

家谱成员关系的元图可视化^{*}

陈国军¹, 蔺想红¹, 段淑凤²

(1. 西北师范大学计算机科学与工程学院, 甘肃 兰州 730070;

2. 石家庄铁道大学信息科学与技术学院, 河北 石家庄 050043)

摘 要:电子家谱用于统计与分析家族成员的情况与关系,其可视化形式是家谱信息化的重要组成部分。由于边的数量随节点数的增加而快速增长,在已有家谱可视化方法中会出现显示效率急速下降问题。针对这一问题,结合元图进行可视化优化,减少不必要存在的边。此外,结合实际情况提出“父母-独生子女”关系等可视化表示,对可视化形式进一步优化。实验表明,在相同家谱数据情况下,优化的家谱元图可视化方法中所需边的数量约为原有元图形式的 1/2,小于通常可视化形式的 1/4,有效提高了家谱可视化效率。同时运用节点图的树形层次结构对家谱进行可视化显示,增强了可视化效果的实用性。

关键词:家谱;系统模型;元图;可视化;层次结构;成员关系

中图分类号:TP393.09

文献标志码:A

doi:10.3969/j.issn.1007-130X.2016.07.023

Visualization of relationship of genealogical members using metagraph

CHEN Guo-jun¹, LIN Xiang-hong¹, DUAN Shu-feng²

(1. College of Computer Science and Engineering, Northwest Normal University, Lanzhou 730070;

2. School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang 050043, China)

Abstract: Electronic genealogy, whose visualization is an important constituent part of genealogical informatization, is used in statistics and analysis of statuses and relationships of members. However, a rapid decline of performance efficiency appears in the existing methods of genealogical visualization due to the rapid increment of edges. We make an optimization for the question by reducing unnecessary edges through metagraph. Besides, the visualization representatives of the “parents-one child” and other relationships are proposed and the visualization forms are optimized further. Experimental results show that the number of the edges in the genealogical metagraphic visualization form is almost half of the original metagraph and less than a quarter of common forms in cases of same genealogical data, and the visualization effect can be significantly improved. The tree-hierarchy of node graph architecture is used to enhance the practicability of visualization in visualization of genealogy.

Key words: genealogy; system model; metagraph; visualization; hierarchy; member relationship

1 引言

家谱作为记述血缘集团世系的载体是维系宗族血缘关系的重要纽带^[1],是一种通过文表形式记

载以血缘关系为连接的家族世系繁衍的重要资料,其与正史、方志并列是中华民族历史研究资料的三大支柱之一。家谱不仅在学术上对开展人文社科类研究有着极其重要的价值,并且对海内外华人华侨寻根问祖有着重要意义,增强了民族认同感与凝

^{*} 收稿日期:2015-07-13;修回日期:2015-09-11

基金项目:国家自然科学基金(61165002,61363058);甘肃省自然科学基金(1010RJZA019)

通信地址:730070 甘肃省兰州市西北师范大学计算机科学与工程学院

Address: College of Computer Science and Engineering, Northwest Normal University, Lanzhou 730070, Gansu, P. R. China

聚力。家谱的存在与流传也同样增强了普通民众的家族归属感,增进了家族内成员之间的感情。随着民众生活水平的提高,家谱愈加得到人们的重视。传统的家谱多书写于纸上,随着科技的进步,信息化时代的电子家谱应运而生。电子家谱中,家谱数据的组成与子代节点和父代节点组成家谱树的特点对应,通常采用树形结构。家谱中前一辈与后一辈之间的关系就是数据结构类型中树的根节点与叶子节点之间的关系。一般的架构方式有兄弟关系架构、夫妻关系架构以及父子关系架构。在家谱中,每个家族成员是最基本的组成部分,是不可再分解的最小组成单元,因此在电子家谱中选定家族成员作为数据的基本单元^[2]。电子家谱涉及家谱数据操作、家谱连接、家谱可视化等众多子系统,其中的子系统又都包含众多要素和下一级子系统,形成了规模庞大的多层次结构。国内外研究人员在家谱系统研究中做了卓有成效的工作,并在实际应用中得到了证实。我们基于实际应用,对现有的各类电子家谱进行了分析,发现其中大多数支持创建和维护一些小型家谱数据库,在大规模家谱应用中存在对家谱可视化能力较弱等问题。由于家谱可视化所使用的方法通常为基于节点-连线的图表形式,因此其可读性会随节点数量的增加而急剧下降^[3]。

在实际使用中,家谱系统可视化展示一般采用树形结构,与家谱中每个成员的子代节点和父代节点分别组成家谱树对应。除基本树形结构外,一些学者对可视化形式进行了优化改进,提出许多优化的树形展示形式,如 MultiTree^[4]、ZoomTree^[5]、DualTrees^[6]、H-Tree^[7]、SpaceTree^[8]等。除此之外,还有多种其它展示形式陆续提出:放射状布局,如 Draper 等^[9]提出的交互式 Fan charts 的展示形式、Kim 等^[10]提出的家谱图 TimeNets 展示形式、Bezerianos 等^[11]提出的基于矩阵的家谱可视化展示方式,以及 Johnson 等^[12]提出的 Tree maps 结构;3D 的展示方式,如 Loh 等^[13]提出 Celestial3D 的应用等。虽然方法众多,但在上述家谱数据可视化形式中,当家族人员数量增多时,图中节点数量增加,从而导致边的数量大斜率线性增加,极大程度上增加了家谱可视化的复杂度。

刘军丹等^[14]针对上述边的数量复杂度问题,提出以元图形式对家谱进行可视化展示,将夫妻关系作为组成元图边的入点集所需满足的二元关系,运用非循环元图的性质,没有表示夫妻关系的无向边,并将表示“父子(女)”“母子(女)”关系的边合并

为表示“父母子(女)”的边,这极大降低了图中边的数量。但是,通过对元图数据与可视化视图还原分析,发现其元图可视化方法没有最大程度上满足家谱的实际需求。

针对上述问题,本文提出一种优化的家谱元图可视化方法,更加简洁地表示出家谱成员关系,并给出家谱数据可视化展示形式的数据比较,同时运用节点连接图的树形层次结构实例化应用,使得展示结果清晰易读,更具实用性。

2 家谱的元图可视化

2.1 家谱的元图表示

Basu 等^[15,16]在普通图基础上将图理论推广,提出“元图”这一新型结构。元图是对有向图和超图的结合和扩充,与其他图相比最主要的区别在于元图描述的是集合与集合之间的关系,而非单个元素之间的有向关系,在图形可视化和形式化分析方面具有很强的能力。目前,对元图的理论研究主要集中在简单路径、元路径、最大元路径、连通性、同构、子图等方面^[15,17]。在信息系统中,元图已被广泛应用于事务处理、决策支持^[18]和工作流分析^[19]等系统分析与建模中,同时在企业组织结构、供应链分析^[20]和专家系统^[21]的构造中也都有实际应用。

定义 1(元图) 元图 S 是定义在产生集 $M = \{m_i, i = 1, 2, \dots, I\}$ 上的一个序列对,即给定的元图 $S = \langle M, E \rangle$ 。其中 $E = \{e_k, k = 1, 2, \dots, K\}$ 表示元图 S 中所有边组成的集合, $e_k = \langle V_k, W_k \rangle$ 为满足对于任意的 $e_k \in E$ 是由产生集 M 的两子集 $V_k, W_k \in M$ (其中 $V_k \cup W_k \neq \emptyset$) 组成的有序对,称 V_k 为 e_k 的入点集(Invertex), W_k 为 e_k 的出点集(Outvertex)。

亲属关系通常由四种途径产生:人口出生产生的血缘关系、婚姻产生的姻亲关系、法律认定的亲属关系(如领养、收养)以及相互认同的亲属关系(如异姓兄弟姐妹)。以上四种亲属关系可以归纳为两类:一类为“父辈-子辈”之间的关系,另一类为同辈之间的关系,由于兄弟姐妹之间的同辈关系由父辈与子辈关系决定的,因此同辈之间的关系在此主要指由于婚姻形成的“夫妻”关系。通过这两类关系,家谱关系网能够完整建立起来。

定义 2(夫妻关系) 给定家谱所有成员信息组成集合 $M = \{M_i, i = 1, 2, 3, \dots, I\}$, 定义 P 为集

合 M 上的夫妻关系,则集合中具有夫妻关系的成员 $m_i, m_j \in M$ 可表示为 $m_i P m_j$ 。

元图中集合之间相连的边表示 V_k 与 W_k 之间的关系,在家谱元图可视化中,元图中的边表示成员间的“父辈-子辈”关系。具有“夫妻”关系的成员在元图表示中被看作一个整体,作为一个节点被表示出来。因此,在家谱的元图可视化中,对于任意的两个家庭成员,根据是否存在“父辈-子辈”关系确定所在集合间是否存在有向边,根据是否为“夫妻”关系确定有向边的入点集^[14]。假设家族成员总数为 a ,具有夫妻关系的对数为 b ,无父母信息显示的成员数为 c ,规定具有配偶的成员可拥有孩子。家谱元图可视化表示中,边的个数等于家谱中孩子节点出现的个数,也就是家族成员总数与只作为父辈节点总数之差,记为:

$$S.|E| = a - c \quad (1)$$

与普遍使用的树形可视化方法比较,元图可视化形式层次鲜明,成员关系清晰。

2.2 优化的元图表示

2.2.1 父母与未婚、已婚子女关系表示的优化

使用元图对家谱进行可视化展示有效地减少了图中边的数量,降低了观察和分析家谱数据的复杂性,提高了可读性。但是,我们通过对元图数据与可视化视图还原分析,发现现有的元图可视化方法所展示的对比结果没有最大程度上满足家谱的实际需求。本节结合实际情况在元图可视化表示基础上进行了改进。

一般情况下,两个家庭之间通过婚姻产生关系,一个家谱图中的关键节点是“夫妻”节点。因此,在家谱中根据是否结婚把子女成员(有父母)分为未婚子女和已婚子女两种类型。于是,对所有成员的“父母子(女)”关系可以用如下形式表示:

定义 3(父母-未婚子女映射) 给定家谱所有成员信息组成集合 $M = \{M_i, i = 1, 2, 3, \dots, I\}$,定义到 M 上的映射 $\varphi: W(M) \rightarrow M$ 表示“父母子(女)”关系。若对于任意的 $\{m_i, m_j\} \in W(M)$,存在 $m_k \in M$,使得 $\varphi(\{m_i, m_j\}) = m_k$ 成立,称 m_k 为 m_i, m_j 的一个子女。当 m_k 同时不具有 P 关系时,称 m_k 为家谱中的一个未婚子女。

性质 1 若在家谱的元图表示形式 $S = \langle M, E \rangle$ 中, m_k 为家谱成员中的未婚子女,则在家谱元图可视化中 m_k 无需通过边与其他成员连接来表示。

证明 设元图 $S = \langle M, E \rangle$ 表示的是家谱图,未

婚子女 m_k 在家谱成员中只与其父母 m_i, m_j 由于血缘关系连接在一起。而在家谱元图可视化中具有“夫妻”关系的成员被看作一个整体并作为一个节点表示出来,因此 m_k 只与 m_i, m_j 所在节点产生关联,其所在家谱图中的关系仅由 m_i, m_j 节点决定,从而 m_k 可作为 m_i, m_j 所在节点的附属部分显示在原图可视化中,而 m_i, m_j 为节点的主要组成部分。故在家谱元图可视化中未婚子女 m_k 无需通过边与其他成员连接来表示。□

定义 4(父母-已婚子女映射) 给定家谱所有成员信息组成集合 $M = \{M_i, i = 1, 2, 3, \dots, I\}$,定义到 M 上的映射 $\varphi: W(M) \rightarrow M$ 表示“父母子(女)”关系。若对于任意的 $\{m_i, m_j\} \in W(M)$,存在 $m_k \in M$,使得 $\varphi(\{m_i, m_j\}) = m_k$ 成立,称 m_k 为 m_i, m_j 的一个子女。当 m_k 同时与其他成员具有 P 关系时,称 m_k 为家谱中的一个已婚子女。

性质 2 若在家谱的元图表示形式 $S = \langle M, E \rangle$ 中, m_k 为家谱成员中具有父母信息的一个已婚子女,即拥有 P 关系,则在家谱元图可视化中 m_k 至少需要一条边与其他成员连接进行表示。

证明 设元图 $S = \langle M, E \rangle$ 表示的是家谱图,在家谱元图可视化中具有“夫妻”关系的成员被看作一个整体并且作为一个节点被表示出来,因此已婚子女 m_k 应与其具有 P 关系的家谱成员组成一个节点。而同时由于 m_k 作为 m_i, m_j 的子女,本应为 m_i, m_j 节点附属部分,故 m_k 需要通过一条边与 m_i, m_j 节点连接来表示。因此,已婚子女 m_k 至少需要一条边与其他成员连接进行表示。□

性质 3 若在家谱的元图表示形式 $S = \langle M, E \rangle$ 中, m_{k1} 与 m_{k2} 为家谱成员中具有父母信息的形成 P 关系的子女,即 $m_{k1} P m_{k2}$,其父母分别为 m_{i1}, m_{j1} 和 m_{i2}, m_{j2} 、如果 m_{i1}, m_{j1} 和 m_{i2}, m_{j2} 除 m_{k1} 与 m_{k2} 外均无其他已婚子女,即 m_{k1} 与 m_{k2} 均为唯一已婚子女,则 m_{k1} 与 m_{k2} 在家谱元图可视化中共需一条边与其他成员连接表示。

证明 设元图 $S = \langle M, E \rangle$ 表示的是家谱图, m_{k1} 与 m_{k2} 为具有夫妻关系并拥有父母 m_{i1}, m_{j1} 和 m_{i2}, m_{j2} 信息的家庭成员,即 $m_{k1}, m_{k2} \in M$ 且 $m_{k1} P m_{k2}$ 。由定义 4、性质 2 可知,一名已婚子女至少需要一条边与其他成员连接表示,因而 m_{k1}, m_{k2} 需要两条边与其他成员连接表示。由于与 m_{k1}, m_{k2} 具有相同父母节点的成员都是未婚子女,由定义 3 及其性质 1 知这些成员作为 m_{i1}, m_{j1} 或 m_{i2}, m_{j2} 的附属部分无需边与其他成员连接表示。因此,在 m_{i1}, m_{j1} 和 m_{i2}, m_{j2} 所有的子女成员中只有 m_{k1} 与

m_{k2} 一对夫妻节点,与 m_{i1} 、 m_{j1} 和 m_{i2} 、 m_{j2} 产生一一对应关系。故 m_{k1} 与 m_{k2} 节点需要一条边与 m_{i1} 、 m_{j1} 和 m_{i2} 、 m_{j2} 节点连接,在家谱元图可视化中 m_{k1} 与 m_{k2} 节点共需一条边与其他成员连接表示。□

假设家族成员总数为 a ,具有夫妻关系的对数为 b ,无父母信息显示的成员数为 c ,规定具有配偶的成员可拥有孩子。由定义 3、定义 4 及其性质 1 和性质 2 可知,此时家谱元图可视化表示中边的个数小于家谱中拥有父母信息显示且具有夫妻关系的成员个数,也就是有父母信息显示且具有夫妻关系的对数的二倍与无父母信息显示的具有夫妻关系的成员个数之差,记为:

$$I.|E| = 2b - c \quad (2)$$

2.2.2 父母与独生子女关系表示的优化

20 世纪 70 年代起中国开始全面实行计划生育,1979 年以后中国开始实行独生子女政策。由于我国多年实行计划生育政策,独生子女数量已经在全国人口中占据了相当大的比重。根据 2010 年第六次全国人口普查主要数据公报显示,我国家庭的平均人口数为 3.10 人。人口学专家易富贤根据最新人口普查数据推断中国现有约 2.18 亿独生子女^[22,23]。与此同时,虽然我国人口具有潜在老龄化危险,但目前大多数家庭迫于经济压力和时间等原因,在很长一段时间内都会保持独生子女家庭或者不计划进行生育。据《中国 2010 年人口普查资料》^[24]显示,中国 2010 年总和生育率仅有 1.18。

定义 5(父母-独生子女映射) 给定家谱所有成员信息组成集合 $M = \{M_i, i = 1, 2, 3, \dots, I\}$,定义到 M 上的映射 $\varphi: W(M) \rightarrow M$ 表示“父母子(女)”关系。若对于任意的 $\{x_i, x_j\} \in W(M)$,存在 $m_k \in M$,使得 $\varphi(\{x_i, x_j\}) = m_k$ 成立,称 m_k 为 m_i 、 m_j 的一个子女。当且仅当 m_k 唯一时,称 m_k 为 m_i 、 m_j 的独生子女。

性质 4 若在家谱的元图表示形式 $S = \langle M, E \rangle$ 中, m_{k1} 、 $m_{k2} \in M$, m_{k1} 和 m_{k2} 为具有父母信息的独生子女,且 $m_{k1} P m_{k2}$,则 m_{k1} 与 m_{k2} 在家谱元图可视化中共仅需一条边与其他成员连接表示。

证明 设元图 $S = \langle M, E \rangle$ 表示的是家谱图, m_{k1} 与 m_{k2} 为具有夫妻关系并拥有父母 m_{i1} 、 m_{j1} 和 m_{i2} 、 m_{j2} 信息的家庭成员,即 m_{k1} 、 $m_{k2} \in M$ 且 $m_{k1} P m_{k2}$ 。由于 m_{k1} 、 m_{k2} 为夫妻关系,所以 m_{k1} 、 m_{k2} 形成一个节点, m_{i1} 、 m_{j1} 与 m_{i2} 、 m_{j2} 可作为一个节点与其他成员相连。由定义 4 和性质 2 可知一名已婚子女至少需要一条边与其他成员连接表示,因而

m_{k1} 、 m_{k2} 需要两条边与其他成员连接表示。由于与 m_{k1} 、 m_{k2} 为独生子女,因此在 m_{i1} 、 m_{j1} 和 m_{i2} 、 m_{j2} 所有的子女成员中只有 m_{k1} 与 m_{k2} 一对夫妻节点,其与 m_{i1} 、 m_{j1} 和 m_{i2} 、 m_{j2} 产生一一对应关系。故 m_{k1} 与 m_{k2} 节点需要一条边与 m_{i1} 、 m_{j1} 和 m_{i2} 、 m_{j2} 节点进行连接,在家谱元图可视化中 m_{k1} 与 m_{k2} 节点共需一条边与其他成员连接表示。□

2.18 亿独生子女约占我国人口总数 14 亿的 15.6%,夫妻双方同为独生子女的比例占 2.4%,夫妻有一方为独生子女的比例占 26.4%,共占 28.8%。由定义 5 及其性质 4 可得,当一对夫妻双方都为独生子女时,其在家谱元图可视化中仅需一条边与其他成员节点连接即可;又由定义 4 和定义 5 及其性质 2~性质 4 可知,当夫妻双方中一方为独生子女,另一方为唯一已婚子女时仅需一条边与其他成员节点相连。因此家谱元图可视化表示中边的个数约为:

$$I.|E| = 0.7 \times (2b - c) \quad (3)$$

由图 1 可以看出,与通常的树形展示形式、原有的元图可视化形式相比,优化的家谱成员关系元图可视化形式具有可视化效果更加简洁明了,家谱主要关系表示简单、鲜明、清晰、易读且便于解释等优点。

3 可视化性能展示

优化后的家谱元图可视化有效地降低了家谱的数据复杂性,更加易于观察与分析。此外通过改进元图可视化的展示形式有效地减少了元图中边的展示数量,一方面没有表示父母-未婚子女的有向边,另一方面将独生已婚子女和唯一已婚子女按照实际情况进行边的合并,提高了家谱数据可视化展示的效果。

假设家族成员总数为 a ,具有夫妻关系的对数为 b ,无父母信息显示的成员数为 c ,规定具有夫妻关系的成员可拥有子女。用普通图进行可视化展示时,所需边的条数为父辈节点同其所有孩子节点间边的条数与夫妻节点间边的条数之和,记为:

$$G.|E| = 2 \times (a - c) + b \quad (4)$$

原有的元图可视化表示中边的个数等于家谱中孩子节点出现的个数,即 $S.|E| = a - c$ 。在优化后的家谱元图可视化表示中所需边的条数约为 $I.|E| = 0.7 \times (2b - c)$ 。

为便于展示优化效果,以文献[14]中一个包含 15 个家庭、50 名家庭成员的家谱关系数据图表为

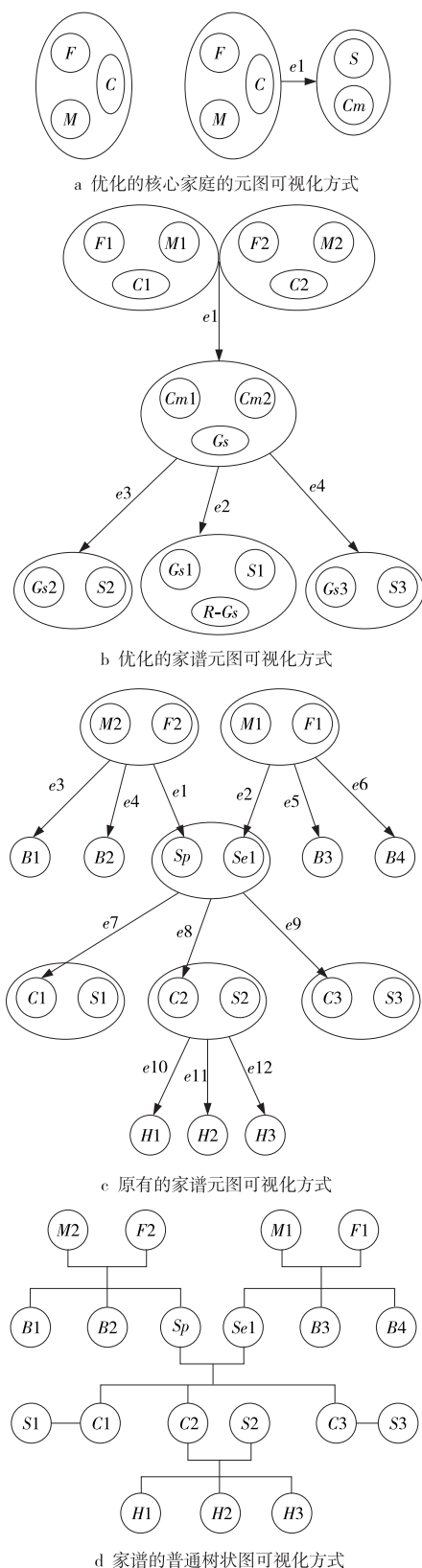


Figure 1 Comparison of improved metagraphic visualization of genealogical member relationship among the improved metagraph, existing metagraph and common form

图1 家谱成员关系的优化元图可视化与原有元图和普通图可视化形式对比例进行比较,通过原有元图可视化形式中数据关系

$S. |E| = a - c$ 以及原表中 $S. |E|$ 和 a 计算其中的 c 值,然后通过公式(3)计算出优化元图可视化表示中所需的边数。表1为该家族家谱优化的元图可视化形式 I 、原有的元图可视化形式 S 以及通常的家谱图可视化形式 G ,在节点数目以及家族中家庭数目不同时所需边数量的对比情况。图2为对比结果的直观线性表示,从图2中可以看出,在家族中节点数量一定的情况下,原有元图可视化展示形式中边的数量约为普通展示形式的 $1/2$,而优化的元图可视化表示形式中边的数量约为原有元图可视化形式的 $1/2$,其有效地减少了可视化表示中所需边的数量。

Table 1 Comparison of the visualization among the improved metagraph, existing metagraph and common form

表1 优化的元图、原有元图和普通图可视化形式对比

a	b	$I. E $	$S. E $	$G. E $
4	1	0	2	5
9	2	1	6	14
11	3	2	7	16
14	4	4	10	21
17	5	5	12	26
19	6	6	13	29
22	7	6	15	34
24	8	8	17	37
28	9	9	20	44
30	10	10	21	47
33	11	11	23	52
38	12	12	27	61
41	13	13	29	64
44	14	13	31	67
50	15	15	37	80

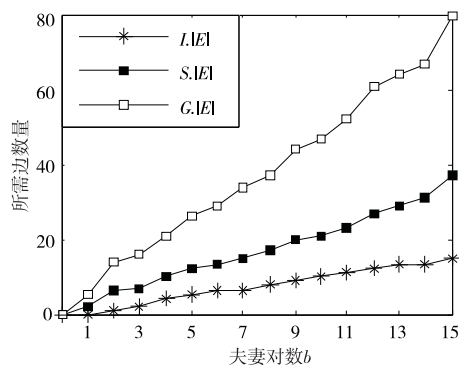


Figure 2 Comparisons of edge number among the three visualizations

图2 可视化图中边的数量对比

此外,统计了山东省某市家族成员数量分别超过20、40、60、80和100人的家庭成员情况,根据统计情况依照公式(1)、公式(3)、公式(4)分别计算出三类可视化形式所需边数,最后求取平均值。家庭

成员情况如表 2 所示, p 表示每组家庭成员数量; g 表示统计组数, 成员数量超过 20 人的家庭数据统计 50 组, 超过 40 人的家庭数据统计 28 组, 超过 60 人家庭数据统计 15 组, 超过 80 人的家庭数据统计 11 组; a 表示在每一家庭成员数量上统计所有组数的成员总数; b 表示具有夫妻关系的总对数; c 表示无父母信息显示的成员总数。每一家庭成员数量上优化的家谱元图可视化表示中所需边的总数为 $Ia \cdot |E|$, 原有元图可视化表示中边的总数为 $Sa \cdot |E|$, 普通图可视化表示中边的总数为 $Ga \cdot |E|$ 。通过实际所需总边数 $Xa \cdot |E|$ 与实际成员总数 a 的比值等于标准家庭每组应用边数 $X \cdot |E|$ 与标准家庭每组成员数量 p (20, 40, 60, 80 或 100) 的比值求出标准家庭每组所需边数 $X \cdot |E| = p \times Xa \cdot |E| / a$ 。图 3 绘制了具有相同家庭成员数量时三种可视化方式所需边数量平均值的对比情况。

Table 2 Statistics of member relationships for different family sizes
表 2 不同规模家庭成员关系统计情况

p	g	a	b	c	$Ia \cdot E $	$Sa \cdot E $	$Ga \cdot E $
20	50	1 036	371	385	334	648	1 668
40	50	2 043	697	651	664	1 399	3 496
60	28	1 694	515	506	451	1 201	2 917
80	15	1 211	422	404	352	798	2 023
100	11	1 125	409	418	285	707	1 817

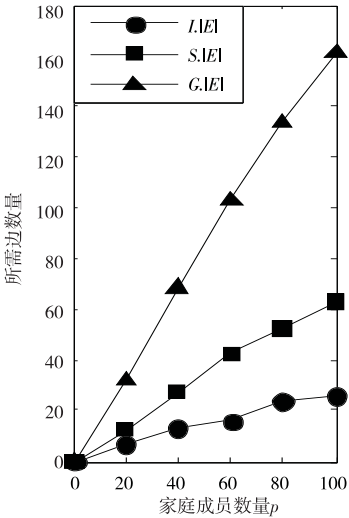


Figure 3 Comparison of edges required in the three visualizations while having the same population

图 3 具有相同成员数量时三种可视化方式所需边数量对比

4 可视化应用实例

优化的家谱元图可视化有效地减少了边的数

量。在实际应用中,应当考虑家谱的具体情况进行可视化显示。一般来说,原始的树形层次结构最符合家谱可视化显示,能够直观地显示出家谱成员之间的关系。但是,由于树形结构分支众多,每一个分支都需要用边进行连接,同时占据一定空间,这不仅需要使用大量的边,并且占据了大量空间,降低了效率。基于空间填充式的显示方法虽然省去了边的使用,同时极大程度上减少了空间的占据,但由于没有了连接边,其层间各部分之间的关系不明显,如用在家谱可视化显示中家谱成员之间的关系不能够直接辨别出。因此,通过分析以上几种层次结构可视化方法,结合家谱可视化实际需要,使用节点连接图的树形结构对家谱进行可视化显示^[25]。

生活中,夫妻之间的关系通常会由于夫妻双方离婚后再婚出现一个成员拥有多个配偶的情况。为了在可视化的展示结果中方便展示,规定按照当前家谱中夫妻关系中拥有“父子关系”或“母子关系”成员的一方为主,其所拥有的多个配偶列表显示,可视化元图中只显示当前最后一个配偶及其子女,但颜色不同于其他。当鼠标点击显示配偶时,所有配偶全部显示。鼠标选择任意一个配偶后,有且仅有被选配偶给予显示;同时,配偶相关子女随之变换显示。这样不仅避免了原有元图可视化中继父母表示的复杂性,也便于使用者更好地理解家庭成员之间的关系。为了区分家谱主要家庭单元,把家谱中本家族男性成员家庭单元显示添加底色。使用双虚线边表示具有过继关系的家谱成员与其亲生父母之间的关系,此外在拥有养父母的家谱成员名字下添加下划线加以表示(如图 4c 中,姓名为“王功”的成员,其亲生父母为“刘兵”和“王婷”,养父母为“王华”和“朱新”)。

在系统的可视化展示结果中,一个成员节点中的矩形节点默认为男左女右,分别代表家族中的男性成员和女性成员,左侧父母节点显示左侧成员(男性成员)的父母家庭信息,右侧父母节点显示右侧成员(女性成员)的父母(如图 4b 中,“刘武”“孙芳”成员节点中左侧成员“刘武”的父母为“刘劲”和“朱云”,右侧成员“孙芳”的父母为“孙明”和“沈燕”)。把女性成员的显示背景添加一定灰度值的底色,这样有灰度底色的子女为女儿,而底色显示白色的子女是儿子。实线的椭圆节点表示具有夫妻关系的两个人员节点组成的入点集,两个成员节点集合之间的实线边和单虚线边分别表示家谱男性子女成员与父母之间和家谱女性子女成员与父

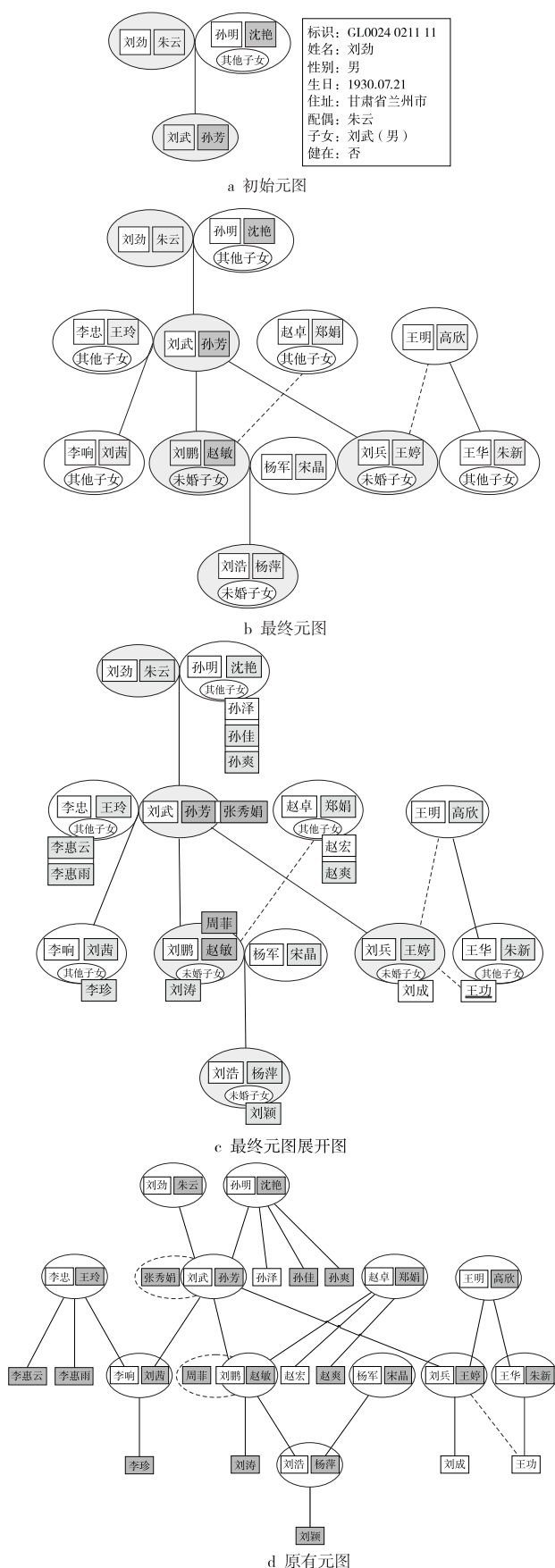


Figure 4 Case diagram of metagraphic visualization of genealogy

图4 家谱元图可视化应用实例图

母之间的“父母子女”关系。将可视化元图中主要家庭单元拥有未婚子女的成员节点的未婚子女部分以标有“未婚子女”标注的椭圆图形代替,其他单元节点的子女以标有“其他子女”标注的椭圆图形代替,当鼠标点击时予以显示;将可扩展成员节点隐藏,当鼠标经过时以颜色变化来表示存在隐藏的节点,从而提高展示空间的利用率。同时,可通过右击家谱成员节点中的成员图形来查看该家谱成员的详细信息(如图4a中姓名为“刘劲”的成员)。图4a到图4b表示“刘武”“孙芳”成员节点通过扩展得到家谱成员关系元图的过程,图4c为最终元图的展开图,图4d为原有元图可视化的对比图。

通过图4b和图4c可以看出,运用节点连接图的树形结构很好地对优化后的家谱元图进行展示。与图4d的原有元图可视化相比,优化的家谱元图可视化在显示时使用边的数量少,结构更加清晰;使用单元节点的数量少,简单整洁;树的层数减少,层内节点数减少,占据的空间也相对减少,取得了良好的显示效果。

5 结束语

家谱成员的数据可视化是电子家谱的重要组成部分,其表现直接关系到电子家谱在实际应用中的效率和交互能力。家谱成员节点的数量和边的数量随着家谱规模的增大而迅速增加,是家谱可视化中的一个重要问题,这会导致展示效率的急速下降,影响数据表现。本文结合实际情况,在原有家谱元图可视化基础上进行优化,通过对已婚子女、未婚子女以及独生子女的论证进一步有效地降低了图中边的数量,提高了家谱成员的展示效率。同时,根据家谱中各成员之间的关系,将家谱元图层次化,增强了可视化效果的实用性和简洁性。

家谱元图可视化的设计提高了电子家谱使用的实用性和易用性,家谱成员关系元图可视化作为其中的重要部分,今后将适用于家谱搜索算法以及家谱成员流动模式的预测研究。

参考文献:

- [1] Wang He-ming. General theory of Chinese genealogy [M]. Shanghai: Shanghai Classics Publishing House, 2010: 4. (in Chinese)
- [2] Chen Guo-jun, Zhang Ting-yu. The design technique of binary-based e-genealogy [J]. Computer Engineering & Software, 2015, 36(3): 95-97. (in Chinese)
- [3] 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.
- [4] Furnas G W, Zacks J. Multitrees: Enriching and reusing hierarchical structure [C] // Proc of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM, 1994: 330-336.
- [5] Wesson M J, Plessis M D, Oosthuizen C. A zoomtree interface for searching genealogical information [C] // Proc of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualization and Interaction in Africa, 2004: 131-136.
- [6] Mc Guffin M J, Balakrishnan R. Interactive visualization of genealogical graphs [C] // Proc of the 2005 IEEE Symposium on Information Visualization. Washington, DC: IEEE Computer Society, 2005: 17-24.
- [7] Tuttle H C, Nonato L G, Silva C. PedVis: A structured, space-efficient technique for pedigree visualization [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2010, 16(6): 1063-1072.
- [8] Splaisant C, Grosjean J, Bederson B B. Space tree: Supporting exploration in large node link tree, design evolution and empirical evaluation [C] // Proc of the IEEE Symposium on Information Visualization. Washington, DC: IEEE Computer Society, 2002: 57-64.
- [9] Draper G M, Riesenfeld R F. Interactive fan charts: A space saving technique for genealogical graph exploration [C] // Proc of the 8th Annual Workshop on Technology for Family History and Genealogical Research, 2008: 1-7.
- [10] Kim N W, Card S K, Heer J. Tracing genealogical data with Time-Nets [C] // Proc of the International Conference on Advanced Visual Interfaces, 2010: 241-248.
- [11] Bezerianos A, Dragicevic P, Fekete J D, et al. GeneaQuilts: A system for exploring large genealogies [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2010, 16(6): 1073-1081.
- [12] Johnson B, Shneiderman B. Tree-Maps: A space-filling approach to the visualization of hierarchical information structures [C] // Proc of the 2nd Conference on Visualization. Washington, DC: IEEE Computer Society, 1991: 284-291.
- [13] Loh B A M, Wilshire S, Emery J, et al. Celestial3D: A novel method for 3D visualization of familial data [J]. *Bioinformatics*, 2008, 24(9): 1210-1211.
- [14] Liu Jun-dan, Zhao Shu-liang, Zhao Jiao-jiao, et al. Meta-graph for genealogical relationship visualization [J]. *Journal of Computer Applications*, 2013, 33(7): 2037-2040. (in Chinese)
- [15] Basu A, Blanning R. Metagraphs [J]. *Omega*, 1995, 23(1): 13-25.
- [16] Basu A, Blanning R. Metagraphs and their applications [M]. Berlin: Springer-Verlag, 2006.
- [17] Basu A, Blanning R W. Cycles in metagraphs [C] // Proc of the 27th Annual Hawaii International Conference on System Sciences. Washington DC: IEEE Computer Society, 1994: 310-319.
- [18] Basu A, Blanning R W. Metagraphs: A tool for modeling decision support systems [J]. *Management Science*, 1994, 40(12): 1579-1600.
- [19] Basu A, Blanning R W. A formal approach to workflow analysis [J]. *Information Systems Research*, 2000, 11(1): 17-36.
- [20] Fan Hong, Liu Jin. Quantitative analysis of metagraph model for supply chain [J]. *Control and Decision*, 2005, 20(10): 1152-1156. (in Chinese)
- [21] Liu Jin. Metagraph and its application in expert system [J]. *Pattern Recognition and Artificial Intelligence*, 1999, 12(1): 6-12. (in Chinese)
- [22] Yi Fu-xian. Big empty nest: Reflection on China's family planning [M]. Beijing: China Development Press, 2012. (in Chinese)
- [23] Guo Zhi-gang. Estimating cohort percentage of only-children using data of population census [J]. *Market & Demographic Analysis*, 2001, 7(1): 5-11. (in Chinese)
- [24] Population Census Office of the State Council, Population Employment Statistical Department of National Bureau of Statistics of China. Census information of China in 2010 [M]. Beijing: China Statistics Press, 2012: 4. (in Chinese)
- [25] Cheng Nie, Huang Kun, Su Xin, et al. Bloom filters based on the tree structure [J]. *Computer Engineering & Science*, 2012, 34(2): 19-24. (in Chinese)

附中文参考文献:

- [1] 王鹤鸣. 中国家谱通论 [M]. 上海: 上海古籍出版社, 2010: 4.
- [2] 陈国军, 张庭玉. 基于二叉树的电子家谱设计方法 [J]. *软件*, 2015, 36(3): 95-97.
- [14] 刘军丹, 赵书良, 赵娇娇, 等. 家谱关系的元图表示 [J]. *计算机应用*, 2013, 33(7): 2037-2040.
- [20] 樊宏, 刘晋. 供应链 Meta 图模型定量分析研究 [J]. *控制与决策*, 2005, 20(10): 1152-1156.
- [21] 刘晋. Meta 图及其在专家系统中的应用 [J]. *模式识别与人工智能*, 1999, 12(1): 6-12.
- [22] 易富贤. 大国空巢: 反思中国计划生育政策 [M]. 北京: 中国发展出版社, 2012.
- [23] 郭志刚. 利用人口普查原始数据对独生子女信息的估计 [J]. *市场与人口分析*, 2001, 7(1): 5-11.
- [24] 国务院人口普查办公室, 国家统计局人口和就业统计司. 中国 2010 年人口普查资料 [M]. 北京: 中国统计出版社, 2012: 4.
- [25] 程聂, 黄昆, 苏欣, 等. 一种基于树形结构的布鲁姆过滤器 [J]. *计算机工程与科学*, 2012, 34(2): 19-24.

作者简介:



陈国军(1991-), 男, 山东德州人, 硕士生, CCF 会员(E200050269G), 研究方向为数字媒体、脉冲神经网络和深度学习。
E-mail: junxiu828@qq.com

CHEN Guo-jun, born in 1991, MS candidate, CCF member(E200050269G), his research interests include digital media, spiking neural networks, and deep learning.