

网络边布局技术研究综述^{*}

邹玲¹, 胡华全^{2,3}, 杨超², 宋汉辰³

¹(虚拟现实技术与系统国家重点实验室(北京航空航天大学), 北京 100191)

²(装备学院 重点实验室, 北京 101416)

³(国防科学技术大学 信息系统工程重点实验室, 湖南 长沙 410073)

通讯作者: 胡华全, E-mail: hhq1986@foxmail.com, http://www.nudt.edu.cn

摘要: 网络边布局技术是网络可视化研究的重要内容, 良好的边布局能够有效地展示网络的整体结构并从中揭示出某些隐含的信息模式. 为了全面地总结和分析现有网络边布局技术, 首先归纳了网络边布局的主要任务, 回顾了指导网络边布局的美学标准. 然后将网络边布局技术归纳为 3 类: 边路由技术、边融合技术和边集束技术, 分别阐述了各类技术中典型布局方法的主要原理和特征, 并重点对边集束技术进行了分类分析. 最后总结了目前研究中还存在的一些问题, 展望了网络边布局技术的发展前景和面临的挑战, 以期对相关领域的研究者提供有益参考.

关键词: 边布局; 边路由; 边融合; 边集束; 网络可视化

中文引用格式: 邹玲, 胡华全, 杨超, 宋汉辰. 网络边布局技术研究综述. 软件学报, 2014, 25(Suppl.(2)): 1-11. <http://www.jos.org.cn/1000-9825/14017.htm>

英文引用格式: Zou L, Hu HQ, Yang C, Song HC. A survey of edge layout for network visualization. Ruan Jian Xue Bao/Journal of Software, 2014, 25(Suppl.(2)): 1-11 (in Chinese). <http://www.jos.org.cn/1000-9825/14017.htm>

A Survey of Edge Layout for Network Visualization

ZOU Ling¹, HU Hua-Quan^{2,3}, YANG Chao², SONG Han-Chen³

¹(State Key Laboratory of Virtual Reality Technology and Systems (BeiHang University), Beijing 100191, China)

²(The Key Laboratory of the Academy of Equipment, Beijing 101416, China)

³(Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, Changsha 410073, China)

Corresponding author: HU Hua-Quan, E-mail: hhq1986@foxmail.com, <http://www.nudt.edu.cn>

Abstract: Edge layout is an important aspect of network visualization and an essential tool for representing the overall structure of network and revealing the hidden information pattern. To survey and analyze the existing edge layout techniques, this paper first gives an overview on the tasks of edge layout of network and the aesthetics standards of graph drawing. Then, it classifies the edge layout techniques into three categories. Edge routing (ER), edge confluent (EC), and edge bundling (EB). Next, it elaborates the principles and characteristics of the typical layout methods, and highlights the analytics of edge bundling technique. Finally, this paper reviews the existing open problems, and offers an outlook on future work and challenges of edge layout.

Key words: edge layout; edge routing; edge confluent; edge bundling; network visualization

信息技术的大规模应用使得网络大数据分析成为亟需解决的问题. 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模式的有效工具而得到广泛应用. 通常将网络建模为一个图 $G=(V,E)$, 其中 V 表示网络的节点集合, E 表示网络边的集合. 网络可视化研究的焦点自然集中于网络节点布局和网络边布局这两个方面. 孙扬等人按照网络节点的布局方法对网络可视化技术进行了综述^[1], 但是没有对网络边布局的研究现状进

* 基金项目: 国家自然科学基金(61103081)

收稿时间: 2013-06-15; 定稿时间: 2013-08-21

行总结分析。

网络边布局研究致力于提高网络可视化图像的可读性,需要解决两个问题:一是边与边的相互交叉,二是边和节点的密集重叠.这两个问题导致了严重的视觉混乱现象.网络边布局的研究目标,即是通过设计新颖的边布局算法尽可能地清除视觉混乱现象,获得可理解的、清晰的网络可视化视图.为此,van Ham 等人尝试着用矩阵描述方法替代节点连接图^[2],其优点在于能够提供整洁的矩阵视图,但是可读性比较差,不如节点连接图显得直观^[3].因此,主流的网络可视化研究仍然采用基于节点连接图的描述方式,在美学标准的指导下开展相关研究.

直观上,网络边的描述方式分为直线和曲线两类.直线边包括普通直线边、折线边以及半线边^[4].当网络规模增加时,直线边描述方式存在严重的视觉混乱现象,难以获取到可理解的清晰网络图像,无法跟踪边的来源和走向.曲线边布局技术被认为是特别有效的一种技术,通过将某些边分组合并在一起,减少了由密集边带来的视觉混乱现象,更高效地利用了屏幕显示空间,在网络可视化社区中迅速得到了广泛应用,在 Graph Drawing, InfoVis, IV 等国际会议和 IEEE TVCG 等国际期刊中成为重要的议题.

1 网络边布局的主要任务和美学标准

网络可视化面临的主要问题是视觉混乱,因此边布局的主要任务就是设计良好的布局算法,通过对边进行布局调整,尽可能地清除视觉混乱现象.

Ellis 等人总结出 8 条视觉混乱清除技术的高层次任务^[5]:避免重叠、保持空间信息、能够定位、可缩放、可调整、能够显示点/线属性、能够区别点/线、能够观察出重叠的密度,同时研究了视觉混乱清除技术的分类方法,认为可以从元素外观、空间形变、时序动画 3 个方面进行文献分类.根据该分类方法,本文所讨论的网络边布局技术应该属于空间变形类.

为了较好地完成边布局任务,使用户可以最大程度地发挥其视觉感知能力,从而能够直观、深入地理解网络结构,研究人员提出了很多美学标准作为指导网络可视化绘制的原则^[6].Sugiyama 等人最早开始研究美学标准问题,为改进网络可视化图像的可读性定义了 5 项美学指导标准^[7]:

- 边交叉数量最少:这是一个统计上的美学标准,边交叉数量越小,越容易追踪边的走向,因此就越能够清晰地体现网络结构.
- 邻接点要相互靠近:具有邻接关系的点相互靠近,不具有邻接关系的点尽量相互远离,可以使边的长度减小,有利于网络结构的展示.
- 直线边:尽量采用直线而不是曲线来绘制网络中的边.事实上,对大规模网络进行可视化展示时,直线边会带来边交叉现象.为了解决该问题,基于曲线边布局的网络可视化研究取得了很好的效果,例如将在后文阐述的边集束技术等.因此我们认为,直线边原则仅适用于指导小规模网络的边布局.
- 边布局要保持整体平衡:与同一节点相连的边要保持比较均匀的夹角,从而保持整体上的布局平衡.
- 节点布局最好具有层次性:节点应该布局在水平或者垂直的不同层次上.

之后,该作者又补充了“布局原则一致性”和“网络边可追溯性”两条美学标准^[8].前者建议在整个网络可视化过程中使用一致的布局指导标准;后者建议清晰展示节点对之间的路径,从而能够较容易地追溯边的来源和走向.

Sindre 等人也提出了一组美学标准^[9],除却一些相似的标准外,新增标准包括:

- 尽量节省屏幕空间的区域最小原则:网络边布局过程中,在保证可视化效果的前提下,应尽量节省屏幕空间,以解决在“小屏幕”中展示“大信息”的问题.
- 尽量将度数较高的节点布局在画布中心的居中原则:节点的度往往反映着节点的重要程度,度越大的节点应该布局在可视化界面的中心;度越小的节点应该布局在可视化界面的边缘.
- 节点密度均匀原则:尽量使每个小区域的节点密度保持均匀,从而体现整体的平衡状态.

Purchase 等人也提出了多项美学指导标准,其中有几项标准实质上是对上述美学标准的不同描述,而新增的美学标准有^[10]:

- 同结构子图应对称布局:具有相同结构的子图,应该围绕可视化界面的中心进行平衡布局.
- 应沿着正交网格线布局:以正交网格线为布局背景,将节点布局在网格交叉点上,边沿着网格线进行布局.

必须指出的是,上述美学指导标准在有的情况下可能会存在冲突.例如,“直线边”原则可能会导致更多的边交叉,即违背了“边交叉数量最小”原则.并且,不同美学标准的重要性有差异.最新的研究通过召集 34 名志愿者进行绘图实验后分析指出^[11]，“移除边交叉”是最重要的美学标准,同时也是众多边布局技术的研究目的.

2 主要研究进展

边布局技术的核心思想是通过调整网络边的位置,或者对网络边进行合并,或者对网络边进行分组,从而缓解网络可视化中的视觉混乱现象.根据边调整的不同方式,本文将网络边布局技术归纳为 3 类:

- (1) 边路由技术(edge routing,简称 ER).通过调整边的路由,使网络边绕过那些重叠节点和密集区域.
- (2) 边融合技术(edge confluent,简称 EC).将相互交叉的边融合绘制成没有自环的光滑曲线.
- (3) 边集束技术(edge bundling,简称 EB).将具有相似走向的边集束在一起.

三者的主要区别在于:ER 只是调整边的路由,没有对网络边进行分组或者合并,其物理隐喻是生活中遇到障碍物时的绕行思想;EC 将多条交叉边融合绘制成一条共有路径,该共有路径与每条入边和出边的连接都是光滑的,其物理隐喻是火车轨道的合并;而 EB 只是将多条具有相似走向的边捆绑在一起,其物理隐喻是将多根电缆线进行捆扎而形成集束.

2.1 边路由技术

边路由技术是图形绘制、大规模集成电路设计(VLSI)及机器人路径设计领域中的热点问题^[12].与图形绘制相比,后两者对于边路由问题的研究主要集中于解决电路布线或机器人行进必须满足的物理约束问题,而这些物理约束与图形绘制中需要的自然的“弯曲”没有明显关系.

事实上,可以将节点与边的重叠现象看作是边的路由路径正好经过节点的布局位置.因此,边路由技术在本质上是要通过调整边的位置,寻找一条新的路由,使得新路由上经过的节点数目达到最少和边交叉数量最少,如图 1 所示.

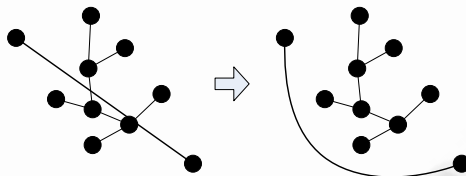


Fig.1 The principle of reducing edge crossing by using edge routing technique

图1 边路由技术减少边交叉原理

Dobkin 等人从图形绘制的角度出发,针对“绕行”穿过一组多边形障碍的问题,引入障碍规避研究中的可视图法(visibility graph),构建一条折线段的最短路径,并且为每一线段选择合适的样条函数,从而在平面内绘制一条光滑的路由曲线.然而,他们并未解决多条边路由跨越障碍的实现问题^[12].

Freivalds 等人提出一种完全不同于通过控制点产生样条曲线的方法,该方法将寻找边路由的过程简化为求解一个微分方程,从而能够直接描述网络绘图的美学^[13].具体方法是在绘图区域中定义一个连续代价函数,通过计算函数的极小值来获取低代价的路由路径.但是该方法的计算速度非常缓慢.

当网络节点具有某种形状时,网络边不仅要“绕过”节点,而且不能与节点的轮廓相交,因此更加具有挑战性.之前的做法是从切线视图中寻找最短路径,然而,完全切线视图的构建需要较大的代价,其时间复杂度至少是节点的二次方.因此,Dwyer 等人利用近似最短路径技术实现快速的边路由^[14],提出一种计算可视图近似值的启发式算法,将算法扩展到能够支持大规模网络的边路由.

边路由倾向于使用数学模型来减少边交叉,但仍然会导致一些边拥塞和边歧义问题.此外,边路由技术并没

有帮助用户识别高层次的宏观上的边模式,因此难以辅助用户理解整个网络的结构.

2.2 边融合技术

边融合技术最早由 Dickerson 等人提出^[15],其中心思想是将网络中的边分组融合在一起,并绘制为一条没有自相交和形状突变的局部单调曲线.事实上,在允许最小数量交叉点的同时,将交叉点隐藏在曲线绘制中.如图 2 所示,通过将边融合绘制到 E 点,图 2(b)中解决了图 2(a)中存在的边交叉问题(将 20 个交叉点转换成 1 个融合绘制点 E).例如, A 点与 C 、 D 两点之间的曲线没有自相交和形状突变,表示有边相连(可达);而 A 点必须经过 E 点才能到达 B 点,明显存在形状突变,从而表示 A 、 B 两点之间没有边相连(不可达).边融合技术实现了“边交叉数量最小”的美学标准,有效解决了网络可视化中的视觉混乱问题.

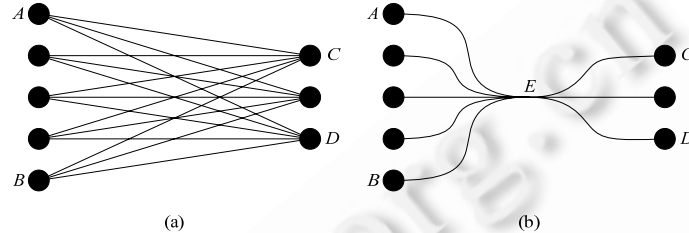


Fig.2 Confluent drawing of $K_{5,3}$

图2 $K_{5,3}$ 边融合绘制图

Dickerson 等人同时提出一种启发式算法^[15],用于检验网络是否适于边融合绘制.通过检验分析,分别列举了适合和不适合边融合绘制的几种典型网络,并从理论上加以证明.其中,适合边融合绘制的网络包含补图、补树、 n 环的补图、区间图等,不支持融合绘制的网络包含 Petersen 图(P 图)、 $P-\{v\}$ 图、四维超立方体等.由此可见,自动实现任意网络的边融合绘制具有较大难度,只有那些通过启发式算法检验的网络才能够被融合绘制,适用范围受到极大限制.由于 Dickerson 等人没有研究检验算法的复杂度问题,因此一般融合检验的复杂度问题仍是一个开放性问题,后续研究往往从融合结构和检验识别算法两个方面进行研究和改进.

在边融合绘制过程中,形成一系列重要的可重用的视觉元素,例如 Λ 连接(Λ -junction)、 Δ 连接(Δ -junction)、开关(switches)和交通环(traffic circles)等,如图 3 所示.其中,开关用于表示一个特殊的点,两条或多条曲线在这个点变换方向.交通环用于表示一个完全子图,所有的局部曲线共享一段公共的圆形轨迹.

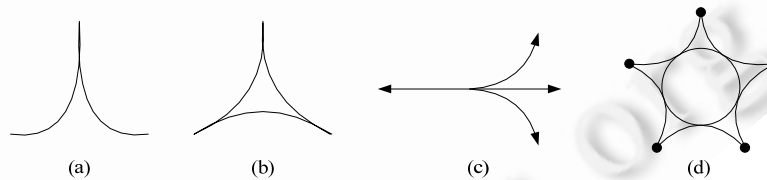


Fig.3 Λ -junction, Δ -junction, switches and traffic circles

图3 Λ 连接、 Δ 连接、开关和交通环

Hui 等人在此基础上定义了强融合图(strongly confluent graphs,简称 SCG)和树融合图(tree confluent graphs,简称 TCG)概念^[16].强融合图将生成局部单调曲线的约束进一步强化为生成光滑曲线,并证明其检验识别是 NP 问题;树融合图定义为当且仅当网络能够用二分的树拓扑加以描述.Eppstein 等人将 TCG 扩展到更加广泛的连接类型,提出基于三分完全连接的“ Δ -融合(Δ -confluent)”概念^[17],若将 Δ -融合的三分完全连接替换为单一节点,则退化为树.

Honciuc 等人提出正交融合(ortho-confluent)技术^[18],它的基础是以网络节点位置为参考点,构建一个局部正交的笛卡尔坐标系统.如图 4 所示.研发出一个正交融合网络绘制系统 ConfluentViz,已用于软件可视化、航线网络可视化、地铁网络可视化等实际应用中.正交融合与上述融合思想的主要区别在于,它的曲线融合绘制点位于正交网格的垂直轴或者水平轴上.

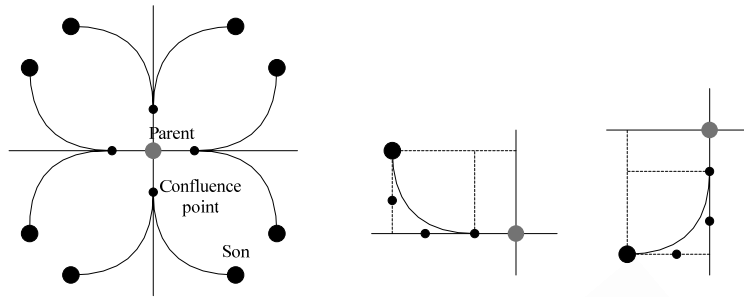


Fig.4 Ortho-Confluent system

图4 正交融合系统

此外,Eppstein 等人还将融合绘制思想应用到多层网络绘制^[19],以减少层与层之间边的交叉,但却放松了条件限制,允许存在少量的边交叉.Graham 等人也利用边融合思想,提出在绘制有向无环图时将多树(multiple tree)中共享相同父节点的一组“哑节点”合并在一起,以减少层间连接的视觉混乱^[20].孙扬等人基于边融合技术提出一种边布局算法,包括网络簇间和簇内的边融合,实现多变元网络边的布局^[21].

2.3 边集束技术

边集束技术被认为是特别有效的一种技术,通过将某些边进行集束,减少了视觉混乱现象,更加高效地利用了屏幕显示空间,在网络可视化社区中迅速得到了广泛应用,出现了一些有代表性的边集束技术,例如:层次边集束(hierarchical edge bundling,简称 HEB)^[24]、基于几何的边集束(geometry-based edge bundling,简称 GBEB)^[25]、力引导边集束(force-directed edge bundling,简称 FDEB)^[28]等.这些算法在不同程度上都改善了网络边布局,特别是对网络整体的结构化信息展示提供了更好的可视化效果.图 5 表示层次边集束技术的效果图,从左至右表示集束强度不断增强.图 6 表示 GBEB 和 FDEB 的效果图.

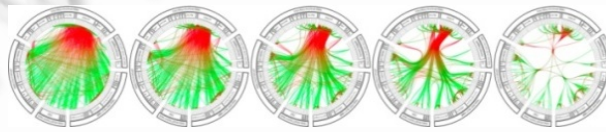


Fig.5 The results of hierarchical edge bundling

图5 层次边集束效果图

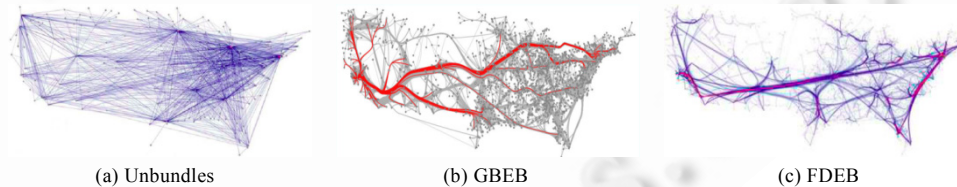


Fig.6 The comparison of unbundles graph, GBEB and FDEB

图6 几种边集束技术的效果图

边集束技术能够应用于大部分常见网络,包括普通网络^[22,24,27]、圆环布局^[30]、层次网络^[24]、并行坐标^[30-32]等.根据边集束算法的差异,本文将众多的边集束方法分成 3 类:基于几何的边集束方法、基于能量的边集束方法、基于图像的边集束方法.

2.3.1 基于几何的边集束方法

Carpendale 等人首先提出边变换技术(edge-displacement technique)^[22],利用变形公式使边弯曲并绕过圆形的感兴趣区域而集束在一起.Phan 等人提出流图布局(flow map layout)用于流量数据的可视化^[23],自动生成一个流量图,从全局上描述网络的骨干.Holten 等人提出的 HEB^[24]要求可视化数据具有层次结构和邻接关系,边的

端点是两个叶节点,边的路径沿着树结构上两节点间的最短路径生成,若多条边共享一段公共路径,则集束在一起,如图 7 所示.然而该技术要求其可视化的数据具有层次结构.Cui 等人基于节点的几何信息提出一种揭示高层次边模式的边集束方法 GBEB^[25],使用均匀的网格结构对边进行采样,然后生成控制网并引导那些几何上相互靠近并且方向相似的边形成弯曲的集束,如图 8 所示.该方法可以提供清晰的可见模式,但是其有效性高度依赖于控制网的质量,而又不能确保总是能自动获取到有效的控制网.

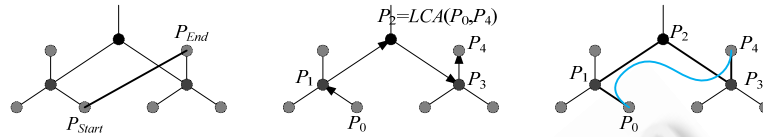


Fig.7 Edge bundling based on hierarchical structure

图7 HEB 利用层次结构实现边集束的示意图

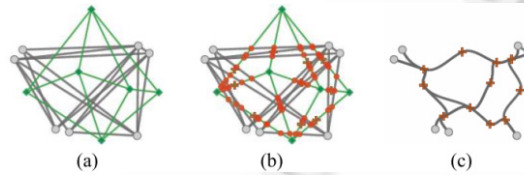


Fig.8 Edge bundling based on geometry

图8 GBEB 利用控制网格实现边集束的示意图

Lambert 等人实现了一种直观的 3D 边集束算法^[26],有效地减少了三维空间中网络的边杂乱现象.该方法首先在虚拟地球的外层空间和内层空间分别增加一层虚拟球面,在双层虚拟节点的控制下进行网格计算,将网络节点所处的空间离散成小区域,然后将区域的边界线作为网络边的新“路径”.在虚拟地球上实现边集束需要保证网络的每条边都至少存在一条不穿越球体的路径.同时提出基于 GPU 的绘制方法,使得用户在感知集束密度的同时保留边的颜色.如图 9 所示,数据来源于 2000 年全球空中交通互联网络,包含 1 524 个机场(节点)和 16 397 个航班(边).

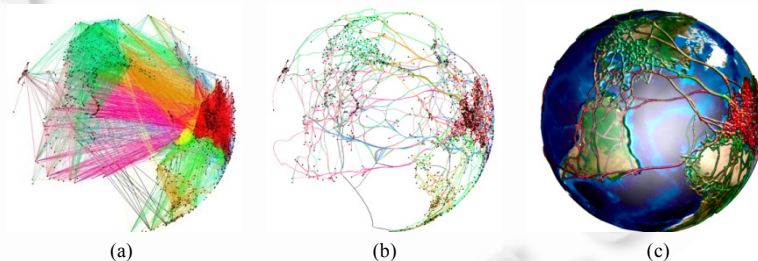


Fig.9 3D edge bundling of global airline network

图9 全球航空网络 3D 边集束技术效果

全局集束方法能够很好地从整体上解决大规模边的可视杂乱问题,但是难以跟踪边的走向.局部集束方法强调网络可视化局部的清晰程度,用户能够比较容易地追踪节点对之间的连接.Luo 等人提出一种无歧义的边集束方法(ambiguity free edge bundling,简称 AFEB)^[37],用于改进网络的局部.该方法更加有效地利用了屏幕显示空间,并且通过交互式界面支持局部的浏览,用户能够追踪边的路径.但是全局上的集束效果并不明显,不用于识别高层次的模式.

此外,Jia 等人将层次边集束(HEB)应用到社会网络可视化(真实的通信网络、合作网络和竞争网络),揭示出一些之前的可视化方法不能发现的隐含信息,并且在 Facebook 上实现了该技术的应用^[27].

2.3.2 基于能量的边集束方法

基于能量的边集束方法将网络边布局看作是一个物理系统的能量再平衡过程,与一般的力引导布局相区别,这里的力不是被施加到网络节点上,而是施加到网络的边上,更确切地说,是施加到网络边的细分点或者控制点上,从而能够调整边的曲线布局位置。

基于几何的边集束方法要求网络具有层次结构,或者需要构建控制网来引导边集束的过程,使得沿着集束方向的曲线曲率可能具有较大的偏差。为此,Holten 等人提出一种力引导边集束方法 FDEB^[28],将网络边建模为灵活的能够相互吸引的弹簧,采用自组织的力引导算法计算边集束。如图 10 所示,为了将直线变成曲线并实现集束,需要将边细分成片段,计算施加在每个细分点上的弹簧力和静电力的合力,当弹簧系统达到能量平衡状态时,完成网络边的集束。与之前的研究相比,FDEB 不要求网络具有层次结构,也不要构建控制网格,从而降低了集束方向上的曲率偏差。

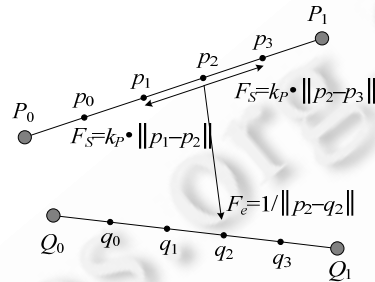


Fig.10 Two interactional edge P and Q

图10 两条相互作用的边 P 和 Q

Kienreich 等人将力引导边集束(FDEB)应用到社交媒体分析结果的可视化(澳大利亚约 30 个政治演说家之间的关系网络),促进了高层次边模式的识别,加强显示了成员之间的语义关系^[29]。

2.3.3 基于图像的边集束方法

基于图像的方法并不是通过显式的边绘制来清除边杂乱现象,而是将网络边布局看作是图像密度的改变,认为集束处的边图像密度较高,而在其他位置的边图像密度较低。文献[33]最早提出 GraphSplating 技术,将网络转换为二维标量场,可以被绘制成颜色编码图、高度场或者轮廓集。绘制采用高斯滤波器对网络节点和边进行卷积操作,形成高度图或者强度图。将能够导致混乱现象的密集边区域,呈现为致密的雪球溅射区域。但是,该技术存在两个问题:一是难以在生成的简化视图中跟踪雪球溅射区域的边的走向,二是需要仔细调节滤波器的带宽,以避免创建不连通的溅射区域,因而可能对用户形成误导。

为了更好地理解网络边集束的连接结构,Telea 等人提出基于图像的边集束方法(image-based edge bundles,简称 IBEB)^[34],该方法的核心思想是用集束的中心来简化网络结构。如图 11 所示,包含 6 个主要步骤:(1) 采用某个现有的边集束算法生成边集束布局;(2) 利用一个反映边位置和数据属性的距离将边分组到聚类层次;(3) 在用户指定的层次上,选择一个聚类集。对每个聚类,围绕它的边创建一个紧致形状;(4) 针对每个形状,构建阴影图元;(5) 以最小化遮挡的顺序绘制形状;(6) 采用语义透镜方法辅助用户浏览。综上所述,IBEB 方法有利于从宏观上识别集束的走向,但是细节表现存在不足,无法为用户提供更多的细节信息。

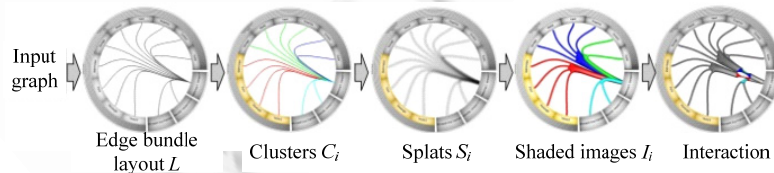


Fig.11 The visualization pipe of IBEB

图11 IBEB 可视化管线

Ersoy 等人提出基于骨架的网络边集束(skeleton-based edge bundling,简称 SBEB)^[35].利用边的骨架(即中轴线)作为位置信息来实现边的集束.将边聚类、距离场和 2D 骨架结合起来,通过不断迭代,使边不断地被吸引到距离场水平集的中心线方向,从而形成集束.与普通边聚类的区别在于,SBEB 方法的整个绘制管线完全是基于图像完成的.实现容易,并且可控性强.

Hurter 等人采用核密度估计方法实现网络边集束(kernel density estimation edge bundling,简称 KDEEB)^[36].该方法适用于普通网络,实现简单,快速稳定.主要步骤是:首先,采用核密度估计将网络转换成密度图;然后,应用图像锐化技术,通过将网络边的卷积移动到高度梯度流方向,合并局部高度,实现边的集束,图 12 展示了核密度估计边集束的主要流程.实现效果如图 13 所示(美国移民数据),从上至下表示迭代次数的增加,左侧表示密度图的演化,右侧是与左侧相对应的集束效果.从图中可见,生成的集束均沿着密度图的“山峰”进行路由.

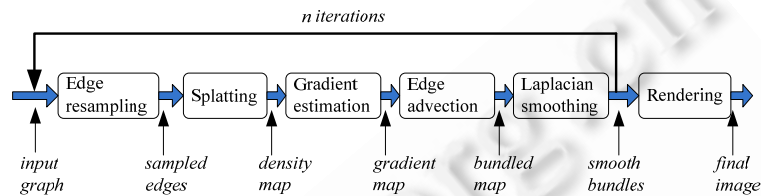


Fig.12 The rendering pipe of KDEEB

图12 KDEEB 绘制管线

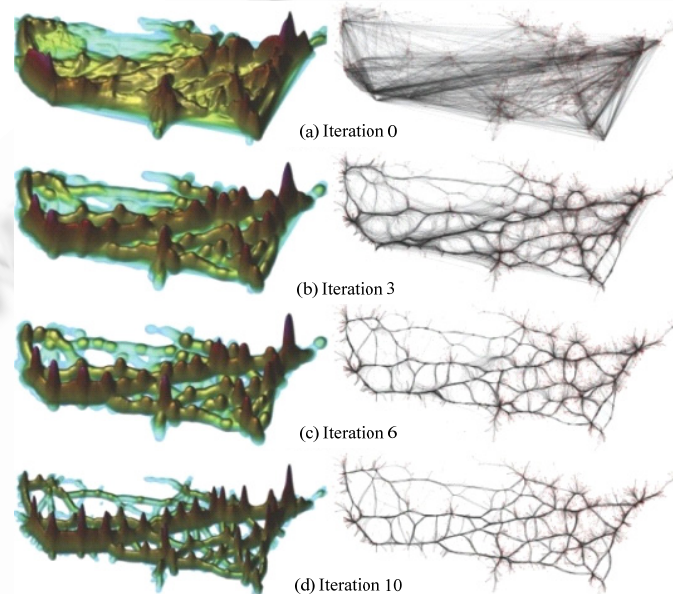


Fig.13 The evolution of density graph and corresponding bundling

图13 密度图的演化以及对应的集束效果

将众所周知的图像处理技术应用到网络边布局研究,是这类方法的最大创新之处,它创造性地提供了一种新的网络边集束研究模式.针对大规模网络的布局往往具有很大的时间复杂度,而一般的网络结构并不具有天然的并行计算优势,我们发现,基于图像的边集束方法有利于网络边布局的并行处理,能够充分利用图形硬件加速处理的优势,实现网络的快速布局和可视化分析.

3 扩展性及应用分析

在扩展性方面,针对大规模网络可视化,边路由技术一般只适用于小规模网络的边布局,在节点数量较少

时,能够通过边路由绕开某些障碍区域.针对大规模网络,很难通过边路由技术减少边交叉,从而不能满足网络可视化的美学标准.边融合技术的最大难题在于检验算法的计算复杂度.如果能够改善算法的计算复杂度,则能相应地将技术扩展到大规模网络;否则,仍然只能适用于中小规模的网络可视化研究.边集束技术能够较好地解决大规模网络可视化问题,一般能够处理数千个节点、上万条边的网络规模.例如文献[26]中的 3D 边集束算法能够处理的数据(2000 年全球空中交通互联网络)包含 1 524 个节点和 16 397 条边;文献[28]中的 FDEB 方法和文献[34]的 IBEB 方法能够处理的数据(美国移民数据)包括 1 715 个节点和 9 780 条边.可见,在扩展性和展示高层次的宏观模式方面,边集束技术要好于边路由技术和边融合技术.但是,在细节的表现能力上,边集束技术稍逊于前两者.

在应用方面,建议根据可视化任务选择合适的网络边布局技术,一般可以从两方面考虑:一是针对中小规模网络,对细节的要求较高,用户需要能够准确地追踪节点对之间的边连接来源与走向,因此最好选择边路由技术和边融合技术,例如科研合作网络的可视化等;二是针对大规模网络,对宏观模式的识别要求较高,用户需要感知高层次的网络模式,因此最好选择边集束技术,例如社交网络可视化、网页链接关系可视化等.值得注意的是,不论是工业界还是学术界,对大数据的关注可谓方兴未艾.在未来很长一段时间内,这种关注还将持续^[38].近年来,网络大数据特别是数十亿顶点规模的网络数据大量涌现.这一趋势对网络可视化研究领域提出了全新的挑战.

4 结束语

边布局技术是网络可视化研究的重要内容,应用十分广泛.本文根据边调整方式的差异将现有边布局技术分为 3 类:边路由技术、边融合技术和边集束技术,阐述了它们的共同点及其差异.从边布局技术实现的基本原理、关键技术和优缺点等方面对现有技术进行了总结分析.得出如下结论:尽管已经进行了大量的、艰苦的研究工作并取得诸多令人鼓舞的研究成果,但网络边布局问题还远远没有被很好地解决,集中体现在以下几个方面:

第一,全局模式识别和局部走向追踪两者之间还没有取得很好的平衡.全局上的可视化专注于使整个网络体现出清晰可见的高层次的边模式,但是网络中两个准确节点之间是否存在连接则无法追踪.相反,局部上的可视化专注于精确体现节点之间的连接,能够准确区分特定边的源点和目标点,但是全局模式体现不足,全局上仍存在一定程度的视觉混乱现象.因此,如何设计平衡的边布局算法仍是有待研究的问题之一.

第二,计算速度和美学标准之间也远未达到实用需求,从网络中准确探测出满足特定条件的社团结构,以及实现美学标准较高的边布局(如交叉数量最少等)往往需要复杂的计算时间.目前,图形硬件已经发展到较高水准,如果能够充分利用 GPU 的通用计算功能实现网络边布局算法的并行加速计算,必然能够提升计算速度,实现良好的美学布局.

第三,随着对网络认识的深入和应用领域的拓展,现有的网络边布局方法已经不能满足多样的需求,需要针对特殊类型的网络研究新型的网络边布局技术.例如,时变网络比传统的静态网络更接近网络的本质状态,针对时变网络建模与分析的研究已经成为研究热点,但是目前尚未见到针对时变网络的边布局技术^[39],相关理论和方法还远未成熟,因此,如何针对新型网络类型设计相应的边布局方法也是当前面临的主要问题之一.

总的来说,随着网络大数据的不断涌现,传统的网络边布局技术需要着重解决扩展性问题,以适应包含海量节点及边的网络结构.同时,随着对网络认识的深入和应用领域的拓展,网络边布局技术还需要关注新的网络类型,例如时变网络、多重网络等.这些都是下一步研究需要重点解决的问题和相应的研究趋势.

致谢 在此,我们向对本文的工作给予支持和建议的同行、实验室参与学术讨论的老师和同学们表示感谢.

References:

- [1] Sun Y, Jiang YX, Zhao X, Xiao WD. Survey on the research of network visualization. *Computer Science*, 2010,37(2):12–18 (in Chinese with English abstract).
- [2] van Ham F. Using multilevel call matrices in large software projects. In: Munzner T, ed. *Proc. of the 9th Annual IEEE Conf. on Information Visualization (INFOVIS 2003)*. Washington: IEEE Computer Society, 2003. 227–232.
- [3] Ghoniem M, Fekete JD, Castagliola P. A comparison of the readability of graphs using node-link and matrix-based representations. In: *Proc. of the IEEE Symp. on Information Visualization (INFOVIS 2004)*. Washington: IEEE Computer Society, 2004. 17–24.
- [4] Becker RA, Eick SG, Wilks AR. Visualizing network data. *IEEE Trans. on Visualization and Computer Graphics*, 1995,1(1):16–28.
- [5] Ellis G, Dix A. A taxonomy of clutter reduction for information visualization. *IEEE Trans. on Visualization and Computer Graphics*, 2007,13(6):1216–1223.
- [6] Shneiderman B, Aris A. Network visualization by semantic substrates. *IEEE Trans. on Visualization and Computer Graphics*, 2006,12(5):733–740.
- [7] Sugiyama K, Tagawa S, Toda M. Methods for visual understanding of hierarchical system structures. *IEEE Trans. on Systems, Man and Cybernetics*, 1981,11(2):109–125.
- [8] Sugiyama K. A cognitive approach for graph drawing. *Cybernetics and Systems*, 1987,18(6):447–488.
- [9] Sindre G, Gulla B, Jokstad H. Onion graphs: Aesthetic and layout. In: *Proc. of the IEEE Symp. on Visual Languages*. Bergen: IEEE Computer Society, 1993. 287–291.
- [10] Purchase HC, Cohen RF, James M. Validating graph drawing aesthetics. In: Brandenburg FJ, ed. *Proc. of the Symp. Graph Drawing (GD'95)*. London: Springer-Verlag, 1996. 435–446.
- [11] Purchase HC, Pilcher C, Plimmer B. Graph drawing aesthetics-created by users, not algorithms. *IEEE Trans. on Visualization and Computer Graphics*, 2012,18(1):81–92.
- [12] Dobkin DP, Gansner ER, Koutsofios E, North SC. Implementing a general-purpose edge router. In: Battista GD, ed. *Proc. of the 5th Int'l Symp. on Graph Drawing (GD'97)*. London: Springer-Verlag, 1998. 262–271.
- [13] Freivalds K. Curved edge routing. In: Freivalds R, ed. *Proc. of the 13th Int'l Symp. on Fundamentals of Computation Theory (FCT'01)*. London: Springer-Verlag, 2001. 126–137.
- [14] Dwyer T, Nachmanson L. Fast edge-routing for large graphs. In: Eppstein D, ed. *Proc. of the 17th Int'l Conf. on Graph Drawing (GD 2009)*. Berlin: Springer-Verlag, 2010. 147–158.
- [15] Dickerson M, Eppstein D, Goodrich MT, Meng JY. Confluent drawings: Visualizing non-planar diagrams in a planar way. In: Liotta G, ed. *Proc. of the 11th Int'l Conf. on Graph Drawing (GD 2003)*. LNCS, 2004. 1–12.
- [16] Hui P, Pelsmaje MJ, Schaefer M, Stefankovic D. Train tracks and confluent drawings. *Algorithmica*, 2004,47(4):465–479.
- [17] Eppstein D, Goodrich MT, Meng JY. Delta-Confluent drawing. In: Healy P, ed. *Proc. of the 13th Int'l Conf. on Graph Drawing (GD 2005)*. Berlin: Springer-Verlag, 2006. 165–176.
- [18] Honciuc I, Croitoru C. On confluent drawings: Visualizing graphs using an ortho-confluent system. *Acta Univ. Sapientiae, Informatica*, 2009,1(2):135–148.
- [19] Eppstein D, Goodrich MT, Meng J Y. Confluent layered drawings. In: Pach J, ed. *Proc. of the 12th Int'l Conf. on Graph Drawing (GD 2004)*. Berlin: Springer-Verlag, 2004. 184–194.
- [20] Kennedy J. Exploring multiple trees through DAG representations. *IEEE Trans. on Visualization and Computer Graphics*, 2007, 13(6):1294–1301.
- [21] Sun Y, Zhao X, Tang JY, Tang DQ, Xiao WD. Multivariate network visualization paradigm. *Ruan Jian Xue Bao/Journal of Software*, 2010,21(9):2250–2261 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3889.htm> [doi: 10.3724/SP.J.1001.2010.03889]
- [22] Carpendale MST, Rong X. Examining edge congestion. In: Tremaine M, ed. *CHI 2001 Extended Abstracts on Human Factors in Computing Systems (CHI EA 2001)*. New York: ACM, 2001. 115–116.
- [23] Phan D, Xiao L, Yeh R, *et al.* Flow map layout. In: Tremaine M, ed. *Proc. of the 2005 IEEE Symp. on Information Visualization (INFOVIS 2005)*. Washington: IEEE Computer Society, 2005. 219–224.
- [24] Holten D. Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *IEEE Trans. on Visualization and Computer Graphics*, 2006,12(5):805–812.
- [25] Cui W, Zhou H, Qu HM, Wong PC, Li XM. Geometry-Based edge clustering for graph visualization. *IEEE Trans. on Visualization and Computer Graphics*, 2008,14(6):1277–1284.

- [26] Lambert A, Bourqui R, Auber D. 3D edge bundling for geographical data visualization. In: Proc. of the 2010 14th Int'l Conf. Information Visualisation (IV 2010). Washington: IEEE Computer Society, 2010. 329–335.
- [27] Jia Y, Garland M, Hart J C. Social network clustering and visualization using hierarchical edge bundles. Computer Graphics Forum, 2011,30(8):2314–2327.
- [28] Holten D, van Wijk JJ. Force-Directed edge bundling for graph visualization. In: Hege HC, ed. Proc. of the 11th Eurographics/IEEE-VGTC Conf. on Visualization (EuroVis 2009). Eurographics Association Aire-la-Ville, 2009. 983–998.
- [29] Kienreich W, Seifert C. An application of edge bundling techniques to the visualization of media analysis results. In: Proc. of the 2010 14th Int'l Conf. Information Visualisation (IV 2010). Washington: IEEE Computer Society, 2010. 375–380.
- [30] Gansner W, Koren Y. Improved circular layouts. In: Kaufmann M, ed. Proc. of the 14th Int'l Conf. on Graph Drawing (GD 2006). Berlin: Springer-Verlag, 2007. 386–398.
- [31] Mcdonnell KT, Mueller K. Illustrative parallel coordinates. In: Vilanova A, ed. Proc. of the 10th Joint Eurographics/IEEE-VGTC Conf. on Visualization (EuroVis 2008). Eurographics Association Aire-la-Ville, 2008. 1031–1038.
- [32] Zhou H, Yuan XR, Qu HM, Cui WW, Chen BQ. Visual clustering in parallel coordinates. In: Vilanova A, ed. Proc. of the 10th Joint Eurographics/IEEE-VGTC Conf on Visualization (EuroVis 2008). Eurographics Association Aire-la-Ville, 2008. 1047–1054.
- [33] van Liere R, de Leeuw W. Graphspatting: Visualizing graphs as continuous fields. IEEE Trans. on Visualization and Computer Graphics, 2003,9(2):206–212.
- [34] Telea A, Ersoy O. Image-Based edge bundles: Simplified visualization of large graphs. In: Melançon G, ed. Proc. of the 12th Joint Eurographics/IEEE-VGTC Conf. on Visualization (EuroVis 2010). Eurographics Association Aire-la-Ville, 2010. 843–852.
- [35] Ersoy O, Hurter C, Paulovich FV, Cantareira G, Telea A. Skeleton-Based edge bundling for graph visualization. IEEE Trans. on Visualization and Computer Graphics, 2011,17(12):2364–2373.
- [36] Hurter C, Ersoy O, Telea A. Graph bundling by kernel density estimation. Computer Graphics Forum, 2012,31(3pt1):865–874.
- [37] Luo SJ, Liu CL, Chen BY, Ma KL. Ambiguity-Free edge-bundling for interactive graph visualization. IEEE Trans. on Visualization and Computer Graphics, 2011,18(5):810–821.
- [38] Feng GD, Xiao YH. Distributed storage of big graphs. Communications of the CCF, 2012,8(11):12–15 (in Chinese with English abstract).
- [39] Hu HQ, Wu LD, Yang C, Song HC. Survey on time-varying network visualization. Journal of System Simulation, 2013,25(9):1–7 (in Chinese with English abstract).

附中文参考文献:

- [1] 孙扬,蒋远翔,赵翔,肖卫东.网络可视化研究综述.计算机科学,2010,37(2):12–18.
- [21] 孙扬,赵翔,唐九阳,汤大权,肖卫东.一种多变元网络可视化方法.软件学报,2010,21(9):2250–2261. <http://www.jos.org.cn/1000-9825/3889.htm> [doi: 10.3724/SP.J.1001.2010.03889]
- [38] 冯国栋,肖仰华.大图的分分布式存储.中国计算机学会通讯,2012,8(11):12–15.
- [39] 胡华全,吴玲达,杨超,宋汉辰.时变网络可视化研究综述.系统仿真学报,2013,25(9):1–7.



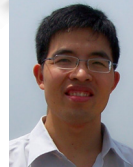
邹玲(1986—),女,博士生,主要研究领域为虚拟现实技术,信息可视化.
E-mail: wld@nudt.edu.cn



杨超(1982—),男,博士,研究员,主要研究领域为多媒体信息与虚拟现实,系统集成.
E-mail: yangchao@nudt.edu.cn



胡华全(1986—),男,博士生,CCF 学生会员,主要研究领域为多媒体信息与虚拟现实,网络可视化,信息可视化.
E-mail: hhq1986@foxmail.com



宋汉辰(1976—),男,博士,副教授,CCF 会员,主要研究领域为网络可视化,信息可视化.
E-mail: songhanchen@nudt.edu.cn