

基于同源控制点的边缘绑定方法

刘梦欣, 张凡, 李天瑞

引用本文

刘梦欣, 张凡, 李天瑞. [基于同源控制点的边缘绑定方法](#)[J]. 计算机科学, 2022, 49(10): 96-102.

LIU Meng-xin, ZHANG Fan, LI Tian-rui. [Edge Bundling Method Based on Homologous Control Points](#)[J].

Computer Science, 2022, 49(10): 96-102.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于四边形最优圈内最短路径的旅行商问题割边方法](#)

Cutting Edge Method for Traveling Salesman Problem Based on the Shortest Paths in Optimal Cycles of Quadrilaterals

计算机科学, 2022, 49(6A): 199-205. <https://doi.org/10.11896/jsjcx.210400065>

[面向无尺度图的 \$\Delta\$ -stepping 算法改进策略](#)

Strategies for Improving Δ -stepping Algorithm on Scale-free Graphs

计算机科学, 2022, 49(6A): 594-600. <https://doi.org/10.11896/jsjcx.210400062>

[面向定位应用的无人机的悬停位置和飞行路径优化](#)

Hover Location Selection and Flight Path Optimization for UAV for Localization Applications

计算机科学, 2021, 48(11): 345-355. <https://doi.org/10.11896/jsjcx.201000105>

[基于 CKSP 的分段路由负载均衡技术](#)

Load Balancing Technology of Segment Routing Based on CKSP

计算机科学, 2020, 47(4): 256-261. <https://doi.org/10.11896/jsjcx.190500122>

[基于增量最短路径优先的域内高效路由保护算法](#)

Efficient Intra-domain Routing Protection Algorithm Based on i-SPF

计算机科学, 2019, 46(8): 116-120. <https://doi.org/10.11896/j.issn.1002-137X.2019.08.019>

基于同源控制点的边缘绑定方法

刘梦欣¹ 张 凡^{1,2,3} 李天瑞^{1,2,3}

1 西南交通大学计算机与人工智能学院 成都 611756

2 四川省制造业产业链协同与信息化支撑技术重点实验室 成都 611756

3 综合交通大数据应用技术国家工程实验室 成都 611756

(mxinl@my.swjtu.edu.cn)

摘 要 对含有大量复杂连接关系的节点连接图进行可视化会造成视觉上的严重混乱,边缘绑定是一种有效降低视觉混乱的方法。以往基于空间邻近性进行边缘绑定的方法会导致独立边缘产生模糊性歧义,给予用户错误的认知,而只专注于图的拓扑结构无法有效解决密集连接造成的视觉干扰问题。基于边缘路径的方法能够较好地利用图中原始节点信息对边缘进行控制绑定,从而避免独立边缘产生模糊性歧义,同时展现数据的高级模式。因此,在边缘路径方法的基础上进行了改进,提出了一种基于同源控制点的边缘绑定方法。该方法结合图的拓扑结构信息计算同源控制点,并以此为基础利用最短路径算法选取边缘控制点,然后结合分级思想对边缘聚合程度进行优化,最后通过 Bézier 曲线对边缘进行平滑处理。将基于同源控制点的边缘绑定方法用于美国迁移数据集和中国铁路线路数据集中,实验结果表明,该方法在改善过度绑定的问题上起到了较好的效果,相比原方法,此方法保留了更多局部数据细节,平衡了整体与局部边缘的绑定程度,可以有效地用于复杂连接图的可视化。

关键词: 节点连接图;同源控制点;最短路径;边缘绑定;图可视化

中图法分类号 TP391.41

Edge Bundling Method Based on Homologous Control Points

LIU Meng-xin¹, ZHANG Fan^{1,2,3} and LI Tian-rui^{1,2,3}

1 School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China

2 Manufacturing Industry Chains Collaboration and Information Support Technology Key Laboratory of Sichuan Province, Chengdu 611756, China

3 National Engineering Laboratory of Integrated Transportation Big Data Application Technology, Chengdu 611756, China

Abstract Edge bundling is an effective method to reduce the visual clutter caused by the visualization of the node-link diagram with a large number of complex connections. Generally, the edge bundling based on spatial proximity will lead to independent edge ambiguity and give users a wrong perception. However, focusing only on topological structure of graphs cannot reduce visual clutter caused by dense connections to a large extent. The method based on edge path can control and bundle the edges by using the original nodes in the graph, avoid independent edge ambiguity, and show the advanced mode of data. Therefore, an edge bundling method based on homologous control points is proposed to improve the edge path method. Based on the topology structure information of the graph, the method can calculate homologous control points and select edge control points by using the shortest path algorithm. Then the degree of edge aggregation is optimized with the thinking of gradation. Finally, the edges are smoothed through Bezier curves and colored according to the direction of the edges. The edge bundling method based on homologous control points is used in the US migration dataset and the Chinese railway line dataset. Experimental results show that this method has a good effect on improving the problem of over-bundling. Compared with the original method, this method retains more local data details, balances the bundling degree between the whole and local edges, and can be effectively used for the visualization of complex connected graphs.

Keywords Node-link diagrams, Homologous control points, Shortest path, Edge bundling, Graph visualization

到稿日期:2022-03-07 返修日期:2022-07-15

基金项目:国家自然科学基金(62176221)

This work was supported by the National Natural Science Foundation of China(62176221).

通信作者:李天瑞(trli@swjtu.edu.cn)

1 引言

图可视化通常被作为探索大量复杂数据的有效手段,其能够结合各个领域数据分析技术,以图形化形式为用户直观地展示数据分析结果^[1],同时可以结合交互技术^[2-4],辅助用户探索数据的更深层次关联。现实生活中的很多真实数据集,如美国移民数据、空中航线网络数据^[5]和道路网络数据^[6]等,一般由大量相互连接的节点组成,而众多边缘的交叉连接会形成视觉上的混乱,从而干扰用户对有效信息的获取,因此需要一种良好的布局帮助用户提高对数据的认知,让用户快速且高效地获取相关信息。边缘绑定技术通过特定方法将大量复杂的连接聚集成束,以减少边缘交叉数^[7],揭示数据隐藏含义,能够有效解决视觉混乱问题。

许多研究工作^[8-10]基于边缘在空间上的邻近程度对其进行聚合,这会导致独立边缘模糊性歧义^[11],即真实独立的两个边缘被聚合成束捆绑在一起,误导用户判断该聚合束上每个节点都是相互连通的。Dickerson 等^[12]于 2003 年提出的 Confluent Drawing 方法则不存在这种问题,该方法只会将具有共享源点和目标点的边缘进行绑定,得到的绑定结果完全符合图的拓扑结构。但正是由于该方法的这一严格限制,图中可以聚合的边缘明显减少,以致于无法聚合的边缘数量较多,不能完全解决视觉混乱问题。Wallinger 等^[11]提出了一种 Edge-Path 边缘绑定方法,能够有效改善视觉混乱问题,同时避免了独立边缘模糊性歧义。虽然该方法对全局数据模式有一个较好的概览,但是缺乏对数据局部细节的描述。因此,本文对 Edge-Path 边缘绑定方法进行改进,并在此基础上提出一种基于同源控制点的边缘绑定方法。同源控制点主要用于控制同一源点出发的边缘,使其能够更大程度地聚合在一起,通过同源控制点对边缘进行聚合,可以在一定程度上放大数据的局部细节,感知图中局部边缘流向。本文提出的方法改善了 Edge-Path 方法产生的边缘过度捆绑问题,能够细致描绘图中节点处边缘的流向,从而提高边缘绑定方法对数据局部特征可视化的能力,同时结合分级思想优化边缘聚合程度,有效改善了较短边缘曲率较大的问题。

2 相关工作

边缘绑定作为目前最有效的一种技术手段,通常被应用于含有大量连接关系的图形中,用于解决复杂连接带来的视觉混乱问题。随着应用场景的多元化,研究者从不同角度探索了边缘绑定技术,包括分层边缘绑定^[13-15]、基于几何的边缘绑定^[8,16-17]和基于力导向的边缘绑定^[9,18-19]技术等,这些方法都在一定程度上改进了图形布局,从而更有效地可视化了整个图形的结构信息。

最初,Phan 等^[13]使用层次聚类将 Flow Map 中具有相同目的地的边缘合并,从而避免了边缘交叉。受到该工作的启发,Holten^[14]提出一种分层边缘绑定技术,该技术可结合树形图帮助用户快速了解层级组织中的邻接关系,但是该方法局限于具有层次结构的图。Zhou 等^[15]则针对一般节点连接图提出了基于能量的层次边缘聚类方法,通过使用 Delaunay

三角剖分分割原始图,基于自适应采样的片段对边缘进行聚类,更好地保留了图的拓扑信息。然而对于大型图而言,该方法生成的捆绑布局会产成较多锯齿形边,误导用户对原始图的分析判断。Cui 等^[8]提出了一种基于几何的边缘绑定技术,该技术利用原始图形结构创建控制网格并生成控制点,经过相同控制点的边缘可以聚类成簇,同时引入局部平滑技术来提高边缘绑定图的有效性。Lambert 等^[16]进一步扩展了该算法,基于四叉树和维诺图的混合方式生成控制网格,有效降低了边缘混乱程度,同时获得了较好的执行性能和较短的执行时间。Luo 等^[17]也利用四叉树结构降低算法的时间复杂度,同时以交互式的方式辅助用户选择局部区域,以帮助其进一步探索边缘密集处区域的特点。

与基于几何体来定义边缘之间聚合性的方法有所不同,Holten 等^[9]引入了物理系统概念,提出了一种基于力导向的边缘绑定算法,该算法可以在保证节点位置不变的情况下将边建模为柔性弹簧,通过相互吸引使边聚集成束,这种方式形成的束曲率变化较小,更易形成平滑束。Wu 等^[20]提出一种基于力导向分段式骨骼布局的边缘绑定方法,有效可视化了数据的离散程度和聚类中心等特征。Nguyen 等^[21]在工作中提出了 4 种不同的力导向方法,通过对社会网络、生物网络、地理网络和聚类图等数据集进行实验,证明了这些方法可以有效地显示图形的骨架结构。但是在很多交互式设备上可视化这些复杂的节点连接图时,上述算法较高的时间复杂度会造成一定的延迟,降低用户的使用感。Wu 等^[18]基于 Web 实现了一种基于纹理的边缘绑定方法,可以有效地降低绑定延迟,同时能够应用于各种设备。很多研究工作^[19,22]将边绑定技术拓展到 3D 空间,同样取得了较好的表现,有效降低了视觉干扰。

这些方法通过不同方式将在空间上较为邻近的边缘聚集在一起,最大程度地减少了边缘混乱现象,从而将图形更深层次的流动模式展示给用户。但是这些方法在一定程度上都受到独立边缘模糊性的影响,从而给用户带来误导性信息,影响用户的分析。Confluent drawing 方法^[12]可以将含有大量交叉点的图形以无交叉的方式绘制,符合图的拓扑结构,因此不会产生这种边缘歧义。Confluent drawing 方法严格的绘制方式^[23]导致其很难被应用于现实世界的实际问题中,因此 Bach 等^[24]将该方法扩展到非平面 Confluent drawing 中,并通过放松捆绑来限制连接的模糊性,但是局部区域边缘连接较密集的地方仍存在模糊性。Zheng 等^[25]在研究工作中表示,可以通过放宽平面性的条件来进一步减少图中交叉点的数量。Wallinger 等^[11]提出一种 Edge-Path 绑定方法,使用最短路径的方法来计算每个边缘的控制点,通过这些控制点拟合所有边缘曲线,可以较好地显示数据的高级模式,同时更好地消除独立边缘模糊性歧义。基于 Edge-Path 方法的优势,本文针对其在局部细节描述上的欠缺进行了改进,提出了同源控制点的概念,以同源控制点作为边缘新的控制点可以清晰地展示各个节点所连接边缘的流向,从而达到放大局部细节的作用。然而我们在进一步探索中发现,同源控制点的选取方法受较长边缘的影响较大,会造成不同程度上的选取偏差,从而

导致捆绑布局中较短边缘曲率过大的问题。因此本文进一步提出分级边缘绑定思想,有效改善了较短边缘的捆绑效果,平衡了整体与局部边缘的捆绑程度。

3 基于同源控制点的捆绑方法

基于同源控制点的捆绑方法的核心点就是边缘控制点的选取,通过相同控制点的控制,具有相似属性(如空间距离邻近、具有同一源点等)的边缘在弯曲时能够经过较为接近的路径,从而在图中形成捆绑束,减少复杂连接图中的边缘混乱现象。基于 Edge-Path 的方法通过寻找每条边缘上两个端点的最短加权路径,并以该路径上的顶点作为控制点,有效控制了边缘的聚合。因此本文利用该思想,通过将最短路径算法与分级思想相结合来选取同源控制点,并将其作为每条边缘的控制点对边缘进行聚合,最终获得了有效的边缘布局。本文方法首先需要计算同源控制点,然后使用最短路径算法为每条边缘选取合适的同源控制点作为边缘控制点,接着以 Bézier 曲线对边缘进行平滑处理,最后根据边缘方向为每条边缘着色。此外,为改善捆绑布局中较短边缘曲率较大的问题,本文结合分级思想对边缘绑定程度进行优化。3.1 节给出了同源控制点的计算方法;3.2 节详细介绍了如何结合同源控制点和最短路径算法计算边缘控制点;3.3 节具体阐述了如何根据边缘控制点对边缘进行渲染;3.4 节描述了分级边缘绑定的具体实现方式。

3.1 同源控制点计算

将从同一节点出发的边缘进行聚合可以在很大程度上减少边缘交叉的数量,因此本文为每个节点构建了一个“同源控制点”作为具有相同源点边的控制点之一,可以有效地将属于相同源点的边进行聚合。本文以具有相同源点边的目标点作为顶点构造多边形,由于各顶点位置可能差异较大,因此需要限制多边形的区域范围,避免边缘长度差距较大导致较短边缘产生较大曲率,从而造成视觉干扰。本文通过同源控制点权重因子 t 控制多边形顶点位置范围,计算式如式(1)所示:

$$D = \text{dist}(\text{src}, \text{crt}) - t \cdot \min(\text{dist}(\text{src}, \text{tgts})) \quad (1)$$

其中, $\text{dist}(\text{src}, \text{crt})$ 表示从源点 src 到当前目标点 crt 的欧氏距离, $\min(\text{dist}(\text{src}, \text{tgts}))$ 表示从源点出发到所有目标点 tgts 的最小欧氏距离。当 $D > 0$ 成立时,当前 crt 节点可作为多边形的顶点,参与同源控制点的计算,否则不予考虑。相比长边,短边弯曲效果明显且不易造成交叉混乱,能够展示源点处的流动方向的更多细节,故本文在对具有相同源点的边缘聚合时更倾向于短边的聚合,通过对最小欧氏距离加权来限制多边形的构建范围以达到短边优先聚合的目的。构建多边形后,本文以其重心作为同源控制点,可以更好地平衡各边缘的位置关系,重心的计算式如式(2)和式(3)所示:

$$C_x = (\sum_{i=1}^n C_{i_x} S_i) / \sum_{i=1}^n S_i \quad (2)$$

$$C_y = (\sum_{i=1}^n C_{i_y} S_i) / \sum_{i=1}^n S_i \quad (3)$$

其中, C_{i_x} 和 C_{i_y} 表示第 i 个三角形重心的坐标, S_i 表示第 i 个三角形的面积, n 表示多边形被划分为 n 个三角形。任意多边形均可划分为多个三角形,求得每个三角形的重心坐标和

面积后,可计算出相应的多边形重心坐标,即每个源点的同源控制点坐标。

3.2 边缘控制点计算

Edge-Path 捆绑方法使用 Dijkstra 算法计算边缘端点间的最短加权路径,以此路径上的节点作为控制点来控制边缘的弯曲。但是,我们在根据实际情况进行研究时发现,相同源点出发的边缘被更好地聚合有利于用户了解当前源点出发边缘的流向,从而判断当前源点的重要性与影响力。如图 1 所示,Edge-Path 捆绑方法将注意力集中在所有边缘的整体流向,忽略了各个源点出发边缘的局部流向,而在引入同源控制点后,具有相同源点的边缘显然更加聚集,同时边缘的汇合与分散更加清晰。因此,本文对 Edge-Path 捆绑方法的控制点计算方法做出了改进,以同源控制点作为边缘控制点,可以更好地促进同一源点出发边缘的聚合。

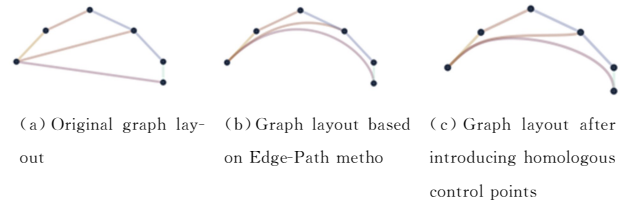


图 1 同源控制点引入前后对比

Fig. 1 Comparison before and after the introduction of homologous control points

算法 1 描述了本文对边缘控制点的计算过程。相比原文算法,本文新增了一个输入参数 t ,用于控制由目标点所构建的多边形生成范围,避免具有相同源点的边缘在聚合时产生较大的弧度。

算法 1 边缘控制点的计算

输入:图 $G=(V,E)$,图布局 D_G ,最大失真阈值 k ,边缘权重因子 d ,重心权重因子 t

输出:捆绑布局 Γ 的所有控制点

for $e \in E$ do

$\text{lock}(e) \leftarrow \text{False}$

$\text{skip}(e) \leftarrow \text{False}$

$\text{weight}(e) \leftarrow (D_G.\text{edgeLength}(e))^d$

$\text{sortedEdges} \leftarrow \text{sortDescending}(E, \text{weight})$

for $\text{node} \in \text{Node}$ do

$\text{sSControlPoints} \leftarrow \text{getsSControlPoint}(\text{node}, t)$

for $e \in \text{sortedEdges}$ do

 If $\text{lock}(e)$ then

 continue

$\text{skip}(e) \leftarrow \text{True}$

$s \leftarrow \text{source}(e)$

$t \leftarrow \text{target}(e)$

 //Dijkstra 算法包含从 G 中跳过的边

$p \leftarrow \text{dijkstraAlgorithm}(G, s, t, \text{weight}, \text{skip})$

 if $p == \text{null}$ then

$\text{skip}(e) \leftarrow \text{False}$

 continue

 if $p.\text{length}() > k * D_G.\text{edgeLength}(e)$ then

$\text{skip}(e) \leftarrow \text{False}$

```

continue
for m ∈ p do
    lock(m) ← True
controlPointsList(e) ← [s, sControlPoints(p), t]
controlPoints(e) ← p.getVertexCoords(controlPointsList)
return controlPoints

```

在算法 1 的描述中, *lock* 和 *skip* 是哈希集合, *lock*(*e*) 表示边缘 *e* 是否参与边缘绑定计算, *skip*(*e*) 表示边缘 *e* 是否参与最短加权路径计算。起初所有边缘都需要参与到计算中, 故需将以上两个集合中的值置为 *False*, 同时以各条边缘欧氏距离的 *d* 次幂作为边缘权值。由于边缘绑定需要的控制点都是由同源控制点组成, 因此需要先获取所有同源控制点, 其计算方式参见 3.1 节内容。获取同源控制点后需要为每条边缘生成对应的同源控制点列表, 为使用 Bézier 曲线进行平滑处理做准备, 生成方式选用 Dijkstra 算法。通过 Dijkstra 算法可以获取原图拓扑结构中最短路径上的所有顶点, 将所有顶点替换为同源控制点, 同时在首尾添加原图中的源点和目标点即可作为每条边缘的控制点。

3.3 边缘渲染

本文对需要捆绑的边缘使用 Bézier 曲线^[26]做平滑处理, 可以避免边缘出现锯齿状, 同时提高视觉效果和布局准确性。Bézier 曲线最初被用于汽车的主体设计, 现在已经被广泛用于各个领域中, 尤其是在二维平面的曲线绘制中占据着重要地位。Zhu 等^[27]根据数据分类结果, 利用 Bézier 曲线绘制部分螺旋线, 有效降低了螺旋线交叉而造成的视觉干扰问题。Bézier 曲线的形状由一组控制点控制, 如果将这组控制点按照顺序依次连接则可形成一个多边形, Bézier 曲线则会尽可能地接近这个多边形。通过控制多边形的顶点可以自由改变 Bézier 曲线的形状, 本文对 Bézier 曲线形状的控制是采用同源控制点。

对于需要捆绑的边缘, 可通过算法 1 获得 *n* 个控制点, 将其作为边缘控制点, 并以其作为 Bézier 曲线的控制点来生成较为平滑的曲线。实际进行绘制时, 本文采用 Bézier 曲线方程生成大量的点集, 将该点集中的点按照顺序依次连接则可以形成一条平滑的曲线。由于每条边缘的控制点数量并不统一, 因此本文通过 *n* 阶 Bézier 曲线方程来生成点集, 每个点的坐标如式(4)和式(5)所示:

$$B(t) = \sum_{i=0}^n C_n^i (1-t)^{n-i} t^i P_i, t \in [0, 1] \quad (4)$$

$$C_n^i = \frac{n!}{(n-i)! \cdot i!} \quad (5)$$

其中, P_i 表示控制点, 由于每条边的控制点数量较少, 通过直线绘制边缘可能出现锯齿状状态, 因此本文先在相邻控制点之间进行插值, 再利用上式生成较为平滑的曲线。

对边缘进行渲染时, 本文使用不同颜色区分来自不同方向的边缘。为防止由于图中颜色过多引起的视觉混乱, 本文只采用两组颜色分别表示边缘方向。如图 2 和图 3 所示, 其中红色和蓝色表示东西方向出发的边缘, 黄色和绿色表示南北方向出发的边缘。两组颜色的搭配使用清晰地显示了相反方向出发边缘的流向, 同时揭示了不同方向上的各个节点与

对应边缘的作用域和影响力。

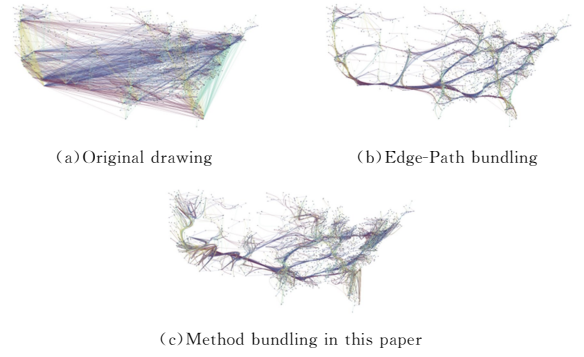


图 2 美国迁移数据图

Fig. 2 USA migration data map

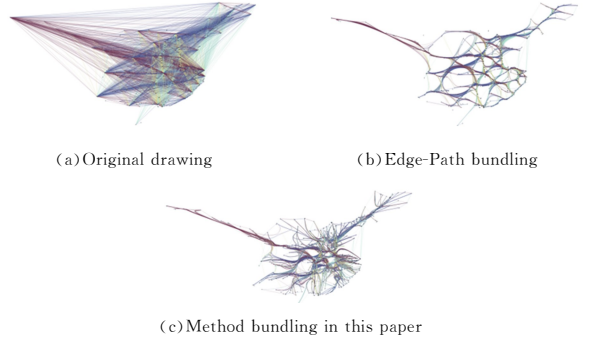


图 3 中国铁路线路图

Fig. 3 China railway route map

3.4 边缘分级绑定

由图 2(c)和图 3(c)可以发现, 基于上述方法的图布局中的边缘具有较大的曲率, 在增强视觉效果的同时会增加一定程度的视觉混乱, 因此本文进一步结合分级优化的思想, 通过对边缘进行分级绑定来优化边缘的弯曲效果, 以平衡局部视图对整体视图的影响。

在计算同源控制点时, 本文方法通过对长边进行限制以减小对较短边缘聚合的影响, 但是为了更好地聚合边缘以减轻布局的整体混乱效果, 通常需要缩小对长边的限制范围以增强对长边的绑定, 从而导致较短边缘产生过大曲率。由于失真阈值 *k* 只限制是否对较长边缘进行绑定, 无法解决本文存在的较短边缘曲率过大的问题, 因此本文舍弃了失真阈值对长边绑定的限制, 根据边缘长度对边缘进行分级绑定, 分级限制如式(6)所示:

$$\begin{cases} (crt_x - arc_x)^2 + (crt_y - arc_y)^2 < \alpha \|e\|, & \|e\| \geq \mu width \\ (crt_x - arc_x)^2 + (crt_y - arc_y)^2 < \beta \|e\|, & \|e\| < \mu width \end{cases} \quad (6)$$

其中, $arc_x = \frac{src_x + tgt_x}{2}$, $arc_y = \frac{src_y + tgt_y}{2}$, *width* 表示布局宽度, μ 表示边缘长度阈值, α 和 β 分别表示各级边缘的绑定权重。利用该式可将边缘划分为两个级别, 通过 α 和 β 可调节各级边缘控制点的使用范围, 从而控制边缘绑定后的曲率变化程度。

本文方法将边缘划分为两个级别, 同时使用圆弧区域来控制边缘控制点的使用范围, 实际应用时可根据数据集

的特点采用不同方式划分边缘级别和控制边缘控制点的使用范围。

4 实验结果和分析

本文数据集采用美国迁移数据集和中国铁路线路数据集。为比较边缘分级绑定的实验效果,本文将未对边缘进行分级绑定的方法称为 HCP 方法,将对边缘进行分级绑定的方法称为 HCPP 方法。HCP 方法涉及 3 个参数,包括最大失真阈值 k 、边缘权重因子 d 和同源控制点权重因子 t 。本文通过大量实验测试了不同参数组合下的边缘绑定效果,为平衡捆绑布局中存在的松散问题和过度绑定问题,最终选择 $k=2$, $d=2$ 和 $t=15$ 作为实验参数值。针对美国迁移数据集进行实验的结果如图 2 所示,针对中国铁路线路数据集进行实验的结果如图 3 所示,图中分别给出了两个数据集的原始图布局、基于 Edge-Path 方法的图布局和基于 HCP 方法的图布局。

为了解决同源控制点范围过大造成的较短边缘曲率过大问题,HCPP 方法利用分级思想对边缘绑定进行优化,并针对不同数据集设置了不同参数,具体参数值如表 1 所列,最终优化结果如图 4 所示。

表 1 各数据集参数设置
Table 1 Parameter setting of each dataset

	d	t	μ	α	β
美国迁移数据集	1.5	15	0.7	2	1
中国铁路线路数据集	2.0	5	0.7	1	3

表 2 基于 Edge-Path 方法和本文方法的定量指标值
Table 2 Quality metrics based on Edge-Path method and the proposed method

	美国迁移数据集								中国铁路线路数据集				
	ink_f	$dist^1$	$dist^2$	amb^1	amb^2	amb^3	amb^4	amb^5	ink_f	$dist^1$	$dist^2$	amb^1	amb^2
Straight-Line	1.00	1.00	1.00	0.69	0.24	0.06	0.02	0.01	1.00	1.00	1.00	0.69	0.01
Edge-Path	0.48	1.07	1.03	0.89	0.24	0.03	0.01	0.01	0.30	1.06	1.05	0.87	0.03
HCP	0.51	1.19	1.15	0.91	0.28	0.04	0.01	0.01	0.35	1.28	1.21	0.86	0.02
HCPP	0.55	1.05	1.02	0.86	0.28	0.06	0.02	0.01	0.52	1.05	1.03	0.85	0.02

实验结果显示,HCPP 方法在失真度和模糊性指标上均优于 Edge-Path 方法,但是在 IR 指标上处于劣势,HCP 方法则在各个指标上均处于劣势,这些结果与捆绑布局的视觉印象相匹配。与 Edge-Path 方法相比,HCP 方法为了更好地展示局部数据,较大程度地加强了较短边缘的捆绑力度,促使较短边缘产生较大曲率,故利用以上指标衡量则处于劣势。HCPP 方法通过分级捆绑边缘以降低边缘曲率,较好地改善了边缘的整体绑定程度,故具有更优的失真度值和模糊性值。虽然 HCPP 方法的 IR 指标不如 Edge-Path 方法,但是 IR 指标是通过比较边缘绑定前后布局相应的灰度图像所占有的像素数来比较捆绑布局对原始布局的视觉混乱降低效果,而在实际可视化时,较少的像素数会造成边缘过度绑定问题,从而带给用户错误的感知。在实际应用时,均衡的捆绑布局更有助于用户对数据进行探索和分析,故 HCPP 方法在实际应用时更具有优势。

4.2 用户实验分析

为比较 Edge-Path 方法和本文提出的方法对复杂连接图

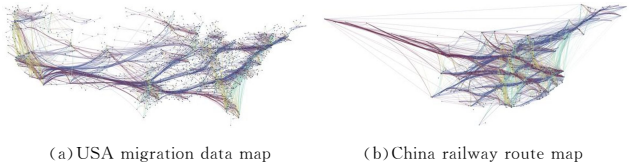


图 4 边缘优化结果
Fig. 4 Edge optimization results

为验证所提方法的有效性,本文分别从定量指标和用户实验两个方面对算法进行定量和定性分析。

4.1 定量指标分析

本文采用 3 个评价指标^[11]对本文方法进行定量分析,包括 Ink Reduction(IR)、失真度(Distortion)和模糊性(Ambiguity)。其中,IR 用捆绑前后图像的灰度位图图像所占像素数的比值来表示,用于衡量捆绑后视觉混乱的降低程度;失真度用捆绑前后两节点间连接线的长度的比值来表示,用于衡量被捆绑边缘的失真程度;模糊性用捆绑布局中节点暗示的错误邻居节点数和暗示的所有邻居节点数的比值来表示,用于衡量捆绑布局中被错误感知的连接边数量和严重程度。所有指标的具体计算方式可参考文献[11]。

对 Edge-Path 方法和本文方法所实现的可视化布局的 3 个评价指标进行计算,实验结果如表 2 所列。其中 $dist^1$ 和 $dist^2$ 分别为失真度的均值和中值。由于本文对 Edge-Path 方法的 IR 指标的计算数据明显优于其原论文所提供的数据,因此该指标数据以本文实验数据为准,而表 2 中关于 Edge-Path 方法的失真度和模糊性数据则以其原论文数据为准。

局部细节的描述能力以及本文算法在实际场景中的适用性,本文设立了以下任务用于用户研究。

任务一:指定任意节点,记录参与者对该节点流向的分析情况和分析时间。

任务二:指定任意两个节点,记录用户对两节点间最短路径的查找结果和查找时间。

任务三:调研用户对两种可视化方法的主观偏好。

本文中,用户实验涉及 20 名参与者,其中 10 名男性,10 名女性,年龄均在 20~40 岁之间,且对图论知识了解较少。将所有参与者按照性别和年龄等情况综合划分为两组,分别完成两种可视化方法对应的实验任务。实验开始前需要对所有参与者进行培训,确保他们已了解相关理论知识和具体实验任务。所有实验任务均通过可视化页面给出,参与者通过点击“开始”和“结束”按钮进行任务计时。两种可视化方法的所有任务均相同,其中,任务一和任务二的实验数据统计如图 5 和图 6 所示。

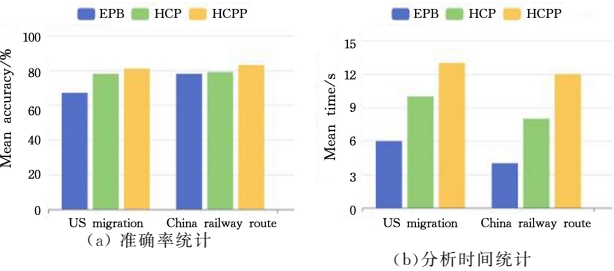


图 5 任务一用户实验数据统计

Fig. 5 User experiment data statistics in task 1

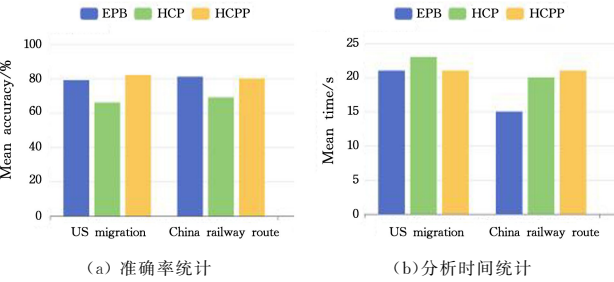


图 6 任务二用户实验数据统计

Fig. 6 User experiment data statistics in task 2

图 5 中的实验数据显示,相较于 Edge-Path 方法,通过 HCP 方法和 HCPP 方法实现的捆绑布局对节点流向分析的准确率更高。虽然参与者对 HCP 方法和 HCPP 方法所生成的布局的分析时间也相对更长,但是通过对参与者的调研发现,该分析时长在可接受范围内,故该部分实验数据证明了本文方法可有效揭示复杂连接图中边缘的局部流向。

图 6 中的实验数据显示,Edge-Path 方法和 HCPP 方法对最短路径的查找有相似的准确率和分析时间,但 HCP 方法表现较差。通过对比图 5 中的实验数据可以发现,HCP 方法由于更注重图形的局部细节,因此在整体布局上表现较差。经过曲率优化后的 HCPP 方法则在准确率上与 Edge-Path 方法不分伯仲,说明本文方法在不损失精准度的同时优化了图形的整体布局结构,大大提高了用户对于局部数据分析的准确度,也为实际场景中的数据分析任务提供了更有利的支持。

对于任务三,本文以访谈的形式对参与者进行调研。参与者表示,从实验结果的对比图中可以看到,本文方法不仅能够显示边缘整体流动趋势,还可以清晰地看见局部节点上的边缘流动趋势。以图 2 中迁移图的西南部分为例,基于 Edge-Path 捆绑方法得到的结果图只能看见迁移的整体趋势,如图 2(b)所示,各个节点大多位于捆绑束中,因此只能了解到该位置的迁移方向为东西方向。然而 HCP 方法得到的结果图中可以明显看见各个位置更详细的迁移方向,如图 2(c)所示,边缘弯曲的方向表示由该节点出发的边缘的主要去向,可以明显看到有一部分位置点倾向于长距离迁移,因此曲线具有较小的曲率,而另一部分位置点则倾向于近距离迁移,因此曲线会产生较大的曲率。图中给出了明显不同的视觉效果用以区分,具有实际意义。从图 4(a)可以看到,在对边缘进行分级绑定后,节点在整体布局中的流向更加清晰,有利于对节点的迁徙路径做进一步的跟踪。

从我国铁路线路数据集的结果图来看,基于本文方法所得到

的路线图,如图 3(c)和图 4(b)所示,可以看见从各个位置出发的线路如何汇聚到主干路,以及如何从主干路分散到各个目的地,而基于 Edge-Path 的方法(见图 3(b))只能显示各个地方的连通性,很难再获取到更细节的信息。实验证明了本文方法不仅能够显示复杂连接图中边缘的整体流向,而且能够提供更细节的局部边缘流向,为用户进行问题分析提供了较大的帮助。

结束语 针对 Edge-Path 边缘绑定方法存在的局部细节难以探索的问题,本文提出了一种新的捆绑方法,可以更有效地显示复杂连接图中的局部边缘流向。该方法通过构建源点多边形来获取同源控制点,并计算源点与目的点之间的最短加权路径,然后以同源控制点替代最短加权路径上的控制点以促进从同一源点出发的边缘更好地聚合,从而展示更详细的边缘流向。此外,为解决同源控制点的选取偏差所造成的较短边缘曲率较大的问题,本文进一步结合分级思想对边缘绑定效果进行优化,有效平衡了捆绑布局中整体与局部边缘的聚合程度。

从边缘的局部流向来看,本文方法能够较好地显示流向细节,但是从整体视觉效果来看,相比 Edge-Path 方法,本文方法会造成一定程度的视觉混乱。引起这部分视觉混乱的主要原因是局部捆绑较为松散,从图 4 可以看到,具有相同源点的边缘在源点附近聚合较慢,导致较多边缘会在邻近位置一起聚合。在未来的工作中,我们会研究相关的控制点计算方法来改善这一问题,提高边缘的聚合速度,同时应用一些交互式技术来辅助用户探索,挖掘更多底层信息。综上所述,基于同源控制点的捆绑方法能够有效地减少复杂连接图中的视觉混乱,展现数据的高级模式,同时有着更强的局部细节描述能力,是一种有效的边缘绑定方法。

参 考 文 献

[1] AUBER D. Tulip—a huge graph visualization framework,in Graph drawing software[M]. Cham:Springer,2004:105-126.

[2] ZHANG H. Design of visual art elements in a sustainable urban transportation system information platform[J]. Aggression and Violent Behavior,2021,12:101719.

[3] DODGE S,TOKA M,BAE C J. DynamoVis 1.0:An exploratory data visualization software for mapping movement in relation to internal and external factors[J]. Movement Ecology, 2021, 9(1):1-17.

[4] DODGE S,NOI E. Mapping trajectories and flows:facilitating a human-centered approach to movement data analytics[J]. Cartography and Geographic Information Science,2021,48(4):353-375.

[5] SCHÖTTLER S,YANG Y,PFISTER H,et al. Visualizing and interacting with geospatial networks:A survey and design space [J]. Computer Graphics Forum,2021,40(6):5-33.

[6] LYU Y,LIU X,CHEN H,et al. OD Morphing:Balancing simplicity with faithfulness for OD bundling[J]. IEEE Transactions on Visualization and Computer Graphics,2019,26(1):811-821.

[7] LI Y,LI T,DU S,et al. Clutter reduction of parallel coordinates based on an approximate measure of line crossing[C]// Data

Science and Knowledge Engineering for Sensing Decision Support: Proceedings of the 13th International FLINS Conference (FLINS 2018). 2018;492-499.

[8] 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.

[9] HOLTEN D,VAN WIJK J J. Force-directed edge bundling for graph visualization[J]. Computer Graphics Forum,2009,28(3):983-990.

[10] YANG H B,ZHOU H. A Distance-Based Edge-Bundling Method[J]. Journal of Graphics,2016,37(3):296-301.

[11] WALLINGER M,ARCHAMBAULT D,AUBER D,et al. Edge-Path Bundling: A Less Ambiguous Edge Bundling Approach[J]. IEEE Transactions on Visualization and Computer Graphics,2021,28(1):313-323.

[12] DICKERSON M,EPPSTEIN D,GOODRICH M T,et al. Confluent drawings: Visualizing non-planar diagrams in a planar way[C]// International Symposium on Graph Drawing. 2003:1-12.

[13] PHAN D,XIAO L,YEH R,et al. Flow map layout[C]// IEEE Symposium on Information Visualization. 2005:219-224.

[14] 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.

[15] ZHOU H,YUAN X,CUI W,et al. Energy-based hierarchical edge clustering of graphs[C]// 2008 IEEE Pacific Visualization Symposium. 2008:55-61.

[16] LAMBERT A,BOURQUI R,AUBER D. Winding roads: Routing edges into bundles[J]. Computer Graphics Forum,2010,29(3):853-862.

[17] LUO S J,LIU C L,CHEN B Y,et al. Ambiguity-free edge-bundling for interactive graph visualization[J]. IEEE Transactions on Visualization and Computer Graphics,2011,18(5):810-821.

[18] WU J,YU L,YU H. Texture-based edge bundling: A web-based approach for interactively visualizing large graphs[C]// 2015 IEEE International Conference on Big Data(Big Data). 2015:2501-2508.

[19] ZIELASKO D,WEYERS B,HENTSCHEL B,et al. Interactive 3D force-directed edge bundling[J]. Computer Graphics Forum,2016,35(3):51-60.

[20] WU B,CAO W Q. A Force-Directed Skeleton-Based Bundling with Clustering in Parallel Coordinates[J]. Journal of Computer-Aided Design & Computer Graphics,2017,29(10):1807-1815.

[21] NGUYEN Q,HONG S H,EADES P. TGI-EB: A new framework for edge bundling integrating topology, geometry and importance[C]// International Symposium on Graph Drawing. 2011:123-135.

[22] HURTER C,PUECHMOREL S,NICOL F,et al. Functional decomposition for bundled simplification of trail sets[J]. IEEE Transactions on Visualization and Computer Graphics,2017,24(1):500-510.

[23] EPPSTEIN D,HOLTEN D,Löffler M,et al. Strict confluent drawing[C]// International Symposium on Graph Drawing. 2013:352-363.

[24] BACH B,RICHE N H,HURTER C,et al. Towards unambiguous edge bundling: Investigating confluent drawings for network visualization[J]. IEEE Transactions on Visualization and Computer Graphics,2016,23(1):541-550.

[25] ZHENG J X,PAWAR S,GOODMAN D F. Further towards unambiguous edge bundling: Investigating power-confluent drawings for network visualization[J]. IEEE Transactions on Visualization and Computer Graphics,2019,27(3):2244-2249.

[26] WEI Z,DING S,WANG Y,et al. From river flow to spatial flow: Flow map via river flow directions assignment algorithm[J]. arXiv:2110.09395,2021.

[27] ZHU L X,LIU T R,TENG F,et al. Edge Bundling Method of Spiral Graph Based on Interval Classification[J]. Computer Science,2019,46(1):107-111.

LIU Meng-xin, born in 1996, postgraduate, is a member of China Computer Federation. Her main research interests include big data analysis and technology.

LI Tian-rui, born in 1969, Ph.D, professor, Ph.D supervisor, is a distinguished member of China Computer Federation. His main research interests include big data intelligence, rough sets and granular computing.

(责任编辑:何杨)