

海量图数据可视化研究*

伍勇, 钟志农, 景宁, 李星

(国防科学技术大学 电子科学与工程学院, 长沙 410073)

摘要: 图的可视化技术可以帮助人们更好地展现和探索数据的内部关系,然而随着数据规模的不断变大,传统的图的可视化技术已不适合海量数据的可视化要求。概括了传统的图的可视化技术,分类阐述了当前主流的海量图可视化布局算法的基本原理和特征,分析比较了相关的软件工具,并对海量图数据可视化技术的研究方向进行了展望。

关键词: 海量图数据; 可视化技术; 布局算法; 交互技术

中图分类号: TP391 文献标志码: A 文章编号: 1001-3695(2012)09-3216-05

doi:10.3969/j.issn.1001-3695.2012.09.004

Research on visualization of massive graph data

WU Yong, ZHONG Zhi-nong, JING Ning, LI Xing

(School of Electronic Science & Engineering, National University of Defense Technology, Changsha 410073, China)

Abstract: Graph visualization techniques can help people to demonstrate and explore the internal relations of the data. With the data size becoming larger and larger, traditional graph visualization techniques are not suitable for massive data visualization. This paper, firstly, summarized traditional graph visualization techniques, then expounded the basic principles and characteristics of visual layout algorithms for massive graph data, and analyzed some visualization software tools. Finally, this paper put forward some suggestions for future research work.

Key words: massive graph data; visualization technology; layout algorithm; interactive technology

0 引言

信息技术的广泛应用已经积累了大量的数据,如何对这些数据进行分析研究以提取有用的信息成为了迫切需要解决的问题,而图的可视化技术便是一种很有效的方法。Palmer 等人^[1]证明,图比其他可视化展现形式更适合探索数据的内部关系。图的可视化作为信息可视化的子领域,它继承了信息可视化的优点,可以将各类抽象的数据信息转换成图形信息,也可以利用各种可视化交互手段,使研究者能真实地观察他们对实际问题的模拟及处理结果。图是一种经典且强有力的展示数据及其内部关系的工具,它可以很好地帮助人们交流、理解现实世界和解决科学问题。随着数据规模的不断扩大,海量图数据的可视化技术成为了本领域研究的热点和难点。本文介绍了图的可视化相关技术,重点分类阐述了适用于海量图数据可视化的布局算法,分析比较了当前主流的海量图数据可视化的软件工具,对海量图数据的可视化技术进行了展望。

1 图的可视化技术

图数据可视化基本过程是:首先将原始数据表示成空间中相互连接的点,将数据之间的关系表示成边;然后依据一定的规则进行布局;最后允许研究者对点和边进行处理以帮助研究

者明白全局和局部的原始数据结构。根据图可视化的基本过程,可以将图的可视化技术分为布局算法和展示交互技术两大部分。

布局算法是图的可视化的核心和基础,对图数据的实时处理具有重要的影响。此外,融入了美学标准的布局算法在一定程度上也能够提高图的展现能力和易读性。国内外有不少学者对图的布局算法进行了分类研究,马里兰大学的 Shneiderman 教授等人^[2]将所有图可视化布局方法分为九类,但其分类存在交叠问题,即一种算法可以同时属于两个类别;而香港大学的 Cui^[3]则将其分为五类,不过她将层次数据的布局算法误包括进图数据的布局算法中。这两种分类的共同特点是都包含了力导引布局算法。力导引布局算法是文献描述最多且应用最广的算法,它能够充分展现图的整体结构及其自同构特征,通用性强,在图的布局算法中占据了主导地位。力导引算法又被称为“弹簧模型”^[4],最早由 Eades 于 1984 年提出,其基本思想是将图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后,弹簧弹力(引力和斥力)的作用会导致钢环的运动,直到系统总能量减少到最小值时停止。而后基于 Eades 的力导引基础算法改进的较为著名的算法有 Kamada 等人^[5]的 KK 算法、Davidson 等人^[6]的 DH 算法以及 Fruchterman 等人^[7]的 FR 算法,他们分别从理论依据、美学标准和展现能力对力导引基础算法进行了改进,使力导引布局算

收稿日期: 2012-03-18; 修回日期: 2012-05-08 基金项目: 国家“863”计划资助项目(2011AA120300); 湖南省自然科学基金资助项目(11JJ4028)

作者简介: 伍勇(1988-),男,湖南邵阳人,硕士研究生,主要研究方向为信息处理和数据挖掘(wuper@qq.com);钟志农(1975-),男,副教授,博士,主要研究方向为信息处理和数据挖掘;景宁(1963-),男,教授,博导,博士,主要研究方向为信息处理和数据挖掘;李星(1985-),男,硕士研究生,主要研究方向为信息处理和数据挖掘。

法日趋完善。国防科技大学的吴鹏等人^[8]将力导引算法进行了改进,提出了适用于社会网络分析与可视化的SAL算法,是一个力导引算法与社会网络分析技术结合的成功案例。不过传统力导引算法的计算复杂度太高,若以 n 表示节点数目,其算法时间复杂度大都为 $O(n^3)$,所能处理的图的规模低于1 000个节点。在Shneiderman教授和Cui的分类系统之外,黄竞伟等人开辟了用遗传算法布局无向图的新方法^[9-11],他们将无向图的布局问题转换为函数优化问题,用遗传算法求目标函数的最优解的近似值,从而达到布局的目的。其方法简单,思路清晰,而且目标函数反映了美学标准,因此布局效果非常好。不过,由于遗传算法本身的复杂度较高,他们的算法也不适用于大规模图的可视化。

图可视化技术的另一重要部分就是展示交互技术。如果说布局算法给了人们数据的形象化表达,那么展示交互技术则给了人们探索数据背后信息的手段。展示交互技术包括边的替换、节点聚类、抽样、缩放和移动、过滤、焦点+背景(focus+context)和动画等方法^[3]。

边的替换主要是指将直边换成曲边,将多条边换成一条边,以减少边的交叉,其改进效果如图1所示。

节点聚类则是根据一定的聚类规则将多个节点替换成一个节点,以体现主体部分,如图2所示。

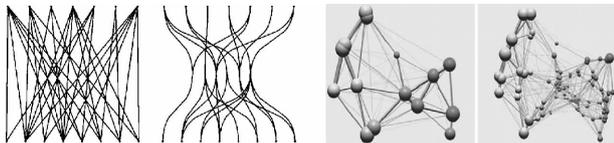


图1 边的替换

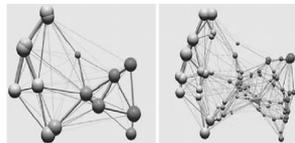


图2 节点聚类

抽样技术则是单纯地为了减少显示的数据量以达到提高显示效果的目的;缩放和移动主要是方便人们浏览图的各个部分;过滤则提供人们浏览特定数据集的权利。

焦点+背景技术则同时提供了细节信息和概要信息,其典型代表是鱼眼(fisheye view)视图,图3展示了原始图和鱼眼视图的效果。

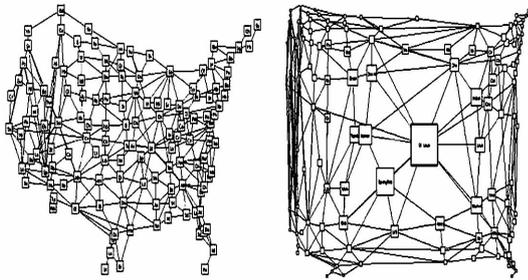


图3 原始图与鱼眼视图

动画是一项新的技术,它可以引导人们的探索过程,而其他交互技术的结合则会取得更好的效果。

图的可视化技术广泛应用于社会网络、互联网通信以及论文引证研究等领域。然而,今天的数据量已经变得非常庞大,传统的图的可视化技术已经很难揭示隐藏在数据背后的信息。一般而言,大规模数据将给图的可视化带来以下几个复杂的问题^[3]:

a)算法复杂。在某些情况下图的规模对算法的要求是非常苛刻的,因为很多图的算法都是NP难或者NP完全问题。因此,一些布局算法在面对拥有巨大数据量的图时根本无能为力。即使算法的时间复杂度可以容忍,长时间的处理过程也不

能满足交互实时性的需求。

b)展现混乱。当数据不断增加时,相应的图的展现将变得混乱并引起视觉困惑,令使用者区分节点和边将越来越困难。在极端情况下,数据量甚至比像素的数目还多。

c)易读性差。即使图的大小在屏幕可处理的能力范围内,怎么有效地将图蕴涵的信息表示成用户所要求的式样也是一个挑战。Ghoniem等人^[12]证明了对于简单如定位一个点或者找到两个点之间的边之类的任务,节点连线图的表现也是很差的,即使所使用的图只有20个节点。

d)迷航。人们很习惯阅读局部详细、全局概要的世界,怎么在大图中漫游而不迷失也是一个挑战。

上述四个问题实质上分别对布局算法和展示交互技术提出了更高的要求,第一个问题要求布局算法降低复杂度,而后三个问题可以通过综合运用展示交互手段予以解决。后文将从布局算法和软件工具两个方向阐述并比较目前为解决上述问题所做的研究工作。

2 适用于海量图数据的布局算法

海量图数据要求布局算法的计算复杂度足够低,通过总结分析当前国际上在海量图数据布局领域研究的成果,可将其分为改进传统力导引算法类、限定图的展现形式类和利用数据属性信息类三类。

2.1 改进传统力导引算法类

改进传统力导引算法主要是用多层级技术对传统力导引算法进行改进,为论述方便起见,后文将此类算法简称为改进类。

第一个能够产生超过1 000个节点良好布局的力导引算法是Hadany等人于2001年提出来的HH算法^[13]。该算法引入了多层级技术作为处理大图的一种方法,是图的可视化技术从处理小规模数据跨越到处理大规模数据的里程碑。多层级技术就是根据一定的分层标准将原始图分成多个展示层级,先绘制抽象度较高的展示层级,忽略图的细节,返回一个粗糙的只保留大体结构的布局;然后添加细节绘制抽象度较低的展示层级并修改布局。实施这个策略最关键的保障就是抽象分层技术能够保留图形的基本属性。Hadany等人建议使用边缘收缩(edge contractions)来计算图的层级序列,以保留图的重要属性。特别地,它的目标是保留聚类的大小、节点的度和图的同伦性这三种拓扑属性。对于高抽象展示层,序列中的每一个图的能量函数可以采用KK算法中的能量函数(即节点间的几何距离正比于节点之间的最小距离)。对于低抽象展示层再布局,他们建议使用Eades、FR、KK算法中三者之一的力导引模型。虽然这种算法的渐进复杂度与KK算法一样,但是分层级展示却可以保证在合理的时间内取得大规模图的良好布局。

Walshaw^[14]在2003年提出了另一种多层级算法。相比于基于KK型算法的细节再布局,该算法将FR算法的网格模型扩展成多层级算法。它的分级策略基于最大独立边集,而低抽象层的优化修正则是基于FR算法的力导引计算。该算法的时间复杂度是 $O(n^2)$,能够在30 s之内处理10 000个节点的图。

Gajer等人^[15]于2004年提出的GGK算法也是一个多层级力导引算法,并且他们还提出了很多新观点。其最重要的贡

献在于,通过一种更简单的分级策略帮助,GGK 算法避免了二次的时间和空间复杂度。相对于从给定大规模原始图中计算一系列的分层图,Gajer 等人创造了一个基于图中心理论和最大独立集的节点过滤器 $v: v_0 \supset v_1 \supset \dots \supset v_k \supset \emptyset$,其中 $v_0 = v(G)$ 表示原始图的节点数目, v_i 是 v_{i-1} 的最大独立子集。通过严格限制每级过滤器的节点数目,并且保证过滤器有 $O(\log n)$ 级,可以使得最终的算法时间复杂度为 $O(n \log^2 n)$ 。除了过滤器,Gajer 等人还介绍了在高维欧氏空间实现初步布局并最终得到二维和三维布局的方法。不过,GGK 算法依赖于基于图理论距离的智能初始分布,而不是随机分布。

为了进一步减少算法复杂度,Hu^[16] 在 2005 年提出的多层次力引导算法中引入了超节点的概念。在这个算法里,将一个节点与它远处一簇节点之间的斥力当做此节点与一个超节点间的斥力来计算,从而大大减少了计算量,使得总的计算复杂度降至 $O(n \log(n))$,已可以处理几十万节点规模的图。

2.2 限定图的展现形式类

限定图的展现形式,顾名思义,就是在对图数据进行布局之前,预先设定好图的最终展现形式,来一个节点布一个节点,不需要考虑节点的分布,从而快速有效地布局海量图数据。为论述方便,后文将此类简称为限定类。限定类算法有的保持着传统的节点连线图的形式,有的则作了相应的改变。

圆形布局法(circular layout)就是前者的代表。该算法的核心是将一个或者一组节点置于中心,与其相连接的点则按一定的顺序置于最邻近的圆周上,然后按这种策略布局余下的节点,直至所有的节点布局完毕^[17]。虽然圆形布局法的展现形式限定了,但是它的实现形式却多种多样,可以是一个中心一个圆周,也可以是一个中心多个圆周;可以只是单个圆周,也可以是多个同心圆周。一种最简单的实现形式是按节点进入布局的时间为序,将所有节点排布在一个圆周上。值得指出的是,圆形布局需要指定节点的排序依据,这个依据可以是人为指定的优先度(如时间等),也可以是数据本身的某个属性(如度、基因序列值等)。圆形布局可以扩展成为径向布局(radial axis layout)^[18],该方法首先对节点进行聚类,而后将聚类的代表节点依序排列成中心圆周,将其余的节点按聚类类别径向辐射排布。2010 年,Groeninger 基于开源软件 Gephi 实现了圆形布局算法和其变体径向布局算法^[19],其运算复杂度甚至达到了 $O(n)$ 。图 4 展示了这两种算法的布局效果。圆形布局算法及其变体径向布局算法对海量图数据的布局效率高,而且在社会网络分析和生物基因研究上有着突出的表现。但是,由于它需要指定节点顺序,而且只能将布局展现为圆形的形式,限制了它对数据其他结构的反应,因此通用性不强。

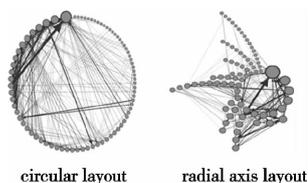


图 4 圆形布局和径向布局

另一种限定图的展示类算法是矩阵布局。相较于传统的节点连线图的形式,该算法将图数据表示成它们的邻接矩阵。首先将节点分别从纵向、横向依次排列一次,用坐标 (i, j) 上的符号表示从节点 i 到节点 j 的边,边的属性由符号的特征(如颜色、形状、大小等)表示。这种邻接矩阵表现形式克服了节点

连线图难以避免的节点重叠、边的交叉以及边通过无关点等问题,而且相对于圆形布局法,它也没有指定节点顺序的必要,因此它的布局效率更高,容易扩展。当数据量不断增大时,矩阵布局需要考虑的是节点的聚类以方便交互手段的加入和导航。Abello 等人^[20] 使用层次聚类机制来显示和导航大型矩阵,其层次地而不是顺序地显示矩阵数据,用户可以将它们当成树一样导航。矩阵布局在社会网络分析和知识引征分析方面应用广泛。法国国家信息与自动化研究所的 Elmqvist 等人^[21] 将矩阵布局法与各种交互手段结合起来,开发出的适用于海量图数据可视化分析的软件工具 ZAME 甚至能够处理整个法国维基百科的数据。不过,邻接矩阵形式的布局难以分辨节点的空间特征,区别聚簇和找寻路径上的节点等此类任务对于它来说无疑是巨大的挑战。

2.3 利用数据属性信息类

充分利用数据的属性信息,从特定的角度进行切入也是解决海量图数据可视化的一种方法。这种思想的实质有点类似于降维的方法,它们不是从整个数据的全部联系进行综合考虑,而是抓住数据的某一个属性来进行数据的可视化布局,因此,数据布局之间的位置差异也仅仅只是单一属性值的差异。这样,大规模数据的快速有效布局就成为了可能。为论述方便,后文将此类算法简称为利用类。

地图布局是反映这一思想的最佳代表。该算法以某幅地图作为背景,将数据按照它们的坐标信息布局在背景图上,将有关系的节点连接起来^[22]。该布局不需要任何布局策略,因此它的复杂度低、布局效率高。Jacomy 以 Gephi 为平台实现了基于地图布局的 GeoLayout 算法^[23],其计算复杂度只有 $O(n)$ 。地图布局的缺陷在于节点交叠和边的交叉现象比较严重,这种情况一方面可以通过交互技术克服,另一方面也可以对算法本身进行改进。香港大学的 Cui 等人^[24] 在布局算法中嵌入边的聚集和替换策略,使得地图布局更加清晰美观,图 5 展示了地图布局的原始效果和改进效果。地图布局算法特别适用于基于地图的海量图数据可视化应用,给人的直观感很强,易于人们理解。

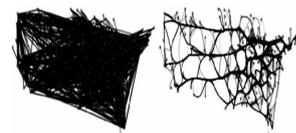


图 5 地图布局及其改进

充分利用属性信息类的另一个代表算法是时间布局法。该算法的表现形式有点类似于平行坐标法,其基本过程是提取节点的时间信息,将其以平行轴的形式依次从左至右排列,在每一个轴上按节点的其他属性由上至下排布节点,最后将有关系的节点用线连接起来。这种布局以节点的时间属性为基础,布局形式简单,处理效率高。Cammarano 等人^[25] 利用时间布局构造了论文引征网络,它的每一个轴代表出版年份,而轴上的文献节点则按引用次数排列,清晰地反映了一段时间内的论文引用情况。时间布局法对于展现图的时间信息是非常有效的,不过对于展现图的其他信息则稍显不力。另外,由于它的布局需要数据具有时间属性,因此通用性也不强。

层次布局法也可以看做是利用数据属性信息的一个例子。该方法首先根据节点的分类属性将屏幕划为几个区域,然后在每个区域中利用其他方法进行布局。一个典型的代表是

Shneiderman 等人^[2]设计的 NVSS 算法,该算法将数据的机构信息作为分类属性,然后在每类中对节点用时间布局法布局,最后连接有关系的节点。NVSS 有效地对诉讼案件网络进行了可视化,直观地揭示了诉讼案件引用网络的结构特点和隐含信息。层次布局法的布局过程简洁有序,由于分层的缘故,在每一层又可以利用不同的方法进行布局,因此灵活性较高。

2.4 各类算法的综合比较

通过上面的分析,表1从提出时间、类型、算法复杂度、适用范围以及能够处理的图的规模对现有适用于海量图数据可视化的代表算法和技术进行了比较。就处理图的规模而言,三类算法较之于传统算法对海量图数据的处理能力已大大提高。而就适用范围而言,由于改进类反映的是图数据内部的结构特征,布局过程不需要其他辅助因素,因此适用范围最广。不过,要将改进的力引导算法运用到具体问题中还需要对其进一步调整,以反映解决问题需要的图的结构特征。限定类限定了图的展现形式,相较于改进类,它的适用范围窄些。而利用类依赖数据的属性信息更严重一些,因此适用范围更窄。而就复杂度而言,改进类仍然比不上后两者。三类方法在处理海量图数据可视化方面都有自己的特点,如何结合这些特点创造更适合海量图数据可视化的技术方法则是今后努力的方向。

表1 部分算法比较

算法名称	提出时间	类型	复杂度	适用范围	图的规模
HH	2001	改进类	$\approx O(n^3)$	广泛	1~1000 节点
Wal	2003	改进类	$O(n^2)$	广泛	1~10000 节点
G GK	2004	改进类	$O(n \log^2 n)$	广泛	1~100000 节点
Hu Yi-fan	2005	改进类	$O(n \log(n))$	广泛	1~500000 节点
Circular	2010	限定类	$O(n)$	一般	1~1000000 节点
Radial axis	2011	限定类	$O(n)$	一般	1~1000000 节点
GeoLayout	2010	利用类	$O(n)$	较窄	1~1000000 节点

3 软件工具

如果说布局算法的研究进展在于降低算法复杂度这一方面,那么展示交互技术的成果则体现于软件工具上。一个好的可视化系统需要在性能、人机交互的丰富程度、支持的算法数量、支持算法扩展与否以及输入输出等方面有所体现。通过对几款软件系统具体分析,表2总结了其主要信息。

表2 部分软件工具信息

比较项	UCINET	Graphviz	Gephi	Tulip	Cytoscape
性能与效率	较高	一般	一般	较高	一般
交互方式与展现能力	一般	一般	好	较好	较好
支持的算法数量	较多	较多	一般	较多	较多
是否支持算法扩展	否	是	是	是	是
输入格式	Excel 等	.dot 等	.dot 等	.tulp,.dot 等	XGMLL 等
输出格式	Excel 等	.gif,.png 等	.png 等	.tulp,.gml 等	XGMLL 等
运行平台	Windows	Linux,Mac,Windows	任何支持 Java1.6 和 OpenGL 的系统	Windows, Linux, Mac OS	任何支持 Java 的系统
开发语言	Java	C	Java	C++	Java

UCINET^[26]是一款经典的功能全面的复杂图分析工具,主要用于分析社群网络,是社群网络分析领域最著名同时也是最为常用的一款网络分析软件包。相对于易用性来说,UCINET 更注重的是速度,因此它的交互界面和交互手段都不尽如人

意,不过与其他可视化工具(如 Pajek、Mage、NetDraw 等)的集成在一定程度上弥补了这方面的缺陷。UCINET 支持多种数据输入输出方式,支持的布局算法和网络分析算法也比较多。不过由于它出现得比较早,所以支持的算法不够新,也不支持算法扩展,它能处理的节点上限为 32 767。

Graphviz^[27]是一款来自 AT&T 研究室的开源作图工具,它支持多种作图算法,并同时支持 Web 界面和交互式图形界面,拥有一系列的辅助工具和类库。Graphviz 利用一种简单的文本语言来描述图形结构,并能够将生成的图形保存成各种格式。此外,它还支持一种表示图的 XML 语言——GXL。

Gephi^[28]是一款应用于各种网络和复杂系统的交互式可视化探索平台,也是一个方便人们探索和理解图的工具。丰富的交互手段是它的特色和优势,它不仅可以通过改变图的各个要素的结构、形状和颜色去显示隐藏在数据背后的信息,而且也拥有一般可视化工具所不具备的动画能力。Gephi 可以通过增加插件来支持算法的扩展,因此它能够支持最新的算法;同时,它还使用三维引擎来实时显示大图网络和加速数据探索过程。它灵活的多任务结构使得处理复杂数据集和产生有价值的可视化结果成为可能。

Tulip^[29]是一个用于关系数据的可视化分析框架,它的目的是为开发者提供完整的库,支持关系数据的信息可视化交互设计。该框架是用 C++ 语言开发的,可以进一步完善算法、交互技术、数据模型和特殊领域可视化。Tulip 允许组件的重用,使得这个框架可以有效地支持原型系统的研究和最终用户软件的设计开发。

Cytoscape^[30]是一个用于复杂网络数据融合、分析和可视化的开源平台,它本来的目的是用于生物信息研究,现在则扩展到一般图数据的可视分析。Cytoscape 的核心提供了一个数据融合和可视化的功能,其他的功能可以通过插件得到。它能够处理 100 000 以上节点和边的图,具备良好的交互手段。

4 结束语

在数据呈爆炸式增长的今天,海量图数据可视化的研究显得尤为急迫。接下来的海量图数据可视化研究可以从下三个方面进行:a)深挖布局算法的本质,与最新技术相结合,提出新的布局算法,以提高海量图数据可视化的规模和速度;b)从具体问题入手,限制图可视化的难度,从特定的方向实现海量图数据的可视化;c)综合运用各种交互手段,提高可视化的展现能力,使其能够实用化。

参考文献:

- [1] PALMER S, ROCK I. Rethinking perceptual organization; the role of uniform connectedness [J]. *Psychonomic Bulletin & Review*, 1994, 1(1): 29-55.
- [2] SHNEIDERMAN B, ARIS A. Network visualization by semantic substrates [J]. *IEEE Trans on Visualization and Computer Graphics*, 2006, 12(5): 733-740.
- [3] CUI Wei-wei. A survey on graph visualization [D]. Hong Kong: Hong Kong University of Science and Technology, 2007.
- [4] EADES P. A heuristic for graph drawing [J]. *Congressus numerantium*, 1984, 42: 149-160.
- [5] KAMADA T, KAWAI S. An algorithm for drawing general undirected graphs [J]. *Information Processing Letters*, 1989, 31(1): 7-15.

- [6] DAVIDSON R, HAREL D. Drawing graphs nicely using simulated annealing[J]. *ACM Trans on Graphics*, 1996, 15(4):301-331.
- [7] FRUCHTMANER T, REINGOLD E M. Graph drawing by force-directed placement[J]. *Software-Practice and Experience*, 1991, 21(11):1129-1164.
- [8] 吴鹏,李思坤. 适于社会网络结构分析与可视化的布局算法[J]. *软件学报*, 2011, 22(10):2467-2475.
- [9] 黄竞伟,康立山. 一个新的无向图画图算法[J]. *软件学报*, 2000, 11(1):138-142.
- [10] 张清国,黄竞伟. 一个无向平面图图画图算法[J]. *小型微型计算机系统*, 2003, 24(6):972-975.
- [11] 张清国,叶俊明,张维,等. 用遗传算法画无向图[J]. *计算机工程与科学*, 2006, 28(6):58-61.
- [12] GHONIEM M, FEKETE J D, CASTAGLIOLA P. On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis[J]. *Information Visualization*, 2005, 4(2):114-135.
- [13] HADANY R, HAREL D. A multi-scale algorithm for drawing graphs nicely[J]. *Discrete Applied Mathematics*, 2001, 113(1):3-21.
- [14] WALSHAW C. A multilevel algorithm for force-directed graph drawing[J]. *Journal of Graph Algorithms and Applications*, 2003, 7(3):253-285.
- [15] GAJER P, GOODRICH M T, KOBOUROV S G. A fast multi-dimensional algorithm for drawing large graphs[J]. *Computational Geometry: Theory and Applications*, 2004, 29(1):3-18.
- [16] HU Yi-fan. Efficient and high quality force-directed graph drawing[J]. *The Mathematica Journal*, 2005, 10(1):37-71.
- [17] BREITKREUTZ B J, STARK C, TYERS M. Osprey: a network visualization system[J]. *Genome Biology*, 2003, 4(3):22-24.
- [18] TOLLIS I G, Di BATTISTA G, EADES P, *et al.* Graph drawing: algorithms for the visualization of graphs[M]. Upper Saddle River: Prentice Hall PTR, 1998.
- [19] Circular layout[EB/OL]. [2012-03-04]. <http://www.gephi.org/plugins/circular-layout>.
- [20] ABELLO J, Van HAM F. Matrix zoom: a visual interface to semi-external graphs[C]//Proc of IEEE Symposium on Information Visualization. Washington DC:IEEE Computer Society, 2004:183-190.
- [21] ELMQVIST N, Do T N, GOODELL H, *et al.* ZAME: interactive large-scale graph visualization[C]//Proc of IEEE Pacific Visualization Symposium. 2008:215-222.
- [22] BECKER R A, EICK S G, WILKS A R. Visualizing network data[J]. *IEEE Trans on Visualization and Computer Graphics*, 1995, 1(1):16-28.
- [23] Map geocoded data with Gephi[EB/OL]. (2010-05-17)[2012-03-04]. <http://www.gephi.org/2010/map-geocoded-data-with-gephi>.
- [24] CUI Wei-wei, ZHOU Hong, QU Hua-min, *et al.* Geometry-based edge clustering for graph visualization[J]. *IEEE Trans on Visualization and Computer Graphics*, 2008, 14(6):1277-1284.
- [25] CAMMARANO M, DONG X, CHAN B, *et al.* Visualization of heterogeneous data[J]. *IEEE trans on Visualization and Computer Graphics*, 2007, 13(6):1200-1207.
- [26] UCINET[EB/OL]. (2010-10-21)[2012-03-04]. <http://www.analytitech.com/ucinet>.
- [27] Graphviz[EB/OL]. [2012-03-04]. <http://www.graphviz.org>.
- [28] Gephi[EB/OL]. [2012-03-04]. <http://www.gephi.org>.
- [29] Tulip[EB/OL]. [2012-03-04]. <http://www.tulip-software.org>.
- [30] Cytoscape[EB/OL]. [2012-03-04]. <http://www.cytoscape.org>.
- [31] EADES P, FENG Qing-wen. Multilevel visualization of clustered graphs[C]//Proc of the 4th Symposium on Graph Drawing. London: Springer-Verlag, 1996:101-112.
- [32] HAREL D, KOREN Y. A fast multi-scale method for drawing large graphs[J]. *Journal of Graph Algorithms and Applications*, 2002, 6(3):179-202.
- [33] QUIGLEY A, EADES P. FADE: graph drawing, clustering, and visual abstraction[C]//Proc of the 8th Symposium on Graph Drawing. London: Springer-Verlag, 2000:197-210.
- [34] CHAN D S, CHUA K S, LECKIE C, *et al.* Visualisation of power-law network topologies[C]//Proc of the 11th IEEE International Conference on Networks. 2003:69-74.
- [35] HAREL D, KOREN Y. Graph drawing by high-dimensional embedding[C]//Proc of the 10th International Symposium on Graph Drawing. London: Springer-Verlag, 2002:207-219.