12]。

3 网络可视化技术

虽然绘图美学标准能够帮助用户理解网络结构,但是它只提供了网络布局的指导准则,没有涉及具体的网络(节点)布局技术,因此研究人员又根据不同领域的任务主题提出了很多种网络可视化方法。Visual Complexity^[13]比较全面地收集了现有的网络可视化应用,包括 600 多种,而且可以采用不同指标对其进行分类,如,任务主题、可视方法及开发时间等。为了能够鲜明对比各方法的优势与不足,本文选取 Ben Shneiderman 提出的分类标准,按照网络节点的布局方法将所有网络可视化技术分为 9 类,包括基于力导引布局(FDA, Force-Directed Algorithm)、基于地图布局(Geographical Map)、基于圆形布局(Circular)、基于相对空间布局(Spatial Calculated)、基于聚类布局(Cluster)、基于时间布局(Time-

oriented)、基于层布局(Substrate-based)、基于手工布局(Hand-made)和基于随机布局(Random)的网络可视化技术^[14]。

3.1 力导引布局

力导引布局方法能够产生相当优美的网络布局,并充分展现网络的整体结构及其自同构特征,所以在网络节点布局技术相关文献中该方法占据了主导地位^[15-27]。该方法最早由 Eades 于 1984 年提出^[15],其基本思想是将网络看成一个顶点为钢环、边为弹簧的物理系统,系统被赋予某个初始状态以后,弹簧弹力(引力和斥力)的作用会导致钢环移动,这种运动直到系统总能量减少到最小值时停止。但是,Eades 在他的实现算法中并没有遵守弹性力学中的胡克定律,而是采用了自己建立的弹簧受力公式。运算复杂度是 FDA 基础算法的主要缺点,虽然 Eades 假设引力只作用于相邻两节点,降低了算法的部分复杂度,但是每次循环都必须计算每对节点间的斥力关系,所以总复杂度仍然是 $O(N^2)$,因此 FDA 的大部分相关工作都在努力改善算法性能^[16,18-21]。

Kamada 和 Kawai 改进了 Eades 的弹簧模型,提出 KK 算法^[16],通过求系统总能量 $E = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{1}{2} k_{ij} (|p_i - p_j| - l_{ij})^2$ 的最小值来确定节点位置,其中 p_i 和 p_j 表示节点 v_i 和 v_j 的位置, l_{ij} 表示 v_i 和 v_j 间弹簧的初始长度, k_{ij} 表示弹性系数。不同于 Eades 的做法,Kamada 和 Kawai 在模型中遵循了胡克定律的偏微分方程,并以此来优化顶点的布局。此外,KK 算法还加入了非相邻节点间理想距离的概念:两个节点间的理想距离与它们之间的最短路径的长度成正比。在系统的最终稳定状态下,节点间的距离都将接近于它们的理想距离。虽然 KK 算法的时间复杂度在理论上仍然为 $O(N^2)$,但是其目标函数的收敛性得到了很大的提高。

Davidson 和 Harel 提出的 DH 算法根据顶点分布、边界靠近、边长及边交叉等约束来建立系统的能量函数。调整约束的权重能量函数可以满足不同的美学标准,而且他们使用模拟退火算法(SA, Simulated Annealing)来求解能量函数的最小值^[17]。SA 算法解得的结果通常十分接近全局能量最小点,并且他们在算法的最后阶段加入针对边交叉及边点交叠的两个罚函数,使得该算法与其他算法相比能够产生效果更好的节点布局,如图 3 所示。但是,由于 SA 算法本身固有复杂性及最优参数选取问题的存在,与 FDA 基础算法相比,该算法甚至更加耗时。

Fruchterman 和 Reingold 基于再次改进的弹性模型提出了 FR 算法^[18]。该算法遵循两个简单的原则:有边连接的节点应该互相靠近;节点间不能离得太近。FR 算法建立在粒子物理理论的基础上,将无向图中的节点模拟成原子,通过模拟原子间的力场来计算节点间的位置关系。算法通过考虑原子间引力和斥力的互相作用,计算得到节点的速度和加速度。依照类似原子或者行星的运动规律,系统最终进入一种动态平衡状态。此外,FR 算法还采用了网格变量(grid-variant)方法进行优化:将布点区域分成若干网格,计算斥力时只考虑节点与相邻网格内的节点间的作用。若以 k 表示节点周围空白区域的理想半径, d 表示节点间的距离,则斥力计算公式为

$$f_r = \frac{k^2}{d} u(2k - d)$$

其中, $u(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$ 。

FR 算法产生的网络布局好像多面体在平面上的投影,如图 4 所示。而且它具有高效、健壮的优点,因此到目前为止仍然是较流行的网络节点布局算法之一。

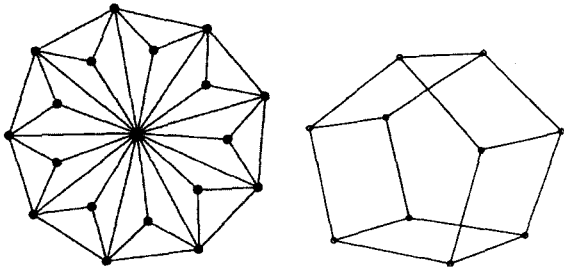


图 3 DH 算法节点布局样图 图 4 FR 算法节点布局样图

为进一步提高 FR 算法的计算效率,各国学者又提出了 GEM(Graph Embedder) 算法^[19]、多尺度(Multi-Scale)算法^[20]以及高维嵌入(High-Dimensional Embedding)算法^[21],它们虽然降低了时间复杂度,但也付出了相当代价。如 GEM 算法的健壮性不及 FR 算法,多尺度算法存在网络直径退化问题,高维嵌入算法对部分真实网络的节点存在交叠等。

国内也有学者在该领域做出贡献。如黄竞伟等人^[22]根据不同的美学标准将一般无向图的画图问题转化为函数优化问题,用遗传算法求目标函数最优解的近似值,从而得到无向图自动画图算法的一个一般框架。该方法的优点在于:算法统一,容易实现,便于修改,并且易于并行化,可以直接用来绘制非连通图。张清国等人^[23]提出了一种基于 D. M. P. 平面性判定算法的新的平面图画图算法。与其它算法相比,该算法具有方法简单、易于实现的优点。

力导引布局方法产生的网络布局优美,因此被应用于大量的网络可视化系统中,如 Vizster^[24], Dig-CoLa^[25], IPSep-CoLa^[26] 及 SocialAction^[27] 等。优秀的人际网络可视工具 Vizster 使用 Friendster 社区^[28] 的人际关系数据,将每一个“朋友”抽象为以其照片为代表的网络节点,将“朋友”间的关系抽象为边,使用力导引方法布局网络节点,并对社区中的人际关系进行可视化分析,发掘其中隐含的“社区行为”,如图 5 所示。

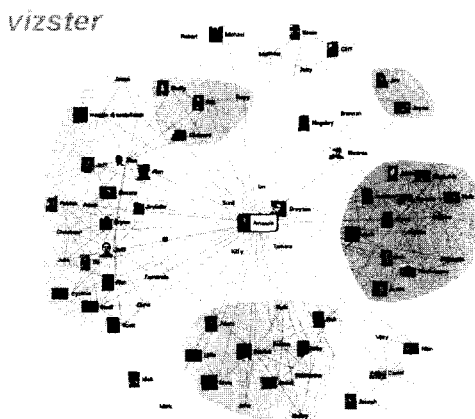


图 5 基于力导引布局的应用

3.2 地图布局

该类方法能够产生用户极易理解的网络布局,它以一幅

世界(大洲、国家、省或市)地图作为背景,根据网络节点的地理坐标将其布局在背景图上,然后根据节点间的连接关系绘制网络边,如图 6 所示,SeeNet 以网络节点表示不同的城市,边表示城市间 Internet 的网络流量,辅助用户深入理解各城市间的网络数据通信量^[29]。地图布局方法的优点是符合人类视觉思维,地图背景可以直观展现网络节点的地理分布,而且能够准确定位网络节点;缺点在于网络节点在背景图上位置固定,节点交叠、边交叉问题严重,不利于可视化节点、边数量较多的网络结构。



图 6 基于地图布局的应用

3.3 圆形布局

该方法在圆心放置一个或一组节点,在同心圆周上按顺序布局其余节点,它能利用通过圆心的十字线产生优良的布局^[30]。该方法需要明确各节点的顺序,指定优先度最高的节点为根节点并将其放置在圆心,与其相关联的节点排列在最近的同心圆周上,逐步迭代直到将所有节点排列完毕;也可按照优先度顺序将不同节点排列在同一圆周上。如图 7 所示, Social Circles 将电邮用户抽象为节点,用户间通信抽象为边,节点尺寸表示用户的通信频率,越靠近中心的用户与其他用户通信越频繁,也就是社会网络中的权威性人物^[31]。圆形布局方法可扩展为放射布局,根据网络节点的排列顺序,将优先度最高的节点放置在网络中心,其余节点按照关联程度沿放射线依次排布^[32],如图 8 所示。圆形布局能够体现网络节点间的层次关系,布局清晰,中心突出,但要求节点存在一定顺序,通用性不强,而且在突出网络层次结构特征的同时,降低了对其他网络拓扑结构特征的表现力度。

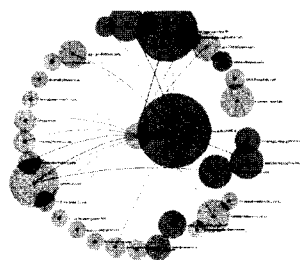


图 7 基于圆形布局的应用

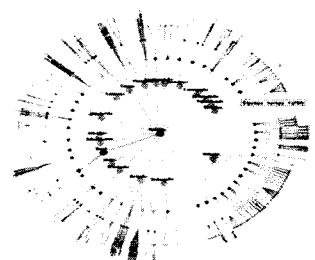


图 8 基于放射布局的应用

3.4 相对空间布局

该方法以“参照体(related entity)”的空间位置为基础,根据网络各节点与“参照体”的关系计算其坐标。“参照体”是以某种方式与网络节点相关联的对象。“参照体”可以是网络节点的一个子集,也可以是散布在网络周围或网络中除节点外的其他对象,如,在 Making Visible the Invisible 中,节点表示

一类事物，“参照体”为节点属性-杜威分类号，它以水平条带的形式排列在网络图的顶部和底部，每一节点都夹在它所属的分类数中间^[33]，如图 9 所示。“参照体”也可以由用户交互确定，这样产生的网络布局就会随用户每次选择不同的“参照体”而改变，如，在 Non-geographic Mapping 中，节点代表城市（未按地理坐标进行布局），边表示城市间有交通相连，用户通过点击选择某节点作为“参照体”，某城市被点选后，其余城市会根据自己与所选城市的关系进行布局，城市间的距离由航班飞行时间决定，地理方位决定了城市间的角度，用户通过移动鼠标获取不同城市与“参照体”间的交通方式^[34]，如图 10 所示。相对空间布局方法主要展现网络节点与“参照体”间的关系，也可以间接体现网络各节点间的关系，且“参照体”选取的质量决定了网络图的效果。



图 9 基于相对空间布局的应用 I

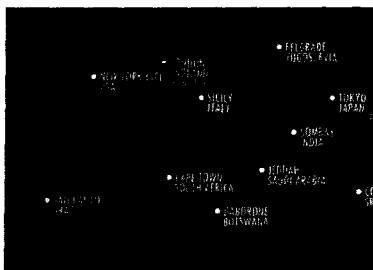


图 10 基于相对空间布局的应用 II

3.5 聚类布局

该方法根据节点的属性及相互间的连接关系，通过人机交互或应用算法（如 MDS、自组织网络（SOM）和 Sammon 映射^[35]等）来聚类分组网络节点。网络节点可以分布在同一平面内（单层聚类），如 Vizster 中不同颜色区域的节点；也可分布在不同平面上（多层聚类），如 Eades 等提出的 MVCG（Multilevel Visualization of Clustered Graphs）方法，根据节点间的连接结构生成聚类，进而抽取网络的层次结构，将同层节点排列在同一平面上。即叶节点分布在最底层，叶节点的聚类节点在上一层，聚类的聚类节点在再上层^[36]，如图 11 所示。聚类布局有助于用户发现网络结构中存在的关系信息、模式信息及聚类信息等隐性知识，一般多与其他布局方法结合使用。

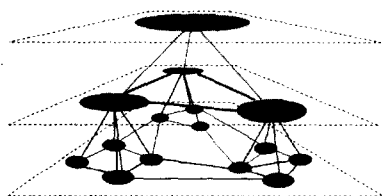


图 11 基于多层聚类布局的应用

3.6 时间布局

该方法的基本原理是根据节点的时间属性对其进行排布，其典型布局是将历史节点排列在屏幕顶（左）端，当前节点摆放在屏幕底（右）端，同一时间的节点放在同一排（列）。如该方法用于文献引用网络时，就形成右（底）部近期文献向左（上）引用右（顶）部历史文献的网络结构^[37]，如图 12 所示，图中纵轴表示文献被引次数，横轴表示文献出版年度，节点代表不同文献，边代表文献间的引用关系。时间布局能够清晰反映网络节点随时间的演化规律，但对于除时间属性外的其他网络结构特性表现力不足，容易产生节点交叠问题。

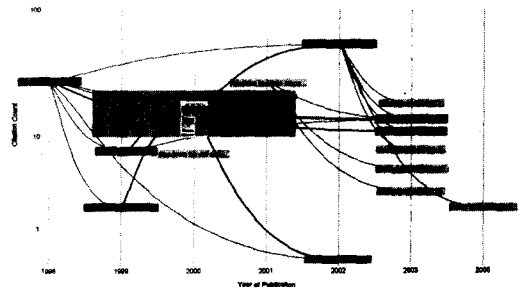


图 12 基于时间布局的应用

3.7 层布局

该方法首先根据节点的分类属性将屏幕划分为几个区域，然后在对应的区域中布局网络节点。Aleks Aris 设计了 NVSS（Network Visualization by Semantic Substrates），其利用该方法可视化诉讼案件网络（英美法系审判诉讼案件需引用历史案例作为依据），节点表示诉讼案件，节点间的边表示案件之间的引用关系。首先按照案件的审理机构（地区法院、中院、高院）将屏幕分层，然后在每层中将案件节点按时间进行排列，最后标注案件之间的引用关系，如图 13 所示。NVSS 直观揭示了诉讼案件引用网络的结构特点及隐含信息，如，地区法院审判案件的引用寿命较短等^[14]。层布局方法有效利用了节点的属性信息，增加了网络图所包含的信息量，有助于用户发现部分趋势及关系信息。但是分层属性选取恰当与否很大程度上决定了网络可视化质量，而且对网络自身结构特征反映不明显。

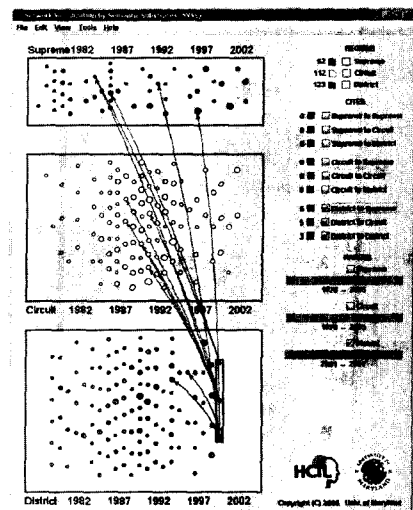


图 13 基于层布局的网络可视化应用

除了上面介绍的方法，现有的节点布局技术还包括手工

布局技术^[38] (包括人工绘制的网络结构,或根据预先计算产生的输入布局网络节点的方法)和随机布局技术,由于其方法较为简单,在此不再赘述。事实上,为更好地反映网络结构特征,使不同布局方法能够取长补短,互为补充,大部分网络可视化应用都结合使用了多种节点布局方法。如 Vizster 及 SocialAction 就同时使用了力引导布局及聚类布局方法, NVSS 在每一层内又使用时间布局方法等。针对不同的网络可视分析任务,如何有效组合各类节点布局方法也是当前网络可视化技术研究的一个热点问题。

4 扩展性

随着大型、超大型及复杂网络的出现,传统的网络可视化技术需解决扩展性问题,以适应包含海量节点及边的网络结构,研究人员提出了用聚类技术、focus+context 技术、边路由聚合技术以及矩阵技术减低或处理网络复杂性的策略。

聚类技术^[39]经常与多尺度(multiscale)可视化方法^[40,41]一起使用,该类方法能够处理超大型网络,大量网络节点可聚类为“元节点”,节点间的连接也可合并为“元连接”,形成保持大型网络拓扑结构的概要视图,“元节点”的尺寸和“元连接”的粗度分别代表了包含节点和边的数量;当用户需要网络节点的详细信息时,根据兴趣程度关注重点区域,选择合适的缩放尺度,展开相应的“元节点”及“元连接”,从而实现大型网络结构的导航。

Focus + Context 技术概括了所有允许用户浏览细节时不丢失背景的交互技术,其基本理念是辅助用户详细地查看关注信息,同时保持对周围背景信息的概览。该技术主要实现形式包括^[42] (1)基于变形的方法(Distortion-oriented):如鱼眼视图^[43]通过模拟鱼眼观察的效果,实现了用户焦点处信息的放大显示,同时挤压焦点周围的信息显示空间,背景信息缩小显示以补偿因焦点区域放大造成的空间的损失;(2)概览的方法(Overview methods):基于 Overview+Detail 的思想在两个窗口分别显示焦点细节和背景信息,背景信息被压缩成一行的像素从而能显示更多信息^[44];(3)筛选的方法(Filtering):在需显示的画面之“上”放置一个类似于透镜的物件,当它通过屏幕某一区域时,用户可透过该物件观察其覆盖区域下经筛选的细节^[45]。

Flow Map 提供了边路由聚合的典型应用,它通过合并经过公共路由的连接来减少网络结构的复杂性^[46]。

基于矩阵的网络结构表示方法避免了网络图(尤其是大型网络)的部分问题,如,节点交叠、边交叉、由节点和边位置固定引起的边在节点底部穿越的问题等。但是该方法难以直观展现网络结构的特性,如查找连接路径上的节点、确认

聚类。因此,研究人员结合节点连接图与矩阵表示法的各自优点提出一种混合网络可视化方法,用以辅助用户理解较为复杂的网络结构^[47,48]。

上面只列举了几种典型的解决网络可视化技术扩展问题的策略。随着复杂网络理论研究的逐步深入,复杂网络可视化技术也得到了研究人员的高度重视^[4]。

结束语 通过对网络信息进行合理、美观的可视化,为抽象信息提供形象生动的描述,网络可视化技术不仅改善了用户对网络结构数据及数据项之间关系的理解,同时能够辅助用户直接操纵信息以增强认知,实现对信息的洞察,避免在海量信息中迷航(disorientation),使其在尽可能短的时间内,以最自然的方式获得和理解网络数据集中隐藏的特征信息及结构信息,从而进行快速有效的决策。本文根据信息可视化大师 Ben Shneiderman 提出的网络可视化技术分类体系,详细分析了目前网络可视化技术的研究状况,介绍了各类技术的代表方法,对每项技术方法的基本原理、适用性及功能作用进行了阐述。下面,将分别从这些技术方法的适用网络规模、节点交叠可能性、能够表达的网络结构特征、对原始数据组织形式的要求、计算复杂度、主要特点、可结合使用的方法及代表应用等方面,对这些技术方法进行比较,如表 1 所列(一般认为,节点数在 1,000~10,000 为大型网络,100~1000 为中型网络,100 以下时为小型网络)。

总的来说,在我们这个信息日益丰富的社会,网络可视化技术的研究和应用已经从根本上改变了大型、复杂、抽象的网络新型的表示和理解方式,并且已经成为引导我们获得新见解和做出快速、准确、有效决策的决定性因素。网络可视化技术一直是信息可视化领域的研究前沿,国外在该领域研究比较活跃,新的研究成果不断出现;国内目前在力引导节点布局算法方面也进行了部分研究,但总体仍处于起步阶段。就网络可视化技术的研究而言,尚存在很多有待解决的开放问题,如网络可视化扩展问题,即如何对现有的网络可视化方法进行拓展改进,使其能够适用于复杂网络数据集;网络可视化方法组合问题,即如何有效组合各类网络可视化方法,互为补充,以更有效地反映隐性的网络结构及特征信息;网络节点动态布局问题,即如何在网络结构改变时(节点增加(减少)、边增加(减少)或聚类扩展(收缩)),保持网络布局的稳定性,维持用户思维图(mental map)的连续性^[49];通用性问题,即如何针对不同领域的任务需求设计通用网络可视化模型及方法,以改变现有方法任务针对性强的缺点;人机交互问题,即如何设计更加直观易用的人机操作界面,以进一步改善用户体验;评估问题,即如何准确评估各类网络可视化模型及方法的有效性等。上述这些问题仍然需要进一步研究。

表 1 网络可视化技术比较

网络可视化技术	适用网络规模	节点交叠可能性	反映的网络结构特征	数据要求	计算复杂度	主要特点	可结合方法	代表应用
力引导布局 (fd)	基础算法 KK 算法 DH 算法 FR 算法	中大型 小	充分展现网络的整体结构及其自同构特征	无要求	高 较高 高 中	布局优美,通用性较强	kx, tx, sb	Vizster, SocialAction
地图布局 (gm)		小型 非常大	网络的地理分布特征	节点需包含地理属性	低	节点位置固定,通用性弱	sx, tx	SeeNet, GeoTime ^[50]

圆形布局 (cx)	同心圆或 放射布局	大型	很小	网络中心, 节点 间层次关系	节点间 存在顺序	低	布局清晰, 中心 突出, 通用性弱	sx, kx, tx, sb	Radial Tree
	单圆	中型	比较小			低			Social Circles
	相对空间(sx)	中小型	适中	节点与“参照 体”间的关系	无要求	较低	“参照体”选取 较为重要	gm, cx	Making Visible the Invisible
聚类布局 (kx)	单层	大型	适中	系、模式及 聚类信息	存在适于聚 类的属性	较低	多与其他方法结 合, 多层聚类视 角选取重要	fd, cx, sx, tx, sb	Vizster
	多层	中型	比较大						MVCG
	时间布局(tx)	小型	非常大	节点随时间 的演化规律	节点需拥有 时间属性	低	多与其他方 法结合使用	fd, gm, cx, kx, sb	GeoTime
	层布局(sb)	中型	适中	节点在不同层 间的联系	节点分层属 性为离散值	低	有效利用了节点 的属性信息	fd, cx, kx, tx	NVSS

参考文献

- [1] Palla G, Derényi I, Farkas I, et al. Uncovering the overlapping community structure of complex networks in nature and society [J]. *Nature*, 2005, 435(7043): 814-818
- [2] Pak Chung Wong, Mackey P, Perrinel K, et al. Dynamic visualization of graphs with extended labels[C]//*Proceedings of InfoVis*. 2005: 73-80
- [3] Ahmed A, Dwyer T, Hong Seok-Hee, et al. Visualization and analysis of large and complex scale-free networks[C]//*Proceedings of Eurographics*. 2005: 1-8
- [4] 王柏, 吴巍, 徐超群, 等. 复杂网络可视化研究综述[J]. *计算机科学*, 2007, 34(4): 17-23
- [5] 张伟明, 罗军勇, 王清贤. 网络拓扑可视化研究综述[J]. *计算机应用研究*, 2008, 25(6): 1606-1610
- [6] Lee Bongshin, Plaisant C, et al. Task taxonomy for graph visualization[C]//*Proceedings of the AVI Workshop on Beyond Time and Errors: novel evaluation methods for information visualization*. 2006: 1-5
- [7] Sugiyama K, Tagawa S, Toda M. Methods for visual understanding of hierarchical system structures[J]. *IEEE Transactions on Systems, Man and Cybernetics*, 1981, 11(2): 109-125
- [8] Sugiyama K. A cognitive approach for graph drawing[J]. *Cybernetics and Systems*, 1987, 18(6): 447-488
- [9] Sindre G, Gulla B, Jokstad H. Onion graphs: aesthetic and layout[C]//*Proceedings of the IEEE Symposium on Visual Languages*. 1993: 287-291
- [10] Purchase H, Cohen R, James M. Validating graph drawing aesthetics [C]//*Proceedings of Graph Drawing*. 1996: 435-446
- [11] Purchase H. Which aesthetic has the greatest effect on human understanding? [C]//*Proceedings of Graph Drawing*. 1997: 248-261
- [12] Ware C, Purchase H, Colpoys L, et al. Cognitive measurements of graph aesthetics[J]. *Information Visualization*, 2002, 1(2): 103-110
- [13] <http://www.visualcomplexity.com/vc/index.cfm>
- [14] Shneiderman B, Aris A. Network visualization by semantic substrates [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2006, 12(5): 733-740
- [15] Eades P. A heuristic for graph drawing[J]. *Congressus Numerantium*, 1984, 42: 149-160
- [16] Kamada T, Kawai S. An algorithm for drawing general undirected graphs[J]. *Information Processing Letters*, 1989, 31: 7-15
- [17] Davidson R, Harel D. Drawing graphs nicely using simulated annealing[J]. *ACM Transactions on Graphics*, 1996, 15(4): 301-331
- [18] Fruchterman T, Reingold E. Graph drawing by force - directed placement[J]. *Software-Practice and Experience*, 1991, 21(11): 1129-1164
- [19] Frick A, Ludwig A, Mehldau H. A fast adaptive layout algorithm for undirected graphs[C]//*Proceedings of the DIMACS International Workshop on Graph Drawing*. 1994: 388-403
- [20] Harel D, Koren Y. A fast multi-scale method for drawing large graphs[C]//*Proceedings of the working conference on Advanced Visual Interfaces*. 2000: 282-285
- [21] Harel D, Koren Y. Graph drawing by high-dimensional embedding[C]//*Proceedings of Graph Drawing*. 2002: 207-219
- [22] 黄竞伟, 康立山, 陈毓屏. 一个新的无向图图画算法[J]. *软件学报*, 2000, 11(1): 138-142
- [23] 张清国, 黄竞伟. 一个无向平面图图画算法[J]. *小型微型计算机系统*, 2003, 24(6): 972-975
- [24] Heer J, Boyd D. Vizster: Visualizing online social networks[C]//*Proceedings of InfoVis*. 2005: 33-40
- [25] Dwyer T, Koren Y. Dig-CoLa: Directed graph layout through constrained energy minimization[C]//*Proceedings of InfoVis*. 2005: 65-72
- [26] Dwyer T, Koren Y, Marriott K. IPSep - CoLa: An incremental procedure for separation constraint layout of graphs[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2006, 12(5): 821-828
- [27] Perer A, Shneiderman B. Balancing systematic and flexible exploration of social networks[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2006, 12(5): 693-700
- [28] <http://www.friendster.com/>
- [29] Becker R, Eick S, Wilks A. Visualizing network data [J]. *IEEE Transactions on Visualization and Computer Graphics*, 1995, 1(1): 16-21
- [30] Breitkreutz B-J, Stark C, Tyers M. Osprey: a network visualization system[J]. *Genome Biology*, 2003, 4(3): R22
- [31] <http://www.marumushi.com/apps/socialcircles/>
- [32] Di Battista G, Eades P, Tamassia R, et al. Graph drawing: algorithms for the visualization of graphs[M]. Upper Saddle River: Prentice Hall, 1998
- [33] <http://www.mat.ucsb.edu/g.legrady/glWeb/Projects/spl/spl.html>
- [34] <http://www.number27.org/projects/maps/traveltime/>
- [35] Martin-Merino M, Munoz A. A new sammon algorithm for sparse data visualization[C]//*Proceedings of Pattern Recognition*. 2004: 477-481
- [36] Eades P, Feng Qingwen. Multilevel visualization of clustered graphs[C]//*Proceedings of Graph Drawing*. 1996: 101-112
- [37] Cammarano M, Dong Xin (Luna), Chan B, et al. Visualization of heterogeneous data[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2007, 13(6): 1200-1207

(下转第 30 页)

mission control in satellite networks[J]. *IEEE Journal on Selected Areas in Communications*, 1999, 17(2): 360-372

- [25] Inzerilli T, Paone E, Pietrabissa A, et al. QoS support for interactive communication with DVB-RCS satellites[C]// *Computers and Communications, Proceedings (ISCC)*, 2004, 2: 897-902
- [26] Acar G, Rosenberg C. Performance study of end-to-end resource management in ATM geostationary satellite networks[C]// *Military Communications Conference*, 2001, 2: 764-769
- [27] Omari A, Seco-Granados G, Vazquez-Castro M A, et al. Analysis of the Efficiency and Delay of Bandwidth Request Algorithms in DVB-RCS[C]// *Satellite and Space Communications International Workshop on*, Sept. 2006: 165-169
- [28] Chisci L, Fantacci R, Pecorella T. Predictive Bandwidth Control for GEO Satellite Networks [C] // *Proc. IEEE International Conf. Communication (ICC2004)*, 2004
- [29] Mitchell P D, Grace D, Tozer T C. Burst targeted demand assignment multiple-access for broadband internet service delivery over geostationary satellite[J]. *IEEE Journal on Selected Areas in Communications*, 2004, 22(3): 546-558
- [30] Chiti F, Fantacci R, Marangoni F. Advanced dynamic resource allocation schemes for satellite systems[C]// *Communications, ICC 2005*, 2005, 3: 1469-1472
- [31] Koutsakis P. Predictive Resource Reservation Access for Multimedia Integration over GEO Satellite Networks[C]// *Proceedings of the Third International Conference on Wireless and Mobile Communications (ICWMC'07)*, 2007
- [32] Omari A, Seco-Granados G, Vazquez-Castro M A, et al. Analysis of the Efficiency and Delay of Bandwidth Request Algorithms in DVB-RCS[C]// *Satellite and Space Communications International Workshop on*, Sept. 2006: 165-169
- [33] Chisci L, Fantacci R, Pecorella T. Dynamic bandwidth allocation in GEO satellite networks: A predictive control approach [J]. *Control Engineering Practice*, 14(9): 1057-1067
- [34] Priscoli F D, Pietrabissa A. Resource management for ATM-based geostationary satellite networks with on-board processing [J]. *Computer Networks*, 2002, 39(1): 43-60
- [35] Priscoli F D, Pietrabissa A. Design of a bandwidth-on-demand (BoD) protocol for satellite networks modelled as time-delay systems[J]. *Automatica*, 2004(40): 729-741
- [36] Pietrabissa A. A multi-model reference control approach for bandwidth-on-demand protocols in satellite networks[J]. *Control Engineering Practice*, 2008(16): 847-860
- [37] Petraki D K, Anastasopoulos M P, Panagopoulos A D, et al. Dynamic Resource Calculation Algorithm in MF-TDMA Satellite Networks[C]// *Mobile and Wireless Communications Summit*, 16th IST, July 2007: 1-5
- [38] Park Jung-Min, Savagaonkar U, Chong E K P, et al. Allocation of QoS connections in MF-TDMA satellite systems: a two-phase approach [J]. *IEEE Transactions on Vehicular Technology*, 2005, 54(1): 177-190
- [39] Panagopoulos A D, Kanellopoulos J D. On the rain attenuation dynamics: spatial-temporal analysis of rainfall rate and fade duration statistics[J]. *Int. J. Sat Commun. and Netw.*, 2003: 595-611
- [40] Yi Qian, Hu R, Abu-Amara H, et al. Performance evaluations on a bandwidth on demand algorithm for a high capacity multimedia satellite network[C]// *Communications, ICC 2000*, 2000, 1: 139-143
- [41] Chai Wei Koong, Karaliopoulos M, Pavlou G. Scheduling for proportional differentiated service provision in geostationary bandwidth on demand satellite networks[C]// *IEEE Globecom*, 2005: 3722-3727
- [42] Lee Ki-Dong, Chang Kun-Nyeong. A real-time algorithm for timeslot assignment in multirate return channels of interactive satellite multimedia networks[J]. *IEEE Journal on Selected Areas in Communications*, 2004, 22(3): 518-528
- [43] Jiang Z, Leung V C M. A Predictive demand assignment multiple access protocol for Internet access for broadband satellite networks[J]. *International Journal of Satellite Communications and Networking*, 2003, 21(4/5): 451-467
- [44] Lee Ki-Dong, Cho Y-H, Lee H J, et al. Optimal scheduling for timeslot assignment in MF-TDMA broadband satellite communications[C]// *Proc. IEEE VTC Fall, Vancouver, Canada, 2002*: 1560-1564
- [45] Qin Yong, Zhang Jun, Zhang Tao, et al. A practical method for improving delay performance of timeslot allocation in DVB-RCS [J]. *Chinese Journal of Aeronautics*, 2009(22): 174-183
- [46] IEEE802.16 IEEE Standards for Local and Metropolitan Area Networks. Part 16: Air Interface for Fixed Broadband Wireless Access Systems[S], 2004

(上接第 18 页)

- [38] Nardi B, Whittaker S, Isaacs E, et al. Integrating communication and information through contactMap[J]. *Communications of the ACM*, 2002, 45(4): 89-95
- [39] Wattenberg M. Visual exploration of multivariate graphs[C]// *Proceedings of CHI*, 2006: 811-819
- [40] Auber D, Chiricota Y, Jourdan F, et al. Multiscale visualization of small world networks[C]// *Proceedings of InfoVis*, 2003: 75-81
- [41] Archambault D, Munzner T, Auber D. Grouse: Feature-based, steerable graph hierarchy exploration[C]// *Proceedings of Eurographics*, 2007: 67-74
- [42] Herman I, Melançon G, Marshall S. Graph visualization and navigation in information visualization: a survey [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2000, 6(1): 24-43
- [43] Sarkar M, Brown M. Graphical fisheye views [J]. *Communications of the ACM*, 1994, 37: 73-83
- [44] Ball T, Eick S. Software visualization in the large [J]. *IEEE Computer*, 1996, 29(4): 33-43
- [45] Bier E A, Stone M C, Pier K, et al. Toolglass and magic lenses: The see-through interface[C]// *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, 1993: 73-80
- [46] Phan D, Xiao Ling, Yeh R, et al. Flow map layout [C]// *Proceedings of InfoVis*, 2005: 219-224
- [47] Henry N, Fekete J-D. MatrixExplorer: a dual-representation system to explore social networks[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2006, 12(5): 677-684
- [48] Henry N, Fekete J-D, McGuffin M. NodeTrix: A hybrid visualization of social networks[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2007, 13(6): 1302-1309
- [49] Frishman Y, Tal A. Online dynamic graph drawing[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2008, 14(4): 727-740
- [50] Kapler T, Wright W. GeoTime information visualization[C]// *Proceedings of InfoVis*, 2004: 25-32