

基于变换的大图点边可视化综述

时磊¹⁾, 廖琦²⁾, 林闯³⁾

¹⁾(中国科学院软件研究所计算机科学国家重点实验室 北京 100190)

²⁾(美国中密歇根大学计算机科学系 美国密歇根省愉悦峰 48859)

³⁾(清华大学计算机科学与技术系 北京 100084)

(shil@ios.ac.cn)

摘要: 大图可视化是信息可视化领域的前沿课题之一,也是在线社会网络、信息安全、电子商务等热点行业大数据分析的重要支撑技术. 基于变换的大图点边可视化方法由于其具有在线处理时间短、可视复杂度低、交互方法灵活多样等优点,近年来在学术界与实际商用系统中得到广泛重视与应用. 文中从图可视化的基本概念及其在大图上的关键挑战出发,梳理了基于变换的大图点边可视化方法的典型分类与主要流程;通过详述 3 类基于变换的大图点边可视化典型方法(图数据抽象、视图变换与视角转换),阐明了不同方案的优缺点与适用场景,并进一步指出了未来工作的可行方向与潜在难点.

关键词: 大图可视化;图数据抽象;视图变换;视角转换

中图法分类号: TP391

Survey on Transformation-based Large Graph Visualization

Shi Lei¹⁾, Liao Qi²⁾, and Lin Chuang³⁾

¹⁾(State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing 100190)

²⁾(Computer Science Department, Central Michigan University, Mount Pleasant, MI 48859 United States)

³⁾(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

Abstract: Large graph visualization is one of the hot topics in the information visualization research field. It is also widely accepted as the fundamental technique of the big data analytics in industries such as online social networks, information security and e-business. The transformation-based large graph visualization methods have been intensively studied recently due to their advantages over the classical drawing methods in the fast processing speed, low visual complexity and versatile interactions available. They are adopted in many real-world systems and applications. In this paper, we start from the basic concept of large graph visualization and its major challenges. We classify this kind of methods into three types (graph abstraction, view transformation and view point transition) and introduce in detail the representative approaches in each type. Both the pros/cons and the practical usage scenarios are talked about for these methods. Future directions are discussed with respect to the potential technical challenges going ahead.

Key words: large graph visualization; graph abstraction; view transformation; view point transition

节点-边图是反映实体间关联关系最重要的一种表示方式,又称为关联图、网络或关联网络. 本文研究主要面向节点数目较高的关联图,统称为大图. 可视化学术界对于大图尚未给出一致的定义,通常认为基本力学模型(force-directed)^[1]和平面图(planar graph)布局^[2]方法均难以实时处理的关联图可视为大图. 除了节点数目高以外,大图还普遍具有无尺度特性、附带高维节点-边信息,并随时间动态变化. 常见的大图包括在线社会网络平台上用户间的好友网络或关注网络,万维网中网页的链接关系网络,知识系统中知识点的关联网络以及顾客与商品之间的二分兴趣网络. 随着信息技术的井喷式发展,尤其是大数据产生、获取、加工、分析等环节的不断进步,基于大图数据的各类应用正逐渐增多. Google 公司的知识图系统^[3]不仅用于优化搜索结果,还提供了搜索内容关联知识点的直观展示; Facebook, Twitter, LinkedIn 等在线社交网站根据已知的海量用户关联图预测未来高概率发生的好友关系,并提前推荐给用户^[4]; Amazon, 淘宝等购物网站利用类似的关联图推荐热点商品^[5].

在上述大图应用中,可视化方法在以下两方面起到重要的作用:

1) 作为可视分析工具,辅助大图特征提取与应用设计:由于大图数据的高度复杂性,其内在规律发现、特征提取与分析算法选择均面临极大的挑战. 可视化方法可从全局直观展示数据,并提供交互式特征发现手段,利于洞察其内在规律、评价特征、完善算法设计;

2) 直观展示上下文,为终端用户诠释大图分析结果:大图上的算法计算通常较为繁重,难以向终端用户直接解释运行原理与规则. 提供分析结果的可

视化,如所推荐用户与当前用户的关联关系网络,可极大地提升应用的可信度与有效性.

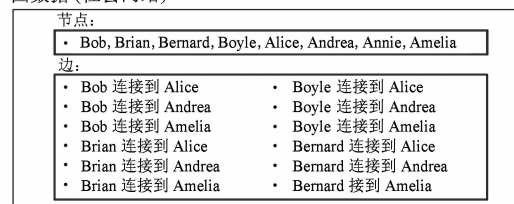
1 图可视化基础与大图上的关键挑战

如图 1 a 所示,图可视化是信息可视化的核心组成部分,它在可视化领域发展壮大的数十年中始终得到高度关注,并产生了大量的研究成果^[6-8]. 本质上,图的点边可视化解决如何将图数据转化为可视关联图的问题,详见图 1 b 所示;其主要研究点包括:1) 图布局. 以何种模型排布图节点并连接边,如何快速、准确地计算节点布局;2) 图渲染. 如何可视化表达图节点与边上信息及节点上下文(如群组、社区等);3) 图交互. 提供哪些交互手段辅助用户可视分析图数据. 本文关注的大图可视化是图可视化领域近期的主要热点课题. 在大图上,经典图可视化方法面临以下两方面挑战:

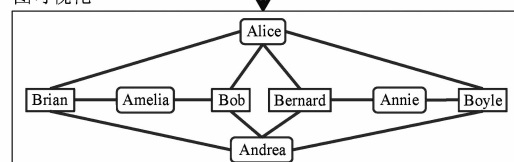
1) 实时快速大图布局算法. 应用基本力学模型的图布局算法^[9-10]计算复杂度显著超过 $O(N^2)$,在节点数大于 10^2 时不再满足实时布局需求. 采用多级粗化-再加工的优化大图布局算法^[11-12]可支持节点数 10^5 级别大图的快速布局(约 1 min). 在近似的时间尺度下(1 min 以内),高维嵌入方法^[13]和引入 Hall 能量模型的 ACE 算法^[14]可进一步布局节点数 10^6 级别大图,然而其仅对类网状图数据达到较好的布局效果. 如何快速、有效地布局百万级或更高节点数大图,仍然是对图可视化研究的关键挑战;

2) 大图可视复杂度简化方法. 即使算法进步允许超大图的快速布局,终端用户仍难以消费和理解过于杂乱的可视表达,尤其是大量边交叉引入的负面影响^[15]. 研究表明,对于节点数超过 20 的图,采用

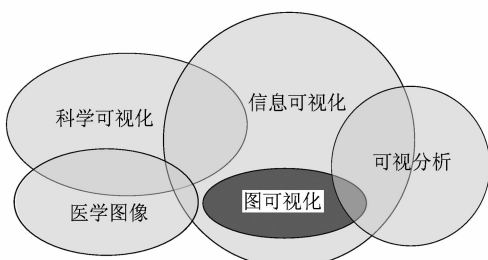
图数据(社会网络)



图可视化



b 概念图解



a 可视化领域所处地位

图 1 图可视化基础

邻接矩阵展示方法在大部分基本图分析任务下性能优于采用节点-边的经典方法^[16]. 如何在数据层面或视图层面简化大图,以兼顾节点-边可视形式带来的直观优势,以及低可视复杂度保证的信息可消费性,是大图布局算法之外重要的研究方向. 本文内容主要概述此类方法.

2 基于变换的大图点边可视化方法

如前文所述,仅仅采用优化布局的方法已不能满足当前大图信息可视化的需求. 大量最新研究^[17-20]集中于复杂大图的简化以及精简视图生成与可视表达,我们统称这些方法为基于变换的大图可视化.

图 2 给出了此类可视化方法的基本分类与处理流程,其中虚线部分表示基于布局算法的传统方法.

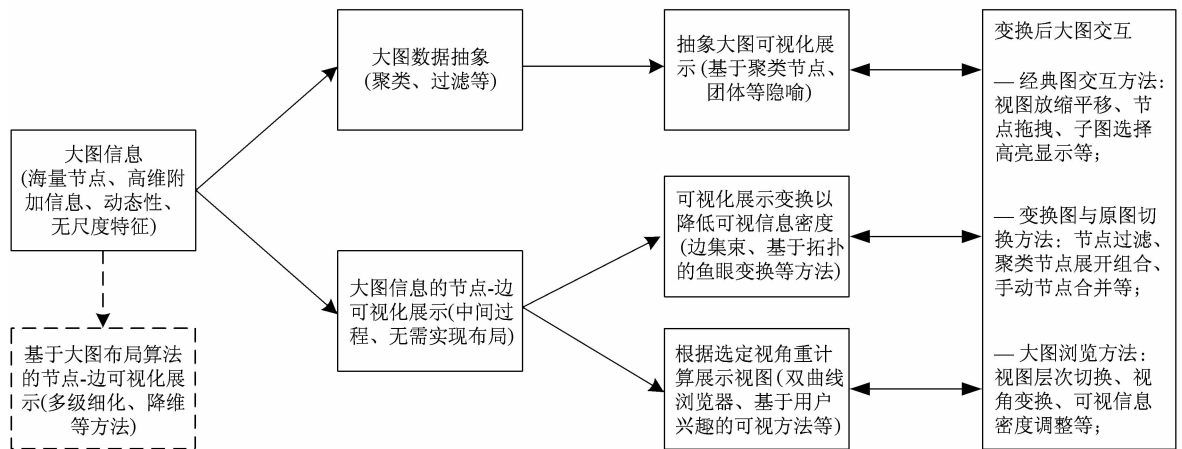


图 2 基于变换的大图可视化方法分类与主要流程

2.1 图数据抽象

图数据抽象最典型的一类方法为图聚类,其相关研究已发展数十年,理论与应用体系均比较完善. 本文内容限于大图可视化常用的图聚类思路及分类,而不试图涵盖所有图聚类方法本身,感兴趣的读者可进一步阅读相关文献^[21]. 如图 3 所示,基于聚类的大图可视化流程分为 3 个步骤. 由于图聚类的核心为将近似节点归并成群组,其首要步骤即是选定量化的节点近似度的指标;之后根据该指标应用聚类算法计算出所有群组以及节点-群组的从属树型关系;最后,采用聚类节点等隐喻实现聚类树的可视化表达,同时支持一定的交互方法以提供用户可视分析能力.

已知文献对图节点相似度做了大量研究^[21-24],其中对图可视化较有价值的主要分为 2 类. 一类为

针对输入的大图信息,其变换方法分为 2 类. 第一类应用各种数据抽象方法,在第一阶段将大图信息变换为较小的图或更易于布局的图类型(如树结构),比较经典的方法如图聚类、基于节点-边权重过滤等. 此类方法在第二阶段基于简化图计算布局并生成视图,通过引入某些特别设计的隐喻表达聚类、群组等新加入信息. 第二类方法在第一阶段仍采用原始大图的可视化表达,但其仅用于中间结果不需实际计算布局. 在第二阶段,此类方法在展示视图基础上实现各种变换,以达到精简视图的相同目标. 其主要方法包括视图变换,如边集束、拓扑鱼眼变换;以及视角转换,如双曲线浏览器、基于用户兴趣的视图生成. 在第三阶段,本文所述的大部分方案均涉及各类大图交互方法,即通过经典图交互、变换图与原图切换、大图浏览模式等交互手段满足大图可视分析的需求.

基于图内聚特性 (cohesive) 的节点相似度,如根据图理论距离^[25]、节点间路径数^[26]的定义. 直观上,按此类方法生成的图聚类结果满足群组内部节点间边密度高、群组之间边密度低的特征. 定量地,Newman 等提出的模块化指标 (modularity)^[27] 是一种评价聚类性能的典型方法. 令 e_{ij} 定义聚类图中群组 i 到群组 j 的边在原图中的比例, $a_i = \sum_j e_{ij}$ 定义从群组 i 出发的边的比例,模块化指标定义为

$$Q = \sum_i (e_{ii} - a_i^2) \quad (1)$$

该指标越大(比如超过 0.3),图聚类结果的内聚性越强. 在社会网络等大图场景下,根据此类相似度生成的群组又称为社区,定义一组连接紧密的人或物通常具有相似的特征或档案.

与之并列,另一类图节点相似度根据其结构等

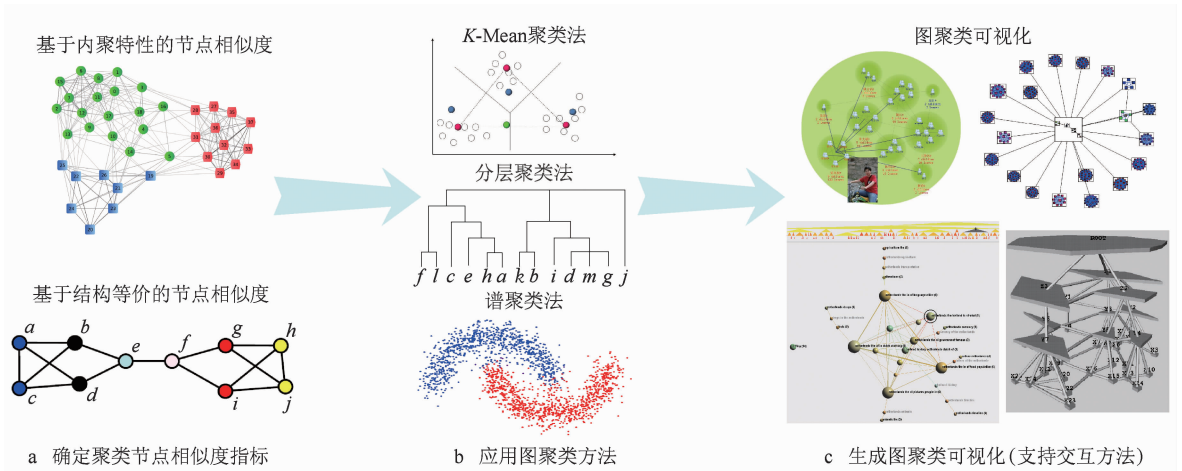


图3 图聚类可视化基本步骤

价性 (structural equivalence) 定义^[23], 并通过比较节点的邻居集合计算; 相似节点在图结构中处于等同或近似的地位. 定量地, 令节点 u 与 v 的邻居节点集合分别为 $\Gamma(u)$ 和 $\Gamma(v)$, 其节点相似度可按 Jaccard 系数计算, 即

$$\omega(u, v) = \frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u) \cup \Gamma(v)|} \quad (2)$$

其中, $\omega(u, v) = 1$ 的节点称为绝对结构等价. 引入图角色的概念, 节点结构等价性还可推广为自同构等价 (automorphic equivalence) 及正则等价 (regular equivalence).

以节点相似度矩阵为输入, 图聚类方法可广泛采用经典聚类算法, 如 K-Mean 方法^[28]、层次聚类方法^[29]、谱聚类^[30]等方法, 本文不再一一赘述. 图聚类方法的输出为节点-群组树型结构, 分为单层树与多层树 (应用层次聚类方法) 2 类. 图聚类可视化即将该树型结构转化为可视化直观表达的方法, 由于单个节点的可视化方式与普通图可视化近似, 其主要由群组及节点-群组关系的可视化表达决定.

Quigley 等^[31]采用基于节点距离的几何聚类方法从大图上生成分层聚类树, 并选取特定层次上的抽象图, 在可视表达上延续经典的节点-边展示方法. Abello 等^[17]进一步利用节点大小和渲染颜色区分聚类群组与原始节点, 并表达群组内节点数目与群组类型. 类似地, 基于结构等价性聚类方法的可视化方案^[32-35]由于不需展示聚类内部结构, 通常也采用特殊设计的聚类节点表示群组, 如颜色饱和度映射群组大小^[33]或采用主题图案 (motif)^[34]. Auber 等^[18]提出以子图的形式同时展示聚类树群组及群组内部结构, 类似的工作还有文献^[20, 36-37]. 由于二维平面的局限性, 以上方法仅支持单层或少数几

层群组-节点树型结构展示, 而利用交互手段作为补充. Eades 等^[38]的工作探索了采用三维可视化方法展示完整树型结构的可能性, 但三维平面的遮挡问题一定程度上影响了此类系统的可用性.

由于聚类可视化的复杂性, 本文所涉及大部分方案均支持一定的用户交互手段, 以实现完整大图的浏览并支持探索性可视分析; 信息可视化领域的经典交互方法在其中都有所体现. Abello 等^[17]采用全局加细节的展示方法, 结合层次结构树图与某个群组节点内部图可视化实现对大图的浏览; 当用户在树图中选取某个群组节点时, 对应的内部结构图也相应展开. Shi 等^[33]允许用户手动拖拽节点合并为群组, 或双击群组以将其展开为原始节点; 提供保持图上下文的方式浏览大图树型结构^[20]. Archambault 等^[39-40]还提出了在大图可视化中灵活修改层次结构、挖掘节点间相关性的一系列交互方法.

图过滤是另一类常用的图数据抽象方法. 最基础的过滤算法利用预先计算或已存在的节点权重, 选取权重较大的若干节点, 并删除剩余节点及连接到被删除节点的所有边. Jia 等方法^[40]与之对应, 选取大图的边权重作为过滤依据; 它根据一种特殊设计的边中心性指标, 过滤后的大图能够更好地维持拓扑结构. Van Ham 等^[42]进一步提出利用最小生成树的方法构建更为清晰的大图可视化, 但视图在一定程度上丢失了原有大图的拓扑结构. 与上述图视图抽象方法均不同的是, Wattenberg^[43]利用了图节点自带信息 (如社会网络中用户的性别) 压缩图结构.

2.2 视图变换

除上述对相似图节点聚类的方法^[17-18, 20, 31], 关联图的边也可在几何上组合, 以降低大图可视化复杂

度,这类方法称为边集束或边捆绑^[19,44-49].边集束方法基本思路是根据图拓扑结构或可视布局为每条边选择若干控制点,通过使原始边布线穿过全局共享的控制点,实现大量边的几何聚集,有效地减少大图可视化尤其是边交叉带来的视觉杂乱.

图 4 所示分级边集束方法^[44]基于以上聚集邻接边的原则,类似于把共同路径的电线和网线捆绑在一起直至临近终点再分开的惯例,使原本错综复杂的电线和电缆系统更易于管理,其中绿色表示呼叫发起方,红色表示接收方.该方法针对树型图结构,利用树的层次关系指导边集束控制点选择.具体地,分

级边集束方法为 2 个具有邻接关系的节点按照其层次结构路径选择一个样条曲线作为控制多边形,组成控制多边形的控制点由起点到终点层次结构路径上的所有点构成.为减轻捆绑边之间分辨不清的问题,可调整曲线参数逐步拉直样条曲线,即在边集束效果和信总损失量之间做权衡.具体实现上,应用样条曲线数学模型建立计算接口,该接口允许用户设计和控制复杂曲线与曲面形状:用户输入若干序列点,即控制点,计算产生曲线.曲线形状逼近此序列控制点,实际穿过每个控制点的曲线称为插值曲线;附近通过但不穿过控制点的曲线称为近似曲线.

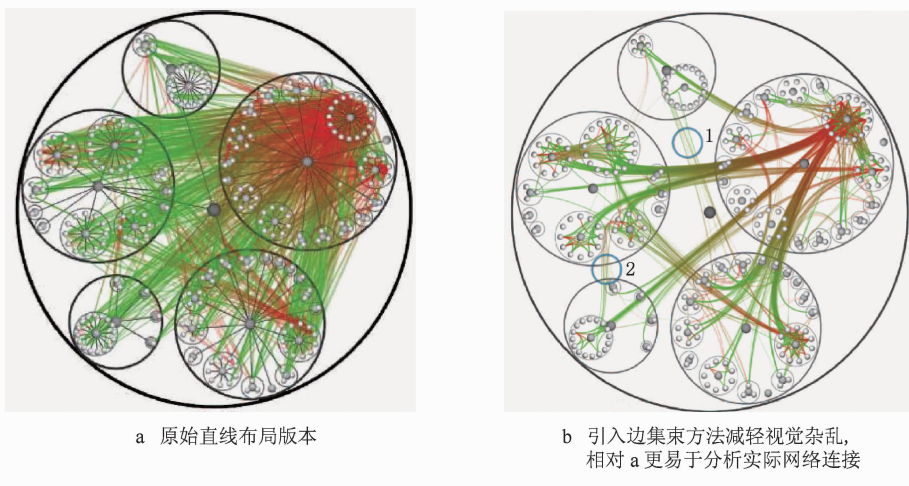


图 4 分级边集束大图可视化方法展示电话呼叫图

变集束方法也可与路线图设计相结合,即节点间连接不仅仅可以是均一的直线或曲线,而是由多段附属道路组成的城市间高速公路. Cui 等研究表明^[19],由普通直线图变换为路线图的方式可以有效地减少视觉杂乱;其中,城市可认为是控制点,道路可认为是连接城市的线段,路径须通过某些城市 and 道路.为合理选择控制点(即道路地图中的城市),可首先从图形结构上生成控制网格(控制点组成的网络);控制网格可手动或自动生成.例如,首先研究边分布以确定一组点,并使这些点间的边穿过原始图中边密集的部分;之后,采用 Delaunay 三角剖分使这些点连接到一起,形成三角控制网格.控制网格中的边与原始图中的边相交,产生一组交叉点;变集束控制点应在这些交叉点的中心.这时,可使用聚类方法,如 K-means 聚类,找到这些交叉点的中点,最终确定控制点.通过使所有边穿过控制点,可达到边集束效果.

综上所述,边集束的共同特征是利用曲线来替代直线绘制边.边集束方法的关键问题在于如何找到有意义的控制点,以引导曲线边延伸的方向.但

是,这种方法附加的潜在挑战是如何既达到边集束效果又不引入歧义性,即不丢失拓扑信息.

基于图拓扑的鱼眼变换^[50]是另一种大图视图变换方案.与大图的多级优化布局方法类似,它首先生成层次树状结构,具体展示图在树的某个层次选取,以控制可视复杂度.其创新点在于,用户可指定一个或多个视图焦点,这些焦点附近采用较深层次的拓扑结构,即更细粒度的展示.远离焦点的部分选取较浅层次的拓扑结构,展示数据更为稀疏.通过附加辐射状变换,还可做到信息密度均一化.这种方案一定程度上也带有视角转换的意思;此外,大图视图变换方案还包括基于节点密度的高效渲染^[50]等方法,本文不再一一赘述.

2.3 视角转换

双曲浏览器^[52-53]是一种直观展示大规模树型(层次)结构的可视化方案,如企业组织结构图或计算机系统目录结构,可显示重要的局部层次结构而避免整体展示带来的复杂度.双曲浏览器的核心思路是焦点加上下文的可视化设计,即分配更多的展示

空间给用户重点关注的层次结构,同时附带刻画所嵌入的整体层次结构背景.算法上,它将树型结构布局在二维双曲平面^[52]或三维双曲空间^[53]上,再映射回传统欧几里德圆形显示区域(称为单位圆).双曲平面是一种非欧几里德几何平面,平行线间发散远离对方,双曲平面上圆面积随半径呈指数增长.因此,这些性质给本文方案带来了显著优势:以指数增长的空间排布树型结构中指数增长的叶子节点,保证了均一的信息密度.

上文提到的基于交互的“全局+细节”大图可视化系统^[17]是另一类视角转换方案.用户通过在全局

视图选取某些聚类群组节点,可在并列细节视图展示群组内部结构.此外,在很多场景下,展示大图全局结构并无必要,且其易受网络带宽和客户端计算资源限制.如图5所示, Van Ham等^[54]支持用户根据兴趣度(degree of interest, DOI)远程交互浏览大图内部结构的方案;其中,兴趣度可由用户搜索的关键词匹配图节点得到,所匹配节点将用户兴趣度根据特定函数与拓扑图结构扩展到全图.使用计算生成的全图兴趣度分布,可进一步生成包含高兴趣度节点、用户收益最大化且尺寸受限的背景拓扑图,如利用贪婪近似算法.

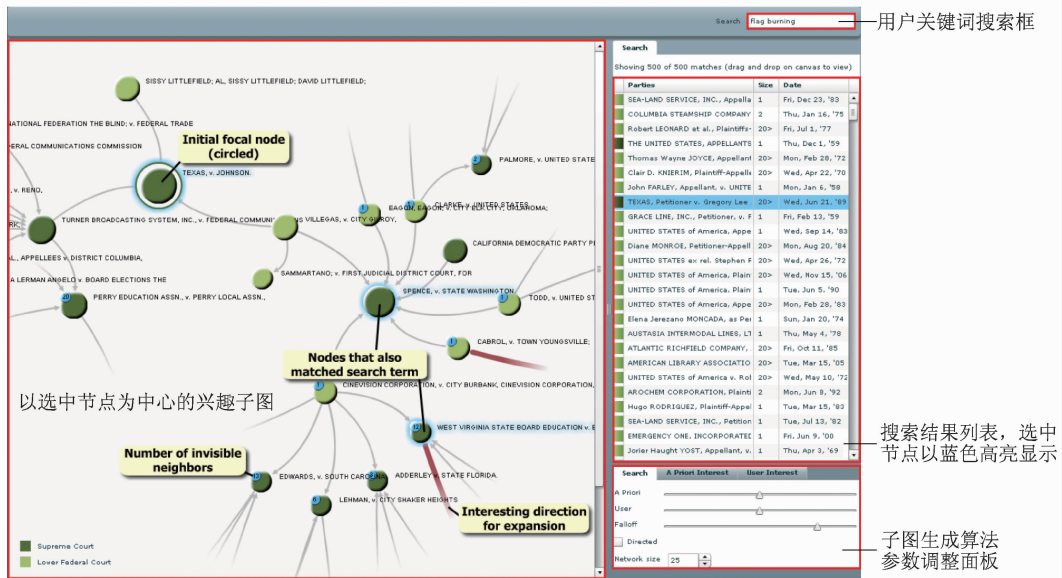


图5 基于用户兴趣度的大图可视化界面

2.4 小结

综上所述,基于变换的大图可视化方法以变换方法为核心、新型交互模式为补充,将大图信息从数据层面或视图层面转换为较小的图或特殊类型图,同时应用各类布局算法与可视表达.与传统基于优化布局的方案相比,它具有在线处理时间短、可视复杂度低、交互方法灵活多样等优点,更加适用于大数据场景下图信息基于浏览器或移动终端的可视化需求.本文所述3类变换方法的优缺点总结如下:

1) 图数据抽象方法(如聚类算法)直接作用于底层数据,可提供较高的视图压缩率,即使对于超大规模的图数据仍可生成极为简洁的可视表达,因此可应用于大图的全局可视总结(又称宏观鸟瞰图).同时,由于数据抽象方法不可避免地改变原始数据的结构,所生成的可视表达在一定程度上异于原始大图的可视表达,这增加了用户理解的难度.

2) 视图变换方法作用于原始大图可视表达基础上,能够较大幅度地保存大图数据客观展示形态,

维持用户对于数据的意象地图(mental map)^[55-56].与之对应,对于大部分实际场景生成的大图,变换后的可视表达仍然存在可视复杂度高的问题,这也给用户的可视分析任务带来了困难.

3) 视角转换方法以大图的局部视图取代全局展示,在更贴近用户的可视分析需求的同时,提供了极为清晰易懂的可视表达.然而,此类方法容易陷于大图的细节而丢失重要的宏观信息,采用全局加细节(或焦点加上上下文)联合展示可在一定程度上解决其局限性.

3 结论与展望

基于变换的大图可视化方法是优化图布局算法外大图可视化研究近期的主要方向,并随着图数据量与内在复杂度的不断增长日益受到重视.本文介绍的大量研究表明,基于变换的大图可视化方法可支持百万量级或更高数目节点的大图(如 ASK-

Graph 系统^[17]和基于双曲面的展示平台), 此类方法也是大部分实际商用系统中采用的方案. 然而, 当前的图变换方法、可视表达以及交互模式均不能完全满足大图上用户可视分析的需求, 这也指引学术界在以下方面开展进一步研究:

1) 突出大图语义的高压缩率通用图变换-抽象方法. 由于用户可视内存的局限性, 变换大图的首要目标是简化程度符合用户可消费、可理解的范围内, 同时保持甚至突出大图整体结构语义. 此外, 常见压缩算法对于不同领域的大图数据存在显著的压缩率差异, 如何设计领域通用的算法是潜在的难点.

2) 符合用户习惯的大图可视表达. 已知文献中已提出若干变换后大图的可视化方案, 如群组聚类节点、子图展示, 但甚少针对大图新型可视表达开展用户调研. 通过此类基于任务的用户调研, 可更加准确地定量地比较变换大图与原始大图可视表达的优缺点, 进一步指导算法与可视化设计.

3) 针对通用大图分析任务的高效交互模式. 本文所述大图可视化方案均配套用户交互方法, 但大部分从普通图分析继承而来或仅支持连续的大图层次浏览; 根据大图特定的分析任务制定新型交互模式, 必然是未来的研究方向.

参考文献 (References):

- [1] Eades P. A heuristic for graph drawing [J]. *Congressus Numerantium*, 1984, 42: 149-160
- [2] Battista G D, Eades P, Tamassia R, *et al.* Graph drawing: algorithms for the visualization of graphs [M]. Upper Saddle River: Prentice Hall PTR, 1998
- [3] Singhal A. Introducing the knowledge graph: things, not strings... is epub [OL]. [2012-12-31]. <http://googleblog.blogspot.com/2012/05/introducing-knowledge-graph-things-not.html>
- [4] Kautz H, Selman B, Shah M. Referral web: combining social networks and collaborative filtering [J]. *Communications of the ACM*, 1997, 40(3): 63-65
- [5] Schafer J B, Frankowski D, Herlocker J, *et al.* Collaborative filtering recommender systems [M] // *Lecture Notes in Computer Science*, Chapter 9, Heidelberg: Springer, 2007, 4321: 291-324
- [6] Herman I, Melancon G, Marshall M S. Graph visualization and navigation in information visualization: A survey [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2000, 6(1): 24-43
- [7] Hu Y. Algorithms for visualizing large networks [M] // *Combinatorial Scientific Computing*, CRC Press, 2012: 525-549
- [8] Von Landesberger T, Kuijper A, Schreck T, *et al.* Visual Analysis of Large Graphs [OL]. [2012-12-31]. <http://www.gris.tu-darmstadt.de/~ttekusov/papers/egstar10.pdf>
- [9] Fruchterman T M J, Reingold E M. Graph drawing by force-directed placement [J]. *Software, Practice & Experience*, 1991, 21(11): 1129-1164
- [10] Kamada T, Kawai S. An algorithm for drawing general undirected graphs [J]. *Information Processing Letters* 1989, 31(1): 7-15
- [11] Gajer P, Kobourov S G. GRIP: Graph drawing with intelligent placement [J]. *Journal of Graph Algorithms and Applications*, 2002, 6(3): 203-224
- [12] Hu Y. Efficient and high quality force-directed graph drawing [J]. *Mathematica Journal*, 2005, 10(1): 37-71
- [13] Harel D, Koren Y. Graph drawing by high-dimensional embedding [J]. *Journal of Graph Algorithms and Applications*, 2004, 8(2): 195-214
- [14] Koren Y, Carmel L, Harel D. ACE: a fast multiscale eigenvector computation for drawing huge graphs [C] // *Proceedings of IEEE Symposium on Information Visualization*. Los Alamitos: IEEE Computer Society Press, 2002: 137-144
- [15] Becker R A, Eick S G, Wilks A R. Visualizing network data [J]. *IEEE Transactions on Visualization and Computer Graphics*, 1995, 1(1): 16-28
- [16] Ghoniem M, Fekete J D, Castagliola P. A comparison of the readability of graphs using node-link and matrix-based representations [C] // *Proceedings of the IEEE Pacific Visualization Symposium*. Los Alamitos: IEEE Computer Society, 2004: 17-24
- [17] Abello J, Van Ham F, Krishnan N. ASK-GraphView: a large scale graph visualization system [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2006, 12(5): 669-676
- [18] Auber D, Chiricota Y, Jourdan F, *et al.* Multiscale visualization of small world networks [C] // *Proceedings of IEEE Symposium on Information Visualization*. Los Alamitos: IEEE Computer Society Press, 2003: 75-81
- [19] Cui W, Zhou H, Qu H, *et al.* Geometry-based edge clustering for graph visualization [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2008, 14(6): 1277-1284
- [20] Shi L, Cao N, Liu S X, *et al.* Himap: adaptive visualization of large-scale online social networks [C] // *Proceedings of the IEEE Pacific Visualization Symposium*. Los Alamitos: IEEE Computer Society Press, 2009: 41-48
- [21] Schaeffer S E. Graph clustering [J]. *Computer Science Review*, 2007, 1(1): 27-64
- [22] Hennig C, Hausdorf B. Design of dissimilarity measures: a new dissimilarity between species distribution areas [M] // *Studies in Classification, Data Analysis, and Knowledge Organization*. Berlin: Springer, 2006: 29-38
- [23] Lorrain F, White H C. Structural equivalence of individuals in social networks [J]. *The Journal of Mathematical Sociology*, 1971, 1(1): 49-80
- [24] Williams W T, Dale M B, Macnaughton-Smith P. An objective method of weighting in similarity analysis [J]. *Nature*, 1964, 201(4917): 327-349

- [25] Edachery J, Sen A, Brandenburg F J. Graph clustering using distance- k cliques [M] // Lecture Notes in Computer Science. Heidelberg: Springer, 1999, 1731: 98-106
- [26] Hartuv E, Shamir R. A clustering algorithm based on graph connectivity [J]. Information Processing Letters, 2000, 76(4-6): 175-181
- [27] Newman M E J, Girvan M. Finding and evaluating community structure in networks [J]. Physical Review E, 2004, 69: No.026113
- [28] Hartigan J A, Wong M A. Algorithm as 136: a k -means clustering algorithm [J]. Applied Statistics, 1978, 28(1): 100-108
- [29] Defays D. An efficient algorithm for a complete link method [J]. The Computer Journal, 1977, 20(4): 364-366
- [30] Capocchia A, Servedio V, Caldarella G, *et al.* Detecting communities in large networks [J]. Physica A: Statistical Mechanics and Its Applications, 2005, 352(2-4): 669-676
- [31] Quigley A, Eades P. FADE: graph drawing, clustering and visual abstraction [M] //Lecture Notes in Computer Science. Heidelberg: Springer, 2001, 1984: 197-210
- [32] Papadopoulos C, Voglis C. Drawing graphs using modular decomposition [J]. Journal of Graph Algorithms and Applications, 2007, 11(2): 343-354
- [33] Shi L, Liao Q, Yang C. Investigating network traffic through compressed graph visualization [OL]. [2012-12-31]. http://lcs.ios.ac.cn/~shil/paper/VAST_Challenge_2012.pdf
- [34] Smith M, Shneiderman B, Milic-Frayling N, *et al.* Analyzing (social media) networks with nodeXL [C] //Proceedings of the 4th International Conference on Communities and Technologies. New York: ACM Press, 2009: 255-264
- [35] Van Ham F, Wattenberg M, Viegas F B. Mapping text with phrase nets [J]. IEEE Transactions on Visualization and Computer Graphics, 2009, 15(6): 1169-1176
- [36] Archambault D, Munzner T, Auber D. Multi-level graph layout by topological features [J]. IEEE Transactions on Visualization and Computer Graphics, 2007, 13(2): 305-317
- [37] Kumar G, Garland M. Visual exploration of complex time-varying graphs [J]. IEEE Transactions on Visualization and Computer Graphics, 2006, 12(5): 805-812
- [38] Eades P, Feng Q. Multilevel visualization of clustered graphics [M] //Lecture Notes in Computer Science. Heidelberg: Springer, 1997, 1190: 101-112
- [39] Archambault D, Munzner T, Auber D. Grouseflocks: steerable exploration of graph hierarchy space [J]. IEEE Transactions on Visualization and Computer Graphics, 2008, 14(4): 900-913
- [40] Archambault D, Munzner T, Auber D. Tugging graphs faster: efficiently modifying path-preserving hierarchies for browsing paths [J]. IEEE Transactions on Visualization and Computer Graphics, 2011, 17(3): 276-289
- [41] Jia Y, Hoberock J, Garland M, *et al.* On the visualization of social and other scale-free networks [J]. IEEE Transactions on Visualization and Computer Graphics, 2008, 14(6): 1285-1292
- [42] Van Ham F, Wattenberg M. Centrality based visualization of small world graphs [J]. Computer Graphics Form, 2008, 27(3): 975-982
- [43] Wattenberg M. Visual exploration of multivariate graphs [C] //Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM Press, 2006: 811-819
- [44] Ersoy O, Hurter C, Paulovich F V, *et al.* Skeleton-based edge bundling for graph visualization [J]. IEEE Transactions on Visualization and Computer Graphics, 2011, 17(12): 2364-2373
- [45] Gansner E R, Hu Y, North S C, *et al.* Multilevel agglomerative edge bundling for visualizing large graphs [C] // Proceedings of the IEEE Pacific Visualization Symposium. Los Alamitos: IEEE Computer Society Press, 2011: 187-194
- [46] Holten D. Hierarchical edge bundles: visualization of adjacency relations in hierarchical data [J]. IEEE Transactions on Visualization and Computer Graphics, 2006, 12(5): 741-748
- [47] Luo S, Liu C, Chen B, *et al.* Ambiguity-free edge-bundling for interactive graph visualization [J]. IEEE Transactions on Visualization and Computer Graphics, 2012, 18(5): 810-821
- [48] Phan D, Xiao L, Yeh R, *et al.* Flow map layout [C] // Proceedings of IEEE Symposium on Information Visualization. Los Alamitos: IEEE Computer Society Press, 2005: 219-224
- [49] Selassie D, Heller B, Heer J. Divided edge bundling for directional network data [J]. IEEE Transactions on Visualization and Computer Graphics, 2011, 17(12): 2354-2363
- [50] Gansner E, Koren Y, North S. Topological fisheye views for visualizing large graphs [J]. IEEE Transactions on Visualization and Computer Graphics, 2005, 11(4): 457-468
- [51] Zinsmaier M, Brandes U, Deussen O, *et al.* Interactive level-of-detail rendering of large graphs [J]. IEEE Transactions on Visualization and Computer Graphics, 2012, 18(12): 2486-2495
- [52] Lamping J, Rao R, Pirolli P. A focus + context technique based on hyperbolic geometry for visualizing large hierarchies [C] //Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Los Alamitos: IEEE Computer Society Press, 1995: 401-408
- [53] Munzner T. H3: laying out large directed graphs in 3D hyperbolic space [C] // Proceedings of IEEE Symposium on Information Visualization. Los Alamitos: IEEE Computer Society Press, 1997: 2-10
- [54] Van Ham F, Perer A. "Search, show context, expand on demand": supporting large graph exploration with degree-of-interest [J]. IEEE Transactions on Visualization and Computer Graphics, 2009, 15(6): 953-960
- [55] Eades P, Lai W, Misue K, *et al.* Preserving the mental map of a diagram [OL]. [2012-12-31]. <http://www.cs.tsukuba.ac.jp/~misue/publications/techreport/iias-rr-91-16e.pdf>
- [56] Misue K, Eades P, Lai W, *et al.* Layout adjustment and the mental map [J]. Journal of Visual Languages & Computing, 1995, 6(2): 183-210